PREDICTING STUDENT PERFORMANCE IN ONLINE ENGLISH LANGUAGE LEARNING DURING CHALLENGING TIMES THROUGH LEARNING ANALYTICS

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

MEHMET ALİ ÇELİKBAĞ

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF DOCTOR OF PHILOSOPHY
IN
COMPUTER EDUCATION AND INSTRUCTIONAL TECHNOLOGY

MARCH 2025

Approval of the thesis:

PREDICTING STUDENT PERFORMANCE IN ONLINE ENGLISH LANGUAGE LEARNING DURING CHALLENGING TIMES THROUGH LEARNING ANALYTICS

submitted by **MEHMET ALİ ÇELİKBAĞ** in partial fulfillment of the requirements for the degree of **Doctor of Philosophy** i**n Computer Education and Instructional Technology, Middle East Technical University** by,

Prof. Dr. Naci Emre Altun
Dean, Graduate School of **Natural and Applied Sciences** ________________

Prof. Dr. Soner Yıldırım
Head of the Department, **Computer Edu. and Inst. Tech.** ________________

Prof. Dr. Ömer Delialioğlu
Supervisor, **Computer Edu. and Inst. Tech., METU** ________________

Assoc. Prof. Dr. Erkan Er
Co-Supervisor, **Computer Edu. and Inst. Tech., METU** ________________

Examining Committee Members:

Prof. Dr. Arif Altun
Computer Edu. and Inst. Tech., Hacettepe University ________________

Prof. Dr. Ömer Delialioğlu
Computer Edu. and Inst. Tech., METU ________________

Prof. Dr. Ceylan Yozgatlıgil
Statistics, METU ________________

Prof. Dr. Tuğba Tokel
Computer Edu. and Inst. Tech., METU ________________

Assoc. Prof. Dr. Gökhan Akçapınar
Computer Edu. and Inst. Tech., Hacettepe University ________________

Date: 03.03.2025

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last name : Mehmet Ali Çelikbağ

Signature :

ABSTRACT

PREDICTING STUDENT PERFORMANCE IN ONLINE ENGLISH LANGUAGE LEARNING DURING CHALLENGING TIMES THROUGH LEARNING ANALYTICS

Çelikbağ, Mehmet Ali
Doctor of Philosophy, Computer Education and Instructional Technology
Supervisor: Prof. Dr. Ömer Delialioğlu
Co-Supervisor: Assoc. Prof. Dr. Erkan Er

March 2025, 164 pages

Learning a foreign language is a complex and significant process that influences academic, business, and social life. The COVID-19 pandemic profoundly impacted education, making online English language learning a new and challenging experience for students. This shift raised important research questions about factors affecting students' academic performance.

Under normal circumstances, university students would begin their language studies face-to-face, but the pandemic necessitated online learning. The study involved 481 students from diverse backgrounds and departments. Various features influenced academic performance at different levels and across language skills, including use of language, writing, and speaking. Gateway and proficiency exams were analyzed to develop a comprehensive understanding of online English language learning during the pandemic. The number of logins, assignment submissions, and attendance in virtual classrooms often played a crucial role in predicting achievement in online English language learning.

To balance data distribution, SMOTE was applied as an oversampling technique, and 10-fold stratified sampling was used to reduce sampling bias. Several classification algorithms were tested, with Logistic Regression and Naïve Bayes performing well in most cases. Additionally, Gradient Boosting, Neural Networks, Random Forest, and SVM were effective in predicting student achievement.

The findings highlight the role of instructional design and technology in facilitating online learning, particularly in times of crisis. Learning analytics considerations and further implications were explored to enhance English language education in higher education. By leveraging technology and data-driven approaches, universities can optimize online learning experiences and better support students in achieving academic success.

Keywords: Online English Language Learning, Learning Analytics, Predicting Student Performance, Online Learning During the COVID-19 Pandemic

ÖZ

ZORLU ZAMANLARDA ÖĞRENME ANALİTİĞİ İLE ÇEVRİMİÇİ İNGİLİZCE DİL ÖĞRENİMİNDE ÖĞRENCİ PERFORMANSINI TAHMİN ETME

Çelikbağ, Mehmet Ali
Doktora, Bilgisayar ve Öğretim Teknolojileri Eğitimi
Tez Yöneticisi: Prof. Dr. Ömer Delialioğlu
Ortak Tez Yöneticisi: Doç. Dr. Erkan Er

Mart 2025, 164 sayfa

Yabancı bir dil öğrenmek, akademik, iş ve sosyal hayatı etkileyen karmaşık ve önemli bir süreçtir. COVID-19 pandemisi eğitimi derinden etkileyerek öğrenciler için çevrimiçi İngilizce öğrenimini yeni ve zorlu bir deneyim hâline getirdi. Bu değişim, öğrencilerin akademik performanslarını etkileyen faktörler üzerine önemli araştırma sorularını gündeme getirdi.

Normal şartlarda üniversite öğrencileri dil eğitimine yüz yüze başlarken, pandemi nedeniyle çevrimiçi eğitime geçmek zorunda kaldılar. Çalışma, çeşitli geçmişlere ve bölümlere sahip 481 öğrenciyi içermiştir. Farklı özellikler, akademik performansı farklı seviyelerde ve dil becerileri arasında (dil kullanımı, yazma, konuşma) etkiledi. Pandemi sürecinde üniversite düzeyinde çevrimiçi İngilizce dil öğrenimini kapsamlı bir şekilde anlamak için giriş ve yeterlilik sınavları analiz edildi. Giriş sayısı, ödev teslimleri ve sanal sınıflardaki devamlılık, çevrimiçi İngilizce dil öğreniminde başarıyı tahmin etmede genellikle kritik bir rol oynamıştır.

Veri dağılımını dengelemek için SMOTE aşırı örnekleme tekniği uygulanmış ve örnekleme yanlılığını azaltmak amacıyla 10 katmanlı ayrıştırılmış örnekleme yöntemi kullanılmıştır. Birçok sınıflandırma algoritması test edilmiş olup, Lojistik Regresyon ve Naïve Bayes çoğu durumda başarılı sonuçlar vermiştir. Ayrıca, Gradient Boosting, Yapay Sinir Ağları, Rastgele Orman ve SVM, öğrenci başarısını tahmin etmede etkili olmuştur.

Bulgular, özellikle kriz dönemlerinde çevrimiçi öğrenmeyi kolaylaştırmada öğretim tasarımı ve teknolojinin rolünü vurgulamaktadır. Öğrenme analitiği ve ileriye dönük çıkarımlar ele alınarak yükseköğretimde İngilizce dil eğitiminin geliştirilmesi amaçlanmıştır. Üniversiteler, teknoloji ve veri odaklı yaklaşımları kullanarak çevrimiçi öğrenme deneyimlerini optimize edebilir ve öğrencilerin akademik başarılarını destekleyebilir.

Anahtar Kelimeler: Çevrimiçi İngilizce Dil Öğrenimi, Öğrenme Analitiği, Öğrenci Performansını Tahmin Etme, COVID-19 Pandemisi Sırasında Çevrimiçi Öğrenme

To my family…

ACKNOWLEDGMENTS

I would like to extend my heartfelt gratitude to my supervisor, Prof. Dr. Ömer Delialioğlu, and my co-supervisor, Assoc. Prof. Dr. Erkan Er, for their invaluable mentorship, insightful guidance, constructive criticism, and unwavering support throughout this research endeavor.

I am also deeply thankful to Prof. Dr. Tuğba Tokel and Prof. Dr. Ceylan Yozgatlıgil for their generous contributions, insightful suggestions, thoughtful comments, and valuable time during the numerous thesis monitoring committee meetings.

My deepest gratitude goes to my family for their unwavering support, patience, and encouragement during the challenging process of pursuing my doctoral degree. Their steadfast belief in me has been a constant source of strength and motivation.

Finally, I would like to express my profound appreciation to my friends and colleagues for their unwavering encouragement, and sincere inspiration throughout this arduous journey. Their support and camaraderie played an instrumental role in helping me overcome challenges and persevere through every phase of this endeavor.

TABLE OF CONTENTS

ABSTRACT v
ÖZ vii
ACKNOWLEDGMENTS x
TABLE OF CONTENTS xi
LIST OF TABLES xvi
LIST OF FIGURES xviii
LIST OF ABBREVIATIONS xx
CHAPTERS
1 INTRODUCTION 1
1.1 Background of the Study 1
1.2 Statement of the Problem 2
1.3 Purpose of the Study 4
1.4 Significance of the Study 8
1.5 Definition of Terms 9
2 LITERATURE REVIEW 11
2.1 Learning Analytics in Contemporary Education 12
2.2 Benefits and Challenges in Learning Analytics 13
2.2.1 Benefits 13
2.2.2 Challenges 14
2.3 Learning Analytics in Online English Language Learning 15
2.4 Factors Affecting Student Performance in Online English Language Learning 19

2.5 Exploring Students’ Engagement Types 20
2.6 Relevance to the Research Study 22
2.6.1 Demographic Influence in Online English Language Learning 22
2.6.2 Quarterly Academic Engagement and Performance 24
2.6.3 Engagement in Virtual Learning Environments 26
2.6.4 Skill-Specific Interactions 27
2.7 Ethical Considerations and Data Privacy in Learning Analytics 29
2.8 Conclusion 30
3 METHODOLOGY 33
3.1 Context and Data 33
3.2 Population, Sampling, and Participants 45
3.3 Research Method 46
3.3.1 Learning Analytics Objectives 47
3.3.2 Frameworks and Dimensions of Learning Analytics 48
3.3.3 Predictive Models and Algorithms 50
3.3.4 Tools and Programming Languages 58
3.4 Data Sources and Instruments 58
3.5 Data Collection Procedure 59
3.6 Data Pre-Processing 61
3.7 Data Mining Tool Selection 62
3.8 Data Analysis 63
3.9 Ethical Issues 64
3.10 Researcher’s Background 65
3.11 Delimitation and Limitations of the Study 67

3.11.1 Delimitation....67

3.11.2 Limitations 67

4 RESULTS 71

4.1 Summary of Descriptive Results 71

4.1.1 Descriptive Results Related to Faculties....71

4.1.2 Descriptive Results Related to Departments....73

4.1.3 Descriptive Results Related to Vocational Schools and Faculties....76

4.1.4 Descriptive Results Related to Language of Instruction at Designated Department....76

4.1.5 Descriptive Results Related to Type of Instruction at Designated Department....77

4.1.6 Descriptive Results Related to Gender....78

4.2 Findings Related to Research Question 1 79

4.2.1 Findings Related to Sub-Research Question 1.1....79

4.2.2 Findings Related to Sub-Research Question 1.2....83

4.3 Findings Related to Research Question 2 91

4.3.1 Findings Related to Sub-Research Question 2.1.1....91

4.3.2 Findings Related to Sub-Research Question 2.1.2....94

4.3.3 Findings Related to Sub-Research Question 2.2.1....95

4.3.4 Findings Related to Sub-Research Question 2.2.2....98

4.3.5 Findings Related to Sub-Research Question 2.3.1....100

4.3.6 Findings Related to Sub-Research Question 2.3.2....103

4.3.7 Findings Related to Sub-Research Question 2.4.1....104

4.3.8 Findings Related to Sub-Research Question 2.4.2....107

4.4 Findings Related to Research Question 3 109
4.4.1 Findings Related to Sub-Research Question 3.1.1 109
4.4.2 Findings Related to Sub-Research Question 3.1.2 111
4.4.3 Findings Related to Sub-Research Question 3.2.1 113
4.4.4 Findings Related to Sub-Research Question 3.2.2 116
4.4.5 Findings Related to Sub-Research Question 3.3.1 119
4.4.6 Findings Related to Sub-Research Question 3.3.2 121
4.5 Summary of the Results 123
5 DISCUSSION AND CONCLUSION 127
5.1 Introduction 127
5.2 Discussion of Descriptive Results 128
5.2.1 Faculties 128
5.2.2 Departments 129
5.2.3 Vocational Schools and Faculties 129
5.2.4 Language of Instruction at Designated Department 130
5.2.5 Type of Instruction at Designated Department 131
5.2.6 Gender 132
5.3 Predictive Basis for Early Warning Systems 132
5.4 Gateway and Proficiency Exams Prediction 134
5.5 Skill-Based Prediction 136
5.6 Methodological Reflections 137
5.7 Conclusion 139
5.8 Practical Implications 140
5.9 Implications for Future Research 141

REFERENCES 143

APPENDICES 161

A. Ethics Approval Form 161

CURRICULUM VITAE 163

LIST OF TABLES

TABLES

Table 3.1. Feature List and Descriptions 36
Table 4.1. Achievement According to Faculties 72
Table 4.2. Achievement According to Departments 74
Table 4.3. Achievement According to Vocational School/Faculty 76
Table 4.4. Achievement According to Language of Instruction 77
Table 4.5. Achievement According to Type of Instruction 78
Table 4.6. Achievement According to Genders 78
Table 4.7. Overall Performance Scores of Classification Algorithms 79
Table 4.8. Confusion Matrix of Classification Algorithm for Early Prediction 81
Table 4.9. Performance Scores of Classification Algorithms Using SMOTE for Gateway Exam in the First Quarter 92
Table 4.10. Confusion Matrix of Classification Algorithm Using SMOTE for Gateway Exam in the First Quarter 93
Table 4.11. Performance Scores of Classification Algorithms for Gateway Exam in the Second Quarter 96
Table 4.12. Confusion Matrix of Classification Algorithm for Gateway Exam in the Second Quarter 97
Table 4.13. Performance Scores of Classification Algorithms Using SMOTE for Gateway Exam in the Third Quarter 101
Table 4.14. Confusion Matrix of Classification Algorithm using SMOTE for Gateway Exam in the Third Quarter 102
Table 4.15. Performance Scores of Classification Algorithms for Proficiency Exam in the Fourth Quarter 105
Table 4.16. Confusion Matrix of Classification Algorithm for Proficiency Exam in the Fourth Quarter 107

Table 4.17. Performance Scores of Classification Algorithms for the Use of English Skills (Grammar/Reading/Listening) in the Second Quarter 109
Table 4.18. Confusion Matrix of Classification Algorithm for the Use of English Skills (Grammar/Reading/Listening) in the Second Quarter 110
Table 4.19. Performance Scores of Classification Algorithms for Writing Skills in the Second Quarter 113
Table 4.20. Confusion Matrix of Classification Algorithm for Writing Skills in the Second Quarter 115
Table 4.21. Performance Scores of Classification Algorithms for Speaking Skills in the Second Quarter 119
Table 4.22. Confusion Matrix of Classification Algorithm for Speaking Skills in the Second Quarter 121

LIST OF FIGURES

FIGURES

Figure 3.1. A View of the Course Interface in the LMS Utilized by the School of Foreign Languages 34
Figure 3.2. A View of the Weekly Course Interface in the LMS Used by the School of Foreign Languages 34
Figure 3.3. An Overview of the Level-Based Material Interface in the LMS Utilized by the School of Foreign Languages 35
Figure 3.4. The LA cycle, from Clow (2012, p. 134). 47
Figure 3.5. Critical dimensions of LA, from Greller and Drachsler (2012, p. 44).. 49
Figure 3.6. The Predictive Model in This Research Study 56
Figure 3.7. The Predictive Model Using SMOTE in This Research Study 57
Figure 4.1. Explanation of the Model for Gateway and Proficiency Exams in the Early Warning System (Q1F Target) 83
Figure 4.2. Explanation of the Model for Gateway and Proficiency Exams in the Early Warning System (Q2F Target) 86
Figure 4.3. Explanation of the Model for Gateway and Proficiency Exams in the Early Warning System (Q3F Target) 87
Figure 4.4. Explanation of the Model for Gateway and Proficiency Exams in the Early Warning System (Q4F Target) 88
Figure 4.5. Explanation of the Model for Gateway and Proficiency Exams in the Early Warning System (Pass Target) 89
Figure 4.6. Explanation of the Model Using SMOTE for Gateway Exam in the First Quarter (Logistic Regression) 94
Figure 4.7. Explanation of the Model for Gateway Exam in the Second Quarter (Naïve Bayes) 99
Figure 4.8. Explanation of the Model Using SMOTE for Gateway Exam in the Third Quarter (Gradient Boosting) 103

Figure 4.9. Explanation of the Model for Proficiency Exam in the Fourth Quarter (Naïve Bayes).. 108
Figure 4.10. Explanation of the Model for the Use of English Skills (Grammar/Reading/Listening) in the Second Quarter (Naïve Bayes)................. 111
Figure 4.11. Explanation of the Model for Writing Skills in the Second Quarter (Naïve Bayes).. 117
Figure 4.12. Explanation of the Model for the Use of Speaking Skills in the Second Quarter (Naïve Bayes) 122
Figure 5.1. Interplay of Explainability-Performance-Fairness, from Gunasekara and Saarela (2024, p. 4). 138

LIST OF ABBREVIATIONS

ABBREVIATIONS

DE: Distance Education

EDM: Educational Data Mining

EFL: English as a Foreign Language

EWS: Early Warning Systems

k-NN: K-Nearest Neighbors

LA: Learning Analytics

LAD: Learning Analytics Dashboard

LMS: Learning Management System

MOODLE: Modular Object-Oriented Dynamic Learning Environment

RQ: Research Question

SGD: Stochastic Gradient Descent

SHAP: SHapley Additive exPlanations

SIS: Student Information System

SMOTE: Synthetic Minority Oversampling Technique

SVM: Support Vector Machines

CHAPTER 1

INTRODUCTION

This study aims to investigate and identify the influential factors that impact student achievement in online English language learning measured by gateway exams and English proficiency exams. By utilizing Learning Analytics (LA), the research aims to discover insights to enhance the online English language learning experience and improve student results at the university level. Additionally, the research endeavors to bring some understanding for the effective design of online language courses, particularly in English, at the university level by identifying major influential factors on student performance. This chapter is organized into the following sections: Background of the Study, Statement of the Problem, Purpose of the Study and Research Questions, and Significance of the Study.

1.1 Background of the Study

The delivery of educational content via the internet, known as online learning, has seen a significant rise in prevalence within educational institutions, especially from the 21st century onwards. The increase in online students also forced education professionals and researchers to investigate the impacts of online learning such as student performance and engagement (Akpen et al., 2024). This trend has further accelerated during the global COVID-19 pandemic, leading to the widespread adoption of online learning platforms and strategies (Elnour et al., 2023). Advances in instructional design and technology have prioritized the integration of information and communication technologies to improve instructional processes, leading to innovative teaching methods and enhanced pedagogical strategies.

The COVID-19 pandemic forced all educational institutions to transition to online learning worldwide; hence, all teaching and learning were performed online as a necessity (Meng et al., 2024). The COVID-19 pandemic, which officially emerged in Türkiye in early 2020, required a quick shift to open and distance education at all levels of education, from K-12 to graduate programs. This sudden change in teaching methods presented significant challenges, especially with completing the spring semester of the 2019-2020 academic year.

During the summer of 2020, stakeholders at all levels had more time for optimal decisions and adapted instructional practices to address the unique demands of the institutions during the pandemic. These changes in educational practices also generated new research questions (RQ) and opportunities for investigation.

This study focuses on the 2020-2021 academic year, during which all aspects of English language learning, instruction, and assessment were conducted online, leaving digital traces on platforms such as Learning Management Systems (LMSs) and virtual classroom software. Students' logs can be monitored to provide appropriate interventions for efficient learning. LA provides critical insights into student behaviors on online learning platforms, aiding in the design of proactive, revised and effective instructional strategies (Martin & Whitmer, 2016; Romero-Zaldivar et al., 2012).

1.2 Statement of the Problem

The COVID-19 pandemic has significantly disrupted education at all levels, both traditional face-to-face and online. This disruption has led to a further acceleration in the adoption of online learning platforms and digital technologies, revealing both opportunities and challenges within the educational landscape (Dhawan, 2020; Basilaia & Kvavadze, 2020). While research in Distance Education (DE) and LA has grown in scope and intensity in recent years, unresolved issues and a pressing need for further investigation persist, particularly in the increasingly prevalent context of online language learning. Such investigations are critical to address disparities in access, optimize instructional methodologies, and enhance student engagement and achievement in online learning environments.

The pandemic-induced shift to open and online learning at all educational levels has created unique opportunities for Educational Data Mining (EDM) and LA research that were previously less feasible. This transition facilitated the collection, storage and analysis of extensive datasets on student behavior, engagement, and performance that have the potential to offer valuable insights for enhancing online learning environments (Müller et al., 2022; Noor, Isa, & Mazhar, 2020). The potential of future pandemics highlights the importance of examining the implications of such events on education at all levels, including the necessity to build resilient educational platforms capable of adapting to remote teaching or online teaching and learning scenarios (Bao, 2020; Hodges et al., 2020).

The convergence of traditional, blended, and online learning modalities has led to a surge in the popularity of online learning environments and a renewed emphasis on scientific research to inform their improvement. This shift has been further eased and accelerated by technological advancements, which enable the integration of diverse instructional models to design and enhance learning (Almeida & Simoes, 2019; Boelens et al., 2017). Acquiring foreign language skills, particularly English, has become increasingly important in the context of computer and internet-based technologies, as proficiency in English is often necessary for accessing global digital resources, traveling, participating and communicating in international affairs, more generally as a way of globalization and more specifically in higher education as a medium of teaching in this context (Crystal, 2003; Coleman, 2006).

A school of foreign languages at a public university tasked with teaching students from basic language skills to advanced academic English provides a valuable case for examining the implications of transitioning from traditional or blended classes to fully online instruction. Institutions with prior experience in blended learning often have structural and pedagogical advantages when adapting to fully online forms of teaching and learning, with the possibility of developing strategies for student engagement and satisfaction in online learning before difficult times (Akpen et al., 2024; Lim, Morris, & Kupritz, 2007). In this sense, this school, with its prior experience with blended learning, offers unique insights into the challenges and opportunities of such transitions, particularly with technology integration and multimedia usage as in virtual classrooms (Lim, Morris, & Kupritz, 2007).

This study explores the influential factors impacting online English language learning, as measured by proficiency exam achievement, through LA during challenging times, particularly during COVID-19. By leveraging predictive analytics, the study seeks to inform future instructional design decisions and prepare for potential challenges arising from future pandemics or natural disasters. Previous research studies have shown the potential of LA to enhance language learning outcomes by identifying key predictors of student success and at-risk students through engagement patterns and task completion rates (Reinders, 2019). The pandemics and natural disasters has highlighted the critical need for flexible instructional strategies that utilize data-driven approaches to improve online learning environments and adapt teaching methods to difficult times (Müller et al., 2022; Chen et al., 2020). Despite its potential, limited research has explored insights related to remote learning and online learning in English language education during extraordinary circumstances in the context of Türkiye.

1.3 Purpose of the Study

This study aims to identify influential features impacting online English language learning through the lens of LA during challenging times. It seeks to develop comprehensive insights for foreign language schools in Türkiye, offering evidence-based recommendations for designing language teaching programs and LMS course structures tailored to the needs of online English language learners. By leveraging LA, the study mainly aims to predict students' academic achievement in online English language learning, particularly during challenging times in higher education, such as those posed by pandemics and earthquakes. Research has shown that the integration of predictive analytics accompanies LA dashboards and can significantly enhance instructional design, foster personalized learning and teaching, and support data-informed decision-making for all stakeholders (Pardo & Siemens, 2014; Viberg et al., 2018; Gašević, Kovanović, & Joksimović, 2017). In this sense, predictive models in LA can identify feature importance and at-risk students, enabling proactive interventions even in the absence of instructors, thus improving retention and learning outcomes in subsequent phases (Akçapınar, Altun, & Aşkar, 2019; Gašević et al., 2017; Macfadyen & Dawson, 2010).

Furthermore, the development of effective, tailored LA dashboards for online English language learning in higher education holds promise for optimizing the learning environment and student experiences by providing actionable insights into student engagement and the development of skills (Ferguson & Clow, 2017; Siemens, 2013). Such dashboards can support instructors in monitoring student progress across progress tests, gateway exams, proficiency exams, key language competencies, which are reading, writing, listening, and speaking, and adapting teaching strategies in real-time for desired learning goals. Nonetheless, this research study focuses on identifying key features and making predictions in online English language learning during difficult times as the foundational phases of LA with plans and strategies for future studies involving stakeholders. Ultimately, this research seeks to contribute to the growing body of literature on LA by advancing methodologies and tools that improve online English language learning outcomes in diverse and dynamic educational contexts in higher education.

The following RQs and sub-questions were addressed within the context of online English language learning:

RQ1: Does engagement in LMS and virtual classrooms in the first quarter of the academic year predict academic performance in future quarters?

- RQ1.1: Does engagement in LMS and virtual classrooms in the first quarter predict academic performance on gateway exams in the first quarter, the second quarter, the third quarter, and the proficiency exam in the fourth quarter?
- RQ1.2: What are the most predictive features that contributed to the best-performing model for the gateway exams and the proficiency exam prediction?

RQ2: Does engagement in LMS and virtual classrooms in different quarters of the academic year predict academic performance across gateway exams and proficiency exams?

- RQ2.1.1: Does engagement in LMS and virtual classrooms predict academic performance on the gateway exam in the first quarter?
- RQ2.1.2: What are the top ten features that contributed to the best-performing model in the prediction performance for the gateway exam in the first quarter?
- RQ2.2.1: Does engagement in LMS and virtual classrooms predict academic performance on the gateway exam in the second quarter?
- RQ2.2.2: What are the most predictive features that contributed to the best-performing model in the prediction performance for the gateway exam in the second quarter?
- RQ2.3.1: Does engagement in LMS and virtual classrooms predict academic performance on the gateway exam in the third quarter?

- RQ2.3.2: What are the most predictive features that contributed to the best-performing model in the prediction performance for the gateway exam in the third quarter?
- RQ2.4.1: Does engagement in LMS and virtual classrooms predict academic performance on the proficiency exam in the fourth quarter?
- RQ2.4.2: What are the most predictive features that contributed to the best-performing model in the prediction performance for the proficiency exam in the fourth quarter?

RQ3: Does engagement in LMS and virtual classrooms predict academic performance related to different English language skills in the second quarter?

- RQ3.1.1: Does engagement in LMS and virtual classrooms predict academic performance in the use of English skills (grammar/reading/listening) in the second quarter?
- RQ3.1.2: What are the most predictive features that contributed to the best-performing model in the prediction performance for the use of English skills (grammar/reading/listening) in the second quarter?
- RQ3.2.1: Does engagement in LMS and virtual classrooms predict academic performance in writing skills in the second quarter?
- RQ3.2.2: What are the most predictive features that contributed to the best-performing model in the prediction performance for writing skills in the second quarter?
- RQ3.3.1: Does engagement in LMS and virtual classrooms predict academic performance in speaking skills in the second quarter?
- RQ3.3.2: What are most predictive features that contributed to the best-performing model in the prediction performance for speaking skills in the second quarter?

1.4 Significance of the Study

As the landscape of higher education increasingly shifts toward digital modalities, understanding how students engage with and succeed in online learning environments has become a pressing concern. The significance of this study lies in its focus on online English language learning—a domain that demands both linguistic and technological competencies—making it a valuable inquiry for the field of instructional design and technology. LA provides a framework for analyzing student behaviors and outcomes, offering stakeholders empirical insights into key areas of engagement and academic performance. While early contributions, such as Romero and Ventura (2007), recognized the potential of data mining in both traditional and distance education, contemporary research has expanded this foundation, demonstrating how LA can inform interventions, enhance effective course design, and support at-risk students through early warning systems and predictive analytics (Macfadyen & Dawson, 2010; Siemens, 2013; Slade & Prinsloo, 2013; Matcha et al., 2020).

The urgency of this research has been amplified by the global disruption caused by the COVID-19 pandemic and other natural disasters which exposed both the possibilities and limitations of online instruction. During these transitions, LA has emerged as a critical tool for identifying key factors influencing access to resources and student performance, offering real-time insights that enable educators to tailor instruction and feedback, as well as plan for the future (Ifenthaler & Yau, 2020; AlTwijri & Alghizzi, 2024; Gašević et al., 2017; Viberg et al., 2018). This is especially crucial in English language learning, where progress is often uneven across skills like grammar, vocabulary, and speaking. Therefore, this study not only contributes to the evidence base by predicting student performance in online English language learning through the lens of LA in higher education, but also offers insights and strategies for institutions aiming to reduce failure rates, enhance English language proficiency in schools of languages, and promote equity in online English instruction across various academic disciplines.

1.5 Definition of Terms

Distance Education: In this research study, distance education is broadly defined but specifically refers to online distance education, where learners and instructors interact remotely through digital platforms.

Engagement: Level of interest and involvement in learning process but it is mostly referred to behavioral engagement in this study.

Language Acquisition: A process in which students learn and develop the ability to understand, produce, and use language, and it is referred to second language acquisition in this research study.

Learning Analytics: It is the collection, measurement, analysis, and reporting of data about learners and their contexts, with the goal of understanding and improving learning outcomes and environments.

Learning Management System (LMS): A software platform designed to support educational administration, delivery, and tracking, providing tools for instructors and learners to manage and engage with digital learning resources.

Modular Object-Oriented Dynamic Learning Environment (MOODLE): An open-source LMS utilized globally, supporting educational needs from K-12 to higher education with features for managing online courses, collaborative learning, and assessments.

Open Learning/Education: Educational practices that emphasize accessibility, flexibility, and inclusivity, often removing barriers such as cost, prerequisites, or location.

Orange Data Mining: A comprehensive open-source data mining software that features intuitive drag-and-drop interfaces and customizable add-on functionalities, allowing users to execute complex data analysis tasks with minimal programming effort.

CHAPTER 2

LITERATURE REVIEW

The literature review was conducted to establish a conceptual framework for the study and to address research gaps in the existing literature on online English language learning and LA at the university level. The review was carefully designed to provide insights into the intersection of language learning and data-driven educational practices, emphasizing the use of technology to enhance student learning experiences.

The search was conducted using specific keywords such as LA, EDM, LAD, online language learning, LA in online language learning, distance language learning, factors influencing online language learning, data mining in education, and data mining in DE. These keywords were selected to ensure a comprehensive exploration of relevant studies, highlighting both theoretical contributions and empirical findings. The systematic search aimed to identify studies that provide a foundation for the proposed research and uncover gaps in understanding the application of LA in language learning contexts.

The literature review consists of the following sections: Learning Analytics in Contemporary Education, Benefits and Challenges in Learning Analytics, Learning Analytics in Online English Language Learning, Factors Affecting Student Performance in Online English Language Learning, Exploring Students' Engagement Types, Relevance to the Research Study, and Ethical Considerations and Data Privacy in Learning Analytics and Conclusion.

2.1 Learning Analytics in Contemporary Education

Enhancing instruction, fostering learning and student success, and supporting data-driven decision-making in educational policies are fundamental objectives of modern education. LA has emerged as a pivotal field in achieving these objectives by leveraging the power of educational data to optimize learning experiences, personalize instruction, and improve institutional outcomes for diverse domains and stakeholders. The rise of e-learning environments, including open and online learning, has accelerated advancements in LA, driving innovative solutions and expanding their application across various educational contexts (Dutt, Ismail, & Herawan, 2016). In this regard, LA is becoming indispensable for understanding and improving the learning process (Ifenthaler, Mah, & Yau, 2019; Siemens & Long, 2011).

Contemporary LA research studies continue to emphasize prediction, personalization, and intervention. For example, predictive models trained on LMS data are widely used to identify at-risk students and trigger timely interventions that support students' retention, performance, success rates (You, 2016; Slade & Prinsloo, 2013; Delen, 2011). Dashboards and visual analytics tools provide instructors and managers with real-time feedback on student engagement, helping them tailor their instructional strategies (Viberg, Hatakka, Bälter, & Mavroudi, 2018). Similarly, students can benefit from personalized feedback and progress tracking, increasing their self-regulation and motivation (Matcha et al., 2020; Romero & Ventura, 2017). In other words, feedback mechanisms facilitated by LA empower students to monitor their progress and make informed decisions, while real-time insights from LA dashboards help instructors adapt teaching methods and techniques to address evolving student needs (Gašević, Dawson, & Siemens, 2015).

The global shift to open and online learning, accelerated by the COVID-19 pandemic, has underscored the critical role of LA in navigating complex educational landscapes. Institutions increasingly rely on LA to inform resource allocation, identify systemic inequities, and scale personalized learning initiatives (Ifenthaler, Mah, & Yau, 2019; Bienkowski, Feng, & Means, 2012). For instance, metrics and clickstream from LMS platforms can reveal engagement trends, predict performance outcomes and offer actionable insights into student behaviors (Macfadyen & Dawson, 2010; Viberg et al., 2018). These developments highlight the growing potential of LA to transform education, addressing the dynamic and evolving needs of both learners and educators.

2.2 Benefits and Challenges in Learning Analytics

With advancements in computer and internet technologies, LA has transformed educational practices, by offering a range of benefits but also introducing significant challenges. As LA becomes more integral to educational settings, particularly in blended and online learning environments, institutions must weigh the potential benefits, such as data-driven decision-making and personalized learning, against challenges like data privacy, ethical concerns, and the need for proper infrastructure and training.

2.2.1 Benefits

LA offers substantial benefits to stakeholders in education, particularly in higher education, where online learning has become increasingly prevalent among adult learners. Siemens and Baker (2012) emphasize that data-driven decision-making has introduced a transformative shift in education, enabling evidence-based planning, interventions, and decision-making. Utilizing learning analytics, educational institutions can enhance both learning experiences and teaching practices, ultimately improving the overall experience for students, instructors and administrators (Ifenthaler & Widanapathirana, 2014).

A significant benefit of LA is its ability to identify at-risk students, enabling institutions to design targeted interventions that promote academic success and retention (Siemens & Long, 2011; Tempelaar, Rienties, & Nguyen, 2021). Through timely alerts and personalized feedback, students can make informed choices about their learning paths, fostering autonomy and creating a tailored learning environment (Sclater, Peasgood, & Mullan, 2016). This predictive capability allows educators to adopt differentiated approaches based on student needs, which can lead to enhanced academic achievement and improved support systems.

Moreover, LA can drive the development of personalized learning pathways, supporting students in setting and achieving their academic goals. By tracking performance data and learning paths, educators can create environments that support adaptive learning, increased engagement, and optimal elements for successful online learning experiences and environments (Ifenthaler & Widanapathirana, 2014; Van Leeuwen, 2019). Thus, LA not only enables more comprehensive and inclusive educational approaches but also promotes a data-driven culture where stakeholders can collaboratively enhance the quality of education and maintain optimal online learning environments.

2.2.2 Challenges

Despite its advantages, LA faces significant challenges, particularly concerning ethical considerations, resource demands, technical proficiency and model reliability. Ethical issues are particularly critical, as educational data is sensitive, and developing clear, enforceable guidelines for its use remains complex across diverse educational contexts (Ferguson, 2012). Concerns about data privacy, consent and potential misuse are persistent in LA (Drachsler & Greller, 2016).

The resource demands of LA are substantial, requiring robust IT infrastructure like servers, high-capacity storage, and powerful computing capabilities (Bienkowski, Feng, & Means, 2012). Although some free online tools can assist with data preprocessing and analysis, these resources may be insufficient for large-scale data, limiting the scope of analysis for researchers without access to advanced infrastructure.

Another challenge arises from the diverse backgrounds of education researchers, many of whom may lack technical training and knowledge in data mining and analytics (Romero & Ventura, 2010). Gaining proficiency in data preprocessing, using analytical tools, and interpreting results can be particularly demanding for educators without technical expertise.

Data accessibility also restricts LA, as IT departments in educational institutions often act as gatekeepers, concerned about data privacy and the security implications of external analysis (Zilvinskis, Willis III, & Borden, 2017). This gatekeeping can limit researchers' access to crucial data, impacting the comprehensiveness and reliability of their analyses.

Ensuring accuracy in predictive models is also challenging. False positives and negatives can lead to type I and type II errors, respectively, affecting the accuracy and trustworthiness of LA insights (Archer & Prinsloo, 2020). Selecting appropriate models and adjusting for acceptable error margins are vital steps, yet achieving a perfect balance remains difficult.

Lastly, predictive analytics, while adept at identifying trends, cannot establish causation. Correlations can suggest patterns but require further research studies, such as experimental or qualitative research studies, to uncover underlying causes (Archer & Prinsloo, 2020; Strang, 2017). Nonetheless, each challenge offers an opportunity for refinement and responsible evolution of LA practices.

2.3 Learning Analytics in Online English Language Learning

Online English language learning has experienced significant growth, driven by technological advancements and the increasing demand for accessible, flexible educational solutions. As a global lingua franca, English plays a vital role in academic, professional, and social contexts, leading to the proliferation of online platforms that cater to diverse linguistic and cultural backgrounds (Crystal, 2003). Online English learning platforms offer distinct advantages, including convenience, accessibility, and flexibility, enabling learners to study at their own pace, access resources anytime, and engage in self-directed learning (Hampel & Stickler, 2015; Hendijani, 2025; Huda, Janattaka, & Prayoga, 2023). This flexibility is particularly advantageous for older adult learners, who often deal with obligations of work, family, and other responsibilities, making traditional classes challenging to attend (Berry & Hudges, 2020).

An education system is the most crucial consideration in LA; other parameters, such as demographic features, can also be added to data to be analyzed to reach a more holistic understanding of the phenomenon (Romero & Ventura, 2007; Chu et al., 2022). Learning platforms integrate diverse materials—videos, interactive multimedia, audio resources, and external links—that stimulate language acquisition through multiple sensory channels (Mayer, 2009; Li & Lan, 2022; Mayer, 2024). The appropriate incorporation of multimedia features fosters comprehension and engagement while it also supports LA interventions and tailored educational plans.

LMSs facilitate self-paced progression and access to resources aligned with individual learner goals besides LA opportunities (Plass & Jones, 2005; Ifenthaler, Mah, & Yau, 2019). LA, through data-driven insights, can recommend resources and interventions necessary for achieving milestones such as proficiency and gateway exams. Moreover, the integration of visual and auditory stimuli has been proven to enhance second-language acquisition (Plass & Jones, 2005; Mayer, 2024), supporting intervention designs aimed at improved student outcomes. In this regard, tracking students' login activity in language learning can help instructors understand their behaviors and identify potential opportunities for instructional improvement (Thomas, Reinders, & Gelan, 2017).

Tools like MOODLE integrated with virtual classrooms, support both asynchronous and synchronous interactions, offering students real-time or on-demand access to instructional content. Asynchronous communication fosters reflective learning, enabling students to process language constructs, grammar, and vocabulary at their own pace—especially beneficial for beginners (Takase, 2024; Hrastinski, 2008).

Asynchronous learning channels in English learning can be especially beneficial when components are appropriately designed (Alfares, 2024). Conversely, synchronous platforms promote both planned and spontaneous interactions, crucial for developing fluency and articulation, mirroring real-life communication scenarios (Blake, 2009; Hrastinski, 2008). In this regard, LA encompasses features that can be seamlessly applied to the research field through the use of computational techniques, taking into account the various conditions and factors that shape the learning experience, including skill development in online English language learning, such as grammar, vocabulary, and speaking. A combination of synchronous and asynchronous communication powered by LA offers an equilibrium, allowing learners to benefit from both domains: reflective and interactive language practices.

Additionally, English language acquisition in online learning platforms can utilize adaptive learning technologies so that the learning experience can be customized to student progress and performance (Yaseen & Alnakeeb, 2023). While this is more common in commercial language learning platforms, it is quite less common in official settings or at language schools of universities. Nonetheless, adaptive learning environments can be created to respond to the needs of students at different levels and skills-based interventions such as speaking and writing thanks to LA. Adaptive learning systems can provide the option of adjusting the level of difficulty, content, and pace of learning material based on each student's needs, thus, promoting a tailored approach to language acquisition (Tseng et al., 2008; Yaseen & Alnakeeb, 2023). For example, a student struggling with grammar might receive additional exercises focused on grammatical structures, while a more advanced student might be introduced to complex vocabulary and reading comprehension tasks that are customized to their contexts. These systems use data on student interaction and performance to identify areas of improvement, which is particularly valuable in language learning, where progression can vary widely across skill sets such as vocabulary, grammar, pronunciation, listening, reading and writing (Brusilovsky & Millán, 2007; Chen et al., 2023). Research suggests that personalized, adaptive learning enhances engagement and motivation, as students feel that the learning path is relevant to their unique contexts, strengths and struggles (Durlach & Lesgold, 2012; Simon & Zeng, 2024).

Language learning is inherently social and often benefits from face-to-face interaction, which can be difficult to replicate online (Byrnes, 2006). While virtual classrooms and forums provide some interaction opportunities, they often struggle to replicate the immediacy and sense of community found in physical classrooms. Furthermore, maintaining motivation in self-directed online environments can be challenging for students, especially in language learning, which requires consistent practice and exposure over time (Ushioda, 2011; Heo & Han, 2022). Effective online English language learning programs often address these challenges by incorporating gamification elements, setting goals, and utilizing LA to monitor engagement and progress, all of which help sustain learner interest and motivation contributing to improved psychological states, such as increased motivation and behavioral engagement (Domínguez et al., 2013; Liu, Wang, & Wang, 2025). In this regard, LA can facilitate and plan interventions related to gamification in online English language learning.

In summary, online English language learning combines flexibility, multimedia resources, adaptive learning, and varied interaction modes to offer a comprehensive approach to language acquisition. However, it requires a thoughtful design to address the unique challenges of engagement, community-building, and motivation. As technology continues to evolve, online English language learning platforms will likely incorporate more advanced analytics, artificial intelligence, and virtual reality to further enhance language acquisition and create more immersive, supportive environments for learners in all levels worldwide.

2.4 Factors Affecting Student Performance in Online English Language Learning

While learning English language through computer and internet technologies has become common practice, the literature has identified various factors that influence students' academic performance in online English language learning. These factors can vary widely, depending on individual and contextual circumstances. One of the most critical factors is self-regulation, which plays a pivotal role in shaping student performance. Online and less formal learning environments demand that students regulate their learning. Research consistently demonstrates that self-regulated learners outperform others in academic contexts (Zimmerman, 2002). Furthermore, self-regulation strongly influences student engagement levels, necessitating measures to enhance this skill among learners (Sun & Rueda, 2012). Through the use of LA, deficiencies in self-regulation can be identified based on LMS usage details and patterns, allowing for targeted interventions to enhance students' academic achievement in online English language learning.

Technological and instructional dimensions must also be addressed when designing online English learning platforms. English language learning is one of the most popular topics on online learning platforms, necessitating a focus on infrastructure, the capacity of service providers, and the perspectives of students and instructors. In this regard, incorporating these perspectives is essential to address key factors, such as self-paced learning environments and the integration of multimedia resources, which have been shown to enhance learner engagement and satisfaction (Chiu, Liang, & Tsai, 2013). A well-designed LMS with ease-of-use features and clear instructional goals can improve the overall learning experience while also providing opportunities for LA to develop effective intervention strategies.

Environmental factors also play a critical role in online English language learning. Students' socio-economic backgrounds often dictate their access to technological equipment and stable internet connections, limiting their ability to participate, especially in synchronous learning activities (Hasan & Bao, 2020). That's to say, this can hinder communication with instructors and peers, particularly in virtual classrooms, and negatively affect skills like speaking in English language learning acquisition. Research studies also show that students; for example, Turkish students, who experience high levels of anxiety resulting from their instructors' attitude, may have difficulty, especially when practicing speaking skills, whether in face-to-face or online English language learning environments (Aydin, 1999 as cited in Horwitz, 2001). These conditions especially apply when learning and teaching are conducted online under extraordinary circumstances rather than in normal contexts typical in online adult learning environments.

Effective course design is another major factor influencing student performance in online English language learning. Within LMS platforms, course design is a critical determinant of student engagement, satisfaction, and achievement, all of which affect students' over experience in learning process. Better course designs indicate better engagement, satisfaction and achievement in general. Within the synchronous part of LMS or discussion forums, the presence of an active instructor may increase cognitive presence, especially through constructive feedback; hence, facilitating student engagement and achievement in online English language learning (Garrison, Anderson, & Archer, 1999). In this regard, addressing both pedagogical and technological elements is essential for fostering student engagement and success. However, these considerations also underscore the complexity of factors affecting student performance in online English language learning.

2.5 Exploring Students' Engagement Types

The COVID-19 pandemic significantly accelerated the adoption of LMSs to address the technological and infrastructural challenges of education. This pivotal integration demonstrated the central role of technology in enhancing student engagement within online English language learning environments. LMS platforms not only provide structured environments where students can access course materials but also actively engage with peers and instructors. Behavioral engagement through LMS usage has been demonstrated through metrics such as course views, unique login days, logins, assignment submissions, and participation in virtual classrooms (e.g., attendance and interaction). These metrics highlight students' frequency of accessing course content and their level of participation in virtual learning spaces. Macfadyen and Dawson (2010) demonstrated that LMS activity metrics, including logins and resource downloads, act as proxies for student engagement and are correlated with academic performance. During extraordinary and challenging circumstances, analyzing LMS metrics can provide essential insights into how students adapt to the online learning paradigm, thereby motivate further and future research into engagement metrics.

The integration of LMS tools with virtual classrooms to enhance emotional engagement proved crucial in addressing the psychological challenges of remote learning. This was particularly significant in online English language instruction during the pandemic. Features such as discussion forums, real-time chat, virtual conferences, and personalized feedback fostered a sense of community comparable to that of face-to-face learning settings and sustained student motivation. Garrison et al. (1999) emphasized the importance of social presence in LMS-based environments, noting that meaningful and timely interactions substantially enhance emotional engagement. For instance, the availability of recorded sessions and consistent virtual classroom attendance enabled students to revisit materials at their own pace, which reduced anxiety and supported their emotional well-being.

Cognitive engagement, a deeper and more critical form of interaction, is notably enhanced through the use of LMS-integrated multimedia tools and skill-specific resources. Metrics such as H5P usage for grammar, listening, and speaking exercises highlight the effectiveness of interactive tools in promoting active learning and supporting targeted skill development in online English language learning. These tools align with findings by Liaw, Huang, and Chen (2007), who argued that positive attitudes towards LMS platforms play a pivotal role in fostering cognitive engagement, active participation, and higher-order thinking. Furthermore, adaptive technologies integrated within LMS frameworks enable personalized learning experiences, effectively addressing individual student needs and skill levels (Kuh et al., 2006).

LMS data analysis during the COVID-19 pandemic also revealed significant inequities in engagement, particularly among students facing technological or socio-economic barriers. The digital divide, evident in metrics such as inconsistent login patterns and low virtual classroom attendance, highlights the necessity of equitable access to technology and reliable internet connectivity (Van Dijk, 2017). Students with limited access to LMS platforms struggled to engage effectively, impairing their language acquisition and preparation for key proficiency exams. This finding aligns with Sun and Rueda (2012), who noted that situational challenges in online learning disproportionately impact behavioral, emotional, and cognitive dimensions of student engagement.

2.6 Relevance to the Research Study

Students' academic performance details and engagement through metrics were captured across four quarters in online English learning during the COVID-19 pandemic. Each quarter's data encapsulates vital metrics, including assignment submissions, virtual classroom participation, resource utilization, and interaction with course materials, which are critical in understanding how students navigate their learning environments. Such data-driven approaches align with the increasing emphasis on leveraging LA to guide educational interventions (Herodotou et al., 2019; Klein et al., 2019).

2.6.1 Demographic Influence in Online English Language Learning

Demographic factors such as age, socioeconomic status, cultural background, geographic location, and prior education play a significant role in shaping students' outcomes and experiences in online English language learning environments. These factors impact access to resources, engagement levels, and overall learning effectiveness. Understanding these influences is critical for designing inclusive and effective online learning platforms that cater to diverse student needs.

Age is widely regarded as a significant factor influencing learners' familiarity and ease with technology. Younger learners are often assumed to be more proficient in navigating online platforms due to their early exposure to digital tools (Prensky, 2001). However, existing literature presents contradictory findings regarding the extent to which age predicts achievement in English language learning, particularly when accounting for its complexity and various dimensions across different age groups, from early childhood to adulthood (Dörnyei, 2014).

Socioeconomic status is another significant factor influencing online English language learning. Students from low-income households may struggle with limited access to reliable internet connections, digital devices, or a conducive learning environment during the COVID-19 pandemic (Rofiah, Sha'ar, & Waluyo, 2022). This digital divide can result in disparities in participation and achievement in online English learning, similar to those observed in other domains. Providing affordable or subsidized resources and ensuring mobile-friendly platforms can help bridge these gaps. Geographic location further amplifies disparities in online learning experiences. Students in rural or remote areas often face inconsistent internet connection and fewer opportunities for real-time interaction with peers or instructors (Greenhow, Lewin, & Willet, 2020).

Gender may also shape students' approaches to online English language learning (Viriya & Sapsirin, 2014). Research has shown that slight differences in motivation, learning strategies, and specific contexts contribute to distinct patterns of language acquisition between male and female students. Studies suggest that while gender may influence learning strategies, it is not a decisive factor in determining learning outcomes (Dörnyei, 2014; Oxford, 2017). Similarly, some studies suggest that, similar to age, gender has little to no influence on learning outcomes in specific language skills and contexts (Ellis, 2015).

Finally, prior education and proficiency levels play a critical role in shaping students' readiness for online English learning. Learners with a strong academic background, prior exposure to English, or other influential factors are generally better prepared to succeed in English as a Foreign Language (EFL) environments (Dörnyei, 2014). In this regard, complete beginners may require additional support through tailored materials, interactive tools, and one-on-one assistance to effectively develop their foundational language skills.

2.6.2 Quarterly Academic Engagement and Performance

Quarterly analysis of student activity can provide a granular and structured analysis of students' academic performance and engagement across four quarters, offering a dynamic perspective on their learning behaviors. By tracking key metrics such as assignment submissions, virtual classroom participation, resource utilization, and interaction with course materials, the dataset enables data-driven insights into student progress and instructional effectiveness. LA frameworks have demonstrated that continuous monitoring of engagement metrics can serve as a predictive tool for student success and retention (Mah, 2016; Shafiq et al., 2022).

Sliced data can highlight patterns of engagement, revealing trends in learning consistency, participation levels, and resource dependency. For instance, fluctuations in virtual classroom attendance and assignment completion rates can serve as early indicators of academic challenges, allowing educators to implement timely interventions (Foung & Chen, 2019). Research suggests that students with inconsistent engagement patterns in digital learning environments often struggle to maintain long-term academic performance and may require interventions, such as LADs, to support their progress (Valle et al., 2021). Similarly, patterns in resource utilization, such as frequent engagement with grammar and vocabulary exercises that enhance writing skills, provide insights into students' specific language learning needs and, when integrated with real-time feedback mechanisms, support educators in designing adaptive learning strategies (Kim et al., 2024; Zhang et al., 2018; Tlili et al., 2021).

By leveraging LA, it is possible to facilitate a proactive approach to instruction, enabling educators to personalize teaching strategies and enhance student outcomes. The integration of quarterly engagement and performance metrics supports longitudinal analysis, allowing institutions to refine curriculum design, identify at-risk students, and optimize pedagogical approaches for sustained learning success (Kieffer, 2012). Furthermore, data-driven curriculum adjustments have been shown to improve learning efficacy by aligning instructional content with student engagement trends (Lin & Hwang, 2018).

The inclusion of quarterly data can enable a longitudinal analysis of student behavior and academic performance, offering insights into learning progression and engagement patterns over time. By tracking key indicators, instructors can identify trends in student engagement, assessing whether participation remains stable, improves, or declines over the course of an academic year. For instance, a decline in login rates may signal disengagement, potentially due to academic difficulties, motivational challenges, or external factors. Conversely, consistent or increasing login frequencies may indicate sustained commitment and active participation in the learning process (Smith, Lange, & Huston, 2012). Research suggests that students with higher engagement metrics in online learning environments tend to achieve better academic outcomes, reinforcing the importance of tracking log-in patterns as an early intervention tool (Moubayed et al., 2018).

Moreover, the transition from one quarter to another provides a clear trajectory of students' academic progress. This longitudinal performance tracking allows institutions to assess whether initial coursework challenges were successfully overcome or if persistent difficulties require targeted support, ensuring the lowest possible failure rate in the English proficiency exam at the end of the academic year. Studies in LA-based performance monitoring highlight the significance of tracking early failures to predict final course outcomes, making such transitions a valuable indicator of student success (Lu et al., 2018).

By leveraging quarterly engagement and performance data, instructors and institutions, with the support of ICT professionals, can develop early-warning systems to identify students at risk of academic decline while reinforcing positive learning behaviors. The ability to track login patterns, performance trends, and pass/fail status over time empowers instructors to implement timely interventions and adaptive learning strategies, ensuring sustained student achievement in online English learning environments. In conclusion, the structured nature of quarterly data not only enhances instructors' ability to provide targeted support but also contributes to institutional decision-making for long-term academic improvements.

2.6.3 Engagement in Virtual Learning Environments

Key engagement indicators, such as virtual classroom attendance reflect students' participation in both synchronous and asynchronous virtual classroom activities. Research has consistently shown that active engagement in virtual learning environments (VLEs) positively influences academic performance and retention rates. In their 2021 study, Suharti, Suherdi, and Setyarini found that utilizing online platforms such as Zoom and Google Classroom in vocational schools enhanced EFL students' behavioral, emotional, and cognitive engagement, which is crucial in the absence of physical classrooms (Consoli & Curle, 2024).

Furthermore, frequent participation in live sessions, combined with extended session durations, indicates active engagement in real-time discussions, which may correlate with an enhanced learning experience and improved performance metrics (Varouchas, Sicilia, & Sánchez-Alonso, 2018). In other words, students who frequently attend live sessions may benefit from real-time feedback, peer interactions, and direct instructor support, all of which contribute to improved comprehension, satisfaction, and academic success. Similarly, the use of asynchronous resources emphasizes students' reliance on self-paced learning strategies and potential LA dashboard designs (Brun et al., 2019; Mirriahi & Dawson, 2013). This aligns with findings that students engaging in asynchronous learning develop higher self-regulation skills, enabling them to review complex materials at their own pace and reinforce learning outcomes, similar to how students using LA tools and dashboards enhance their learning through self-monitoring and strategic adjustments (Matcha et al., 2020). The ability to revisit lectures and course content asynchronously offers significant benefits in language learning environments, providing flexibility while enhancing comprehension and retention through repeated exposure to spoken and written material (Gorjian et al., 2011; Chafouk & Marjanei, 2024).

Tracking engagement across synchronous and asynchronous modalities in virtual learning environments allows instructors to identify students' preferred learning strategies and detect areas needing intervention. For example, a decline in live session participation alongside an increase in recorded lecture downloads may suggest schedule conflicts or shifts in students' learning preferences. Analyzing virtual engagement metrics provides valuable insights into students' learning behaviors, participation trends, and academic success predictors in online English language learning. By combining real-time participation indicators with asynchronous LA, instructors can optimize virtual instruction to accommodate diverse learning needs and enhance overall student outcomes.

2.6.4 Skill-Specific Interactions

Tracking skill-specific interactions can enable a data-driven understanding of students' focus areas in English language learning. These interactions can provide valuable insights into students' self-directed, skill-based learning behaviors in English language acquisition, enabling educators to tailor instructional strategies to individual needs when an intervention plan or a LAD is available.

The segmentation of skill interactions allows for precise analysis of students' focus areas and learning gaps. For instance, despite frequent engagement with writing-related materials and proficiency in writing, the student may struggle to speak English effectively at an adequate level. In such cases, targeted interventions in LA, such as speaking workshops, peer reviews, and personalized instructor feedback, can help bridge the gap between engagement and performance by providing students with necessary feedback on relevant English language skills (Klein et al., 2019; Lin & Hwang, 2018; Yin & Hwang, 2018).

Furthermore, tracking skill-specific interactions longitudinally can reveal progress trends and evolving learning preferences. For example, a student who initially focuses on grammar resources—a common approach, especially in the first and second quarters of the academic year— but later shifts towards writing and speaking exercises may indicate a transition from foundational skill-building to more advanced language application. Understanding these patterns can allow instructors to align instructional materials with students' evolving learning trajectories and provide adaptive learning pathways for continuous English language skill development (Lin & Hwang, 2018; Yin & Hwang, 2018).

Additionally, research suggests that students who engage more frequently with interactive and multimodal resources—such as listening exercises, speaking practice modules, and real-world language simulations—tend to achieve higher overall proficiency in second-language learning (Tlili et al., 2021). This highlights the importance of tracking not just raw engagement numbers but also the diversity of resources used, ensuring that students receive balanced exposure to all language skills and effectively prepare for English proficiency exams.

In conclusion, skill-specific tracking capabilities offer educational professionals deep insights into student learning behaviors, strengths, and weaknesses. By analyzing engagement alongside achievement outcomes, educators can implement targeted interventions, optimize course content across different EFL levels, and enhance student learning experiences in second language acquisition. This data-driven approach to language LA ensures that instructional strategies are personalized, effective, and aligned with students' evolving needs, particularly during challenging times to achieve optimal learning outcomes.

2.7 Ethical Considerations and Data Privacy in Learning Analytics

The integration of LA in education, more specifically in online English language learning can raise ethical and data privacy concerns that require careful attention to avoid harm to students, instructors, or institutions. LA heavily depends on collecting, storing, and analyzing vast amounts of student data to facilitate potential interventions. Siemens (2013) highlights that using educational data may create ethical and privacy problems if not utilized properly while this is of particular importance for LA as it deals with increasingly high detailed volumes of educational data to improve learning and teaching. Without clear ethical guidelines and inattention, there is a persistent risk of misusing educational data, creating inequities, or damaging institutions (Slade & Prinsloo, 2013).

Data privacy remains one of the central concerns in LA. The use of personal data must comply with Türkiye's law number 6698, known as KVKK (Kişisel Verilerin Korunması Kanunu – Personal Data Protection Law), which mandates responsibilities for institutions and researchers in the collection, analysis, and application of all kinds of data, including educational ones. Bienkowski, Feng, and Means (2012) argue that while combining institutional data, if distributed, can improve prediction precision and intervention strategies, such practices must ensure contextual understanding and adherence to privacy standards.

Moreover, Pardo and Siemens (2014) highlight that institutions are often trusted by students; nonetheless, they can face challenges while balancing the benefits of LA without compromising student privacy and securing the use of educational data with stakeholders. To illustrate, the use of predictive analytics to identify at-risk students can lead to bias among students or instructors so it may be necessary to plan applications of LA in education settings carefully, perhaps after investigating the learning environment and educational data for a few years in sequence.

Another ethical consideration in LA is the potential for algorithmic bias. Algorithms used in LA are only as unbiased as the train data and features on which they are run. Due to legal limitations, it might be necessary to eliminate demographic components of data and this may lead to further misrepresentation of students in LA, and it can result in inaccurate predictions to some extent (Slade & Prinsloo, 2013). Nonetheless, it can be required to eliminate some dimensions of data that cannot be intervened while LA seeks opportunities for possible interventions.

To address ethical and privacy challenges, institutions must establish clear policies and frameworks aligned with ethical principles and privacy laws for the ethical use of LA. Techniques such as data anonymization and regular audits of algorithms are critical the initial phase of LA for protecting personal information while maintaining data utility (Slade & Prinsloo, 2013). As noted by Kay et al., (2012), institutions continually need to balance risks and awards in the context of LA regarding their responsibilities for using data in view of corporate regulations and national laws. In this regard, it can be asserted that transparency, stakeholder collaboration, and inclusivity in data interpretation are imperative to minimize unethical practices and uphold ethical standards (Rubel & Jones, 2016; West, Huijser, & Heath, 2016).

2.8 Conclusion

This comprehensive exploration of online English language learning and the application of LA within a university context underscores the dynamic interplay between technology and pedagogy in optimizing educational experiences. The literature demonstrates how LA enhances instructional design, fosters student engagement, and addresses persistent challenges in online English learning environments.

LA supports a data-driven approach by offering predictive capabilities that enable institutions to identify at-risk students and intervene early through tailored support mechanisms. Early warning systems and predictive models—often built on LMS data—enable timely responses to declining engagement or academic performance, enhancing retention and learning outcomes. These systems can empower instructors to deliver personalized feedback, adjust course content, and implement targeted interventions that align with individual student needs.

The review also highlights the importance of aligning LA practices with pedagogical goals, ensuring ethical data use, and integrating behavioral and cognitive insights into instructional design. While challenges such as technical infrastructure, algorithmic limitations, and data privacy concerns remain, successful implementation depends on stakeholder collaboration, capacity building, and iterative refinement of LA practices.

In summary, LA holds transformative potential for online English language learning by enabling adaptive, responsive, and student-centered education. Future research should continue to expand the use of predictive analytics, refine early warning mechanisms, and ensure ethical, inclusive practices that support sustainable growth in this evolving field.

CHAPTER 3

METHODOLOGY

This chapter comprises Research Design, Sampling and Participants, Data Sources and Instruments, Data Collection Procedure, Data Analysis, Ethical Issues, Researcher’s Background, and Delimitation and Limitation of the Study.

3.1 Context and Data

In this study, only students who began learning English as complete beginners, starting at level one, were included. The research focused on their progress during the fall and spring terms, which were divided into two quarters per term, resulting in a total of four quarters.

The LMS utilized by the School of Foreign Languages was pivotal in facilitating the online learning experience. Figure 3.1 presents a snapshot of the general course interface, offering an overview of how course materials and interactions were structured. Figure 3.2 highlights the weekly course interface, showcasing the detailed organization of activities and lessons for each week, allowing students to track their progress systematically. Lastly, Figure 3.3 provides an overview of the level-based material interface, demonstrating how the LMS supported differentiated learning paths for students at various proficiency levels. These features collectively illustrate the LMS's role in creating a structured and accessible learning environment during an unprecedented time.

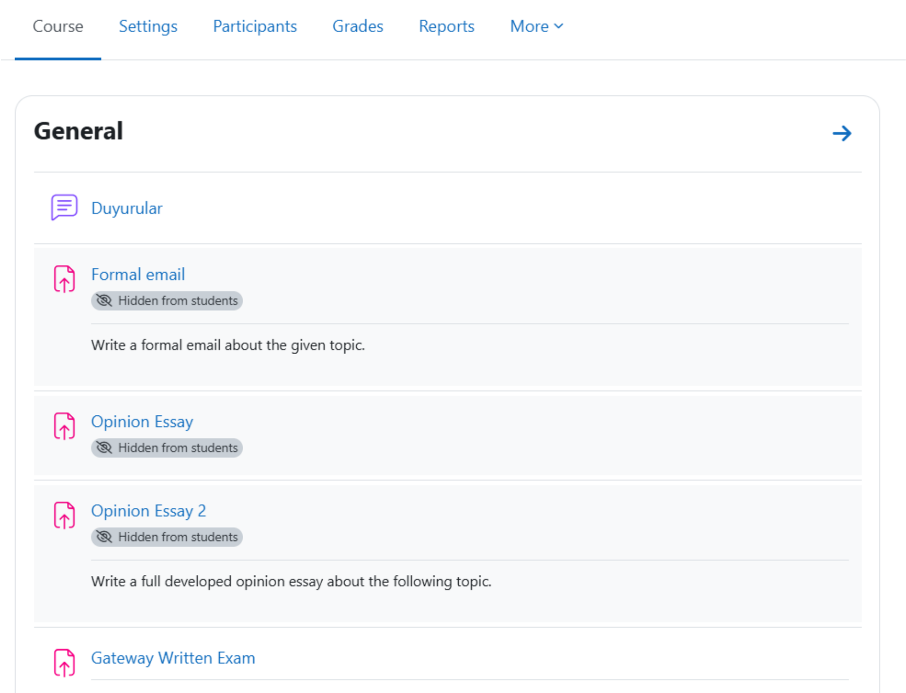

Figure 3.1. A View of the Course Interface in the LMS Utilized by the School of Foreign Languages

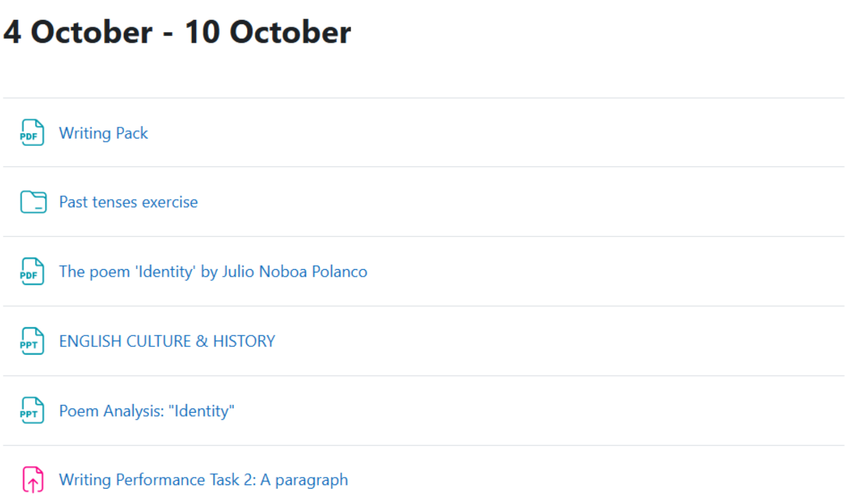

Figure 3.2. A View of the Weekly Course Interface in the LMS Used by the School of Foreign Languages

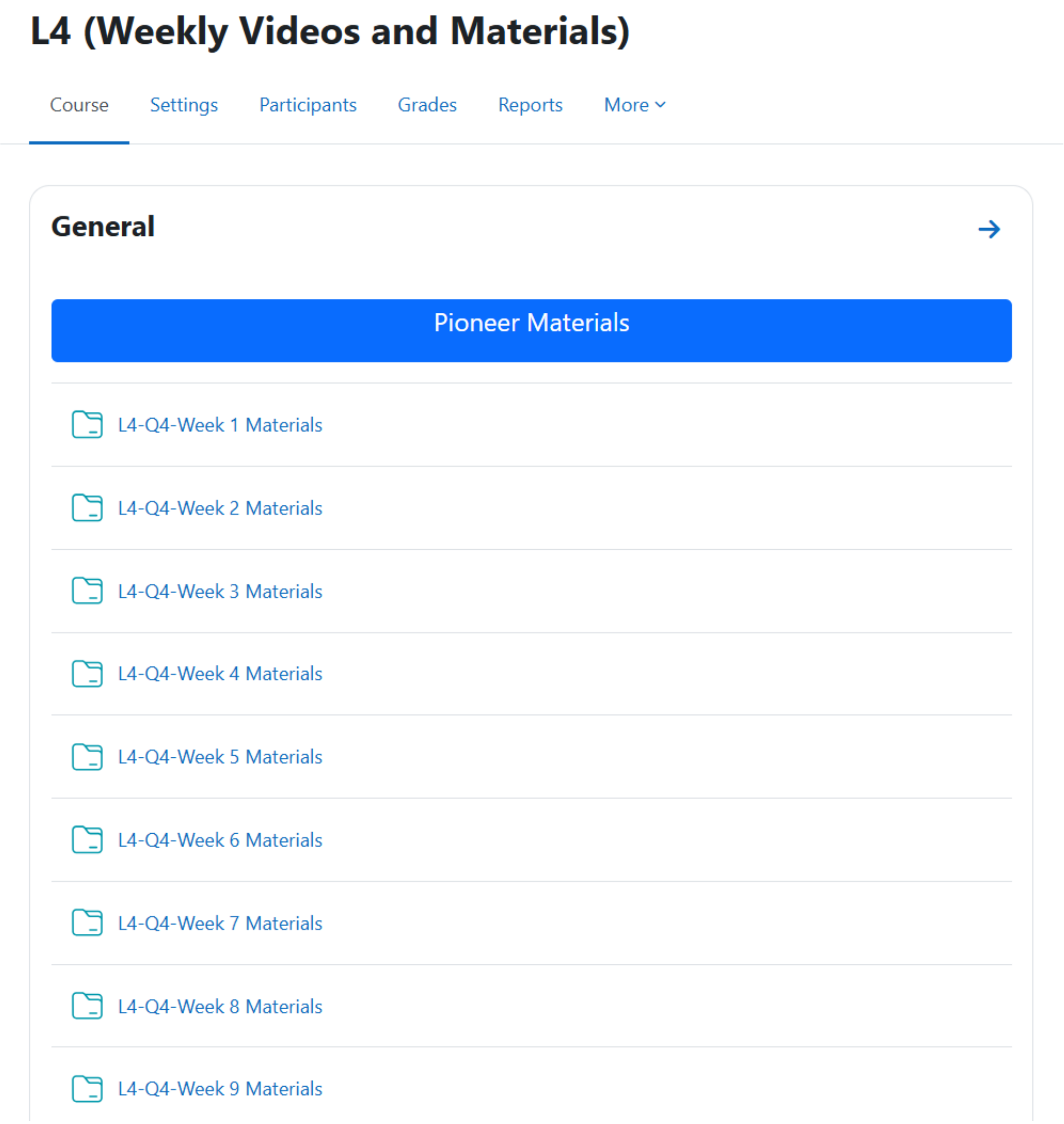

Figure 3.3. An Overview of the Level-Based Material Interface in the LMS Utilized by the School of Foreign Languages

Table 3.1. Feature List and Descriptions

Feature	Description
Faculty	The faculty where the student is registered.
FacultySchool	Indicates whether the student is registered at a faculty or vocational school.
Department	The department where the student is enrolled.
LanguageOfInstruction	The primary language of instruction in the designated department.
TypeOfInstruction	Indicates whether the designated department provides distance or face-to-face instruction.
Gender	The gender of the student (e.g., male, female).
Age	The age of the student at the time of the study.
Q1_Assignment_Submissions	The number of assignment submissions in the first quarter.
Q1_VClass_Rec_Downloads	Number of virtual classroom recording downloads in the first quarter.
Q1_VClass_Joins	Number of virtual classroom joins in the first quarter.
Q1_CourseViews_UniqueDays	Number of unique course views days in the first quarter.
Q1_FolderDownloads_UniqueDays	Number of unique folder download days in the first quarter.
Q1_UniqueFolderDownloads	Number of unique folder downloads in the first quarter.

Table 3.1. (continued)

Q1_FolderViews_UniqueDays	Number of unique folder view days in the first quarter.
Q1_UniqueFolderViews	Number of unique folder views in the first quarter.
Q1_URL_Interaction_UniqueDays	Number of unique URL interactions days in the first quarter.
Q1_URL_NonUniqueInteractions	Number of non-unique URL interactions in the first quarter.
Q1_LoginDays	Number of login days in the first quarter.
Q1_TotalLogins	Number of total logins in the first quarter.
Q1_URL_Skill_Grammar_Unique	Number of unique URL visits related to grammar skill in the first quarter.
Q1_URL_Skill_Vocabulary_Unique	Number of unique URL visits related to vocabulary skill in the first quarter.
Q1_URL_Skill_Speaking_Unique	Number of unique URL visits related to speaking skill in the first quarter.
Q1_URL_Skill_Listening_Unique	Number of unique URL visits related to listening skill in the first quarter.
Q1_URL_Skill_Writing_Unique	Number of unique URL visits related to writing skill in the first quarter.
Q1_URL_Skill_Reading_Unique	Number of unique URL visits related to reading skill in the first quarter.
Q1_URL_Skill_General_Unique	Number of unique URL visits related to general skill in the first quarter.
Q1_URL_Skill_Total_Unique	Number of unique URL visits in the first quarter.

Table 3.1. (continued)

Q1_Joined_Live_Session_Duration	The total duration of time joined live sessions in the first quarter.
Q1_Pass_Fail_Status	Indicates the pass or fail status at the end of the first quarter.
Q1_Skill_Writing	Reflects the achievement status of writing skill at the end of the first quarter.
Q1_Skill_Grammar_Reading_Listening	Reflects the achievement status of grammar, reading and listening skill at the end of the first quarter.
Q1_Skill_Speaking	Reflects the achievement status of speaking skill at the end of the first quarter.
Q2_Assignment_Submissions	The number of assignment submissions in the second quarter.
Q2_VClass_Rec_Downloads	Number of virtual classroom recording downloads in the second quarter.
Q2_VClass_Joins	Number of virtual classroom joins in the second quarter.
Q2_CourseViews_UniqueDays	Number of unique course views days in the second quarter.
Q2_FolderDownloads_UniqueDays	Number of unique folder download days in the second quarter.
Q2_UniqueFolderDownloads	Number of unique folder downloads in the second quarter.
Q2_FolderViews_UniqueDays	Number of unique folder view days in the second quarter.

Table 3.1. (continued)

Q2_UniqueFolderViews	Number of unique folder views in the second quarter.
Q1_URL_UniqueInteractions	Number of unique URL interactions days in the second quarter.
Q2_LoginDays	Number of login days in the second quarter.
Q2_TotalLogins	Number of total logins in the second quarter.
Q2_Joined_Live_Session_Duration	The total duration of time joined live sessions in the second quarter.
Q2_Pass_Fail_Status	Indicates the pass or fail status at the end of the second quarter.
Q2_Skill_Writing	Reflects the achievement status of writing skill at the end of the second quarter.
Q2_Skill_Grammar_Reading_Listening	Reflects the achievement status of grammar, reading and listening skill at the end of the second quarter.
Q2_Skill_Speaking	Reflects the achievement status of speaking skill at the end of the second quarter.
Q3_Assignment_Submissions	The number of assignment submissions in the third quarter.
Q3_VClass_Rec_Downloads	Number of virtual classroom recording downloads in the third quarter.
Q3_VClass_Joins	Number of virtual classroom joins in the third quarter.

Table 3.1. (continued)

Q3_CourseViews_UniqueDays	Number of unique course views days in the third quarter.
Q3_FolderDownloads_UniqueDays	Number of unique folder download days in the third quarter.
Q3_UniqueFolderDownloads	Number of unique folder downloads in the third quarter.
Q3_FolderViews_UniqueDays	Number of unique folder view days in the third quarter.
Q3_UniqueFolderViews	Number of unique folder views in the third quarter.
Q3_H5P_InteractionDays	Number of unique H5P interactions days in the third quarter.
Q3_Unique_H5P_Interactions	Number of unique H5P interactions in the third quarter.
Q3_NonUnique_Total_H5P_Interactions	Number of non-unique H5P interactions in the third quarter.
Q3_LoginDays	Number of login days in the third quarter.
Q3_TotalLogins	Number of total logins in the third quarter.
Q3_H5P_Skill_UseOfEnglish_Unique	Number of unique H5P interactions related to use of English skill in the third quarter.
Q3_H5P_Skill_Grammar_Unique	Number of unique H5P interactions related to grammar skill in the third quarter.

Table 3.1. (continued)

Q3_H5P_Skill_Writing_Unique	Number of unique H5P interactions related to writing skill in the third quarter.
Q3_H5P_Skill_Listening_Unique	Number of unique H5P interactions related to listening skill in the third quarter.
Q3_H5P_Skill_Reading_Unique	Number of unique H5P interactions related to reading skill in the third quarter.
Q3_H5P_Skill_Integrated_Unique	Number of unique H5P interactions related to integrated skill in the third quarter.
Q3_H5P_Skill_Speaking_Unique	Number of unique H5P interactions related to speaking skill in the third quarter.
Q3_H5P_Skill_Total_Unique	Number of unique H5P interactions related to skills in the third quarter.
Q3_Joined_Live_Session_Duration	The total duration of time joined live sessions in the third quarter.
Q3_Pass_Fail_Status	Indicates the pass or fail status at the end of the third quarter.
Q3_Skill_Writing	Reflects the achievement status of writing skill at the end of the third quarter.
Q3_Skill_Grammar_Reading_Listening	Reflects the achievement status of grammar, reading and listening skill at the end of the third quarter.

Table 3.1. (continued)

Q3_Skill_Speaking	Reflects the achievement status of speaking skill at the end of the third quarter.
Q4_Assignment_Submissions	The number of assignment submissions in the fourth quarter.
Q4_VClass_Rec_Downloads	Number of virtual classroom recording downloads in the fourth quarter.
Q4_VClass_Joins	Number of virtual classroom joins in the fourth quarter.
Q4_CourseViews_UniqueDays	Number of unique course views days in the fourth quarter.
Q4_FolderDownloads_UniqueDays	Number of unique folder download days in the fourth quarter.
Q4_UniqueFolderDownloads	Number of unique folder downloads in the fourth quarter.
Q4_FolderViews_UniqueDays	Number of unique folder view days in the fourth quarter.
Q4_UniqueFolderViews	Number of unique folder views in the fourth quarter.
Q4_H5P_InteractionDays	Number of unique H5P interactions days in the fourth quarter.
Q4_Unique_H5P_Interactions	Number of unique H5P interactions in the fourth quarter.
Q4_NonUnique_Total_H5P_Interactions	Number of non-unique H5P interactions in the fourth quarter.

Table 3.1. (continued)

Q4_LoginDays	Number of login days in the fourth quarter.
Q4_TotalLogins	Number of total logins in the fourth quarter.
Q4_H5P_Skill_UseOfEnglish_Unique	Number of unique H5P interactions related to use of English skill in the fourth quarter.
Q4_H5P_Skill_Grammar_Unique	Number of unique H5P interactions related to grammar skill in the fourth quarter.
Q4_H5P_Skill_Writing_Unique	Number of unique H5P interactions related to writing skill in the fourth quarter.
Q4_H5P_Skill_Listening_Unique	Number of unique H5P interactions related to listening skill in the fourth quarter.
Q4_H5P_Skill_Reading_Unique	Number of unique H5P interactions related to reading skill in the fourth quarter.
Q4_H5P_Skill_Integrated_Unique	Number of unique H5P interactions related to integrated skill in the fourth quarter.
Q4_H5P_Skill_Speaking_Unique	Number of unique H5P interactions related to speaking skill in the fourth quarter.
Q4_H5P_Skill_Total_Unique	Number of unique H5P interactions related to skills in the fourth quarter.

Table 3.1. (continued)

Q4_LoginDays	Number of login days in the fourth quarter.
Q4_TotalLogins	Number of total logins in the fourth quarter.
Q4_H5P_Skill_UseOfEnglish_Unique	Number of unique H5P interactions related to use of English skill in the fourth quarter.
Q4_H5P_Skill_Grammar_Unique	Number of unique H5P interactions related to grammar skill in the fourth quarter.
Q4_H5P_Skill_Writing_Unique	Number of unique H5P interactions related to writing skill in the fourth quarter.
Q4_H5P_Skill_Listening_Unique	Number of unique H5P interactions related to listening skill in the fourth quarter.
Q4_H5P_Skill_Reading_Unique	Number of unique H5P interactions related to reading skill in the fourth quarter.
Q4_H5P_Skill_Integrated_Unique	Number of unique H5P interactions related to integrated skill in the fourth quarter.
Q4_H5P_Skill_Speaking_Unique	Number of unique H5P interactions related to speaking skill in the fourth quarter.
Q4_H5P_Skill_Total_Unique	Number of unique H5P interactions related to skills in the fourth quarter.

Table 3.1. (continued)

Q4_Joined_Live_Session_Duration	The total duration of time joined live sessions in the fourth quarter.
Q4_Proficiency_Skill_Writing	Reflects the achievement status of writing skill of proficiency exam at the end of the fourth quarter.
Q4_Proficiency_Skill_Grammar_Reading_Listening	Reflects the achievement status of grammar, reading and listening skill of proficiency exam at the end of the fourth quarter.
Q4_Proficiency_Skill_Speaking	Reflects the achievement status of speaking skill at the end of the fourth quarter.
Proficiency_Pass_Fail_Status	Indicates the pass or fail status of the proficiency exam at the end of the fourth quarter.

3.2 Population, Sampling, and Participants

The population for this study consisted of students enrolled at the School of Foreign Languages at a public university in Türkiye. The research specifically targeted students participating in online classes and online exams across language learning during the COVID-19 pandemic, with a particular emphasis on those studying English over two academic semesters. These students formed the accessible population as their performance data, spanning multiple assessments, was readily available for analysis within the institutional LMS.

To facilitate data collection, convenience sampling was employed, targeting all enrolled students in English courses at the School of Foreign Languages. This method was selected due to the accessibility and willingness of participants within the structured online learning and assessment environment, which facilitated systematic data collection with the least affordance during the challenging period (Etikan, Musa, & Alkassim, 2016). Participants were invited to join the study through LMS pop-up notifications and a dedicated webpage. These platforms provided clear information about the research's objectives, the extent of data collection, and the data usage policy. Participation was strictly voluntary, ensuring that only data from students who provided informed consent were included in the analysis. Ethical principles, such as transparency, confidentiality, and data security, were rigorously upheld throughout the study, in accordance with institutional guidelines and ethical standards (Resnik, 2018).

The final sample comprised 481 students, representing a significant proportion of the target population. This group encompassed a diverse range of academic programs, ensuring heterogeneity and broad applicability of the findings. The robust dataset enabled a comprehensive analysis of English language performance, focusing on progression and learning outcomes within the pandemic's remote/online education context. The inclusion of participants from varied academic backgrounds provided valuable insights into group-specific engagement with online English language learning, facilitating a nuanced understanding of the challenges and opportunities presented by this mode of instruction (Bozkurt et al., 2020).

By employing this sampling approach and engaging a diverse participant base, the study enhanced its generalizability within the institutional context while maintaining rigorous adherence to ethical standards and methodological rigor.

3.3 Research Method

In this study, LA procedures were utilized to predict student performance and provide insights into online English language learning during challenging times. LA encompasses several interconnected processes, including data collection, pre-processing, metric selection, analysis, visualization, and interventions (Clow, 2012).

These components form the basis of the LA cycle, a well-regarded framework in the field (Khalil, Prinsloo, & Slade, 2023; Ameloot et al., 2024).

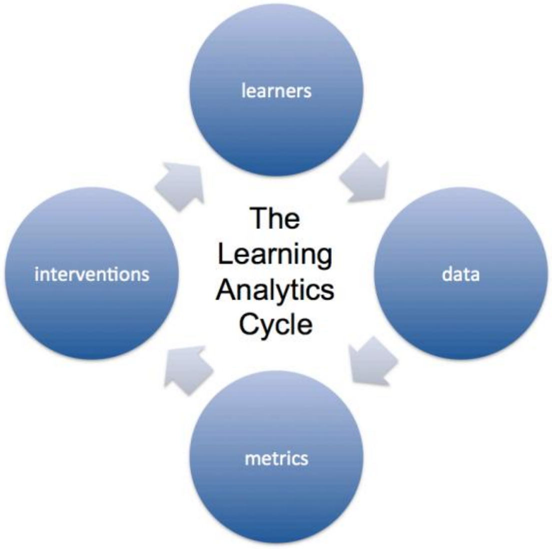

Figure 3.4. The LA cycle, from Clow (2012, p. 134).

Since interventions are the ultimate goal of LA, predictive modeling plays a crucial role by enabling education professionals to forecast student success and implement early interventions. In this regard, this research study adopted the classification method within predictive goals and modeling. The two primary classification approaches in predictive modeling are binary classification, which is used for pass/fail or dropout predictions, and multi-class classification, which is employed to segment students into different performance levels (Andersson, Arvemo, & Gellerstedt, 2016; Jayaprakash et al., 2014). Both classification approaches were employed in this study to answer the RQs effectively.

3.3.1 Learning Analytics Objectives

According to Verbert, Manouselis, Drachsler, and Duval (2012), key objectives of learning and knowledge analytics include predicting learner performance, modeling learner behavior, recommending relevant resources, enhancing reflection and awareness, and detecting undesirable learning patterns. This research study aligns with these objectives, with a particular focus on students' performance prediction and providing individual and organizational feedback.

Further refining these objectives, Seufert, Meier, Soellner, and Rietsche (2019) conducted a systematic literature review, categorizing common LA research goals into performance prediction, formative feedback, social LA, and the effective utilization of LA tools. The categories also align closely with objectives of this study, emphasizing predictive insights and reflective practices that benefits learners, instructors, administrations, and institutions in a larger aspect.

In this context, the study prioritizes predictive analytics to forecast student performance. It also underscores the need for systematic implementation to enhance decision-making processes for diverse stakeholders.

3.3.2 Frameworks and Dimensions of Learning Analytics

The LA life cycle framework provides a generalizable model for structuring LA processes. However, alternative frameworks provide more detailed insights tailored to specific research contexts. For instance, Greller and Drachsler's (2012) essential dimensions of LA highlight six interconnected components: objectives, stakeholders, data, instruments, external constraints, and internal limitations. These dimensions also emphasize the interplay between technical and contextual factors influencing the deployment of LA.

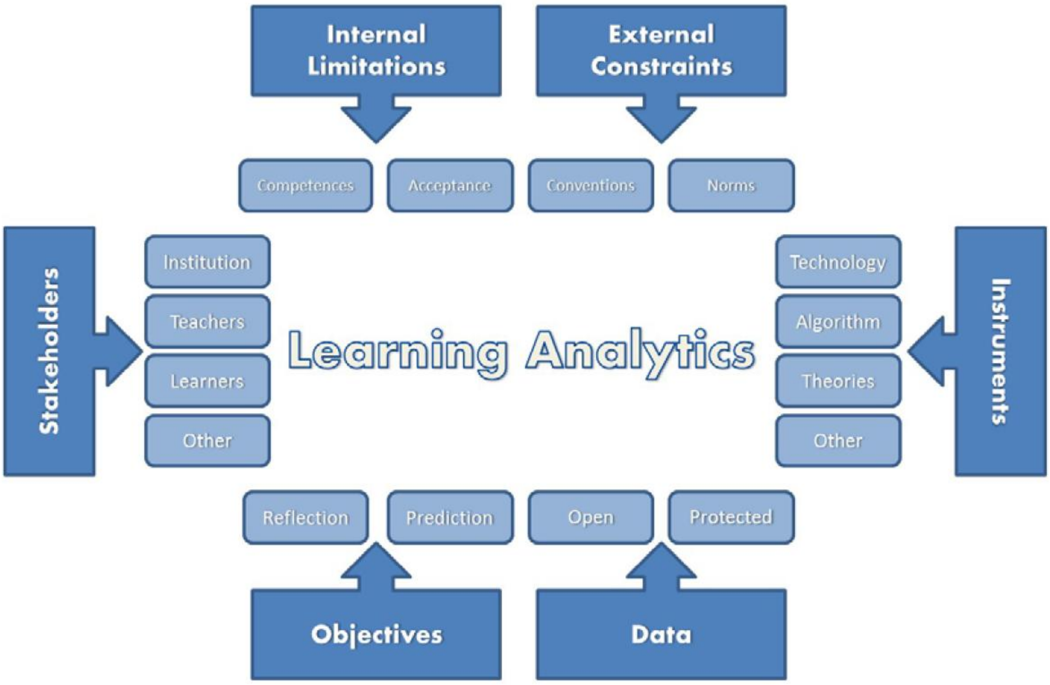

Figure 3.5. Critical dimensions of LA, from Greller and Drachsler (2012, p. 44)

The relevance of these dimensions to this study is also evident in its focuses on internal, external considerations, instruments, data, objectives, and stakeholders. Internal constraints, such as monitoring student engagement, align with the operational goals of the learning environment, while external limitations, including ethical considerations, ensure compliance with institutional and societal standards. Ethical clearance was obtained from the institutional ethics committee, reinforcing adherence to protocols for confidentiality, informed consent, and responsible data usage (Greller & Drachsler, 2012). Furthermore, the instruments utilized in this study included advanced machine learning algorithms, such as Naïve Bayes, Random Forest, and Logistic Regression, to ensure robust data analysis and prediction accuracy. Data protection measures were implemented using secure software platforms across the institution, adhering to stringent privacy standards. The primary objectives of the study centered on prediction of learning outcomes and fostering reflective practices among stakeholders. This research study encompassed a wide range of stakeholders, including learners, instructors, administrators, and institutional policymakers, recognizing their critical roles in the implementation and application of LA.

3.3.3 Predictive Models and Algorithms

Classification algorithms formed a cornerstone of this study, providing a systematic approach to predicting student performance through the analysis of diverse and extensive datasets. Widely recognized methodologies in EDM and LA, such as regression, clustering, and classification, are foundational, and among these, classification stands out as one of the most frequently applied techniques, offering significant utility for outcome prediction and student profiling (Baker & Yacef, 2009; Romero & Ventura, 2020).

To address the study's predictive objectives, multiple machine learning algorithms were employed, including Naïve Bayes, Support Vector Machines (SVM), Neural Networks, Logistic Regression, k-NN, and Random Forest. These algorithms were selected based on their established effectiveness in educational research and their capacity to handle varied data complexities (Shamsi & Lakshmi, 2016; Akçapınar, Altun, & Aşkar, 2019; Altabrawee, Ali, & Ajmi, 2019). Each model's performance was rigorously evaluated using cross-validation techniques and hyperparameter optimization, ensuring robustness and mitigating overfitting risks.

Algorithm selection was guided by the specific characteristics of the dataset, such as its size, dimensionality, and distribution. Evaluation metrics, including accuracy, precision, recall, and F1 score, were employed to benchmark algorithm performance; hence, this approach allowed for a comprehensive comparison, enabling the identification of the most effective models for further analysis, reporting and application (Roy & Garg, 2017).

Predictive models derived from these algorithms play a critical role in educational contexts by enabling the prediction of student outcomes, the early identification of at-risk students, and the design of targeted learning interventions. By employing these tools, this study not only improved predictive accuracy but also contributed to the optimization of learning processes and evidence-based decision-making for educators and administrators.

To address the RQs, the study employed, within models, a diverse set of classification algorithms, encompassing both traditional and advanced machine learning techniques. These machine learning algorithms with models included:

- **AdaBoost** is an ensemble learning algorithm that improves weak classifiers by sequentially adjusting weights, giving more focus to misclassified instances in each iteration (Freund & Schapire, 1997). This results in a more robust model that reduces errors. In LA, AdaBoost is used for predicting student dropouts, evaluating academic performance, and enhancing adaptive learning environments. It is particularly effective in handling noisy data, where traditional classifiers struggle.
- **CN2 Rule Induction** is a rule-based classification algorithm that generates IF-THEN rules from data, allowing for interpretable decision-making (Clark & Niblett, 1989). This is particularly valuable in LA, where understanding student behaviors and performance is essential. By analyzing past academic records, CN2 can identify students at risk of failing and provide actionable recommendations to educators and administrators.
- **Gradient Boosting** builds predictive models sequentially by optimizing for errors in models using gradient descent (Friedman, 2001). It is highly effective for structured data and is widely applied in LA to predict student success rates, personalize learning pathways, and assess engagement trends. Compared to AdaBoost, Gradient Boosting provides finer control over learning rates and feature interactions, making it a preferred choice for detailed educational data analysis.
- **k-Nearest Neighbors (k-NN)** is a non-parametric algorithm that classifies data points based on their proximity to k nearest neighbors (Cover & Hart,

1967). It is commonly used in LA to cluster students based on similar learning behaviors, predict course completion likelihood, and provide recommendations for peer-assisted learning. Since k-NN is sensitive to feature scaling and distance metrics, preprocessing steps like normalization are crucial when applying it to educational datasets.

- **Logistic Regression** is a statistical model that estimates the probability of an event occurring, typically used for binary classification tasks (Hosmer Jr, Lemeshow, & Sturdivant, 2013). In education, it predicts student outcomes such as pass/fail rates, identifies influential factors affecting academic success, and assesses the effectiveness of learning interventions. Its simplicity and interpretability make it an essential tool for educational research and policy-making.
- **Naïve Bayes** is a probabilistic classifier based on Bayes' Theorem, assuming independence between features (Lewis, 1998). Despite its simplistic assumption, it performs well in text classification tasks, making it useful for sentiment analysis of student feedback, email classification, and academic performance predictions. It is often applied in detecting at-risk students by analyzing historical performance trends.
- **Neural Networks** are deep learning models designed to recognize complex patterns by mimicking the structure of the human brain (LeCun, Bengio, & Hinton, 2015). In learning analytics, they are used for adaptive learning, student behavior prediction, and identifying cheating patterns in online assessments. By processing large volumes of data, Neural Networks can uncover hidden trends that traditional statistical models might miss, enhancing personalized education strategies.
- **Random Forest** is an ensemble method that constructs multiple decision trees and aggregates their outputs to improve accuracy and reduce overfitting (Breiman, 2001). In LA, it is widely applied to predict student retention, and provide recommendations for personalized course content. Its ability to

handle large datasets with missing values makes it highly effective in analyzing educational data.

- **Support Vector Machines (SVM)** is a supervised learning algorithm that finds the optimal hyperplane to separate classes in high-dimensional space (Cortes & Vapnik, 1995). In education, SVM is used for early detection of struggling students, classifying learning behaviors, and predicting academic achievement. The algorithm is particularly effective when dealing with complex, high-dimensional data such as student interaction logs from online learning platforms.
- **Stochastic Gradient Descent (SGD)** is an optimization technique that updates model parameters iteratively, allowing efficient training on large datasets (Bottou, 2010). In LA, it can be used to train deep learning models in real-time, optimize adaptive learning environments, and enhance recommendation systems. Its fast convergence and ability to handle sparse data make it suitable for dynamic educational datasets.
- **Decision Trees** are hierarchical models that split data into branches based on feature conditions, allowing for intuitive and interpretable decision-making (Quinlan, 1986). In LA, decision trees are used to identify factors influencing student success, predict dropouts, and recommend interventions. They provide a clear visualization of decision processes, making them an effective tool for educators seeking to understand student learning patterns.

These algorithms within models were evaluated and explained using the following key performance metrics to ensure a comprehensive assessment of their predictive capabilities:

- **Area under the Curve (AUC)** measures a model's ability to distinguish between different classes by evaluating the trade-off between the true positive rate and false positive rate across various threshold values. A higher AUC indicates better classification performance and is considered one of the most effective methods for evaluating a model's discriminative ability

(Bradley, 1997). In LA, AUC is commonly used to assess how effectively models can predict outcomes such as student success or dropout risk.

- **Classification Accuracy (CA)** calculates the percentage of correct predictions over total instances (Tharwat, 2020). It is commonly used to evaluate student performance predictions but may be misleading when class distributions are imbalanced (e.g., when most students pass, leading to high accuracy but poor identification of failing students). The classification accuracy formula is given as follows:

$$\mathrm{CA} = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN}}$$

- **F1 Score or F-measure** is a crucial evaluation metric in machine learning that balances precision and recall, especially in imbalanced datasets. It is defined as the harmonic mean of precision and recall, providing a single measure of a model's performance (Christen, Hand, & Kirielle, 2023). In LA, the F1 Score is particularly useful when both false positives and false negatives have significant consequences. For instance, when predicting which students are struggling, a balance is needed between identifying those at risk and avoiding unnecessary interventions for those performing well. F1 Score can be calculated using the following formula:

$$\mathrm{F1} = 2 \text{ x } \frac{\text{Precision x Recall}}{\text{Precision} + \text{Recall}}$$

- **Precision** focuses on the accuracy of positive predictions, making it valuable in detecting students who will succeed (Powers, 2020). A high precision score ensures that only students who are truly excelling are classified as successful. A low precision score means that many struggling students are mistakenly classified as passing, which can lead to ineffective interventions. The formula of precision is shown below:

$$\text{Precision} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}}$$

- **Recall** measures how well the model identifies actual positive cases, such as correctly detecting all well-performing students (Sokolova & Lapalme, 2009). A high recall ensures that fewer successful students are overlooked. The recall can be calculated as follows:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

- **Matthews Correlation Coefficient (MCC)** provides a comprehensive assessment of classification performance, considering all aspects of the confusion matrix, and yielding a coefficient value between -1 and +1 (Baldi et al., 2000). It is particularly effective for imbalanced educational datasets, such as predicting dropout rates in small at-risk student groups.
- **Synthetic Minority Over-sampling Technique (SMOTE**) is a data preprocessing technique used to address class imbalance in datasets. It works by generating synthetic examples of the minority class (e.g., failing students) to balance the number of samples between classes (Chawla et al., 2002). This helps improve model performance, especially in cases where models tend to favor the majority class. In LA, SMOTE is often applied to increase the representation of underperforming or at-risk students in training data, enhancing the predictive accuracy of classifiers.
- **SHapley Additive exPlanations (SHAP)** is a model-agnostic interpretability method that explains the impact of each feature on a model's prediction using game theory (Lundberg & Lee, 2017). It assigns a contribution value (SHAP value) to each feature, helping to understand how the model arrived at a specific prediction. In LA, SHAP is valuable for identifying which factors (e.g., attendance, quiz scores, login frequency) most influence student outcomes, providing transparency and supporting data-driven decision-making.

The evaluation process also involved cross-validation and hyperparameter tuning to enhance model robustness and generalizability. The Synthetic Minority

Oversampling Technique (SMOTE) was also utilized to balance the minority class in the dataset.

In addition to performance metrics, model interpretability was prioritized to facilitate actionable insights. SHapley Additive exPlanations (SHAP) values were utilized to interpret the predictions, providing a transparent understanding of feature importance and contributions, and enabling informed decision-making (Lundberg & Lee, 2017). This approach ensured that both the accuracy and usability of the predictive models were addressed, aligning with the study's goal of balancing performance with interpretability.

The most effective algorithm based on these criteria was selected for further analysis and reporting, offering a detailed examination of their contributions to the study's objectives.

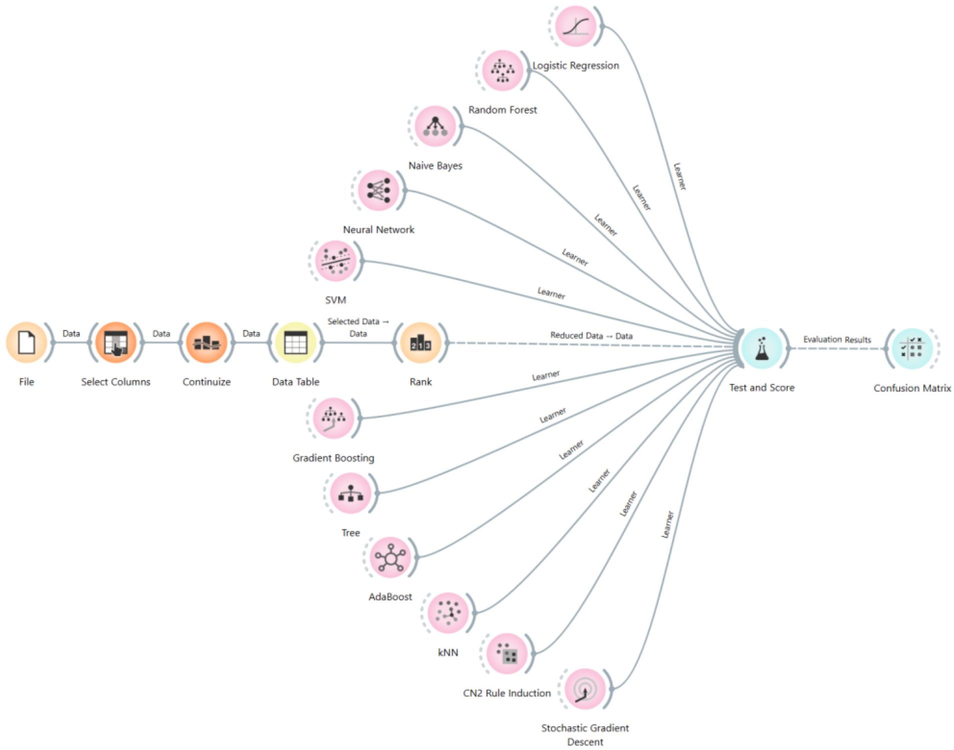

Figure 3.6. The Predictive Model in This Research Study

The first model incorporating machine learning algorithms (Figure 3.6) was utilized when predictive analytics performed better without training data processed with SMOTE.

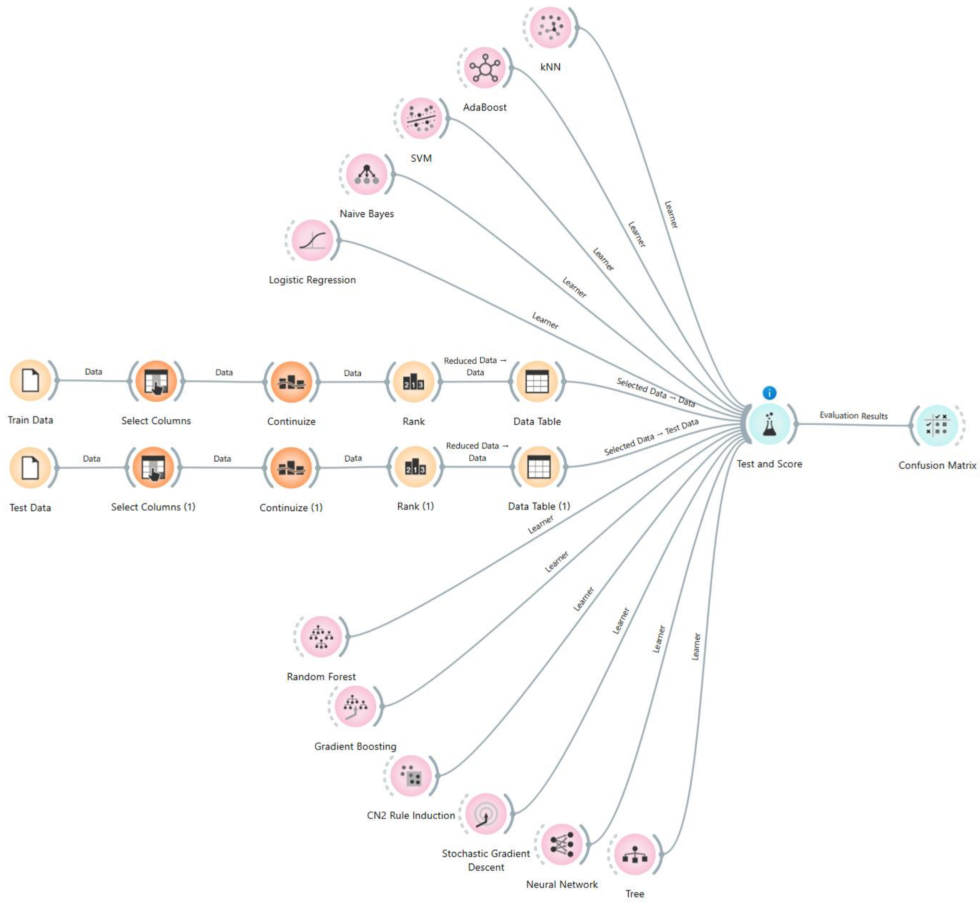

Figure 3.7. The Predictive Model Using SMOTE in This Research Study

The second model, incorporating classification algorithms (Figure 3.7), was utilized when predictive analytics performed better with training data processed using SMOTE.

3.3.4 Tools and Programming Languages

This study employed a combination of programming and visual tools to optimize data preprocessing, analysis, and model evaluation. In addition to programming-based solutions, user-friendly platforms like WEKA and Orange Data Mining are integrated into the workflow for their intuitive interfaces in research studies. WEKA, renowned for its suite of machine learning algorithms and visualization tools, simplifies exploratory data analysis and model evaluation. Similarly, Orange Data Mining, which was a key tool in this research study, offers an accessible yet powerful environment for implementing classification algorithms, statistical methods, and visual workflows (Holmes, Donkin, & Witten, 1994; Demšar et al., 2013).

The combination of these tools ensures a balance between analytical rigor and accessibility, enabling the study to cater to both technical and non-technical researchers. In this study, this hybrid approach not only streamlined the research process but also ensured comprehensive and interpretable results, supporting robust decision-making.

3.4 Data Sources and Instruments

The data for this study were derived from multiple sources, including activity logs from MOODLE and Blackboard Collaborate Ultra, as well as demographic and academic records from the Student Information System (SIS) at the School of Foreign Languages in a public university. These datasets offered a holistic view of student activities, engagement patterns, and academic outcomes, providing critical insights into the online learning process (Tempelaar, Rienties, & Giesbers, 2015).

The data comprised MOODLE and Blackboard Collaborate Ultra logs, capturing key interactions such as resource views, assignment submissions, and attendance in virtual classroom sessions. Academic performance metrics included scores from various assessments, such as progress tests, gateway exams, and proficiency exams. These assessments utilized diverse formats, including online exams conducted via

MOODLE, hand-written writing exams, and speaking exams administered through Microsoft Teams. The inclusion of such diverse assessment data enabled a detailed analysis of students' language proficiency and skill development (Romero & Ventura, 2013).

To ensure data accuracy and consistency, MOODLE logs detailing resource usage and activity timelines were exported and organized systematically using Microsoft Excel. Exam scores collected through MOODLE, SIS and other platforms were manually integrated into the dataset, ensuring coherence and completeness. Similarly, demographic and academic details, such as program enrollment, age, and gender, were retrieved from the SIS and cross-referenced with activity logs and exam data. This careful integration provided comprehensive student profiles, enabling meaningful contextual analysis (Gašević, Dawson, & Siemens, 2015).

The combination of these datasets facilitated a robust analysis of student engagement, performance, and progression within the context of online English language learning. By synthesizing log data with demographic and academic performance information, the study applied advanced LA techniques to uncover actionable insights. These insights can inform pedagogical strategies and guide targeted interventions to enhance teaching practices and support student success (Tempelaar, Rienties, & Giesbers, 2015; Tempelaar et al., 2012).

3.5 Data Collection Procedure

The data collection procedure for this study integrated multiple datasets to enable a holistic analysis of student performance in online English language learning. Data collection began with a consent process initiated through a pop-up notification in the LMS, which transparently outlined the study's purpose and sought participant approval for using demographic data from the SIS, activity logs from MOODLE and Blackboard Collaborate Ultra, and assessment performance data. This approach ensured clarity and accessibility, aligning with recommended best practices in LA research (Siemens & Long, 2011).

Structured Query Language (SQL) queries were employed to extract relevant data from the MOODLE database (MySQL). The data were subsequently organized in Microsoft Excel to facilitate integration with additional datasets. To uphold data integrity and ethical standards, only the information from participants who explicitly provided consent was included in the study, minimizing biases and ensuring transparency (Ferguson, 2012).

Data integration involved systematically merging multiple sources, including MOODLE activity logs, Blackboard Collaborate Ultra interaction records, and scores from various assessments, such as online gateway exams, proficiency exam, and their components. Unique identifiers, such as student numbers and MOODLE IDs, were meticulously matched across platforms to ensure accurate and consistent data linkage. Preprocessing steps, including data cleaning, formatting, and standardization, were conducted to ensure compatibility with analysis tools and methodologies. For example, MOODLE logs capturing interactions with resources, assignment submissions, and virtual classroom participation were consolidated with exam performance data to create a unified dataset suitable for in-depth analysis (Avella et al., 2016).

This comprehensive data collection process enabled the examination of key learning behaviors and academic outcomes in an online education context. By integrating LMS logs and virtual classroom logs with performance data, also demographic data in descriptive analysis, the study leveraged LA to identify patterns and provide actionable insights to support evidence-based teaching practices. Such data-driven methodologies are increasingly utilized in educational research to enhance learning experiences, improve student outcomes, and optimize algorithmic interventions (Arnold & Pistilli, 2012; Viberg et al., 2018).

3.6 Data Pre-Processing

The data utilized in this study were collected from diverse platforms, including MOODLE databases, the SIS, Virtual Classroom logs, and exam software. Due to the heterogeneous structure of these datasets, a rigorous pre-processing approach was essential to prepare them for analysis. The university's reliance on multiple LMSs, each employing distinct virtual classroom tools, necessitated the filtering of relevant logs. Linux (Debian) commands were employed to efficiently extract logs corresponding to students who had provided consent. Such technical methods can be employed in LA for managing and harmonizing large-scale, multi-source data.

MOODLE logs were exported in CSV format to ensure compatibility with data mining and analysis tools. Subsequent pre-processing steps were conducted using Google Colab, leveraging Python scripts for cleaning, managing missing values, and extracting relevant features. The use of Python’s versatile libraries, such as Pandas and NumPy, facilitated efficient data manipulation, which is critical for preparing datasets for machine learning algorithms. Google Colab's collaborative interface streamlined these processes, while Microsoft Excel was employed for initial exploration, anomaly detection, and manual cleaning to ensure data consistency.

The integration of datasets required aligning identifiers such as student numbers, MOODLE IDs, and other unique keys to accurately merge demographic data, interaction logs, and academic performance metrics. Meticulous attention was paid to maintaining data integrity during this process, as errors in merging could significantly undermine the validity of subsequent analyses. To ensure compliance with ethical standards, personally identifiable information was anonymized during pre-processing. This step adhered to institutional and legal guidelines for handling sensitive data, following established best practices in data privacy and ethical research (Slade & Prinsloo, 2013).

Once pre-processed, the dataset was imported into Orange, a data mining tool renowned for its intuitive interface and robust machine learning capabilities.

Features were aligned and validated within Orange to ensure readiness for advanced analysis. Orange's capabilities facilitated the exploration of the dataset and the application of classification and prediction models, aligning with the study's LA objectives (Demšar et al., 2013). The combination of automated tools and manual validation ensured that the data were meticulously prepared for meaningful analysis, providing a solid foundation for generating actionable insights—a cornerstone of effective LA (Tempelaar, Rienties, & Giesbers, 2015).

3.7 Data Mining Tool Selection

After the pre-processing phase, which extensively utilized Linux commands (Debian), Google Colab, and MS Excel, the Orange data mining tool was selected for the analysis phase. Orange was chosen for its intuitive visual interface, which simplifies complex analytical workflows through a drag-and-drop system of widgets. These widgets enable a variety of essential data mining tasks, such as clustering, classification, and predictive modeling, making Orange an ideal tool for analyzing student engagement and predicting performance across diverse educational datasets (Demšar et al., 2013).

One of Orange's key strengths lies in its accessibility. Comprehensive tutorials and documentation support a smooth learning curve, making it a practical choice for researchers with varying levels of expertise in data mining and computer technologies. Additionally, Orange integrates seamlessly with Python, providing enhanced flexibility by allowing custom scripts to extend its capabilities. This combination enables researchers to leverage Orange's visual interface while taking advantage of Python's computational power, creating a robust and adaptable analytical environment.

During the tool selection process, alternative platforms such as Weka and RapidMiner were also considered. While Weka offers a wide range of algorithms and is accompanied by thorough documentation, its interface is less conducive to visual workflows, especially when compared to Orange (Holmes, Donkin, & Witten, 1994). Similarly, RapidMiner was noted for its advanced analytics capabilities, but its steep learning curve and licensing costs rendered it less suitable for the study's requirements.

Using Orange, the study efficiently generated visual and analytical outputs, which facilitated the analysis of student data and the derivation of actionable insights. This combination of intuitive design and robust functionality allowed the researcher to focus on interpreting results rather than managing technical complexities. Orange’s balance of usability and advanced features aligned with the study’s objective of employing LA to enhance educational outcomes (Romero & Ventura, 2020).

3.8 Data Analysis

The quantitative data collected from the LMS, exam scores, and demographic sources were analyzed using descriptive statistics to summarize and interpret student performance. Key metrics such as frequencies, and percentages were calculated to examine central tendencies and variability in the data. These descriptive statistics were visually represented through tables to enhance clarity and facilitate interpretation. Visualizations provided insights into student engagement, exam outcomes, and demographic patterns within the dataset, aligning recommended practices in educational data analysis (Romero & Ventura, 2020).

Beyond descriptive analysis, advanced classification algorithms, including decision trees and k-nearest neighbors (k-NN), were employed through the Orange data mining tool. These algorithms were selected for their effectiveness in identifying patterns and classifying students based on performance indicators such as LMS interaction logs and exam scores. Decision trees, known for their interpretability, were used to explore the relationships between variables, while k-NN provided robust predictions of student performance by analyzing proximity in feature spaces (Ahmed, 2024). This approach enabled the study to predict academic outcomes, expectedly future ones, and identify students at risk of underperforming, addressing key research objectives.

The analysis process was structured to align with the study's RQs, with specific variables selected for each sub-question. For instance, LMS engagement data, such as login frequencies and resource access patterns, were analyzed to determine their contribution to exam performance. Predictive models, developed using the selected classification algorithms, generated actionable insights by identifying students likely to face challenges in their academic progress, thus informing the design of early intervention strategies (Macfadyen & Dawson, 2010).

3.9 Ethical Issues

Ethical considerations were prioritized at every stage of the study to ensure that the research adhered to the highest standards of integrity and responsibility. This commitment to ethical rigor aligns with the frameworks outlined in literature on ethical and privacy principles in LA (Pardo & Siemens, 2014; Slade & Prinsloo, 2013). Permission for the study was obtained from the Institutional Ethics Board at Middle East Technical University, ensuring compliance with institutional and national ethical guidelines. Furthermore, the research design adhered to the principles of transparency and accountability, as advocated in ethical oversight frameworks (Willis, Slade, & Prinsloo, 2016; Resnik, 2018). Data collection activities were coordinated with relevant academics and administrative officers to maintain transparency and ensure alignment with institutional policies.

Throughout the research study, efforts were made to eliminate identifying information to protect participants' privacy and confidentiality. This approach was informed by the understanding that privacy is a fundamental right and a critical pillar of ethical research practices within educational data (Rubel & Jones, 2016). Personal data such as names, student IDs, and other sensitive identifiers were anonymized or pseudonymized during the pre-processing and analysis phases. Additionally, data access was restricted to the researcher only, further safeguarding participants' information.

To minimize the risk of harm, the researcher carefully designed the study to avoid any potential negative impact on students, instructors, managers, or the institution. Participation in the study was voluntary, and informed consent was obtained prior to data use. The informed consent process was structured to uphold student autonomy and promote their understanding of the research purpose, as recommended in informed consent models (Slade & Prinsloo, 2013). Participants were provided with clear information about the purpose of the study and their rights to withdraw without penalty. This also aligns with ethical standards outlined by organizations such as the American Educational Research Association (AERA) and ensures respect for participant autonomy and avoidance for harm to stakeholders (AERA, 2011).

The research adhered to data protection laws applicable in Türkiye, such as the Personal Data Protection Law (KVKK), to ensure compliance with legal frameworks governing the use of personal data. Compliance with such regulatory frameworks parallels the international call for ethical considerations of educational data (West et al., 2016; Hoel & Chen, 2019). All findings were reported in a way that preserved anonymity and did not disclose individual or institutional identities, thereby preventing any reputational harm or unintended consequences. This ethical responsibility is emphasized in discussions of privacy-preserving LA (Gursoy et al., 2016).

3.10 Researcher’s Background

The researcher serves as both a lecturer and a Linux system administrator at the Faculty of Open and Distance Education in a public university in Türkiye. His responsibilities include overseeing and managing the technological infrastructure that supports online learning, with a particular focus on open-source LMSs and other relevant software. This dual role has enabled him to develop a comprehensive understanding of the technological and operational aspects of digital education platforms.

Prior to transitioning to academia, he worked as a Linux system administrator and team leader for large-scale, nationwide IT projects, where he gained extensive experience in open-source solutions, strategic project management, and team leadership.

His primary research interests lie in DE technologies, educational free software (GNU GPL), and instructional design and technology. With a multidisciplinary focus, the researcher applies his expertise across various domains, including education, engineering, tourism, and aviation at the higher education level. This diverse academic and professional background uniquely positions him to bridge the gap between technology and pedagogy, enabling him to develop and implement innovative solutions for online and blended learning environments.

This blend of technical and pedagogical expertise played a pivotal role in this study. The researcher’s technical acumen supported critical tasks such as data extraction, preprocessing, and analysis, ensuring the integrity and efficiency of the research sub-processes. Simultaneously, his pedagogical insights significantly contributed to interpreting findings, particularly those related to student engagement, academic performance, and the broader implications for DE.

Throughout the study, the researcher adhered rigorously to ethical guidelines, safeguarding data security and maintaining objectivity to mitigate any potential bias arising from his dual responsibilities. This approach ensured the credibility, reliability, and validity of the study's outcomes.

3.11 Delimitation and Limitations of the Study

3.11.1 Delimitation

This study was delimited to one public university in Türkiye, a prominent institution recognized for its wide range of academic programs and diverse student population. Focusing on a single institution ensured consistency in data collection and minimized variability arising from differences in institutional policies and practices. However, it is acknowledged that Schools of Foreign Languages in other universities may employ varying instructional strategies, assessment methods, and administrative policies, which could influence language learning outcomes and paths in distinct ways. For instance, the availability of resources, class sizes, and administrative structures may vary significantly; thereby affecting both student engagement, academic performance, and overall learning and teaching experience.

While many institutions adopt common frameworks such as the Common European Framework of Reference for Languages (CEFR) for language proficiency, institutional differences may still pose challenges for generalizing the findings. Future research could expand this study to include multiple universities, enabling a comparative analysis of institutional practices and validating the broader applicability of the findings. Such an extension would provide valuable insights into the interplay between institutional factors and student outcomes in online foreign language learning, specifically English.

3.11.2 Limitations

The generalizability of the study's quantitative results is inherently limited by the focus on a single public university. Although the findings reveal trends and patterns in online English language learning, their primary purpose is to inform the design, feedback mechanisms, and evaluation of online courses within a specific institutional context rather than to establish universal principles. As highlighted by Romero and

Ventura (2020), such insights serve as a valuable starting point for improving online foreign language education, particularly during challenging periods.

A key limitation of this study is the reliance on voluntary participation, which introduces potential selection bias. Students who consented to participate may have been more motivated or engaged than those who declined, thereby limiting the randomness of the sample. While the study included 481 students, which constitutes a substantial sample size for this institution, some data mining models may require larger datasets for robust predictive accuracy. Consequently, the analysis relied on simpler classification models, which, while effective, may lack the complexity needed to capture intricate patterns in larger datasets.

The absence of comprehensive proctoring measures significantly undermines the validity and reliability of the study's findings as a critical limitation. Limited to monitoring through student cameras and microphones, this approach offers inadequate oversight during assessments, thereby increasing the risk of academic dishonesty (Holden, Norris, & Kuhlmeier, 2021). Although students completed their writing tasks by hand and submitted them via the LMS, this method alone does not guarantee authenticity or prevent external assistance. Moreover, unproctored or limited proctored testing conditions may not accurately reflect standardized assessment environments, potentially compromising the generalizability of the results. Consequently, the lack of robust proctoring protocols diminishes the confidence with which these findings can be interpreted and applied to broader educational contexts (Fidas et al., 2023).

Additionally, the transferability of findings to other contexts may be constrained by institutional differences in technological infrastructure, teaching practices, and administrative policies. For example, disparities in LMS capabilities and teacher training can significantly affect the outcomes of online learning initiatives (Kamraju et al., 2024). The unique circumstances of the COVID-19 pandemic also posed a limitation, as students’ engagement and learning behaviors during this period may not fully reflect their typical educational experiences. Future studies incorporating larger, multi-institutional datasets and conducted under more stable conditions would yield more generalizable and comprehensive insights.

CHAPTER 4

RESULTS

This chapter presents the study's results through three RQs that shaped this study following general descriptive statistics for participant demographics and English proficiency exam achievement. Subheadings in this section try to answer each RQ.

4.1 Summary of Descriptive Results

General descriptive data provided information related to achievement in the English proficiency exam, categorized by faculties, departments, registered study level (vocational school or faculty), language of instruction in the program, type of instruction in the program, and gender.

4.1.1 Descriptive Results Related to Faculties

The data presented in Table 4.1 provides an overview of student performance across faculties in the English proficiency exam.

Table 4.1. Achievement According to Faculties

Faculty	Fail	Pass	Total
Agriculture	15	7	22
Divinity	8	6	14
Engineering	101	94	195
Languages and History-Geography	18	14	32
Medicine	1	7	8
Open and Distance Education	17	5	22
Pharmacy	5	15	20
Politics	19	25	44
Science	29	25	54
Veterinary Medicine	12	21	33
Vocational School	10	1	11
Other	17	9	26
Total	252	229	481

The descriptive analysis of pass and fail rates across various faculties in online English proficiency exam revealed significant disparities. Students of faculties such as Medicine and Pharmacy demonstrated the highest pass rates, at 87.5% (7 out of 8) and 75% (15 out of 20), respectively. These strong results likely reflected the rigorous entry requirements and selective nature of these faculties, which positively influenced their students' achievements in learning English.

In contrast, Vocational School recorded the lowest pass rate at just 9.1% (1 out of 11), with only one student passing. This stark disparity suggested significant challenges, possibly due to a curriculum that was not sufficiently aligned with the language learning needs of students in such programs. Similarly, Open and Distance Education also exhibited a relatively low pass rate of 22.7% (5 out of 22), highlighting further disparities in performance.

Faculties such as Engineering and Science fell within the middle range of pass rates for the English proficiency exam at the end of two semesters, with 48.2% (94 out of 195) and 46.3% (25 out of 54), respectively. These results suggested that while many students from these faculties achieved proficiency in English, a substantial number still required additional support to meet the language requirements successfully.

4.1.2 Descriptive Results Related to Departments

The data presented in Table 4.2 provides a descriptive overview of student performance across various departments in the English proficiency exam. A total of 481 students were eligible to participate in the exam at the beginning of the academic year, with 252 failing and 229 passing at the end of two semesters, resulting in an overall pass rate of approximately 47.6%.

Table 4.2. Achievement According to Departments

Department	Fail	Pass	Total
Biology	6	6	12
Biomedical Engineering	14	15	29
Business Administration	24	13	37
Chemical Engineering	13	27	40
Chemistry	4	5	9
Computer Engineering	4	19	23
Divinity	8	6	14
Electrical and Electronics Engineering	14	10	24
Energy Engineering	11	2	13
Fisheries and Aquaculture	8	4	12
Food Engineering	25	14	39
Geography	8	5	13
Geology Engineering	7	1	8
Mathematics	11	13	24
Medicine	1	7	8
Pharmacy	5	15	20
Philosophy	5	3	8
Physics Engineering	13	5	18
Politics and Economy	4	7	11
Veterinary Medicine	12	21	33
Other	55	31	86
Total	252	229	481

Several departments demonstrated exceptional outcomes. Medicine department achieved the highest pass rate at 87.5% (7 out of 8), followed by Pharmacy with a 75% pass rate (15 out of 20). Similarly, Computer Engineering students showcased strong performance, with 82.6% (19 out of 23) of students passing the exam. These results suggest that rigorous academic programs, selective admission criteria, and students’ established study habits likely contributed to their students' success.

Students of departments such as Veterinary Medicine and Mathematics exhibited moderate performance, with pass rates of 63.6% (21 out of 33) and 54.2% (13 out of 24), respectively. Biomedical Engineering students, with a pass rate of 51.7% (15 out of 29), also fell into this category. These results indicate that while a significant proportion of students achieved proficiency, a substantial number required additional support to meet the proficiency standard in English. Students of Politics and Economy department, with a pass rate of 63.6% (7 out of 11), further reflected this middle tier of achievement.

In contrast, several departments reported significantly lower pass rates, raising concerns about their student preparedness or the suitability of online learning environment for their needs. For instance, students of Geology Engineering and Energy Engineering department recorded pass rates of 12.5% (1 out of 8) and 15.4% (2 out of 13), respectively. Similarly, Fisheries and Aquaculture, and Geography showed low pass rates of 33.3% (4 out of 12) and 38.5% (5 out of 13). Business Administration department also struggled, with only 35.1% (13 out of 37) of students passing. Students at these departments may benefit from tailored interventions, such as enhanced language instruction and additional support tailored to their struggles.

Furthermore, students in Food Engineering department displayed a mixed outcome, with a pass rate of 35.9% (14 out of 39), indicating potential gaps in student preparation or engagement. Physics Engineering students similarly exhibited a modest pass rate of 27.8% (5 out of 18), reflecting challenges that require attention.

The "Other" category, encompassing a range of smaller departments, accounted for the largest number of participants (86 students) but recorded a relatively low pass rate of 36.0% (31 out of 86). The variability within this category suggests the need for a deeper analysis to identify specific departments that require additional support.

4.1.3 Descriptive Results Related to Vocational Schools and Faculties

The data presented in Table 4.3 illustrates the performance of students categorized by vocational schools and faculties in the English proficiency exam. Out of a total of 481 students, the majority, 470 students (97.7%), were enrolled in faculties, while only 11 students (2.3%) were from vocational schools. When combined, the data reveals an overall pass rate of 47.6% across both groups, with 229 students passing and 252 failing the exam.

Table 4.3. Achievement According to Vocational School/Faculty

Vocation School/Faculty	Fail	Pass	Total
Vocational Schools	10	1	11
Faculties	242	228	470
Total	252	229	481

The results show a stark contrast in performance between the two groups. In Vocational Schools, only 1 student passed the English proficiency exam out of 11, resulting in a pass rate of 9.1% and a fail rate of 90.9%. At faculty level, among faculty students, 228 out of 470 passed, yielding a significantly higher pass rate of 48.5%, with a fail rate of 51.5%.

4.1.4 Descriptive Results Related to Language of Instruction at Designated Department

The data presented in Table 4.4 provides a detailed analysis of student performance in the English proficiency exam, categorized by the language of instruction students will follow in their academic programs after graduating from the English preparation school. Among the 481 students who participated, 99 were preparing for Turkish-medium programs, while 382 were preparing for English-medium programs.

Overall, 252 students failed and 229 passed, resulting in an overall pass rate of 47.6%.

Table 4.4. Achievement According to Language of Instruction

Language of Instruction	Fail	Pass	Total
Turkish	56	43	99
English	196	186	382
Total	252	229	481

Students preparing for English-medium programs demonstrated a slightly higher pass rate, with 186 out of 382 students passing, yielding a success rate of 48.7%. In contrast, students preparing for Turkish-medium programs had a pass rate of 43.4%, with only 43 out of 99 students succeeding.

4.1.5 Descriptive Results Related to Type of Instruction at Designated Department

The data presented in Table 4.5 provides an analysis of student performance in the English proficiency exam based on the type of instruction their academic programs will follow after completing the English preparation school. Importantly, students who did not meet the required proficiency level or failed to reach the English proficiency exam by the end of two academic semesters were also labeled as fail. This adds another layer to the performance outcomes, as it highlights the critical nature of timely achievement in the language preparation school.

Table 4.5. Achievement According to Type of Instruction

Type of Instruction	Fail	Pass	Total
Distance	17	5	22
Face to Face	235	224	459
Total	252	229	481

Students preparing for DE programs demonstrated lower pass rates compared to those preparing for face-to-face programs. Of the 22 students in the DE category, only 5 passed, resulting in a pass rate of 22.7%, while the remaining 77.3% failed. Conversely, students preparing for face-to-face programs achieved a 48.8% pass rate, with 224 out of 459 students passing.

4.1.6 Descriptive Results Related to Gender

The data presented in Table 4.6 provides an analysis of student performance in the English proficiency exam based on gender. The results highlight distinct patterns of achievement between female and male students, shedding light on potential gender-related differences in language proficiency outcomes.

Table 4.6. Achievement According to Genders

Gender	Fail	Pass	Total
Female	129	142	271
Male	123	87	210
Total	252	229	481

Female students demonstrated higher levels of success rates in the exam. Out of 271 female students, 142 passed, yielding a pass rate of 52.4%. The remaining 129 students failed, resulting in a fail rate of 47.6%. In contrast, male students recorded a lower pass rate of 41.4%, with only 87 out of 210 students passing and 123 failing.

This disparity may suggest that female students were better prepared for the exam or more engaged in their language studies, leading to superior performance.

4.2 Findings Related to Research Question 1

RQ1: Does engagement in LMS and virtual classrooms in the first quarter of the academic year predict academic performance in future quarters?

4.2.1 Findings Related to Sub-Research Question 1.1

RQ1.1: Does engagement in LMS and virtual classrooms in the first quarter predict academic performance on gateway exams in the first quarter, the second quarter, the third quarter, and the proficiency exam in the fourth quarter?

Table 4.7. Overall Performance Scores of Classification Algorithms

Model	AUC	CA	F1	Prec.	Recall	MCC
Logistic Regression	0.693	0.547	0.462	0.448	0.547	0.279
Neural Network	0.675	0.518	0.459	0.428	0.518	0.240
Naïve Bayes	0.664	0.443	0.427	0.426	0.443	0.205
SVM	0.660	0.547	0.457	0.419	0.547	0.275
Random Forest	0.649	0.468	0.411	0.379	0.468	0.158
k-NN	0.633	0.501	0.436	0.399	0.501	0.204
Gradient Boosting	0.631	0.464	0.417	0.395	0.464	0.156
SGD	0.598	0.536	0.429	0.419	0.536	0.269
CN2 rule inducer	0.574	0.351	0.352	0.355	0.351	0.081
Tree	0.569	0.391	0.379	0.369	0.391	0.101
AdaBoost	0.551	0.372	0.374	0.375	0.372	0.103

The analysis of first-quarter engagement metrics from the LMS and virtual classrooms demonstrated their moderate effectiveness in predicting academic performance in gateway and proficiency exams across all four quarters of the academic year. Machine learning models, including Logistic Regression, Neural Networks, Naïve Bayes, and Support Vector Machines (SVM), were employed to evaluate the predictive potential of these engagement metrics.

Logistic Regression exhibited the highest Area under the Curve (AUC) value at 0.693, Neural Networks followed by (0.675) and Naïve Bayes (0.664). Decision Trees, while more interpretable, showed a comparatively lower AUC of 0.569. Logistic Regression achieved the highest Matthews Correlation Coefficient (MCC) of 0.279, outperforming Neural Networks (0.240) and Decision Trees (0.101). These results are consistent with previous studies highlighting the superior performance of Logistic Regression and Neural Networks in educational contexts (Cheewaprakobkit, 2015).

The precision and recall metrics further supported these findings. Logistic Regression achieved a recall of 0.547, Neural Networks slightly trailing at 0.518, with both models outperforming Decision Trees (0.391). Precision scores for Logistic Regression and Neural Networks were similarly aligned at 0.448 and 0.428, respectively, underscoring their moderate predictive capabilities. These metrics suggest that engagement data such as login frequency, assignment submissions, and interaction with digital learning materials are effective indicators of academic outcomes but not definitive predictors considering comparably low prediction accuracy in this example (Junco & Clem, 2015). In this regard, the results of algorithms and models highlight the potential of using LMS and virtual classroom engagement metrics for early academic performance prediction to some extent, confirming the validity of established machine learning models in this context.

Table 4.8. Confusion Matrix of Classification Algorithm for Early Prediction

	Logistic Regression					
	Q1F	Q2F	Q3F	Q4F	P	Total
Q1F	40	10	0	0	18	68
Q2F	14	15	0	0	66	95
Q3F	3	7	0	0	26	36
Q4F	1	3	0	1	48	53
P	8	11	0	3	207	229
Total	66	46	0	5	365	481

The classification model including logistic regression classifier was applied to predict student outcomes across multiple categories, including quarterly failures (Q1F, Q2F, Q3F, Q4F) and passing (P) status in the English proficiency exam. The study utilized 10-fold cross-validation to ensure robust evaluation of the model. This method divided the dataset into ten subsets, with each subset serving as a test set while the remaining nine subsets were used for training, cycling through all subsets. This approach reduced variability and improved the reliability of the model's performance metrics. Additionally, data normalization was conducted as part of the preprocessing to standardize feature scales, preventing any single variable from disproportionately influencing the model.

The confusion matrix provided critical insights into the model's performance. The logistic regression demonstrated substantial success in predicting the "Pass" (P) category, where 207 of 229 actual instances were correctly classified, achieving an accuracy rate of 90.4%. However, its performance in classifying failure categories was inconsistent. For instance, in Q1F, only 40 out of 68 instances were correctly classified (58.8%), while in Q2F, just 15 out of 95 instances were accurately predicted (15.8%). A significant issue was the systemic bias toward predicting positive outcomes, as evidenced by the high rate of misclassification of failure instances as P or other failure categories, particularly in Q3F (100.0% misclassified) and Q4F (98.7% misclassified). These results reflected similar challenges in LA, where algorithmic models often struggle to capture nuanced student behaviors, particularly in underrepresented categories.

Normalization and cross-validation were critical to ensuring that the model operated efficiently and yielded reliable results. Within the context of LA, these practices align intending to create scalable and generalizable systems for predicting student outcomes. LA emphasizes the integration of data from various sources to inform decision-making processes and improve educational experiences. By employing normalization, the logistic regression model aligned feature values to a uniform scale, reducing potential biases during training. Similarly, 10-fold cross-validation enhanced the credibility of the findings by minimizing the effects of overfitting, a common concern in predictive analytics for education.

Despite these methodological strengths, this section of the study highlighted areas where the logistic regression model required improvement. The inability to accurately classify failure categories reduced its utility for early interventions, which are central to the objectives of LA. Addressing these limitations would involve incorporating more behavioral and interactional data into the model to capture a holistic view of student performance. Moreover, adopting hybrid or ensemble approaches that combine multiple algorithms could further reduce biases and improve the robustness of predictions across all categories.

In conclusion, the logistic regression model, supported by 10-fold cross-validation and data normalization, provided valuable insights into student performance prediction. While it excelled in identifying students likely to pass, it demonstrated limitations in classifying failure categories. These findings underscore the need for further refinements, such as integrating more dynamic behavioral data and leveraging advanced modeling techniques to improve prediction accuracy. The results align with broader trends in LA, emphasizing the importance of blending methodological rigor with an understanding of the complexities of student learning behaviors to enhance predictive capabilities (Kovačić, 2012; Perrotta & Williamson, 2018).

4.2.2 Findings Related to Sub-Research Question 1.2

RQ1.2: What are the most predictive features that contributed to the best-performing model for the gateway exams and the proficiency exam prediction?

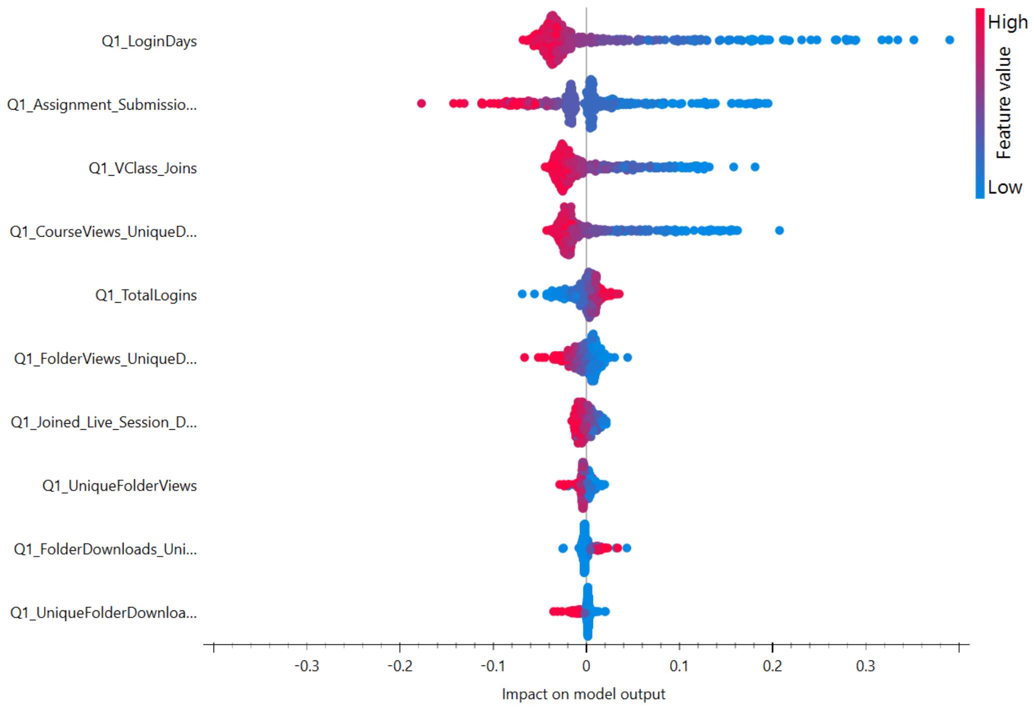

Figure 4.1. Explanation of the Model for Gateway and Proficiency Exams in the Early Warning System (Q1F Target)

The model for the Q1F target demonstrated comparably good performance, correctly predicting 40 out of 68 instances (58.8%) for predicting failure in the first quarter. In Figure 4.1., the SHAP summary plot provided an analysis of the impact of various features on a machine learning model used to predict Q1F (failure in the first quarter) in a gateway and proficiency exam early warning system. The analysis demonstrated that behavioral and engagement-related features were critical in determining the likelihood of failure. Among these, Q1_LoginsDays emerged as the most influential feature. Higher values of this feature (indicated by red markers) were associated with a reduced risk of Q1F, while lower values (blue markers) increased the likelihood of failure. This finding was consistent with prior research in LA, which had highlighted the role of consistent course engagement in improving academic success.

Other important features included Q1_Assignment_Submissions and Q1_VClass_Joins, which reflected students' participation and academic activity. Higher submission rates and frequent virtual classroom joins (red markers) correlated with a lower risk of Q1F, while low participation and submission rates (blue markers) increased failure probabilities. These results aligned with LA studies that had emphasized student activies and submission behaviors as key predictors of student success (Kokoç, Akçapınar, & Hasnine, 2021; Kovačić, 2012; Amrieh et al., 2015). Similarly, Q1_CourseViews_UniqueDays and Q1_FolderViews_UniqueDays, which captured overall engagement with the platform, showed that students who accessed course material more frequently were less likely to fail. These observations supported previous findings about the importance of sustained interaction with educational platforms in reducing students' risk of failure (Delen, 2011).

Features such as Q1_Joined_Live_Session_Duration and Q1_UniqueFolderViews also contributed positively to the predictions. Students who consistently accessed and utilized educational resources had a lower risk of Q1F, while low engagement with these resources was linked to higher failure probabilities. Although smaller-impact features, such as Q1_UniqueFolderDownloads, had less overall influence on the predictions, they still offered insights into specific aspects of student behavior. These patterns were consistent with earlier discussions in LA, which had emphasized resource utilization as a critical component of academic performance in a long period of EDM and LA literature (Sukhija et al., 2015).

The findings had significant implications for the development of early warning systems in education. The analysis demonstrated that features such as virtual class attendance, assignment submissions, and login activity could effectively identify students at risk of failure. These insights underscored the importance of designing targeted interventions to increase engagement and participation. For example, automated alerts could be generated for students with low interaction levels, enabling educators to intervene before students encountered significant challenges. This may foster constructive stress that promotes greater engagement with the LMS. These strategies align with emphases that value behavioral metrics in predicting and improving academic outcomes.

In conclusion, the SHAP analysis provided a clear understanding of the behavioral and engagement features that influenced the likelihood of Q1F. The findings reinforced earlier studies in LA, which had demonstrated the importance of active participation, resource usage, and consistent engagement in improving academic performance. By utilizing these insights, educational institutions were better equipped to design and implement effective early warning systems to reduce failure rates and improve student success during critical early academic periods.

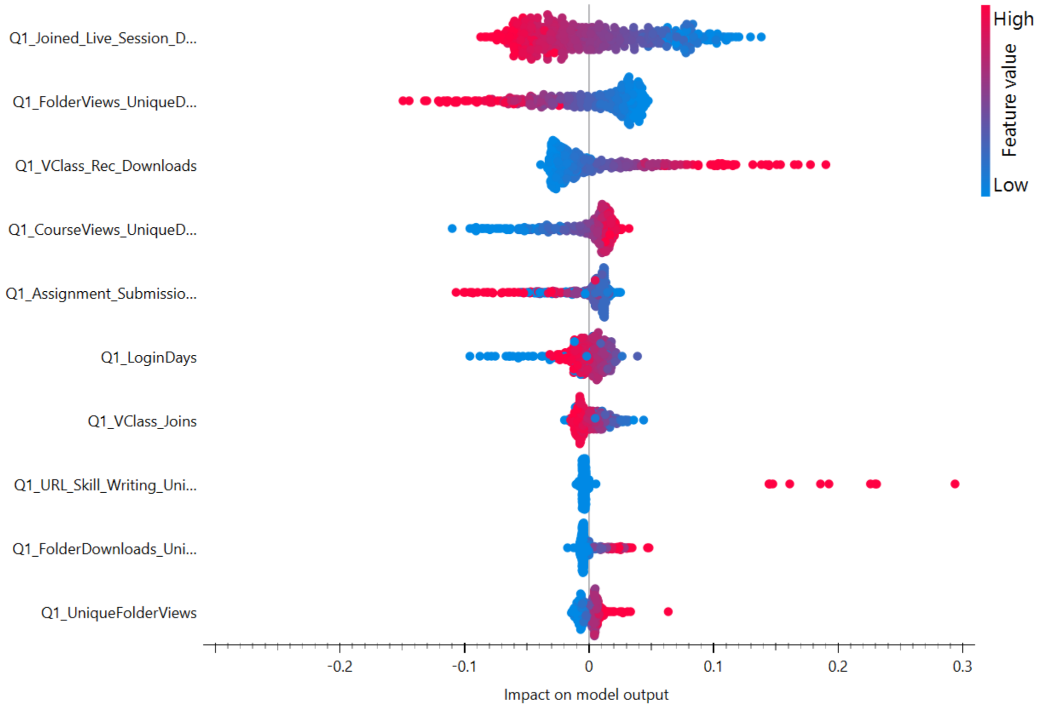

Figure 4.2. Explanation of the Model for Gateway and Proficiency Exams in the Early Warning System (Q2F Target)

The model for the Q2F target demonstrated poor performance, correctly predicting 15 out of 95 instances (15.8%) for predicting failure in the second quarter. In Figure 4.2., the SHAP summary plot provided an analysis of the impact of various features on a machine learning model used to predict Q2F (failure in the second quarter) in a gateway and proficiency exam early prediction. Features such as; Q1_Joined_Live_Session_Duration and Q1_FolderViews_UniqueDays, were critical in determining the percentage of failure rate.

The findings align with previous research indicating that algorithmic adjustments and preprocessing improvements, such as data balancing, feature engineering or tool selection, are critical for enhancing predictive outcomes (Mishra et al., 2014; Naik & Samant, 2016).

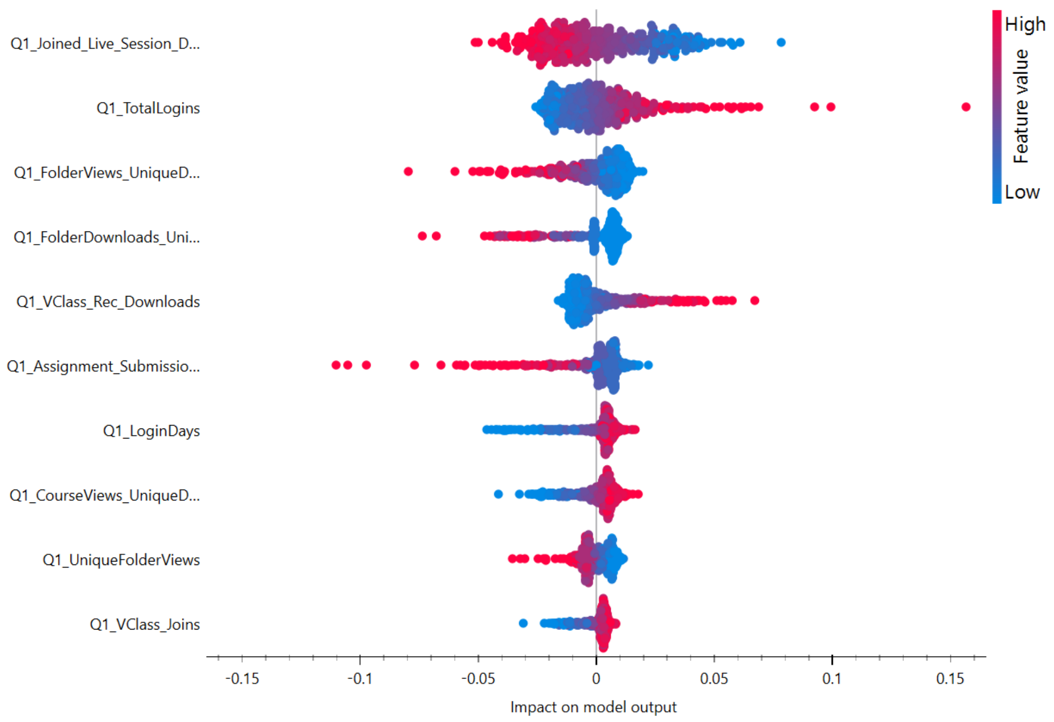

Figure 4.3. Explanation of the Model for Gateway and Proficiency Exams in the Early Warning System (Q3F Target)

For the Q3F target, results exhibited no predictive accuracy, incorrectly identifying all instances. The low accuracy suggested challenges in the model's ability to generalize effectively, likely due to suboptimal feature engineering or imbalanced data. To improve performance, researchers recommended refining feature selection, balancing the dataset, and employing ensemble algorithms to better capture complex patterns. These observations aligned with previous studies that emphasized the importance of robust predictors and algorithmic optimization in early warning systems (Leppänen et al., 2017; Kabra & Bichkar, 2011; Vandamme et al., 2007).

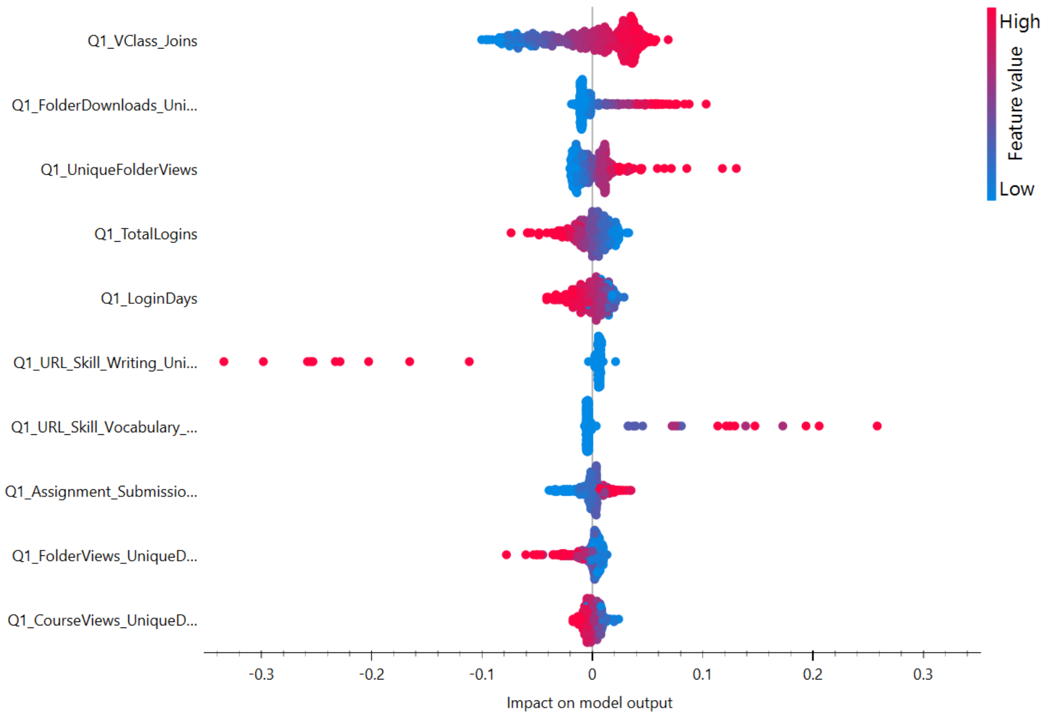

Figure 4.4. Explanation of the Model for Gateway and Proficiency Exams in the Early Warning System (Q4F Target)

For the Q4F target, the model demonstrated very limited predictive performance, accurately identifying only 1 out of 53 instances (1.9%). This low accuracy highlighted fundamental challenges in the model's structure and its ability to generalize from the data. These challenges might stem from an imbalanced dataset, inadequate feature selection, or an inappropriate algorithm. The findings suggest a need for refinement, including better feature engineering, balanced data, and possibly adopting more robust predictive models such as ensemble techniques or neural networks. Similar limitations in early warning systems have been addressed in previous studies by focusing on improving predictors and leveraging more suitable machine learning approaches (Leppänen et al., 2017; Cheewaprakobkit, 2015; Vandamme et al., 2007).

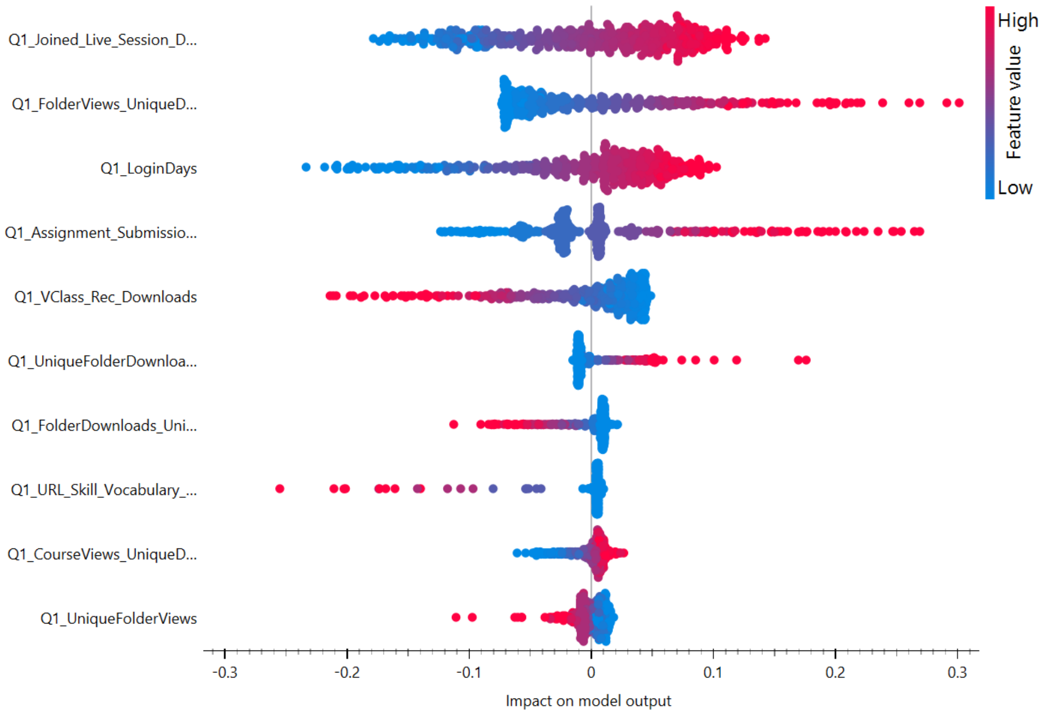

Figure 4.5. Explanation of the Model for Gateway and Proficiency Exams in the Early Warning System (Pass Target)

On the other hand, for the Pass target, the model demonstrated notable predictive performance, accurately predicting 207 out of 229 instances (90.4%). This high accuracy indicates the model's effectiveness in leveraging engagement metrics to identify students likely to succeed. The feature importance analysis provided key insights into the predictors of success, with several metrics standing out as significant contributors to the model's output.

Among the features, Q1_Joined_Live_Session_Duration and Q1_FolderViews_UniqueDays emerged as the most influential. These metrics reflect students' sustained interaction with course materials and consistent paritipation in virtual classrooms, both of which are associated with strong academic outcomes. Frequent attendance likely indicates higher levels of student engagement and time spent in the learning environment, while unique interaction days with course materials suggest active exploration and usage of available resources. In this regards, live session participation was a critical factor, with longer durations positively correlating with success. This finding underscores the importance of interactive and synchronous learning experiences, which provide opportunities for direct instructor and peer engagement.These findings are consistent with existing research, which highlights the importance of continuous engagement in predicting academic success (Kahu & Nelson, 2018).

Other significant contributors included Q1_LoginDays and Q1_Assignment_Submission. The first highlights the value of consecutive logins into LMS, suggesting that students, who more frequently login, perform better. Similarly, Q1_UniqueFolderDownloads was an influential factor, positively correlating with student achievement in online English learning. On the contrary, features such as Q1_CourseViews_UniqueDays showed minimal impact, indicating that these activities are less critical to achieving the pass target.

Despite its strong performance, the model leaves room for improvement. Features with low contributions, such as Q1_VClass_Rec_Downloads, Q1_FolderDownloads_UniqueDays and Q1_UniqueFolderViews, could be revisited or excluded to streamline the model and reduce noise. Additionally, refining feature selection by incorporating temporal metrics, such as the timing of engagement closer to exams, could further improve its predictive capabilities. Ensuring that data is balanced and exploring ensemble methods could also enhance the model's robustness and applicability across different student populations.

The implications of this analysis align with prior studies that emphasize the critical role of engagement and interaction metrics in predicting academic success. The model's strong predictive performance demonstrates its potential as a valuable tool in early warning systems, enabling educators to proactively identify and support students if further improvements can be employed for failing students. By focusing on key engagement metrics, such systems can inform targeted interventions to ensure better academic outcomes for all students and stakeholders.

4.3 Findings Related to Research Question 2

RQ2: Does engagement in LMS and virtual classrooms in different quarters of the academic year predict academic performance across gateway exams and proficiency exams?

4.3.1 Findings Related to Sub-Research Question 2.1.1

RQ2.1.1: Does engagement in LMS and virtual classrooms predict academic performance on the gateway exam in the first quarter?

When classification algorithms are applied to normal datasets, their performance can be influenced by class imbalances, leading to biases toward majority classes. By employing SMOTE, the data is balanced by generating synthetic examples for the minority class, allowing algorithms to better learn the decision boundaries. This sometimes results in improved performance metrics such as AUC, F1 Score, and recall, as seen in the comparison of normal and SMOTE-enhanced data, where models like Logistic Regression and SVM demonstrated notable gains in handling imbalanced class distributions effectively.

Table 4.9. Performance Scores of Classification Algorithms Using SMOTE for Gateway Exam in the First Quarter

Model	AUC	CA	F1	Prec.	Recall	MCC
Logistic Regression	0.847	0.807	0.824	0.866	0.807	0.513
k-NN	0.847	0.731	0.761	0.853	0.731	0.434
Naive Bayes	0.843	0.834	0.841	0.851	0.834	0.488
SVM	0.828	0.841	0.848	0.860	0.841	0.518
Gradient Boosting	0.825	0.821	0.821	0.821	0.821	0.391
Neural Network	0.799	0.524	0.572	0.797	0.524	0.210
Tree	0.784	0.807	0.810	0.813	0.807	0.363
SGD	0.754	0.793	0.809	0.839	0.793	0.436
Random Forest	0.743	0.807	0.810	0.813	0.807	0.363
AdaBoost	0.739	0.793	0.807	0.832	0.793	0.416
CN2 Rule Induction	0.571	0.297	0.266	0.857	0.297	0.170

The performance comparison of various classification algorithms applied to the Gateway Exam data in the first quarter using the post-SMOTE training data revealed interesting insights. According to Table 4.9., Logistic Regression and k-NN achieved the highest AUC score of 0.847, highlighting their strong discriminative ability between classes. Regarding classification accuracy, Naive Bayes stood out with a score of 0.834, slightly falling behind Support Vector Machines (SVM) at 0.841. In terms of F1 Score, which balances precision and recall, SVM emerged as the top performer with a value of 0.848, followed closely by Logistic Regression.

Precision was highest for Logistic Regression, achieving a remarkable 0.866, indicating its effectiveness in minimizing false positives. SVM, on the other hand, excelled in recall (0.841), showcasing its ability to correctly identify true positives, while Logistic Regression followed closely with a recall score of 0.807. The Matthews Correlation Coefficient (MCC), which evaluates the overall quality of binary classifications, was also highest for SVM (0.518), reflecting its balanced performance across all metrics.

Despite having a relatively low AUC (0.799) and classification accuracy (0.524), the Neural Network algorithm demonstrated a high recall (0.797), suggesting a tendency to favor positive predictions. However, CN2 Rule Induction performed poorly across all metrics, with the lowest AUC (0.571), accuracy (0.297), and MCC (0.170), making it unsuitable for this dataset.

Overall, SVM consistently outperformed other algorithms with strong metrics in F1, recall, and MCC, making it a reliable choice for balanced predictions. Logistic Regression also proved to be a robust and practical option due to its simplicity and high precision. Naive Bayes performed reliably, especially in accuracy and AUC, while algorithms like Neural Networks and CN2 Rule Induction showed limitations in this context. The results highlight the importance of selecting algorithms based on task-specific metric priorities, such as recall for reducing false negatives or precision for minimizing false positives. The model incorporating Logistic Regression was selected for reporting, as it demonstrated superior performance in identifying students at risk of failure.

Table 4.10. Confusion Matrix of Classification Algorithm Using SMOTE for Gateway Exam in the First Quarter

	Logistic Regression		
	F	P	Total
F	21	5	26
P	23	96	119
Total	44	101	145

The confusion matrix (Table 4.10) for Logistic Regression after applying SMOTE provides detailed insights into its classification performance. For Logistic

Regression, the model correctly classified 21 out of 26 negative instances (80.8%), while misclassifying 5 instances (19.2%) as positive. For the positive class, it correctly identified 96 out of 119 instances (80.7%), with 23 instances (19.3%) misclassified as negative. This indicates a balanced performance across both classes, with comparable accuracy for negatives and positives.

4.3.2 Findings Related to Sub-Research Question 2.1.2

RQ2.1.2: What are the most predictive features that contributed to the best-performing model in the prediction performance for the gateway exam in the first quarter?

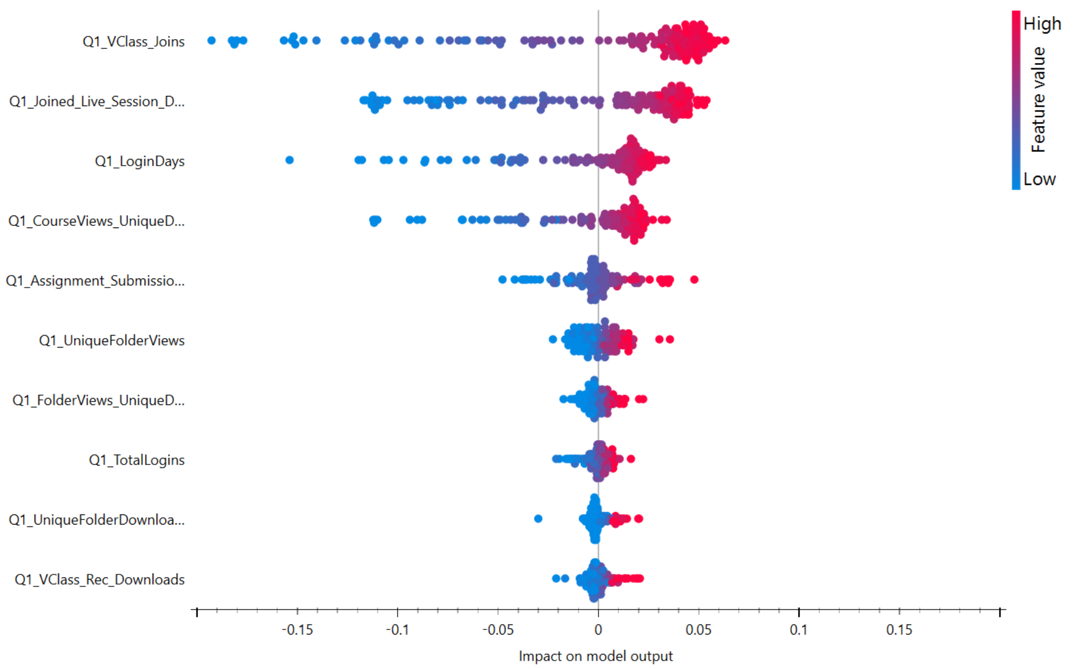

Figure 4.6. Explanation of the Model Using SMOTE for Gateway Exam in the First Quarter (Logistic Regression)

Figure 4.6 illustrates the SHAP summary plot, explaining the logistic regression model's predictions using SMOTE for the Gateway Exam in the first quarter. The most influential feature was Q1_VClass_Joins, where higher values (indicated by red dots) positively impact the model's predictions, showing that increased virtual class participation correlates with better outcomes. Similarly, Q1_Joined_Live_Session_Days also played a significant role, emphasizing the importance of live session engagement. Moderate contributions were observed from features like Q1_LoginDays and Q1_CourseViews_UniqueD, where higher login days and unique course views positively affected predictions. On the other hand, features such as Q1_TotalLogins, Q1_UniqueFolderDownloads, and Q1_VClass_Rec_Downloads had a lower impact, with some negative contributions for lower feature values (blue dots), indicating less direct influence on the outcome. Notably, Q1_Assignment_Submissions showed a clear positive effect for higher values, reinforcing the role of academic engagement. Overall, the plot highlights that active participation metrics, particularly those involving live sessions and interactions, were key drivers of model predictions, while features like downloads and total logins had a more limited influence. This underscores the significance of engagement-focused activities in predicting student success, with SMOTE effectively balancing the dataset to make these insights more apparent.

4.3.3 Findings Related to Sub-Research Question 2.2.1

RQ2.2.1: Does engagement in LMS and virtual classrooms predict academic performance on the gateway exam in the second quarter?

Table 4.11. Performance Scores of Classification Algorithms for Gateway Exam in the Second Quarter

Model	AUC	CA	F1	Prec.	Recall	MCC
Logistic Regression	0.780	0.835	0.818	0.825	0.835	0.480
Neural Network	0.776	0.811	0.794	0.794	0.811	0.402
Naïve Bayes	0.774	0.746	0.758	0.781	0.746	0.373
Random Forest	0.756	0.809	0.796	0.793	0.809	0.405
Gradient Boosting	0.740	0.782	0.762	0.758	0.782	0.303
SVM	0.735	0.814	0.781	0.802	0.814	0.383
k-NN	0.709	0.792	0.769	0.768	0.792	0.324
CN2 rule inducer	0.660	0.714	0.716	0.718	0.714	0.205
AdaBoost	0.652	0.765	0.761	0.758	0.765	0.315
SGD	0.646	0.818	0.790	0.807	0.818	0.406
Tree	0.637	0.765	0.760	0.756	0.765	0.310

The analysis of whether engagement in LMS and virtual classrooms predicted academic performance on the gateway exam in the second quarter highlighted the effectiveness of several machine learning algorithms. In Table 4.11., Logistic Regression demonstrated the highest performance with an AUC of 0.780, classification accuracy (CA) of 0.835, and recall of 0.835. These results underscored its ability to predict academic outcomes with high precision and sensitivity. Neural Networks and Random Forest also exhibited strong performances, with AUCs of 0.776 and 0.756, respectively, indicating their reliability in predicting student success. Notably, Naïve Bayes achieved an AUC of 0.774 while maintaining a balance between type I and type II errors, which was particularly valuable in contexts where minimizing both false positives and false negatives equitably was essential (Cheewaprakobkit, 2015; You, 2016).

Naïve Bayes' balance of error types ensured that predictions did not disproportionately favor or penalize specific groups of students. This characteristic was especially significant in academic evaluations, where false positives (students predicted to succeed but failed) and false negatives (students predicted to fail but succeeded) could have substantial consequences. While Logistic Regression achieved the highest accuracy and recall, its focus on precision might not have offered the same fairness across error types as Naïve Bayes. On the other hand, Gradient Boosting, with an AUC of 0.740, and other complex models like SVM and k-NN, demonstrated lower effectiveness compared to simpler models like Logistic Regression and Naïve Bayes. These findings emphasized the importance of evaluating models not only by accuracy but also by their balance in error rates.

The comparative evaluation of these algorithms reinforced the value of model selection based on specific academic contexts and requirements. Logistic Regression remained the most accurate, but Naïve Bayes' ability to manage errors made it a compelling choice for settings where fairness in predictive outcomes was prioritized. Gradient Boosting and other complex models, while often effective in other domains, demonstrated limitations in this context, reinforcing the need for careful evaluation of their applicability to academic performance prediction tasks (Perrotta & Williamson, 2018; Kabra & Bichkar, 2011).

Table 4.12. Confusion Matrix of Classification Algorithm for Gateway Exam in the Second Quarter

	Naïve Bayes		
	F	P	Total
F	60	35	95
P	70	248	318
Total	130	283	413

The confusion matrix (Table 4.12.) for the classification algorithm provided additional insights into the performance of Naïve Bayes in predicting academic performance on the gateway exam in the second quarter. The model including Naïve Bayes classifier demonstrated a balanced approach. It correctly classified 248 students as passers (82.8%) and 60 as failers (76.5%). However, it misclassified 70 students who passed as failers (17.2%) and 35 failers as passers (23.5%). The algorithm's ability to capture a balance between type I and type II errors was evident in these results. Naïve Bayes effectively reduced false positives and negatives, a crucial aspect in academic evaluations to ensure fair and equitable prediction outcomes.

4.3.4 Findings Related to Sub-Research Question 2.2.2

RQ2.2.2: What are the most predictive features that contributed to the best-performing model in the prediction performance for the gateway exam in the second quarter?

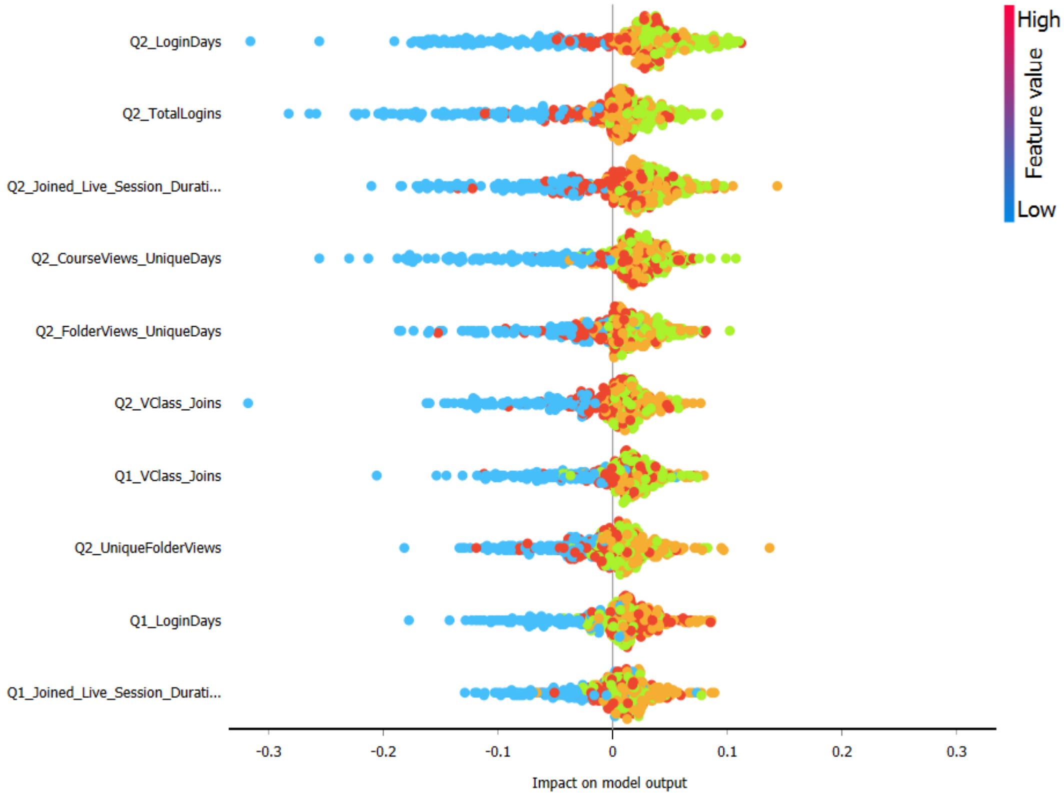

Figure 4.7. Explanation of the Model for Gateway Exam in the Second Quarter (Naïve Bayes)

The analysis of the Naïve Bayes model for predicting performance in the gateway exam in the second quarter, as illustrated in Figure 4.7., highlights the top features and their contributions to the model's predictions.

The feature with the highest impact was Q2_LoginDays, representing the number of days students logged into the LMS during the second quarter. Students with frequent logins were positively associated with better academic outcomes, emphasizing the importance of consistent engagement with digital learning platforms. Similarly, Q2_TotalLogins, which captured the total login frequency, further reinforced this trend, demonstrating the predictive power of cumulative engagement metrics.

Another key feature was Q2_Joined_Live_Session_Duration, which reflects the time students spent in live sessions during the second quarter. Higher participation in live sessions positively influenced the model's predictions, underscoring the importance of synchronous learning environments in supporting academic performance (Banna et al., 2015) Additionally, features related to material access, such as Q2_CourseViews_UniqueDays and Q2_FolderViews_UniqueDays, highlighted the value of diverse and sustained access to course resources.

Participation metrics, including Q2_VClass_Joins and Q1_VClass_Joins, indicated the significance of active engagement in virtual classrooms over both quarters. Interestingly, Q1_Joined_Live_Session_Duration, which measured live session duration in the first quarter, had a relatively lower positive impact compared to its second-quarter counterpart, suggesting that second-quarter engagement was more crucial for success on the gateway exam.

Finally, features like Q2_UniqueFolderViews and Q1_LoginDays demonstrated the continued importance of accessing diverse materials and maintaining regular engagement across the academic timeline. The balance of feature impacts and their directional effects in the Naïve Bayes model indicated its ability to equitably capture contributions from various aspects of student behavior, further emphasizing its strength in balancing type I and type II errors (Cheewaprakobkit, 2015).

These findings align with the principles of LA, which emphasize the use of diverse engagement metrics to predict academic success. Unlike Logistic Regression, which focused more on precision, the Naïve Bayes model provided a more balanced perspective, ensuring equitable treatment of students at risk of failing or succeeding in the gateway exam.

4.3.5 Findings Related to Sub-Research Question 2.3.1

RQ2.3.1: Does engagement in LMS and virtual classrooms predict academic performance on the gateway exam in the third quarter?

Table 4.13. Performance Scores of Classification Algorithms Using SMOTE for Gateway Exam in the Third Quarter

Model	AUC	CA	F1	Prec.	Recall	MCC
Neural Network	0.810	0.896	0.896	0.896	0.896	0.442
Logistic Regression	0.773	0.771	0.808	0.872	0.771	0.276
SVM	0.760	0.365	0.429	0.911	0.365	0.202
Random Forest	0.745	0.844	0.857	0.875	0.844	0.323
Gradient Boosting	0.741	0.917	0.913	0.910	0.917	0.514
Naive Bayes	0.734	0.917	0.913	0.910	0.917	0.514
Tree	0.733	0.917	0.913	0.910	0.917	0.514
k-NN	0.580	0.438	0.527	0.838	0.438	0.067
AdaBoost	0.578	0.323	0.384	0.870	0.323	0.112
SGD	0.545	0.740	0.777	0.827	0.740	0.067
CN2 Rule Induction	0.535	0.167	0.138	0.907	0.167	0.088

The performance of classification algorithms (Table 4.13.) using the SMOTE dataset for the gateway exam in the third quarter highlights distinct variations in their effectiveness. The Neural Network stood out as the top performer, achieving the highest AUC (0.810) and balanced metrics across accuracy (0.896), F1 score (0.896), precision (0.896), and recall (0.896), though its MCC (0.442) indicates moderate consistency. Similarly, Gradient Boosting, Naive Bayes, and Tree exhibited strong and identical performance, achieving high accuracy (0.917), F1 score (0.913), precision (0.910), and recall (0.917), along with the highest MCC scores (0.514), indicating robust predictive reliability. Logistic Regression also demonstrated a balanced prediction approach, with an AUC of 0.773, accuracy of 0.771, and an F1 score of 0.808. Its high precision (0.872) and reasonable recall (0.771) reflect a good balance in classifying both positive and negative outcomes, supported by an MCC of 0.276, which, while lower than the top performers, shows moderate consistency.

In contrast, algorithms like Random Forest showed solid performance but slightly lagged behind, with an accuracy of 0.844 and F1 score of 0.857, though its AUC (0.745) and MCC (0.323) were less impressive. On the other hand, k-NN, SGD, SVM, AdaBoost, and CN2 Rule Induction struggled significantly. k-NN and SGD showed limited effectiveness, with accuracies of 0.438 and 0.740, respectively, and low MCC values. SVM demonstrated high precision (0.911) but struggled with accuracy (0.365) and recall (0.365), leading to poor consistency (MCC of 0.202). AdaBoost and CN2 Rule Induction performed the weakest, with low accuracy (0.323 and 0.167, respectively) and poor generalization. Overall, Neural Network, Gradient Boosting, Naive Bayes, and Logistic Regression proved effective in balancing predictions and achieving high performance on the SMOTE-enhanced dataset, with Logistic Regression offering a relatively well-balanced approach across metrics. Since it proved to be the most effective in identifying failing students after training on the post-SMOTE dataset, the confusion matrix (Table 4.14.) of Gradient Boosting was reported.

Table 4.14. Confusion Matrix of Classification Algorithm using SMOTE for Gateway Exam in the Third Quarter

Gradient Boosting			
	F	P	Total
F	5	5	10
P	3	83	86
Total	8	388	96

The confusion matrix for Gradient Boosting in the third quarter highlights its classification performance when using the SMOTE-balanced data for the Gateway Exam. Gradient Boosting showed strong performance for the positive class, correctly classifying 83 out of 86 positive instances (96.5%), with only 3 instances (3.5%)

misclassified as negative. However, it struggled with the negative class, correctly identifying only 5 out of 10 instances (50.0%), while misclassifying the remaining 5 instances (50.0%) as positive.

4.3.6 Findings Related to Sub-Research Question 2.3.2

RQ2.3.2: What are the most predictive features that contributed to the best-performing model in the prediction performance for the gateway exam in the third quarter?

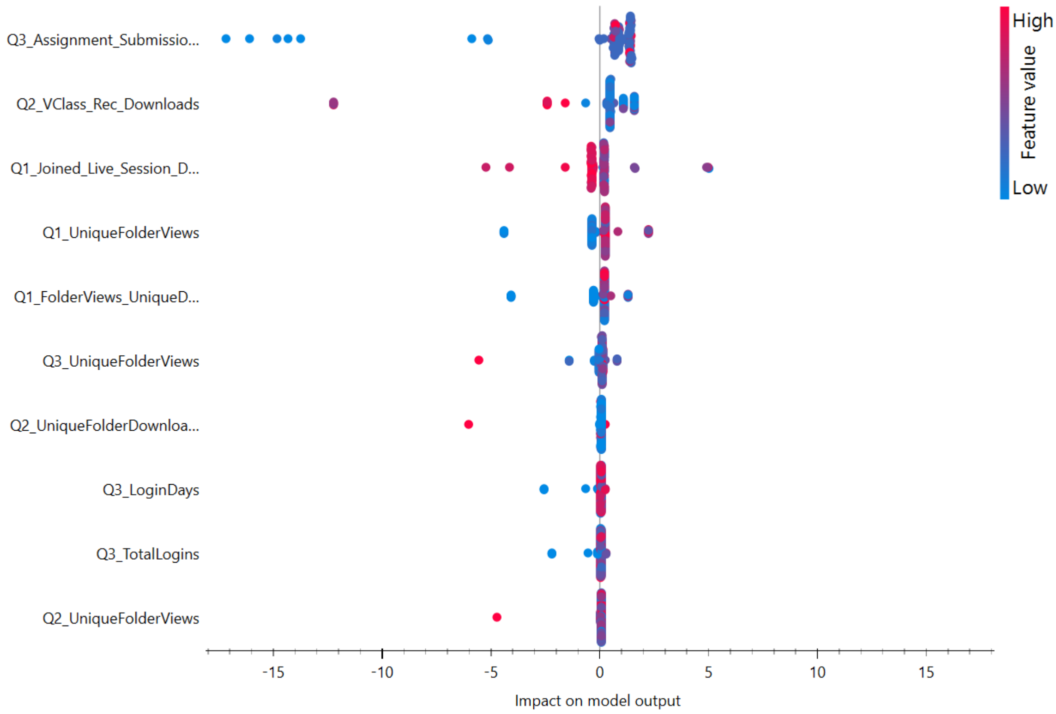

Figure 4.8. Explanation of the Model Using SMOTE for Gateway Exam in the Third Quarter (Gradient Boosting)

Figure 4.8. illustrated the SHAP summary plot for the Gradient Boosting model using the SMOTE-applied dataset for the Gateway Exam in the third quarter. Among the features, Q3_Assignment_Submissions, Q1_UniqueFolderViews, and Q3_LoginDays emerged as positively impactful, with higher feature values having strongly contributed to the model's positive predictions. These features highlighted the importance of student engagement through assignment submissions, accessing unique folders, and consistent logins during the third quarter. Conversely, Q1_Joined_Live_Session_Days acted as a negative predictor, where higher values reduced the likelihood of positive outcomes, suggesting potential limitations in how the model interpreted earlier live session participation.

Features such as Q2_VClass_Rec_Downloads, Q3_TotalLogins, and Q2_UniqueFolderDownloads appeared neutral, with their contributions remaining close to zero on the x-axis, reflecting minimal influence on the model's predictions. However, the SHAP plot was challenging to evaluate fully, as the overlapping dots and clustering around zero made it difficult to interpret the exact magnitude and direction of certain features' impacts. This ambiguity suggested that clearer visualization or additional feature-specific analyses might have been necessary to better understand the Gradient Boosting model's decision-making process.

4.3.7 Findings Related to Sub-Research Question 2.4.1

RQ2.4.1: Does engagement in LMS and virtual classrooms predict academic performance on the proficiency exam in the fourth quarter?

Table 4.15. Performance Scores of Classification Algorithms for Proficiency Exam in the Fourth Quarter

Model	AUC	CA	F1	Prec.	Recall	MCC
Random Forest	0.589	0.809	0.749	0.750	0.809	0.119
CN2 rule inducer	0.554	0.720	0.719	0.718	0.720	0.076
Naïve Bayes	0.544	0.511	0.562	0.724	0.511	0.073
Neural Network	0.541	0.762	0.725	0.700	0.762	0.016
Gradient Boosting	0.532	0.787	0.740	0.720	0.787	0.069
k-NN	0.526	0.794	0.719	0.657	0.794	-0.065
AdaBoost	0.490	0.677	0.683	0.689	0.677	-0.020
Logistic Regression	0.486	0.777	0.710	0.654	0.777	-0.092
SGD	0.486	0.766	0.714	0.678	0.766	-0.046
SVM	0.458	0.812	0.728	0.659	0.812	0.000
Tree	0.452	0.670	0.678	0.686	0.670	-0.028

The performance scores of classification algorithms (Table 4.15.) for predicting academic performance on the proficiency exam in the fourth quarter suggest a moderate relationship between engagement in LMS and virtual classrooms and academic success. The evaluation of the algorithms highlights both their strengths and limitations in leveraging engagement data for prediction.

Random Forest emerged as the top-performing model with an Area under the Curve (AUC) of 0.589 and a Classification Accuracy (CA) of 0.809. It also achieved a respectable F1-score of 0.749 and balanced precision (0.750) and recall (0.809), indicating its ability to effectively identify both high-performing and at-risk students. However, its Matthews Correlation Coefficient (MCC) of 0.119 suggests that its overall predictive reliability was limited.

The CN2 rule inducer model ranked second in performance with an AUC of 0.554 and a CA of 0.720. Its F1-score of 0.719 and precision (0.718) were slightly lower than Random Forest, reflecting a reasonable balance between positive and negative predictions. Its MCC of 0.076, however, indicates limited correlation between its predictions and actual outcomes.

Naïve Bayes demonstrated lower predictive performance, with an AUC of 0.544 and a CA of 0.511. While it achieved a relatively high precision of 0.724, its recall (0.511) and F1-score (0.562) suggest that it struggled to consistently identify true positives. Its MCC of 0.073 also highlights its limited effectiveness in leveraging engagement data for accurate predictions.

Other models, such as Neural Network, Gradient Boosting, and k-NN, showed moderate performance. Neural Network achieved an AUC of 0.541 and a CA of 0.762, with an F1-score of 0.725 and recall of 0.762. Gradient Boosting had slightly lower scores, with an AUC of 0.532 and CA of 0.787. k-NN achieved a CA of 0.794, but its precision of 0.657 and MCC of -0.065 indicate inconsistencies in its predictions.

Interestingly, Logistic Regression, traditionally a strong performer in predictive tasks, had a lower AUC of 0.486 and a CA of 0.777. Its MCC of -0.092 and precision of 0.654 highlight its limitations in this context. Similarly, SGD and SVM struggled, with SVM achieving the highest CA of 0.812 but a low MCC of 0.000.

The analysis suggests that while engagement in LMS and virtual classrooms provided valuable data for predicting academic performance, the predictive power of the algorithms was moderate overall. Random Forest stood out as the most reliable model, although its limited MCC suggests room for improvement. The varying performances across models highlight the complexity of using engagement data to predict academic outcomes, emphasizing the need for further refinement of algorithms and feature selection processes.

Table 4.16. Confusion Matrix of Classification Algorithm for Proficiency Exam in the Fourth Quarter

	Naïve Bayes		
	F	P	Total
F	32	21	53
P	117	112	229
Total	149	133	282

The confusion matrix (Table 4.16) for the Naïve Bayes algorithm in predicting proficiency exam performance at the end of the fourth quarter revealed moderate classification performance. Naïve Bayes demonstrated a balanced performance between predicting true negatives and true positives. It correctly identified 32 students (60.4%) of actual negatives as true negatives, showcasing a strong ability to classify low-performing students. However, Naïve Bayes misclassified 21 students (39.6%) of actual negatives as positives. For actual positives, Naïve Bayes correctly classified 112 students (48.9%) while misclassifying 117 students (51.1%) as negatives, indicating a struggle with accurately identifying high-performing students. This performance reflects Naïve Bayes' capacity to provide an even classification, though its ability to predict positive outcomes was weak.

Naïve Bayes' ability to detect true negatives makes it more suitable for scenarios where identifying and supporting at-risk students is a priority. The results underscore the importance of selecting an algorithm that aligns with the specific goals of the educational analysis.

4.3.8 Findings Related to Sub-Research Question 2.4.2

RQ2.4.2: What are the most predictive features that contributed to the best-performing model in the prediction performance for the proficiency exam in the fourth quarter?

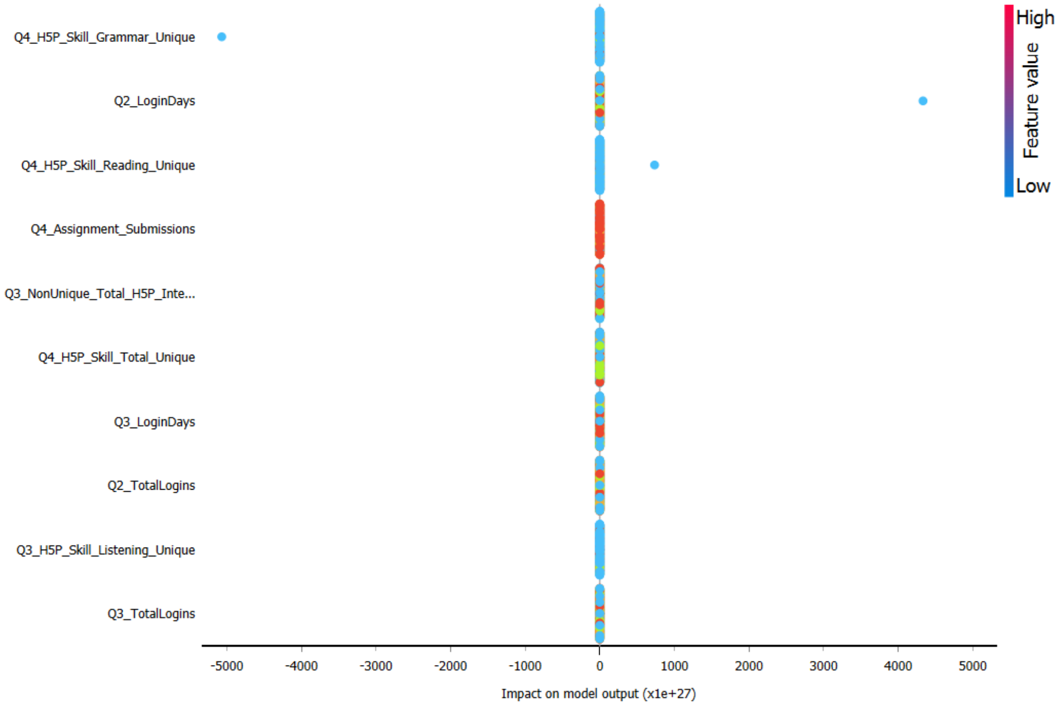

Figure 4.9. Explanation of the Model for Proficiency Exam in the Fourth Quarter (Naïve Bayes)

The visualization of the Naïve Bayes model for the proficiency exam in the fourth quarter (Figure 4.9.) shows that all features are centered, indicating that their impacts on the model's output are balanced around zero. This suggests that no single feature strongly influenced the predictions, and the contributions of all features were distributed evenly. The balanced nature of the feature impacts reflects the algorithm's inherent design to handle data probabilistically, without assigning disproportionately high weight to specific variables. As a result, the interpretation of individual feature importance is limited, emphasizing the model’s focus on overall patterns rather than dominant predictors.

4.4 Findings Related to Research Question 3

RQ3: Does engagement in LMS and virtual classrooms predict academic performance related to different English language skills in the second quarter?

4.4.1 Findings Related to Sub-Research Question 3.1.1

RQ3.1.1: Does engagement in LMS and virtual classrooms predict academic performance in the use of English skills (grammar/reading/listening) in the second quarter?

Table 4.17. Performance Scores of Classification Algorithms for the Use of English Skills (Grammar/Reading/Listening) in the Second Quarter

Model	AUC	CA	F1	Prec.	Recall	MCC
Logistic Regression	0.753	0.845	0.814	0.838	0.845	0.413
SVM	0.741	0.826	0.786	0.805	0.826	0.309
Random Forest	0.726	0.847	0.829	0.833	0.847	0.442
Neural Network	0.726	0.833	0.808	0.813	0.833	0.370
Naïve Bayes	0.688	0.738	0.754	0.779	0.738	0.292
Gradient Boosting	0.679	0.821	0.798	0.796	0.821	0.331
k-NN	0.652	0.806	0.776	0.772	0.806	0.251
CN2 rule inducer	0.643	0.738	0.738	0.738	0.738	0.171
SGD	0.639	0.840	0.814	0.825	0.840	0.397
AdaBoost	0.581	0.731	0.733	0.735	0.731	0.159
Tree	0.566	0.743	0.738	0.734	0.743	0.155

In Table 4.17., among the models, Logistic Regression achieved the highest AUC at 0.753, indicating its strong ability to differentiate between students' performance levels. This was closely followed by SVM with an AUC of 0.741, demonstrating a slightly lower but still effective level of predictive power. Random Forest and Neural Network both recorded AUCs of 0.726, highlighting their comparable performance in identifying patterns in engagement data relevant to English language skills.

Moderate-performing models included Naïve Bayes, which achieved an AUC of 0.688, and Gradient Boosting, which recorded an AUC of 0.679. These models showed reasonable predictive power but were outperformed by the top models. k-NN, with an AUC of 0.652, also fell into this mid-range category, showing limited effectiveness compared to higher-performing algorithms.

Lower-performing models such as CN2 Rule Inducer (AUC: 0.643), AdaBoost (AUC: 0.581), and Tree (AUC: 0.566) reflected weaker discrimination capabilities. These models struggled to effectively distinguish between performance levels in English language skills based on engagement data.

Table 4.18. Confusion Matrix of Classification Algorithm for the Use of English Skills (Grammar/Reading/Listening) in the Second Quarter

	Naïve Bayes		
	S	U	Total
Satisfactory	261	71	332
Unsatisfactory	37	44	81
Total	298	115	413

The confusion matrix for Naïve Bayes (Table 4.18.) in predicting academic performance in English language skills (grammar, reading, and listening) during the second quarter provides a detailed view of their classification capabilities. The performance of the model varied significantly in identifying students with satisfactory and unsatisfactory outcomes.

Naïve Bayes displayed a balanced, though low accurate, classification. For satisfactory outcomes, it correctly classified 261 students (78.6%) but misclassified 71 students (21.4%) as unsatisfactory. For unsatisfactory outcomes, it correctly identified 44 students (54.3%) and misclassified 37 students (45.7%) as satisfactory. While Naïve Bayes showed good effectiveness in identifying unsatisfactory outcomes compared to other algorithms, its overall performance for satisfactory students was moderate.

4.4.2 Findings Related to Sub-Research Question 3.1.2

RQ3.1.2: What are the most predictive features that contributed to the best-performing model in the prediction performance for the use of English skills (grammar/reading/listening) in the second quarter?

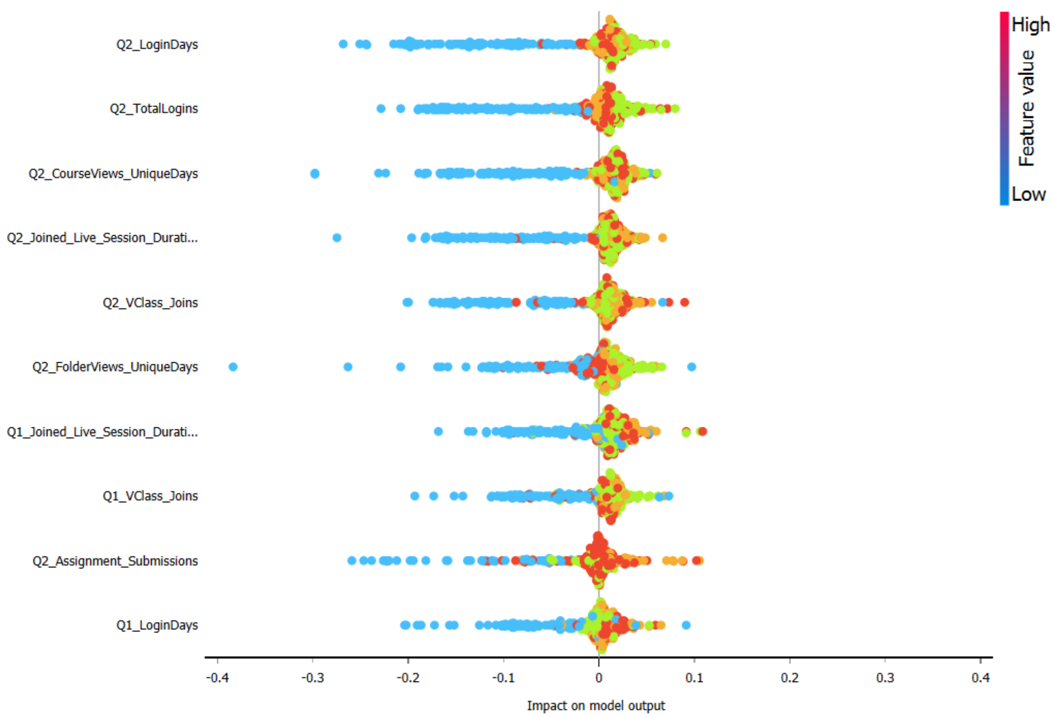

Figure 4.10. Explanation of the Model for the Use of English Skills (Grammar/Reading/Listening) in the Second Quarter (Naïve Bayes)

In the Naïve Bayes model (Figure 4.10.) for predicting English skills (grammar, reading, and listening) in the second quarter, all features contributed positively to the predictions, with varying levels of importance. Q2_LoginDays emerged as the most influential feature, indicating that frequent logins during the second quarter were closely tied to better performance in English skills. This highlights the importance of regular and consistent interaction with the LMS.

The second most important feature was Q2_CourseViews_UniqueDays, which demonstrated the value of sustained engagement with course content over unique days. This metric underscored that students who accessed course materials consistently were more likely to achieve better outcomes. Following this, Q2_Joined_Live_Session_Duration, the duration of live sessions attended during the second quarter, also ranked highly. This suggests that active participation in synchronous learning activities positively impacted academic performance.

Q2_VClass_Joins, representing the number of virtual class sessions joined in the second quarter, was another significant contributor. This feature emphasized the importance of direct participation in interactive instructional activities. Similarly, Q2_FolderViews_UniqueDays, which reflected the number of unique days students accessed folders, highlighted the benefit of engaging with diverse resources throughout the quarter.

Features from the first quarter also played an important role. Q1_Joined_Live_Session_Duration showed that participation in live sessions early in the academic year had a lasting positive effect on performance. Likewise, Q1_VClass_Joins demonstrated that early engagement with virtual class sessions contributed to sustained academic success. These findings reinforce the importance of building strong engagement habits early on.

Other impactful features included Q2_Assignment_Submissions, which stressed the value of completing assignments on time in the second quarter. This feature underscored the role of coursework in achieving strong academic outcomes. Q1_LoginDays, capturing login activity in the first quarter, also contributed positively, demonstrating that early and consistent engagement set the foundation for success. Lastly, Q1_FolderViews_UniqueDays, the number of unique days students accessed folders in the first quarter, further emphasized the long-term benefits of early interaction with learning resources.

In summary, the Naïve Bayes model showed that consistent and meaningful engagement, both early and ongoing, significantly contributed to performance in English language skills. Features such as login activity, assignment submissions, live session participation, and resource access all played a critical role in driving academic success, reflecting the value of sustained and purposeful engagement with LMS and virtual classroom resources.

4.4.3 Findings Related to Sub-Research Question 3.2.1

RQ3.2.1: Does engagement in LMS and virtual classrooms predict academic performance in writing skills in the second quarter?

Table 4.19. Performance Scores of Classification Algorithms for Writing Skills in the Second Quarter

Model	AUC	CA	F1	Prec.	Recall	MCC
Naive Bayes	0.729	0.678	0.681	0.686	0.678	0.325
Logistic Regression	0.719	0.678	0.663	0.665	0.678	0.271
Random Forest	0.703	0.685	0.677	0.676	0.685	0.300
k-NN	0.685	0.666	0.662	0.660	0.666	0.268
Neural Network	0.685	0.673	0.669	0.666	0.673	0.282
Gradient Boosting	0.674	0.685	0.678	0.676	0.685	0.301
SVM	0.667	0.673	0.664	0.663	0.673	0.272

CN2 rule inducer	0.644	0.625	0.623	0.621	0.625	0.185
SGD	0.639	0.690	0.678	0.679	0.690	0.302
AdaBoost	0.587	0.617	0.617	0.616	0.617	0.175
Tree	0.556	0.615	0.611	0.608	0.615	0.158

The performance of classification algorithms (Table 4.19.) in predicting academic performance in writing skills during the second quarter demonstrates the moderate predictive power of engagement in LMS and virtual classrooms. Among the evaluated models, Naïve Bayes emerged as the top-performing algorithm with an AUC of 0.729 and a Classification Accuracy (CA) of 0.678. Its F1-score of 0.681 and Matthews Correlation Coefficient (MCC) of 0.325 indicate its strong ability to capture patterns in the data compared to other models. The balance between precision (0.686) and recall (0.678) further highlights its effectiveness in predicting writing skill performance.

Gradient Boosting and Random Forest also performed well, both achieving a CA of 0.685. Gradient Boosting had an AUC of 0.674, an F1-score of 0.678, and an MCC of 0.301, while Random Forest recorded an AUC of 0.703, an F1-score of 0.677, and an MCC of 0.300. These models demonstrated strong recall values (0.685) and balanced performance, making them reliable predictors of writing skills.

SGD showed comparable performance, with the highest CA of 0.690 and an F1-score of 0.678. Its MCC of 0.302 reflected its reliability, despite a slightly lower AUC of 0.639. This model's strong recall (0.690) indicates its effectiveness in identifying true positives, though it slightly lagged in its overall discrimination power.

Models such as Logistic Regression and Neural Network also delivered solid performance. Logistic Regression had an AUC of 0.719, a CA of 0.678, and an MCC of 0.271, demonstrating balanced metrics for precision (0.665) and recall (0.678). Neural Network achieved an AUC of 0.685 and a CA of 0.673, with a respectable MCC of 0.282, underscoring its potential for predicting outcomes in writing skills.

Lower-performing models included SVM, which recorded an AUC of 0.667 and a CA of 0.673, and k-NN, with an AUC of 0.685 and a CA of 0.666. While these models showed moderate predictive capabilities, their MCC values of 0.272 and 0.268, respectively, indicated limitations in overall reliability. Models like CN2 Rule Inducer, AdaBoost, and Tree were less effective, with AUCs ranging from 0.556 to 0.644 and MCC values below 0.200.

In conclusion, engagement in LMS and virtual classrooms provided moderate predictive power for writing skills in the second quarter, with Naïve Bayes, Logistic Regression, Gradient Boosting, and Random Forest emerging as the most reliable models. These algorithms demonstrated a strong balance of accuracy, precision, and recall, highlighting the role of active and consistent engagement in predicting writing skill performance. However, the performance of lower-ranked models underscored the need for further refinement in feature selection and modeling to improve predictions in this domain.

Table 4.20. Confusion Matrix of Classification Algorithm for Writing Skills in the Second Quarter

	Naïve Bayes		
	S	U	Total
Satisfactory	186	75	261
Unsatisfactory	58	94	152
Total	244	169	413

The confusion matrix for Naïve Bayes (Table 4.20.) in predicting writing skills during the second quarter provides insights into the performance of these models. Both algorithms demonstrated distinct strengths and limitations in identifying satisfactory (S) and unsatisfactory (U) outcomes.

Naïve Bayes classified 186 students (71.3%) with satisfactory writing skills correctly, misclassifying 75 students (28.7%) as unsatisfactory. For students with unsatisfactory performance, it correctly identified 94 students (61.8%) but misclassified 58 students (38.2%) as satisfactory. This balance between true positives and true negatives highlights Naïve Bayes' ability to predict unsatisfactory outcomes reasonably well, though its accuracy for satisfactory predictions was somewhat lower.

4.4.4 Findings Related to Sub-Research Question 3.2.2

RQ3.2.2: What are the most predictive features that contributed to the best-performing model in the prediction performance for writing skills in the second quarter?

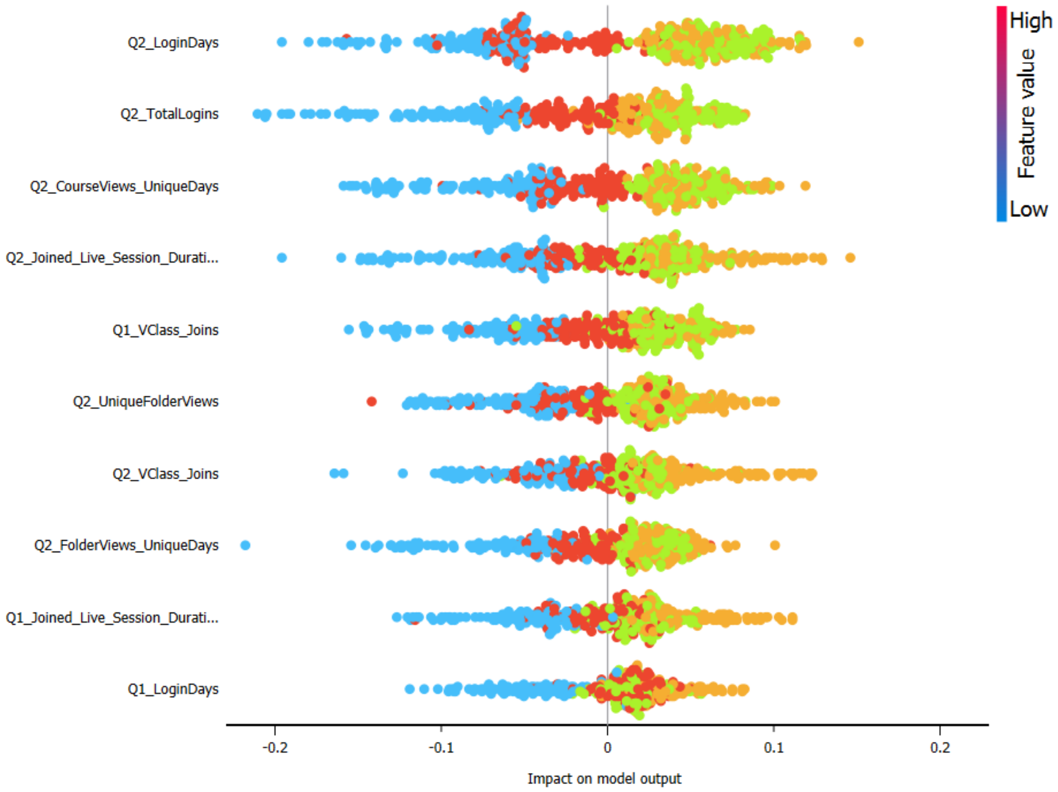

Figure 4.11. Explanation of the Model for Writing Skills in the Second Quarter (Naïve Bayes)

The Naïve Bayes model (Figure 4.11.) for predicting writing skills in the second quarter identified several key features that significantly contributed to its predictive performance. Among these, Q2_LoginDays emerged as the most influential. Frequent logins during the second quarter were strongly correlated with improved writing skills, emphasizing the importance of regular engagement with the LMS. Similarly, Q2_TotalLogins, which represented the total number of logins during the second quarter, was another critical contributor. High login counts reflected active participation and interaction with the platform, which positively influenced students' academic outcomes.

Another major factor was Q2_CourseViews_UniqueDays, which measured the number of unique days students accessed course materials. This metric highlighted the value of sustained and consistent engagement with learning content, demonstrating its significant role in enhancing writing skills. Q2_Joined_Live_Session_Duration, representing the time spent in live sessions during the second quarter, also contributed positively. Active participation in synchronous learning environments supported a deeper understanding and application of writing concepts.

Features from the first quarter also played an important role. For instance, Q1_VClass_Joins, which tracked the number of virtual class sessions attended in the first quarter, had a lasting positive effect on students' writing performance in the second quarter. This early engagement provided a strong foundation for success. Similarly, Q1_Joined_Live_Session_Duration, which measured the time spent in live sessions during the first quarter, demonstrated that early participation in interactive learning activities had a cumulative impact on students' academic performance.

Other significant contributors included Q2_UniqueFolderViews and Q2_VClass_Joins. Accessing unique folders and joining virtual class sessions during the second quarter reflected the importance of exploring diverse resources and participating in instructional activities. Both metrics demonstrated the positive effect of active engagement with a variety of learning materials and activities. Additionally, Q2_FolderViews_UniqueDays, which measured the number of unique days folder materials were viewed, reinforced the importance of consistent and varied interaction with course resources.

Lastly, Q1_LoginDays, capturing login frequency during the first quarter, further underscored the critical role of early engagement in shaping later success. This feature highlighted how building strong engagement habits early in the academic year supported improved performance in writing skills during the second quarter.

In summary, the Naïve Bayes model demonstrated that both early and ongoing engagement with LMS resources and virtual classrooms played a vital role in predicting writing skills. Features such as frequent logins, consistent access to course materials, participation in live sessions, and interaction with diverse resources collectively underscored the value of sustained and meaningful engagement in achieving academic success.

4.4.5 Findings Related to Sub-Research Question 3.3.1

RQ3.3.1: Does engagement in LMS and virtual classrooms predict academic performance in speaking skills in the second quarter?

Table 4.21. Performance Scores of Classification Algorithms for Speaking Skills in the Second Quarter

Model	AUC	CA	F1	Prec.	Recall	MCC
Logistic Regression	0.793	0.743	0.732	0.735	0.743	0.407
Gradient Boosting	0.775	0.722	0.716	0.714	0.722	0.368
Naïve Bayes	0.774	0.719	0.721	0.723	0.719	0.390
SVM	0.765	0.743	0.734	0.735	0.743	0.410
Neural Network	0.764	0.714	0.708	0.706	0.714	0.350
Random Forest	0.761	0.729	0.717	0.719	0.729	0.373
k-NN	0.746	0.709	0.704	0.702	0.709	0.343
CN2 rule inducer	0.706	0.649	0.652	0.657	0.649	0.243
SGD	0.679	0.731	0.722	0.722	0.731	0.383
AdaBoost	0.613	0.639	0.643	0.647	0.639	0.223
Tree	0.589	0.668	0.663	0.660	0.668	0.250

The predictive performance of classification algorithms (Table 4.21.) for speaking skills in the second quarter highlights the effectiveness of engagement in LMS and virtual classrooms. Among the evaluated models, Logistic Regression emerged as one of the top-performing algorithms. It achieved the highest Area under the Curve (AUC) at 0.793 and a Classification Accuracy (CA) of 0.743, alongside strong metrics for F1-score (0.732), precision (0.735), recall (0.743), and a Matthews Correlation Coefficient (MCC) of 0.407. This performance indicates its strong ability to predict speaking skill outcomes based on LMS engagement data.

SVM also performed well, matching Logistic Regression in CA (0.743) and recall (0.743), with an AUC of 0.765 and a slightly higher MCC of 0.410. These metrics demonstrate its strong predictive reliability. Gradient Boosting followed closely with an AUC of 0.775 and a CA of 0.722, alongside balanced F1 (0.716) and MCC (0.368), indicating its potential as a robust predictive model. Similarly, Naïve Bayes achieved an AUC of 0.774, a CA of 0.719, and an MCC of 0.390, reflecting its ability to handle probabilistic relationships in the data effectively.

Other models, such as Neural Network and Random Forest, performed moderately. Neural Network recorded an AUC of 0.764, a CA of 0.714, and an MCC of 0.350, demonstrating reliable, albeit slightly lower, performance. Random Forest achieved a CA of 0.729, an F1-score of 0.717, and an MCC of 0.373, indicating a good balance of precision and recall but slightly lower discrimination capability (AUC: 0.761) compared to Logistic Regression and Gradient Boosting.

Lower-performing models included k-NN, CN2 Rule Inducer, AdaBoost, and Tree. k-NN achieved an AUC of 0.746 and a CA of 0.709, while CN2 Rule Inducer recorded an AUC of 0.706 and a CA of 0.649, reflecting limited predictive capabilities. AdaBoost and Tree had the lowest AUCs (0.613 and 0.589, respectively) and CAs below 0.670, indicating their inadequacy in predicting speaking skills in the second quarter.

In conclusion, Naïve Bayes, Logistic Regression and SVM emerged as the most effective models for predicting speaking skill performance, demonstrating high accuracy, precision, and recall. Models like Gradient Boosting also showed strong performance, making them viable alternatives for analyzing LMS engagement data. The results emphasize the importance of choosing advanced and reliable algorithms to predict speaking skill outcomes accurately.

Table 4.22. Confusion Matrix of Classification Algorithm for Speaking Skills in the Second Quarter

	Naïve Bayes		
	S	U	Total
Satisfactory	207	62	269
Unsatisfactory	54	90	144
Total	261	152	413

The confusion matrix (Table 4.22.) for Naïve Bayes in predicting speaking skills during the second quarter provides insights into its performance in classifying satisfactory (S) and unsatisfactory (U) outcomes.

Naïve Bayes showed a balanced approach in classifying both satisfactory and unsatisfactory outcomes, albeit with lower overall accuracy for satisfactory predictions. It correctly classified 207 students (77.0%) as satisfactory, while misclassifying 62 students (23.0%) as unsatisfactory. For unsatisfactory outcomes, Naïve Bayes correctly identified 90 students (62.5%) and misclassified 54 students (37.5%) as satisfactory.

4.4.6 Findings Related to Sub-Research Question 3.3.2

RQ3.3.2: What are the most predictive features that contributed to the best-performing model in the prediction performance for speaking skills in the second quarter?

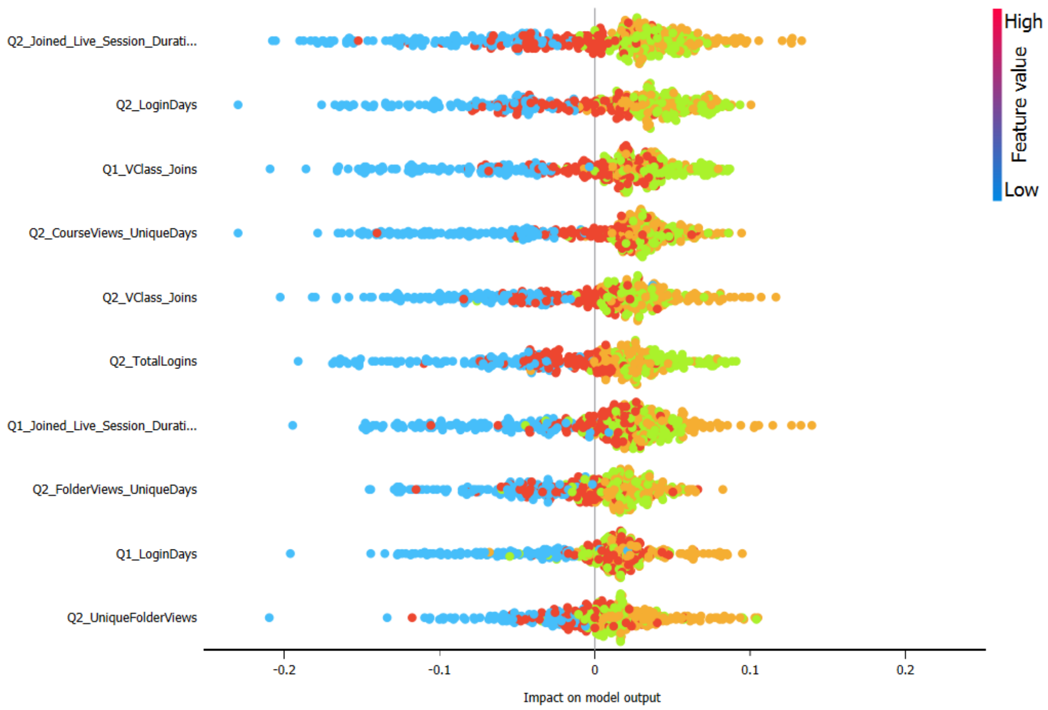

Figure 4.12. Explanation of the Model for the Use of Speaking Skills in the Second Quarter (Naïve Bayes)

In the SHAP graph (Figure 4.12.), the Naïve Bayes model for predicting speaking skills in the second quarter identified several key features contributing to its predictive performance. Among the positively impactful features, Q2_Joined_Live_Session_Duration emerged as a critical factor. Longer participation in live sessions during the second quarter strongly correlated with improved speaking skills, emphasizing the importance of active engagement in real-time learning activities. Similarly, Q2_LoginDays, which represents the number of login days during the second quarter, had a significant positive influence, highlighting the role of regular LMS engagement in supporting skill development.

Other positively contributing features included Q1_VClass_Joins, which captured participation in virtual class sessions during the first quarter. This early engagement provided a foundation for continued success in the second quarter.

Q2_CourseViews_UniqueDays, which tracks the number of unique days course materials were accessed during the second quarter, also positively influenced the model's output, emphasizing the value of consistent interaction with learning resources.

Additional features contributing positively were Q2_VClass_Joins, which measured virtual class participation in the second quarter, and Q2_TotalLogins, representing overall login activity during the same period. These features reinforced the importance of consistent participation and overall engagement in LMS activities. Similarly, Q1_Joined_Live_Session_Duration, which captured live session participation in the first quarter, contributed positively, highlighting the lasting benefits of early interactive engagement.

Resource-related features also had positive impacts. Q2_FolderViews_UniqueDays, measuring the number of unique days folder materials were accessed, and Q2_UniqueFolderViews, reflecting the diversity of folder materials accessed in the second quarter, demonstrated the importance of consistent and varied engagement with resources. Lastly, Q1_LoginDays, representing the login frequency during the first quarter, showed that early and sustained engagement positively influenced speaking skill outcomes.

In conclusion, the Naïve Bayes model underscored the importance of consistent LMS usage, active participation in live sessions, and timely access to diverse learning materials in predicting speaking skills. Early engagement combined with sustained interaction in the second quarter played a crucial role in supporting skill development. The model highlighted how active and consistent learning behaviors are central to improving academic performance in speaking skills.

4.5 Summary of the Results

The analysis of student performance in the English proficiency exam revealed significant disparities across various categories, including faculties, departments, instructional methods, language of instruction, gender, and predictive analytics. These findings highlight the complex interplay of institutional, demographic, and behavioral factors that shape language learning outcomes and provide insights into potential areas for targeted interventions.

Faculties and departments demonstrated clear differences in achievement. Faculties such as Medicine and Pharmacy performed exceptionally well, likely due to rigorous academic programs and selective admission policies that attract high-performing students. In contrast, vocational schools faced significant difficulties in attaining English proficiency. Similarly, department-level analysis revealed high performance in fields like Computer Engineering, while others, such as Geology and Energy Engineering, faced substantial challenges in meeting English proficiency requirements during online English language learning.

The type of instruction students would follow in their academic programs after completing the preparation school also influenced outcomes. Students preparing for face-to-face academic programs outperformed those transitioning to DE programs.

Language of instruction at designated programs further emerged as a significant factor in exam performance. Students preparing for English-medium programs generally achieved better results than those preparing for Turkish-medium programs. This difference likely stems from the higher proficiency threshold required for English-medium programs and increased exposure to English in their academic contexts, which may have better motivated and prepared these students for the exam.

Gender-based differences were also apparent. Female students outperformed their male counterparts, aligning with research suggesting that female learners tend to have stronger engagement and study habits in language acquisition. Male students faced greater challenges, highlighting the need for targeted strategies to enhance their language learning outcomes and engagement with the preparation program.

In addition to analyzing observed performance, the study incorporated predictive modeling to identify key factors influencing success and failure. Various machine learning algorithms, including Logistic Regression, Naïve Bayes, and Neural Networks, were employed to predict student outcomes based on engagement metrics. Key predictors of success included frequent logins, timely assignment submissions, and active participation in virtual classes, indicating that consistent engagement and interaction with course materials were critical for achieving proficiency in English. Predictive models also identified students at risk of failing, offering valuable opportunities for early intervention.

Overall, the results reflect the multifaceted nature of student achievement in the English proficiency exam, influenced by a combination of institutional, behavioral, and demographic factors. The inclusion of predictive analytics provided deeper insights, enabling the identification of at-risk students and highlighting actionable pathways for enhancing learning outcomes through tailored support and interventions. These findings underscore the importance of leveraging both traditional and data-driven approaches to foster equitable and effective language learning environments.

CHAPTER 5

DISCUSSION AND CONCLUSION

This section of the study includes sub-sections of introduction, discussion of descriptive results, predictive basis for early warning systems, gateway and proficiency exams prediction, skill-based prediction, methodological reflections, conclusion, practical implications, and implications for future research.

5.1 Introduction

This chapter synthesizes the findings of the study on predictive modeling in online English language learning at a public university in Türkiye, particularly during the COVID-19 pandemic. The study focused on the unique challenges posed by the abrupt shift to online learning environments, where students faced varying levels of preparedness and resource availability. The unprecedented situation, enabling online learning across all domains, which was not typical under normal circumstances. For instance, the adoption of online platforms for practical courses in domains such as engineering and the arts proved both innovative and challenging. While theoretical disciplines like social sciences and literature adapted relatively smoothly, hands-on learning often faced difficulties due to the lack of physical interaction and resources, highlighting the need for tailored solutions in different areas. However, it also introduced challenges distinct from those encountered in traditional online learning settings, requiring adaptive strategies to address these new dynamics. By employing a variety of machine learning algorithms, including Logistic Regression, Gradient Boosting, Naïve Bayes, and Neural Networks, the research aimed to identify the most effective methods for predicting student performance in online English language learning. This discussion integrates insights from the literature review, methodology, and results, linking them to practical implications and future directions in the context of online English learning and teaching in Türkiye.

5.2 Discussion of Descriptive Results

This section presents a discussion of the descriptive results obtained in the study, specifically students' success rates based on faculty, department, vocational school, language of instruction, type of instruction, and gender.

5.2.1 Faculties

The wide variation in pass rates in English proficiency exam across faculties highlighted the role of institutional and individual factors, including students' motivation and preparedness. Faculties with higher pass rates likely benefited from their selective admission policies and the presence of high-performing students. Conversely, faculties with lower pass rates required targeted interventions, such as intensive English language training programs, to help bridge the gap and improve outcomes for underperforming students.

Descriptive data results revealed significant opportunities for improvement in faculty based intervention with lower achievement rates. Efforts could focus on identifying best practices from high-performing faculty students and tailoring them to address the specific needs of low performing students of other faculties. This approach would ensure more equitable language learning opportunities across the institution. Additionally, providing supplemental support, repetition activities, and ongoing mentorship for struggling student groups could enhance their outcomes effectively, helping to reduce performance disparities.

5.2.2 Departments

The descriptive analysis of English proficiency exam outcomes across various departments revealed significant disparities. Departments like Medicine, Computer Engineering, and Pharmacy exhibited high achievement levels in English proficiency exam, potentially reflecting rigorous academic standards and students' prior exposure to English, given the international nature of their fields. In contrast, departments such as Geology Engineering and Energy Engineering reported notably low pass rates, indicating potential challenges in student preparedness or the effectiveness of language support mechanisms.

To address these disparities, implementing predictive learning analytics, early warning systems, and field-specific interventions can be instrumental. Studies have demonstrated that such systems can accurately predict at-risk students early in the academic term, enabling timely interventions. For instance, Akçapınar, Altun and Aşkar (2019) found that the k-Nearest Neighbors (kNN) algorithm could predict unsuccessful students with an accuracy rate of 89% by the end of the term and 74% as early as the third week. By leveraging data from LMS data and other educational sources, such as categorical data from the student information system, institutions can identify students who may require additional language support and tailor interventions accordingly. This proactive approach not only enhances individual student outcomes but also contributes to the overall effectiveness and equity of English language learning programs within university-based schools of languages.

5.2.3 Vocational Schools and Faculties

The extremely low pass rate in vocational school students may highlight potential issues such as inadequate preparation, misaligned curricula, or a lack of emphasis on English language learning. The descriptive data results suggest a need for targeted interventions, such as specialized language support tailored to the unique needs of vocational students.

Faculty students performed, in percentage, better than vocational school students. Despite outperforming vocational schools, the 51.5% fail rate among faculty students indicates that a large portion of students in faculties also faced challenges in reaching or achieving English proficiency exam at the end of two academic semesters. This suggests a need for broader initiatives to enhance English language acquisition for students registered at faculties.

5.2.4 Language of Instruction at Designated Department

The descriptive results highlighted the challenges faced by Turkish-medium students in adapting to the English preparation school curriculum. These students may experience reduced exposure to practical and immersive English language use, which could hinder their ability to perform well on the proficiency exam. In contrast, English-medium students are more likely to experience curricula tailored to the demands of their future academic environment, providing them with better preparation for the exam. This disparity underscores the importance of aligning English preparation programs with the specific needs of both groups.

To address these challenges, targeted interventions can be implemented to support Turkish-medium students. These may include emphasizing the broader benefits of English proficiency, even for those in Turkish-medium programs, such as its importance in global communication and career opportunities. Tailored resources, such as additional practice sessions and bilingual workshops, could also enhance their motivation and proficiency. For English-medium students, preparation programs could further align language instruction with the academic language requirements of their future disciplines, ensuring they are well-equipped for success.

In conclusion, while students preparing for English-medium programs outperformed their Turkish-medium counterparts in percentage of pass rate, the gap in performance highlights the need for differentiated support strategies. By addressing these disparities, institutions can ensure that all students, regardless of their future language of instruction, receive the resources and guidance necessary to excel in English proficiency and beyond.

5.2.5 Type of Instruction at Designated Department

Despite all students receiving online instruction during the pandemic, the mode of instruction in their future academic programs appears to have influenced their motivation and performance. Students bound for face-to-face programs likely placed greater emphasis on achieving English proficiency, understanding its importance for interactive, in-person academic environments. In contrast, students preparing for DE programs may have struggled to maintain engagement and motivation, compounded by the challenges of adapting to the remote learning model required by their future studies.

The low pass rates among DE students underscore the need for additional support and resources. Enhancing their self-directed learning skills through tools such as virtual study groups, regular feedback, and interactive online activities could help bridge the gap. For face-to-face program students, further emphasis on collaborative and participatory learning methods during preparation could strengthen their outcomes.

To conclude, the inclusion of students who failed to reach the English proficiency exam within the fail category provides a more comprehensive view of the challenges faced during the English preparation school. While students transitioning to face-to-face programs performed better in percentage, the struggles of DE students highlight the necessity for targeted interventions. By addressing these disparities, institutions can better equip all students with the language proficiency needed for their academic and professional success, regardless of their future instructional medium type.

5.2.6 Gender

The higher success rate among female students aligns with broader educational research, which frequently indicates that females excel in language acquisition and tend to outperform males in language-based assessments. This may be attributed to stronger engagement, consistent study habits, or greater motivation toward academic success in language-related contexts. On the other hand, male students faced notable challenges, as evidenced by their higher fail rate of 58.6% in English proficiency exam. This highlights the need for targeted strategies to support male students in developing their English language proficiency skills.

It is also worth noting the gender imbalance in participation rates. Female students made up 56.3% of the participants, while male students represented 43.7%. This disproportion might reflect broader enrollment patterns in programs requiring English proficiency or gendered preferences for certain academic fields. The larger proportion of female participants further emphasizes the importance of addressing disparities in outcomes to ensure equitable success rates across genders.

5.3 Predictive Basis for Early Warning Systems

Early warning systems (EWSs) in LA represent a critical development in educational technology, particularly in online English language learning in this context. These systems leverage the vast data generated in online learning environments to identify at-risk students and enable timely interventions. So-called early alert or warning systems rely on predictive analytics to track student interaction and engagement patterns, offering insights into factors contributing to academic success or failure (Jayaprakash et al., 2014).

In online English language courses, EWSs can monitor participation in activities such as reading comprehension exercises, vocabulary activities and quizzes, writing tasks, and interactive resource engagements and discussions. These systems assess engagement metrics, such as login frequency, time spent on specific tasks, and completion rates. For example, Keržić et al. (2019) demonstrated that quiz performance is strongly predictive of final exam outcomes, suggesting that similar metrics could be adapted to evaluate English language proficiency growth over time. In this study, features such as virtual class attendance, assignment submissions, and login activity played an important role in identifying students at risk of failure—particularly in predicting first- and second-quarter failure based on first-quarter data. The ability to identify disengaged students early allows instructors to implement targeted interventions, such as personalized feedback, additional resources, or adaptive learning paths.

Furthermore, LA tools integrated into EWSs support customization and scalability, particularly in large online cohorts. For instance, tools like LA dashboards provide students and instructors with feedback and real-time data visualizations of student progress, enabling them to prioritize support for those in need (Park & Jo, 2019). This capacity is especially relevant in language learning, where diverse learner needs and proficiency levels demand tailored educational strategies.

Ethical considerations also play a significant role in deploying EWS in online English language learning platforms similar to other educational contexts. The collection and analysis of student data must align with privacy regulations and ensure that interventions are supportive rather than discouraging, denunciatory or punitive (West et al., 2018). Transparency in how data is used fosters trust and encourages students to actively participate in their LA journey.

EWSs in LA hold transformative potential for online English language learning. However, this study had limited predictive capability, particularly in classifying students at risk of failure. This study’s first research question aimed to use first-quarter engagement data to establish a predictive foundation for an early warning system. While the model faced challenges in predicting students at risk of failure—particularly those who failed during the third and fourth quarters—it effectively identified those who successfully passed the English proficiency exam. By identifying at-risk students and facilitating timely, data-driven interventions, these systems could enhance learning outcomes and engagement. In this regard, the study faced challenges stemming from the time constraints of predictive analytics and the diverse backgrounds of the students. Designing efficient interventions is impossible without effective predictive models, aligning with Larrabee Sønderlund et al.'s (2019) argument that dynamic and precise predictive models are essential for successful early interventions. Therefore, it may be more effective to focus on predicting student achievement in the upcoming quarter rather than in later ones, in order to maintain higher prediction accuracy.

5.4 Gateway and Proficiency Exams Prediction

Predicting outcomes for gateway and proficiency exams presented unique challenges due to the diverse difficulty levels and specific skill requirements inherent to these assessments. Logistic Regression emerged as a robust tool, particularly in its ability to identify high-performing students effectively. This model's simplicity and transparency can provide some actionable insights for educators at the School of Foreign Languages in Türkiye, facilitating targeted interventions for students at risk of underperformance within online English language learning contexts.

Naïve Bayes added value by excelling in scenarios requiring a balance between precision and recall. Its capacity to minimize false positives proved advantageous in contexts where over-identification of at-risk students posed fewer consequences than overlooking those genuinely in need. Meanwhile, ensemble methods such as Gradient Boosting also offered competitive accuracy in some cases.

The integration of the SMOTE addressed class imbalances effectively in specific cases, particularly enhancing recall for underrepresented groups in the first and third quarter Gateway Exams. However, the technique exhibited limited effectiveness in other scenarios, indicating its dependency on dataset characteristics and prediction tasks. These findings are critical for this school of foreign languages, where disparities in demographic and domain backgrounds significantly affect student performance.

Validity and academic integrity emerged as critical concerns, particularly given the limited proctoring during the pandemic. Unauthorized collaboration and use of external resources likely skewed performance metrics, introducing noise into predictive models (Jenkins et al., 2023; Janke et al., 2021). Future systems must integrate more secure proctoring methods, including AI-based monitoring and hybrid examination setups, to enhance academic integrity and model reliability (Nigam et al., 2021).

In the predictive models, this study focused exclusively on scale features, intentionally excluding demographic and categorical variables (e.g., gender, faculty, department) to maintain generalizability across broader populations. While this approach facilitated scalable insights, it limited the understanding of subgroup-specific dynamics. Future research should consider incorporating these variables to derive nuanced predictions and models while adhering to privacy and ethical considerations (Chatti & Muslim, 2019).

The findings also highlight the importance of context-specific model selection (e.g., using post-SMOTE training data). Institutions, along with all stakeholders, must prioritize models that align with their educational objectives, whether focusing on precision, recall, or interpretability. Accuracy in predictive analytics should not be viewed in isolation but must consider Type I (false positives) and Type II (false negatives) errors to foster ethical, equitable, and emotionally considerate interventions.

In addressing the second research question, which involved predicting student performance across four quarters separately, the study achieved its goals to some extent—particularly in the first, second, and third quarters. However, the predictive capacity was limited in the fourth quarter. In this regard, it can be asserted that predictive accuracy tends to decrease as students approach the final phase of passing the English proficiency exam. Another important reason for the decrease in predictive accuracy may be one of the main limitations of this study—the limited proctoring capability during the examination process. Further investigation—especially through qualitative research—may be necessary to gain deeper insights into these findings.

5.5 Skill-Based Prediction

The skill-based prediction models offered valuable insights into specific language competencies, including grammar-reading-listening, writing, and speaking in the second quarter. Logistic Regression excelled in predicting foundational skills like grammar-reading-listening, achieving high precision and recall. Since the model utilizing Naïve Bayes demonstrated strong performance, particularly in terms of a balanced confusion matrix, it was selected for further reporting in this study on skill-based prediction. Improved balance and accuracy in predictive models can lead to more effective intervention plans, particularly for students at risk of failure.

However, the predictive capacity for higher-order language skills, such as writing and speaking, could be improved in this study. Where applicable, it may be beneficial to further differentiate LMS resources by specific language skills—particularly during the third and fourth quarters. In this regard, these skills (e.g., writing) can be designed for higher-order thinking level such as according to Bloom's taxonomy (Ganapathy & Kaur, 2014). Models utilizing Naïve Bayes and Logistic Regression demonstrated moderate predictive capabilities in these areas, highlighting the complexity of modeling nuanced language skills. The results suggest that the development of richer feature sets e.g., better logging capacity by designing on LMS, incorporating data on student interactions, peer feedback, and contextual language use, could enhance the accuracy of these models. Furthermore, splitting data into smaller partitions for the improved predictive models can be necessary so that further individualized support for students can be designed.

Additionally, the study revealed the importance of skill-specific predictive modeling. For example, speaking skills benefited from algorithms that accounted for the participation in live sessions of online classes. Improving coordination with stakeholders—such as scheduling 15-minute weekly online meetings—can increase the likelihood of skill-based applications of learning analytics in online English language learning and teaching within schools of foreign languages. On the other hand, stakeholders should be more willing to cooperate for possible applications of LA. This can be achieved when the benefits are clearly stated with supporting data and theoretical understanding, fostering a shared commitment to implementation. In collaboration with stakeholders, all educational resources could be better labeled according to the four core skills of English language acquisition—listening, reading, writing, and speaking, especially in the third and fourth quarters. Such labeling would enhance the potential for more precise skill-based intervention plans, allowing educators to address specific deficiencies and tailor learning pathways to individual or group needs (Brinkhuis et al., 2018).

5.6 Methodological Reflections

The study's methodology, which combined advanced machine learning techniques with rigorous evaluation metrics, demonstrated the importance of aligning analytical approaches with research objectives. Although this study included data from four quarters, dividing the data into finer categories, such as eight timelines, could enable more granular applications of predictive LA. This approach may enhance the ability to detect patterns and intervene effectively. While the SMOTE-enhanced datasets were tested across all cases, its positive contributions were restricted to gateway exams in the first and third quarters, suggesting that oversampling techniques should be applied selectively based on context.

The findings also point to several areas for methodological improvement. This study included beginner students from L1 to L4, acknowledging that students starting to learn English online at different levels may require different tailored considerations.

Future methodologies should account for these differences to enhance prediction accuracy and intervention strategies. Maintaining optimal accuracy in prediction has been particularly challenging due to the abrupt changes in learning environments and the variability in student access to online resources. For instance, while Naïve Bayes showed potential in skill-based predictions, their lower interpretability highlighted the need for more transparent modeling approaches. The interpretability and explainability of predictive models are critically important for planning effective interventions.

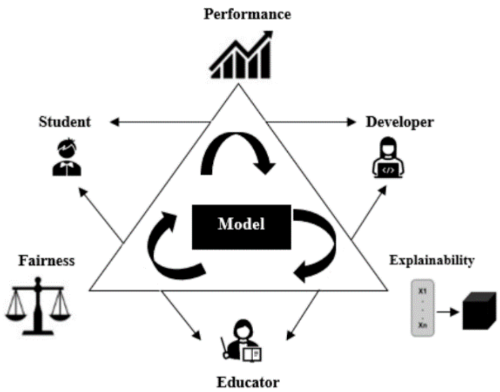

Figure 5.1. Interplay of Explainability-Performance-Fairness, from Gunasekara and Saarela (2024, p. 4).

The validity of the results in this study could be enhanced by addressing concerns related to cheating and data integrity in limited or non-proctored examination settings, although some level of proctoring was implemented and maintained throughout the examination processes. Developing secure performance data collection and monitoring mechanisms will be critical for ensuring robust conclusions.

Moreover, the sampling strategy employed in this study was based on convenience sampling, which may limit the generalizability of the findings. To provide a more comprehensive overview, future studies should aim to include the entire student population at the university. By encompassing a broader sample, predictive models can more effectively account for the diversity of students' online learning experiences. Furthermore, involving a broad range of stakeholders—such as administrators, instructors, and students—in both data collection and evaluation processes can offer a more comprehensive understanding of LA's effectiveness and areas for improvement.

While this study employed ensemble machine learning algorithms such as Gradient Boosting and AdaBoost, incorporating additional techniques—such as stacking—could further enhance predictive performance. Moreover, utilizing alternative software platforms that support a broader range of ensemble algorithms may improve accuracy across different prediction categories.

5.7 Conclusion

This study highlights the transformative potential of predictive modeling in the context of online English language learning in Türkiye. By employing a range of machine learning algorithms, it offers valuable insights for improving student performance. The findings underscore the importance of aligning predictive analytics with educational goals, by prioritizing models that balance accuracy, interpretability, explainability, and equity.

Through targeted interventions, enhanced curriculum design, and personalized learning pathways, predictive analytics can significantly improve educational outcomes. These approaches are consistent with the primary goals of LA, which ultimately focus on drawing robust causal inferences and identifying the most effective interventions to optimize and enhance learning outcomes, as noted by Motz et al. (2018). As the field of LA continues to evolve, future research should focus on advancing methodological rigor, expanding datasets, and enhancing model interpretability, ensuring that these tools remain accessible and impactful for all stakeholders.

5.8 Practical Implications

The findings of this study hold significant implications for educational practice in Türkiye. EWSs can be integrated into LMSs to provide real-time feedback to both educators and students within schools of foreign languages. Additionally, a real-time dashboard could be developed to inform students, instructors, and administrators when predictive accuracy improves, fostering a collaborative environment for data-driven decision-making. By leveraging predictive analytics, institutions can design targeted interventions and allocate resources more effectively to support at-risk students. In line with the aspirations of LA as a field, these interventions should ultimately aim to identify specific, causally effective actions that optimize and enhance learning outcomes (Motz et al., 2018). This is particularly important in Türkiye, where disparities in access to technology and internet connectivity—especially during pandemics or natural disasters—can significantly impact online learning outcomes.

For gateway and proficiency exams, predictive models can help students overcome learning bottlenecks and assist educators in identifying the areas where students struggle most in their efforts to succeed. Moreover, skill-based predictions can inform personalized learning pathways by tailoring instructional strategies to meet individual student needs. Resources should be more effectively categorized and labeled according to specific English language skills. These personalized pathways can be successfully implemented by leveraging adaptive technologies, aligning instructional strategies with predictive insights, and fostering collaboration among instructors and administrators to share best practices.

In-service training programs should be offered to stakeholders—particularly instructors—to strengthen their information and communication technology (ICT) skills and promote the effective use of LMS and other educational technologies. For example, it was observed that some instructors collected writing assignments through alternative methods—such as email—rather than utilizing the LMS. This would not only promote more effective use of the LMS but also generate additional data for predictive LA, thereby enhancing the accuracy of predictions related to writing skills and other language skill performances in gateway exams.

Decision-makers at the university should prioritize the implementation of LA and invest in its development by allocating the necessary human and technological resources whenever feasible. Although it may not be possible to keep all low-performing students on track, there is substantial potential to support many of them in achieving their learning goals in online English language learning—particularly in gateway and English proficiency exams. Interpretable and scalable models can be developed across institutions that adopt similar teaching approaches and technological infrastructures. Additionally, the integration of secure proctoring tools can enhance the reliability of LA outcomes. By collaborating to share technological infrastructure and expertise, institutions can reduce the costs associated with maintaining adaptable learning environments and ensuring secure examination settings.

5.9 Implications for Future Research

Future research should build on this study by examining the generalizability of predictive LA models across schools of foreign languages at different universities in Türkiye. This implication could involve incorporating and comparing data from similar settings across different universities, thereby enriching the analysis and applicability of predictive LA by capturing insights from a broader spectrum of educational environments. On the other hand, differences in educational contexts—such as institutional resources, student demographics, and curriculum design—can significantly influence the applicability of predictive models. For instance, schools with limited technological infrastructure may require simpler predictive models, whereas resource-rich institutions might benefit from more complex, data-intensive approaches. Such variations should be carefully considered to ensure both the effectiveness and equity of LA applications. Case studies and qualitative research are needed to investigate the underlying causes of performance deficiencies in online

English language learning. Such approaches can more effectively address issues of causation and guide the development of targeted interventions (Motz et al., 2018; Gardner & Brooks, 2018). Cross-institutional studies could offer valuable insights into the adaptability of predictive models by addressing variations in curriculum design, student demographics, and technological infrastructure.

Additionally, further exploration of feature engineering techniques could enhance model performance for complex skills such as writing and speaking. Developing richer datasets that incorporate qualitative and contextual information may provide a more holistic understanding of student performance. Furthermore, improving the interpretability of complex models will be essential for their widespread adoption in educational settings. Specific strategies—such as employing Explainable AI (XAI) tools, simplifying model architectures, and integrating user-friendly visualization techniques—can help bridge the gap between predictive power and usability. These approaches enable instructors and administrators, including those in schools of foreign languages in Türkiye, to confidently interpret and apply predictive models, thereby fostering broader acceptance and practical implementation.

REFERENCES

AERA Code of Ethics: American Educational Research Association Approved by the AERA Council February 2011. (2011). *Educational Researcher, 40*(3), 145-156. https://doi.org/10.3102/0013189X11410403

Ahmed, E. (2024). Student Performance Prediction Using Machine Learning Algorithms. *Applied Computational Intelligence and Soft Computing, 2024*(1), 4067721. https://doi.org/10.1155/2024/4067721

Akçapınar, G., Altun, A., & Aşkar, P. (2019). Using learning analytics to develop early-warning system for at-risk students. *International Journal of Educational Technology in Higher Education, 16*(1). https://doi.org/10.1186/s41239-019-0172-z

Akpen, C. N., Asaolu, S., Atobatele, S., Okagbue, H., & Sampson, S. (2024). Impact of online learning on student's performance and engagement: a systematic review. *Discover Education, 3*(1), 1-15. https://doi.org/10.1007/s44217-024-00253-0

Alfares, N. (2024). Is synchronous online learning more beneficial than asynchronous online learning in a Saudi EFL setting: teachers' perspectives. *Frontiers in Education, 9.* https://doi.org/10.3389/feduc.2024.1454892

Almeida, F., & Simoes, J. (2019). The Role of Serious Games, Gamification and Industry 4.0 Tools in the Education 4.0 Paradigm. *Contemporary Educational Technology*, *10*(2), 120–136. https://doi.org/10.30935/cet.554469

Altabrawee, H., Ali, O. a. J., & Ajmi, S. Q. (2019). Predicting students' performance using machine learning techniques. *Journal of University of Babylon for Pure and Applied Sciences, 27*(1), 194–205. https://doi.org/10.29196/jubpas.v27i1.2108

AlTwijri, L., & Alghizzi, T. M. (2024). Investigating the integration of artificial intelligence in English as foreign language classes for enhancing learners' affective factors: A systematic review. *Heliyon, 10(*10), e31053. https://doi.org/10.1016/j.heliyon.2024.e31053

Ameloot, E., Rotsaert, T., Ameloot, T., Rienties, B., & Schellens, T. (2024). Supporting students' basic psychological needs and satisfaction in a blended learning environment through learning analytics. *Computers & Education, 209*, 104949. https://doi.org/10.1016/j.compedu.2023.104949

Amrieh, E. A., Hamtini, T., & Aljarah, I. (2015, November). Preprocessing and analyzing educational data set using X-API for improving student's performance. In *2015 IEEE Jordan conference on applied electrical engineering and computing technologies (AEECT*) (pp. 1-5). IEEE.

Andersson, U., Arvemo, T., & Gellerstedt, M. (2016). How well can completion of online courses be predicted using binary logistic regression?. In *IRIS39, Information Systems Research Seminar in Scandinavia, Ljungskile, August 7-10, 2016* (pp. 1-12).

Archer, E., & Prinsloo, P. (2020). Speaking the unspoken in learning analytics: troubling the defaults. *Assessment & Evaluation in Higher Education, 45*(6), 888-900. https://doi.org/10.1080/02602938.2019.1694863

Arnold, K. E., & Pistilli, M. D. (2012, April). Course signals at Purdue: Using learning analytics to increase student success. In *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge* (pp. 267-270). https://doi.org/10.1145/2330601.2330666

Avella, J. T., Kebritchi, M., Nunn, S. G., & Kanai, T. (2016). Learning analytics methods, benefits, and challenges in higher education: A systematic literature review. *Online Learning, 20*(2), 13-29. https://doi.org/10.24059/olj.v20i2.790

Baker, R. S., & Yacef, K. (2009). The State of Educational Data Mining in 2009: A Review and Future Visions. *Journal of Educational Data Mining, 1*(1), 3–17. https://doi.org/10.5281/zenodo.3554657

Baldi, P., Brunak, S., Chauvin, Y., Andersen, C. A., & Nielsen, H. (2000). Assessing the accuracy of prediction algorithms for classification: an overview. *Bioinformatics, 16*(5), 412-424. https://doi.org/10.1093/bioinformatics/16.5.412

Banna, J., Grace Lin, M. F., Stewart, M., & Fialkowski, M. K. (2015). Interaction matters: Strategies to promote engaged learning in an online introductory nutrition course. *Journal of Online Learning and Teaching, 11*(2), 249–261.

Bao, W. (2020). COVID-19 and online teaching in higher education: A case study of Peking University. *Human Behavior and Emerging Technologies, 2*(2), 113-115. https://doi.org/10.1002/hbe2.191

Basilaia, G., & Kvavadze, D. (2020). Transition to online education in schools during a SARS-CoV-2 coronavirus (COVID-19) pandemic in Georgia. *Pedagogical Research, 5*(4). https://doi.org/10.29333/pr/7937

Berry, G. R., & Hughes, H. (2020). Integrating work–life balance with 24/7 information and communication technologies: The experience of adult students with online learning. *American Journal of Distance Education, 34*(2), 91-105. https://doi.org/10.1080/08923647.2020.1701301

Bienkowski, M., Feng, M., & Means, B. (2012). Enhancing Teaching and Learning through Educational Data Mining and Learning Analytics: An Issue Brief. *Office of Educational Technology, US Department of Education.*

Blake, R. J. (2009). The use of technology for second language distance learning. *The Modern Language Journal, 93*, 822-835. https://doi.org/10.1111/j.1540-4781.2009.00975.x

Boelens, R., De Wever, B., & Voet, M. (2017). Four key challenges to the design of blended learning: A systematic literature review. *Educational Research Review, 22*, 1-18. https://doi.org/10.1016/j.edurev.2017.06.001

Bottou, L. (2010, August). Large-scale machine learning with stochastic gradient descent. In *Proceedings of COMPSTAT'2010: 19th International Conference on Computational Statistics* (pp. 177-186). COMPSTAT 2010: Paris, France. https://doi.org/10.1007/978-3-7908-2604-3_16

Bozkurt, A., Jung, I., Xiao, J., Vladimirschi, V., Schuwer, R., Egorov, G., ... & Paskevicius, M. (2020). A global outlook to the interruption of education due to COVID-19 pandemic: Navigating in a time of uncertainty and crisis. *Asian Journal of Distance Education, 15*(1), 1-126. https://doi.org/10.5281/zenodo.3878572

Bradley, A. P. (1997). The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recognition, 30*(7), 1145-1159. https://doi.org/10.1016/S0031-3203(96)00142-2

Breiman, L. (2001). Random forests. *Machine learning, 45,* 5-32. https://doi.org/10.1023/A:1010933404324

Brinkhuis, M. J. S., Savi, A. O., Hofman, A. D., Coomans, F., Van Der Maas, H. L. J., & Maris, G. (2018). Learning as it happens: A decade of analyzing and shaping a Large-Scale Online Learning System. *Journal of Learning Analytics, 5*(2). https://doi.org/10.18608/jla.2018.52.3

Brun, A., Bonnin, G., Castagnos, S., Roussanaly, A., & Boyer, A. (2019). Learning analytics made in France: The METAL project. *The International Journal of Information and Learning Technology, 36*(4), 299-313. https://doi.org/10.1108/IJILT-02-2019-0022

Brusilovsky, P., & Millán, E. (2007). User Models for Adaptive Hypermedia and Adaptive Educational Systems. In *Springer eBooks* (pp. 3–53). https://doi.org/10.1007/978-3-540-72079-9_1

Byrnes, H. (2006). *Advanced Language Learning : The Contribution of Halliday and Vygotsky.* https://doi.org/10.5040/9781474212113

Chafouk, O., & Marjanei, D. (2024). Comparing the Impact of Synchronous and Asynchronous Online Learning on Student Retention and Satisfaction. *European Journal of Education Studies, 11*(11). https://doi.org/10.46827/ejes.v11i11.5605

Chatti, M. A., & Muslim, A. (2019). The PERLA Framework: Blending Personalization and Learning Analytics. *The International Review of Research in Open and Distributed Learning, 20*(1). https://doi.org/10.19173/irrodl.v20i1.3936

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic Minority Over-sampling Technique. *Journal of Artificial Intelligence Research, 16,* 321–357. https://doi.org/10.1613/jair.953

Cheewaprakobkit, P. (2015). Predicting student academic achievement by using the decision tree and neural network techniques. *Catalyst, 12*(2), 34-43.

Chen, Y., Li, M., Huang, C., Cukurova, M., & Ma, Q. (2023). A Systematic Review of Research on Immersive Technology-Enhanced Writing Education: The Current State and A Research Agenda. *IEEE Transactions on Learning Technologies, 17,* 919–938. https://doi.org/10.1109/tlt.2023.3341420

Chen, X., Zou, D., Cheng, G., & Xie, H. (2020). Detecting latent topics and trends in educational technologies over four decades using structural topic modeling: A retrospective of all volumes of Computers & Education. *Computers & Education, 151*, 103855. https://doi.org/10.1016/j.compedu.2020.103855

Chiu, YL., Liang, JC. & Tsai, CC. Internet-specific epistemic beliefs and self-regulated learning in online academic information searching. *Metacognition Learning 8*, 235–260 (2013). https://doi.org/10.1007/s11409-013-9103-x

Christen, P., Hand, D. J., & Kirielle, N. (2023). A review of the F-measure: its history, properties, criticism, and alternatives. *ACM Computing Surveys, 56*(3), 1-24. https://doi.org/10.1145/3606367

Chu, Y., Hosseinalipour, S., Tenorio, E., Cruz, L., Douglas, K., Lan, A., & Brinton, C. (2022). Mitigating biases in student performance prediction via Attention-Based Personalized Federated Learning. *arXiv (Cornell University).* https://doi.org/10.48550/arxiv.2208.01182

Clark, P., & Niblett, T. (1989). The CN2 induction algorithm. *Machine Learning, 3*(4), 261-283. https://doi.org/10.1023/A:1022641700528

Clow, D. (2012). The Learning Analytics Cycle: Closing the Loop Effectively. In S. B. Shum, D. Gasevic, & R. Ferguson (Eds.), *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge (LAK12)* (pp. 134-138). New York: ACM. https://doi.org/10.1145/2330601.2330636

Coleman, J. A. (2006). English-medium teaching in European higher education. *Language Teaching, 39*(1), 1-14. https://doi.org/10.1017/S026144480600320X

Consoli, S., & Curle, S. (2024). Online Learner Engagement: Insights from a University Language Centre. *RELC Journal*. https://doi.org/10.1177/00336882241269603

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning, 20*, 273-297. https://doi.org/10.1007/BF00994018

Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. *IEEE Transactions on Information Theory, 13*(1), 21-27. https://doi.org/10.1109/TIT.1967.1053964

Crystal, D. (2003). *English as a global language.* Cambridge University Press.

Delen, D. (2011). Predicting student attrition with data mining methods. *Journal of College Student Retention: Research, Theory & Practice*, *13*(1), 17-35. https://doi.org/10.2190/CS.13.1.b

Demšar, J., Curk, T., Erjavec, A., Gorup, Č., Hočevar, T., Milutinovič, M., ... & Zupan, B. (2013). Orange: data mining toolbox in Python. *The Journal of Machine Learning Research, 14*(1), 2349-2353.

Dhawan, S. (2020). Online learning: A panacea in the time of COVID-19 crisis. *Journal of Educational Technology Systems, 49*(1), 5-22. https://doi.org/10.1177/0047239520934018

Domínguez, A., Saenz-De-Navarrete, J., De-Marcos, L., Fernández-Sanz, L., Pagés, C., & Martínez-Herráiz, J. (2013). Gamifying learning experiences: Practical implications and outcomes. *Computers & Education, 63*, 380–392. https://doi.org/10.1016/j.compedu.2012.12.020

Dörnyei, Z. (2014). *The psychology of the language learner: Individual differences in second language acquisition*. Routledge.

Drachsler, H., & Greller, W. (2016, April). Privacy and analytics: it's a DELICATE issue a checklist for trusted learning analytics. In *Proceedings of the sixth international conference on learning analytics & knowledge* (pp. 89-98).

Durlach, P. J., & Lesgold, A. M. (2012). Adaptive Technologies for Training and Education. *Cambridge University Press eBooks*. https://doi.org/10.1017/cbo9781139049580

Dutt, A., Ismail, M. A., & Herawan, T. (2017). A systematic review on educational data mining. *IEEE Access*, *5*(2017), 15991-16005. https://doi.org/ 10.1109/ACCESS.2017.2654247

Ellis, R. (2015). *Understanding Second Language Acquisition (Second Edition)*. Oxford: Oxford University Press.

Elnour, A. A., Abou Hajal, A., Goaddar, R., Elsharkawy, N., Mousa, S., Dabbagh, N., ... & Sadeq, A. (2023). Exploring the pharmacy students' perspectives on off-campus online learning experiences amid COVID-19 crises: A cross-sectional survey. *Saudi Pharmaceutical Journal, 31*(7), 1339-1350. https://doi.org/10.1016/j.jsps.2023.05.024

Etikan, I., Musa, S. A., & Alkassim, R. S. (2016). Comparison of convenience sampling and purposive sampling. *American Journal of Theoretical and Applied Statistics, 5*(1), 1-4. https://doi.org/10.11648/j.ajtas.20160501.11

Ferguson, R. (2012). Learning analytics: drivers, developments and challenges. *International Journal of Technology Enhanced Learning, 4*(5-6), 304-317. https://doi.org/10.1504/IJTEL.2012.051816

Ferguson, R., & Clow, D. (2017, March). Where is the evidence?. In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference* (pp. 56-65). https://doi.org/10.1145/3027385.3027396

Fidas, C. A., Belk, M., Constantinides, A., Portugal, D., Martins, P., Pietron, A. M., Pitsillides, A., & Avouris, N. (2023). Ensuring academic integrity and trust in online learning environments: A longitudinal study of an AI-Centered proctoring system in tertiary educational institutions. *Education Sciences, 13*(6), 566. https://doi.org/10.3390/educsci13060566

Foung, D., & Chen, J. (2019). A learning analytics approach to the evaluation of an online learning package in a Hong Kong University. *Electronic Journal of e-Learning, 17*(1), 11-24.

Freund, Y., & Schapire, R. E. (1997). A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences, 55*(1), 119-139. https://doi.org/10.1006/jcss.1997.1504

Friedman, J. H. (2001). Greedy Function Approximation: A Gradient Boosting Machine. *The Annals of Statistics, 29*(5), 1189-1232. http://www.jstor.org/stable/2699986

Ganapathy, M., & Kaur, S. (2014). ESL Students' Perceptions of the use of Higher Order Thinking Skills in English Language Writing. *Advances in Language and Literary Studies, 5*(5). https://doi.org/10.7575/aiac.alls.v.5n.5p.80

Gardner, J., & Brooks, C. (2018). Student success prediction in MOOCs. *User Modeling and User-Adapted Interaction, 28,* 127-203. https://doi.org/10.1007/s11257-018-9203-z

Garrison, D. R., Anderson, T., & Archer, W. (1999). Critical inquiry in a text-based environment: Computer conferencing in higher education. *The Internet and Higher Education, 2*(2-3), 87-105. https://doi.org/10.1016/S1096-7516(00)00016-6

Gašević, D., Dawson, S., & Siemens, G. (2015). Let's not forget: Learning analytics are about learning. *TechTrends, 59,* 64-71. https://doi.org/10.1007/s11528-014-0822-x

Gašević, D., Kovanović, V., & Joksimović, S. (2017). Piecing the learning analytics puzzle: a consolidated model of a field of research and practice. *Learning Research and Practice,* *3*(1), 63–78. https://doi.org/10.1080/23735082.2017.1286142

Gorjian, B., Moosavinia, S. R., Kavari, K. E., Asgari, P., & Hydarei, A. (2011). The impact of asynchronous computer-assisted language learning approaches on English as a foreign language high and low achievers' vocabulary retention and recall. *Computer Assisted Language Learning, 24*(5), 383–391. https://doi.org/10.1080/09588221.2011.552186

Greller, W., & Drachsler, H. (2012). Translating Learning into Numbers: A Generic Framework for Learning Analytics. *Educational Technology & Society, 15*(3), 42–57.

Greenhow, C., Lewin, C., & Staudt Willet, K. B. (2020). The educational response to Covid-19 across two countries: a critical examination of initial digital pedagogy adoption. *Technology, Pedagogy and Education, 30*(1), 7–25. https://doi.org/10.1080/1475939X.2020.1866654

Gunasekara, S., & Saarela, M. (2024). Explainability in educational data mining and learning analytics: An umbrella review. In *International Conference on Educational Data Mining*. International Educational Data Mining Society.

Gursoy, M. E., Inan, A., Nergiz, M. E., & Saygin, Y. (2016). Privacy-Preserving Learning Analytics: Challenges and Techniques. *IEEE Transactions on Learning Technologies,* *10*(1), 68–81. https://doi.org/10.1109/tlt.2016.2607747

Hampel, R., & Stickler, U. (2015). Developing Online Language Teaching. *Palgrave Macmillan UK eBooks*. https://doi.org/10.1057/9781137412263

Hasan, N., & Bao, Y. (2020). Impact of "e-Learning crack-up" perception on psychological distress among college students during COVID-19 pandemic: A mediating role of "fear of academic year loss". *Children and Youth Services Review, 118,* 105355. https://doi.org/10.1016/j.childyouth.2020.105355

Hendijani, Z. D. (2025). Learning English through Online Platforms: An Analysis of Language Learning Methods via Websites and Educational Platforms like Coursera and edX Compared to Traditional Teaching Methods. *TMP Universal Journal of Research and Review Archives, 4*(1). https://doi.org/10.69557/ujrra.v4i1.151

Heo, J., & Han, S. (2022). Challenges and problems on self-directed learning readiness in non–face-to-face educational settings during COVID-19. *OBM Neurobiology, 06(*04), 1-11. https://doi.org/10.21926/obm.neurobiol.2204141

Herodotou, C., Hlosta, M., Boroowa, A., Rienties, B., Zdrahal, Z., & Mangafa, C. (2019). Empowering online teachers through predictive learning analytics. *British Journal of Educational Technology, 50*(6), 3064-3079. https://doi.org/10.1111/bjet.12853

Hodges, C. B., Moore, S., Lockee, B. B., Trust, T., & Bond, M. A. (2024). The Difference between Emergency Remote Teaching and Online Learning. In *BRILL eBooks* (pp. 511–522). https://doi.org/10.1163/9789004702813_021

Hoel, T., & Chen, W. (2019). Privacy engineering for learning analytics in a global market. *International Journal of Information and Learning Technology, 36*(4), 288–298. https://doi.org/10.1108/ijilt-02-2019-0025

Holden, O. L., Norris, M. E., & Kuhlmeier, V. A. (2021). Academic Integrity in Online Assessment: A Research review. *Frontiers in Education, 6*. https://doi.org/10.3389/feduc.2021.639814

Holmes, G., Donkin, A., & Witten, I. H. (1994, November). Weka: A machine learning workbench. In *Proceedings of ANZIIS'94-Australian New Zealnd Intelligent Information Systems Conference* (pp. 357-361). IEEE. https://doi.org/10.1109/ANZIIS.1994.396988

Horwitz, E. (2001). Language anxiety and achievement. *Annual Review of Applied Linguistics, 21*, 112-126. https://doi.org/10.1017/S0267190501000071

Hosmer Jr, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied Logistic Regression*. John Wiley & Sons. https://doi.org/10.1002/9781118548387

Hrastinski, S. (2008). Asynchronous and synchronous e-learning. *Educause quarterly, 31*(4), 51-55.

Huda, M. C., Janattaka, N., & Prayoga, N. A. (2023). Exploring the Use of Online Resources for English Language Learning: Students' Perspectives. *English Journal Literacy UTama, 8*(1), 794–802. https://doi.org/10.33197/ejlutama.v8i1.231

Ifenthaler, D., Mah, D. K., & Yau, J. Y. K. (2019). Utilising Learning Analytics for Study Success: Reflections on Current Empirical Findings. In *Utilizing Learning Analytics to Support Study Success* (pp. 27-36). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-64792-0_2

Ifenthaler, D., & Widanapathirana, C. (2014). Development and validation of a learning analytics framework: Two case studies using support vector machines. *Technology, Knowledge and Learning, 19*, 221-240. https://doi.org/10.1007/s10758-014-9226-4

Ifenthaler, D., & Yau, J. Y. K. (2020). Utilising learning analytics to support study success in higher education: a systematic review. *Educational Technology Research and Development, 68*(4), 1961-1990. https://doi.org/10.1007/s11423-020-09788-z

Janke, S., Rudert, S. C., Petersen, Ä., Fritz, T. M., & Daumiller, M. (2021). Cheating in the wake of COVID-19: How dangerous is ad-hoc online testing for academic integrity?. *Computers and Education Open, 2*, 100055. https://doi.org/10.1016/j.caeo.2021.100055

Jayaprakash, S. M., Moody, E. W., Lauría, E. J., Regan, J. R., & Baron, J. D. (2014). Early alert of academically at-risk students: An open source analytics initiative. *Journal of Learning Analytics, 1*(1), 6-47. https://doi.org/10.18608/jla.2014.11.3

Jenkins, B. D., Golding, J. M., Le Grand, A. M., Levi, M. M., & Pals, A. M. (2023). When Opportunity Knocks: College Students' Cheating Amid the COVID-19 Pandemic. *Teaching of Psychology, 50*(4), 407-419. https://doi.org/10.1177/00986283211059067

Junco, R., & Clem, C. (2015). Predicting course outcomes with digital textbook usage data. *Internet and Higher Education, 27*, 54-63. https://doi.org/10.1016/j.iheduc.2015.06.001

Kabra, R. R., & Bichkar, R. S. (2011). Performance prediction of engineering students using decision trees. *International Journal of Computer Applications, 36*(11), 8-12.

Kahu, E. R., & Nelson, K. (2018). Student engagement in the educational interface: Understanding the mechanisms of student success. *Higher Education Research & Development, 37*(1), 58-71. https://doi.org/10.1080/07294360.2017.1344197

Kamraju, M., Krishnaiah, J., Durgesham, G., Shaba, N., Begum, S. A., Fatima, N., & Madhuri, Y. (2024). Exploring the Impact of Online Education on Higher Education. *ASEAN Journal of Educational Research and Technology, 3*(1), 27-36.

Kay, D., Korn, N., & Oppenheim, C. (2012). Legal, Risk and Ethical Aspects of Analytics in Higher Education.

Keržić, D., Aristovnik, A., Tomaževič, N., & Umek, L. (2019). Assessing the impact of students' activities in e-courses on learning outcomes: a data mining approach. *Interactive Technology and Smart Education, 16*(2), 117-129. https://doi.org/10.1108/ITSE-09-2018-0069

Khalil, M., Prinsloo, P., & Slade, S. (2023). The use and application of learning theory in learning analytics: A scoping review. *Journal of Computing in Higher Education, 35*(3), 573-594. https://doi.org/10.1007/s12528-022-09340-3

Kieffer, M. J. (2012). Early oral language and later reading development in Spanish-speaking English language learners: Evidence from a nine-year longitudinal study. *Journal of Applied Developmental Psychology, 33*(3), 146-157. https://doi.org/10.1016/j.appdev.2012.02.003

Kim, M., Kim, S., Lee, S., Yoon, Y., Myung, J., Yoo, H., ... & Lee, T. Y. (2024). LLM-Driven Learning Analytics Dashboard for Teachers in EFL Writing Education. *arXiv preprint*. https://doi.org/10.48550/arXiv.2410.15025

Klein, C., Lester, J., Rangwala, H., & Johri, A. (2019). Technological barriers and incentives to learning analytics adoption in higher education: Insights from users. *Journal of Computing in Higher Education, 31*, 604-625. https://doi.org/10.1007/s12528-019-09210-5

Kokoç, M., Akçapınar, G., & Hasnine, M. N. (2021). Unfolding students' online assignment submission behavioral patterns using temporal learning analytics. *Educational Technology & Society, 24*(1), 223-235. https://www.jstor.org/stable/26977869

Kovačić, Z. J. (2012). Predicting student success by mining enrolment data. *Research in Higher Education Journal, 15*, 1.

Kuh, G. D., Kinzie, J. L., Buckley, J. A., Bridges, B. K., & Hayek, J. C. (2006). *What matters to student success: A Review of the Literature* (Vol. 8). Washington, DC: National Postsecondary Education Cooperative.

Larrabee Sønderlund, A., Hughes, E., & Smith, J. (2019). The efficacy of learning analytics interventions in higher education: A systematic review. *British Journal of Educational Technology, 50*(5), 2594-2618. https://doi.org/10.1111/bjet.12720

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521*(7553), 436-444. https://doi.org/10.1038/nature14539

Leppänen, L., Leinonen, J., Ihantola, P., & Hellas, A. (2017, September). Predicting academic success based on learning material usage. In *Proceedings of the 18th Annual Conference on Information Technology Education* (pp. 13-18). https://doi.org/10.1145/3125659.3125695

Lewis, D. D. (1998, April). Naive (Bayes) at forty: The independence assumption in information retrieval. In *European Conference on Machine Learning* (pp. 4-15). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/BFb0026666

Li, P., & Lan, Y.-J. (2022). Digital Language Learning (DLL): Insights from Behavior, Cognition, and the Brain. *Bilingualism: Language and Cognition, 25*(3), 361–378. https://doi.org/10.1017/S1366728921000353

Liaw, S. S., Huang, H. M., & Chen, G. D. (2007). Surveying instructor and learner attitudes toward e-learning. *Computers & Education*, *49*(4), 1066-1080. https://doi.org/10.1016/j.compedu.2006.01.001

Lim, D. H., Morris, M. L., & Kupritz, V. W. (2007). Online vs. blended learning: Differences in instructional outcomes and learner satisfaction. *Journal of Asynchronous Learning Networks, 11*(2), 27-42. http://doi.org/10.24059/olj.v11i2.1725

Lin, C. J., & Hwang, G. J. (2018). A learning analytics approach to investigating factors affecting EFL students' oral performance in a flipped classroom. *Journal of Educational Technology & Society, 21*(2), 205-219.

Liu, H., Wang, Y., & Wang, H. (2025). Exploring the mediating roles of motivation and boredom in basic psychological needs and behavioural engagement in English learning: a self-determination theory perspective. *BMC Psychology, 13*(1). https://doi.org/10.1186/s40359-025-02524-3

Lu, O. H. T., Huang, A. Y. Q., Huang, J. C. H., Lin, A. J. Q., Ogata, H., & Yang, S. J. H. (2018). Applying Learning Analytics for the Early Prediction of Students' Academic Performance in Blended Learning. *Journal of Educational Technology & Society, 21*(2), 220-232. http://www.jstor.org/stable/26388400

Lundberg, S. M., & Lee, S. (2017). A Unified Approach to Interpreting Model Predictions. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.1705.07874

Macfadyen, L. P., & Dawson, S. (2010). Mining LMS data to develop an "early warning system" for educators: A proof of concept. *Computers & Education*, *54*(2), 588-599. https://doi.org/10.1016/j.compedu.2009.09.008

Mah, D. K. (2016). Learning analytics and digital badges: Potential impact on student retention in higher education. *Technology, Knowledge and Learning, 21*, 285-305. https://doi.org/10.1007/s10758-016-9286-8

Martin, F., & Whitmer, J. C. (2016). Applying learning analytics to investigate timed release in online learning. *Technology, Knowledge and Learning*, *21*, 59-74. https://doi.org/10.1007/s10758-015-9261-9

Matcha, W., Uzir, N. A., Gasevic, D., & Pardo, A. (2020). A Systematic Review of Empirical Studies on Learning Analytics Dashboards: A Self-Regulated Learning Perspective. IEEE Transactions on Learning Technologies, 13(2), 226–245. https://doi.org/10.1109/TLT.2019.2916802

Mayer, R. E. (2009). *Multimedia Learning* (2nd ed.). Cambridge: Cambridge University Press.

Mayer, R. E. (2024). The past, present, and future of the Cognitive Theory of Multimedia Learning. *Educational Psychology Review, 36*(1). https://doi.org/10.1007/s10648-023-09842-1

Meng, W., Yu, L., Liu, C., Pan, N., Pang, X., & Zhu, Y. (2024, January). A systematic review of the effectiveness of online learning in higher education during the COVID-19 pandemic period. In *Frontiers in Education* (Vol. 8, p. 1334153). Frontiers Media SA. https://doi.org/10.3389/feduc.2023.1334153

Mirriahi, N., & Dawson, S. (2013, April). The pairing of lecture recording data with assessment scores: a method of discovering pedagogical impact. In *Proceedings of the Third International Conference on Learning Analytics and Knowledge* (pp. 180-184). https://doi.org/10.1145/2460296.2460331

Mishra, T., Kumar, D., & Gupta, S. (2014, February). Mining students' data for performance prediction. In *Fourth International Conference on Advanced Computing & Communication Technologies* (pp. 255-262). https://doi.org/10.1109/ACCT.2014.105

Motz, B. A., Carvalho, P. F., de Leeuw, J. R., & Goldstone, R. L. (2018). Embedding experiments: Staking causal inference in authentic educational contexts. *Journal of Learning Analytics, 5*(2), 47-59. https://doi.org/10.18608/jla.2018.52.4

Moubayed, A., Injadat, M., Shami, A., & Lutfiyya, H. (2018, March). Relationship between student engagement and performance in e-learning environment using association rules. In *2018 IEEE world engineering education conference (EDUNINE)* (pp. 1-6). IEEE. https://doi.org/10.1109/EDUNINE.2018.8451005

Müller, U., Hancock, D. R., Wang, C., Stricker, T., Cui, T., & Lambert, M. (2022). School leadership, Education for Sustainable Development (ESD), and the impact of the COVID-19 pandemic: Perspectives of principals in China, Germany, and the USA. *Education Sciences, 12*(12), 853. https://doi.org/10.3390/educsci12120853

Naik, A., & Samant, L. (2016). Correlation review of classification algorithm using data mining tool: WEKA, Rapidminer, Tanagra, Orange and Knime. *Procedia Computer Science, 85*, 662-668. https://doi.org/10.1016/j.procs.2016.05.251

Nigam, A., Pasricha, R., Singh, T., & Churi, P. (2021). A Systematic Review on AI-based Proctoring Systems: Past, Present and Future. *Education and Information Technologies, 26*(5), 6421–6445. https://doi.org/10.1007/s10639-021-10597-x

Noor, S., Isa, F. M., & Mazhar, F. F. (2020). Online teaching practices during the COVID-19 pandemic. *Educational Process: International Journal, 9*(3), 169-184. https://doi.org/10.22521/edupij.2020.93.4

Oxford, R. L. (2017). *Teaching and researching language learning strategies: Self-regulation in context.* Routledge.

Pardo, A., & Siemens, G. (2014). Ethical and privacy principles for learning analytics. *British Journal of Educational Technology*, *45*(3), 438-450. https://doi.org/10.1111/bjet.12152

Park, Y., & Jo, I. H. (2019). Factors that affect the success of learning analytics dashboards. *Educational Technology Research and Development, 67*(6), 1547-1571. https://doi.org/10.1007/s11423-019-09693-0

Perrotta, C., & Williamson, B. (2018). The social life of learning analytics: Cluster analysis and the 'performance' of algorithmic education. *Learning, Media and Technology,* *43*(1), 3-16. https://doi.org/10.1080/17439884.2016.1182927

Plass, J. L., & Jones, L. C. (2005). Multimedia Learning in Second Language Acquisition. In R. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (pp. 467-488). Cambridge: Cambridge University Press.

Powers, D. M. (2020). Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation. *arXiv preprint arXiv:2010.16061.*

Prensky, M. (2001). Digital Natives, Digital Immigrants Part 1. *On The Horizon the International Journal of Learning Futures,* *9*(5), 1–6. https://doi.org/10.1108/10748120110424816

Quinlan, J. R. (1986). Induction of decision trees. *Machine Learning*, *1*, 81-106. https://doi.org/10.1007/BF00116251

Reinders, H. (2018). Learning analytics for language learning and teaching. *Jalt Call Journal*, *14*(1), 77-86.

Resnik, D. B. (2018). The ethics of research with human subjects: *Protecting People, Advancing Science, Promoting Trust* (Vol. 74). Springer. https://doi.org/10.1007/978-3-319-68756-8

Rofiah, N. L., Sha'ar, M. Y. M. A., & Waluyo, B. (2022). Digital divide and factors affecting English synchronous learning during Covid-19 in Thailand. *International Journal of Instruction, 15*(1), 633-652.

Romero, C., & Ventura, S. (2007). Educational data mining: A survey from 1995 to 2005. *Expert Systems with Applications,* *33*(1), 135-146. https://doi.org/10.1016/j.eswa.2006.04.005

Romero, C., & Ventura, S. (2010). Educational data mining: a review of the state of the art. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (applications and reviews), 40*(6), 601-618. https://doi.org/10.1109/TSMCC.2010.2053532

Romero, C., & Ventura, S. (2013). Data mining in education. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 3*(1), 12-27. https://doi.org/10.1002/widm.1075

Romero, C., & Ventura, S. (2017). Educational data science in massive open online courses. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 7*(1), e1187. https://doi.org/10.1002/widm.1187

Romero, C., & Ventura, S. (2020). Educational data mining and learning analytics: An updated survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 10*(3), e1355. https://doi.org/10.1002/widm.1355

Romero-Zaldivar, V. A., Pardo, A., Burgos, D., & Kloos, C. D. (2012). Monitoring student progress using virtual appliances: A case study. *Computers & Education, 58*(4), 1058-1067. https://doi.org/10.1016/j.compedu.2011.12.003

Roy, S., & Garg, A. (2017, October 26-28). *Analyzing performance of students by using data mining techniques a literature survey* [Paper presentation]. 4th IEEE Uttar Pradesh Section International Conference on Electrical, Computer and Electronics (UPCON), GLA University, Mathura, India (pp. 130-133). IEEE. https://doi.org/10.1109/UPCON.2017.8251035

Rubel, A., & Jones, K. M. (2016). Student privacy in learning analytics: An information ethics perspective. *The Information Society, 32*(2), 143-159. https://doi.org/10.1080/01972243.2016.1130502

Sclater, N., Peasgood, A., & Mullan, J. (2016). Learning analytics in higher education. *London: Jisc. Accessed October, 30*(2024), 176.

Seufert, S., Meier, C., Soellner, M., & Rietsche, R. (2019). A pedagogical perspective on big data and learning analytics: A conceptual model for digital learning support. *Technology, Knowledge and Learning, 24*, 599-619. https://doi.org/10.1007/s10758-019-09399-5

Siemens, G. (2013). Learning analytics: The emergence of a discipline. *American Behavioral Scientist, 57*(10), 1380-1400. https://doi.org/10.1177/0002764213498851

Shafiq, D. A., Marjani, M., Habeeb, R. A. A., & Asirvatham, D. (2022). Student retention using educational data mining and predictive analytics: a systematic literature review. *IEEE Access, 10*, 72480-72503. https://doi.org/10.1109/ACCESS.2022.3188767

Shamsi, M. S., & Lakshmi, J. (2016). A Comparative Analysis of classification data mining techniques : Deriving key factors useful for predicting students performance. *arXiv (Cornell University).* https://doi.org/10.48550/arxiv.1606.05735

Siemens, G., & Baker, R. S. D. (2012, April). Learning analytics and educational data mining: towards communication and collaboration. In *Proceedings of the 2nd international conference on learning analytics and knowledge* (pp. 252-254). https://doi.org/10.1145/2330601.2330661

Siemens, G., & Long, P. (2011). Penetrating the fog: Analytics in learning and education. *EDUCAUSE Review, 46*(5), 30.

Simon, P. D., & Zeng, L. M. (2024). Behind the Scenes of Adaptive Learning: A scoping review of teachers' perspectives on the use of adaptive learning technologies. *Education Sciences, 14*(12), 1413. https://doi.org/10.3390/educsci14121413

Slade, S., & Prinsloo, P. (2013). Learning analytics: Ethical issues and dilemmas. *American Behavioral Scientist, 57*(10), 1510-1529. https://doi.org/10.1177/0002764213479366

Smith, V. C., Lange, A., & Huston, D. R. (2012). Predictive modeling to forecast student outcomes and drive effective interventions in online community college courses. *Journal of Asynchronous Learning Networks, 16*(3), 51-61.

Sokolova, M., & Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. *Information Processing & Management, 45*(4), 427-437. https://doi.org/10.1016/j.ipm.2009.03.002

Strang, K. D. (2017). Beyond engagement analytics: Which online mixed-data factors predict student learning outcomes?. *Education and Information Technologies, 22*(3), 917-937. https://doi.org/10.1007/s10639-016-9464-2

Suharti, D. S., Suherdi, D., & Setyarini, S. (2021, April). Exploring students' learning engagement in EFL online classroom. *In Thirteenth Conference on Applied Linguistics (CONAPLIN 2020)* (pp. 139-149). Atlantis Press. https://doi.org/10.2991/assehr.k.210427.022

Sukhija, K., Jindal, M., & Aggarwal, N. (2015, October). The recent state of educational data mining: A survey and future visions. In *2015 IEEE 3rd International Conference on MOOCs, Innovation and Technology in Education (MITE)* (pp. 354-359). IEEE. https://doi.org/10.1109/MITE.2015.7375344

Sun, J. C. Y., & Rueda, R. (2012). Situational interest, computer self-efficacy and self-regulation: Their impact on student engagement in distance education. *British Journal of Educational Technology, 43* (2), 191-204. https://doi.org/10.1111/j.1467-8535.2010.01157.x

Takase, N. (2024). Synchronous and asynchronous online communication for developing foreign language speaking skills. *AsiaCALL Online Journal, 15*(2), 60-76. https://doi.org/10.54855/acoj.241525

Tempelaar, D. T., Niculescu, A., Rienties, B., Gijselaers, W. H., & Giesbers, B. (2012). How achievement emotions impact students' decisions for online learning, and what precedes those emotions. *The Internet and Higher Education, 15*(3), 161-169. https://doi.org/10.1016/j.iheduc.2011.10.003

Tempelaar, D. T., Rienties, B., & Giesbers, B. (2015). In search for the most informative data for feedback generation: *Learning analytics in a data-rich context. Computers in Human Behavior, 47*, 157-167. https://doi.org/10.1016/j.chb.2014.05.038

Tempelaar, D., Rienties, B., & Nguyen, Q. (2021). Dispositional Learning Analytics for Supporting Individualized Learning Feedback. *Frontiers in Education, 6.* https://doi.org/10.3389/feduc.2021.703773

Tharwat, A. (2021). Classification assessment methods. *Applied Computing and Informatics, 17*(1), 168-192. https://doi.org/10.1016/j.aci.2018.08.003

Thomas, M., Reinders, H., & Gelan, A. (2017). Learning Analytics in online language learning. In *Routledge eBooks* (pp. 197–212). https://doi.org/10.4324/9781315205618-13

Tlili, A., Hattab, S., Essalmi, F., Chen, N. S., Huang, R., Chang, M., & Solans, D. B. (2021). A smart collaborative educational game with learning analytics to support English vocabulary teaching. *International Journal of Interactive Multimedia and Artificial Intelligence, 6*(6), 215-224. https://doi.org/10.9781/ijimai.2021.03.002

Tseng, J. C., Chu, H. C., Hwang, G. J., & Tsai, C. C. (2008). Development of an adaptive learning system with two sources of personalization information. *Computers & Education, 51*(2), 776-786. https://doi.org/10.1016/j.compedu.2007.08.002

Ushioda, E. (2011). Language learning motivation, self and identity: Current theoretical perspectives. *Computer Assisted Language Learning, 24*(3), 199-210. https://doi.org/10.1080/09588221.2010.538701

Valle, N., Antonenko, P., Dawson, K., & Huggins-Manley, A. C. (2021). Staying on target: A systematic literature review on learner-facing learning analytics dashboards. *British Journal of Educational Technology, 52*(4), 1724-1748. https://doi.org/10.1111/bjet.13089

Van Dijk, J. (2017). Digital divide: Impact of access. *The International Encyclopedia of Media Effects, 1,* 1–11. https://doi.org/10.1002/9781118783764.wbieme0043

Van Leeuwen, A. (2019). Teachers' perceptions of the usability of learning analytics reports in a flipped university course: When and how does information become actionable knowledge?. *Educational Technology Research and Development, 67*, 1043-1064. https://doi.org/10.1007/s11423-018-09639-y

Vandamme, J. P., Meskens, N., & Superby, J. F. (2007). Predicting academic performance by data mining methods. *Education Economics, 15*(4), 405. https://doi.org/10.1080/09645290701409939

Varouchas, E., Sicilia, M. A., & Sánchez-Alonso, S. (2018). Towards an integrated learning analytics framework for quality perceptions in higher education: A 3-tier content, process, engagement model for key performance indicators. *Behaviour & Information Technology, 37*(10-11), 1129-1141. https://doi.org/10.1080/0144929X.2018.1495765

Verbert, K., Manouselis, N., Drachsler, H., & Duval, E. (2012). Dataset-Driven Research to Support Learning and Knowledge Analytics. *Educational Technology & Society, 15*(3), 133-148.

Viberg, O., Hatakka, M., Bälter, O., & Mavroudi, A. (2018). The current landscape of learning analytics in higher education. *Computers in Human Behavior, 89*, 98-110. https://doi.org/10.1016/j.chb.2018.07.027

Viriya, C., & Sapsirin, S. (2014). Gender Differences in Language Learning Style and Language Learning Strategies. *Indonesian Journal of Applied Linguistics, 3*(2), 77. https://doi.org/10.17509/ijal.v3i2.270

West, D., Huijser, H., & Heath, D. (2016). Putting an ethical lens on learning analytics. *Educational Technology Research and Development, 64*, 903-922. https://doi.org/10.1007/s11423-016-9464-3

West, D., Tasir, Z., Luzeckyj, A., Si Na, K., Toohey, D., Abdullah, Z., Searle, B., Farhana Jumaat, N., & Price, R. (2018). Learning analytics experience among academics in Australia and Malaysia: A comparison. *Australasian Journal of Educational Technology, 34*(3). https://doi.org/10.14742/ajet.3836

Willis, J. E., Slade, S., & Prinsloo, P. (2016). Ethical oversight of student data in learning analytics: a typology derived from a cross-continental, cross-institutional perspective. *Educational Technology Research and Development, 64*(5), 881-901. https://doi.org/10.1007/s11423-016-9463-4

Yaseen, M. G., & Alnakeeb, S. S. (2023). Exploring the evolution of AI integration in English as a Foreign Language Education: A Scopus-Based Bibliometric Analysis (1997-2023). *Mesopotamian Journal of Computer Science*, 149-164. https://doi.org/10.58496/mjcsc/2023/019

Yin, C., & Hwang, G. J. (2018). Roles and Strategies of Learning Analytics in the e-Publication Era. *Knowledge Management & E-Learning, 10*(4), 455-468. https://doi.org/10.34105/j.kmel.2018.10.028

You, J. W. (2016). Identifying significant indicators using LMS data to predict course achievement in online learning. *Internet and Higher Education*, *29*, 23-30. https://doi.org/10.1016/j.iheduc.2015.11.003

Zhang, J., Zhang, X., Jiang, S., Ordóñez de Pablos, P., & Sun, Y. (2018). Mapping the study of learning analytics in higher education. *Behaviour & Information Technology*, *37*(10-11), 1142-1155. https://doi.org/10.1080/0144929X.2018.1529198

Zilvinskis, J., Willis III, J., & Borden, V. M. (2017). An overview of learning analytics. *New Directions for Higher Education*, *2017*(179), 9-17. https://doi.org/10.1002/he.20239

Zimmerman, B. J. (2002). Becoming a Self-Regulated Learner: An Overview. *Theory into Practice*, *41*(2), 64-70. https://doi.org/10.1207/s15430421tip4102_2

APPENDICES

A. Ethics Approval Form

UYGULAMALI ETİK ARAŞTIRMA MERKEZİ
APPLIED ETHICS RESEARCH CENTER

ORTA DOĞU TEKNİK ÜNİVERSİTESİ
MIDDLE EAST TECHNICAL UNIVERSITY

DUMLUPINAR BULVARI 06800
ÇANKAYA ANKARA/TURKEY
T: +90 312 210 22 91
F: +90 312 210 79 59
ueam@metu.edu.tr
www.ueam.metu.edu.tr

Sayı: 28620816 /

22 MART 2021

Konu : Değerlendirme Sonucu

Gönderen: ODTÜ İnsan Araştırmaları Etik Kurulu (İAEK)

İlgi : İnsan Araştırmaları Etik Kurulu Başvurusu

Sayın Prof. Dr. Ömer DELİALİOĞLU

Danışmanlığını yürüttüğünüz Mehmet Ali ÇELİKBAĞ'ın "İngilizce yeterlilik sınavları bağlamında çevrim içi yabancı dil öğreniminde veri madenciliği" başlıklı araştırması İnsan Araştırmaları Etik Kurulu tarafından uygun görülmüş ve **075-ODTU-2021** protokol numarası ile onaylanmıştır.

Saygılarımızla bilgilerinize sunarız.

Dr.Öğretim Üyesi Ali Emre TURGUT
İAEK Başkan Vekili

CURRICULUM VITAE

PERSONAL INFORMATION

Surname, Name: Çelikbağ, Mehmet Ali
Nationality: Turkish

EDUCATION

Degree	Institution	Year of Graduation
PhD	METU, Computer Education and Instructional Technology	2025
MA	Bilkent University, Comp. and Inst. Tech. Teacher Education	2012
BA	Bilkent University, Comp. and Inst. Tech. Teacher Education	2011
AD	Muğla Sıtkı Koçman University, Tourist Guidance	2023
AD	Ankara Yıldırım Beyazıt University, Aircraft Technology	2020

WORK EXPERIENCE

Year	Place	Enrollment
2014-Present	Ankara University	Lecturer
2012-2014	TUBİTAK ULAKBİM	Linux System Administrator

FOREIGN LANGUAGES

Advanced English, Intermediate Polish, Basic Greek, Basic Russian

PUBLICATIONS

1. Afacan Adanır, G., Muhametjanova, G., Celikbag, M. A., Omuraliev, A., İsmailova, R. (2020). Learners' preferences for online resources, activities, and communication tools: A comparative study of Turkey and Kyrgyzstan. *E-Learning and Digital Media.* *17*(2):148-166. https://doi.org/10.1177/2042753019899713.
2. Çelikbağ, M. A., Delialioğlu, Ö. (2021). Proctored vs Unproctored Online Exams in Language Courses: A Comparative Study. *Proceedings of the 29th International Conference on Computers in Education (ICCE)* (pp. 533-538). https://icce2021.apsce.net/proceedings/volume1

HOBBIES

Computer and Internet Technologies, Learning and Teaching, Travelling, Photography, Basketball, Tennis, Aviation, History, Mythology, Archeology