MACHINE LEARNING BASED INDOOR AIR POLLUTANT SOURCE RECOGNITION WITH GAS RESISTANCE AND MULTI-SENSOR ARRAY ELECTRONIC NOSES

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

BY

MEHMET YİĞİTCAN YEŞİLATA

IN PARTIAL FULFILMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
BUILDING SCIENCE IN ARCHITECTURE

FEBRUARY 2022

Approval of the thesis:

MACHINE LEARNING BASED INDOOR AIR POLLUTANT SOURCE RECOGNITION WITH GAS RESISTANCE AND MULTI-SENSOR ARRAY ELECTRONIC NOSES

submitted by **MEHMET YİĞİTCAN YEŞİLATA** in partial fulfillment of the requirements for the degree of Master of Science **in** Building Science in Architecture Department**, Middle East Technical University** by,

Prof. Dr. Halil Kalıpçılar
Dean, Graduate School of **Natural and Applied Sciences** __________________

Prof. Dr. Fatma Cana Bilsel
Head of the Department, **Building Science in Architecture** __________________

Asst. Prof. Dr. Mehmet Koray Pekeriçli
Supervisor, Architecture Department, METU __________________

Examining Committee Members:

Assoc. Prof. Dr. Ayşe Tavukçuoğlu
Architecture Department, METU __________________

Asst. Prof. Dr. Mehmet Koray Pekeriçli
Architecture Department, METU __________________

Asst. Prof. Dr. Aktan Acar
Architecture Department, TOBB ETU __________________

Date: 10.02.2022

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.

Yeşilata, Mehmet Yiğitcan

Signature :

ABSTRACT

MACHINE LEARNING BASED INDOOR AIR POLLUTANT SOURCE RECOGNITION WITH MULTI-SENSOR ARRAY ELECTRONIC NOSE

Yeşilata, Mehmet Yiğitcan
M.Sc. in Building Science, Department of Architecture
Supervisor: Asst. Prof. Dr. Mehmet Koray Pekeriçli

10 February 2022, 101 pages

Indoor Air Quality (IAQ) is closely linked to health and well-being. Humans spend the majority of their time indoors. Breathable air that is free of harmful pollutants can result in an improved quality of life, a decreased risk of respiratory infections, and a decreased risk of developing chronic conditions. Cleaning chemicals, construction operations, smoking, perfumes, building materials and outdoor pollutants can all contribute to indoor air pollution. Detecting the sources of pollution is essential in order to improve interior air quality. A sensing device known as an electronic nose collects various sensor data to detect scents or flavors. Two different types of electronic nose are used in this research to collect data. The first one has 8 VOC sensors that operate in response to the gas resistance of the MOX layer. Each sensor is simultaneously heated by a different heater profile, and their reaction to the gas produces an 8 dimensional sensitivity that is proportional to the sensor's gas resistance at that temperature. The second one collects different types of parameters that affect indoor air quality like PM2.5, PM10, CO2, Formaldehyde, Ethanol, H2, TVOC, Temperature and Humidity with a multi sensor array. The experiment was conducted in a 130-liter box. Throughout the experiment, nine different materials were measured, including office air, smoking, cleaning materials, alcohol-sanitizer, curry, coffee, painted tile, stone wool, and varnished wood with two different electronic noses. Data has been modelled with Naive Bayes, kNN, Random Forest Classifiers to predict pollutant sources in the box. Random Forest algorithm with 10 trees gives the best result with data collected from multi sensor array based electronic nose. The algorithm gives 0.95 accuracy with 0.94 precision on the sensor data. The field experiments demonstrate that detecting pollutant sources using a multisensor array-based electronic nose is feasible without requiring a complex infrastructure, and that this technology can be used as a reliable solution for pollutant source detection because it provides higher accuracy and precision results than existing approaches in the industry.

Keywords: Indoor Air Quality, Electronic Nose, Odor Detection, Machine Learning

ÖZ

MACHINE LEARNING BASED INDOOR AIR POLLUTANT SOURCE RECOGNITION WITH MULTI-SENSOR ARRAY ELECTRONIC NOSE

Yeşilata, Mehmet Yiğitcan
Tez Yöneticisi : Dr. Öğr. Üyesi Mehmet Koray Pekeriçli
10 Şubat 2022, 101 sayfa

İç mekan hava kalitesi (IAQ), sağlık ve konfor ile yakından ilişkilidir. İnsanlar zamanlarının çoğunu iç mekanlarda geçirirler. Zararlı kirleticiler içermeyen solunabilir hava, yaşam kalitesinin iyileşmesini, solunum yolu enfeksiyonu riskinin azalmasını ve kronik rahatsızlıkların gelişme riskinin azalmasını sağlayabilir. Temizlik kimyasalları, inşaat işlemleri, sigara, parfüm, yapı malzemeleri ve dış ortam kirleticilerinin tümü iç mekan hava kirliliğine katkıda bulunabilir. İç hava kalitesini iyileştirmek için kirlilik kaynaklarının tespiti esastır. Elektronik burun olarak bilinen bir algılama cihazı, kokuları veya tatları algılamak için çeşitli sensör verilerini toplar. Bu araştırmada veri toplamak için iki farklı tipte elektronik burun kullanılmıştır. İlki, MOX katmanındaki gaz direncine duyarlı olarak çalışan sekiz VOC sensörüne sahiptir. Her sensör eş zamanlı olarak farklı bir ısıtıcı profili tarafından ısıtılır ve sensörlerin gaza tepkileri, her sensörün o sıcaklıktaki gaz direnciyle orantılı olarak sekiz boyutlu bir veri üretir. İkincisi, çoklu sensör dizisi ile PM2.5, PM1 0, CO2, formaldehit, etanol, H2, TVOC, sıcaklık ve nem gibi iç hava kalitesini etkileyen farklı türde parametreleri ölçer. Deney gerçekleştirilirken 130lt kapasiteli bir kutu kullanıldı. Deney boyunca ofis havası, sigara, temizlik malzemeleri, alkol-dezenfektan, köri, kahve, boyalı kiremit, taş yünü ve vernikli ahşap olmak üzere iki farklı elektronik burun ile dokuz farklı malzeme ölçülmüştür. Kutudaki kirletici kaynakları tahmin etmek için veriler Naive Bayes, kNN ve Random Forest sınıflandırıcıları ile modellenmiştir. Çok sensörlü elektronik burundan toplanan verilerle 10 ağaçlı Random Forest algoritması en iyi sonucu verir. Algoritma, sensör verileri üzerinde 0.95 doğruluk ile 0.94 kesinlik sonuçlarını vermektedir. Saha deneyleri, karmaşık bir altyapı gerektirmeden çok sensörlü dizi tabanlı bir elektronik burun kullanarak kirletici

kaynaklarının tespit edilmesinin mümkün olduğunu ve bu teknolojinin endüstrideki mevcut yaklaşımlardan daha yüksek doğruluk ve kesinlik sonuçları sağladığı için kirletici kaynağı tespitinde güvenilir bir çözüm olarak kullanılabileceğini göstermektedir.

Anahtar Kelimeler: İç Hava Kalitesi, Elektronik Burun, Koku Tespiti, Makine Öğrenmesi

ACKNOWLEDGMENTS

First, I would like to express my deepest gratitude to my supervisor Asst. Prof. Dr. Koray Pekeriçli, for his guidance and support throughout this study.

I would like to thank my mentor and my father Bülent Yeşilata who never stops believing in me, and always supports me through the challenges of life. I would also like to thank my mom Mine Yeşilata who raised me with love and respect, always keeping me motivated for life. I'm grateful for having a lovely sister Serra Yeşilata who is one of my closest friends and biggest supporters.

I would like to thank Yeşim Erdoğan who has been always there for me, emotionally supporting me when I really needed to, never stopping loving me.

I would like to thank Fatih Aykut Yıldırım and Hasan Basri Tosun who are not only my colleagues, but also my companions that we share the joy and the pain together in life.

Lastly, I would like to thank my young and talented colleagues Zelal İrem Aldağ , Mustafa Algun and Carlos Garcia-Maurino who supported me during the experiment phase of the research. Also, I would like to thank my sweet friends İdil Güneş Mete and Furkan Ginaz Almus who supported me in having a well-written proposal.

TABLE OF CONTENTS

1 INTRODUCTION 1

1.1 Motivation 1

1.2 Aims and Objectives 3

2 LITERATURE REVIEW 5

2.1 Importance of Indoor Air Quality 5
2.1.1 Well-being and Health 5
2.1.2 Performance 7
2.1.3 Energy 8
2.1.4 Thermal Comfort 10
2.1.5 Building Envelope and Interior Design 11

2.2 Indoor Air Quality Parameters 12
2.2.1 CO2 12
2.2.2 VOC 12
2.2.3 PM2.5 13
2.2.4 PM10 14
2.2.5 Formaldehyde 15
2.2.6 Ozone 15
2.2.7 NO_2 15
2.2.8 Temperature and Humidity 16

2.3 Indoor Air Pollutant Sources 16
2.3.1 Building Materials 17
2.3.2 Office Equipment 17
2.3.3 Chemical Reactions 19
2.3.4 Air Conditioning 19
2.3.5 Smoking 19

2.4 Indoor Air Quality Sensing Instruments 20
2.4.1 Metal Oxide Gas Sensors (MOX) 20
2.4.2 Optical Gas Sensors 21
2.4.3 Electrochemical Gas Sensors 22

2.5 Indoor Air Quality Standards 22

2.6 Electronic Nose (EN) 24
2.6.1 History of Electronic Nose 25
2.6.2 Components of Electronic Nose 26
2.6.3 Electronic Nose Applications for Pollutant Source Recognition 29

3 MATERIAL AND METHOD 33

3.1 Gas Resistance Based Electronic Nose 33
3.1.1 Gas Measurement Principle 34
3.1.2 Configuration 35

3.2 Multi Sensor Array Based Electronic Nose 37
3.2.1 Hardware Development 37
3.2.2 Software Development 47
3.2.3 Calibration 48

3.3 Chosen Materials for Detection *50*

4 RESULTS AND DISCUSSION **57**

4.1 Experimental Results for the Gas Resistance Based Electronic Nose *57*
4.1.1 Transient Variation of Temperature 57
4.1.2 Transient Variation of Humidity 59
4.1.3 Transient Variation of Pressure 61
4.1.4 Transient Variation of Gas Resistance 62

4.2 Analysis of the Results with Different Classification Models for the Gas Resistance Based Electronic Nose *64*
4.2.1 The Naive Bayes (NB) Classification Model 64
4.2.2 The kNN Classification Model 66
4.2.3 The Random Forest (RF) Classification Model 68

4.3 Experimental Results for the Multi Sensor Array Based Electronic Nose *70*
4.3.1 Transient Variation of CO_2 70
4.3.2 Transient Variation of Formaldehyde (HCHO) 72
4.3.3 Transient Variation of PM2.5 74
4.3.4 Transient Variation of TVOC (Total Volatile Organic Compounds) 75
4.3.5 Transient Variation of Other parameters (H_2, Ethanol, Temperature, Humidity) . 77

4.3 Analysis of the Results with Different Classification Models for the Multi Sensor Array Based Electronic Nose *78*
4.4.1 The Naive Bayes (NB) Classification Model 78
4.4.2 The kNN Classification Model 81
4.4.3 Random Forest Classification Model 84

5 CONCLUSION **87**

REFERENCES **93**

LIST OF TABLES

TABLES

Table 2.1 Different Indoor Air Quality Standards for Different Parameters 22

Table 3.1 Overview of Measurement Capabilities 45

Table 3.2. Used Gases for Calibration of the TVOC Sensor 49

Table 3.3. Chosen Materials with Their Activity Classes 52

Table 4.1. Analysis of Experimental Temperature Measurements with the Gas Resistance Based Electronic Nose 58

Table 4.2. Analysis of Experimental Humidity Measurements with the Gas Resistance Based Electronic Nose 60

Table 4.3. Analysis of Gas Resistance Measurements with the Gas Resistance Based Electronic Nose 62

Table 4.4. Classification Report by the NB Classification Model for Experiments with the Gas Resistance Based Electronic Nose 64

Table 4.5. Classification Report by the kNN Classification Model for Experiments with the Gas Resistance Based Electronic Nose 66

Table 4.6. Classification Report by the RF n_tree=10 Classification Model for Experiments with the Gas Resistance Based Electronic Nose 69

Table 4.7. Analysis of Experimental CO2 (ppm) Measurements with the Multi Sensor Array Based Electronic Nose 71

Table 4.8. Analysis of Experimental HCHO (ppm) Measurements with the Multi Sensor Array Based Electronic Nose 73

Table 4.9. Analysis of Experimental PM2.5 (μg/m3) Measurements with the Multi Sensor Array Based Electronic Nose 74

Table 4.10. Analysis of Experimental TVOC (ppb) Measurements with the Multi Sensor Array Based Electronic Nose 76

Table 4.11. Classification Report by the NB Classification Model for Experiments with the Multi Sensor Array Based Electronic Nose 79

Table 4.12. Classification Report by the kNN Classification Model (k=8) for Experiments with the Multi Sensor Array Based Electronic Nose 82

Table 4.13. Classification Report for Experiment I n_tree=10, RFC model of Multi Sensor Array Based Electronic Nose 85

Table 5.1 Material- High Concentration Parameter Comparison 91

LIST OF FIGURES

FIGURES

Figure 2.1. Hybrid air distribution systems 10
Figure 2.2. Metal Oxide Gas Sensor Structure 20
Figure 2.3. Optical Gas Sensor Structure 21
Figure 2.4. Electrochemical Gas Sensor Structure 22
Figure 2.5. Analogy between Biological nose and Electronic Nose 25
Figure 2.6. The Logic of Electronic Nose 27
Figure 2.7. Classification with k-NN 28
Figure 2.8. Classification with SVM 28
Figure 3.1. BME Board x8 Schematic 34
Figure 3.2. Working Principle of BME688 35
Figure 3.3. Heater Profile Schedule of Gas Sensor 36
Figure 3.4. Arduino Mega2560 38
Figure 3.5. Working Principle of NDIR Technology 39
Figure 3.6. Sensirion SCD30 40
Figure 3.7. Sensirion SPS30 Particulate Matter 40
Figure 3.8. Internal Structure of SPS30 41
Figure 3.9. Dust Prevention of SPS30 42
Figure 3.10. DFRobot Gravity HCHO Sensor 43
Figure 3.11. SGP30 VOC Sensor 44
Figure 3.12. NodeMCU 45
Figure 3.13. Multi Sensor Array Based Electronic Nose 46
Figure 3.14. System Architecture of Multi Sensor Array Based Electronic Nose 47
Figure 3.15. Experimental Setup with Different Materials 53
Figure 4.1. Temperature Change Over Time for the Gas Resistance Based Electronic Nose 58
Figure 4.2. Temperature Change Levels of Different Materials for the Gas Resistance Based Electronic Nose 59
Figure 4.3. Humidity Change Over Time for the Gas Resistance

Based Electronic Nose 60
Figure 4.4. Humidity Change Levels of Different Materials
for the Gas Resistance Based Electronic Nose 61
Figure 4.5. Pressure Change Over Time for Gas Resistance
Based Electronic Nose 62
Figure 4.6. Gas Resistance Parameter Over Time
for the Gas Resistance Based Electronic Nose 63
Figure 4.7. Gas Resistance Change Levels of Different Materials
for the Gas Resistance Based Electronic Nose 63
Figure 4.8. Confusion Matrix by the NB Model Classification Model
for Experiments with the Gas Resistance Based Electronic Nose 65
Figure 4.9. The kNN Model with Different Number of Neighbors Accuracy on
Training and Test Set for the Gas Resistance Based Electronic Nose 66
Figure 4.10. Confusion Matrix by the kNN Model Classification Model for
Experiments with the Gas Resistance Based Electronic Nose 67
Figure 4.11. Feature Importance Scores by the RFC Model
for the Gas Resistance Based Electronic Nose. 68
Figure 4.12. Confusion Matrix by the RF n_tree=10 Classification Model
for Experiments with the Gas Resistance Based Electronic Nose 70
Figure 4.13. CO2 Change Over Time
for the Multi Sensor Array Based Electronic Nose 72
Figure 4.14. CO2 Change Levels of Different Materials
for the Multi Sensor Array Based Electronic Nose 72
Figure 4.15. HCHO Change Over Time for the Multi Sensor Array Based Electronic
Nose 73
Figure 4.16. HCHO Change Levels of Different Materials for the Multi Sensor Array
Based Electronic Nose 74
Figure 4.17. PM2.5 Change Over Time for the Multi Sensor Array Based Electronic
Nose 75
Figure 4.18. PM2.5 Change Levels of Different Materials for the Multi Sensor Array
Based Electronic Nose 75

Figure 4.19. TVOC Change Over Time for the Multi Sensor Array Based Electronic Nose 76
Figure 4.20. TVOC Change Levels of Different Materials for the Multi Sensor Array Based Electronic Nose 76
Figure 4.21. All Binary Combinations of Different Parameters on Material Classification 78
Figure 4.22. Confusion Matrix by the NB Classification Model for Experiments with the Multi Sensor Array Based Electronic Nose 80
Figure 4.23. The kNN Model with Different Number of Neighbors Accuracy on Training and Test Set for the Multi Sensor Array Based Electronic Nose 81
Figure 4.24. Confusion Matrix by the kNN (k=8) Classification Model for Experiments with the Multi Sensor Array Based Electronic Nose 83
Figure 4.25. Feature Importance Scores for Experiment I RFC Model of Multi Sensor Array Based Electronic Nose 84
Figure 4.26. Classification Report for Experiment I n_tree=10, RFC model of Multi Sensor Array Based Electronic Nose 86

CHAPTER 1

INTRODUCTION

1.1 Motivation

Indoor air pollution refers to the undesirable physical, chemical, and biological characteristics of air in the building. Indoor air might be polluted by numerous kinds of contaminants which might come from cleaning products, construction activities, tobacco smoke, perfumes, water-damaged building elements and outdoor pollutants. It is one of the biggest environmental risks that affects performance, health and comfort. Indoor air pollution might be 100 times higher than outdoor pollution in some cases and most people spend 90% of their time in buildings. Exposure to poor indoor air can cause different health issues such as skin irritation, nausea, headaches, sick building syndrome and even cancer in some cases. According to The World Health Organization (WHO) Reports, about 700.000 people died just because of breathing poor air inside buildings (Air Pollution, 2020).

Indoor air quality gains much more importance in post-COVID world. After the SARS-CoV-2 has rapidly spread in the whole world, building occupants have become much more sensitive to the air they inhale. Transmission of the virus through the air is directly related to indoor air quality inside buildings. Changes to building operations, especially HVAC operations, can reduce airborne transmission. Ventilation and filtration are the key operations that can potentially reduce the risk of transmission. However, these two operations may not be sufficient on their own to provide good indoor air quality inside buildings.

Many different factors such as building materials, bad filtration, chemicals spreading from office furniture, smoking, food and beverages can cause bad air quality. Detecting the source of pollutants is crucial to optimize indoor air quality and retrofit the building in some cases. Apart from ventilation and filtration, facility management should be aware of the sources of indoor air pollution to make buildings safer in terms of virus spread.

Researchers should focus on quality problems of indoor air to adopt legislation, inspection and creating mechanisms that act in real time to improve public health. There must be mechanisms for data acquisition from buildings and analysis tools that correctly analyze the health of the environment in the spaces of public use. Mechanisms for data acquisition of indoor air and analysis tools that accurately process the data and quality of air in public spheres are needed.

As for the present technology, some sensing devices have succeeded in providing reliable, quick and continuous information about surrounding air. Furthermore, new technologies like cloud computing provide an excessive amount of computation power to analyze the data collected from these sensors and allow building intelligent algorithms based on indoor air quality. After the artificial intelligence evolution in the world, many different data analysis and processing tools have been released to create smart data-driven applications. All of these technologies and improvements could be helpful for creating a better environment inside buildings in post-COVID world.

Air pollutants can cause an increase on the levels of different types of sensing parameters such as CO, CO_2, NO_2, Ozone, *etc*. Every pollutant inside the building has different parameter patterns indicators. For example, Ozone can be increased by electric arcing, electronic air cleaners, some copiers, and printers. Furthermore, increase in VOC levels can be caused by air fresheners, furniture, office equipment and cleaning agents. Combustion equipment, engines, faulty heating systems are some other devices that can cause a rise of CO. A concentration of living creatures (i.e., animals, plants) can increase the level of CO_2.

The sensing instrument measuring many different parameters can help to analyze different pollutant sources. Electronic nose is a concept about detecting odors or flavors. Data collected from the sensing instrument may be used together with the electronic nose concept. An electric nose commonly identifies odors by detecting the "fingerprint" of a substance compound across several sensors monitored by pattern recognition applications. Fingerprint represents measurements of the sensor array during the response time to a particular pollutant source. Every different pollutant source creates a different fingerprint on the sensor array while polluting indoor air.

In this study, the fingerprints of different indoor pollutant sources will be analyzed. A scalable and high-performance algorithm will be able to detect the source of pollution by analyzing fingerprints of building materials and equipment.

1.2 Aims and Objectives

There are many studies on pattern recognition of gas sensors. However, most of the studies focused on outdoor odor sensing and there is not a sufficient amount of study on understanding the complex nature of indoor air quality.

By the end of this study it is expected to:

● Understanding the relationship between sensing gas parameters and pollution sources

● Investigate the indoor air quality characteristics of different building materials

● Investigate the effects of systems operating inside buildings on indoor air quality

● Analysis of big data to recognize the pattern of pollution source

CHAPTER 2

LITERATURE REVIEW

2.1 Importance of Indoor Air Quality

2.1.1 Well-being and Health

The impact of indoor air quality on human health was known throughout history. Greeks and Romans realized that sick people's numbers lowered when they increased ventilation in mines. Medieval age was very stagnant in the scope of indoor air quality. Boyle (1627–1691) and Hooke (1635–1703) found that supply air was essential for lungs. The role of oxygen in breathing was shown by Lavoisier in 1781. He also pointed out the relation between oxygen consumption and concentration of CO_2.
Max Pettenkofer found that some undesirable sensations due to stale air were not related to temperature, humidity or CO_2. He stated that they are related to the presence of organic materials exhaled by humans, and while poor indoor air does not have to make people sick immediately, it weakens the immune system in time.

Minimum required ventilation rate was firstly defined by Thomas Tredgoldin in 1836. He calculated that minimum fresh air should be 2 l/s per person due to breathing and candle burning. ASHRAE stated their first ventilation recommendation in 1895 as 15 l/s per person. During that era, there were two dominant ideas relevant to ventilation. Architects and engineers were concerned with providing thermal comfort and physicians were concerned with preventing the spread of disease.
There have been difficulties in assessing the exposure of specific pollutants to humans due to lack of monitoring devices until the last few decades. With the spread of measurement devices in the last few decades, international scientific communities,

health institutions and governments have started to focus on indoor air quality for improving comfort and health of building occupants. For that purpose, IAQ related standards, regulations and policies have been developed during the last decades. A typical human's 90% of time is spent in buildings, therefore IAQ has a massive impact on health and life quality.

In 2020, World Health Organization (WHO) reported that about 3.8 million people dies because of indoor air pollution (WHO, 2020). Indoor air pollution is a combination of particulate and gas components. The mixture of indoor air is highly dependent on sources, ventilation and emission rates. (Hamanaka and Mutlu, 2018)

Over the last few decades, it is realized that many different symptoms and illnesses are related to indoor air quality in buildings. Exposure of different types of pollutants is very common, although their effects were not inspected during long years. WHO categorizes indoor air related illnesses in two categories: Sick Building Syndrome (SBS) and Building Related Illness (WHO, 2020).

WHO defined the concept of Sick Building Syndrome as a collection of nonspecific symptoms including eye, nose and throat irritation, mental fatigue, headaches, nausea, dizziness and skin irritations, which seem to be linked with occupancy of certain workplaces in the 1980s (WHO, 1983).

Greer defined SBS as a group of non-specific symptoms with a temporal connection to a particular building, but with no specific or obvious cause (Greer, 2007) WHO stated that 20% of total building occupants showed symptoms of SBS (WHO, 2010) Air-tightness, high temperature and inadequate ventilation increase the effect of SBS in buildings.

One of the biggest challenges in SBS is the non-specific presence of it. This leads to high variability of symptoms in very diverse parts of the human body. SBS can show general (hoarseness of voice, allergies, flu-like symptoms, respiratory diseases, nausea, dizziness, headache, fatigue, and inability to concentrate), mucous (eye, throat and nose irritations or coughing), dermal (itching skin, face, hands or scalp) symptoms (Amin *et al.*, 2015).

Building Related Illnesses can be identified if there are symptoms that are clearly related to exposure to poor air in buildings. BRI symptoms can be associated with flu, including fever, chills, chest tightness, muscle aches, and cough (Tran et al., 2020).

2.1.2 Performance

Poor indoor air quality can affect academic and work performance. It has health and cognitive effects. While it causes school and work absenteeism, it also causes difficulty in concentrating, impaired memory, and slowed down mental processing.

Bako-Biro *et al.* (2012) researched the association between ventilation and student performance in primary schools of the United Kingdom. For three weeks, they measured CO_2 and other components of air in two chosen classrooms at each school. Ventilation was provided through the windows, which were also used for arranging the ventilation rate. They measured the student performance with computer aided tests and paper based tasks. They observed that due to the intervention, fresh air supply has increased from 0.3-05 to 13-16 l/s per person which led to an increase in student work rate by ~7%.

Haverinen-Shaughnessy *et al.* (2011) focused on larger size of experimental data for assessing relation between classroom ventilation rates and academic achievement. They observed 100 US primary schools in two different school districts in Southwest USA. 87 classrooms had unsatisfactory ventilation rates according to ASHRAE Standard 62. They found a linear relationship between ventilation and performance within the range of 0.9-7.1 l/s per person. The performance of students increased by 2.9 % (95 %CI 0.9-4.8 %) for math and 2.7 % (0.5-4.9 %) for reading when ventilation rate was increased by 1 l/s per person

Fang *et al.* (2004) observed the office environment for indicating the effect of the indoor environment on work performance. They created environments with different temperatures (20 °C, 23 °C, 26 °C), different humidity levels (40, 50, 60 %) and different ventilation rates (3.5 l/s, 10 l/s). Occupants were exposed to each condition for 280 minutes. They performed in simulated office work during each exposure and repeatedly marked a set of visual-analog scales to reflect their perception of environmental conditions and extent of SBS symptoms at the time. Researchers observed that temperature and humidity changed the perception of ventilation rate.

The effect of lowering the ventilation rate from 10 to 3.5 l/s per person on perceived air quality might be mitigated by lowering the temperature and humidity levels from 23°C /50 %RH to 20°C / 40 % RH.. Ventilation isn't the only factor that influences a person's performance when it comes to indoor air quality. The impact of PM2.5 and PM10 levels on academic performance was studied by Alves et al. (2013). They measured the PM level on a regular basis and discovered that the school couldn't reach the acceptable PM level roughly 70 % of the time. The findings clearly show that these schools have a significant degree of particulate matter exposure. Continual PM10 measurements imply that students' physical activity, which is thought to be higher in younger children, leads to a continuous process of sediment resuspension.

2.1.3 Energy

Due to the global climate crisis, energy efficiency has become very important throughout the world. As buildings consume 40 % of the world's energy, issues related to energy efficiency in buildings play an important role in responding to the climate crisis (IEA, 2020).

Ventilation components have had a huge impact on the total energy consumption of the building. With the new ventilation standards, the impact rate of ventilation on the total energy consumption of buildings is expected to increase. However, new building regulations mandate the construction of more airtight buildings, which will definitely affect indoor air quality (Pagliano, 2010).

Although it is generally believed that energy efficiency has a negative impact on air quality, it can be improved through some policies and decisions. Consistently, success can be achieved by measuring indoor air quality and ensuring that the building's ventilation rate is never lower than the minimum required ventilation rate, which is a key parameter to provide energy efficiency and indoor air quality at the same time (Bipat, 2020).

The complaints and consciousness about indoor air quality is increasing as time goes by although ventilation and air distribution systems are being more advanced. Four types of ventilation systems are widely used in buildings. Mixing ventilation, displacement ventilation, personalized ventilation, and hybrid air distribution are examples of these methods.

Mixing ventilation (MV) is the method that mixes the contaminated room air with provided fresh supply air. Displacement ventilation (DV) is the method that displaces all contaminated room air with fresh supply air by supplying air at low velocity to create an upward air movement as it warms up by heat sources of the room. Personalized ventilation does not aim to ventilate the whole space, it is used to supply fresh air directly to occupants, *i.e.* office desks or hospital beds.

DV and MV methods can provide efficient air supply, although they have two main problems. Firstly, they cannot be used in heating mode, secondly, the penetration of the air into the depth of room is poor. Hybrid air supply systems combine the characteristics of these methods.

These systems can be classified with two different types: impinging jet (IJ) and the confluent jet (CJ) as can be seen on Figure 2.1. IJ has a vertical duct that supplies air to spread air over the floor. CJ is used to create circular apertures with the same flow (Awbi, 2017).

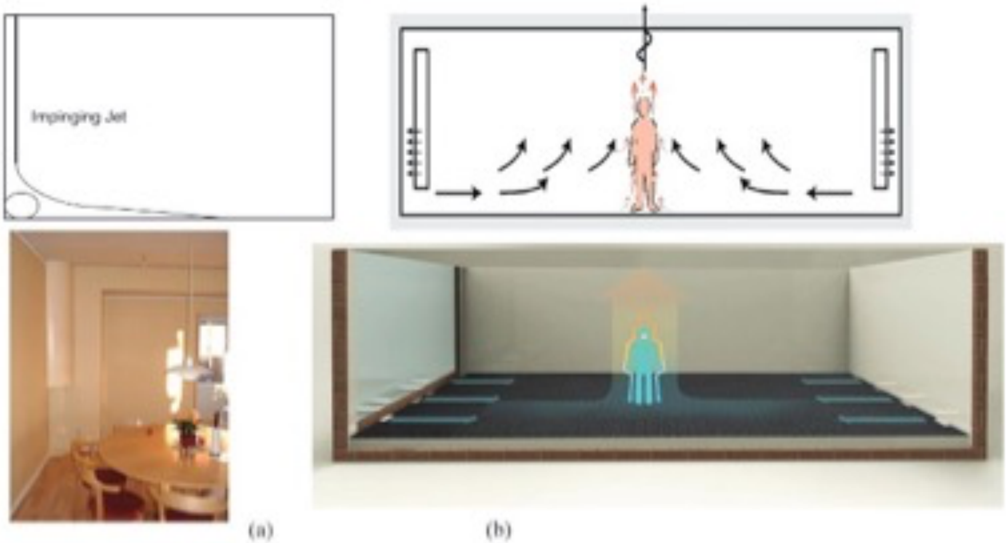

Figure 2.1 Hybrid air distribution systems; (a) impinging jet; (b) confluent jets

2.1.4 Thermal Comfort

It is a very critical fact that there is a very strong correlation between thermal comfort and indoor air quality. The PMV method, invented by Fanger, is the most often used method for assessing people's thermal sensibility. PMV indicates thermal feeling by the use of six distinct parameters: air temperature, air humidity, air velocity, mean radiant temperature, clothing insulation, and human activity. The PMV index value varies between 3 and +3, corresponding to human sensations ranging from chilly to hot (Fanger, 1972).

Huzienga *et al*. (2006) made a large dataset by getting thermal comfort feedback of building occupants. The results indicate that in only 11% of the buildings studied at least 80% of residents rated their thermal comfort as satisfactory. Air quality scores were slightly higher, with 26 % of buildings reporting occupant satisfaction of at least 80 %.

2.1.5 Building Envelope and Interior Design

Healthy building concept requires many different points of views while the building is being designed. Acoustic design, mechanical stability and basic safety for occupants are crucial parts of building design. However, those are not enough to ensure indoor environmental quality for building occupants. There are dozens of different factors that affect the well being of occupants. Sustainable design is progress for achieving ecological balance between construction and nature. This progress can combine natural and passive solutions with intelligent systems that give flexibility for controlling environmental quality with resource consciousness (Loftness et al, 2007).

Choosing the correct specifications for coatings, adhesives, surface finishing, and furnishings is important for designing buildings with sustainability principles especially to provide better indoor air quality. Architects must determine the mixture of those compounds, consequently their emission rates.

The process of modeling the VOC emissions of used building materials and furnishings is critical and time consuming. As material emissions account for a significant portion of indoor pollution, a VOC emission model is required prior to entering the building phase. Those models are mostly constructed by using diffusion principles. Diffusion principles are generated by Fick’s law.

According to Fick's diffusion law, diffusive flux moves according to the concentration gradient from a high-concentration area to a low-concentration area. Flores (2014) and Yang *et al.* (1998) developed a widely used numerical model that utilizes internal diffusion principles to predict VOC emissions. The model requires four parameters: initial VOC concentrations in the material, a solid-phase diffusion coefficient, a material-air partition coefficient, and the material's age.

2.2 Indoor Air Quality Parameters

2.2.1 CO2

CO_2, which is present effortlessly in the environment, is a particle created by living bodies through expiration (deep breathing). The concentration of CO_2 in indoor airflow in buildings will depend on human air and occupancy exchange rates (Occupational Health & Safety, 2016).

Levels of CO_2 concentration and their effects on humans:
250-350 parts per million (ppm): ambient (normal) outdoor air level
350-1,000 ppm: average amount found in heavily inhabited environments with frequent air exchange
1,000-2,000 ppm: threshold associated with tiredness and poor air quality
2,000–5,000 ppm: level associated with headaches, tiredness, and stagnant, stale, stuffy air; poor focus, lack of attention, moderate nausea, and elevated heart rate may also be present.
>5,000 ppm: This suggests unusual air conditions in which excessive levels of several other gases may also be present. Toxicity or even oxygen starvation is a possibility. This is the permissible exposure level for common workplace hazards.
>40,000 ppm: This level is immediately harmful due to oxygen deprivation.

Carbon Dioxide can have dangerous effects on the body, leading to health issues including inflammation, reduced cognitive performance and also kidney bone issues (Jacobson et al. 2019).

2.2.2 VOC

Volatile Organic Compounds are organic substances with a low boiling point. Their volatility is proportional to the quantity of molecules in the air. VOC concentrations can be 2 to 5 times higher than outdoor levels. However, they can be 1000 times higher than outdoor levels during some events. (Fleder, 2020)

Organic emissions can vary between buildings and building materials. Although the VOC amount can vary from 0.2 to 1.7 mg/m^3 in old buildings, it can vary from 0.5 to 19 mg/m^3 in new buildings. From insulations and wood panels to varnishes and sealants, numerous building materials emit VOCs. Dry materials like acoustic ceiling tiles, solid wood, oriented strand board(OSB) and wet materials like floor wax, caulking sealant, wood stain emit a significant amount of VOC in buildings. The adverse health effects because of VOC emission can happen above 3 mg/m^3. (Lafond, 2020).

Breathing VOCs may irritate the eyes, throat and nose. Also, they can cause nausea and difficulty in breathing, and may damage the main nervous system along with other organs. Even a number of VOCs are able to cause cancer (Lafond, 2020).

2.2.3 PM2.5

PM (Particulate Matter) is a term that refers to a mixture of solid particles and liquid droplets found in the air. The pathogenicity of PM is driven by the size, origin, composition, solubility and the ability of theirs to create reactive oxygen. PM2.5 represents particles and droplets with diameters that are generally 2.5 micrometers and smaller (Samoli et al. 2005).

PM2.5 particle concentration at indoors is directly related to the concentration of outdoor particles. As a result, factors such as ventilation, whether through infiltration, mechanical ventilation, or natural ventilation, as well as the magnitude of the air exchange rate (AER), all contribute to the building's PM2.5 concentration. Other characteristics and the airtightness of the structure, in addition to the layout and operation of the HVAC system, may increase particle amount indoors (Allen *et al.* 2012).

On average, 47 % of indoor PM2.5 concentrations are formed by soil particles suspended in the air with 13 % and 34 % by a mixture of organic (skin flakes, clothing fibers, potential condensation of VOCs) and calcium-rich particles (from chalk and building deterioration). 53 % of indoor PM2.5 concentrations are caused by outdoor emission sources. The major outdoor sources are motor vehicles (26 %), followed by biomass burning (17 %), soil dust (7 %), road dust (3 %), and marine aerosols (1 %) (Amato et al, 2014)

PM2.5, which has very small diameters, mostly has big surfaces despite their size. This may cause transmission of toxic stuff, passing through the filtration of nose hair and reaching till the end of the respiratory system with airflow. The diffusion of these particles into other parts of the body through air exchange can damage the whole body (Brunekreef and Holgate, 2002).

2.2.4 PM10

PM10 particles are produced by a variety of causes, including dust from untreated roads, smoke from fires, sea salt, automobile and truck exhausts, and industry. If PM10 levels are too high, people with lung or heart problems may have additional symptoms. Symptoms are able to include wheezing, chest tightness, or maybe a problem in breathing ("PM10 Particles in the Air").

PM10 and PM2.5 particles are often formed by distinct emission sources and chemical processes. Combustion of gasoline, oil, or diesel fuel is a significant producer of PM2.5 and a significant source of PM10. Additionally, PM10 contains dust from construction sites, landfills, and farms, wildfires and brush/waste burning, industrial sources, wind-blown dust from open spaces, pollen, and bacterium particles.

In recent years, researches show that PM2.5 has a much more adverse effect on our health than PM10. That's why PM10's effect on Indoor Pollution Calculation methods has become less important in recent years. Guttikunda (2020) stated that new experiments are evidencing that smaller particles, that it will go deeper into the lungs and harm the human body. This is also one of the important reasons for WHO to push for all countries to have standards for PM2.5 instead of PM10 (Guttikunda, 2019).

2.2.5 Formaldehyde

Formaldehyde is a crucial substance used commonly to manufacture numerous household and building materials products. It's a product of combustion and some other natural processes. Formaldehyde is able to cause irritation of the epidermis, throat, nose and eyes. A human can contract cancer with high levels of exposure to formaldehyde (Environmental Protection Agency, 2016).

2.2.6 Ozone

The ozone layer in the lower atmosphere forms when sunlight hits some air pollutants. Sunlight catalyzes these pollutants, and they are converted into ozone. Ozone exposure could easily worsen currently existing lung conditions like emphysema, asthma, along with persistent bronchitis (Air Filters for Clean Air, 2018).

Home air cleaners are supposed to work by changing the charge of molecules. Therefore, the molecules are attracted to one another and combine harmlessly. However, the method used in these cleaners utilized to ionize air molecules similarly changes several of the oxygen in the environment into ozone. Studies show that these products are ineffective in removing interior air contaminants (Salonen et al. 2018).

2.2.7 NO_2

Nitrogen dioxide is a gaseous substance. At room temperature, the primary route of exposure is through inhalation. Indoor sources of heat include tobacco smoke and gas, wood, oil, kerosene, and also coal-burning appliances. The primary impact of inhaling raised amounts of nitrogen dioxide will be the higher likelihood of breathing issues. Nitrogen dioxide is able to decrease immunity to lung infection. This can result in problems like coughing, bronchitis, wheezing, colds and flu (Adamkiewicz, 2010).

2.2.8 Temperature and Humidity

Temperature and humidity affect us in two ways. Firstly, they may cause an uncomfortable environment. Secondly, they may accelerate bacteria and mold growth. When cooling operations are going on, dry air diffuses into the room and is warmed. The relative humidity drops 20 % immediately. This kind of drop in humidity makes airborne viral particles to travel rather easily. Research suggests that relative humidity of 40-60 % is perfect for preventing the transmission of viruses (Medical News Today, 2020).

2.3 Indoor Air Pollutant Sources

Over the last half-century, considerable developments have occurred in both construction materials and consumer products used within. Composite wood, foam padding, polymeric flooring, synthetic carpets, scented cleaning solutions, and agents for plastic objects have become ubiquitous. Precisely the same holds true for electrical and mechanical appliances like washer/dryers, computers and TVs. These products and materials produce several chemical substances like solvents, unreacted monomers (Weschler, 2009).

In the 1970s, indoor pollutants that originated outdoors received high interest, particularly sulfur dioxide, nitrogen oxides, airborne particles and ozone, *etc.* with the increasing number of researches that show outdoor pollution impacts interior spaces (Andersen, 1967). Following that, attention was focused on contaminants that were easily detectable and quantifiable, such as formaldehyde, radon, asbestos, and tobacco smoke, as well as nonpolar volatile organic chemicals (Lioy *et al.* 1985). Over time, pollutants along with other semi-volatile organic compounds have been measured inside. As far as better analytical tools have been created and instrument sensitivities improved, knowledge on indoor air pollutants improved massively.

2.3.1 Building Materials

Numerous building components given below release chemical compounds into the indoor air.

Composite-wood: Following World War II, plywood started replacing solid wood on residential building construction. When plywood was originally introduced to the market, its adhesive resin was predominantly urea-formaldehyde, with rather high formaldehyde emission rates (Hodgson et al. 2002).

PVC pipes: They have partly changed copper piping in several indoor plumbing programs, which includes empty, waste and vent programs, in addition to water distribution systems. PVC piping is rigid and does not release plasticizers in the same way as flexible PVC products do (*e.g.,* vinyl flooring and wall covering). Nonetheless, organotin compounds are frequently used as stabilizers in PVC pipes, and these semi-volatile components are likely to diffuse into indoor areas over time (Berens, 1985).

2.3.2 Office Equipment

Numerous office equipment given below release chemical compounds into the indoor air.

Carpeting: In a health threat assessment of carpeted floors, two elements are of interest. One of them is that carpets may act as a repository for indoor air pollutants including soil, dust particles, allergens and additional natural contaminants that can build up in the carpets. The second one is that carpets might produce volatile organic compounds (VOCs) which could result in irritation and smell of mucous membranes, particularly in individuals that are sensitive. Nevertheless, assessments of new carpets indicate that emissions happen to be lowered and have a shorter length. Pollution over an extended period depends partially on the deposition of soil, care and cleaning agents as well as on cleaning procedures (Becher *et al.* 2018).

Paints: In the 1950s, water-based (latex) paints began to supplant solvent-based (oil) paints in indoor areas. Paints based on water produce fewer volatile organic pollutants than paints based on solvents. Their expansion accelerated significantly during the ecologically conscious 1970s, resulting in reduced VOC emissions during interior painting. The type of binder used in water-based paint determines the residual monomers and breakdown products that the paint emits. Initially, styrene-butadiene latex was the predominant binder in water-based color; this was later superseded by acrylic latex and vinyl acrylic (Weschler, 2009). 'Green' or' natural' paints released in last few years, dependent on organic ingredients, have a large market share. These substances are often unsaturated organic molecules (*e.g.*, limonene, linseed oil, and a variety of other terpenoids) that are reactive with ozone. The oxidation products produced include secondary organic aerosols.

Furnishing: The great bulk of furnishings (including composite and upholstered wood furniture) have potentially harmful components which could off gas chemical compounds and release particulates into the fresh air we inhale (The Best of What's Healthy, Local and Green in Maine, 2013).

Electronic Devices: Photocopiers generate ozone, styrene, formaldehyde, and a variety of other aldehydes, as well as semivolatile natural compounds (VOCs) from heat transfer fluids. The shells of personal computers and printers, as well as the circuit boards, release flame retardants and plasticizers, particularly brominated flame retardants. Certain laser printers have been identified as sources of ozone and particulates in the air. Computer monitors have evolved from cathode ray tubes to flat panel screens; the latter frequently generate fewer organic pollutants than the former (Lee *et al.* 2001).

2.3.3 Chemical Reactions

Chemical reactions inside buildings are proven to take place between "oxidizing" compounds including ozone and many common indoor compounds like terpenes that are readily oxidized, airborne, or on surfaces like carpets, typically through the development and response of hydroxyl radicals. By-products of these inside air chemical reactions (*e.g.*, hydroperoxides, acrolein, formaldehyde, ultrafine particles and fine) are hypothesized to trigger sensory irritation of airways and eyes or maybe skin irritation ("VOCs from Indoor Chemical Reactions and Health").

2.3.4 Air Conditioning

The biological contaminants, which include bacteria, fungi, pathogenic bacteria, mites and viruses, can build up in the air-cooling system and may boost in fan coil unit and heat sinks where it is ideal for the growth of theirs. Air conditioning filtration screens and especially heat sinks are susceptible to the collection of dust. When the cooling is operating, the humid environment and also ideal temperature created by condensation, which comes with dust accrued in the heat sink, provides the ideal living condition for bacterias, other microorganisms and viruses to reproduce quickly ("Be aware of air conditioning pollution").

2.3.5 Smoking

Tobacco smoking is a major contributor to indoor air pollution. According to research on toxicological substances of tobacco smoking, PM-fractions (PM10, PM2.5 and PM1) are increased dramatically after smoking in the room (Hamanaka and Mutlu, 2018).

2.4 Indoor Air Quality Sensing Instruments

Indoor air quality monitoring devices can assist in identifying and improving indoor air quality. Local and distributed monitoring of chemical substance concentrations is critical for safety and security purposes, as well as for efficiently controlling heating, ventilation, and air conditioning (HVAC) systems (Vito *et al.* 2011).

Recently, numerous new technologies for monitoring environmental parameters have been introduced, all with the goal of improving the efficiency of indoor air quality. The availability of inexpensive, low-power embedded sensors, radios, and processors makes it much easier to construct a sensor bundle that detects a variety of parameters while engaging with the physical environment of applications such as air quality control via wireless communication protocols (Al-Haija *et al.* 2013).

Low-cost analog air quality sensors are received increasing attention in recent years. However, analog air quality sensors require expensive and continuous re-calibration. Over the past three years, low-cost & low-power digital air quality sensors have been developed. These low-cost & low-power air quality sensors are user friendly, they require very less maintenance effort and they can continuously monitor during 10+ years (Chojer *et al*, 2020).

2.4.1 Metal Oxide Gas Sensors (MOX)

In principle, detecting the pollutants with MOX sensors is quite simple: The sensor consists of two parts. The sensing material (Metal Oxide Plate) located at the top, and the heater located at the bottom part of the sensor as can be seen on Figure 2.3. The heater is heated by the sensor until temperatures between 150-400°C. The electricity flow starts with the increasing temperatures. According to sensitivity of the sensing material to the particular pollutant, the resistivity of the sensing material determines the level of the pollutant. This kind of sensor is the most popular type at indoor air quality sensing (Dey, 2018).

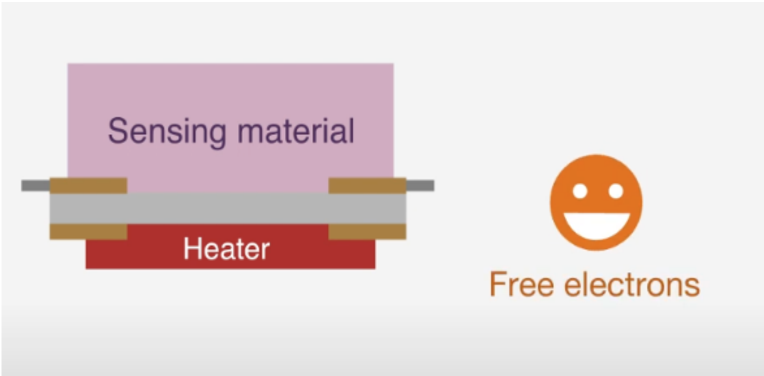

Figure 2.2. Metal Oxide Gas Sensor Structure (Figaro, 2020)

2.4.2 Optical Gas Sensors

Optical sensors measure the particles in flight. They are usually a low-cost way of getting real-time data about particles in the air we breathe. The working principle of these sensors is: A sensor and a beam of light are located at a particular angle to each other. When the particle passes in front of the light, light is reflected back to the sensor. The sensor generates a signal when it senses a reflected light. If the airflow is constant around the sensor, the length of this signal can be used to detect the size of the particle (Badura *et al.* 2018).

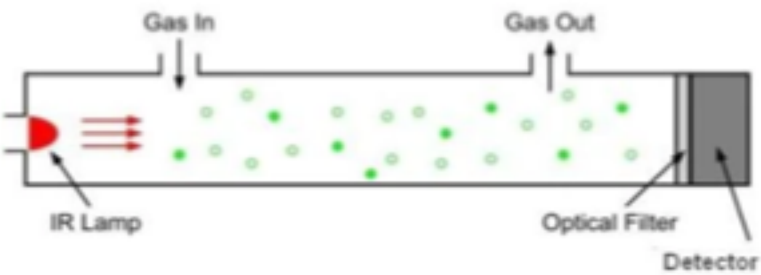

Figure 2.3. Optical Gas Sensor Structure (Figaro, 2020)

2.4.3 Electrochemical Gas Sensors

Electrochemical sensors work by responding with the analyte and creating an electric signal. Nearly all electrochemical gas detectors are amperometric sensors, producing the current that's linearly proportional to the pollutant concentration. The working principle of amperometric sensors is the measurement of the current-voltage connection in an electrochemical cell wherein equilibrium isn't developed. The current is quantitatively associated with the speed of the electrolytic process in the sensing electrode whose voltage generally is kept constant by another electrode (called guide electrode) ("Electrochemical Gas Sensors").

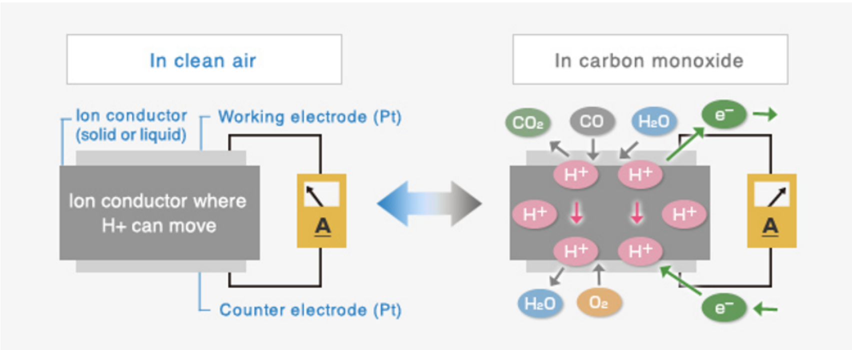

Figure 2.4. Electrochemical Gas Sensor Structure (Figaro, 2020)

2.5 Indoor Air Quality Standards

Table 2.1 Different Indoor Air Quality Standards for Different Parameters

Parameter	**WHO**	**OSHA**	**EC**	**ASHRAE**	**EPA**	**WELL**
Location (Country)	Worldwide	US	EU	USA	USA	USA
O_3	100 µg/m^3 **8-hour mean**	0.1 ppm **8-hour mean**	120 µg/m^3 **8-hour mean**	N/A	0.070 ppm **8-hour mean**	51 ppb **8-hour mean**

CO	7 mg/m^3 **24-hour mean** 10 mg/m^3 **8-hour mean** 100 mg/m^3 **1-hour mean** 100 mg/m^3 **15-min mean**	50 ppm **8-hour mean**	10 mg/m3 **8-hour mean**	9 ppm **8-hour mean**	9 ppm **8-hour mean** 35 ppm **1-hour mean**	N/A
CO_2	N/A	5000 ppm **8-hour mean**	N/A	1000 ppm **8-hour mean**	N/A	N/A
SO_2	20 µg/m^3 **24-hour mean** 500 µg/m^3 **10-minute mean**	5 ppm **8-hour mean**	350 µg/m^3 **1-hour mean** 125 µg/m^3 **24-hour mean**	N/A	75ppb **1-hour mean**	N/A
NO_2	40 µg/m^3 **Annual mean** 200 µg/m^3 **1-hour mean**	5 ppm	40 µg/m^3 **Annual mean** 200 µg/m^3 **1-hour mean**	N/A	100 ppb **1-hour mean** 53 ppb **Annual mean**	100 ppb
PM2.5	10 µg/m^3 **Annual mean** 25 µg/m^3 **24-hour mean**	5 mg/m^3 **8-hour mean**	25 µg/m^3 **Annual mean**	N/A	35 µg/m^3 **24-hour mean** 12 µg/m^3 **Annual mean**	15 µg/m^3 **Annual mean**

Table 2.1 (continued)

Parameter	**WHO**	**OSHA**	**EC**	**ASHRAE**	**EPA**	**WELL**
PM10	20 µg/m^3 **Annual mean** 50 µg/m^3 **24-hour mean**	15 mg/m^3 **8-hour mean**	40 µg/m^3 **Annual mean** 50 µg/m^3 **24-hour mean**	N/A	150 µg/m^3 **24-hour mean**	40 µg/m^3 **Annual mean** 50 µg/m^3 **24-hour mean**

Radon	4 pCi/L	N/A	N/A	N/A	4 pCi/L	0.15 Bq/L [4 pCi/L]
t-VOCs	500 µg/m^3	5 ppm **8-hour mean**	N/A	N/A	N/A	500 mg/m^3
Reference	(WHO, 2020)	(Osha, 2020)	(EC, 2020)	(ASHRAE, 2020	(EPA, 2020)	(WELL, 2020)

2.6 Electronic Nose (EN)

Electronic nose classifies odors by analyzing fingerprints of chemical compounds by using a sensor array with a pattern recognition algorithm. There are applications of electronic nose in many different industries including agricultural, environmental, food, manufacturing, and the military. However, there are still some challenges for developing an electronic nose to be used in a wide variety of applications. Huge amount of data from different sensors that can convert interactions of molecules into reliable patterns should be collected.

Different pollutant sources result in diverse responses in sensors, and these responses provide different signal characteristics on various indoor pollutants in buildings. The algorithm evaluates the signal patterns and compares different pollutant characteristics to find the source of them. (Caballero *et al.* 2003).
Even though the program part could be believed as the "brain", the hardware part could be viewed as the receptors of an electronic nose system. The program section mostly has information processing algorithms that identify as well as classifying every scent detected utilizing digital signatures of the sensed chemicals. The hardware portion is essentially a sensor array. Since the key goal of the EN is classifying and detecting several scents, the sensing array must cover various kinds of specific sensors, wherein each sensor is liable for detecting a unique substance. Ideal sensor selection is one of the most important parts of the electronic nose concept. Sensors may vary according to the application it will be used.

The selection of ideal sensors to a certain process is a vital thing in this particular technology. It may be concluded here that choosing appropriate hardware along with cost-efficient application pieces is extremely important in developing a profitable EN system for a specific issue (Karakaya *et al.* 2019).

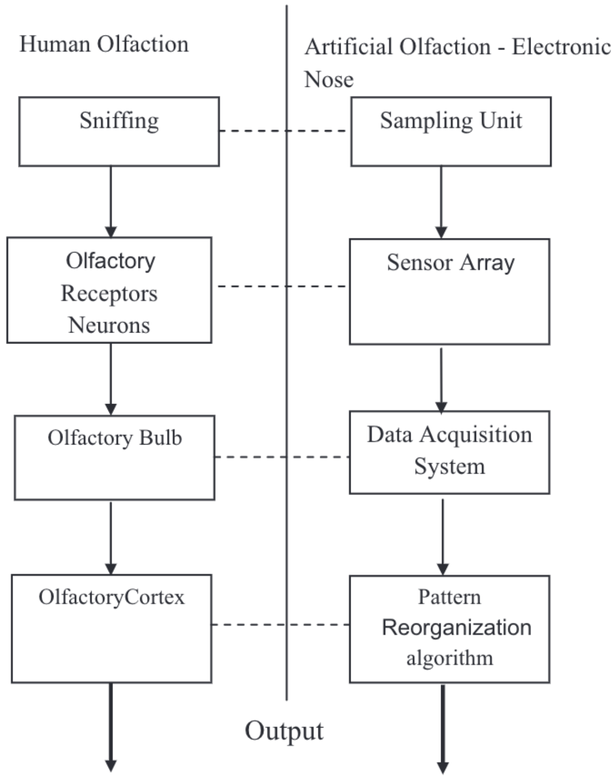

Figure 2.5. Analogy between Biological nose and Electronic Nose (Karakaya *et al.* 2019)

2.6.1 History of Electronic Nose

The term Electronic Nose was firstly used by Wilkens, Hatman, and Buck (1964). The chemical sensor array for odor pattern recognition was firstly developed by Persaud (1982). Gardner and Bartlett (1988) defined the word "e-nose" as "a device capable of detecting simple and complicated scents that consists of a partial specificity electronic chemical sensor array and an appropriate pattern recognition system." The working principle of the electronic nose is not based on receptors like traditional ways of chemical sensing, but it is based on the distribution pattern processing. It is mostly defined as an "intelligent chemical sensor array system that mimics the mammalian olfactory system" (Park *et al.*, 2019).

Hatfield(1994) developed an integrated circuit (IC) to reduce size and power consumption of electronic nose. For a highly integrated EN, the chemoresistive sensors are well suited for IC integration owing to their simple electrical characteristics and interface circuitry. Therefore, chemoresistive sensors are widely used in electronic noses. Recent studies shows that chemoresistive sensors will be able to give 2D outputs according to gas present in the air. This will increase the data amount to detect pollutant source correctly (Park *et al.*, 2019).

2.6.2 Components of Electronic Nose

An EN involves each hardware and software elements that is briefly depicted within a smell category program within Fig. 2.7. Initially, the launched background gas is absorbed through the sensor array. The detection of the input signal can be analyzed based on the modification in voltage, frequency, current, opposition variables based on the kinds of elements within the sensor array. Since unique forms of receptors are generally used around sensor arrays, the acquired signals must be preprocessed to recognize all those modifications then and properly prepared to digitize them to be able to develop a dataset. Hence, the sensed indicators are properly manipulated, or converted, filtered, amplified, *e.g.*, in the order to be readily utilized in additional stages.

The prepared signals are later examined in conditions of the specific properties of theirs in the data-gathering stage. Furthermore, adequate information is acquired from these signals as well as the obtained information is preprocessed based on the demands of the used pattern recognition algorithm. Finally, the air signal is classified together with the pattern recognition phase. (Karakaya *et al.* 2019).

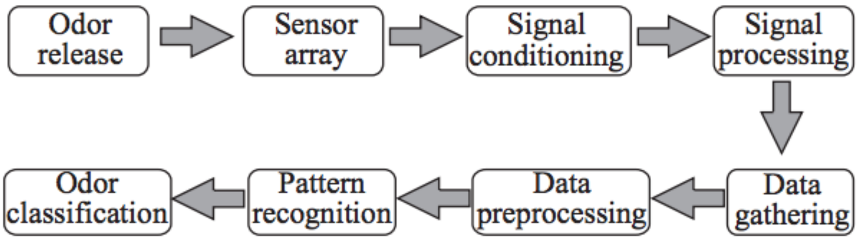

Figure 2.6. The Logic of Electronic Nose (Karakaya *et al.* 2019)

Sensor Array: A sensor array is a combination of sensors, each one of which changes power into an electrical response that could be represented by a sequence of bits. The other part of the sensor array is analyzing these electrical responses with digital signal processing algorithms (SAM Lab).

Since the key goal of an EN is sensing somewhat more than a single substance, this particular goal could be accomplished with increased accuracy just by merging many unique sensors in the array. Since every different application of EN requires a different type of combinations in the sensor array, the following conditions should be considered (Gebicki, 2016):

- Response time
- Sensitivity to different temperature and humidity ranges
- Sensitivity to different chemical compounds
- Stability of measurement
- Size
- Ease of Use
- Price

Machine Learning Algorithms: Machine learning is used for the procedure of recognizing patterns of the different pollutant sources. Pattern recognition could be described as the classification of information based on knowledge probably gained and on statistical info extracted from collected sensor data (GeeksforGeeks, n.d).

K- Nearest Neighbor: The k-NN algorithm is a non-parametric classification approach. The input data set contains the k nearest training data points in the feature space. The output is determined by neighboring votes. The object's estimated class is determined by the class that is most prevalent among its k nearest neighbors (Altman, 1992).

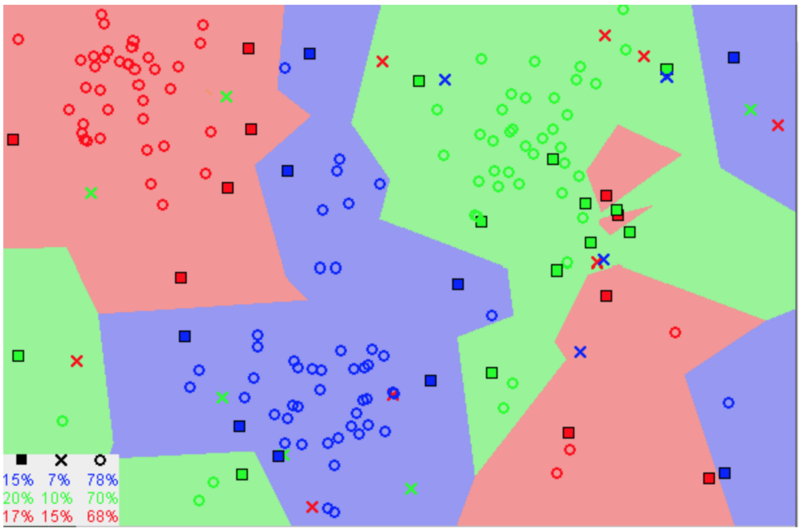

Figure 2.7. Classification with k-NN (Chugh *et al.*, 2020).

Support Vector Machines: Support vector machine (SVM) is an algorithm that is used for classification. SVM is a supervised learning algorithm, it looks at data and sorts the data into one of two categories. SVM gives a map of sorted data which is separated by margins between the two as far as possible as output.

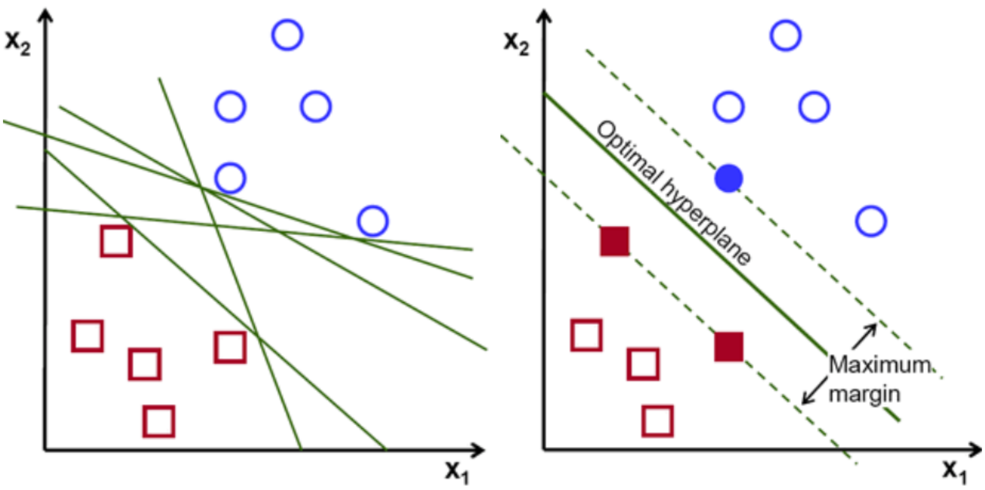

Figure 2.8. Classification with SVM (Yuan *et al.*, 2017).

Artificial Neural Networks: Neural networks are an algorithm that is used to recognize underlying relationships in the dataset with a method that mimics how the human brain works. Neural networks refer to connected neurons. These neurons can adapt their parameters to changing input, this provides generation of the best possible result without needing to design any output criteria.

2.6.3 Electronic Nose Applications for Pollutant Source Recognition

Air Quality Analysis Applications:

The electronic nose can be used in many different applications of environmental pollution monitoring. It is used in different mediums such as atmosphere, gasoline, water, and indoor. Negri and Reich (2001) modeled an electronic nose composed of sensors to address the gasses that pollute the atmosphere. They analyze the concentration and composition of carbon monoxide, ethanol, isobutane, and methane. They show that it is possible to understand environmental pollutants when they are detected in the presence of other sensible gases.

Kim *et al.* (2013) developed an integrated chemical sensor array to detect and identify the environmental pollutants in diesel and gasoline. There is a low noise floor analog front end to achieve better signal processing performance on the system. They present methods to detect, digitize, and classify pollutants at a certain set of analytes. The analog front end reads data from eight conductometric and amperometric electrochemical sensors. Required features are extracted from the data and then pattern classification methods have been applied to the detection of the pollutants in gasoline.

Jie *et al.* (2017) provided a very accurate electronic nose application. The sensor array is composed of the QS-01 from FIS, the TGS2600 and TGS2602 from FIGARO, and the temperature and humidity sensor SHT10. The pattern recognition technique is based on the Back Propagation (BP) neural network, a well-known machine learning technology. The comparison of the suggested performance e-nose's to that of previous e-nose solutions demonstrates the improvement.

Herrero *et al.* (2016) developed a portable and waterproof electronic nose with database connection to classify water pollutants in the cloud environment. The device is hand-held, battery-powered, and wireless. It uses four sensors to sample VOC levels. Artificial Neural Networks are used to classify pollutants in water. Results show that the proposed prototype can discriminate the samples measured (Blank water, acetone, toluene, ammonia, formaldehyde, hydrogen peroxide, ethanol, benzene, dichloromethane, acetic acid, xylene, and dimethylacetamide) with a 94% classification success rate.

Zampolli *et al.* (2004) developed a dedicated, mini, cost-effective electronic nose by using metal oxide sensors and signal processing techniques to analyze data of electronic nose. The nose developed to quantify carbon monoxide and nitrogen oxide in mixtures with relative humidity and VOC at indoor environments. The study stated that it is possible to discriminate concentrations as low as 20 ppb for NO_2 and 5 ppm for CO in the test gas environment thanks to the fuzzy logic system.
Saad *et al.* (2015) developed a scalable wireless indoor air quality monitoring system and deploy these sensors to different rooms. They prepared controlled experimental setups to classify the sources influencing IAQ in various environments like ambient air, chemical presence, fragrance presence, foods and beverages and human activity. They developed sensor module cloud (SMC) to collect data of these experimental setups. After they constructed a labeled dataset to identify the source of pollution, they proposed an algorithm based on Artificial Neural Networks to classify the patterns of different sources. On average, the system was about 99.1% correct. Overall, it can be concluded that the system delivered a high classification rate based on ANN.

Health Applications:

Human exhaled breath is a mixture of 3000+ different VOCs. This gives the opportunity to use electronic nose technology in medical applications. Based on this hypothesis, there have been many published studies to assess the role of electronic nose technology for diagnosing various respiratory and systemic diseases (Dragonieri *et al.*, 2017).

Tozlu *et al.*(2021) investigated if the electronic nose can be used to detect coronary artery diseases. The electronic nose is purposed for analyzing the air from exhaled respiratory. They collected data from 33 patients diagnosed with myocardial infarction that underwent a primary percutaneous coronary intervention, 22 patients with stable coronary artery disease and 26 patients without heart disease. They developed an electronic nose which includes the sensor array that has 19 different gas sensors on it. Statistical parameters like mean, derivative variance, kurtosis, skewness have been extracted. Collected data splitted as training data(64 %) and test data (36 %). Patients who have myocardial infarction were separated from healthy patients with 97 % accuracy. Patients who have stable coronary artery disease were separated from the healthy patients with 81 % accuracy (Tozlu *et al.*, 2021).

Only 15 % of lung cancer patients can be treatable. Kort *et al.* (2017) focused on electronic nose technology that measures and analyzes different volatile organic compounds in the breath of lung cancer patients. They have done breath analysis on 210 suspected patients where approximately half will have a confirmed diagnosis and the other half will have a rejected diagnosis of lung cancer. Furthermore, they analysed 150 healthy control subjects. They have done pre-processing, data compression and neural networks to create a model that predicts the result. They reached 85 % accuracy to detect lung cancer cases (Kort *et al.*, 2017).

Food Applications:

Jia *et al.* (2019) used an electronic nose to detect and recognize if the apple is fresh or moldy. Apples were divided into two seperate groups, group A containing mixed apples with/without different molds and group B containing only fresh apples. The electronic nose took different VOC gas samples from every different sample in those groups. Four pattern recognition methods, including backpropagation neural network (BPNN), radial basis function neural network (RBFNN), linear discriminant analysis (LDA), backpropagation neural network (BPNN), support vector machines (SVM) to analyze freshness of apples. Results show that gas sensors gave a strong signal response to the characteristic flavor of apples. The accuracy of the four pattern recognition methods showed that BPNN gave best results for the training and testing sets for both Groups A and B, with prediction accuracies of 96.3 % and 90.0 % (Group A), 77.7 % and 72.0 % (Group B), respectively (Jia *et al*., 2019)

Zheng *et al.* (2009) recognized successfully four different long grain rice samples with an electronic nose that consists of 32 different polymer sensors. They reached over 80 % accuracy by using CRYnose-320 technology. Song *et al.* (2013) investigated if the electronic nose that contains 18 different metal oxide semiconductor gas sensors can be used for measuring and modelling flavor quality changes of refined chicken fat during controlled oxidation. Partial least squares regression (PLSR) was used to predict relationships between the chemical parameters, GC–MS data, free fatty acid profiles and electronic nose responses for controlled oxidation of refined chicken fat. The results showed that peroxide value (PV) and acid value (AV) were predicted by the electronic nose responses with 73 % accuracy.

CHAPTER 3

MATERIAL AND METHOD

3.1 Gas Resistance Based Electronic Nose

BME688 gas sensor is a kind of electronic nose that can detect different gases by measuring their unique electronic fingerprint and therefore distinguish different gas compositions. However, the sensor should be trained with different gas compositions to be teached by Machine Learning. BME688 provides unique capabilities to detect different pollutant sources:

- High sensitivity and selectivity towards various gases
- Sensitivity and selectivity can be adjusted during operation
- Low power consumption
- Small footprint
- Additional sensing capabilities for temperature, relative humidity and barometric pressure

BME Board x8 (which is also referred to as BME Development Kit) is an experimental board with eight BME688 sensors.This board allows you to test and collect data with multiple configurations concurrently. This intrinsically increases the accuracy and reduces development time also. It stores data on the external Micro SD Card that can be inserted on the board itself (Bosch, 2021).

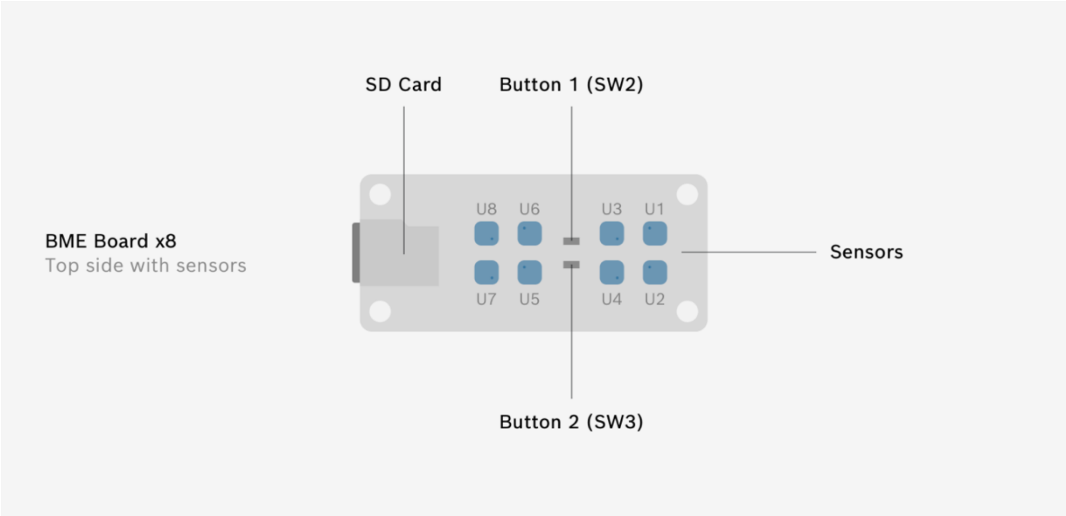

Figure 3.1. BME Board x8 Schematic (Bosch, 2021).

3.1.1 Gas Measurement Principle

The BME680 is a metal oxide-based sensor that detects volatile organic compounds (VOCs) by adsorption on its sensitive layer. Most volatile chemicals that pollute indoor air respond to BME680 (exception CO_2). The BME680 can detect the total amount of VOCs in the air, such as outgassing from paint, furniture, or rubbish, as well as high VOC levels from cooking, consumption of food, and exhaled breath or perspiration.

BME680 will output resistance values as a raw signal, which will alter owing to variable VOC concentrations. The lower the resistance, the higher the concentration of lowering VOCs, and vice versa. Because this raw signal is impacted by factors other than VOC concentration, it is difficult to interpret.

BME688 gas sensor's measuring principle is based on a metal oxide, such as SnO_2, that has been heated to temperatures above 300°C. As ionosorbed oxygen species, O_2 molecules are adsorbed on the surface of metal oxide agglomerates and trap electrons (*e.g.* O-). Reducing gases, such as CO, combine with oxygen species to produce gaseous CO_2, which raises electrical conductivity. The pressure sensor employs a novel piezoelectric MEMS (micro-electro-mechanical-systems) technology to transform pressure changes into an electrical charge, which is then used to determine temperature. The humidity sensor is a polymer that changes resistance in response to variations in humidity.

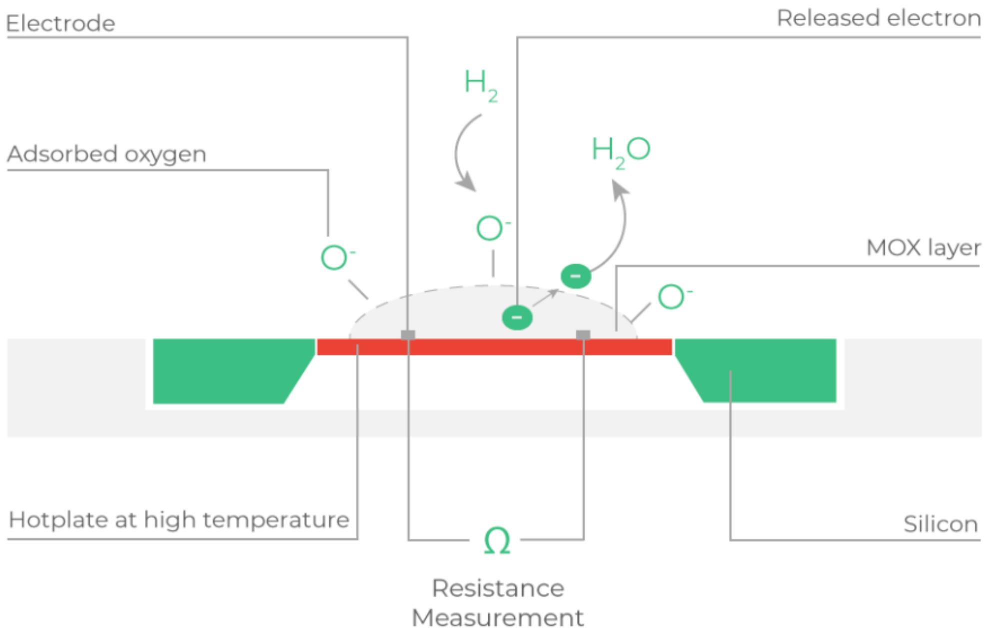

Figure 3.2. Working Principle of BME688.

3.1.2 Configuration

Each sensor at the board can operate with a specific setting which consists of Heater Profile (HP) and Duty Cycle (RDC). Hence, the term is stated as HP/RDC Combination.

Heater Profile: The gas sensor portion of the BME688 is placed on a hotplate to change its temperature. During scanning cycles, this procedure occurs. Depending on the heater profile, the gas sensor component can be heated to one of ten specified temperatures (HP). The gas sensor part's surface chemistry varies depending on the heater profile used during scanning cycles. As a result, the sensor's sensitivity and selectivity vary according on the gas's heater profile.

Heater profile has a significant effect on the performance of the gas sensor. Choosing the right Heater Profile depending on the use case can be crucial for getting good performance from the sensor. It is one of the most important features of BME688. This allows the sensor to adopt to individual gases by adopting the measurement scheme to the gases' individual

Temperature profile with 10 steps which is run through by the sensor during a Scanning Cycle. Each of the 10 steps consists of:

- A temperature
- A duration (timebase is 140 ms)

At the end of each step, a measurement is recorded and stored in the data.

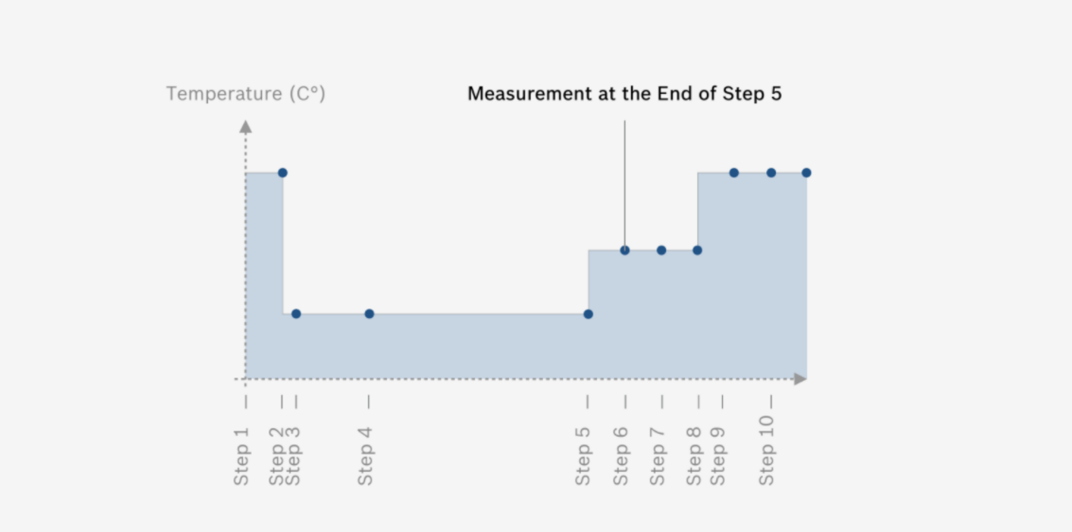

Figure 3.3. Heater Profile Schedule of Gas Sensor (Bosch, 2021).

Duty Cycles: Duty cycle refers to a combination of a Scanning Cycle and a Sleeping Cycle. Scanning cycle can be described as one iteration of the Heater Profile. Hence, each Scanning Cycle can be determined from the period of the Heater Profile used. Furthermore, Sleeping Cycle can be described as the sensor's not working mode and it doesn’t perform any measurement during this cycle. The duration of the whole Duty Cycle defines the algorithm's response time.

3.2 Multi Sensor Array Based Electronic Nose

Since the first electronic nose is working based on TVOC related gas resistance, I decided to develop a new electronic nose which can measure different parameters. Due to the fact that parameters like PM, CO_2, HCHO can be quite important to detect pollutant source, those parameters should be measured.

3.2.1 Hardware Development

Microcontroller

Arduino Mega2560 is chosen as a microprocessor due to its computational power and large range of input-output opportunities compared to Arduino Uno. The Arduino board is a microcontroller board based on the Atmega 2560 microprocessor that is open-source. The processing or wiring language is executed by this board's growing environment. With their simple to use platform, these boards have re-energized the automation sector, allowing anybody with a little or no technical background to begin learning the essential skills to program and run the Arduino board. These boards are used to link to applications on your PC such as MaxMSP, Processing, and Flash, or to expand independent interactive items. This page provides an overview of the Arduino Mega 2560 board, including its pin diagram and specs.

The ATmega2560 microcontroller is used in microcontroller boards such as the "Arduino Mega." It has 54 digital input/output pins (16 analog inputs, 14 PWM outputs), 4 hardware serial ports (UARTs), an ICSP header, a power jack, a USB connection, and a RST button. It also has a crystal oscillator of 16 MHz, an ICSP header, a power jack, a USB connection, and a RST button. This board primarily contains all of the components required to support the microcontroller. As a result, this board's power may be supplied by connecting it to a PC through a USB connection, a battery, or an AC-DC converter. A base plate can be used to shield this board from an unexpected electrical discharge (ElProCus, 2019).

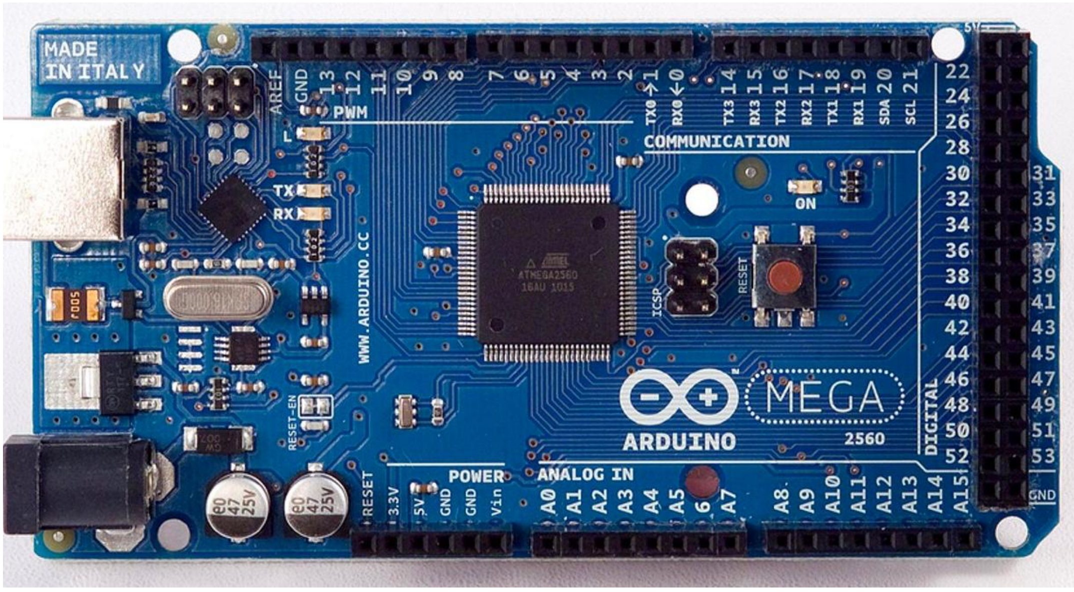

Figure 3.4. Arduino Mega2560 (ElProCus, 2019).

<u>Sensors</u>

Carbon Dioxide Sensor: The great accuracy of the Sensirion SCD30 sensor's carbon dioxide detecting technique NDIR was a major factor in the decision. When it comes to NDIR sensors, the fact that CO_2 molecules absorb certain infrared wavelengths is what makes them so effective. As the quantity of CO_2 rises, so does the amount of radiation it can absorb. 4.3 m is the wavelength that CO_2 absorbs the most compared to other gases in the atmosphere.

A light bulb emitting infrared radiation is put on one side of the tube. On the other side, two sensors with optical filters are installed. An 4.3 m band-pass filter is utilized to measure the intensity of radiation Id from the first sensor (CO_2). It employs a band-pass filter to measure the radiation intensity I0 at a wavelength where gas molecules are least likely to absorb it (the reference sensor) (usually 4 m). The Beer-Lambert

Law is used to calculate the CO_2 content from the intensity of light detected at these wavelengths.

Id/I0=e^KCL

A CO_2 absorption coefficient of K equals the CO_2 concentration of C and a path length of L equals the distance between a radiation source and light detectors. The reference radiation intensity is I0.

Fluctuations in radiation intensity are eliminated by the reference sensor. This indicates that as the intensity changes, Id / I0 remains constant since Id changes in the same manner in electronic form (SoS, 2021).

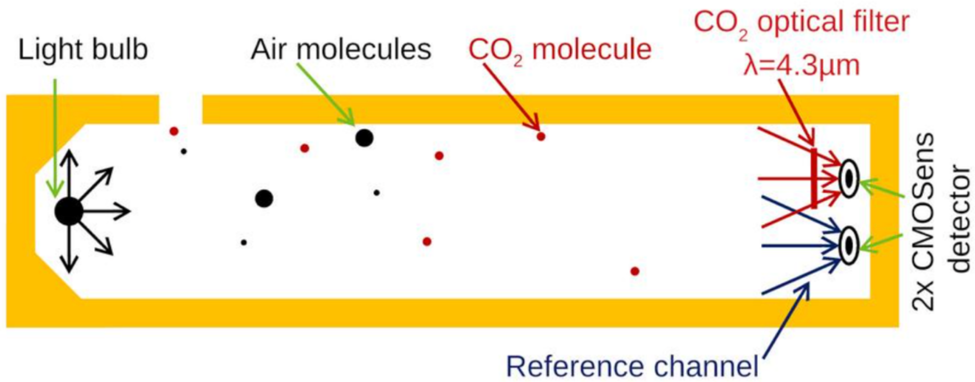

Figure 3.5. Working Principle of NDIR Technology (SoS, 2021).

For concentrations between 400 and 10000 ppm, the accuracy of SCD30 is around 30ppm which makes it great compared to its low price ($30). The SCD30 contains built-in temperature and humidity sensors, as well as controls to adjust the current altitude, in order to increase accuracy. For even more precision, the SCD30 also collects measurements from the ambient pressure.

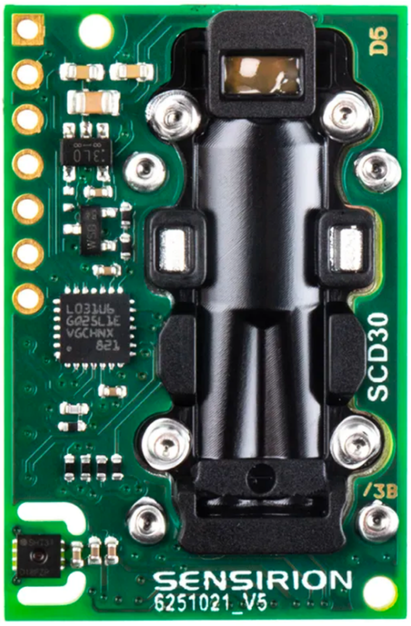

Figure 3.6. Sensirion SCD30 (SoS, 2021).

Particulate Matter Sensor: Along with its dust collection resistance and very precise calculation approach, Sensirion SPS30 was chosen as the particulate matter sensor.

Figure 3.7. Sensirion SPS30 Particulate Matter (SPS30, 2020).

Laser scattering is the basis for the Sensirion SPS30's operation. A fan within the sensor creates a regulated airflow. An internal feedback loop between the microprocessor and fan maintains the sensor's airflow at a constant rate. The airflow carries particulate matter (PM) from the sensor's entrance to its exit. Particles in the air travel across a concentrated laser beam, generating light scattering in tandem with the photodiode. Sensirion's patented algorithms operate on the SPS30 internal microprocessor to transform the scattered light into a mass/number concentration output (SPS30, 2020).

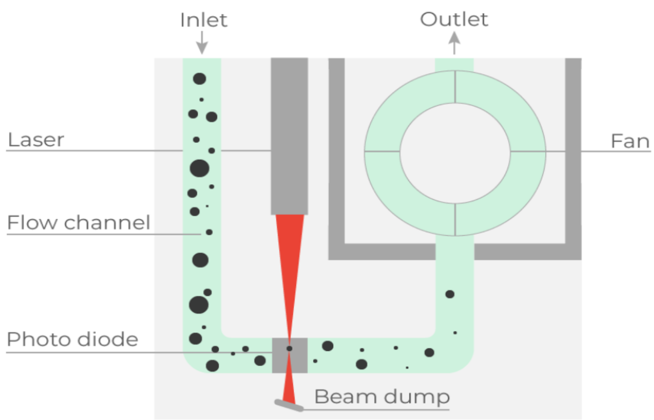

Figure 3.8. Internal Structure of SPS30.

A correct front-end electronics design and the manufacturer's algorithms make a significant impact in the calculation of mass concentration from scattered light. Assuming a constant mass density in calibration, most low-cost PM sensors use the measured particle count as a multiplier to determine mass concentration. There are many distinct particle kinds with many different optical characteristics in normal life, from "heavy" home dust to "light" combustion particles. This assumption only works if the sensor monitors a single particle type (for example tobacco smoke). No matter what sort of particle is being detected, Sensirion's unique algorithms employ an innovative technique to accurately estimate mass concentration. In addition, this method provides an accurate estimate of the bin sizes. Unlike most current consumer PM sensors on the market, the PM4.0 is an extra in output. It enables new applications for improved aerosol identification accuracy to enable new applications to be developed depending on particle composition (SPS30, 2020).

Dust accumulates on the critical optical sections of the PM sensor, including the laser, the photodiode, and the beam-dump, and this causes the sensor's output to fluctuate. Sensirion created and integrated a patented, proprietary flow path technology into the SPS30 that prevents dust and dirt from building up on the optical components.

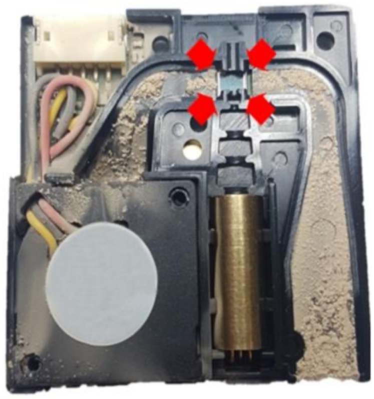

Figure 3.9. Dust Prevention of SPS30 (SPS30, 2020).

Formaldehyde Sensor: DFRobot Gravity HCHO sensor has been chosen as formaldehyde sensor due it's fast response and high accuracy capability. This sensor may be used to determine the concentration of HCHO (Formaldehyde) in the air. The sensor module is capable of precisely detecting and measuring HCHO. Numerous benefits exist, including strong anti-jamming capabilities, high stability, up to 0.01ppm sensitivity, long life , and rapid reaction capacity (Gravity, 2021).

The reagent or colorimetric card is the most often used solution for measuring HCHO in air. The colorimetric detector utilizes a solid phase colorimetric reagent composed of AHMT, ZnO, KIO_4, and agar that discolors from white to purple when exposed to a certain concentration of HCHO gas. Although the color degree is proportional to the HCHO concentration levels, visual assessment often requires subjective judgment, which results in inaccuracy. Additionally, this approach will take a significant amount of time and will be utilized just once. And the conclusion is not particularly precise, leaving you with just probabilities (Sekine *et al.*, 2016). The second solution is a VOC gas sensor. It can detect HCHO in general terms, since this sensor measures the whole content of VOC gas, not only HCHO. The DFRobot Gravity Formaldehyde sensor module is capable of properly detecting and measuring the concentration of HCHO gas.

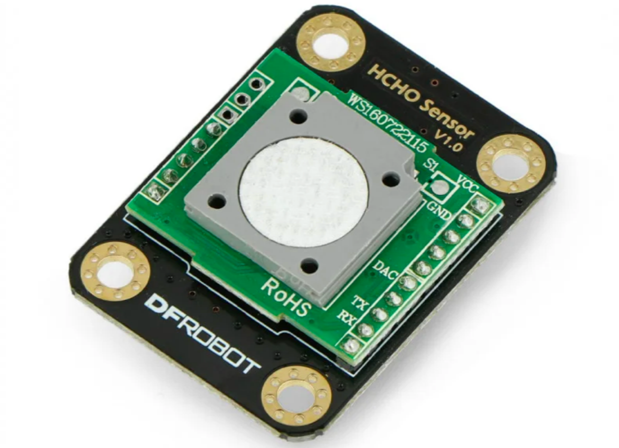

Figure 3.10. DFRobot Gravity HCHO Sensor (Gravity, 2021).

VOC, Ethanol, H2 Sensor: SGP30 has been chosen as a VOC sensor due to its wide range of output gas capabilities(H2 and Ethanol). The SGP30 is a digital multi-pixel gas sensor that is ideal for use in air purifiers, demand-controlled ventilation, and Internet of Things (IoT) applications. CMOSens technology combines a digital I2C interface, a thermally controlled micro hotplate, and two preprocessed indoor air quality signals on a single chip. The SGP30 is the first metal-oxide gas sensor to have several detecting components on a single chip, providing more comprehensive information regarding air quality. The sensing element shows resistance to polluting gases encountered in real-world applications, resulting in exceptional long-term consistency and minimal drift (Zuo, 2021). The SGP30 provides two complimentary air quality values using a dynamic baseline correction method and on-chip calibration parameters. Both the H2 Signal and the Ethanol Signal can be used to calculate the gas concentrations c relative to a reference concentration cref using the formula:

$\ln(C/Cref)=(Sref-Sout)/a$ with $a = 512$,

Sref: the H2 Signal or Ethanol Signal output at the reference concentration

Sout : Sout H2 or Sout EthOH.

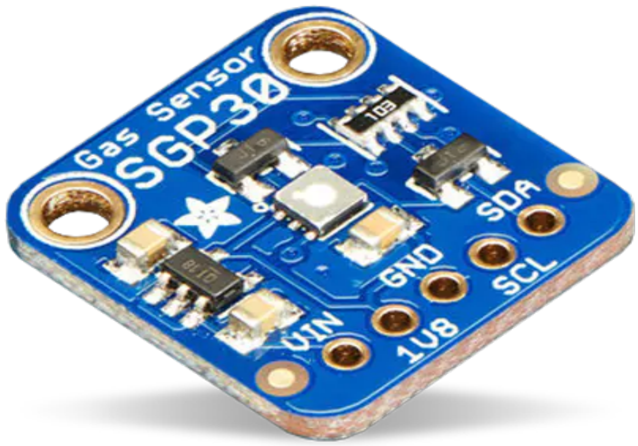

Figure 3.11. SGP30 VOC Sensor (Sensirion, 2021).

Communication Module

NodeMCU has been chosen as a communication module due to it's easily configurable hardware and open source software. NodeMCU (Node MicroController Unit) is based on the ESP8266, a cheap System-on-a-Chip (SoC). An operating system and SDK are also included with the ESP8266, which was created and produced by the company Espressif Systems. That makes it a great solution for Internet of Things (IoT) applications of all types.

Although the ESP8266 is difficult to access and utilize as a chip, it is also quite powerful. Simply turning it on or sending a keystroke to its computer requires soldering proper analog voltages to its pins. Low-level machine instructions must also be written in order for the chip to understand them. With the ESP8266 embedded controller chip, mass-produced electrical devices may achieve this degree of integration.
It is capable of either hosting the application or offloading all Wi-Fi networking operations from another application processor to another application processor. ESP8266 NodeMCU can be coupled with sensor-specific devices via its GPIOs with minimum work up-front and minimal burden during runtime when combined with the robust on-board processing and storage capabilities.

Figure 3.12. NodeMCU.

Table 3.1 Overview of Measurement Capabilities

Target Parameter	Sensor	Sensor type	Measurement Range
PM2.5	SPS30	Optical Laser	0-1000 ug/m^3
PM10.0	SPS30	Optical Laser	0-1000 ug/m^3
HCHO	DFRobot	MOX	0-5 ppm
TVOC	SGP30	MOX	0-10000 ppb
eCO_2	SGP30	MOX	400-5000 ppm
Raw Ethanol	SGP30	MOX	0-20000 ppm
Raw H_2	SGP30	MOX	0-20000 ppm
CO_2	SCD30	NDIR	400-2500 ppm
Temperature	SCD30	Thermal	0-60 C
Humidity	SCD30	Thermal	0-100 %RH

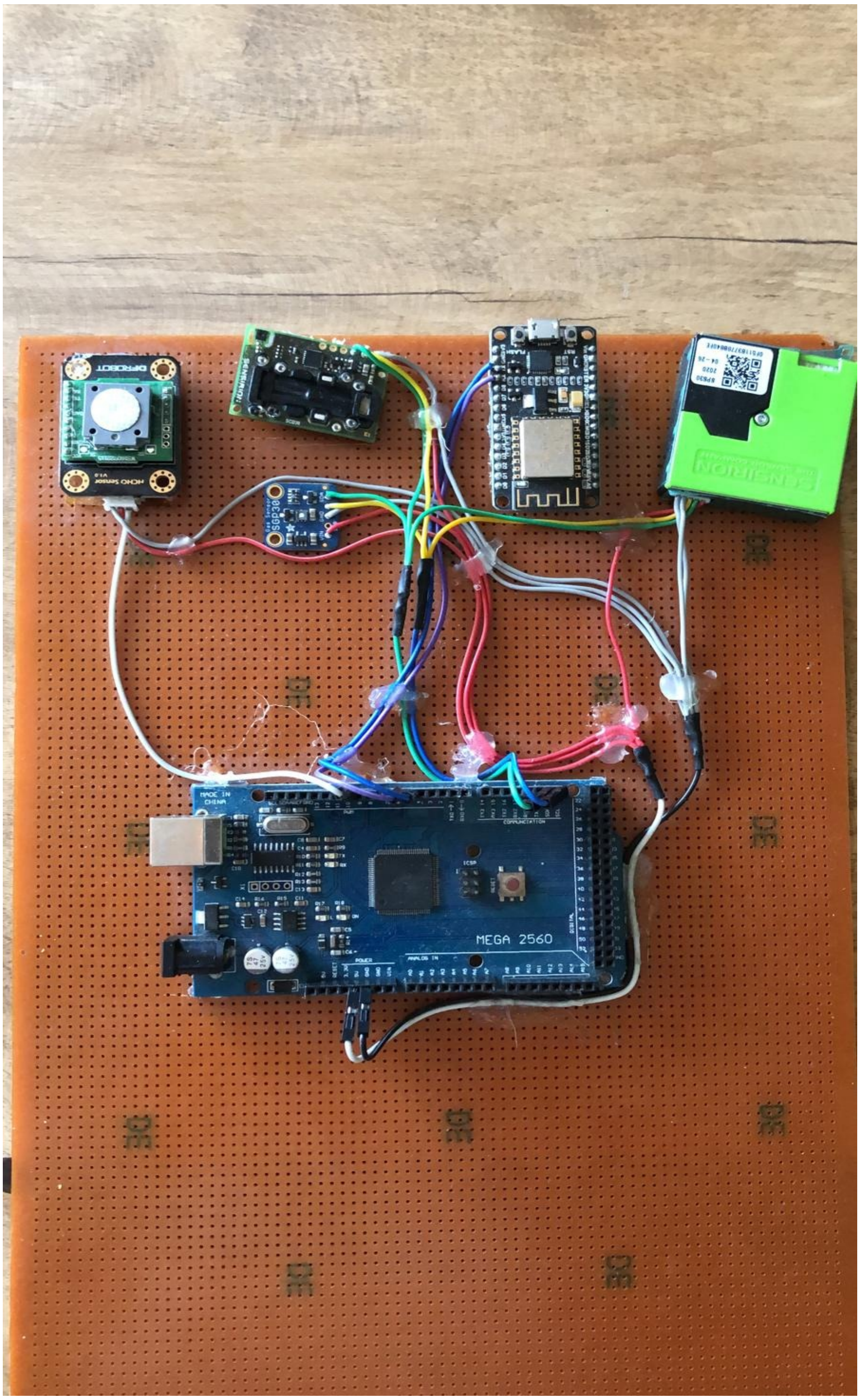

Figure 3.13. Electronic Nose II developed by the author for the study.

3.2.2 Software Development

Open source libraries of sensors and the communication module that are developed for the Arduino environment have been used on the software design. The sensors are programmed to send data every 10 seconds and after every 10 seconds the communication module transmits the data to the database via WiFi connection. Arduino communicates with sensors and the communication module via different serial communication methods like UART and I2C.

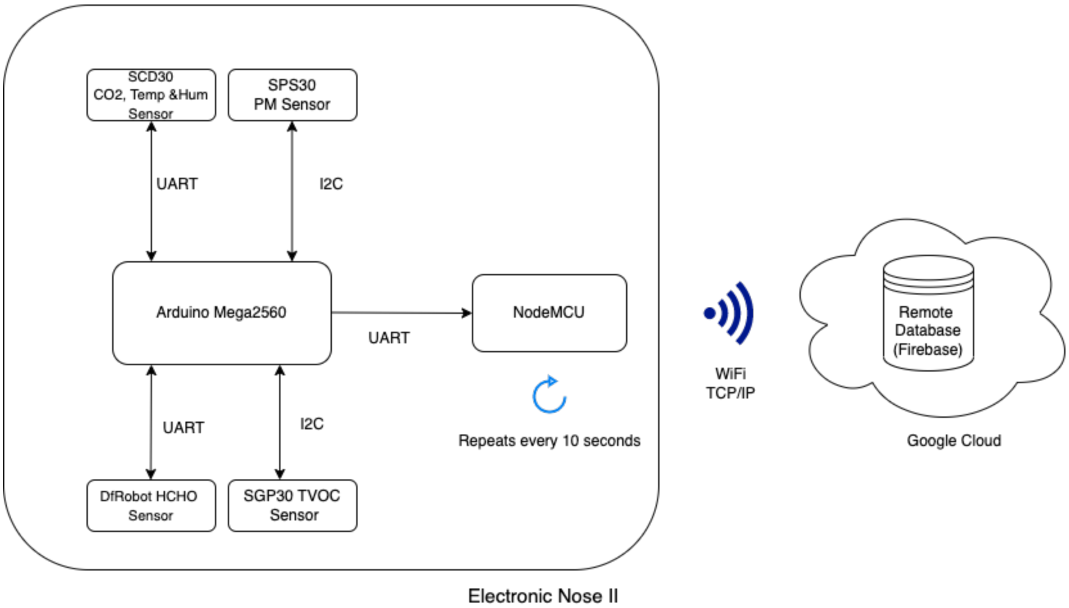

Figure 3.14. System Architecture of Electronic Nose II.

The Database: Firebase real-time database has been chosen due to it's simple data storage mechanism and open source software that gives opportunity to directly store data in the database without needing any server-side application. Firebase Realtime Database is a cloud-based database that stores data in JSON format. The data is synced in real time with each client connected. The Firebase Realtime Database is a NoSQL database that enables to store and sync data in real-time amongst users. It's a sizable JSON object that developers may manipulate in real time. The Firebase database gives the application with the current value of the data and modifications to that data through a single API. The Realtime database enables users to collaborate. It includes mobile and web SDKs that enable us to develop apps without the need of servers. When users go offline, the Real-time Database SDKs serve and save updates using the device's local cache. When the device is connected to the Internet, the local data is immediately synced (Firebase, 2021).

3.2.3 Calibration

Calibration is a process that confirms a sensor's accuracy by comparing its results to established reference values. Calibrations of air quality sensors are performed in a controlled environment under the same circumstances as the sensor's intended use case to guarantee that the sensor functions properly and is within specification across the whole operating range. To calibrate an air quality sensor, a collection of reference values with numerous data values is required.

PM Sensor: Particle counters function by collecting an air sample using a fan or pump and then measuring the quantity and size of the particles in the sample using light diffractometers. Dust may accumulate in the sensor over time, possibly interfering with airflow by settling on the fan and obstructing the device's optics. Since any error can make a big impact on the accuracy of the experiment although SPS30 has a dust prevention mechanism, the PM sensor was opened and cleaned to remove all the dust in the sensor. Then the sensor was put in a controlled and cleaned environment to make sure it measures no dust in the environment.

TVOC Sensor: MOS sensors operate by heating a sheet of metal oxide particles. The sheet absorbs oxygen, the oxygen interacts with the target gas on the sheet, and the consequent change in the electrical resistance of the sheet may be utilized to measure the TVOC concentration.
The sensor that was used in this experiment has production level calibration. The manufacturer calibrates the sensors using the ISO16000-29 standard "Test procedures

for VOC detectors." Apart from ethanol (EtOH), the sensors are also calibrated with breath-VOC (b-VOC), the composition of which is specified in the table below.

Table 3.2. Used Gases for Calibration of the TVOC Sensor

Molar fraction	**Compound**	**Production Tolerance**	**Certified accuracy**
5 ppm	Ethane	20 %	5 %
10 ppm	Isopropene	20 %	5 %
10 ppm	Ethanol	20 %	5 %
50 ppm	Acetone	20 %	5 %
15 ppm	Carbon Monoxide	10 %	2 %

CO_2 Sensor: Carbon dioxide sensors, on the other hand, are perhaps an exception to this rule. For periods when the building is empty, the carbon dioxide content in the external air may be utilized, which is normally about 400 parts per million.
The majority of carbon dioxide sensors are non-dispersive infrared (NDIR) sensors that detect CO_2 concentrations by gas spectrometry. NDIR sensors are not subject to deterioration in the same manner as TVOC sensors are, since no oxidation occurs on the sensor's surface. Additionally, some NDIR sensors feature a cover over the air intake to avoid dust accumulation.

NDIR sensors, like any other sensor, will drift with time. Unlike particle matter and volatile organic compounds, however, we may utilize outside air as a baseline for adjustment. Because humans generate the great majority of carbon dioxide released inside buildings, when a building is vacant, the mixing of outside and interior air returns indoor CO_2 levels to those found in fresh air. Then CO_2 values acquired between occupancy periods may be calibrated back to 400 ppm to account for any drift during the sensor's lifespan.

An Automatic Baseline Calibration algorithm was implemented into the software of the electronic nose. The algorithm continuously monitors the sensor's lowest value over a predefined time period and gradually normalizes for any long-term deviation found as compared to the predicted fresh air CO_2 concentration of 400 parts per million (or 0.04 percent vol). Carbon dioxide levels in most indoor applications fall to near-outdoor levels at some point within a week. By collecting measurements over a seven-day period and comparing the minimum number to the sensor's 400 point, the sensor may determine whether the zero point has to be adjusted.

3.3 Chosen Materials for Detection

Increased Indoor Air Pollution levels may cause specific health problems, whereas excessively high levels may be caused by situations such as air fresheners, combustion appliances, or water damage. Different parameters may originate from a variety of sources and have a variety of health consequences, including tobacco smoke, fire, and combustion equipment. However, for the sake of this research, focus will be on the attitude toward people. There are undoubtedly more activities and sources, but they are confined to five because of their prevalence in the indoor environment.

Office Air: The office air is contaminated if it has an excessive amount of dust from carpets and furniture, or if it contains an excessive amount of ozone and beverages, or if it contains scents. The initial state of indoor air sources from office machinery.

Combustion: The second need for IAP sources is combustion activity. The surrounding air is polluted by combustion. Ambient air is the air that occurs naturally in an interior setting without the presence of additional sources of indoor air pollution. Activities such as tobacco smoke create harmful gases such as HCHO. Combustion may contaminate indoor CO_2, and particulate matter at a greater concentration than ambient air, which may be harmful to human health if it contains an excessive amount of dust from carpets and furniture or contains an excessive amount of ozone from the workplace. Smoking was considered as a surrogate for combustion for the purposes of this research.

Chemicals: The existence of chemical products or chemicals as a cause of indoor air pollution is the third requirement. Chemical goods, such as chemical cleaners that are often used in homes and businesses, may emit toxic levels of VOCs. Excessive VOC levels may result in respiratory disorders such as lung cancer. Thus, chemical cleaning products and cologne will be used as proxies for the existence of chemicals in this research.

Food and Beverage: The presence of food and drinks is the fourth criteria for sources of indoor air pollution. Cooking activities and some foods and drinks release volatile organic compounds (VOCs), which may create an unpleasant odor within a building. VOCs have been linked to eye discomfort, headaches, and nausea in certain individuals. Thus, pizza, curry and coffee will be used as proxies for food and beverages in this research.

Building Materials: Building materials lead to the release of air pollutants in interior spaces; as a result of the apparent interaction between the many kinds of construction materials and the room temperature, occupants are exposed to several contaminants concurrently. Different types of insulation, paints, and wood items made from these components may all contribute to pollution in a structure.

The presence of VOCs in indoor air demonstrates the tremendous complexity of the processes and interactions involved. Differences in the adsorption capabilities of different construction materials, particularly in the case of heat insulation materials and wall paints, are to be anticipated. Thus, painted tile, varnished wood and stone wool used as proxies for building materials in this research.

Table 3.3. Chosen Materials with Their Activity Classes

Chosen Material	Activity
Office Air	Office Air
Tobacco Smoke	Combustion
Cleaning Material	Chemicals
Cologne	Chemicals
Pizza	Food and Beverages
Curry	Food and Beverages
Coffee	Food and Beverages
Painted Tile	Building Material
Stone Wool	Building Material
Varnished Wood	Building Material

3.4 Experimental Setup

The experiment was conducted in a box which has 130 liters of volume. On the lid of the box, some holes are opened on the two corners to enable some airflow to appear. Inside the box, there were two different electronic noses mentioned in previous sections, which are gas resistance based and multi sensor array based.

While experimenting, the timestamps were being recorded on a table to be used in reading and analyzing the data. Data from the multi sensor array based electronic nose setup comes in UNIX epoch time which is the total seconds elapsed since January 1, 1970 (midnight UTC/GMT). On the other hand, data from the gas resistance based electronic nose sensor came in DateTime format. Thus, recording timestamps is crucial to match between two data sources.

Each substance was put in the box for 30 minutes and between each substance, the box was carried out to outside for 10 minutes to refresh the air inside the box. The temperature of the experiment surroundings remained constant throughout the experiment. Nine different substances from five categories are measured. Those categories are ambient air, chemicals, building materials, food and beverages, and combustion.

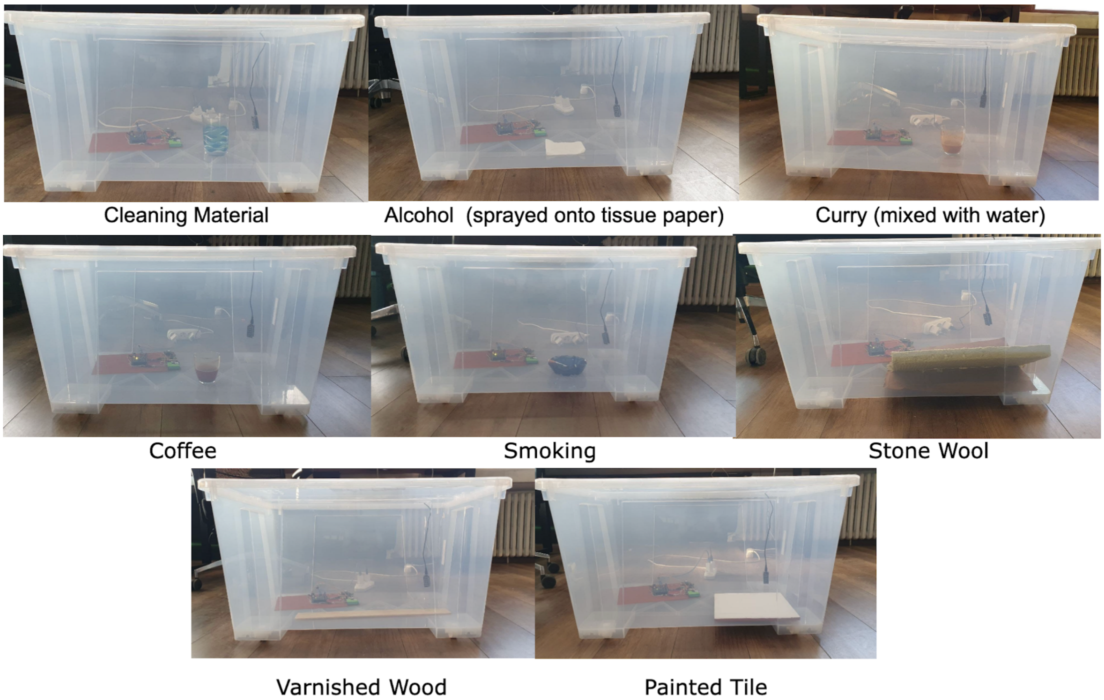

Figure 3.15. Experimental Setup with Different Materials

3.5 Data Analysis

3.5.1 Storage

The data collected from the gas resistance based electronic nose was downloaded from the SD card attached to it, while the data collected from the multisensor array-based electronic nose was downloaded from Google Cloud's Firebase Realtime Database.

3.5.2 Development Environment:

Downloaded data imported to Python Jupyter Notebook. Jupyter is an interactive development environment for notebooks, code, and data that is accessible via the web. Its adaptable user interface enables users to configure and organize data science, scientific computing, and machine learning operations.

3.5.3 Data Visualization:

Firstly, Plotly data visualization tool has been used to to identify and analyze data trends, outliers, and patterns. Violin plot has been used to analyze distribution of the data. Violin plot displays the distributions of numeric data for one or more groups using density curves. Each curve's width is proportional to the approximate frequency of data points in each zone.

Furthermore, line plots have been used to analyze parameter behavior over time. Also, pair plot has been used for analyzing relations between binary parameter distribution. Pairplot function allows the users to create an axis grid via which each numerical variable stored in data is shared across the X- and Y-axis in the structure of columns and rows. We can create the scatter plots in order to display the pairwise relationships in addition to the distribution plot displaying the data distribution in the column diagonally.

3.5.4 Modelling:

Before modelling, the data for both electronic noses splitted into training and test with 70 % and 30 % ratio. Training set is the set of data for training the model and teaching it how to discover hidden features/patterns in the data. The test set is a distinct set of data used to validate the model following training.

Naive Bayes, K-Nearest Neighbors and Random Forest Classifier models have been used to classify materials based on sensor data. Python's sklearn machine learning library has been used to apply these algorithms on the data.
The Naive Bayes Classifier is a simple and effective classification method that enables the development of fast machine learning models capable of making accurate predictions. It is a probabilistic classifier based on Bayes Theorem, which implies that it makes predictions based on an object's probability.

K-Nearest Neighbour is one of the simplest machine learning algorithms. It is based on the approach of supervised learning. The kNN method makes an assumption about the similarity between the new case/data and the existing cases and assigns the new case to the category that is the most similar to the existing categories. The K-NN algorithm retains all available data and classifies each new data point based on its similarity to the previous data point. k denotes the number of neighbors should be calculated while class is detecting. Based on the number of k, the model accuracy can be optimized. In this research, k numbers are optimized to have best accuracy.

CHAPTER 4

RESULTS AND DISCUSSION

4.1 Experimental Results for the Gas Resistance Based Electronic Nose

During the experiment, every scanning cycle took around 160 seconds. Every material has shown different gas resistance during those scanning cycles. Pressure remained constant for every parameter. Also, temperature and humidity slightly changed based on the material.

4.1.1 Transient Variation of Temperature

Temperature and gas permeability has a relation with following formula:

$P = P0.\exp(-Ep/RT)$

where P0 denotes the pre-exponential factor (($cm^3(STP)\cdot cm)/(cm^2\cdot s\cdot cmHg)$), Ep denotes the permeation activation energy (J/mol), T denotes the temperature (K), and R denotes the ideal gas constant (8.314 kJ/(molK)).

Therefore, big temperature differences between material experiments could affect gas resistance measured by the electronic nose. However, as can be seen at Table 4.1, the maximum difference between the mean of two materials is 2.57 °C which is between Coffee and Empty Box.

Table 4.1. Analysis of Experimental Temperature Measurements with the Gas Resistance Based Electronic Nose

Material	*Mean*	*Standard Deviation*	*Min*	*Max*
Empty Box	30.80	2.08	23.24	34.28
Alcohol	32.55	1.09	28.94	34.83
Cleaning Material	32.51	1.07	28.71	34.64
Painted Tile	32.74	1.05	29.43	34.92
Varnished Wood	31.91	2.32	22.73	35.2
Stone Wool	33.37	1.01	30.42	35.57
Coffee	33.44	1.06	29.43	35.39
Curry	32.51	1.12	29.02	35.51
Smoking	32.89	1.07	29.43	35.4

96 % of whole data points are between 30°-34°C degrees which can be seen on Figure 4.1. Also temperature averagely fluctuates 3.1°C degrees for every 160 seconds which is caused by the internal scanning cycle of the electronic nose.

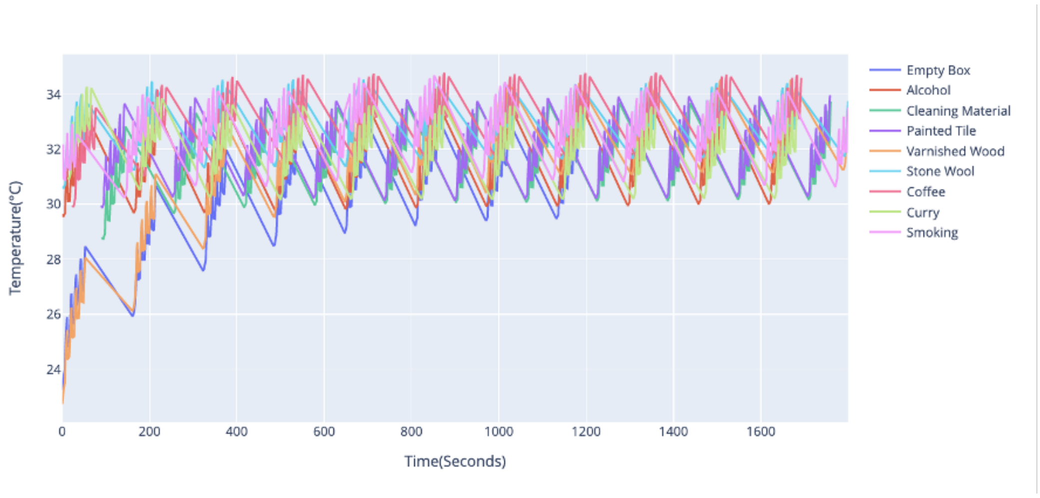

Figure 4.1. Temperature Change Over Time for the Gas Resistance Based Electronic Nose

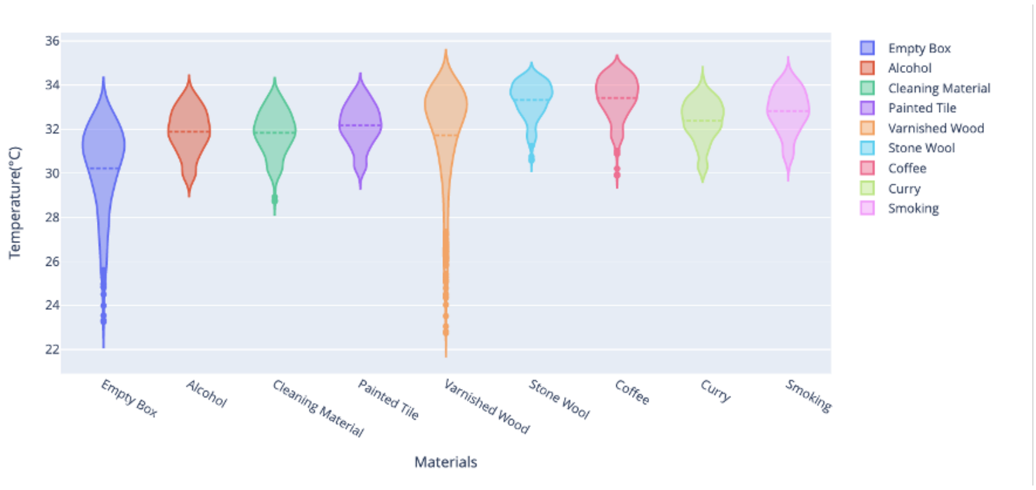

Figure 4.2. Temperature Change Levels of Different Materials for the Gas Resistance Based Electronic Nose.

Due to the lack of significant differences in the temperature patterns of various materials, it was decided not to include temperature as a parameter while developing models on top of these datasets. Temperature differences in the ambient air during the experiment can overfit the model, resulting in misunderstandings about the prediction.

4.1.2 Transient Variation of Humidity

Shooshtari *et al.*(2021) discovered that increasing the relative humidity from 10 % to 80 % reduces the electrical conductivity of the sensor by about 4 %. Therefore, humidity has an important effect on the gas resistivity of the electronic nose.

All materials except painted tile have the same humidity pattern. Painted tile created extra humidity due to the chemical composition of the paint. The paint chosen was water based paint which increased the humidity level during the experiment. The painted tile has the maximum mean and maximum data point with 22.1 % and 32.9 %. Also, it has the biggest variance with standard deviation of 3.58. On the other hand, stone wool has the minimum mean and minimum data point with 14.55 % and 12.68 %. Since stone wool does not have any water based composition, it did not create extra humidity during the experiment.

Table 4.2. Analysis of Experimental Humidity Measurements with the Gas Resistance Based Electronic Nose

Material	*Mean*	*Standard Deviation*	*Min*	*Max*
Empty Box	15.95	2.18	12.78	26.07
Alcohol	16.73	1.42	14.09	22.09
Cleaning Material	15.57	1.18	13.10	19.74
Painted Tile	22.12	3.58	13.27	32.90
Varnished Wood	15.98	2.65	13.19	32.09
Stone Wool	14.55	1.05	12.68	18.07
Coffee	19.66	1.63	15.75	25.55
Curry	19.27	1.77	14.74	25.60
Smoking	19.36	1.84	15.45	25.63

89 % of the data fluctuated between 15-23 % relative humidity during the experiment. Averagely, relative humidity fluctuated 3.9 % during every scanning cycle of the electronic nose.

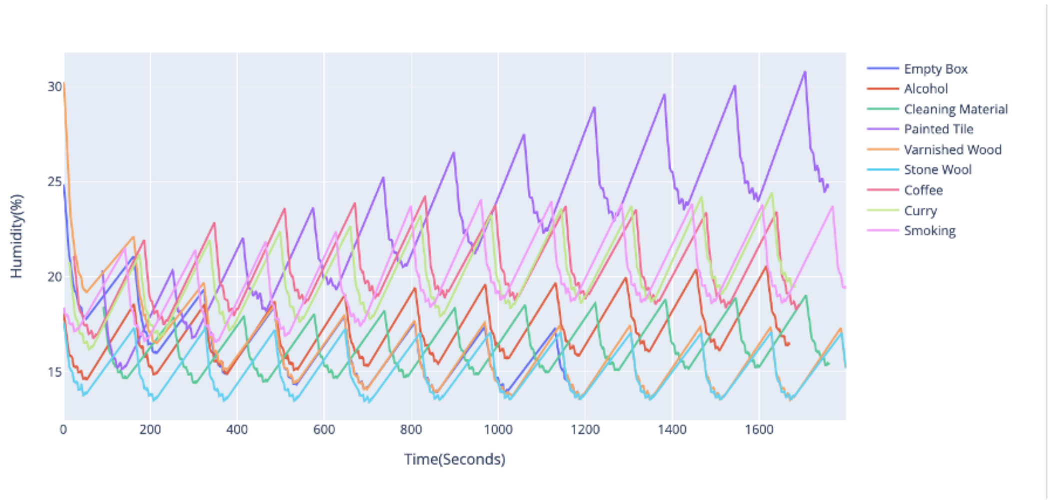

Figure 4.3. Humidity Change Over Time for the Gas Resistance Based Electronic Nose

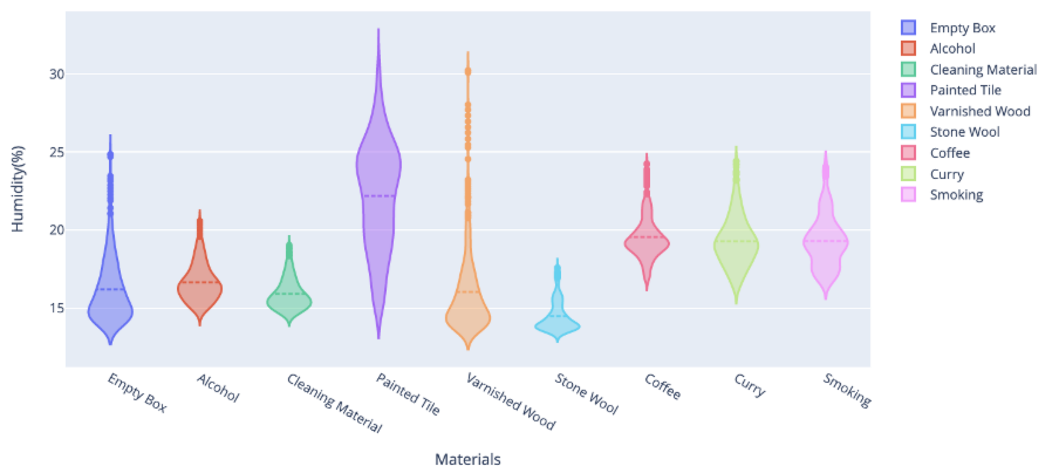

Figure 4.4. Humidity Change Levels of Different Materials for the Gas Resistance Based Electronic Nose

4.1.3 Transient Variation of Pressure

Dabill *et al.*(1996) indicate that pressure, both continuous and transient, may have an effect on the sensors’ performance in certain cases. For instance, transitory pressure fluctuations influence the gas sensors, resulting in signal spikes and increasing the likelihood of false alarms. However, as can be seen at Figure 4.5, pressure remains constant during the whole experiment at 900 hPa level.

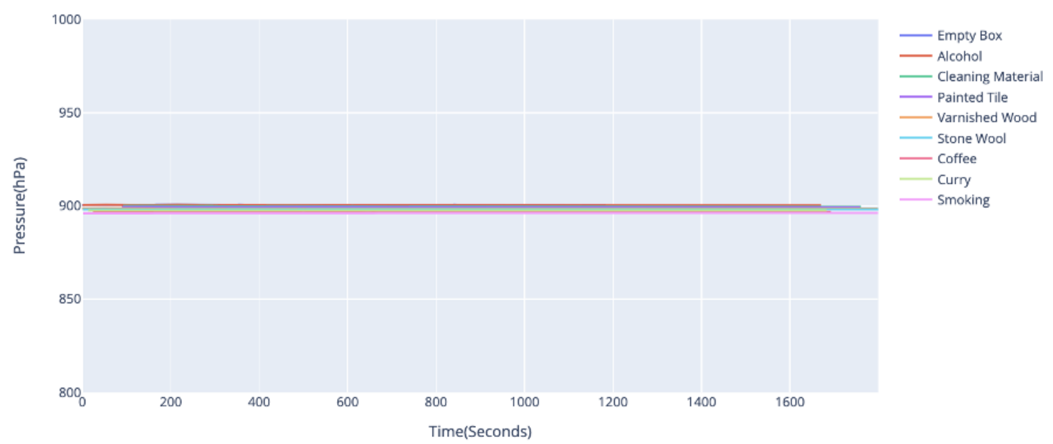

Figure 4.5. Pressure Change Over Time for Gas Resistance Based Electronic Nose

4.1.4 Transient Variation of Gas Resistance

Gas resistance is the electrical resistance of the gas sensor's MOX layer, which changes in response to the ionization of the air. The resistance of a thin metal oxide surface layer, which is changed by adsorbing volatile gas molecules from the surrounding environment, is measured in this sensor.

Gas concentrations in pure, unpolluted air (no unfavorable gases) match to the sensor's maximum resistance output of 5000K ohms. Normally, the sensor produces a value for the gas resistance in the range of 0.5K ohms to 5000K ohms and beyond. For resistance values between 0.5K and 5000K ohms, a linear relationship is assumed and the output is scaled.

Due to the sensor's heater profile throughout the scanning cycle, gas resistivity variations are extremely large. While the sensor became colder as it was heated, the gas resistance became very low at the lowest temperature.

Table 4.3. Analysis of Gas Resistance Measurements with the Gas Resistance Based Electronic Nose

Material	Mean	Standard Deviation	Min	Max
Empty Box	720.6	926.26	0.8	4000
Alcohol	6.8	12	0.7	109.4
Cleaning Material	468	745.1	6.8	2834.4
Painted Tile	267.5	472.4	2.7	2744.1
Varnished Wood	283.3	463.8	1.89	2186.7
Stone Wool	693.9	1116.8	7.7	3713.5
Coffee	754.8	1241.2	6.8	4228
Curry	192.7	328.8	2.2	1273.9
Smoking	26.3	145.3	0.7	1549.6

As predicted, the empty box, stone wool, and coffee have greater averages than the rest. These materials contain much less chemicals and volatile organic compounds (VOCs) than others. Also, materials like alcohol and smoking have lower averages than others due to their heavy chemical composition.

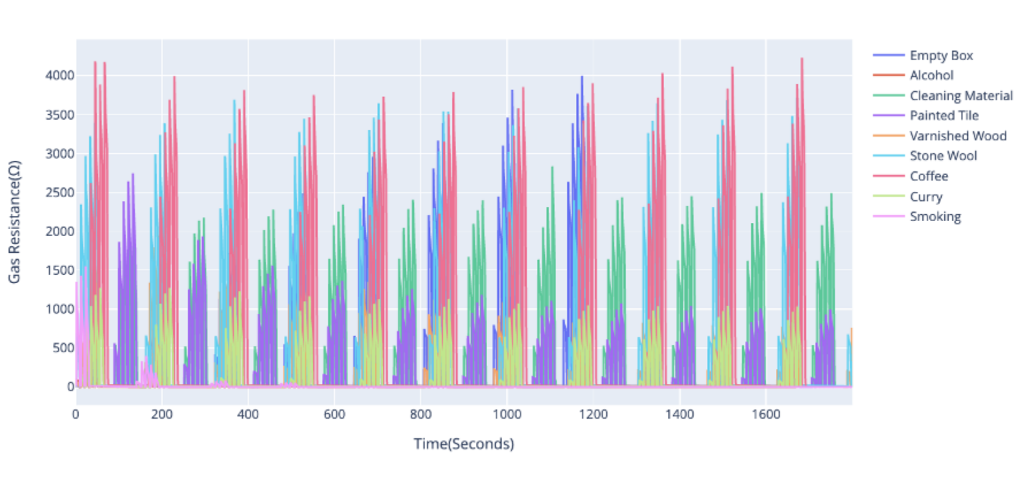

Figure 4.6. Gas Resistance Parameter Over Time for the Gas Resistance Based Electronic Nose

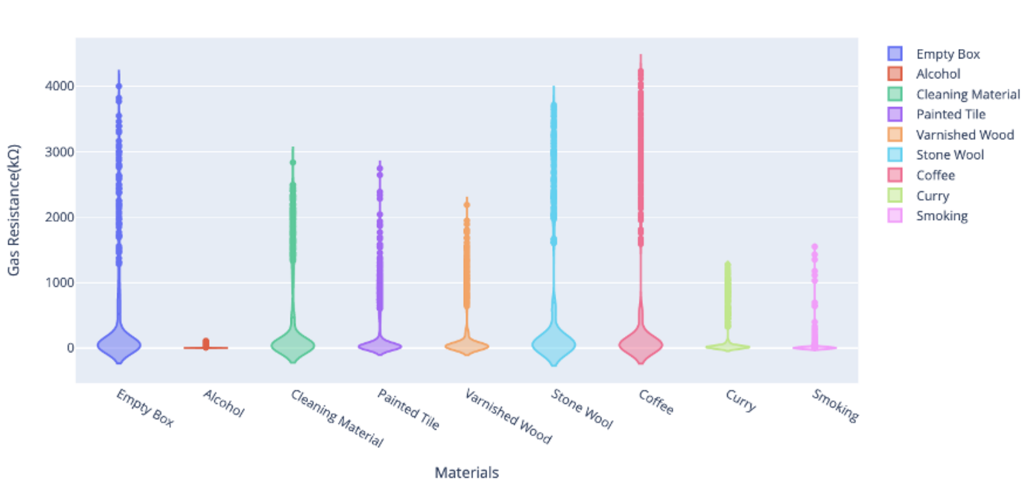

Figure 4.7. Gas Resistance Change Levels of Different Materials for the Gas Resistance Based Electronic Nose.

4.2 Analysis of the Results with Different Classification Models for the Gas Resistance Based Electronic Nose

4.2.1 The Naive Bayes (NB) Classification Model

I began by using Naive Bayes modeling to classify the data. After developing the model on the training set, predictions were made on the test set. The model's accuracy on the test set is quite low, at 0.35. Additionally, the model's accuracy on the training set was verified to rule out the possibility of overfitting. The training set's accuracy is also quite low, at 0.34. Table 4.4 shows the recall, precision, and f1 score for each parameter.

Table 4.4. Classification Report by the NB Classification Model for Experiments with the Gas Resistance Based Electronic Nose

Material	Precision	Recall
Empty Box	0.49	0.10
Alcohol	0.27	0.97
Cleaning Material	0.35	0.27
Painted Tile	0.75	0.49
Varnished Wood	0.36	0.16
Stone Wool	0.46	0.49
Coffee	0.66	0.22
Curry	0.61	0.14
Smoking	0.18	0.27

Despite the low precision of each material, alcohol has a high recall, which means that it can be easily separated from other materials if the material is alcohol. However, it can be seen from the confusion matrix in Figure 4.8. 51 % of the total data classified as alcohol which caused a lot of misclassification and decreased accuracy.

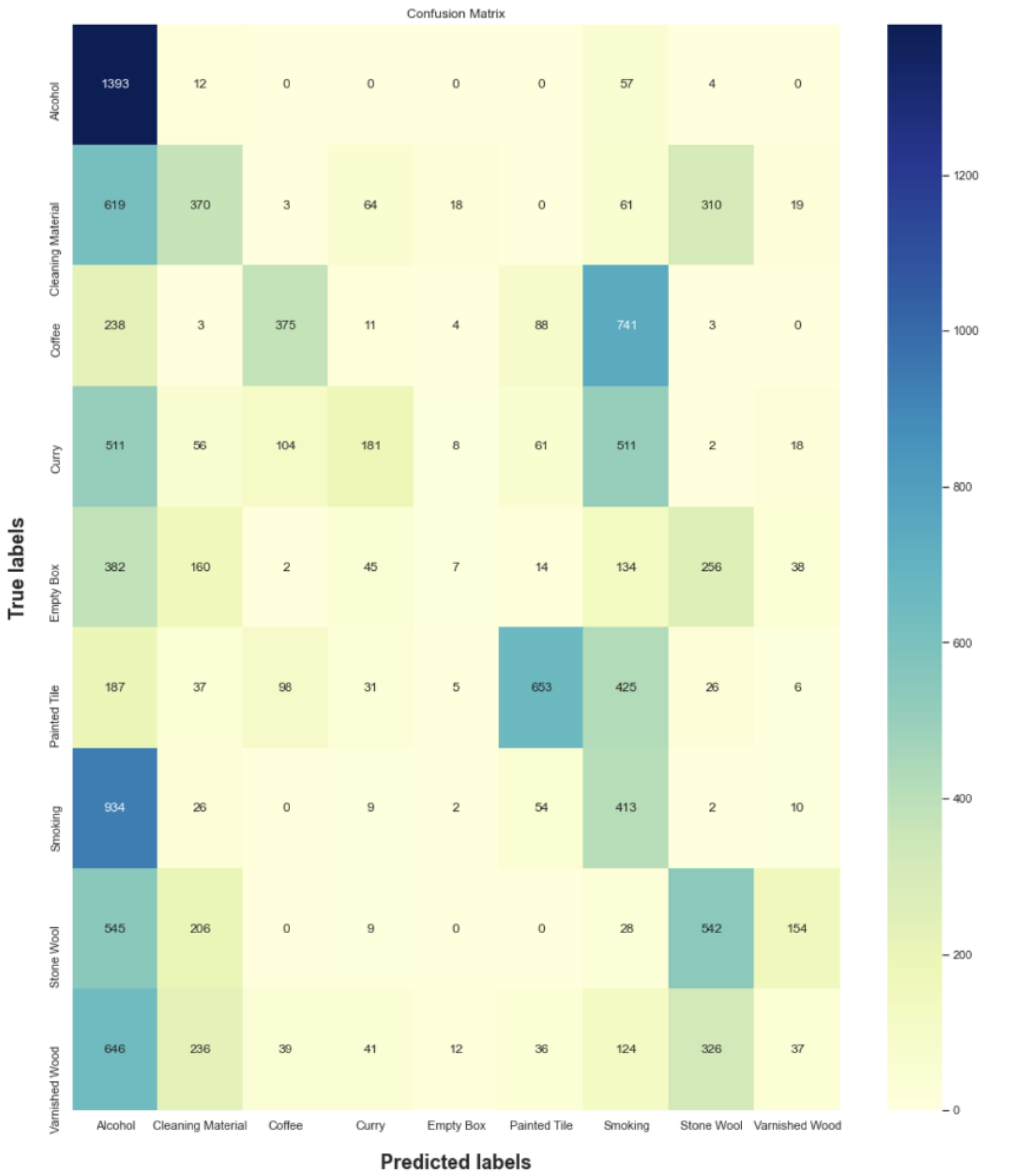

Figure 4.8. Confusion Matrix by the NB Model Classification Model for Experiments with the Gas Resistance Based Electronic Nose.

(Y-Axis represents True labels of the data and X-Axis represent predicted labels by the model)

4.2.2 The kNN Classification Model

The materials were then classified using the k-nearest neighbors algorithm. On the dataset, between two and fifteen neighbors were applied. The highest accuracy is achieved with the k=2 model. With a k=2 model, the accuracy on the training set is 0.79 and on the test set is 0.57 as can be seen on Figure 4.9.

Stone wool is the best classified material; approximately 80 % of painted tile predictions are true. With a recall rating of 0.91, stone wool has the highest recall value. The empty box has the lowest precision; only 34 % of predictions of an empty box are accurate. Additionally, it has the lowest recall value with 0.29.

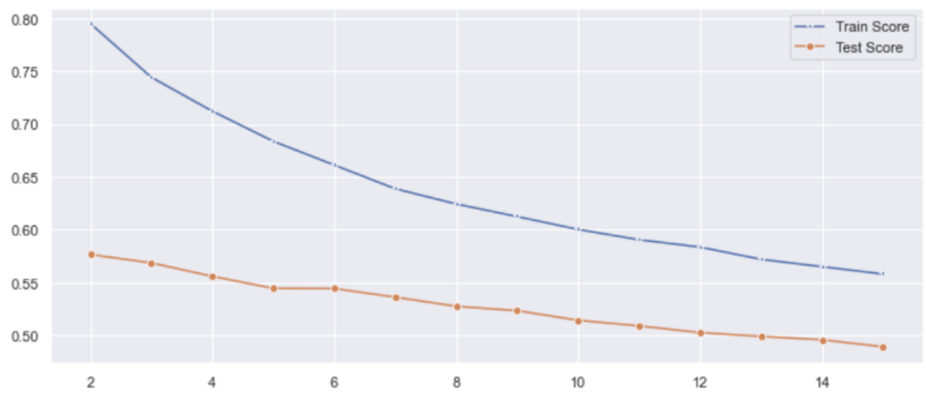

Figure 4.9. The kNN Model with Different Number of Neighbors Accuracy on Training and Test Set for the Gas Resistance Based Electronic Nose.

Table 4.5. Classification Report by the kNN Classification Model for Experiments with the Gas Resistance Based Electronic Nose

Material	Precision	Recall
Empty Box	0.34	0.29
Alcohol	0.55	0.72
Cleaning Material	0.45	0.65
Painted Tile	0.79	0.54
Varnished Wood	0.79	0.33
Stone Wool	0.80	0.91

Coffee	0.73	0.76
Curry	0.46	0.46
Smoking	0.46	0.58

Painted tile has the highest precision, 95 % of painted tile predictions are accurate. Stone wool has the highest recall value with 0.91.

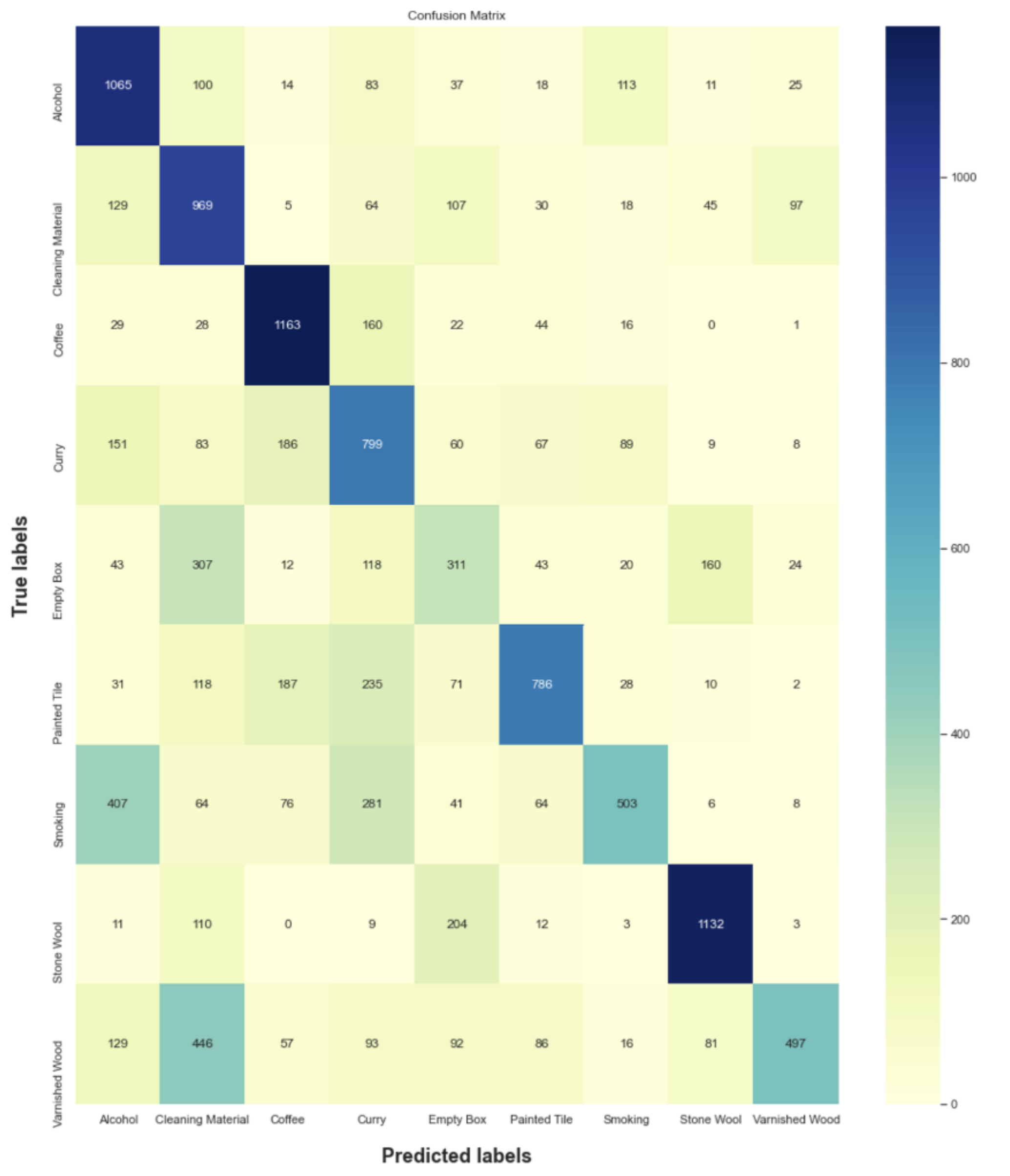

Figure 4.10. Confusion Matrix by the kNN Model Classification Model for Experiments with the Gas Resistance Based Electronic Nose.

As illustrated in Figure 4.10 above, Alcohol is frequently misclassified as tobacco smoke. 14% of alcohol samples were misidentified as smoking. Smoking is mostly misclassified as alcohol or curry. 25.2 % of tobacco smoke samples were misclassified as alcohol and 8.1% of them were misclassified as curry. On the other hand, 23.9 % of curry samples were misclassified as smoking.

4.2.3 The Random Forest (RF) Classification Model

Following to the kNN classification, the materials are classified using the Random Forest Classifier. The mode has been constructed using ten, fifty, and one hundred decision trees. There are no significant differences between the three models. As a result, I chose ten trees. The classification accuracy of 10, 50, and 100 tree models are 0.82, 0.81, and 0.81.

The importance of features has been determined by calculating the information gain of each feature using gini impurity. Gas resistance has the highest importance score of 0.43, followed by humidity at 0.36.

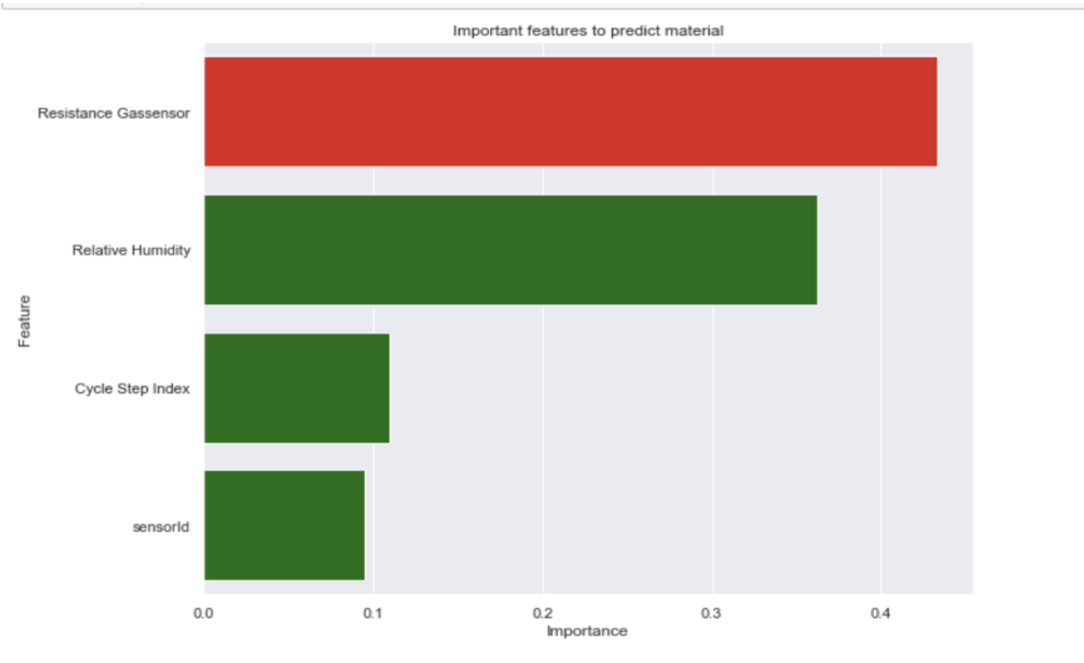

Figure 4.11. Feature Importance Scores by the RFC Model for the Gas Resistance Based Electronic Nose.

All materials have higher precision and recall values than kNN and Naive Bayes models. painted tile and varnished has relatively low recall values than others as can be seen at Table 4.6.

Table 4.6. Classification Report by the RF n_tree=10 Classification Model for Experiments with the Gas Resistance Based Electronic Nose

Material	Precision	Recall
Empty Box	0.58	0.43
Alcohol	0.93	0.95
Cleaning Material	0.64	0.77
Painted Tile	0.82	0.74
Varnished Wood	0.80	0.66
Stone Wool	0.84	0.92
Coffee	0.93	0.94
Curry	0.76	0.86
Tobacco Smoke	0.90	0.85

All materials have higher precision and recall values than kNN and Naive Bayes models. Empty box and varnished wood have relatively low recall values than others as can be seen at Table 4.7.

As illustrated in Figure 4.13 below, most of the data points are classified accurate. The most misclassified material is varnished wood. 15.1% samples of painted tile misclassified as cleaning material and 13.7% samples of them misclassified as stone wool.

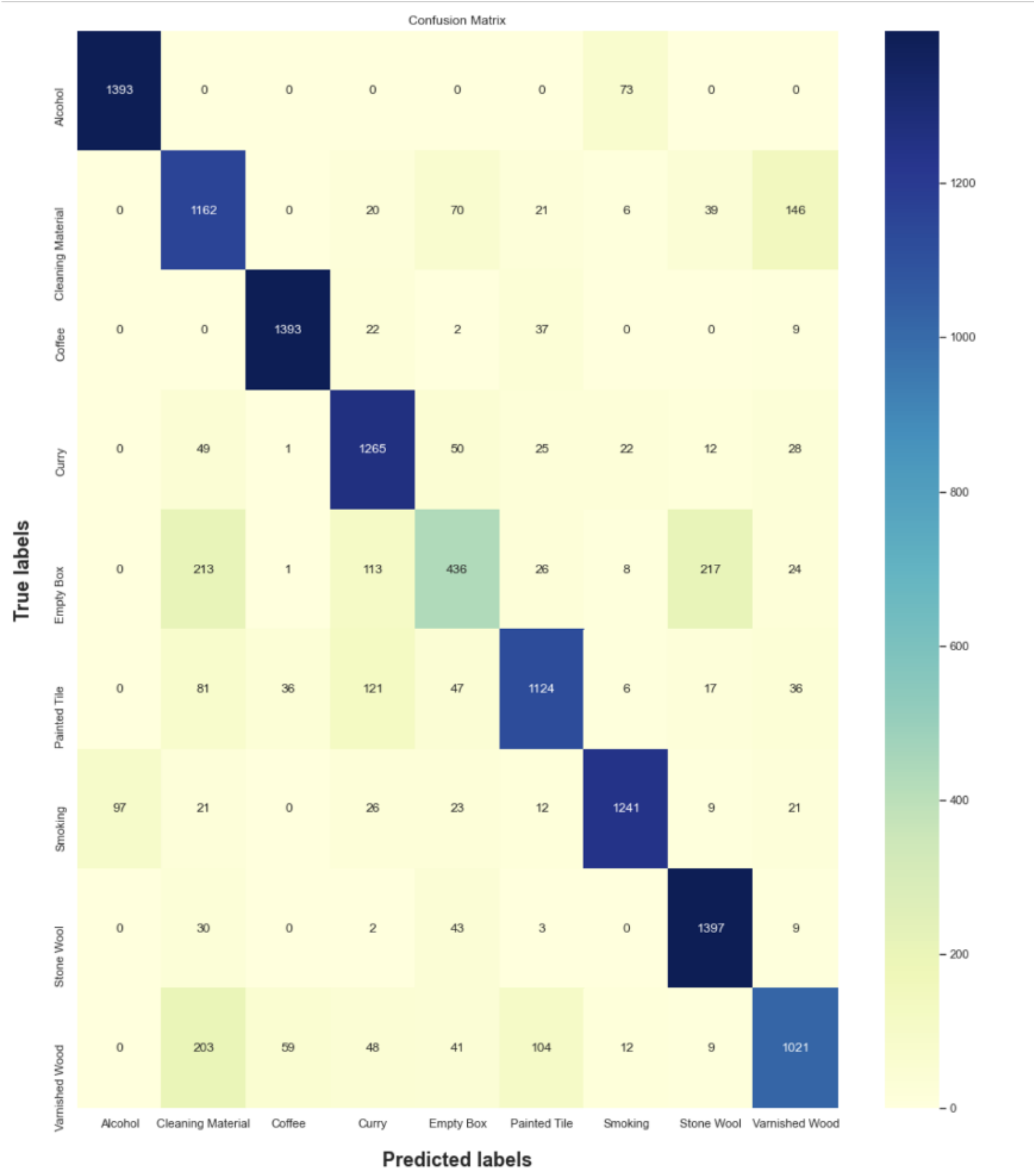

Figure 4.12. Confusion Matrix by the RF n_tree=10 Classification Model for Experiments with the Gas Resistance Based Electronic Nose.

4.3 Experimental Results for the Multi Sensor Array Based Electronic Nose

4.3.1 Transient Variation of CO_2

CO_2 is most often created by the air we breath, although it may accumulate inside in poorly ventilated places. CO_2 levels in the building are determined by a variety of factors: The number of persons who can fit into a place, the length of time that the place has been inhabited, the quantity of fresh outside air that enters the region, the area's dimensions, the extent to which combustion by-products contaminate the indoor air (*e.g.*, idling vehicles near air intakes, leaky furnaces, tobacco smoke).

Smoking had the highest average concentration at 1711 ppm, as predicted. Other materials have an average of 490 to 585 ppm. Except for smoking, all data is in the range of 480-600 ppm. Except for the smoking, the shift in CO_2 levels is produced by the indoor ambient air of the room, which is inhabited by two or three persons during the experiment.

Table 4.7. Analysis of Experimental CO_2 (ppm) Measurements with the Multi Sensor Array Based Electronic Nose

Material	**Mean**	**Standard Deviation**	**Min**	**Max**
Empty Box	514.9	3.1	508	521
Alcohol	534.6	4.4	529	5444
Cleaning Material	503.2	7.9	490	517
Painted Tile	553.5	3.4	543	559
Varnished Wood	560	2.4	554	565
Stone Wool	585.6	5.02	565	595
Coffee	551.1	3.4	544	557
Curry	490	4.6	479	498
Tobacco Smoke	1711.7	472.1	618	2107

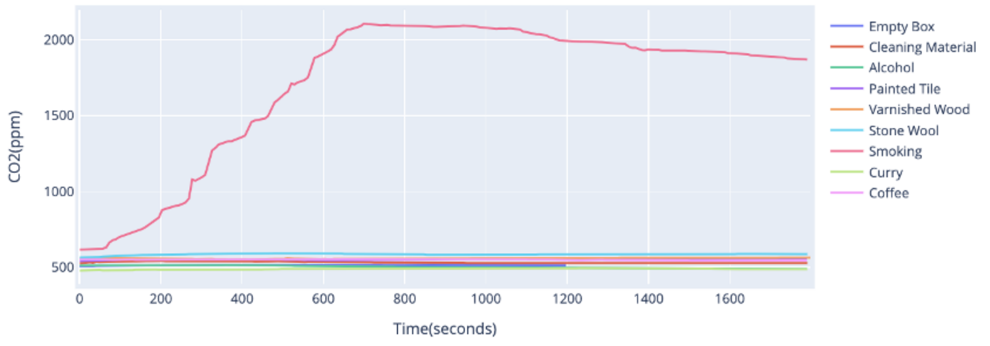

Figure 4.13. CO_2 Change Over Time for the Multi Sensor Array Based Electronic Nose

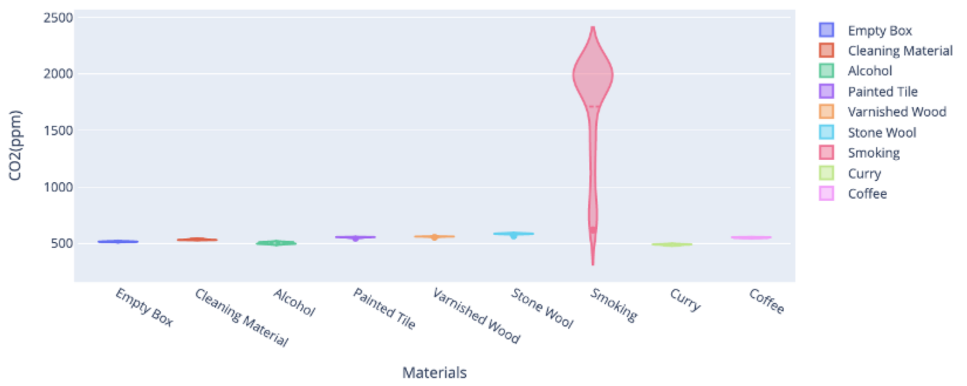

Figure 4.14. CO_2 Change Levels of Different Materials for the Multi Sensor Array Based Electronic Nose.

4.3.2 Transient Variation of Formaldehyde (HCHO)

Formaldehyde is found in resins used to make composite wood products (hardwood plywood, particleboard, and medium-density fiberboard), building materials and insulation, household products such as glues, permanent press fabrics, paints and coatings, lacquers and finishes, and paper products; preservatives used in some medicines, cosmetics, and other consumer products such as dishwashing liquids and fabric softeners; and fertilizers and pesticides. As a result of combustion and some other natural processes, it may also be found in the following: Emissions from unvented, fuel-burning equipment, such as gas stoves or kerosene space heaters; and smoking smoke.

Increased amounts of HCHO were caused by chemicals such as alcohol, cleaning materials, and combustion activities such as smoking. Alcohol has the highest average at 4.7 ppm, followed by smoking at 4.3 ppm. Foods and beverages such as curry and coffee produced a trace amount of HCHO. During the studies, the varnished wood, stone wool, and empty box did not produce any HCHO.

.

Table 4.8. Analysis of Experimental HCHO (ppm) Measurements with the Multi Sensor Array Based Electronic Nose

Material	Mean	Standard Deviation	Min	Max
Empty Box	0	0	0	0
Alcohol	4.7	1.1	0	5
Cleaning Material	3.4	1.7	0.7	5
Painted Tile	0.3	0.3	0	0.9
Varnished Wood	0	0	0	0
Stone Wool	0	0	0	0
Coffee	0.1	0.02	0.06	0.15
Curry	0.3	0.02	0.25	0.33
Smoking	4.3	1.5	0.2	5

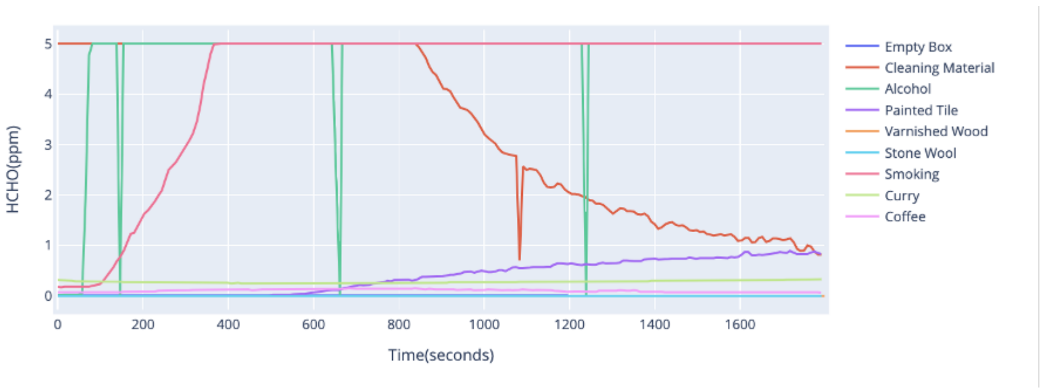

Figure 4.15. HCHO Change Over Time for the Multi Sensor Array Based Electronic Nose

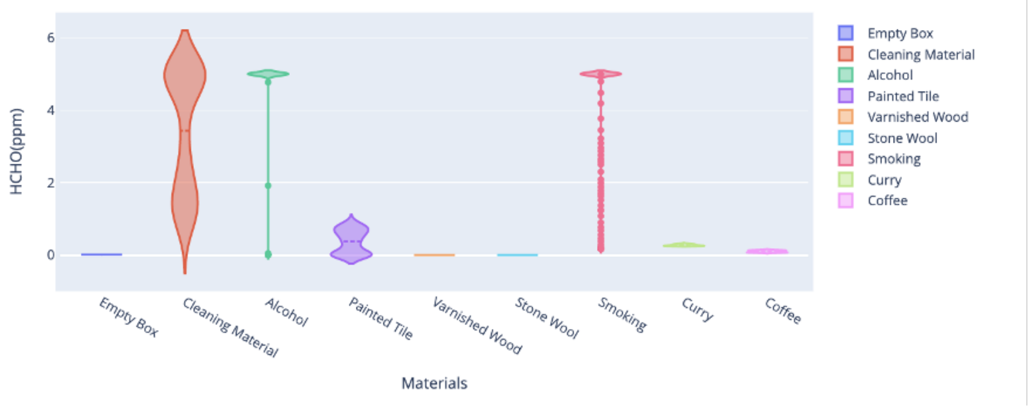

Figure 4.16. HCHO Change Levels of Different Materials for the Multi Sensor Array Based Electronic Nose.

4.3.3 Transient Variation of PM2.5

PM observed inside will comprise particles from the outside that have migrated indoors as well as particles from interior sources. Indoor PM may be created in a variety of ways, including cooking and combustion activities.

Only smoking smoke created a significant quantity of particulate matter throughout the trial, as predicted. Additionally, curry is the second largest source of particulate particles in the experiment.

Table 4.9. Analysis of Experimental PM2.5 (µg/m^3) Measurements with the Multi Sensor Array Based Electronic Nose

Material	Mean	Standard Deviation	Min	Max
Empty Box	10.4	0.6	8.6	12.1
Alcohol	9.9	0.5	8.7	11.1
Cleaning Material	11.0	1.1	9.4	13.4
Painted Tile	11.9	0.3	10.4	13.6
Varnished Wood	8	1.2	5.9	12
Stone Wool	7.2	0.8	5.7	10.4
Coffee	9.2	0.5	8.1	10.8
Curry	48.8	2.7	40	54.5
Tobacco Smoke	54116	13675	283	61389

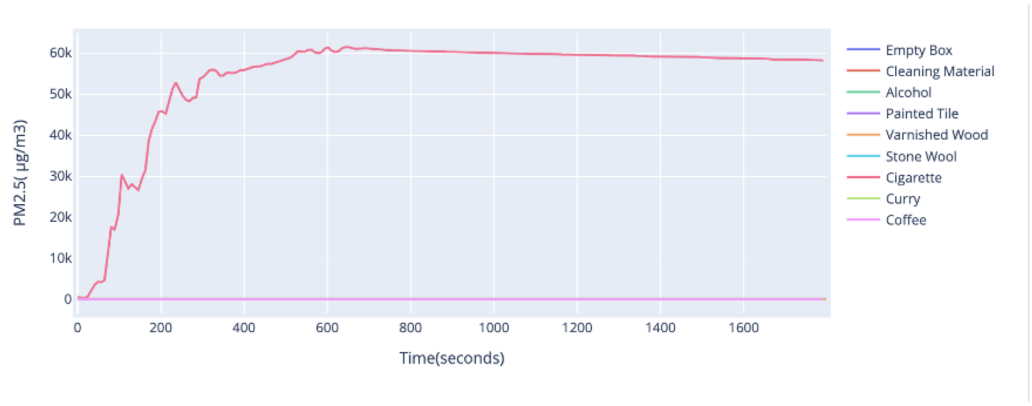

Figure 4.17. PM2.5 Change Over Time for the Multi Sensor Array Based Electronic Nose

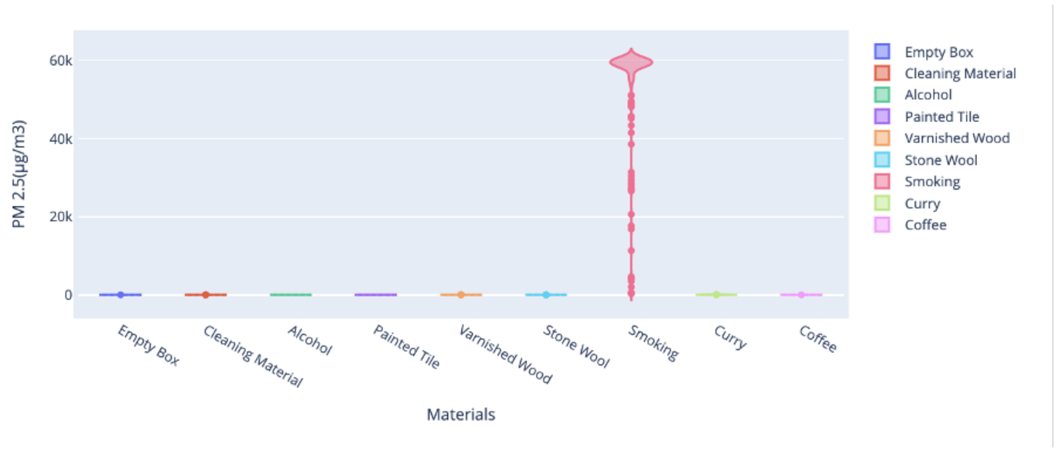

Figure 4.18. PM2.5 Change Levels of Different Materials for the Multi Sensor Array Based Electronic Nose.

4.3.4 Transient Variation of TVOC (Total Volatile Organic Compounds)

Due to the chemical nature of alcohol, it created the highest degree of voc. While cleaning materials, painted tile, and varnished wood all contain VOCs, they are not as volatile as alcohol. As a result, the alcohol generated sparks on the sensor. Smoking creates the second highest level of VOCs, followed by curry.

Table 4.10. Analysis of Experimental TVOC (ppb) Measurements with the Multi Sensor Array Based Electronic Nose

Material	Mean	Standard Deviation	Min	Max
Empty Box	1.3	2.4	0	5
Alcohol	29846.9	16495.5	0	50484
Cleaning Material	61.0	31.1	0	115
Painted Tile	71.9	35.3	0	124
Varnished Wood	93	85.1	0	225
Stone Wool	47.2	20.8	0	80.4
Coffee	4.3	5.9	0	25
Curry	875.4	351.6	194	1327
Tobacco Smoke	14169.8	6135	234	25526

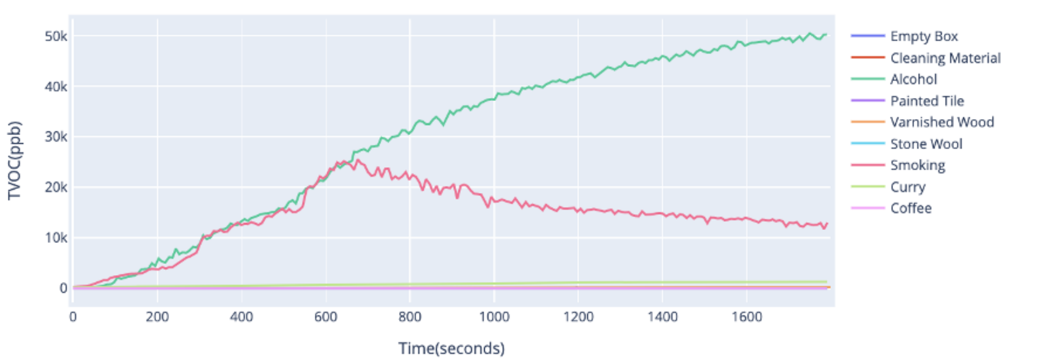

Figure 4.19. TVOC Change Over Time for the Multi Sensor Array Based Electronic Nose

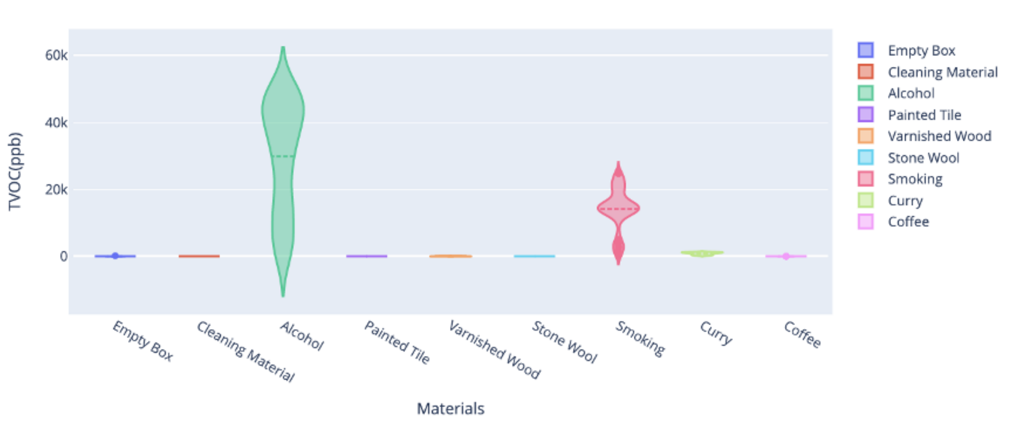

Figure 4.20. TVOC Change Levels of Different Materials for the Multi Sensor Array Based Electronic Nose.

4.3.5 Transient Variation of Other parameters (H_2, Ethanol, Temperature, Humidity)

Increased pollution results in a drop in H_2 and ethanol levels. Smoking and alcohol contain the least ethanol and hydrogen peroxide. Additionally, ethanol and hydrogen had the most varied distributions, and all materials responded differently to these factors, as seen in Figure 4.22.

During the experiment, only coffee substantially elevated the temperature. It raised the ambient temperature from 23 to 25 °C during the experiment. The temperature of the experiment surroundings remained constant throughout the experiment.
As we saw with the electronic nose based on gas resistance, painted tile increased humidity from 20 % to 40 %. Additionally, curry and coffee elevated humidity levels slightly. Figure 4.22 illustrates the impact of all the binary combinations on the material classifications.

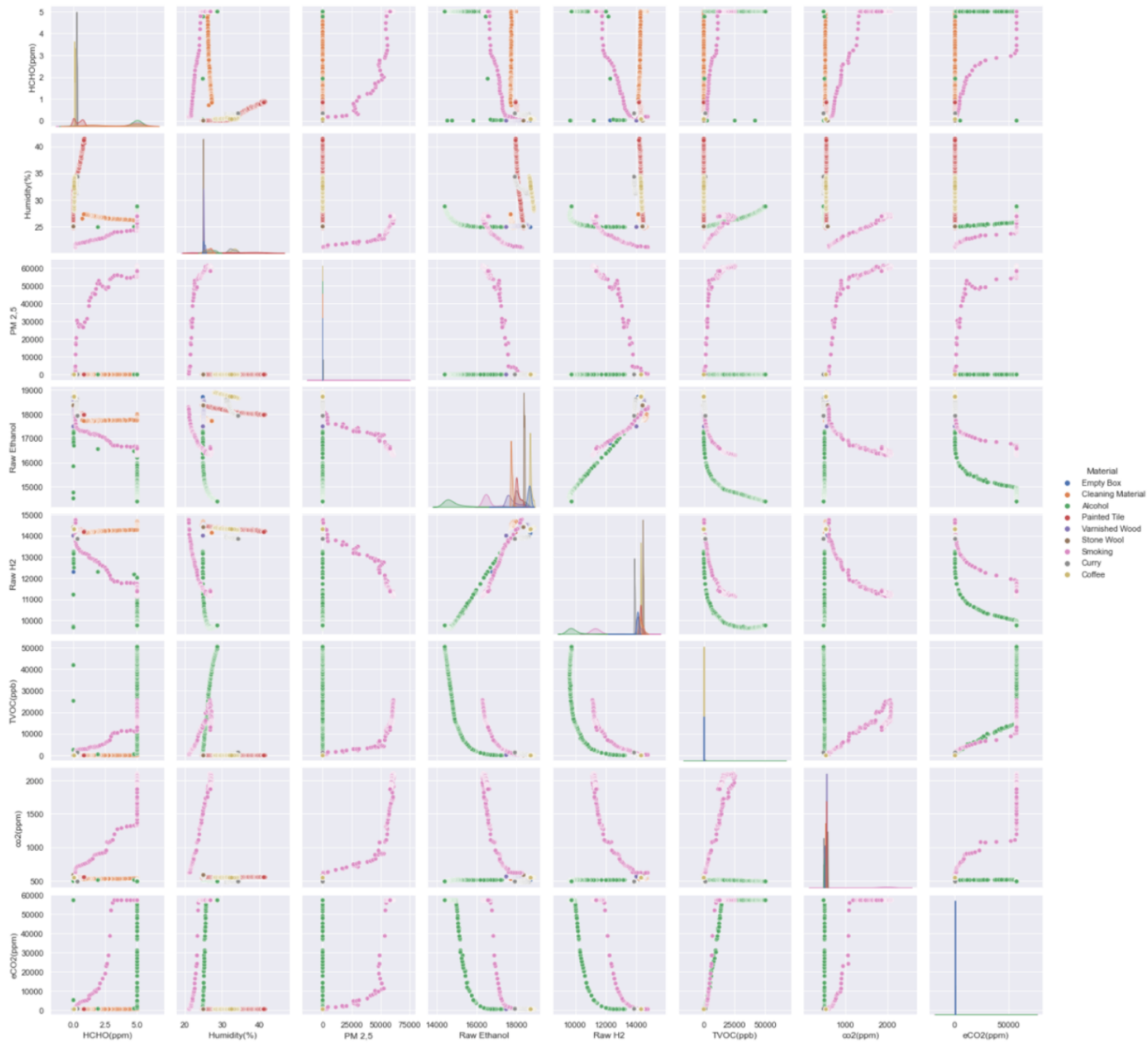

Figure 4.21. All Binary Combinations of Different Parameters on Material Classification.

4.3 Analysis of the Results with Different Classification Models for the Multi Sensor Array Based Electronic Nose

4.4.1 The Naive Bayes (NB) Classification Model

I again started by classifying the data using Naive Bayes modeling. Predictions were made on the test set after the model was developed on the training set. The model's precision is far greater than that of the model developed for the gas resistance-based electronic nose. The model's accuracy for the training and test sets is 0.94 and 0.92, respectively.

Materials such as stone wool or an empty box have lower precision values than other materials because they did not produce any peaks on any parameter. Varnished wood has lower recall than other materials with 0.64 recall score. Only varnished wood is significantly misclassified. 18.9 % of varnished wood samples misclassified as stone wool and 17.5 % of samples misclassified as empty box.

Table 4.11. Classification Report by the NB Classification Model for Experiments with the Multi Sensor Array Based Electronic Nose

Material	**Precision**	**Recall**
Empty Box	0.61	0.98
Alcohol	0.99	0.96
Cleaning Material	1.00	0.97
Painted Tile	0.95	0.85
Varnished Wood	0.98	0.64
Stone Wool	0.83	0.95
Coffee	0.95	0.95
Curry	1.00	1.00
Tobacco Smoke	1.00	1.00

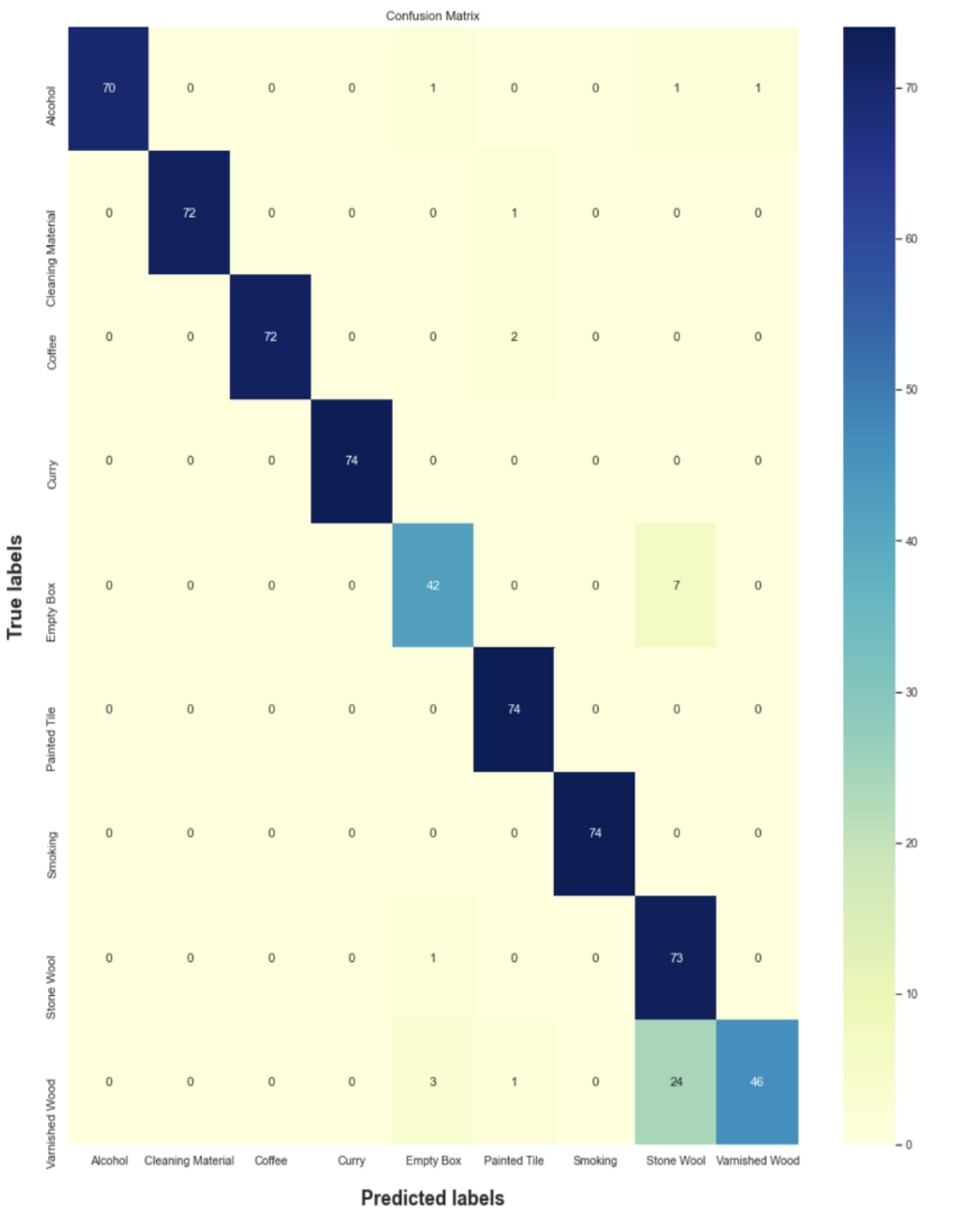

Figure 4.22. Confusion Matrix by the NB Classification Model for Experiments with the Multi Sensor Array Based Electronic Nose.

4.4.2 The kNN Classification Model

The materials were then classified according to their k-nearest neighbors distance using the k-nearest neighbors method. Between two and fifteen neighbors were applied to the dataset. The k=8 model achieves the maximum accuracy, as illustrated in Figure 4.25. The accuracy of a k=8 model is 0.79 on the training set and 0.74 on the test set.

Despite the fact that k=2 has the highest accuracy on the training set, it has a poor accuracy on the test set with 0.71 due to overfitting on the training set.

As can be seen from Table 4.12, kNN has great classification accuracy on chemicals and combustion activities like alcohol, cleaning material and smoking. However, it does not have good performance on classifying building materials like stone wool, varnished wood.

The most common misunderstanding occurs when varnished wood is classified as stone wool. 32% of varnished wood samples were incorrectly identified as stone wool. Except for this classification, the algorithm works nearly perfect. Almost every sample was correctly classified by k=2 kNN model as can be seen in Figure 4.26.

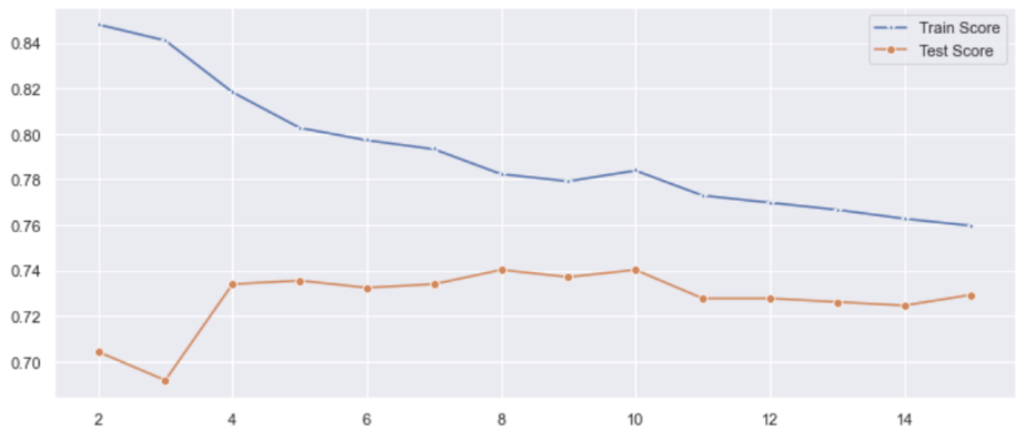

Figure 4.23. The kNN Model with Different Number of Neighbors Accuracy on Training and Test Set for the Multi Sensor Array Based Electronic Nose.

Table 4.12. Classification Report by the kNN Classification Model (k=8) for Experiments with the Multi Sensor Array Based Electronic Nose

Material	Precision	Recall
Empty Box	0.61	0.98
Alcohol	0.99	0.95
Cleaning Material	0.99	0.92
Painted Tile	0.84	0.49
Varnished Wood	0.50	0.58
Stone Wool	0.54	0.58
Coffee	0.66	0.73
Curry	0.71	0.74
Tobacco Smoke	1.00	0.96

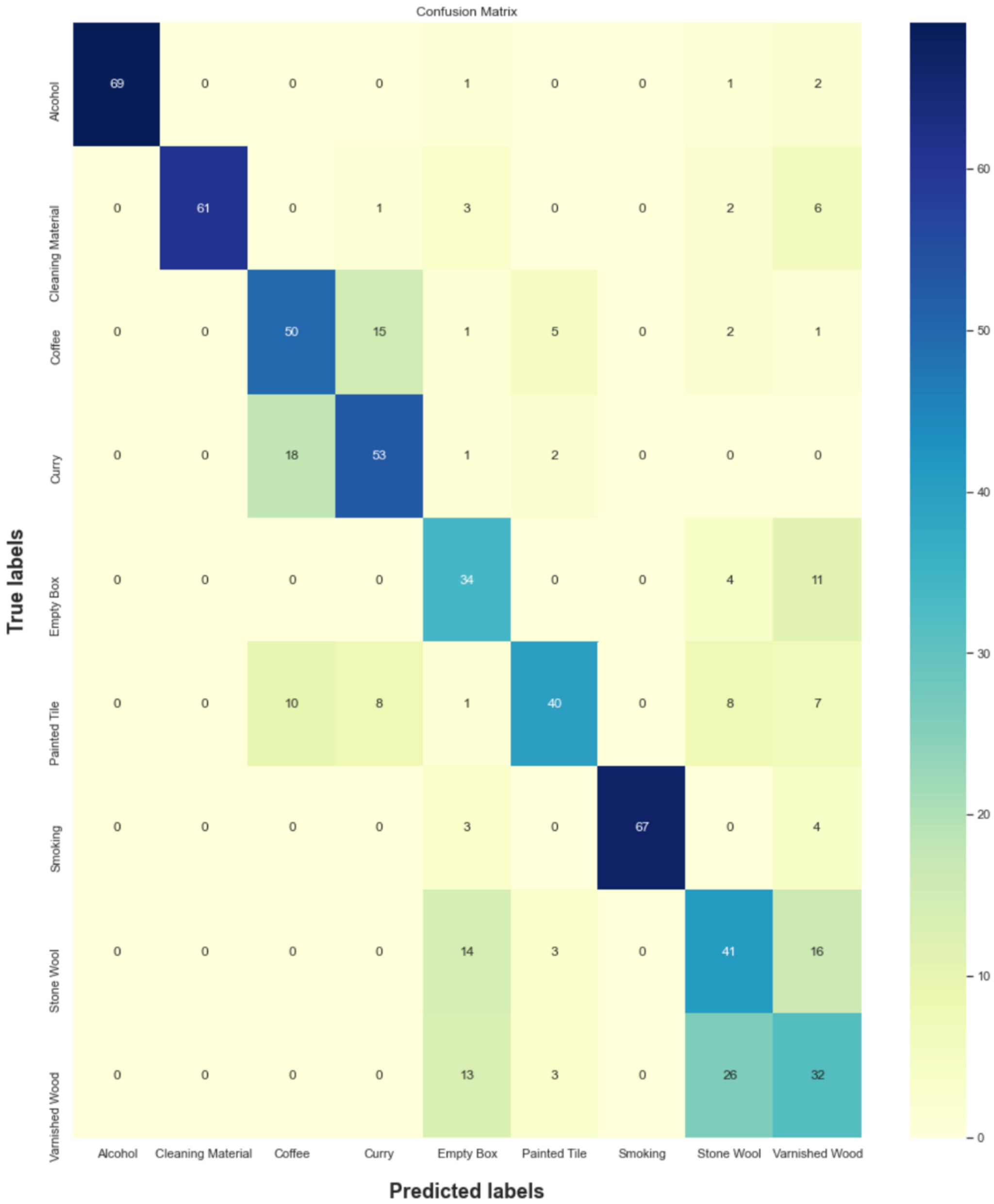

Figure 4.24. Confusion Matrix by the kNN (k=8) Classification Model for Experiments with the Multi Sensor Array Based Electronic Nose.

4.4.3 Random Forest Classification Model

The items are classified using the Random Forest Classifier following the kNN classification. Ten, fifty, and one hundred decision trees were used to develop the mode. Between the three models, there are no apparent differences. As a result, I decided to build the model with ten trees. The classification accuracy of 10, 50, and 100 tree models is 0.95, 0.94, and 0.94, respectively.

The importance of features was determined by computing the information gain associated with each feature using the gini impurity. PM 2.5 scores the highest at 0.23, followed by Formaldehyde at 0.22 and TVOC at 0.13. The parameters with the least significance are Raw Ethanol (0.01), and Temperature (0.02).

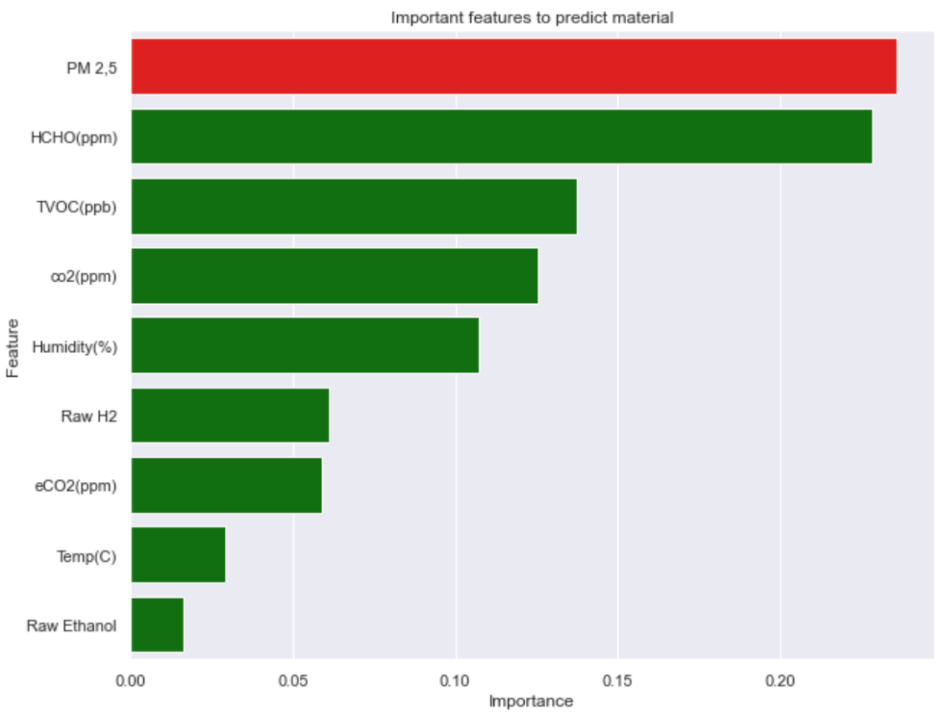

Figure 4.25. Feature Importance Scores for Experiment I RFC Model of Multi Sensor Array Based Electronic Nose

Table 4.13. Classification Report for Experiment I n_tree=10, RFC model of Multi Sensor Array Based Electronic Nose

Material	precision	recall
Empty Box	0.61	0.98
Alcohol	1.00	0.96
Cleaning Material	0.97	0.97
Painted Tile	0.86	0.95
Varnished Wood	0.91	0.80
Stone Wool	0.91	0.97
Coffee	0.99	0.99
Curry	0.99	1.00
Tobacco Smoke	1.00	0.99

The Random Forest algorithm combined with a multisensor array-based electronic nose provides the most accurate results. Except for varnished wood, the algorithm performs perfectly.

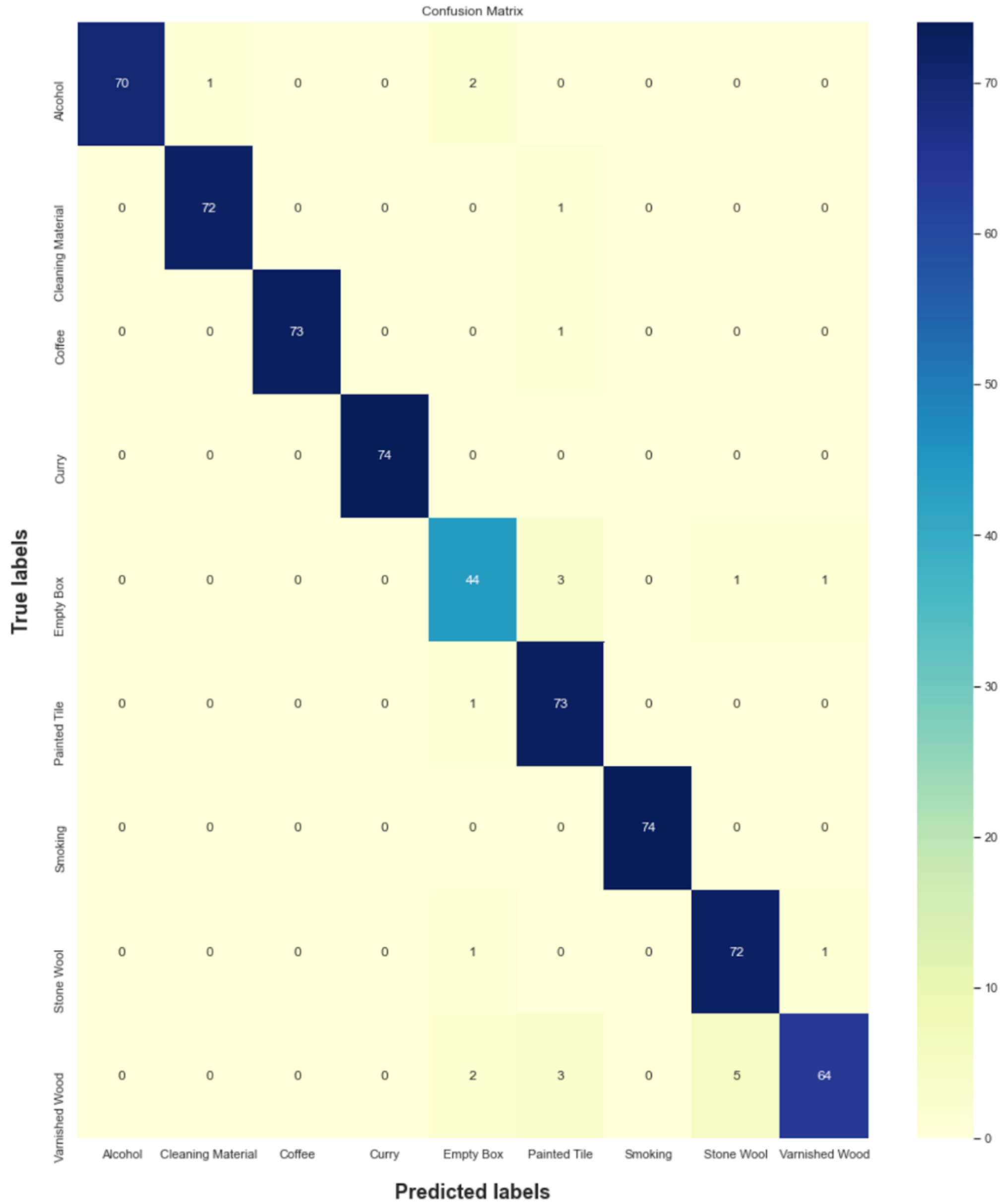

Figure 4.26. Classification Report for Experiment I n_tree=10, RFC model of Multi Sensor Array Based Electronic Nose

CHAPTER 5

CONCLUSION

5.1 Summary of the Research

Cleaning chemicals, construction operations, smoking, perfumes, building materials and outdoor pollutants can all contribute to indoor air pollution. An environmental danger that affects health and well-being is one of the biggest. In extreme circumstances, indoor pollution can be 100 times worse than outdoor pollution, and most people spend 90 % of their time indoors. Skin irritation, nausea, headaches, sick building syndrome, and even cancer have all been linked to poor indoor air quality. Poor air quality can be caused by a wide range of things, including construction materials, poor filtering, chemicals released from office furniture, smoking, food, and beverages. In order to improve interior air quality and, in some circumstances, retrofit the structure, finding the sources of pollution is essential.

Different measures of indoor air quality have varying implications on human health. Temperature has a direct effect on occupant performance, with each degree above 25°C resulting in a 2 % decrease in productivity. Inadequate thermal comfort can also contribute to symptoms of Sick Building Syndrome (SBS), such as headaches, itchy skin, dry or painful eyes, clogged or runny noses, and rashes. Humidity has a direct impact on the health and comfort of occupants, as well as the presence of biological pollutants such as mold spores. There is a strong correlation between increased indoor CO_2 levels and increased symptoms of Sick Building Syndrome (SBS). Additionally, elevated CO_2 levels can result in up to an 11 % drop in productivity, a 23 % impairment in decision making, and a 29 % reduction in information consumption. Short-term exposure to high levels of VOCs has a number of deleterious consequences, including irritation of the eyes and respiratory tract, headaches, dizziness, vision disturbances, and memory impairment. Additionally, there has been a correlation between increased VOC concentrations in indoor air and allergies, asthma, and other respiratory health issues. PM2.5 are dangerous because they can enter the lungs and circulation. Short-term exposure can irritate the lungs, cause coughing, and contribute to cardiovascular problems. Prolonged exposure can result in an early death from heart disease, lung disease, and cancer.

Electronic nose concept has been decided to use in this research to detect pollutant sources in the building. Electronic nose is a sensor device that uses electrical signals to detect scents or flavors. It consists of sensor arrays and pattern recognition systems to replicate human senses. Electronic noses are used in different applications such as outdoor air, indoor air, food, health applications.

Two distinct electronic noses were used in the study. One was built by Bosch and is called the BME688 Breakout Board, while the other was developed specifically for this purpose. On the BME Board, there are eight separate BME688 sensors that operate dependent on the MOX layer's gas resistance. Each sensor is simultaneously heated with a separate heater profile, and their reaction to the gas provides an eight-layer sensitivity dependent on the gas resistance of each sensor. The other electronic nose was built as a result of the BME688 sensor's limited sensitivity to VOC gases, despite the fact that other parameters may generate a vast amount of information as discovered during the literature review.

The experiment was carried out in a box with a volume of 130 liters. On the lid of the box, two holes are opened to allow for some airflow. Nine distinct materials are measured (Office Air, Tobacco Smoke, Cleaning Material, Alcohol-Sanitizer, Curry, Coffee, Painted Tile, Stone Wool, and Varnished Wood) among five categories. Ambient air, chemicals, building materials, food and beverages, and combustion are the categories.

To classify materials based on sensor data, Naive Bayes, K-Nearest Neighbors, and Random Forest Classifier models were applied. Different metrics have been chosen to evaluate the models, including accuracy, precision, and recall.

5.2 Main Results and Discussion

The main objective of this research is to detect indoor pollutant sources by analyzing the sensor data that is collected by electronic noses. 95 % accuracy has been achieved after the evaluations on the hardware development and machine learning modelling development.

- To begin, data from an electronic nose based on gas resistance have been evaluated. Only painted tile, due to its water-based composition, increased humidity levels. No material experienced a change in pressure. Temperature changed throughout the experiment due to the electronic nose's scanning cycle. Each material showed the same temperature pattern. Due to the lack of significant variations in the temperature patterns of various materials, temperature was removed as a parameter while creating models on top of these datasets. Temperature changes in the ambient air during the experiment can cause the model to overfit, leading to misinterpretations of the prediction.

- Then gas resistance values have been analyzed. Gas resistance refers to the electrical resistance of the gas sensor's MOX layer, which varies in response to air ionization. Reduced resistance levels are indicative of polluted air. The empty box, stone wool, and coffee all have higher averages. Additionally, materials such as alcohol and smoking have lower averages than others due to their chemical composition.

- Data has been modelled with Naive Bayes, kNN, Random Forest Classifiers to predict pollutant sources in the box. Naive Bayes has very low accuracy with 0.34. Different k numbers have been tried to find optimum k. k=2 gives the best accuracy on the test set with 0.57. Then Random Forest algorithms have been applied with different numbers of trees. n_tree=10 gives the best results with 0.82 accuracy. Based on the RFC model's feature importance, gas resistance is the most important feature as expected with 0.43 importance score.

- Then, the data that is collected from multi sensor array based electronic nose has been analyzed. CO_2, HCHO, PM2.5, TVOC, Ethanol, H2, Temperature, Humidity and eCO_2 values have been analyzed. Smoking had the highest average CO_2 levels at 1711 ppm. Other materials range between 490 and 585 ppm. Alcohol has the greatest average concentration of HCHO at 4.7 ppm, followed by smoking at 4.3 ppm. HCHO was created in trace amounts by foods and beverages such as curry and coffee. Only tobacco smoke generated a considerable amount of PM2.5 (54000 ug/m^3) during the experiment. Due to the chemical composition of alcohol, TVOC produced the maximum degree of voc at 29000 ppb. Smoking produces the second greatest level of VOCs at 14000 ppb, followed by curry at 875 ppb.

Table 5.1 Material- High Concentration Parameter Comparison

Material	High Concentration Parameter
Alcohol	HCHO, TVOC, Ethanol, H2, Humidity
Cleaning Material	HCHO, Humidity
Painted Tile	HCHO, Humidity
Varnished Wood	TVOC
Stone Wool	-
Coffee	TVOC, Humidity, HCHO
Curry	TVOC, Humidity, HCHO
Tobacco Smoke	HCHO, TVOC, Ethanol, H2, CO2, PM2.5

- Modelling on the data that is collected by a multi sensor array based electronic nose gives much more accurate results than the models with gas resistance. Naive Bayes gives 0.92 accuracy, kNN with k=8 gives 0.74 accuracy and Random Forest Classifier gives 0.95 accuracy. Based on the Random Forest feature importance calculation PM 2.5 scores the highest at 0.23, followed by Formaldehyde at 0.22 and TVOC at 0.13.

- The majority of electronic noses used in this application are based on VOC calculations. However, indoor air pollutants are far more complex than this, and hence utilizing a lot of possible parameters affecting indoor air quality yields the best results when determining the source of a pollutant in a building.

REFERENCES

Abelmann, S. (2013, November 27). German Committee on Indoor Guide Values. Retrieved April 02, 2021, from https://www.umweltbundesamt.de/en/topics/health/commissions-working-groups/german-committee-on-indoor-guide-values#-how-is-indoor-environment-defined

Adamkiewicz, G. (2010). *WHO guidelines for indoor air quality:: Selected pollutants*. Geneva: World Health Organization.

Air pollution. (n.d.). Retrieved December 03, 2020, from https://www.who.int/health-topics/air-pollution

Air Quality Standards. (n.d.). Retrieved April 02, 2021, from https://ec.europa.eu/environment/air/quality/standards.htm

Al-Haija, Q. A., Al-Qadeeb, H., & Al-Lwaimi, A. (2013). Case Study: Monitoring of AIR Quality in King Faisal University Using a Microcontroller and WSN. *Procedia Computer Science, 21*, 517-521.

Allen, R. W., Adar, S. D., Avol, E., Cohen, M., Curl, C. L., Larson, T., . . . Kaufman, J. D. (2012). Modeling the Residential Infiltration of Outdoor PM 2.5 in the Multi-Ethnic Study of Atherosclerosis and Air Pollution (MESA Air). *Environmental Health Perspectives, 120*(6), 824-830.

Altman, N. S. (1992). An Introduction to Kernel and Nearest-Neighbor Nonparametric Regression. *The American Statistician, 46*(3), 175.

Amato, F., Rivas, I., Viana, M., Moreno, T., Bouso, L., Reche, C., . . . Querol, X. (2014). Sources of indoor and outdoor PM2.5 concentrations in primary schools. *Science of The Total Environment, 490*, 757-765.

Amin, N. D., Akasah, Z. A., & Razzaly, W. (2015). Architectural Evaluation of Thermal Comfort: Sick Building Syndrome Symptoms in Engineering Education Laboratories. *Procedia - Social and Behavioral Sciences, 204*, 19-28.

Andersen, I. (1972). Relationships between outdoor and indoor air pollution. *Atmospheric Environment (1967), 6*(4), 275-278.

Apr 01, 2. (n.d.). Carbon Dioxide Detection and Indoor Air Quality Control. Retrieved June 09, 2020, from https://ohsonline.com/Articles/2016/04/01/Carbon-Dioxide-Detection-and-Indoor-Air-Quality-Control.aspx?Page=2

Awbi, H. B. (2017). Ventilation for Good Indoor Air Quality and Energy Efficiency. *Energy Procedia, 112*, 277-286.

Badura, M., Batog, P., Drzeniecka-Osiadacz, A., & Modzel, P. (2018). Optical particulate matter sensors in PM2.5 measurements in atmospheric air. *E3S Web of Conferences, 44*, 00006.

Becher, R., Øvrevik, J., Schwarze, P., Nilsen, S., Hongslo, J., & Bakke, J. (2018). Do Carpets Impair Indoor Air Quality and Cause Adverse Health Outcomes: A Review. *International Journal of Environmental Research and Public Health, 15*(2), 184.

Berens, A. R. (1985). Prediction of Organic Chemical Permeation Through PVC Pipe. *Journal - American Water Works Association, 77*(11), 57-64.

Bipat, C. (n.d.). Energy Efficiency and Indoor Air Quality: Friends or Foes? Retrieved January 23, 2021, from https://www.ny-engineers.com/blog/energy-efficiency-and-indoor-air-quality

BME688 software. (n.d.). Retrieved January 09, 2022, from https://www.bosch-sensortec.com/software-tools/software/bme688-software/

Brunekreef, B., & Holgate, S. T. (2002). Air pollution and health. *The Lancet, 360*(9341), 1233-1242.

Caballero, B., Trugo, L. C., & Finglas, P. M. (2003). *Encyclopedia of food sciences and nutrition*. San Diego, CA: Academic.

Chojer, H., Branco, P., Martins, F., Alvim-Ferraz, M., & Sousa, S. (2020). Development of low-cost indoor air quality monitoring devices: Recent advancements. *Science of The Total Environment, 727*, 138385.

Chugh, R. S., Bhatia, V., Khanna, K., & Bhatia, V. (2020). A comparative analysis of classifiers for Image Classification. *2020 10th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*. doi:10.1109/confluence47617.2020.9058042

Dabill, D., Lamont, D., & Groves, J. (1996). The effect of pressure on portable gas monitoring equipment during compressed air tunnelling. *The Annals of Occupational Hygiene*. doi:10.1093/annhyg/40.1.11

Department of Labor logo UNITED STATESDEPARTMENT OF LABOR. (n.d.). Retrieved April 02, 2021, from https://www.osha.gov/annotated-pels/table-z-1

Dey, A. (2018). Semiconductor metal oxide gas sensors: A review. *Materials Science and Engineering: B, 229*, 206-217.

Dragonieri, S., Pennazza, G., Carratu, P., & Resta, O. (2017). Electronic Nose Technology in Respiratory Diseases. *Lung, 195*(2), 157-165.

Electrochemical Gas Sensors. (n.d.). Retrieved June 12, 2020, from https://www.membrapor.ch/electrochemical-gas-sensors/

Electronic, S. (n.d.). SCD30 is more than just the NDIR CO2 sensor. Retrieved December 13, 2021, from https://www.soselectronic.com/articles/sensirion/scd30-is-more-than-just-the-ndir-co2-sensor-2152

ElProCus, J. (2019, August 01). Arduino Mega 2560 board: Specifications, and Pin Configuration. Retrieved December 13, 2021, from https://www.elprocus.com/arduino-mega-2560-board/

Environment Protection Authority Victoria. (2019, November 13). PM10 particles in the air. Retrieved June 10, 2020, from https://www.epa.vic.gov.au/for-community/environmental-information/air-quality/pm10-particles-in-the-air

Fanger. (1972). Thermal comfort: Analysis and applications in environmental engineering. *Applied Ergonomics, 3*(3), 181.

Firebase realtime database | firebase documentation. (n.d.). Retrieved December 17, 2021, from https://firebase.google.com/docs/database

FLEDER, R. M. (2020). *ELEMENTARY PRINCIPLES OF CHEMICAL PROCESSES*. S.l.: JOHN WILEY.

Flores, R. M. (2014). Coalification, Gasification, and Gas Storage. *Coal and Coalbed Gas,* 167-233.

Formaldehyde's Impact on Indoor Air Quality. (2016, September 06). Retrieved June 10, 2020, from https://www.epa.gov/indoor-air-quality-iaq/formaldehydes-impact-indoor-air-quality

The Furniture We Breathe: Home Furnishing & Indoor Air Quality. (n.d.). Retrieved June 08, 2020, from https://thesunriseguide.com/the-furniture-we-breathe-home-furnishing-indoor-air-quality/

F_129. (n.d.). Be aware of air conditioning pollution. Retrieved June 08, 2020, from http://en.people.cn/90782/7909500.html

Gebicki, J. (2016). Application of electrochemical sensors and sensor matrixes for measurement of odorous chemical compounds. *TrAC Trends in Analytical Chemistry, 77*, 1-13.

Gravity: Formaldehyde (HCHO) sensor. (n.d.). Retrieved December 21, 2021, from https://www.dfrobot.com/product-1574.html

Hamanaka, R. B., & Mutlu, G. M. (2018). Particulate Matter Air Pollution: Effects on the Cardiovascular System. *Frontiers in Endocrinology, 9*.

He, J., Xu, L., Wang, P., & Wang, Q. (2017). A high precise E-nose for daily indoor air quality monitoring in living environment. *Integration, 58*, 286-294.

Herrero, J. L., Lozano, J., Santos, J. P., & Suárez, J. I. (2016). On-line classification of pollutants in water using wireless portable electronic noses. *Chemosphere, 152*, 107-116.

The Hidden Hazards Of Indoor Air Pollution From Ozone. (2018, February 05). Retrieved June 10, 2020, from https://cleanair.camfil.us/2018/02/05/the-hidden-hazards-of-indoor-air-pollution-from-ozone/

Hodgson, A. T., Beal, D., & Mcilvaine, J. E. (2002). Sources of formaldehyde, other aldehydes and terpenes in a new manufactured house. *Indoor Air, 12*(4), 235-242.

Home. (n.d.). Retrieved April 02, 2021, from https://www.who.int/

How humidity may affect COVID-19 outcome. (n.d.). Retrieved June 10, 2020, from https://www.medicalnewstoday.com/articles/how-humidity-may-affect-covid-19-outcome

How humidity may affect COVID-19 outcome. (n.d.). Retrieved June 22, 2020, from https://www.medicalnewstoday.com/articles/how-humidity-may-affect-covid-19-outcome

Huzienga, C., Abbaszadeh, S., Zagreus, L., & Arens, E. (2006). Air Quality and Thermal Comfort in Office Buildings: Results of a Large Indoor Environmental Quality Survey.

Iea. (n.d.). Buildings – Topics. Retrieved January 23, 2021, from https://www.iea.org/topics/buildings

Jacobson, T. A., Kler, J. S., Hernke, M. T., Braun, R. K., Meyer, K. C., & Funk, W. E. (2019). Direct human health risks of increased atmospheric carbon dioxide. *Nature Sustainability, 2*(8), 691-701.

Jasinski, G., Wozniak, L., Kalinowski, P., & Jasinski, P. (2018). Evaluation of the Electronic Nose Used for Monitoring Environmental Pollution. *2018 XV International Scientific Conference on Optoelectronic and Electronic Sensors (COE)*.

Jones, A. (1999). Indoor air quality and health. *Atmospheric Environment, 33*(28), 4535-4564.

Karakaya, D., Ulucan, O., & Turkan, M. (2019). Electronic Nose and Its Applications: A Survey. *International Journal of Automation and Computing, 17*(2), 179-209.

Kim, H., Konnanath, B., Sattigeri, P., Wang, J., Mulchandani, A., Myung, N., . . . Bakkaloglu, B. (2011). Electronic-nose for detecting environmental pollutants: Signal processing and analog front-end design. *Analog Integrated Circuits and Signal Processing, 70*(1), 15-32.

Kort, S., Brusse-Keizer, M., Gerritsen, J., & Van der Palen, J. (2017). Data analysis of electronic nose technology in lung cancer: Generating prediction models by means of Aethena. *Journal of Breath Research, 11*(2), 026006.

Lafond, A., & Lafond, A. (2018, September 26). Volatile Organic Compounds (VOCs) Exposure in Building Materials. Retrieved March 19, 2021, from https://foobot.io/guides/volatile-organic-compounds-in-building-materials.php

Lee, S., Lam, S., & Fai, H. K. (2001). Characterization of VOCs, ozone, and PM10 emissions from office equipment in an environmental chamber. *Building and Environment, 36*(7), 837-842.

Lin, Y., Yu, H., Wan, F., & Xu, T. (2017). Research on classification of Chinese text data based on SVM. *IOP Conference Series: Materials Science and Engineering, 231*, 012067. doi:10.1088/1757-899x/231/1/012067

Lioy, P. J., Avdenko, M., Harkov, R., Atherholt, T., & Daisey, J. M. (1985). A Pilot Indoor-Outdoor Study of Organic Particulate Matter and Particulate Mutagenicity. *Journal of the Air Pollution Control Association, 35*(6), 653-657.

Loftness, V., Hakkinen, B., Adan, O., & Nevalainen, A. (2007). Elements That Contribute to Healthy Building Design. *Environmental Health Perspectives, 115*(6), 965-970.

Mendell, M. J., & Heath, G. A. (2005). Do indoor pollutants and thermal conditions in schools influence student performance? A critical review of the literature. *Indoor Air, 15*(1), 27-52.

NAAQS Table. (2021, February 10). Retrieved April 02, 2021, from https://www.epa.gov/criteria-air-pollutants/naaqs-table#1

Negri, R., & Reich, S. (2001). Identification of pollutant gases and its concentrations with a multisensor array. *Sensors and Actuators B: Chemical, 75*(3), 172-178.

Operating principle. (n.d.). Retrieved January 09, 2022, from https://www.figaro.co.jp/en/technicalinfo/principle/mos-type.html

Pagliano, L. (2016, May 26). Directive 2010/31/EU on the energy performance of buildings (recast) - 19 May 2010. Retrieved January 23, 2021, from https://www.buildup.eu/en/practices/publications/directive-201031eu-energy-performance-buildings-recast-19-may-2010#:~:text=On%2019%20May%202010%2C%20a,the%202002%20Directive%20it%20replaces.

Park, S. Y., Kim, Y., Kim, T., Eom, T. H., Kim, S. Y., & Jang, H. W. (2019). Chemoresistive materials for electronic nose: Progress, perspectives, and challenges. *InfoMat, 1*(3), 289-316.

Saad, S., Andrew, A., Shakaff, A., Saad, A., Kamarudin, A., & Zakaria, A. (2015). Classifying Sources Influencing Indoor Air Quality (IAQ) Using Artificial Neural Network (ANN). *Sensors, 15*(5), 11665-11684.

Saini, J., Dutta, M., & Marques, G. (2020). A comprehensive review on indoor air quality monitoring systems for enhanced public health. *Sustainable Environment Research, 30*(1).

SakilAnsariCheck out this Author's contributed articles., SakilAnsari, & Check out this Author's contributed articles. (2018, May 04). Pattern Recognition: Introduction. Retrieved June 12, 2020, from https://www.geeksforgeeks.org/pattern-recognition-introduction/

Salonen, H., Salthammer, T., & Morawska, L. (2018). Human exposure to ozone in school and office indoor environments. *Environment International, 119*, 503-514.

SAM Lab. (n.d.). Retrieved June 12, 2020, from http://people.ee.duke.edu/~jk/

Samoli, E., Analitis, A., Touloumi, G., Schwartz, J., Anderson, H. R., Sunyer, J., . . . Katsouyanni, K. (2005). Estimating the Exposure–Response Relationships between Particulate Matter and Mortality within the APHEA Multicity Project. *Environmental Health Perspectives, 113*(1), 88-95.

Sankaran, S., Khot, L. R., & Panigrahi, S. (2012). Biology and applications of olfactory sensing system: A review. *Sensors and Actuators B: Chemical, 171-172*, 1-17.

Sekine, Y., Katori, R., Tsuda, Y., & Kitahara, T. (2016). Colorimetric monitoring of formaldehyde in indoor environment using built-in camera on mobile phone. *Environmental Technology, 37*(13), 1647-1655.

Shooshtari, M., Salehi, A., & Vollebregt, S. (2021). Effect of humidity on gas sensing performance of carbon nanotube gas sensors operated at room temperature. *IEEE Sensors Journal, 21*(5), 5763-5770. doi:10.1109/jsen.2020.3038647

Standard 55 – Thermal Environmental Conditions for Human Occupancy. (n.d.). Retrieved April 02, 2021, from https://www.ashrae.org/technical-resources/bookstore/standard-55-thermal-environmental-conditions-for-human-occupancy

Technical Overview of Volatile Organic Compounds. (2017, April 12). Retrieved June 09, 2020, from https://www.epa.gov/indoor-air-quality-iaq/technical-overview-volatile-organic-compounds

Thu, M. Y., Htun, W., Aung, Y. L., Shwe, P. E., & Tun, N. M. (2018). Smart Air Quality Monitoring System with LoRaWAN. *2018 IEEE International Conference on Internet of Things and Intelligence System (IOTAIS)*.

Tozlu, B. H., Şimşek, C., Aydemir, O., & Karavelioglu, Y. (2021). A High performance electronic nose system for the recognition of myocardial infarction and coronary artery diseases. *Biomedical Signal Processing and Control, 64*, 102247.

Unesda, Ministers, N., & Week, E. (n.d.). Household air pollution, the forgotten health hazard. Retrieved June 22, 2020, from https://euobserver.com/health/139592

V2.wellcertified.com. (n.d.). Retrieved April 02, 2021, from https://v2.wellcertified.com/v2.2/en/air/feature/1#

Vito, S. D., Fattoruso, G., Liguoro, R., Oliviero, A., Massera, E., Sansone, C., . . . Francia, G. D. (2011). Cooperative 3D Air Quality Assessment with Wireless Chemical Sensing Networks. *Procedia Engineering, 25*, 84-87.

VOC sensor SGP30 / SGPC3 (NRND). (n.d.). Retrieved January 09, 2022, from https://www.sensirion.com/en/environmental-sensors/gas-sensors/sgp30/

VOCs from Indoor Chemical Reactions and Health. (n.d.). Retrieved June 08, 2020, from https://iaqscience.lbl.gov/voc-indoor

Volatile Organic Compounds. (n.d.). Retrieved June 10, 2020, from https://www.lung.org/clean-air/at-home/indoor-air-pollutants/volatile-organic-compounds

Weschler, C. J. (2009). Changes in indoor pollutants since the 1950s. *Atmospheric Environment, 43*(1), 153-169.

Wilkens, W. F., & Hartman, J. D. (1964). AN ELECTRONIC ANALOG FOR THE OLFACTORY PROCESSES? *Annals of the New York Academy of Sciences, 116*(2 Recent Advanc), 608-612.

The World Air Quality Index project. (n.d.). Why is PM2.5 often higher than PM10? Is PM10 still a relevant measure? Retrieved March 19, 2021, from https://aqicn.org/faq/2013-02-02/why-is-pm25-often-higher-than-pm10/

Wp-Admin. (n.d.). The Truth About IoT Implementations - Wireless vs. Wired. Retrieved June 10, 2020, from https://blog.senseware.co/2017/10/10/iot-implementations-wireless-vs-wired

Yang, X., & Chen, Q. (1998). Prediction of short-term and long-term VOC emissions from SBR bitumen-backed carpet under different temperatures.

Zampolli, S., Elmi, I., Ahmed, F., Passini, M., Cardinali, G., Nicoletti, S., & Dori, L. (2004). An electronic nose based on solid state sensor arrays for low-cost indoor air quality monitoring applications. *Sensors and Actuators B: Chemical, 101*(1-2), 39-46.

Zuo, B. (n.d.). Grove - VOC and eCO2 gas sensor(sgp30). Retrieved December 23, 2021, from https://wiki.seeedstudio.com/Grove-VOC_and_eCO2_Gas_Sensor-SGP30/