Network Based NLP Approach to Reveal Disease Comorbidities. COVID Case Study

Emre Taylan Duman emretaylanduman@gmail.com Turkey Gebze Technical University

Pinar Pir pinarpir@gtu.edu.tr Turkey Gebze Technical University

SARS-CoV-2 virus, which caused the pandemic COVID-19 , infects the host cells via ACE2 receptor protein binding to spike proteins. Viral binding also alters the vascular effects of the ACE/ACE2 ratio, which results in a number of complications. These complications may have detrimental consequences on patients with chronic diseases such as diabetes, heart disease, tension-related disorders, and kidney disease, which already have an unbalanced ratio of ACE/ACE2. Due to this mechanism, COVID-19 is extremely comorbid, and there may be other comorbidities with yet unidentified etiologies.

In this work, possible comorbidities of COVID-19 examined by a network-based representation of miRNA-gene-disease interactions to reveal possibly similarly altered mechanisms in between diseases. The original algorithm was implemented as the mpDisNet package which is a word-embedding based deep learning technique. Algorithm itself was modified and optimized to increase True-Positive discovery rate by removal of highly frequent diseases such as cancer and expanding the test data by manual curation of comorbid diseases. Original approach was also modified by integration of miRNA expression profiles from SARS-CoV-2 infected cell line and transcription factors regulated by these miRNAs from related databases.

The main aim this work is discovery of possible unknown comorbidities by connecting the diseases by their miRNA mediated regulatory interactions. The algorithm can predict the majority of COVID-19's known comorbidities reported in CDC website with high scores; (Diabetes (0,996), Heart Diseases (0,989), Schizophrenia (0,994), and Hypertension (0,994)), as well as total of 131 diseases that have yet to be discovered to be comorbid with COVID-19. These potential comorbid diseases should be investigated further in order to raise COVID-19 awareness and prevention. Our approach can be used for informing the comorbidity research for the next possible pandemic outbreak.