CHAPTER 5

DISCUSSION AND FUTURE WORK

5.1 Summary

This research aims to identify an effect of air pollution on people's health in Indonesian cities using descriptive epidemiology and Predictive analysis. The analysis adopts both primary and secondary research data focused on air quality and the health consequences. The exploratory data analysis (EDA) indicated structure and relations where the impact of pollutants like PM2.5 and PM10, as well as their impact on respiratory and cardiovascular diseases. Exploratory, graphical and summary description methods gave the first view of cyclicity and spatial distribution of pollutants and health impact. The machine learning models that were implemented in this study were shown to have the possibility of estimating subsequent levels of air pollution and the related health risks, with the preliminary findings showing promising potential.

The descriptive analysis provided insights on areas with high pollution levels and the related health effects to enhance the understanding of the current state. The predictive analysis and machine learning allowed for predicting the future events regarding the pollution levels and the related health concerns which could help in managing and preventing the potential adverse effects. A statistical analysis was conducted to establish the validity and reliability of the results obtained. In the case of this project, the use of Microsoft Power BI together with Python was helpful in managing the various data and ensuring a deep analysis of the complex dataset was achieved.

5.2 Future Works

Therefore, the future research should extend the sample to more detailed and fine grain data covering the other areas in Indonesia's cities to increase the reliability and the external validity of the findings. Using current air quality measurements and health data, the models will be able to provide better and timelier estimates for public health interventions. Improvement of the level of differentiation of the results by applying more fine-grained machine learning algorithms, such as deep learning and with the help of ensemble methods, and the refinement of the methods applied for hyperparameters setting and features selection will enhance the predictive accuracy. Moreover, including socio-economic and demographic variables in the models will help to have a better view of how effectively air pollution will affect all the groups of the population.

The actual development of easy-to-use tools and dashboards derived from the predictive models can help decision-makers in the organizations to make immediate decisions on resource usage. Specific health promotion interventions and policies related to pollution and health risks can be developed based on the findings of this research. Thus, this study provides the necessary and sufficient information on the consequences of air pollution on the health of the people of Indonesia's urban areas and the necessity to focus on corresponding actions and measures. Further research should extend the present analysis in regard to scope and depth and create collaborative partnerships in order to achieve pertinent public health outcomes.