ABSTRACT

Air quality problems have become a global problem due to an important impact on human and environmental health. The goal of this research is to improve the accuracy of the air quality prediction using the method of learning with the machine. How to select the features that are used to receive data are used to identify the most involved parameters such as carbon monoxide (CO), sulfur dioxide (SO2), nitrogen dioxide (NO2), the temperature is similar to the industrial area and measurement. Efficiency, such as AUC and MCC K-Nearest, two, the main algorithm called neighbors (KNN) is used to show that the decision-making plan is the best results with AUC 0.9380 and CA at 0.8820 while KNN also. There is significant improvement after selecting this method of choosing this method not only But improving the analysis of efficiency data but still has more accurate predictions This research is a trend for further development, including the application of all the methods or the combination of additional parameters to solve more complex challenges in the future.

Keywords: Air Quality Prediction, Feature Selection, Information Gain, K-Nearest Neighbors, Decision Tree