ABSTRACT

Air is very important for the sustainability of living things, air allows living things to do their activities properly. However, with the increase in air pollution over time due to the large growth in the industrial sector and the number of people owning motorized vehicles. In 2019 Indonesia was included in the worst air pollution point that had reached the red point which indicated the unhealthy air in DKI Jakarta and the deterioration of the air made DKI Jakarta Province the 5th worst air pollution in the world IQair. To find out and monitor and detect air quality information, classification can be done. Classification is used because it can monitor air quality information based on ISPU data processing that already has a target label. Classification is carried out using the DKI Jakarta Province air pollution ISPU dataset from 2019 to 2022. This research will classify ISPU data using the K-Nearest Neighbor algorithm. By using 5 attributes PM10, SO2, NO, O3, CO2, and categories as target labels in this study. The results of the study show that the KNN algorithm gets the highest accuracy in the initial test with a ratio of 80:20 with neighborhood K = 5 with an accuracy value of 90.98%. The second test with hyperparameter tuning resulted in the highest accuracy at a ratio of 80:20 with k = 7 neighborhood with a combination of parameters weight "distance", p "1" of 91.37%, precision 82.87%, recall 85.22%, and f1-score 84.03%. and validation of the algorithm using K-Fold Cross Validation with the number of folds 10 resulted in an average of 89.43%.

Keywords - ***Air, DKI Jakarta, ISPU, Classification, KNN***