

ABSTRACT

USED CAR CLUSTERING IN BANDUNG CITY USING K- PROTOTYPES ALGORITHM

By

FIE ALFAIN NURIL HAQUE

1202184172

Since the case of the Coronavirus Disease 2019 pandemic or commonly referred to as Covid-19, the use of public transportation has slowly begun to become an option as transportation in an effort to reduce the spread of the corona virus cluster, therefore some people prefer to buy private vehicles. However, due to the increasing price of cars, some people prefer to buy used cars. On the used car buying and selling platform, OLX Autos Indonesia, the demand for used cars increased by 15% to 20%. Therefore, this study was conducted to determine the characteristics of the cluster formed from the used car sales dataset taken from AtapData (atapdata.ai). AtapData is an open data site in Indonesia that can be used for research related to Data Science. This cluster model was created using the K-Prototypes algorithm, Silhouette Score and Davies Bouldin Index to evaluate the resulting cluster results. This clustering model will produce three clusters. The results of the three clusters will have one thing in common, namely brands that dominate sales, including Toyota, Honda, Daihatsu, Nissan, and Mitsubishi. Clustering evaluation using the Silhouette Score method produces a value of 0.7744140503593034. And for the evaluation of the Davies-Bouldin Index it produces a value of 0.4999221950856398.

Keywords: *Clustering, machine learning, K-Prototypes, Silhouette Score, Davies-Bouldin Index*