

CHAPTER 3

RESEARCH METHODOLOGY

3.1. Gething Data from Kaggle

The car price data set was obtained from Kaggle in a 2.2 MB CSV format. This data has 19237 rows and 18 columns consisting of id, manufacture, year of production, model, category, car price, etc. This data is drawn from five manufacturers and production years ranging from 1939 to 2020.

3.2. Proccesing Data

The orange data mining application with the Neural Network and K-Nearest Neighbor algorithms is used to process this data. In the orange application, first select the data that we will use with the CSV file import. After that in the transform section there are two steps, namely selecting columns and sample data, in the select columns step is to determine the columns that will be parameters when making predictions later. Here the target is the price column, and in the data sampler step, data is collected in the range of 70%, 80%, 90% of the original data which will later become training data. Next is modeling which consists of two algorithms, namely the Neural Network and K-Nearest Neighbor algorithms. In the Evaluate + report process we can see the results of the predictions of the two algorithms above and can conclude which algorithm has the smaller MSE number.

3.3. Modeling

In this Project the modeling process was carried out to obtain training data, with that we can make predictions on the dataset. This study uses two different modeling namely Neural Network and K-Nearest Neighbor algorithms. Before getting the training data, we have to do the transform-data sampler process. We will train and test the data to get the final prediction results.

3.4. Result Report

The processed data will enter the evaluation step and report the results. In this step, a prediction will be made to be able to see the comparative value between the two algorithm models. We can make predictions by comparing the training data with the original data in the predictions section of the orange data mining application. The model with the lowest MSE value is the most effective at predicting car prices in this step.

