CHAPTER 5
IMPLEMENTATION AND RESULTS

5.1. Implementation

This project is implemented with Orange. This section will explain how to use the application.

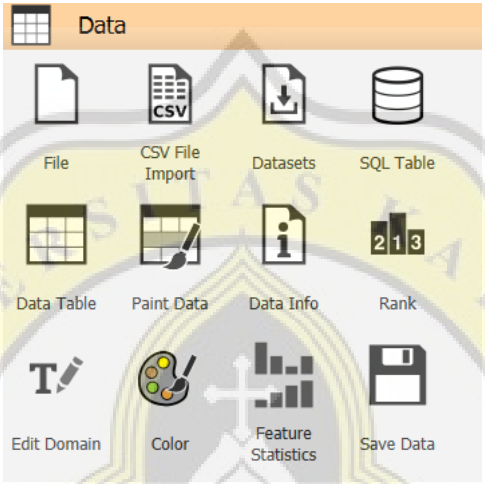

Figure 5.1 Data Type

The initial stage is to select the type of data used. In this research, the dataset obtained from kaggle.com is in the form of CSV and has a regression model. The next step is to import the CSV file.

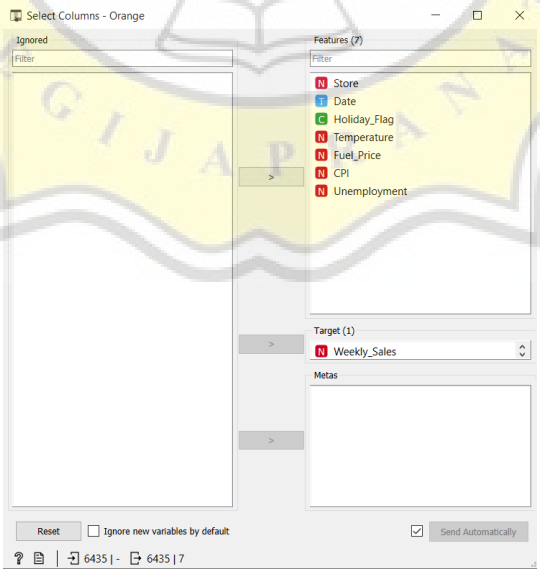

Figure 5.2 Target Data

In the picture above, there are 8 attributes. One of the eight attributes is selected as the target data. The target data used is Weekly Sales. The target is used as a reference in prediction.

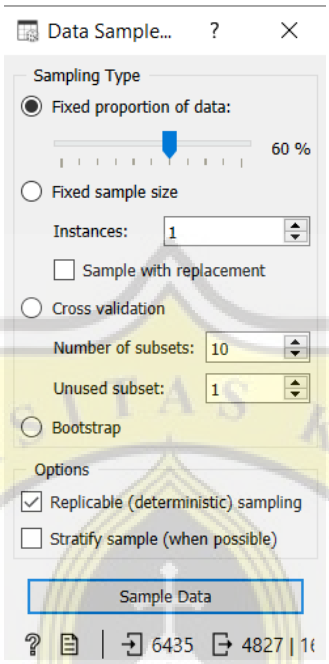

Figure 5.3 Data Sampler

The next process is to determine the proportion of data. In this study, the proportion of data used was 60%, 70%, and 80%, this data is called Data Training. The rest of the data is called Data Testing and is processed without using the algorithm.

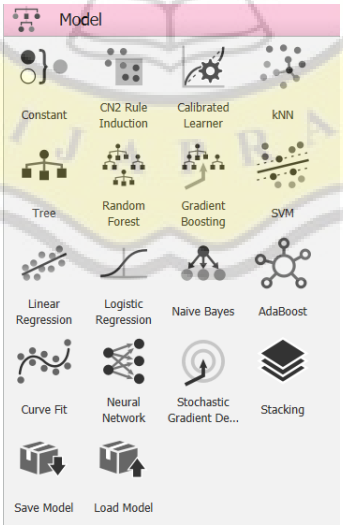

Figure 5.4 Model

In this study, the models used are Random Forest and Tree because these two models are the best of the other models.

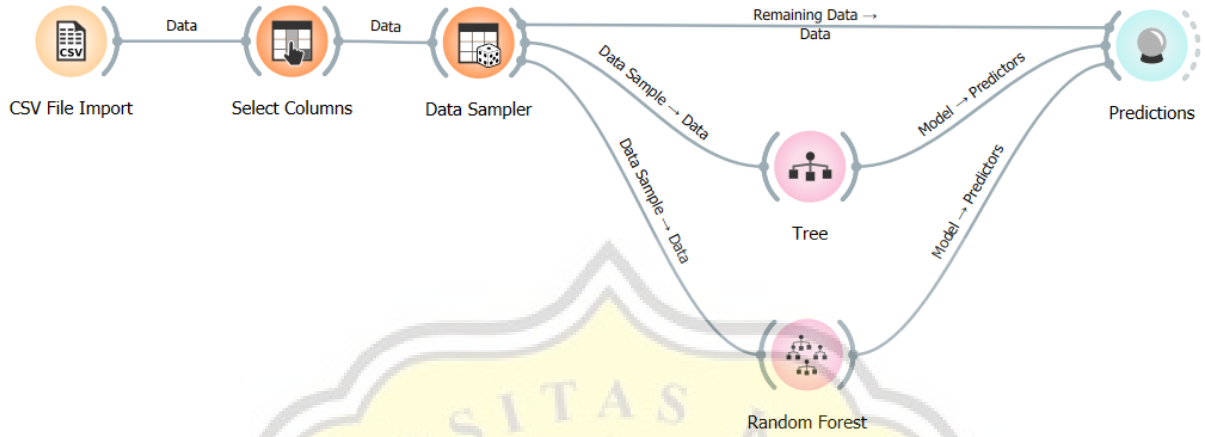

Figure 5.5 Final

As shown in Figure 5.5, it is the final result where the data has been processed by importing CSV files, determining target data, determining the proportion of data used, and then processing using the Random Forest and Tree algorithms.

5.2. Results

Table 5.1. Final Result

Model	MSE	RMSE	MAE
Random Forest	22546715369.754	150155.637	81058.773
Tree	30978136497.637	176006.069	93138.857

Table 5.1 shows the final result of the test on the Data Sampler which has 60% Data Training and 40% Data Testing. Where it is shown that the Random Forest algorithm is better than the Tree algorithm because it has a low value.

Table 5.2. Final Result

Model	MSE	RMSE	MAE
Random Forest	25555783722.165	159861.764	82556.176
Tree	34721071172.919	186335.910	93394.325

Table 5.2 shows the final result of the test on the Data Sampler which has 70% Data Training and 30% Data Testing. Where it is shown that the Random Forest algorithm is better than the Tree algorithm because it has a low value.

Table 5.3. Final Result

Model	MSE	RMSE	MAE
Random Forest	19734575289.577	140479.804	76120.199
Tree	29335988982.031	171277.520	90702.232

Table 5.3 shows the final result of the test on the Data Sampler which has 80% Data Training and 20% Data Testing. Where it is shown that the Random Forest algorithm is better than the Tree algorithm because it has a low value.

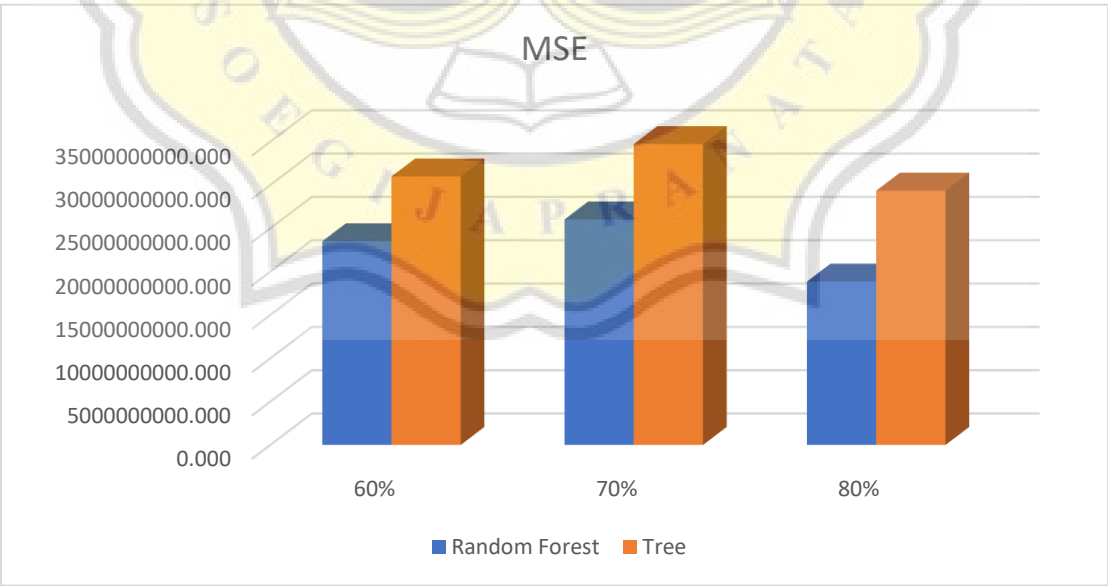

Figure 5.6 MSE

Based on the illustration in Figure 5.6, it can be seen that in the MSE results, the Random Forest algorithm gets lower results compared to the Tree algorithm. Therefore, it can be concluded that the Random Forest result is the best because it has a low value, which means that the errors were not as many as Tree.

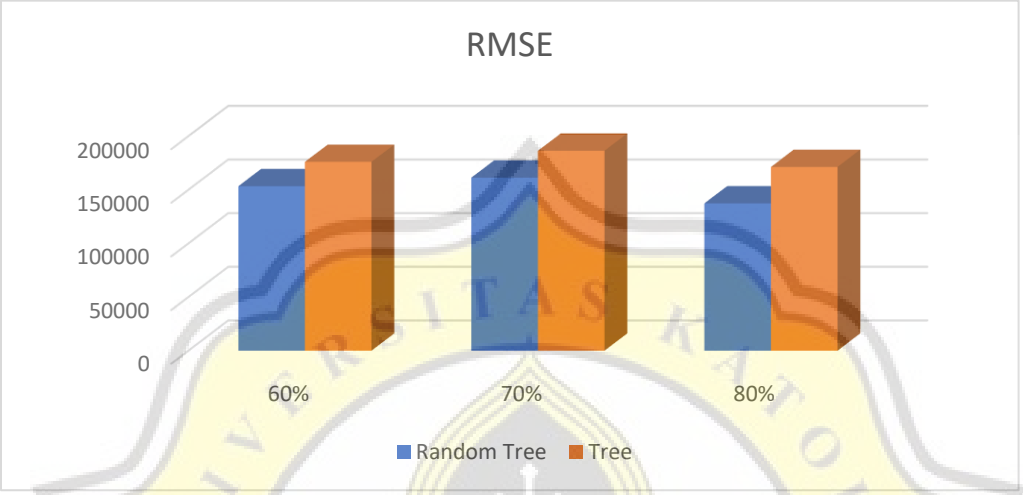

Figure 5.7 RMSE

Based on the illustration in Figure 5.7, it can be seen that in the RMSE results, the Random Forest algorithm gets lower results compared to the Tree algorithm. Therefore, it can be concluded that the Random Forest result is the best because it has a low value, which means that the errors were not as many as Tree.

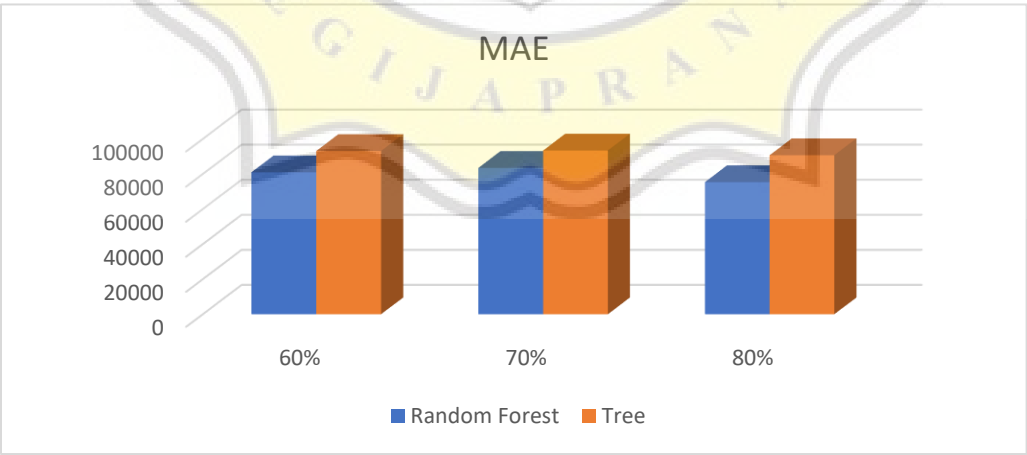

Figure 5.8 MAE

Based on the illustration in Figure 5.8, it can be seen that in the MAE results, the Random Forest algorithm gets lower results compared to the Tree algorithm. Therefore, it can be concluded that the Random Forest result is the best because it has a low value, which means that the errors were not as many as Tree.