ABSTRACT

Competition in the telecommunications industry is becoming increasingly fierce, making churn prevention a key strategy for maintaining company revenue. One of the main challenges faced by PT. XYZ is the high churn rate of its Fiber To The Home (FTTH) service customers. This study aims to build a churn prediction model using the random forest algorithm to help the company identify customers who are likely to churn at an early stage. Customer data from 2017 to 2023 was used through stages of data preprocessing, feature selection, model training, and performance evaluation using a test dataset comprising 20% of the total data (2,462 test data). The evaluation results show that the random forest model is capable of predicting churn with an accuracy of 92%, with a precision of 0.98, recall of 0.84, and F1-score of 0.90 in recognizing the churn class. In addition, the results are supported by visualizations in the form of pie charts, scatter plots, and histograms of customer churn probability segmentation. It is expected that these prediction results can serve as a basis for PT. XYZ to develop more effective and targeted churn prevention strategies.

Keywords— customer churn, fiber to the home, machine learning, prediction, random forest algorithm.