

ABSTRACT

ANALYSIS AND FRAUD DETECTION ON CALL DATA USING NAÏVE BAYES ALGORITHM IN PT XYZ

By

ILNA NUZLA PUTRI

NIM : 1202164374

Telecommunications fraud is one of the acts or activities that use telecommunications that is carried out illegally and intentionally in various forms of fraud, payment or incorporation by certain people or organizations that require services that are accompanied by service fees and payments made with assistance carried out illegally.

The purpose of this study is to submit a number of calls submitted as SIMBox fraud owned by PT XYZ who have the task in the matter of fraud. This research was conducted using data mining and using the Naïve Bayes algorithm.

Data mining is a technique that utilizes large amounts of data to obtain information that can be used in making important decisions. Naïve Bayes is an algorithm that can be used to predict phone numbers that are predicted to be SIMBox fraud which can be categorized as fraud and not fraud. By using data mining, specifically in the classification for predictions using the Naïve Bayes algorithm, predictions can be made on telephone numbers called fraud from call data. The test results using the Naïve Bayes algorithm an average value of 87.2% with an average macro precision value of 90%, the average value of macro recall is 86% and the average macro f1-score is 87%. While the lowest value is 85.2%. What is meant by the application of the Naïve Bayes algorithm is one of the best algorithms to be applied in predicting fraudulent call data at PT XYZ.

Keywords: Telecom Fraud, Data Mining, Bernoulli Naïve Bayes, SIMBox fraud prediction.