



**PROJECT REPORT**  
**FEATURE SELECTION USING MUTUAL**  
**INFORMATION FOR COMPARING NAIVE BAYES AND**  
**C4.5 IN PREDICTING CUSTOMER LOYALTY IN**  
**TELECOMMUNICATIONS COMPANIES**

**STEFANUS STANLEY PRASETYO MULYONO**  
**21.K1.0004**

**Faculty of Computer Science**  
**Soegijapranata Catholic University**  
**2024**

## ABSTRACT

*Customer loyalty is crucial for telecommunications companies, significantly impacting their sustainability and profitability. Retaining loyal customers and understanding the cause in customer churn is essential for creating successful retention strategies. While numerous studies have employed various data mining algorithms to predict customer loyalty, there is a need to analyze the effectiveness of various algorithms in this context. This study uses a feature selection method called Mutual Information to select the important features that have an impact on customer loyalty. While also applying data processing methods such as data cleansing and data transformation. The preprocessed data is divided into training sets and testing sets. After that, it compares the C4.5 and Naive Bayes algorithms based on the features selected by Mutual Information method to pick the most suitable for predicting customer loyalty based on customer status and to see which algorithm outperforms the other in terms of precision, recall, and accuracy. Both C4.5 and Naive Bayes algorithms are implemented and their performances evaluated using a confusion matrix across different data splits. This research provides telecommunications companies with insights into the advantages and disadvantages of these algorithms, enabling them to choose the most effective one for predicting customer loyalty. The findings will help develop more focused retention strategies, ultimately improving customer satisfaction and reducing churn, while also identifying the key factors influencing customer retention.*

**Keyword:** Customer Loyalty, Mutual Information, Naive Bayes, C4.5, Telecommunications companies