

***FEATURE SELECTION ANALYSIS OF CLASSIFICATION DATA USING  
C4.5 AND CHI SQUARE METHODS***

*Supervisor (1 People)*

**Nevarida Sindyka Diliansa Putri**  
***Study Program of Informatics Engineering***  
***Majoring in Information Technology***

***ABSTRACT***

*Data mining is part of the Knowledge Discovery in Database (KDD) process. With data mining we can perform description, prediction, grouping, association, forecasting, sorting and classification. Classification is a form of data analysis by representing a model of important data through certain features. Some of the algorithms in the classification include SVM, K-NN, Naive Bayes, Random Forest, Logistics Regression and Decision Tree. Decision Tree is a structure that is used to assist the decision-making process. Chi-square is one type of comparative test (Independency test) to compare two unpaired (Independent) categorical variables or proportions. With the results obtained, namely with 19 parameters, then feature selection is carried out with evidence that these parameters have an influence on the cervical cancer data class, which is 9 parameters and obtains an accuracy result of 100%.*

***Keywords:*** *Data mining, Classification, Decision Tree, Chi-Square, feature selection*