

ABSTRACT

QUESTION CLASSIFICATION BASED ON TOPIC CATEGORIES USING SUPPORT VECTOR MACHINES AND NAIVE BAYES METHODS (Case Study: DWBI Courses SI Information System)

By :

CHIKA ENGGAR PUSPITA

1202174292

Question classification is a computer science model, which aims to analyze questions and can label each question based on existing categories. Questions can be collected from several materials or topics that are many and different. With the classification of quiz questions, it can also help students in making decisions to determine the type of questions based on the topic category. Therefore, the researcher intends to create a classification model for quiz questions Data Warehouse and Business Intelligence that can be grouped into topics Data Warehouse, Business Intelligence, Data Analytics, and Performance Measurement. These topics were obtained from quiz questions in the Data Warehouse and Business Intelligence (DWBI) course in the 2018 Telkom University Information Systems SI study program.

One way to overcome this problem is to approach machine learning. Researchers also discussed how to deal with data that is not balanced by using technique oversampling with SMOTE and data validation using K-Fold Cross-Validation that the end result will be a comparison between datasets using SMOTE with datasets that do not use SMOTE

In this study, researchers used a comparison of algorithms, machine learning namely Naïve Bayes and Support Vector Machines using SMOTE and methods Cross-Validation. The results of this study show the best accuracy results and are very helpful. The results obtained in the method cross-validation before SMOTE produced an accuracy rate of 82.02% for the results after going through the SMOTE stage of 94.79% in the algorithm Naïve Bayes, while the algorithm Support Vector Machine produced an accuracy of 81.39% in the process before SMOTE and for the results after going through SMOTE by 96.52%.

Keywords— Question Classification; Machine Learning; Naive Bayes; Support Vector Machine; Cross-Validation