

ABSTRACT

Education, as a fundamental aspect of human life, continues to evolve alongside advances in Science and Technology. One of the key innovations in education is e-learning, which allows students to learn without being confined to a physical classroom. However, with the increasing number of quiz questions that vary in topics, particularly in the subject of Science, which covers various scientific concepts, managing questions manually becomes inefficient. Therefore, a system that can automatically and efficiently organize and classify these questions is needed to enhance students' understanding, especially in Science subjects. This study aims to implement the Support Vector Machine (SVM) method for multi-label classification of science questions at the junior high school level. The classification process includes data cleaning, case folding, tokenization, stopword removal, stemming, and text feature extraction using TF-IDF. The modelling uses a problem transformation approach with the label powerset method to convert multi-label questions into a multi-class format, enabling binary classification by SVM. The model is evaluated using a confusion matrix to analyze classification performance and K-Fold Cross Validation to ensure accuracy and generalization of the model. The results show that SVM can be applied for multi-label question classification with an accuracy of 69% and F1-score of 76%. The confusion matrix analysis reveals that the model has some classification errors, indicating room for further improvement. Nonetheless, the SVM model demonstrates good potential in this application. This study also identifies several areas for improvement, including enhancing data quality and optimizing model parameters. Therefore, the SVM method is worth considering in educational systems for developing question banks and technology-based evaluation systems, although further improvements in the model and data are necessary.

Keywords— confusion matrix, e-learning, k-fold cross validation, multi-label class, question bank, support vector machine.