

ABSTRAK

Pendidikan, sebagai aspek fundamental kehidupan manusia, terus berkembang seiring dengan kemajuan Ilmu Pengetahuan dan Teknologi (IPTEK). Salah satu inovasi penting dalam pendidikan adalah *e-learning*, yang memungkinkan siswa belajar tanpa terikat ruang kelas. Namun, dengan semakin banyaknya soal kuis yang beragam topiknya, terutama dalam mata pelajaran IPA yang mencakup berbagai konsep ilmiah, pengelolaan soal secara manual menjadi tidak efisien dan memerlukan banyak waktu. Oleh karena itu, diperlukan sistem yang dapat mengorganisasi dan mengelompokkan soal secara otomatis dan efisien, sehingga dapat meningkatkan pemahaman siswa, khususnya pada mata pelajaran IPA. Penelitian ini bertujuan untuk mengimplementasikan metode *Support Vector Machine* (SVM) dalam klasifikasi soal *multi-label* pada mata pelajaran IPA tingkat SMP. Proses klasifikasi mencakup pembersihan data, *case folding*, *tokenization*, *stopword removal*, *stemming*, dan ekstraksi fitur teks menggunakan TF-IDF. Pemodelan menggunakan pendekatan *problem transformation* dengan metode *label powerset* untuk mengubah soal dengan *multi-label* menjadi bentuk *multi-class* sehingga bisa dilakukan klasifikasi biner oleh SVM. Evaluasi model dilakukan menggunakan *confusion matrix* untuk menganalisis performa klasifikasi dan *K-Fold Cross Validation* untuk memastikan keakuratan dan generalisasi model. Hasil penelitian menunjukkan bahwa SVM dapat diterapkan untuk klasifikasi soal *multi-label* dengan akurasi 69% dan *F1-score* sebesar 76%. Analisis *confusion matrix* mengungkapkan bahwa model memiliki beberapa kesalahan klasifikasi, mengindikasikan ruang untuk perbaikan lebih lanjut. Meskipun demikian, model SVM menunjukkan potensi yang baik. Penelitian ini juga mengidentifikasi beberapa area untuk perbaikan, termasuk peningkatan kualitas data dan pemilihan parameter model yang lebih optimal. Oleh karena itu, metode SVM layak dipertimbangkan dalam sistem pendidikan untuk pengembangan bank soal dan sistem evaluasi berbasis teknologi, meskipun diperlukan perbaikan lebih lanjut pada model dan data.

Kata kunci— bank soal, *confusion matrix e-learning*, klasifikasi *multi-label*, *k-fold cross validation*, *support vector machine*.