

BAB I PENDAHULUAN

I.1 Latar Belakang

Pendidikan adalah suatu hal yang sangat penting dalam kehidupan manusia, dan setiap individu memiliki hak untuk memperoleh pendidikan sebagai langkah dalam pengembangan dirinya. Secara umum, pendidikan memiliki arti suatu proses kehidupan dalam mengembangkan diri setiap individu untuk dapat hidup dan melangsungkan kehidupan. Sehingga menjadi seorang yang terdidik itu sangat penting (Yayan Alpian dkk., 2019).

Seiring berkembangnya Ilmu Pengetahuan dan Teknologi (IPTEK), metode pembelajaran berkembang menjadi *e-learning*. *E-learning* merupakan suatu konsep dalam proses belajar mengajar yang dapat dilakukan oleh siswa kapan saja dan tidak terpacu pada sebuah ruang kelas. Pada *E-learning*, siswa dapat meningkatkan kemampuan secara mandiri dalam memahami materi pembelajaran melalui kuis. Soal-soal kuis tersebut kemudian dapat dikumpulkan dan disimpan ke dalam bank soal (Dan & Mailangkay, 2016). Namun, dengan semakin banyaknya soal yang beragam topiknya, pengelolaan soal secara manual menjadi tidak efisien dan memakan banyak waktu. Hal ini terutama berlaku untuk mata pelajaran dengan topik yang luas dan kompleks, seperti IPA, yang mencakup topik seperti manusia, hewan, dan tumbuhan, serta istilah-istilah ilmiah yang seringkali sulit dipahami oleh siswa (Yulistiawati, 2019).

Klasifikasi soal berdasarkan topik adalah salah satu cara untuk mempermudah pengelolaan bank soal. Namun, proses ini menjadi tantangan ketika dilakukan secara manual karena rentan terhadap ketidakakuratan dan inkonsistensi hasil klasifikasi (Aninditya dkk., 2019). Selain itu, proses klasifikasi soal berdasarkan topik dapat yang dilakukan secara manual akan menghabiskan banyak waktu seiring dengan soal yang banyak dan beragam (Ariandi dkk., 2023). Penelitian lain menyebutkan bahwa klasifikasi otomatis dapat memangkas waktu pemrosesan hingga 180 kali lebih cepat dibandingkan dengan klasifikasi manual dalam skenario industri farmasi (Unnikrishnan dkk., 2020). Hal ini menegaskan pentingnya efisiensi waktu dalam pengelolaan data yang besar dan beragam, yang juga relevan dalam konteks pengelolaan bank soal pada platform *e-learning*. Oleh

karena itu, diperlukan penerapan teknik klasifikasi otomatis untuk mengelompokkan soal-soal berdasarkan topiknya secara lebih efisien dan akurat. Penggunaan sistem klasifikasi soal dalam platform *e-learning* akan mempermudah pengajar dan siswa dalam mengelola dan menavigasi bank soal yang kompleks, serta membantu dalam proses pengambilan keputusan terkait topik soal (Muhaimin dkk., 2023).

Selain itu, banyak soal evaluasi yang tidak hanya berhubungan dengan satu topik, tetapi juga dapat mencakup beberapa topik sekaligus (*multi-label*). Hal ini menimbulkan tantangan dalam proses klasifikasi karena setiap soal dapat memiliki lebih dari satu label topik. Oleh karena itu, metode klasifikasi *multi-label* diperlukan untuk menangani kasus di mana satu soal harus diklasifikasikan ke dalam beberapa topik sekaligus. Klasifikasi *multi-label* ini akan memungkinkan sistem untuk mengelompokkan soal-soal tersebut sehingga dapat membantu guru dalam mengelola bank soal yang kompleks. Tantangan utama dalam klasifikasi teks *multi-label* adalah ruang label yang sangat besar, yang meningkat secara eksponensial dengan jumlah label kandidat. Selain itu, metode klasifikasi *multi-label* sebelumnya sering kali tidak mampu menangani peningkatan jumlah label spesifik dan contoh pelatihan dengan efisien, sehingga sulit untuk diimplementasikan pada skala besar (Taha & Tiun, 2016).

Pada klasifikasi soal secara otomatis, tentunya terdapat banyak teknik atau algoritma yang dapat digunakan yaitu Naïve Bayes, *Support Vector Machine*, *Decision Tree Support*, *Neural Network*, dan lain-lain (Wibawa dkk., 2018). Dalam penelitian ini, proses klasifikasi soal berdasarkan topik akan difokuskan pada mata pelajaran IPA dengan menggunakan metode *Support Vector Machine* (SVM). SVM atau *Support Vector Machine* adalah algoritma *machine learning* yang dikembangkan oleh Vapnik, Guyon, dan Boser. SVM pertama kali diperkenalkan pada *Annual Workshop on Computational Learning Theory* sekitar tahun 1992. Algoritma SVM ini memiliki tujuan untuk mencari *hyperplane* terbaik yang dapat membagi dua kelas pada *input space* (Handayanto dkk., 2019). Walaupun demikian, SVM juga dinilai dapat menjadi *classifier* karena dapat mengklasifikasikan soal dari area yang beragam dengan akurasi yang tinggi fitur

BOW yang menjadi salah satu indikator dalam klasifikasi teks seperti soal-soal (Sangodiah dkk., 2015).

Penelitian sebelumnya dalam jurnal berjudul "*Question Classification using Machine Learning Approaches*" (Panicker & Venkitakrishnan, 2012) telah menggunakan metode mesin pembelajaran. Hasil penelitian tersebut mengungkapkan bahwa algoritma *Support Vector Machine* (SVM) secara konsisten unggul dibandingkan dengan *Naïve Bayes Classifier* dalam konteks klasifikasi soal, menunjukkan bahwa SVM adalah alat klasifikasi yang efektif dalam konteks *e-learning*.

Selain itu, SVM memiliki kemampuan yang baik dalam generalisasi, yang berarti algoritma ini dapat mengklasifikasikan data baru yang belum pernah dilihat sebelumnya dengan tingkat akurasi yang tinggi. Dalam konteks klasifikasi *multi-label*, SVM telah diperluas dengan teknik seperti *Binary Relevance* (BR) dan *Label Powerset* (LP) untuk menangani lebih dari satu label per *instance* yang menjadikannya pilihan yang fleksibel dan kuat dalam klasifikasi *multi-label* (Boutell dkk., 2004).

I.2 Perumusan Masalah

Berdasarkan latar belakang yang telah dibahas, berikut merupakan rumusan masalah yang akan diidentifikasi pada penelitian ini antara lain:

1. Bagaimana implementasi klasifikasi soal *multi-label* berdasarkan topik pada mata pelajaran IPA tingkat SMP menggunakan metode *Support Vector Machine*?
2. Bagaimana kinerja yang dihasilkan dari penerapan metode *Support Vector Machine*?
3. Bagaimana karakteristik soal di setiap topik pada mata pelajaran IPA?

I.3 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah diuraikan, maka tujuan dari penelitian ini yaitu:

1. Mengetahui hasil implementasi klasifikasi soal *multi-label* berdasarkan topik pada mata pelajaran IPA tingkat SMP menggunakan metode *Support Vector Machine*.
2. Mengetahui hasil kinerja dari penerapan metode *Support Vector Machine* dalam proses pengklasifikasian soal *multi-label* berdasarkan topik pada mata pelajaran IPA tingkat SMP.
3. Mengetahui karakteristik dan kata kunci yang dimiliki setiap topik pada mata pelajaran IPA.

I.4 Batasan Penelitian

Adapun batasan dalam melakukan penelitian ini adalah sebagai berikut:

1. Dataset bersumber dari bank soal Pinteran.
2. Soal yang diuji menggunakan mata pelajaran IPA.
3. Soal yang diuji hanya menggunakan jenis pilihan ganda, berbentuk teks, berbahasa Indonesia, dan tidak mengandung persamaan rumus.
4. *Tools* yang digunakan adalah *Python*.

I.5 Manfaat Penelitian

Manfaat penelitian ini adalah sebagai berikut:

1. Bagi pengajar, penelitian ini diharapkan dapat membantu pengajar dalam melakukan klasifikasi soal berdasarkan kategori topik.
2. Bagi pengembang *e-learning*, penelitian ini diharapkan dapat memberikan informasi mengenai klasifikasi soal khususnya menggunakan algoritma *Support Vector Machine*.
3. Bagi peneliti lain, penelitian ini diharapkan dapat memberikan informasi bagi peneliti ataupun calon peneliti lain untuk diterapkan ke dalam model yang lebih luas dan lebih kompleks serta sebagai bahan acuan untuk referensi pustaka akademik.