

ABSTRAK

Penentuan peminatan siswa di tingkat sekolah menengah atas (SMA) merupakan langkah penting dalam proses pendidikan, karena berpengaruh terhadap pengembangan potensi akademik siswa serta arah karier di masa depan. SMAN 2 Probolinggo, seperti banyak sekolah lainnya, menghadapi tantangan dalam mengoptimalkan pemilihan peminatan bidang studi yang sesuai dengan kemampuan dan minat siswa. Oleh karena itu, diperlukan sistem pendukung keputusan yang dapat membantu pihak sekolah dalam mengklasifikasikan siswa berdasarkan kemampuan mereka. Penelitian ini bertujuan untuk menerapkan *algoritma C4.5* dalam membangun model klasifikasi berbasis *Decision Tree* guna menentukan peminatan siswa di SMAN 2 Probolinggo. penelitian ini menggunakan metode *Decision Tree* dengan algoritma C4.5. Data yang digunakan meliputi nilai akademik seperti Matematika, Fisika, Kimia, Biologi, Ekonomi, Sosiologi, Geografi, dan Bahasa Inggris. Model diuji dengan parameter seperti *max_depth*, *min_samples_split*, dan *random_state*. Penelitian ini juga mengevaluasi akurasi model untuk memastikan hasil yang optimal.

Hasil penelitian menunjukkan bahwa model *Decision Tree* dengan *max_depth=10*, *min_samples_split=2*, dan *random_state=42* memberikan akurasi terbaik sebesar 94,29%. Model ini mampu mengklasifikasikan siswa ke dalam 4 bidang peminatan, yaitu Paket Bidang Minat A, B, C, dan D dengan tingkat akurasi yang tinggi.

Kata kunci: *Algoritma C4.5*, klasifikasi, *decision tree*, peminatan siswa, sistem pendukung keputusan.

ABSTRACT

The determination of students' specialization at the senior high school (SMA) level is a crucial step in the educational process, as it influences their academic potential and future career direction. SMAN 2 Probolinggo, like many other schools, faces challenges in optimizing the selection of study specializations that align with students' abilities and interests. Therefore, a decision support system is needed to assist the school in classifying students based on their capabilities. This study aims to implement the *C4.5 algorithm* in building a *Decision Tree-based classification model* to determine students' specialization at SMAN 2 Probolinggo.

This research employs the *Decision Tree method* using the *C4.5 algorithm*. The dataset includes academic scores in subjects such as Mathematics, Physics, Chemistry, Biology, Economics, Sociology, Geography, and English. The model is tested using parameters such as *max_depth*, *min_samples_split*, and *random_state*. Additionally, this study evaluates the model's accuracy to ensure optimal results.

The findings indicate that the *Decision Tree* model with *max_depth=10*, *min_samples_split=2*, and *random_state=42* achieves the highest *accuracy of 94.29%*. This model successfully classifies students into four specialization categories: Specialization Package A, B, C, and D, with a high level of accuracy.

Keywords: *C4.5 Algorithm, Classification, Decision Tree, Student Specialization, Decision Support System.*