

BAB 1. PENDAHULUAN

1.1 Latar Belakang

Pendidikan merupakan komponen penting dalam membentuk sumber daya manusia yang berkualitas, baik dari aspek pengetahuan, keterampilan, maupun karakter (Maulana dkk., 2024). Seiring dengan perkembangan zaman, dunia pendidikan dituntut untuk mampu beradaptasi dengan kemajuan teknologi informasi dan komunikasi. Transformasi digital tidak hanya terjadi pada proses pembelajaran, tetapi juga pada sistem evaluasi yang digunakan untuk mengukur capaian belajar peserta didik (Puteri dkk., 2025). Pemanfaatan teknologi dalam sistem evaluasi pembelajaran menjadi salah satu solusi strategis untuk meningkatkan efektivitas dan efisiensi pengolahan data hasil belajar. Sistem evaluasi berbasis digital memungkinkan proses penilaian dilakukan secara lebih cepat, akurat, dan terdokumentasi dengan baik sehingga dapat meningkatkan kualitas proses evaluasi pembelajaran secara keseluruhan (Wati dkk., 2026).

Salah satu aspek penting dalam sistem evaluasi pembelajaran adalah pengelolaan bank soal. Bank soal merupakan kumpulan butir soal yang disusun secara sistematis dan terorganisir, serta dilengkapi dengan kunci jawaban dan informasi pendukung lainnya untuk memudahkan penyusunan instrumen evaluasi pembelajaran (Qhorifadillah dkk., 2022). Pengelolaan bank soal yang baik bertujuan untuk mempermudah guru dalam menyusun instrumen evaluasi pembelajaran, menjaga kualitas soal, serta meningkatkan efisiensi dalam pelaksanaan penilaian. Selain berfungsi sebagai media penyimpanan soal, bank soal juga dapat menjadi sumber data yang mendukung pengembangan kualitas instrumen evaluasi melalui analisis karakteristik setiap butir soal. Namun, pada praktiknya, pengelolaan bank soal di sejumlah sekolah masih belum didukung oleh sistem yang mampu melakukan analisis karakteristik soal secara optimal sehingga informasi yang dihasilkan dari proses evaluasi belum dapat dimanfaatkan secara maksimal.

Berdasarkan hasil wawancara dengan Bapak Muhammad Fathur Rozi selaku Kepala TU di MA Manbaul Ulum, diperoleh informasi bahwa sekolah telah menggunakan Aplikasi Madrasah CBT (Computer Based Test) sebagai *platform* pelaksanaan ujian, terutama pada kegiatan evaluasi tertentu seperti ujian kenaikan kelas. Penggunaan aplikasi tersebut telah membantu sekolah dalam melaksanakan proses evaluasi secara digital. Namun, karena aplikasi Madrasah CBT merupakan sistem pihak ketiga yang bersifat berbayar, sekolah memiliki keterbatasan dalam melakukan pengembangan maupun penyesuaian fitur sesuai kebutuhan. Sistem tersebut juga belum dirancang secara khusus untuk mendukung pengelolaan bank soal dan analisis tingkat kesulitan soal yang menjadi fokus dalam penelitian ini. Saat ini, sistem yang digunakan belum menyediakan fitur analisis butir soal yang mampu menyajikan informasi mengenai persentase jawaban benar siswa dan rata-rata waktu pengerjaan pada setiap butir soal. Penentuan tingkat kesulitan soal pun masih didasarkan pada pertimbangan dan pengalaman guru, sehingga berpotensi menimbulkan perbedaan antara tingkat kesulitan yang ditetapkan dengan kondisi sebenarnya berdasarkan hasil pengerjaan siswa. Kondisi tersebut menyebabkan data hasil pengerjaan siswa belum dapat dimanfaatkan secara optimal, sehingga guru mengalami kesulitan dalam memperoleh informasi yang objektif mengenai karakteristik setiap butir soal, khususnya tingkat kesulitan soal yang digunakan dalam evaluasi pembelajaran.

Analisis kualitas butir soal merupakan aspek penting dalam evaluasi pembelajaran karena dapat digunakan untuk menilai kualitas dan kelayakan instrumen tes yang digunakan (Muatikasari dkk., 2025). Informasi yang diperoleh dari analisis tersebut dapat dijadikan dasar dalam mempertahankan, memperbaiki, atau mengganti soal sehingga kualitas instrumen evaluasi dapat terus ditingkatkan. Oleh karena itu, diperlukan suatu sistem bank soal berbasis *website* yang tidak hanya berfungsi sebagai media penyimpanan soal, tetapi juga mampu melakukan analisis tingkat kesulitan soal secara otomatis dan objektif berdasarkan data hasil pengerjaan siswa.

Salah satu metode yang dapat digunakan untuk mendukung analisis tingkat kesulitan soal adalah *K-Means Clustering*. *K-Means* merupakan teknik dalam data

mining yang termasuk ke dalam kategori *unsupervised learning* dan digunakan untuk mengelompokkan data berdasarkan tingkat kemiripan karakteristik tertentu (Pradana dkk., 2025). Metode ini bekerja dengan mengelompokkan data berdasarkan kedekatan terhadap pusat *cluster (centroid)* yang diperbarui secara iteratif hingga diperoleh kelompok yang stabil. *K-Means* memiliki keunggulan dalam mengidentifikasi pola dalam data tanpa memerlukan label sebelumnya sehingga efektif digunakan untuk proses pengelompokan data secara objektif (Hendrastuty, 2024). Dalam penelitian ini, *K-Means* digunakan untuk mengelompokkan soal berdasarkan atribut persentase jawaban benar dan rata-rata waktu pengerjaan siswa. Persentase jawaban benar digunakan untuk menggambarkan tingkat keberhasilan siswa dalam menjawab soal, sedangkan rata-rata waktu pengerjaan digunakan sebagai indikator tambahan untuk menggambarkan tingkat kompleksitas soal yang dihadapi siswa. Berdasarkan kedua atribut tersebut, diperoleh kategori tingkat kesulitan soal secara lebih sistematis dan berbasis data.

Selain *K-Means Clustering*, penelitian ini juga menerapkan algoritma *Naïve Bayes* untuk proses klasifikasi tingkat kesulitan soal. *Naïve Bayes* merupakan algoritma *supervised learning* yang bekerja berdasarkan teori probabilitas dengan memanfaatkan data latih untuk memprediksi kelas suatu data baru (Imanuel dkk., 2024). Algoritma ini dipilih karena memiliki proses komputasi yang sederhana, mampu bekerja dengan baik pada jumlah data yang relatif terbatas, serta memiliki tingkat akurasi yang baik dalam proses klasifikasi. Dalam penelitian ini, hasil pengelompokan yang diperoleh dari *K-Means* digunakan sebagai label kelas pada data latih yang kemudian dimanfaatkan oleh *Naïve Bayes* untuk mengklasifikasikan tingkat kesulitan soal baru. Kombinasi kedua metode tersebut memungkinkan sistem untuk mengelompokkan soal ke dalam kategori mudah, sedang, dan sulit secara objektif, serta mengklasifikasikan tingkat kesulitan soal baru secara otomatis berdasarkan pola data yang telah terbentuk.

Berdasarkan uraian permasalahan dan kebutuhan yang telah dijelaskan, penelitian ini bertujuan untuk mengembangkan sistem bank soal berbasis *website* yang mampu mengelompokkan dan mengklasifikasikan tingkat kesulitan soal

menggunakan metode *K-Means Clustering* dan *Naïve Bayes* di MA Manbaul Ulum. Sistem yang dikembangkan diharapkan dapat membantu guru dalam mengelola bank soal secara lebih mandiri dan efektif, menyediakan informasi hasil pengelompokan dan klasifikasi tingkat kesulitan soal ke dalam kategori mudah, sedang, dan sulit secara objektif, serta mendukung penyusunan instrumen evaluasi yang lebih terukur dan berkualitas berdasarkan data hasil pengerjaan siswa.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan, maka rumusan masalah dalam penelitian ini adalah sebagai berikut:

- a. Bagaimana merancang dan membangun sistem bank soal berbasis *website* yang dilengkapi fitur analisis data hasil pengerjaan siswa, seperti rata-rata waktu pengerjaan dan persentase keberhasilan per butir soal, di MA Manbaul Ulum?
- b. Bagaimana menerapkan algoritma *K-Means Clustering* untuk mengelompokkan soal ke dalam kategori mudah, sedang, dan sulit berdasarkan data hasil pengerjaan siswa secara objektif?
- c. Bagaimana memanfaatkan algoritma *Naïve Bayes* untuk mengklasifikasikan soal baru ke dalam kategori tingkat kesulitan yang sesuai secara otomatis berdasarkan parameter persentase jawaban benar dan rata-rata waktu pengerjaan?

1.3 Batasan Masalah

Untuk membatasi ruang lingkup agar penelitian ini lebih terarah dan fokus, maka batasan-batasan dalam penelitian ini adalah sebagai berikut:

- a. Sistem yang dikembangkan berbasis *website* dan menggunakan metode *K-Means* dan *Naive Bayes*.
- b. Jenis soal yang digunakan terbatas pada soal pilihan ganda untuk mata pelajaran Bahasa Indonesia dan Matematika di MA Manbaul Ulum.
- c. Data yang digunakan untuk pengelompokan dan klasifikasi soal meliputi persentase jawaban benar dan rata-rata waktu pengerjaan soal per butir.
- d. Pengujian sistem dilakukan menggunakan data soal dan hasil pengerjaan 22 siswa kelas XII IPA A di MA Manbaul Ulum.

1.4 Tujuan

Adapun tujuan dari penelitian ini adalah:

- a. Merancang dan membangun sistem bank soal berbasis *website* yang dilengkapi fitur analisis data hasil pengerjaan siswa, sehingga guru dapat mengelola soal secara terstruktur, efisien, dan berbasis data yang objektif.
- b. Menerapkan algoritma *K-Means Clustering* untuk mengelompokkan soal ke dalam kategori tingkat kesulitan mudah, sedang, dan sulit berdasarkan persentase jawaban benar dan rata-rata waktu pengerjaan.
- c. Menggunakan hasil pengelompokan *K-Means* sebagai data latih dalam algoritma *Naïve Bayes* untuk mengklasifikasikan tingkat kesulitan soal baru secara otomatis.

1.5 Manfaat

Penelitian ini diharapkan dapat memberikan manfaat sebagai berikut:

- a. Memberikan kemudahan bagi guru di MA Manbaul Ulum dalam menyusun, menyimpan, dan mengelola soal secara digital, terstruktur, dan terdokumentasi dengan baik.
- b. Membantu guru dalam menentukan tingkat kesulitan soal secara objektif berbasis data hasil pengerjaan siswa, sehingga tidak lagi bergantung pada penilaian subjektif semata dan distribusi soal dalam ujian menjadi lebih seimbang.
- c. Meningkatkan efektivitas dan efisiensi proses evaluasi pembelajaran melalui sistem bank soal yang terintegrasi dengan metode *K-Means Clustering* dan *Naïve Bayes*.
- d. Memberikan kemudahan bagi siswa dalam mengikuti latihan ujian berbasis *website* tanpa ketergantungan pada *platform* pihak ketiga, sehingga persiapan ujian dapat dilakukan secara lebih fleksibel dan terstruktur.