

BAB I

PENDAHULUAN

1.1 Latar Belakang

Penilaian hasil pembelajaran diperlukan untuk mengevaluasi tingkat pemahaman siswa dalam memahami materi pembelajaran. Banyak cara dapat dilakukan untuk mengevaluasi hasil belajar siswa diantaranya melalui pemberian pekerjaan rumah, ulangan harian dan ujian. Tingkat pemahaman siswa pada materi yang disampaikan bergantung dengan bahan penilaian yang dibuat oleh guru. Siswa dianggap mengerti materi pembelajaran yang disampaikan jika siswa berhasil mengerjakan evaluasi atau ujian dengan baik. Pada zaman globalisasi sekarang ini kemajuan teknologi, komunikasi dan informasi sangat berdampak di kehidupan manusia termasuk diantaranya pendidikan. Kualitas proses pembelajaran harus meningkat seiring dengan perkembangan teknologi dan informasi sesuai tuntutan di era globalisasi modern saat ini [1]. Pengaruh perkembangan teknologi pada proses pembelajaran seperti adanya sistem pembelajaran daring dan bentuk ujian berbasis daring melalui perangkat komputer atau ponsel cerdas. Tetapi di dalam ujian berbasis daring sebagian besar terbatas pada bentuk soal pilihan ganda yang dianggap kurang efektif karena siswa dapat mengira-ngira jawaban yang tertera. Hal ini menyebabkan guru kurang dapat menilai atau mengevaluasi hasil belajar siswa tersebut [2].

Tipe soal uraian dianggap dapat menguji kemampuan siswa dalam menganalisis, berpikir dan memecahkan persoalan dengan gagasan yang dimiliki masing-masing siswa. Maka dari itu seorang guru sanggup mengetahui seberapa jauh kualitas pemahaman setiap siswa tersebut dalam memahami materi pembelajaran. Salah satu kelemahan pada tipe soal uraian yaitu sulitnya menilai jawaban siswa dan dibutuhkan banyak waktu dalam proses penilaian, karena guru harus mencocokkan hasil jawaban siswa satu persatu dengan kunci jawaban yang benar. Hal ini menuntut guru tetap konsisten dalam mengoreksi banyak jawaban [3]. Untuk mengatasi masalah tersebut, sistem penilaian otomatis (*Automatic Essay Scoring*) perlu dibuat agar mempercepat dan memudahkan proses penilaian. Pada penelitian yang dilakukan ini berfokus pada penilaian soal uraian secara otomatis. Cara kerja sistem penilaian otomatis soal uraian adalah mencari kemiripan antara

jawaban masing-masing siswa dengan kunci jawaban dari buku menerapkan algoritma *text similarity* seperti *winnowing*.

Penelitian yang dilakukan ini akan menerapkan metode *winnowing* karena metode ini mempunyai tingkat akurasi lebih tinggi ketika dibandingkan dua algoritma lainnya sesuai dalam referensi [4] yang membandingkan tiga algoritma *text similarity* yaitu *winnowing*, *rabin-karp* dan *knuth-morris-pratt*. Disimpulkan bahwa dari 20 dokumen yang diuji dengan jumlah kata sebanyak 200 kata. Penelitian ini menyimpulkan bahwa metode *winnowing* dapat mendeteksi kesamaan kata tingkat tinggi dalam dokumen, metode *rabin karp* pada level menengah, dan metode *knuth morris pratt* dapat mendeteksi kesamaan kata pada level rendah. Dari hasil penghitungan rata-rata kemiripan metode paling optimal dari ketiganya adalah metode *winnowing*, yang menghasilkan akurasi 94,3%. Ini juga menunjukkan akurasi tertinggi, dengan hanya perbedaan 1,19% dari pemeriksaan manual oleh manusia, sebesar 93,11%. Dengan tingginya akurasi algoritma *winnowing* dibanding *rabin-karp* dan *knuth morris* maka algoritma ini akan digunakan dalam sistem penilaian otomatis yang dirancang.

Algoritma *winnowing* berfokus pada menghitung nilai *hash* yang kemudian dilakukan pencocokan tingkat kesamaan dokumen menggunakan hasil pemilihan nilai *hash*. Oleh karena hanya mencocokkan nilai *hash*, algoritma *winnowing* memiliki kekurangan yaitu hanya melakukan pencocokan nilai *hash* tanpa mempedulikan arti dari kata. Apabila terdapat dua kata berbeda namun kata-kata tersebut memiliki makna sama, pada algoritma *winnowing* hal tersebut akan diabaikan. Untuk mengatasi kekurangan algoritma *winnowing* diperlukan *synonym recognition* yaitu pengenalan kata melalui pendekatan sinonim. *Synonym recognition* bertujuan untuk mengenali kata berbeda yang memiliki makna sama [5].

Pembuatan sistem penilaian soal uraian otomatis menggunakan metode *winnowing* memerlukan beberapa tahap yaitu: tahap *text preprocessing*, yang tersusun dari beberapa tahapan yaitu *text cleaning*, *case folding*, *tokenizing* dan *stemming*. Selanjutnya tahap *synonym recognition*, tahap *winnowing* untuk menentukan nilai *fingerprint* yang mewakili dokumen, terakhir *jaccard coefficient similarity* untuk menghitung tingkat kesamaan jawaban antara kunci jawaban dan

jawaban siswa. Data pada penelitian diambil di SMP Negeri 1 Megaluh mata pelajaran biologi. Proses penilaian di SMP Negeri 1 Megaluh masih dikerjakan menggunakan cara manual dengan jumlah siswa yang banyak, diperlukan waktu yang lama dalam proses penilaian. Dengan menerapkan sistem penilaian otomatis diharapkan bisa menolong para guru dalam melakukan penilaian dan mempersingkat waktu yang dibutuhkan dalam menilai jawaban siswa.

1.2 Rumusan Masalah

Berdasarkan dari uraian latar belakang maka dapat dipaparkan rumusan masalah sebagai berikut:

1.2.1 Permasalahan

Penilaian otomatis soal uraian dengan *input* jawaban yang memiliki banyak variasi kata menjadi kendala dalam metode *winnowing*. Metode ini memiliki kekurangan yaitu tidak dapat mengenali kata berbeda dengan makna sama. Kekurangan ini dapat mempengaruhi hasil perbandingan kesamaan dokumen kunci jawaban dan jawaban siswa. Selain itu diperlukan konfigurasi parameter dalam metode *winnowing* agar memperoleh hasil terbaik.

1.2.2 Solusi Permasalahan

Penelitian ini mengusulkan beberapa pengujian dengan menggunakan kombinasi nilai *parameter* berbeda pada metode *winnowing* agar menghasilkan nilai *RMSE* terkecil. Lalu untuk mengatasi kekurangan metode *winnowing* dalam mengenali arti kata diusulkan penerapan *synonym recognition* untuk mengatasi variasi kata agar kata berbeda dengan makna sama dapat dikenali oleh metode *winnowing*.

1.2.3 Pertanyaan Penelitian

1. Berapa nilai parameter pada metode *winnowing* agar memperoleh nilai *RMSE* terkecil pada sistem penilaian otomatis soal uraian?
2. Bagaimana perbandingan nilai *RMSE* sistem penilaian otomatis dengan metode *winnowing* tanpa menggunakan *synonym recognition* dan dengan menggunakan *synonym recognition*?

1.3 Batasan Masalah

1. Aplikasi hanya menguji data berupa teks, tidak menguji jawaban dalam bentuk persamaan matematika, gambar dan suara.
2. Data soal dan jawaban siswa menggunakan mata pelajaran biologi berbahasa Indonesia.
3. Perancangan sistem menggunakan bahasa pemrograman *python* dengan bantuan *library nltk* dan *sastrawi*.

1.4 Manfaat dan Tujuan

1.4.1 Tujuan

Penelitian ini bertujuan sebagai berikut:

1. Memperoleh nilai parameter pada metode *winnowing* yang menghasilkan nilai *RMSE* paling kecil pada sistem penilaian otomatis soal uraian.
2. Mengetahui pengaruh *synonym recognition* terhadap nilai *RMSE* metode *winnowing* dalam menentukan nilai siswa.

1.4.2 Manfaat

Diharapkan dapat membantu para guru dalam menilai jawaban siswa dan dapat dijadikan sebagai rujukan penelitian lainnya agar terbentuk sistem yang memiliki hasil lebih baik.

1.5 Sistematika Proposal

Sistematika penulisan proposal skripsi ini adalah sebagai berikut:

Bab I berisi tentang pendahuluan yang dilakukan pada penelitian ini. Bab ini membahas latar belakang permasalahan yang ada, solusi yang diusulkan untuk memecahkan permasalahan, tujuan serta manfaat dilakukannya penelitian, batasan permasalahan yang ada di dalam penelitian, serta urutan sistematika penulisan.

Bab II berisi tentang kajian pustaka yang memaparkan tentang materi serta algoritma yang dijadikan landasan teori di dalam penelitian ini. Di antaranya materi yang dibahas adalah algoritma *winnowing*, *synonym recognition* dan penelitian terkait.

Bab III membahas mengenai metode usulan berisi penjelasan tentang *dataset* yang akan digunakan, alur kerja sistem, tahapan pengujian dan skenario uji coba.

Bab IV membahas tentang implementasi serta pengujian berisi implementasi dari metode *winnowing* untuk memecahkan permasalahan yang kemudian dilakukan pengujian untuk mengukur hasilnya.

Bab V berisi kesimpulan dari implementasi yang sudah dilakukan, dan juga saran agar dapat mengembangkan penelitian berikutnya.