

BAB I

PENDAHULUAN

1.1. Latar Belakang

Mencontek merupakan perilaku yang sering dijumpai pada siswa yang ada di Indonesia baik pada tingkat sekolah dasar, menengah, hingga perguruan tinggi. Salah satu penelitian dengan subjeknya adalah 239 mahasiswa di Universitas Islam Nasional Walisongo, sebanyak 98% dari sampel penelitian menyatakan bahwa pernah mencontek (Hadjar, 2017). Mencontek biasanya dilakukan ketika mengerjakan pekerjaan rumah, ulangan harian, ujian tengah semester, ujian akhir semester dan sebagainya.

Menurut Anderman dan Murdock dalam (Cahyo & Solicha, 2017), menyontek dapat mengurangi fungsi dari penggunaan data asesmen sebagai indikator pencapaian belajar siswa dan juga sumber acuan bagi guru dalam melakukan tindakan serta pemberian *feedback*. Dari pernyataan sebelumnya, dapat diketahui bahwasanya hal-hal seperti pemberian tugas atau ujian merupakan suatu cara ataupun upaya guru agar siswa belajar dan mengevaluasi pembelajaran.

Evaluasi pembelajaran adalah suatu proses atau kegiatan yang sistematis, berkelanjutan dan menyeluruh dalam rangka pengendalian, penjaminan dan penetapan kualitas (nilai dan arti) pembelajaran terhadap berbagai komponen pembelajaran, berdasarkan pertimbangan dan kriteria tertentu, sebagai bentuk pertanggungjawaban guru dalam melaksanakan pembelajaran (Arifin, 2012). Setiap pembelajaran memerlukan evaluasi untuk mengetahui kadar pemilikan dan pemahaman didik terhadap materi pembelajaran, baik dalam aspek kognitif, psikomotorik, maupun afektif. (Ratnawulan & Rusdiana, Evaluasi Pembelajaran , 2014).

Untuk melakukan evaluasi pembelajaran, guru biasanya menggunakan alat evaluasi. Alat evaluasi adalah sesuatu yang digunakan dalam proses atau sistem evaluasi. Alat evaluasi juga dapat dikenal dengan istilah instrumen evaluasi. Banyak alat evaluasi yang digunakan dalam kegiatan evaluasi seperti, tes lisan, tes tulisan, observasi, dan wawancara. Secara garis besar, alat evaluasi yang dapat

digunakan digolongkan menjadi 2 macam, yaitu tes dan nontes. Tes merupakan suatu teknik atau cara yang digunakan dalam rangka melaksanakan kegiatan pengukuran, yang didalamnya terdapat berbagai pertanyaan, pernyataan, atau serangkaian tugas yang harus dikerjakan atau dijawab oleh peserta didik untuk mengukur aspek perilaku peserta didik. Sedangkan non tes dapat digunakan jika ingin mengetahui kualitas proses dan produk dari suatu pekerjaan serta hal-hal yang berkenaan dengan domain afektif, seperti sikap, minat, bakat, motivasi, dan lain-lain. (Arifin, 2012).

Tes tertulis merupakan salah satu dari teknik tes. Tes tertulis dapat dibedakan menjadi dua bentuk, yaitu tes objektif dan tes subjektif (tes uraian). Tes objektif adalah tes yang dalam pemeriksaannya dapat dilakukan secara objektif. Hal ini dimaksudkan untuk mengatasi kelemahan-kelemahan dari tes bentuk subjektif. Contoh dari tes objektif adalah tes benar-salah, tes pilihan ganda, menjodohkan, dan tes isian. Sedangkan tes subjektif yang pada umumnya berbentuk esai (uraian). Tes berbentuk esai adalah sejenis tes kemajuan belajar yang memerlukan jawaban yang bersifat pembahasan atau uraian kata-kata (Arikunto, 2013).

Tes tertulis sering digunakan untuk kegiatan evaluasi baik seperti ulangan harian, ujian tengah semester, ujian akhir semester ataupun ujian nasional. Saat ini, alat evaluasi untuk tes tertulis biasanya menggunakan media kertas. Namun sekarang sudah ada yang dilaksanakan dengan media berbasis digital ataupun menggunakan komputer. Contoh ujian berbentuk digital ini adalah Ujian Nasional Berbasis Komputer (UNBK). Di Kota Bandung, salah satu sekolah, yaitu Sekolah Menengah Kejuruan Pekerjaan Umum Negeri Kota Bandung sudah menggunakan atau memanfaatkan ujian berbasis digital atau teknologi yang disebut Ujian Dalam Jaringan (UDJ).

Pada umumnya, dalam ujian menggunakan media kertas, selalu ada tes subjektif. Namun, sistem atau aplikasi yang digunakan dalam UNBK ataupun UDJ, soal yang digunakan hanya berbentuk soal pilihan ganda. Dengan menggunakan teknologi, tes uraian dapat diimpelentasikan seiring dengan banyaknya pengembangan sistem untuk penilaian tes uraian. Tujuan dari pengembangan

penilaian tes uraian salah satunya adalah mengurangi subjektifitas penilaian yang tidak bisa dilakukan oleh penilaian oleh manusia.

Salah satu metode yang sering digunakan dalam penilaian uraian otomatis adalah *term frequency* dan *cosine similarity* yang sesuai dengan beberapa teori berikut. Cara yang biasa digunakan untuk mengecek kemiripan antara dua dokumen adalah dengan memeriksa kata-kata yang terkandung pada dua dokumen tersebut dengan cara menghitung frekuensi kata-kata pada dua dokumen tersebut. Salah satu metodenya adalah *term frequency* (TF). (Cavalcanti, Pires, Cavalcanti, & Pires, 2012) Hasil perhitungan frekuensi tersebut akan menghasilkan *vector space model* (VSM). *Similarity measurment* (pengukuran kemiripan) sering digunakan untuk *document retrieval* dan *information retrieval system* dengan menghitung vektor dari *vector space model* pada setiap dokumen dan *cosine similarity* adalah salah satu pengukuran untuk membandingkan kemiripan antara dua vektor. (Ali, et al., 2018).

Beberapa penelitian yang sejenis tentang penilaian uraian otomatis dalam bahasa Indonesia yaitu tentang aplikasi penilaian ujian uraian otomatis menggunakan metode *term frequency* dan *cosine similarity*. Dalam penelitian ini, digunakan *term frequency* (tf) sebagai metode untuk menentukan bobot kata yang didasarkan pada jumlah frekuensi kata dalam sebuah dokumen. Lalu digunakan metode *Cosine Similarity* untuk melakukan penilaian jawaban ujian esai secara otomatis. Aplikasi ini diujicoba pada ujian uraian dalam Bahasa Inggris. Hasil uji coba menunjukkan kesesuaian nilai sistem dengan nilai yang diberikan oleh pengajar adalah rata-rata 89,48%. (Fitri & Asyikin, 2015)

Selanjutnya, ada penelitian tentang penilaian ujian bertipe essay menggunakan metode *text similarity*. Dalam penelitian ini, digunakan algoritma Nazief-Adriani dalam proses *stemming*, lalu penilaiannya menggunakan metode *tf-idf*, serta *cosine similarity*. Adapun sistem ini digunakan dalam LMS *moodle* dan diujicoba dengan menggunakan berbagai sumber data soal mata pelajaran Geografi tingkat SMA. Hasil uji coba menunjukkan bahwa hasil penilaian sistem tidak berbeda signifikan dengan hasil penilaian dari pakar. Hal tersebut dibuktikan dengan pengujian menggunakan uji T berpasangan dengan derajat kebebasan (α) sebesar 5%, dimana nilai $t_{tabel} > t_{hitung}$ ($2.045 > 0.3755$) serta nilai signifikansi $>$ derajat

kebebasan ($0.71 > 0.05$) serta rata-rata penilaian dari nilai pakar sebesar 82.11667 dan rata-rata nilai sistem sebesar 81.88633. (Sulistyo, Saptono, & Asshidiq, 2015)

Selanjutnya, ada penelitian tentang perbandingan algoritma TF/IDF dan BLEU untuk penilaian jawaban esai otomatis. Dalam penelitian ini, metode penilaian jawaban esai otomatis menggunakan algoritma TF/IDF dan algoritma BLEU sebagai perbandingannya. Sistem ini diujicoba dengan data berupa 10 soal, kunci jawaban untuk masing-masing soal dan jawaban siswa dari 22 siswa yang masing-masing siswa mengerjakan 10 soal pada tes esai mata pelajaran Pendidikan Agama Islam. Hasil uji coba menunjukkan bahwa rata-rata korelasi TF/IDF-penilaian guru mencapai 0.70 dengan nilai korelasi tertinggi 0.98, sedangkan rata-rata korelasi BLEU-penilaian guru mencapai 0,63 dengan nilai korelasi tertinggi mencapai 0,97. Secara keseluruhan nilai yang dihasilkan algoritma TF/IDF lebih tinggi dari BLEU. (Nugroho, Pribadi, Arief, & Sukamta, 2014)

Selanjutnya, ada penelitian tentang penilaian otomatis terhadap jawaban esai pada soal berbahasa Indonesia menggunakan *Latent Semantic Analysis* (LSA). Dalam penelitian ini, digunakan metode LSA untuk menangani kesamaan katanya, sedangkan *Euclidean Distance* untuk mengukur kemiripan antara kunci jawaban dan jawaban siswa. Selain itu digunakan perhitungan suatu kata dengan menggunakan perhitungan TF-IDF. Sistem ini diujicobakan pada soal esai berbahasa Indonesia, khususnya beberapa mata pelajaran Produktif pada Program Keahlian TKJ Kelas X. Data yang digunakan terdiri dari 25 soal dengan jumlah responden sebanyak 25 siswa yang diambil secara acak dari 64 siswa kelas X TKJ. Output yang dihasilkan adalah berupa skor nilai akhir siswa yang berkisar antara angka 0-4. Hasil uji coba yang dilakukan menunjukkan sistem ini mampu melakukan proses penilaian secara otomatis terhadap jawaban esai berbahasa Indonesia dengan nilai korelasi sebesar 71%. (Yustiana, 2015)

Berdasarkan studi literatur diatas, banyak penelitian yang menunjukkan bahwasanya metode *term frequency* dan *cosine similarity* dapat digunakan sebagai cara penilaian terhadap soal berbentuk uraian. Namun peneliti akan mencoba mengimplementasikan *metode term frequency*, *n-gram*, dan *cosine similarity*. Penambahan *n-gram* digunakan untuk melihat atau memperhatikan bagaimana

Regi Ismayana Pratama, 2019

RANCANG BANGUN PENDETEKSI KEMIRIPAN JAWABAN ANTAR SISWA PADA SISTEM UJIAN DIGITAL BERBENTUK SOAL URAIAN

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

susunan kata pada suatu dokumen atau jawaban siswa. *N-gram* merupakan urutan dari n item dari urutan yang diberikan. Item tersebut dapat berupa fonem, suku kata, huruf, atau yang lain, yang sesuai dengan pengaplikasian. (Oduntan, Adeyanju, Olabiyisi, & Omidiora, 2015). Metode dengan penambahan *n-gram* akan diimplementasikan untuk mengetahui kemiripan jawaban antar siswa. Sehingga selain dapat digunakan untuk menilai jawaban siswa dari jawaban guru, sistem dapat memberitahukan guru bahwa ada jawaban siswa yang mirip dengan jawaban siswa lainnya.

Dengan melihat paparan sebelumnya, penulis berkeinginan untuk melakukan penelitian berjudul “Rancang Bangun Pendeteksi Kemiripan Jawaban Antar Siswa Pada Sistem Ujian Digital Berbentuk Soal Uraian”.

1.2. Rumusan Masalah

Dengan melihat latar belakang yang telah disampaikan sebelumnya, maka didapat beberapa rumusan masalah. Diantaranya:

- 1) Bagaimana rancang bangun sistem ujian digital berbentuk soal uraian?
- 2) Bagaimana tingkat akurasi penilaian jawaban siswa menggunakan metode *term frequency* dan *cosine similarity* dibandingkan dengan penilaian guru/manusia?
- 3) Bagaimana penerapan metode *term frequency*, *n-gram*, dan *cosine similarity* untuk menunjang penilaian kemiripan jawaban siswa?

1.3. Tujuan Penelitian

Berdasarkan rumusan masalah yang telah dipaparkan sebelumnya, maka tujuan penelitian ini adalah:

- 1) Mengetahui cara merancang dan membangun sistem ujian digital berbentuk soal uraian.
- 2) Mengetahui perbandingan tingkat akurasi penilaian jawaban antara sistem dan manusia menggunakan metode *term frequency* dan *cosine similarity*.
- 3) Mengetahui bagaimana penerapan metode *term frequency*, *n-gram*, dan *cosine similarity* dalam menunjang penilaian jawaban siswa terhadap menyontek.

1.4. Batasan Masalah

Batasan Masalah dalam penelitian ini adalah sebagai berikut:

- 1) Sistem Ujian hanya untuk soal tipe uraian.
- 2) Sistem Ujian yang dikembangkan hanya untuk soal uraian berbahasa Indonesia.
- 3) Sifat soal yang digunakan berupa teks dan tidak bersifat matematis, gambar, suara, ataupun video.
- 4) Soal yang digunakan harus memiliki kunci jawaban atau pertanyaan tanggapan terbatas.
- 5) Sistem hanya memberikan hasil kemiripan jawaban antar siswa saja, tidak sampai memberikan tindakan lanjut.

1.5. Manfaat Penelitian

Penelitian ini diharapkan mampu memberikan manfaat bagi semua pihak yang terkait, diantaranya sebagai berikut:

- 1) Peneliti
Mengetahui cara membuat sistem dan mengetahui bagaimana hasil dari dampak penggunaan sistem.
- 2) Siswa
Diharapkan adanya dampak positif bagi siswa terutama dalam kejujuran siswa pada ujian.
- 3) Guru
Diharapkan dapat terbantu dalam melaksanakan evaluasi pembelajaran.
- 4) Instansi Pendidikan
Diharapkan adanya pertimbangan menggunakan aplikasi yang digunakan dalam evaluasi pembelajaran.

1.6. Struktur Penulisan

Sistematika penulisan penelitian skripsi ini adalah:

BAB 1 PENDAHULUAN

Bab ini menjelaskan tentang latar belakang permasalahan sehingga ada penelitian ini, perumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah dan sistematika penulisan penelitian.

Regi Ismayana Pratama, 2019

RANCANG BANGUN PENDETEKSI KEMIRIPAN JAWABAN ANTAR SISWA PADA SISTEM UJIAN DIGITAL BERBENTUK SOAL URAIAN

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

BAB 2 KAJIAN PUSTAKA

Bab ini menjelaskan landasan teoritis terhadap topik atau permasalahan yang diangkat dalam penelitian. Teori yang dijelaskan dalam bab ini yaitu mengenai evaluasi pembelajaran, *text processing*, *term frequency*, *n-gram*, dan *cosine similarity*.

BAB 3 METODE PENELITIAN

Bab ini menjelaskan langkah-langkah penelitian yang akan dilakukan, dimulai dari desain penelitian, fokus penelitian, alat dan bahan yang digunakan untuk penelitian dan metode penelitian.

BAB 4 TEMUAN DAN PEMBAHASAN

Bab ini menjabarkan hasil penelitian yang telah dilakukan dan analisisnya. Semua pertanyaan mengenai masalah yang diangkat dalam tema skripsi dibahas di sini. Yaitu tentang proses pengumpulan data, pengembangan model, implementasi sistem, studi kasus, desain eksperimen, dan analisis.

BAB 5 SIMPULAN DAN SARAN

Bab ini menjelaskan tentang penyimpulan hasil penelitian dan pembahasan yang telah dilakukan serta pengusulan hal-hal yang dapat dimanfaatkan dari hasil penelitian dan rekomendasi bagi peneliti selanjutnya.