

BAB I

PENDAHULUAN

1.1 Latar Belakang

Pendidikan adalah bagian dari tumbuh kembang manusia yang berguna untuk merangsang pengetahuan, potensi, dan kemampuan yang dimiliki. Pelaksanaan pembelajaran pendidikan dilengkapi oleh proses evaluasi hasil yang disebut ulangan. Proses ini dilakukan dengan cara pemberian soal-soal yang telah dipelajari kepada peserta didik untuk memberikan jawaban yang sesuai. Setelah itu tenaga pendidik akan menilai berdasarkan keakuratan jawaban peserta didik.

Perkembangan teknologi saat ini memungkinkan kita untuk memberikan kemampuan memahami bahasa manusia kepada komputer untuk diolah secara lanjut. Keilmuan ini disebut Pengolahan Bahasa Alami atau *Natural Language Processing* yang merupakan cabang dari keilmuan *machine learning*. Elizabeth D. Liddy (2001)^[1] menuliskan dalam jurnalnya bahwa “*Natural Language Processing* adalah sekumpulan teknik komputasi yang didasari teori untuk menganalisis dan merepresentasikan teks alamiah pada satu atau lebih tingkatan analisis linguistik dengan tujuan mencapai kemampuan pengolahan bahasa setingkat manusia untuk berbagai tugas dan penerapan.”

Natural Language Processing dapat dimanfaatkan di bidang pendidikan sebagai alat bantu untuk tenaga pendidik dalam menjalankan tugasnya. Salah satunya adalah

sistem ulangan daring dengan fitur penilaian otomatis. Sistem ini telah diuji coba oleh penulis lain salah satunya pada penelitian milik Yohana Fithri (2018)^[2] yang menggunakan teknik *Latent Semantic Analysis* dan mendapatkan nilai keakuratan sebesar 83.3% ketika dibandingkan dengan penilaian dari tenaga pendidik yang terkait.

Selain teknik *Latent Semantic Analysis* terdapat juga teknik lain seperti *Word Embedding*. Arliyanti et al. (2020)^[3] dalam jurnalnya menjelaskan cara kerja *word embedding* yang berupa mengubah kata dari *corpus* menjadi bentuk *dense vector* yang merepresentasikan kata dalam ruang *vector*. *Word embedding* ini mampu menangkap makna semantik dan sintaktik dari sebuah kata.

Word2vec adalah salah satu algoritma dari *word embedding* yang dikembangkan oleh Tomas Mikolov dan tim risetnya pada tahun 2013. *Word2vec* utamanya menggunakan *continuous bag-of-words* dan *continuous skip-gram* untuk mengubah setiap kata ke dalam bentuk *vector* (Mikolov et al., 2013)^[4]. Algoritma ini termasuk ke dalam *unsupervised learning*, yaitu sistem akan menemukan sendiri pola dari sejumlah data yang kita berikan kepada sistem. Lalu sistem akan menyimpan hasil pembelajaran pola yang dilakukan untuk digunakan ke dalam aplikasi ke depannya dan disebut dengan *model*.

Setelah itu, pola yang tersimpan dalam *model* akan digunakan dengan konsep, algoritma, atau rumus lainnya. Untuk menghitung kemiripan kalimat digunakan metode *cosine similarity*. *Cosine similarity* adalah sebuah metode untuk mengukur kesamaan antar dua buah *vector* dalam sebuah ruang dimensi. *Cosine similarity*

memiliki nilai antara 0 dan 1, dimana nilai 1 menyatakan bahwa dua *vector* yang dibandingkan sama. (Dewa Ayu et al., 2016)

Dengan teori yang telah disebutkan diatas, memungkinkan untuk algoritma tersebut diimplementasikan ke dalam ranah pendidikan. Dengan membuat sebuah sistem ulangan daring dengan penilaian essay otomatis untuk menguji tingkat keakuratan algoritma *Word2vec* yang dibantu dengan *Smooth-Inverse Frequency* dalam menangkap arti kalimat. Serta pengujian *Cosine Similarity* dalam pengkalkulasian kemiripan dua kalimat yang telah diolah dengan *Word2vec + Smooth-Inverse Frequency*.

1.2 Rumusan Masalah

Dilihat dari observasi dan apa yang tertulis pada latar belakang diatas, masalah yang ingin diselesaikan pada penelitian ini adalah: Seberapa efektif *word2vec* dan *cosine similarity* dalam menghitung penilaian jawaban jika dibandingkan dengan penilaian tenaga pendidik?

1.3 Batasan Masalah

Dari masalah yang ada, dibuat batasan-batasan agar perancangan terarah dan dapat diaplikasikan dengan baik seperti :

1. Sistem ulangan daring merupakan aplikasi website.
2. Kata yang dapat diolah oleh *word2vec* adalah kata yang sudah dipelajari dan tersimpan dalam model.

3. Proses penghitungan kemiripan kalimat dilakukan menggunakan python yang dibuat menjadi web API.

1.4 Tujuan dan Manfaat

1.4.1 Tujuan

1. Mempermudah tugas tenaga pendidik dalam melaksanakan proses pelaksanaan ulangan dan juga pada proses penilaiannya.
2. Menguji coba metode *word2vec* dan *cosine similarity* untuk menghitung kemiripan antar dua kalimat jawaban.

1.4.2 Manfaat

1. Bagi penulis

Menerapkan pembelajaran teori yang sudah diterima selama proses perkuliahan pada sebuah sistem yang akan dibuat untuk meneliti keakuratan sebuah metode dan keefektifannya pada dunia nyata.

2. Bagi pembaca

Hasil penelitian ini dapat dijadikan sebagai referensi penelitian selanjutnya yang berkaitan dengan penerapan metode *word2vec* dan *cosine similarity*.

1.5 Metodologi Penulisan

1.5.1 Metode Pengumpulan Data

Metode yang digunakan untuk mengumpulkan data yang digunakan pada tugas akhir ini adalah:

a) Metode Wawancara

Merupakan sebuah metode untuk mendapatkan data dengan memberikan beberapa pertanyaan yang terkait dengan penelitian dan digunakan sebagai acuan dalam penelitian.

b) Metode Studi Pustaka

Studi pustaka adalah mengambil data-data melalui buku, jurnal, karya ilmiah dan karya tulis lain yang berkaitan dengan penelitian.

c) Metode Studi Dokumen

Merupakan metode untuk mendapatkan data dari dokumen-dokumen resmi berupa daftar nilai, catatan rapat, dan surat-surat resmi lainnya.

1.5.2 Metode Perancangan Sistem

Metode yang digunakan pada saat pengembangan sistem adalah *classic waterfall* atau yang disebut juga *linear sequential model / classic life cycle*. Pada metode ini satu fase dapat baru dapat dimulai setelah menyelesaikan fase sebelumnya. (Mei Prabowo, 2020: 37)

1.6 Sistematika Penulisan

Adanya sistematika penulisan ini ditujukan untuk mempermudah bagi pembaca untuk mencari isi yang diinginkan. Penulisan dipisah menjadi 5 bab yang dapat dijelaskan garis besarnya seperti berikut :

BAB 1 Pendahuluan

Berisikan tentang informasi umum dari karya tulis, seperti latar belakang, rumusan masalah, batasan masalah, tujuan dan manfaat, serta sistematika penulisan.

BAB 2 Landasan Teori

Berisikan tentang pengertian singkat dari teori-teori yang diimplementasikan kedalam rancangan sistem.

BAB 3 Analisis dan Rancangan Sistem

Berisikan tentang desain, rancangan, dan rangkaian serta pengimplementasian teori dari sistem yang dibuat.

BAB 4 Implementasi Hasil

Berisikan tentang data yang dihasilkan dari pengimplementasian sistem kedalam skenario nyata.

BAB 5 Kesimpulan dan Saran

Berisikan kesimpulan dari keseluruhan karya tulis serta saran dari penulis untuk pengembangan sistem lebih lanjut.

