

IMPLEMENTASI LIBRARY DEEP LEARNING KERAS PADA SISTEM UJIAN ESSAY ONLINE

Ekojono¹, Faisal Rahutomo², Dhiana Novita Sari³

^{1,2,3} Program Studi Teknik Informatika, Jurusan Teknologi Informasi, Politeknik Negeri Malang

¹ekojono2@gmail.com, ²faisal@polinema.ac.id, ³dhiana.novita17@gmail.com

Abstrak

Ujian Essay merupakan salah satu alat evaluasi dari sistem pembelajaran di instansi pendidikan. Ujian essay sangat penting dilakukan karena dapat melatih siswa dalam mengolah kemampuan berpikir kritis saat menjawab soal yang dipaparkan dalam beberapa kalimat. Dalam penerapannya, ujian ini memungkinkan hasil jawaban siswa lebih variatif dan mengharuskan para pengajar dalam bidang mata pelajaran tersebut untuk mengoreksi jawaban secara manual. Dari hal itu menimbulkan masalah jika mengoreksi secara manual membutuhkan waktu lama. Berdasarkan permasalahan tersebut maka diperlukan sebuah sistem otomatis yang dapat menilai jawaban essay siswa. Pengembangan tentang sistem penilaian essay otomatis masih terus dilakukan dengan tujuan untuk memperoleh nilai akurasi kemiripan antara jawaban dan kunci jawaban yang semakin baik. Pada penelitian ini menggunakan algoritma Siamese LSTM untuk menguji tingkat kemiripan jawaban dan kunci jawaban. Data yang digunakan berjumlah 2162 dari 4 kategori soal yaitu politik, lifestyle, teknologi dan olahraga dari penelitian tentang “Analisis Aspek-Aspek Ujian Esai Daring Berbahasa Indonesia”. Tiap kategori memiliki 10 soal dan dijawab kurang lebih oleh 50 siswa. Dari hasil uji coba dengan menggunakan algoritma Siamese LSTM didapatkan nilai error untuk kategori politik sebesar 61%. Untuk kategori lifestyle memiliki error percentage sebesar 112%. Kategori olahraga sebesar 162% dan untuk kategori teknologi mendapatkan error percentage sebesar 175%.

Kata kunci : Essay Online, Deep Learning, Siamese LSTM, Keras

1. Pendahuluan

Ujian merupakan salah satu bagian dari proses pembelajaran untuk mengevaluasi seberapa jauh tingkat pemahaman siswa dalam menguasai materi. Ujian biasanya dilakukan dalam beberapa cara yaitu dengan memberikan ujian pilihan ganda, ujian isian singkat atau ujian essay pada siswa. Di era teknologi saat ini telah banyak diterapkannya ujian berbasis komputer, hanya saja hal tersebut berlaku pada ujian pilihan ganda (multiple choice) dan untuk ujian seperti essay berbasis komputer (ujian essay online) masih belum diimplementasikan. Manfaat dari ujian essay sendiri yaitu dapat melatih daya pikir siswa dan meningkatkan kemampuan tata bahasa selama mengisi jawaban yang terdiri dari beberapa kalimat. Oleh karena itu ujian essay sangatlah penting dilakukan. Masalah yang dihadapi ketika menerapkan ujian essay yaitu proses pengoreksian manual yang cukup lama oleh guru apalagi tenaga guru yang jumlahnya lebih sedikit daripada siswa. Sehingga hal ini menghabiskan banyak tenaga dan waktu.

Salah satu cara yang harus dilakukan yaitu menerapkan adanya ujian essay online untuk mengotomatisasi pengoreksian jawaban. Saat ini

penelitian tentang essay online telah banyak dilakukan seperti Analisis Aspek-Aspek Ujian Esai Daring Berbahasa Indonesia yang menggunakan 3 algoritma yaitu Cosine Similarity menghasilkan nilai error sebesar 59.49%, Euclidean Distance dengan nilai error 332.90%, serta Jaccard dengan nilai error 52.31% [1]. Berdasarkan penelitian tersebut maka diperlukan suatu algoritma pembelajaran baru yang digunakan untuk menghitung tingkat keakurasian dalam menghitung kemiripan jawaban dan kunci jawaban seperti algoritma Deep learning yang saat ini sangat populer di kalangan data scientist. Algoritma yang akan digunakan adalah varian dari RNN (Recurrent Neural Network) yang merupakan algoritma yang bekerja pada data sequence seperti audio, video dan teks yaitu algoritma Siamese LSTM. Algoritma ini bekerja untuk menghitung tingkat kemiripan teks yang terdapat pada sebuah dokumen atau kalimat. Seperti penelitian tentang “Siamese Recurrent Architectures for Learning Sentence Similarity” yang menghasilkan nilai MSE atau error sebesar 0.2286 [2].

Dalam mengimplementasikan Siamese LSTM di system ujian essay online ini menggunakan library keras yang diterapkan pada bahasa pemrograman python untuk memperdalam pengetahuan tentang

Deep learning dan mengetahui tingkat error yang dihasilkan apakah kecil ataupun semakin besar [3].

2. Landasan Teori

Penelitian terkait tentang sistem ujian essay online menggunakan algoritma kemiripan Cosine Similarity, Jaccard dan Euclidean Distance yang dilakukan oleh Trisna Roshinta Ari dan Faisal Rahutomo mengenai analisis aspek-aspek ujian esai daring berbahasa Indonesia [1].

Penelitian terkait tentang kemiripan teks menggunakan Siamese Long Short Term Memory (Siamese LSTM) yang sudah dilakukan oleh J.Mueller mengenai pembelajaran kemiripan teks menggunakan arsitektur Siamese Network [2].

2.1 Ujian Essay

Ujian essay merupakan sebuah alat evaluasi dalam pendidikan untuk mengetahui seberapa mendalam tingkat pemahaman siswa dalam menyerap materi pada saat proses pembelajaran. Ujian essay pun dapat mengembangkam pemikiran kritis siswa terhadap persoalan yang akan dijawabnya melalui sebuah kalimat panjang [4]. Kelemahan dalam implementasi ujian essay adalah sulitnya dalam penentuan skor, karena tidak sembarang orang bisa melakukan penilaian terhadap jawaban essay, umumnya, jawaban yang essay lebih bervariasi sehingga sulit dalam mencocokkan dengan kunci jawaban. Bisa saja jawaban dari ujian essay tersebut benar, tetapi karena variasi jawaban, kata yang dituliskan bisa saja memiliki kesamaan arti dan makna, tetapi karena kunci jawaban berbeda penulisan kemungkinan jawaban essay tersebut tidak cocok sangat besar. Berbeda dengan ujian pilihan yang bisa dikoreksi oleh siapapun karena terdapat kunci jawaban yang sama persis. Disamping itu jumlah tenaga pengajar saat ini lebih sedikit dibandingkan dengan siswanya, sehingga untuk proses penilaian essay membutuhkan waktu dan tenaga yang besar.

2.2 Teks Mining

Text mining digunakan untuk menggali informasi pada dokumen yang tidak terstruktur [5]. Pada Machine learning untuk melakukan text mining diperlukan beberapa tahapan seperti ekstraksi fitur terhadap data input seperti text preprocessing yang meliputi Case folding, Tokenizing, Filtering dan Stemming. Setelah itu dilakukan proses pemodelan deep learning untuk melatih setiap dataset yang diolah untuk dipelajari pola dari data tersebut.

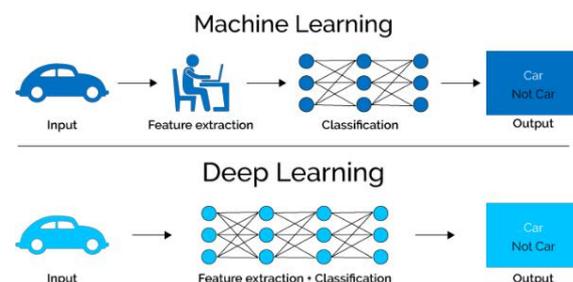
Case folding dan Tokenizing Case folding merupakan proses dalam text preprocessing untuk mengubah format masukan data teks menjadi huruf besar atau huruf kecil semua. Tetapi biasanya yang

sering dilakukan yaitu mengubah teks menjadi huruf kecil semua [6]. Tokenizing merupakan proses pemisahan kalimat menjadi penggalan kata. Biasanya dilakukan split atau pemisahan kata ketika terdapat spasi pada kalimat tersebut [7]. Pada proses tokenisasi secara garis besar merupakan proses untuk memecah sekumpulan karakter atau kata dalam suatu kalimat menjadi sebuah satuan kata. Karakter – karakter yang digunakan sebagai pemisah dari susunan kata dalam kalimat seperti whitespace, enter, tabulasi dan spasi. Untuk karakter seperti petik tunggal (‘), titik dua (:), titik (.) dan lainnya memiliki peran yang cukup banyak dalam memisahkan kata. Proses tokenisasi akan semakin rumit jika harus memperhatikan struktur bahasa.

Filtering dan Stemming. Filtering biasanya dilakukan untuk membuang kata-kata atau term yang memiliki tingkat kemunculan tinggi dan tidak memiliki makna khusus seperti kata “di”, “ke”, “dari”, “pada”, “yang”, “adalah” atau biasa disebut juga dengan stopword removal. Stemming merupakan proses mengubah kata berimbuhan, memiliki awalan dan akhiran agar menjadi bentuk kata dasar [8]. Pada penggunaan stoplist atau kata yang masuk dalam stopword tidak selalu meningkatkan nilai dari information retrieval menjadi semakin baik. Masih belum terdapat suatu kesimpulan bahwa penggunaan stopword removal dapat meningkatkan nilai information retrieval.

2.3 Deep Learning

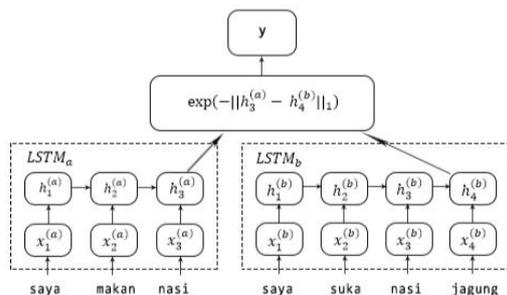
Deep learning merupakan sebuah revolusi dari pengembangan kecerdasan buatan yang memungkinkan untuk melakukan pembelajaran yang mendalam [9]. Konsep Deep learning sama seperti Neural network, hanya saja didalam deep learning, tidak diperlukan adanya deskripsi fitur atau biasa disebut dengan Feature extraction / Feature Selection. Perbedaan proses antara Neural network dengan Deep learning adalah dalam algoritma deep learning terdapat proses input, kemudian proses feature extraction sudah dilakukan secara bersamaan dengan proses klasifikasi untuk menghasilkan sebuah output. Deep learning dan machine learning dapat digambarkan dalam gambar 1.



Gambar 1. Perbedaan *deep learning* dan *machine learning*

2.4 Siamese LSTM

Arsitektur dari Siamese LSTM adalah untuk mempelajari kemiripan antar kalimat satu dengan yang lainnya. Tujuan umum dari Siamese LSTM adalah untuk embandingkan apakah dua kalimat memiliki kesamaan atau tidak [2][10][11].Arsitektur neural network tersebut memiliki dua neural network yang sama (shared weight) seperti LSTMa = LSTMb yang ditunjukkan pada gambar 2



Gambar 2. Pemodelan Siamese LSTM

Proses pertama yaitu mengubah kata menjadi representasi vector (embedding vector) kemudian menuju ke neural network. Dua representasi vector tersebut akan menuju ke sub-neural network (shared weight). Kemudian menghitung kemiripan antar dua data tersebut dengan fungsi berikut

$$(h^{(a)}, h^{(b)}) = \exp(-||h^{(a)} - h^{(b)}||) \in [0,1]. \quad (1)$$

Keterangan:

$h^{(a)}, h^{(b)}$: Representasi kemiringan gambar

Table 1. Spesifikasi iPhone 6s dan iPhone X

2.5 Word2Vec

Word2Vec merupakan sebuah algoritma word embedding yang digunakan untuk mengubah kata menjadi sebuah bentuk representasi vektor angka dengan jumlah dimensi tertentu. Pada penelitian ini, hasil dari pengubahan kata pada kunci jawaban dan jawaban siswa menjadi representasi vektor akan dijadikan input awal dalam proses training data dengan deep learning. Dalam merepresentasikan kata, word2vec menggunakan neural network untuk menghitung contextual and semantic similarity dari setiap kata input yang berbentuk one hot encoded vectors. Hasil dari kemiripan kontesktual dan semantic ini kemudian dapat merepresentasikan hubungan suatu kata dengan kata yang lainnya [12].

2.6 Library. Keras

Keras merupakan library deep learning jaringan saraf tiruan tingkat tinggi yang ditulis dengan Bahasa python dan mampu berjalan diatas

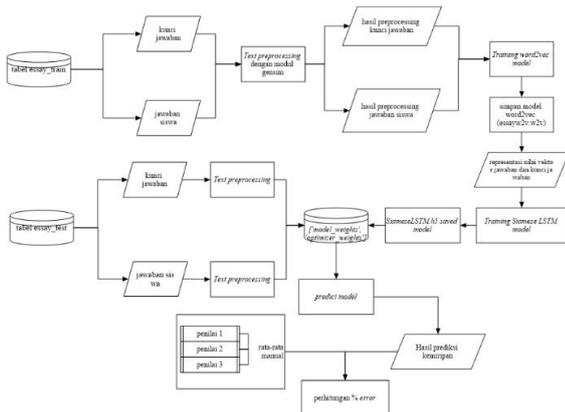
TensorFlow, CNTK dan Theano. Library ini menyediakan fitur yang digunakan dengan fokus mempermudah pengembangan lebih dalam tentang deep learning. Library ini dikembangkan untuk memungkinkan pengujian yang cepat pada CPU (Central Processing Unit) dan GPU (Graphic Processing Unit) serta mendukung algoritma convolutional neural network dan recurrent neural network ataupun kombinasi dari keduanya [13].

2.7 Error Rate

Kemiripan dari jawaban dan kunci jawaban akan dihitung nilai presentasi error-nya. Nilai presentasi error menunjukkan seberapa besar perbedaan antara system dengan manual [1]. Nilai error yang kecil menunjukkan bahwa penilaian semakin baik. Dihitung dengan persamaan 2 :

3. Implementasi Dan Pengujian

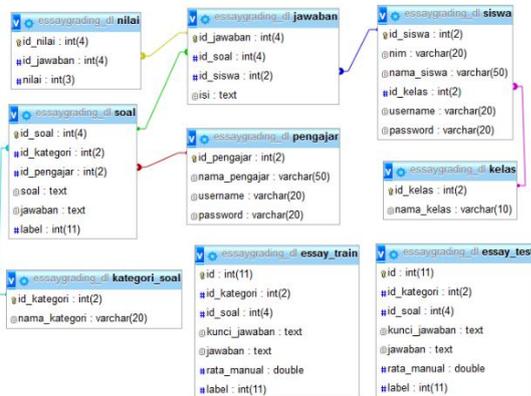
Dalam makalah ini penulis mengembangkan aplikasi dengan arsitektur sistem dari sistem ujian essay online yang akan peneliti buat, terdapat beberapa proses yaitu dimulai dari proses membaca data dari database, data berupa kunci jawaban dan jawaban siswa, setelah itu data tersebut dilakukan text preprocessing dan kemudian hasil preprocessing dijadikan kedalam sebuah list kedalam satu dokumen yang sama. Proses selanjutnya yaitu melatik dokumen kedalam word2vec model dan hasil model pelatihan word2vec disimpan pada di komputer dengan nama essayw2v.w2v dan setelah proses pelatihan word2vec model menghasilkan vektor jawaban dan kunci jawaban. Tujuan dari menyimpan model word2vec kedalam disk computer adalah supaya tidak memuat seluruh korpus (vocab) kedalam RAM sehingga proses menjadi lama. Kemudian hasil vector kunci jawaban dan jawaban siswa dijadikan sebagai data masukan pada pemodelan Siamese LSTM dan dilakukan proses training model Siamese LSTM. Pada saat proses pelatihan model selesai kemudian Siamese LSTM model disimpan kedalam komputer dengan nama SiameseLSTM.h5 isi dari file model SiameseLSTM.h5 merupakan bobot model dan bobot optimizer selanjutnya untuk proses pengujian yaitu sistem membaca data dari database berupa kunci jawaban, jawaban siswa dan label. Setelah itu dilakukan preprocessing dan memanggil kembali model Siamese LSTM yang tersimpan untuk melakukan prediksi kemiripan. Kemudian setelah selesai hasil prediksi kemiripan akan dihitung persentase error nya berdasarkan nilai rata-rata manula dari penilai 1, penilai 2 dan penilai 3.



Gambar 3. Arsitektur Sistem

A. Implementasi Database

Implementasi perancangan database pada sistem yang dibangun ini memiliki 9 tabel seperti pada gambar dibawah ini.



Gambar 4. Physical Data Model

Dalam database yang dirancang memiliki tabel – tabel sebagai berikut :

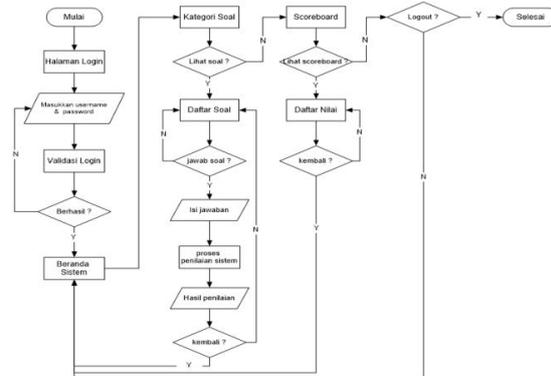
- tabel kategori_soal yang digunakan untuk menyimpan data kategori soal.
- tabel pengajar digunakan untuk menyimpan data pengajar pada database essay grading_dl.
- tabel soal yang digunakan untuk menyimpan data soal. tabel kelas yang digunakan untuk menyimpan data kelas
- tabel siswa yang digunakan untuk proses masuk ke sistem menggunakan data username dan password
- tabel jawaban yang digunakan untuk menyimpan data jawaban terutama pada kolom isi.
- tabel nilai yang digunakan untuk menyimpan data nilai setelah proses isi jawaban oleh metode Siamese LSTM berhasil dan otomatis menambahkan data nilai ke dalam tabel nilai.

- tabel essay_train yang digunakan untuk menyimpan data yang akan diproses dalam proses training dan validating dengan metode Siamese LSTM.
- tabel essay_test yang digunakan untuk menyimpan data yang akan diproses dalam proses testing dengan metode Siamese LSTM.

B. Implementasi Diagram

1) Flowchart sistem

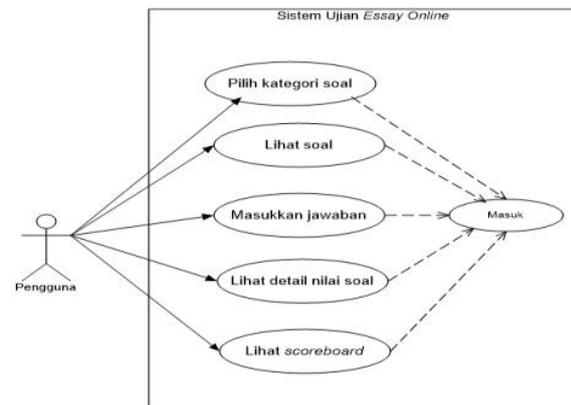
Flowchart atau diagram alir sistem merupakan model statis yang menggambarkan alur dari jalannya sistem mulai dari awal hingga akhir. Tujuan dari pembuatan sistem ujian essay online ini adalah untuk mendapatkan hasil prediksi dengan nilai error terkecil. Proses prediksi kemiripan antara kunci jawaban dengan jawaban siswa menggunakan metode Siamese LSTM ini menggunakan data masukan yaitu data jawaban siswa yang sudah melalui proses text-preprocessing terlebih dahulu. Berikut diagram alir untuk keseluruhan proses prediksi kemiripan kunci jawaban dengan jawaban siswa menggunakan Deep Learning model Siamese LSTM berbasis Keras. Diagram alir yang dibangun seperti Gambar 5.



Gambar 5. Flowchart System

2) Use Case System

Use case adalah gambaran fungsionalitas dari suatu sistem, sehingga pengguna sistem dapat memahami dan



Gambar 6. Use Case System

Pada gambar 6 terdapat aktor yaitu pengguna pertama kali harus masuk kedalam sistem dengan memasukkan username dan password. Kemudian ketika berhasil masuk ke sistem, pengguna dapat melihat halaman beranda yang terdapat menu kategori soal. Pengguna dapat klik menu kategori soal untuk melihat detail soal. Setelah itu pengguna dapat menjawab soal sesuai id soal yang dipilih. Ketika sudah berada pada halaman jawab soal, pengguna dapat mengisi jawaban dan klik tombol simpan untuk memproses jawaban tersebut agar mendapatkan nilai. Ketika jawaban berhasil diproses maka pengguna mendapatkan nilai dari hasil isian jawabannya yaitu dengan melihat detail nilai per soal atau dengan melihat detail keseluruhan nilai.

C. Implementasi Sistem

Implementasi sistem pada sistem ujian essay online ini yaitu dengan menampilkan halaman form jawab soal seperti pada gambar 7.



Gambar 7. Halaman input jawaban soal

Hasil dari proses penilaian essay dari sistem akan menghasilkan tampilan seperti gambar 8.



Gambar 8. Halaman Nilai

4. Hasil Dan Pembahasan

A. Analisis Penilaian Sistem

Analisis penilaian sistem dilakukan untuk mengetahui nilai dari hasil perhitungan dengan sistem menggunakan model deep learning. Pada analisis penilaian sistem ini diambil contoh soal yang memiliki jawaban pasti dan memiliki jawaban siswa yang isinya sama serta memiliki tambahan karakter yang berbeda seperti penambahan angka untuk penomoran di awal jawaban. Hasil penilaian sistem terdapat pada gambar 9 dibawah ini.

Soal	Kunci jawaban	Jawaban Siswa	Nilai
Sebutkan 5 butir sila Pancasila.	4. Ketuhanan Yang Maha Esa, - kemanusiaan yang adil dan beradab, - persatuan Indonesia, - kerakyatan yang dipimpin oleh hikmat kebijaksanaan dalam permusyawaratan perwakilan, - keadilan sosial bagi seluruh rakyat Indonesia	ketuhanan yang maha esa, kemanusiaan yang adil dan beradab, persatuan Indonesia kerakyatan yang dipimpin oleh hikmat kebijaksanaan dalam permusyawaratan dan perwakilan, keadilan sosial bagi seluruh rakyat Indonesia	94
		1.Ketuhanan Yang Maha Esa 2.Kemanusiaan yang adil dan beradab. 3.Persatuan Indonesia 4.Kerakyatan yang dipimpin oleh hikmat kebijaksanaan dalam permusyawaratan perwakilan 5.Keadilan sosial bagi seluruh rakyat Indonesia.	100
		1)ketuhanan yang maha esa 2)kemanusiaan yang adil dan beradab 3)persatuan indonesia 4)kerakyatan yang dipimpin oleh hikmat kebijaksanaan dalam permusyawaratan perwakilan 5)keadilan sosial bagi seluruh rakyat Indonesia	96

Gambar 9. Hasil Penilaian Sistem

Dari hasil penilaian sistem sudah dilakukan, diambil dari soal “Sebutkan 5 butir sila Pancasila.” Yang memiliki kunci jawaban “- Ketuhanan Yang Maha Esa, - kemanusiaan yang adil dan beradab, - persatuan Indonesia, - kerakyatan yang dipimpin oleh hikmat kebijaksanaan dalam permusyawaratan perwakilan, - keadilan sosial bagi seluruh rakyat Indonesia” yang merupakan kategori soal jawaban pasti dan telah diujikan dengan data jawaban siswa baru. hasil skor jawaban siswa tersebut menunjukkan angka yang berbeda pada nomor ke 1 nilai yang dihasilkan adalah 94 oleh sistem, untuk nomor ke 2 nilai yang dihasilkan adalah 100 oleh sistem, dan nomor ke 3 nilai yang dihasilkan adalah 96 oleh sistem dari semua jawaban tersebut merupakan jawaban yang benar berdasarkan penilaian secara manual oleh manusia. Yang membedakan adalah pada nomor ke 1 stuktur penulisan jawaban tidak terdapat angka penomoran didepannya, untuk nomor 2 terdapat penomoran didepan jawaban dan untuk yang ke 3 terdapat angka penomoran yang diikuti kurang akhir. Hal ini dapat terjadi karena pola data pelatihan dari training data jawaban siswa lebih banyak melatih data yang memiliki penomoran angka dengan titik didepan jawaban pada soal “Sebutkan 5 butir sila Pancasila.” daripada struktur penulisan data seperti nomor 1 dan nomor 3. Hal terpenting dalam melakukan prediksi dengan model deep learning adalah, semakin banyak pola data yang dilatih maka hasil prediksi akan semakin baik.

B. Analisis Akurasi Sistem

Pengujian akurasi sistem dilakukan untuk mengukur tingkat keakurasian dari hasil perhitungan nilai pada sistem. Penrhitungan keakurasian dilakukan dengan menghitung persentase error (error percentage) yang merupakan perbandingan antara rata nilai manual oleh manusia dengan penilaian sistem. Data yang digunakan merupakan data dari penelitian sebelumnya tentang esai daring berbahasa

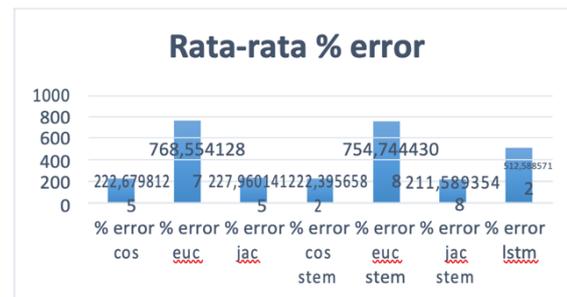
Indonesia yang mana terdapat 2162 data jawaban siswa namun yang akan diuji hanya 400 data jawaban siswa, karena dari data 2162 dibagi menjadi data training dan data testing dengan menggunakan metode Siamese LSTM.

Pengujian keakurasian pada penelitian ini dilakukan dengan menghitung nilai kemiripan antara kunci jawaban dengan jawaban siswa, kemudian hasil dari nilai tersebut dihitung persentasenya. Jika nilai error semakin kecil maka penilaian sistem akan semakin akurat. Setelah persentase error didapatkan, kemudian membandingkannya dengan metode – metode yang digunakan pada penelitian terdahulu yang menggunakan metode pembobotan Term Frequency – Inverse Document Frequency (TFIDF). Tujuan dari penelitian ini yang menggunakan metode LSTM yaitu untuk mengetahui apakah tingkat keakurasian kemiripan menggunakan LSTM semakin baik atau sebaliknya. Pada penelitian terdahulu menggunakan metode kemiripan vector yaitu Cosine Similarity, Euclidean Distance, Jaccard, Cosine Stemming, Euclidean Stemming, dan Jaccard Stemming. Hasil persentase error dari metode tersebut akan dihitung dan dilakukan perhitungan peringkat yang mana, jika nilai error terkecil akan mendapatkan peringkat pertama. Berikut merupakan detail dari perhitungan keakurasian sistem dari penelitian ini dan penelitian terdahulu tentang essay online.

Kategori	Rata-rata % error Cos	Rata-rata % error euc	Rata-rata % error jac	Rata-rata % error cos stem	Rata-rata % error euc stem	Rata-rata % error jac stem	Rata-rata % error LSTM
Politik	51,5 6185	101,65	53,4 4274	53,5 7117	101,1446	46,2 8487	61,7 9948 249
Lifestyle	87,8 3905	270,4502	64,4 4825	88,5 7796	267,3701	60,6 2417	112,8165
Olahraga	37,8 1563	262,1049	56,7 2716	37,3 1524	254,5963	53,6 4608	162,2908
Teknologi	45,4 6328	134,349	53,3 4199	42,9 3129	131,6334	51,0 3424	175,6818
Rata-rata	222,6798	768,5541	227,9601	222,3957	754,7444	211,5894	512,5886
Peringkat	3	7	4	2	6	1	5

Pada tabel 1 berisi 8 kolom yaitu kolom kategori yang merupakan kategori dari soal yang digunakan. Untuk % error cosine merupakan kolom dari persentase error algoritma cosine. Untuk % error euclidean merupakan kolom dari persentase error algoritma euclidean. Untuk % error jaccard merupakan kolom dari persentase error algoritma jaccard. Untuk % error cosine stemming merupakan kolom dari persentase error algoritma cosine stemming. Untuk % error Euclidean stemming merupakan kolom dari persentase error algoritma Euclidean stemming. Untuk % error jaccard stemming merupakan kolom dari persentase error

algoritma jaccard stemming. Untuk % error Siamese LSTM merupakan kolom dari persentase error algoritma Siamese LSTM. Untuk baris rata – rata merupakan jumlah rata – rata error dari tiap metode dan untuk baris peringkat merupakan hasil peringkat dari metode dengan persentase error terkecil yang mendapatkan peringkat teratas. Akurasi pengujian kemiripan teks jawaban siswa dan kunci jawaban untuk semua kategori soal pada metode LSTM menempati posisi ke-5 dengan tingkat nilai rata – rata percentage error 512%. Untuk hasil akurasi terbaik yaitu dari penelitian terdahulu pada metode Jaccard Stemming dengan nilai error 211%.



Pada gambar 10 menunjukkan grafik akurasi pengujian kemiripan teks jawaban siswa dan kunci jawaban untuk semua kategori soal pada metode LSTM menempati posisi ke-5 dengan tingkat nilai rata – rata percentage error 512%, untuk metode cosine tanpa stemming memiliki nilai error 222,67%, Euclidean 768%, jaccard 227%, cosine stemming 222,3%, Euclidean stemming 754% dan untuk hasil akurasi terbaik yaitu dari penelitian terdahulu pada metode Jaccard Stemming dengan nilai error 211%.

5. Kesimpulan dan Saran

Berdasarkan hasil pengujian dan pembahasan dalam menghitung tingkat kemiripan kunci jawaban dengan jawaban siswa dalam sistem ujian essay online menggunakan metode Siamese LSTM maka hasil pengujian menghasilkan nilai akurasi pada kategori soal Politik, Lifestyle, Olahraga yang menempati peringkat ke 5, sedangkan pada kategori soal Teknologi berada pada peringkat ke 7 dari total 7 penerapan metode dan penelitian terdahulu dan penelitian menggunakan metode Siamese LSTM. 4. Nilai akurasi dengan metode Siamese LSTM pada seluruh kategori soal mencapai nilai error sebesar 512 %. Sehingga metode ini tingkat keakurasiannya masih kurang baik dari metode penelitian tentang Analisis Aspek-Aspek Ujian Esai Daring berbahasa Indonesia. Untuk metode cosine tanpa stemming memiliki nilai error 222,67%, Euclidean 768%, jaccard 227%, cosine stemming 222,3%, Euclidean stemming 754% dan untuk hasil akurasi terbaik yaitu pada metode Jaccard Stemming dengan nilai error 211%.

Daftar Pustaka

- T. Roshinta Ari and F. Rahutomo. (2016). Analisis Aspek-Aspek Ujian Esai daring Berbahasa Indonesia. Vol. 54, no. 8, pp 543–547.
- J. Mueller. (2015). Siamese Recurrent Architectures for Learning Sentence Similarity. No. 2014, pp 2786–2792.
- G. Santoso, Aditya and Ariyanto.(2018). Implementasi Deep Learning Berbasis Keras Untuk Pengenalan Wajah. *Emit. J. Tek. Elektro*, vol. 18, no. 01, pp. 15–21.
- Y. Harika *et al.*(2017). AUTOMATED ESSAY GRADING USING FEATURES SELECTION. *Int. Res. J. Eng. Technol.*
- F. Rahutomo, P. Y. Saputra, C. Febriawan, and P. Putra. (2018). Implementasi Explicit Semantic Analysis Berbahasa Indonesia Menggunakan Corpus Wikipedia Indonesia. *J. Inform. Polinema*, vol. 4, no. 4, pp. 252–257.
- I. F. Rozi, E. N. Hamdana, M. Balya, and I. Alfahmi. (2018). Pengembangan Aplikasi Analisis Sentimen Twitter Menggunakan Metode Naïve Bayes Classifier (Studi Kasus Samsat Kota Malang). *J. Inform. Polinema*, pp. 149–154.
- M. F. Kurniawan, F. Rahutomo, and R. Rismanto. (2015). Analisa Komputasi Kemunculan Dan Kepunahan Kosakata Bahasa Indonesia Berdasarkan Corpus. *J. Inform. Polinema*, vol. 3, pp. 12–16.
- I. T. Hapsari, B. S. Andoko, and C. Rahmad. (2015). Aplikasi Information Retrieval untuk Pencarian Dokumen Laporan Penelitian. *J. Inform. Polinema*, vol. 1, no. 3, pp 23–28.
- A. Ahmad. (2017). Mengenal Artificial Intelligence, Machine Learning, Neural Network, dan Deep Learning. *J. Teknol. Indones.*.
- P. Neculoiu, M. Versteegh, M. Rotaru, and T. B. V Amsterdam.(2016). Learning Text Similarity with Siamese Recurrent Networks, July.
- C. Baziotis, N. Pelekis, and C. Doukeridis. (2017). DataStories at SemEval-2017 Task 6: Siamese LSTM with Attention for Humorous Text Comparison,” pp. 390–395.
- C. McCormick. (2016). Word2Vec Tutorial - The Skip-Gram Model. *Chris McCormick's Blog*. .
- F. Chollet. (2017). Keras 2015. URL [http//keras.io](http://keras.io).