

Klasifikasi Kategori Feedback EDOM Primakara University dengan Algoritma RNN LSTM

Rizky Aditya Ichwanto^{1*}, Ketut Queena Fredlina², I Gede Juliana Eka Putra³

^{1,2,3} Program Studi Pendidikan Informatika, Fakultas Teknologi Informasi dan Desain, Universitas Primakara,
Indonesia

*Corresponding Author e-mail: raditya2678@gmail.com

Abstract: This research aims to apply the Long Short-Term Memory (LSTM) algorithm in sentiment analysis of the Student Evaluation of Teaching (EDOM) at Primakara University, with the goal of improving higher education quality through the evaluation of faculty performance. Primary data was collected through interviews with the Quality Assurance Agency, while secondary data was obtained from EDOM for the academic years 2020/2021 to 2022/2023. A sentiment classification model was constructed using LSTM, initially dividing the data into 20 categories. To balance the data distribution, these categories were then merged into 6 main categories. The model was trained and tested using cross-validation, achieving an accuracy of 97%. However, when the model was tested with 100 new EDOM data points, the accuracy decreased to 74%, which is suspected to be caused by the emergence of new vocabulary that was not recognized or stored in the trained machine learning model. This decline in accuracy indicates the limitations of the model in handling new EDOM data that differs from the training data, and highlights the importance of periodic model updates or the use of out-of-vocabulary techniques to improve model performance in the future.

Key Words: Machine Learning, Deep Learning, LSTM, Classification, EDOM.

Abstrak: Penelitian ini bertujuan untuk menerapkan algoritma Long Short-Term Memory (LSTM) dalam analisis sentimen Student Evaluation of Teaching (EDOM) di Universitas Primakara, dengan tujuan untuk meningkatkan mutu pendidikan tinggi melalui evaluasi kinerja dosen. Data primer dikumpulkan melalui wawancara dengan Badan Penjaminan Mutu, sedangkan data sekunder diperoleh dari EDOM tahun akademik 2020/2021 sampai dengan 2022/2023. Model klasifikasi sentimen dibangun menggunakan LSTM, awalnya membagi data menjadi 20 kategori. Untuk menyeimbangkan distribusi data, kategori-kategori tersebut kemudian digabung menjadi 6 kategori utama. Model dilatih dan diuji menggunakan cross-validation, mencapai akurasi sebesar 97%. Namun, ketika model diuji dengan 100 titik data EDOM baru, akurasinya menurun hingga 74%, yang diduga disebabkan oleh munculnya kosakata baru yang tidak dikenali atau tersimpan dalam model machine learning yang dilatih. Penurunan akurasi ini menunjukkan keterbatasan model dalam menangani data EDOM baru yang berbeda dari data pelatihan, dan menyoroti pentingnya pembaruan model berkala atau penggunaan teknik yang tidak lazim untuk meningkatkan kinerja model di masa mendatang.

Kata Kunci: Machine Learning, Deep Learning, LSTM, Klasifikasi, EDOM.

Pendahuluan

Pendidikan merupakan hak setiap individu yang harus dipenuhi. Hal ini tertuang dalam UUD 1945 pasal 31 ayat 1 yang menyatakan bahwa setiap warga negara berhak mendapatkan pendidikan. Begitu pentingnya pendidikan bagi masyarakat, membuat pemerintah terus gencar memajukan sistem pendidikan demi mencapai generasi emas yang diharapkan dapat membantu pembangunan infrastruktur Indonesia dengan mewajibkan 12 tahun mengenyam bangku Pendidikan. Saat ini, dunia pendidikan dihadapkan pada kondisi di mana harus mempersiapkan lulusan yang memiliki kompetensi atau kapabilitas untuk bersaing dalam masa kini dan masa depan (Kim, Raza, & Seidman, 2019). Oleh karena itu meningkatkan kualitas pendidikan sangat diperlukan, terutama di perguruan tinggi. Sesuai dengan ketentuan yang tertuang dalam Undang-Undang Nomor 20 Tahun 2003 dan Peraturan Pemerintah Nomor 19 Tahun 2005 tentang Standar Nasional Pendidikan, “setiap penyelenggara pendidikan berkewajiban menjamin mutu”, salah satunya meningkatkan mutu pelayanan dalam hal kinerja dosen.

Kinerja adalah evaluasi atas prestasi kerja pribadi seorang karyawan yang menggambarkan atau secara sistematis menggambarkan kelebihan dan kelemahan yang terkait dengan pekerjaannya (Sevsa, 2018). Kinerja tenaga menjadi faktor penting atas kualitas



pendidikan di dalam lembaga pendidikan. Sehingga kualitas pendidikan bergantung pada ketersediaan tenaga pendidik untuk saling bekerja sama dan mendukung dalam melaksanakan tugas sesuai bidangnya (Sanda, Warman, Pitriyani, & Yesepa, 2022).

Primakara University sebagai lembaga pendidikan telah berusaha untuk memberikan pelayanan dan kualitas pendidikan yang terbaik bagi mahasiswa, salah satu diantaranya adalah menerapkan sistem EDOM (Evaluasi Dosen Oleh Mahasiswa) untuk menilai kinerja atau kualitas pembelajaran yang dilakukan dosen selama satu semester dan dilakukan setiap akhir semester (Direktorat Pengembangan dan Pelayanan Sistem Informasi UI, 2023). Tujuan dari EDOM itu sendiri adalah untuk memperoleh informasi dari mahasiswa tentang kinerja dosen selama proses mengajar (Direktorat Pengembangan dan Pelayanan Sistem Informasi UI, 2023). Dengan adanya EDOM, pihak kampus dapat melihat bagaimana kinerja dosen selama proses mengajar dari perspektif mahasiswa. Ditambah lagi bahwa selesainya EDOM berarti siswa telah berpartisipasi untuk membantu meningkatkan kualitas pembelajaran. EDOM berguna bagi guru untuk menilai sendiri apakah masih terdapat kesenjangan dan juga untuk mengembangkan potensi yang dimiliki. Bagi pimpinan perguruan tinggi, fakultas, dan jurusan (program studi), hasil EDOM dapat dijadikan acuan dalam pengembangan program yang bertujuan untuk meningkatkan kualitas proses pengajaran dan kinerja guru. Dan yang terpenting, bagi siswa, mereka dapat memperoleh manfaat dari peningkatan kualitas proses pembelajaran yang berkelanjutan. Namun, menganalisis data yang sangat banyak dalam waktu singkat merupakan tantangan yang dihadapi saat ini.

Maka dari itu, penelitian ini dilakukan untuk memudahkan dalam menganalisis EDOM dengan cara memanfaatkan machine learning untuk melakukan klasifikasi opini mahasiswa secara otomatis dengan algoritma long-short term memory (LSTM). Algoritma long-short term memory (LSTM) merupakan salah satu pengembangan lebih lanjut dari algoritma recurrent neural network (RNN). Algoritma ini mampu mengolah data input secara berulang, dan biasanya data yang digunakan berupa data time series atau data yang berurutan (Arinata, Satwika, & Utami, 2022).

Adapun penelitian terkait, sudah dilakukan oleh Arinata, I. G. B., dkk. (Arinata, Satwika, & Utami, 2022) yaitu menggunakan sentimen analisis untuk mengetahui opini mahasiswa yang diklasifikasikan kedalam 3 kategori seperti positif, negatif, dan netral. Namun penelitian ini akan melakukan klasifikasi feedback mahasiswa yang terdapat pada EDOM (Evaluasi Dosen Oleh Mahasiswa) ke dalam beberapa kategori yang terkait dengan aspek - aspek pembelajaran seperti, kualitas mengajar, kedisiplinan dosen, kualitas materi yang dibawakan, dan lain - lain. Untuk data awal yang digunakan yaitu rekapan data EDOM Primakara University dari tahun ajaran 2020/2021 sampai 2022/2023 dan dilabeli kategori satu per satu untuk data latih machine learning. Pada penelitian yang lain oleh Rusli, M., dkk. (Rusli, 2020), diketahui bahwa data kuesioner mahasiswa dilabeli secara otomatis melalui kata kunci dengan metode ekstraksi fitur. Namun hal tersebut kurang efektif dikarenakan pada simpulan (Rusli, 2020), kata kunci tersebut terdiri dari satu kata saja dan tidak bisa menggabung beberapa kata menjadi satu makna.

Dengan menggunakan algoritma LSTM, model machine learning dapat melakukan klasifikasi dengan cara memahami makna setiap kata dikarenakan dalam suatu kalimat, berisi kata - kata yang berurutan, dan ini sesuai dengan algoritma yang dipakai pada penelitian ini.

Adapun penelitian terkait, sudah dilakukan oleh Arinata, I. G. B., dkk. (Arinata, Satwika, & Utami, 2022) yaitu menggunakan sentimen analisis untuk mengetahui opini mahasiswa yang diklasifikasikan kedalam 3 kategori seperti positif, negatif, dan netral. Namun penelitian ini

akan melakukan klasifikasi *feedback* mahasiswa yang terdapat pada EDOM (Evaluasi Dosen Oleh Mahasiswa) ke dalam beberapa kategori yang terkait dengan aspek - aspek pembelajaran seperti, kualitas mengajar, kedisiplinan dosen, kualitas materi yang dibawakan, dan lain - lain. Untuk data awal yang digunakan yaitu rekapan data EDOM Primakara University dari tahun ajaran 2020/2021 sampai 2022/2023 dan dilabeli kategori satu per satu untuk data latih machine learning. Pada penelitian yang lain oleh Rusli, M., dkk. (Rusli, 2020), diketahui bahwa data kuesioner mahasiswa dilabeli secara otomatis melalui kata kunci dengan metode ekstraksi fitur. Namun hal tersebut kurang efektif dikarenakan pada simpulan (Rusli, 2020), kata kunci tersebut terdiri dari satu kata saja dan tidak bisa menggabungkan beberapa kata menjadi satu makna.

Dengan menggunakan algoritma LSTM, model *machine learning* dapat melakukan klasifikasi dengan cara memahami makna setiap kata dikarenakan dalam suatu kalimat, berisi kata - kata yang berurutan, dan ini sesuai dengan algoritma yang dipakai pada penelitian ini.

Metode Penelitian

Penelitian ini menggunakan metode supervised learning dengan pendekatan Long Short-Term Memory (LSTM) untuk melakukan klasifikasi opini mahasiswa dalam Evaluasi Dosen oleh Mahasiswa (EDOM). Metode ini dipilih karena kemampuannya dalam memproses data teks sekuensial dan mempertahankan hubungan semantik dalam suatu kalimat. Proses penelitian ini mencakup beberapa tahapan utama, yaitu pengumpulan data, pengolahan dan anotasi data, pembangunan arsitektur model, pelatihan model, evaluasi model, serta implementasi program.

Dalam penelitian ini, alat yang digunakan terdiri dari perangkat keras dan perangkat lunak. Perangkat keras yang digunakan meliputi laptop merk Asus dengan RAM 8GB DDR4, penyimpanan 1 TB HDD serta 240 GB SSD, kartu grafis Nvidia GeForce MX110, dan prosesor Intel Core i3 7280U (7th Gen). Sementara itu, perangkat lunak yang digunakan dalam penelitian ini adalah Windows 11 Professional 64 bit, Google Chrome, Google Spreadsheet, dan Google Colab sebagai platform untuk membangun, melatih, serta menguji model machine learning berbasis TensorFlow dan Keras.

Sumber data yang digunakan dalam penelitian ini terdiri dari data primer dan data sekunder. Data primer diperoleh dari wawancara tidak langsung dengan Lembaga Penjaminan Mutu di Primakara University, yang memberikan data EDOM dari tahun ajaran 2021/2022 dengan 1.027 data komentar yang terbagi dalam 76 label, yang kemudian dikategorikan menjadi 22 kategori positif dan 54 kategori negatif. Data sekunder diperoleh dari rekapan data EDOM tahun ajaran 2020/2021 hingga 2022/2023, dengan total 3.100 data mentah yang digunakan sebagai sumber data tambahan untuk pelatihan dan evaluasi model.

Jenis data yang digunakan dalam penelitian ini terdiri dari data kuantitatif dan data kualitatif. Data kuantitatif berupa statistik dari jumlah komentar yang diklasifikasikan ke dalam berbagai kategori sentimen, sementara data kualitatif berupa opini atau umpan balik mahasiswa yang akan digunakan sebagai input dalam model klasifikasi.

Proses penelitian dimulai dengan identifikasi masalah, yang dilakukan melalui wawancara dengan Lembaga Penjaminan Mutu untuk memahami tantangan dalam analisis data EDOM yang masih dilakukan secara manual. Setelah itu, dilakukan studi literatur guna memahami berbagai pendekatan yang telah digunakan dalam penelitian sebelumnya, khususnya terkait penerapan LSTM dalam klasifikasi opini berbasis teks. Selanjutnya, pengumpulan data dilakukan dengan mengambil data EDOM dari tiga tahun ajaran terakhir, yang kemudian diproses lebih lanjut dalam tahap pengolahan dan anotasi data.

Pada tahap pengolahan data, beberapa teknik preprocessing teks diterapkan agar data siap digunakan dalam model machine learning. Proses ini meliputi cleaning text untuk menghapus tanda baca dan angka, case folding untuk mengubah teks menjadi huruf kecil, tokenizing untuk memecah kalimat menjadi kata-kata, stemming untuk mengubah kata berimbuhan menjadi bentuk dasar, serta encoding untuk mengonversi kata-kata ke dalam bentuk numerik. Setiap komentar kemudian diklasifikasikan ke dalam 20 kategori awal, yang kemudian dikonsolidasikan menjadi 10 kategori positif (P1-P10) dan 10 kategori negatif (N1-N10) agar model dapat menangkap aspek pembelajaran secara lebih spesifik.

Dalam proses anotasi data, validasi dilakukan oleh Lembaga Penjaminan Mutu guna memastikan bahwa setiap komentar dikategorikan secara objektif. Jika ditemukan bahwa suatu kategori memiliki jumlah data yang kurang dari 50 data feedback, kategori tersebut akan digabungkan dengan kategori lain agar distribusi data tetap merata. Sebaliknya, jika kategori “dan lain-lain” (P10 dan N10) memiliki jumlah feedback yang berlebihan (>300), maka dilakukan analisis ulang untuk menentukan apakah perlu dibuat kategori baru.

Setelah data siap, dilakukan tahap pembangunan arsitektur model menggunakan TensorFlow dan Keras. Model yang dikembangkan memiliki beberapa lapisan utama, yaitu embedding layer untuk mengonversi teks menjadi vektor numerik, LSTM layer untuk memproses data secara sekuensial, dense layer untuk mengekstraksi pola, serta softmax activation layer yang digunakan untuk mengklasifikasikan opini mahasiswa ke dalam berbagai kategori.

Pada tahap pelatihan model, dilakukan K-Fold Cross Validation (K=5) untuk meningkatkan generalisasi model terhadap data yang beragam. Model dilatih dalam 120 epoch dengan menggunakan Adam Optimizer serta early stopping untuk mencegah overfitting. Namun, dalam proses pelatihan ditemukan bahwa dataset yang digunakan mengalami ketidakseimbangan data (imbalanced data), di mana beberapa kategori memiliki jumlah data yang jauh lebih kecil dibandingkan kategori lain. Untuk mengatasi hal ini, diterapkan teknik Synthetic Minority Over-sampling Technique (SMOTE) guna menyeimbangkan distribusi data antara kategori mayoritas dan minoritas.

Setelah model selesai dilatih, dilakukan evaluasi model menggunakan multiclass confusion matrix, yang mengukur beberapa metrik evaluasi, seperti akurasi, precision, recall, dan F1-score. Hasil evaluasi menunjukkan bahwa model mencapai akurasi 97% dalam skenario validasi, namun saat diuji dengan 100 data EDOM terbaru, akurasi menurun menjadi 74%. Penurunan akurasi ini diduga terjadi karena adanya kosakata baru (out-of-vocabulary) yang belum dikenal oleh model selama pelatihan, sehingga penelitian ini merekomendasikan pembaruan model secara berkala agar mampu mengakomodasi perubahan pola bahasa yang digunakan mahasiswa dalam memberikan feedback.

Langkah terakhir dalam penelitian ini adalah implementasi model ke dalam sistem otomatisasi klasifikasi EDOM. Model yang telah dikembangkan diintegrasikan ke dalam sebuah program berbasis Python, yang memungkinkan input berupa file Excel berisi komentar mahasiswa, kemudian mengeluarkan hasil klasifikasi dalam format Excel dengan label kategori sentimen yang sesuai.

Penelitian ini dilakukan di Primakara University dan berlangsung dari September 2023 hingga April 2024. Dengan metode yang digunakan, penelitian ini diharapkan dapat memberikan solusi yang efektif dalam otomatisasi analisis EDOM, sehingga dapat meningkatkan efisiensi dalam evaluasi kinerja dosen. Selain itu, model yang dikembangkan

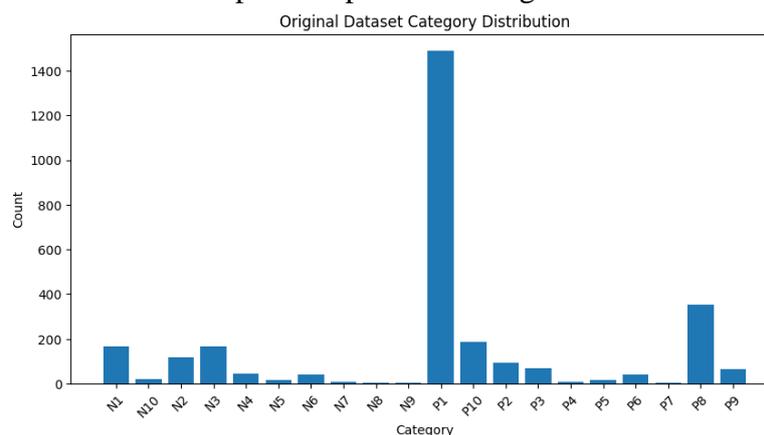
juga dapat menjadi referensi bagi penelitian selanjutnya yang ingin mengimplementasikan LSTM dalam klasifikasi opini berbasis teks dengan kategori yang lebih spesifik.

Hasil dan Pembahasan

Pengolahan dan Anotasi Data

Pada tahap awal penelitian, dilakukan pengolahan dan anotasi data terhadap dataset yang diperoleh dari Evaluasi Dosen oleh Mahasiswa (EDOM) di Primakara University. Dataset ini terdiri dari 3.100 data mentah, yang berasal dari periode akademik 2020/2021 hingga 2022/2023. Setelah dilakukan proses pembersihan dan penghapusan data duplikat, jumlah data yang valid adalah 2.912 data. Data ini kemudian dikategorikan berdasarkan aspek pembelajaran yang telah ditentukan dalam 20 kategori awal, yang terdiri dari 10 kategori positif (P1-P10) dan 10 kategori negatif (N1-N10).

Proses anotasi dilakukan dengan membaca setiap komentar mahasiswa secara manual dan mengklasifikasikannya ke dalam kategori yang sesuai. Untuk memastikan objektivitas, anotasi data divalidasi oleh Lembaga Penjaminan Mutu di Primakara University. Data yang mengalami ketidakseimbangan distribusi kemudian ditinjau ulang. Jika terdapat kategori dengan jumlah data yang sangat sedikit (< 50 data), kategori tersebut digabungkan dengan kategori lain yang memiliki kesamaan semantik. Sebaliknya, jika terdapat kategori “dan lain-lain” (P10 dan N10) yang memiliki jumlah data yang terlalu banyak (> 300 data), dilakukan analisis ulang untuk menentukan apakah diperlukan kategori baru.



Gambar 1 Jumlah Data Komentar Per-Kategori

Distribusi awal data per kategori ditunjukkan pada Gambar 1, yang mengindikasikan bahwa beberapa kategori memiliki jumlah data yang dominan dibandingkan kategori lainnya. Untuk mengatasi ketidakseimbangan ini, dilakukan metode Synthetic Minority Over-sampling Technique (SMOTE), yang memungkinkan data kategori minoritas diperbanyak secara sintesis agar distribusi data lebih merata.

Membangun Arsitektur Model

Setelah data diproses, model klasifikasi berbasis Long Short-Term Memory (LSTM) dikembangkan menggunakan TensorFlow dan Keras. Model ini dirancang untuk menangani data teks sekuensial dan mengklasifikasikan opini mahasiswa secara lebih akurat. Arsitektur model yang digunakan terdiri dari beberapa lapisan utama:

1. Embedding Layer – Mengubah teks menjadi vektor numerik berdimensi 128 untuk diproses lebih lanjut.
2. LSTM Layer – Memproses data teks secara sekuensial dan mempertahankan hubungan semantik antar kata dalam kalimat.

3. Dense Layer – Lapisan fully connected yang digunakan untuk mengolah fitur lebih lanjut.
4. Softmax Activation Layer – Lapisan terakhir yang berfungsi untuk menentukan probabilitas kategori pada setiap input teks.

Dalam tahap ini, dilakukan eksperimen terhadap beberapa hyperparameter, seperti jumlah unit LSTM, ukuran vektor embedding, jumlah epoch, dan learning rate untuk mendapatkan kombinasi terbaik yang menghasilkan akurasi tinggi.

Pelatihan Model

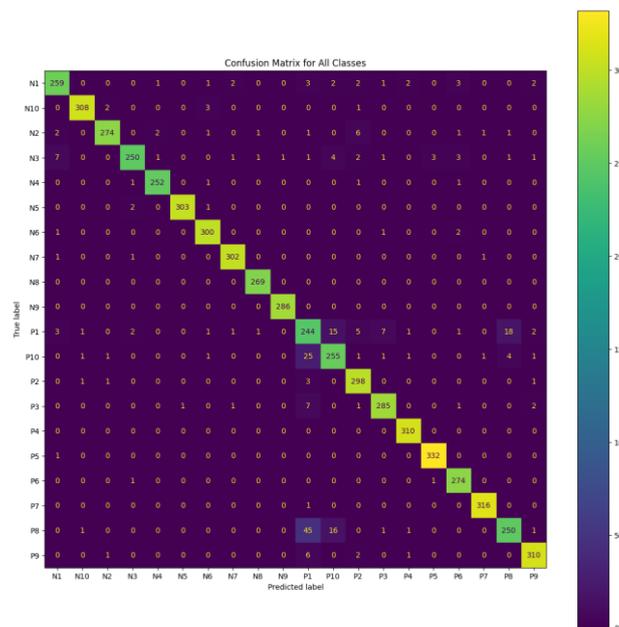
Setelah arsitektur model dibangun, dilakukan pelatihan model dengan menggunakan 5-Fold Cross Validation agar model dapat mempelajari pola dari data dengan lebih baik dan menghindari overfitting. Model dilatih selama 120 epoch dengan Adam Optimizer serta diterapkan metode early stopping, sehingga model hanya dilatih selama model masih mengalami peningkatan performa.

Hasil pelatihan model menunjukkan bahwa akurasi pada data latih mencapai 97%, dengan akurasi pada data uji sebesar 97%. Namun, ditemukan bahwa ketika model diuji menggunakan 100 data baru, akurasi menurun menjadi 74%. Hal ini mengindikasikan bahwa model mengalami kesulitan dalam menangani kosakata baru yang belum dikenali dalam data latih (out-of-vocabulary problem).

Evaluasi Model

Untuk mengevaluasi performa model, digunakan multiclass confusion matrix yang mengukur beberapa metrik evaluasi seperti:

- Akurasi: Mengukur seberapa akurat model dalam mengklasifikasikan data dengan benar.
- Precision: Menghitung proporsi prediksi yang benar dibandingkan dengan total prediksi dalam setiap kelas.
- Recall: Mengukur sejauh mana model mampu menemukan semua data dalam setiap kategori.
- F1-score: Rata-rata harmonik dari precision dan recall.



Gambar 2 Output Multiclass Confusion Matrix

Hasil evaluasi menunjukkan bahwa model memiliki performa yang baik pada kategori dengan jumlah data besar, tetapi kurang optimal pada kategori yang memiliki data sedikit. Gambar 2 menunjukkan multiclass confusion matrix yang menggambarkan akurasi model dalam tiap kategori.

Berikut adalah hasil evaluasi pada beberapa kategori utama:

Tabel 1 Hasil Evaluasi

Kategori	Akurasi (%)	Precision (%)	Recall (%)	F1-Score (%)
P1 - Metode Pengajaran	99.43	94.53	93.17	93.84
P2 - Kelas & Suasana	99.65	98.21	94.48	96.31
P3 - Kepemahaman Materi	99.43	97.28	93.27	94.94
N1 - Metode Pengajaran Kurang	99.21	91.34	92.17	91.75
N2 - Kelas Tidak Nyaman	98.97	89.54	90.12	89.83

Dapat dilihat bahwa kategori dengan jumlah data lebih banyak memiliki akurasi lebih tinggi dibandingkan dengan kategori dengan jumlah data lebih sedikit. Hal ini menunjukkan bahwa penggunaan metode SMOTE dalam menangani ketidakseimbangan data sudah memberikan dampak positif, tetapi masih terdapat beberapa kategori yang perlu ditingkatkan.

Implementasi Model ke dalam Sistem

Setelah model berhasil dilatih dan diuji, tahap selanjutnya adalah implementasi model dalam program berbasis Python. Model ini dikembangkan menjadi sebuah sistem otomatisasi yang memungkinkan pengguna mengunggah file Excel berisi komentar EDOM, kemudian sistem akan melakukan klasifikasi otomatis terhadap setiap komentar dan menyimpannya kembali dalam format Excel.

Cara kerja sistem otomatisasi klasifikasi EDOM:

1. Input Data: Mahasiswa atau administrator kampus mengunggah file Excel berisi komentar EDOM.
2. Preprocessing Data: Sistem membersihkan dan menyiapkan teks sebelum diklasifikasikan.
3. Klasifikasi dengan Model LSTM: Model yang telah dilatih digunakan untuk menentukan kategori setiap komentar.
4. Output Data: Hasil klasifikasi disimpan kembali dalam file Excel dengan kolom tambahan yang menunjukkan kategori sentimen dari setiap komentar.

Keuntungan dari sistem ini adalah:

- Mempercepat proses evaluasi dosen, yang sebelumnya dilakukan secara manual.
- Mengurangi subjektivitas dalam analisis komentar mahasiswa.
- Memungkinkan analisis berbasis data yang lebih akurat dalam menentukan kualitas pengajaran dosen.

Pembahasan

Berdasarkan hasil penelitian ini, dapat disimpulkan bahwa penggunaan LSTM dalam klasifikasi opini mahasiswa terhadap EDOM menghasilkan akurasi tinggi dan dapat membantu proses evaluasi dosen secara lebih cepat dan objektif. Namun, terdapat beberapa kendala yang perlu diperhatikan:

1. Penurunan akurasi ketika diuji dengan data baru – Model mengalami kesulitan dalam mengenali kosakata baru yang tidak terdapat dalam data latih. Solusi yang dapat diterapkan adalah pembaruan model secara berkala dengan dataset terbaru.
2. Ketidakseimbangan data pada beberapa kategori – Meskipun metode SMOTE telah diterapkan, beberapa kategori dengan jumlah data kecil masih memiliki performa lebih rendah dibandingkan kategori lainnya. Salah satu solusi yang dapat diterapkan adalah penggunaan teknik transfer learning atau word embeddings yang lebih luas seperti BERT atau Word2Vec.
3. Kelemahan dalam menangani opini netral – Model ini tidak menggunakan kategori netral, sehingga opini mahasiswa yang bersifat netral dapat diklasifikasikan secara kurang akurat. Penelitian di masa depan dapat mempertimbangkan untuk menambahkan kategori netral dalam klasifikasi.

Hasil penelitian ini menunjukkan bahwa metode LSTM sangat efektif dalam mengklasifikasikan opini berbasis teks, dan penerapannya dalam analisis EDOM di Primakara University dapat memberikan manfaat signifikan dalam peningkatan kualitas pengajaran dosen.

Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa penggunaan Long Short-Term Memory (LSTM) dalam klasifikasi sentimen Evaluasi Dosen oleh Mahasiswa (EDOM) mampu menghasilkan model dengan tingkat akurasi yang tinggi dalam mengelompokkan opini mahasiswa terhadap dosen. Model ini awalnya dikembangkan dengan 20 kategori sentimen, namun karena adanya ketidakseimbangan data, dilakukan oversampling menggunakan metode SMOTE untuk menyeimbangkan proporsi data di setiap kategori.

Model LSTM yang telah dikembangkan dilatih dan diuji menggunakan validasi silang (cross-validation) dan berhasil mencapai akurasi sebesar 95,40% berdasarkan F1-score, yang menunjukkan bahwa model memiliki performa yang baik dalam klasifikasi teks. Namun, saat model diuji menggunakan data baru, akurasinya menurun menjadi 74%. Penurunan ini disebabkan oleh kemunculan kosakata baru yang tidak terdapat dalam kamus encoding model, sehingga model kesulitan dalam mengklasifikasikan opini yang menggunakan istilah atau gaya bahasa yang berbeda dari data latih.

Sebagai bagian dari implementasi, model yang telah dilatih diterapkan dalam sebuah sistem otomatisasi klasifikasi EDOM, yang memungkinkan pengguna untuk melakukan klasifikasi opini mahasiswa secara otomatis melalui aplikasi berbasis Python. Sistem ini mampu mempercepat proses analisis EDOM yang sebelumnya dilakukan secara manual, namun tetap memiliki tantangan dalam menangani kata-kata baru dan variasi bahasa yang digunakan oleh mahasiswa dalam feedback mereka.

Agar penelitian ini dapat ditingkatkan lebih lanjut, diperlukan beberapa langkah perbaikan dalam pengembangan model dan sistem yang digunakan. Salah satu langkah yang dapat diterapkan adalah peningkatan kualitas dataset dengan mengumpulkan lebih banyak data dan melatih model secara berkala agar tetap relevan dengan kosakata terbaru yang digunakan oleh mahasiswa. Selain itu, dataset harus seimbang antar kategori agar model tidak bias terhadap kategori tertentu yang memiliki jumlah data lebih banyak.

Selain SMOTE yang telah digunakan dalam penelitian ini, metode lain seperti data augmentation berbasis teks atau synthetic text generation dapat dieksplorasi untuk meningkatkan keberagaman data dan membantu model belajar dari variasi bahasa yang lebih

luas. Model juga masih mengalami kesulitan dalam mengenali kata-kata dengan kesalahan ejaan (typo), sehingga diperlukan algoritma koreksi ejaan dalam preprocessing untuk meningkatkan akurasi pemrosesan teks sebelum dilakukan klasifikasi.

Selanjutnya, penggunaan metode encoding yang lebih canggih perlu dipertimbangkan. TextVectorizer dari TensorFlow yang digunakan dalam penelitian ini masih memiliki keterbatasan dalam memahami konteks kata dalam kalimat. Oleh karena itu, metode Word Embeddings seperti Word2Vec, FastText, atau BERT dapat digunakan untuk meningkatkan pemahaman model terhadap makna kata berdasarkan konteksnya serta memperluas kamus encoding, terutama dalam bahasa Indonesia.

Dengan menerapkan langkah-langkah tersebut, diharapkan model dapat semakin akurat, robust, dan adaptif terhadap perubahan pola bahasa yang digunakan oleh mahasiswa dalam memberikan umpan balik pada EDOM. Hal ini akan meningkatkan efektivitas sistem dalam melakukan klasifikasi opini mahasiswa secara otomatis dan mendukung peningkatan kualitas evaluasi kinerja dosen secara lebih objektif dan efisien.

Referensi

- Alharbi, F. (2021). Sampling strategies for tackling imbalanced data in human activity recognition [Doctoral dissertation, Goldsmiths, University of London].
- Amrustian, M. A., Widayat, W., & Wirawan, A. M. (2022). Analisis sentimen evaluasi terhadap pengajaran dosen di perguruan tinggi menggunakan metode LSTM. *Jurnal Media Informatika Budidarma*, 6(1), 535. <https://doi.org/10.30865/mib.v6i1.3527>
- Andrade, E. de O., Viterbo, J., Vasconcelos, C. N., Guérin, J., & Bernardini, F. C. (2019). A model based on LSTM neural networks to identify five different types of malware. *Procedia Computer Science*, 159, 182–191. <https://doi.org/10.1016/j.procs.2019.09.173>
- Arinata, I. G. B., Satwika, I. P., & Utami, N. W. (2022). Analisis Sentimen pada EDOM STMIK Primakara Menggunakan Metode Long Short-Term Memory [Undergraduate thesis].
- Chen, B., Huang, Q., Chen, Y., Cheng, L., & Chen, R. (2018). Deep neural networks for multi-class sentiment classification. 2018 IEEE 20th International Conference on High Performance Computing and Communications; IEEE 16th International Conference on Smart City; IEEE 4th International Conference on Data Science and Systems (HPCC/SmartCity/DSS), 854–859. <https://doi.org/10.1109/HPCC/SmartCity/DSS.2018.00142>
- Cholissodin, I., Sutrisno, S., Soebroto, A. A., Hasanah, U., & Febiola, Y. I. (2020). AI, Machine Learning & Deep Learning. Fakultas Ilmu Komputer, Universitas Brawijaya.
- Danukusumo, K. P. (2017). Implementasi Deep Learning Menggunakan Convolutional Neural Network untuk Klasifikasi Citra Candi Berbasis GPU [Undergraduate thesis].
- Direktorat Pengembangan dan Pelayanan Sistem Informasi UI. (2023, November 29). Evaluasi Dosen Oleh Mahasiswa. Retrieved from <https://edom.ui.ac.id/>
- Eke, H., Petrovski, A., & Ahriz, H. (2020). Detection of false command and response injection attacks for cyber-physical systems security and resilience. 13th International Conference on Security of Information and Networks, 1–8. <https://doi.org/10.1145/3433174.3433615>
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.
- Hasiholan, A., Cholissodin, I., & Yudistira, N. (2022). Analisis sentimen tweet Covid-19 varian Omicron pada platform media sosial Twitter menggunakan metode LSTM berbasis

- multi fungsi aktivasi dan GLOVE. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 6(10), 4653–4661.
- Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8), 1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
- Kandhro, I. A., Wasi, S., Kumar, K., Rind, M., & Ameen, M. (2019). Sentiment analysis of student's comment by using long-short term model. *Indian Journal of Science and Technology*, 12(8), 1–16. <https://doi.org/10.17485/ijst/2019/v12i8/141741>
- Kim, S., Raza, M., & Seidman, E. (2019). Improving 21st-century teaching skills: The key to effective 21st-century learners. *Research in Comparative and International Education*, 14(2), 174–191. <https://doi.org/10.1177/1745499919829214>
- Kristiawan, K., Somali, D. D., Linggan Jaya, T. A., & Widjaja, A. (2020). Deteksi buah menggunakan supervised learning dan ekstraksi fitur untuk pemeriksa harga. *Jurnal Teknik Informatika dan Sistem Informasi*, 6(3). <https://doi.org/10.28932/jutisi.v6i3.3029>
- Larasati, K. D., & Primandari, A. H. (2021). Forecasting Bitcoin price based on blockchain information using long-short term method. *Parameter: Journal of Statistics*, 1(1), 1–6. <https://doi.org/10.22487/27765660.2021.v1.i1.15389>
- Mac, H., Tran, D., Tong, V., Nguyen, L. G., & Tran, H. A. (2017). DGA botnet detection using supervised learning methods. *Proceedings of the Eighth International Symposium on Information and Communication Technology*, 211–218. <https://doi.org/10.1145/3155133.3155166>
- Markoulidakis, I., Kopsiaftis, G., Rallis, I., & Georgoulas, I. (2021). Multi-class confusion matrix reduction method and its application on net promoter score classification problem. *Proceedings of the 14th PErvasive Technologies Related to Assistive Environments Conference*, 412–419. <https://doi.org/10.1145/3453892.3461323>
- Rusli, M. (2020). Ekstraksi fitur menggunakan model Word2Vec pada sentiment analysis kolom komentar kuisisioner evaluasi dosen oleh mahasiswa. *KLIK - Kumpulan Jurnal Ilmu Komputer*, 7(1), 35. <https://doi.org/10.20527/klik.v7i1.296>
- Salam, N. F. (2021). Klasifikasi Teks Sentimen Review E-Commerce Menggunakan Algoritma Long Short-Term Memory (LSTM) [Undergraduate thesis].
- Sanda, Y., Warman, W., Pitriyani, A., & Yesepa, Y. (2022). Peningkatan mutu perguruan tinggi melalui manajemen pendidik dan tenaga kependidikan. *Jurnal Akuntabilitas Manajemen Pendidikan*, 10(1), 85–94. <https://doi.org/10.21831/jamp.v10i1.47855>
- Sevsa, B. A. (2018). Analisis Sentimen Komentar pada Indeks Kinerja Dosen Fakultas Sains dan Teknologi UIN Sunan Kalijaga Menggunakan Naïve Bayes Classifier [Undergraduate thesis].
- Srikanth, J., & Shanmugam, A. D. (2023). A deep LSTM-RNN classification method for Covid-19 Twitter review based on sentiment analysis. *Scalable Computing: Practice and Experience*, 24(3), 315–326.
- Sugiyono. (2018). *Metode penelitian kuantitatif, kualitatif, dan R&D*. Bandung: Alfabeta.
- Suh, S. (2021). Improving classification performance under imbalanced data conditions using generative adversarial networks [Doctoral dissertation, Technische Universität Kaiserslautern].
- Supartini, I. A. M., Sukarsa, I. K. G., & Srinadi, I. (2017). Analisis diskriminan pada klasifikasi desa di Kabupaten Tabanan menggunakan metode K-Fold cross validation. *E-Jurnal Matematika*, 6(2), 106.

- TensorFlow. (2023, November 29). Introduction to TensorFlow. Retrieved from <https://www.tensorflow.org/learn>
- TensorFlow. (2023, November 29). TensorFlow API Documentation. Retrieved from https://www.tensorflow.org/api_docs
- Witantoa, K. S., ERa, N. A. S., Karyawatia, A. E., Arya, I., Kadyanana, I., & Astutia, L. G. (2022). Implementasi LSTM pada analisis sentimen review film menggunakan Adam dan RMSprop optimizer. *Jurnal Elektronik Ilmu Komputer Udayana*, 2301, 5373.