

# Analisis Perbandingan Algoritma *Machine Learning* Dan *Deep Learning* Untuk Sentimen Analisis Teks Umpan Balik Tentang Evaluasi Pengajaran Dosen

<sup>1</sup>Hadiguna Setiawan, <sup>2</sup>Dhani Ariatmanto

<sup>1,2</sup>Program Studi Magister Teknik Informatika, Universitas Amikom Yogyakarta

<sup>1</sup>hadi@students.amikom.ac.id, <sup>2</sup>dhaniani@amikom.ac.id

## Article Info

### Article history:

Received, 2024-05-28

Revised, 2024-06-20

Accepted, 2024-06-29

### Kata Kunci:

KNN,  
BiLSTM,  
SVM,  
LSTM  
Evaluasi dosen

### Keywords:

KNN,  
BiLSTM,  
SVM,  
LSTM  
Evaluation of lecturer

## ABSTRAK

Evaluasi kinerja dosen sangat penting karena membantu dalam memantau dan memastikan bahwa dosen memenuhi tugas secara efektif dalam menjaga integritas dan mengajarkan materi perkuliahan. Dengan melakukan penilaian kinerja dosen berdasarkan kriteria seperti pengajaran, dapat mengidentifikasi area untuk perbaikan dan memberikan dukungan jika diperlukan. Penelitian ini bertujuan untuk mengetahui tingkat akurasi machine learning dan deep learning yang dikombinasikan dengan word-embedding untuk analisis teks evaluasi kinerja pengajaran dosen dengan menggunakan teknik pra-proses. Dataset terdiri dari 663 data positif, 552 data negatif, 465 data netral. Berhasil menghasilkan hasil eksperimen, nilai akurasi pelatihan untuk setiap metode klasifikasi antara lain KNN sebesar 74.75%, SVM sebesar 65.78%, RF sebesar 98.58%, LSTM sebesar 95.64% dan Bi-LSTM sebesar 95.91%. Untuk nilai akurasi pengujian untuk setiap metode klasifikasi antara lain KNN sebesar 59.82%, SVM sebesar 62.88%, RF sebesar 69.37%, LSTM sebesar 70.81% dan Bi-LSTM sebesar 72.25%. Metode yang paling unggul dalam mengolah data 663 data positif, 552 data negatif, 465 data netral dengan menerapkan metode word-embedding yaitu BiLSTM dengan akurasi pelatihan 95.91% dan akurasi pengujian sebesar 72.25%.

## ABSTRACT

Evaluation of lecturer performance is very important because it helps in monitoring and ensuring that lecturers fulfill their duties effectively in maintaining integrity and teaching lecture material. By assessing lecturer performance based on criteria such as teaching, it can identify areas for improvement and provide support if needed. This study aims to determine the accuracy level of machine learning and deep learning combined with word-embedding for text analysis of lecturer teaching performance evaluation using preprocess techniques. The dataset consisted of 663 positive data, 552 negative data, and 465 neutral data. Successful in the results of the experiment, the training accuracy value for each classification method included KNN of 74.75%, SVM of 65.78%, RF of 98.58%, LSTM of 95.64% and Bi-LSTM of 95.91%. The test accuracy value for each classification method includes KNN of 59.82%, SVM of 62.88%, RF of 69.37%, LSTM of 70.81% and Bi-LSTM of 72.25%. The most superior method in processing data of 663 positive data, 552 negative data, and 465 neutral data by applying the word-embedding method, namely BiLSTM with a training accuracy of 95.91% and a testing accuracy of 72.25%.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-nc-nd/4.0/) license.



## Penulis Korespondensi:

Hadiguna Setiawan,  
Program Studi Magister Teknik Informatika  
Universitas Amikom Yogyakarta  
Email: hadi@students.amikom.ac.id

## 1. PENDAHULUAN

Evaluasi kinerja dosen berdasarkan persepsi mahasiswa memiliki sejarah panjang untuk mendukung pengembangan kualitas universitas. Evaluasi kinerja dosen ini diukur melalui instrumen kuesioner dengan komentar kualitatif terbuka tergantung pada kebijakan universitas [1]–[3]. Data yang dikumpulkan dari instrumen ini digunakan untuk berbagai tujuan yang berbeda termasuk umpan balik untuk meningkatkan kualitas pengajaran atau sebagai masukan untuk proses pengukuran kinerja dosen atau akhirnya untuk menanggapi persyaratan pemerintah [4], [5].

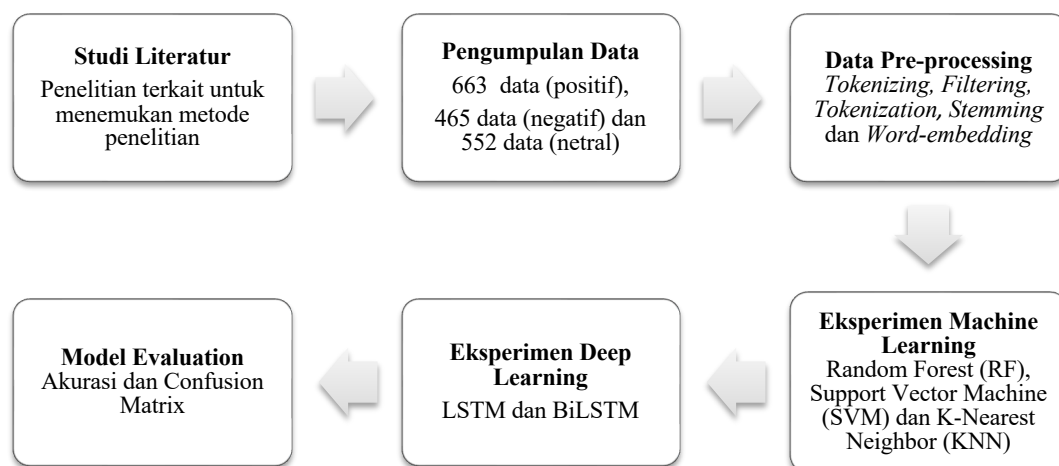
Analisis sentimen mengacu pada studi analisis teks, proses bahasa alami, linguistik komputasi untuk mengidentifikasi, mengekstrak, dan mempelajari informasi sub-kata sifat secara ilmiah dari data tekstual [6], [7]. Tugas penting dari analisis sentimen adalah untuk menentukan polaritas teks yang diberikan pada tingkat fitur, kalimat, dan dokumen. Analisis sentimen membantu menganalisis data teks dan mengekstrak beberapa wawasan penting yang akan membantu mendukung keputusan bisnis [8]–[10].

Penelitian sebelumnya telah melakukan analisis umpan balik evaluasi pembelajaran dengan menggunakan algoritma Model Naïve Bayes (NB), Recurrent Neural Network (RNN), Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Support Vector Machine (SVM), Random Forest (RF) [11]–[16]. Berdasarkan penelitian oleh [13], algoritma Support Vector Machine (SVM) dan mendapatkan akurasi sebesar tertinggi dibandingkan dengan algoritma lainnya. Namun, penelitian lain mencoba untuk membandingkan pengklasifikasi pembelajaran mesin tradisional yaitu, KNN, SVM terhadap arsitektur deep learning Long Short-Term Memory (LSTM) [17]. Berdasarkan hasil pada penelitian lainnya diketahui bahwa arsitektur Bi-LSTM berkinerja lebih baik daripada arsitektur LSTM untuk pemrosesan dataset berkenaan ujaran kebencian [18].

Riset ini akan melakukan analisis data teks sebanyak tentang umpan balik proses pembelajaran dari sistem internal perguruan tinggi dengan menggunakan teknik praproses dan membandingkan kinerja dari K-Nearest Neighbor (KNN), Support Vector Machine (SVM), Random Forest (RF), Long Short – Term Memory Neural Network (LSTM) dan Bidirectional LSTM (BiLSTM).

## 2. METODE PENELITIAN

Penelitian ini bersifat eksperimental. Eksperimen dilakukan terhadap data evaluasi dari *feedback* pembelajaran di kelas. Dataset yang telah memiliki label sentiment diberikan sebuah label positif, netral dan negatif menggunakan metode *machine learning* dan *deep learning* yang menerapkan *word-embedding* dan *dropout*. Pengumpulan data dilakukan di Fakultas Ilmu Komputer, Universitas Sjakhyakirti (UNISTI), Palembang, Sumatera Selatan. Data yang diambil ada evaluasi dari *feedback* pembelajaran di kelas. Dataset yang telah memiliki label sentiment diberikan sebuah label positif, netral dan negatif. Adapun alur penelitian dapat dilihat pada **Gambar 1**.



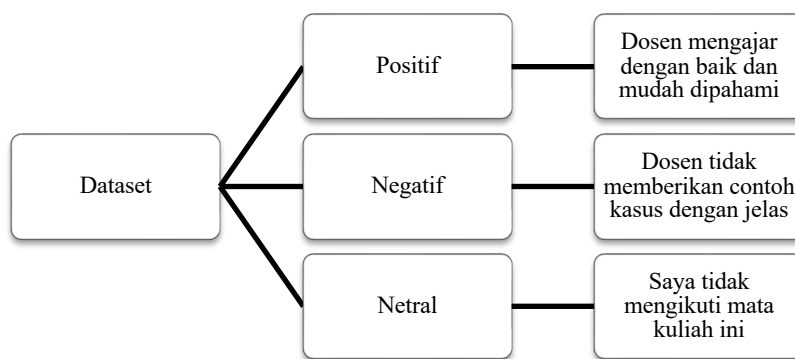
Gambar 1 Alur Penelitian

Penelitian ini bertujuan untuk mengetahui tingkat akurasi *machine learning* dan *deep learning* yang dikombinasikan dengan *word-embedding* untuk analisis teks evaluasi kinerja pengajaran dosen. Dataset tersebut dimasukkan ke sistem dan dilakukan proses preprocessing melewati beberapa proses yaitu filtering, stemming, dan tokenization. Pembagian data dilakukan untuk membagi data yang akan dianalisis menjadi data latih atau training dan data uji atau testing dengan rasio 8 berbanding 2. Adapun skenario eksperimen yang akan dilakukan dapat dilihat pada **Tabel 1**.

Tabel 1 Eksperimen Penelitian

Eksperimen	Model	Word-Embedding
Eksp-1	K-Nearest Neighbor (KNN)	√
Eksp-2	Support Vector Machine (SVM)	√
Eksp-3	Random Forest (RF)	√
Eksp-4	LSTM	√
Eksp-5	BiLSTM	√

Data penelitian dikumpulkan menggunakan layanan Google Form dalam bentuk kuesioner yang dapat diisi oleh mahasiswa kelas. Dataset yang telah dikumpulkan dalam bentuk teks akan proses untuk diberikan label positif, netral dan negatif. Label ini digunakan untuk mengetahui presentase tingkat kepuasan mahasiswa terhadap pengajaran yang dilakukan dosen. Data akan dibagi menjadi data pelatihan dan data pengujian. Teknik analisis data yang digunakan adalah *tokenizing*, *filtering*, *stemming* dan *word-embedding*. Dataset terdiri dari 663 data positif, 552 data negatif, 465 data netral. Adapun contoh kalimat positif, netral dan negatif untuk pelabelan data penelitian dapat dilihat pada Gambar 2.



Gambar 2 Data Penelitian

### 3. HASIL DAN ANALISIS

Dalam melakukan eksperimen ini diterapkan teknik praproses untuk menyiapkan data yang lebih baik untuk dapat dianalisis. Teknik praproses digunakan untuk menghilangkan noise dari data, membuatnya lebih mudah dianalisis, dan mengurangi beban komputasi pada saat diproses menggunakan algoritma klasifikasi. Selain itu, teknik praproses diterapkan untuk membantu meningkatkan akurasi dan efisiensi model klasifikasi. Terdapat empat teknik praproses yang diterapkan pada penelitian ini, yaitu case folding, filtering, tokenisasi, stop-word removal dan stemming. Dalam penelitian ini, teknik praproses stopword removal dan stemming menggunakan perpustakaan NLTK dan Sastrawi. Hasil dari teknik praproses dapat dilihat pada Gambar 3.

Dataset	Case folding	Tokenisasi	Stop-word removal	Stemming
Sangat siap dan mudah di mengerti	sangat siap dan mudah di mengerti	['sangat', 'siap', 'dan', 'mudah', 'di', 'meng...']	['mudah', 'mengerti']	['mudah', 'erti']
Sangat baik dalam menerangkan materi ataupun p...	sangat baik dalam menerangkan materi ataupun p...	['sangat', 'baik', 'dalam', 'menerangkan', 'ma...']	['menerangkan', 'materi', 'praktek']	['terang', 'materi', 'praktek']
Menyiapkan materi yang akan di bahas	menyiapkan materi yang akan di bahas	['menyiapkan', 'materi', 'yang', 'akan', 'di', '...']	['materi', 'bahas']	['materi', 'bahas']
Penyampaiannya baik dan mudah di pahami	penyampaiannya baik dan mudah di pahami	['penyampaiannya', 'baik', 'dan', 'mudah', 'di', '...']	['penyampaiannya', 'mudah', 'pahami']	['sampil', 'mudah', 'paham']

Gambar 3 Hasil Praproses Dataset

Setelah melakukan tahap praproses, data diproses menggunakan K-Nearest Neighbor (KNN), Support Vector Machine (SVM), Random Forest (RF), Long Short – Term Memory Neural Network (LSTM) dan Bidirectional LSTM (BiLSTM). Dataset terdiri dari 663 data positif, 552 data negatif, 465 data netral.

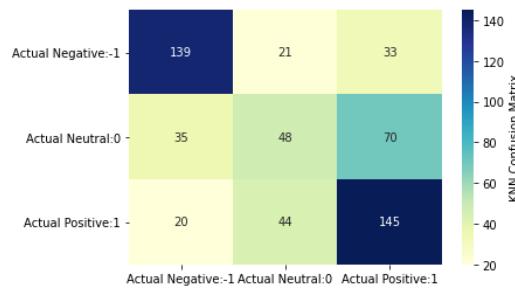
Penelitian ini bertujuan untuk menerapkan metode *word-embedding* pada metode *machine learning* (KNN, RF, SVM) dan deep learning (LSTM dan BiLSTM).

Pada eksperimen pertama dilakukan dengan menggunakan metode *word-embedding* dan metode KNN. Metode KNN adalah salah satu metode yang paling populer namun masih memiliki banyak kekurangan dalam mengklasifikasikan teks, seperti kompleksitas perhitungan yang besar, tidak ada perbedaan antara kata-kata karakteristik, tidak mempertimbangkan asosiasi antara kata kunci dan sebagainya. Hasil evaluasi nilai precision, recall dan F1-score dapat dilihat pada **Tabel 2**.

Tabel 2 Evaluasi KNN berdasarkan precision, recall dan F1-score

Kelas Data	Precision	Recall	F1-score
Negatif -1	0.72	0.72	0.72
Netral 0	0.42	0.31	0.36
Positif 1	0.58	0.69	0.63

Berdasarkan hasil evaluasi, KNN mendapatkan nilai akurasi pelatihan sebesar 74.75% dan akurasi pengujian sebesar 59.82%. Nilai ini didapatkan berdasarkan nilai prediksi benar yang dilakukan model KNN terhadap 663 data positif, 552 data negatif, 465 data netral. Hasil prediksi benar untuk setiap kelas dapat dilihat pada **Gambar 4**.



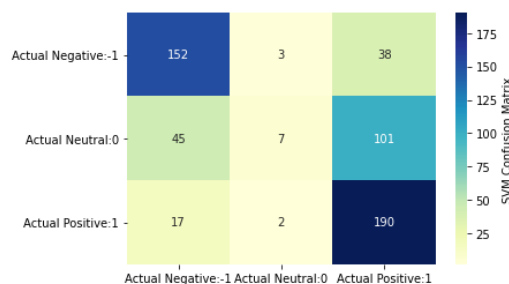
Gambar 4 Confusion matrix KNN

Pada eksperimen kedua dilakukan dengan menggunakan metode *word-embedding* dan metode SVM. Metode SVM telah digunakan secara ekstensif untuk klasifikasi teks. Penelitian sebelumnya menunjukkan bahwa pengklasifikasi SVM berkinerja baik ketika ukuran set fitur kecil, dengan eksperimen menunjukkan efektivitasnya dalam menganalisis dokumen teks bahasa Inggris. Hasil evaluasi nilai precision, recall dan F1-score dapat dilihat pada **Tabel 3**.

Tabel 3 Evaluasi SVM berdasarkan precision, recall dan F1-score

Kelas Data	Precision	Recall	F1-score
Negatif -1	0.71	0.79	0.75
Netral 0	0.58	0.05	0.08
Positif 1	0.58	0.91	0.71

Berdasarkan hasil evaluasi, SVM mendapatkan nilai akurasi pelatihan sebesar 65.78% dan akurasi pengujian sebesar 62.88%. Nilai ini didapatkan berdasarkan nilai prediksi benar yang dilakukan model SVM terhadap 663 data positif, 552 data negatif, 465 data netral. Hasil prediksi benar untuk setiap kelas dapat dilihat pada **Gambar 5**.



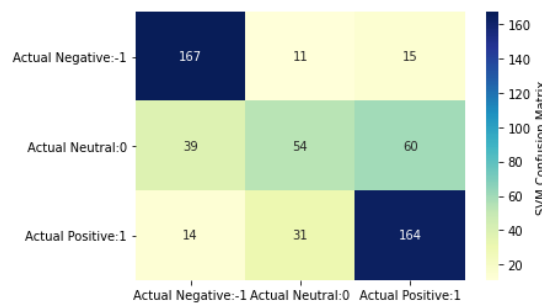
Gambar 5 Confusion matrix SVM

Pada eksperimen ketiga dilakukan dengan menggunakan metode *word-embedding* dan metode RF. Metode Random Forest (RF) telah banyak digunakan dalam tugas klasifikasi teks di berbagai domain dan menunjukkan efektivitasnya dalam meningkatkan efisiensi dan akurasi klasifikasi. Studi sebelumnya menyoroti ketahanan dan fleksibilitas RF dalam tugas klasifikasi teks, menjadikannya metode yang cukup baik untuk menganalisis dan mengkategorikan data tekstual secara efektif. Hasil evaluasi nilai precision, recall dan F1-score dapat dilihat pada **Tabel 4**.

Tabel 4 Evaluasi RF berdasarkan precision, recall dan F1-score

Kelas Data	Precision	Recall	F1-score
Negatif -1	0.76	0.87	0.81
Netral 0	0.56	0.35	0.43
Positif 1	0.69	0.78	0.73

Berdasarkan hasil evaluasi, RF mendapatkan nilai akurasi pelatihan sebesar 98.58% dan akurasi pengujian sebesar 69.37%. Nilai ini didapatkan berdasarkan nilai prediksi benar yang dilakukan model RF terhadap 663 data positif, 552 data negatif, 465 data netral. Hasil prediksi benar untuk setiap kelas dapat dilihat pada **Gambar 6**.



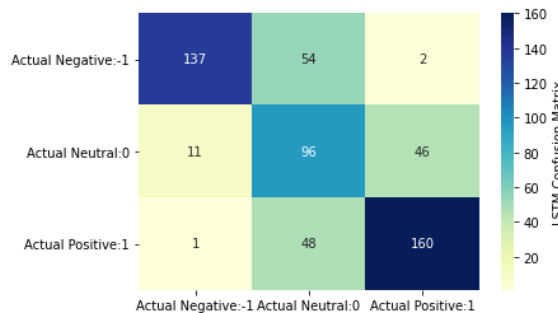
Gambar 6 Confusion matrix RF

Pada eksperimen keempat dilakukan dengan menggunakan metode *word-embedding* dan metode LSTM. Metode LSTM telah banyak digunakan dalam tugas klasifikasi teks karena kemampuannya untuk menangkap dependensi jangka panjang dalam data berurutan. Perbaikan terbaru dalam model LSTM termasuk menggabungkan mekanisme perhatian dan subsampling untuk meningkatkan ekstraksi fitur dan akurasi klasifikasi. Hasil evaluasi LSTM berdasarkan nilai precision, recall dan F1-score dapat dilihat pada **Tabel 5**.

Tabel 5 Evaluasi LSTM berdasarkan precision, recall dan F1-score

Kelas Data	Precision	Recall	F1-score
Negatif -1	0.92	0.71	0.80
Netral 0	0.48	0.63	0.55
Positif 1	0.77	0.77	0.77

Berdasarkan hasil evaluasi, LSTM mendapatkan nilai akurasi pelatihan sebesar 95.64% dan akurasi pengujian sebesar 70.81%. Nilai ini didapatkan berdasarkan nilai prediksi benar yang dilakukan model LSTM terhadap 663 data positif, 552 data negatif, 465 data netral. Hasil prediksi benar untuk setiap kelas dapat dilihat pada **Gambar 7**.



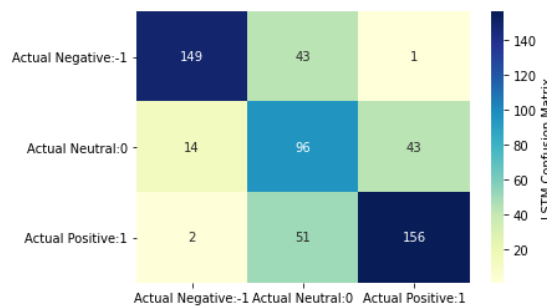
Gambar 7 Confusion matrix LSTM

Pada eksperimen kelima dilakukan dengan menggunakan metode *word-embedding* dan metode LSTM. Metode Model BiLSTM (*Bidirectional Long Short-Term Memory*) telah banyak digunakan dalam tugas klasifikasi teks untuk menangkap informasi kontekstual secara efektif. Berbagai penelitian telah mengusulkan pendekatan inovatif yang menggabungkan BiLSTM dengan teknik lain untuk meningkatkan akurasi klasifikasi teks. Hasil evaluasi BiLSTM berdasarkan nilai precision, recall dan F1-score dapat dilihat pada **Tabel 6**.

Tabel 6 Evaluasi BiLSTM berdasarkan precision, recall dan F1-score

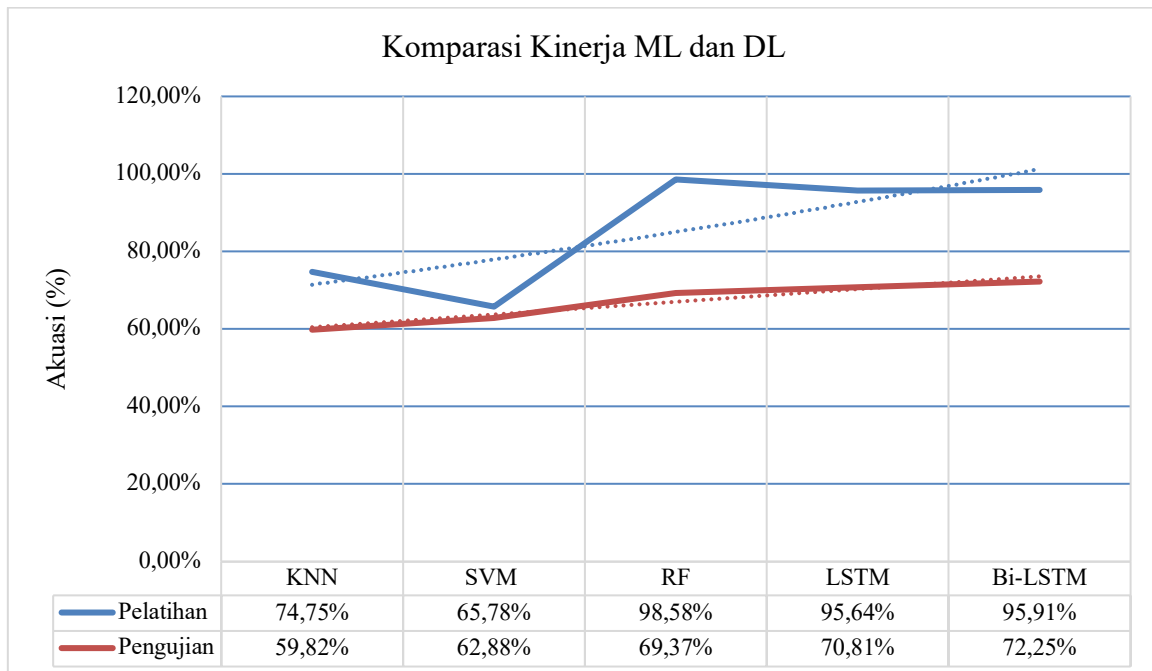
Kelas Data	Precision	Recall	F1-score
Negatif -1	0.90	0.77	0.83
Netral 0	0.51	0.63	0.56
Positif 1	0.78	0.75	0.76

Berdasarkan hasil evaluasi, BiLSTM mendapatkan nilai akurasi pelatihan sebesar 95.91% dan akurasi pengujian sebesar 72.25%. Nilai ini didapatkan berdasarkan nilai prediksi benar yang dilakukan model BiLSTM terhadap 663 data positif, 552 data negatif, 465 data netral. Hasil prediksi benar untuk setiap kelas dapat dilihat pada **Gambar 8**.



Gambar 8 Confusion matrix BiLSTM

Berdasarkan hasil evaluasi, BiLSTM mendapatkan nilai akurasi tertinggi dengan nilai pelatihan sebesar 95.91% dan akurasi pengujian sebesar 72.25%. LSTM mendapatkan nilai akurasi tertinggi kedua dengan nilai akurasi pelatihan sebesar 95.64% dan akurasi pengujian sebesar 70.81%. Adapun hasil perbandingan hasil kinerja metode *machine learning* (KNN, RF, SVM) dan deep learning (LSTM dan BiLSTM) dengan menerapkan metode *word-embedding* dapat dilihat pada **Gambar 9**.



Gambar 9 Komparasi Kinerja ML dan DL

#### 4. KESIMPULAN

Penelitian ini bertujuan untuk mengetahui kinerja metode *machine learning* (KNN, RF, SVM) dan deep learning (LSTM dan BiLSTM) dengan teknik praproses digunakan adalah *tokenizing*, *filtering*, *stemming* dan *word-embedding* dalam mengklasifikasikan dataset evaluasi pembelajaran yang terdiri dari 663 data positif, 552 data negatif, 465 data netral. Berdasarkan hasil eksperimen, BiLSTM mendapatkan nilai akurasi tertinggi dengan nilai pelatihan sebesar 95.91% dan akurasi pengujian sebesar 72.25%. Metode LSTM mendapatkan nilai akurasi tertinggi kedua dengan nilai akurasi pelatihan sebesar 95.64% dan akurasi pengujian sebesar 70.81%. Metode RF mendapatkan nilai akurasi pelatihan sebesar 98.58% dan akurasi pengujian sebesar 69.37%. Metode SVM mendapatkan nilai akurasi pelatihan sebesar 65.78% dan akurasi pengujian sebesar 62.88%. Terakhir, metode KNN mendapatkan nilai akurasi pelatihan sebesar 74.75% dan akurasi pengujian terendah yaitu sebesar 59.82%.

#### UCAPAN TERIMA KASIH

Terima kasih kepada Program Studi Magister Teknik Informatika, Universitas Amikom Yogyakarta yang telah mendukung penelitian ini.

#### REFERENSI

- [1] E. Sarmigi, M. Alfian, M. Ravico, M. S. Tiara, L. Angela, and F. Asbupel, *Instrumen Penelitian Dan Monitoring & Evaluasi (Money) Di Perguruan Tinggi*. Penerbit Adab, 2023.
- [2] L. Lohman, "Evaluation of university teaching as sound performance appraisal," *Stud. Educ. Eval.*, vol. 70, p. 101008, 2021.
- [3] B. C. E. Oguguo, F. A. Nannim, J. J. Agah, C. S. Ugwuanyi, C. U. Ene, and A. C. Nzeadibe, "Effect of learning management system on Student's performance in educational measurement and evaluation," *Educ. Inf. Technol.*, vol. 26, pp. 1471–1483, 2021.
- [4] S. Winarsih, "Kebijakan dan implementasi manajemen pendidikan tinggi dalam meningkatkan mutu pendidikan," *Cendekia J. Kependidikan Dan Kemasyarakatan*, vol. 15, no. 1, pp. 51–66, 2017.
- [5] J. M. Ramirez-Hurtado, A. G. Hernandez-Diaz, A. D. Lopez-Sanchez, and V. E. Perez-Leon, "Measuring online teaching service quality in higher education in the COVID-19 environment," *Int. J. Environ. Res. Public Health*, vol. 18, no. 5, p. 2403, 2021.
- [6] S. Lestari and S. Saepudin, "Analisis sentimen vaksin sinovac pada twitter menggunakan algoritma Naive Bayes," in *Prosiding Seminar Nasional Sistem Informasi dan Manajemen Informatika Universitas Nusa Putra*, 2021, vol. 1, no. 01, pp. 163–170.
- [7] M. U. Albab, E. Utami, and D. Ariatmanto, "Comparison of Algorithms for Sentiment Analysis of Operator Satisfaction Level for Increasing Neo Feeder Applications in PDDikti Higher Education LLDIKTI Region VI Semarang Central Java," *Sink. J. dan Penelit. Tek. Inform.*, vol. 8, no. 4, pp. 2099–2108, 2023.
- [8] S. K. Trivedi and A. Singh, "Twitter sentiment analysis of app based online food delivery companies," *Glob. Knowledge, Mem. Commun.*, vol. 70, no. 8/9, pp. 891–910, 2021.
- [9] D. A. Dzulhijjah, K. Kusriani, and K. A. Yuana, "Comparative Analysis of RNN, LSTM, Bi-LSTM Performance for Location and Time Entity Recognition in Forest Fire Texts," in *2024 2nd International Conference on Software Engineering and Information Technology (ICoSEIT)*, 2024, pp. 181–186.
- [10] A. Y. Ferianto, D. Ariatmanto, and M. Hayaty, "Perbandingan Metode Word Embedding Untuk Analisis Sentimen Pada Data Ulasan Marketplace," *J. ICT Inf. Commun. Technol.*, vol. 22, no. 2, pp. 220–225, 2022.
- [11] A. C. Adi, D. P. Lestari, F. S. S. Elsa, and Y. Sabui, "Online School Sentiment Analysis in Indonesia on Twitter Using The Naïve Bayes Classifier and Rapid Miner Tools," *Int. J. Innov. Sci. Res. Technol.*, vol. 71, 2022.
- [12] N. H. Harani and C. Prianto, "Sentiment Analysis of Student Emotion During Online Learning Using Recurrent Neural Networks (RNN)," *IJISTECH (International J. Inf. Syst. Technol.)*, vol. 5, no. 3, pp. 299–307, 2021.
- [13] S. F. C. Haviana and B. S. W. Poetro, "Deep Learning Model for Sentiment Analysis on Short Informal Texts," *Indones. J. Electr. Eng. Informatics*, vol. 10, no. 1, pp. 82–89, 2022.
- [14] A. R. Isnain, J. Supriyanto, and M. P. Kharisma, "Implementation of K-Nearest Neighbor (K-NN) Algorithm For Public Sentiment Analysis of Online Learning," *IJCCS (Indonesian J. Comput. Cybern. Syst.)*, vol. 15, no. 2, pp. 121–130, 2021.
- [15] D. F. Sengkey, A. Jacobus, and F. J. Manoppo, "Effects of kernels and the proportion of training data on the accuracy of SVM sentiment analysis in lecturer evaluation," *IAES Int. J. Artif. Intell.*, vol. 9, no.

- 4, p. 734, 2020.
- [16] H. Suparwito, A. M. Polina, and M. Budiraharjo, "Student Perceptions Analysis of Online Learning: A Machine Learning Approach," *Indones. J. Inf. Syst.*, vol. 4, no. 1, pp. 64–75, 2021.
- [17] T. L. Nikmah, M. Z. Ammar, Y. R. Allatif, R. M. P. Husna, P. A. Kurniasari, and A. S. Bahri, "Comparison of LSTM, SVM, and naive bayes for classifying sexual harassment tweets," *J. Soft Comput. Explor.*, vol. 3, no. 2, pp. 131–137, 2022.
- [18] E. A. A. Zahra, Y. Sibaroni, and S. S. Prasetyowati, "Classification of Multi-Label of Hate Speech on Twitter Indonesia using LSTM and BiLSTM Method," *JINAV J. Inf. Vis.*, vol. 4, no. 2, pp. 170–178, 2023.