IMPLEMENTATION OF CHATBOT FOR MERDEKA BELAJAR KAMPUS MERDEKA PROGRAM USING LONG SHORT-TERM MEMORY

Muhammad Rahaji Jhaerol[1], Sudianto Sudianto[2]

[12]Informatics Engineering, Institut Teknologi Telkom Purwokerto
email: 19102144@ittelkom-pwt.ac.id[1], sudianto@ittelkom-pwt.ac.id[2]

Abstract

Good service can help the organization improve efficiency and effectiveness in operations. Optimal service can also improve the customer experience and provide added value to an organization that provides services. One of the services that can be optimized is the Merdeka Belajar Kampus Merdeka (MBKM) program which is a learning program organized by the Ministry of Education, Culture, Research, and Technology (Kemendikbudristek), especially MBKM services at the Institut Teknologi Telkom Purwokerto (ITTP). The problem is that the MBKM service at ITTP is not optimal due to inaccessibility to anyone and so many programs available. Thus, resulting in not optimal services provided. Therefore, this study aims to implement a Chatbot service in the MBKM program at ITTP. The method used in building a Chatbot service is the Deep Learning Long Short-Term Memory (LSTM) algorithm. LSTM is a type of artificial neural network architecture that matches text data. The results show an accuracy score of 100% and a loss of 0.121%. Meanwhile, the results of the further evaluation are in the form of average weights consisting of precision, recall, and F1-score, respectively of 100%, 100%, and 100%.

Keywords *:* academic services, ai-chatbot, Deep Learning, LSTM, MBKM

INTRODUCTION

The development of information and communication technology has helped a lot and provided rapid changes for science [1], such as in several fields, namely health [2]–[4], agriculture [5]–[8], business [9], and education [10], [11]. The use of this technology aims to improve more optimal services so that it is expected to make consumers or users can be helped in obtaining information with technology, especially in universities.

Higher Education is one of the sectors that utilize technological developments. One of them is utilizing social media as a service medium and information dissemination [12]. Social media and university websites are also a source of information for students. However, existing services tend not to be optimal because admin operators on social media and websites cannot serve 24 hours. This results in the flexibility of the service being suboptimal and seemingly long. So, technology is needed to accommodate all forms of services, such as the Merdeka Belajar Kampus Merdeka (MBKM) program, which is widely followed in every university.

Merdeka Belajar Kampus Merdeka (MBKM) is one of the components of the Merdeka Belajar policy of the Ministry of Education, Culture, Research, and Technology (Kemendikbudristek), which provides opportunities for all students to hone their abilities according to their talents and interests by entering the workforce directly as a step to prepare for a career. Therefore, to automate the MBKM program's services, artificial intelligence (AI)-based technology is needed [13]. Artificial intelligence is a science that studies how to create a system/computer that can do as humans do [14]. One example is a Chatbot or chatterbot that is used to solve the problem of the question and answer service [15]. Chatbots are designed to mimic human conversational skills using voice, text, or both [16].

Nowadays, the popularity of Chatbots in many sectors is continually increasing because they have great potential to automate user services and reduce human effort or intervention. When used in the field of education, Chatbot technology is beneficial, primarily when used in Colleges to assist students in obtaining relevant information. With

the help of Chatbot technology, campus services have now become of a higher caliber.

Previous research has been conducted by several researchers [17] entitled "Dinus Intelligent Assistance (DINA) Chatbot for University Admission Services". The research solved problems related to student admissions services by creating a Chatbot called DINA (Dinus Intelligent Assistance). DINA uses a Deep Learning-based knowledge approach and gets good results by successfully answering questions from every student's perspective.

Another study by [18] entitled "NEU-chatbot: Chatbot for admission of National Economics University" also solved problems related to the administration of National Economics University students in Vietnam using a chatbot called NEU-Chatbot. The NEU-Chatbot is applied to the official admissions fan page of the National Economics University on the Facebook platform, a well-known social network in Vietnam. NEU-chatbots get good results, with almost 98.61% of 1000 clients giving happy questions with the answers of these chatbots.

Another study conducted by [19] entitled "A Chatbot Using LSTM-based Multi-Layer Embedding for Elderly Care" created chatbots using Long Short-Term Memory (LSTM). The dataset used is in the form of daily conversations of people in the Elderly. The experiment results showed that the proposed method achieved an accuracy of 79.96% for the top 1, 93.14% for the top 5, and 94.85% for the top 10 message pairs to match. The results show that the proposed system performance in matching message pairs is much better than the base system.

Therefore, this study aims to build an AI Chatbot using the Long Short-Term Memory algorithm for MBKM program services at the Institut Teknologi Telkom Puwokerto—selection of LSTM due to the relevant algorithm for the Chatbot [19]. LSTM is a recurrent neural network with state memory and multi-layer cell structures [20]. LSTM is often used for translation, modeling, and time series prediction. It can capture long-term dependencies in sequential data. They have also been used in various other applications, including speech recognition, natural language processing, and machine translation.

MATERIALS AND METHODS
A. Academic Services

The Merdeka Belajar Kampus Merdeka Program aims to improve the quality of higher education in Indonesia by encouraging creativity and innovation among students and increasing student involvement in research and development activities [21]. This program is called the Merdeka Belajar Kampus Merdeka (MBKM) program. MBKM at Institut Teknologi Telkom Purwokerto (ITTP) in the last two years has experienced an increase in student involvement in the MBKM program to thousands of students, so this is an essential part of the process to help and assist students in involvement in the MBKM program.

Currently, academic services about MBKM still use the hotline and visit in person. Of course, this process needs to be improved with services that can be accessed anytime and anywhere to ask about the MBKM program.

B. Chatbot

A Chatbot is computer software created to mimic text and audio-based intellectual conversations with one or more humans. A chatbot can also be interpreted as a computer system that allows humans to interact with computers using natural human language [22]. Chatbot comes from two words, namely "chatbot". Chat can be likened to a written media communication activity in the computer world. A bot is a computer program with several data points that, when given input, will output an output in response. The advantages of using a Chatbot increase efficiency in answering questions and saving time. Chatbots can also help improve the customer experience by providing quick and accurate responses to desired requests and queries.

C. Long Short-Term Memory

Long Short-Term Memory (LSTM) is a recurrent neural network with state memory and multilayer cell structures [20]. Sepp Hochreiter and Jürgen Schmidhuber first proposed LSTM in 1997. This algorithm can be considered a Recurrent Neural Network (RNN) algorithm in development. The RNN algorithm uses the previous step's results as input for the running step. The downside of the RNN algorithm is that it cannot predict words already stored in long-term memory. The strength of RNN algorithms—namely, their ability to make more precise predictions based on the latest data—is retained.

In contrast, LSTM algorithms are created to address these weaknesses. The LSTM architecture includes an input gate, forget gate, and output gate where x_t refers to the current input, c_t and $c_{(t-1)}$ respectively indicating the status of the new and previous cells, and h_t and $h_{(t-1)}$ are current and previous outputs

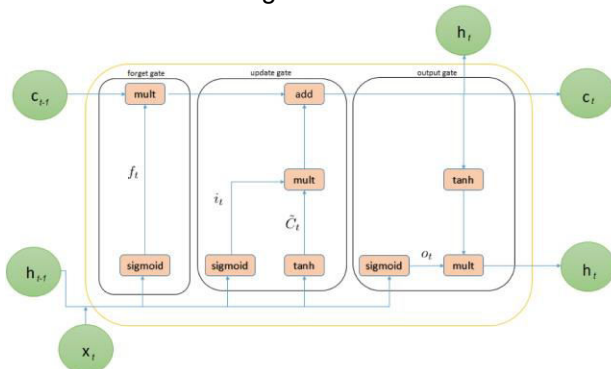respectively [23]. The internal structure of the LSTM is shown in Figure 1.

Figure 1. Internal Structure of LSTM [23]

Memory processing is performed by a section known as a gate, and the information collected by the LSTM algorithm will then be stored by the cell. This gate is implemented using the sigmoid activation function and a particular layer type called "forget gate", which determines which information should be discarded from the previous time step. LSTM also has an "input gate" that specifies which information from the current time step should be passed on to the next time step, and an "output gate" specifies which information from the current time step should be used to calculate the final LSTM result.

LSTM is often used for translation, modeling, and time series prediction. It can capture long-term dependencies in sequential data. They have also been used in various other applications, including speech recognition, natural language processing, and machine translation.

D. Methodology

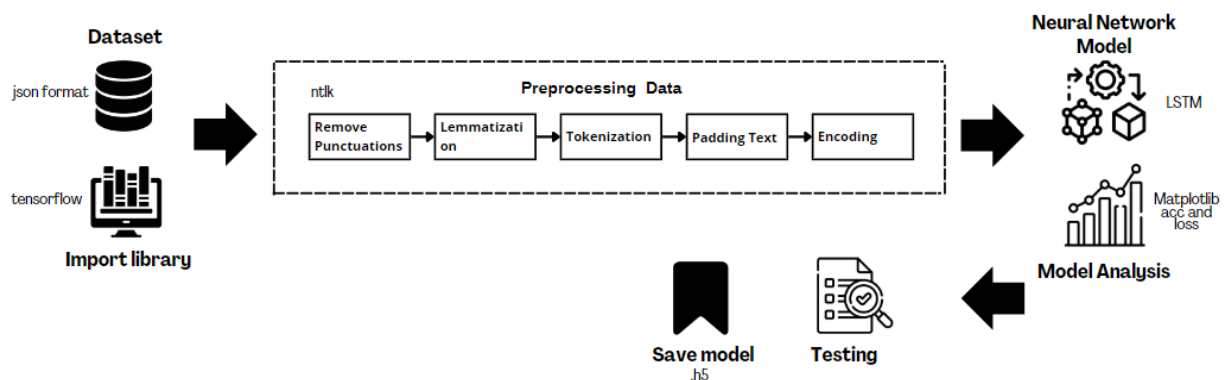The method for creating a Chatbot using the Long Short-Term Memory (LSTM) algorithm, is shown in Figure 2.

1. Dataset

The dataset used contains questions often asked to special academic services for the Merdeka Belajar Kampus Merdeka (MBKM) program and the Merdeka Campus website FAQ [21]. Table 1 shows the datasets used in building a Chatbot.

2. Library

The libraries needed in creating Chatbots use Long Shor Term Memory (LSTM) such as NumPy for mathematical computing, Matplotlib for visualization of data models, NLTK for text processing, Pandas for reading data, and Tensorflow for models on data using LSTM algorithms.

3. Preprocessing

Data preprocessing is performed for data manipulation or deletion before use. The stages of preprocessing include:
   a. Remove Punctuations

In the initial preprocessing stage for text data, special characters such as exclamation points (!), commas (,), periods as stop(.), question marks (?), and others are omitted [24]–[26]. This stage will facilitate the processing of data to be processed. Table 2 shows the differences in data after the removal of the levy.

   b. Lemmatization

Lemmatization, also known as "lematization" is the process of removing only the inflectional suffix from words and restoring the dictionary form (words present in the dictionary) from a word known as "lemma" [27]. Table 3 shows how to turn a word into a base word.

Figure 2. Stages of Creating Chatbot

Table 1. Chatbot Datasets

No	Questions	Response
1	Apa itu kampus merdeka?	Kampus Merdeka adalah kebijakan yang dikeluarkan oleh Kemendikbudristek dengan memberikan hak kepada Mahasiswa untuk mengambil mata kuliah di luar program studi selama 1 semester dan berkegiatan di luar perguruan tinggi selama 2 semester. Perguruan tinggi diberikan kebebasan untuk menyediakan kegiatan Kampus Merdeka yang sesuai dengan kebutuhan dan minat mahasiswanya
2	Apa tujuan Kampus Merdeka?	Tujuan kebijakan Kampus Merdeka adalah memberikan kesempatan kepada mahasiswa untuk memilih mata kuliah yang akan mereka tempuh berdasarkan keinginan sendiri serta mendorong mahasiswa meningkatkan soft skills serta hard skills agar siap bersaing dalam dunia global.
3	Apa persyaratan mengikuti Kampus Merdeka?	Persyaratan mengikuti program Kampus Merdeka, yaitu merupakan mahasiswa yang berasal dari Program Studi terakreditasi dan terdaftar aktif pada PDDikti.
4	Apa saja jenis program Kampus Merdeka?	Jenis kegiatan yang tersedia di program Kampus Merdeka, yaitu: Magang Bersertifikat, Studi Independen, Kampus Mengajar, Indonesian International Student Mobility Awards (IISMA), Pertukaran Mahasiswa Merdeka, Membangun Desa (KKN Tematik), Proyek Kemanusiaan, Riset atau Penelitian, dan Wirausaha
5	Apa itu Studi Independen?	Salah satu program Kampus Merdeka Kemendikbud yang bertujuan untuk mewujudkan gagasan mahasiswa dalam mengembangkan diri melalui aktivitas di luar perkuliahan dan menambah kompetensi spesifik serta praktis yang akan digunakan dalam dunia industri

Table 2. Data Discrepancies after Defunding Removal

Before	After
"ini-= adalah percobaan ? * () menghapus tanda baca pada ^^* kalimat"	ini adalah percobaan menghapus tanda baca pada kalimat

Table 3. Data Differences After Lematization

Before	After
Walk, walker, walked, walking	walk

c. Tokenization

Assigning a character string to a specific document unit is a tokenization process. Tokenization is segmenting sentences into "Tokens" and omitting specific components, including punctuation [28]–[30]. Table 4 shows how to break a sentence into token parts.

Tabel 4. Decomposition of Sentences into Tokens

Before	After
Apa itu Kampus Merdeka?	"apa" "itu" "kampus" "merdeka"
Apa persyaratan mengikuti kampus merdeka?	"apa" "persyaratan" "mengikuti" "kampus" "merdeka"

d. Padding Text

Padding is used to sort the exact text. Each sequence in padding will be made equal in length by adding a value of 0 via a suffix or prefix until it reaches the maximum length of the sequence [31]. In addition, padding can shorten the sequence to its maximum length. Table 5 shows the differences after padding text.

file storage can be transient direct or in Google Drive.

Table 5. Differences After Padding Text

Before	After
[3 4 2]	[0 0 0 3 4 2]
[1 5 3 2 7 9]	[1 5 3 2 7 9]

e. Encoding

Encoding translates categorical data, such as characters or text, into numeric or integer data according to the applied data labels. Encoding converts text in a column of data tags in this operation into numeric data using computer binaries of 0 and 1 [31].

4. Long Short-Term Memory

In this Chatbot, neural networks consist of embedding layers, one of the most potent things in natural language processing, or NLP [32], [32]. The output of the embedding layer is the text data input of the recurrent layer with the LSTM (Long Shot Term Memory Gate Layer) gate layer. Then the output is flattened, and a solid layer is used with the Softmax activation function where the implementation of this Chatbot has data labels for more than two classes [19], [20], [33].

5. Model Analys

After running model training with the Long Short-Term Memory (LSTM) algorithm and knowing the accuracy results in the last step, thus, the next stage is to analyze the model with visualization of the accuracy and loss plot to see the accuracy results of the LSTM model training algorithm [34].

6. Testing

After knowing the accuracy and loss results on the model set with the LSTM algorithm, so, the next stage is to test or test the model on a chatbot that has been previously trained using a confusion matrix [35] with training data parameters. Machine learning classification performance, where the output maybe two or more classes, is evaluated using a confusion matrix.

7. Save Model

After testing, the Chatbot has been adjusted to the sentences and answers. Thus, the chatbot model can be saved in .h5 or .pkl (pickle) format using AI Chatbot applications with websites or Android systems [36]. Model

RESULT AND DISCUSSION

The training process using the Long Short-Term Memory algorithm has 53 classes in the dataset as follows:

	classes
0	ITTPBot
1	biaya_studi_independen
2	cek_data_pddikti
3	daftar_lowongan_magang
4	data_tidak_ada_pddikti
5	gaji_magang
6	goodbye
7	greeting
8	jenis_kampus_merdeka
9	jurusan_magang
10	jurusan_studi_independen
11	lowongan_magang
12	lupa_password_akun_mbkm
13	manfaat_IISMA
14	manfaat_PMMB
15	manfaat_kampus_mengajar
16	manfaat_magang
17	manfaat_pejuang_muda
18	manfaat_pertukaran_mahasiswa
19	manfaat_studi_independen
20	mitra_IISMA
21	mitra_PMMB

22	mitra_magang
23	mitra_studi_independen
24	pelaksanaan_studi_independen
25	penjelasan_IISMA
26	penjelasan_PMMB
27	penjelasan_kampus_mengajar
28	penjelasan_kampus_merdeka
29	penjelasan_magang
30	penjelasan_pejuang_muda
31	penjelasan_pertukaran_mahasiswa
32	penjelasan_studi_independen
33	perguruan_tinggi_magang
34	perguruan_tinggi_studi_independen
35	periode_dan_pendaftaran_IISMA
36	periode_dan_pendaftaran_PMMB
37	periode_dan_pendaftaran_kampus_mengajar
38	periode_dan_pendaftaran_magang
39	periode_dan_pendaftaran_pejuang_muda
40	periode_dan_pendaftaran_studi_independen
41	persyaratan_IISMA
42	persyaratan_kampus_mengajar
43	persyaratan_kampus_merdeka
44	persyaratan_magang
45	persyaratan_pejuang_muda
46	persyaratan_pertukaran_mahasiswa
47	persyaratan_studi_independen
48	pilihan_studi_studi_independen
49	program_kampus_merdeka
50	seleksi_magang
51	terimakasih
52	tujuan_kampus_merdeka

Figure 3. Labels on Datasets

Furthermore, a model architecture with four layers is created, embedding with input_dim vocabulary+*1*, LSTM with ten hidden units,

flattened and dense with softmax activation. The compiled model uses loss sparse_categorical_crossentropy, adam optimizer, and accuracy metrics. Then the model in training using 400 epochs is obtained, and visualization of the training results is shown in Figure 4.
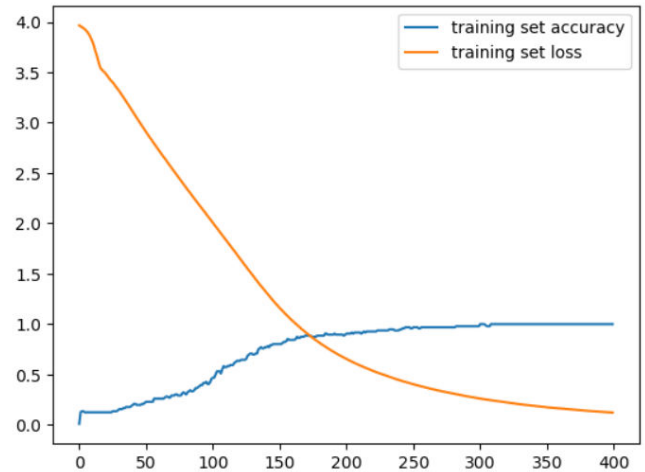
Figure 4. Visualization Accuracy and Loss Training Model

Then the results of the evaluation of the training model based on Figure 4 and Table 6 showed that the accuracy value calculated from 400 epochs was 100%, and the loss was 0.121%. The next stage is to test the model using a confusion matrix. The results of the classification report are shown in Figure 6.

Table 6. Train Model Evaluation

	Accuracy	Loss
Train	1.000	0.121

These results show that the model that has been trained is good because there is no overfitting or underfitting. Furthermore, a model evaluation of the training data was carried out to find precision, recall, and F1-score values, as shown in Figure 5.
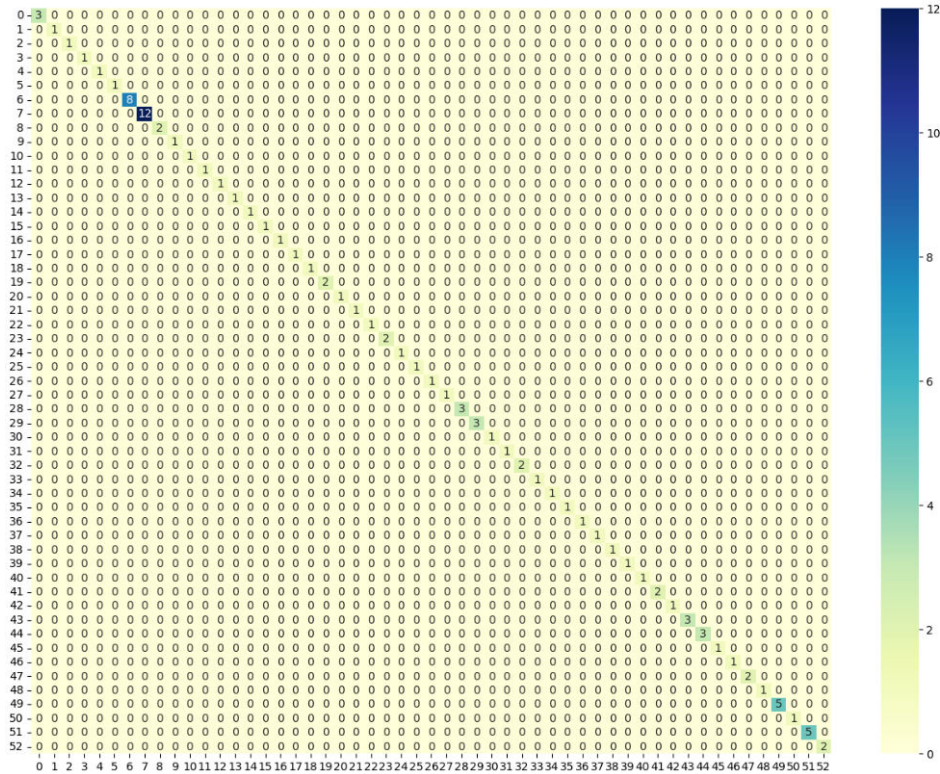
Figure 5. The Results of Confussion Matrix

In addition, from the confusion matrix results with the Long Short-Term Memory Model, the corresponding predicted values in the categories/labels could be seen in Figures 5 and 6. The average value of precision, recall, and F1-score weights are 1.00, 1.00, and 1.00, respectively.

Figure 7. Validation Chatbot

Figure 6. Precision, recall, and F1-Score values in each category using Long Short-Term Memory

The final stage is model testing, as shown in Figure 7. It can be seen that the chatbot can answer several questions related to the

explanation, purpose, and program of the independent campus. Not only that, the chatbot can also answer questions related to the Merdeka Belajar Kampus Merdeka program.

CONCLUSION

Implementing the Long Short-Term Memory (LSTM) algorithm in the application of academic service chatbots, especially in the Merdeka Belajar Kampus Merdeka (MBKM) program, through several stages such as collecting datasets, preprocessing data, making models, analyzing models, and testing models have been successfully carried out. The LSTM algorithm can perform classification by obtaining an accuracy score of 100% and a loss of 0.121%. Then further evaluation in the form of average weights consisting of precision, recall, and F1-score respectively of 100%, 100%, and 100%.

Suggestions for the following research, Chatbot is not only for the MBKM program but for other academic services by adding many datasets. In addition, MBKM Chatbot can be developed based on the knowledge base and uses relationships between words in answering a question.

ACKNOWLEDGMENT

REFERENCES

[1] S. Yoga, "Perubahan Sosial Budaya Masyarakat Indonesia Dan Perkembangan Teknologi Komunikasi," *albayan*, vol. 24, no. 1, Mar. 2019, doi: 10.22373/albayan.v24i1.3175.

[2] T. Widodo, S. Maghfiroh, S. H. B. Ginting, A. Aryaputra, and S. Sudianto, "Prediction of Covid-19 Cases in Central Java using the Autoregressive (AR) Method," *Data Science*, no. 1, 2023.

[3] A. Yani, "Utilization of Technology in the Health of Community Health," *PROMOTIF*, vol. 8, no. 1, p. 97, Jun. 2018, doi: 10.31934/promotif.v8i1.235.

[4] T. K. Putri, M. L. Arnumukti, K. Khatimah, E. Zalsabila, and S. Sudianto, "Diabetes Diagnostic Expert System using Website-Based Forward Chaining Method," *Data Science*, vol. 3, no. 1, 2023.

[5] L. S. Safitri, "Pemanfaatan Teknologi Informasi dan Komunikasi Dalam Mengakses Kredit Bagi Petani," *wbaj*, pp. 125–142, Apr. 2020, doi: 10.37950/wbaj.vi.750.

[6] Sudianto, Y. Herdiyeni, A. Haristu, and M. Hardhienata, "Chilli quality classification using deep learning," in *2020 International Conference on Computer Science and Its Application in Agriculture, ICOSICA 2020*, 2020. doi: 10.1109/ICOSICA49951.2020.9243176.

[7] S. Sudianto and R. D. Wahyuningrum, "Identifikasi Sebaran Nitrogen pada Tanaman Padi Berbasis Pengetahuan Fenologi dan Remote Sensing," *Jurnal Nasional Pendidikan Teknik Informatika (Janapati)*, vol. 11, no. 3, pp. 166–175, 2022.

[8] R. M. S. Adi and S. Sudianto, "Prediksi Harga Komoditas Pangan Menggunakan Algoritma Long Short-Term Memory (LSTM)," vol. 4, no. 2, pp. 1137–1145, 2022, doi: 10.47065/bits.v4i2.2229.

[9] A. Basry, E. M. Sari, and J. S. Raya, "Penggunaan Teknologi Informasi dan Komunikasi (TIK) pada Usaha Mikro Kecil dan Menengah (UMKM)," vol. 2, no. 3, 2018.

[10] I. K. R. Arthana, G. R. Dantes, and N. Dantes, "Optimalisasi Pemanfaatan Teknologi Informasi dan Komunikasi (TIK) dalam Bidang Pendidikan Melalui Penerapan Smart School," *Jurnal Widya Laksana*, vol. 7, no. 1, 2018.

[11] S. Sudianto, "Analisis Kinerja Algoritma Machine Learning Untuk Klasifikasi Emosi," vol. 4, no. 2, pp. 1027–1034, 2022, doi: 10.47065/bits.v4i2.2261.

[12] N. Meilinda, "Social Media on Campus: Studi Peran Media Sosial sebagai Media Penyebaran Informasi Akademik pada Mahasiswa di Program Studi Ilmu Komunikasi FISIP UNSRI," *JSM*, vol. 2, no. 1, p. 53, Jul. 2018, doi: 10.26740/jsm.v2n1.p53-64.

[13] D. Anggraini, "Kecerdasan Buatan (AI) dan Nilai CO-Creation dalam Penjualan B2B (Business-To-Business)," *JUST-SI*, vol. 1, no. 2, Dec. 2020, doi: 10.25126/justsi.v1i2.7.

[14] Y. Devianto and S. Dwiasnati, "Kerangka Kerja Sistem Kecerdasan Buatan dalam Meningkatkan Kompetensi Sumber Daya Manusia Indonesia," *InComTech*, vol. 10, no. 1, p. 19, Apr. 2020, doi: 10.22441/incomtech.v10i1.7460.

[15] K. A. Nugraha and D. Sebastian, "Chatbot Layanan Akademik Menggunakan K-Nearest Neighbor," *JSI*, vol. 7, no. 1, pp.

11–19, Mar. 2021, doi: 10.34128/jsi.v7i1.285.

[16] J. Casas, M.-O. Tricot, O. Abou Khaled, E. Mugellini, and P. Cudré-Mauroux, "Trends & Methods in Chatbot Evaluation," in *Companion Publication of the 2020 International Conference on Multimodal Interaction*, Virtual Event Netherlands, Oct. 2020, pp. 280–286. doi: 10.1145/3395035.3425319.

[17] H. Agus Santoso *et al.*, "Dinus Intelligent Assistance (DINA) Chatbot for University Admission Services," in *2018 International Seminar on Application for Technology of Information and Communication*, Semarang, Sep. 2018, pp. 417–423. doi: 10.1109/ISEMANTIC.2018.8549797.

[18] T. T. Nguyen, A. D. Le, H. T. Hoang, and T. Nguyen, "NEU-chatbot: Chatbot for admission of National Economics University," *Computers and Education: Artificial Intelligence*, vol. 2, p. 100036, 2021, doi: 10.1016/j.caeai.2021.100036.

[19] M.-H. Su, C.-H. Wu, K.-Y. Huang, Q.-B. Hong, and H.-M. Wang, "A chatbot using LSTM-based multi-layer embedding for elderly care," in *2017 International Conference on Orange Technologies (ICOT)*, Singapore, Singapore, Dec. 2017, pp. 70–74. doi: 10.1109/ICOT.2017.8336091.

[20] K. Smagulova and A. P. James, "A survey on LSTM memristive neural network architectures and applications," *Eur. Phys. J. Spec. Top.*, vol. 228, no. 10, pp. 2313–2324, Oct. 2019, doi: 10.1140/epjst/e2019-900046-x.

[21] M. Tohir, "Buku Panduan Merdeka Belajar - Kampus Merdeka," 2020, doi: 10.31219/osf.io/ujmte.

[22] A. S. Lokman and M. A. Ameedeen, "Modern Chatbot Systems: A Technical Review," in *Proceedings of the Future Technologies Conference (FTC) 2018*, vol. 881, K. Arai, R. Bhatia, and S. Kapoor, Eds. Cham: Springer International Publishing, 2019, pp. 1012–1023. doi: 10.1007/978-3-030-02683-7_75.

[23] Md. Z. Islam, Md. M. Islam, and A. Asraf, "A combined deep CNN-LSTM network for the detection of novel coronavirus (COVID-19) using X-ray images," *Informatics in Medicine Unlocked*, vol. 20, p. 100412, 2020, doi: 10.1016/j.imu.2020.100412.

[24] S. Sudianto, J. A. Marseli, N. Nugroho, R. W. A. Rumpoko, and Z. Akhmad, "Comparison of Support Vector Machines and K-Nearest Neighbor Algorithm Analysis of Spam Comments on YouTube Covid Omicron," *JTI*, vol. 15, no. 2, pp. 110–118, doi: https://doi.org/10.15408/jti.v15i2.24996.

[25] S. Chandra Ayunda Apta, N. Trivetisia, N. A. Winanti, D. P. Martiyaningsih, T. W. Utami, and S. Sudianto, "Analisis Komparasi Algoritma Machine Learning untuk Sentiment Analysis (Studi Kasus: Komentar YouTube 'Kekerasan Seksual')," *Jurnal Pengembangan IT*, vol. 7, no. 2, pp. 80–84, 2022.

[26] W. Afandi, S. N. Saputro, A. M. Kusumaningrum, H. Ardiansyah, M. H. Kafabi, and S. Sudianto, "Klasifikasi Judul Berita Clickbait menggunakan RNN-," *Jurnal Pengembangan IT*, vol. 7, no. 2, pp. 85–89, 2022.

[27] S. Symeonidis, D. Effrosynidis, and A. Arampatzis, "A comparative evaluation of pre-processing techniques and their interactions for twitter sentiment analysis," *Expert Systems with Applications*, vol. 110, pp. 298–310, Nov. 2018, doi: 10.1016/j.eswa.2018.06.022.

[28] M. Mhatre, D. Phondekar, P. Kadam, A. Chawathe, and K. Ghag, "Dimensionality reduction for sentiment analysis using pre-processing techniques," in *2017 International Conference on Computing Methodologies and Communication (ICCMC)*, Erode, Jul. 2017, pp. 16–21. doi: 10.1109/ICCMC.2017.8282676.

[29] S. Sudianto, A. D. Sripamuji, I. R. Ramadhanti, R. R. Amalia, J. Saputra, and B. Prihatnowo, "Penerapan Algoritma Support Vector Machine dan Multi-Layer Perceptron pada Klasifikasi Topik Berita," *Jurnal Nasional Pendidikan Teknik Informatika: JANAPATI*, vol. 11, no. 2, pp. 84–91, 2022.

[30] S. Sudianto, P. Wahyuningtias, H. W. Utami, U. A. Raihan, and H. N. Hanifah, "Comparison Of Random Forest And Support Vector Machine Methods On Twitter Sentiment Analysis (Case Study : Internet Selebgram Rachel Vennya Escape From Quarantine) Perbandingan Metode Random Forest Dan Support Vector Machine Pada Analisis Sentimen Twitt," *Jutif*, vol. 3, no. 1, pp. 141–145, 2022.

[31] N. R. Bhowmik, M. Arifuzzaman, and M. R. H. Mondal, "Sentiment analysis on Bangla text using extended lexicon dictionary and deep learning algorithms," *Array*, vol. 13, p. 100123, Mar. 2022, doi: 10.1016/j.array.2021.100123.

[32] E. Kasthuri and S. Balaji, "Natural language processing and deep learning chatbot using long short term memory algorithm," *Materials Today: Proceedings*, p.

S221478532103056X, May 2021, doi: 10.1016/j.matpr.2021.04.154.

[33] R. Jing, "A Self-attention Based LSTM Network for Text Classification," *Journal of Physics: Conference Series*, vol. 1207, no. 1, 2019, doi: 10.1088/1742-6596/1207/1/012008.

[34] A. A. Rizal and S. Soraya, "Multi Time Steps Prediction dengan Recurrent Neural Network Long Short Term Memory," *matrik*, vol. 18, no. 1, pp. 115–124, Nov. 2018, doi: 10.30812/matrik.v18i1.344.

[35] P. Muangkammuen, N. Intiruk, and K. R. Saikaew, "Automated Thai-FAQ Chatbot using RNN-LSTM," in *2018 22nd International Computer Science and Engineering Conference (ICSEC)*, Chiang Mai, Thailand, Nov. 2018, pp. 1–4. doi: 10.1109/ICSEC.2018.8712781.

[36] S. R. Kodandaram, "Improving the Performance of Neural Networks," *International Journal of Scientific Research*, vol. 7, no. 5, 2021.