Virtual Assistant for Thesis Technical Guide Using Artificial Neural Network

[1]Mohammad Ovi Sanjaya, [2]Saiful Bukhori, [3]Muhammad 'Ariful Furqon
[1,2,3]Departement of Informatics, Faculty of Computer Science, Universitas Jember
Email: [1]movi.sanjaya12@gmail.com, [2] saiful.ilkom@unej.ac.id, [3]ariful.furqon@unej.ac.id

Article Info

Article history:
Received Jun 12th, 2023
Revised Jul 28th, 2023
Accepted Aug 15th, 2023

Keyword:
Artificial Neural Network
Chatbot
Confussion Matrix
Machine Learning Model
Stochastic Gradient Descent

ABSTRACT

This study focuses on finding best practice for Artificial Neural Network (ANN) implementation in the information system for student's thesis technical instructions. The machine learning model applied sequential model, it means ANN only use 1 input layer, a hidden/dense layer and 1 output layer. The Stochastic Gradient Decent (SGD) method was applied into data training process. The results of this study are chatbot applications, and model testing using the confusion matrix. The result of model evaluation are 99.49% accuracy and 91% in F-1 score.

Corresponding Author:
Mohammad Ovi Sanjaya,
Departement of Informatics, Faculty of Computer Science,
Universitas Jember,
Jl. Kalimantan Tegalboto No.37, Krajan Timur, Sumbersari, Kec. Sumbersari, Kabupaten Jember, Jawa Timur.
Email: movi.sanjaya12@gmail.com

DOI: http://dx.doi.org/10.24014/ijaidm.v6i2.23473

1. INTRODUCTION

Thesis is a term used in Indonesia to describe a research paper as an explanation of the results of undergraduate research, in which a problem/phenomenon of a particular discipline is discussed in accordance with the rules of that science. The author developed a Chatbot to answer questions about technical instructions for thesis technical instructions for the Faculty of Computer Science, University of Jember with an Artificial Neural Network (ANN) approach which can help provide information related to thesis technical instructions according to the 2022 Thesis Technical Manual, Faculty of Computer Science, University of Jember. The author uses chatbot media because it is more practical and has a quick response in answering questions.

Chatbots are becoming more popular as a way to interact with computers. However, chatbots that can understand natural language and generate meaningful responses are still under development [1]. The chatbot was tested on a collection of human-human conversations, and the results demonstrated that it can generate responses that are both pertinent and informative. The chatbot can also handle a variety of user queries, including questions, requests, and instructions [2].

Research on chatbot application architecture in the university realm by Villegas-Ch, et al. in 2020, using concepts related to the smart campus and how to implement it using AI itself. Then we get a chatbot model using AI for learning needs. The development of chatbots depends on several parameters such as the level of knowledge and infrastructure prevailing in universities [3] [4]. In this research, the chatbot architectural model obtained from research conducted by Villegas, et al. Processing of input and output data uses a series of natural language processing (NLP) and machine learning processes. In Those research does not discuss machine learning methods and the results of chatbot trials with certain methods. So that technical performance is not known regarding the chatbot application being developed.

Previous research conducted by Anran Jiao, Nankai University, in 2020 found that the test results used a confusion matrix, chatbot implementation using the RASA Natural Language Understanding (NLU) and Recurrent Neural Network (RNN) methods had an accuracy rate of 1.0, entity integrity was 0.95, and sentence integrity was of 0.92 [5]. Based on these test results it can be concluded that the level of accuracy of the response results from the implementation of RASA NLU and RNN method are high. Therefore this research was conducted to evaluate the results of ANN implementation in chatbot applications with the architecture that has been discussed in the literature review.

ANN is is a series of units or nodes that are interconnected through a processing element whose functionality is similar to how neurons work in animals [6]. Although it is a multi-layered algorithm, the chance of gradual damage is low. Rather, it happens over a period of time, giving plenty of time to correct the slips. ANN stores information throughout the network, leaving very little room for the entire system to crash due to the lack of a few bits of information. It is this property that makes the ANN more fault tolerant than others [7]. They are popular for their multitasking abilities, as they use a multi-layered system where information is stored in each node, developing the ability to generate results by comparing one event to previous ones. In one of ANN can out perfome some models [7] [8] [9].

In this study the authors created a machine learning model using the ANN method to determine the quality of the final model. Author uses Rectified Linear Unit (ReLU) as input and dense layer's activation function, while in the output layer using softmax method. Machine learning model was generated using Stochastic Gradient descent (SGD) method and applying Nesterov Accelerated Gradient to boost the quality of data training process and the quality of it's model [10] [11]. Also this research has come in to mind because ANN is rarely been considered in text classification or sentiment analysis literature [7].

The author hopes that this research could help future ressearch by implementing ANN best practice when it comes with chatbot case. In this study the authors used accuracy and F-1score as parameters for measuring the quality of machine learning models to help finds it's best practice.

2. RESEARCH METHOD

Users interact with the system by inputing text (Figure 1). The text data from the chat will be passed through the UX interface to the integrator. Integrator is part of information processing. First, information extraction is carried out, in this process is about breaking the input information into a format so that it can be read by the machine, after that the information is processed with the ANN data model using business logic according to the information requested by the user. ANN training is the process of producing a model using a knowledge base database that has been prepared based on data taken from the 2022 Faculty of Computer Science Thesis Manual.

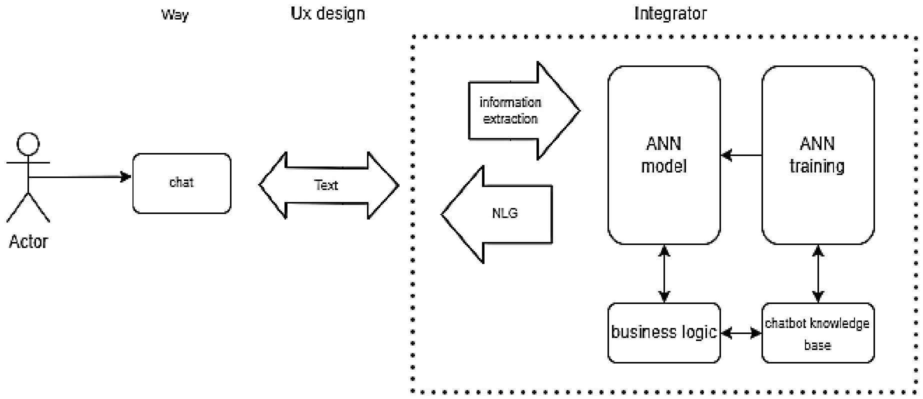


Figure 1. Chatbot Architecture Framework

At this stage, the implementation of ANN is carried out according to the design as the basis for making the system. The ANN implementation uses the Python programming language. The ANN implementation uses 128 nodes in the input layer, then 64 nodes in the dense/hidden layer, and the last one selects as many nodes as the length of the TrainY data (Figure 2). TrainY is a data tag in an intents knowledge-base. The number of nodes used in each layer is the result of the trial and error method. Trial and error is a method that repeating data training proses without using a certain patent so that an optimal result is obtained. The optimal results can be seen from the results of the data training output log by comparing the loss values and accuracy.

The next process is the requirements are changed to a chatbot knowledge base design that will be implemented in the system. This process focuses on the chatbot knowledge base that is used to be able to answer user questions. The data entered as a Knowledge-base is taken from the Thesis Technical Manual of the Faculty of Computer Science, University of Jember 2022. The chatbot's knowledge base include 106 intent's clases and 254 keyword total for each classes contain around 2 or 3 keywords. The example of chatbot's knowledge base structure is shown by Figure 3.

In pre-processing state, author uses several Natural Language Processing (NLP) to help the computer can understand what talking about. NLP can be divided into two main tasks: natural language understanding (NLU) and natural language generation (NLG). NLU is the task of understanding the meaning of text, while

NLG is the task of generating text [12]. NLU is a field of artificial intelligence that enables machines to understand natural language. NLU systems are able to extract meaning from text by identifying concepts, entities, emotions, keywords, and other linguistic features [12].

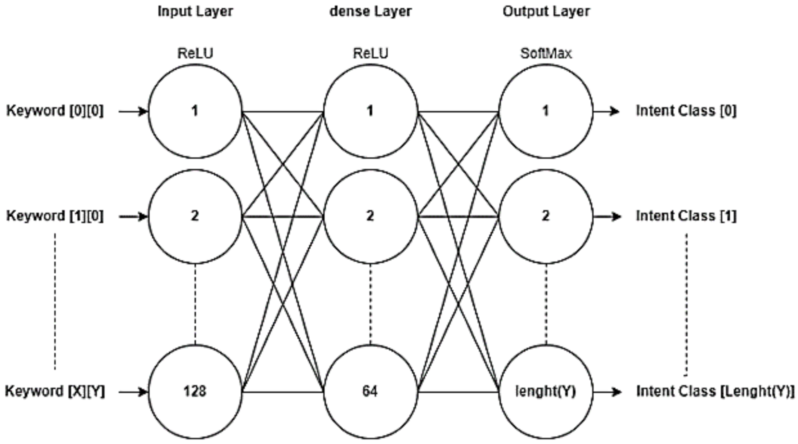
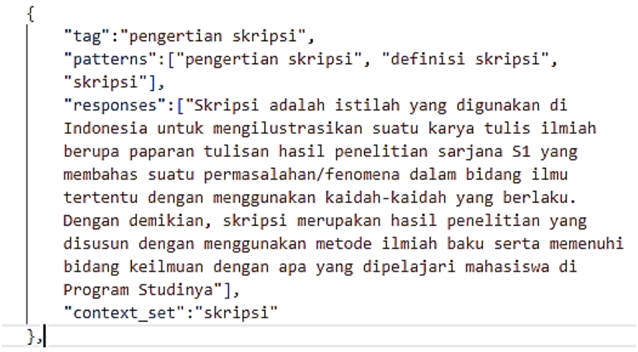


Figure 2. ANN Design



```
{
    "tag":"pengertian skripsi",
    "patterns":["pengertian skripsi", "definisi skripsi",
    "skripsi"],
    "responses":["Skripsi adalah istilah yang digunakan di
    Indonesia untuk mengilustrasikan suatu karya tulis ilmiah
    berupa paparan tulisan hasil penelitian sarjana S1 yang
    membahas suatu permasalahan/fenomena dalam bidang ilmu
    tertentu dengan menggunakan kaidah-kaidah yang berlaku.
    Dengan demikian, skripsi merupakan hasil penelitian yang
    disusun dengan menggunakan metode ilmiah baku serta memenuhi
    bidang keilmuan dengan apa yang dipelajari mahasiswa di
    Program Studinya"],
    "context_set":"skripsi"
},
```

Figure 3. Chatbot's Knowledge Base Structure

In NLU processing state, author using tokenization, lemmatization, and bag of words. Lemmatization is the process of reducing various tenses into one form to facilitate analysis. Lemmatization is a process of reducing a word to its base form, also known as its lemma. It does not remove the suffix of a word, but rather uses a vocabulary to identify the base form of the word [13]. An example is "eat", "eating", "ate", and "eats" are forms of the word "eat", so all these words are classified as the word "eat". Tokenization is the process of dividing large continuous text into discrete units [12]. An example is "this is a nice sentence". Each word in the sentence is separated into "this", "is", "a", "nice", and "sentence". Each word that has been separated is called a token.

Chatbot knowledge base was processed by bag of words (BOW) algorithm. BOW is an NLP algorithm that converts a sentence of text into a vector form. BOW is widely used in the text weighting process [9]. This algorithm is mainly used to extract feature sets from text during data pre-processing. The value of each word is calculated by the frequency of each word in a sentence (Table 1).

Table 1. Bag of Words Structure

index	word[0]	word[1]	word[2]	word[3]	…	word[n-1]	word[n]	intents
Keyword[0]	0	1	2	0	…	0	0	Intent[0]
Keyword [1]	1	0	0	1	…	0	1	Intent[1]
…	…	…	…	…	…	…	…	…
Keyword [n-1]	1	0	0	0	…	1	0	Intent[n-1]
Keyword [n]	0	0	1	0	…	0	0	Intent[n]

The bag-of-words model would represent a sentences as a vector model, each of which would contain the number of times each word appears in the sentence. For example, the vector for the first sentence would be something like [1, 1, 1, 0, 0], etc. where the 1s represent the number of times each word appears in the sentence.

ANN is a combination of several artificial neurons, which are known as perceptons, which are arranged in several layers. The minimal form of a neural network is composed of one percepton or single percepton (Equation 1).

$$y = f(z) \tag{1}$$

A single percepton will process the input by multiplying the input value (x) with the weight (w) and then adding up the bias (b). The results of input processing (z) will then be passed to the activation function (f). the output of a single percepton is the result of the activation function (y) (Equation 2). Activation Function is a function that mimics whether a biological neuron is active or not based on the input received. Each node receives the number of input products and initialization of the weights randomly along with the bias statistics for each layer [14]. In this research the activation function will be ReLU and Softmax.

$$z = \sum_{i=1}^{m} w_i x_i + b \tag{2}$$

Rectified Linear Unit (ReLU) is a non-linear function that outputs directly if the input is positive but if it is negative it will output zero (Equation 3) [15].

$$f(x) = \max(0, x) = (x + |x|)/2 = \begin{cases} x & if\ x > 0 \\ 0 & otherwise. \end{cases} \tag{3}$$

The ReLu function acts as a linear function for positive values and acts as non-linear when values are negative. ReLU is not included in the linear function because in a linear function, if simplified it can form the formula to y = ab + x, which forms a straight line in the graph [15].

Softmax is an activation function implemented in the outer layer to determine the output. The softmax function gets the raw output vector from the results of the training data and then returns the predicted value of the probability score (Equation 4) [17][18].

$$f(z)_i = \frac{e^{z_i}}{\sum_{j=1}^{N} e^{z_j}} \tag{4}$$

So that the i-th entry at the output of the softmax function is a prediction of the input probability of class i. The output of the softmax function has a value between 0 and 1. While in this research is multi-class case, all of the neuron's output of the function if added up will produce a value of 1 [18].

While the training process uses SGD method and apllying Nesterov Accelerated Gradient. Stochastic Gradient Descent (SGD) is an iterative function used to optimize neural networks. SGD function to find the local minimum value of a function. SGD works according to parameter updates for each training data sample $x^{(i)}$and $y^{(i)}$ (Equation 5) [19].

$$\theta = \theta - \eta . \nabla_{\theta} J(\theta; x^{(i)}; y^{(i)}) \tag{5}$$

Gradient group derivation is performed for redundant calculations in the computation of the dataset each time there is an update of the parameters. SGD can eliminate redundancy by performing updates at a time [19]. SGD frequently performs updates with high variance which causes the objective function to fluctuate greatly [19].

Nesterov accelerated gradient (NAG) is a type of gradient descent that can be used to make the algorithm converge more quickly [20]. NAG works by taking a "look ahead" at where the parameters will be after the next gradient update, and then using that information to calculate the current update. This helps to prevent the algorithm from overshooting the minimum, which can happen with vanilla gradient descent. This has the effect of preventing the algorithm from overshooting the minimum, which can happen with vanilla gradient descent. NAG is also more robust to noise in the gradients, which can make it a better choice for problems where the gradients are not very smooth [20].

The chatbot system testing is using the confusion matrix approach. The purpose of this test is to assess the quality of the ANN model that has been made. The confusion matrix is a predictive analysis tool that displays and compares true values with model predictive values, which can be used to generate evaluation metrics such as precision, accuracy, recall, and F1-score. In the confusion matrix there are four variables in the classification process, namely, False Positive (FP), True Positive (TP), True Negative (TN), and False Negative (FN) (Figure 4). TP is correctly detected positive data. TN is the number of negative data correctly detected. FP is negative data but detected as positive data. While the FN data is positive but detected as negative data [21].

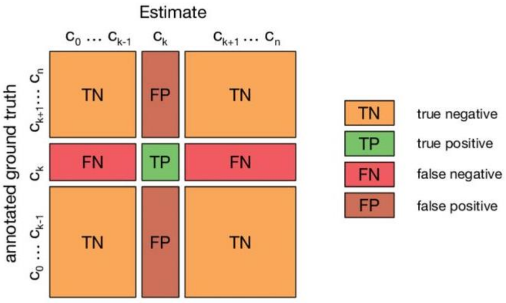


Figure 4. Confussion Matrix Table Plotting

Classification on intent, entity extraction, and responses given by chatbots in the system can be measured with precision, recall, F-1 score, and accuracy. The test data is obtained from the results of input testing in the chatbot application. Precision is the ratio of positive correct predictions compared to all predictions that have a positive value (Ecuation 6) [22].

$$\frac{\mathrm{TP}}{(\mathrm{TP+FP})} * 100\% \tag{6}$$

Recall is the ratio of correct positive predictions or TP compared to all correct positive data (Equation 7) [22].

$$\frac{\mathrm{TP}}{(\mathrm{TP+FN})} * 100\% \tag{7}$$

F-1 score is a comparison of the average precision and recall which is weighted (Equation 8) [22]. The F1 score can be deciphered as a consonant cruel of the accuracy and review, where an F1 score comes to its best esteem at 1 and most noticeably awful score at 0. For imbalance data like information retrieval or chatbot, author use this type of measurement because the accuracy rate may not be sufficient [23] [24].

$$2 * \frac{Recall*Precission}{Recall+Precission} * 100\% \tag{8}$$

Accuracy is one metric for evaluating classification models. Informally, accuracy is the fraction of predictions our model got right (Equation 9) [22].

$$\frac{\mathrm{TP+TN}}{(\mathrm{TP+FP+FN+TN})} * 100\% \tag{9}$$

A confusion matrix is a table that summarizes the performance of a classifier. It shows how many instances were correctly classified, incorrectly classified, and so on. It provides a deeper insight into how the classifier is operating, and it can reveal weaknesses that can be addressed to improve the model's performance. The confusion matrix is a more comprehensive way of evaluating classifier performance than simply computing precision, recall, or other evaluation measures [22]. Overall, the confusion matrix is a powerful tool for evaluating classifier performance and identifying areas for improvement.

3. RESULTS AND ANALYSIS

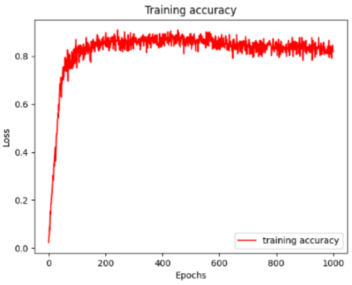
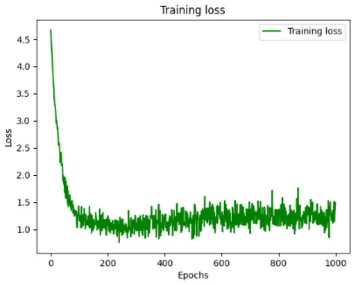
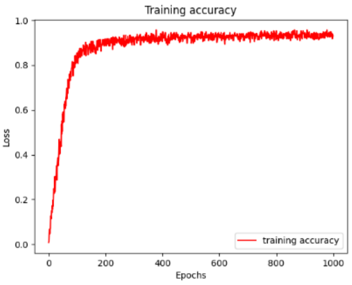
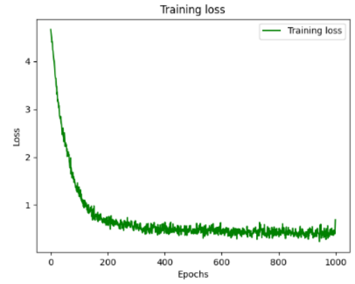
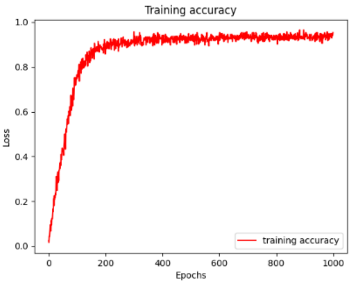
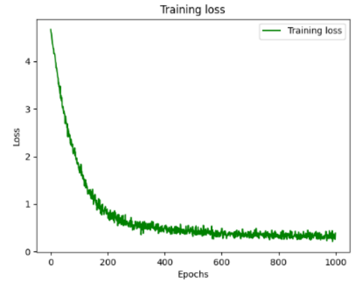
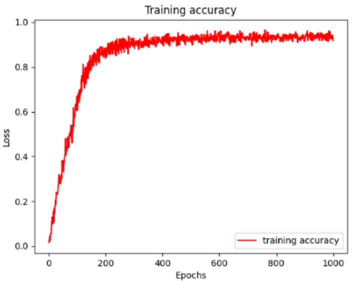
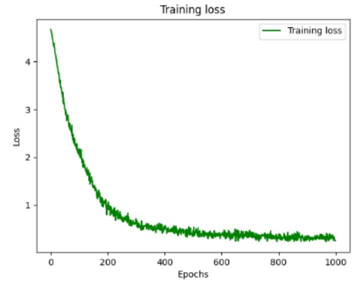
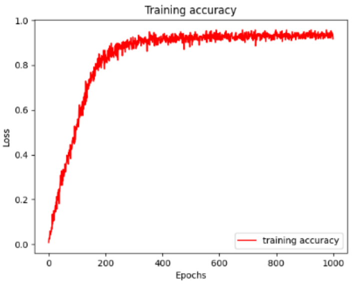
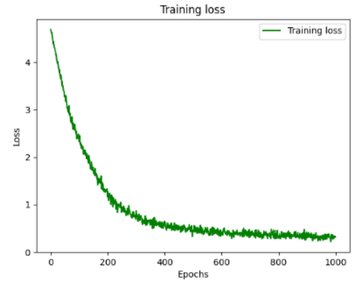
This section focuses on machine learning model training result, system testing that has been implemented in accordance with the system design chapter.

3.1. Data Training Result

Data training uses the SGD function which has learning rate, decay, momentum, and nesterov parameters. The learning rate used in this study is the default value of the SGD function in Tensorflow. The default value is 0.01. The selection of the default value is because if this value is a stable and optimal value. Changing the learning rate value can have an impact on the instability of the data training process [16]. The decay rate uses the default value of the SGD function. The decay rate used is 1e-6, where e is the weight of the learning rate in the ongoing training.

The momentum value used in this implementation is 0.6. This value is obtained from the method of trial and error until the optimal value is obtained. The optimal value obtained is seen from the development of the loss value in each epoch (Table 2). By using the true value for the nesterov parameter, SGD uses the Nesterov Accelerated Gradient when conducting data training.

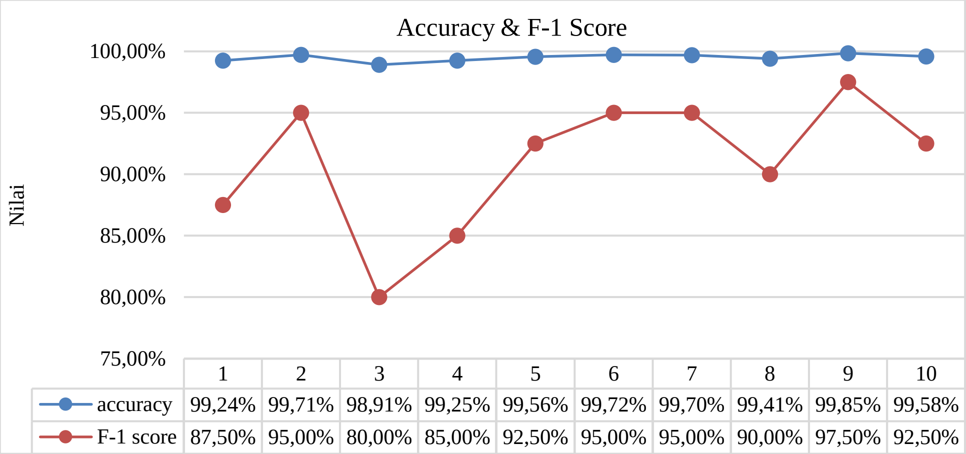
Table 2. Accuracy and Losss Growth Graph

Momentum	Accuracy Graph	Loss Graph
0.9		
0.8		
0.7		
0.6		
0.5		

The 0.6 and 0.5 momentum value’s result is slightly the same, the result of 0.5 is have more stable growth in loss value but in the accuracy growth is little bit more unstable than 0.6 value of momentum. The main purpose of momentum parameter is to speed up the learning or training process [15]. From this data training result author uses 0.6 because it is more stable both accuracy and loss growth for each epoch.

3.2. Model Evaluation

Model testing uses the confusion matrix method. Testing using the confusion matrix aims to determine the accuracy of the machine learning model. At this stage the test is run ten times on the grounds beacause each test run the results obtained may be different from the previous test. After testing ten times, the measurement variables were recorded and the average value for each variable will be calculated (Figure 5).



	1	2	3	4	5	6	7	8	9	10
accuracy	99,24%	99,71%	98,91%	99,25%	99,56%	99,72%	99,70%	99,41%	99,85%	99,58%
F-1 score	87,50%	95,00%	80,00%	85,00%	92,50%	95,00%	95,00%	90,00%	97,50%	92,50%

Figure 5. 40 : 60 Accuracy and F-1 Score Graph

The test uses the "train_test_split" function which divides the data into training and test sets. Using the “train_test_split” function can reduce overfitting, which occurs when a model performs well on the training data but fails to generalize to new data. Overfitting is an unwanted machine learning behavior that occurs when a machine learning model provides accurate predictions for training data but not for new data.

Table 3. Train, Test, Split Testing

Test : Train	Accuracy (%)	F-1 Score (%)
10 : 90	97.77	88
20 : 80	98.66	86.5
30 : 70	99.17	88.33
40 : 60	99.49	91

The high value of the measurement variable is also influenced by the application of SGD optimization in the data training process. So that the overall error can be optimized. The model's ability to correctly classify keywords in the knowledge-base data is 99.49% through accuracy calculations. The model's ability to classify keywords correctly when the model predicts that keywords are out of context but keywords are detected in context is 91% through precision calculations. While the value of the model's ability to classify keywords correctly for data that is in the knowledge-base is 91% through recall calculations. While the F-1 score is a comparison between precision and recall which is weighted. Assessment using the F-1 score is used in the chatbot model because the amount of data for each class is not balanced or is dynamic. This dynamic is influenced by the development of keywords in the chatbot when there are updates. The results of the calculation of the F-1 score are 91% so that the model can be considered to have very good quality.

3.3. Functional Testing

Chatbot testing was carried out in the Telegram application. Testing by providing a sample of questions regarding the thesis technical instructions so that it can indicate that the chatbot has been successfully implemented on Telegram (Table 4). This test focuses on the input and output provided by the chatbot. In addition, this test is to see whether the response given by Chabot is the same as the design that has been made.

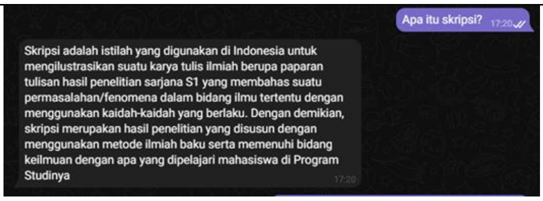
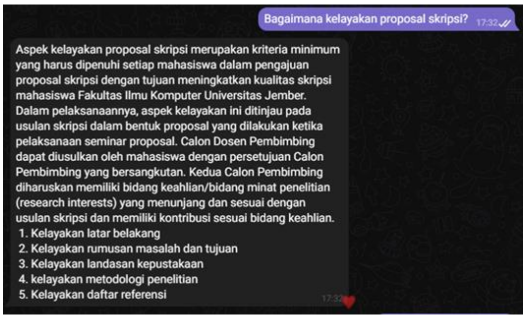
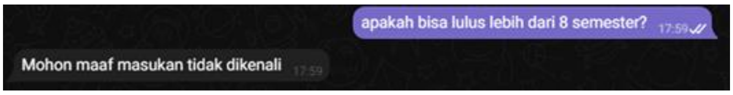
Each category in the chatbot knowledge-base mindmap creates 1 sample question and then inputs it into the chatbot application. After input, the chatbot will reply to questions according to input. This shows that the chatbot system that was built has succeeded in running according to design (Table 5).

Table 4. Functional Testing

Testing	Chatbot implementation result in Telegram.
Purpose	To find out if there is a discrepancy between the input and the output provided by the chatbot.
Testing Procedures	1. Enter sample questions on the chatbot. 2. Documenting the results of chatbot answers.

Expected Test Results	Chatbot can run according to its function.
Obtained Test Results	Chatbot can run according to its function.

Table 5. Functional Testing Result

No.	Knowledge-base	Screenshoot
1	Thesis ("Skripsi")	 Apa itu skripsi? 17:20 Skripsi adalah istilah yang digunakan di Indonesia untuk mengilustrasikan suatu karya tulis ilmiah berupa paparan tulisan hasil penelitian sarjana S1 yang membahas suatu permasalahan/fenomena dalam bidang ilmu tertentu dengan menggunakan kaidah-kaidah yang berlaku. Dengan demikian, skripsi merupakan hasil penelitian yang disusun dengan menggunakan metode ilmiah baku serta memenuhi bidang keilmuan dengan apa yang dipelajari mahasiswa di Program Studinya 17:20
2	Thesis proposal	 Bagaimana kelayakan proposal skripsi? 17:32 Aspek kelayakan proposal skripsi merupakan kriteria minimum yang harus dipenuhi setiap mahasiswa dalam pengajuan proposal skripsi dengan tujuan meningkatkan kualitas skripsi mahasiswa Fakultas Ilmu Komputer Universitas Jember. Dalam pelaksanaannya, aspek kelayakan ini ditinjau pada usulan skripsi dalam bentuk proposal yang dilakukan ketika pelaksanaan seminar proposal. Calon Dosen Pembimbing dapat diusulkan oleh mahasiswa dengan persetujuan Calon Pembimbing yang bersangkutan. Kedua Calon Pembimbing diharuskan memiliki bidang keahlian/bidang minat penelitian (research interests) yang menunjang dan sesuai dengan usulan skripsi dan memiliki kontribusi sesuai bidang keahlian. 1. Kelayakan latar belakang 2. Kelayakan rumusan masalah dan tujuan 3. Kelayakan landasan kepustakaan 4. kelayakan metodologi penelitian 5. Kelayakan daftar referensi 17:32
3	Out of Context	 apakah bisa lulus lebih dari 8 semester? 17:59 Mohon maaf masukan tidak dikenali 17:59

4. CONCLUSION

Implementation of ANN in the process of determining answers to user questions can be carried out and has been implemented with optimal settings or best practices according to the results of the evaluation using the confusion matrix. In this case study the ANN used has 3 layers consisting of an input layer that uses the ReLU activation function, a hidden layer which uses the ReLU activation function, and an outer layer that uses Softmax. The results of the evaluation of the machine learning model performed using the confusion matrix produce accuracy with a value of 99.49%, F-1 score with a value of 91%The results of the evaluation are classified as very good, and it's result is higher than chatbot either using RASA NLU or RNN method. Even this result is higher than same chatbot using ANN method. It is affected by SGD and Nesterov accelerated gradient due to data training process.

REFERENCES

[1] Ali, N. Chatbot: A conversational agent employed with named entity recognition model using artificial neural network. 2020: arXiv preprint arXiv:2007.04248. 20

[2] Vamsi, G. Krishna; RASOOL, Akhtar; HAJELA, Gaurav. Chatbot: A deep neural network based human to machine conversation model. In: 2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT). IEEE, 2020. p. 1-7.

[3] Villegas-Ch W, Arias-Navarrete A, Palacios-Pacheco X. Proposal of an Architecture for the Integration of a Chatbot with Artificial Intelligence in a Smart Campus for the Improvement of Learning. Sustainability. 2020; 12(4): 1500.

[4] Villegas-Ch, W., García-Ortiz, J., Mullo-Ca, K., Sánchez-Viteri, S., & Roman-Cañizares, M. Implementation of a virtual assistant for the academic management of a university with the use of artificial intelligence. Future Internet. 2021;13(4). https://doi.org/10.3390/FI13040097

[5] Jiao, Anran. An Intelligent Chatbot System Based on Entity Extraction Using RASA NLU and Neural Network. Journal of Physics: Conference Series. 2020; 1487(12014): 1742-6596.

[6] Gurney, K. 1997. *An Introduction to Neural Networks*. London: UCL Press.

[7] M., Rodrigo, Valiati, J.F., P., Wilson. Document-level sentiment classification: An empirical comparison between SVM and ANN. Expert Systems with Applications. 2013; 40(2); https://doi.org/10.1016/j.eswa.2012.07.059.

[8] Kalarani, P., Brunda, S. Selva. Sentiment analysis by POS and joint sentiment topic features using SVM and ANN. Soft computing. 2019; 23(15); 7076-7079; https://doi.org/10.1007/s00500-018-3349-9.

[9] K. Kamran, J.M. Kiana, H. Mojtaba, M. Sanjana, B. Laura, B. Donald. Text Classification Algorithm: A survey. Information. 2022;13(2), 83; https://doi.org/10.3390/info13020083

[10] M., Luis, Granz, Maximilian, Landgraf, Tim. Chaotic Dynamics are Intrinsic to Neural Network Training with SGD. Advances in Neural Information Processing Systems. 2022; 35; 5219-5229; ISBN: 9781713871088.

[11] Wang, Junxiang, Li, Hongyi, Zhao, Liang. Accelerated Gradient-free Neural Network Training by Multi-convex Alternating Optimization. Neurocomputing. 2022; 487; 130-143; https://doi.org/10.1016/j.neucom.2022.02.039

[12] Khyani, Divya, et al. An interpretation of lemmatization and stemming in natural language processing. Journal of University of Shanghai for Science and Technology, 2021, 22.10: 350-357.

[13] Khurana, Diksha, et al. Natural language processing: State of the art, current trends and challenges. Multimedia tools and applications, 2023, 82.3: 3713-3744.

[14] Hinkelmann, K. Neural Network, P.7. Zürich: University of Applied Sciences Northwestern Switzerland. 2018: 48.

[15] Brownlee, J. A Gentle Introduction to the Rectified Linear Unit (ReLU). 2019; https://machinelearningmastery.com/rectified-linear-activation-function-for-deep-learning-neural-networks/.

[16] Chen, B. Learning Rate Schedule in Practice: an example with Keras and TensorFlow 2.0. 2020; https://towardsdatascience.com/learning-rate-schedule-in-practice-an-example-with-keras-and-tensorflow-2-0-2f48b2888a0c.

[17] Goodfellow Ian, Bengio Yoshua, Courville Aaron. Deep Learning. Cambridge: MIT Press. 2016; 180.

[18] Jiang, M., Liang, Y., Feng, X. et al. Text classification based on deep belief network and softmax regression. Neural Comput & Applic.2018; 29; 61–70; https://doi.org/10.1007/s00521-016-2401-x

[19] Bottou, Leon. Stochastic Gradient Descent Tricks. Lecture Notes in Computer Science. 2015; 7700(25). ISBN : 978-3-64-235288-1, 978-3-64-235289-8.

[20] Wu, Wei, et al. Learning dynamics of gradient descent optimization in deep neural networks. Science China Information Sciences, 2021, 64: 1-15.

[21] Narkhede, S. Understanding Confusion Matrix. 2018; https://towardsdatascience.com/understanding-confusion-matrix-a9ad42dcfd62.

[22] Krstinic, Damir, et al. Multi-label classifier performance evaluation with confusion matrix. Computer Science & Information Technology, 2020, 1.

[23] Maratea, Antonio; Petrosino, Alfredo; Manzo, Mario. Adjusted F-measure and kernel scaling for imbalanced data learning. Information Sciences, 2014, 257: 331-341.

[24] George, S.; Srividhya, V. Performance evaluation of sentiment analysis on balanced and imbalanced dataset using ensemble approach. Indian Journal of Science and Technology, 2022, 15.17: 790-797.

BIBLIOGRAPHY OF AUTHORS

Mohammad Ovi Sanjaya, currently a student in the Informatics, Faculty of Computer Science, University of Jember. Research interest: Natural Language Processing and artificial intelligence implementation in game design. Research experience including: Chatbot development using natural language understanding and feedforward neural network.

Saiful Bukhori, professor of Artificial Intelligence at the at the University of Jember. Research interest: artificial intelligence with research experience including: Development of graph mining for predicting payment system networks in RTGS based on Clearing Houses, Intelligent Agent for Serious Game of RTGS using Forest Fire Model, Parrondo's paradox based strategies in the serious game of RTGS using Forest Fire Model, Serious Game Supply Chain Management Agribusiness as a Production Planning using Cournot Model, Serious Game Relationship Between Socio-Economic and Territorial Conditions

Muhammad 'Ariful Furqon received bachelor's degree in Informatics Education from Universitas Negeri Malang in 2016. In 2017, he continued studies at the Department of Information Systems, Institut Teknologi Sepuluh Nopember. He graduated from Institut Teknologi Sepuluh Nopember in 2019. In 2021, he worked as lecturer at Informatics Department at Universitas Jember. His research focus on Knowledge Graph, Machine learning, and Natural language processing.