Journal of Intelligent Systems and Internet of Things
Vol. 13, No. 02, PP. 293-302, 2024
DOI: https://doi.org/10.54216/JISIoT.130223

A Predictive Analysis of IMDb Movie Reviews Using LSTM and ANN Models

Noor alhuda A. Salih[1], Osama A. Qasim[2], Mohammed S. Noori[2], Rabei Raad Ali[2,*], Khawla Ahmad Wali[3]

[1]Presidency of Thi-Qar University, Thi-Qar, Iraq
[2]Department of Computer Engineering Technology, Northern Technical University, 41000, Mosul, Iraq
[3]Al Turath University. English Department, Baghdad, Iraq
Emails: nooralhooa.s@utq.edu.iq; osama.hassani@ntu.edu.iq; moh.sami@ntu.edu.iq; rabei@ntu.edu.iq; Khawla.ahmad@uoturath.edu.iq

Abstract

The Machine Learning domain has made a major process with the progression of state-of-the-art technologies. Since current algorithms often don't provide palatable learning performance, it is necessary to continually upgrade them. This paper has illustrated the comparison of the Long Short-Term Memory (LSTM) model and the Artificial Neural Networks (ANN) model in the prediction of the Internet Movie Database (IMDb) website. These evaluations were then related to sentiment assessment approaches to evaluate their predicted accuracy and performances. The results demonstrate that the ANN model outperforms the LSTM model with a high accuracy rate in terms of the prediction accuracy and loss indicators for the IMDb movie review's sentiment analysis task in terms of the prediction accuracy and loss indicators for the IMDb movie review's sentiment analysis task. The accuracy of prediction on the test dataset of the ANN model is 83.5 % and the LSTM model is 83.5%. Therefore, it can be concluded that the standard artificial neural network model that was utilized is an appropriate technique for sentiment assessment tasks in IMDb rating text data.

Received: November 02, 2023 Revised: March 19, 2024 Accepted: July 14, 2024

Keywords: Long Short-Term Memory (LSTM); Artificial Neural Networks (ANN); Internet Movie Database (IMDb); Prediction Accuracy

1. Introduction

Today, in the age of digital media and online streaming platforms, the ability to predict target audience sentiment towards the film has become valuable to the production house, filmmakers, studios, and streaming offerings. As a result of the sheer amount of human-generated content in terms of writing reviews for a movie, sentiment analysis uses machine learning as an effective tool to predict expert audiences' choices and movie fulfillment. In this study, the machine learning models on movie reviews from the IMDb platform are the Long Short-Term Memory and Artificial Neural Network (ANN), which was used to predict movie reviews and results. Long Short-Term Memory (LSTM) is a model of about Recurrent Neural Network. This can be a good example of sentiment evaluation, human analysis, and opinion. It can work with text data accordingly. A more efficient approach to system analysis is represented by ANN as a benchmark which is used to compare its results with the commonly used LSTM.
This study examines the level of accuracy and efficiency in the predictions of the machine learning algorithm, which is examined by reviewing a dataset containing the opinions of a group of film critics, as well as the opinions of the filmmakers and the viewing audience and their critical reviews on the Internet Movie Database (IMDb) website. We will also consider the comparative merits of the model in terms of offering advice to viewers and film industry

professionals to make good sentiment-based ratings. Studio, streaming platforms, and film producers' practical advice will block the audience from obtaining prediction sentiment when releasing the film or a film creation work.
IMDb will align its focus on two popular machine learning models, LSTM and ANN. These two models will enable the generation of movie reviews from this platform. By applying those models to the IMDb movie reviews dataset, the system seeks to evaluate and compare the fashions' predicted accuracy, robustness, and efficiency. We hope to learn more about the advantages and disadvantages of each model through this comparison study and offer insightful advice to practitioners and scholars working in the fields of sentiment analysis and predictive analytics for the motion picture industry.
The paper is organized as follows throughout the remainder of it. A review of prior research on this topic is done in Section 2 of the literature. Section 3 explains the methodology, standards, and algorithms used in the work process. Section 4 Experimental results and the values of the suggested criteria and standards were also given and debated, along with the results of the processing algorithms. Section 5 talks about the outcomes and how to assess them. Section 6 clarifies the findings reached.

2. Related Work

In [1], the authors investigate the software used in supervised machine learning approaches, specifically K-nearest associates (KNN), Support Vector Machine (SVM), and Random Forest Model (RF) to develop a movie recommendation system. We assess the models in terms of accuracy, recall, and F1 score using Movie Lens datasets. The Random Forest Model performs better than the other models, according to the results. Additionally, they discussed the historical background of devices learning about advice systems, examined the efficacy of various models and designs, and outlined the obstacles and future research objectives in the field.
In [2], the authors examined the effectiveness of several sentiment analysis techniques, such as lexicon-based, machine learning, Bi-LSTM, BERT, and GPT-3 models by using Sentiment140 datasets and IMDB movie reviews. The objective is to decide the maximum attainable method for assessing sentiment in tweets concerning the WHO Framework Convention on Tobacco Control's Ninth Conference of the Parties (COP9) in 2021. After a partial annotation of COP9-associated tweets, strategies have been assessed using a diploma evaluation system on famous sentiment evaluation datasets. The findings suggest that BERT produced the best F1 rankings on broadly used datasets, however, GPT-three fared higher on COP9 tweets. This demonstrates how useful pre-trained fashions are for sentiment analysis in quite a few fields.
In [3], the authors explore lexicon-based and BERT neural network patterns, in addition to sentiment assessment strategies for categorizing IMDb film evaluations as positive or negative feelings. A quantitative analysis is performed at the Kaggle movie review statistics from IMDb. Standard criteria which include accuracy, precision, recall, and F1 rating score are used to teach and evaluate the BERT neural network layout and the lexicon-primarily based approach. The results show that the BERT neural network model beats the lexicon-primarily based approach, with the highest accuracy (ninety.67%) and F1 rating (0.88). The BERT model has greater precision and don't forget, making it the greenest technique for sentiment analysis in IMDb film evaluations.
In [4], the author's study dives into sentiment evaluation, an essential vicinity of Natural Language Processing, specifically for studying social sentiments across numerous platforms. With social media being any such popular manner to convey feelings, sentiment evaluation is critical for enhancing patron experiences, safety, and hassle-fixing. This paper investigates numerous deep learning methods, such as LSTM, BiLSTM, and CNN, in conjunction with Glove and Word2Vec word embeddings. The effects display tremendous enhancements in LSTM and BiLSTM models with Glove (50d) embeddings, with accuracies of 90.36% and 91 sixty-eight%. The examine additionally compares the excellent-acting version, BiLSTM with a batch size of 128, to unlabeled testing information, yielding encouraging consequences. Additionally, word clouds are created to depict sentiments from each fine and negative evaluation within the unlabeled dataset.
In [5], the authors look at provides a predictive analytic method to foresee a film's overall performance primarily based on quite a few characteristics. The movie's forged, narrative, container office sales, and audience and critic critiques are all vital concerns. This examination makes use of device studying strategies consisting of Random Forest, Decision Tree, K-Nearest Neighbors (KNN), NLP, XGBoost Classifier, and Deep Neural Network to carry out a whole analysis on the IMDB dataset. The outcomes show that the XGBoost Classifier has exceptional accuracy, providing beneficial insights for forecasting film success. Through this study, we were inspired to base the results on other forecasting techniques and movie review datasets.
In [6], the authors focused on JPEG file cluster classification to recognize the JPEG file or non- JPEG file type of every cluster in JPG file recovery. Three features (entropy, byte frequency distribution, and rate of change approach to derive cluster features) with a support vector machine (SVM) are applied to extract fingerprints for each file cluster

to be considered. The SVM is applied using two functions which are polynomial and the radial basis function. The results displayed that the accuracy of the SVM classifier with the polynomial function is higher than the SVM classifier with the radial basis function. The accuracy of the SVM with the polynomial is 96 %, and the SVM with the radial basis is 57 %.

3. Methodology

The IMDB Movie Reviews dataset is collected and preprocessed to easily tokenize the textual content records before translating them to a suitable numerical illustration. LSTM and ANN models are then created and implemented utilizing a deep mastering framework. The LSTM model has LSTM layers with dropout regularization, while the ANN model has dense hidden layers with ReLU activation capabilities. Both fashions are constructed with the applicable loss functions and optimizers. After education on the education dataset, the fashions are assessed based on parameters like accuracy, precision, recall, and F1 rating. Cross-validation and hyper parameter adjustment are used to enhance the version's overall performance [6]. The findings are evaluated, evaluating the efficacy of LSTM and ANN models for sentiment evaluation of IMDB movie reviews and highlighting applicable topics for further studies as in Figure 1.

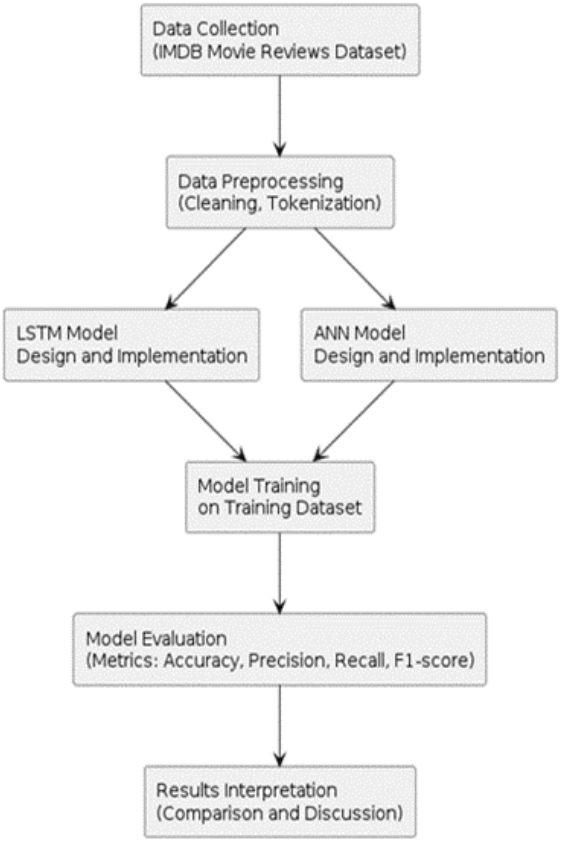

Figure 1. Predictive structure of IMDB movie reviews using LSTM and ANN models

IMDb (Internet Movie Database) provides numerous non-commercial datasets that include information approximately movies, actors, directors, and different associated entities. These datasets are meant for research functions and are free to be had for educational and non-business use [7]. For typical machine learning tasks such as sentiment analysis using the IMDB Movie Review dataset, the actual value (also known as ground truth or true score) is the sentiment assigned to each movie review in the test data set. This sensitivity is usually binary, indicating whether the evaluation is positive or negative as in Table 1.

Table 1: Movie Review and Actual Sentiment for the dataset.

MOVIE REVIEW	ACTUAL SENTIMENT
REVIEW 1	Positive
REVIEW 2	Negative
REVIEW 3	Positive
REVIEW 4	Positive

Overall, IMDb non-business datasets function as treasured resources for studying films, and exploring trends within the film industry. In our study, we specialize in reviews and evaluations of films by fans, both positive and negative, and conducting studies in related fields as in Figure 2.

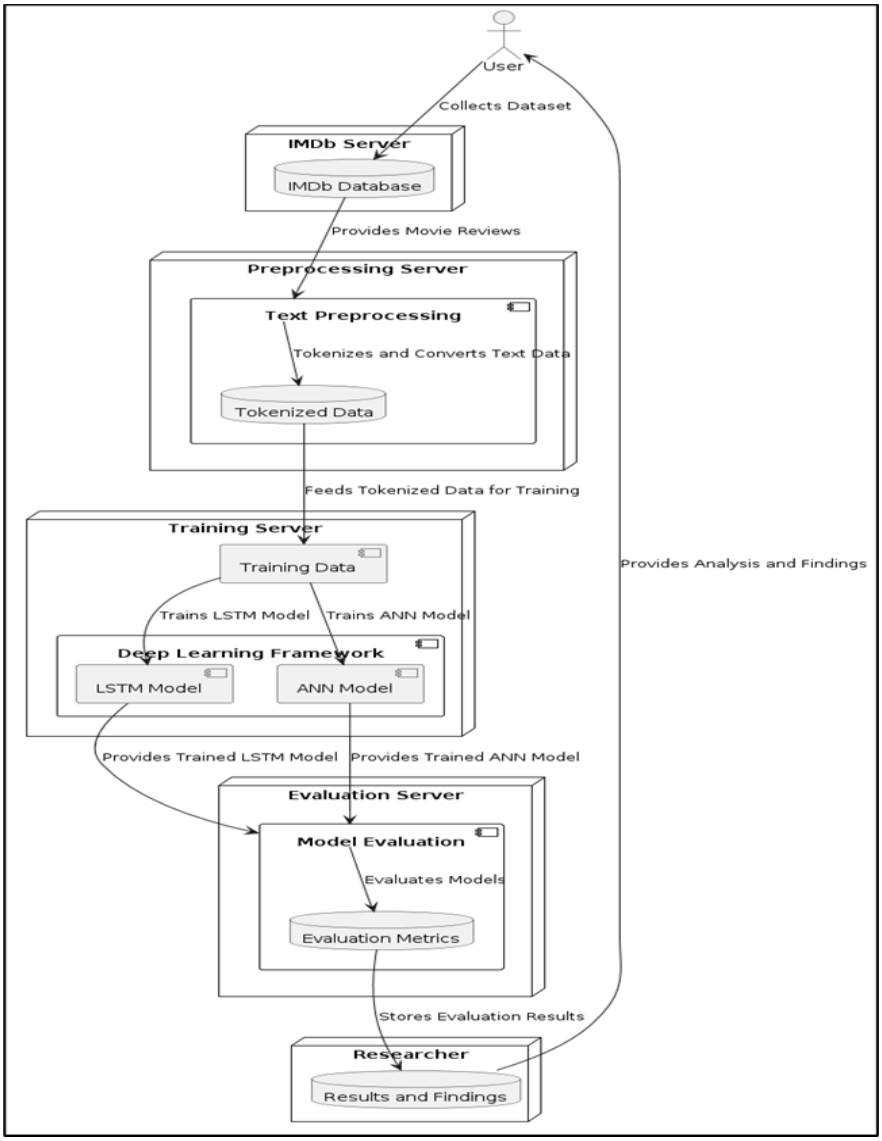

Figure 2. Stages of analyzing IMDb using ANN and LSTM

- IMDb Datasets: IMDb gives various datasets containing established data about films, inclusive of titles, release years, genres, ratings, and cast members. These datasets are to be had in exceptional formats, inclusive of plain textual content documents and TSV (tab-separated values) files, making them without difficulty on hand for evaluation [8].
- Content: The datasets normally include comprehensive information about films, together with their titles, launch dates, genres, plot summaries, consumer rankings, and opinions. They may additionally include information about the solid and group worried inside the production of each movie, which includes actors, administrators, writers, and producers.
- Access: IMDb offers these datasets at no cost on their internet site, allowing researchers and lovers to download and examine them for numerous functions. Users are required to conform to IMDb's terms of use, which normally consist of restrictions on industrial use and redistribution of the datasets.
- Applications: Researchers, data scientists, and film fanatics use IMDb datasets for a huge variety of functions, including sentiment evaluation of movie critiques, recommendation systems, trend evaluation, and academic research in fields along with statistics mining, gadget studying, and natural language processing.

There are four metrics communally provide a comprehensive evaluation of the model's performance in terms of accuracy, precision, recall, and error prediction

1- Accuracy measures the overall correctness of predictions made by a model. Acc is defined as the ratio of correctly predicted instances to the total number of cases in the dataset as presented in the following equation.

$$\text{Accuracy} = \frac{no.correct\ predictions}{\text{Total no. predictions}} \quad (1)$$

2- Precision measures the proportion of true positive predictions (correctly predicted positive instances) between all positive predictions made by the model. It is presented as:

$$\text{Precision} = \frac{True\ Positive}{True\ Positive + Fals\ Positive} \quad (2)$$

3- Recall, also known as sensitivity or true positive rate, quantifies the model's ability to correctly identify all positive instances. It is defined as:

$$\text{Recall} = \frac{True\ Positive}{True\ Positive + Fals\ Negatives} \quad (3)$$

4- Error Prediction refers to the estimation of the model's prediction error, often quantified using the following equation:

$$\text{Error Prediction} = \frac{1}{n}\sum_{i=1}^{n}(y_i - x_i)^2 \quad (4)$$

where y_i are the actual values, x_i are the predicted values, and n is the number of samples.

4. Artificial Neural Networks (ANN)

ANN are computational models primarily based on the shape and function of biological neural networks seen in the human mind. In 1943, Warren McCulloch, a neurophysiologist, and Walter Pitts, a young mathematician, created the first neural network models. They prepared a paper entitled The Logical Calculus of the Ideas Immanent in Nervous Activity about how neurons could operate [9].
They are made from connected nodes, or artificial neurons, organized into an enter layer, one or more hidden layers, and an output layer. Each neuron enters alerts and analyzes them with the usage of a weighted sum, after which applies an activation function to generate an output signal [10]. ANN changed into employed as one of the prediction models for IMDb movie evaluations.
The ANN version became educated on a dataset that included characteristics accrued from movie critiques, inclusive of textual statistics, user rankings, and style records. These traits were obtained through the input layer and then transmitted to the hidden layers, wherein the weights were changed below the training data. The output layer generated predictions approximately the emotion or other aspects of the film's opinions [11].

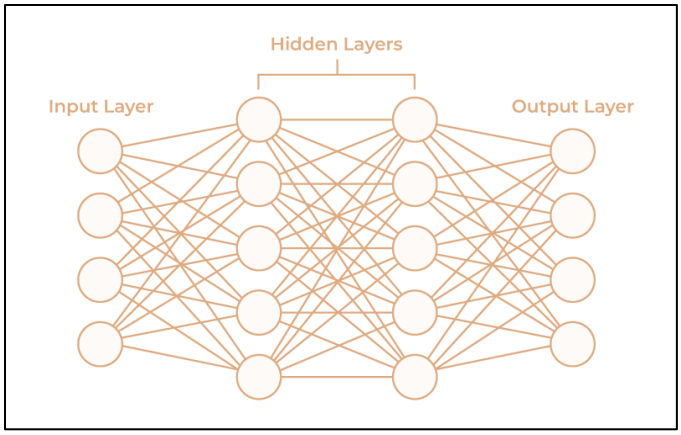

Figure 3. The ANN architecture

The ANN version was educated using backpropagation, an iterative optimization manner that modifies the weights of neuron connections to reduce the error among projected and real outputs. After schooling, the ANN version was tested on a separate take look at the dataset to decide its accuracy in predicting IMDb film evaluations [12].

5. Long Short-Term Memory (LSTM)

LSTM is a structure designed to overcome the vanishing gradient hassle and capture lengthy-time period dependencies in sequential statistics. It includes memory cells and gates that adjust the float of data, allowing it to keep critical data over extended sequences [13]. Hochreiter and Schmid Huber devised long short-term memory (LSTM) networks in 1997, which have established accuracy records in a variety of application sectors. In several speech packages and different machine learning tactics, LSTM outperformed traditional fashions, beginning to revolutionize speech recognition in 2007 [14]. The operations within an LSTM unit can be shown in Figure 4.

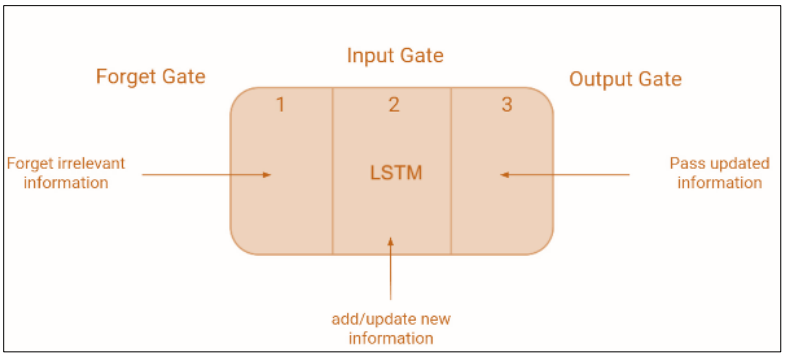

Figure 4. The LSTM architecture

- Forget Gate: Controls the flow of information from the previous cell state, allowing the LSTM to forget irrelevant information. It is calculated in the following equation.

$$\text{Forget Gate } (f_t) = \sigma\left(w_f \,.\, [h_{t-1}, x_i] + b_f\right) \quad 5$$

- Input Gate: Regulates the update of the cell state by selectively adding new information to it. It is calculated in the following equation.

$$\text{Input Gate } (i_t) = \sigma\left(w_i \,.\, [h_{t-1}, x_i] + b_i\right) \quad 6$$

Update Gate $(x_t))$ = $tanh\ (w_c\ .\ [h_{t-1}, x_i] + b_c)$ 7

- Cell State: Represents the memory of the LSTM unit and is updated over time through the forget and input gates. It is calculated in the following equation.

  Cell State (C_t) = $f_t.C_{t-1} + i_t.x_t$ 8

- Output Gate: Determines the output based on the current cell state and selectively exposes parts of the cell state.

  Output Gate (o_t) = $\sigma\left(w_o\ .\ [h_{t-1}, x_i] + b_f\right)$ 9

  Hidden State Update (h_t) = $o_t.\ tanh\ (C_t)$ 10

where: x_t is the input at the time step, h_{t-1} is the hidden state of the previous time step, W and b are weight matrices and bias vectors that are learned during training, σ denotes the sigmoid activation function, and tanh denotes the hyperbolic tangent activation function.
This work used LSTM to analyze the sentiment of IMDB film evaluations. The text records turned into preprocessed, tokenized, and transformed into numerical sequences that the LSTM version should use. The LSTM structure became set up with a couple of layers to extract skills from the sequential data correctly. The version learned to investigate the sentiment styles in the textual content using categorized IMDB film assessment information. The LSTM modified its settings through education to reduce prediction mistakes and maximize its capacity to categorize critiques as positive or negative [15].

6. Testing Results

The IMDb movie reviews dataset was used to train and evaluate the LSTM models. Larger datasets allow for more robust training and evaluation of the models, resulting in better performance. An epoch is defined as one complete run on the training data set for neural network training.
The model learns from the full dataset at each time point and updates its weights and biases to minimize the loss function. Typically, training is repeated many times until the model converges to an adequate level of performance or additional training significantly improves model performance [16]. Figure 5 shows the confusion matrix setting.

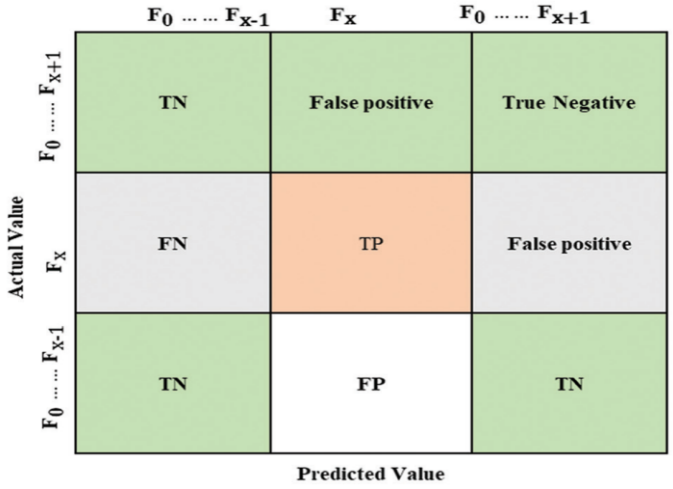

Figure 5. Confusion matrix [21]

The IMDb dataset is divided into testing and training data. The data in the dataset is divided into 70% for testing and 30% for training. Table 2 illustrates the classification setting.

Table 2: The prediction setting.

Units number (neurons)	Dropout rate	Batch Size	Epochs no.	Software
LSTM				
150	0.2	64	30	Python
ANN				
3	150	Rela	64	Python

In this study, we applied the 10-fold cross-validation learning approach in the production models to calculate the performance. Table 3 summarizes the confusion matrix results of the LSTM and ANN prediction models.

Table 3: The confusion matrix of LSTM and ANN

Model	Positive	Negative
LSTM	425	75
	90	410
ANN	430	70
	85	415

The overall evaluation of the performance and average of classification accuracy, precision, recall, and error prediction is shown in Table 4.

Table 4: The LSTM and ANN model evaluation

METRIC	LSTM	ANN
Accuracy	0.835	0.845
Precision	0.85	0.86
Recall	0.825	0.835
Error Prediction	0.165	0.155

The results illustrate that the highest accuracy 83.5 % is found in the LSTM model while the higher accuracy 84.5 % is found in ANN. Subsequently, the comparisons find that the LSTM model outperforms the ANN model in the predictive analysis of IMDb movie reviews as shown in Figure 6 and Table 5 The prediction analysis

Table 5: The prediction analysis

Measure	LSTM		ANN	
Positive	425	90	430	85
Negative	90	410	70	415
Accuracy	83.5%		84.5%	
Precision	85%		86%	
Recall	82.5%		83.5%	
Error Prediction	16.5%		15.5%	

Both LSTM and ANN models showed improved accuracies by tuning the parameters. Through the above results, this process emphasizes the importance of improving the approach to discover more optimal configurations for sentiment analysis in the IMDb movie review dataset.

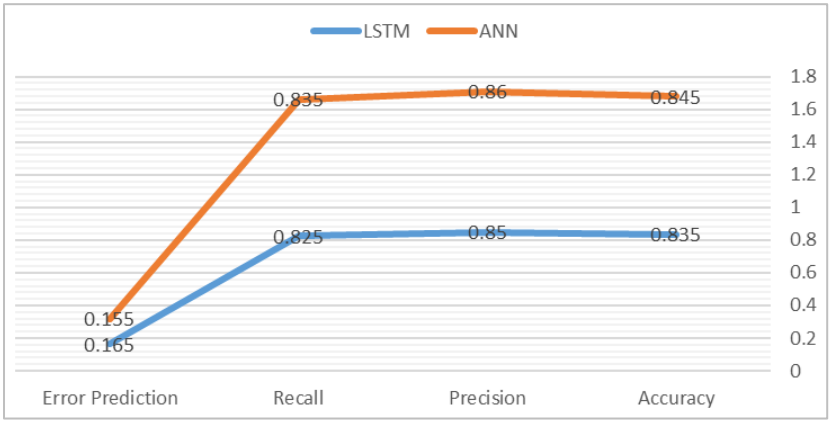

Figure 6. Measures comparison of all ANN model and LSTM model

In addition to continuing adjustments repeatedly based on the results obtained from each adjustment process to further improve the model's performance. The ANN model also has better test accuracy 84.5 % than LSTM 83.5 %. The higher accuracy indicates that the ANN model predicted better on the test data set than LSTM. Overall, according to these measures, ANN in the test loss and accuracy beat the LSTM model.

7. Conclusion

The ANN model plays higher than the LSTM model in conditions of accuracy and test loss. The ANN version outperforms the LSTM version, which had a test loss of 0.3501 and a test accuracy of 0.835, with a test lack of zero. These outcomes lead us to agree that sentiment analysis obligations on IMDB film critiques. The ANN model performs higher than the LSTM technique. The ANN version did better than the LSTM model in the context of sentiment evaluation on IMDB film reviews. This implies that well-known artificial neural networks are properly ideal for handling the complexities of sentiment evaluation obligations and the use of textual records from IMDb movie opinions and reviews. The LSTM and ANN models were improved through hyper parameters, where the LSTM model achieved an accuracy of 84.5%, which indicates improved performance compared to the initial 86.50%. As for the ANN model, it improved to an accuracy of 88.50%, which also represents an improvement over the initial 83.5%. Subsequently, it can be said that the ANN model is the best after and before optimization. This study confirms the effective role of hyper parameters in developing machine learning models and improving the accuracy of sentiment analysis.

References

[1] R. R. Ali, W. S. Al-Dayyeni, S. S. Gunasekaran, S. A. Mostafa, A. H. Abdulkader, & E. H. Rachmawanto, "Content-based feature extraction and extreme learning machine for optimizing file cluster types identification, " In Future of Information and Communication Conference (pp. 314-325). Cham: Springer International Publishing, 2022.

[2] S. Elmitwalli and J. J. F. i. b. D. Mehegan, "Sentiment analysis of COP9-related tweets: a comparative study of pre-trained models and traditional techniques," vol. 7, p. 1357926, 2024.

[3] P. S. S. V. Domadula and S. S. Sayyaparaju, "Sentiment analysis of IMDB movie reviews: a comparative study of Lexicon based approach and BERT Neural Network model," ed, 2023.

[4] D. Soam and S. Thakur, "Sentiment Analysis Using Deep Learning: A Comparative Study," in 2022 Second International Conference on Computer Science, Engineering and Applications (ICCSEA), 2022: pp. 1-6.

[5] N. Darapaneni et al., "Movie success prediction using ml," in 2020 11th IEEE Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON), 2020: IEEE, pp. 0869-0874.

[6] L R. R. Ali, N. Z. Waisi, Y. Y. Saeed, M. S. Noori, & E. H. Rachmawanto, (2023). "Intelligent Classification of JPEG files by Support Vector Machines with Content-based Feature Extraction," Full-Length Article, 11(1), 01-1.

[7] Abd, B. Obaidi, M. "Enhancing Healthcare Data Classification: Leveraging Machine Learning on ChatGPT-Generated Datasets," Journal of International Journal of Advances in Applied Computational Intelligence, vol. 5, no. 2, pp. 34-45, 2024. DOI: https://doi.org/10.54216/IJAACI.050203

[8] I. F. Ashari, R. Banjarnahor, D. R. Farida, S. P. Aisyah, A. P. Dewi, and N. Humaya, "Application of data mining with the K-means clustering method and Davies Bouldin index for grouping IMDB movies," Journal of Applied Informatics and Computing, vol. 6, no. 1, pp. 07-15, 2022.

[9] A. N. A., Al-Masri & H. Mokayed, (2021). "Intelligent fault diagnosis of gears based on deep learning feature extraction and particle swarm support vector machine state recognition, " Journal of Intelligent Systems and Internet of Things, 4(1), 26-40.

[10] K. Alakkari, A. A. Subhi, H. Alkattan, A. Kadi, A. Malinin, I. Potoroko,... & E. S. M. El-kenawy, (2023). "Forecasting COVID-19 Infection Using Encoder-Decoder LSTM and Attention LSTM Algorithms," Journal of Intelligent Systems & Internet of Things, 8(2).

[11] W. T. JHandoko, & A. N. Handayani, (2023). "Forecasting Solar Irradiation on Solar Tubes Using the LSTM Method and Exponential Smoothing," J. Ilm. Tek. Elektro Komput. dan Inform, 9(3), 649-660.

[12] A. T. Hoang et al., "A review on application of artificial neural network (ANN) for performance and emission characteristics of diesel engine fueled with biodiesel-based fuels," Sustainable Energy Technologies and Assessments, vol. 47, p. 101416, 2021.

[13] M., M. M., S. "Hybrid Metaheuristics with Deep Learning Assisted Intrusion Detection on Cyber-Physical Smart Grid Environment," Journal of International Journal of Wireless and Ad Hoc Communication, vol. 8, no. 2, pp. 53-66, 2024. DOI: https://doi.org/10.54216/IJWAC.080206

[14] C. A. Sari, I. P. Sari, E. H. Rachmawanto, E. Proborini, R. R. Ali, & I. Rizqa, "Papaya fruit type classification using LBP features extraction and naive bayes classifier, " In 2020 international seminar on application for technology of information and communication (iSemantic) (pp. 28-33). IEEE, 2020.

[15] R. Bolboacă and P. Haller, "Performance analysis of long short-term memory predictive neural networks on time series data," Mathematics, vol. 11, no. 6, p. 1432, 2023.

[16] V. Brown et al., "Core outcome set for early intervention trials to prevent obesity in childhood (COS-EPOCH): agreement on “what” to measure," International Journal of Obesity, vol. 46, no. 10, pp. 1867-1874, 2022.

[17] M. Ala’raj, M. F. Abbod, and M. Majdalawieh, "Modelling customers credit card behaviour using bidirectional LSTM neural networks," Journal of Big Data, vol. 8, no. 1, p. 69, 2021.

[18] R. Asha and S. K. KR, "Credit card fraud detection using artificial neural network," Global Transitions Proceedings, vol. 2, no. 1, pp. 35-41, 2021.

[19] Jeena, R. M., K. A., P. A., T. "Earthquake Location Forecasting In Map Using XGBOOST Algorithm," Journal of Journal of Cognitive Human-Computer Interaction, vol. 5, no. 1, pp. 42-45, 2023. DOI: https://doi.org/10.54216/JCHCI.050104

[20] M. Alhayani, N. Alallaq, and M. Al-Khiza’ay, "Optimize One Max Problem by PSO and CSA," in International Congress on Information and Communication Technology, 2023: Springer, pp. 829-839.