# Sentiment Analysis to Evaluate Different Deep Learning Approaches

**Sheikh Muhammad Saqib and Tariq Naeem**
Saqibsheikh414@gmail.com tariknaim@hotmail.com
Institute of Computing and Information Technology, Gomal University. PAKISTAN

**Abstract**
The majority of product users rely on the reviews that are posted on the appropriate website. Both users and the product's manufacturer could benefit from these reviews. Daily, thousands of reviews are submitted; how is it possible to read them all? Sentiment analysis has become a critical field of research as posting reviews become more and more common. Machine learning techniques that are supervised, unsupervised, and semi-supervised have worked very hard to harvest this data. The complicated and technological area of feature engineering falls within machine learning. Using deep learning, this tedious process may be completed automatically. Numerous studies have been conducted on deep learning models like LSTM, CNN, RNN, and GRU. Each model has employed a certain type of data, such as CNN for pictures and LSTM for language translation, etc. According to experimental results utilizing a publicly accessible dataset with reviews for all of the models, both positive and negative, and CNN, the best model for the dataset was identified in comparison to the other models, with an accuracy rate of 81%.
***Keyword:*** *Deep Learning, LSTM, CNN, RNN, GRU, Machine Learning, Supervised and Un-Supervised Learning*

## 1. Introduction

Customer feedback can significantly affect how people view your brand. If efforts aren't made to address negative reviews, they attract a lot of attention and give potential new customers cause for concern about the level of customer service you provide. However, a positive review can be highlighted, which will have the opposite impact of luring in new customers and boosting the confidence and contentment of your current clientele [1][2][3]. Sentiment analysis of these reviews and opinions is an extremely difficult and popular area of research, and many original subproblems have been addressed [4][5][6][7]. Today's consumers are particularly concerned about the quality of any goods; therefore, they check the manufacturer's website before making a purchase. Here, all customer reviews of the product are displayed for both customers and the business owner to see [8][9][10]. These reviews take a long time to read and are challenging. Researchers have therefore put a lot of effort into separating these reviews into positive and negative ones [11]. Many studies on sentiment analysis have been conducted utilizing supervised and unsupervised learning techniques [12][13][14][15]. In machine learning, manually designed features are frequently over-specified, insufficient, and expensive to design and validate [16]. Deep learning characteristics can be quickly and easily modified. A extremely adaptable, all-encompassing, and teachable framework for encoding linguistic, visual, and global information is provided by deep learning with the use of neural network [17]. LSTM (Long Short-Term Memory), CNN (Convolutional Neural Network), RNN (Recurrent Neural Networks), and GRU (Gated Recurrent Unit) with particular purpose are the four main models for deep learning.

Time-series data processing, forecasting, and categorization all require LSTM. In contrast to standard feed-forward neural networks, LSTM has feedback connections. It can handle both discrete data streams, like speech or video, as well as single data items, like photographs [18].
Recurrent neural networks identify patterns in data and utilize them to anticipate the following most likely scenario. Deep learning and the creation of models that mimic the neuronal activity of the human brain both require RNNs [19].

A CNN is a particular type of network design for deep learning algorithms that is utilized for tasks like image recognition and pixel data processing. Although there are different kinds of neural networks for deep learning, CNNs are the preferred network architecture for identifying and recognizing objects [20].

A Gated Recurrent Unit (GRU) is a component of a particular recurrent neural network architecture that aims to exploit connections through a series of nodes to carry out machine learning tasks related to memory and grouping, for example, in speech recognition [21].
This paper suggested novel architectures for LSTM, CNN, RNN, and GRU deep learning classifiers based on text dataset as product reviews to determine

Manuscript received November 5, 2023
Manuscript revised November 20, 2023
https://doi.org/10.22937/IJCSNS.2023.23.11.11

the polarity of reviews. The performance of all four models was compared over 50 epochs with 20 batches, and CNN performed the best, achieving an accuracy rate of 82%.

## 2. Literature Review

Filtering out objective reviews is not necessary for sentiment analysis, however doing so will also improve the precision of the analysis. There are numerous studies that examine sentence polarity in relation to figuring out the sentiment of a review or comment [22][23][24][25][26]. Sentiment orientation states that an opinion will be exactly favorable or negative depending on the situation [27]. A sentiment is a person's opinion, evaluation, or feeling about a good or service [26], feature [28][29], or both [30][31][32]. The majority of research on reviews or blogs relies on sentiment analysis with binary categorization, or positive or negative classes [33][34]. The majority of work on reviews or blogs relies on sentiment analysis with binary categorization, or positive or negative classifications [33][34]. As text categorization is carried out utilizing methodologies that are score-based, deep learning-based, and machine learning-based [35][36] [37][38][39]. Machine learning and deep learning techniques employ training data, whilst other techniques use different rules based on attributes and entities. In score-based systems, orientation of opinion as favorable or unfavorable has been taken into account [37]. Work of [40][41][42] employs a combination strategy using lexical resources and SentWordNet to calculate ratings for slangs. The polarity of opinion has also been identified using a lexicon of positive and negative words using supervised [43][44][45][46] and unsupervised [47][48] approaches with increased accuracy. Latent semantic indexing has been applied to improve supervised and unsupervised methodologies in order to increase machine intelligence [49][50]. Many studies have been conducted to extract aspects and conduct aspect-based sentiment analysis in order to determine the polarity of opinions based on those aspects [51][52][53]. In addition to machine learning, deep learning has also been used extensively for sentiment analysis across a variety of dimensions [54][55][56][57][58]. In the work of [59], word2vec was utilized to reduce the number of parameters by taking a large number of words into account. Authors [60] looked into how altering convolutional neural network hyperparameters affected performance throughout numerous runs. In [61], the k-max pooling-based OpCNN model was introduced while taking the Chinese word order issue into account. The LSTM neural network was used to implement sentiment classification on tweets, identifying whether they were favorable or negative [62][63].

## 3. Proposed Work

The entire proposed work is depicted in Figs. 1 and 2. The method for transforming the text dataset into model-readable form is depicted in Fig. 1. Each review is divided into pieces in Process 1 before stop words and punctuation are eliminated. Finding the vocabulary size is part of Process 2, which is required for the One-hot encoding that will be covered later. Vocabulary size is the total number of distinct terms in the dataset. Long sentences from the entire dataset are used in Process 3 to calculate the embedding vector. Table 2 displays the obtained vocabulary size as well as the length of a long phrase. On train data, four models—LSTM, CNN, RNN, and Gru—will be trained, and test data will be used to see how well they perform.

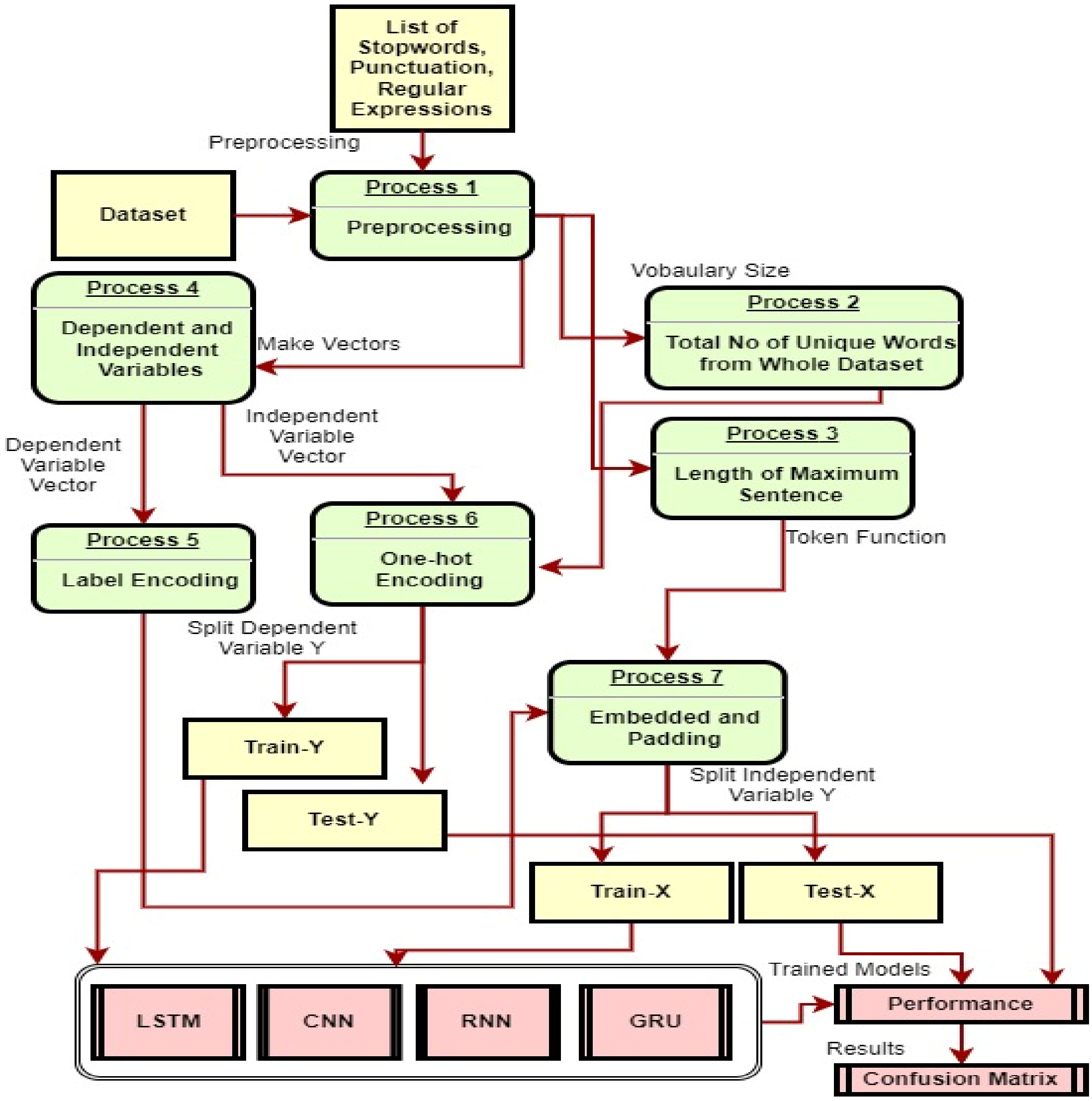
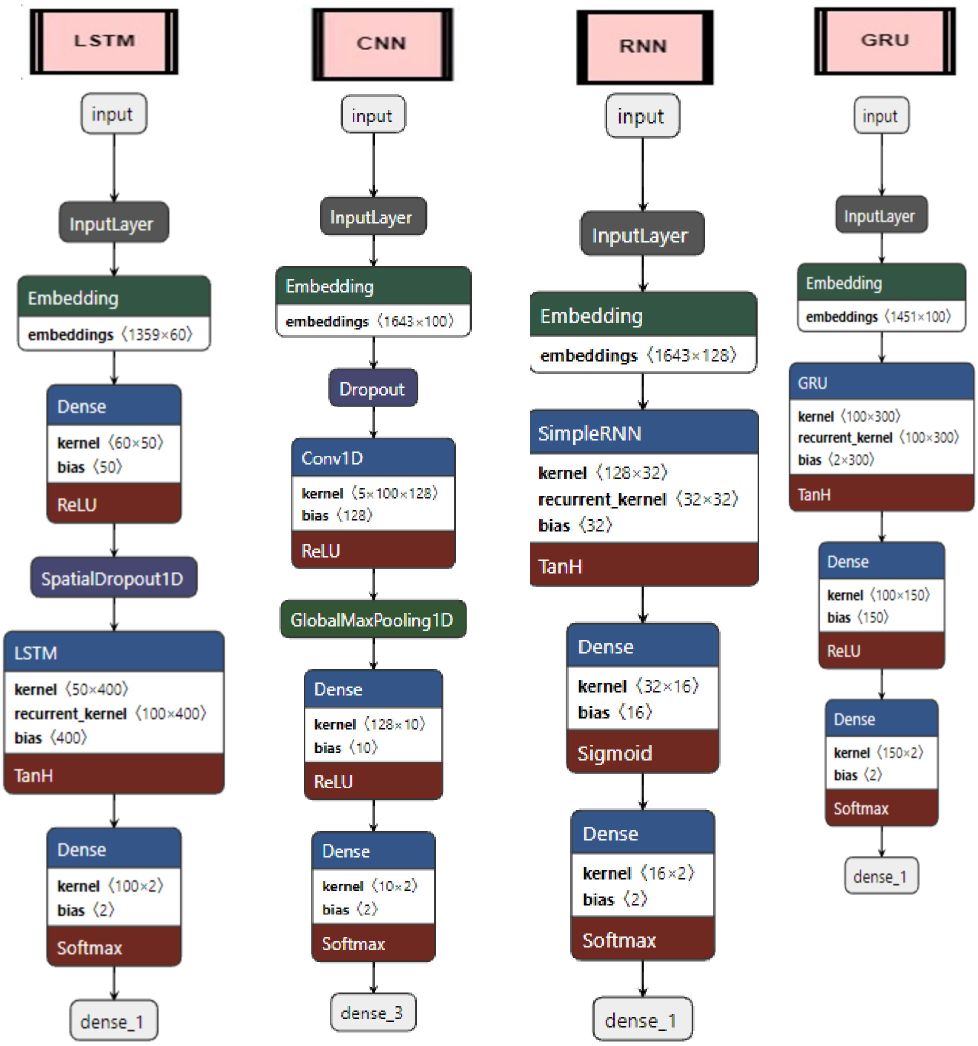



Fig-1: Proposed Work

LSTM as shown in Fig-2 (a), contains embedding layers with (1356×60) input shape. Then LSTM layer with 100 units and 0.2 dropout. In CNN, one convolutional layer with embedding layer is added as shown in Fig-2 (b) and in RNN, one SimpleRNN layer with embedding layer is added as shown in Fig-2 (c). GRU model has one GRU layer with embedding layer as shown in Fig-2 (d). Rest of work with respect to dense layers in all four models with activation functions are depicted in Fig-2.

LSTM has embedding layers with (1356×60) input shape, as shown in Fig. 2(a). LSTM layer with 100 units and 0.2 dropout follows. As indicated in Fig. 2(b), a convolutional layer with an embedding layer is added to the CNN, and Fig. 2(c), a SimpleRNN layer with an embedding layer is added to the RNN (c). As shown in Fig. 2 (d), the GRU model has a single GRU layer with an embedding layer. Complete Fig. 2 shows the remaining work with regard to dense layers in all four models with activation functions.




(a) (b) (c) (d)

Fig-2: Implemented Model LSTM, CNN, RNN and GRU

## 4. Used Dataset

This dataset [64] contains reviews in the form of positive and negative opinions in roughly 2000 sentences. It has two columns: one for text and the other for classes. Based on the text's review, it might be classified as negative or positive. Table 1 displays a sample of the dataset:

Table-1: 5-Sentences as Sample Dataset

| Class | Text |
|---|---|
| Positive | "their rotating beers on tap is also a highlight of this place" |
| Negative | "worst thai ever" |
| Positive | "if you stay in vegas you must get breakfast here at least once" |
| Positive | "I our server was great and we had perfect service" |
| Positive | "the pizza selections are good" |

## 5. Convert Dataset into Model Readable Form

The method for transforming the text dataset into model-readable form is depicted in Fig. 1. Each review will be divided into pieces in Process 1 before stop words and punctuation are eliminated. Finding the vocabulary size is part of Process 2, which is required for the One-hot encoding that will be covered later. Vocabulary size is the total number of distinct terms in the dataset. Long sentences from the entire dataset are used in Process 3 to calculate the embedding vector. Table 2 displays the obtained vocabulary size as well as the length of a long phrase.

Table-2: Parameters Required for Embedded Vector

| Name | Values |
|---|---|
| Vocabulary Size | 1643 |
| Size of Lengthy Sente | 25 |

Then, for vectors, Process-4 identifies dependent and independent variables. Using label encoding, which encodes negative values as 0 and positive values as 1, Process 5 will convert the dependent vector from text to integer values. Using the vocabulary size specified in Process-6 and Table-3, one-hot encoding will transform an independent vector into an integer.

Table-3: Sample of Numeric Vector

| | X1 | X2 | X3 | X4 | X5 | X6 | X7 | X8 | X9 | X10 | X11 | X12 | X13 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1594 | 25 | 1067 | | | | | | | | | | |
| 2 | 1340 | 25 | 1456 | 853 | 1607 | 947 | 754 | 940 | 1408 | | | | |
| 3 | 1441 | 1387 | 1607 | 1283 | 401 | 866 | 847 | 732 | 419 | 803 | 383 | 853 | 140 |
| 4 | 1607 | 496 | 1607 | 855 | 754 | 1428 | 853 | 8 | 1607 | 1342 | | | |
| 5 | 524 | 377 | 1016 | 853 | 112 | 282 | 262 | | | | | | |

In Process-7 padding (from Process-3) has been implemented by using the length of a lengthy phrase to construct an embedded vector comprising 25 columns as long as the long sentence, and a sample with a few columns is provided in Table-4. Padding converts all rows to be the same size.

Table-4: Embedded Vector Sample

| | X1 | X2 | X3 | X4 | X5 | X6 | X7 | X8 | X9 | X10 | X11 | X12 | X13 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1594 | 25 | 1067 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | 1340 | 25 | 1456 | 853 | 1607 | 947 | 754 | 940 | 1408 | 0 | 0 | | 0 |
| 3 | 1441 | 1387 | 1607 | 1283 | 401 | 866 | 847 | 732 | 419 | 803 | 383 | 853 | 140 |
| 4 | 1607 | 496 | 1607 | 855 | 754 | 1428 | 853 | 8 | 1607 | 1342 | 0 | 0 | 0 |
| 5 | 524 | 377 | 1016 | 853 | 112 | 282 | 262 | 0 | 0 | 0 | 0 | 0 | 0 |

Data from Processes 6 and 7 will be divided into four vectors: Train-X, Train-Y, Test-X, and Test-Y. Table 5 depicts the forms of these vectors.

| Vectors | Shape |
|---|---|
| Train-X | (705, 25) |
| Train-X | (705, 1) |
| Test-X | (235, 25) |
| Test-Y | (235, 1) |

Table-5: Shapes of Training and Testing Data

## 6. Models Performance

Confusion matrices are a widely used measurement when attempting to solve classification issues. Both binary classification and multiclass classification issues can be solved with it. The formula below is used to determine a model's correctness (via a confusion matrix). The correctly categorized TP values, FP values in the appropriate class when they should be in another class, FN values in another class when they should be in the appropriate class, and correctly classified TN values in the other class are all represented in the confusion matrix. According to these values, the performance metrics accuracy (ACC) from Eq-1, precision (P) from Eq-2, recall (R) from Eq-3, and F-score from Eq-4 are most typically used for classification [65].

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \quad \text{------------------ (1)}$$

$$P = \frac{TP}{TP + FP} \quad \text{-------------------(2)}$$

$$R = \frac{TP}{TP + FN} \quad \text{-------------------(3)}$$

$$F - Score = 2 * \frac{P * R}{P + R} \quad \text{-------------------(4)}$$

For the equations above, expected, and actual classes of data are needed. Predicted values are established for each model following training across all LSTM, CNN, RNN, and GRU models. A sample of such values is provided in Table 6.

Table-6: Predicted and Actual Values of Test Data

| Parameters | Predicted Values | Actual Values |
|---|---|---|
| LSTM | [1.95033645e-05, 9.99980450e-01]<br>[9.99982119e-01, 1.78446626e-05]<br>[1.97696518e-05, 9.99980211e-01]<br>[9.99979258e-01, 2.07688045e-05]<br>[2.05380838e-05, 9.99979496e-01] | [0., 1.]<br>[0., 1.]<br>[1., 0.]<br>[1., 0.]<br>[1., 0.] |
| CNN | [9.97400641e-01, 2.59935507e-03]<br>[2.88111892e-06, 9.99997139e-01]<br>[9.99999285e-01, 6.95826031e-07]<br>[9.99079943e-01, 9.20102349e-04]<br>[9.99951124e-01, 4.88810365e-05] | [0., 1.]<br>[0., 1.]<br>[1., 0.]<br>[1., 0.]<br>[1., 0.] |
| RNN | [0.9970348 , 0.00296524]<br>[0.01525643, 0.9847436 ]<br>[0.99838865, 0.00161139]<br>[0.4114155 , 0.5885844 ]<br>[0.9983367 , 0.00166332] | [0., 1.]<br>[0., 1.]<br>[1., 0.]<br>[1., 0.]<br>[1., 0.] |
| GRU | [9.99709666e-01, 2.90305848e-04]<br>[3.05698393e-03, 9.96943057e-01]<br>[9.99939203e-01, 6.08045812e-05]<br>[9.99845743e-01, 1.54176683e-04]<br>[9.99923229e-01, 7.68146638e-05] | [0., 1.]<br>[0., 1.]<br>[1., 0.]<br>[1., 0.]<br>[1., 0.] |

Equations for all models' accuracies are calculated using the aforementioned data and are displayed in Table 7 with CNN model receiving the highest score. Table-8 also displays precision, recall, and f-1score.

Table-7: Accuracies of All models Based on Training and Testing Data

| Parameters | LSTM | CNN | RNN | GRU |
|---|---|---|---|---|
| Training Data Accuracy | 100% | 100% | 100% | 99% |
| Testing Data Accuracy | 76% | 81% | 69% | 79% |

Table-8: Performance Parameters for all Models

| Models | Precision | Recall | F1-Score |
|---|---|---|---|
| LSTM | 0.76 | 0.76 | 0.76 |
| CNN | 0.82 | 0.82 | 0.82 |
| RNN | 0.69 | 0.69 | 0.69 |
| GRU | 0.79 | 0.79 | 0.79 |

## 7. Conclusion

Online surveys, particularly reviews from product- related websites, are a simple way to identify issues relating to a product's qualities. Modern deep learning algorithms may now quickly identify the review class from a text corpus. Four deep learning models have been built in this paper to determine if a given review is good or negative. Researchers from various fields have worked on sentiment analysis, which requires quick processing. The dataset used in this work includes product reviews that have been categorized as favorable and negative. Four deep learning models—LSTM, CNN, RNN, and GRU—have been deployed. To determine the best model for this dataset, the same embedded vector has been provided to all mentioned deep learning models. All models split the dataset into 25% training data and 75% testing data, using the same 50 epochs and 20 batch sizes. Fig. 3 displays the losses and accuracies of all models.

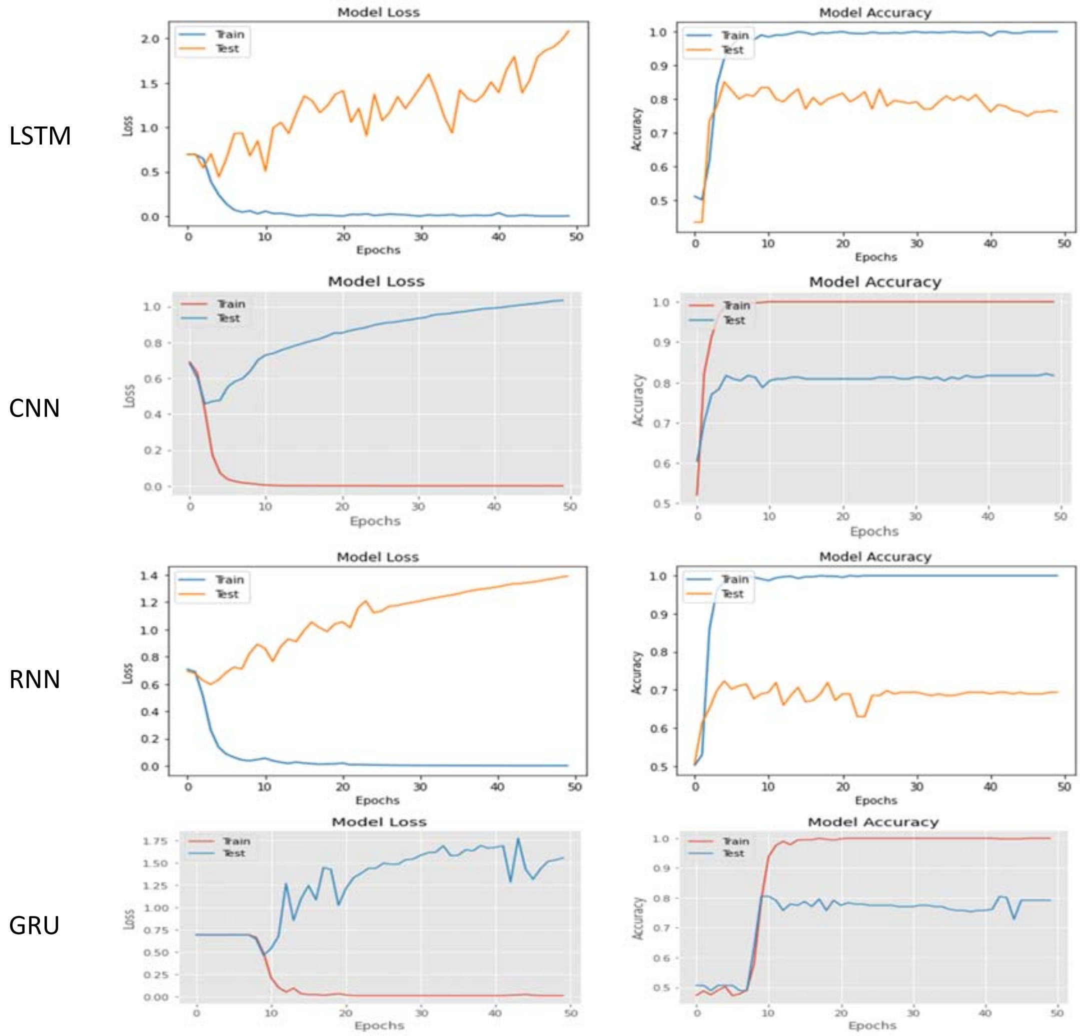



Fig-3: Losses and Accuracies of all Models

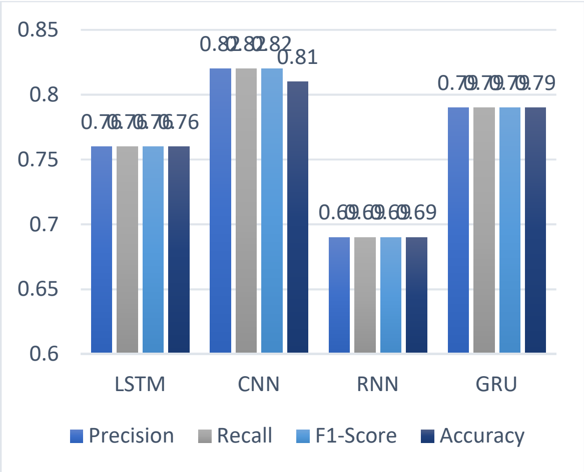



Fig-4: Performance Measures of LSTM, CNN, RNN and GRU

Following the training of all models on the sentiment dataset, it is found that CNN not only has the highest accuracy but also the highest precision, recall, and f1-score, measuring 81%, 82%, 82%, and 82% correspondingly. RNN has the lowest accuracy, precision, recall, and f1-score among the four models by 69%. As their accuracies and other metrics are 76% and 79%, respectively, LSTM and GRU are in a medium position between CNN and RNN. Performance of all models with respect to precision, recall, f1-score, and accuracy are shown in Fig-4.

## References


[1] J. Linkov, "Who Makes the Most Reliable New Cars?," *Consumer Reports*, 2022. https://www.consumerreports.org/.

[2] I. Roldós, "12 Product Review Examples to Get Eyes in 2022," *Monkey Learn*, 2022. https://monkeylearn.com/blog/product-reviews-examples/.

[3] M. Hameed, F. Tahir, and M. A. Shahzad, "Empirical comparison of sentiment analysis techniques for social media," *Int. J. Adv. Appl. Sci.*, vol. 5, no. 4, pp. 115–123, 2018.

[4] N. Archak, A. Ghose, and P. G. Ipeirotis, "Show me the Money ! Deriving the Pricing Power of Product," *Proc. 13th ACM SIGKDD Int. Conf. Knowl. Discov. data Min. - KDD '07*, pp. 56–65, 2007, doi: 10.1145/1281192.1281202.

[5] S. Muhammad and F. Masud, "MMO: Multiply-Minus-One Rule for Detecting & Ranking Positive and Negative Opinion," *Int. J. Adv. Comput. Sci. Appl.*, vol. 7, no. 5, pp. 122–127, 2016, doi: 10.14569/IJACSA.2016.070519.

[6] Y. Chen and J. Xie, "Online Consumer Review: Word-of-Mouth as a New Element of Marketing Communication Mix," *Manage. Sci.*, vol. 54, no. 3, pp. 477–491, 2008, doi: 10.1287/mnsc.1070.0810.

[7] I. El Alaoui, Y. Gahi, R. Messoussi, Y. Chaabi, A. Todoskoff, and A. Kobi, "A novel adaptable approach for sentiment analysis on big social data," *J. Big Data*, vol. 5, no. 1, 2018, doi: 10.1186/s40537-018-0120-0.

[8] D. Gräbner, M. Zanker, G. Fliedl, and M. Fuchs, "Classification of Customer Reviews based on Sentiment Analysis," *Inf. Commun. Technol. Tour. 2012*, pp. 460–470, 2012, doi: 10.1007/978-3-7091-1142-0_40.

[9] S. M. Saqib, F. M. Kundi, and S. Ahmad, "Unsupervised Learning Method for Sorting Positive and Negative Reviews Using LSI (Latent Semantic Indexing) with Automatic Generated Queries," *IJCSNS Int. J. Comput. Sci. Netw. Secur.*, vol. 18, no. 1, pp. 56–62, 2018.

[10] Q. Sun, J. Niu, Z. Yao, and H. Yan, "Exploring eWOM in online customer reviews: Sentiment analysis at a fine-grained level," *Eng. Appl. Artif. Intell.*, vol. 81, pp. 68–78, 2019, doi: 10.1016/j.engappai.2019.02.004.

[11] T. N. Sheikh Muhammad Saqib, Fazal Masud Kundi, Shakeel Ahmad, "Automatic classification of product reviews into interrogative and noninterrogative: Generating real time answer," *Int. J. Adv. Appl. Sci.*, vol. 6, no. 8, pp. 23–31, 2019, doi: 10.21833/ijaas.2019.08.004.

[12] M. Ahmad, S. Aftab, M. S. Bashir, N. Hameed, I. Ali, and Z. Nawaz, "SVM optimization for sentiment analysis," *Int. J. Adv. Comput. Sci. Appl.*, vol. 9, no. 4, pp. 393–398, 2018, doi: 10.14569/IJACSA.2018.090455.

[13] M. Abbas, K. Ali Memon, and A. Aleem Jamali, "Multinomial Naive Bayes Classification Model for Sentiment Analysis," *IJCSNS Int. J. Comput. Sci. Netw. Secur.*, vol. 19, no. 3, p. 62, 2019, [Online]. Available: http://paper.ijcsns.org/07_book/201903/20190310.pdf.

[14] J. Khairnar and M. Kinikar, "Sentiment Analysis Based Mining and Summarizing Using SVM-MapReduce," *Int. J. Comput. Sci. Inf. Technol.*, vol. 5, no. 3, pp. 4081–4085, 2014.

[15] B. Liu, "Sentiment Analysis and Opinion Mining," *Synth. Lect. Hum. Lang. Technol.*, vol. 5, no. 1, pp. 1–167, 2012, doi: 10.2200/S00416ED1V01Y201204HLT016.

[16] V. S. D. F. X. Christopher, "Evolving Trends in Conversational Systems with Natural Language Processing," *Int. J. Comput. Intell. Informatics*, vol. 8, no. 3, pp. 123–129, 2018.

[17] A. See, "Natural Language Processing with Deep Learning: Natural Language Generation," pp. 1–39, 2019.

[18] Debasish Kalita, "An Overview on Long Short Term Memory (LSTM)," *March 11, 2022*, 2022, [Online].

Available: https://www.analyticsvidhya.com/blog/2022/03/an-overview-on-long-short-term-memory-lstm/.

[19] A. Tsantekidis, N. Passalis, and A. Tefas, "Recurrent neural networks," 2022. doi: 10.1016/B978-0-32-385787-1.00010-5.

[20] Rahul Awat, "Definition: convolutional neural network," 2022. [Online]. Available: https://www.techtarget.com/searchenterpriseai/definition/convolutional-neural-network.

[21] S. Kostadinov, "Understanding GRU Networks," 2017. [Online]. Available: https://towardsdatascience.com/understanding-gru-networks-2ef37df6c9be.

[22] T. Chen, R. Xu, Y. He, and X. Wang, "Improving sentiment analysis via sentence type classification using BiLSTM-CRF and CNN," *Expert Syst. Appl.*, vol. 72, pp. 221–230, 2017, doi: 10.1016/j.eswa.2016.10.065.

[23] F. Wu, Y. Huang, and Z. Yuan, "Domain-specific sentiment classification via fusing sentiment knowledge from multiple sources," *Inf. Fusion*, vol. 35, pp. 26–37, 2017, doi: 10.1016/j.inffus.2016.09.001.

[24] T. Naeem, F. M. Kundi, and S. M. Saqib, "Extracting Words' Polarity With Definition and Examples," *Int. J. Comput. Sci. Inf. Secur.*, vol. 15, no. 2, pp. 180–190, 2017, doi: 10.1036/0071395415.

[25] wang maoquan, C. Shiyun, X. yufei, and Z. lu, "EICA at SemEval-2017 Task 4: A Simple Convolutional Neural Network for Topic-based Sentiment Classification," *Proc. 11th Int. Work. Semant. Eval.*, pp. 737–740, 2017, [Online]. Available: http://www.aclweb.org/anthology/S17-2124.

[26] D. Bollegala, T. Mu, and J. Y. Goulermas, "Cross-Domain Sentiment Classification Using Sentiment Sensitive Embeddings," in *IEEE Transactions on Knowledge and Data Engineering*, 2016, vol. 28, no. 2, pp. 398–410, doi: 10.1109/TKDE.2015.2475761.

[27] B. Liu, *Sentiment Analysis and Opinion Mining*, vol. 5, no. 1. 2012.

[28] J. Jin, P. Ji, and R. Gu, "Identifying comparative customer requirements from product online reviews for competitor analysis," *Eng. Appl. Artif. Intell.*, vol. 49, pp. 61–73, 2016, doi: 10.1016/j.engappai.2015.12.005.

[29] L. Shu, H. Xu, and B. Liu, "Lifelong Learning CRF for Supervised Aspect Extraction," in *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics*, 2017, pp. 148–154, doi: 10.18653/v1/P17-2023.

[30] F. e-M. K. Khan, B.B. Baharudin, A. Khan, "Mining opinion from text documents," *Adv. Res. Comput. Commun. Eng.*, vol. 3, no. 7, pp. 217–222.

[31] B. Pang and L. Lee, "Opinion mining and sentiment analysis," *Found. Trends Inf. Retr.*, vol. 2, no. 1–2, pp. 1–135, 2008, doi: 10.1561/1500000011.

[32] M. Z. Asghar, A. Khan, S. Ahmad, and F. M. Kundi, "A Review of Feature Extraction in Sentiment Analysis," *J. Basic. Appl. Sci. Res*, vol. 4, no. 3, pp. 181–186, 2014.

[33] S. Wang, D. Li, Y. Wei, and H. Li, "A Feature Selection Method Based on Fisher 's Discriminant Ratio for Text Sentiment Classification Suge," *Springer-Verlag Berlin Heidelb.*, pp. 88–97, 2009.

[34] A. Esmin and S. Matwin, "Hierarchical Classification Approach to Emotion Recognition in Twitter Hierarchical Classification Approach to Emotion Recognition in Twitter," 2016, no. March, doi: 10.1109/ICMLA.2012.195.

[35] W. Y. and L. H. A. Wang S, Li D, Song X, "A feature selection method based on improved fisher's discriminant ratio for text sentiment classification," *Expert Syst. Appl.*, vol. 38, no. 7, pp. 8696-8702., 2011.

[36] L. C. H. and C. H. Chen LS, "A neural network based approach for sentiment classification in the blogosphere.," *J. Informetr.*, vol. 5, no. 2, pp. 313–322, 2011.

[37] F. M. Kundi, A. Khan, S. Ahmad, and M. Z. Asghar, "Lexicon-Based Sentiment Analysis in the Social Web," *J. Basic. Appl. Sci. Res*, vol. 4, no. 6, pp. 238–248, 2014.

[38] J. Fiaidhi, O. Mohammed, S. Mohammed, S. Fong, and T. Kim, "Opinion Mining over Twitterspace : Classifying Tweets Programmatically using the R Approach," *Digit. Inf. Manag. (ICDIM), Seventh Int. Conf. on. IEEE*, pp. 313–319, 2012.

[39] A. Jeyapriya and C. S. K. Selvi, "Extracting aspects and mining opinions in product reviews using supervised learning algorithm," in *2nd International Conference on Electronics and Communication Systems, ICECS 2015*, 2015, pp. 548–552, doi: 10.1109/ECS.2015.7124967.

[40] M. Z. A. Fazal Masud Kundi, Shakeel Ahmad, Aurangzeb Khan, "Detection and Scoring of Internet Slangs for Sentiment Analysis Using SentiWordNet Fazal," *Life Sci. J.*, vol. 11, no. 1, pp. 66–72, 2014.

[41] D. K. Kirange, R. R. Deshmukh, and M. D. K. Kirange, "Aspect Based Sentiment analysis SemEval-2014 Task 4," *Asian J. Comput. Sci. {&} Inf. Technol.*, vol. 4, no. 8, pp. 72–75, Aug. 2014, doi: 10.15520/ajcsit.v4i8.9.

[42] Deepak Kumar Gupta and Asif Ekbal, "Supervised Machine Learning for Aspect based Sentiment Analysis," in *Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014)*, 2014, pp. 319–323.

[43] T. Shaikh and D. Deshpande, "A Review on Opinion Mining and Sentiment Analysis," in *IJCA Proceedings on National Conference on Recent Trends in Computer Science and Information Technology*, 2016, no. 2, pp. 6–9.

[44] A. Alghunaim, M. Mohtarami, S. Cyphers, and J. Glass, "A Vector Space Approach for Aspect Based Sentiment Analysis," *Proc. NAACL-HLT 2015*, pp. 116–122, 2015.

[45] M. Cuadros, S. Sebastian, G. Rigau, E. H. Unibertsitatea, and S. Sebastian, "V3: Unsupervised Aspect Based Sentiment Analysis for SemEval-2015 Task 12," no. SemEval, pp. 714–718, 2015.

[46] S. Rosenthal, N. Farra, and P. Nakov, "SemEval-2017 Task 4 : Sentiment Analysis in Twitter," *Proc. 11th Int. Work. Semant. Eval.*, vol. 3, no. 4, pp. 502–518, 2017.

[47] S. M. Street and B. Liu, "Identifying Noun Product Features that Imply Opinions," *Proc. 49th Annu. Meet. Assoc. Comput. Linguist. Hum. Lang. Technol.*, no. Issue: 2009, pp. 575–580, 2011, [Online]. Available:

http://www.aclweb.org/anthology-new/P/P11/P11-2101.pdf.

[48] S. S. Htay and K. T. Lynn, "Extracting product features and opinion words using pattern knowledge in customer reviews," *Sci. World J.*, 2013, doi: 10.1155/2013/394758.

[49] H. J. Alyamani, S. Ahmad, A. H. Syed, S. M. Saqib, and Y. D. Al-Otaibi, "LSI Authentication-Based Arabic to English Text Converter," *Int. J. Fuzzy Log. Intell. Syst.*, vol. 21, no. 4, pp. 409–422, 2021, doi: 10.5391/IJFIS.2021.21.4.409.

[50] A. O. Shakeel Ahmad, Sheikh Muhammad Saqib, "LSI Based Search Technique: Using Extracted Keywords and Key Sentences," *VAWKUM Trans. Comput. Sci.*, vol. 14, no. 2, pp. 1–8, 2017.

[51] S. M. Saqib and F. M. Kundi, "Semi Supervised Method for Detection of Ambiguous Word and Creation of Sense : Using WordNet," vol. 9, no. 11, pp. 353–359, 2018.

[52] L. Zhuang, F. Jing, and X.-Y. Zhu, "Movie review mining and summarization," in *Proceedings of the 15th ACM international conference on Information and knowledge management - CIKM '06*, 2006, pp. 43–50, doi: 10.1145/1183614.1183625.

[53] S. Gojali and M. L. Khodra, "Aspect based sentiment analysis for review rating prediction," 2016, doi: 10.1109/ICAICTA.2016.7803110.

[54] N. C. Dang, M. N. Moreno-García, and F. De la Prieta, "Sentiment analysis based on deep learning: A comparative study," *Electron.*, vol. 9, no. 3, 2020, doi: 10.3390/electronics9030483.

[55] P. Cen, K. Zhang, and D. Zheng, "Sentiment Analysis Using Deep Learning Approach," *J. Artif. Intell.*, vol. 2, no. 1, pp. 17–27, 2020, doi: 10.32604/jai.2020.010132.

[56] S. M. Saqib, S. Ahmad, A. H. Syed, T. Naeem, and F. M. Alotaibi, "Grouping of Aspects into Relevant Category Based on WordNet Definitions," vol. 19, no. 2, pp. 113–119, 2019.

[57] K. Dashtipour, M. Gogate, A. Adeel, H. Larijani, and A. Hussain, "Sentiment analysis of persian movie reviews using deep learning," *Entropy*, vol. 23, no. 5, pp. 1–16, 2021, doi: 10.3390/e23050596.

[58] L. Zhang, S. Wang, and B. Liu, "Deep learning for sentiment analysis: A survey," *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.*, vol. 8, no. 4, 2018, doi: 10.1002/widm.1253.

[59] R. Johnson and T. Zhang, "Effective use of word order for text categorization with convolutional neural networks," *NAACL HLT 2015 - 2015 Conf. North Am. Chapter Assoc. Comput. Linguist. Hum. Lang. Technol. Proc. Conf.*, no. 2011, pp. 103–112, 2015, doi: 10.3115/v1/n15-1011.

[60] Y. Zhang and B. Wallace, "A Sensitivity Analysis of (and Practitioners' Guide to) Convolutional Neural Networks for Sentence Classification," in *Pro_ceedings of the 8th International Joint Conference on Natural Language Processing*, 2017, p. 253263, [Online]. Available: http://arxiv.org/abs/1510.03820.

[61] S. Zhao, Z. Xu, L. Liu, M. Guo, and J. Yun, "Towards Accurate Deceptive Opinions Detection based on Word Order-preserving CNN," *Available Online: http://arxiv.org/abs/1711.09181*, pp. 1–8, 2018.

[62] X. Wang, Y. Liu, C. Sun, B. Wang, and X. Wang, "Predicting polarities of tweets by composing word embeddings with long short-Term memory," *ACL-IJCNLP 2015 - 53rd Annu. Meet. Assoc. Comput. Linguist. 7th Int. Jt. Conf. Nat. Lang. Process. Asian Fed. Nat. Lang. Process. Proc. Conf.*, vol. 1, pp. 1343–1353, 2015, doi: 10.3115/v1/p15-1130.

[63] D. Tang, B. Qin, and T. Liu, "Document modeling with gated recurrent neural network for sentiment classification," *Conf. Proc. - EMNLP 2015 Conf. Empir. Methods Nat. Lang. Process.*, no. September, pp. 1422–1432, 2015, doi: 10.18653/v1/d15-1167.

[64] "https://www.kaggle.com/yelp-dataset/yelp-dataset." .

[65] A. Kulkarni, "Confusion Matrix Confusion matrix is a very popular measure used while solving classification problems.," *Data Democracy, Elviewer*, 2020. https://www.sciencedirect.com/topics/engineering/confusion-matrix#:~:text=A confusion matrix is a,malignant tissue is considered cancerous.