Rohan Karna (M25MAC005)
Akshay Gupta (M25MAC017)

Theoretical Analysis of Deep Learning Architectures for Natural Language Processing

1. Introduction

The natural language processing is this beneficial connection between human linguistic speech and the ability to understand it through computers, which creates the possibility to increase the ability of a machine to handle and react to written content intelligently. One of the most eminent uses in this field is sentiment analysis where the sentiment amounting or approach of evaluative lexical material is determined and then classified as a positive, negative, or neutral sentiment. With recent advances in deep learning approaches, the accuracy of this task can be significantly increased, with neural designs capable of automatically identifying significant patterns and contextual features on the corpora, without the need to design features manually.

The current research compares deep learning architecture with traditional machine learning architectures, such as Recurrent Neural Networks (RNNs), Long Short-Term Memory networks (LSTMs), Gated Recurrent Units (GRUs), Deep Neural Networks (DNNs), Text Convolutional Neural Networks (TextCNN), Naive Bayes classifiers, and random forests, to estimate their effectiveness in emotive text representation and identify which architectural structure can deliver the most effective results.

1.1 Application of Natural Language Processing

The field of Natural Language Processing (NLP) has become an area of the work of artificial intelligence, promoting the ability of computational systems to interpret and produce human language. Its uses are diverse, including chatbot, email filtering tools, language translators, intelligent personal assistants, document analysis and predictive typographic tools and monitoring social media content.
Among them one of the applications with which this project will focus on is sentiment analysis.

1.2 Sentiment Analysis

Sentiment analysis is an approach where text information is tested to identify and categorize the mood that is expressed by the writer or speaker. This action makes it possible to get the understanding of opinions, emotions, and attitudes expressed through a message, which makes it useful in application in sales (e.g., product reviews), customer support, and social media monitoring.
In this project, sentiment analysis is applied to classify text data into three sentiment categories:

1. **Positive:** expressing satisfaction, approval, or favorable opinion,
2. **Negative:** expressing dissatisfaction, disapproval, or criticism, and
3. **Neutral:** expressing an unbiased or factual tone without strong emotion.

The exploration of how different learning frameworks are understood as the sources of these sentiments based on natural text and how deep learning frameworks outperform the earlier schemes by independently acquiring contextual and sequential language patterns.

1.3 Limitation of Sentiment Analysis

Although sentiment analysis is an effective tool to explain the opinion of the population and emotional positions, it still has considerable limitations. It is often not able to recognize sarcasm or hidden connotations in written text and is ambiguous when the lexical elements take on more than one meaning based on contextual effects. Moreover, one sentence can have an ambivalent affect, thus making the task of assigning a definite sensation expression rather difficult. The level of the training data must determine the empirical results, and currently models that generally distinguish single, primitive categories of effect of positive, negative, or neutral; thus, ignoring more subtle emotions and even tonal nuances discerned instinctively by human beings

2. Methodology

2.1 Flowchart of workflow:

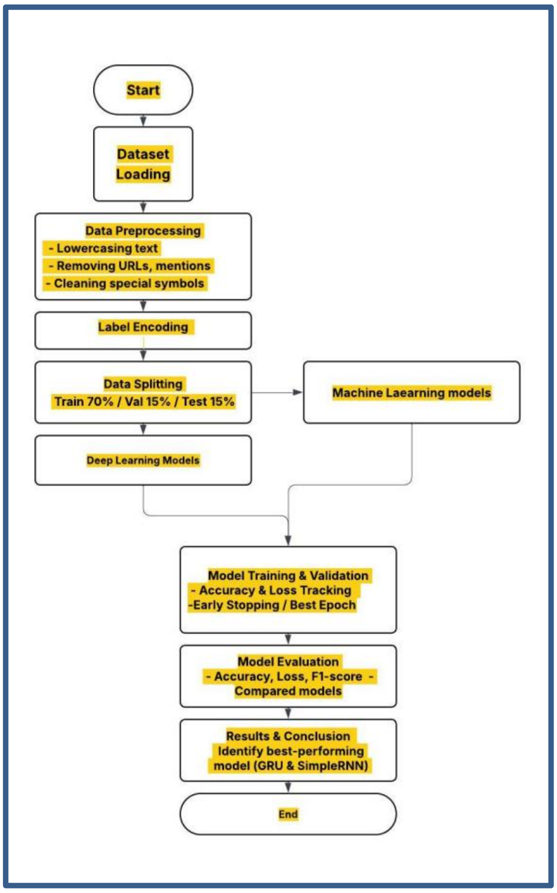

Fig-Diagram of workflow

2.2 Architectures Used:

Deep Learning Architectures	Machine Learning Architectures
Recurrent Neural Network (RNN)	Naïve Bayes
Long Short Term Memory (LSTM)	Random Forest
Gated Recurrent Unit (GRU)	
Text- Convolutional Neural Network (TCNN)	
Deep Neural Network (DNN)	

After applying deep learning and machine learning methods, this project uses an approach to perform a comparative study of sentiment analysis. Text detection Recurrent neural networks (RNN), long short-term memory networks (LSTM), gated recurrent units (GRU), Convolutional Neural Networks on text (TextCNN) and densely connected neural networks (DNN) are evaluated in comparison to more traditional algorithms, including Naïve Bayes and Random Forest, to find out their ability to detect sentiment in textual data.

3. Data Preprocessing and Model Evaluation

3.1 Dataset Description:

Twitter US Airline Sentiment Dataset that is used in the current research was sourced on Kaggle and was initially collected by Crowdflower. The corpus consists of tweets to the big airlines in the U.S. and they have been manually marked as positive, negative, and neutral. It has an average of 14,640 observations and 15 attributes, including the text of the tweet and supplementary metadata. In preprocessing, the attributes that were considered to be relevant were only taken and those columns that had high redundancy and a high number of missed values were removed. The textual data was subjected to cleaning protocols of converting all text to lower case, removal of hyperlinks, user references, and superfluous characters, creating a new variable named text clean. After that, sentiment annotations were employed numerically as 0 0= negative, 1 = neutral, and 2 = positive. Eventually, the two variables such as the text clean and the text-encoded label were used as input in order to train and test both deep-learning and traditional machine learning sentiment analyzers

.3.2 Data Splitting:

The cleaned dataset is divided into three subsets:

1. 70% training data,
2. 15% validation data, and
3. 15% test data.

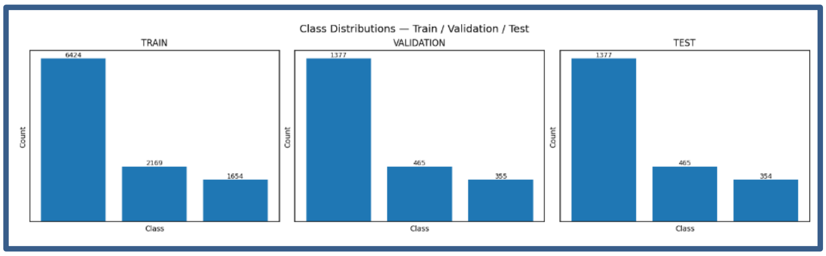

Fig-Class Distribution Across splits

3.3 Model Evaluation:

The procedure focuses on measuring the ability of each model to analyze twitters and categorize them into Positive, neutral and negative sentiments.

3.3.1 Accuracy and Loss Plot

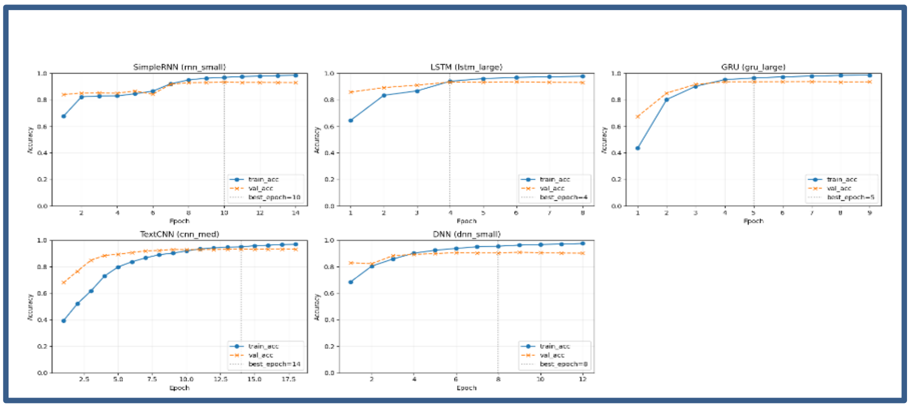

Fig-Accuracy Vs Epoch Plot of Deep Learning Architectures

Rohan Karna (M25MAC005)
Akshay Gupta (M25MAC017)

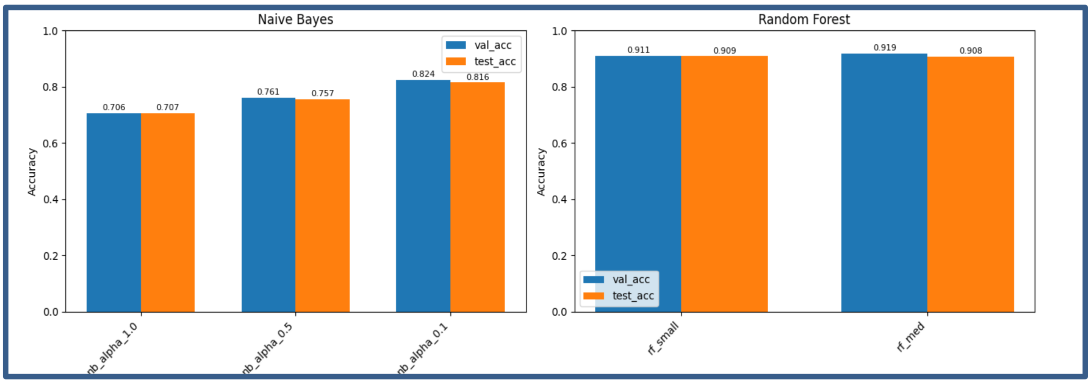

Fig-Accuracy Comparison of ML architectures for different value of parameters

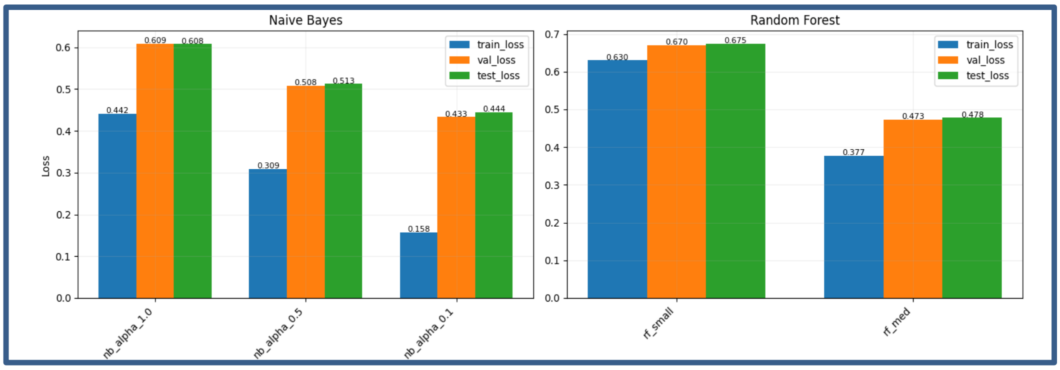

Fig-Loss Comparison of ML architectures for different value of parameters

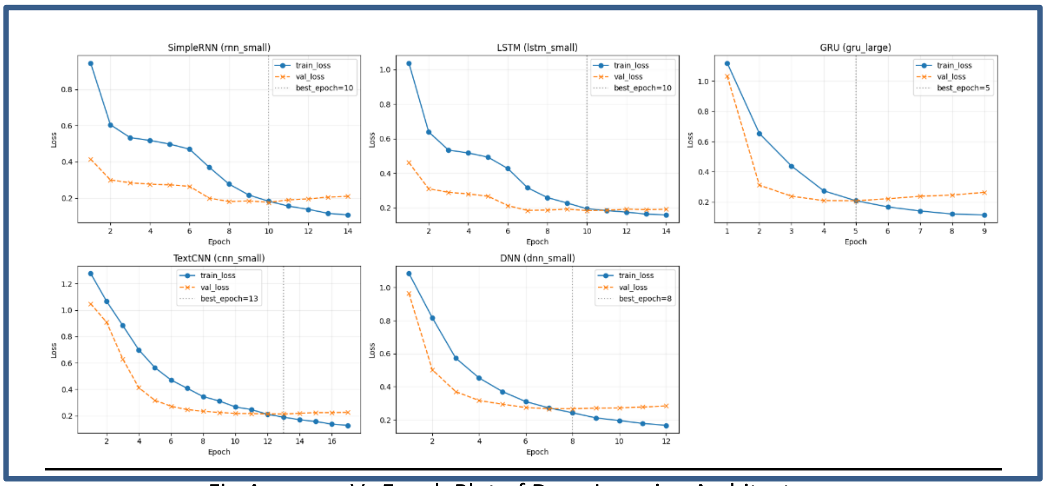

Fig-Accuracy Vs Epoch Plot of Deep Learning Architectures

Training and validation plot analysis show that deep learning models such as RNN, LSTM,GRU,TextCNN and DNN exhibited steady convergence in terms of low validation losses and high accuracy thus portraying strong learning abilities. GRU and TextCNN achieved the best performance faster and had little overfitting in this group. Nonely, machine learning models (Naive Bayes and Random Forest) were the most consistent but mediocre models, with the first one slightly beating the second in validation and test performances. Overall, deep learning-based architectures were more effective at contextual relationship modeling in textual data and the classical models provided a reliable baseline of operations at lower cost to the computation.

4. Result&Conclustion

4.1 Comparative Performance Table:

family	model_name	train_acc	val_acc	test_acc	train_loss	val_loss	test_loss	best_epoch
GRU	gru_large	0.977	0.936	0.938	0.1	0.208	0.209	5
SimpleRNN	rnn_small	0.98	0.934	0.94	0.079	0.179	0.17	10
LSTM	lstm_large	0.964	0.934	0.933	0.151	0.198	0.199	4
TextCNN	cnn_large	0.979	0.932	0.927	0.169	0.272	0.284	8
RandomForest	rf_med	0.981	0.919	0.908	0.377	0.473	0.478	NaN
DNN	dnn_med	0.968	0.904	0.908	0.156	0.279	0.3	10
NaiveBayes	nb_alpha_0.1	0.953	0.824	0.816	0.158	0.433	0.444	NaN

The table gives a brief analysis of the quantitative performance measure of all of the rated models. Out of the studied architectures, the Simple RNN has the best test accuracy of (94.03) closely matched by GRU (93.80) and LSTM (93.30). Other models, including Naive Bayes and the Random Forest, had lower accuracies, which can be explained by the fact that they are not as effective as capturing sequential dependencies created by text data.

4.2 Conclusion:

Implementation of various deep learning and classical machine learning models on the sentiment classification in the twitter US Airline Sentiment Dataset was found to be successful. After intensive testing, the gated recurrent unit (GRU) and the simple recurrent neural network (SimpleRNN) models turned out to be the most useful with higher predictive power and more reliability to capture contextual effects of textual data. Traditional classifiers like Naive Bayes and Random Forest showed a reasonable level of performance; however, they lacked ability to extract sequential dynamics which are required to be present in the case of language. Generally, the results highlight the benefits of deep learning paradigms, in particular, recurrent networks, strike an appropriate trade-off between predictive accuracy and computational efficiency, and, thus, make them specifically appropriate to be deployed in real-world sentiment analysis and a wide range of natural language processing tasks.

4.3 Reference:

[1] A. Pisote, S. Mangate, Y. Tarde, H. A. Inamdar, S. Ashok Nangare and V. Borate, "A Comparative Study of ML and NLP Models with Sentimental Analysis," 2025 International Conference on Advancements in Power, Communication and Intelligent Systems (APCI), Kannur, India, 2025, pp. 1-5, doi: 10.1109/APCI65531.2025.11136837. keywords: {Sentiment analysis;Analytical models;Technological innovation;Accuracy;Video on demand;Social networking (online);Natural language processing;Multilingual;Web sites;Context modeling;Multilingual Sentiment Analysis;Marathi;Natural Language Processing;Text Summarization;Lexicon-based approaches},

[2] M. S. Islam and K. Masudul Alam, "Sentiment Analysis on Bangla Food Reviews Using Machine Learning and Explainable NLP," 2023 26th International Conference on Computer and Information Technology (ICCIT), Cox's Bazar, Bangladesh, 2023, pp. 1-6, doi: 10.1109/ICCIT60459.2023.10441309. keywords: {Radio frequency;Sentiment analysis;Machine learning algorithms;Reviews;Social networking (online);Logic gates;Random forests;Sentiment Analysis;Food Review Analysis;CNN;Bi-GRU;Explainable NLP},

Project Contribution

Name	Contribution
Rohan Karna	Deep Learning Architectures
Akshay Gupta	Machine Learning Architectures