Comparative Analysis of Machine Learning and Deep Learning Algorithms for Sentiment Polarity Detection.

-Project By
Raveesh R- 21011101099- 21110282- 3rd year AI&DS-B
Samyuktha S- 21011101106 – 21110237- 3rd year AI&DS-B

Abstract:

This research paper presents a comparative analysis of machine learning and deep learning algorithms for sentiment polarity detection. Using a subset of the Amazon reviews dataset, we preprocess the textual data and explore traditional machine learning methods such as SVM, Naïve Bayes and Random Forest Classification. Additionally, we propose hybrid models combining TF-IDF representations with CNN, BiLSTM, and LSTM networks. The best-performing models are identified based on accuracy, with CNN achieving 85% and the hybrid CNN-LSTM model obtaining 82.23%. Future work is proposed to explore advanced deep learning techniques like GRU, Transformers and BERT Model to achieve accuracy levels beyond 90%.

Introduction:

Sentiment analysis, also known as opinion mining, plays a vital role in understanding public perceptions and customer feedback. This research paper focuses on sentiment polarity detection and compares traditional machine learning algorithms with state-of-the-art deep learning techniques. The Amazon reviews dataset serves as the basis for the analysis. After preprocessing the text data, various machine learning methods and hybrid models are employed for sentiment classification. The superiority of CNN(85% accuracy) and the potential of the hybrid CNN-LSTM model(82.23%) are highlighted. Future work suggests exploring advanced deep learning techniques to further enhance accuracy.

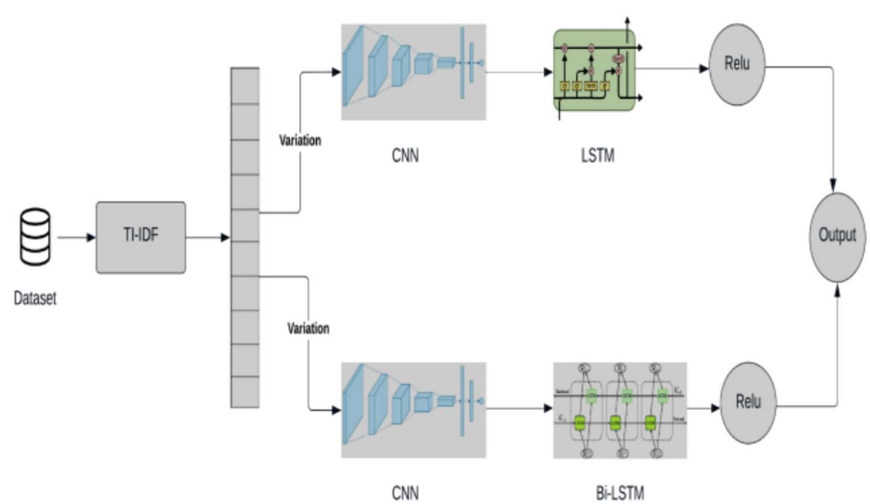
Dataset:

For this research paper the experimental dataset is the amazon reviews dataset, which includes the reviews from Amazon selected from Hugging Face.
The dataset contains 36 million reviews out of which 18 million are positive comments and the remaining are negative comments. We will be using a subset of the dataset for our research purposes to reduce the computational complexity.
Link: https://huggingface.co/datasets/amazon_polarity

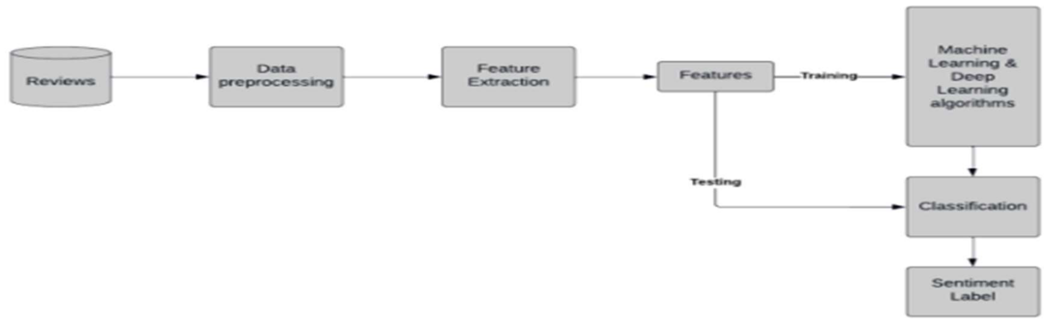
Dataset Sample:

label (class label)	title (string)	content (string)
1 (positive)	"Stuning even for the non-gamer"	"This sound track was beautiful! It paints the senery in your mind so well I would recomend it even to people who hate vid. game music! I have played the game Chrono…
1 (positive)	"The best soundtrack ever to anything."	"I'm reading a lot of reviews saying that this is the best 'game soundtrack' and I figured that I'd write a review to disagree a bit. This in my opinino is Yasunori…
1 (positive)	"Amazing!"	"This soundtrack is my favorite music of all time, hands down. The intense sadness of "Prisoners of Fate" (which means all the more if you've played the game) and the…
1 (positive)	"Excellent Soundtrack"	"I truly like this soundtrack and I enjoy video game music. I have played this game and most of the music on here I enjoy and it's truly relaxing and peaceful.On disk…
1 (positive)	"Remember, Pull Your Jaw Off The Floor After Hearing it"	"If you've played the game, you know how divine the music is! Every single song tells a story of the game, it's that good! The greatest songs are without a doubt,…

Hybrid deep learning model architecture:



Proposed workflow:



Workflow explanation:

1) Data Preprocessing:

The text data undergoes cleaning and preprocessing to eliminate irrelevant information, punctuation, and special characters. It is then tokenized into words or sub word units and converted to lowercase to ensure consistency.

2) Machine Learning Methods:

a. SVM (Support Vector Machine): The SVM algorithm is employed for sentiment classification, effectively determining whether sentiments are positive or negative.
b. Naive Bayes Classifier: This method predicts the probability of data points belonging to specific sentiment categories, aiding in sentiment classification.
c. Random Forest: Utilizing a combination of decision trees, the Random Forest algorithm is used to evaluate the likelihoods of different sentiment classes.

3) TF-IDF Vectorization:

It is a technique used to convert text data into numerical vectors, representing the importance of words in a document relative to the entire corpus. It stands for "Term Frequency-Inverse Document Frequency," and it helps capture the significance of words in individual documents while considering their rarity across the entire dataset.

4) Hybrid Methods:

a. TF-IDF followed by CNN with Rectified Linear Unit (ReLU) activation.
b. TF-IDF followed by BiLSTM with Rectified Linear Unit (ReLU) activation.
c. TF-IDF followed by LSTM with Rectified Linear Unit (ReLU) activation.
d. TF-IDF followed by CNN, then LSTM with Rectified Linear Unit (ReLU) activation.

e. TF-IDF followed by CNN, then BiLSTM with Rectified Linear Unit (ReLU) activation.

5) Deep Learning Models:

a. CNN (Convolutional Neural Network): Employ a single convolutional layer for sentiment analysis, primarily utilized in image recognition tasks.
b. BiLSTM (Bidirectional Long Short-Term Memory): Process input sequences in both forward and backward directions to capture contextual information for sentiment analysis in natural language processing tasks.
c. ReLU (Rectified Linear Unit): An activation function used to facilitate the network in learning non-linear relationships within the data.

6) Evaluation:

The performance of each method and model will be assessed using appropriate metrics, including accuracy, precision, recall, and F1 score. These metrics will help gauge the effectiveness and efficiency of the implemented approaches in sentiment analysis.

7) Model Selection:

Choose the best-performing method or model based on evaluation results.

8) Deployment:

Once the chosen sentiment analysis model is finalized, it will be deployed to analyse new and unseen customer text data. This deployment allows the model to process real-time inputs and provide accurate sentiment analysis for previously unseen customer reviews or feedback, contributing to better decision-making and insights for businesses.

9) Continuous Improvement:

A crucial aspect of the process involves closely monitoring the performance of the sentiment analysis model and actively gathering feedback. This feedback-driven approach enables the identification of areas that require improvement.

Result analysis:

The study compared machine learning and deep learning models with SVM and CNN achieving the highest accuracy, respectively. Naive Bayes outperformed Random Forest, and hybrid models performed better than BiLSTM. Deep learning showed superior learning capacity, with CNN being the best-performing model overall, with an accuracy of 85%.

The table below shows the accuracy of various models used:

Algorithms	Accuracy
Naive Bayes	75%
Random Forest	74.5%
BiLSTM	78%
LSTM	80.67%
CNN-BiLSTM	80.13%
Support Vector Machine	82%
CNN-LSTM	82.23%
CNN	**85%**

Future works:

In recent years, sentiment analysis with emoticons has become more important due to the increased popularity of emojis in modern media. There are opportunities for developing comprehensive emoji libraries, conducting cross-lingual mood analysis, and comparing emojis and text. Several algorithms, including rule-based approaches, emotion lexicon-based approaches, emoji embedding-based approaches, and hybrid machine learning approaches, can be further explored to achieve these tasks.

Proposed Arch:

To increase the accuracy of the sentiment analysis model to more than 90%, incorporating more advanced deep learning and hybrid methods can be beneficial. Here are some techniques that we can implement:

1. Gated Recurrent Unit (GRU):
   GRU is a variation of recurrent neural networks (RNNs) that can effectively capture long-range dependencies in sequential data. Replacing LSTM with GRU in the model architecture might lead to better performance.
2. Transformers:
   Transformers, particularly the Transformer architecture introduced in the "Attention is All You Need" paper, have shown remarkable success in various natural language processing tasks. Implementing a transformer-based model can potentially improve the model's understanding of context and relationships between words.
3. BERT (Bidirectional Encoder Representations from Transformers):
   Pre-trained BERT models have achieved state-of-the-art results in many NLP tasks. Fine-tuning a pre-trained BERT model on the sentiment analysis task can bring significant improvements to the accuracy.

In addition to this, the quality and size of the dataset also plays a major role. The dataset used for training the sentiment analysis model should be representative and diverse, covering various sentiments and emotions. A larger dataset can provide the model with more examples to learn from, potentially leading to better generalization.

References:

Reference papers:

1: Sentiment Analysis of Comment Texts Based on BiLSTM

2. Sentiment Analysis With Comparison Enhanced Deep Neural Network

3. A Sentiment Polarity Categorization Technique for Online Product Reviews

4. Sentiment Analysis Algorithm Based on BERT and Convolutional Neural Network