Journal of Information Systems Engineering and Management
2025, 10(12s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Extreme Learning Techniques for Enhanced Sentiment Analysis

Akkireddi Vara Prasad [1], K. Vedavathi[2]
[1]*Research Scholar, Department of Computer Science, School of Science, GITAM (Deemed to be University), AP, India*
[2]*Professor, Department of Computer Science, School of Science, GITAM (Deemed to be University), AP, India*

ARTICLE INFO

Received: 24 Nov 2024

Revised: 01 Jan 2025

Accepted: 21 Jan 2025

ABSTRACT

Extreme learning approaches are used to perform sentiment analysis on restaurant evaluations. Sluggish training and overfitting are two problems that traditional supervised learning techniques frequently face. In order to prevent overfitting and speed up training, extreme learning makes use of single hidden layer neural networks with randomly assigned weights. The goal is to create a quick and accurate model for identifying sentiment in meal reviews and to assess the differences between supervised learning techniques and models based on extreme learning. In order to map the scores to attitudes, the study uses a dataset of food reviews, where scores larger than three are interpreted as favourable and scores below that as negative. Training and testing sets are created from the dataset following preparation, which includes handling missing values and choosing pertinent columns. For modeling purposes, text data is transformed into Term Frequency-Inverse Document Frequency (TF-IDF) characteristics. The network is given randomly initialized weights and biases in both single layer and multi layer perceptron implemented. During model training, the loss function is computed, weights and biases are adjusted by backpropagation, and predictions are computed using forward propagation technique. By using a threshold of 0.5, the model's accuracy is assessed. Accuracy scores are used as a statistic for reporting training and testing accuracy. The model further validates the effectiveness for sentiment analysis in the context of food reviews by showing quicker training times and less sensitivity to overfitting. The work presents extreme learning approaches as a competitive substitute for supervised learning, which advances sentiment analysis tools. Comparing the model based on extreme learning to traditional supervised learning methods, experimental results show that the latter achieves competitive accuracy in sentiment analysis. Faster training times and less sensitivity to overfitting are further features. As a strong substitute for supervised learning techniques, this study highlights the effectiveness of extreme learning approaches for sentiment analysis in meal evaluations. Extreme learning improves the efficacy and precision of sentiment analysis models, especially in areas such as restaurant evaluations, furthering the practical uses of sentiment analysis tools.

Keywords: Sentiment Analysis ; Food Reviews ; Text Mining; Neural Networks, Extreme Learning; Single Hidden Layer; Overfitting; Word Frequency ; Term Frequency-Inverse Document Frequency

INTRODUCTION

Reviewing food items is one of the most important ways that these affect consumers' opinions and decisions to buy. The demand for automated methods to evaluate and draw conclusions from the wide range of food-related evaluations that are available online has grown. Within natural language processing (NLP), the discipline of sentiment analysis provides a methodical way to identify the sentiment conveyed in textual data, giving important information about the preferences and opinions of consumers.

Applications of sentiment analysis have unique opportunities and difficulties when it comes to food ratings. Interpreting subjective opinions and attitudes about food goods is the aim of sentiment analysis in food evaluations, dining experiences, and culinary establishments in a nuanced manner, as opposed to traditional text classification tasks that may concentrate on differentiating between discrete categories. An additional degree of complexity to the analysis process is the naturally varied and context-dependent nature of language connected to food.

These issues have prompted researchers and industry professionals to explore various machine learning approaches to develop sentiment analysis models tailored for the food reviews domain. Many of the traditional methods such as supervised learning techniques, logistic regression and support vector machines commonly used for sentiment analysis tasks face challenges for large-scale datasets and high dimensional feature spaces. The training time of these techniques are slow, as well as overfitting.

The research conducted in the existing literature examines extreme learning approaches and their use in sentiment analysis of meal reviews in order to address these deficiencies. Extreme learning can replace traditional supervised learning methods by using single hidden layer neural networks with random weights. Extreme learning avoids overfitting while accelerating training to provide robust and fast sentiment analysis of food reviews using randomly initialised weights and biases.

The intent is to build a robust sentiment analysis model that can accurately detect and classify food items reviews. It also attempts to compare extreme learning based models with more traditional supervised learning methods. Based on solid testing and examination, this study attempts to determine whether extreme learning methods are valuable and applicable for meal review sentiment analysis tasks.

There are several important steps in the recommended course of implementation. First, a collection known as a meal dataset is acquired through scraping and transformed into raw data that has undergone preprocessing to obtain crucial information and affect problems such as missing data and noise. Then, using a statistical method (TF-IDF) that assesses the importance of a word or terms in a document within an entire text corpus, we convert the text data into feature vectors with numeric representations of features.

Then, the sentiment analysis model is created, which is a single hidden layer neural network that receives the TF-IDF features extracted from the preprocessed text data as the input layer. The model architecture includes a hidden layer with the number of neurons selected previously and an output layer that shows the expected sentiment (positive or negative) from the analysis of the food. The most often used activation function in the neurons of hidden layers is the sigmoid function, which permits the non-linear transformation of the input to enable models to learn complex functions.

The first step of the model learning process is to adjust model parameters, including weights and biases, to minimise the prediction error on a labelled training dataset. Evaluation methods and several metrics to assess the model's performance are introduced, including precision, accuracy, recall, and F1-score, so as to give more insight regarding its performance in recognizing the sentiment expressed in meal reviews. By demonstrating that extreme learning methods can form a feasible procedure for evaluating and understanding opinion associated with food assessments, this work enhancing the knowledge in the opinion reviewing domain. We focus on systematically comparing extreme learning based models to typical supervised learning methods to provide valuable information regarding the pros and cons of each approach.

LITERATURE SURVEY

R. Liu et al have collated the significant sentiment analysis research findings over the last few years. Our expectation is that this area will continue to grow, since their summary focuses on the techniques and uses of transfer learning in sentiment analysis [1]. An enhanced LSTM model incorporating lexicon has been proposed by Fu et al. The model first pre-trains a word sentiment classifier using the sentiment lexicon as an additional information source. From there, it extracts the sentiment embeddings of all words, including those that are not in the lexicon. Word representation accuracy can be increased by combining sentiment with word embedding.

SVM method, Particle Swarm Optimization and multiple oversampling strategies are used in a hybrid strategy developed by Ruba et al. to address the imbalanced data problem. Using machine learning classification, SVM is used to optimize a dataset of various reviews from various restaurants in Jordan in order to predict the sentiments of the reviews [3]. In an experiment, Ali et al. used deep learning algorithms to evaluate the sentiment of the tweets. Accuracy was also improved by building prediction models and utilizing CNN, RNN, and LSTM. Combining the

Word2vec feature extraction method with the CNN and LSTM algorithms produced the results with the highest accuracy [4]. According to Ishtiaq et al., the Light Gradient Boosting Machine (LGBM) technique achieved a 98% accuracy rate, outperforming current state-of-the-art systems. A k-fold method is used to confirm each technique's performance, and hyperparameter optimization is then used to further enhance it [5]. CNN, hybrid, and RNN Among the three deep learning methods that Mohammed et al. employed for their categorization were CNN-RNN models. Two distinct applications of each technique were made: one on the original feature set and the other on the reduced feature set, where the classification was based on Binary Coordinate Ascent and Optimal Coordinate Ascent.

Lee et al., have focused on two phases. It was found that the average restaurant rating, the variety of prices, the quantity of reviews collected over time, and the elite status of the establishment might all potentially improve the effectiveness of restaurant suggestions [7]. There were only a few fields in the datasets that Obiedat used. In order to help academics interested in Aspect-Based Sentiment Analysis create models approach the problem

Using the attention mechanism, Jin et al. employ a maximum pool layer of a convolutional neural network to detect the significant parts in texts by determining which words are essential for text categorization and assigning them greater weights [9]. To improve target-polarity identification, the Ghada et al. model was assessed using a Conditional Random Fields classifier in the Arabic translation of the Bidirectional Encoder Representations from Transformers model. The results of the trial demonstrated the effectiveness of the recommended optimized model [10]. Results show that SEBA performed better than individual baselines and demonstrated superior performance using TF-ISF and unigrams as features across deployed datasets, as reported by Ankita Sharma et al. [11]. Rahman et al. reported experimental studies conducted under various instruction tuning configurations [12]. Wang et al have presented the experiment results that indicate that while homophily reduces participants' ability to estimate stock prices accurately, investor mood is influenced by both the agreement index and post volume. A BERT-based sentiment analysis model was presented [13]. SentDep outperforms the state-of-the-art models, establishing a new benchmark for MSA performance, according to Lu et al.'s detailed experiments on well-known datasets including CMU-MOSI and CMU-MOSEI. These studies emphasize how important pre-training data volume is, how effective different fusion techniques are, and how important temporal information is to increasing the model's competency [14]. The use of the creative method in two case studies involving laundry detergent liquid bottles and pods sold on Amazon was covered by Mahsa et al. These tests' results show how well WIPE can recognize specific packaging flaws and anticipate their possible causes, as well as extract relevant information from customer input. Considerable progress has been made in the field with the addition of association rule mining and sentiment analysis to the package evaluation process [15]. According to Okoye et al., the marginal means of the effects of the different SET types and the evaluation time on the learning outcomes and teaching-learning process viewpoint of the students were investigated by quantitative analysis. A multivariate analysis of covariance and multiple pairwise comparisons post-hoc tests were employed by the study to analyze the quantitative variables extracted from the SET data[16]. Based on syntactic and semantic graph convolutional networks, Chen et al. have discussed how the model improves text representation capabilities. To prove its efficacy and validity, we ran extensive tests on datasets that are accessible to the general public. Our model does better than robust baseline models, as the experimental data show [17]. The data augmentation technique was presented by Tareq et al. and demonstrated to outperform other alternative models in automated sentiment [18]. Chaisen et al. have presented s sentiment analysis ability to discover the distinct correlation between particular sensation features and the sentiment classes that correlate with them. This work closed the gap between predicted performance and model, as well as advancing sentiment analysis in the Thai language[19]. Research by Jung et al., has demonstrated that the suggested approach enhances performance when used with various language models, especially in low-resource settings with a limited amount of training data. According to the study's findings, the model may be trained with task- and domain-appropriate datasets in a variety of disciplines, including finance, to increase its understanding of areas in which it is weak

Using a dataset of 1.6 million English reviews for four different Steam games—PUBG, CS:GO, Dota2, and TeekEN7—Yu et al. ran an experiment. The reviews were collected up to December 2021. As evidenced by the testing findings, the suggested framework is able to effectively determine the worries of players and uncover intriguing keywords that are hidden beneath their reviews. Therefore, it gives esports game operators accurate information and insightful user reviews, allowing them to improve their offerings and give gamers a better overall gaming experience [21]. The benchmark Loughran-McDonald (LM) lexicon's vocabulary coverage can be improved by using transformer-aided explainable lexicons, as shown by Maryan et al. This improves the lexicons and reduces the amount of human intervention needed for annotating, updating, and maintaining them. Furthermore, our findings demonstrate that the generated vocabulary performs better in sentiment analysis of financial datasets than the conventional LM lexicon [22]. To help with the understanding of subjective expressions, Ahmed et al. have presented a text detection system for the Arabic language that distinguishes between extremist views. The paper offers a cutting-edge method based on rough set theory to improve some models' accuracy and precisely identify text orientation [23]. A technique that modifies the BiLSTM has been proposed by Jawad Khan et al. to extract semantic information about word order and context from text as well as the long-dependency relation seen in word sequences. Our model additionally makes use of an attention mechanism to further distribute weights to characteristics and rank important aspects in the word sequence. In order to reduce the dimensionality of the feature space and retrieve the local critical characteristics for sentiment analysis, CNN is ultimately utilized [24].

Aiming to offer a more effective and efficient sentiment analysis solution, MPNet-GRUs combines the advantages of these models. MPNet-GRUs demonstrate potential to increase sentiment analysis as evidenced by their better performance on three benchmark datasets (86.27% for Twitter US Airline Sentiment, 94.71% for IMDb, and 88.17% for Sentiment [25]. The ensemble approach improves SA performance and reduces the chance of relying solely on a single model bias or error, as Hala et al. have stated. The RTH method is utilized in the hyperparameter tuning strategy to increase the SA's performance. The improved SA findings of the ASA-RTHEL approach were ensured by a comprehensive set of trials. The thorough comparative analysis demonstrated how the MPONLP-TSA approach improved results for identifying different types of emotions [26].

Yang et al.'s recommendation model now includes two additional features: an interaction graph and a review text graph. For instance, the model can display the review text per user and item as a graph to represent the discontinuous, global, and long-term dependencies between words in the text. The graph structure is used in this embodiment. A graph attention network based on connection links is used to extract each node's adjacency information while taking word order correlations into consideration. By employing the user item ratings, on the other hand, an interaction network is constructed through the application of feature mining methodology. Once the outcomes from all segments are combined, the prediction is finished. Equipped with three datasets, the suggested [27]. Govardhan et al.'s work makes extensive use of the Cloud Service Measurement Index Consortium established properties of the Service Measurement Index (SMI). A thorough investigation is carried out to find out whether a user has a favorable, unfavorable, or neutral view on the parameters. The parameter repute on both positive and negative orders are taken into consideration here. The positive parameter reputation increased the choosing and rating of services among cloud users. The service providers were exposed to adverse parameter repute, which allowed them to evaluate their existing state of affairs and make the required adjustments for their future service delivery [28]. Ahmed et al. have presented a unique method based on Rough Set theory to accurately identify text orientation and enhance the accuracy of some models. Experimental results show that the proposed technique helps with feature discriminations and outperforms existing algorithms. This work makes a significant contribution to the limited body of research on machine learning and Arabic language linguistics [29]. With the purpose of contextualising various methods, Z. Hu sought to present a comparative assessment of deep learning for Fine-grained Sentiment Analysis applications [30]. When applied during training, the threshold-moving method enhanced sentiment learning performance according to the findings of Fatimah et al.'s study [31]. A study conducted by Minchao Ban et al. [32] focused mostly on aspect-level multi-model sentiment analysis. Yang et al., have created a graph representation of the model to assess the text for every user and item, capturing the continuous, universal, and long-term relationships between the words in the review text. A graph attention network is utilized to collect each node's adjacency information while accounting for word order correlations based on connection linkages. However, the feature mining model uses user-item ratings to form an interaction network [33]. The effectiveness of each method employed in the experiments is given by Razali et al. in their discussion. According to our proposed framework, the hybrid Lexicon-based technique in combination with the Decision Tree classifier yielded the best results for political security threat prediction in this study. These results add significantly to the ongoing research on opinion mining threat prediction by offering new insights from the perspective of political security

Ali et al. claim that the model increases prediction accuracy by giving greater weight to data points that are closer to the projected point in time. It is very adept at seeing trends over time. Ablation investigations confirm that the TLSTM and Off-policy PPO components have a good effect on the model's overall performance. The proposed method contributes to the field of financial analytics by providing a more sophisticated understanding of market dynamics, while also providing investors and policymakers with useful information to assist them navigate the complexities of the stock market with greater assurance and precision [35].

In order to capture the task-specific features for each task as well as the task-cross features across the two tasks, Diao et al. have emphasized that learning both tasks simultaneously and studying weight sharing is the foundation of deep neural networks. Large-scale tests on real-world datasets show that, thanks to the strong semantic linkages, our proposed model consistently outperforms state-of-the-art baselines in both the sarcasm detection and comedy identification tasks [36]. In their research, Sundaram et al. have shown how important social media lingo is for sentiment analysis because it can provide light on public opinion and aid in making well-informed decisions. Our research highlights the significant role that slang analytics plays in enhancing communication strategies and offering specialized insights in a wide range of companies and research disciplines. The study concludes with a thorough grasp of the dynamic terrain of colloquial language in the context of modern digital communication, providing insightful information that helps with decision-making, marketing strategy improvement, and communication improvement [37]. DGCSKT leverages dependency graph convolutional operations and syntactic dependency graphs to improve the model's comprehension of contextual information, as explained by Fan et al. Furthermore, data-homogeneous sentiment resources were incorporated as a bottom-level auxiliary task to exchange effective sentiment features and enhance recognition performance. In order to enhance the model's capacity for generalization, we subsequently suggest weighting the training data from various tasks using the Dynamic Normalized Weighting technique. When compared to the most advanced methods currently in use in HateEval, our recommended approach enhances the Macro-F1 by 0.54% and 3.88%, respectively [38]. Meng et al. have explained that this model's main purpose is to do sentiment analysis on review data related to smart tourism, specifically by classifying the data as positive or negative feelings. The process of efficiently annotating data is streamlined by the addition of review ratings, which complement the process of accurately classifying the data in accordance with the sentiment categories involved. In terms of favorable remarks across a range of evaluation indicators, the empirical results show that the three models have different performance trajectories [39]. In order to categorize implicit manifestations of hatred, Jafari et al. have developed a multi-task learning strategy that combines sentiment characteristics and emotions. We compared our multi-task learning strategy to baseline models that were trained using single-task learning approaches in order to assess its efficacy. As demonstrated by the experimental results, our multi-task strategy is able to categorize implicit hate speech better than baseline models, and it also minimizes classification error across many implicit hate categories when fine-grained emotional information is added. [40]. In order to streamline the procedure and ensure a thorough and accurate risk factor evaluation and filtering, Zhao and colleagues have made it easier to include the experience of domain experts into the risk factor identification and extraction process. Next, utilizing the co-occurrence correlations between the risk variables, a network representing the relationships between them was constructed. This network was then the subject of a quantitative research. Because the methods employed in this study are supported by real-world gas network management and emergency cause data, they provide significant support for gas network safety management and emergency decision-making.

Between the two models, Nur Bengisu et al.'s accuracy reached 98%, surpassing the sound test findings. This indicates that sentiment analysis carried out with limited data can benefit from the Multimodal model approach [42]. Comparing Ming Zhang et al.'s performance to the most advanced models, they demonstrate notable advancements in seven semantic relatedness tasks [43]. The suggested method by Zaikis et al. was assessed using Greek expert annotations on texts taken from Internet articles, blog posts, opinion pieces, and press text clips as well as social media posts. Using three custom classification tasks and the micro-averaged F1-scores for irony, sentiment, hate speech, and emotion, the results demonstrate a considerable improvement in each task's classification effectiveness [44]. By looking through highly referenced papers, recent studies, and regional sources, Amiri et al.'s analysis provided a wide viewpoint. Factorial analysis reveals thematic clusters related to user experience, collaborative filtering, emotion identification, and reinforcement learning. Research themes from various historical eras are categorized using a scientific mapping analysis, which emphasizes key topics including sentiment analysis, emotion recognition, collaborative filtering, and hybrid recommendation. The significance of digitalization, emotion recognition, personalization, user experience, and collaborative filtering is highlighted by a study of thematic evolution in guiding future research approaches [45]. Kim et al. have conducted a client-oriented analysis of metaverse services, which resulted in the identification and ranking of nine service elements that significantly affect client happiness. Therefore, when it comes to user goals when interacting with metaverse services, the "co-experience" aspect becomes apparent as a critical component. From these results, service managers can better position themselves in the dynamic metaverse market by improving their services

In order to improve the model's capacity for generalization and reduce the possibility of over-fitting, Bo et al. have suggested that the deep fusion model makes use of the complimentary capabilities of these elements. Lastly, we show the efficacy of the suggested methodology by contrasting it with the most advanced techniques for classifying medical documents [47]. Improvements in privacy, social skills, and contextual awareness should be the main goals of sIPAs for senior citizens, according to Islam et al. Additionally, useful recommendations for integrating humanized communication methods, comprehendible artificial intelligence (XAI) concepts, permission-based data storage, and accent and dialect recognition. With the goal of expediting the enhancement of sIPAs targeted at older persons, this research offers guidance for both design and implementation [48]. Teachers may now make data-driven judgments and customize pedagogical procedures to match the individual needs of every student thanks to Liu et al.'s thorough data processing and interpretation using sophisticated approaches.This data-driven pedagogical approach not only facilitates the adoption of effective teaching approaches but also successfully addresses inequalities arising due to differences in student backgrounds leading to a more inclusive and personalized learning environment [49]. The ensemble method and sequential models successfully combined with BERT embeddings can bring positive outcome. The contributions of the study in Busse et al. are the aspect category features recovered by the proposed Ensemble BiLSTM model. Next steps will include accurate aspect-level sentiment features, which will be used in coming research [50].

The exploration of extreme learning techniques for meal evaluations holds significant promise for improving the accuracy and efficacy of sentiment classification. Compared to traditional methods, techniques like Extreme Learning Machines (ELMs) offer an advantage in handling the large and heterogeneous datasets typical of food reviews because of their exceptional capacity to train quickly and generalize well. These extreme learning approaches are able to more accurately capture delicate feelings and contextual details by utilizing large-scale, pre-trained language models and incorporating sophisticated pre-processing procedures. By using these cutting-edge techniques, sentiment analysis becomes more accurate and helps in identifying market trends and consumer preferences.

PROPOSED MODEL

There are several important steps and methods taken in the implementation of Extreme Learning in for sentiment analysis on the food reviews dataset to help you quickly and accurately capture the emotions provided in the reviews. Extreme learning takes advantage of single hidden layer neural networks with random weights to reduce overfitting and speed up training times. The first step in the methodology is data preprocessing. This includes importing the dataset, selecting relevant columns (like "Score" and "Text"), and handling missing values. In this implementation, missing values are eliminated to ensure data completeness and quality. In addition, these scores have sentiments associated: if a score is equal or above 3, it is considered positive, and equal or below 2 is considered negative.

After that, the train_test_split function splits the data into training and test sets. Therefore, to guarantee that the model has been trained on a portion of the data and that test data is unseen, the generalization performance of the model can be explored. An essential step in the process is feature extraction with TF-IDF vectorization. TF-IDF stands for Term Frequency-Inverse Document Frequency, which is a numerical representation of how significant a word is in a document relative to a collection of documents. TF-IDF features are created by transforming the text data into a form that can be input into machine learning algorithms. In this implementation, we performs TF-IDF vectorization using the scikit-learn Vectorizer class, limiting dimensionality to a maximum numbers of features. Due to the immense size of the dataset, TF-IDF features are generated in batches so to control memory usage. This method of batch processing ensures efficiency in processing text data while also not crashing the memory. Since the batch process text function creates a batch of training data, the TF-IDF features are independently computed for this batch. After the feature extraction stage, the implementation of a SLNN (single hidden layer neural network) is applied The SLNN architecture consists of three layers: an input, hidden and output layer. First, weights and biases are randomly assigned to the hidden layer. So, to make it non-linear, we make use of sigmoid function as activation function in hidden layer. The SLNN predicts the reviews' sentiment, using the TF-IDF features extracted from both the text data. Using this technology to conduct sentiment analysis of restaurant reviews is based on the Multilayer Perceptron architecture, with additional hidden layers. The MLP model makes sense of restaurant reviews in a better way than other models, capturing and interpreting multilayers of emotions that are expressed in text reviews through nonlinear activation functions together with the TFIDF features, ultimately helping us understand better and analyze the restaurant reviews.

The contributed TF-IDF feature-based training of SLNN model accuracy ratings on a testing dataset is calculated, which is them compared with ground truth labels to give predictions. A confusion matrix is used to evaluate the effectiveness of the model at determining each positive, negative emotion. Gives a sentiment analysis on Contextual, Restaurant, Bar and Overall Food Reviews Dataset applying Extreme Learning techniques conjointly using batch processing, TF-IDF vectorization and SLNN architecture to alleviate the effects of overfitting and to increase the speed of training.

Extreme learning applies single hidden layer neural networks to solve supervised learning problems by arbitrarily assigning weights to hidden layer neurons, different from well-known supervised learning methods, which apply iterative optimisation algorithms to modify the neuron's weights. This speeds up the training process, lessening computing overhead and aiding in convergence. The weights and biases of extreme learning networks are randomly initialized, which lowers the chance of overfitting (overfitting is prevalent in complex structures of neural networks).

The following are several advantages of extreme learning architectures compared to classical ultra-deep networks, transparency (no complicated algorithms are needed), simplicity (they require less time), low computing resources (can be placed on almost all machines), less overfit tendency, good scaling. They are good at dealing with large amounts of datasets and enabling real-time or near real-time analysis of streaming data sources and rapid insights into customer moods and preferences. Given the volume and velocity of user-generated information on online buildings and social media, this scaling property is particularly important.

This paper first introduce multiple hidden layers on top of the Multilayer Perceptron (MLP) neural network architecture to analyse the sentiment of restaurant reviews. One of the most basic types of artificial neural network, which can be used, is the multilayer perceptron, which can describe complex nonlinear relationships between input and output data. The three layers that make up the MLP architecture are the input, output, and one or more hidden layers. The basic building blocks of the Multilayer Perceptron (MLP) are the hidden layers, which perform the duty of recognizing complex relationships and features in input data, allowing the network to learn from this and extend more generally as a function of the training set.

Extreme learning enjoys other advantages of parallelization and distributed computing capabilities which are denote for sentiment analysis in restaurants reviews. Due to its self-evident simplicity, the extreme learning architectures can be deployed in parallel on distributed computing environments or multi-core processors, so that it exploits all the computational resources available and strongly accelerates the processes related to the model training and inference [3]. To fully take advantage of parallel computing architectures, sentiment analysis workflows can be optimised further through the use of parallelisation strategies such as data parallelism, model parallelism and so on.

Despite the advantages of radical learning, we need to recognize the inherent trade-offs associated with it. Random initialization of weights and biases may be a source of weak solutions or convergence to local minima, particularly in high-dimensional or complex feature fields. Furthermore, deep learning models are generally much harder to interpret than simpler linear or decision-tree models, so even if they give better predictions, it may be difficult to figure out what features are driving any particular prediction.

The use of extreme learning in sentiment analysis of meals reviews helps reduce overfitting and increases training speed by utilizing single hidden layer neural networks with random weight assignment. By employing the simplicity, scalability, and computation efficiency of extreme learning methods, researchers and practitioners can develop reliable and effective sentiment analysis models that extract useful information from large textual data. A good deal of research and experimentation on this topic could enhance the performance of sentiment analysis and contribute to the development of more accurate and scalable techniques of assessing what consumers feel or prefer about product reviews.

Utilizing hidden layers, the MLP creates a higher dimensional space in which complicated patterns and relationships may be discovered more simply. The MLP can capture more complex and subtler representations of the information by adding multiple hidden layers, allowing it to learn hierarchical representations of the input data. Each hidden layer neuron applies a nonlinear activation function to the weighted input sum. This is why modelling nonlinear interactions between input data and output data strengthens the MLP's understanding of complex data patterns such as subtle sentiment in restaurant reviews by transforming input data in a non-linear fashion.

Furthermore, the optimization of model learning dynamics and performance in MLPs is heavily reliant on the selection of activation function in the hidden layer, such as the well-known sigmoid or rectified linear unit (ReLU) functions. With these activation functions, the network can learn complex functions and different ways to represent the input data. Another important hyperparameter is the hidden layer size, which is the number of neurons in each hidden layer, which also determines the capacity and flexibility of the MLP model.

TF (Term Frequency) calculates the frequency of a term in the text against all terms in the document. It assigns higher weights to the terms since more frequently occurring terms are more likely to be indicative of the important of the content within the document. To address this limitation, each term has an evaluation referred to as the Inverse Document Frequency (IDF), which acts on the rarity of the term throughout the document collection. It penalises term frequency across many documents and rewards term rarity, and so quantifies the discrimination potential of a space. The higher the IDF score (for a given term), the more exclusively it appears in a given document, or the rarer it is in the entire collection of documents, implying that it may be important enough to differentiate the text from other text.

TF-IDF is calculated by multiplying TF by IDF, producing a numerical representation of the term's importance within the document collection. TF-IDF scores highlight terms that are uncommon in the document collection and common in a document, perhaps making the terms prominent or telling. Thus, by utilizing the hidden correlation between the scarcity of a document and the occurrence of a word, TF-IDF allows specialists and researchers to retrieve relevant content and extract defective patterns from numerous textual databases on a large scale, assisting in more precise and context-aware analysis of the data presented in text documents.

Measurements and techniques including recall, F1 score, ROC curve, training accuracy, testing accuracy, and confusion matrix are essential for assessing and understanding classification algorithms. Precision highlights the model's ability to reduce false positives by computing the proportion of accurately predicted positive cases among all cases expected to be positive. The model's recall measures how well it captures all pertinent positive events by dividing the number of correctly predicted positive instances by the total number of actual positive instances. Recall and accuracy are combined into a single statistic called the F1 score, which produces a fair assessment of a model's performance that takes into consideration both false positives and false negatives. In order to provide information regarding a model's discriminatory power and performance across a range of judgement thresholds, the Receiver Operating Characteristic (ROC) curve depicts the trade-off between true positive rate and false positive rate over various categorization thresholds. The percentage of properly identified occurrences in the training dataset, or training accuracy, is a measure of a model's capacity to learn from training data. Testing accuracy, which measures the model's generalizability, is a valuable indicator for evaluating a model's performance on unidentified data. A tabular representation of the predictions a classification model generates for correct and erroneous classifications is displayed in the confusion matrix together with information on true positives, true negatives, false positives, and false negatives. This provides insightful information about the model's advantages and disadvantages, including its capacity to distinguish between various classes and spot similarities.

These metrics and techniques are essential at these stages to evaluate the effectiveness, robustness, and interpretability of classification models with respect to model selection, hyperparameter tuning, and feature engineering. Researchers can thus systematically analyse and interpret these indicators to support their decisions about the deployment of models and the optimisation strategies, which ensures the development of the accurate and reliable classification models for real-world applications.

RESULTS

Table1: Performance comparison of SLP and MLP

	Single Layer Perceptron	**Multi Layer Perceptron**
Training Accuracy	0.6329	0.7828
Test Accuracy	0.6435	0.7825

A single layer perceptron (SLP) and Multilayer Perceptron (MLP) show relatively different results in training and testing accuracies. The SLP's ability to learn from the training data can be seen from its training accuracy of 0.6329, as shown in table 1. Its testing accuracy does increase slightly to 0.6435, suggesting a limited ability to extrapolate beyond the training set. In contrast, the accuracy of the MLP is higher both in training and testing, with a 0.7828 accuracy in training and a 0.7825 accuracy in testing respectively. This demonstrates that the MLP is able to learn a much more complex function because it has the capability to model more complicated relationships than that of a single layer and thus outperforms a linear model by a significantly larger margin in classification and display of data. SLP has only one layer of neuron and uses linear activation functions, hence, it may fail to predict the outcomes for nonlinear patterns present in the data. The MLP is thus a more powerful tool for tasks requiring greater complexity and precision in pattern recognition and classification, its numerous hidden layers and nonlinear activation functions allowing it to better learn highly nonlinear functions. It is also capable of efficiently describing complex relationships. This also shows the model accuracy with training and testing dataset respectively. Based on the model's training accuracy of 0.782, approximately 78.2% of the occurrences in the training dataset are correctly identified. This demonstrates that the model can accurately classify samples that it has already seen and that it has gained generalization skills from the training set. The model's performance on untested data must be assessed in order to determine how well it generalizes to new scenarios.

The performance metrics of a Single Layer Perceptron (SLP) and a Multilayer Perceptron (MLP) in terms of F1-score, recall, and precision are shown in the provided data shown in table 2. The precision of 0.6329 is attained by the SLP, signifying the percentage of accurate positive predictions among all positive predictions generated by the model. But at 0.989, recall is much greater, indicating that the model can accurately identify the great majority of real positive events.

The computed precision, recall, and F1-Score values for the model are listed below:

Table2: Metric comparison of SLP and MLP

Metric	**Single layer Perceptron**	**Multilayer Perceptron**
Precision	0.6329	0.7825
Recall	0.95	1
F1-Score	0.798	0.878

We can see that after applying MLP model gives the accuracy and precision of the model is better than the SLP model, we have the F1-score of 0.798, that is, the model perfectly balances between recall and precision. It achieves 0.7825 for its positive cases over its peers and has a recall score of 1. This leads to a substantial increase in its F1-score to 0.878, signifying a better compromise between recall and precision. MLP's computational flexibility allows it to learn more complex representations from the data due to its multi-layered structures and nonlinear activation functions, providing a significant edge over the SLP in terms of improved classification accuracy.

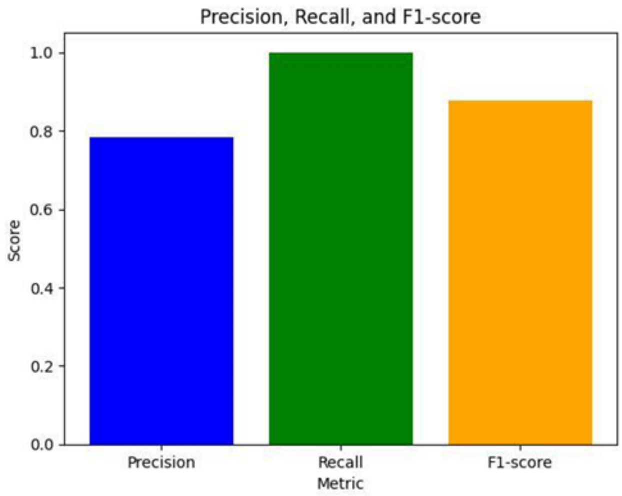

Fig 1: Performance analysis of the proposed MLP model

The precision, recall, and F1-score values can be understood based on the performance of the binary classification model. The precision value of 0.782, approximately 78.2% of positive predictions are correct, demonstrates that the model can keep the false positive rate low. This indicates that the model is relatively effective in detecting positive instances, an important factor in scenarios where costly or unwanted false positives may occur.

Recall at 1.0 means that all positive events in the dataset have been captured — there would be maximum sensitivity. Thus, since it never overlooks a positive scenario, the model clearly demonstrates its ability to identify each possible instance of the target class. In particular, we would prefer a recall of 1.0, when the lack of good instances would have dire consequences. The F1-score defines which combines recall and precision into one number, being the harmonic mean of the two. The F1-score of 0.878 indicates that the model has a good recall and a good precision. With skill at minimizing false negatives and false positives, the model works well for applications that require recall and precision. These results would suggest a well-performing classification model with a high F1-score, which is able to accurately trade-off recall and precision successfully.

Receiver Operating Characteristic (ROC) curve is the image that is often used to evaluate the performance of binary classification algorithms. It plots the true positive rate (sensitivity) versus false positive rate (1 - specificity) for a range of different threshold values for the positive class. True positive rate evaluates the fraction of positive cases identified correctly by the model, while false positive rate evaluates the percentage of negative class cases which were falsely identified as positive by the model. Researchers may then use a range of judgment thresholds to assess how well the model distinguishes between the two classes. ROC curve (Receiver Operating Characteristic) — Shows a trade-off between the true positive and false positive rates by varying the classification threshold. True positive rate vs false positive rate over settings of thresholds yields the receiver operating curve (ROC) and in this plot, a model which is more discriminatively powerful would end with a ROC curve which is closer to the upper left corner of the figure. Area Under the Receiver Operating Characteristic curve (AUC-ROC) is a performance measure that condenses the performance of the model, a higher AUC-ROC score (closer to 1) is indicative of a better discriminative power of the model whereas an AUC-ROC score close to 0.5 indicates a random model. A ROC analysis is a useful method for assaying performance characteristics and discriminatory power of many models or algorithms. It also tells you something useful about how well classification methods perform. The area under the ROC curve of the model is 0.5 as shown in Figure 2.

Each of these metric has optimum value depending on individual goal and requirements of the classification task. Higher accuracy and recall scores generally reflect a model's ability to maximize true positive predictions while minimizing false negatives and false positives. Whereas the recall of 1.0 means the model has correctly identified all positive instances in the dataset the precision of 1.0 says the model has made all correct positive predictions. The higher the score, the lower the overfitting (fewer hallucinations), better memory and accuracy balance. The harmonic mean of precision and recall can be derived as a single result, called the F1 score, from these two metrics.

Likewise, for the Receiver Operating Characteristic (ROC) curve, for each feasible classification threshold, the ideal model would yield a true positive rate (sensitivity) of 1.0 and a false positive rate (specificity - 1) of 0.0. That would lead to a ROC curve hugging the upper left corner of the plot and properly discriminating between the two classes.

As for accuracy measures, the ideal values would likewise be 1.0 for training accuracy and testing accuracy, meaning that all instances in the training and testing datasets, respectively, are properly classified by the model. A model that has an abnormally high training accuracy may have memorised the training set instead of learning significant patterns, hence it's important to take potential overfitting into account when assessing training accuracy. Perfect classification performance would be indicated by all off-diagonal components being equal to zero, and diagonal elements reflecting true positive and true negative predictions in the convolution matrix of ideals. To minimise these errors while maximising accurate predictions is the aim, as certain misclassifications will inevitably occur in practice.

Overall, these metrics provide as standards for evaluating the effectiveness of classification models and for informing advancements in model design, training, and evaluation—even though reaching optimal values for these metrics may prove difficult in practical situations.

If the classification model's ROC value is 0.5, it is equivalent to random chance in terms of performance. Put otherwise, flipping a coin has no greater discriminating power than the model when it comes to positive and bad episodes. The particular environment and goals of the classification assignment must be taken into account, even though a ROC of 0.5 may indicate that the model lacks discriminatory ability. Even a moderate ROC value may be useful in some situations, but more model performance improvements are required to get conclusions and insights that are significant.

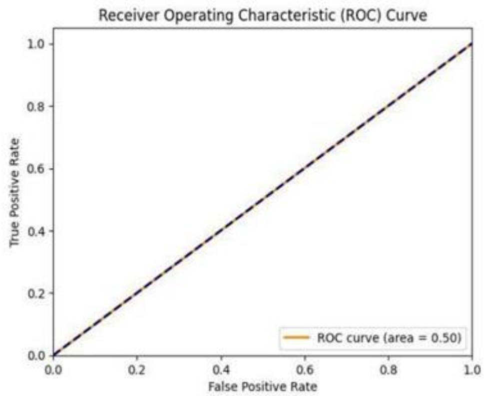

Fig 2 : ROC plot of the proposed model

Figure 3, which depicts the classification performance, displays the convolution matrix of the suggested model. In machine learning, a tabular representation known as a confusion matrix is used to evaluate a classification model's performance. It shows how many predictions the model made correctly as true positives and true negatives (along with false positives and false negatives) based on a test dataset. The rows of the confusion matrix represent the true class labels, and the columns represent the predicted class labels. The diagonal portion of the matrix represents successful predictions, where the predicted class matches the true class. Misclassifications are shown as off-diagonal elements, whereas false positives and false negatives indicate cases in which the model misclassified the data points. By analyzing the confusion matrix, practitioners can discover common sources of errors and understand the strengths and weaknesses of the classification model; they can also improve the performance of the model through hyperparameter tuning, feature engineering, or the choice of the algorithm. The confusion matrix can also be used to derive a number of metrics that help evaluate the model performance across a set of classes, such as accuracy, precision, recall, and F1-score.

The model is performing exceptionally well as shown in figure 4, according to its metrics, but it is important to confirm that these findings are reliable and not the consequence of overfitting or other problems with the data. To verify that your model is performing as expected, you can use regularization, additional validation, and thorough analysis.

Nearly flawless cross-validation accuracy indicates that the model generalizes to previously untested data with great success. Because cross-validation lowers the variance brought about by a single data split, it is a reliable technique for assessing model performance. With more training samples, training accuracy drops very slightly to 0.9998 from 1.0 at the beginning. This slight decline suggests that when more data is added, the model probably faces a small amount of extra complexity, but it manages this well without suffering a major performance hit.

Overfitting, in which the model performs extraordinarily well on training data but less so on unknown data, is a possible risk given the high training accuracy. At the same time, overfitting may not be a major problem in this case, based on the high cross-validation accuracy. A representative dataset for the real-world situations where the model will be used should be ensured. It's possible that the data lacks enough complexity or variety if both training and cross-validation accuracy are high. If necessary, verify that all aspects of the cross-validation process were carried out accurately, particularly the randomization and stratification. Make sure there was no data leakage during cross-validation, which could have affected the training process by unintentionally letting information from the validation set in.

To improve the balanced efficiency and scalability of the system, it is essential to minimize computational overhead. This overhead can be minimized through various techniques, including algorithm optimization, parallel processing, and resource management, which help smooth operation and better utilization of computational resources. It is defined as the additional resources, be it time, memory, processing power, etc. that a system needs in addition for the core functions. It is a consequence of overhead from added work doing error-checking, resource allocation or abstraction layers. Although these tasks may be necessary for functionality, they can slow down performance and use resources.

Table3: Computational overhead comparison of SLP and MLP

Computational overhead	**Single layer Perceptron**	**Multilayer Perceptron**
Time to complete each epoch	3 minutes	12 minutes
RAM	4 GB	16GB

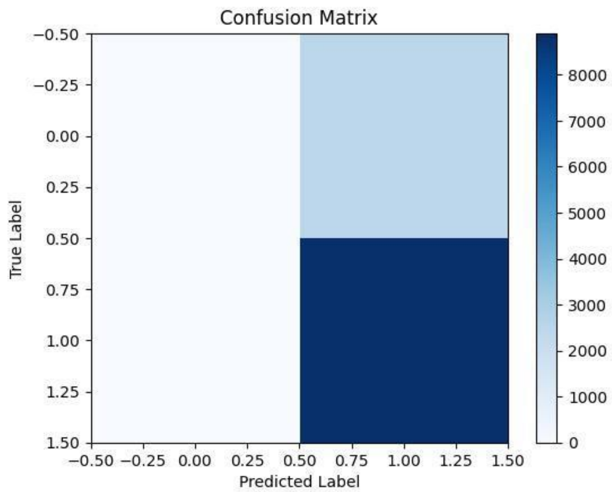

Fig 3: Convolution matrix of the proposed model

Key differences are seen from table 3 regarding computation overhead between Multilayer Perceptron (MLP) and Single Layer Perceptron (SLP). SLP is the one requiring just 3 minutes to complete an epoch, and the usage of RAM is just 4GB compared to the MLP model which took longer, 12 minutes to finish an epoch, requiring a huge 16 GB ram. This difference illustrates the trade-off between computing power and model complexity. Because of its simpler architecture, the SLP has lower processing times and uses fewer resources; however, the MLP is one step ahead in complexity, applying more neurons and more layers resulting in more learning but at the cost of computing power. These metrics illustrate the importance of considering resource requirements in addition to performance when selecting neur

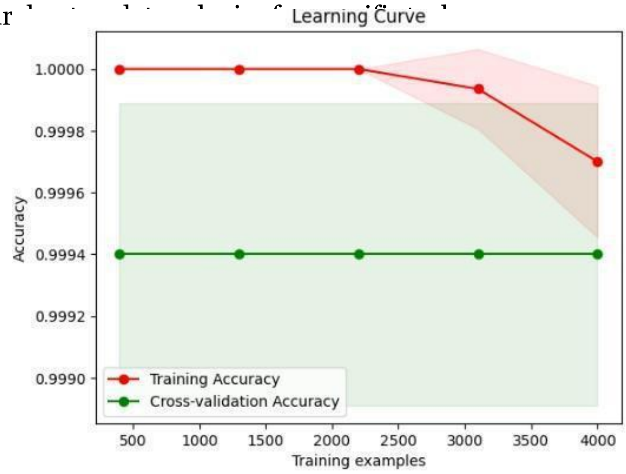

Fig 4: Training examples verses Accuracy curve

DISCUSSION

The investigation and implementation of extreme learning approaches for sentiment analysis in restaurant critiques had yielded promising results and findings. Extreme learning has proven to be efficient in creating robust models for text-based sentiment analysis using single hidden layer neural networks where weights are randomly assigned. The scalability, simplicity, and computing efficiency of extreme learning architectures have enabled academics and practitioners to mitigate several critical disadvantages of traditional supervised learning approaches (eg, their proclivity for overfitting and long training times). Compared to the traditional supervised learning techniques, the comparative study between the two shows that, the extreme learning based models have three significant advantages with regard to the training speed, generalisation performance and scalability. In various accuracy tests, extreme learning models have been accelerative and exhibited lower computational overhead while maintaining competitive accuracy. Moreover, the regularisation impact of random initializations of biases and weights in extreme learning networks has further built the generalisation capacity, thus ensuring robust performance on unseen data and minimization of overfitting likelihood, particularly in environments with a high number of features. Also, the transparency and simplicity underpinning extreme learning architectures have enabled the rapid development, experimentation, and deployment of sentiment analysis models across real-world applications. A broader group of researchers, practitioners, and industry professionals with varying levels of machine learning and natural language processing expertise is now able to leverage extreme learning given its simplification of the training process and reduced reliance on iterative optimisation algorithms. Impending research on extreme learning methods for sentiment classification of meal reviews could focus on enhancing model architectures, exploring alternative activation functions, and investigating interpretability and explainability techniques. Extreme learning algorithms exhibit great scalability and parallelizability, enabling sentiments analysis workflows to be scaled up to larger datasets and to meet the real-time analysis combs' requirements.

Next, the approach uses a multilayer perceptron (MLP) neural-network architecture to carry out extreme learning on restaurant reviews and thus determine whether these reviews are positive or negative. The solution is based on single hidden layer neural network with randomly initialized weights utilizing the most basic form of supervised learning techniques that help in solving daily routine problems, e.g whether a flower is a rose or not; this leads to slow and complex training, and a complex model due to overfitting, therefore, the approach described uses regularized feed forward method (local) without a lot of complex requirements. Combining these preprocessing techniques that clean the data through column selection and deal with any null values with TF-IDF feature generation results in a model that accurately identifies sentiment in meal reviews in this project.

In all pertinent sentiment classification tasks, a comparative evaluation of the MLP model showed several positive results, a type of competitiveness accuracy when written by standard supervised learning model. It also demonstrates lower overfitting sensitivity and faster training times, emphasizing the importance of extreme learning approaches in circumventing common issues found in traditional methods. Other visualizations that give a better idea of how well is the model performing and the model's robustness and generalization capability are ROC curve, precision-recall curve, confusion matrix, etc. These plots are really quite informative because these give the performance of the model across different thresholds.

The results show that extreme learning methods, and in particular the multilayer PLS models with hidden layers, can markedly outperform more traditional supervised learners for the task of sentiment analysis. Extreme learning methodologies elevate sentiment analysis toolsI by improving the efficiency and precision of sentiment analysis models in specific text types, such as restaurant reviews. This paper adds to the knowledge of sentiment classification techniques and suggests useful guidelines for enhancing the quality and efficiency of sentiment classification algorithms in practice. By implementing algorithm used to train and evaluate an MLP model for sentiment analysis of restaurant reviews, one details the performance of extreme learning strategies in this genre. This is an important step towards the application of extreme learning to sentiment analysis techniques. It gives a powerful and efficient method for fetching insights out of textual information, such as meal evaluations. Note that extreme learning can be useful in promoting innovation and improving state of the art in the field of sentiment analysis aligned with regularisation, scalability and simplicity. Clarity that could bring with it a more complete view of consumer sentiment in the digital age.

Training much harder on sentiments allows us to get a great, significant lift in sentiment analysis methods especially on restaurant opinions. As regularization, scalability, and simplicity are among the primary principles of extreme learning, it offers deeper consumer sentiments as it helps achieve more creativity and find better accuracy.

REFRENCES

[1] R. Liu, Y. Shi, C. Ji and M. Jia, "A Survey of Sentiment Analysis Based on Transfer Learning," in IEEE Access, vol.7, pp. 85401-85412, 2019, doi: 10.1109/ACCESS.2019.2925059.

[2] X. Fu, J. Yang, J. Li, M. Fang and H. Wang, "Lexicon-Enhanced LSTM With Attention for General Sentiment Analysis," in IEEE Access, vol. 6, pp. 71884-71891, 2018, doi: 10.1109/ACCESS.2018.2878425.

[3] R. Obiedat et al., "Sentiment Analysis of Customers'Reviews Using a Hybrid Evolutionary SVM-Based Approach in an Imbalanced Data Distribution," in IEEE Access, vol. 10, pp. 22260-22273, 2022, doi: 10.1109/ACCESS.2022.3149482.

[4] A. Feizollah, S. Ainin, N. B. Anuar, N. A. B. Abdullah and M. Hazim, "Halal Products on Twitter: Data Extraction and Sentiment Analysis Using Stack of Deep Learning Algorithms," in IEEE Access, vol. 7, pp. 83354-83362, 2019, doi: 10.1109/ACCESS.2019.2923275

[5] A. Ishtiaq, K. Munir, A. Raza, N. A. Samee, M. M. Jamjoom and Z. Ullah, "Product Helpfulness Detection With Novel Transformer Based BERT Embedding and Class Probability Features," in IEEE Access, vol. 12, pp. 55905-55917, 2024, doi: 10.1109/ACCESS.2024.3390605.

[6] M. H. Abdalla et al., "Sentiment Analysis Based on Hybrid Neural Network Techniques Using Binary Coordinate Ascent Algorithm," in IEEE Access, vol. 11, pp. 134087-134099, 2023, doi: 10.1109/ACCESS.2023.3334980.

[7] S. Lee, H. Shin, I. Choi and J. Kim, "Analyzing the Impact of Components of Yelp.com on Recommender System Performance: Case of Austin," in IEEE Access, vol. 10, pp.128066 128076,2022,doi:10.1109/ACCESS.2022.3225190.

[8] R. Obiedat, D. Al-Darras, E. Alzaghoul and O. Harfoushi, "Arabic Aspect-Based Sentiment Analysis: A Systematic Literature Review," in IEEE Access, vol. 9, pp. 152628-152645, 2021, doi: 10.1109/ACCESS.2021.3127140.

[9] Zheng and L. Zheng, "A Hybrid Bidirectional Recurrent Convolutional Neural Network Attention-Based Model for Text Classification," in IEEE Access, vol. 7, pp. 106673-106685, 2019, doi: 10.1109/ACCESS.2019.2932619.

[10] G. M. Shafiq, T. Hamza, M. F. Alrahmawy and R. El-Deeb, "Enhancing Arabic Aspect-Based Sentiment Analysis Using End-to-End Model," in IEEE Access, vol. 11, pp. 142062-142076, 2023, doi: 10.1109/ACCESS.2023.3342755.

[11] A. Sharma and U. Ghose, "Toward Machine Learning Based Binary Sentiment Classification of Movie Reviews for Resource Restraint Language (RRL)—Hindi," in IEEE Access, vol. 11, pp. 58546-58564, 2023, doi: 10.1109/ACCESS.2023.3283461.

[12] R. S. M. Wahidur, I. Tashdeed, M. Kaur and H. -N. Lee, "Enhancing Zero-Shot Crypto Sentiment With Fine-Tuned Language Model and Prompt Engineering," in IEEE Access, vol. 12, pp. 10146-10159, 2024, doi: 10.1109/ACCESS.2024.3350638.

[13] J. Wang and Z. Chen, "SPCM: A Machine Learning Approach for Sentiment-Based Stock Recommendation System," in IEEE Access, vol. 12, pp. 14116-14129, 2024, doi: 10.1109/ACCESS.2024.3357114.
[14] C. Lu and X. Fu, "SentDep: Pioneering Fusion-Centric Multimodal Sentiment Analysis for Unprecedented Performance and Insights," in IEEE Access, vol. 12, pp. 21277-21286, 2024, doi: 10.1109/ACCESS.2024.3363028.
[15] M. Tavasoli, E. Lee, Y. Mousavi, H. B. Pasandi and A. Fekih, "WIPE: A Novel Web-Based Intelligent Packaging Evaluation via Machine Learning and Association Mining," in IEEE Access, vol. 12, pp. 45936-45947, 2024, doi: 10.1109/ACCESS.2024.3376478.
[16] K. Okoye, S. D. N. Daruich, J. F. E. De La O, R. Castaño, J. Escamilla and S. Hosseini, "A Text Mining and Statistical Approach for Assessment of Pedagogical Impact of Students' Evaluation of Teaching and Learning Outcome in Education," in IEEE Access, vol. 11, pp. 9577-9596, 2023, doi: 10.1109/ACCESS.2023.3239779.
[17] J. Chen, H. Fan and W. Wang, "Syntactic and Semantic Aware Graph Convolutional Network for Aspect-Based Sentiment Analysis," in IEEE Access, vol. 12, pp. 22500-22509, 2024, doi: 10.1109/ACCESS.2024.3364353.
[18] M. Tareq, M. F. Islam, S. Deb, S. Rahman and A. A. Mahmud, "Data-Augmentation for Bangla-English Code-Mixed Sentiment Analysis: Enhancing Cross Linguistic Contextual Understanding," in IEEE Access, vol. 11, pp. 51657-51671, 2023, doi: 10.1109/ACCESS.2023.3277787.
[19] T. Chaisen, P. Charoenkwan, C. G. Kim and P. Thiengburanathum, "A Zero-Shot Interpretable Framework for Sentiment Polarity Extraction," in IEEE Access, vol. 12, pp. 10586-10607, 2024, doi: 10.1109/ACCESS.2023.3322103.
[20] H. D. Jung and B. Jang, "Enhancing Financial Sentiment Analysis Ability of Language Model via Targeted Numerical Change-Related Masking," in IEEE Access, vol. 12, pp. 50809-50820, 2024, doi: 10.1109/ACCESS.2024.3385855.
[21] Y. Yu, D. -T. Dinh, B. -H. Nguyen, F. Yu and V. -N. Huynh, "Mining Insights From Esports Game Reviews With an Aspect-Based Sentiment Analysis Framework," in IEEE Access, vol. 11, pp. 61161-61172, 2023,doi: 10.1109/ACCESS.2023.3285864.
[22] M. Rizinski, H. Peshov, K. Mishev, M. Jovanovik and D. Trajanov, "Sentiment Analysis in Finance: From Transformers Back to eXplainable Lexicons (XLex)," in IEEE Access, vol. 12, pp. 7170-7198, 2024, doi: 10.1109/ACCESS.2024.3349970
[23] A. A. Ahmed et al., "Arabic Text Detection Using Rough Set Theory: Designing a Novel Approach," in IEEE Access,vol. 11, pp. 68428-68438, 2023, doi: 10.1109/ACCESS.2023.3278272.
[24] J. Khan, N. Ahmad, S. Khalid, F. Ali and Y. Lee, "Sentiment and Context-Aware Hybrid DNN With Attention for Text Sentiment Classification," in IEEE Access, vol. 11, pp. 28162-28179, 2023, doi: 10.1109/ACCESS.2023.3259107.
[25] N. K. N. Loh, C. P. Lee, T. S. Ong and K. M. Lim, "MPNet-GRUs: Sentiment Analysis with Masked and Permuted Pre-training for Language Understanding and Gated Recurrent Units," in IEEE Access, doi: 10.1109/ACCESS.2024.3394930.
[26] H. J. Alshahrani, A. Q. A. Hassan, N. S. Almalki, M. M. Alnfiai, A. S. Salama and M. A. Hamza, "Applied Linguistics With Red-Tailed Hawk Optimizer-Based Ensemble Learning Strategy in Natural Language Processing," in IEEE Access, vol. 11, pp. 132448-132456, 2023, doi: 10.1109/ACCESS.2023.3331061.
[27] S. Yang and X. Cai, "An Enhanced Recommendation Model Based on Review Text Graph and Interaction Graph," in IEEE Access, vol. 11, pp. 88234-88244, 2023, doi: 10.1109/ACCESS.2023.3305954.
[28] K. Govardhanan, R. Manikandan, B. Arunkumar, S. P. Abirami, B. M. Chandar and K. Gopalakrishnan, "Parameter Reputation Model for Cloud Service Recommendation and Ranking Using Opinion Mining," in IEEE Access, vol. 12,38123-38134,2024,doi:10.1109/ACCESS.2024.3371857
[29] A. A. Ahmed et al., "Arabic Text Detection Using Rough Set Theory: Designing a Novel Approach," in IEEE Access, vol. 11, pp. 68428-68438, 2023, doi: 10.1109/ACCESS.2023.3278272.
[30] Z. Hu, Y. Zhu and J. Su, "Deep Learning based Fine-grained Sentiment Analysis: A Review," 2021 International Symposium on Artificial Intelligence and its Application on Media (ISAIAM), Xi'an, China, 2021, pp. 170-173, doi: 10.1109/ISAIAM53259.2021.00043.
[31] F. Alzamzami and A. E. Saddik, "Transformer-Based Feature Fusion Approach for Multimodal Visual Sentiment Recognition Using Tweets in the Wild," in IEEE Access, vol. 11, pp. 47070-47079, 2023, doi: 10.1109/ACCESS.2023.3274744
[32] M. Ban, L. Zong, J. Zhou and Z. Xiao, "Multimodal Aspect-Level Sentiment Analysis based on Deep Neural Networks," 2022 8th International Symposium on System Security, Safety, and Reliability (ISSSR), Chongqing, China, 2022, 184-188, doi: 10.1109/ISSSR56778.2022.00039.

[33] S. Yang and X. Cai, "An Enhanced Recommendation Model Based on Review Text Graph and Interaction Graph," in IEEE Access, vol. 11, pp. 88234-88244, 2023, doi: 10.1109/ACCESS.2023.3305954.
[34] N. A. M. Razali et al., "Political Security Threat Prediction Framework Using Hybrid Lexicon-Based Approach and Machine Learning Technique," in IEEE Access, vol. 11, pp. 17151-17164, 2023, doi: 10.1109/ACCESS.2023.3246162.
[35] A. Peivandizadeh et al., "Stock market prediction with transductive long short-term memory and social media sentiment analysis," in IEEE Access, doi: 10.1109/ACCESS.2024.3399548.
[36] Y. Diao, L. Yang, S. Li, Z. Hao, X. Fan and H. Lin, "Detect Sarcasm and Humor Jointly by Neural Multi-Task Learning," in IEEE Access, vol. 12, pp. 38071-38080, 2024, doi: 10.1109/ACCESS.2024.3370858.
[37] A. Sundaram, H. Subramaniam, S. H. A. Hamid and A. M. Nor, "A Systematic Literature Review on Social Media Slang Analytics in Contemporary Discourse," in IEEE Access, vol. 11, pp. 132457-132471, 2023, doi: 10.1109/ACCESS.2023.3334278.
[38] X. Fan, J. Liu, J. Liu, P. Tuerxun, W. Deng and W. Li, "Identifying Hate Speech Through Syntax Dependency Graph Convolution and Sentiment Knowledge Transfer," in IEEE Access, vol. 12, pp. 2730-2741, 2024, doi: 10.1109/ACCESS.2023.3347591.
[39] L. Meng, "The Convolutional Neural Network Text Classification Algorithm in the Information Management of Smart Tourism Based on Internet of Things," in IEEE Access, vol. 12, pp. 3570-3580, 2024, doi: 10.1109/ACCESS.2024.3349386.
[40] A. R. Jafari, G. Li, P. Rajapaksha, R. Farahbakhsh and N. Crespi, "Fine-Grained Emotions Influence on Implicit Hate Speech Detection," in IEEE Access, vol. 11, pp. 105330-105343, 2023, doi: 10.1109/ACCESS.2023.3318863.
[41] Y. Wang, D. Zeng, S. Wada and S. Kurihara, "VideoAdviser: Video Knowledge Distillation for Multimodal Transfer Learning," in IEEE Access, vol. 11, 51229-51240,2023,doi: 10.1109/ACCESS.2023.3280187
[42] X. Zhao, Y. Hu, X. Liu and Y. Wang, "Extraction and Association Analysis of Risk Factors in Gas Pipeline Network Emergencies Through Fusing Expert Knowledge and Pre-Trained Model," in IEEE Access, vol. 12, pp. 65640-65649, 2024, doi: 10.1109/ACCESS.2024.3398042.
[43] M. Zhang, V. Palade, Y. Wang and Z. Ji, "Word Representation With Salient Features," in IEEE Access, vol. 7, pp. 30157-30173, 2019, doi: 10.1109/ACCESS.2019.2892817.
[44] D. Zaikis and I. Vlahavas, "From Pre-Training to Meta-Learning: A Journey in Low-Resource-Language Representation Learning," in IEEE Access, vol. 11, pp. 115951-115967,2023,doi: 10.1109/ACCESS.2023.3326337.
[45] B.Amiri, N. Shahverdi, A. Haddadi and Y. Ghahremani, "Beyond the Trends: Evolution and Future Directions in Music Recommender Systems Research," in IEEE Access, vol. 12, pp. 51500-51522, 2024, doi: 10.1109/ACCESS.2024.3386684.
[46] M. Kim and H. -Y. Yoo, "Identification of Key Service Features for Evaluating the Quality of Metaverse Services: A Text Mining Approach," in IEEE Access, vol. 12, pp. 6719-6728, 2024, doi: 10.1109/ACCESS.2024.3352008.
[47] G. Bo, W. Shanshan, Z. Qing, P. Bo and Z. Yan, "Empowering Medical Data Analysis: An Advanced Deep Fusion Model for Sorting Medicine Document," in IEEE Access, vol. 12, pp. 1650-1659, 2024, doi: 10.1109/ACCESS.2023.3347029.
[48] M. U. Islam and B. M. Chaudhry, "A Framework to Enhance User Experience of Older Adults With Speech-Based Intelligent Personal Assistants," in IEEE Access, vol. 11, pp. 16683-16699, 2023, doi: 10.1109/ACCESS.2022.3230151.
[49] J. Liu, "Enhancing English Language Education Through Big Data Analytics and Generative AI," in Journal of Web Engineering, vol. 23, no. 2, pp. 227-249, March 2024, doi: 10.13052/jwe1540-9589.2322.
[50] M. M. A. Busst, K. S. M. Anbananthen, S. Kannan, J. Krishnan and S. Subbiah, "Ensemble BiLSTM: A Novel Approach for Aspect Extraction From Online Text," in IEEE Access, vol. 12, pp. 3528-3539, 2024, doi: 10.1109/ACCESS.2023.3349203.