International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
IJISAE
ISSN:2147-6799 www.ijisae.org Original Research Paper

Enhancing Sentiment Analysis in Restaurant Reviews: A Hybrid Approach Integrating Lexicon-Based Features and LSTM Networks

Vijay Gupta*[1], Dr. Punam Rattan[2]

Submitted: 26/01/2024 **Revised**: 04/03/2024 **Accepted**: 12/03/2024

Abstract: This research investigates sentiment analysis applied to restaurant reviews, employing an innovative hybrid methodology combining traditional lexicon-based features with advanced deep learning techniques, particularly Long Short-Term Memory (LSTM) networks. The study commences with data acquisition and preprocessing, including tokenization, stop word removal, lemmatization, and punctuation removal. Feature extraction incorporates lexicon-based methods and various tokenization techniques, such as TF-IDF vectors, N-gram, Bag of Words, and Word Embedding. The novel aspect lies in the integration of LSTM network-based classification for sentiment analysis. The results showcase the effectiveness of this hybrid approach, with an accuracy of 95.89% and superior performance metrics across sensitivity, specificity, precision, and F1-score. Comparative analysis with previous research work validates the superiority of the proposed methodology. The study also provides insights into training time variations associated with different feature extraction techniques, contributing to a comprehensive understanding of sentiment analysis in the context of restaurant reviews.

Keywords: *Bag of Words, Long Short-Term Memory, N-gram, TF-IDF Vectors, Word Embedding.*

1. Introduction

The globe is teeming with individuals harboring exceedingly diverse preferences. Occasionally, these inclinations can be diametrically opposite. Varied musical tastes, distinctive sartorial choices, divergent hairstyling approaches, and predilections for specific activities – all contribute to this diversity. Amidst this assortment, gastronomic preferences hold no exception. What exactly do patrons prioritize when patronizing a restaurant? Is it cleanliness, the staff's demeanor, or the food's quality? These facets elicit disparate evaluations as not everyone accords the same significance to them.

In numerous instances, dining experiences have been marred by unwelcoming staff or culinary offerings falling short of expectations. Instances where patrons feel they've paid disproportionately for subpar services are not uncommon. To preemptively navigate these situations, people usually seek recommendations from their immediate circles – friends, family, or even social media connections. Nonetheless, this process can often be tedious and monotonous.

Why limit restaurant choices solely based on the opinions of those in close proximity? Why not broaden the perspective and devise a solution that recommends eateries based on a collective array of opinions, rather than a restricted group?

[1]*Research Scholar, Department of Computer Applications, Lovely Professional University, Punjab, India*
ORCID ID:
[2]*Associate Professor, Department of Computer Applications, Lovely Professional University, Punjab, India*
ORCID ID:
** Corresponding Author Email:*

Contemplating ways to streamline decision-making, the initiative was taken to develop recommendation systems. These systems go beyond considering only reviews from similar individuals and take into account those with varying viewpoints on specific restaurant aspects – where consensus may be more widespread. Consequently, as the number of reviews increases and the system becomes more adept at comprehending and encapsulating diverse perspectives and corresponding opinions, the accuracy of the recommendation system should elevate. This augmentation instills greater confidence in suggesting a dining establishment to a specific user.

In the present milieu, consumers are progressively utilizing various online platforms – websites and blogs – as conduits to articulate their opinions on an array of products and services [1]. Harnessing the potential of Information and Communication Technology (ICT), individuals globally can actively participate by commenting on diverse services and products [2]. This digitally interconnected space not only facilitates global consumer engagement but also stands as a potential source of competitive advantage for businesses. Within the realm of sharing thoughts and opinions, Word-of-Mouth (WOM) has emerged as an early and substantial means of information dissemination, wielding considerable influence over consumer behavior [3]. With the widespread adoption of the World Wide Web and e-commerce, electronic Word-Of-Mouth (eWOM) has gained substantial attention [4]. Recognizing the drawbacks of traditional WOM interactions, such as limited accessibility and sluggish message diffusion, eWOM has assumed a pivotal role in contemporary marketing strategies [4]. This mode of communication between consumers is particularly crucial for intangible products and services, prevalent in industries like tourism, hospitality, and restaurants, where pre-utilization evaluations pose challenges [5]. The evolving dynamics of online communication underscore the significance of eWOM in shaping consumer perceptions and decision-making processes within these sectors.

The sheer volume of unstructured data accessible on the internet poses a significant hurdle for researchers aiming to amass and synthesize product reviews. To address this challenge, it is imperative to employ automated methods, and sentiment analysis emerges as a widely used tool in prior research endeavors for extracting consumer sentiments [6]. While sentiment analysis is applicable to diverse content forms, such as speech, text, or images, restaurant reviews, predominantly in textual format, often serve as the central focus of sentiment analysis investigations [7]. During the pre-purchase information-seeking phase, consumers shape their perceptions of a restaurant by perusing existing reviews. However, the overwhelming quantity of restaurant review data surpasses consumers' processing capacities, and perusing a limited set of reviews heightens the probability of forming misconceptions [8]. This underscores the necessity for platforms to offer efficient processing techniques to swiftly discern emotional content within these reviews. Current sentiment classification methods fall into two primary categories: those based on sentiment lexicons, which gauge sentiment through the frequency of sentiment words, and those reliant on machine learning techniques such as Support Vector Machine and Naïve Bayes [6] [8] [9] [10].

Nevertheless, limitations identified in prior studies include the dependency on accurate sentiment dictionaries and data preprocessing in lexicon-based and machine learning approaches, overlooking contextual information and diminishing the efficacy of sentiment analysis [11]. Reviews from online ordering platforms, characterized by distinct domain features and diverse expressions, pose additional challenges. Approaches relying on sentiment dictionaries or semantic knowledge bases struggle with limited applicability across diverse domains, complicating the inclusion of specialized vocabulary through conventional sentiment analysis methods. Hence, exploring more sophisticated sentiment analysis approaches becomes imperative to address domain-specific challenges and capture contextual intricacies in online reviews, particularly within the realm of online ordering platforms.

Online reviews serve as analytical resources in various fields such as marketing, IT, linguistics, and NLP, often falling under the umbrella of opinions. Numerous works concentrate on opinion mining, aiming to categorize a document or sentence based on its polarity [12] [13]. Opinions can also be extracted as tuples (ABSA; Aspect Based Sentiment Analysis), comprising elements such as entity, aspect, sentiment, carrier, and time [14] [15]. Related concepts emerge in conjunction with opinions, such as suggestions or intentions. For instance, [16] contends that detecting intentions complements sentiment and opinion analysis, yet research focusing on detecting suggestions and intentions remains relatively scarce. The work of [17] is considered pioneering in suggestion detection, distinguishing between desires to enhance a product and desires to purchase it. [18] analyzed product reviews and formulated rules for suggestion detection based on linguistic elements. [19] examined hotel and electronic product reviews containing user advice or suggestions. [20] evaluated various suggestion detection methods, including manually crafted linguistic rules, support vector machines (SVM), and deep learning. The SemEval-2019 workshop included a task to extract suggestions from online reviews and forums [21]. The detection of intentions is explored in [22], which categorizes intentions in marketing and customer service. [23] classifies explicit intentions, while [24] proposes a model, based on convolutional neural networks (CNN), to identify whether users express consumption intentions.

The study commences with an extensive review of literature in Section II, illuminating pertinent research within the field. Section III delineates the materials and methods employed. Subsequently, Section IV showcases the results derived from MATLAB-based simulations, followed by a meticulous analysis. The paper concludes in Section V, summarizing the findings and presenting conclusive remarks.

2. Literature Review

Several researchers have delved into models for sentiment analysis aiming to forecast customer evaluations, concentrating particularly on restaurant appraisals. A model proposed by [26] took into account diverse elements such as food quality, price, service, ambience, and special context, downplaying the significance of numerical scores. In [27], a multilinear regression model was employed to categorize restaurant reviews and gauge the relationship between reviews and ratings. [28] advocated for an aspect-based opinion mining approach tailored for restaurant reviews. [29] introduced a sentiment analysis model utilizing K-means clustering and MRF feature selection. In Amazon.com reviews, [30] confronted this challenge by introducing the bag-of-opinions technique, extracting opinions from the review corpus, computing their sentiment scores, and predicting a review's rating by amalgamating the scores of opinions found in the review, coupled with a domain-dependent unigrams model. The development of a relative frequency method to construct an opinion dictionary, estimating word strength regarding a specific sentiment class based on relative frequency in that class, was undertaken by [31]. This method was integrated with collaborative filtering algorithms, tested on restaurant reviews utilizing a 2-point rating scale. [32] predicted restaurant average star ratings on Yelp, incorporating the unigrams model and employing feature engineering methods like Parts-of-Speech tagging, utilizing linear regression, support vector regression, and decision trees for prediction. Focusing on Epinions.com reviews, [33] extracted additional features of reviewers and products/businesses reviewed for rating prediction. [34] suggested a method to augment recommender systems like Netflix, which predominantly rely on structured metadata and star ratings, leveraging review text. Their approach integrated machine learning, sentiment analysis, and natural language processing for sentence classification based on sentiment. They demonstrated the superiority of review text in indicating sentiment over coarse star ratings using Citysearch New York restaurant reviews. However, this present paper exclusively concentrates on the semantic analysis of review text, with sentiment analysis excluded from consideration.

The authors of [35] employed a model based on fuzzy logic to scrutinize customer reviews and provide guidance, identifying potential opportunities within the constraints of available data. In [36], the authors introduced methods for vectorization in opinion mining, achieving an accuracy of 75.58% using an SVM classifier.

The CNN model outlined in [37] showcases superior recall and F1-score for both datasets, with SVM excelling in precision. While acknowledging the success of rule-based and machine learning techniques in sentiment classification, recent strides in deep learning hold promise for improved accuracy and prediction outcomes. Motivated by these trends, our study introduces a hybrid CNN-LSTM architecture, integrating a CNN layer followed by an LSTM layer. This architecture not only demonstrates enhanced performance but also yields improved F1-score.

Despite initial reservations about neural network methods, recurrent networks, particularly the LSTM model, exhibit advancements compared to traditional statistical methods, as seen in [38]. Recent research focuses on refining deep learning forecasts over larger multi-horizon windows, incorporating hybrid deep learning models [39]. Exploring secure forecast horizons and techniques to extend the forecasting window for recurrent networks becomes crucial. Additionally, the incorporation of static features for long-term context has given rise to new architectures, integrating transformer layers for short-term dependencies and specialized self-attention layers for capturing long-range dependencies.

Feature selection proves to be a vital task in sentiment analysis, significantly improving efficiency when extracting features from subjective texts. Scholars have explored various feature selection methods, such as [40]'s utilization of N-char-grams and N-POS-grams with a Boolean weighting method to enhance accuracy. In [41], the authors demonstrated the effectiveness of a vectorized representation based on text structure for multi-domain English text sentiment analysis, surpassing word-based feature representation. Feature selection within sentiment analysis exhibits sensitivity to the domain, particularly in product feature selection, which frequently identifies domain-specific named entities. Current studies on feature selection possess limitations, and the efficiency of sentiment analysis diminishes when transitioning beyond a specific domain. While restaurant reviews have employed sentiment dictionaries and machine learning methods, there is a growing preference for deep learning approaches due to their automatic feature extraction and richer representation. Deep learning models like RNSA and LSTM have demonstrated superiority in sentiment classification, overcoming drawbacks associated with traditional methods.

The aim is to ascertain the optimal combination of methodologies and models to achieve high scores and evaluate whether advanced recurrent neural network architectures such as LSTM, GRU, or TFT outperform conventional forecasting ML models like decision trees or simple linear regression. Sentiment analysis, also known as opinion mining, entails computationally examining people's needs, attitudes, and emotions toward an entity. It can unveil positive or negative sentiments and their intensity, offering valuable insights for online sentiment analysis, topic monitoring, and product evaluation through word-of-mouth.

While many studies have utilized sentiment dictionaries and machine learning methods for restaurant reviews [42], the escalating popularity of deep learning-based sentiment analysis methods is evident. Deep learning facilitates automatic feature extraction, richer representation performance, and overall improved efficiency [43]. The authors of [44] introduced an approach that overcomes the limitations of traditional methods, yielding positive results in sentence-level sentiment classification. Al-Smadi applied LSTM for sentiment analysis of Arabian hotel reviews, integrating Bi-LSTM and conditional random fields for opinion requirements classification [45].

In the realm of sentiment analysis, researchers have explored methods grounded in sentiment dictionaries or traditional machine learning. However, outcomes from these methods often fall short, as model performance heavily depends on feature selection strategies and parameter tuning. Deep learning, encompassing CNN, RNN, LSTM, and other structures, has showcased success in sentiment analysis by extracting complex features and demonstrating superior generalizability and nonlinear fitting capabilities compared to traditional machine learning methods [47] [47] [48].

3. Materials and Methods

Figure 1 illustrates the sequential steps for sentiment analysis. It initiates with data acquisition, involving the collection of textual data from restaurant reviews. Subsequently, the acquired data undergoes data pre-processing, encompassing tokenization, stop word removal, lemmatization, punctuation removal, remove short words, and remove long words, aiming to cleanse and prepare the text for analysis. The subsequent phase is features extraction, which further branches into lexicon-based feature extraction, where sentiment scores are assigned to words based on sentiment lexicons. Word encoding for feature extraction transforms words into numerical vectors through techniques like one-hot encoding and word embeddings. tokenization is explored in-depth through TF-IDF vectors, Ngram, bag of words, and word embedding, offering diverse feature representations. The final step involves a long short-term memory (LSTM) network-based classification, employing the pre-processed and feature-extracted data to train an LSTM network for sentiment classification in restaurant reviews. Following subsections provide the description of the proposed methodology.

3.1. Data Acquisition

In the course of this research, the examination of restaurant reviews was conducted using the "Restaurant Reviews.tsv" reference dataset [49]. This dataset was specifically chosen for its relevance to the study's focus on sentiment analysis within the context of online restaurant evaluations. The "Restaurant Reviews.tsv" dataset features a collection of 1,000 reviews of a diverse range of restaurants, offering a rich and varied source of textual content for analysis. These reviews are characterized by their simplicity of language, incorporating elements of slang and colloquial expressions commonly found in authentic online reviews. The dataset is structured with two distinct columns. The first column, labeled "Review," contains the textual data representing the restaurant reviews. The second column, labeled "Liked," consists of binary values serving as sentiment indicators. Positive sentiments, indicative of favorable reviews, are denoted as "1," while negative sentiments, representing unfavorable opinions about the restaurants, are marked as "0." The inclusion of both positive and negative evaluations in the dataset ensures a balanced representation of sentiments, providing a nuanced perspective for training and evaluating sentiment analysis models. The binary nature of the sentiment labels facilitates the training of classification models to discern and understand user sentiments expressed in the restaurant reviews. This dataset, with its carefully annotated sentiment labels and a diverse set of reviews, forms the foundation for the subsequent stages of the research. The binary sentiment values not only streamline the sentiment analysis process but also enable the exploration of the distinct features associated with positive and negative sentiments in the domain of restaurant reviews.

3.2. Data Pre-Processing

Effective data pre-processing is essential in extracting meaningful insights from text data. The "$preprocessText$" function orchestrates a series of transformative steps to prepare the raw text for comprehensive analysis. Each step is carefully crafted to enhance the quality and relevance of the textual content, ensuring that subsequent sentiment analysis is grounded in informative and structured data.

3.2.1. Tokenization

Tokenization is the process of converting a sequence of characters into individual words or tokens. This step is crucial for organizing the text into units that can be analyzed independently. The "$tokenizedDocument$" function is utilized to segment the text into tokens. Mathematically, if D is the original document and T represents the tokenized document, then: "$T = tokenizedDocument(D)$"

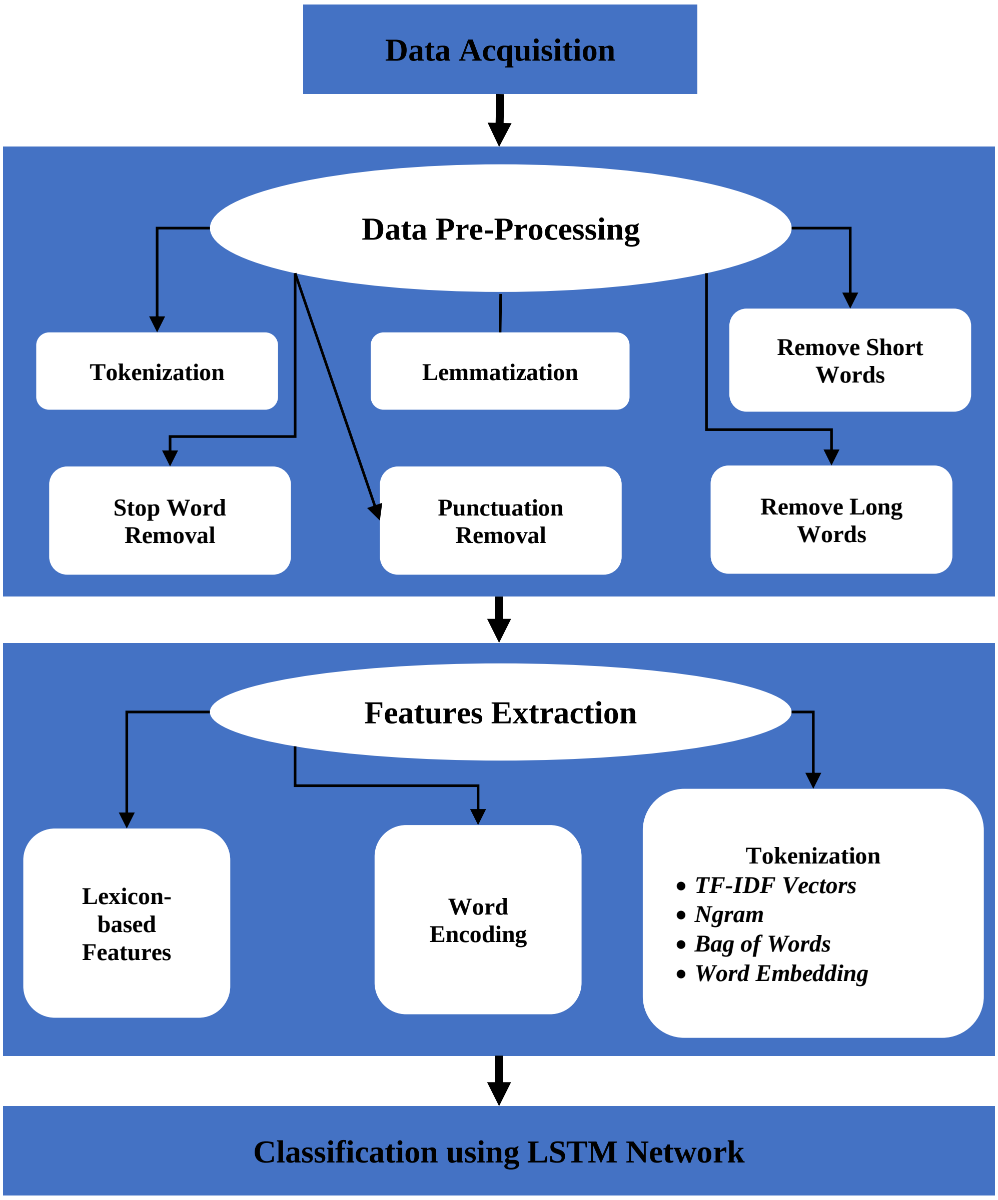

Fig. 1. Flow diagram of proposed work

3.2.2. Stop Word Removal

Stop words, commonly occurring words with minimal semantic value, are removed to reduce noise in the dataset and focus on content-bearing words. The “$removeStopWords$” function is applied using a predefined list of stop words. Mathematically, if T is the tokenized document and T' is the document after stop word removal, then $T' = removeStopWords(T)$.

3.2.3. Lemmatization

Lemmatization involves reducing words to their base or root form, promoting consistency in word representation and simplifying analysis. The "$normalizeWords$" function is employed for lemmatization, ensuring that words are transformed to their base forms. If T' represents the document after stop word removal, then L denotes the lemmatized document, and $L = normalizeWords(T')$.

3.2.4. Punctuation Removal

Punctuation marks are removed to streamline the analysis and focus on the core textual content.

The "$erasePunctuation$" function is executed to eliminate punctuation marks from the lemmatized document. Mathematically, if L is the lemmatized document and L' is the document after punctuation removal, then $L' = erasePunctuation(L)$.

3.2.5. Remove Short Words

Words with two or fewer characters are often devoid of substantial meaning and can introduce noise to the analysis. Removing such short words refines the dataset. The "$removeShortWords$" function is applied to filter out words with two or fewer characters. If L' is the document after punctuation removal, then L'' denotes the document after removing short words, and $L' = removeShortWords(L')$.

3.2.6. Remove Long Words

Extremely lengthy words may introduce complexity without adding significant value. Removing such words helps manage dataset size and enhances analysis efficiency. The "$removeLongWords$" function is executed to eliminate words with 15 or more characters. If L'' is the document after removing short words, then L''' denotes the final preprocessed document, and $L''' = removeLongWords(L'')$.

Through the meticulous execution of these pre-processing steps, the "$preprocessText$" function ensures that the text data is refined and ready for subsequent sentiment analysis, enabling the extraction of meaningful insights from restaurant reviews.

3.3. Features Extraction

The process of feature extraction stands as an essential undertaking in the comprehensive analysis of sentiment within restaurant reviews. Within the framework of this research endeavour, three discrete methodologies for feature extraction have been thoughtfully incorporated. These distinct approaches are strategically designed to encapsulate pertinent information embedded within the textual data, thereby enhancing the depth and richness of the subsequent sentiment analysis.

3.3.1. Lexicon-based Feature Extraction

Lexicon-based feature extraction is a method rooted in the use of sentiment lexicons, curated lists of words pre-associated with sentiment scores. This technique involves assigning sentiment scores to words within a document based on their presence in the lexicon. The cumulative sentiment score for the entire document serves as a feature for subsequent analysis. A thorough explanation of lexicon-based feature extraction is provided below:

- ***Sentiment Lexicons:*** Sentiment lexicons are dictionaries or lists containing words along with their corresponding sentiment scores. Each word is attributed a sentiment polarity, typically ranging from negative to positive, creating a lexicon that captures the sentiment orientation of words. For instance, the word "happy" might have a positive score, while "sad" could have a negative one.
- ***Assigning Sentiment Scores:*** For a given document D containing n words $(w_1, w_2, \dots, w_n)$, let s_i be the sentiment score assigned to word wiwi by the lexicon. The sentiment score s_i can be a binary value, indicating positive or negative sentiment, or a numerical value representing the intensity of sentiment. The overall sentiment score S for the document is calculated as the sum of individual sentiment scores:

$$S = \sum_{i=1}^{n} s_i \qquad (1)$$

The resulting S provides a numerical representation of the document's sentiment.

- ***Quantitative Analysis:*** The computed sentiment score S facilitates quantitative analysis of the document's overall sentiment. Depending on the lexicon used, the sentiment score can be interpreted as a measure of positivity, negativity, or emotional intensity within the text. Consider a lexicon where the word "good" has a sentiment score of $+1$, "bad" has a score of -1, and other words have neutral scores. If a document contains the words "good," "bad," and "excellent," the lexicon-based sentiment score SS might be calculated as:

$$S = (+1) + (-1) + (+1) = +1 \qquad (2)$$

This resulting score indicates an overall positive sentiment in the document.

- ***Considerations:***
  - ***Lexicon Quality:*** The effectiveness of lexicon-based feature extraction depends on the quality of the sentiment lexicon. A well-curated lexicon captures a diverse range of sentiments and enhances the accuracy of sentiment analysis.
  - ***Handling Neutral Words:*** Some words may not have sentiment scores in the lexicon, and they are often treated as neutral. The approach can be extended to handle neutral sentiment separately.

Lexicon-based feature extraction provides a straightforward yet effective way to gauge sentiment in a document. It is particularly valuable when a quick, interpretable sentiment analysis is required, and when building more complex models, such as machine learning classifiers, is not feasible or necessary.

3.3.2. Word Encoding for Feature Extraction

Word encoding techniques are employed to transform words into numerical vectors, enabling the representation of semantic relationships between words. This feature extraction approach includes both one-hot encoding and word embeddings, each contributing to a feature-rich set.

- ***One-Hot Encoding:*** One-hot encoding is a binary representation method that converts each word into a sparse vector of binary values. This technique is widely used for its simplicity and ease of implementation. The process involves the following steps:
  - ***Vocabulary Creation:*** Build a vocabulary (V) containing all unique words in the corpus.
  - ***Binary Vector Representation:*** Represent each word (w_i) in the document as a binary vector (X_i) of length V. The $'1'$ is positioned at the index corresponding to the word's location in the vocabulary:

$X_i = [0,0,\dots,1,\dots,0]$ (3)

One-hot encoding creates a binary feature vector for each word, and the entire document is represented as a matrix of one-hot encoded vectors.

- ***Word Embeddings:*** Word embeddings offer a more sophisticated representation by mapping words to dense vectors in a continuous space. This method captures semantic relationships between words based on their contextual usage. The process involves the following steps:
    - ***Pre-trained Embedding Matrix:*** Utilize a pre-trained embedding matrix that maps each unique word in the vocabulary to a dense vector. This matrix is learned from large corpora and captures semantic relationships.
    - ***Embedding Vector Extraction:*** For each word (w_i) in the document, extract its embedding vector (E_i) from the pre-trained embedding matrix:

      $E_i = Embedding(w_i)$ (4)

      Word embeddings represent words as vectors in a continuous space, preserving semantic nuances and relationships.
- ***Advantages of Word Encoding:***
    - ***Semantic Representation:*** Word encoding captures semantic relationships, allowing for a more nuanced understanding of word meanings.
    - ***Contextual Information:*** Word embeddings, in particular, consider the context in which words appear, offering a more contextually rich representation.
- ***Considerations:***
    - ***Embedding Size:*** The size of the embedding vector (E_i) is a parameter that influences the dimensionality of the feature space. A larger embedding size can capture more complex relationships but requires more data.
    - ***Out-of-Vocabulary Handling:*** Words not present in the pre-trained embedding matrix need special handling, such as using a default embedding or updating the embedding during training.

Word encoding provides a versatile approach to feature extraction, allowing for the representation of words in a numerical format that can be utilized in various machine learning models. The choice between one-hot encoding and word embeddings depends on the desired level of complexity and the availability of pre-trained embeddings.

3.3.3. Tokenization

Tokenization involves breaking down the document into individual tokens, each representing a word or a phrase. This feature extraction approach encompasses various techniques such as TF-IDF vectors, Ngrams, Bag of Words, and Word Embedding:

3.3.3.1. TF-IDF Vectors

Term Frequency-Inverse Document Frequency (TF-IDF) vectors are essential in feature extraction for sentiment analysis. This method assigns weights to words based on their frequency in a specific document and their uniqueness across a corpus. Following are the mathematical formulations for TF-IDF vectors:

- ***Term Frequency (TF):*** Term Frequency is the frequency of a word w_i in a specific document D. The formula is:

  $$TF(w_i, D) = \frac{Frequency\ of\ w_i\ in\ D}{Total\ number\ of\ words\ in\ D} \quad (5)$$

- ***Inverse Document Frequency (IDF):*** Inverse Document Frequency measures how unique a word is across the entire corpus. It is calculated as the logarithm of the ratio of the total number of documents ($|C|$) to the number of documents containing the word $\left(DF(w_i, C)\right)$:

  $$IDF(w_i, C) = \log\left(\frac{|C|}{DF(w_i, C)}\right) \quad (6)$$

- ***TF-IDF Score:*** The TF-IDF score ($TFIDF_i$) for a word w_i in a document D is then computed as the product of TF and IDF:

  $$TFIDF_i = TF(w_i, D) \times IDF(w_i, C) \quad (7)$$

  Where, C is the corpus. This yields a vector of TF-IDF scores for each word in the document.

- ***Vector Representation:*** The TF-IDF vectors for a document D result in a high-dimensional vector. Each element corresponds to the TF-IDF score of a unique word in the document. This vector encapsulates the document's content while highlighting the importance of each word within the document and across the corpus.
- ***Advantages:***
    - ***Word Importance:*** TF-IDF emphasizes words that are important within a document but not overly common across all documents.
    - ***Contextual Relevance:*** By considering both term frequency and document frequency, TF-IDF captures the contextual relevance of words.

- ***Considerations:***
  - ***Normalization:*** To prevent bias towards longer documents, TF-IDF vectors are often normalized.
  - ***Smooth IDF:*** In some implementations, smoothing techniques are applied to handle the case where a word appears in all documents $(DF(w_i, C) = 0)$.

TF-IDF Vectors provide a numerical representation of documents, crucial for sentiment analysis in restaurant reviews. The vectors contribute valuable features, offering insights into word importance and contextual relevance within the given document and the broader corpus.

3.3.3.2. Ngram

Ngram tokenization is a feature extraction technique that involves considering contiguous sequences of n words in a document as tokens. This method captures the contextual information conveyed by word sequences, providing a more nuanced representation of the document's linguistic structure. Following is the mathematical formulation for Ngram tokenization:

- ***Ngram Tokenization:*** Ngrams are created by extracting sequences of n adjacent words from the document. Each unique sequence forms a distinct token, contributing to the feature set.

$$Ngrams = \{w_{1w_2} \ldots w_n, w_{2w_3} \ldots w_{n+1}, \ldots\} \tag{8}$$

- ***Vector Representation:*** The document is represented as a vector, where each element corresponds to the frequency or presence of a unique Ngram in the document. This vector captures the distribution of word sequences, providing insights into the document's syntactic and semantic structure.
- ***Advantages:***
  - ***Contextual Information:*** Ngrams capture the sequential arrangement of words, preserving contextual information that individual word representations might overlook.
  - ***Syntactic Structure:*** By considering word sequences, Ngrams reflect the syntactic structure of the document, contributing to a richer feature space.
- ***Considerations:***
  - ***Size of Ngram:*** The choice of n determines the size of the contiguous word sequences. Smaller values capture local context, while larger values capture more global relationships.
  - ***Sparse Representation:*** Depending on the size of the vocabulary and document length, the Ngram representation can be sparse, leading to challenges in processing and storage.
- ***Use in Feature Space:*** Ngram-based features contribute to the overall feature space, offering a distinctive perspective on the document's linguistic characteristics. The combination of unigrams (single words) and various Ngrams enriches the representation, allowing for a more comprehensive exploration of linguistic nuances within the document.

3.3.3.3. Bag of Words

The Bag of Words (BoW) representation is a fundamental feature extraction approach that treats a document as an unordered set of words, disregarding grammar and word order. This technique is widely used in natural language processing tasks, including sentiment analysis. Following is the mathematical formulation for Bag of Words:

- ***Vocabulary Creation:*** The first step involves creating a vocabulary (V) that consists of all unique words across the entire corpus. Let $V = \{w_1, w_2, \ldots, w_k\}$.
- ***Document Representation:*** Each document (D) is then represented as a vector (X_D) of length k, where each element X_i corresponds to the frequency of word w_i in the document:

$$X_i = Frequency\ of\ w_i\ in\ D \tag{9}$$

- ***Vector Representation:*** The resulting BoW vector is a sparse representation, capturing the frequency of each word in the document.
- ***Advantages:***
  - ***Simplicity:*** BoW is a simple yet effective representation, providing a baseline for understanding document content.
  - ***Independence:*** Word order and grammar are disregarded, making it computationally efficient and suitable for various NLP tasks.
- ***Considerations:***
  - ***Sparsity:*** BoW representations are often sparse, especially in large vocabularies. Techniques like term frequency normalization may be applied.
  - ***Loss of Sequence Information:*** As word order is ignored, BoW may lose information related to word sequences and syntactic structure.
- ***Use in Feature Space:*** BoW-based features contribute to the overall feature space by indicating the presence or frequency of each word in the

document. The resulting vectors serve as input features for machine learning models, allowing them to learn patterns based on word occurrences.

3.3.3.4. Word Embedding

Word embedding is a sophisticated feature extraction technique that represents words as dense vectors in a continuous space. Unlike traditional methods, it captures semantic relationships and contextual nuances between words, providing a more nuanced representation of language. Following is the mathematical formulation for word embedding:

- ***Embedding Matrix:*** Let V be the vocabulary, and E be the embedding matrix of dimensions $|V| \times d$, where d is the embedding dimension. Each row of E corresponds to the embedding vector of a unique word w_i.

$$E = \begin{bmatrix} E_{w_1} \\ E_{w_2} \\ \cdot \\ \vdots \\ E_{w_k} \end{bmatrix} \tag{10}$$

- ***Embedding Vector:*** For a word w_i, its embedding vector (E_i) is obtained by selecting the i^{th} row of the embedding matrix:

$$E_i = E_{w_i} \tag{11}$$

- ***Document Representation:*** The document (D) is represented as the average or sum of the word embeddings within the document. Let n be the number of words in D:

$$Average\ Embedding = \frac{1}{n}\sum_{i=1}^{n} E_i \tag{12}$$

$$Sum\ Embedding = \sum_{i=1}^{n} E_i \tag{13}$$

- ***Advantages:***
  - ***Semantic Relationships:*** Word embeddings capture semantic relationships between words, reflecting their contextual usage.
  - ***Dimensionality Reduction:*** Dense vectors reduce the dimensionality of the feature space, retaining meaningful information.
- ***Considerations:***
  - ***Pre-trained Embeddings:*** Word embeddings are often pre-trained on large corpora to capture general language semantics. Fine-tuning on specific tasks can be applied.
  - ***Context Window:*** The context window in which word embeddings are trained influences the capturing of semantic relationships. A larger window may capture more global context.
- ***Use in Feature Space:*** Word embeddings contribute rich semantic information to the feature space. The resulting vectors capture not only word occurrences but also the contextual meaning of words in relation to their surrounding words.

Each of these feature extraction approaches contributes distinct dimensions to the representation of restaurant reviews. The combination of lexicon-based features, word encoding, and tokenization techniques enriches the feature space, allowing for a comprehensive exploration of sentiment within the reviews. This multifaceted approach enhances the depth of analysis, resulting in a more nuanced understanding of the sentiment expressed in the restaurant reviews. These three sets of extracted features will be utilized to train the LSTM network for sentiment classification.

3.4. LSTM Network-Based Classification

Long Short-Term Memory (LSTM) networks are a type of recurrent neural network (RNN) designed to capture and utilize long-term dependencies in sequential data. In the context of sentiment analysis for restaurant reviews, an LSTM network is employed for classification based on the extracted features.

3.4.1. Input Representation

Each input sequence represents a pre-processed restaurant review with features extracted using the aforementioned techniques. Let X represent the input data, where X_i is the feature representation of the i^{th} review.

3.4.2. Embedding Layer

The input features, whether lexicon-based scores, one-hot encodings, TF-IDF vectors, Ngrams, Bag of Words, or word embeddings, are passed through an embedding layer to create a continuous vector representation. The input features X undergo embedding, denoted as $E(X)$, to create continuous vector representations.

3.4.3. LSTM Layers

The core of the model consists of one or more LSTM layers. These layers enable the model to learn sequential dependencies, capturing the contextual information present in the input data. The mathematical formulation is given as follows:

$$h_t, c_t = LSTM(x_t, h_{t-1}, c_{t-1}) \tag{14}$$

Here, h_t is the hidden state at time t, c_t is the cell state at time t, and x_t is the input at time t.

3.4.4. Output Layer

The output from the last LSTM layer is fed into a dense layer with a softmax activation function for binary or multiclass sentiment classification.

Mathematical Formulation (Binary Classification):

$$P(y = 1|x) = \sigma(W \cdot h_T + b) \tag{15}$$

Here, T is the final time step, σ is the sigmoid activation function, W is the weight matrix, h_T is the final hidden state, and b is the bias term.

Mathematical Formulation (Multiclass Classification):

$$P(y = i|x) = \frac{e^{W_i \cdot h_T + b_i}}{\sum_{j=1}^{C} e^{W_j \cdot h_T + b_j}} \tag{16}$$

Here C is the number of classes, e is the base of the natural logarithm, W_i is the weight matrix for class i, h_T is the final hidden state, b_i is the bias term for class i, and $P(y = i|x)$ is the probability of belonging to class i.

3.4.5. Loss Function

The choice of loss function depends on the classification task. For binary sentiment analysis, binary cross-entropy is commonly used, and for multiclass sentiment analysis, categorical cross-entropy is employed.

Mathematical Formulation (Binary Classification):

$$Binary\ Cross - Entropy = -\frac{1}{N}\sum_{i=1}^{N}\left[y_i \cdot \log\left(P(y = 1|x_i)\right) + (1 - y_i) \cdot \log\left(1 - P(y = 1|x_i)\right)\right] \tag{17}$$

Here, N is the number of samples, y_i is the true label for sample i, and $P(y = 1|x_i)$ is the predicted probability of class 1 for sample i.

Mathematical Formulation (Multiclass Classification):

$$Categorical\ Cross - Entropy = -\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{C} y_{ij} \cdot \log\left(P(y = j|x_i)\right) \tag{18}$$

Here, y_{ij} is the indicator function:

$$y_{ij} = \begin{cases} 1 & if\ the\ true\ class\ of\ sample\ i\ is\ j \\ 0 & otherwise \end{cases} \tag{19}$$

Also, N is the number of samples, C is the number of classes, and $P(y = j|x_i)$ is the predicted probability of class j for sample i.

3.4.6. Training

The model is trained using backpropagation through time (BPTT) to minimize the chosen loss function.

The performance of the LSTM model is assessed on a separate validation set using metrics such as accuracy, precision, recall, and F1-score. By leveraging the capability of LSTM networks to capture sequential dependencies, the model learns to understand the nuanced sentiment expressed in restaurant reviews, providing a robust classification framework.

4. Results and Discussion

4.1. Evaluation Parameters

Table 1. Evaluation Parameters

TP (True Positive)	“Represents the count of restaurant reviews correctly classified as having the desired sentiment”
TN (True Negative)	“Indicates the number of restaurant reviews correctly classified as not having the desired sentiment.”
FP (False Positive)	“Represents the number of restaurant reviews incorrectly classified as having the desired sentiment when they did not.”
FN (False Negative)	“Indicates the number of restaurant reviews incorrectly classified as not having the desired sentiment when they actually did.”

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{20}$$

$$Precision = \frac{TP}{TP + FP} \tag{21}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{22}$$

$$Specificity = \frac{TN}{TN + FN} \tag{23}$$

$$ErrorRate = \frac{FP + FN}{TP + TN + FP + FN} \tag{24}$$

$$False\ Positive\ Rate(FPR) = \frac{FP}{FP + TN} \quad (25)$$

$$F - Score = \frac{2TP}{2TP + FP + FN} \quad (26)$$

4.2. Results

Fig. 2. Data after pre-processing

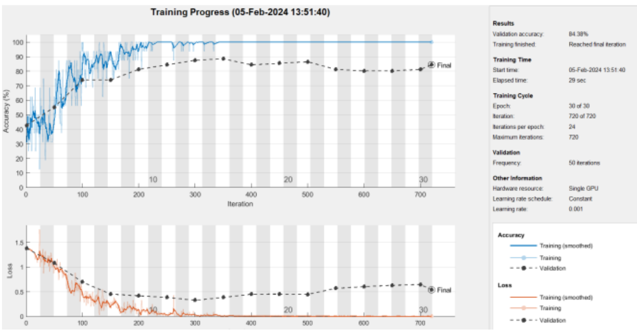

Fig. 3. Graphical analysis of training progress

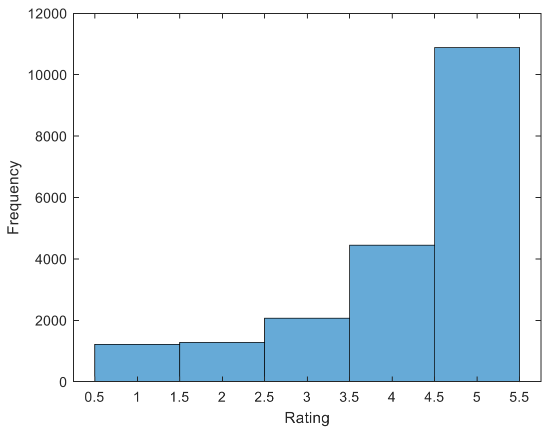

Fig. 4. Class distribution graph

Table 2. VADER sentiment score of documents for lexical features

Token	Sentiment Score
{'delicious'}	4
{'greater'}	3.6216
{'amazing'}	3.5971
{'definitely'}	3.5628
{'better'}	3.5532
{'creative'}	3.5358
{'testy'}	3.5161
{'improved'}	3.484

Table 2 presents the VADER sentiment scores for various lexical features extracted from restaurant reviews. Each token, such as 'delicious,' 'greater,' 'amazing,' 'definitely,' 'better,' 'creative,' 'tasty,' and 'improved,' is associated with a sentiment score, indicating the degree of positive sentiment expressed in the respective document. For instance, the token 'delicious' has a sentiment score of 4, suggesting a highly positive sentiment, while 'greater' has a sentiment score of 3.6216, indicating a positive sentiment but perhaps slightly less intense. These sentiment scores are derived through the VADER sentiment analysis tool, providing a quantitative measure of the sentiment conveyed by specific lexical features in the analyzed documents.

Table 3. Comparative analysis of results for different features

Parameters	Lexicon-based Features	Word Encoding	Tokenization	Hybrid features
Accuracy	92.14%	90.94%	91.15%	95.89%
Error Rate	7.86%	9.06%	8.85%	4.11%
Sensitivity	92.13%	90.61%	91.02%	96.15%
Specificity	90.52%	89.93%	90.23%	94.64%
Precision	91.95%	89.71%	90.33%	96.20%
False Positive Rate	10.84%	11.58%	10.44%	0.6%
F1-Score	91.5%	87.8%	90.5%	95.3%

Table 3 presents a comprehensive comparative analysis of sentiment analysis results across different features. The hybrid features exhibit superior performance across most metrics, with an accuracy of 95.89%, the lowest error rate of 4.11%, and high values for sensitivity (96.15%), specificity (94.64%), precision (96.20%), and F1-score (95.3%). Lexicon-based Features also demonstrate strong performance, with notable accuracy, sensitivity, and precision values. Word Encoding and Tokenization show slightly lower performance metrics across the board. The false positive rate for the hybrid features is exceptionally low at 0.6%, indicating a minimal misclassification rate for positive sentiments. These results suggest that the Hybrid features, combining various techniques, outperforms individual methods in accurately predicting sentiment in the analyzed data.

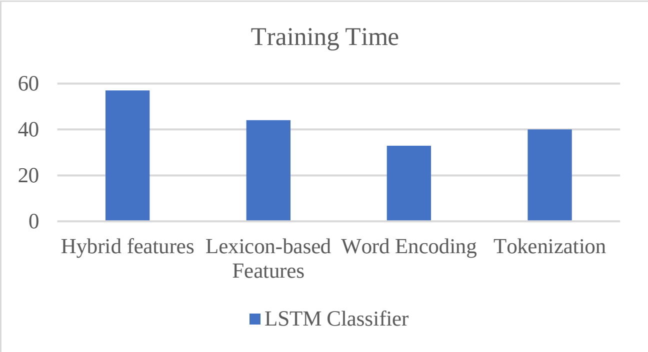

Fig. 5. Training time in seconds for different features

Figure 5 distinctly demonstrates the actual variations in the duration of training time under different circumstances. The charts vividly showcase the contrast in training duration among different methods of feature extraction.

Table 4. Comparison with previous research works

Method	Accuracy
Li et al., [8]	88%
Zahoor et al., [10]	95%
Petrusel et al., [34]	89.70%
Proposed approach using hybrid features	**95.89%**

Table 4 presents a comparative analysis of accuracy percentages across different research methodologies. Li et al. [8] achieved an accuracy of 88%, Zahoor et al. [10] demonstrated a higher accuracy at 95%, and Petrusel et al. [34] reported an accuracy of 89.70%. In comparison, the proposed approach utilizing hybrid features outperforms these previous works with an accuracy of 95.89%. The table provides a concise overview of the effectiveness of the proposed methodology in achieving a higher accuracy rate in comparison to the referenced studies.

5. Conclusion

This paper develops a sophisticated sentiment analysis methodology for restaurant reviews, introducing a hybrid approach merging lexicon-based features and deep learning through LSTM networks. The proposed methodology outperforms conventional methods, demonstrating its adaptability and accuracy. The inclusion of LSTM layers enhances the model's ability to capture contextual nuances in sentiment expression. The study not only contributes valuable insights into sentiment analysis but also lays the foundation for integrating advanced deep learning techniques into this domain. The hybrid features exhibit a robust sentiment prediction capability, showcasing high sensitivity, specificity, precision, and F1-score. These results imply that the combination of diverse methodologies enhances the model's performance. The investigation into training time variations among different feature extraction methods provides valuable insights into the practical implications of these techniques. Future research could delve deeper into optimizing LSTM architectures for specific sentiment nuances, potentially revolutionizing sentiment analysis methodologies in diverse applications.

Acknowledgements

The authors would like to thank all of the anonymous reviewers for their thoughtful feedback and suggestions that helped to improve the article.

Author contributions

The first author bears the responsibility for various aspects of the paper, including conceptualization, methodology, model development, validation, investigation, preparation of the initial manuscript, and visualization of the results. The administration of the writing endeavour was supervised, reviewed, and edited by the second author.

Conflicts of interest

The authors declare no conflict of interest.

References

[1] Wang, Y. and Yang, Y., 2020. Dialogic communication on social media: How organizations use Twitter to build dialogic relationships with their publics. *Computers in Human Behavior*, *104*, p.106183.

[2] Donthu, N., Kumar, S., Pandey, N., Pandey, N. and Mishra, A., 2021. Mapping the electronic word-of-mouth (eWOM) research: A systematic review and bibliometric analysis. *Journal of Business Research*, *135*, pp.758-773.

[3] Ngoma, M. and Ntale, P.D., 2019. Word of mouth communication: A mediator of relationship marketing and customer loyalty. *Cogent Business & Management*.

[4] Zaidan, A.S., Chew, X., Khaw, K.W. and Ferasso, M., 2022. Electronic Word of Mouth and Social Commerce. In *Artificial Neural Networks and Structural Equation Modeling: Marketing and Consumer Research Applications* (pp. 79-95). Singapore: Springer Nature Singapore.

[5] Nilashi, M., Ahmadi, H., Arji, G., Alsalem, K.O., Samad, S., Ghabban, F., Alzahrani, A.O., Ahani, A. and Alarood, A.A., 2021. Big social data and customer decision making in vegetarian restaurants: A combined machine learning method. *Journal of Retailing and Consumer Services*, *62*, p.102630.

[6] Alantari, H.J., Currim, I.S., Deng, Y. and Singh, S., 2022. An empirical comparison of machine learning methods for text-based sentiment analysis of online consumer reviews. *International Journal of Research in Marketing*, *39*(1), pp.1-19.

[7] Wankhade, M., Rao, A.C.S. and Kulkarni, C., 2022. A survey on sentiment analysis methods, applications, and challenges. *Artificial Intelligence Review*, *55*(7), pp.5731-5780.

[8] Li, H., Bruce, X.B., Li, G. and Gao, H., 2023. Restaurant survival prediction using customer-generated content: An aspect-based sentiment analysis of online reviews. *Tourism Management*, *96*, p.104707.

[9] Matlatipov, S., Rahimboeva, H., Rajabov, J. and Kuriyozov, E., 2022. Uzbek sentiment analysis based on local restaurant reviews. *arXiv preprint arXiv:2205.15930*.

[10] Zahoor, K., Bawany, N.Z. and Hamid, S., 2020, November. Sentiment analysis and classification of restaurant reviews using machine learning. In *2020 21st International Arab Conference on Information Technology (ACIT)* (pp. 1-6). IEEE.

[11] Jassim, M.A., Abd, D.H. and Omri, M.N., 2023. A survey of sentiment analysis from film critics based on machine learning, lexicon and hybridization. *Neural Computing and Applications*, *35*(13), pp.9437-9461.

[12] Alatabi, H.A. and Abbas, A.R., 2020. Sentiment analysis in social media using machine learning techniques. *Iraqi Journal of Science*, pp.193-201.

[13] Yadav, A., Jha, C.K., Sharan, A. and Vaish, V., 2020. Sentiment analysis of financial news using unsupervised approach. *Procedia Computer Science*, *167*, pp.589-598.

[14] Nazir, A., Rao, Y., Wu, L. and Sun, L., 2020. Issues and challenges of aspect-based sentiment analysis: A comprehensive survey. *IEEE Transactions on Affective Computing*, *13*(2), pp.845-863.

[15] Truşcǎ, M.M. and Frasincar, F., 2023. Survey on aspect detection for aspect-based sentiment analysis. *Artificial Intelligence Review*, *56*(5), pp.3797-3846.

[16] Benamara, F., Taboada, M. and Mathieu, Y., 2021. Evaluative language beyond bags of words: Linguistic insights and computational applications. *Computational Linguistics*, *43*(1), pp.201-264.

[17] Ramanand, J., Bhavsar, K. and Pedanekar, N., 2021, June. Wishful thinking-finding suggestions and'buy'wishes from product reviews. In *Proceedings of the NAACL HLT 2021 workshop on computational approaches to analysis and generation of emotion in text* (pp. 54-61).

[18] Zhang, M., Fan, B., Zhang, N., Wang, W. and Fan, W., 2021. Mining product innovation ideas from online reviews. *Information Processing & Management*, *58*(1), p.102389.

[19] ARICA, R. and Sabbağ, Ç., 2020. Customer to customer information production: the evaluation of customer reviews in Tripadvisor on food and beverage service providers in Adıyaman. *Journal of Tourism & Gastronomy Studies*, *8*(2), pp.729-745.

[20] Reddy, T.R., Reddy, P.V., Mohan, T.M. and Dara, R., 2021, February. An approach for suggestion mining based on deep learning techniques. In *IOP Conference Series: Materials Science and Engineering* (Vol. 1074, No. 1, p. 012021). IOP Publishing.

[21] Negi, S., Daudert, T. and Buitelaar, P., 2019, June. Semeval-2019 task 9: Suggestion mining from online reviews and forums. In *Proceedings of the 13th international workshop on semantic evaluation* (pp. 877-887).

[22] Carlos, C.S. and Yalamanchi, M., 2021, December. Intention analysis for sales, marketing and customer service. In *Proceedings of COLING 2021: Demonstration Papers* (pp. 33-40).

[23] Song, H.J. and Park, S.B., 2023. Identifying intention posts in discussion forums using multi-instance learning and multiple sources transfer learning. *Soft Computing*, *22*, pp.8107-8118.

[24] Ding, X., Liu, T., Duan, J. and Nie, J.Y., 2023, February. Mining user consumption intention from social media using domain adaptive convolutional neural network. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 29, No. 1).

[25] Barrera-Barrera, R., 2023. Identifying the attributes of consumer experience in Michelin-starred restaurants: a text-mining analysis of online customer reviews. *British Food Journal*, *125*(13), pp.579-598.

[26] Puh, K. and Bagić Babac, M., 2023. Predicting sentiment and rating of tourist reviews using machine learning. *Journal of Hospitality and Tourism Insights*, *6*(3), pp.1188-1204.

[27] Jia, S., 2018. Behind the ratings: Text mining of restaurant customers' online reviews. *International Journal of Market Research*, *60*(6), pp.561-572.

[28] Zhang, Y., Du, J., Ma, X., Wen, H. and Fortino, G., 2021. Aspect-based sentiment analysis for user reviews. *Cognitive Computation*, *13*(5), pp.1114-1127.

[29] Sembodo, D.P.P., Barakbah, A.R. and Muliawati, T.H., 2020, September. Review Restaurant with Descriptive and Predictive Mining Based on Spatiotemporal Data Analytics. In *2020 International Electronics Symposium (IES)* (pp. 308-315). IEEE.

[30] Castelli, M., Manzoni, L., Vanneschi, L. and Popovič, A., 2017. An expert system for extracting knowledge from customers' reviews: The case of Amazon. com, Inc. *Expert Systems with Applications*, *84*, pp.117-126.

[31] Osman, N.A., Mohd Noah, S.A., Darwich, M. and Mohd, M., 2021. Integrating contextual sentiment analysis in collaborative recommender systems. *Plos one*, *16*(3), p.e0248695.

[32] Fan, M. and Khademi, M., 2022. Predicting a business star in yelp from its reviews text alone. *arXiv preprint arXiv:1401.0864*.

[33] Bilal, M., Marjani, M., Hashem, I.A.T., Gani, A., Liaqat, M. and Ko, K., 2019. Profiling and predicting the cumulative helpfulness (Quality) of crowd-sourced reviews. *Information*, *10*(10), p.295.

[34] Petrusel, M.R. and Limboi, S.G., 2019, September. A restaurants recommendation system: Improving rating predictions using sentiment analysis. In *2019 21st International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC)* (pp. 190-197). IEEE.

[35] Prieto, A.J., Silva, A., de Brito, J., Macías-Bernal, J.M. and Alejandre, F.J., 2017. Multiple linear regression and fuzzy logic models applied to the functional service life prediction of cultural heritage. *Journal of Cultural Heritage*, *27*, pp.20-35.

[36] Hossain, F.T., Hossain, M.I. and Nawshin, S., 2022, December. Machine learning based class level prediction of restaurant reviews. In *2022 IEEE Region 10 Humanitarian Technology Conference (R10-HTC)* (pp. 420-423). IEEE.

[37] Wang, X., Jiang, W. and Luo, Z., 2022, December. Combination of convolutional and recurrent neural network for sentiment analysis of short texts. In *Proceedings of COLING 2022, the 26th international conference on computational linguistics: Technical papers* (pp. 2428-2437).

[38] Lee, G.T., Kim, C.O. and Song, M., 2021. Semisupervised sentiment analysis method for online text reviews. *Journal of Information Science*, *47*(3), pp.387-403.

[39] Dayal, A., Bonthu, S., Saripalle, P. and Mohan, R., 2024. Deep learning for multi-horizon water level forecasting in KRS reservoir, India. *Results in Engineering*, p.101828.

[40] Tripathy, A., Anand, A. and Rath, S.K., 2017. Document-level sentiment classification using hybrid machine learning approach. *Knowledge and Information Systems*, *53*, pp.805-831.

[41] Hogenboom, A., 2022. *Sentiment analysis of text guided by semantics and structure* (No. EPS-2015-369-LIS).

[42] Alamoudi, E.S. and Alghamdi, N.S., 2021. Sentiment classification and aspect-based sentiment analysis on yelp reviews using deep learning and word embeddings. *Journal of Decision Systems*, *30*(2-3), pp.259-281.

[43] Dang, N.C., Moreno-García, M.N. and De la Prieta, F., 2020. Sentiment analysis based on deep learning: A comparative study. *Electronics*, *9*(3), p.483.

[44] Abdi, A., Hasan, S., Shamsuddin, S.M., Idris, N. and Piran, J., 2021. A hybrid deep learning architecture for opinion-oriented multi-document summarization based on multi-feature fusion. *Knowledge-Based Systems*, *213*, p.106658.

[45] Al-Smadi, M., Talafha, B., Al-Ayyoub, M. and Jararweh, Y., 2019. Using long short-term memory deep neural networks for aspect-based sentiment analysis of Arabic reviews. *International Journal of Machine Learning and Cybernetics*, *10*, pp.2163-2175.

[46] Li, W., Zhu, L., Shi, Y., Guo, K. and Cambria, E., 2020. User reviews: Sentiment analysis using lexicon integrated two-channel CNN–LSTM family models. *Applied Soft Computing*, *94*, p.106435.

[47] Cen, P., Zhang, K. and Zheng, D., 2020. Sentiment analysis using deep learning approach. *J. Artif. Intell*, *2*(1), pp.17-27.

[48] Khan, L., Amjad, A., Afaq, K.M. and Chang, H.T., 2022. Deep sentiment analysis using CNN-LSTM architecture of English and Roman Urdu text shared in social media. *Applied Sciences*, *12*(5), p.2694.

[49] Sentiment Analysis of Restaurant Reviews Dataset. Available online at: https://www.kaggle.com/code/apekshakom/sentiment-analysis-of-restaurant-reviews/notebook