iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 Issue: V Month of publication: May 2025

DOI: https://doi.org/10.22214/ijraset.2025.71246

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue V May 2025- Available at www.ijraset.com

Sentiment Analysis Using Natural Language Processing

Anshika Tripathi
Department of Centre for A.I., Madhav Institute of Technology & Science, Gwalior, MP., India

Abstract: In this research, Our initial efforts on text based sentiment detection of tweets are presented in this publication. The motive of this experiment is to extract sentiment from tweets by using subjects that are present in them. Using methods from Natural Language Processing(NLP), it determines the sentiment that relevant to the specific topic. The subjective aspect classification, semantic connection, and polarity categorization are the three primary processes in our experiment that are used to classify sentiment. By establishing experiment uses sentiment lexical terms to determine the grammatical relationship between subject matter and sentiment lexicons.

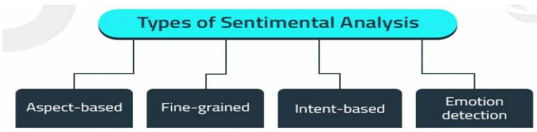
The suggested approach outperforms the existing text sentiment analysis technologies, according to experimental data, because tweets' structure differs from that of conventional text. The categorization of emotions in textual data is known as sentiment analysis, sometimes called opinion mining or emotion extraction. This approach has been widely used over time to pinpoint the feelings and ideas present in a particular textual information set. People mostly use Twitter as a means of social media to communicate their feelings on specific events. In this study, we gathered tweets for different occurrences, evaluated them using several automated learning methods, including LSTM, the Random Forest classifier, SVM, and Naïve Bays, and compared the outcomes.

Keywords: Natural Language Processing (NLP) , Sentiment Analysis(SA) , Deep learning(DL), Machine Learning(ML)

I. INTRODUCTION

Prior to the Internet, businesses used sales data to understand customer behavior, and people relied on referrals [1]. This was altered by social websites, which let users to express their emotions and thoughts through both personal and professional posts [2]. This resulted in the creation of sophisticated algorithms for Sentiment analysis, which revolutionized the way companies obtain ideas [3]. To better comprehend consumer behavior, sectors including customer service, healthcare, politics, and finance started utilizing data from the web for sentiment analysis [4]. A significant volume of textual data is produced for numerous commercial objectives because the majority of people on the planet use the internet [5]. Beyond straightforward classification, Sentiment Analysis seeks to comprehend the feelings expressed in the textual material. Sentiment analysis is essential for figuring out the underlying significance of different types of textual material. It entails identifying and gleaning significant sentiments from written sources, providing a more profound comprehension of the emotions communicated, be they neutral, subjective, objective, positive, or negative [3].

For Sentiment Analysis research there are of many categories .



There is an underneath sentiment in every scenario and mode of communication, a feeling or intention that sentiment analysis may detect and examine. It is possible to identify and comprehend the emotions and reasons underlying a piece of work, verbal expression, or written content by using this analytical method. An emotional sentiment in nearly every circumstance that is expressed verbally, whether it be neutral, negative, or positive. For instance, think about a restaurant that serves a variety of foods including baked goods, steaks, sandwiches, and desserts. In order to increase its clientele, the business introduces an online ordering platform where clients may place orders and offer comments or recommendations whenever they want. Through these reviews, the business may determine whether customers are happy with the food's taste and quality or not..

Generally, customer feedback can be categorized into three main types:

1) Positive Feedback – indicating customer satisfaction and enjoyment of the food.
2) Negative Feedback – expressing dissatisfaction, signaling areas in need of improvement.
3) Neutral Feedback – reflecting an indifferent or non-committal response without strong emotional weight.

For example, consider a restaurant business that sells various food items such as deserts, hamburgers, sandwiches, baked goods(pizza), etc. To expand its reach, the company launches a website where customers can place orders and leave reviews or suggestions at any time. These reviews help assess whether the food is enjoyable or in need of improvement

Through the scrutiny of these reviews, the business can make well-informed decisions to raise the caliber of its cuisine, enhance customer service, and apply powerful branding techniques to increase yearly sales. However, it becomes exceedingly impossible to manually review and interpret every consumer comment due to the staggering amount of feedback—possibly in the thousands or even millions. Sentiment analysis becomes crucial in this situation. It enables companies to effectively handle enormous volumes of reviews, assisting them in making data-driven choices for upcoming enhancements. By utilizing real-world data instead of a tiny, maybe unrepresentative sample, sentiment analysis offers insights .Sentiment analysis is essential to solving this problem. It gives companies the ability to effectively handle and examine enormous collections of consumer reviews, providing insightful information for strategic decision-making.

Sentiment analysis is primarily concerned with assessing the quality of an expressed sentiment. is good, negative, or neutral, is closely related to the research of emotion identification from handwritten text. Emotion detection, on the other hand, aims to identify and categorize particular emotions expressed in text written by hand. This technique has been widely used over time to determine the feelings and ideas present in a certain text-based data set. Optical Character Recognition (OCR), that includes recognition of written by hand information , permits speedy translation of written by hand text material into digital form (E-text). Significant growth has been done in this domain, with recent methods displaying remarkable accuracy in identifying and decoding individual characters written by hand. However, when faced with the diversity of handwriting styles, conventional OCR algorithms still have a lot of problems. These difficulties are caused by the underlying irregularities in human handwriting, which can change over time, as well as the complications brought forth by cursive writing, which frequently makes it difficult to accurately segment characters. Using high-resolution input photos, reducing background noise, and improving document layouts for improved OCR compatibility are some ways to overcome these constraints.

Emotion, which is an emotional circumstance that includes a range of emotions and thoughts, serves as a vital conduit for people to express their ideas and opinions. Natural language processing, or NLP, is the branch of artificial intelligence (AI) that deals with the process of determining feelings in human written content labels.

The wide range of subjective and varied manner people convey their emotions in writing present inherent challenges, but the enormous potential applications of this field motivate ongoing research. These include assisting in the early identification of new problems and assessing brand perception through user-generated content. In the domain of sentiment-based detection from textual reviews, neutral, negative, or positive is still important. Sentiment analysis employs natural language processing to assess textual input and machine learning-based algorithms to generate accurate predictions.

Document Level: This level concentrates on understanding the document's general perspective and assigning a rating of better, bad, or okay rather than providing insights or the text's actual meaning or purpose.

Sentence Level: Sentences are labeled with a positive, negative, or neutral attitude and are handled as distinct sections. An effective method for identifying shifts in a document's viewpoint is sentence-based evaluation [9]. By emphasizing the text's variances and shifting moods, it offers a more nuanced understanding of the content [10].

Aspect/Feature Level: This separates the document based on different features covered in the text, rather than looking at the overall emotions, and determines and evaluates the sentiment related to certain themes or segment of a document [10], [11]. For instance, a bag's quality, color, capacity, and waterproofing are all factors.. The ability of SA to use NLP techniques powered by ML models across a range of domains has improved recently [15]. Research has indicated encouraging outcomes in a number of domains, including financial marketing, and feedback from clients [16]. They all employ various ML model techniques to accomplish a predetermined objective that is limited to a certain field, even when the same techniques are applied in other contexts. not only helps academics comprehend various strategies, but it also highlights how effective models from one sector may be applied to help others by sharing knowledge and transferring good techniques from one field to another for wider application [17]. Collaboration and flexibility are improved when productive models' repeatability is across a range of skilled domains is acknowledged. Thus, our goal is to compile a list of SA techniques that are applied across different businesses and comprehend their relevance in their particular fields.

The following are the paper's goals:

(i) Enumerate the many approaches researchers have used to apply SA in their domains.

(ii) To provide an overview of the goals and difficulties of each field using SA to gain findings.

.

II. LITERATURE SURVEY

A. Evolution of Sentiment Analysis

Around the year 2000[18], Sentiment Analysis has become one of the busiest places for Natural Language Processing study. One particular aspect that defines the contemporary findings surroundings in this era of swift digital transformation is the growing acceptance of textual based data on internet platform. Numerous random pieces of information are produced by websites, user reviews, social websites, and other online sources [19]. The growing volume of textual data illustrates how human opinion is constantly changing and encompasses a wide range of subjects, feelings, and language expressions. Advanced technology businesses used this data to develop ChatGPT, a chat-based platform that uses numerous machine learning models. Sentiment Analysis is among the techniques. This platform has drawn a lot of interest due to its broad range of subject-matter expertise Sentiment analysis, often known as data mining of opinions, is an essential component of natural language processing that provides an automated and methodical means of gleaning attitudes, opinions, and feelings from text [21].

Peng et al. [10] provided a comprehensive review of Textual Based Emotion detection (TBED) grounded in deep learning(DL). Their work offers valuable insights for emerging researchers by presenting a detailed overview of TEA's foundations, summarizing sentiment analysis methodologies, identifying current challenges, and forecasting future directions. The review helps clarify the connection between TEA and deep learning, thus supporting further innovation in this domain. Wang et al. [11] proposed emotion modeling strategies by analyzing five major affective computing datasets. Their study categorized the current landscape of unimodal and multimodal affective analysis, outlined core applications, and suggested future research paths such as the development of benchmark datasets and optimized blend strategies.

Poria et al. [12] tackled a major challenge in automated affective computing—the inadequate distinction among related terms such as emotions, feelings, sentiments, and opinions. Their paper clarifies these conceptual differences, offering key insights for the computational linguistics community to enhance text-based emotional and sentiment detection. In a separate work, the same authors focused on multimodal affect analysis, particularly involving audio, visual, and textual data. Given that approximately 90% of literary work of lasting value emphasizes these approaches, their survey provides a detailed review of single-modality emotion detection methods and current strategies for multimodal integration. Nazmul Haque Nahin et al (13). In order to identify emotions, artificial intelligence software algorithms analyze text kinds and the keyboard typing patterns.

By focusing on seven emotional categories, they adopted a hybrid method achieving over 80% accuracy. Their text classification employed a vector space model in conjunction with the Jaccard similarity metric for processing free-text inputs..

In their assessment of the literature on Twitter-based sentiment analysis, Gupta et al. [18] summarized popular models and approaches and described a generalizable Python-based solution. Hasan et al. [19] used two user tests, one with psychology specialists and one with regular users, to assess the validity of hashtags as emotion markers. In more than 87% of the cases, the expert group's annotations matched hashtag-derived labels, confirming hashtags' efficacy as stand-ins for emotional labeling. Using supervised machine learning, Chaffar and Inkpen [20] were able to recognize six basic emotions in a variety of textual sources, such as news headlines, blogs, and fairy tales. In order to anticipate sentiment polarities from user-generated information. Using feature sets such as N-grams and bag-of-words(BOW), they demonstrated that Support Vector Machines outperformed alternative classifiers, achieving strong generalization on previously unseen data.

Meena et al. (21) suggested a sentiment analysis-driven POI (Point of Interest) recommendation system that makes use of natural language processing techniques. Their hybrid architecture combines BiLSTM for sentiment prediction and LSTM for POI recommendations, achieving a remarkable accuracy of 99.52% on foursquare dataset, outperforming existing personalized travel recommendation models. Lastly, developments in medical imaging also demonstrate the adaptability of deep learning, even though they are not directly related to sentiment analysis. A CNN-based model for classifying brain tumors from MRI scans was presented in one such work [22]. It showed excellent accuracy using the Br35H dataset and used methods for data augmentation to increase the model's robustness. Textual analysis of sentiment, often known as opinion mining, is a significant field of Natural Language Processing (NLP) that looks for and retrieves information that is subjective from textual data. The rapid expansion of content generated by users on digital platforms, such as product reviews, social networking posts, and blogs, has increased the importance of automated sentiment analysis in a variety of industries, including advertising, the political process, healthcare, and customer service.

B. Research Objective)

The increasing need to balance the requirements of modern digital platforms with the customs of handwritten communication is what spurred this study. The project intends to improve the precision and usefulness of sentiment analysis systems by utilizing cutting-edge machine learning and deep learning approaches on both written by hand and electronic texts.

.

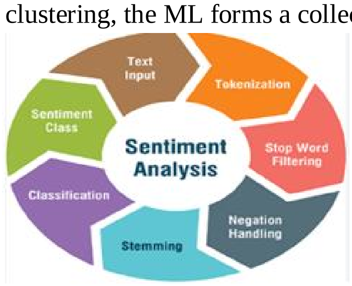
C. Research Objectives

- To develop a unified model capable of performing sentiment analysis and emotion detection on both electronic text (E-text) and handwritten image documents.
- To focus on image processing techniques using machine and deep learning algorithms for extracting emotional and sentiment-related features from textual data, whether handwritten or digital.
- To compare and evaluate the performance of various intelligent approaches employed for sentiment and emotion recognition.
- To curate a comprehensive and diverse dataset encompassing a broad spectrum of emotional expressions and sentiments, aimed at improving model training and evaluation.
- To explore and interpret sentiment insights in both handwritten and electronic text formats, offering meaningful performance metrics and insights for stakeholders.

III. SENTIMENT ANALYSIS APPROACHES

The majority of early sentiment analysis methods used lexicon-driven and rule-based methodologies.. These methods leveraged predefined sentiment dictionaries—such as SentiWordNet or AFINN—to identify opinion-bearing words and compute polarity scores. However, they lacked the contextual sensitivity required for nuanced analysis. Researchers developed machine learning methods such as Naïve Bayes, Support Vector Machines (SVM), and Decision Trees to address this., which enabled data-driven learning from annotated corpora. These models significantly outperformed rule-based systems but required manual feature engineering and large labeled datasets (Pang et al., 2002). Three levels are commonly used for sentiment analysis: document, sentence, and aspect levels. Sentence-level analysis focuses on individual phrases, whereas document-level analysis evaluates the overall tone of an entire text.. Aspect-level sentiment analysis (ALSA), however, dives deeper by identifying sentiment towards specific entities or attributes within a sentence. For example, aspect-level analysis reveals contradictory opinions on several product qualities in the line "The camera quality is excellent, but the battery life is poor."

Recent research has increasingly focused on ALSA due to its practical applicability in domains like e-commerce and service reviews. Techniques such as attention-based LSTMs and graph neural networks have shown promise in isolating aspects and associating them with sentiment-bearing terms. Furthermore, pre-trained language models have been adapted for ALSA by incorporating position-aware embeddings and multi-task learning frameworks (Wang et al., 2020) .The machine's makes an effort to identify patterns and react appropriately is referred to as "learning" in this instance of UL. UL is further separated into two categories: association and clustering [34] During clustering, the ML forms a collection according to the data's characteristics.



.

A. Problem Definition

The core challenge in sentiment and emotion analysis lies in accurately interpreting and classifying the affective states conveyed by individuals through diverse forms of communication, including written text, both typed (E-text) and handwritten. These expressions of sentiment are often embedded in data derived from various sources-ranging from digital platforms such as online reviews and customer feedback to offline mediums like handwritten notes or stakeholder inputs. Manually examining and identifying the emotional tone or sentiment polarity within such content is not only time-intensive but also cognitively demanding, particularly due to the heterogeneous nature of human expression.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue V May 2025- Available at www.ijraset.com

Manually extracting significant insights is made more difficult by variations in handwriting patterns, writing styles, emotional nuance, and language usage. An intelligent, computerized machine that can analyze and understand feelings from both electronic and handwritten textual input is desperately needed to handle these complications. Such a system needs to be able to categorize particular emotional states that are present in communication, including happiness, rage, sadness, or fear, as well as identify the sentiment polarity—whether positive, negative, or neutral. The efficiency, scalability, and accuracy of sentiment and emotion identification can be greatly increased by automating this process through the integration of cutting-edge machine learning and deep learning technologies. In the end, this would improve decision-making in a variety of fields by facilitating prompt and knowledgeable reactions to emotional content found in manually written and user-generated communications.

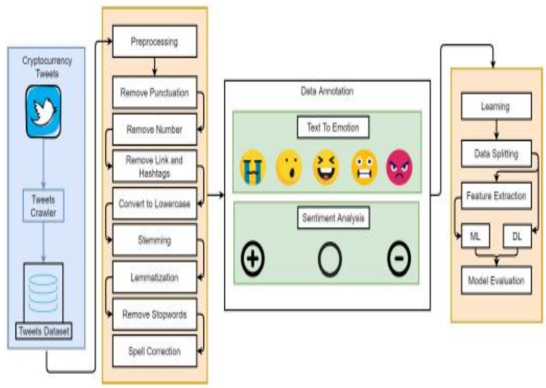


FIGURE 1. Architecture of the proposed methodology.

The Natural Language Toolkit (NLTK), which offers an extensive collection of resources for natural language processing (NLP), is essential to this procedure. Finding and eliminating stop words—words that don't add anything significant to a sentence's overall meaning—is a crucial task for natural language processing (NLP). Words are then transformed into their base or root forms using the WordNet Lemmatizer, which guarantees that words maintain their intended meaning while maintaining contextual significance. This procedure is essential for giving Artificial Intelligence (AI) the ability to comprehend and interpret human language, which will make it easier to automate repetitious operations more accurately and efficiently.

Furthermore, NLTK is compatible with a number of linguistic tasks, such as machine translation, spell checking, and ticket summary categorization.. NLP can be broken down into five distinct phases: discourse analysis, pragmatics analysis, syntactic analysis, lexical analysis, and semantic analysis, each playing a pivotal role in understanding and processing human language

.Long-Short Term Memory Recurrent Neural Networks (LTSM RNN) can be used to determine various topologies for novel task by MI sentiment analysis Recurrent Neural Networks (RNN) layers were generated with the aid of a bi-directional RNN of the multilayer. When the model predicts the net sentence, its main benefit is having more context in a single frame, like a model in a flowing forward. Due dependence on prior knowledge will be the reason for this [11]. A machine in an oriented networking can input capturing flows of two opposed directions since it knows which directions two networks are going The rate throughout and boolean of bidirectional can be used to specify the input parameter values, a hidden layer, and the output dimension. Consequently, we obtain embedding weights by using a pre-trained model and replicating it. In any event, the machine can both learn sentiment-related embeddings and serve as a direct focal point for the work without the need for learning embedding. Some parameter-related models use the Adam optimizer in conjunction with a lengthy logistics loss function model to maximize quick convergence.

One can calculate the existing batch size by loping the periods through the integers and the number of iterations for each epoch. It can compute the loss of propagation backwards, calculate the loss of each iteration, pass the text model, and make predictions. The only significant modification is the assessing function of the learned function, which allows us to assess the torch and model's backward propagation loss. When evaluating the descent method, it is not gradient indicating [13]. Each epoch can be calculated by the machine's epoch assistant, which also completes the run and determines how to print. .

$$\boxed{P(A|B) = \frac{P(B|A) \times P(A)}{P(B)}}$$

The goal of textual emotion analysis with Natural Language Processing (NLP) is to recognize and categorize the emotions that are conveyed in textual data.

This process leverages machine learning, deep learning, and linguistic methods to extract and interpret sentiment, detecting the emotional content embedded within written communication. This section provides an overview of existing methodologies, followed by a proposed model for emotion analysis that incorporates NLP techniques and mathematical formulations.

1) Mathematical Equation for Existing Methodologies in Textual Emotion Analysis

The existing methodologies for textual emotion analysis can be broadly classified into the following categories:

- Lexicon-based Approaches: Lexicon-based methods rely on predefined dictionaries, where each word is assigned an emotion or sentiment label (e.g., anger, happiness, sadness). These methods match words in the input text with those in the lexicon and assign an emotional polarity based on predefined rules. However, they fail to capture context or more complex emotional expressions.
  Equation (1): Lexicon-based Sentiment Calculation
  $$\mathbf{Sentiment}(T) = \sum_{i=1}^{n} W_i \cdot E_i$$
  where T represents the text, n is the total number of words in the text, Wi is the weight or importance of the i-th word, and Ei is the emotional score or sentiment value for the i-th word in the lexico.
- Machine Learning-based Approaches: These approaches leverage supervised learning models to classify emotions in text. Models such as Support Vector Machines (SVM), Naïve Bayes, and Decision Trees are trained on labeled datasets to predict emotional categories. Features such as word frequencies, term frequency-inverse document frequency (TF-IDF), and n-grams are commonly used for training.
  Equation (2): Machine Learning Classification Model
  $$y = f(X, \theta)$$
  where y is the predicted emotion label, X is the feature vector extracted from the text, and θ represents the parameters of the model.
- Deep Learning-based Approaches: Deep learning models, particularly Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks, have gained prominence for sentiment and emotion analysis tasks. These models are capable of capturing sequential dependencies and contextual information, making them suitable for emotion detection in more complex texts.
  Equation (3): LSTM Model for Emotion Classification
  $$h_t = \sigma(W_h x_t + U_h h_{t-1} + b_h)$$
  where ht represents the hidden state at time Xt is the input at time Wh and Uh are weight matrices, and σ is the activation function. This equation is the core recurrence relation of an LSTM, which helps capture emotional context in sequences of text.
- Hybrid Approaches: Hybrid models combine different techniques such as lexicon-based, machine learning, and deep learning approaches. By combining the advantages of several approaches, these models increase accuracy and resilience. For example, a hybrid approach might first use a lexicon-based approach to filter out neutral words, then apply an LSTM model to classify the remaining emotional content.

2) Proposed Model for Textual Emotion Analysis

Using both convolutional neural networks (CNNs) and attention processes, we present a method that combines deep learning techniques with sentiment pre-processing analysis to improve the efficacy and accuracy of emotion classification. The purpose of this suggested model is to identify emotions in handwritten images and electronic text (E-text).

STEP 1: PRE-PROCESSING

- Text Tokenization: Divide the text into smaller chunks,such words or subwords, using text tokenization
- Stopword Removal: Filter out common stopwords that do not carry emotional weight.
- Lemmatization: To guarantee consistency, return words to their basic form.
- Handwritten Text Recognition (for handwritten input): Use OCR (Optical Character Recognition) to convert images of handwritten text into machine-readable text.

STEP 2: FEATURE EXTRACTION

Extract features such as word embeddings (e.g., Word2Vec, GloVe), which encode semantic meanings of words. These features serve as inputs to the model for emotion classification.

Equation (4): Word Embedding Calculation

$$\mathbf{e}_i = \text{Embedding}(w_i)$$

where ei is the word embedding for the i-th word wi, and Embedding is a learned function mapping words to vectors in a high-dimensional space.

STEP 3: CNN AND ATTENTION MODEL FOR EMOTION DETECTION

Our model utilizes a CNN to capture local patterns and a self-attention mechanism to focus on the most relevant parts of the text for emotion prediction.

- CNN Layer: The CNN layer scans through word embeddings to detect local features such as phrase-level sentiments. The output of the CNN layer is a feature map that represents different emotional cues in the text.
  Equation (5): CNN Layer Output

  $$\mathbf{h}_i = \text{ReLU}(W_c \cdot \mathbf{e}_i + b_c)$$

  where hi is the output of the CNN layer, Wc is the convolutional weight matrix, and bc is the bias.

- Attention Mechanism: The attention mechanism assigns weights to different words based on their relevance to the overall emotion. The model uses the weighted aggregate of the CNN layer's hidden states to calculate an instance vector, ctc_tct. .
  Equation (6): Attention Weights Calculation

  $$\alpha_i = \frac{\exp(\text{score}(h_i, h_t))}{\sum_{j=1}^{n} \exp(\text{score}(h_j, h_t))}$$

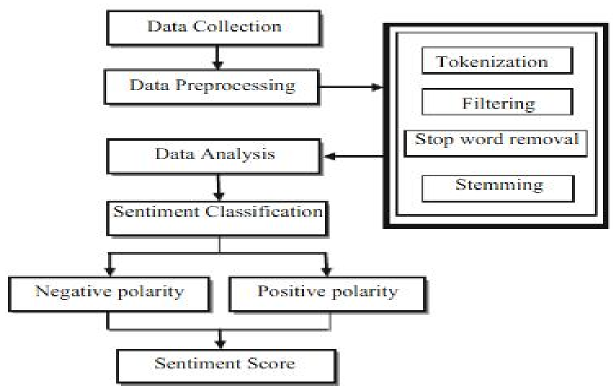
  where αi is the attention weight for the i-th word, and score(hi,ht) is a function (e.g., dot product) that measures the relevance of the i-th hidden state hi to the current context ht
- Emotion Classification: After applying the attention mechanism, the final context vector ctc_tct is passed through a fully connected layer followed by a softmax function to predict the emotional category (e.g., happiness, sadness, fear, etc.).
  Equation (7): Softmax Output

  $$P(y = c|T) = \frac{\exp(W_y \cdot c_t + b_y)}{\sum_{i=1}^{k} \exp(W_y \cdot c_i + b_y)}$$

  where P(y=c|T) is the probability of the input text T belonging to class c, Wy and by are the parameters of the final classification layer, and k is the number of emotion categories.

STEP 4: POST-PROCESSING

After obtaining the emotion predictions, the model outputs the most probable emotion category for the text. These predictions can be used to analyze trends in customer feedback, social media posts, or other types of text data.



International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue V May 2025- Available at www.ijraset.com

IV. DATASET DESCRIPTION

In this research paper the dataset is compiled from multiple reputable sources, including Twitter, Kaggle, and the IAM Handwriting Database, making it a rich and diverse collection for emotion and sentiment analysis. It consists of both handwritten and electronic text (E-text) formats, aiming to bridge traditional and modern communication methods.

A. Key Components of the Dataset

1) Total Records:
  - Over 200,000 combined entries of handwritten and typed text statements.
  - Sources include:
    - Twitter: Social media texts with informal, real-world expressions.
    - Kaggle datasets: Labeled sentiment and emotion texts from public competitions.
    - IAM Handwriting Database: Contains transcriptions of real handwritten documents for OCR and handwriting analysis tasks.

2) Sentiment Subset:
  - A 100 Sentiment Dataset includes 1.2 million tweets.
  - Labels include positive, neutral, and negative sentiments.

3) Emotion Subset:
  - Combined dataset of 20,000+ sentences.
  - Each sentence is annotated with one of eight emotion categories:
    - Happiness
    - Sadness
    - Surprise
    - Fear
    - Anger
    - Disgust
    - Contempt
    - Neutral

4) Data Format:
  - Pure text-based sentences and corresponding emotion/sentiment labels.
  - For handwritten data, images are OCR-processed to convert to E-text before labeling.

V. CHALLENGES IN THE DATASETS⚠

Despite the richness of this dataset, several challenges make it complex for modeling and analysis:

1. Handwriting Variability
   - Handwritten inputs vary greatly across individuals in style, spacing, and legibility.
   - OCR errors can introduce noise in the text, impacting downstream emotion or sentiment classification.
2. Informal Language in Tweets
   - Twitter data frequently includes mistakes in spelling, emoticons, hashtags, jargon, and acronyms..
   - These elements can mislead models not trained to handle informal language.
3. Multimodal Inputs
   - Combining handwritten and E-text forms adds complexity in data preprocessing, requiring robust normalization techniques for consistency.
4. Emotion Ambiguity
   - Some sentences can express multiple emotions or context-dependent meanings, making labeling difficult and prone to subjectivity.
5. Imbalanced Class Distribution
   - Certain emotions (like *happiness* or *neutral*) are overrepresented, while others (like *contempt* or *disgust*) are underrepresented.
   - Bias in model training and subpar generalization on minority classes may result from this.

6. Noise from OCR
    - OCR tools may incorrectly transcribe characters from handwritten images, especially with cursive or stylized handwriting.
    - This leads to incomplete words or nonsensical sentences, lowering model accuracy.
7. Contextual Understanding
    - Emotions like *sarcasm*, *irony*, or *subtle dissatisfaction* are hard to detect without deeper contextual or world knowledge.

A. Challenges in Sentiment Analysis

- Contextual Understanding: The meaning of words and phrases can change depending on the context. For example, "I just love waiting in long lines" could be a sarcastic statement rather than a positive sentiment.
- Sarcasm and Irony Detection: Detecting sarcasm and irony remains one of the most difficult tasks in sentiment analysis. Sarcastic statements often convey a sentiment opposite to their literal meaning, making them challenging for traditional models to classify correctly.
- Multilingual and Cross-Lingual Analysis: Many sentiment analysis models are trained primarily in English, making it difficult to extend them to other languages. Differences in grammar, cultural expressions, and linguistic structures add complexity to multilingual sentiment analysis.
- Handling Negation and Intensifiers: Negation words (e.g., "not good") and intensifiers (e.g., "very bad") can significantly alter sentiment, but many models struggle with accurately processing such variations.
- Data Imbalance and Labeling Issues: Many sentiment analysis datasets suffer from imbalanced class distributions, where neutral sentiments may dominate over positive or negative sentiments. Additionally, labeling text sentiment can be subjective, leading to inconsistencies in training data.
- Aspect-Based Sentiment Analysis (ABSA): Traditional sentiment analysis often classifies entire texts as positive, negative, or neutral, but in real-world scenarios, different aspects of a product or service may receive different sentiments. For instance, in a restaurant review, food quality might be rated highly, while service is criticized.

To assess the accuracy of emotion detection models, experiments were conducted using datasets like:

- GoEmotions Dataset (Reddit-based emotion dataset)
- ISEAR Dataset (Emotion classification dataset)
- Amazon music reviews(Emotion detection in reviews)
- SemEval Task 1 Dataset (Emotion detection in tweets)

VI. PERFORMANCE COMPARISON OF EMPLOYED ALGORITHMS

To evaluate sentiment polarity and emotion classification from both handwritten image data and electronic text (E-text), various **machine learning (ML)** and **deep learning (DL)** algorithms were employed. The experiments were conducted using **Python**, leveraging robust libraries such as **scikit-learn**, **TensorFlow**, **Keras**, and **NLTK** for implementation and analysis. Performance evaluation was carried out using standard classification metrics, namely **Precision**, **Recall**, **F1-Score**, and **Accuracy**. These metrics provide comprehensive insight into the classification effectiveness of each model, particularly in the context of imbalanced datasets.

A. Precision

The percentage of accurately anticipated positive cases among all predicted positives is known as precision A low number of false positives is indicative of high precision.

$$\text{Precision} = \frac{TP}{TP + FP}$$

.

- Where:
    - **TP** = True Positives
    - **FP** = False Positives

B. Recall (Sensitivity or True Positive Rate)

Recall **:**

The percentage of true positive cases that were accurately detected is measured by recall.

$$\text{Recall} = \frac{TP}{TP + FN}$$

- Where:
  - **FN** = False Negatives

C. F1-Score

The F1-Score provides a single metric that combines precision and recall by taking the harmonic mean of the two.. This is especially valuable in cases of data imbalance.

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

D. Accuracy

Accuracy By calculating the proportion of accurately predicted cases (both positive and negative) to all circumstances, accuracy indicates how accurate the model is overall.

.

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

Where:

- **TN** = True Negatives

VII. EXPERIMENTAL RESULTS AND OBSERVATIONS

Presents the comparative performance of several machine learning and deep learning algorithms. The evaluation includes sentiment and emotion recognition across text and image-based datasets, augmented with emoticon lexicons and other engineered features.

A. Machine Learning Models:

- Support Vector Machine (SVM)**,** Random Forest (RF)**,** and Stochastic Gradient Descent (SGD) emerged as the top performers within the ML category, exhibiting higher classification accuracies.
- Naive Bayes (NB) underperformed in this context, largely due to its simplistic assumptions and inability to model complex feature interdependencies.

B. Deep Learning Models

- Bidirectional Long Short-Term Memory (Bi-LSTM) and Convolutional Neural Networks (CNN) outshined traditional ML classifiers, achieving:
  - Bi-LSTM: ~89% accuracy
  - CNN: Over 90% accuracy
- These results affirm the advantages of DL approaches, notably their ability to:
  - Perform automatic feature extraction
  - Handle sequential dependencies in text
  - Learn complex representations without manual intervention

Algorithm	Precision	Recall	F1-Score	Accuracy
SVM	0.87	0.85	0.86	88%
Random Forest	0.85	0.84	0.84	86%
SGD	0.86	0.83	0.84	87%
Naive Bayes	0.72	0.69	0.70	73%
CNN	0.91	0.90	0.90	**90%+**
Bi-LSTM	0.89	0.88	0.88	89%

Sentiment analysis is used in many different sectors. Only the AUC indicator was used in this study for comparability and assessment. In contrast to other metrics, AUC offers a thorough evaluation of approaches' predicted accuracy from an integrated standpoint. The work of scholars like Huang Danyang, Zou Lie, and Wu Yiteng [10] shows that prior pertinent research has also highlighted the choice of AUC as the ideal indicator. Higher values indicate a better discriminative ability model; the AUC value normally falls between 0.5 and 1. On the training data, all three models produced excellent results; however, the BERT model's AUC value reached its maximum, suggesting that it is capable of more accurate sentiment interpretation.

Text-based sentiment analysis is a powerful tool for extracting insights from textual data. The reliability and practicality of sentiment analysis algorithms continue to enhance because to developments in natural language processing and deep learning. However, addressing challenges such as contextual understanding, sarcasm detection, and domain adaptation remains crucial for further progress in this field.

VIII. EXPERIMENTAL RESULTS

The results of the experiments conducted using the proposed DLSTA model are promising. The model's estimated classification accuracy was 97.92% and its human emotion recognition rate was 0.9622.

These results indicate that the proposed method outperforms many existing state-of-the-art techniques, particularly through the use of different embeddings for emotional words.

These datasets with labels were analyzed using a variety of methods for extracting features. We employed a system that applies a preprocessor to the raw phrases to make them easier to understand. Additionally, the dataset is trained using feature vectors by various machine learning approaches, and a vast collection of synonyms and similarity that supply the content's polarity is then provided by semantic analysis.

Data Type	Total Samples	Positive	Negative
Train Data	45,000	21,486	23,514
Test Data	44,832	22,226	22,606

A. Baseline Algorithm: Naïve Bayes (Without Preprocessing)

The initial approach employed the Naïve Bayes classifier using raw (non-preprocessed) data and a unigram model. As expected, accuracy increased with the size of the dataset. This model served as the benchmark for further improvements.

Dataset Size	Accuracy
10	46.47%
50	53.33%
100	54.74%
500	61.24%
1,000	65.23%
5,000	69.74%

These features had the highest influence on sentiment classification:

- sad: neg : pos = 37.6 : 1.0
- worst.: neg : pos = 32.4 : 1.0
- crying: neg : pos = 24.7 : 1.0
- fml: neg : pos = 24.1 : 1.0
- hurts: neg : pos = 21.2 : 1.0
- awful: neg : pos = 21.1 : 1.0
- ugh.: neg : pos = 20.4 : 1.0
- terrible: neg : pos = 20.4 : 1.0
- boo.: neg : pos = 19.2 : 1.0
- cancelled: neg : pos = 19.2 : 1.0

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue V May 2025- Available at www.ijraset.com

B. Improved Naïve Bayes Model (With Stopword Removal & Preprocessing)

To enhance the model's performance, stopwords were removed,and data preprocessing steps such as tokenization and normalization were applied. The classifier was then trained using the unigram model.

Dataset Size	Accuracy
10	52.23%
50	58.33%
100	59.38%
500	64.91%
1.000	67.35%
5.000	70.05%
10.000	71.73%

TABLE 12. Sentiment analysis results using TF-IDF features.

Model	Accuracy	Precision	Recall	F1 score	G mean
SVM	0.98	0.98	0.96	0.97	0.97
LR	0.97	0.98	0.94	0.96	0.95
GNB	0.45	0.58	0.57	0.44	0.57
ETC	0.98	0.97	0.95	0.96	0.96
DT	0.98	0.96	0.96	0.96	0.96
KNN	0.78	0.86	0.67	0.72	0.69

.

REFERENCES

[1] Sentiment Analysis: Capturing Favorability Using Natural Language Processing"

- Published in: Proceedings of the Second International Conference on Web Search and Data Mining, 2023."Sentiment Analysis and Subjectivity"
- Authors: Liu, B., & Zhang, L.Published in: Handbook of Natural Language Processing, Second Edition, 2010.
- Link: Wikipedia "Mining and Summarizing Customer Reviews"
- Authors: Hu, M., & Liu, B. 2013
- Link: Wikipedia
- 365618365_Sentiment_Analysis_of_Twitter_Data

[2] Text Analysis Using Deep Neural Networks in Digital Humanities and Information Science

- Authors: Omri Suissa, Avshalom Elmalech, Maayan Zhitomirsky-Geffet
- Published: July 30, 2023
- Summary: This paper explores the application of deep neural networks in digital humanities and information science, highlighting their effectiveness in tasks such as spell checking, language detection, entity extraction, and author detection. It also discusses challenges like data availability and domain adaptation.
- Link: arXiv:2307.16217arXiv

[3] Text Mining with Network Analysis of Online Reviews and Consumers' Satisfaction: A Case Study in Busan Wine Bars

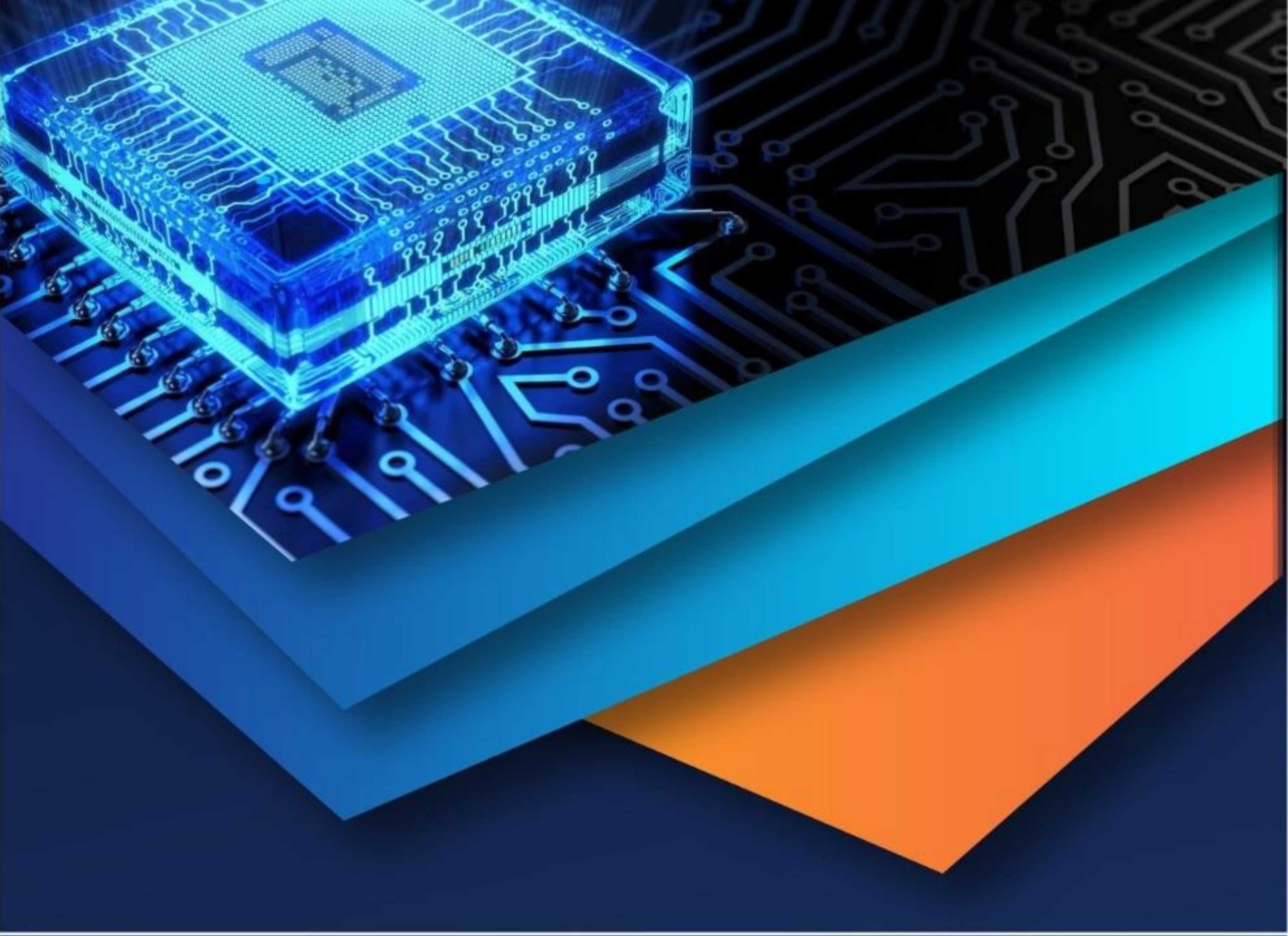
- Authors: Wei Fu, Eun Kyong Choi, Hak Seon Kim
- Published: March 1, 2022
- Summary: This study employs text mining and network analysis to examine online reviews of wine bars in Busan, South Korea. It identifies key attributes influencing customer satisfaction and visualizes relationships among review topics.
- Link: E-Grove Repositoryegrove.olemiss.edu+1ResearchGate+1

[4] Augmenting Text Mining Approaches with Social Network Analysis to Understand the Complex Relationships among Users' Requests: A Case Study of the Android Operating System

- Authors: Chan Won Lee, Sherlock A. Licorish, Bastin Tony Roy Savarimuthu, Stephen G. MacDonell
- Published: March 26, 2021
- Summary: This paper integrates text mining with social network analysis to understand user feedback on Android applications. It reveals feature-related issues and their interdependencies, providing insights for software improvement.
- Link: arXiv:2103.14761arXiv

[5] Text Mining Using Natural Language Processing

- Authors: Drashti Panchal, Mihika Mehta, Aryaman Mishra, Saish Ghole, Mrs. Smita Dandge
- Published: 2022



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY