SSRG International Journal of Electronics and Communication Engineering
ISSN: 2348-8549/ https://doi.org/10.14445/23488549/IJECE-V13I2P103

Volume 13 Issue 2, 30-39, February 2026

Original Article

Punjabi Text Sentiment Analysis using Conjunction, Disjunction, and Negation Words

Surbhi Sekhri[1], Khushboo Bansal[2]

[1,2]*Department of Computer Science and Engineering, Desh Bhagat University, Mandi Gobindgarh, Punjab, India*

[1]*Corresponding Author : surbhilamba49@gmail.com*

Received: 15 November 2025 Revised: 17 December 2025 Accepted: 18 January 2026 Published: 20 February 2026

Abstract *- In modern times, the digital world has expanded in such a way that people can express opinions and thoughts with great ease. Earlier, English was the sole language to communicate over the internet. However, in today's scenario, users on the internet have more personal choices to share content in vernacular languages. This leads to an increase in the usage of Punjabi content on the web for blogs, sharing opinions, providing feedback, and recommendations. Hence, Sentiment Analysis in Punjabi text is one of the growing study areas that aims to understand and clarify emotions and intentions into positive, negative, or neutral polarity. This paper comprehensively examines the evolving techniques, detailing foundational processes like data gathering and pre-processing. Also, a dataset in the Punjabi language was compiled and pre-processed, followed by the development of a lexicon for conjunctive, disjunctive, and negation words. The paper also elaborates on challenges connected with sentiment analysis in Punjabi, such as a lack of digital resources like wordnets and labelled corpora, morphological variations, and the problem of code-mixing with English.*

Keywords *- Conjunction, Disjunction, Machine Learning, Negation, NLP, Sentiment analysis*.

1. Introduction

The goal of sentiment analysis, a prominent field of study in Natural Language Processing (NLP), is to recognize and categorize the attitudes, feelings, and opinions expressed in textual data. Sentiment analysis is essential to decision-making processes in fields including business intelligence, healthcare, politics, and social research due to the explosive expansion of social media platforms, online reviews, and user-generated material. Low-resource regional languages have received less attention than high-resource languages like English, which have been the subject of much research.

Despite being one of the most commonly spoken languages, especially in the Indian subcontinent, Punjabi still lacks sophisticated linguistic and computational tools. Sentiment analysis in Punjabi is difficult due to its complex morphology, script differences, lack of standardized datasets, and scarcity of annotated corpora. Because of language and structural differences, sentiment analysis methods created for English cannot be used for Punjabi.

While some research has tried sentiment categorization for regional Indian languages, the majority of these studies either concentrate on Hindi or use Conventional Machine Learning Techniques with low accuracy. Punjabi sentiment analysis is the subject of very few studies, and those that do frequently use tiny datasets, simple feature extraction methods, or traditional classifiers, which leads to less-than-ideal results. Furthermore, the research mostly lacks a comprehensive comparison of various models and their efficacy on Punjabi textual data.

The current study focuses on creating an efficient sentiment analysis framework for Punjabi literature in order to overcome these constraints. In order to increase classification accuracy, the proposed work will evaluate Punjabi textual data utilizing sophisticated pre-processing methods and appropriate Machine Learning Algorithms. This paper strengthens NLP research for low-resource languages by addressing language-specific issues and comparing performance to current approaches.

The following research gaps have been found based on the literature review:

- Annotated Punjabi sentiment datasets are scarce.
- Absence of research using sophisticated or optimal ML methods for Punjabi text
- Inadequate comparison of various sentiment categorization methods for the Punjabi language
- Studies on Punjabi sentiment analysis have found lower accuracy scores.

These deficiencies demonstrate the necessity of a reliable and effective sentiment analysis method designed especially for the Punjabi language.

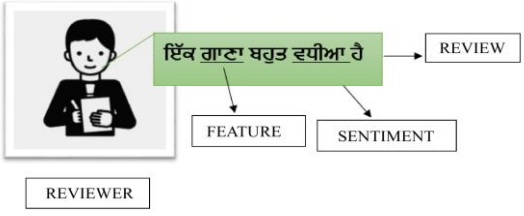

Fig. 1 Sentiment analysis components

The following contributions are what make the current work novel:

- Creation of a sentiment analysis model tailored to Punjabi textual data
- Utilizing efficient pre-processing and feature extraction methods appropriate for the structure of the Punjabi language
- Comparative analysis of sentiment classification algorithms' performance to determine the best strategy Evidence of increased performance and accuracy in comparison to previous Punjabi sentiment analysis research
- This study improves the efficacy and dependability of sentiment categorization for Punjabi language text by offering a structured and comparative framework, in contrast to previous works that rely on conventional or constrained methodologies.

In the light of this context, sentiment analysis is used by marketers and various organizations to get information regarding brand reputation, customers' preferences, product development, market research, and improving customer services by fetching feedback from sources like social media, surveys, and blogs. This is the reason why sentiment analysis with artificial intelligence has gained more attention over the last 10 years or more from scholars for studying not only the English language but other languages as well. Figure 1 shows various components of a sentiment given by a reviewer.

Sentiment Analysis can be conducted at three distinct stages: Aspect level, Sentence level, and Document level. Aspect-level sentiment analysis assesses customer reviews on one feature of an entity to determine sentiments. Also, it gives a detailed breakdown of what a user likes or dislikes about different parts of products or services. Sentence Level Analysis analyzes a paragraph from a product review and assigns a feeling to each sentence. This level is closely related to subjectivity classification. Document Level SA helps in briefing the sentiment of an article or a user review by considering the communal judgement stated across all sentences in the text. It helps to determine the overall satisfaction level from a single review of a product or service. Figure 2 provides an example of different levels of sentiment analysis in Punjabi content.

Subjectivity classification, which entails identifying the clause within the sentence, is done in the first steps of the Punjabi text analysis process. The assertive text is then classified as either positive or negative. To determine if a text contains a perspective (subjectivity) or not (objectivity), subjectivity classification is helpful. A "subjective sentence" in Punjabi refers to a statement based on personal opinion or feeling, like "ਇਹ ਗੀਤ ਬਹੁਤ ਵਧੀਆ ਹੈ।" (This film is boring).

An "Objective Sentence" in Punjabi refers to writing that presents information in an unbiased, factual, and undistorted manner, like "ਫਿਲਮ ਦਾ ਚੱਲਣ ਦਾ ਸਮਾਂ 180 ਮਿੰਟ ਸੀ |" (The film duration is 180 minutes.) Sometimes the text does not directly contain positive or negative sentiment. However, the whole sentence is evaluated for evaluating the emotional pitch of the sentence, like ਮੈਂ ਇੱਕ ਪੀਜ਼ਾ ਖਰੀਦਿਆ ਪਰ ਇਸਦਾ ਸੁਆਦ ਚੰਗਾ ਨਹੀਂ ਹੈ। (I have bought one pizza, but it does not taste well.)

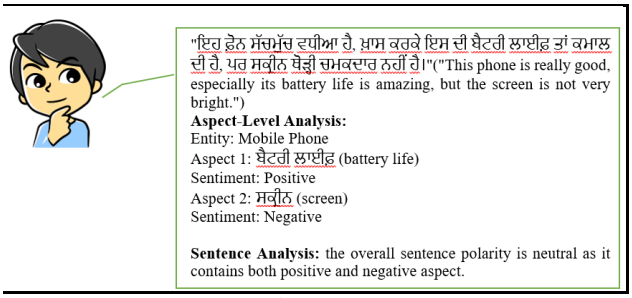

Fig. 2 Example of sentiment analysis levels

2. Punjabi Language & Digital Platforms

Punjabi is one of the world's ancient languages. Approximately 70% of Punjab's population has internet access, and nearly all of them utilize social networking sites ranging from texting a personal message to a friend to commenting in a community group, preferring to use them in Punjabi. Hence, digital data in the Punjabi language is also increasing exponentially.

2.1. Digital Engagement in the Punjabi Context

2.1.1. Punjabi Speakers Worldwide

Punjabi is regarded as the first language of approximately 150 million people [1]. Punjabi is the 10th most spoken language in the world [2].

2.1.2. Wikipedia & Online Content

Punjabi Wikipedia attracts several million page views monthly, with peak traffic of about 7.2 million views in June 2023. Engagement is robust in the United States, Canada, and India [3].

2.1.3. Social Media Usage

In Punjab (India), it was reported that ~70% of the population uses the internet, and among internet users, the majority prefer social media in the Punjabi language [2].

2.2. Localization and Punjabi Language

In the twenty-first century, "localization" is proving to be a revolutionary concept. This entails translating a language into another that even a six-year-old can comprehend. Although "localization" is beneficial for language users, it is now required for many digitally oriented sectors. Since every industry has embraced the digital age, almost everything may now be purchased or sold and paid for from the comfort of your own home. Numerous applications about payments and shopping have been created. Several businesses are therefore attempting to "localize" these applications in an effort to build a more intimate relationship with the users. Google is a well-known brand in this industry. The company is localizing a lot of its items. Ordinary merchants can now advertise their goods and services in the local language, and customers can purchase them in their native tongue. A big thanks to

"localization." Online payment platforms like Paytm and BHIM, which are available in Punjabi, are also used as payment-related apps. People who only speak the local language, Punjabi, can now easily use smartphones thanks to Android phones' ability to select their native language. As a result, we can state that excellent work on "localization" is being done in Punjabi.

2.3. Punjabi Language and The Realm

Neither the Indian Punjab nor the Pakistani Punjab is the only place where Punjabi is spoken. This language has acquired a distinct identity globally. Approximately 140 million people worldwide speak Punjabi. In Canada, Punjabi has emerged as a second language. In the same vein, Punjabi is widely spoken in England, the United States, Australia, New Zealand, Spain, Italy, Germany, and Arab nations. We may observe that members of the Punjabi community are being elected to many national parliaments today, which is causing the Punjabi language to advance daily. The Punjabi language contains the resources needed for a wide range of topics, including science, technology, and medicine, and much effort is being made in this area.

3. Pipeline of Sentiment Analysis

The pre-processing of text, feature extraction, and classification are some of the crucial stages in analyzing sentiments. The basic text input is cleaned up in the pre-processing step by eliminating superfluous information such as stop words, symbols, and numerals.

The text is then converted into features using techniques such as Term Frequency-Inverse Document Frequency (TF-IDF), Global Vectors (GloVe), fastText, and Word2vec. Using either more Advanced Deep Learning Models like Recurrent Neural Networks (RNN) and Long Short-Term Memory (LSTM) or more traditional Machine Learning Techniques like Logistic Regression, Naïve Bayes (NB), and Support Vector Machines (SVM), the processed text is categorized into different sentiment categories during the feature extraction stage.

The pipeline [4] of sentiment analysis includes:

3.1. Text Input

A passage of text is used as the starting point for sentiment analysis. A sentence, paragraph, page, or whole collection of textual material could be an example of this.

3.2. Text Pre-Processing

The text frequently goes through pre-processing procedures before analysis in order to eliminate superfluous information, fix typos, and tokenize the text into distinct words or phrases. This stage helps prepare the text for further examination.

3.3. Feature Extraction

In order to represent the information pertinent to sentiment, sentiment analysis algorithms extract features from the pre-processed text. Individual words, word combinations (n-grams), or more complex representations like word embedding can all be considered features.

3.4. Sentiment Categorisation

This stage splits the text into predetermined sentiment classes, such as positive, negative, or neutral, with the help of ML algorithms. This task is typically performed using supervised learning models, in which the algorithm is trained on labelled data to identify patterns related to each sentiment class.

3.5. Output

The classification or score representing the sentiment indicated in the input text is the sentiment analysis's ultimate output. It sheds light on the dominant viewpoints, attitudes, or feelings that the text expresses.

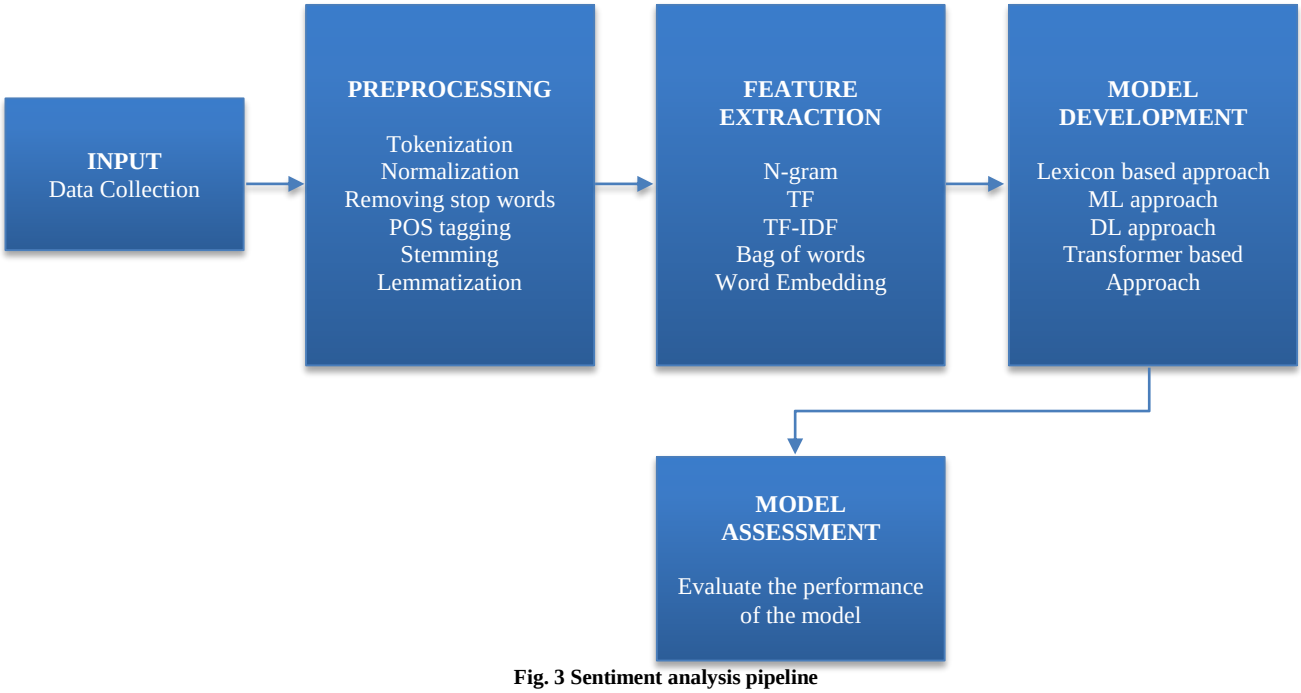

Fig. 3 Sentiment analysis pipeline

4. Literature Review

Punjabi language sentiment analysis is a novel research field. There is a vast amount of data available for processing in various languages, but very little research is being la done on Punjabi contexts. The study papers in the following literature review offer comprehensive details that could aid in a more profound comprehension of current systems.

According to [5], a person's personality can be characterised as a set of requirements that demand a propensity on their behaviour; this tendency is constant over time and situations. It is important since it refers to a person's priorities in a properly established manner. The method that aids in characterising people is sentiment analysis. In this study, we used personality keywords to extract the personality from user-submitted reviews about any individual written in Hindi.

The study in [6] is an early example of opinion mining for an Indian language and the first known effort for the Hindi language. The authors employ three methods: i.) building a Hindi classifier model, ii.) training a model on an annotated English corpus, and iii.) translating Hindi documents into English. iv.) Making use of "SentiWordNet," a majority-based classifier for Hindi. The findings indicate that using an annotated corpus of Hindi yields efficient outputs than the other two methods, which is 78.14%.

The consequences of negative emotion detection in SA in movie reviews were investigated by Dadvar et al. [7]. When negative terms, for instance, do not appear in the sentences, their polarity is investigated. To determine how it might affect the polarity identification of the sentences, several negation scopes that influence classification accuracy are explored. The findings demonstrate that using different window sizes does not significantly alter categorization accuracy.

The Sentiment Analyzer (SA) is a tool designed to extract sentiment about a specific topic from various online text sources, as outlined by Jeonghee Yi et al. [8]. It employs NLP techniques to identify sentiment related to the subject in each reference, rather than categorizing the sentiment of the respective document concerning that topic. The SA relies on two key language resources for its analysis: a sentiment lexicon and a sentiment pattern database. Its algorithms were tested using online product reviews, as well as a range of broader documents such as news articles and general websites.

Today, everyone on social media is using more than one language to convey their message for their ease; likewise, English-Punjabi mixed code language, according to Mukhtiar Singh et al [9]. So, there is a need to analyse this data appropriately to make the online communication better; they used the trigram and five-gram approach to predict the performance of the system. Trigram shows 83% accuracy to predict the mix-code data, while five-gram has 82% accuracy.

In [10], a bootstrap method was suggested, wherein a sentiment lexicon for Hindi without an annotated corpus was introduced. The suggested method minimized the incorrect polarity terms by extracting sentiment words from the Hindi "WordNet." A Hindi Sentiment Lexicon-based multi-module sentiment analysis system was applied to the film and product sectors. The average accuracy, according to the results, was over 85%.

The authors in [11] concluded that sentiment analysis is a powerful technique that uncovers and interprets subjective information from diverse sources. By leveraging computational linguistics, text analysis, and natural language processing, it reveals valuable insights that can drive informed decisions and improve public sentiment's understanding. Better products have been created, user opinions have been understood, and business decisions have been implemented and managed thanks to sentiment analysis. It is a method for finding consumer reviews of any product. Positive or negative feedback might be found in people's reviews. Sentiment analysis is a field that is constantly developing with several applications. Its primary goal is to enable computers to recognise and produce human-like emotions. The increase in User-Generated Content (UGC) in Hindi on websites, in business, academia, and other contexts, has made it possible to effectively study and mine the data. They put out an algorithm that takes the review out of a review sentence and also determines the subject of the opinion. Additionally, it utilises a database dictionary to determine the review's sentiment score.

Table 1. Comparative analysis of various algorithms

Author	Description	Algorithm Used	Dataset	Evaluation
Kaur and Sahni [12]	Punjabi poetry is classified into four categories.	Ten models were trained and evaluated for poetry classification.	240 Punjabi poems were manually obtained from various websites.	58.79% accuracy with SVM and 50.63% accuracy with Hyperpipes
Hentschel and Pal [13]	Online crowdsourcing of Punjabi language text, audio, and video data	A website is proposed that contains crowdsourced Punjabi language content	Eight Punjabi speakers from various fields were interviewed to provide the text and videos in the initial archive.	

Kaur and Sharma [14]	An automated graph-based method for the ontological development of the Punjabi domain	The domain ontology of Punjabi text is generated by the suggested system.	A total of 1000 Punjabi documents about politics, entertainment, sports, health, and agriculture are manually retrieved from the internet.	
Nidhi and Gupta [15]	Punjabi text classification using a hybrid approach and domain-based ontology	Suggested a hybrid, ontology-based pre-processing algorithm and trained and evaluated a machine learning algorithm for classification.	180 news articles in the Punjabi language were taken from various digital sources to create the corpus.	85% classification accuracy for hybrid and ontology-based methods, 71% for centroid-based methods, and 64% for naïve bayes classification
Bhatia and Sharma [16]	Used the morphological processing of Punjabi text to construct a converter from Punjabi to the Universal Network of Languages (UNL).	The converter was designed using a rule-based conversion mechanism.	To define morphological norms, Punjabi sentences from various books and novels, and internet-based articles were used.	The conversion performance of the suggested system was notable.
Gupta [17]	Punjabi words are automatically normalized for Human Language Technology applications.	Punjabi noun normalization using a rule-based algorithm	From internet news sources, 50 Punjabi news stories were painstakingly gathered.	Significant standardization of Punjabi nouns with a spelling variance of 1.562%
Kaur et al [12]	carried out optical character recognition for Gurmukhi text in Punjabi.	Using Matlab's toolbox, algorithms for image acquisition, recognition, and classification	Images of handwritten Punjabi text are manually extracted from Punjabi authors.	Produce good outcomes while classifying Punjabi characters.
Grover S., Verma A. [18]		Naïve Bayes, SVM, Rule-Based Engine	Standard Punjabi dataset "HC Corpora" Linguistics:	
Songbo and Jin [19]	Sentence-level sentiment analysis is done	Classifiers such as Centroid, K-nearest, Winnow, Naïve Bayes, and SVM	Chn- sentiCorp	With SVM and IG, the micro F1 is 90.60%, and the macro F1 is 90.43%.
Goularas and Kamis [20]		Random forest and SVM		Accuracy = 95% with random forest.

5. Methodology

Punjabi sentiment analysis is a recent analysis that needs to be understood for better online communications. For that, an application based on ASP.NET with SQL SERVER as the end is used. The main feature of this application is that it maintains the directory of conjunctive, disjunctive, negation words, positive, and negative keywords. Then, Punjabi text is provided in the textbox, and after the submit button, it gives a result by categorizing words of a sentence in positive, negative, or neutral polarity. The Lexicon approach is used to discover the emotive tone of the sentence. For this, the first collection of words is given scores (Positive, Negative, or Neutral) with a pre-assigned sentiment score. Then, Sentiment words are identified following the identification of conjunction and disjunction words. Analyse connected words. Finally, the whole sentence is evaluated for its opinion mining.

5.1. Collecting Datasets

The dataset used for testing the system consists of 40 reviews (1748 words) retrieved from a popular newspaper website, www.jagbani.com. On these 1748 words, when the test is applied, 68 words come out to be positive, and 172 are negative.

5.2. Specification Required

Hardware Specification Processor- Any Intel or AMD Ryzen 7 9700X processor supporting the SSE2 instruction set.

RAM- Minimum 1024 Mb.

Operating system used- Windows XP, Windows 7, and Windows 8. b) Software Specification.

Programming Environment- C#, ASP.NET

SQL Server is used for handling the database of the system.

5.3. Algorithm

Step Number 1: Initially, the Punjabi Reviews are collected in one paragraph.

Step Number 2: Then make a repository of Punjabi words in an array.

Step Number 3: Set the flag value as “POSITIVE”.

Step Number 4: The word is selected from the paragraph.

Step Number 5: Suppose if the outcome is “POSITIVE”.

(A) Change the flag outcome to TRUE Termination of Step 4 Loop

Step Number 6: Suppose that the flag outcome is “TRUE”.

(B) Every letter of the Punjabi paragraph word with one one letter of all the words of the database.

Step Number 7: If the outcome is “TRUE”.

Step Number 8: Display the sentence as “positive” and mark the complete sentence “positive” from a paragraph.

Stop the Step number 5 if statement.

Stop the Step number 6 if statement.

Step Number 9: If the sentence is negative, then highlight the whole sentence from a paragraph and count the number of words.

Step Number 10: The result is shown in graph form and shows the total scores of a sentence.

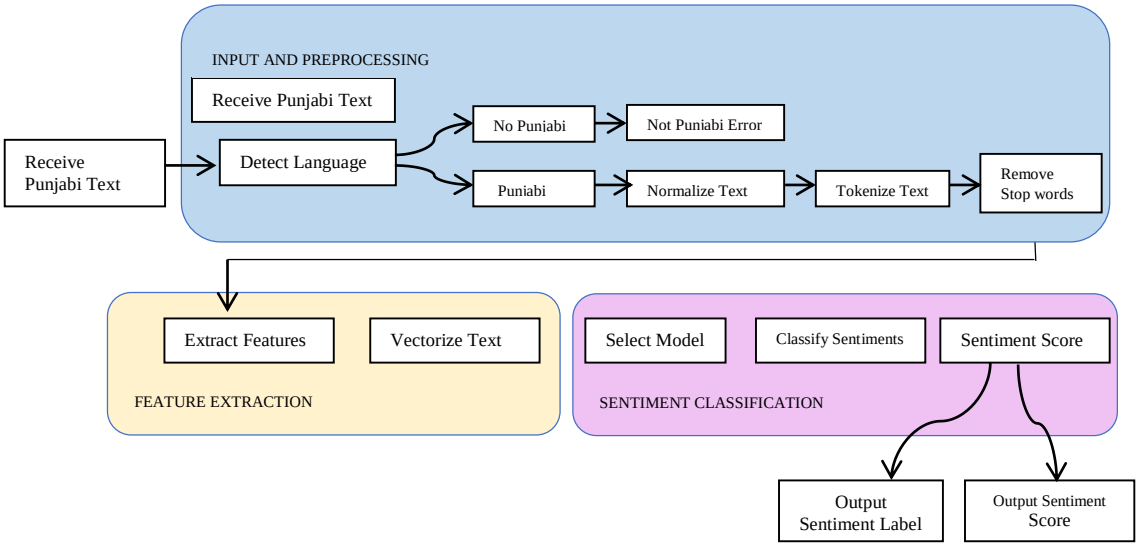

Fig. 4 Proposed flowchart

5.4. Conjunction, Disjunction, and Negation Handling

5.4.1. Conjunctive Words

These words help to connect words and phrases in a sentence. Words like ਅਤੇ (And), ਜਾਂ (Or), ਪਰ (But), ਕਿ (That / Because) , ਜਿਵੇਂ (As / Like), ਸਗੋਂ (Rather). To symbolize “and boy came” in logic, use the disjunctive formula A∧B,

Where ‘A’ stands for “the girl” and ‘B’ for “boy came.”Similarly, in Punjabi “ਮੁੰਡਾ ਅਤੇ ਕੁੜੀ ਆਏ ਸੀ। "| These words helps to provide alternative within sentence.

5.4.2. Disjunctive Words

Disjunction, also known as V, is a logical connective that means “or” when read aloud. For example, the English sentence “it is sunny, or it is warm” in logic uses the disjunctive formula AvB, where ‘A’ stands for “it is sunny” and ‘B’ for “it is warm.” Similarly, Punjabi text –“ਇਹ ਧੁੱਪ ਵਾਲਾ ਹੈ” ਜਾਂ “ਇਹ ਨਿੱਘਾ ਹੈ” | These words help to provide alternatives within a sentence. Words like- ਜਾਂ(Or), ਜਾਂ ਫਿਰ(else), ਨਹੀਂ ਤਾਂ(Otherwise).

5.4.3. Negation Handling

Some words are classified as negation words like- "ਨਸੀਂ", "ਨਾਂਸ", "ਨਾ", "ਨਾਮਾਤਰ", "ਨਾਮੁਿਨਕਨ", "ਨਸੀਂ". These words flip the meaning and polarity of the given opinionated sentence. A list of these words is maintained in the dictionary.

6. Results & Discussions

The fundamental goal of sentiment analysis is to identify positive and negative words from selected paragraphs. Existing research on them has primarily focused on their junction with opinion. However, we should note that all of these ideas and definitions are adequately defined.

6.1. Login Form

Figure 5 shows the screenshot of the login form that requires user authentication by entering the valid username and password to access the forms of application.

6.2. Add Keyword Form

As shown in Figure 6, in this form, we add the conjunctive, disjunctive, and negation words in the dictionary/database for sentiment analysis

6.3. View Keywords Form

As shown in Figure 7, in this form, we can view all the keywords that we have added in the Add keywords form.

6.4. Add New Sentiment Words

Figure 8 shows the screenshot of the Add sentiment words form, which is used to enter the new sentiment words in the dictionary/database for sentiment analysis. Admin can add the words, assign the status positive or negative, and define the score of the words.

Fig. 5 Login form

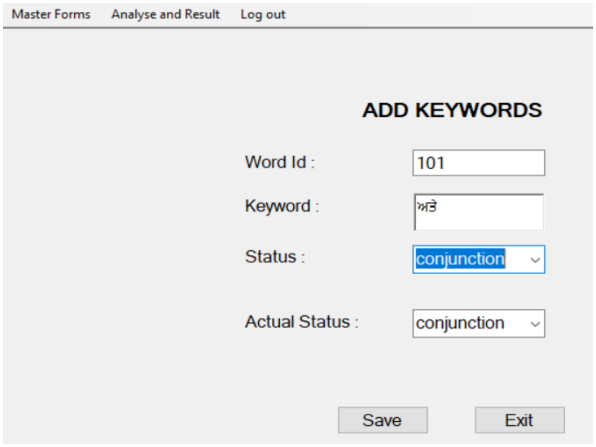

Fig. 6 Add keywords form

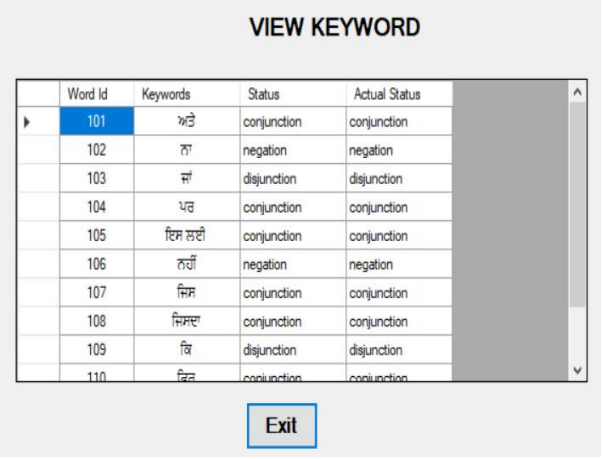

Fig. 7 View keyword form

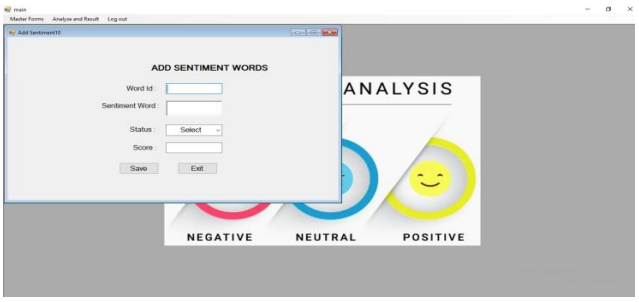

Fig. 8 Add new sentiment words

6.5. View Sentiment Words

Figure 9 shows the screenshot of View sentiment words. This form is used to view the entered sentiment word in the dictionary.

After clicking on the submit button, we got the following results:

ਨੇਤਾ ਜੀ ਨੇ ਲੋਕਾ ਦੀ ਭਲਾਈ ਲਈ ਕੋਈ ਵੀ ਕੰਮ ਨਹੀਂ ਕੀਤਾ । ਉਹ ਲਾਲਚੀ ਅਤੇ ਧੋਖੇਵਾਜ ਸੀ। ਉਸ ਨੇ ਲੋਕਾਂ ਤੋ ਧੋਖੇ ਨਾਲ ਪੈਸਾ ਇਕੱਠਾ ਕੀਤਾ ।

From the above result, we have found that there are three negative words (ਲਾਲਚੀ, ਧੋਖੇਵਾਜ, ਧੋਖੇ), three words are keywords (ਵੀ, ਅਤੇ, ਨਹੀਂ). In which ਵੀ and ਅਤੇ are conjunctive words, and ਨਹੀਂ is a negation word. Identified Negative words are highlighted in Red Colour. Identified Keywords are highlighted in Blue.

The result from Case 1 is shown in Figure 10. It shows the number of words, keywords, positive words, negative words, total positive score, total negative score, final score, and sentiment of the recently entered sentence.

7. Graphical Results Achieved

The graphical result shown in Figure 11 gives a bifurcation of a given sentence. It shows 2 pie charts in which the former depicts the total number of words (positive, negative, keywords) in a sentence, and the later chart represents the overall score of the sentence subject to the negative and positive polarisation of the sentence. Figure 12 represents tabular information of the given sentence, and the whole sentence polarity is calculated.

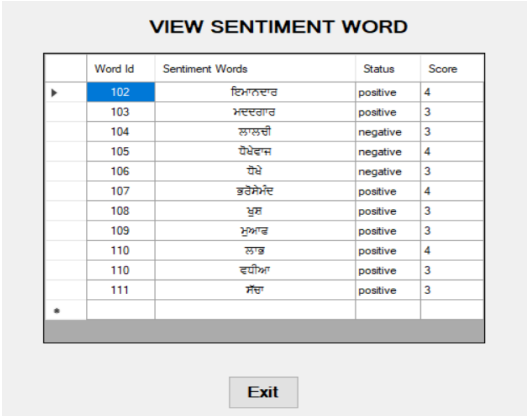

Fig. 9 View sentiment words

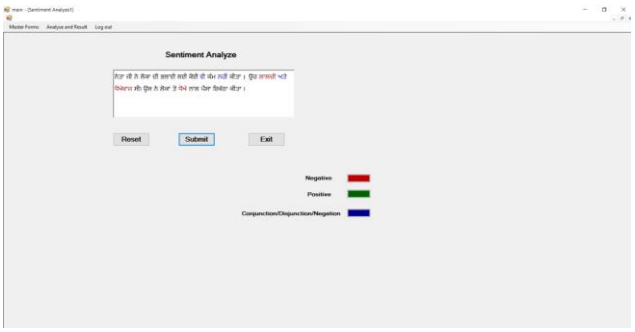

Fig. 10 Snapshot of sentiment analysis

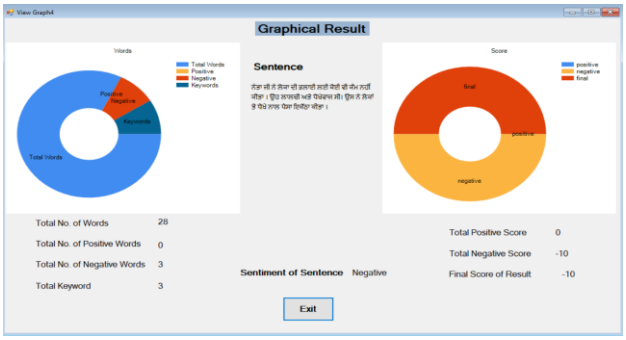

Fig. 11 Snapshot of graphical result

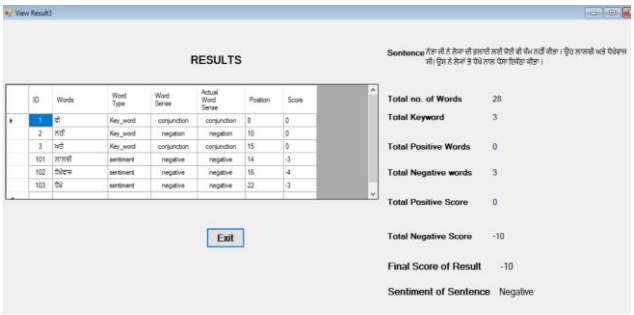

Fig. 12 Snapshot of polarity calculation

8. Parameters for Evaluation of the Proposed Algorithm

Evaluation parameters that tell how well a classification model works include parameters like precision, recall, and F-measure. Precision is essential for pattern identification and information retrieval using binary classification.

8.1. Precision

The proportion of optimistic predictions that are genuinely positive. A higher precision yields fewer false positives, whereas a lower precision yields more false positives. This is frequently at conflict with recollection, because decreasing recall is a simple method to enhance precision. In a practical sense, it answers the question: "Of all the instances the classifier labelled as positive, how many were actually correct?"

$$Precision = \frac{Number\ of\ true\ Positives}{Number\ of\ True\ Prositives+FalsePositives}$$

8.2. Recall

Recall is an important classification parameter that highlights the model's capacity to find all pertinent occurrences. It is sometimes referred to as sensitivity or the actual positive rate.

Improving recall can typically reduce precision since precision becomes increasingly complex as the sample space grows.

$$Recall = \frac{Number\ of\ true\ Positives}{Number\ of\ True\ Prositives+False\ Negatives}$$

8.3. F-Measure Metric

It is a metric for assessing a binary classification model's performance in statistical analysis and machine learning. It gives a balanced measure that is particularly helpful for issues with imbalanced classes by combining two additional metrics, precision and recall, into a single score. It is used to determine that the F-measure is approximately as valuable as accuracy.

$$F1 = \frac{2 * Precision * Recall}{Precision + Recall}$$

9. Challenges of Sentiment Analysis in the Punjabi Language

In spite of significant progress in the past few years, sentiment analysis still tackles several difficulties. Some of the existing problems in analysing sentiment are as follows:

Context-Specific Language: Sentiment analysis methods sometimes struggle to understand subtleties of context-specific language, such as irony, sarcasm, or metaphorical phrases. Sentiment analysis still struggles to accurately capture these nuances, which have a significant influence on the sentiment conveyed in context.

Domain-specific sentiment analysis: Language usage and sentiment expressions can vary significantly throughout domains, including social media, product reviews, and healthcare. As a result, it is possible that sentiment analysis models trained on one large dataset will not perform well with another.

Absence of labelled data: Sentiment analysis algorithms often require a considerable portion of labelled data for training. However, obtaining labelled data can be expensive, inefficient, and resource-intensive. The development of accurate and trustworthy sentiment analysis algorithms is still hindered by the lack of categorized data, especially for specific subjects or languages.

Sometimes, there are chances that sentences have similar lexical features but different sentiments

ਇਹ ਫ਼ਿਲਮ ਚੰਗੀ ਨਹੀਂ ਹੈ। (This movie is not good.)

ਇਸ ਤੋਂ ਵਧੀਆ ਕੋਈ ਫ਼ਿਲਮ ਨਹੀਂ ਹੋ ਸਕਦੀ। (No movie can be better than this). Sentences may have different style of writing but expresses same sentiment.

ਸੈਮਸੰਗ ਫੋਨ ਬਹੁਤ ਹੀ ਬੇਕਾਰ ਹੈ। (Samsung phone is extremely useless.)

ਮੇਰੇ ਪੈਸੇ ਸੈਮਸੰਗ ਫੋਨ 'ਤੇ ਬਰਬਾਦ ਹੋ ਗਏ। (My money was wasted on Samsung phone.)

ਮੈਂ ਸੈਮਸੰਗ ਦੀ ਬਜਾਏ ਆਈਫੋਨ ਖਰੀਦ ਸਕਦਾ ਸੀ। (I could have bought Iphone instead of Samsung.)

10. Conclusion and Future Scope

Currently, sentiment analysis is the most important field in feedback analysis. On the internet, it is nearly impossible to assess manual reviews provided by individuals about any product, service, decision, or politics. In the internet era, everyone voiced their thoughts on the web via feedback or social media in their native language. In addition, a significant amount of information has been generated in Punjabi on the internet. We created an application that evaluates the feelings of reviews by detecting the properties of the item or entity. Our primary goal is to create a method that works in Punjabi and analyzes the feelings, thoughts, or sentiments of the reviews. We tested the developed approach on more than 50 paragraphs. The developed application's results are 88% accurate when compared to manual results. The limitation of this strategy is that the application must be educated on a significant amount of data prior to testing. Pre-defined training data is not available for the Punjabi language. This study is highly beneficial to researchers who have worked on NLP applications employing the Punjabi language. In this study, we detect the qualities and negatives in a sentence to determine its polarity (negative or positive). In the future, scholars may work on synonym words combined with conjunction and disjunction terms to improve the results.

References

[1] Punjabi Language, Wikipedia. [Online]. Available: https://en.wikipedia.org/wiki/Punjabi_language

[2] 21st Century and Punjabi Language - Billion Voices Blog Magnon Sancus, 2021. [Online]. Available: https://www.magnonsancus.com/blog/english/21st-century-and-punjabi language/?utm_source=chatgpt.com

[3] Shahid Minhas, and Abiodun Salawu, "Wikipedia and Indigenous Language Preservation: Analysis of Setswana and Punjabi Languages," *Frontiers in Communication*, vol. 10, pp. 1-12, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[4] Abdullah Alsaeedi, and Mohammad Zubair Khan, "A Study on Sentiment Analysis Techniques of Twitter Data," *International Journal of Advanced Computer Science and Applications*, vol. 10, no. 2, pp. 361-374, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[5] M. Dadvar, C. Hauff, and Franciska M.G. de Jong, "Scope of Negation Detection in Sentiment Analysis," *11th Dutch-Belgian Information Retrieval Workshop, DIR 2011*, Amsterdam, Netherlands, pp. 16-20, 2011. [Google Scholar] [Publisher Link]

[6] Pooja Mehta, and Sharnil Pandya, "A Review on Sentiment Analysis Methodologies, Practices and Applications," *International Journal of Scientific & Technology Research*, vol. 9, no. 2, pp. 601-609, 2020. [Google Scholar] [Publisher Link]

[7] Mukhtiar Singh, Vishal Goyal, and Sahil Raj, "Sentiment Analysis of English-Punjabi Code-Mixed Social Media Content to Predict Elections," *Advances in Information Communication Technology and Computing*, pp. 81-90, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[8] Amandeep Kaur, and Vishal Gupta, "Proposed Algorithm of Sentiment Analysis for Punjabi Text," *Journal of Emerging Technologies in Web Intelligence*, vol. 6, no. 2, pp. 180-183, 2014. [CrossRef] [Google Scholar] [Publisher Link]

[9] S. Soumya, and K.V. Pramod, "Sentiment Analysis of Malayalam Tweets Using Machine Learning Techniques," *ICT Express*, vol. 6, no. 4, pp. 300-305, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[10] Raksha Sharma, and Pushpak Bhattacharyya, "A Sentiment Analyzer for Hindi Using Hindi Senti Lexicon," *Proceedings of the 11th International Conference on Natural Language Processing*, pp. 150-155, 2014. [Google Scholar] [Publisher Link]

[11] Manju Venugopalan, and Deepa Gupta, "Sentiment Classification for Hindi Tweets in a Constrained Environment Augmented Using Tweet Specific Features," *International Conference on Mining Intelligence and Knowledge Exploration*, Hyderabad, India, pp. 664-670, 2015. [CrossRef] [Google Scholar] [Publisher Link]

[12] Anterpreet Kaur, Parminder Singh, and Kamaldeep Kaur "Punjabi Dialects Conversion System for Majhi, Malwai and Doabi Dialects," *ICCMS '17: Proceedings of the 8th International Conference on Computer Modeling and Simulation*, Canberra Australia, pp. 125-128, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[13] Jasmine Hentschel, and Joyojeet Pal, "Sada Vehra: A Framework for Crowdsourcing Punjabi Language Content," *ICTD '15: Proceedings of the Seventh International Conference on Information and Communication Technologies and Development*, Singapore, pp. 1-4, 2015. [CrossRef] [Google Scholar] [Publisher Link]

[14] Rajveer Kaur, and Saurabh Sharma, "Semi-Automatic Domain Ontology Graph Generation System in Punjabi," *ICTCS '16: Proceedings of the Second International Conference on Information and Communication Technology for Competitive Strategies*, Udaipur, India, pp. 1-5, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[15] Nidhi, and Vishal Gupta, "Domain Based Classification of Punjabi Text Documents," *Proceedings of COLING 2012: Demonstration Papers*, pp. 297-304, 2012. [Google Scholar] [Publisher Link]

[16] Parteek Bhatia, and R.K Sharma, "Role of Punjabi Morphology in Designing Punjabi-UNL Enconverter," *ICAC3 '09: Proceedings of the International Conference on Advances in Computing, Communication and Control*, Mumbai India, pp. 562-566, 2009. [CrossRef] [Google Scholar] [Publisher Link]

[17] Vishal Gupta, "Automatic Normalization of Punjabi Words," *International Journal of Engineering Trends and Technology*, vol. 6, no. 7, pp. 353-357, 2013. [CrossRef] [Google Scholar] [Publisher Link]

[18] Sheeba Grover, and Amandeep Verma, "Design for Emotion Detection of Punjabi Text using Hybrid Approach," *2016 International Conference on Inventive Computation Technologies (ICICT)*, Coimbatore, India, pp. 1-6, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[19] Songbo Tan, and Jin Zhang, "An Empirical Study of Sentiment Analysis for Chinese Documents," *Expert Systems with Applications*, vol. 34, no. 4, pp. 2622-2629, 2008. [CrossRef] [Google Scholar] [Publisher Link]

[20] Dionysis Goularas, and Sani Kamis, "Evaluation of Deep Learning Techniques in Sentiment Analysis from Twitter Data," *2019 International Conference on Deep Learning and Machine Learning in Emerging Applications (Deep-ML)*, Istanbul, Turkey, pp. 12-17, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[21] Partha Mukherjee et al., "Effect of Negation in Sentences on Sentiment Analysis and Polarity Detection," *Procedia Computer Science*, vol. 185, pp. 370-379, 2021. [CrossRef] [Google Scholar] [Publisher Link]