ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20231227064230

Original scientific paper
Received: 2023-12-27, Accepted: 2024-04-18

Contextual Ambiguity Framework for Enhanced Sentiment Analysis

Pinky Yadav*, Indu Kashyap, Bhoopesh Singh Bhati

Abstract: Negation is a universal linguistic phenomenon that affects the performance of Natural Language Processing (NLP) applications, especially opinion mining data. Many words exists in sentences that have multiple interpretations or sentiments depending on how they are placed with respect to the negation word in the sentence. A cutting-edge framework is designed to tackle the nuanced challenge of detecting contextual ambiguity through negation in sentiment analysis. The approach uniquely combines advanced natural language processing techniques with deep linguistic insights, enabling the accurate interpretation of sentiment in complex sentences where negation plays a key role. The framework identifies negation cues and their scope, then assesses their impact on sentiment, considering contextual dependencies and word semantics. The model's innovation lies in context-sensitive algorithms that adeptly handle different sentence structures and idiomatic expressions, a notable advancement over traditional sentiment analysis tools. Particularly effective in interpreting sarcastic or ironic statements, the framework significantly outperforms existing models in accuracy, especially in negation-heavy contexts. This advancement enhances sentiment analysis applications like social media monitoring and customer feedback analysis, offering a more nuanced understanding of public opinion.

Keywords: contextual ambiguity; linguistic nuances; machine learning; natural language processing; sentiment interpretation; text analytics

1 INTRODUCTION

Contextual ambiguity can be defined as the presence of such words in sentences that have multiple interpretations or sentiments depending on how they are placed and what the grammar usage has been. It is a very challenging task in sentiment analysis as the models need to consider the surrounding words, phrases, or sentences to accurately determine the sentiment. Consider the word "sick" in the following sentences:

a) "The new album by the band is sick! I love it."
b) "I feel sick today; I can't go to work."

In sentence (a) "sick" is used in a positive context to mean something is excellent. In sentence (b) "sick" is used in a negative context to indicate feeling unwell. The sentiment of the word "sick" varies based on the surrounding context. Contextual ambiguity poses challenges for sentiment analysis models because they must consider the broader context to accurately determine the sentiment of a text. Simple keyword-based approaches may not work well in these cases. To resolve contextual ambiguity, sentiment analysis models need to consider the surrounding words, phrases, and even the tone of the text to make an accurate sentiment prediction [15]. They should analyze the entire sentence or paragraph, not just isolated words. Some sentiment analysis models use sentiment lexicons and predefined rules to help disambiguate sentiment in context. Lexicons assign sentiment scores to words based on their typical usage, and rules define how to handle certain context-specific cases. In some cases, domain-specific or user-defined rules can be added to sentiment analysis models to handle contextual ambiguity unique to a particular application or industry [16]. Therefore, this research work delves into the world of contextual ambiguity in sentiment analysis, examining the challenges it poses and its significance. It also explores the impact of negation words on sentiment analysis, illustrating how they can reverse or modify sentiment and complicate the task. The objective is to present the complexities of this problem and provide a new approach that addresses these challenges effectively. Addressing Contextual Ambiguity offers various compelling motivations such as Enhanced Communication, Improved User Experience and Effective Informational Retrieval. It also leads to precision in decision making, Advancements in Artificial Intelligence and Ethical considerations.

1.1 Contextual Factors

In addition to providing an objective account of events, texts frequently communicate the sentiments of authors or people involved in the recounted event. The emotional disposition is conveyed through the selection and organization of language. Sentiment is supported by context, which gives statements complex interpretations that are frequently dependent on language, situational, or cultural nuances. The same phrase might express completely different feelings in different contexts, so ignoring this context can result in a serious misreading of sentiments. While certain words consistently exhibit either positive or negative connotations, others are prone to undergo contextual shifts in valence as a result of neighboring words and the overall structure of the text. Instances of valence-shifting encompass the subsequent examples:

- The concept of negation refers to the logical operation of denying or contradicting a statement or proposition. The predominant form of valence-shifting, seen by the phrase "She doesn't like this laptop," involves the utilization of the negation "doesn't" to reverse the affective connotation of the verb "like."
- Defective verbs are lexical items employed to convey various notions such as capacity, certainty, permission, request, competence, suggestion, order, obligation, or advice. These verbs are inherently incomplete and require the presence of another verb that possesses semantic content and modifiers to convey a complete meaning. When a statement contains a word that is deficient, the emotional significance of the sentence is sometimes diminished due to the

inadequacy of the feeling word or phrase used in conjunction with the verb. The statement "The school should reduce the tuition fee" expresses a negative evaluation, whereas the statement "The school reduces the tuition fee" presents a factual observation.

- Words with heightened or lessened intensity: These lexical choices amplify or reduce the emotional significance of the associated term. In the given instance, the phrase "She studies hard" exhibits a heightened emotional connotation in comparison to "She studies quite hard," while possessing a diminished emotional connotation when contrasted with the statement "She studies very hard." The observed correlation between emotions (e.g. intensity) and the presence of reinforcing or mitigating words (e.g. "very" or "quite") suggests a causal relationship.
- The phenomena of contrast is observed when the text from the provided link displays a shift in emotional valence, indicated by the use of phrases. Typically, the emotional inclination of the entire sentence is only evident on one side. An instance of an insulting remark can be observed in the given statement, "This laptop is beautiful, but the price is too expensive." The term "too expensive" contributes to the pejorative nature of the remark, mostly owing to the presence of the linking word "but" earlier in the sentence.
- Incompatibility or inconsistency is evident in numerous instances where the emotional inclination of a text is largely reliant on the surrounding context.
- Valence-shifting is a phenomenon that occurs when the semantic interpretation of a word is subject to alteration within a particular context [11]. Machine learning techniques like bag-of-words and n-grams fail to account for the impact of negation structures and other sentiment valence shifter structures. This limitation is evident in statements like "I like this hotel but the price is quite high" and "I like this hotel but the price is too high." Based on a bag-of-words model, it is probable that the sentiment value of the two lines is equivalent, as they both contain the emotive words "like" and "high." Nevertheless, the bag-of-words approach fails to account for the presence of words such as "quite" and "too," which significantly influence the sentiment conveyed by the two statements. Sequence mining techniques are employed to uncover patterns related to valence-shifting. These patterns encompass several linguistic phenomena, including negation, contrast, intensification, and attenuation of polarization. Semantic Oriented CALculator (SO-CAL) is an innovative system that effectively addresses valence-shifting through the utilization of rule-based procedures and a mix of methodologies. This system incorporates rules and collections of words that have been annotated with emotional attributes. The researchers [14] and [15] utilize a dependency tree that incorporates comprehensive syntactic structure information in order to establish syntactic rules for assessing the influence of negation and other valence-shifting structures on the emotional polarity of sentences or the entirety of a document. In this analysis, straightforward yet efficient strategies are examined for employing rules to identify instances of contextual valence changing. The utilization of neural network methods, particularly the Long short-term memory (LSTM) network [16], has facilitated the advancement of deep learning models. In this regard, the attention mechanism has emerged as a potent approach for effectively representing context.

Although previous research work have investigated the role of gesture in conveying denial, there is a need for more investigation in the context of the interplay between scope-bearing elements, particularly the relationship between different types of quantification and negation. Interactions of this nature present a communication issue due to the production of semantically ambiguous phrases that can be interpreted in numerous ways, with the scope of negation varying depending on the intended interpretation. An example is provided in which the negator and quantifier are underlined. All the magnolias won't bloom [3]. The relevant interpretations are captured by the logical representations in the following: $\forall x$ [magnolia $(x) \rightarrow \neg$bloom (x)], $\neg\forall x$ [magnolia $(x) \rightarrow$ bloom (x)]

To demonstrate the lack of consistency in empirical findings regarding the English language [10] conducted perception experiments. These experiments focused on the prosodic cues used to disambiguate socially ambiguous sentences. Specifically, sentence-final falling intonation was used for sentences such as "all>not," while sentence-final rising intonation was used for sentences like "not>all." The results showed that participants were able to interpret the intended meaning with a relatively high degree of success, ranging above 63% depending on the sentence type. This success rate was significantly higher than chance. Again, these rates were significantly higher than chance. In contrast, the initial production experiment [2] demonstrated that speakers tend to provide prosodic cues with limited reliability in a semi-naturalistic environment. This context involved the speakers reading scripted ambiguous lines that were embedded within disambiguating paragraph-length texts. In the context of the production investigation, it was observed that speakers commonly employed descending intonational contours when articulating negation combined with quantification utterances. According to the findings of Syrett et al. [17], it was determined that there is no direct correspondence between intonation contour and scopal interpretation in the context of language production. Therefore, it was concluded that the patterns of intonation cannot be considered as either a necessary or sufficient condition for resolving the ambiguity in sentences with multiple scopal interpretations [12].

The remaining sections of the paper are organized as follows: The next section discusses prior work on sentiment analysis for contextual ambiguity. Section 3 elaborates on the proposed methodology and model building process. Section 4 offers the results. Finally, Section 5 discusses the conclusions of the research, and outlines future directions.

2 LITERATURE REVIEW

The field of sentiment analysis has evolved significantly since its inception. Early foundational work by Pang and Lee [6] outlined basic approaches for classifying sentiments in texts, primarily using bag-of-words models and basic machine learning techniques. These techniques laid the groundwork for more advanced sentiment analysis. Negation handling is a well-acknowledged challenge in sentiment analysis. Wiegand et al. [2] highlighted the complexities that negation introduces in determining sentiment orientation, emphasizing the need for sophisticated linguistic analysis. Similarly, Councill et al. [3] demonstrated the impact of negation on sentiment analysis accuracy. Contextual ambiguity in sentiment analysis has been explored extensively. Studies by Cambria et al. [4] emphasized the importance of context in interpreting sentiments, especially in idiomatic and sarcastic expressions. This is further supported by research of Reyes and Rosso [5], who focused on the role of irony and sarcasm in sentiment analysis.

The integration of advanced NLP techniques has significantly improved sentiment analysis. Socher et al. [6] introduced deep learning approaches, which leverage neural networks for better understanding of contextual nuances. This is further elaborated by Zhang et al. [7], who explored the use of deep learning for negation and speculation identification.

Arabic language poses challenges for automatic processing due to its multiple dialects, ambiguous syntax, and limited high-quality datasets. Duwairi et al. [21] introduced a novel framework for augmenting Arabic sentences using the language's rich morphology, synonymy lists, and grammatical rules. The approach, focused on sentiment analysis, significantly increased the size of initial datasets and improved accuracy by 42 % through reliable rule-based augmentation. Aoumeur et al. [22] introduced a new dataset, CASAD, derived from art books and labeled by human experts. Unlike previous methods relying on word frequency, this method employs word embedding techniques Word2Vec to extract deep relations in formal Arabic language features. The dataset is evaluated using machine learning algorithms such as SVM, LR, NB, KNN, LDA, and CART, with statistical methods for validation and reliability. Results show that the Logistic Regression with Word2Vec approach achieves the highest accuracy in predicting topic-polarity occurrences in classical Arabic texts.

Alshutayri et al. [23] explores sentiment analysis on social media, focusing on Arabic Twitter data. It employs machine learning algorithms like Naïve Bayes, Logistic Regression, and Support Vector Machines, with bigrams and unigrams as features. Logistic Regression achieved the highest accuracy at 63.40 %. Additionally, a deep learning approach using Long Short-Term Memory (LSTM) neural network achieved a higher accuracy of 70 %, surpassing related works in the field. The effectiveness of these methods in analyzing sentiments expressed in Arabic tweets is demonstrated. Cumaoğlu et al. [24] introduced a new dataset gathered from Twitter, focusing on Arab opinions about Turkey across various topics. The dataset is multi-dialectic Arabic and covers fields like the Turkish economy, tourism, food, and politics. A deep learning-based Arabic Sentiment Analysis (ASA) approach using Word2Vec and Bidirectional Encoder Representations for feature extraction is employed. Bidirectional long short-term memory, Convolutional neural networks, feedforward neural networks, and a transformer auto classifier based on AraBERT are applied for binary classification. The results show that AraBERT outperforms Word2Vec, and the transformer auto classifier achieves the highest accuracy in classifying positive and negative emotions in the dataset.

Refai et al. [25] addressed the challenge of dataset adequacy in Arabic language models by proposing a new Data Augmentation (DA) method using AraGPT-2. Existing approaches for Arabic DA are limited, relying on traditional methods like paraphrasing. The proposed method leverages AraGPT-2 for augmentation and evaluates generated sentences using various metrics. The augmented dataset is then tested on sentiment classification tasks using the AraBERT transformer. Results demonstrate improved performance across different sentiment Arabic datasets, with F1 scores increasing by 4 % in AraSarcasm, 6 % in ASTD, 9 % in ATT, and 13 % in MOV.

Díaz et al. [25] worked on negation and speculation detection in natural language processing. It defines these phenomena, discusses the need for their processing, reviews existing research, and provides a comprehensive list of resources and tools. This research introduced new datasets and scripts, offering an overview of the current state of the art in negation and speculation detection. Mahany et al. [1] introduced NSAR, the first Arabic corpus designed for the purpose of review analysis, which has been annotated with negation and speculation markers. The dataset consists of 3,000 words that span various categories, revealing that 29% of these sentences entail negation, while just 4% of them contain conjecture. The presence of a high level of inter-annotator agreement serves as evidence of the dependability of the data. This particular resource holds significant importance in the field of Arabic Natural Language Processing (NLP), since it provides valuable insights into the prevailing linguistic characteristics seen in reviews.

Significant progress has been achieved in the field of supervised relation extraction (RE) using neural networks (NNs) in recent years, particularly with the introduction of deep learning models such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) [18]. Within this particular paradigm, the task of relation extraction (RE) is approached as a classification problem. In order to accomplish this objective, informative features are carefully constructed using training data, and a variety of classification models are then trained. The taught classifiers are subsequently utilized to make predictions about relationships, either using the pipeline approach or the joint learning approach. The pipeline approach involves the sequential execution of named entity recognition (NER) and relation extraction (RE) operations [19].

The first step in the research experiment involves conducting Relation Extraction (RE) on sentences that contain identified entity pairs. Afterwards, entities that have relationships are merged into triples and presented as the predicted outcome. Nevertheless, this methodology involves the division of Named Entity Recognition (NER) and Relation Extraction (RE) activities, hence rendering the performance of RE vulnerable to the outcomes of NER. On the other hand, the collaborative learning approach entails the concurrent detection and extraction of many sorts of relationships between items [20]. Both methodologies commence by acquiring the fundamental vector representation of characters, subsequently extracting sentence attributes through the use of diverse neural network models. The procedure is finalized by employing nonlinear classifiers for the purpose of relation classification.

Zheng et al. focused specifically on the detection of irony and sarcasm in text, noting that these linguistic devices often invert the literal sentiment, posing a significant challenge for automated systems. Their research highlighted the need for models to recognize contextual cues that indicate sarcasm or irony.

In the work by Poria et al. [3], the authors demonstrated that the sentiment polarity of a word or phrase can change dramatically based on the context in which it appears. This research underlines the need for sentiment analysis models to incorporate broader contextual analysis rather than relying solely on individual word sentiments.

The advancement in deep learning techniques has provided new opportunities for addressing contextual ambiguity. Socher et al. [4] showed how recursive neural networks could be used to understand the compositional effects of sentiment in sentences, accounting for the influence of context.

The emergence of contextual word embeddings, like those introduced by Devlin et al. [5] in BERT (Bidirectional Encoder Representations from Transformers), represents a significant leap in capturing the contextual nuances of language for sentiment analysis. These models are designed to understand the meaning of a word in the context of the surrounding text, thereby providing a more accurate sentiment interpretation.

3 METHODOLOGY

The aim is to enhance the performance of collaborative recommender systems by addressing challenges related to word ambiguity, data sparsity, and mistake rate. Textual evaluations are to be incorporated into the matrix of ratings provided by users for items. Textual reviews have the potential to mitigate the issue of data sparsity and enhance the quality of suggestions. The aim is to address the issue of domain sensitivity in recommender systems by including contextual information in the sentiment analysis model. The majority of sentiment-based recommender systems commonly employ conventional techniques that lack semantic analysis. Consequently, these systems encounter challenges related to data sparsity and domain sensitivity. There have been limited recommendations on the inclusion of domain sensitivity information in order to enhance sentiment accuracy and address the issue of data sparsity [10]. The manner in which domain sensitivity information can enhance the quality of suggestions remains unclear.

3.1 Data Collection

The Amazon Movie Reviews dataset is a rich and extensive collection of user-generated reviews sourced from the Amazon website, one of the world's largest online retailers. This dataset forms a critical part of the web data used in sentiment analysis and natural language processing research. It encompasses 8 million reviews, offering insights into consumer opinions, preferences, and viewing experiences. Each review typically includes a textual comment along with a star rating, providing a dual perspective of qualitative and quantitative data. The sheer volume and variety of the reviews make this dataset an invaluable resource for training and testing machine learning models, especially those focused on sentiment analysis.

3.2 Data Preprocessing

Data preprocessing is an essential stage in any text analysis task. It involves transforming the raw data into a format that can be easily analyzed. Several preprocessing steps are used, including Data Cleaning, Tokenization, Removing stopwords, Removing punctuations, and Lowercasing.

3.3 Proposed Algorithm

There are two primary stages that are involved. The process initially starts with conversion of textual reviews into numerical ratings. The incorporation of sentiment grading into collaborative combination is a secondary consideration. Context-based sentiment analysis is a computational approach that involves the conversion of domain-specific textual reviews into numerical ratings. The CF method utilizes the extended user-item matrix obtained from the previous phase in order to calculate the recommendation value.

Inputs:

R: A set of textual reviews.
S: Traditional numerical ratings.
Algorithm Steps:
Textual Review Preprocessing and Numerical Conversion

Step 1.1 Text Preprocessing:
- Segment phrases and annotate with part-of-speech (POS) tags.
- Apply custom tagging for domain-specific features.

Step 1.2 Opinion Word Extraction:
- Extract opinion words using predefined linguistic rules.

Step 1.3 Sentiment Scoring:
- Apply sentiment analysis techniques to convert reviews into numerical sentiment scores.

Step 1.4 Merge Scores:
- Combine S with sentiment scores to create an augmented dataset.
- Extended User-Item Matrix Construction

Step 2.1 Matrix Formation:
- Construct an extended user-item matrix incorporating both traditional ratings and sentiment scores.
- Collaborative Filtering with Sentiment Analysis

Step 3.1 Score Generation using OWs:

Step 3.2 Score Generation using Negation

Step 3.3 CNN/K-Means based distance Clustering of the scores:
- Compute review-review similarity using cosine similarity on the extended matrix.
- A probabilistic model, denoted as p(rai, nri) is utilised for generating representation space.

Step 4.1 Contrastive Loss:
- Applying the NTXent Loss values to the clusters.

Step 4.2 Identifying Class for each cluster:
- Attract puts them in same class, repel puts them in unknown class.

Step 4.3 Domain-Specific Tailoring:
- Unsupervised method puts the reviews in same class.
- Unsupervised method requires user guidance

Step 5.1 Comparative Analysis:
- Compare the performance of the scores against general-learned scores.

Step 5.2 Accuracy Assessment:
- Evaluate sentiment analysis accuracy and recommendation relevance.

System Output:
- Final polarity of the statement based on the context, whilst understanding contextual ambiguity.

End

The Collaborative Combination (CC) algorithm incorporates both rating and sentiment rating. A sample of 1000 textual evaluations sourced from Amazon's domains pertaining to TV and Film series was utilized for the purpose of conducting a comparative analysis. The presentation of experimental data within the domain of television and film is focused. The commencement of dataset analysis is initiated. The input datasets consist of textual reviews and numerical ratings. The conversion process commences by conducting pre-processing on textual data to prepare it for textual review input. Opinion words are extracted by the use of specific rules, which are defined for the contextual ambiguous words. The pre-processing endeavor encompasses the tasks of phrase segmentation, part-of-speech (POS) annotation, and the incorporation of custom tags. Traditional sentiment analysis involves the transformation of reviews into numerical ratings. The sentiment analysis of user movie reviews is employed to ascertain a discernible inclination toward specific ratings. A training set is utilized in machine learning to estimate a probabilistic model, denoted as p(rai, nri). The emotion score of each review was computed through the utilization of unsupervised machine learning techniques. This methodology yields quantitative ratings. The creation of an extended user-item matrix is achieved through the merging of the input data.

Fig. 1 illustrates the augmented user-item rating matrix, wherein sentiment ratings are incorporated to reflect user preferences. However, it should be noted that the rating matrix exhibits sparsity. The task of accurately predicting and providing recommendations for dependable products necessitates considerations beyond just ratings. Sentiment analysis is utilized to examine the contextual information present in user ratings. The enlarged user-item matrix is augmented with a sentiment rating derived from sentiment analysis.

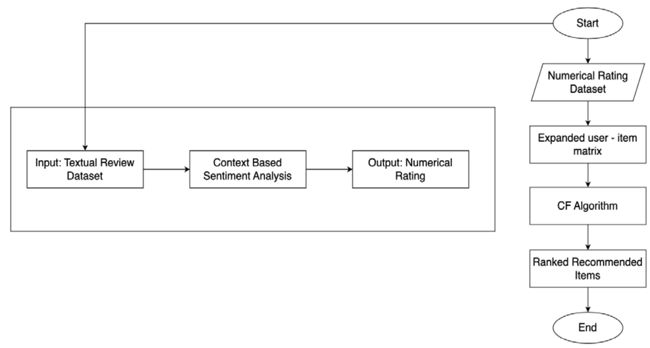

Figure 1 Flow chart for ranked recommendation

The feedback provided by all users is utilized in the computation of suggestions. Typically, a matrix is employed to represent the relationship between users and items.

The items in the user-item matrix are assigned ratings by users. In order to identify the closest neighbors, the neighborhood-based approach initially identifies a cohort of users who exhibit the highest degree of similarity to the active user. The predictive power of active user ratings can be determined by calculating the weighted aggregate of their ratings. The comparison of users is conducted through the utilization of cosine similarity.

The evaluations for second and third-process adjectives are consolidated. Subsequently, the scores of noun and verb phrases obtained from the second stage are utilized. The sentiment lexicon has been carefully calibrated and is responsive to the characteristics of the corpus. The efficacy of the model was enhanced through testing on Amazon TVs, films, and technological devices. The sentiment lexicon generator incorporates two sentiment ratings in addition to the lexicon. The scores of each word are computed and then averaged. The opinion words and negative words are identified and are kept for training the model. The model mentioned in Fig. 1 is used to depend on the opinion words and negative words one by one to create two different scores of the statement showing dependency on the respective kind of words being used in the review.

In order to identify which scoring provides better correlation with the actual score, K means based distance is calculated, via applying the NT-Xent loss, which is a loss function used in Contrastive Learning, which follows the concept of attracting similar scores in one class, and repelling the different scores in different classes, requiring supervision from the user, which thus combines both, supervised learning as well as unsupervised learning. From the attract domain, where the dimensionality is increased to 1000 due to the application of the NT Xent loss, the final polarity of the review can be found, which is now dependent on both, contextual words as well as negative words. The implementation of contextual rules enhances the level of accuracy. The inclusion of valence shifters such as negations, intensifiers, and diminishers yielded enhanced outcomes. In order to evaluate the accuracy of this approach, the lexicons generated using domain sensitivity were compared to the general lexicon. The primary investigation employed a dataset consisting of Movie Review information.

The language developed by Bing Liu served as the baseline for the experiment. The Bing Liu lexicon is a sentiment classification tool that is applicable across several domains. The newly generated vocabulary was evaluated against the existing lexicon in order to assess the performance of the model. The experimental results indicate that the lexicon generated in this work exhibits superior performance compared to general lexicons. This can be attributed to the fact that the lexicon was constructed using a corpus from the same domain. The lexicon was utilized to compute a numerical sentiment score for each review paper.

The electronic product review dataset underwent a similar testing procedure. The accuracy of both models is assessed. The subsequent empirical illustration highlights the significance of employing a sentiment lexicon that is attuned to the specific domain in order to enhance the performance of the model. The classification of the statement "I would suggest that you go read the book instead of watching the movie" as neutral within a general-purpose vocabulary can be attributed to its absence of sentiment-carrying terms. Nonetheless, a vocabulary that is sensitive to the domain of movies and has been specifically designed for this purpose would correctly classify the sentence as having a negative polarity. This is because the writer intends to express a negative sentiment towards the movie by suggesting that the audience should "go read the book" instead. Based on empirical evidence, the lexicon that is attuned to the specific subject in question deems the term "book" to possess negative connotations. A privacy-preserving mechanism used refers to any technique or approach designed to protect individuals' sensitive information while still allowing useful analysis or processing to take place. These mechanisms are crucial in situations where data needs to be shared or analyzed while ensuring that the privacy of individuals is respected and upheld. Privacy-preserving mechanism encompass a broad range of techniques and tools designed to enhance privacy in various contexts. This can include tools for anonymizing web browsing, secure messaging protocols, and privacy-focused operating systems or applications. These mechanisms can be applied in various contexts, including healthcare, finance, social media, and government, to ensure that sensitive data is handled responsibly and ethically while still allowing for meaningful analysis and processing.

The process of domain adaptation involves the refinement of the sentiment lexicon in order to assign domain-specific polarities to phrases, hence enhancing the accuracy of the sentiment analysis model. The numerical scores derived from dataset documents can be utilized as input for collaborative filtering recommender systems subsequent to converting textual data into sentiment ratings. The collaborative filtering system utilized user sentiment to make predictions, rankings, and recommendations for movies. The advantages of proposed model includes a) Contextual Comprehension These embeddings grasp word meanings based on their context within a sentence. b) Ambiguity Resolution: Contextual embeddings help clarify polysemous words by capturing their various senses in different contexts. c) Capturing Word Relationships: Contextual word embeddings capture intricate dependencies between words in a sentence, incorporating both preceding and following words. d) Transfer Learning: Pre-trained on large text datasets using unsupervised methods, contextual word embeddings capture general language patterns and semantics. e) Leading Performance: Contextual word embeddings have demonstrated superior performance on a wide array of NLP benchmarks and tasks, such as sentiment analysis, named entity recognition, question answering, and machine translation.

4 RESULTS

The research utilized datasets from Amazon that consisted of movie ratings and textual evaluations. Additionally, records pertaining to electrical devices sold on Amazon were also employed. Various hyperparameters used in the training Process are Learning rate, Training Epochs/Batch size, Number of layers, Negation Sensitivity, Majority Voting Threshold and Dropout rate. Three different models were employed: the basic ratings-based collaborative filtering (CF) model, an enhanced CF model incorporating sentiment ratings, and a model using contextual sentiment ratings. The experiment baselines include Rating CF and sentiment CF. The experiment utilized the Bing Liu lexicons as the foundational framework. The contextCF will augment the lexicons by including the newly developed domain lexicon. The present research work utilizes benchmark datasets sourced from Amazon. Polarity Identifier system leads to more nuanced interpretations of emotion, whether they have to do with figurative language like irony or with identifying which entities or attributes of entities are truly being discussed in an opinionated remark. This is evident in the development of resources as well. For example, we have observed that polarity has been added to corpora at the sentence level, as opposed to the previous practice of adding it at the entity or global document levels. As a result, the Identifier System must constantly improve in accuracy and sophistication. As a result, they are trying to undertake deeper semantic processing, which has some success with linguistic data that is cleaner.

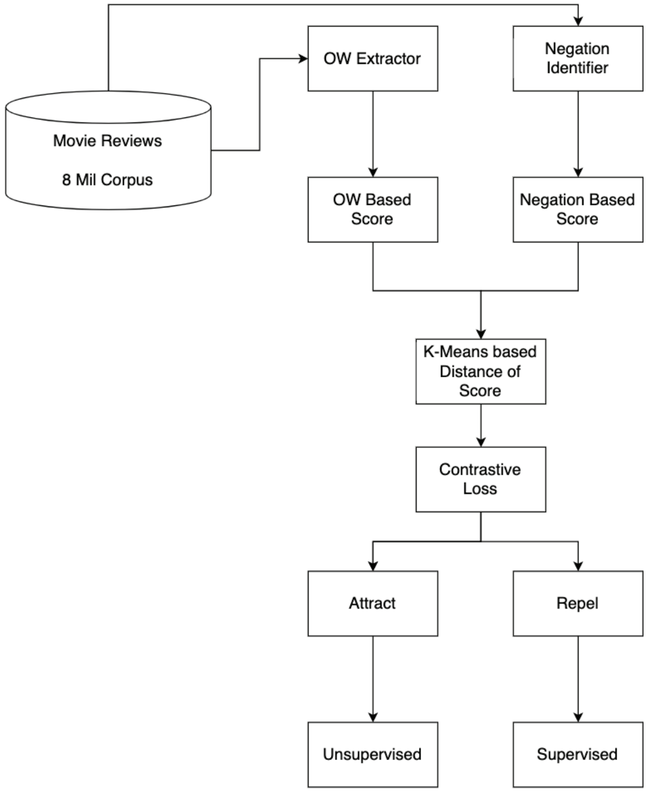

Figure 2 Polarity identifier systems

The experiments were conducted on two distinct domains in order to substantiate the efficacy of the proposed technique. The first dataset has a total of 50,000 ratings and 1000 written reviews, with a scaling range of 1 to 5. These ratings and reviews pertain to various categories such as televisions, movies, and electrical devices. In order to mitigate the risk of overfitting, the dataset is divided into three subsets, namely training, evaluation, and testing data [23].

Table 1 Comparison of the proposed algorithm with RCF & SCF

Model	Amazon TV and Movies			Amazon Electronic Products		
	Accuracy	RMSE	MAE	Accuracy	RMSE	MAE
Rating CF	80.80	3.14	2.47	89.8	2.49	3.34
Sentiment CF	84.35	2.57	3.43	92.7	3.04	3.80
Proposed	90.70	3.45	4.33	99.3	3.95	3.79

The dataset was partitioned into two sets: 80% for training purposes and 20% for testing purposes. The test set comprises items that were subjected to random testing for each user, whereas the training set is utilized to train the recommendation model. The evaluation of proposals generated by a recommender system should be based on their anticipated accuracy and the value they provide to users. The performance of recommender systems is significantly influenced by the level of accuracy they exhibit. The evaluation of recommender system performance is typically conducted using metrics such as Mean Absolute Error (*MAE*) and Root Mean Square Error (*RMSE*). The evaluation of recommender systems commonly employs standard measurements such as *RMSE* and *MAE* to gauge the accuracy of predictions [19].

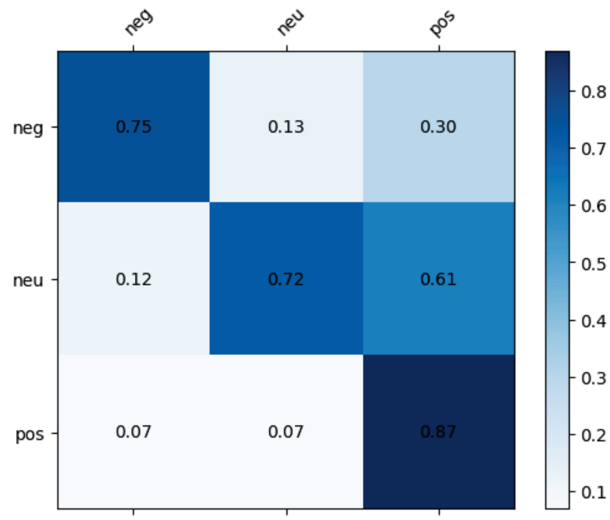

Figure 3 Confusion matrix

5 CHALLENGES

Despite the progress made in addressing contextual ambiguity and negation in sentiment analysis, several challenges remain. These include:

- Contextual Sensitivity: Capturing precise contextual information and understanding nuanced expressions of sentiment remains a challenge. Models that over-rely on context may struggle when the surrounding text is insufficient to determine sentiment.
- Ambiguity Resolution: Resolving ambiguity in sentiment analysis is an ongoing challenge, particularly in creative and informal texts. Handling phrases that can be both positive and negative is a complex task.
- Cross-Linguistic Variability: Different languages and cultures may have unique linguistic nuances, making sentiment analysis in multilingual settings challenging.

Real-Time Sentiment Analysis: In applications like social media monitoring, real-time sentiment analysis is crucial. Balancing the need for rapid analysis with accurate results poses a constant challenge.

Upon examination, it may be concluded that a variety of unique obstacles significantly affect how accurate predictive models are. Sentiment analysis models must be updated to reflect new language trends and issues. Since political discourse is constantly changing. If outdated characteristics are relied upon, sentiment analysis models may become significantly inaccurate. Separating fleeting from persistent sentiment characteristics presents additional challenges that frequently lead to overfitting or the omission of emergent sentiment patterns, which eventually deteriorates model accuracy, particularly when training and test data distributions are different. Due to subjective interpretations and the ambiguity of political attitudes, the addition of human annotation to the active learning process brings an additional layer of variability and potential bias, introducing noise.

6 CONCLUSION AND FUTURE WORK

Contextual ambiguity and negation are critical challenges in the field of sentiment analysis. The complexities are explored that arise when dealing with the interplay of these two factors and examined existing approaches to mitigate them. From machine learning models to sentiment lexicons and rule-based systems, there are diverse strategies that can be employed to handle these challenges. It is clear that context-aware sentiment analysis is essential for accurate sentiment interpretation. As sentiment analysis continues to find applications in various domains, from marketing and customer service to public opinion analysis, addressing contextual ambiguity and negation becomes increasingly vital. The research underscores the need for further research and development in this field.

Given the advancements and challenges identified in the "Negation-Aware Contextual Ambiguity Framework for Enhanced Sentiment Analysis," future research can be directed towards several promising avenues:

1) Integration of Cross-Lingual Capabilities: Expanding the framework to support multiple languages, especially those with complex linguistic structures and varying syntax. This would involve training the model on diverse datasets across different languages, addressing the challenges of negation and contextual ambiguity in a multilingual context.
2) Incorporating Context-Aware Word Embeddings: Further exploration of advanced word embedding techniques, such as contextual embeddings (e.g. BERT, GPT-3), to enhance the understanding of context in sentiment analysis. Future work could involve fine-tuning these models specifically for detecting nuanced sentiment expressions influenced by negation.
3) Dealing with Evolving Slang and Internet Jargon: The dynamic nature of language, especially on social media platforms, poses a continuous challenge. Future research should focus on continuously updating the model to understand evolving slang, abbreviations, and internet jargon which could carry sentiments influenced by negation.
4) Exploring the Role of Emojis and Multimedia Content: As sentiment analysis extends beyond text to include multimedia content, future work could explore integrating analysis of emojis, images, and videos, which often accompany text in digital communication and can influence the interpretation of sentiment.
5) Real-Time Sentiment Analysis Applications: Developing and testing real-time sentiment analysis systems, especially in dynamic environments like social media and customer service platforms, where immediate sentiment interpretation can provide valuable insights for businesses and organizations.
6) Ethical Considerations and Bias Mitigation: Future research must also consider the ethical implications of sentiment analysis, particularly in ensuring that the models do not perpetuate biases present in training data. This involves developing methods to detect and mitigate biases related to gender, race, and cultural background in sentiment analysis models.
7) Explainable AI in Sentiment Analysis: Enhancing the transparency and explainability of AI models used in sentiment analysis. This is crucial for users to trust and understand the basis on which the model makes its sentiment predictions, especially in sensitive applications like mental health analysis and political sentiment tracking.
8) Hybrid Models Combining Rule-Based and ML Approaches: Investigating the synergy between rule-based and machine learning approaches in sentiment analysis to leverage the strengths of both methodologies, especially in handling negation and contextual ambiguities.

7 REFERENCES

[1] Mahany, A., Khaled, H., Nouh, E., Aljohani, N. & Ghoniemy, S. (2022). Annotated Corpus with Negation and Speculation in Arabic Review Domain: NSAR. *International Journal of Advanced Computer Science and Applications, 13*(7). https://doi.org/10.14569/IJACSA.2022.0130706

[2] Reyes, A. & Rosso, P. (2014). Making objective decisions from subjective data: Detecting irony in customer reviews. *Decision Support Systems, 53*(4), 754-760. https://doi.org/10.1016/j.dss.2012.05.027

[3] Poria, S., Cambria, E. & Gelbukh, A. (2016). Aspect extraction for opinion mining with a deep convolutional neural network. *Knowledge-Based Systems*, 108, 42-49. https://doi.org/10.1016/j.knosys.2016.06.009

[4] Socher, R., Perelygin, A., Wu, J., Chuang, J., Manning, C. D., Ng, A. & Potts, C. (2013). Recursive deep models for semantic compositionality over a sentiment treebank. *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 1631-1642.

[5] Devlin, J., Chang, M. W., Lee, K. & Toutanova, K. (2018). BERT: Pre-training of deep bidirectional transformers for language understanding. *arXiv* preprint arXiv:1810.04805.

[6] Pang, B. & Lee, L. (2008). Opinion Mining and Sentiment Analysis. *Foundations and Trends in Information Retrieval, 2*(1-2), 1-135. https://doi.org/10.1561/9781601981516

[7] Wiegand, M., Balahur, A., Roth, B., Klakow, D. & Montoyo, A. (2010). A survey on the role of negation in sentiment analysis. *Proceedings of the Workshop on Negation and Speculation in Natural Language Processing*, 60-68.

[8] Councill, I. G., McDonald, R. & Velikovich, L. (2010). What's great and what's not: learning to classify the scope of negation for improved sentiment analysis. *Proceedings of the Workshop on Negation and Speculation in Natural Language Processing*, 51-59.

[9] Cambria, E., Poria, S., Bajpai, R. & Schuller, B. (2013). SenticNet 3: A common and common-sense knowledge base for cognition-driven sentiment analysis. *Proceedings of the AAAI Conference on Artificial Intelligence, 27*(1). https://doi.org/10.1609/aaai.v28i1.8928

[10] Zhang, Y., Jin, R. & Zhou, Z. H. (2018). Understanding AI-Enabled Sentiment Analysis: Techniques, Applications, and Challenges. *Annals of Data Science, 5*(2), 155-177.
[11] Maas, A. L., Daly, R. E., Pham, P. T., Huang, D., Ng, A. Y. & Potts, C. (2011). Learning Word Vectors for Sentiment Analysis. *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, 142-150.
[12] Liu, B. (2012). Sentiment Analysis and Opinion Mining. *Synthesis Lectures on Human Language Technologies, 5*(1), 1-167. https://doi.org/10.1007/978-3-031-02145-9
[13] Duwairi, R. & Abushaqra, F. (2021). Syntactic-and morphology-based text augmentation framework for Arabic sentiment analysis. *PeerJ Computer Science*, 7, e469. https://doi.org/10.7717/peerj-cs.469
[14] Mahany, A., Khaled, H., Elmitwally, N. S., Aljohani, N. & Ghoniemy, S. (2022). Negation and Speculation in NLP: A Survey, Corpora, Methods, and Applications. *Applied Sciences, 12*(10), 5209. https://doi.org/10.3390/app12105209
[15] Wang, J., Yue, K. & Duan, L. (2023). Models and Techniques for Domain Relation Extraction: A Survey. *Journal of Data Science and Intelligent Systems*. https://doi.org/10.47852/bonviewJDSIS3202973
[16] Maña López, M. J. & Cruz Díaz, N. P. (2019). Negation and Speculation Detection. *Negation and Speculation Detection*, 1-105. https://doi.org/10.1075/nlp.13
[17] Syrett, C., Ebbels, S., Nicoll, H., Davies, K., White, S. & Zuniga-Montanez, C. (2023). Supporting adjective learning by children with Developmental Language Disorder: Enhancing metalinguistic approaches. *International Journal of Language & Communication Disorders, 58*(2), 629-650. https://doi.org/10.1111/1460-6984.12792
[18] Cai, R., Zhang, X. D. & Wang, H. F. (2016). Bidirectional recurrent convolutional neural network for relation classification. *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*, 1, 756-765. https://doi.org/10.18653/v1/P16-1072
[19] Li, F., Zhang, M. S., Fu, G. H. & Ji, D. H. (2017). A neural joint model for extracting bacteria and their locations. *Proceedings of the 21st Pacific-Asia Conference on Knowledge Discovery and Data Mining*, 15-26. https://doi.org/10.1007/978-3-319-57529-2_2
[20] Zheng, S. C., Wang, F., Bao, H. Y., Hao, Y. X., Zhou, P. & Xu, B. (2017). Joint extraction of entities and relations based on a novel tagging scheme. *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics*, 1, 1227-1236. https://doi.org/10.18653/v1/P17-1113
[21] Duwairi, R. & Abushaqra, F. (2021). Syntactic-and morphology-based text augmentation framework for Arabic sentiment analysis. *PeerJ Computer Science*, 7, e469. https://doi.org/10.7717/peerj-cs.469
[22] Aoumeur, N. E., Li, Z. & Alshari, E. M. (2023). Improving the polarity of text through word2vec embedding for primary classical Arabic sentiment analysis. *Neural processing letters*, 1-16. https://doi.org/10.1007/s11063-022-11111-1
[23] Cumaoğlu, İ. (2022). Sentiment analysis of Arabs in Turkey using deep learning on social media data. *Doctoral dissertation*.
[24] Refai, D. S. (2022). A New Data Augmentation Approach for Boosting Arabic Text Classification Performance Using Transformers and Similarity Measures. *Doctoral dissertation*, Princess Sumaya University for Technology (Jordan).
[25] Díaz, N. P. C. & López, M. J. M. (2019). *Negation and speculation detection (Vol. 13)*. John Benjamins Publishing Company.

Authors' contacts:

Pinky Yadav
(Corresponding author)
FET, MRIIRS (Manav Rachna International Institute of Research and Studies),
Sector – 43, Aravalli Hills, Delhi – Surajkund Road, Faridabad – 121004,
(Haryana), India
pinky.yadavphd@gmail.com

Indu Kashyap, Dr.
FET, MRIIRS (Manav Rachna International Institute of Research and Studies),
Sector – 43, Aravalli Hills, Delhi – Surajkund Road, Faridabad – 121004,
(Haryana), India

Bhoopesh Bhati, Dr.
Indian Institute of Information Technology, Sonepat (IIIT, Sonepat),
IIT Delhi Techno Park, Rajiv Gandhi Education City, Rai, Sonipat, Haryana – 131001,
India