FUZZY-RULES GRADING AND SENTIMENT COMPOUND BASED DICTIONARY FOR ESTIMATING THE DEGREE OF SENTIMENT POLARITY

Lijimol George (1), P Sumathy (2), A Vadivel (3)*

(1) (2) Department of Computer Science, Bharathidasan University, Khajamalai, Tiruchirappalli, Tamil Nadu; (3) Department of Artificial Intelligence and Data Science, Drone Technologies Lab, MURTI Research Centre, GITAM School of Technology, GITAM Deemed to be University, Bangalore,
India

* Corresponding Author, e-mail: sumathy.p@bdu.ac.in

Abstract: This paper proposes a fuzzy rule-based sentiment analysis framework for airline reviews. A domain-specific sentiment compound dictionary is developed by incorporating intensifiers, negators, and sentence position indicators. Sentences are classified using decision trees constructed from Parts of Speech (PoS) tagged components, resulting in eight distinct sentiment classes. Proposed to handle ambiguity in sentiment class boundaries, Fuzzy membership functions are applied, enabling graded polarity scoring. The model is evaluated on Air India and Twitter US Airline Sentiment datasets and achieved 88% and 91% as accuracy respectively. Comparative results with VADER, GINS and BERT based models confirm the proposed method's performance is encouraging.

Key words: sentiment analysis, sentiment dictionary, airline review, airline tweets, fuzzy functions

1. INTRODUCTION

Natural Language Processing (NLP) is one of the emerging areas to understand, analyze and interpret human languages. Recent research highlights its effectiveness across diverse applications, including multilingual authorship attribution [1], spam email detection [2], and sentiment analysis [3-4]. Sentiment Analysis aims to analyze and find the emotion in the text/speech, etc. The research intense increases with volume of online data, content in social media and reviews that are related to online shopping [5]. However, understanding and measuring the sentiments from large volume of online data is challenging. Traditional methods often rely on sentiment lexicons or machine learning techniques. However, they struggle to handle linguistic nuances and contextual sentiment effectively. To address this limitation, in this paper, sentiment compound-based dictionary is developed that leverages fuzzy membership functions and rules to estimate the degree of sentiment polarity. Sentiment analysis is tested on airline reviews by integrating a domain-specific sentiment polarity dictionary with advanced linguistic and computational methods. Combination of domain-specific dictionary design, hierarchical linguistic modelling, and fuzzy logic represents a significant advancement in sentiment analysis techniques.

The novelty of this paper is the proposal of a sentiment compound-based dictionary framework that can be tuned to a specific application domain, say sentiments of passengers of airlines based on their reviews. Rules through decision tree and fuzzy membership functions are derived to estimate nuanced degrees of sentiment polarity of the review sentences. Domain-specific linguistic constructs including negators, intensifiers, and their positional contexts are uniquely integrated using fuzzy membership functions to handle ambiguity and imprecision effectively. The proposed fuzzy membership functions capture nuanced variations in sentiment intensity, which conventional lexicon-based and machine learning approaches typically overlook. The imprecision and ambiguity inherent in natural language are distinctly addressed, providing a graded sentiment polarity rather than traditional discrete categories. Key contributions include:

- *Domain-Specific Dictionary Construction:* The paper introduces a sentiment polarity dictionary framework that can be tailored any application domain, enhancing the relevance and accuracy of sentiment analysis for the domain, say sentiments of the passengers through their review sentences.
- *Integration of PoS Tagging and Rules of Decision Tree for Sentiment Classification*: Sentences are decomposed into components using PoS tagging, and if-then-else rules are constructed using decision tree to ensure a structured and interpretable method for sentiment classification.
- *Capturing imprecision in the Boundary of the Sentiment Classes with Fuzzy Membership Functions*: Propose fuzzy membership functions that measure the degree of sentiment polarity to address ambiguity and impreciseness in sentiment boundaries.

This paper arranges the manuscript as follows, section 2 discusses the contemporary research, section 3 present the fuzzy functions and sentiment component-based dictionary. The result is analysed in section 4 and conclusion and further wore is presented in section 5.

2. RELATED WORKS

This section presents the review of well-known sentiment analysis methodologies such as lexicon-based, machine learning based, and the combined approaches.

The dictionaries based on lexicon use sentiments that are defined already for assigning score for polarity to option-bearing terms. While general-purpose lexicons are widely used, domain-specific resources often yield better accuracy [6-7]. These methods typically lack contextual awareness, as they evaluate isolated terms without considering their placement or syntactic roles. Recent studies focus on constructing tailored lexicons using both manual and automated methods, to improve domain relevance. Fuzzy sentiment models have gained prominence for managing linguistic ambiguity effectively [8-9]. They enhance lexicon-based methods by enabling graded polarity assignment using fuzzy sets. Recent advancements integrate fuzzy logic with deep learning models such as BERT and CNN frameworks like LeBERT and MultiLexANFIS [10-11]. While fuzzy lexicon-based sentiment analysis offers benefits in handling linguistic nuances, it is essential to consider its limitations in terms of contextual understanding, training data availability, and computational efficiency. Hybrid approaches combine lexicons with machine learning or deep learning models to improve interpretability and accuracy. Several ML approaches have been used for sentiment and emotion analysis such as Random Forest (RF), Decision Tree

(DT), Bayesian Network (BN) and Support Vector Machine (SVM) [12-16]. However, they fail to highlight the important role of emotional words. Deep Learning models like CNN, LSTM, and attention-based architectures outperform traditional ML methods and struggle with explainability and data demands [17-19]. Hybrid dictionaries like GINS [22] and Senti-N-Gram combine n-gram analysis with fuzzy logic and intensifier weights to refine sentiment scoring [20-21]. These methods balance rule-based clarity and data-driven precision, especially in domain-specific applications.

It is imperative from the reviews presented above; almost all the approaches are having challenges in identifying the contextual sentiment. However, the degree of sentiment and emotions present in sentence has been missed by most of these methods. This issue can been handled by constructing a Decision Tree based on the tokens that are extracted from the sentences. The if-then-else construct of the tree formulates eight distinct classes of positive, negative and neutral. The degree of sentiment and emotions present in the sentences are captured through eight fuzzy functions. Sentiment compound-based dictionary is developed for estimating the degree of sentiment polarity. The effectiveness of the dictionary is evaluated on Air India service and US Airline service datasets and observed that the performance is encouraging.

3. PROPOSED APPROACH

A sentiment compound-based dictionary, along with fuzzy membership functions, is proposed for measuring the level of sentiment polarity, say positive, negative, neutral etc., of a review submitted by airline passenger, and they are presented in detail below.

3.1. Preprocessing Phase

Two airline review datasets are used for evaluation, and they are evaluated on 'Air India Review Dataset (AIRD)', created based on reviews available on the Airline Quality website, and 'Twitter US Airline Sentiment Dataset (TUSD)' from Kaggle. The AIRD dataset is created using a web scraping tool to extract reviews from the Airline Quality website 1201reviews. The user details, time and plan of travel, title of review, rate for seat comfort, food and entertainment are considered as parameters. However, the proposed work has considered details of the user, title and comments of the reviewer. Review comments spanning from January 2023 to January 2024, are collected. The 'Twitter US Airline Sentiment Dataset (TUSD)' from Kaggle, consists of 14,640 tweets collected through experimental data collection. These tweets are about six major American airlines: American Airlines, United, US Airways, Southwest, Delta, and Virgin America. The tweets are labelled as positive, negative, and neutral. Though there are lot parameters available with the review, Tweet- id, text and sentiment label is considered for processing. The URL, Quotes, HML and other tags, spam, empty fields, special characters etc present in the review are considered as noise removed at the pre-processing stage and stored as repository for future processing.

The Figure 1 presents various phases of proposed approach. The overall approach is divided into five phases, and they are Data collection, preprocessing, feature extraction, sentence classification and grading sentiment polarity based on intensity. The dictionary that uses compound of terms for inferring the meaning/sentiment of a sentence is basic concept of the approach. The sentiment term is identified, and it is considered as pivotal along with negators, intensifiers and so on. The organizing principles that are being used are the

Decision Tree and Fuzzy logic. The Decision Tree is constructed based on the Domain Term and captures the effect of sentiment compound. The details are explained in the sub sections.

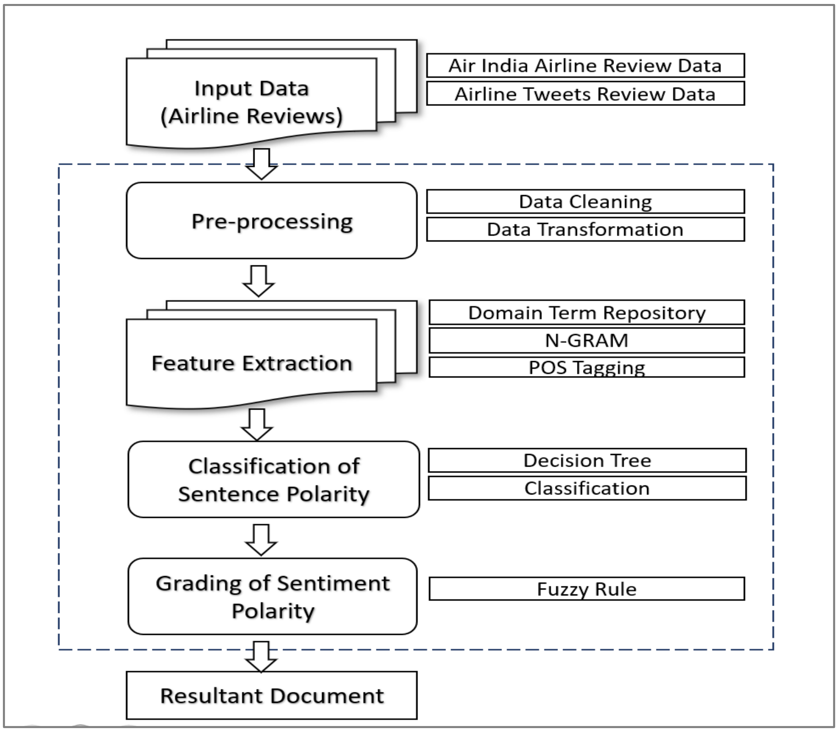


Figure 1. *Framework of proposed approach*

3.2. Creating Domain Term Repository for Feature Extraction

In this paper, each airline review is considered as a group of sentences. Consider a sentence '*S*' and it contains various components, and is represented as

$$S = \{comp_1, comp_2, \dots, comp_i, \dots comp_n\} \quad (1)$$

In Eq. (1), $comp_i$ represents various components of a sentence. Sentences are generally processed to understand their characteristics. Based on this understanding, they are used for domain-specific applications, with each component ($comp_i$) mapped to Part of speech (PoS) tags. The characteristic of the sentence depends on the PoS tags. The review on Airline is fetched using web scraping tool and the tags, URLs etc. associated with review are removed in the pre-processing stage. The domain term repository is created, and the dictionary is developed using the Domain Terms, intensifiers, and negators from corpus as an alternative to human annotators for creating a domain-specific dictionary automatically. Domain terms are identified through POS tagging and NLP techniques from the airline review corpus. Intensifiers are considered from Brooke's thesis and negators are sourced from the Cambridge Grammar reference on negation. The list of intensifiers, negators, and domain-specific terms is combined and then generated the domain-dependent, n-gram-based sentiment dictionary. Finally, the features are extracted in the form of PoS tags and N-grams. The sentiment polarity is measured to capture the degree of positive, negative and neutral sentiment. Since airline is the domain of analysis, the terms related to the domain is denoted as *Domain Term (DT)*. As a result, the Sentence (*S*) is represented in terms of *DT*, the term immediately appearing before or after with *DT* as *DT*$\pm$*1* and the terms appearing after certain distance is (*DT*$\pm$*n*). Also, Verb, Noun, Adjective, Intensifiers, Negators, Connectors, etc are identified as the components of the sentences. All these components are used for understanding the polarity of sentence sentiment. In this work, the stand ford NIP tool is used for tagging the sentences, since the component-based sentence are realized as N-gram.

3.3. Classification of Sentence Polarity

The PoS tagged features are considered along with Intensifiers (*INTF*) and Negations (*NG*) for classifying the sentence polarity. Sentences/reviews of a user is oriented as positive, negative, or neutral and these are represented in Eqs. (2)-(4)

$$l^P: S_P \in S \rightarrow \{true\ |false\} \tag{2}$$
$$l^N: S_N \in S \rightarrow \{true\ |false\} \tag{3}$$
$$l^{NL}: S_{NL} \in S \rightarrow \{true\ |false\} \tag{4}$$

In Eqs. (2)-(4) , l^P, l^N *and* l^{NL} represents the labelling of positive, negative and neutral sentences. The S_P, S_N and S_{NL} are set of sentences labelled as positive, negative and neutral. The Eqs. 2-4 can be visualized in terms of PoS tags, and the relationships between PoS tags are shown in Figure 2. The POS feature of the DT term is used for fixing the root node of the decision tree. The root mode is always noun/verb Explained with DTIN and DTIN at second level of Discussion Tree. With the presence and absence of Negators. The Level 2 represents the presence or absence of Negators in the sentence at $DT \pm n$. The final level validates the presence of Intensifiers in the sentence *at* $DT \pm n$. It is observed from the Figure 2 that there are eight classes of sentences in which four sentences are /positive (s_P), three are negative (s_N) and one is neutral(s_{NL}).

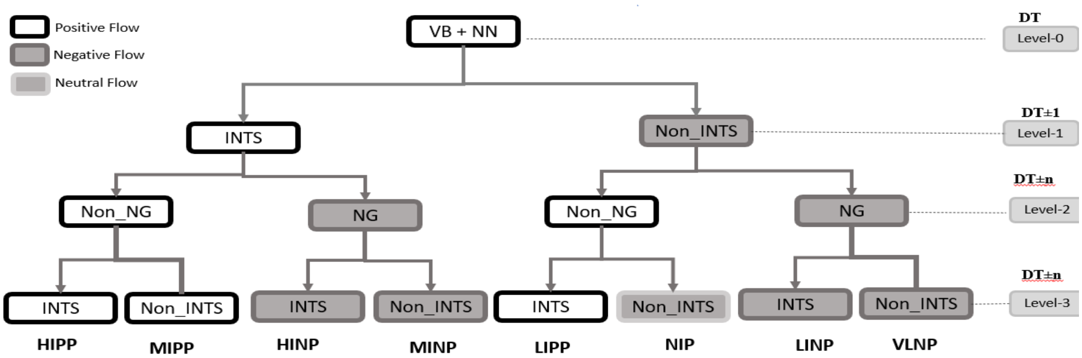


Figure 2: Classification of Sentences using PoS tags

The level 2 of the tree determines that whether a sentence is negative or positive. As a result, Eqs. (2)-(4) can be re-written as follows.

$$l^P: S_P \in S \rightarrow \{< DT = VB||DT = NN, DT \pm 1 = INTS\ ||\ DT \pm 1 = Non\_INTS\ , DT \pm n = Non\_NG\ \ >\} = True \tag{5}$$
$$l^N: S_N \in S \rightarrow \{< DT = VB||DT = NN, DT \pm 1 = INTS\ ||\ DT \pm 1 = Non\_INTS\ , DT \pm n = NG\ \ >\} = True \tag{6}$$

$$l^{NL}: S_{NL} \in S \rightarrow \{< DT = VB || DT = NN, DT \pm 1 = Non\_INTS \ , DT \pm n = Non\_NEG \ , DT \pm n = Non\_INTS >\} = True \quad (7)$$

The sentences clustered using Eqs. (5) - (7) are in the form of predicates and defines the hard boundary between classes. However, the classes are overlapping intense of their boundaries and thus there is inconsistency in clearly defines it. The if-then-else construct from the decision tree generates eight distinct classes and out of which three are positive, one is neutral and the remaining four are negative. The detail on each class is presented in Table 1. The Table explains the rules to derive the intensified polarity of High Intensified Positive Polarity *(HIPP),* Medium Intensified Positive Polarity *(MIPP),* Low Intensified Positive Polarity *(LIPP),* Neutral Intensified Polarity *(NIP),* Very Low Negative Polarity *(VLNP),* Low Intensified Negative Polarity *(LINP),* Medium Intensified Negative Polarity *(MINP) and* High Intensified Negative Polarity *(HINP)* classes.

Table 1. Name of the Classes and Rules

Details	***Rules***
High Intensified Positive Polarity (HIPP)	$(ET \pm 1 = INTF)$^$(ET \pm n = Non\_NG)$^$(ET \pm n \ = INTF)$
Medium Intensified Positive Polarity (MIPP)	$(ET \pm 1 = INTF)$^$(ET \pm n = Non\_NG)$^$(ET\pm = Non\_INTF)$
Low Intensified Positive Polarity (LIPP)	$(ET \pm 1 = Non\_INTF)$^$(ET\pm = Non\_NG)$^$(ET \pm n = INTF)$
Neutral Intensified Polarity (NIP)	$(ET \pm 1 = Non\_INTF)$^$(ET\pm = Non\_NG)$^$(ET\pm = Non\_INTF)$
Very Low Negative Polarity (VLNP)	$(ET \pm 1 = Non\_INTF)$^$(ET \pm n = NG)$^$(ET\pm = Non\_INTF)$
Low Intensified Negative polarity (LINP)	$(ET \pm 1 = Non\_INTF)$^$(ET \pm n = NG)$^$(ET \pm n = INTF)$
Medium Intensified Negative Polarity (MINP)	$(ET \pm 1 = INTF)$^$(ET \pm n = NG)$^$(ET \pm n = Non\_INTF)$
High Intensified Negative Polarity (HINP)	$(ET \pm 1 = INTF)$^$(ET \pm n = NG)$^$(ET \pm n = INTF)$

3.4. Grading the Sentiment Polarity based on Fuzzy rules

Sentences are classified as positive, negative, and neutral classes based on the intensity of a sentence. In this section, the grade of each class is calculated using fuzzy functions for measuring the degree of polarity and reduce the ambiguity present in the boundary of the class. The sentences that share common characteristics can be modelled based on fuzzy linguistic variables and fuzzy functions are given below.

$$\mu_{HIPP}(x) = e^{\{-(x_{HIPP}-1)^2\}} \quad (8)$$

μ_{HIPP} the membership function defined on sentences *HIPP* and x_{HIPP} $(w1 * A_{IPP}) * \{T|F\} + (w2 * B_{IPP}) * \{T|F\} + (w3 * C_{IPP}) * \{T|F\} + (w4 * D_{IPP}) * \{T|F\}$ is the parameter assigned to each sentence *x*. In x_{HIPP}, *w1, w2, w3, w4* are heuristic weight, empirically derived based on the characteristic of the *DT* and the values of the weights are 0.1,0.4,0.3 and 0.2 respectively. A_{IPP}, B_{IPP}, C_{IPP} and D_{IPP} represents the if-then-else construct for positive sentence that are derived in terms of rules from the *DT*. As a result,

$$A_{IPP} = \{(s = DT \ \wedge DT \in NN|VB)\} \quad (9)$$

$$B_{IPP} = \{(ET \ \pm \ 1 = INTS\} \quad (10)$$

$$C_{IPP} = \{(ET \ \pm \ n = Non\_NG \ \} \quad (11)$$

$$D_{IPP} = \{(ET \ \pm \ n = INTS \ \} \quad (12)$$

The value of x_{HIPP} ranging from 0 to 1. The fuzzy function for *MIPP* is

$$\mu_{MIPP}(x) = e^{-\{(x_{MIPP}-1)^2\}} \quad (13)$$

In Eq. (13), $X_{MIPP} = (w1 * A_{IPP}) * \{T|F\} + (w2 * B_{IPP}) * \{T|F\} + (w3 * C_{IPP}) * \{T|F\} + (w4 * D_{IPP}) * \{T|F\}$ The value of A_{IPP}, B_{IPP}, C_{IPP} & D_{IPP} are defined in Eqs (9)-(12) that are the if-then-else construct derived from *DT* and weight values are same as mentioned in Eq.(8). The Eqs. (14) and (15) are the fuzzy function for *LIPP* and *NIP*.

$$\mu_{LIPP}(x) = e^{-\{(x_{LIPP}-1)^2\}} \tag{14}$$

$$\mu_{NIP}(x) = e^{-\{(x_{NNP}-1)^2\}} \tag{15}$$

In Eqs. (14) and (15)

$$X_{LIPP} = (w1 * A_{IPP}) * \{T|F\} + (w2 * B_{IPP}) * \{T|F\} + (w3 * C_{IPP}) * \{T|F\} + (w4 * D_{IPP}) * \{T|F\}$$

$$X_{NIP} = (w1 * A_{IPP}) * \{T|F\} + (w2 * B_{IPP}) * \{T|F\} + (w3 * C_{IPP}) * \{T|F\} + (w4 * D_{IPP}) * \{T|F\}$$

The polarity of positive sentences ranging from 0 to 1. The grading of sentiment polarity of positive sentences is explained in the Table 2. For example, all four levels are true for *HIPP* sentence, so that x_{HIPP} and $\boldsymbol{\mu}_{HIPP}$ are 1. Similarly, the condition *ET*$\pm$*n =INTS* is false and other conditions are true for *MIPP* sentences, hence the $\boldsymbol{\mu}_{\text{MIPP}}(x)$ is 0.961. In case of *LIPP* sentences, *ET*$\pm$*1=INTS* is false and hence the $\boldsymbol{\mu}(x)$ is 0.852. Accordingly, the output of *NIP* sentences results in 0.697. It is noticed from the Table 2 that the value of $\mu(x)$ is changing based on the rules captured through A_{IPP}, B_{IPP}, C_{IPP} and D_{IPP}. The change in value captures the degree of positive polarity of sentences.

Table 2: Grading of Sentiment Polarity of Positive Sentences

	S=DT	***ET±1=INTS***	***ET±n=Non_NG***	***ET±-n=INTS***	***µ (x)***
	A_{IPP}	$\boldsymbol{B}_{IPP}$	$\boldsymbol{C}_{IPP}$	$\boldsymbol{D}_{IPP}$	
HIPP	*0.1*	*0.4*	*0.3*	*0.2*	*1*
MIPP	*0.1*	*0.4*	*0.3*	*0.0*	*0.961*
LIPP	*0.1*	*0.0*	*0.3*	*0.2*	*0.852*
NIP	*0.1*	*0.0*	*0.3*	*0.0*	*0.697*

The fuzzy function for negative classes *VLNP, LINP, MINP, HINP* are presented below Eqs. (16) to (19).

$$\mu_{VLNP}(x) = -\left(e^{-\{(x_{NP}-1)^2\}}\right) \tag{16}$$

$$\mu_{LINP}(x) = -(e^{-\{(x_{LINP}-1)^2\}}) \tag{17}$$

$$\mu_{MINP}(x) = -\,(e^{-\{(x_{MINP}-1)^2\}}) \tag{18}$$

$$\mu_{HINP}(x) = -\,(e^{-\{(x_{HINP}-1)^2\}}) \tag{19}$$

In Eqs. (16)-(19), the x_{VLNP}, x_{LINP}, x_{MINP}, x_{HINP} are calculated as given below.

$$X_{VLNP} = (w1 * A_{INP}) * \{T|F\} + (w2 * B_{INP}) * \{T|F\} + (w3 * C_{INP}) * \{T|F\} + (w4 * D_{INP}) * \{T|F\}$$

$$X_{LINP} = (w1 * A_{INP}) * \{T|F\} + (w2 * B_{INP}) * \{T|F\} + (w3 * C_{INP}) * \{T|F\} + (w4 * D_{INP}) * \{T|F\}$$

$$X_{MINP} = (w1 * A_{INP}) * \{T|F\} + (w2 * B_{INP}) * \{T|F\} + (w3 * C_{INP}) * \{T|F\} + (w4 * D_{INP}) * \{T|F\}$$

$$X_{HINP} = (w1 * A_{INP}) * \{T|F\} + (w2 * B_{INP}) * \{T|F\} + (w3 * C_{INP}) * \{T|F\} + (w4 * D_{INP}) * \{T|F\}$$

The A_{INP}, B_{INP}, C_{INP} and D_{INP} represents the if-then-else construct for negative sentences that are derived in terms of rules from the *DT*. As a result,

$$A_{INP} = \{(s = DT \wedge DT \in NN|VB)\} \quad (20)$$

$$B_{INP} = \{(ET \pm 1 = INTS\} \quad (21)$$

$$C_{INP} = \{(ET \pm n = NG\} \quad (22)$$

$$D_{INP} = \{(ET \pm n = INTS\} \quad (23)$$

The grade value of negative sentences is in the range of [-1,0]. In the Table 3, the grade values of the ***μ****(x)* for various negative sentences are presented. It is observed that the grade value of *HINP* start with -1 and its value is keep reducing. This captures the degree of negative polarity of a sentence.

Table 3: Grading of Sentiment Polarity of Negative Sentences

	S=DT	**ET+-1=INTS**	**ET+-n=NG**	**ET+-n=INTS**	***μ*(x)**
	A_{INP}	$\mathbf{B}_{INP}$	$\mathbf{C}_{INP}$	$\mathbf{D}_{INP}$	
HINP	0.1	0.4	0.3	0.2	-1.0
MINP	0.1	0.4	0.3	0.0	-0.961
LINP	0.1	0.0	0.3	0.2	-0.852
VLNP	0.1	0.0	0.3	0.0	-0.697

It is observed from Table 2 and 3 that the degree/grade of positive and negative sentences are captured by the fuzzy functions. In Table 2, the grading of *HINP* is 1 and grading in Table 3 is -1. It is imperative that the value of ***μ****(x)* is ranging from -1 to +1. The rest of the sentence in Table 2 are graded such that the values are minute to handle the overlapping the characteristics all positive category sentences. Similarly, the grading value of negative sentences are capturing the negatively polarized sentences. It is observed from Figure 2 that the fuzzy grade for positive sentences have the range of values form 0 to 1 and -1 to 0 for negative sentences. The range of *HIPP* is 0.961 to 1, 0.852 to 0.960 for *LIPP*, 0.697 to 0.851 for *LIPP* and 0 to 0.696 is for *NIP* sentence classes. Similarly, the fuzzy grade membership values for HIPP class are 0.961 to -1, -0.852 to -0.960 to *MINP*, -0.697 to -0.851 to *LINP* and 0 to -0.696 to *VLNP* respectively. The comprehensive pseudocode underlying the proposed approach is detailed in Algorithm 1, providing a clear outline of the step-by-step procedures employed.

Algorithm 1: Fuzzy_Sentiment()

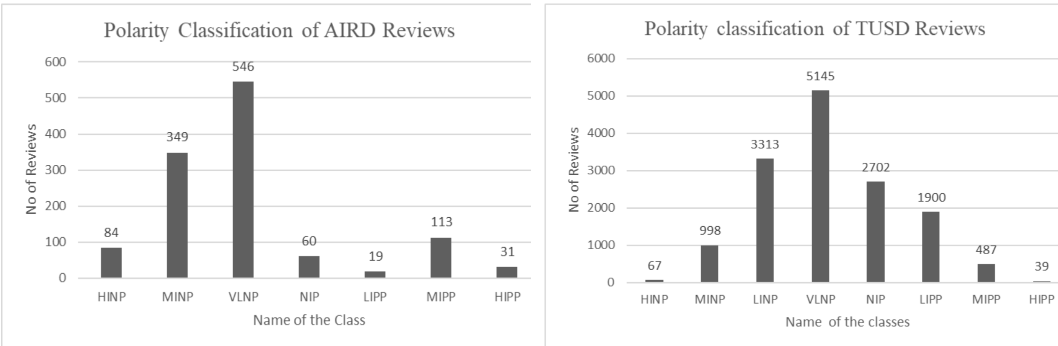
```
Input: Airline Review Dataset
Output: Grade of Sentiment Polarity
D = Set of Review Documents
#Step 1: Data Pre-processing and Feature Extraction

do
for each  di € D   (pre-process as per 3.2)
S=Sentences in di
Extract Si € S
            comp =  decompose (Si )
            Construct Domain Term Repository (as per 3.3)
            Extract features (POS Tags, N-grams)
end do
```

#Step 2: Classification of sentence polarity
lP, lN and lNL = Predicates for positive, negative and neutral sentences
Define SP , SN and SNL -> derive the if-then-else rule and classify the S (as per 3.4)

#Step 3: Grading the Sentiment Polarity
Define the Fuzzy membership (Eqn. 8-23) (as per 3.5)
Grade the Sentiment class of Si € S

4. EXPERIMENTAL RESULTS AND DISCUSSION

The experiment is performed on two airline review dataset and result are explained below. There are 1201 AIRD reviews that are labelled as positive or negative based on the 'Recommendation' field of the review document. The result of classification of AIRD reviews is shown in the Figure 3 (a) with eight classes based on the sentence polarity. Based on this result, 45% of the reviews are classified to *VLNP* and 29% of documents are belonging to *MINP*. The performance of TUSD reviews are also evaluated and sentence polarity classification is applied to the dataset and the result is shown in the Figure 3(b). The sentiment polarity is graded using eight fuzzy functions and result is evaluated. The performance is measured in terms of precision, Recall, F1 score and Accuracy and is presented in the Figure 4. The performance on AIRD is presented in Fig. 4(a) and Fig. 4(b) presents the performance on TUSD reviews. The accuracy on AIRD is 88% and 91% on TUSD dataset, respectively. Similarly, Precision, Recall and F1-score on AIRD and TUSD is presented and found to be encouraging.



(a) (b)

Figure 3. Sentiment Polarity Classification (a) AIRD Reviews (b) TUSD Reviews

	precision	recall	f1-score	support
0.0	0.92	0.95	0.93	314
1.0	0.56	0.47	0.51	47
accuracy			0.88	361
macro avg	0.74	0.71	0.72	361
weighted avg	0.88	0.88	0.88	361

accuracy= 0.8836565096952909

(a)

	precision	recall	f1-score	support
0	0.89	0.92	0.90	1647
1	0.93	0.91	0.92	2025
accuracy			0.91	3672
macro avg	0.91	0.91	0.91	3672
weighted avg	0.91	0.91	0.91	3672

0.9120370370370371

(b)

Figure 4. Classification Accuracy (a) AIRD Reviews (b) TUSD Reviews

The accuracy of the proposed approach is assessed with the Friedman test and Nemenyi post-hoc analysis and both have p-values above the 0.05 threshold. The result is provided in Table 4. Consequently, no significant performance differences are found across datasets, confirming the method’s stable and encouraging accuracy.

Table 4: Friedman test and Nemenyi post-hoc results

```
Friedman test statistic: 6.0, p-value: 0.11161022509471268

Nemenyi post-hoc test results:
                      VADER      GINS      BERT  Proposed_Approach
VADER              1.000000  0.865978  0.408050           0.092643
GINS               0.865978  1.000000  0.865978           0.408050
BERT               0.408050  0.865978  1.000000           0.865978
Proposed_Approach  0.092643  0.408050  0.865978           1.000000
```

Performance Comparison. The performance of the proposed approach is further consolidated by comparing it with VADER, GINS dictionaries and Bidirectional Encoder Representations from Transformers (BERT) and is presented in Table 5.

Table 5: Performance comparison of the proposed approach with contemporary approaches

Dataset	***Accuracy***			
	VADER	***GINS***	***BERT***	***Proposed Approach***
Dataset-1 Air India Review Dataset (AIRD)	*0.63*	*0.81*	*0.85*	*0.88*
Dataset-2 Twitter US Airline Sentiment Dataset (TUSD)	*0.72*	*0.87*	*0.90*	*0.91*

The accuracy of VADER is low compared all methods on AIRD and TUSD dataset. The accuracy of the proposed approach is 0.88 and 0.91 on AIRD and TUSD dataset respectively and the performance of proposed method is encouraging.

5.CONCLUSION AND FUTURE WORK

This paper introduces a novel sentiment dictionary designed to handle sentiment compounds to classify distinct sentiment classes using domain-specific terms, negators, and the co-occurrence of other sentence components and generated eight fuzzy functions to delineate boundaries between these classes. The dictionary is evaluated on AIRD and TUSD datasets. The accuracy on AIRD and TUSD datasets 88% and 91% respectively. Performance of the proposed dictionary is evaluated on VADER, GINS and BERT and found performance improvements. As future work, the dictionary could be dynamically adapted in real time to reflect evolving language usage and domain-specific trends. Integrating advanced contextual embedding techniques, such as transformer-based models, may further enhance the system's performance.

6.ACKNOWLEDGEMENT

This research was not supported by any funded Project.

7. REFERENCES

[1] Arta Misini, Arbana Kadriu, Ercan Canhasi. Authorship classification techniques: Bridging textual domains and languages. *International Journal on Information Technologies and Security*, vol.16, no.1, 2024, pp. 27–38. https://doi.org/10.59035/UKBE1226.

[2] Bandar Alshawi, Amr Munsh, Majid Alotaibi, Ryan Alturki, Nasser Allheeib. Classification of SPAM mail utilizing machine learning and deep learning techniques. I*nternational Journal on Information Technologies and Security*, vol.16, no.2, 2024, pp. 71–82. https://doi.org/10.59035/FPKO7430.

[3] M. Cobo, I. J. Pérez, A. Velez-Estevez, F. J. Cabrerizo. Uncovering the conceptual evolution of sentiment analysis research field during the period 2017–2021. *IEEE International Conference on Evolving and Adaptive Intelligent Systems (EAIS)*, 2022, pp. 1–8. https://doi.org/10.1109/eais51927.2022.9787534.

[4] S. Sushma, S. K. Nayak, M. V. Krishna. A Comprehensive review of sentiment analysis: Trends, challenges, and future directions. *IEEE International Conference on Data, Information, and Communication (ICDICI),* 2024, pp. 1175–1181. https://doi.org/10.1109/icdici62993.2024.10810919.

[5] Rimpy, Dhankhar, A., Dhankhar, A. Sentimental analysis of social networks: A comprehensive review (2018–2023). *Multidisciplinary* Reviews, vol.7, no.7, 2024: 2024126. https://doi.org/10.31893/multirev.2024126

[6] Q. He. Recent works for sentiment analysis using machine learning and lexicon based approaches. *IEEE International Conference on Advanced Engineering, Management and Computational Sciences (AEMCSE)*, 2022, pp. 422–426. https://doi.org/10.1109/AEMCSE55572.2022.00090.

[7] A. Rufaida, A. Permanasari, N. Setiawan. Lexicon-based sentiment analysis using inset dictionary: A systematic literature review. *International Conference on Applied Engineering, ICAE 2022*, Batam, Indonesia, October 2022 https://doi.org/10.4108/eai.5-10-2022.2327474

[8] Y. Wang, F. Subhan, S. Shamshirband, M. Z. Asghar, I. Ullah, A. Habib, F. M. Kundi. Fuzzy-based sentiment analysis system for analyzing student feedback and satisfaction. *Computers, Materials & Continua*, vol.62, 2020, pp. 631–655.

[9] Vashishtha. Srishti , Gupta. Vedika , Mittal. Mamta. Sentiment analysis using fuzzy logic: A comprehensive literature review. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 2023. https://doi.org/10.1002/widm.1509.

[10] E. Gerasimenko, V. Stetsenko. Sentiment analysis of text reviews using tone dictionaries and fuzzy set cardinality. Izvestiya of the Southern Federal University. *Engineering Sciences,* vol. 5, pp. 106–116, 2022. doi: 10.18522/2311-3103-2022-5-106-116.

[11] M. James. Sentiment analysis on text reviews using Lexicon Selected-Bert Embedding (LeBERT) model with convolutional neural network. *Research Square*, 2022. https://doi.org/10.21203/rs.3.rs-2330887/v1.

[12] Veneta Yosifova. Application of the compositional machine learning dynamic classification algorithm for natural language sentiment analysis. *International Journal on Information Technologies and Security*, vol.16, no.4, 2024, pp. 61–72. https://doi.org/10.59035/AESI6882.

[13] M. S. M. Prasanna, S. G. Shaila, A. Vadivel, M. Mittal. Sarcastic sentiment detection and polarity classification of tweets using supervised learning. *Lecture Notes in Networks and Systems, Springer*, 2022. https://doi.org/10.1007/978-981-16-7118-0_15.

[14] M. S. M. Prasanna, S. G. Shaila. Polarity classification on Twitter data for classifying sarcasm using clause pattern for sentiment analysis. *Multimedia Tools and Applications*, vol.82, 2023, pp. 32789–32825. https://doi.org/10.1007/s11042-023-14909-w.

[15] S. Annadurai, M. Arock. Real and fake emotion detection using enhanced boosted support vector machine algorithm. *Multimedia Tools and Applications,* vol.82, 2023, pp. 1333–1353. https://doi.org/10.1007/s11042-022-13210-6.

[16] A. Swaminathan, A. Vadivel, M. Arock. FERCE: Facial expression recognition for combined emotions using FERCE algorithm. *IETE Journal of Research*, vol.68, 2022, pp. 3235–3250. https://doi.org/10.1080/03772063.2020.1756471.

[17] Lujin Huang. Deep learning for text sentiment analysis: A survey. *Applied and Computational Engineering*, vol.104, no.1, 2024, pp. 134–138. https://doi.org/10.54254/2755-2721/104/20241153.

[18] Nikhil Suryawanshi. Sentiment analysis with machine learning and deep learning: A survey of techniques and applications. *International Journal of Science and Research Archive*, vol.12, no.2, 2024, pp. 005–015. https://doi.org/10.30574/ijsra.2024.12.2.1205.

[19] Arjun Kumar, Nitesh Kumar Gupta, Anurag Shrivastava. A critical review on sentiment analysis based on deep learning techniques. *International Journal for Multidisciplinary Research*, vol.6, no.5, 2024. https://doi.org/10.36948/ijfmr.2024.v06i05.28572.

[20] C. J. Hutto, E. E. Gilbert. VADER: A parsimonious rule-based model for sentiment analysis of social media text. *Proceedings of the International AAAI Conference on Web and Social Media*, vol.8m no.1, 2014, pp.216-225. https://doi.org/10.1609/icwsm.v8i1.14550

[21] A. Dey, M. Jenamani, J. J. Thakkar. Senti-N-Gram: An n-gram lexicon for sentiment analysis. *Expert Systems with Applications*, vol. 103, pp. 92–105, 2018. https://doi.org/10.1016/j.eswa.2018.03.004

[22] H. Zargari, M. Zahedi, M. Rahimi. GINS: A global intensifier-based N-gram sentiment dictionary. *Journal of Intelligent & Fuzzy Systems*, vol. 40, 2021, pp. 1–14. https://doi.org/10.3233/JIFS-202879

Information about the authors:

Lijimol George – is a Research Scholar in Bharathidasan University and working as a Data Scientist in Tata Consultancy Services Her research interests include artificial intelligence, Sentiment Analysis and Large Language Models.

P Sumathy, PhD– is an Assistant Professor in the Department of Computer Science in Bharathidasan University. Her research interests include artificial intelligence, Image Processing, Sentiment Analysis.

Vadivel A., PhD – is a Professor and Head of the Department of Artificial Intelligence and Data Science, Drone Technologies Lab, MURTI Research Centre, GITAM School of Technology, GITAM Deemed to be University. His research interests include Medical Image Analysis, Video Analytics, Pattern and Sentiment Analysis

Manuscript received on 18 May 2025