DOI 10.15622/ia.23.1.2

G. JANARDANA NAIDU, M. SESHASHAYEE

SENTIMENT ANALYSIS FRAMEWORK FOR TELUGU TEXT BASED ON NOVEL CONTRIVED PASSIVE AGGRESSIVE WITH FUZZY WEIGHTING CLASSIFIER (CPSC-FWC)

Naidu G., Seshashayee M. **Sentiment Analysis Framework for Telugu Text Based on Novel Contrived Passive Aggressive with Fuzzy Weighting Classifier (CPSC-FWC).**

Abstract. Natural language processing (NLP) is a subset of artificial intelligence demonstrating how algorithms can interact with individuals in their unique languages. In addition, sentiment analysis in NLP is better in numerous programs, including evaluating sentiment in Telugu. Several unsupervised machine-learning algorithms, such as k-means clustering with cuckoo search, are used to detect Telugu text. However, these techniques struggle to cluster data with variable cluster sizes and densities, slow search speeds, and poor convergence accuracy. This study developed a unique ML-based sentiment analysis system for Telugu text to address the shortcomings. Initially, in the pre-processing stage, the proposed Linear Pursuit Algorithm (LPA) removes words in white spaces, punctuation, and stops. Then, for POS tagging, this research proposed a Conditional Random Field with Lexicon weighting; following that, a Contrived Passive Aggressive with Fuzzy Weighting Classifier (CPSC-FWC) is proposed to classify the sentiments in Telugu text. Consequently, the method we propose produces efficient outcomes in terms of accuracy, precision, recall, and f1-score.

Keywords: machine learning, natural language processing, polarity, sentiment analysis, Telugu.

1. Introduction. NLP is both a study and a topic which examines how computers can comprehend and adapt natural language text or speech to do useful tasks [1 – 4]. NLP scientists try to learn further about how people understand and use language to ensure suitable technology and methods can be developed to assist computers in perceiving and manipulating natural languages to do the required tasks. Sentiment analysis is an NLP technique that examines a person's emotions, sentiments, and opinions on various items, including products, films, events, reports, and businesses [5]. The main goal of sentiment analysis is to identify the polarity of a text within a given resource. Positive, negative, and neutral polarities are all conceivable.

In addition, there are three layers of sentiment analysis for text: sentence level, document level, and aspect level [6]. The goal of sentence-level analysis is to determine the polarity importance of each sentence in the examined material. Lexicon-based methods, along with Machine Learning (ML) [7] techniques, are the two basic approaches to Sentiment Analysis (SA). The lexicon-based technique analyses the text data by applying a sentiment lexicon for keyword matching. Word sentiment information can be found in lexicons. In [15] the aythors suggested SenticNet, and in study [14] suggested SentiWordNet as examples of the lexicon. Lexicon alone fails to deliver high accuracy, yet when combined with the Semantic Rule [16, 17], good results are obtained [19]. Semantic rules are used to deal with language exceptions like negation. It positively affects the categorisation of polarities. However, the method necessitates the human definition of Semantic Rules.

Furthermore, the polarity value is established by document-level analysis according to content evaluation. To determine the polarity of each text feature, aspect-level analysis (word-by-word) is performed [18]. While document categorization is less common than short text analysis along with social network evaluation, it could prove quite beneficial in a variety of situations, including the research of political views [20] displayed in the media, the analysis of feedback from consumers [21], with news coverage. Then there's "NLP," a discipline of AI that shows how techniques may connect with people utilizing their native languages. Considering the proliferation of fabricated reports then the growth of internet platforms, mining Telugu news data and classifying it depending on public sentiment is critical. The Stanford Sentiment Treebank – 2 (SST-2) dataset is extensively utilized for sentiment analysis (SA), and Bidirectional Encoder Representations from Transformers (BERT) outperform previous modern methods in this field [8]. Bataa and Wu [11] used BERT and transferable learning methodologies to achieve a new state-of-the-art Japanese SA.

Despite its efficacy in basic sentiment categorization, aspect-based sentiment analysis (ABSA), an enhanced SA task, showed less significant improvement when BERT was directly applied [9]. In paper [9] the authors created an auxiliary phrase to use BERT's powerful representation more by transforming Aspect Based Sentiment Analysis (ABSA) from a single-sentence classification issue to a sentence pair classification challenge. Furthermore, study [12] suggested that the BERT might profit from the addition of a pooling component, which might be performed as a Long Short Term Memory (LSTM) or an attention mechanism to use the BERT's intermediate levels for ABSA [13].

The amount of variables used in models has increased dramatically due to the introduction of Deep Learning (DL) [12]. A substantially bigger dataset is necessary to train the model's variables entirely while avoiding overfitting. Creating large-scale labelled datasets, nevertheless, is a substantial obstacle for the bulk of NLP positions owing to the relatively high annotation costs, especially for semantically and syntactically related jobs. In earlier research [22], we studied the efficacy of Convolutional Neural Network (CNN) and LSTM methods for lengthy text and observed that combining Doc2vec and CNN models generated marginally superior outcomes than RNN. It is required so that CNN may use the Doc2Vec models to identify the polarities of the entire work.

Telugu is India's second most frequently spoken language, behind Hindi. Following Ethnologue, Telugu is the fifteenth most frequently spoken dialect globally, with 85 million native Telugu speakers globally [23]. Several Telugu-language electronic newspapers publish news daily, including Eenadu, Sakshi, Andhrajyothy, Vaartha, and Andhrabhoomi, among others. Telugu material is prevalent on most news websites, web journals, Twitter accounts, and other social media platforms. Therefore, it is important to analyse the feelings underlying Telugu news.

Natural language processing has benefited considerably from data mining techniques [24]. Application for Knowledge Discovery in Real Time, such as Clinical Analysis [25], to proceed with comprehending the prediction methods, Association Rule Mining in the Business of Marketing [26], and the System of Education, demand a lean towards information disclosure techniques. The advent of ML and DL in NLP simplified and practicalized the difficult and tedious work of constructing perceives [23]. The following is the primary contribution of this study:

After collecting the Annotated Corpus of Telugu Sentiment Analysis (ACTSA) Telugu annotation dataset, the following steps were taken to ensure good quality input data for our machine learning models.

In the pre-processing stage Split () function is utilized to split the sentences into words based on the white spaces and punctuation. Then, we removed the words with no information (stop words) and special characters by using our proposed Linear Pursuit Algorithm (LPA). We balanced the dataset by using our proposed Adaptive Synthetic Sampling (ADASYN) approach.

– Then, this research proposed a Conditional Random Field (CRF) with Lexicon weighting for POS tagging.

– Then, to classify the Telugu text, this research proposed Contrived Passive Aggressive with Fuzzy Weighting Classifier (CPSC-FWC). Consequently, the suggested CPSC-FWC framework classifies emotion as positive, negative, or neutral.

– This research article is organised as follows: The third part discusses the proposed machine learning-based approach after a study of the present sentiment analysis for various article phrases in Part 2. Section 4 examines the implementation outcomes briefly, and Section 5 concludes this study piece.

2. Literature Survey. Utilising Telugu SentiWordNet, in paper [27] the authors presented a two-phase sentiment analysis for

Telugu news phrases. It first identifies subjectivity categorization, categorising statements as subjective or objective. Objective sentences are considered to have neutral feelings because they lack any emotive value. As part of the Sentiment Classification procedure, the subjective statements are separated into positive and negative sentences. However, the accuracy of this study is lacking, as is the ability to develop another technique for creating this SentiwordNet dynamic in the future.

Study [28] adopted a rule-based methodology to create SentiPhraseNet. SentiPhraseNet was used to extract the sentiment in this case, and the results were confirmed using the annotated corpus data set ACTSA. Sentiphrasenet does not include all possible phrases.

Paper [29] used ML methods with a knowledge-based strategy to develop a Telugu-language Word Sense Disambiguation (WSD) framework. Lexical Knowledge Base (LKB) is the information source for developing the WSD system. Disambiguation of words is still in the early stages, and little research has been published. Telugu has greater potential for word sense disambiguation than every other regional dialect nowadays. An unsupervised method can develop an upcoming word sense disambiguation algorithm for the local Telugu language.

In paper [30] the authors used ML classifiers to create an effective framework for classifying Telugu news data. In the authors' work, ML classifiers, including Multinomial Nave Bayes, Random Forest, Passive Aggressive Classifier, Perceptron, and Support Vector Machine (SVM), address the challenge of categorizing news sentiments in Telugu. A dataset is acquired via the open-source Kaggle platform. DL methodologies, including Recurrent Neural Network (RNN) and LSTM framework, will be used to discover the attitudes. It is also possible to extend the work to examine the news sentiments in Malayalam, Kannada, Tamil, and Urdu.

In study [31] the authors investigated numerous unsupervised machine-learning techniques for categorising Telugu text into either negative or positive classifications. Cuckoo search's superiority over K-means in locating the cluster's centroid is the explanation behind this. Next, research might look at more unsupervised techniques to find a viable solution for sentiment analysis in Telugu.

In [32] the authors planned to broaden a unique Evolving C4.5 ML with Spider Monkey Optimization (EC4.5-ML-SMO) to categorize sentiment evaluation in the Telugu language successfully. Spider Monkey Optimization is a contemporary-inspired technique defined by Swarm Intelligence Approaches. This approach is being employed in the present study to improve the accuracy of sentiment categorization. Furthermore, the bioinspired structure has its fitness that is determined by its behaviour. As a result, the current study evolved a novel hybrid system learning model for validating the client summary in the Telugu dataset. Iteration, on the other hand, requires more time.

Study [33] systematically classified views in Telugu using a lexicon-based methodology with ML in sentiment analysis. First, we detected subjective phrases from the Telugu corpus using a Lexicon-based technique called Telugu SentiWordNet. Second, we classified the sentiment in the corpus using ML methods such as SVM, Naive Bayes, and Random Forest.

Hence, the existing research utilized a rule-based methodology to create SentiPhraseNet. However, Sentiphrasenet does not include all possible phrases. Then, to classify the Telugu text, various unsupervised ML algorithms such as k-means clustering and cuckoo search algorithm are used; however, k-means has difficulty clustering data with varying cluster sizes and density, and the cuckoo method has the drawbacks of slow search speed and low convergence accuracy. This study created a special network, which is covered in more detail in the following section, to overcome the above-mentioned disadvantages.

3. Proposed Approach. The following are the steps of an innovative ML-based sentiment analysis tool for Telugu text. Stage I: we collected the ACTSA Telugu annotated dataset, and then the datasets were pre-processed using the proposed Linear Pursuit Algorithm. Hence we removed the white spaces, punctuation, and stop words and balanced the dataset by using our proposed Adaptive Synthetic Sampling (ADASYN) approach. In Stage II, a Conditional Random Field with Lexicon weighting is proposed for parts of speech tagging. Then, the processed data are fed into the proposed Contrived Passive Aggressive with Fuzzy Weighting Classifier (CPSC-FWC) is described in Stage III. Therefore, the proposed CPSC-FWC framework classifies emotion as positive, negative, or neutral. Figure 1 displays the proposed approach's design.

Dataset Description. In this part, we will peek at the places where the initial information was received. Our source data was collected from five distinct Telugu news websites: Andhra Bhoomi, Andhra-Jyothi, Eenadu, Kridajyothi, and Saakshi. The writers gathered over 453 news pieces and whittled them down to 321.

Annotators labelled all of the statements, with every statement annotated by precisely two annotators (Table 1).

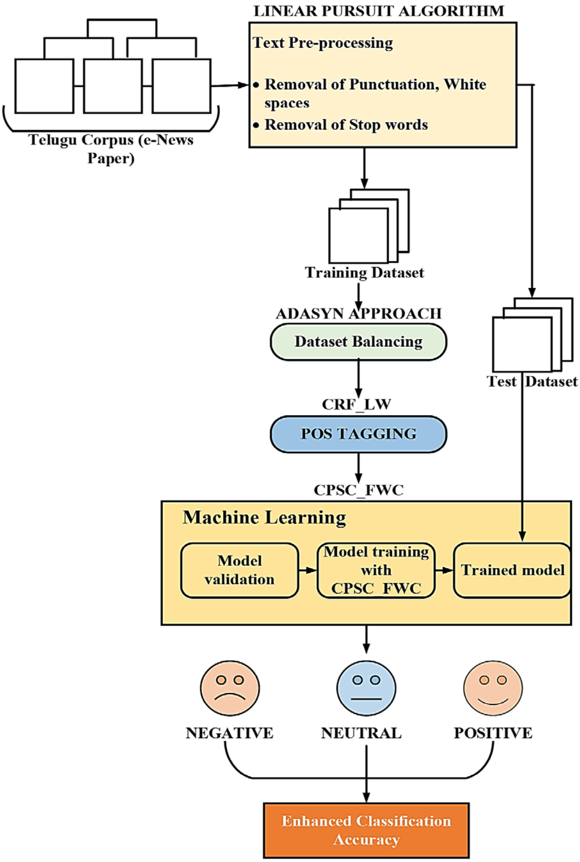

Fig. 1. Architecture of the proposed approach

Table 1. Example Annotations

ID	Original Sentence	English Translation	Class
1	అమెరికా అధ్యక్షుడు డొనాల్డ్ ట్రంప్ పారిస్ నుంచి అమెరికాను ఉపసంహరించుకున్నారు. సహచరుడు ఒప్పందం	US President Donald Trump withdrew the US from the Paris Climate Agreement	Negative
2	అవసరం లేదు ఇందులో ఎవరికైనా ఏదైనా అభ్యంతరం	There is no need for any objection to any- one in this	Neutral
3	భారతదేశ ప్రధాన మంత్రి కశ్మీర్ అల్లర్లపై నరేంద్ర మోడి తీవ్రంగా స్పందించారు	India's Prime Minister Narendra Modi has re- acted severely to the Kashmir riots	Negative
4	మంత్రి సంతోషించారు ఫలితాలపై	The minister is happy on the results	Positive

Sentences deemed ambiguous by at least one annotator were removed from the corpus to avoid unclear sentences. Table 2 displays several corpus annotation examples.

Table 2. Agreement for Sentences in ACTSA

Annotator	Positive	Negative	Neutral	Total
Positive	1463	31	103	1597
Negative	23	1421	116	1560
Neutral	112	127	2427	2666
Total	1598	1579	2646	5823

Our data statistics, from raw data collecting to the last words. The writers gathered 453 news stories and whittled them down to 321 associated with their study area. The writers have 11952 sentences in their initial information. The researchers then examined the statements for subjectivity and deleted 4327 objective sentences, leaving 7812 sentences. The annotators were given these sentences to annotate. There were 1802 sentences eliminated because a minimum of one annotator identified them as questionable. Of the remaining 5823 sentences, 512 were in dispute and were referred to a third independent annotation. Following the third annotation, 413 sentences were deleted if there was a dispute or if they were objective. The annotated corpus requested by ACTSA is made up of the other 5405 sentences. Table 3 gives statistics for our whole corpus. Following the third annotation, 413 contested or objective sentences were eliminated. The annotated corpus requested by ACTSA is made up of the remaining 5405 sentences. Table 3 gives statistics for our whole corpus.

Table 3. Statistics about the data

News articles	321
Cleaned Sentences	11952
Objective Sentences (Removed)	4327
Uncertain Sentences (Removed)	1802
Disagreement Sentences Classified Removed	512 99 413
Positive sentences	1489
Negative sentences	1441
Neutral sentences	2475
Total sentences	5405

In the pre-processing stage, the information received was cleaned, for example, by deleting punctuation and removing phrases with non-Telugu terms, excess spaces, URLs, and other trash values. Later, Sentence Segmentation is conducted after all of this info is separated into discrete sentences. The initial data corpus is split between 90% training and 10% test datasets.

Pre-processing. We took the following measures to secure high-quality input data for our ML models (Table 4).

Table 4. Example of the pre-processed Telugu sentence

Original Telugu Sentence	**Pre-processed Sentence**	**Polarity (1, 0, -1)**
పరువునష్టాన్ని కలిగించే చర్యలను 1860 నాటి భారత శిక్షాస్మృతిలోని సెక్షన్ 499 నిర్వచించగా, అందుకు విధించగలిగిన శిక్షను, జరిమానాను సెక్షన్ 500 పేర్కొన్నది.	'పరువునష్టాన్ని' 'కలిగించే' 'చర్యలను' '1860' 'నాటి' 'భారత' 'శిక్షాస్మృతిలోని' 'సెక్షన్' '499' 'నిర్వచించగా, ' 'అందుకు' 'విధించగలిగిన' 'శిక్షను, ' 'జరిమానాను' 'సెక్షన్' '500' 'పేర్కొన్నది'	0
వాహనంలో ఉన్న పోలీసు సిబ్బందికి బుల్లెట్ ప్రూఫ్ కవచాలు లేవన్నది అటుంచితే, వారు ఏమాత్రం అప్రమత్తంగా లేరన్నది కూడా వాస్తవం ఎవరైనా	'వాహనంలో' 'ఉన్న' 'పోలీసు' 'సిబ్బందికి' 'బుల్లెట్' 'ప్రూఫ్' 'కవచాలు' 'లేవన్నది' 'అటుంచితే' 'వారు' 'ఏమాత్రం' 'అప్రమత్తంగా' 'లేరన్నది' 'కూడా' 'వాస్తవం'	1
దేశంపట్ల కానీ, ప్రజలపట్ల కానీ బాధ్యత కనిపించదు	'దేశంపట్ల' 'కానీ,' 'ప్రజలపట్ల' 'కానీ' 'బాధ్యత' 'కనిపించదు'	-1

In the pre-processing stage Split () function is utilized to split the sentences into words based on the white spaces and punctuation. Here, we will separate the text by all commonly used characters, such as "." and "/n," then, given two sentences, the model will determine whether they should be combined. As a result, the model will provide us with a new sentence-by-sentence breakdown of the text.

Then, we removed the words with no information (stop words) and special characters by using our proposed Linear Pursuit Algorithm (LPA) (Algorithm 1, Table 5). It searches all the elements in the datasets sequentially. First, it checks for the first sentence; if the sentence has to stop words, it will remove them. Otherwise, it continues for n sentences; by doing this, we obtain an accurate dataset with less memory usage. The most fundamental search method is linear search, also called sequential search. This type of search entails searching the complete list (Telugu dataset) for a match for a particular component. If a match is detected, the stop words of the matched target component are sent back. On the contrary, if the component cannot be found, it returns NULL. The following is a step-by-step procedure for doing the Linear Pursuit Procedure:

– First, read the search sentence (Stop words) in the dataset.
– The stop words are compared to the initial phrase in the array in the next stage.
– If both are matched, the Linear Pursuit function will be terminated and the message "Stop word found" will be shown.
– If none is found, compare the stop words to the next phrase in the dataset.
– Steps 3 and 4 should be repeated until the search result (Stop words) contrasts the dataset's end sentence.
– If the final sentence in the list fails to match, the Linear Search Function is halted, and the message "Sentence not found" is shown.

Algorithm 1. Linear Pursuit Algorithm
Linear Pursuit (Telugu Sentence Sen, Value b) //Sen is the name of the dataset, and b is the stop words element.
Step 1: Set i = 0 // i is the index of a which starts from 0
Step 2: if i > n, then go to step 7 // n is the number of stop words in the sentence
Step 3: if Sen [i] = b, then go to step 6
Step 4: Set i = i+1
Step 5: Go to Step 2
Step 6: Display stop word is found
Step 7: Display stop word is not found
Step 8: Terminate

Table 5. Example of Linear Pursuit Algorithm

Original Telugu Sentence	**Stop words removed sentence**	**Polarity (1, 0, -1)**
2008లో ముంబయిలోని తాజ్మహల్ హోటల్, ఒబెరా ట్రిడెంట్లపై జరిగిన ముష్కరుల దాడి సందర్భంలో కూడా ఇదే విధమైన మీడియా అత్యుత్సాహం అపారనష్టం తెచ్చింది.	2008లో ముంబయిలోని తాజ్మహల్ హోటల్, ఒబెరా ట్రిడెంట్లపై జరిగిన ముష్కరుల దాడి సందర్భంలో కూడా ఇదే విధమైన మీడియా అత్యుత్సాహం అపారనష్టం తెచ్చింది	-1
మీరు విడిపోదామని నిర్ణయించుకున్న పక్షంలో బ్రిటన్ మునిగిపోతుంది	మీరు విడిపోదామని నిర్ణయించుకున్న పక్షంలో బ్రిటన్ మునిగిపోతుంది	-1
వరుసగా అదే పనిచేస్తే డిడిఎస్ రంగంలోకి దిగుతుంది.	వరుసగా అదే పనిచేస్తే డిడిఎస్ రంగంలోకి దిగుతుంది	0

This study proposed an Adaptive Synthetic Sampling (ADASYN) method for learning from unbalanced data sets. The goal here is twofold: minimizing bias and learning adaptively. [Algorithm ADASYN] describes the proposed solution for the three-class (positive, negative, and neutral) classification issue.

Input: Stop words removed training dataset D_{Str} with m samples $\{x_i, y_i\}$, $i = 1, \ldots, m$, where x_i is a sentence in the n-dimensional dataset X and $y_i \in Y = \{1, -1\}$ is the class identity label (positive, negative, and neutral) associated with x_i. Define m_s and m_l as the number of samples from the minority & majority classes correspondingly. Consequently, $m_s \leq m_l$ and $m_s + m_l = m$.

Process Flow:

Determine the degree of class imbalance:

$$d = m_s/m_l, \text{ Where d } \epsilon \text{ (0, 1).} \tag{1}$$

If $d < d_{th}$ then (d_{th} is a pre-determined highest allowed degree of class imbalance ratio).

a) Determine the amount of synthetic data samples necessary for the minority class:

$$G = (m_l - m_s) \times \beta, \quad (2)$$

where $\beta \epsilon$ [0, 1] is a parameter utilized to set the desired balance level once the synthetic data has been generated. $\beta = 1$ After the generalization process, a properly balanced data set is obtained.

b) For every sample $x_i \in minority\ class$, find K nearest neighbours depending on Euclidean distance in n dimensional space and estimate the proportion r_i described as:

$$r_i = \Delta_i / K, i = 1, \dots, m_s, \quad (3)$$

where Δ_i is the amount of samples in K nearest neighbours of x_i that belong to the majority class, consequently $r_i \epsilon$ [0, 1];

c) Normalize r_i allowing to $\widehat{r_i} = r_i / \sum_{i=1}^{m_s} r_i$, so that $\widehat{r_i}$ is a density distribution ($\sum_i \widehat{r_i} = 1$).

d) Determine the number of synthetic data samples required for every minority sample x_i:

$$g_i = \widehat{r_i} \times G. \quad (4)$$

G is the total synthetic data samples required for the minority class, as stated in Equations (2).

e) Generate g_i synthetic samples of data for every minority class data item x_i using the following stages.

Make the Loop from 1 to g_i.

Select one minority data sample at random x_{zi}, from K nearest neighbours for data x_i.

Create a synthetic data sample (s_i).

$$s_i = x_i + (x_{zi} - x_i) \times \lambda, \quad (5)$$

where $(x_{zi} - x_i)$the n-dimensional difference is a vector, and λ is a random number: $\lambda \epsilon$ [0, 1].

End Loop.

The basic concept behind the ADASYN method is to utilize a density distribution ($\widehat{r_i}$) as a criterion to mechanically determine the number of synthetic samples necessary for every minority data sample. As a result, this study generates synthetic data based on data density, eliminating the bias created by class imbalance. Hence, our dataset has negative, positive and neutral samples as 1400, 1400, and 1400. Then, for Parts of Speech (POS) tagging, this research proposed a novel Conditional Random Field with Lexicon weighting described in the following section.

POS Tagging. Part-of-speech tagging, often known as POS, provides a unique label to each token in a text to recognize its part of speech and, in certain situations, extra grammatical meanings. Text classifier algorithms then use this labelling information. For POS tagging, this research proposed a Conditional Random Field with Lexicon weighting, which is utilized to identify similar entities. These lexicons (e.g., positive and negative) are used as features to increase the models' accuracy.

A CRF is a sequence-modelling technique that recognizes entities or patterns in text, including POS tags (Table 6). This technique not only shows that attributes are connected, but it also takes into account potential implications while learning a pattern. The CRF equation is as follows: Y is the hidden state (such as a chunk of speech), and X is the observed variable (in our case, the entity or other words within it). A CRF stands for Discriminative Probabilistic Classifier. The difference between discriminative and generative mathematical structures is that discriminative models try to reflect conditional probability distributions. However, generative models do not, i.e., P(y|x), and generative models try to model a joint probability distribution, i.e., P(x,y). As demonstrated in Equation (6), the CRF framework may offer a conditional probability of a potential output sequence for an input sequence supplied by x.

$$p(y|x)] = \frac{1}{\underbrace{Z\,(x)}_{\text{Normalization}}} \prod_{t=1}^{T} exp\left\{\sum_{k=1}^{K} \underbrace{F_k T_k}_{\text{Weight}}\ \underbrace{(y_t, y_{t-1}, X_t)}_{\text{Feature,}}\right\}, \tag{6}$$

where (y_t, y_{t-1}, X_t) is utilized to represent a feature function, $(F_k T_k)$ is employed to signify the lexicon weight vector, then to close $Z\,(x)$ is used to represent the normalization factor. $T_k \rightarrow F_k$ is such that every term in T is mapped to a term in F with or without stemming. Following that, n-grams with the odds of particular words occurring in specific sequences might enhance auto-completion system forecasts.

Table 6. Example of POS Tagged Sentence

Telugu Sentence	English Translation	POS Tagged
ఆడిట్	Audit	NN
నిర్వహణ	Management	NN
ఆడిటర్	Auditor	NN
ఒక	A	QFNUM
కొత్త	New	JJ
చేపట్టే	Undertaking	VRB
ముందు	Before	PREP
సక్రమ	Correct	JJ
పద్ధతి	Method	NN
లో	In	PREP
కార్య	Work	JJ
ప్రణాళికను	Plan	NN
రూపొందించాలి	Should be made	VFM
.		SYM

Then, these tagged sentences are fed into the proposed Contrived Passive Aggressive with Fuzzy Weighting Classifier (CPSC-FWC) (Figure 2).

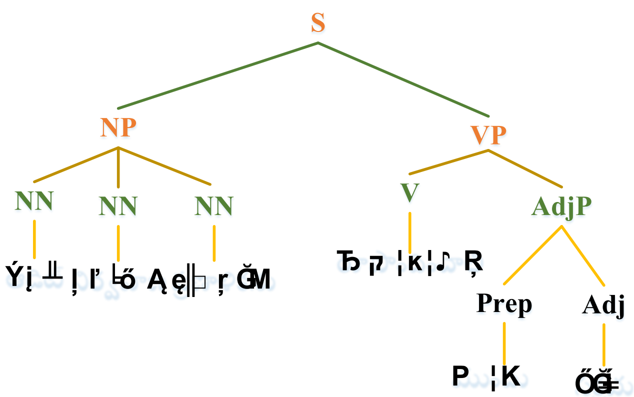

Fig. 2. Tree for POS tagged sentence

Contrived Passive Aggressive with Fuzzy Weighting Classifier (CPSC-FWC). For sentiment classification, the existing research utilized regularization parameters. However, it leads to dimensionality reduction in the dataset, and it leads to error. Moreover, the loss functions are very sensitive to outliers, so there is a need to change the regularization and loss functions. Therefore, this research proposed a Contrived Passive Aggressive with Fuzzy Weighting Classifier (CPSC-FWC) to overcome the limitations. In this proposed model, the membership function is evaluated based on the passive-aggressive classifier, and then, for each sample, the weights of vectors are evaluated based on the fuzzy weighting classifier. The weight vector (W_t) can be changed as the membership is introduced using this simple, efficient method.

$$W_{t+1} = W_t + \mu_t \tau_t y_t x_t, \qquad (7)$$

where μ_t is the membership degree corresponding to the t-th sample, τ_t is the Lagrange multiplier, C is a positive parameter that governs how aggressive the update is, x_t *and* y_t is the input and target data to the t-th sample. This approach might be adapted to other online learning techniques because the membership computation remains independent of the framework update. Furthermore, such a generalized system may be used with kernelized weighing systems. The generalized Fuzzy PA technique is summarized below:

Algorithm 2. Contrived Passive Aggressive with Fuzzy Weighting Classifier (CPSC-FWC)
Input: Data stream (x_1, y_1), (x_2, y_2),…, (x_N, y_N), Parameter C > 0
Output: The vector of weights W_N
Initialization: $W_1 = (0, \dots, 0)$
For t = 1; N - 1
Calculate the membership μ_t of the sample x_t
Update the vector of weights, $W_{t+1} = W_t + \mu_t \tau_t y_t x_t$
End

At last, in our proposed CPSC-FWC, the Bi-direction Long Short Term Memory (Bi-LSTM) model is used to tune the hyperparameter. This table lists the layer names and the hyperparameter settings for every layer. The hypertune BiLSTM model comprises six layers in total.

In this framework, there are additional settings for the hyperparameters. Compared to the unidirectional LSTM structure, the proposed framework model is a superior option for managing huge time series sequences and preventing information loss. The hyperparameter influences network convergence and accuracy significantly. Bi-LSTM was used to choose network hyperparameters. The variety of epochs and the learning rate are standard Bi-LSTM training hyperparameters. The network speed is modified according to weight updates (CPSC-FWC) and the learning rate (Figure 3). The learning strategy that retrains the network multiple times using the complete dataset represents the entire epochs. As a result, the network can produce more precise outcomes.

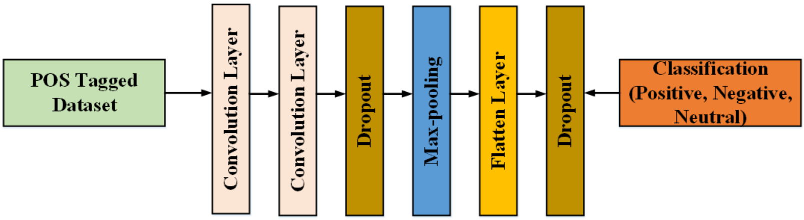

Fig. 3. Architecture for Bi-LSTM

Consequently, the proposed CPSC-FWC framework categorizes emotion as positive, negative, or neutral (Figure 4). The following section discusses the proposed model performance.

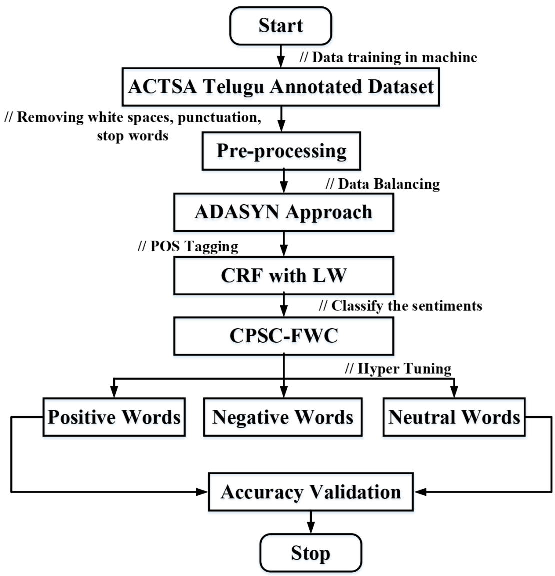

Fig. 4. Overall flowchart of the proposed approach

4. Result and Discussion. This part examines the outcomes of an experiment on the ACTSA dataset using the proposed CPSC-FWC classifier.

Tool	: PYTHON 3
OS	: Windows 7 (64-bit)
Processor	: Intel Premium
RAM	: 8GB RAM

Performance Metrics and Comparison Analysis. After developing all of the techniques mentioned above, we must evaluate the effectiveness of these frameworks.

Performance Measures. The various performance metrics of the novel Machine Learning based sentiment analysis framework for Telugu text are described in this section.

Various metrics, including accuracy, F1 Score, and precision, are utilized to assess the efficacy of our proposed strategy. The proposed method's performance evaluation measures are shown in Figure 5. The results were 78% accuracy, 75% precision, and 79.9% F1 score. By implementing a unique Contrived Passive Aggressive with Fuzzy Weighting Classifier (CPSC-FWC), Accuracy, F1 score, and precision are all improved by our proposed approach.

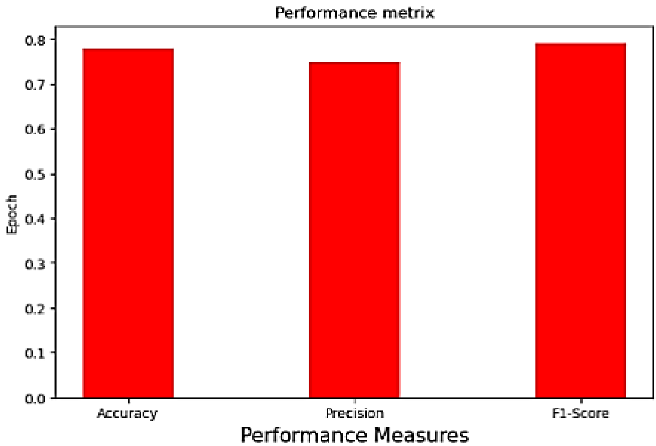

Fig. 5. Performance metrics of the proposed approach

Error Metrics of the proposed approach. The mean squared error (MSE) is used to calculate the degree of error in statistical frameworks. It is determined as the average squared variance between actual and anticipated values.

$$MSE = \frac{1}{n_z}\sum_{i=1}^{n_z}(y_{zi} - \widehat{y_{zi}})^2, \quad (8)$$

$$MAE = \frac{1}{n_z}, \quad (9)$$

where, n_z is the number of data points, y_{zi} are the observed values, $\widehat{y_{zi}}$ is the forecasted value.

Figure 6 illustrates the MSE and MAE for the proposed Contrived Passive Aggressive with Fuzzy Weighting Classifier (CPSC-FWC) model. The obtained MSE and MAE values are 0.26 and 0.17, respectively, using our proposed approach.

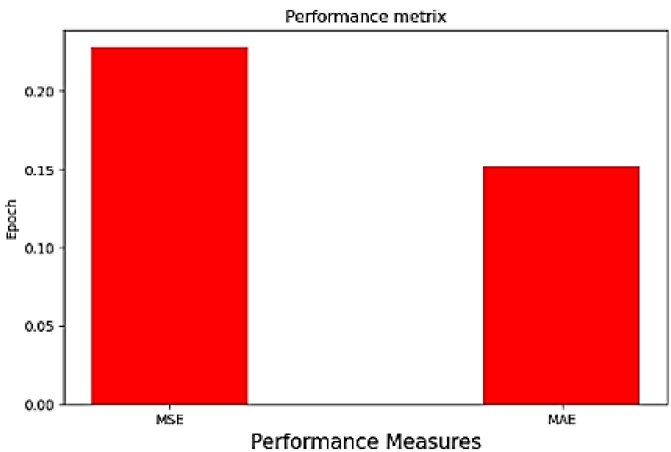

Fig. 6. Error measures of the proposed approach

Comparison of Accuracy. The entire performance validation of a system's learning approach is completed by assessing the precision of classification according to true positive (TP), true negative (TN), false positive (FP), and false negative (FN) results. Table 7 shows the evaluation validation of accuracy for sentiment type.

$$Accuracy = \frac{TN + TP}{TN + TP + FN + FP}. \quad (10)$$

Table 7. Comparison of Accuracy

Techniques	Accuracy (%)
K means Clustering	51
Cuckoo Search Algorithm	58
Rule-based approach	70
K-Nearest Neighbour	49
Ada Boost Classifier	73
Proposed Approach	78

The overall accuracy comparison is shown in Figure 7. The accuracy of the proposed technique improves by using Contrived Passive Aggressive with Fuzzy Weighting Classifier (CPSC-FWC). When compared to the baseline, the method we propose achieves greater accuracy as K means Clustering, Cuckoo Search Algorithm, Rule-based approach, K-Nearest Neighbour, and Ada Boost Classifier such as 51%, 58%, 70%, 49% and 73%. As a result, our novel technique has an accuracy of 78%, which is higher than baseline approaches.

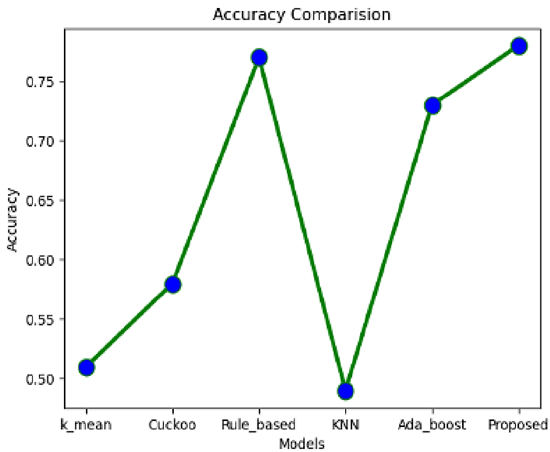

Fig. 7. Comparison of Accuracy

Comparison on Precision. The accuracy of the information processed is calculated by dividing the entire amount of sentiment sentences by the amount of precise particular sentiment predictions. Table 8 shows the precision validation evaluation for sentiment type.

$$Precision = \frac{TP}{TP + FP}. \tag{11}$$

Table 8. Comparison of Precision

Techniques	Precision (%)
K means Clustering	51
Cuckoo Search Algorithm	55
Rule-based approach	70
Ada Boost Classifier	71
Proposed Approach	75

The overall precision comparison is shown in Figure 8. The precision of the proposed technique improves by using Contrived Passive Aggressive with Fuzzy Weighting Classifier (CPSC-FWC). The proposed approach attains higher precision when compared to the baseline as K means Clustering, Cuckoo Search Algorithm, Rule-based approach, and Ada Boost Classifier, which have the precision of 51%, 55%, 70%, and 71%, respectively. As a result, this novel technique has a 75 % precision higher than baseline approaches.

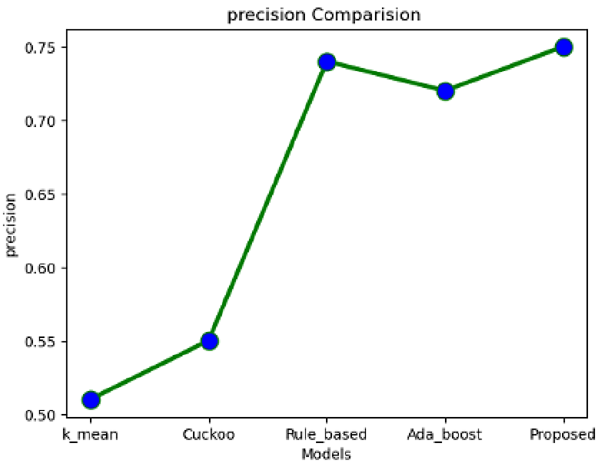

Fig. 8. Comparison of Precision

Comparison on F1_score. The f1_score has been verified to evaluate the mean average for accuracy and recall, allowing the f1_score to be compared. The f1_score average is produced via the average of accuracy and precision. F1_score is used to test categorization accuracy. Excellent accuracy and precision result in a higher f1_score rate (Table 9).

$$\mathrm{F1}_{\mathrm{score}} = 2 \times \frac{Precision \times Recall}{Precision + Recall}. \tag{12}$$

Table 9. Comparison of F1_score

Techniques	F1_score (%)
K means Clustering	67.4
Cuckoo Search Algorithm	69.9
Rule-based approach	79
Ada Boost Classifier	78.6
Proposed Approach	79.9

The overall f1_score comparison is shown in Figure 9.

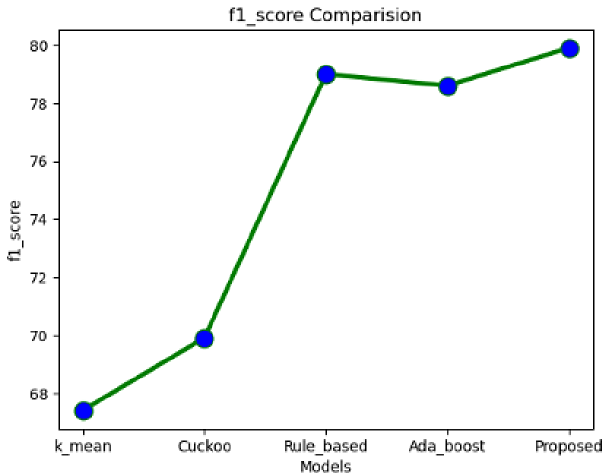

Fig. 9. Comparison of F1_score

The f1_score of the proposed technique improves by using Contrived Passive Aggressive with Fuzzy Weighting Classifier (CPSC-FWC). The proposed approach attains a higher f1_score when compared to the baseline as K means Clustering, Cuckoo Search Algorithm, Rule-based approach, and Ada Boost Classifier such as 67.4%, 69.9%, 79%, and 78.6%. As a result, the novel technique has a precision of 79.9%, which is higher than the baseline approaches.

5. Conclusion. In this research, we developed a unique ML-based sentiment analysis framework for Telugu text. First, in the pre-processing stage, Linear Pursuit Algorithm (LPA) is proposed, which aids in removing the white spaces, punctuation, and stop words in the ACTSA Telugu dataset. Following that, we balanced the dataset by consuming our proposed Adaptive Synthetic Sampling (ADASYN) method, decreasing the bias presented by the class imbalance. Then, for POS tagging, this research proposed a Conditional Random Field with Lexicon weighting to improve the accuracy of the proposed Contrived Passive Aggressive with Fuzzy Weighting Classifier (CPSC-FWC) models. As a result, our proposed approach provides an accuracy of 78%, a precision of 75% and an f1_score of 79.9% when compared to the existing approach such as K means Clustering, Cuckoo Search Algorithm, Rule-based approach, and Ada Boost Classifier. The goal for years to come is to develop those classifiers so that they may adapt effectively to large-scale datasets. Consequently, DL models such as multi-layer feed-forward neural networks, CNN, RNN, and ensemble DL models have become an unavoidable avenue of future study.

References

1. Chowdhary K.R., Chowdhary K.R. Natural language processing. Fundamentals of artificial intelligence. 2020. pp. 603–649.
2. Eisenstein J. Introduction to natural language processing. MIT Press. 2019. 536 p.
3. Raina V., Krishnamurthy S., Raina V., Krishnamurthy S. Natural language processing. Building an Effective Data Science Practice: A Framework to Bootstrap and Manage a Successful Data Science Practice. 2022. pp. 63–73.
4. Nguyen H.V., Tan N., Quan N.H., Huong T.T., Phat N.H. Building a Chatbot System to Analyze Opinions of English Comments. Informatics and Automation. 2023. vol. 22. no. 2. pp. 289–315. DOI: 10.15622/ia.22.2.3.
5. Qiu X., Sun T., Xu Y., Shao Y., Dai N., Huang X. Pre-trained models for natural language processing: A survey. Science China Technological Sciences. 2020. vol. 63. no. 10. pp. 1872–1897.
6. Song L., Xin C., Lai S., Wang A., Su J., Xu K. CASA: Conversational aspect sentiment analysis for dialogue understanding. Journal of Artificial Intelligence Research. 2022. vol. 73. pp. 511–533.
7. Wang Y., Chen Q., Ahmed M.H., Chen Z., Su J., Pan W., Li Z. Supervised Gradual Machine Learning for Aspect-Term Sentiment Analysis. Transactions of the Association for Computational Linguistics. 2023. vol. 11. pp. 723–739.
8. Liu B. Sentiment analysis and opinion mining. Springer Nature, 2022. 167 p.
9. Talaat A.S. Sentiment analysis classification system using hybrid BERT models. Journal of Big Data. 2023. vol. 10. no. 1. pp. 1–18.
10. Hoang M., Bihorac O.A., Rouces J. Aspect-based sentiment analysis using Bert. Proceedings of the 22nd nordic conference on computational linguistics. 2019. 187–196.

11. Bataa E., Wu J. An investigation of transfer learning-based sentiment analysis in Japanese. arXiv preprint arXiv:1905.09642. 2019.
12. Lv H., Liu J., Wang H., Wang Y., Luo J., Liu Y. Efficient hybrid generation framework for aspect-based sentiment analysis. Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics. 2023. pp. 1007–1018.
13. Chen C., Teng Z., Wang Z., Zhang Y. Discrete opinion tree induction for aspect-based sentiment analysis. Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics. 2022. vol. 1. pp. 2051–2064.
14. Esuli A., Sebastiani F. Determining the semantic orientation of terms through gloss classification. Proceedings of the 14th ACM international conference on information and knowledge management. 2005. pp. 617–624.
15. Cambria E., Havasi C., Hussain A. SenticNet 2: A semantic and affective resource for opinion mining and sentiment analysis. Proceedings of the Twenty-Fifth International Florida Artificial Intelligence Research Society Conference. 2012. pp. 202–207.
16. Xiaomei Z., Jing Y., Jianpei Z., Hongyu H. Microblog sentiment analysis with weak dependency connections. Knowledge-Based Systems. 2018. vol. 142. pp. 170–180.
17. Appel O., Chiclana F., Carter J., Fujita H. Successes and challenges in developing a hybrid approach to sentiment analysis. Applied Intelligence. 2018. vol. 48. pp. 1176–1188.
18. Yin C., Chen S., Yin Z. Clustering-based Active Learning Classification towards Data Stream. ACM Transactions on Intelligent Systems and Technology. 2023. vol. 14. no. 2. pp. 1–18.
19. Naseri S., Dalton J., Yates A., Allan J. CEQE to SQET: A study of contextualized embeddings for query expansion. Information Retrieval Journal. 2022. vol. 25. no. 2. pp. 184–208.
20. Sobkowicz P., Kaschesky M., Bouchard G. Opinion mining in social media: Modeling, simulating, and forecasting political opinions in the web. Government information quarterly. 2012. vol. 29. no. 4. pp. 470–479.
21. Hu Y.H., Chen Y.L., Chou H.L. Opinion mining from online hotel reviews–a text summarization approach. Information Processing and Management. 2017. vol. 53. no. 2. pp. 436–449.
22. Yousfi S., Rhanoui M., Mikram M. Comparative study of CNN and LSTM for opinion mining in long text. Journal of Automation, Mobile Robotics and Intelligent Systems. 2020. pp. 50–55.
23. Ethnologue Languages of the World [online]. Available at: https://www.ethnologue.com/statistics/size (accessed 01.09.2023).
24. Sultana J., Rani M.U., Farquad M.A.H. Knowledge discovery from recommender systems using deep learning. International Conference on Smart Systems and Inventive Technology (ICSSIT). 2019. pp. 1074–1078.
25. Sultana J., Jilani A.K. Predicting breast cancer using logistic regression and multi-class classifiers. International Journal of Engineering and Technology. 2018. vol. 7. no. 4(20). pp. 22–26.
26. Sultana J., Nagalaxmi G. How Efficient is Apriori: A Comparative Analysis. International Journal of Current Engineering and Scientific Research. 2015. pp. 2393–8374.
27. Naidu R., Bharti S.K., Babu K.S., Mohapatra R.K. Sentiment analysis using telugu sentiwordnet. International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET). 2017. pp. 666–670.
28. Garapati A., Bora N., Balla H., Sai M. SentiPhraseNet: An extended SentiWordNet approach for Telugu sentiment analysis. International Journal of Advance Research, Ideas and Innovations in Technology. 2019. vol. 5. no. 2. pp. 433–436.

29. Koppula N., Rani B.P., Srinivas Rao K. Graph-based word sense disambiguation in Telugu language. International Journal of Knowledge-based and Intelligent Engineering Systems. 2019. vol. 23. no. 1. pp. 55–60.
30. Sultana J. Telugu News Data Classification Using Machine Learning Approach. Handbook of Research on Advances in Data Analytics and Complex Communication Networks. 2022. pp. 181–194.
31. Janardana Naidu G., Seshashayee M. Sentiment analysis for Telugu text using cuckoo search algorithm. Smart Computing Techniques and Applications: Proceedings of the Fourth International Conference on Smart Computing and Informatics. 2021. vol. 2. pp. 253–257.
32. Suryachandra P., Venkata P., Reddy S. Machine Learning Approach to Classify the Sentiment Value of Natural Language Processing in Telugu Data. Journal of Engineering and Applied Sciences. 2020. vol. 15. pp. 3593–3598.
33. Tammina S. A hybrid learning approach for sentiment classification in Telugu language. International Conference on Artificial Intelligence and Signal Processing (AISP). 2020. pp. 1–6.

Naidu G Janardana — Research scholar, Department of computer science, Gandhi Institute of Technology and Management GITAM (Deemed to be University). Research interests: computer science. jana.766@gmail.com; Gandhi Nagar, Rushikonda, 530045, Visakhapatnam, Andhra Pradesh, India; office phone: +91(0891)284-0501.

M Seshashayee — Ph.D., Dr.Sci., Assistant professor, Department of computer science, Gandhi Institute of Technology and Management GITAM (Deemed to be University). Research interests: computer science. mseshashayee@gmail.com; Gandhi Nagar, Rushikonda, 530045, Visakhapatnam, Andhra Pradesh, India; office phone: +91(0891)284-0501.

УДК 004.7 **DOI** 10.15622/ia.23.1.2

Д. НАЙДУ, М. СЕШАШАЯИ

СИСТЕМА АНАЛИЗА ТОНАЛЬНОСТИ ТЕКСТА НА ТЕЛУГУ НА ОСНОВЕ НОВОГО ПАССИВНО-АГРЕССИВНОГО КЛАССИФИКАТОРА С НЕЧЕТКИМ ВЗВЕШИВАНИЕМ

Найду Д., Сешашаяи М. **Система анализа тональности текста на телугу на основе нового пассивно-агрессивного классификатора с нечетким взвешиванием.**

Аннотация. Обработка естественного языка (NLP) – это разновидность искусственного интеллекта, демонстрирующая, как алгоритмы могут взаимодействовать с людьми на их уникальных языках. Кроме того, анализ настроений в NLP лучше проводится во многих программах, включая оценку настроений на телугу. Для обнаружения текста на телугу используются несколько неконтролируемых алгоритмов машинного обучения, таких как кластеризация k-средних с поиском с кукушкой. Однако эти методы с трудом справляются с кластеризацией данных с переменными размерами и плотностью кластеров, низкой скоростью поиска и плохой точностью сходимости. В ходе этого исследования была разработана уникальная система анализа настроений на основе машинного обучения для текста на телугу, позволяющая устранить указанные недостатки. Первоначально, на этапе предварительной обработки, предлагаемый алгоритм линейного преследования (LPA) удаляет слова в пробелах, знаках препинания и остановках. Затем для маркировки POS в этом исследовании было предложено условное случайное поле с лексическим взвешиванием; После этого предлагается надуманный пассивно-агрессивный классификатор с нечетким взвешиванием (CPSC-FWC) для классификации настроений в тексте на телугу. Следовательно, предлагаемый нами метод дает эффективные результаты с точки зрения точности, воспроизводимости и показателя f1.

Ключевые слова: машинное обучение, обработка естественного языка, полярность, анализ настроений, телугу.

Литература

1. Chowdhary K.R., Chowdhary K.R. Natural language processing. Fundamentals of artificial intelligence. 2020. pp. 603–649.
2. Eisenstein J. Introduction to natural language processing. MIT Press. 2019. 536 p.
3. Raina V., Krishnamurthy S., Raina V., Krishnamurthy S. Natural language processing. Building an Effective Data Science Practice: A Framework to Bootstrap and Manage a Successful Data Science Practice. 2022. pp. 63–73.
4. Nguyen H.V., Tan N., Quan N.H., Huong T.T., Phat N.H. Building a Chatbot System to Analyze Opinions of English Comments. Informatics and Automation. 2023. vol. 22. no. 2. pp. 289–315. DOI: 10.15622/ia.22.2.3.
5. Qiu X., Sun T., Xu Y., Shao Y., Dai N., Huang X. Pre-trained models for natural language processing: A survey. Science China Technological Sciences. 2020. vol. 63. no. 10. pp. 1872–1897.
6. Song L., Xin C., Lai S., Wang A., Su J., Xu K. CASA: Conversational aspect sentiment analysis for dialogue understanding. Journal of Artificial Intelligence Research. 2022. vol. 73. pp. 511–533.
7. Wang Y., Chen Q., Ahmed M.H., Chen Z., Su J., Pan W., Li Z. Supervised Gradual Machine Learning for Aspect-Term Sentiment Analysis. Transactions of the Association for Computational Linguistics. 2023. vol. 11. pp. 723–739.
8. Liu B. Sentiment analysis and opinion mining. Springer Nature, 2022. 167 p.

9. Talaat A.S. Sentiment analysis classification system using hybrid BERT models. Journal of Big Data. 2023. vol. 10. no. 1. pp. 1–18.
10. Hoang M., Bihorac O.A., Rouces J. Aspect-based sentiment analysis using Bert. Proceedings of the 22nd nordic conference on computational linguistics. 2019. 187–196.
11. Bataa E., Wu J. An investigation of transfer learning-based sentiment analysis in Japanese. arXiv preprint arXiv:1905.09642. 2019.
12. Lv H., Liu J., Wang H., Wang Y., Luo J., Liu Y. Efficient hybrid generation framework for aspect-based sentiment analysis. Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics. 2023. pp. 1007–1018.
13. Chen C., Teng Z., Wang Z., Zhang Y. Discrete opinion tree induction for aspect-based sentiment analysis. Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics. 2022. vol. 1. pp. 2051–2064.
14. Esuli A., Sebastiani F. Determining the semantic orientation of terms through gloss classification. Proceedings of the 14th ACM international conference on information and knowledge management. 2005. pp. 617–624.
15. Cambria E., Havasi C., Hussain A. SenticNet 2: A semantic and affective resource for opinion mining and sentiment analysis. Proceedings of the Twenty-Fifth International Florida Artificial Intelligence Research Society Conference. 2012. pp. 202–207.
16. Xiaomei Z., Jing Y., Jianpei Z., Hongyu H. Microblog sentiment analysis with weak dependency connections. Knowledge-Based Systems. 2018. vol. 142. pp. 170–180.
17. Appel O., Chiclana F., Carter J., Fujita H. Successes and challenges in developing a hybrid approach to sentiment analysis. Applied Intelligence. 2018. vol. 48. pp. 1176–1188.
18. Yin C., Chen S., Yin Z. Clustering-based Active Learning Classification towards Data Stream. ACM Transactions on Intelligent Systems and Technology. 2023. vol. 14. no. 2. pp. 1–18.
19. Naseri S., Dalton J., Yates A., Allan J. CEQE to SQET: A study of contextualized embeddings for query expansion. Information Retrieval Journal. 2022. vol. 25. no. 2. pp. 184–208.
20. Sobkowicz P., Kaschesky M., Bouchard G. Opinion mining in social media: Modeling, simulating, and forecasting political opinions in the web. Government information quarterly. 2012. vol. 29. no. 4. pp. 470–479.
21. Hu Y.H., Chen Y.L., Chou H.L. Opinion mining from online hotel reviews–a text summarization approach. Information Processing and Management. 2017. vol. 53. no. 2. pp. 436–449.
22. Yousfi S., Rhanoui M., Mikram M. Comparative study of CNN and LSTM for opinion mining in long text. Journal of Automation, Mobile Robotics and Intelligent Systems. 2020. pp. 50–55.
23. Ethnologue Languages of the World [online]. Available at: https://www.ethnologue.com/statistics/size (accessed 01.09.2023).
24. Sultana J., Rani M.U., Farquad M.A.H. Knowledge discovery from recommender systems using deep learning. International Conference on Smart Systems and Inventive Technology (ICSSIT). 2019. pp. 1074–1078.
25. Sultana J., Jilani A.K. Predicting breast cancer using logistic regression and multi-class classifiers. International Journal of Engineering and Technology. 2018. vol. 7. no. 4(20). pp. 22–26.
26. Sultana J., Nagalaxmi G. How Efficient is Apriori: A Comparative Analysis. International Journal of Current Engineering and Scientific Research. 2015. pp. 2393–8374.

27. Naidu R., Bharti S.K., Babu K.S., Mohapatra R.K. Sentiment analysis using telugu sentiwordnet. International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET). 2017. pp. 666–670.
28. Garapati A., Bora N., Balla H., Sai M. SentiPhraseNet: An extended SentiWordNet approach for Telugu sentiment analysis. International Journal of Advance Research, Ideas and Innovations in Technology. 2019. vol. 5. no. 2. pp. 433–436.
29. Koppula N., Rani B.P., Srinivas Rao K. Graph-based word sense disambiguation in Telugu language. International Journal of Knowledge-based and Intelligent Engineering Systems. 2019. vol. 23. no. 1. pp. 55–60.
30. Sultana J. Telugu News Data Classification Using Machine Learning Approach. Handbook of Research on Advances in Data Analytics and Complex Communication Networks. 2022. pp. 181–194.
31. Janardana Naidu G., Seshashayee M. Sentiment analysis for Telugu text using cuckoo search algorithm. Smart Computing Techniques and Applications: Proceedings of the Fourth International Conference on Smart Computing and Informatics. 2021. vol. 2. pp. 253–257.
32. Suryachandra P., Venkata P., Reddy S. Machine Learning Approach to Classify the Sentiment Value of Natural Language Processing in Telugu Data. Journal of Engineering and Applied Sciences. 2020. vol. 15. pp. 3593–3598.
33. Tammina S. A hybrid learning approach for sentiment classification in Telugu language. International Conference on Artificial Intelligence and Signal Processing (AISP). 2020. pp. 1–6.

Найду Дж Джанардана — ученый-исследователь, факультет компьютерных наук, Институт технологий и менеджмента Ганди. Область научных интересов: информатика. jana.766@gmail.com; Ганди Нагар, Рушиконда, 530045, Вишакхапатнам, Эндра Прадеш, Индия; р.т.: +91(0891)284-0501.

Сешашаяи М — Ph.D., Dr.Sci., доцент, факультет компьютерных наук, Институт технологий и менеджмента Ганди. Область научных интересов: информатика. mseshashayee@gmail.com; Ганди Нагар, Рушиконда, 530045, Вишакхапатнам, Эндра Прадеш, Индия; р.т.: +91(0891)284-0501.