Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 12, No.4, December : 2024

A DUAL WEIGHTED HYBRID OPTIMIZATION MODEL FOR SENTIMENT ANALYSIS IN ARTIFICIAL INTELLIGENCE

Sunil Kumar Pal[1], Amarnath Awasthi [2]
Assistant Professor Axis Institute of Higher Education,Kanpur,U.P, INIDA

Abstract— A specialized activity in natural language processing (NLP), opinion mining is used to glean insights from unstructured online data. User’s aspect-based internet evaluations have become a reliable source of information for company and consumer purchase decisions in the age of artificial intelligence. Data preprocessing, aspect sentiment extraction, aspect term grouping, and aspect sentiment classification are the four main components of the novel AI-driven aspect/feature-based opinion mining methodology proposed in this research. To standardize user evaluations, the preprocessing stage uses sophisticated natural language processing (NLP) techniques such tokenization, lemmatization, stemming, and stop-word removal. The two stages of aspect sentiment extraction are rule-based lexicon extraction and part-of-speech (POS) tagging. The traits that have been detected are divided into major and secondary components using semantic similarity score. A weighted function is then used to further improve these qualities. A new hybrid optimization model called the Particle with Levy Updated Sparrow Search Algorithm is presented in order to improve the weighting parameters. The Sparrow Search Algorithm (SSA) and conventional Particle Swarm Optimization (PSO) are used in this model to improve classification accuracy and efficiency. The modified characteristics are then processed by a neural network classifier, which divides opinions into three categories: positive, neutral, and negative feelings. By integrating feature extraction, semantic analysis, and optimization strategies for accurate sentiment classification, this AI-enhanced platform exemplifies a methodical approach to opinion mining.

Keywords— *Aspect term Extraction, Machine Learning, Reinforcement Learning, Pattern Recognition, Artificial Neural Networks, Cognitive Computing.*

I. Introduction

People all around the world use the Internet every day to express their opinions and observations on different goods and services. These viewpoints have a significant impact on how decisions are made. As a result, the Internet has developed into a vast collection of various viewpoints and comments. However, for successful analysis, these views must be divided into domain-specific aspects like "items," "services," and "environments" because they are frequently expressed in complicated phrases and unstructured text forms. This methodical assessment makes it easier to obtain important data. A review that reads, everything about the establishment was so lively and colorful—ample menu choice, helpful service, great decor—I LOVED the entire old Europe decor vibe—everything was so vibrant and colorful, for example, incorporates several elements into a single sentence without following a set order. It is impossible to manually handle such large amounts of data, which makes it challenging to extract important insights. This problem is addressed by data mining techniques, which provide ways to uncover hidden information in unstructured text and display it as correlations and patterns. Businesses and customers like gain from these insights as they make it possible for them to accurately gauge public opinion on a good or service. Opinion mining, another name for sentiment analysis, is a popular method for determining how the general public feels about goods and services. Document-level, sentence-level, and aspect/feature-level analysis are the three levels at which opinion mining functions. Sentiment analysis at the document level establishes if a document's general tone is favorable or negative. By examining the proportion of favorable and unfavorable reviews, it gives consumers a comprehensive overview of opinions. Conversely, phraselevel sentiment analysis assesses each sentence in a text to ascertain if it expresses a favorable or unfavorable attitude. By utilizing word frequency data and choosing pertinent attributes, this method improves the level of detail in opinion extraction. Sentiment analysis at the aspect/feature level categorizes reviews according to particular characteristics. Finding implicit aspects and aspect words has received little attention, despite the fact that there are several techniques for aspect-based opinion classification. Unrelated terms frequently make aspect-based categorization more difficult and lower the accuracy of machine learning-based classification methods found in user evaluations. The development of an automated sentiment analysis framework that integrates optimal machine learning approaches is therefore urgently needed.

The planned study's main contributions are as follows:-

- Particle using Levy Updated Sparrow Search Algorithm (PLUSA) weight optimization for main and secondary characteristics.
- The Sparrow Search Algorithm (SSA) and Particle Swarm Optimization (PSO) approaches are combined in this hybrid model. The Evaluation of performance using a variety of evaluation indicators.

The next section of this paper is organized as follows:

- Section 2 reviews related works on sentiment classification.
- Section 3 elaborates on the proposed approach, including an overview, preprocessing techniques such as tokenization, lemmatization, stemming, and stop word removal, aspect-based sentiment extraction, and the proposed neural network model for sentiment analysis. Section 4 presents the results obtained using the proposed framework, while Section 5 provides the conclusion.

II. LITERATURE REVIEW

Different researchers worked in the field of aspect-based sentiment analysis (ABSA). Some of their contributions and research studies are mentioned in this section. In 2021, Pathan et al. [1] have projected a Sentiment Intensity Lexicon-integrated "Attention-based position-aware Bidirectional Long Short-Term Memory network" for Aspect-Based Opinion Mining (ABOM). To get the aspect vector closer to its nearest sentimentally and semantically related neighbors, they made adjustments to its pre-trained vector. The model's aspect weights were determined by fusing word embedding, location data and lexical sentiment intensity ratings as external knowledge. The projected model has been tested with SemEval 2014 dataset. In 2019, using the Deep Learning Method for Recommender System (AODR), Da'u et al. [2] projected a weighted Aspect-Based Opinion Mining to extract the aspect of the product reviews. In AODR deep learning has been used by the authors have extracted the product's aspects as well as its underlying weighted user opinions from the text reviews. Then, to enhance the recommender system's performance, the extended Collaborative Filtering (CF) technique has been utilized. In 2019, a recommendation system that uses ABOM to improve the recommendation process has been proposed by Da'u et al. [3]. The ABOM and rating prediction are indeed the two major phases of the projected model. The Multichannel Deep Convolutional Neural Network (MCNN) has been utilized in the ABOM phase for extracting the aspects by computing users' sentiment polarities. For the overall rating Aspect-specific ratings were fed into a Tensor Factorization (TF) algorithm to estimate the overall rating. Finally, a comparative analysis was conducted to verify the effectiveness of the proposed approach. In 2019, Asghar et al. [4] have projected a new integrated framework with hybrid sentiment classification module for aspect-based opinion mining for assisting the summary generating process. An improvement in terms of precision, recall, and F-measure were observed when the projected model was compared to the current models. In 2021, Lai et al. [5] have projected a novel rating prediction method using the attention-based deep learning model named as Gated Recurrent Unit (GRU) with semantic aspects. A two phase approach has been projected, wherein the aspect features of review semantics and word attention mechanism were used. From users' reviews, they have extracted the important words using the bidirectional GRU neural network. Then, users' reviews were split into tokens and aspect-based semantic vectors were generated with Latent Dirichlet Allocation. On the basis of these aspect based semantic vectors, the user preference ratings were predicted with the XGBoost method.

III. PROPOSED METHODOLOGY

The methodology of the proposed work is discussed in this section. The overview of the ABOM. [9] which include pre-processing steps, aspect term sentiment extraction, aspect grouping and neural based opinion mining [16] is utilized in this section. Fig. 1 depicts the architecture of the proposed work.

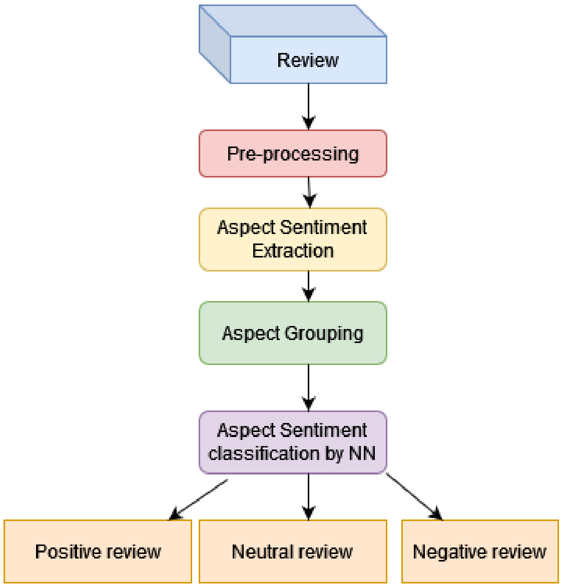


Fig 1 Architecture of the proposed work

A. *Proposed Aspect Based Opinion Mining: An Overview*

This research work introduces a novel ABOM [16] approach with four primary modules: pre-processing, aspect term sentiment extraction, aspect grouping, and classification of aspect sentiment. Let the collected reviews be denoted as. This initially enters the pre-processing phase, wherein is exposed to major pre-processing stages like Tokenization, Stemming, Lemmatization and Stop word removal methods. The data obtained from the pre-processing stage is the tokenized information and it is pointed from, the aspect sentiments are extracted by feeding it as input to the aspect-based sentiment extraction phase. This stage encapsulates two major phases: POS tagging and rule-based lexicon extraction [14][15]. The extracted aspects, as well as sentiments/ opinions from are pointed as and, respectively. The extracted aspects alone move to the stage of aspect grouping [17]. The weights of the retrieved are used in the recommended process for aspect grouping. Initially, the semantic similarity score for is determined, and the aspects with a greater similarity score are considered as secondary, while those with a lower similarity score are considered as primary aspect. A weight function is multiplied by both the primary aspect and secondary aspect to provide a more exact evaluation. Using a newly established hybrid optimization model referred to as PLUSA, the weighting factors of primary and secondary aspects are fine-tuned to make the classification more appropriate[18][19]. The dual weighting factor, the weight function of the primary aspect, and the weight function of the secondary aspect are fine-tuned. Finally, the features are the weight-optimized primary aspect, weight-optimized secondary aspect, and extracted, which are provided as input to the sentiment classification step. The classification of the sentiment [16] [17] is done by NN which classifies the sentiment as positive, neutral or negative.

B. Data Pre-Processing

Sentiment analysis has been widely fed by natural language utterances. There may well be incomplete texts as well as incorrect phrases when creating material using natural language. As a result, to perform sentiment classification, such phrases must be pre-processed in their native state. In data mining, the pre-processing is undergone to transform the collected raw data or reviews[17] into useful as well as efficient information. High accessibility to computing facilities encourages the development of vast volumes of electronic data in the current context. The expansion of data persuades academics to do critical analysis to extract the most potential patterns for better decision-making. Pre-processing of text to a more organized format is required for this type of analysis. Therefore, the collected reviews are preprocessed through the following steps:-

- Tokenization: It is the fundamental step of pre-processing which transforms raw data into words and phrases known as tokens before converting them to vectors.
- Stemming: It is a normalization approach, wherein the tokenized words are converted to their root words. The stemming is effective at lowering the number of calculations necessary; hence it is preferred in this study. The term 'Eating' stems from the word 'eat'.
- Lemmatization: The issue of stemming is resolved in lemmatization phase, wherein the cutting prefix/suffix from the word sometimes may not produce a meaningful word. Lemmatization is used to discover the tokens' base words. The lemmatization is done using morphological analysis and vocabulary to determine the word's basic form. Lemmatization is similar to stemming, except it considers the word's context. For example, the lemma of the 'raining' token is 'rain.'
- Removal of Stop word: The terms that are frequently used in the language are known as stop words and these words do not have any meaning. Stop words such as "the," "a," "an," and "in" are often used in the review and should be deleted to save computational load. At the end of the pre-processing stage, the tokenized data obtained is indicated as from, the opinion and aspects are retrieved in the aspect sentiment extraction phase.

C. Aspect Sentiment Extraction

Phase 1: POS-tagging is considered the salient phase in sentiment analysis. Using POS tagging [4], the aspects and opinions are collected from database. The technique of assigning parts of speech to words is known as POS tagging. Verbs, Nouns, Adverbs, Pronouns, Adjectives, Conjunctions, and their sub-groups are included in the POS tagger [4]. The POS tag will help to identify the nouns, adjectives, and verbs in the review sentence. Aspect and sentiment words were used to indicate the nouns/noun phrases and adjectives that resulted from the POS tag. Verbs and adverbs can also be used as sentiment words in specific instances. The Penn Treebank POS Tag was used as the basis for this work's POS tag.

For example, if the statement “The photographs are absolutely incredible" is fed to the POS tagger, the POS tagger will consider “pictures” as a noun and “amazing” as an adjective.

Phase 2: To extract the aspects effectively, rule-based lexicons are utilized as in [4]. In the example, “The photographs are absolutely fantastic”, the words “wonderful” and “images” refer to the opinion and aspect [8] [17], respectively. The retrieved aspects and related sentiments after POS tagging are stored for further processing. The abbreviation for the extracted opinions. Here denotes the extracted aspects, which are subsequently subjected to an aspect grouping strategy.

D. Proposed Aspect Grouping and Neural Network based Opinion Mining

The word “data” is plural, not singular. Grouping of the relevant aspects is the next step after extracting the aspects. Users usually use different words or phrases to express different points of view on the same subject. For example, in the sentence, “The battery life is very long”, the words "battery," "battery life," and "battery use" all refer to the same character in the mobile dataset "battery". To generate an accurate aspect-based opinion evaluation [6][18], these terms pertaining to the aspects ought to be grouped. The weights of the retrieved, which are derived depending on its SSC, are used in the proposed approach for aspect grouping. The SSC is calculated using the cosine similarity function in the Google news vector library, which determines the cosine of the angle between two vectors. Eq. (1) shows the cosine formula mathematically.

$$\cos = \frac{\sum_{i}^{n} A_i . B_i}{\sqrt{\sum_{i}^{n} (A_i .)^2} * \sqrt{\sum_{i}^{n} (B_i .)^2}} \quad (1)$$

where the topic document is represented by the notation and the review document is designated by the notation .

The aspects with a greater similarity score are considered to be secondary aspects, while those with a lower similarity score are said to be primary aspects, based on the computed findings. If the computed similarity score is more than 0.2, it is regarded as a secondary component in this study, while the others are considered primary. The product's exact review is provided in the secondary aspect. The weight function that comes inside the limit [0, 1] is multiplied by the obtained cosine similarity score to make the grouping more exact. Furthermore, a newly presented Particle with Levy Updated Squirrel Search Algorithm (PLUSA), which is a hybridization of traditional PSO and SSA, is employed to adjust the weight function of the primary aspect Wt^p and the weight function of the secondary aspect Wt^s .

E. NN for Opinion Mining-sentiment analysis

In this work, the NN classifier [10] is used to provide the final classification results. The NN [11] is trained with extracted aspect features. The input, hidden and output layer together makes up the NN. The input, hidden and output neurons in the input, hidden and output layers are denoted as $i = 1,2,\cdots N_i$ $h = 1,2,\cdots N_h$ and $o = 1,2,\cdots N_o$, respectively[7]. Here, N_i , N_h and N_o denotes the number of input, hidden and output neurons, respectively. The network model of NN is expressed mathematically in Eq. (2) –Eq. (4), respectively. In addition, $wg^N_{bias.h}$, $wg^r_{bias.o}$, $wg^N_{inp,h}$ and $wg^r_{h,o}$ denotes the $h's$ bias weight, $o's$ bias weight, weight from i^{th} to h^{th} and weight from h^{th} to o^{th} , respectively. The error function $F(er)$ which is the discrepancy between the expected output Out_{pre} and actual outputs Out_{act} is shown in Eq. (12). In Eq. (2) and Eq. (3), activation function is represented as AF .

$$HIDDEN = AF\left(wg^N_{bias,h} + \sum_{i=1}^{N_i} wg^N_{i,h} .F \right) \quad (2)$$

$$Out_{pre} = AF\left(wg^r_{bias.o} + \sum_{h=1}^{N_h} wg^r_{hid.o} .HIDDEN \right) \quad (3)$$

$$F(er) = \underset{\{wg^N_{bias.h}, wg^N_{i,h}, wg^r_{bias.o}, wg^r_{h.o}\}}{\arg\min} \sum_{out=1}^{N_{out}} \left| Out_{act} - Out_{pre} \right| \quad (4)$$

F. Proposed Hybrid optimization Model for Primary and Secondary Weight Tuning

In this work, a novel optimization model has been introduced for optimizing the weight of NN. The SSA and PSO model has been combined to formulate a novel optimization model named to as Particle with Levy Updated Sparrow Search Algorithm (PLUSA). Both the PSO model and the SSA are effective in solving complicated optimization algorithms with more convergence. Interestingly, on hybridizing these two algorithms, the convergence of the solutions still increases and aids in achieving the global optima. The foraging, anti-predation, and collective knowledge of sparrows served as inspiration for the development of the SSA. Furthermore, the PSO model was created using inspiration from the movements of schooling fish and bird flocks. On hybridizing the SSA and PSO, the PLUSA model has been developed. The various steps followed in the PLUSA model are manifested below:

Step 1-The population of N search agents pop is initialized. In addition, the maximal count of iterations, $\max^{iter}$ maximal count of producers PD and the maximal count of sparrows that perceive the danger SD are triggered.

Step 2- Initialize the position of search agent as per Eq. (5).

$$Y = \begin{bmatrix} Y_{1,1} & Y_{1,2} & Y_{1,3} & \cdots & Y_{1,d} \\ Y_{2,1} & Y_{2,2} & Y_{2,3} & \cdots & Y_{2,d} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ Y_{n,1} & Y_{n,2} & Y_{n,3} & \cdots & Y_{n,d} \end{bmatrix} \quad (5)$$

Here, n, d denotes the count of sparrows and dimension of the variables, respectively.

Step 3- Set the current iteration, $itr = 0$

Step 4- If $itr > \max^{iter}$ do

Step 5- Compute the fitness of every search agent. The fitness can be expressed as per Eq. (6).

$$Fit(Y) = \begin{bmatrix} Fit(Y_{1,1}) & Fit(Y_{1,2}) & Fit(Y_{1,3}) & \cdots & Fit(Y_{1,d}) \\ Fit(Y_{2,1}) & Fit(Y_{2,2}) & Fit(Y_{2,3}) & \cdots & Fit(Y_{2,d}) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ Fit(Y_{n,1}) & Fit(Y_{n,2}) & Fit(Y_{n,3}) & \cdots & Fit(Y_{n,d}) \end{bmatrix} \quad (6)$$

Here, $Fit(Y)$ denotes the fitness of the search agent.

Step 6- Rank the search agent based on their fitness score.

Step 7- Find the current best individual as well as the current worst individual.

Step 8- Set the alarm value $R2 = rand(1)$

Step 9- For $i = 1 : PD$

Step 10- Update the position of the solutions using the newly proposed levy based particle updating model of PSO algorithm, which is shown in Eq. (6). This is the phase where we contribute. The mathematical expression concerning the newly projected levy based update expression is shown in Eq. (7).

$$Y_{i,d}^{t+1} = Levy(\beta) * Y_{i,d}^{t} + V_{i,d}^{t+1} \quad (7)$$

Here, $Y_{i,d}^{1}$ points to the search agent's position corresponding to the current iteration and $V_{i,d}$ represents the velocity of the search agent. Mathematically, the velocity function can be calculated as per Eq. (8)

$$V_{i,d}^{t+1} = \alpha . V_{i,d}^{t} + C_1 * \chi_1 * \left(Y_{i,d}^{pbest} - Y_{i,d}\right) + C_2 * \chi_2 * \left(Y^{pbest} - Y_{i,d}^{pbest}\right) (8)$$

Here, C_1, C_2 are the acceleration coefficient, and χ_1 as well as χ_2 are randomly generated values. In addition, the inertial weight is pointed as α and Y^{pbest} is the best position of the search agent. In this research work, α is varied from 0.2, 0.4, 0.6 and 0.8, and the respective optimal outcomes acquired are noted.

Step 11- End for

Step 12- For $i = (PD + 1) : n$

Step 13- Update the position of the solutions using the new updated model of the scrounger's position update as shown in Eq. (9). In fact, certain scrounger more frequently monitors the producers. They instantly leave their present place to compete for food once they discover that the producer has located nice food. If they win, they will be able to instantly obtain the producer's food.

$$X^{t+1} = (1 - \varpi) + \varpi . rand \quad (9)$$

Where in, $\varpi = \dfrac{\max^{iter} + 1}{\max^{iter}}$

Here, $rand$ is a random number generated using the tent map. The tent map helps to improve the solutions' convergence.

Step 14- End for

Step 15- For $i = 1 : SD$

Step 16- Update the position of the solutions using Eq. (10). These sparrows, which are presumably aware of the threat, make up 10% to 20% of the total population in the simulation experiment. These sparrows' beginning placements are produced at random in the population.

$$Y_{i,j}^{iter+1} = \begin{cases} Y_{best}^{iter} + \gamma \left| Y_{i,j}^{iter} - Y_{best}^{iter} \right| & if\ f_i > f_g \\ Y_{i,j}^{iter} + K \dfrac{Y_{i,j}^{iter} - Y_{worst}^{iter}}{(f_i - f_w) + \varpi} & if\ f_i = f_g \end{cases} \quad (10)$$

Here, γ denotes the step size, f_i, the fitness value corresponding to the present sparrow and f_w and f_g denotes the fitness of the worst and best search agents, respectively.

Step 17- End for

Step 18- Acquire the new current location

Step 19- If the newly acquired location is superior than the existing one, then update it

Step 20- $itr = itr + 1$

Step 21- End while

Step 22- Return the position of global fitness f_g and the best search agent X_{best}

Finally, the weight optimized primary aspect $\left(Wt_{opt}^{p}\right)$, weight optimized secondary aspect $\left(Wt_{opt}^{s}\right)$ as well as extracted $R^{opinion}$ are the features F fed as the input to the classification phase. This is shown mathematically in Eq. (11).

$$F = Wt_{opt}^{p} + Wt_{opt}^{s} + R^{opinion} \quad (11)$$

The algorithm for PLUSA model is as follows:

Initialize pop, $\max^{iter}$ PD, SD

Initialize the search agent's position

set $itr = 0$

If $itr > \max^{iter}$ do

Determine each search agent's level of fitness

The fitness can be expressed as per Eq. (4).

Rank the search agent based on their fitness score.

Find the current worst individual as well as current best individual.

set $R2 = rand(1)$

For $i = 1 : PD$

Update the position of the solutions using Eq. (6)

End for

For $i = (PD + 1) : n$

Update the position of the solutions using Eq. (8)

End for

For $i = 1 : SD$

Update the position of the solutions using Eq.(9)

End for

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 12, No.4, December : 2024

Acquire the new current location
If the newly acquired location is superior than
the existing one, then update it
$itr = itr + 1$
End while
Return the position of the best search agent X_{best}
and global fitness f_g

IV. RESULTS AND DISCUSSION

The findings of the research project are covered in this section. Section *A* discusses the simulation procedure. The analysis on proposed aspect classification model performance on dataset is discussed in section *B* The analysis on Projected Model for varying Acceleration Coefficient α is covered in section *C* Section *D* utilized the statistical analysis of the proposed work.

A. *Simulation procedure*

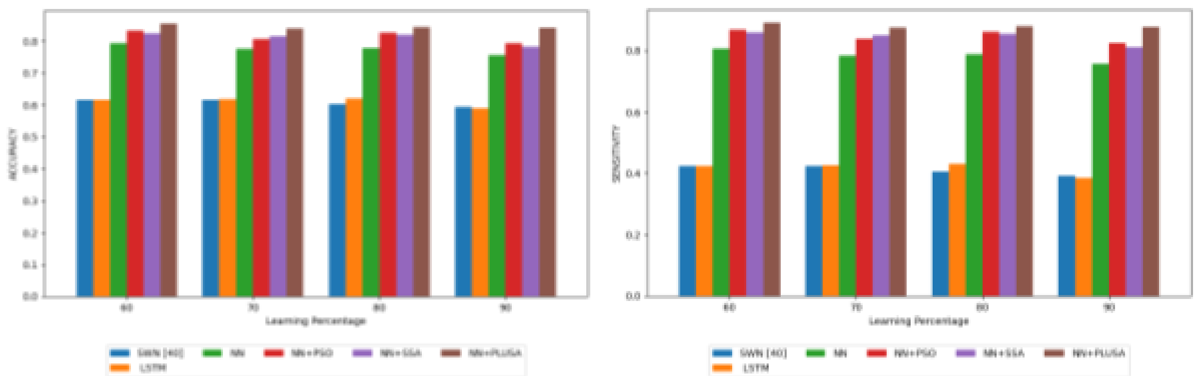
The implementation of the proposed work is performed in Python. The Semeval 2014 dataset of laptop domain is used in this work. The proposed work (PLUSA+ NN) is used to classify the reviews. We had utilized sequential model with relu activation function and batch size of 100. The performance of PLUSA+ NN was tested against known like Sentiwordnet (SWN) [4], Long-Short Term Memory (LSTM)), NN, PSO+NN, SSA+NN models by changing the LP to 60, 70, 80, and 90, respectively to validate its performance. The proposed work's performance is measured on the basis of different performance metrics which include Accuracy, Specificity, Precision, False Positive Rate (FPR), False Negative Rate (FNR), Negative Predictive Value (NPV), FDR, F1-score, and Mathews Correlation Coefficient (MCC).

B. *Analysis on Proposed Aspect Classification Model's Performance for Laptop dataset*

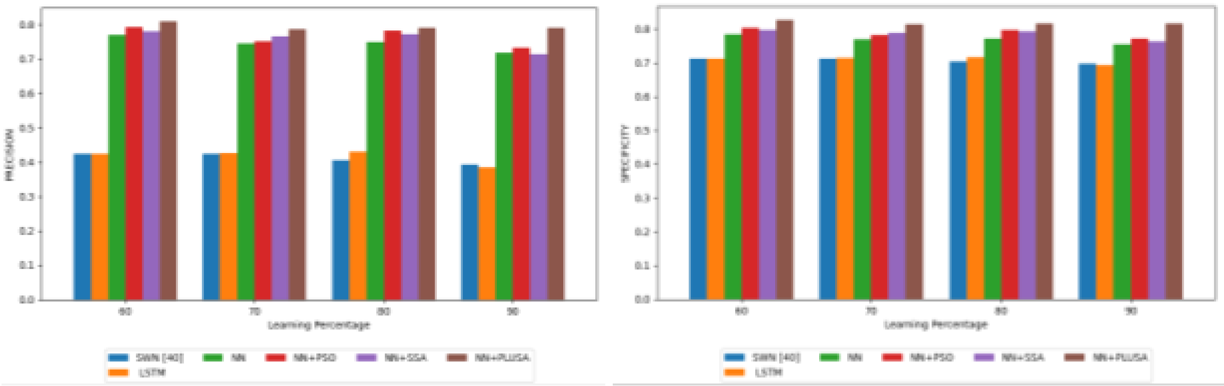
The comparison of the suggested model with the standard models is shownin Fig. 2 - Fig. 4. The comparison in terms of accuracy, sensitivity, precision and specificity is depicted in fig 2. The performance of the models in terms of FNR and FPR is represented in Fig 3.

Fig 4 shows the model's performance in terms of MCC and NPV. The analysis is conducted for several TP values, namely 60, 70, 80, and 90. On examining the graphs, the accuracy score of developed approach at 60th TP is 85.48, whereas, at 90th TP, the generated model now has an improved accuracy of 84.35. Likewise, for the positive measures, the values increase with increase in TP. Similarly, the negative measures of adopted model expose better value than other models. Especially, excellent outputs are attained at 90th TP for proposed as well as extant methods.

However, the developed model has exposed more effective outputs than the compared schemes for all TPs. Thus, the analysis established the enhanced efficacy of PLUSA + NN classifier method with the better classified outputs.

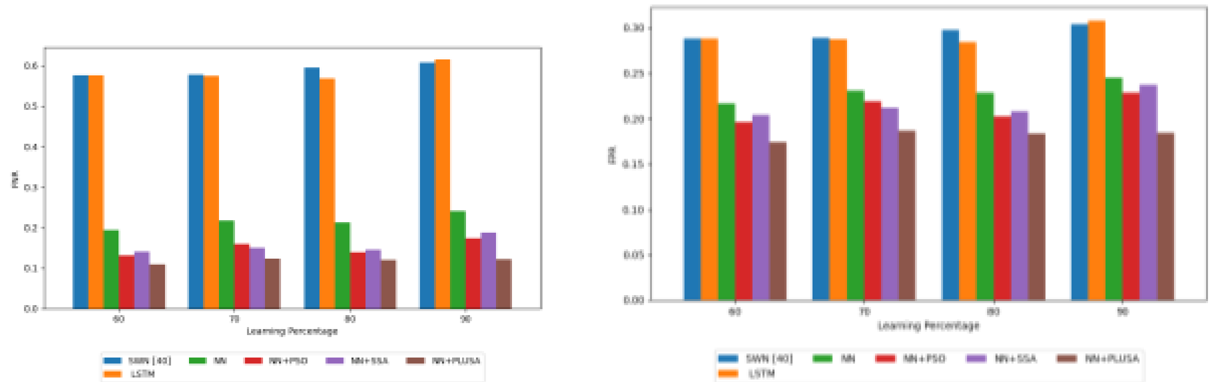


(a) (b)



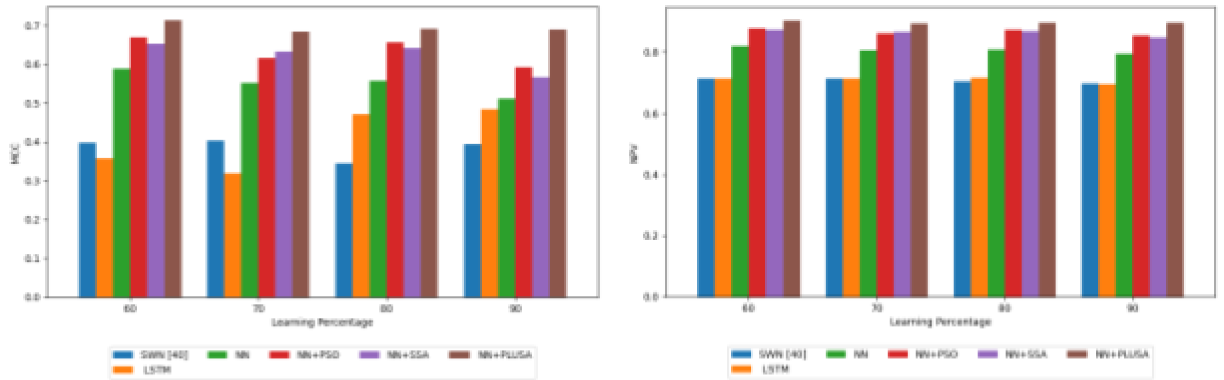
(c) (d)

Fig 2. Performance of developed approach over existing approaches regarding (a) Accuracy, (b) Sensitivity, (c) Precision and (d) Specificity



(b) (b)

Fig 3. Performance of developed approach over existing approaches regarding (a) FNR and (b) FPR



(a) (b)

Fig 4. Performance of developed approach over existing approaches regarding (a) MCC and (b) NPV

C. Analysis on Projected Model for varying Acceleration Coefficient α

In this research work, α is varied from 0.2, 0.4, 0.6 and 0.8, and the respective optimal outcomes acquired are noted. The results acquired are manifested in Table 1 – Table 4, respectively. On observing the outcome, the proposed model has achieved the highest accuracy value at $\alpha = 0.8$

TABLE I. PERFORMANCE ANALYSIS OF PROPOSED MODEL AT VARYING VALUES OF ACCELERATION COEFFICIENT AT LEARNING PERCENTAGE=60

Measures	**α= 0.2**	**α= 0.4**	**α= 0.6**	**α= 0.8**
Sensitivity	0.919510	0.920188	0.918755	0.920591

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 12, No.4, December : 2024

Specificity	0.919088	0.919298	0.918857	0.919424
Accuracy	0.919205	0.919545	0.918829	0.919749
Precision	0.813065	0.814458	0.811516	0.815287
F-Measure	0.863018	0.864101	0.861812	0.864745
MCC	0.809098	0.810293	0.807769	0.811005
NPV	0.967569	0.967658	0.967472	0.967711
FPR	0.080912	0.080702	0.081143	0.080576
FNR	0.080490	0.079812	0.081245	0.079409

TABLE II. PERFORMANCE ANALYSIS OF PROPOSED MODEL AT VARYING VALUES OF ACCELERATION COEFFICIENT AT LEARNING PERCENTAGE=70

Measures	α= 0.2	α= 0.4	α= 0.6	α= 0.8
Sensitivity	0.920791	0.919077	0.919940	0.918985
Specificity	0.919487	0.918955	0.919221	0.918927
Accuracy	0.919851	0.918989	0.919421	0.918943
Precision	0.815699	0.812175	0.813949	0.811987
F-Measure	0.865065	0.862326	0.863705	0.862179
MCC	0.811360	0.808334	0.809856	0.808173
NPV	0.967738	0.967513	0.967626	0.967502
FPR	0.080513	0.081045	0.080779	0.081073
FNR	0.079209	0.080923	0.080060	0.081015

TABLE III. PERFORMANCE ANALYSIS OF PROPOSED MODEL AT VARYING VALUES OF ACCELERATION COEFFICIENT AT LEARNING PERCENTAGE=80

Measures	α= 0.2	α= 0.4	α= 0.6	α= 0.8
Sensitivity	0.919918	0.917990	0.921715	0.918269
Specificity	0.919214	0.918626	0.919780	0.918710
Accuracy	0.919409	0.918452	0.920323	0.918589
Precision	0.813903	0.809949	0.817608	0.810521
F-Measure	0.863669	0.860592	0.866546	0.861037
MCC	0.809816	0.806425	0.813000	0.806915
NPV	0.967623	0.967374	0.967861	0.967410
FPR	0.080786	0.081374	0.080220	0.081290
FNR	0.080082	0.082010	0.078285	0.081731

TABLE IV. PERFORMANCE ANALYSIS OF PROPOSED MODEL AT VARYING VALUES OF ACCELERATION COEFFICIENT AT LEARNING PERCENTAGE=90

Measures	α= 0.2	α= 0.4	α= 0.6	α= 0.8
Sensitivity	0.915920	0.915920	0.913545	0.913545
Specificity	0.918016	0.918016	0.917342	0.917342
Accuracy	0.917449	0.917449	0.916329	0.916329
Precision	0.805730	0.805730	0.800919	0.800919
F-Measure	0.857298	0.857298	0.853533	0.853533
MCC	0.802810	0.802810	0.798696	0.798696
NPV	0.967117	0.967117	0.966832	0.966832
FPR	0.081984	0.081984	0.082658	0.082658
FNR	0.084080	0.084080	0.086455	0.086455

D. Stastical Analysis

The developed optimization algorithm has been examined five times due to its stochastic character, and the best outcomes are reported in terms of mean, median, best, worst, and standard deviation performance. The results achieved are shown in Table 5. On observing the acquired outcomes the proposed model has achieved the highest mean value as 91.8% (in terms of accuracy). Therefore, the projected model is said to be highly applicable for ABOM.

V. CONCLUSION

This study outlined four key components of a novel aspect-based opinion mining approach. Some methods for eliminating noisy data from user reviews include tokenization, lemmatization, stemming, and stop word removal. Aspect sentiment extraction was then carried out in two phases: rule-based lexicon extraction and POS tagging. The evaluated semantic similarity score is used to categorize the obtained characteristics into major and minor features. By assigning dual weights to the review's components, the mining becomes more accurate. Consequently, the major and secondary traits are multiplied by the weight function. To enhance the categorization in terms of opinion, the suggested model also adjusts the major and secondary weighting factors. A Levy Updated Squirrel Search for a New Particle.

REFERENCES

[1] Pathan, A. F., & Prakash, C. (2021). An attention-based position-sensitive framework for aspect-focused opinion analysis utilizing bidirectional long short-term memory networks. Journal of King Saud University-Computer and Information Sciences.

[2] Da'u, A., Salim, N., Rabiu, I., & Osman, A. (2020). Enhanced aspect-focused opinion evaluation through weighted modeling and deep learning for recommendation systems. Expert Systems with Applications, 140, 112871.

[3] Da'u, A., Salim, N., Rabiu, I., & Osman, A. (2020). Leveraging aspect-driven opinion mining with deep learning methodologies for recommender systems. Information Sciences, 512, 1279-1292.

[4] Asghar, M. Z., Khan, A., Zahra, S. R., Ahmad, S., & Kundi, F. M. (2019). A heuristic pattern-driven framework for aspect-oriented opinion mining. Cluster Computing, 22(3), 7181-7199.

[5] Lai, C. H., Liu, D. R., & Lien, K. S. (2021). Integrating XGBoost and aspect-focused review evaluation with attention neural networks for predicting user preferences. International Journal of Machine Learning and Cybernetics, 12(5), 1203-1217.

[6] Zuo, Y., Wu, J., Zhang, H., Wang, D., & Xu, K. (2018). A complementary model for aspect-oriented opinion mining. IEEE Transactions on Knowledge and Data Engineering, 30(2), 249-262.

[7] Abas, A. R., El-Henawy, I., Mohamed, H., & Abdellatif, A. (2020). A deep learning architecture for fine-grained aspect-based opinion evaluation. IEEE Access, 8, 128845-128855.

[8] Sindhu, I., Daudpota, S. M., Badar, K., Bakhtyar, M., Baber, J., & Nurunnabi, M. (2019). Evaluating faculty teaching performance through aspect-based opinion mining on student feedback. IEEE Access, 7, 108729-108741.

[9] Shams, M., Khoshavi, N., & Baraani-Dastjerdi, A. (2020). LISA: A language-agnostic framework for aspect-based sentiment analysis. IEEE Access, 8, 31034-31044.

[10] Ishaq, A., Asghar, S., & Gillani, S. A. (2020). Employing CNN and GA for hybridized aspect-based sentiment analysis. IEEE Access, 8, 135499-135512.

[11] Liu, H., Chatterjee, I., Zhou, M., Lu, X. S., & Abusorrah, A. (2020). A comprehensive review of deep learning methods for aspect-based sentiment analysis. IEEE Transactions on Computational Social Systems, 7(6), 1358-1375.

[12] Rida-E-Fatima, S., et al. (2019). A multi-layer dual-attention deep learning architecture using refined word embeddings for aspect-based sentiment analysis. IEEE Access, 7, 114795-114807.

[13] Kastrati, Z., Imran, A. S., & Kurti, A. (2020). A weakly supervised model for analyzing aspect-based sentiment in MOOCs reviews. IEEE Access, 8, 106799-106810.
[14] Wang, R., Zhou, D., Jiang, M., Si, J., & Yang, Y. (2019). Opinion mining advancements: From stance evaluation to product aspect analysis. IEEE Access, 7, 41101-41124.
[15] Cerqueira, T. L., Bertoni, F. C., & Pires, M. G. (2020). Optimizing opinion classification via fuzzy rule-based systems and instance genetic selection. IEEE Latin America Transactions, 18(07), 1215-1221.
[16] Zhang, B., Bai, Y., Zhang, Q., Lian, J., & Li, M. (2020). Mining influential opinion leaders in social networks through phased-clustering approaches. IEEE Access, 8, 31539-31550.
[17] Yu, L., Wang, L., Liu, D., & Liu, Y. (2019). Advanced computational intelligence models for fine-grained opinion mining in online reviews. IEEE Access, 7, 116900-116910.
[18] Lin, Y., Zhao, W., Li, Y., Zhang, H., & Zhou, Y. (2019). Identifying opinion target and term pairs from reviews using adaptive crowd labeling techniques. IEEE Access, 7, 100491-100500.
[19] Uddin, G., & Khomh, F. (2021). Automatic opinion extraction regarding APIs from Stack Overflow discussions. IEEE Transactions on Software Engineering, 47(3), 522-559.

***