International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
IJISAE
ISSN:2147-6799 www.ijisae.org Original Research Paper

Aspect Based Sentiment Classification with Various Feature Extraction and Selection Techniques using Machine Learning

[1]Ms. Anuradha N. Nawathe, [2]Dr. Avinash S. Kapse, [3]Dr. V. M. Thakare, [4]Dr. Arvind S. Kapse

Submitted: 26/10/2023 **Revised**: 18/12/2023 **Accepted**: 26/12/2023

Abstract: Aspect-based sentiment classification has various applications, such as analysing product reviews, social media sentiment analysis, or customer feedback analysis. It provides a more fine-grained understanding of sentiment by capturing the sentiment expressed towards specific aspects, enabling businesses to gain deeper insights into customer opinions and preferences. After detection of aspects, it works on Sentiment classification (SA) which focuses on determining the sentiment polarity associated with each aspect. It involves assigning sentiment labels to each aspect. This step often utilizes machine learning algorithms. A properly selected feature set is key for aspect prediction. This thesis primarily offers three systems. In the first system, a novel methodology called "two phase weighted correlation feature selection" is proposed for identifying the important elements of the items under consideration. Linguistically linked features indicate the aspects as well as the sentiments towards it. On addition, the suggested system looked at the impact of linguistic rule-based features in the aspect prediction job. Several previous techniques treated ABSA as a binary class or multiclass problem, with each review phrase predicting only one class. Since a review may discuss various characteristics of an entity, the second method suggested in this thesis uses multilabel classifiers to handle ABSA. In the third system, sentiments for extracted aspects are determined using a RNN-based approach for an end-to-end ABSA. The different machine learning algorithm provides 90% accuracy while the deep learning provides 96% average accuracy with various cross validation. The thesis presents novel approaches for each subtask that outperform existing methods. Evaluation measures such as F1-score, accuracy per label, and hamming loss are used to assess the performance of these subtasks.

Keywords: *Aspect classification, feature extraction, sentiment classification, emotion detection, NLP, data parsing, social media*

1. Introduction

With the introduction of Web 2.0, web users now can share their opinions on a variety of social networking and retail websites. As a result, there has been an exponential increase in User Generated Content (UGC), which can aid in making sensible decisions regarding a product, brand, national or international event. Many firms, for example, can use UGC to continuously monitor their brand reputations, benefits, drawbacks, and adversaries (Nguyen H. and Nguyen M. 2019). However, mining UGC is difficult due to the number, variety, veracity, and speed with which it is generated online constantly. Researchers have been constantly working to design modelling frameworks that can aid in the mining of these information corpora utilizing Natural Language Processing (NLP), SA, Artificial Intelligence (AI), Information Extraction, and Information Retrieval (Al-Smadi et al. 2019). A few technological requirements and standards must be developed to make this context understandable and transparent.

AI models are being utilized to simulate the human brain, and so NLP is an important aspect of AI. Building numerous models for computers to interpret and interact with human language is part of NLP. SA is a sub-task in NLP that is focused on identifying a user's emotion, attitude conveyed towards an entity or an aspect. Splitting a document or sentence into tokens or words (tokenization), converting the tokens into their root forms (lemmatization), identifying various parts of speech (POS Tagging), and finally extracting sentiment polarity and magnitude of the opinion expressed about the entity or aspect are some of the standard techniques in NLP and SA. Information retrieval is extremely vital in SA, especially when it comes to finding the most important and relevant documents in a large text corpus. Information extraction allows you to extract the furthermost significant data from a given piece of text, such as the major points in a consumer review. Furthermore, The notion of subjectivity utilised while annotating texts has a big impact on the results. However, it was shown that removing objective statements from a document prior to polarity classification enhanced performance. Subsequent that, sentiment magnitude relates to the strength of detected sentiment, or how powerful or intense the shown sentiment is. The final subtask entails detecting various

[1]*Author and Research Scholar Sant Gadge Baba Amravati University, Amravati.*
anunawathe77@gmail.com
[2]*Co-Author and PhD Guide Sant Gadge Baba Amravati University, Amravati*
askapse@gmail.com
[3]*Co-Author, Professor Sant Gadge Baba Amravati University, Amravati Horizon vilthakare@gmail.com*
[4]*Co-Author, Professor Information Science and Engineering New College of Engineering, Bengaluru,*
arvind.kapse2021@gmail.com

emotions such as sadness, rage, excitement, and happiness.

This procedure aims to find and extract object characteristics that the perspective holder has remarked on, as well as to assess if the opinion is positive, negative, or neutral. Grouping feature synonyms create a feature-based overview of several reviews. Aspect extraction is a job that involves identifying aspects of an object and may be thought of as a knowledge extraction task in general. The sentiment categorization of various features indicates if whether views are positive, negative, or neutral. While lexicon-based techniques rely on a list of essential piece sentiment phrases as their primary source, the major challenge for active learning is determining the breadth of each emotion expression and whether it covers the aspect in words. All prevalent phrases throughout evaluations must be identified and sorted by criteria like "occurs immediately after emotion word" to find all aspect words contained in a statement (e.g. great food). Then we may create a list of terms that come up often. Another option is to plan ahead of time and look for all of the details in the reviews. Food, service, value, and décor are all possible aspects of a restaurant.

2. Literature Review

Authors Xinzhi Wang et al. Researchers' interest has been piqued by [1] suggested sentiment analysis. When it comes to artificial intelligence, most prior efforts have focused on sentiment recognition itself rather than sentiment recognition error mining. One reason why sentiment detection often falls short is because emotions tend to be intertwined. Using natural language text sourced from online news, we want to bridge the gap between sentiment detection and sentiment correlation mining. Human feeling cognitive bias is the primary source of the association between emotions, manifested as confusion and the development of emotion. Three different types of characteristics and two different deep neural-network models are provided to harvest sentiment correlation from sentiment recognition via text. An orthogonal foundation is used to derive the law of emotional confusion. We examine the sentiment development law from three angles: one-step transfer, limited-step transfer, and shortest-path transfer. Objective and subjective texts of varied lengths (long and short) are used to verify the method: 1) titles of news stories, 2) article bodies, and 3) comments. The findings of the experiment demonstrate that in subjective remarks, emotions are often misunderstood as rage. Love-anger and sadness-anger are two emotions that tend to circulate in response to comments. Fear and happiness may be easily identified in the news as love in the text. These results may provide light on the ways in which applications like network public sentiment analysis, social media communication, and human-computer interface might benefit from including emotional interaction.

Authors Renata L. Rosa et al. Proposed online social networks (OSN) [2] give useful data about people' perspectives on many topics. This information may then be analysed by applications like monitoring and recommendation systems (RS). In this research, we provide a Knowledge-Based Recommendation System (KBRS) that employs an emotional health monitoring system to identify signs of sadness and stress in its users. The monitoring findings trigger the ontology- and sentiment-based KBRS, which in turn sends users experiencing psychological disorders messages designed to make them happy, calm, relax them, or motivate them. Additionally, the solution provides a method for sending alert messages to approved individuals in the event that the monitoring system detects a disruption related to depression. Based on empirical evidence, the suggested KBRS was able to get a 94% extremely pleased user rating, whereas an RS using neither a sentiment measure nor ontologies was only able to achieve a 69% rating. Furthermore, subjective test results showed that the suggested approach makes little use of the resources available in modern mobile electronic devices (memory, processing, and energy).

When Hong-HanShuai et al. Online social networks (OSNs) have become a regular part of many people's life, as suggested by [3] due to the meteoric rise in popularity of social networking and messaging applications. The primary goal of social network mining studies is to uncover the insights hidden in data that can be used to better people's daily lives. Online social networks (OSNs) may have the opposite effect of what they intend, reducing users' in-person social interactions even as their virtual networks grow. As the use of mobile social networking applications spreads like a virus, new words have emerged to explain the need to keep checking them, such as Phubbing and Nomophobia.

Chih-Hua Tai et al. According to [4], people's daily routines have been drastically altered due to the popularity of social networking platforms like Facebook and Twitter. Many individuals now document their everyday activities online, including their regular use of social media for sharing thoughts, views, and ideas and interacting with loved ones. The resulting deluge of data opens up exciting prospects for learning about human behaviour and motivates the development of several novel tools for bettering the lives of individuals. Many studies in recent years have tried to decipher people's emotions by poring into their tweets, Facebook postings, and even their diaries. There may be a connection between words and sentiments, they have speculated. There have been a number of studies looking at how persons with certain diagnoses (such as serious depression and PTSD) express themselves in language and emotion. This allows for the development of important new fields like service satisfaction surveying and mental health treatment. Motivated by these kinds of ideas, we made the observation that people's online posts/diaries not only indicate their level of happiness with life and their relationships, but also hint to the presence of mental illness in their lives.

Logical and emotional stress is becoming a risk to peoples physical health, as discussed by Huijie Lin and Jia Jia [5]. More and more individuals are stressed out by the hectic pace of modern life. In 2010, New business reported on a global study that found that more than half of all respondents reported feeling much more stressed than they were in 2008. Stress is a normal, everyday part of life, but it may have negative effects on your body and mind if it becomes too intense or lasts too long. Existing studies have linked chronic stress to a wide range of illnesses, including severe depression, sleeplessness, and more. Suicide has surpassed traffic accidents as the leading cause of mortality among Chinese adolescents, and the Chinese Centre for Disease Control and Prevention attributes most of this to the effects of severe stress. These facts demonstrate that rising stress levels pose a serious threat to people's health and well-being. Consequently, it is crucial to identify signs of stress before they balloon into major issues. Conventional methods of detecting mental or emotional distress rely on either in-person interviews, self-report questionnaires, or wearable sensors. However, conventional approaches are often reactive, costly in terms of both time and resources, and emotionally charged. Is there a way to identify stress that is both immediate and preventative? The proliferation of social media is impacting not only people's daily lives but also medical and health-related studies. More and more individuals are comfortable talking about themselves and their feelings online because to the rise of social media platforms like Twitter and Sina Weibo2. These social media data provide a timely reflection of users' actual states and sentiments, opening up new possibilities for representing, measuring, modelling, and mining users' behaviour patterns across massive social networks; the theoretical foundation for this kind of social information can be found in the field of psychology. For instance, studies have shown that users who are under stress are less likely to engage in social activities, and there has been growing interest in using social media data to improve both mental and physical health diagnostic and treatment methods. The use of Twitter data for real-time illness monitoring has been suggested, and community-generated health data has been used to try to close the lexical gap between patients and doctors. There have also been studies that use the content of users' tweets on Twitter and other social media sites to gauge the emotional state of those individuals.

The work of Yefeng Wang et al. 25% of adult Americans have been identified with at least one psychological disorder, which is much higher than the global average. About 65% of all individuals in the United States now use social media sites like Facebook and Twitter to research and discuss health topics. Research on mental diseases has made use of social networking platforms that let users record and share their everyday lives with peers. Twitter's potential as a depression predictor was previously investigated by De Choudhury et al. The use of antidepressants was also addressed in the research, however the impact of nutritional supplements was disregarded. The Twitter habits of people who have self-diagnosed mental problems were compared to those of people in a control group by Coppersmith et al. Studies have focused on analysing tweet language patterns, but the impact of nutritional supplement usage among those with mental problems has been ignored. In light of these considerations, we set out to examine the Twitter data on the link between mental illness and nutritional supplement use.

Md. Araftur Rahman et al. The online social network (OSN) suggested by et al. [7] is a place where people may simply and quickly disseminate information. Users of OSNs are encouraged to be open and honest about themselves, their lives, their passions, and their relationships with others. The rapid expansion of OSNs like Facebook, Twitter, LinkedIn, Google+, and others may be directly attributed to the maturation of Internet infrastructure and software. Facebook, with over 1.86 billion monthly active users, is the biggest of these sites. Users that sign up for an OSN account may communicate with one another via the sharing of personal profiles, messages, status updates, images, and videos. Making friends on OSNs involves sharing private information with others, some of whom may be complete strangers. The open and welcoming nature of OSNs is enticing to its users, who feel comfortable sharing details about themselves and their networks. Sharing private information online, such as one's whereabouts, habits, or other details about oneself, is also risky.

Dr. Madhuri Siddula et al. Privacy in online social networks has emerged as a major issue recently. A social network's many problems, such as user anonymity, link transparency, and link characteristics, are all referred to by this term. Each facet of a social network's privacy is massive and broken down into a number of smaller issues. For instance, the concept of user privacy encompasses a wide range of related issues, such as the protection of sensitive user data and the anonymity of users' whereabouts. The purpose of this study is to provide a foundational introduction and first step for future academics interested in the topic of privacy in social networks. Among the concepts and techniques we offer for protecting user privacy are naïve anonymization, perturbation, and the creation of a whole new network. To demonstrate, previous efforts by a number of scholars, whereby social networks were described as network graphs including users and nodes.

Professor Yefeng Wang and colleagues. [9] sixty-five percent of all American people use social media sites like Facebook and Twitter for health-related news and advice. Research on mental diseases has made use of social networking platforms that let users record and share their everyday lives with peers. Twitter's potential as a depression predictor was previously investigated by De Choudhury et al. The use of antidepressants was also addressed in the research, however the impact of nutritional supplements was disregarded. The Twitter habits of people who have self-diagnosed mental problems were compared to those of people in a control group by Coppersmith et al. The effect of nutritional supplement usage among people with mental problems has been disregarded in these investigations, which have instead focused on studying patterns of language use in tweets. In light of these considerations, we set out to examine the Twitter data on the link between mental illness and nutritional supplement use.

Prof. Amir Hossein Yazdavar, Prof. et al. [10] show that depression is a serious global public health issue and a leading cause of disability. In 2011, around five percent of the global population had a depressive episode, as revealed by the World Mental Health Survey done in seventeen different countries. Additionally, it annually affects around 16 million people (6.7% of the population). Suicide and other harmful actions, such as substance abuse, are possible outcomes of severe depression that goes untreated or is inadequately managed. Suicide victims have been diagnosed with depression at a rate of above 90%. Identifying clinical depression by survey-based approaches, such as phone or online surveys, is part of a worldwide effort to combat the disorder. These methods mostly suffer from sampling biases (a limited set of respondents) and underrepresentation. In addition, there are issues with survey data because of delays between the collecting of data and the release of results, which means that the data only reflects the attitudes of participants as of a single point in time. In contrast, the explosive expansion of social media has led to an unprecedented quantity of data being willingly shared as individuals open up about their everyday experiences with mental health issues on sites like Twitter. This opens up possibilities for gaining fresh insight on these groups. Dr. Rafiqul Islam and co-authors [11] People's electronic communication and interaction has been revitalised by the advent of new media and communication tools, most notably online social networks. Social media users benefit from the ability to share their thoughts and opinions on any topic they choose, while professionals in the health field can learn more about a person's emotional state based on how they responded to a given topic.

Individuals, families, and, by extension, society as a whole are profoundly affected by mental illness, as Budhaditya Saha et al. [12] noted. Researchers studying textual indicators of mental health issues have a gold mine in social networks, where people with mental diseases may connect with others who share their experiences through online groups. It's not uncommon for people to have many mental health issues at once. Our efforts to categorise depression-related online networks centre on this co-occurring mental health condition. To do this, we have combed over 620,000 postings from 80,000 people across 247 social media forums. We have used topic and psycholinguistic feature extraction to provide our model with information about what people are talking about in their postings. In order to categorise online groups dealing with co-occurring mental health issues, we have developed a collaborative modelling framework using a machine learning approach. Finally, we conducted empirical validation of the model on the crawling dataset, finding that our model is superior than state-of-the-art baselines developed recently.

According to research given by Chun-Hao Chang et al. [13], mental health issues impact millions of individuals across all demographics of age, ethnicity, and location. There is a disturbingly high rate of undiagnosed cases and incorrect diagnoses due to the difficulty in detecting mental problems in people under distress. Our goal in this study is to develop prediction models that can tell if a user is suffering from one of two types of mental condition by analysing their language and behaviour patterns, especially as they appear in social media. Subconscious Crowdsourcing, a unique data gathering technique, allows for the creation of these prediction models by facilitating the acquisition of a more comprehensive and timely patient dataset. Our results imply that using trustworthy patient datasets to extract particular language patterns and social interaction aspects may significantly advance the study and diagnosis of mental diseases. According to a proposal in [14], depression is often considered the leading cause of disability across the world and one of the leading causes of death by suicide.

It was suggested by Kardelen Cepni et al. [15] that A new paradigm, the Social Internet of cars (SIoV) makes it possible for cars to form online friendships and communities. Social sensing is the practise of gathering information about a physical occurrence by the use of information generated by people in vehicles through the use of online social networks (OSNs). In this research, we investigate the Facebook user wall interaction-based core social sensing mechanism, the comment thread network (CTN), for SIoV. Facebook CTN becomes a social sensing medium for estimating an occurrence through social consensus when users view the contents their fellow commuters have posted about it and respond by commenting on or like such posts. This work is the first to examine Facebook CTN's social sensing capabilities, i.e. the precision of collective observations for SIoV. Since user perceptions and Facebook usage might influence observation signals, accuracy is dependent on both user traits and OSN features. We examine how Facebook CTN performs over a wide range of user actions, user connections, Facebook capabilities, and network sizes. According to the findings, the accuracy of the estimate signal, i.e. social consensus, is diminished by CTN's polarised weighting of the observations and its usage of less trustworthy post kinds. Additionally, user selection is anticipated to play a significant role in social sensing.

3. Research Methodology

3.1 system description

Similar to the well-known semantic gap, Aspect Based Sentiment Classification (ABSA) has an affective gap between low-level characteristics and the emotional content of a text conveying a certain sentiment. There is a large body of work that attempts to solve this problem by using techniques like neural networks and custom-built features. It provides a concise summary of some of the most seminal research on the topic. Many jobs involving social media analytics rely heavily on the results of sentiment analysis of user-generated material published online. Textual sentiment analysis has been heavily used by researchers to create algorithms for predicting political elections, measuring economic indicators, etc. [2][8]. Images and videos are becoming more popular means of communication among social media users. Predicting sentiment from visual material is complimentary to textual sentiment analysis since it may assist better extract user feelings towards events or subjects, such as those in text tweets. It has been said that a picture is worth a thousand words. The value increases exponentially when it comes to communicating feelings and thoughts. There is an abundance of evidence to back this up: The best images include strong emotional clues that make it easy for the viewer to relate to the subject of the shot. More and more individuals are turning to picture sharing sites like Twitter and Facebook to broadcast their feelings of elation, anger, and boredom. Healthcare, anthropology, communication studies, marketing, and many branches of computer science, including computer vision, are just some of the many fields that can benefit from automatically inferring emotion and sentiment information from such ever-growing, massive amounts of user-generated photos. Consider this: The state of one's emotional health affects many facets of their daily existence. Self-empathy is one such trait that is fostered, since it promotes a heightened awareness of one's own emotions. Additionally, it boosts confidence and toughness, making it easier to recover from mental and physical illness.

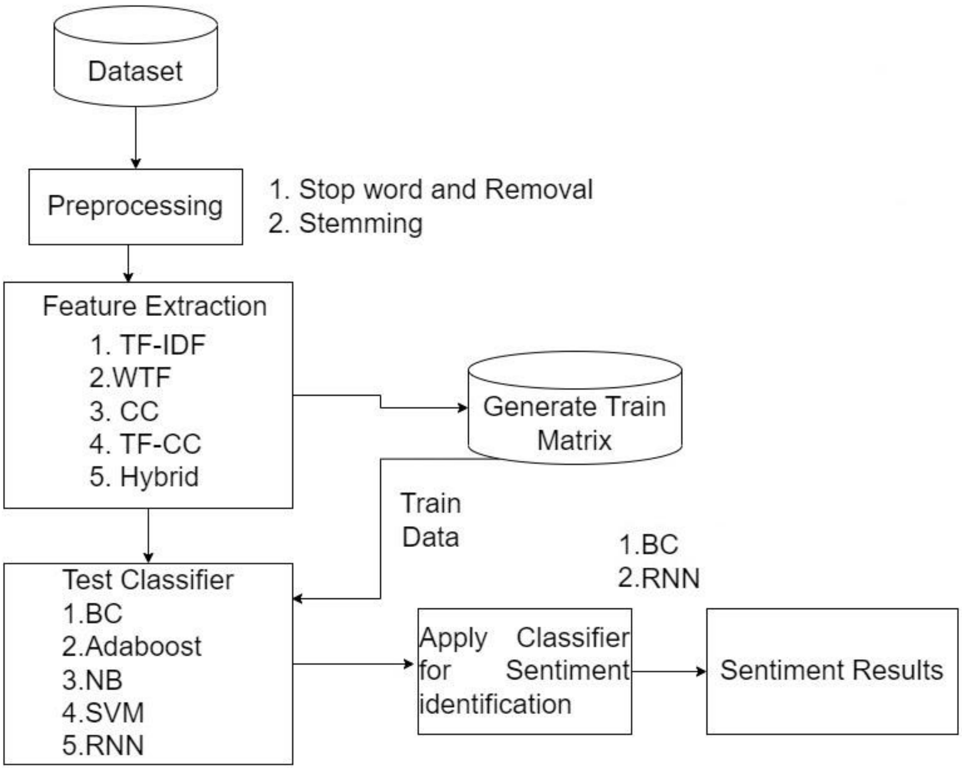

Fig 3.1: proposed system architecture for aspect based sentiment classification

1) Feature extraction:-

Preprocessing techniques are used to each phrase in the training dataset, including tokenization, casing-down, stop-word filtering, and stemming. You may find the often used stop word dictionary at https://gist.github.com/larsyencken/1440509. In the preprocessing phase, stemming and lemmatization play important roles in normalising features. The stemming technique returns all the impacted words to their original, stem form in the text. 'Studying' and 'Studies' become'study' and'studi,' respectively, as an example. Lemmatization is the process of reducing all word forms to their lemma. "studying" and "studies" will become "study" as a lemma, for instance. That's why we value lemma characteristics more highly than stemmed ones. In this study, we use feature selection techniques on data that has lemmas extracted as its features.[2][1].

2) Feature selection:-

In the system, we compare and contrast many feature selection methods and suggest a hybrid method. Methods for selecting relevant features are broken down into:

Feature selection in text refers to the process of selecting a subset of relevant features or attributes from a larger set of text-based data. It is a crucial step in natural language processing (NLP) and machine learning tasks involving text analysis, such as sentiment analysis, document classification, information retrieval, and text summarization.

Feature selection is important for several reasons:

1. Dimensionality reduction: Text data can be high-dimensional, especially when represented using techniques like bag-of-words or word embeddings. Feature selection helps reduce the number of features, which can improve computational efficiency and prevent overfitting.

2. Improved model performance: By selecting the most informative and relevant features, feature selection can enhance the performance of machine learning models. Irrelevant or noisy features can negatively impact the accuracy and generalizability of models.

There are various techniques for feature selection in text:

The choice of feature selection technique depends on the specific task, dataset, and desired outcome. It is often a combination of domain knowledge, experimentation, and evaluating the impact of feature selection on model performance.

a) Frequency (TF):-

Feature selection in this method is driven by the frequency of individual terms. The frequency of the terms describing each attribute is determined for each category of characteristics. For feature selection, a cut off is determined. In each area of consideration, features with a term frequency of 2 or higher are chosen. The output is a matrix with rows for each category and columns for each word (feature). A binary train matrix is constructed from this matrix, with 1 representing the frequency of terms that are not zero[7][8].

b) Weighted Term Frequency (WTF):-

The relative importance of each phrase is determined by (1) in this method. The importance of a term is its conditional probability, where X_(t,k) is the number of times term t appears in category k and X_t is the sum of the numbers of times term t appears in all categories. The significance of a phrase t grows if the frequency with which it appears in category k is higher than in other categories. For each class of consideration, a cutoff is determined. To create a binary train matrix, we choose the terms (features) whose weight is larger than the threshold. Kim Schouten et al. [2] also do the task of calculating the weight of a phrase. This paper, like [2], presents a correlation-based hybrid method for feature selection to reduce feature redundancy. Weights are used for feature selection and then to build a binary train matrix, as shown in [2], which is used to identify the aspect category of a test phrase.

$$\text{weight}(t) = \frac{X_{t,k}}{\mathrm{X}_t} \quad \text{----} \qquad (1)$$

c) Term Frequency with Correlation Coefficient (TF+CC):-

Features included in a classifier should be informative without being too redundant. The approach makes use of the word frequency matrix discovered in step (i). Relevant but unnecessary features may be gained by utilising this matrix. Therefore, in order to cut down on repetition, we determine how each characteristic correlates with the other features in the same dimension. One popular method for doing so is by the use of the Pearson correlation coefficient.

$$C0weight\ [t_i] = \frac{\mathrm{n}\ (\sum \mathrm{X}[]\mathrm{Y}[]) - (\sum \mathrm{X}[]) * (\sum \mathrm{Y}[])}{\sqrt{[\mathrm{n} \sum \mathrm{X}^2 - (\sum \mathrm{X})^2]}\ \sqrt{[\mathrm{n} \sum \mathrm{Y}^2 - (\sum \mathrm{Y})^2]}} \; -- \quad (2)$$

The term-to-term correlation is calculated using Eq. (2), where x[''] and y[''] are term frequency vectors including t_(i) and t_(i+1) respectively with respect to each aspect category. Each word t's correlation with the other terms in its category is averaged. A binary train matrix is constructed from terms with a correlation value of less than or equal to 0.85.

d) Weighted Term Frequency with Correlation Coefficient (WTF+CC):-

In this method, the weighted matrix generated in (ii) is utilised to create a new matrix that details how much weight a given phrase has in relation to each category of aspects. The weight of a word from each aspect category[1][2] is represented by the vectors x[] and y[] in Eq. (2), where x[] and y[] are vectors of terms t_(i) and t_(i+1) respectively. As described in (iii), the process culminates in the creation of a binary train matrix.

This study contributes by proposing a supervised method for extracting aspect categories, one that uses feature selection and correlation analysis to cut down on irrelevant information. Based on the data, we find that the f-score for the weighted word frequency with correlation method is much higher than the other methods[7, 8]. Here, characteristics chosen using weighted term frequency are preferable. By determining the association between characteristics in a given aspect category, redundancy may be eliminated.

Train module: Apply appropriate machine learning or deep learning algorithms to train the module. This involves feeding the preprocessed text data and corresponding labels into the chosen algorithm and optimizing its parameters using techniques like gradient descent or backpropagation. The specific training process will depend on the chosen algorithm and the task at hand.

Test Module: Module testing refers to the process of evaluating the performance and functionality of individual modules within a larger text processing system. It involves assessing how well each module performs its specific task and verifying if it meets the desired requirements and objectives. Module testing is an important step in ensuring the overall quality and reliability of the text processing system.

Prepare a representative set of test data that covers various scenarios and edge cases relevant to the module's task. This data should be distinct from the data used for training to assess the module's generalization ability. Verify if the module processes the input correctly and produces the expected output. Compare the module's output against the ground truth labels or annotations for evaluation. This can be done manually or by using automated testing frameworks.

Assess the module's performance using the defined evaluation metrics. This can involve calculating the accuracy, precision, recall, or other relevant metrics based on the module's output and the ground truth labels. It is important to analyze both the overall performance and performance on specific subsets of data (e.g., different domains, sentiment classes, etc.).

Evaluation and fine-tuning: Evaluate the trained module's performance using appropriate metrics such as accuracy, precision, recall, F1-score, or task-specific metrics. If the performance is not satisfactory, fine-tune the module by adjusting hyperparameters, trying different architectures, or augmenting the training data.

4. Results and Discussion

Experiment with various feature extraction selection techniques machine learning and deep learning classification algorithms

4.1 Experimental analysis with various machine learning methods

Table 1: classification accuracy with Term frequency approach feature extraction methods

	Accuracy	Precision	Recall	F-score
Binary classifier	0.6564	0.7600	0.7515	0.7857
Naive Bayes	0.7614	0.8683	0.8555	0.8619
Adaboost	0.8371	0.9583	0.8660	0.9098
SVM	0.8514	0.9867	0.8605	0.9193
RNN	**0.9420**	**0.9930**	**0.9560**	**0.9910**

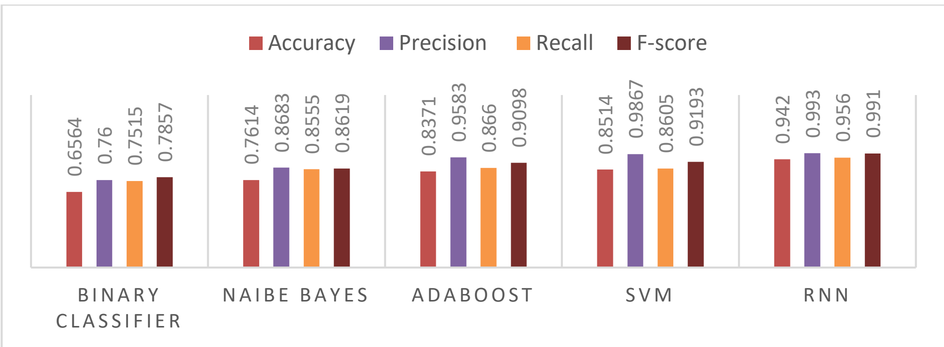

Fig 2 : classification accuracy with Term frequency approach feature extraction methods

Table 2 : classification accuracy with TF-CC approach feature extraction methods

Algorithm	Accuracy	Precision	Recall	F-score
Binary classifier	0.7610	0.8510	0.8641	0.8620
Naive Bayes	0.7520	0.8720	0.8490	0.8740
Adaboost	0.8235	0.8930	0.8588	0.9152
SVM	0.8440	0.9030	0.8730	0.9170
RNN	**0.9350**	**0.9145**	**0.9430**	**0.9836**

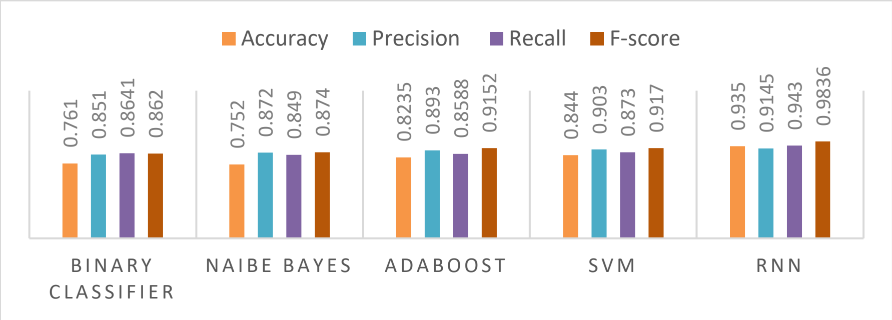

Fig 3: classification accuracy with TF-CC approach feature extraction methods

Table 3: machine learning based classification results using all dependency features

Algorithm	Accuracy	Precision	Recall	F-score
Binary classifier	0.9050	0.8735	0.9020	0.9090
Naive Bayes	0.8930	0.8770	0.9180	0.914
Adaboost	0.8760	0.8995	0.9085	0.9140
SVM	0.8850	0.8890	0.9360	0.9410
RNN	**0.9060**	**0.9130**	**0.9240**	**0.9630**

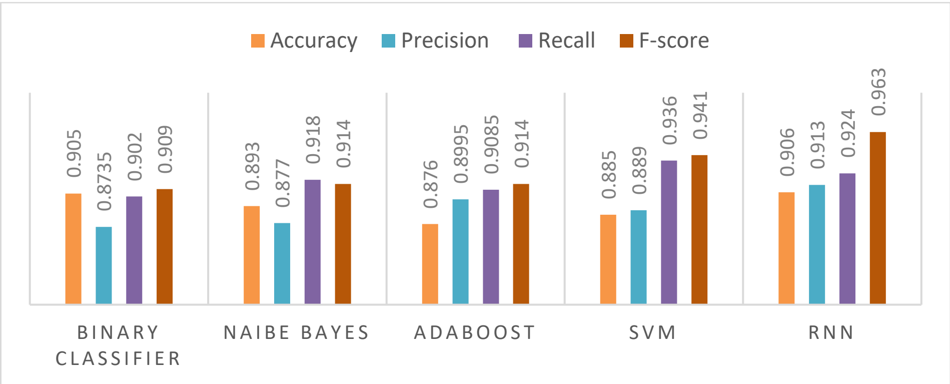

Fig 4: machine learning based classification results using all dependency features

Table 4: machine learning based classification results using all dependency features

Algorithm	Accuracy	Precision	Recall	F-score
Binary classifier	0.9098	0.9041	0.9160	0.9240
Naive Bayes	0.8995	0.8943	0.9230	0.9360

Adaboost	0.8870	0.9120	0.9180	0.9380
SVM	0.8971	0.9210	0.9370	0.9480
RNN	**0.9460**	**0.9975**	**0.9490**	**0.9685**

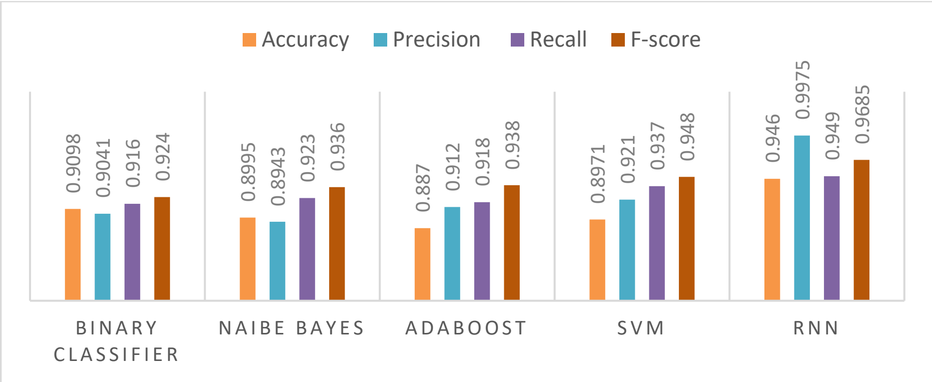

Fig 5: machine learning based classification results using all dependency features

Table 5: machine learning based classification results using hybrid feature selection methods

	Accuracy	**Precision**	**Recall**	**F-score**
Binary classifier	0.8050	0.8430	0.9040	0.9230
Naive Bayes	0.7930	0.8360	0.9230	0.9350
Adaboost	0.8560	0.9120	0.9050	0.9560
SVM	0.8950	0.9040	0.9230	0.9650
RNN	**0.9660**	**0.9360**	**0.9560**	**0.9880**

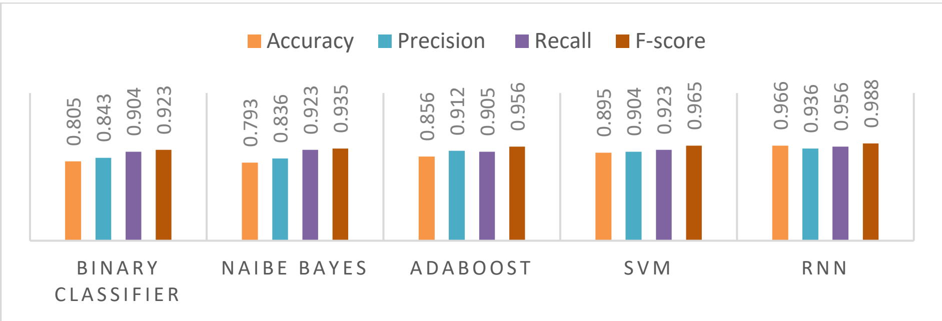

Fig 6: machine learning based classification results using hybrid feature selection methods

4.2 Various experiments using Deep Learning Methods

In our proposed methodology of predicting student academic performance using recurrent neural network-long short-term memory classification model, real-time large dataset of student is used. Three experimentations are performed to obtain, accuracy, precision, recall and f-score with various cross validation as follows

- Experimentation using RNN-LSTM (sigmoid) model
- Experimentation using RNN-LSTM (Tanh) model
- Experimentation using RNN-LSTM (ReLU) model

4.2.1 Experimentation using RNN-LSTM (sigmoid)

In this experiment of RNN-LSTM (sigmoid) model, accuracy, precision, recall and f-score with various cross validation are obtained. Figure 7, 8 and 9 illustrates the validation of model using 5-fold, 10-fold and 15-fold cross validation respectively using RNN-LSTM (sigmoid). Consider the following figure 7 which depicts the validation of model with 5-fold cross validation using RNN-LSTM (sigmoid) classifier.

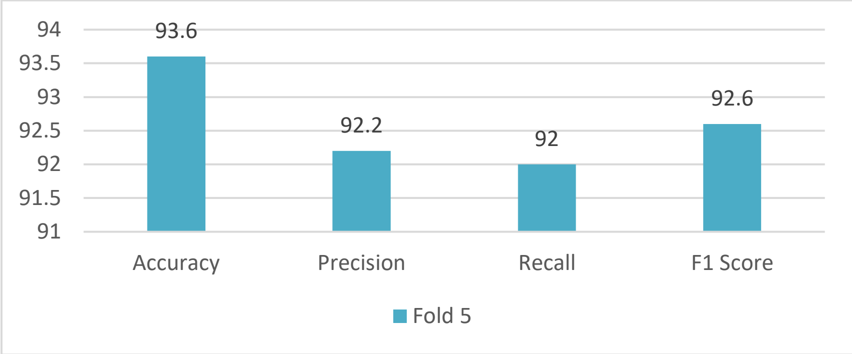

Fig 7: Validation of model with 5 fold cross validation using RNN-LSTM (sigmoid) classifier

Experimental findings of figure 7 shows by using 5-fold cross validation, the accuracy, precision, recall and f-score of RNN-LSTM (sigmoid) model is 93.6, 92.2, 92 and 92.6 respectively. Consider the following figure 8 which depicts the validation of model with 10-fold cross validation using RNN-LSTM (sigmoid) classifier.

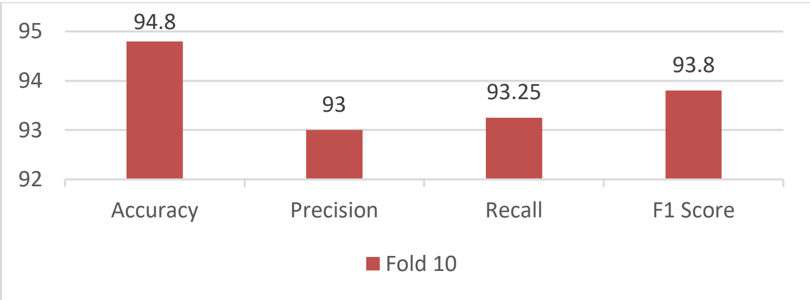

Fig 8: Validation of model with 10 fold cross validation using RNN-LSTM (sigmoid) classifier

Experimental findings of figure 8 shows by using 10 fold cross validation, the accuracy, precision, recall and f-score of RNN-LSTM (sigmoid) model is 94.8, 93, 93.25 and 93.8 respectively. Consider the following figure 9 which depicts the validation of model with 15 fold cross validation using RNN-LSTM (sigmoid) classifier.

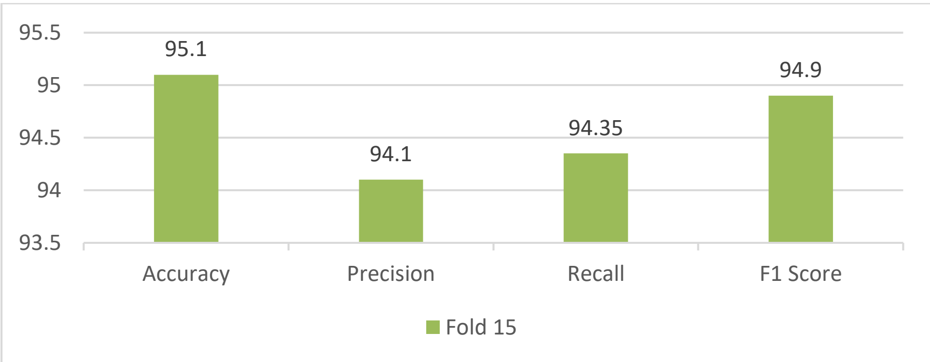

Fig 9: Validation of model with 15 fold cross validation using RNN-LSTM (sigmoid) classifier

Experimental findings of figure 9 shows by using 15 fold cross validation, the accuracy, precision, recall and f-score of RNN-LSTM (sigmoid) model is 95.1, 94.1, 94.3 and 94.9 respectively. As per experimental findings, 15-fold cross-validation has obtained the better average classification accuracy of 95.10%.

5.2.2 Experimentation using Recurrent Neural Network (Tan h) model

In this experiment of RNN-LSTM (Tan h) model, accuracy, precision, recall and f-score with various cross validation are obtained. Figure 10, 11 and 12 illustrates the validation of model using 5-5-fold, 10-fold and 15-fold cross validation respectively using RNN-LSTM (Tan h). Consider the following figure 10 which depicts the validation of model with 5 fold cross validation using RNN-LSTM (tan h) classifier.

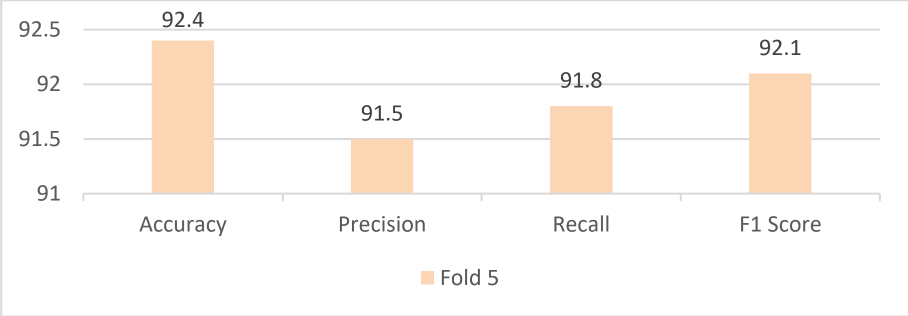

Fig 10: Validation of model with 5 fold cross validation using RNN-LSTM (Tan h) classifier

Experimental findings of figure 10 shows by using 5-fold cross validation, the accuracy, precision, recall and f-score of RNN-LSTM (Tan h) model is 92.4, 91.5, 91.8 and 92.1 respectively. Consider the following figure 11 which depicts the validation of model with 10-fold cross validation using RNN-LSTM (Tan h) classifier.

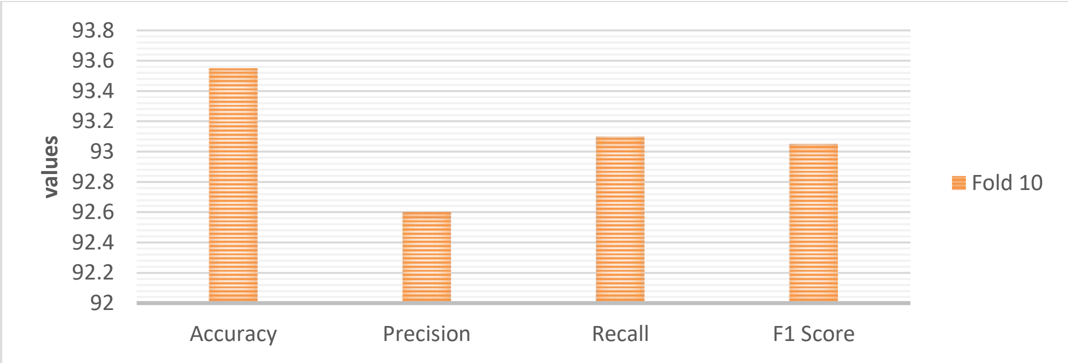

Fig 11: Validation of model with 10-fold cross validation using RNN-LSTM (Tan h) classifier

Experimental findings of figure 11 shows by using 10-fold cross validation, the accuracy, precision, recall and f-score of RNN-LSTM (Tan h) model is 93.55, 92.6, 93.1 and 93.05 respectively. Consider the following figure 12 which depicts the validation of model with 15-fold cross validation using RNN-LSTM (Tan h) classifier.

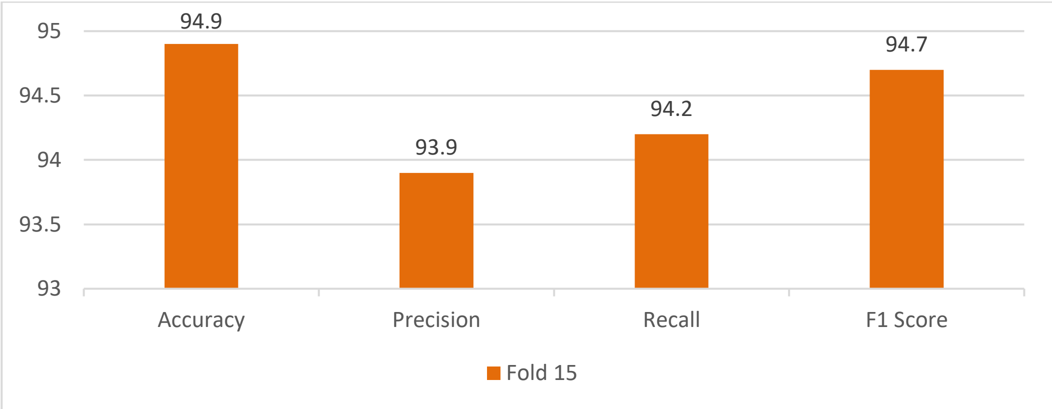

Fig 12: Validation of model with 15 fold cross validation using RNN-LSTM (Tan h) classifier

Experimental findings of figure 12 shows by using 15-fold cross validation, the accuracy, precision, recall and f-score of RNN-LSTM (Tan h) model is 94.9, 93.9, 94.2 and 94.7 respectively. As per experimental findings, 15-fold cross-validation has obtained the better average classification accuracy of 94.9%.

4.2.3 Experiment using Recurrent Neural Network (ReLU)

In this experiment of RNN-LSTM (ReLU) model, accuracy, precision, recall and f-score with various cross validation are obtained. Figure 13, 14 and 15 illustrates the validation of model using 5 fold, 10 fold and 15 fold cross validation respectively using RNN-LSTM (ReLU). Consider the following figure 13 which depicts the validation of model with 5-fold cross validation using RNN-LSTM (ReLU) classifier.

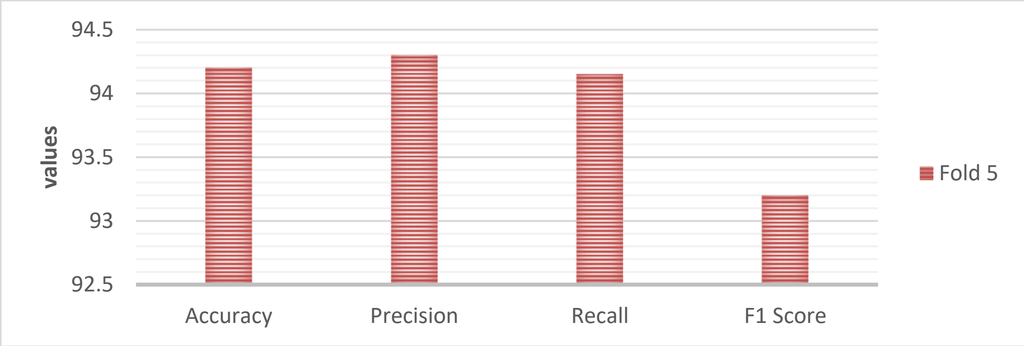

Fig 13: Validation of model with 5 fold cross validation using RNN-LSTM (ReLU) classifier

Experimental findings of figure 13 shows by using 5 fold cross validation, the accuracy, precision, recall and f-score of RNN-LSTM (ReLU) model is 92.4, 91.5, 91.8 and 92.1 respectively. Consider the following figure 14 which depicts the validation of model with 10 fold cross validation using RNN-LSTM (ReLU) classifier.

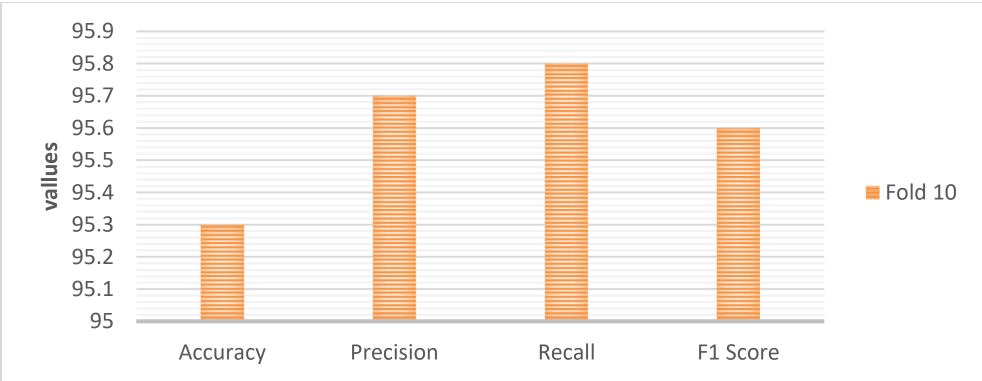

Fig 14: Validation of model with 10 fold cross validation using RNN-LSTM (ReLU) classifier

Experimental findings of figure 14 shows by using 10 fold cross validation, the accuracy, precision, recall and f-score of RNN-LSTM (ReLU) model is 93.55, 92.6, 93.1 and 93.05 respectively. Consider the following Figure 15 which depicts the validation of model with 15 fold cross validation using RNN-LSTM (ReLU) classifier.

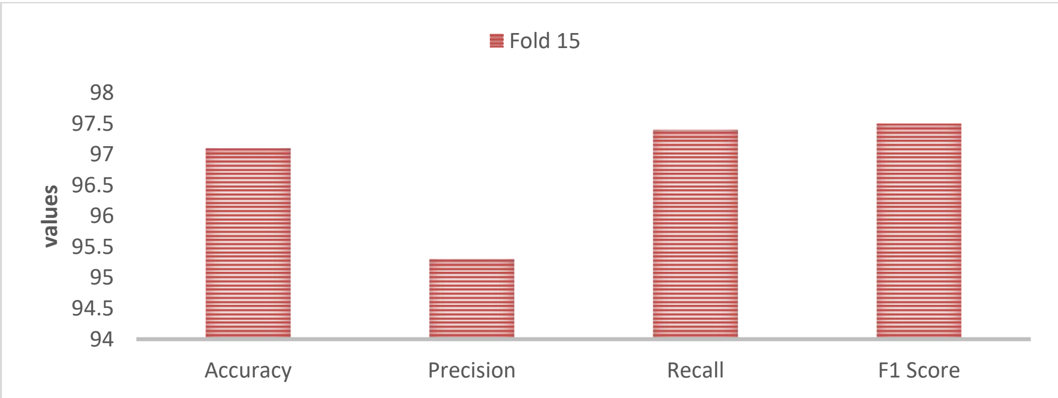

Fig 15: Validation of model with 15 fold cross validation using RNN-LSTM (ReLU) classifier

Experimental findings of figure 15 shows by using 15 fold cross validation, the accuracy, precision, recall and f-score of RNN-LSTM (ReLU) model is 94.9, 93.9, 94.2 and 94.7 respectively. As per experimental findings, 15-fold cross-validation has obtained the better average classification accuracy of 97.1%. In order to forecast student's performance, we used a minimum of 3 hidden layers. As per empirical findings, we conclude that RNN-LSTM with ReLu gives better detection accuracy than that of the RNN-LSTM (Tan h) and RNN-LSTM (sigmoid) function.

5. Conclusion

The prposed sentiment classification approach has divided into 3 different phases. The first method makes advantage of the co-occurrence link between feature categories and aspect types. Both feature selection and class prediction make use of the co-occurrence matrix. The method relies on a mixed-type feature collection.

This feature set combines characteristics based on lemmas with those based on grammatical relations. Specifically, the "multivariate filter approach" is used to pick out the best properties of the lemma. It picks characteristics based on importance and association. By using this technique, we can choose relevant traits while leaving out unnecessary ones. Aspects and emotions are represented via linguistic connection based attributes. In this method, the feature space is condensed by selecting characteristics based on linguistic relations using restrictive dependency constraints. Two-stage WCFS is the second strategy suggested in this thesis. With this method, we may choose n features at random from the whole dataset. From each class of characteristics, a subset is chosen. If the dataset is imbalanced, this is a helpful approach. It lessens the odds of picking characteristics from the more popular group. Class prediction utilises BR, CC, and LP multilabel classifiers with SVM, NB, and RF base classifiers. ABSA is a multilabel, not a multiclass, prediction issue. With the BR-SVM classifier. The third strategy we proposed an using an ABSA throughout the whole process. ABSA was sometimes treated as two distinct issues by many early systems, aspect prediction and sentiment analysis. This comprehensive approach predicts emotions for the retrieved features. We present a hybrid model that combines linguistic rule-based features with two-phase WCFS-based features and multilabel classifiers. An end-to-end ABSA is offered, as is an enhanced deep learning solution, as well as a RNN with LSTM. Using the enhanced RNN-LSTM system, the accuracy of the experiments was determined to be around 96%. The findings of the experiments confirm that the hybrid model using RNN-LSTM is comparable to the various machine learning algorithms and it provides higher accuracy than others.

Future Scope

The primary goal of all three methods we offered was to improve ABSA performance. A FS technique, using ML classifiers, and DL based hybrid approach are proposed to accomplish this. Under future work we can included a number of areas where we can do better. The method developed in this thesis is evaluated using a variety of unbalanced datasets, including those for laptop, restaurant, and medicine reviews. Additionally, it may be evaluated on a variety of huge datasets consisting of several classes.

References

[1] Xinzhi Wang , Luyao Kou , Vijayan Sugumaran , Xiangfeng Luo, and Hui Zhang, " Emotion Correlation Mining Through Deep Learning Models on Natural Language Text", IEEE TRANSACTIONS ON CYBERNETICS, May 2010.

[2] Renata L. Rosa, Gisele M. Schwartz, Wilson V. Ruggiero, and Demostenes Z. Rodr ıguez, Senior Member, IEEE, "A Knowledge-Based Recommendation System that includes Sentiment Analysis and Deep Learning", IEEE Transactions on Industrial Informatics Vol: 15 , April 2019.

[3] I.-R. Glavan, A. Mirica, and B. Firtescu, "The use of social media for communication.", Official Statistics at European Level. Romanian Statistical Review, vol. 4, Dec. 2016, pp. 37–48.

[4] M. Al-Qurishi, M. S. Hossain, M. Alrubaian, S. M. M. Rahman, and A. Alamri, "Leveraging analysis of user behavior to identify malicious activities in large-scale social networks," IEEE Transactions on Industrial Informatics, vol. 14, no. 2, pp. 799–813, Feb 2018, pp. 799–813.

[5] R. L. Rosa, D. Z. Rodr´ıguez, and G. Bressan, "Music recommendation system based on user‟s sentiments extracted from social networks," IEEE Transactions on Consumer Electronics, vol. 61, no. 3, pp. 359–367, Oct 2015. pp. 359–367.

[6] R. Rosa, D. Rodr, G. Schwartz, I. de Campos Ribeiro, G. Bressan et al., "Monitoring system for potential users with depression using sentiment analysis," in 2016 IEEE International Conference on Consumer Electronics (ICCE). Sao Paulo, Brazil: IEEE, Jan 2016, pp. 381–382.

[7] B. Weiner and R. L. Greene, "Handbook of personality assessment," in John Wiley and Sons, N.J, EUA, 2008.

[8] H. Lin, J. Jia, J. Qiu, Y. Zhang, G. Shen, L. Xie, J. Tang, L. Feng, and T. S. Chua, "Detecting stress based on social interactions in social networks," IEEE Transactions on Knowledge and Data Engineering, vol. 29, no. 9, Sept 2017, pp. 1820–1833.

[9] J. T. Hancock, K. Gee, K. Ciaccio, and J. M.-H. Lin, "I‟m sad you‟re sad: Emotional contagion in cmc," in Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work, 2008, pp. 295– 298.

[10] B. Liu,"Many facets of sentiment analysis, a practical guide to sentiment analysis," Springer International Publishing, Jan 2017, pp. 11–39,

[11] Y. P. Huang, T. Goh, and C. L. Liew, "Hunting suicide notes in web 2.0 - preliminary findings," in Ninth IEEE International Symposium on Multimedia Workshops (ISMW 2007), Dec 2007, pp. 517– 521.

[12] W. H. Organization, "World health statistics 2016: Monitoring health for the sustainable development goals, world health statistics annual," World Health Organization, p. 161, 2016.

[13] Hu M, Liu B 2004 ,“Mining and summarizing customer reviews.” ,In: Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining , doi:10.1145/1014052.1014073, pp. 168-177.

[14] Ahmad K, Cheng D and Almas Y 2007, “Multi-lingual sentiment analysis of financial news streams. In: 1st International Workshop on Grid Technology for Financial Modeling and Simulation”,SISSA Medialab , pp. 001

[15] Benkhelifa R, Bouhyaoui N, Laallam F Z 2019 , “A Real-Time Aspect-Based Sentiment Analysis System of YouTube Cooking Recipes. In: Machine Learning Paradigms: Theory and Application” , Springer, Cham , doi:10.1007/978-3-030-02357-7_11, pp. 233-251.