The Ciência & Engenharia - Science & Engineering Journal
ISSN: 0103-944X
Volume 11 Issue 1, 2023
pp: 1939 –1951

Various Feature Extraction and Selection Techniques for Lexicon Based and Machine learning Sentiment Classification

Ms. Shital A. Patil

Research Scholar, Department of Computer Engineering, SSBT, COET, Jalgaon,

shital.a.patil2014@gmail.com

Dr. Krishnakant P. Adhiya

Professor, Department of Computer Engineering, SSBT, COET, Jalgaon,

adhiyakp@gmail.com

Dr. Girishkumar K. Patnaik

Professor, Department of Computer Engineering, SSBT, COET, Jalgaon,

patnaik.girish@gmail.com

Abstract

Introduction: Natural Language Processing (NLP) is a kind of software that gives computers the ability to comprehend human languages. Words are often used as the fundamental unit for grammatical and semantic analysis on a deeper level, and the major objective of most natural language processing (NLP) projects is word segmentation

Objectives: Proposed System uses the feature extraction method for generating the hybrid feature to get good accuracy for classification in order to solve the practical problem of huge structural differences between different data modalities in a multi-modal environment, where traditional machine learning methods cannot be directly applied to solve the problem.

Methods: In order to do so, this paper introduces the method for generating the hybrid feature. In this System, we utilise a various feature extraction and selection technique from large text. The data has collected from students' feedback and these feature extraction techniques has applied

Results: Each technique provides different feature extraction while NLP based dependency features provides homogeneous feature set with relationship.

Conclusions: In extensive experimental analysis NLP based features obtains higher precision over the other feature extraction techniques in classification

Keywords: Feature extraction, feature selection, NLP, text processing, sentiment analysis, dependency features, classification, Aspect Based Sentiment analysis(ABSA),Sentiment Analysis(SA)

1. Introduction

Sentiment Analysis (SA) methods use as input a collection of texts that is student's feedback. The systems try to identify the primary characteristics and calculate the averaged emotion of the texts per characteristic. Although many ABSA methods have been suggested, most of which are development models [1], there is no defined task breakdown for ABSA and no recognized assessment metrics for the embedding process ABSA systems must accomplish. This work introduces a new task involving sentiment called identification and classification techniques, which consists of three subtopics: word retrieval, term polarization estimation, and classification techniques. The first subtask identifies single and inter phrases identifying characteristics. The second subtask calculates the average sentiments per word or cluster multiple phrases, and the third subtask performs sentiments categorization using machine learning approaches.

During the implementation, datasets of students' feedback were created for each of the subtasks mentioned earlier. New assessment techniques are presented for each process step, which is more suitable than primary evaluation methods. The paper presents new weighting versions of precision, recall, including average accuracy, as well as an explanation of why they are superior to the conventional variants for assessing techniques. In addition, the paper offers an enhanced version of a prominent unsupervised approach [2], which includes an additional pruning step that eliminates potential words. The new trimming step is based on newly popular word-to-continuous-space-vector mapping techniques [3]. Finally, the system demonstrates that the new enhanced approach, with the additional trimming step, is substantially superior to the traditional technique using the information and analysis techniques provided.

The rest of paper demonstrates detail description of entire manuscript, section 2 describes SA literature survey of various techniques for feature extraction and machine learning for sentiment classification. Objectives are discussed in section 3. In section 4 proposed system implantation has been described with numerous feature extraction techniques from student feedback dataset. The section 5 defines partially implementation results with all feature extraction and selection techniques with binary classification algorithm and finally section 6 describes conclusion and future work of system.

2. Related Work

For the SA subtask, many studies employed readily available lexicons like WordNet, SentiWordNet and SenticNet. Some authors contributed by proposing lexicons for certain topics. The disadvantage of lexicon-based approaches is that they can become application-specific or fail to recognize new characteristics.

Muslimah Wook, et al. [1] proposed a system opinion mining-based student feedback analysis using machine learning and a lexicon-based approach. The findings of the pedagogical evaluations are revealed using a lexicon-based method and an information extraction technique included in this system. One popular approach to quantifying textual data, linguistic analysis, may be used to determine the overall tone of a student's comments. Word polarity was determined using the Vader Lexicon, an English attitude word collection. The instructor's teaching evaluation findings were obtained by evaluating these sentiment polarities, which contained intensifier words derived from students' input. Opinion findings were improved because of the system's ability to comprehend novel elements like

The Ciência & Engenharia - Science & Engineering Journal
ISSN: 0103-944X
Volume 11 Issue 1, 2023
pp: 1939 –1951
capitalization and emoji. In the end, this brand-new method will supply valuable data to universities to enhance the quality of classroom instruction and syllabi.

Ganpat Singh Chauhan et al. [2] aspect-based sentiment classification for improvement of teaching-learning for students. Every day, an enormous quantity of explanations and background is provided on social networking sites. To extract emotion from free-form text, sentiment analysis is often used. The sentimental analysis of social media sites has been linked to a reduction in using more conventional methods of soliciting comments and suggestions. Most efforts have been put towards processing user comments, mostly consisting of classifying the positive or negative sentiments using lexicon-predicated or machine-learning algorithms at the document level. Findings indicate that sentiment analysis is useful for gathering student feedback on various topics, although it is seldom used in the classroom.

Laishram Kirtibas Singh et al. [3] describe a literature study of student feedback analysis. The students' responses may largely gauge the state of the workplace and the teacher's effectiveness; by listening to and learning from their comments, teachers can refine their practices and better serve their students. Learners sometimes need a better method for providing feedback, whether it's given verbally or via counselling. Using a student processing model, whether online or offline, makes grading and analysis of student feedback simpler and has the potential to enhance classroom instruction. Surveys and data processing provide significantly more timely and precise feedback. Emotion analysis is a method for determining if a student's emotional state is good, negative, or neutral.

Aadesh N.D. [4] proposed a system of student feedback-based sentiment classification using machine learning techniques SVM, NB and ANN Classifier, and Random Forest are only some of the machine learning algorithms suggested for doing content analysis on student comments. These methods of machine learning are also compared to one another. Based on the data, the Multinomial Naive Bayes Classifier achieves the highest accuracy in practice. They may use the results of this study to enhance the quality of education and the campus overall. Student feedback on a conference, workshop, etc., may also be gathered using this method.

Zemun Kastrati et al. [5] proposed a weekly supervised learning approach for aspect-class student feedback datasets. To successfully identify the component categories highlighted in the unlabeled students' reviews, the proposed method leverages weakly supervised annotating of MOOC-related characteristics and proliferates the weak supervision signal. As a result, it drastically lessens the primary bottleneck of all deep learning techniques: the necessity for manually labelled data. Experiments are conducted using two different quality education datasets: one with around 105k student ratings acquired from Coursera and another with 5989 students' input from more conventional classroom settings. Experimental findings show that our suggested approach achieves excellent outcomes in aspect category recognition and aspect emotion classification.

The Ciência & Engenharia - Science & Engineering Journal
ISSN: 0103-944X
Volume 11 Issue 1, 2023
pp: 1939 –1951
Saida Ulfa et al. [6] proposed a system of student feedback learning for sentiment classification. This research examined how sentiment analysis methods may be used for online student feedback. Twelve papers discussed the use of sentiment analysis to analyze student responses. ZemunKastrati et al. [7] proposed a system of student feedback for sentiment analysis using NLP and deep learning techniques. Several machine learning and deep learning methods are compared on a dataset consisting of student comments. Several natural language processing methods were analyzed and rated during the sorting process. These included the elimination of stop words, the creation of lemmas, the extraction of stems, and the analysis of dependency parsing. It covers the key obstacles currently in the fields of emotion classification and recommendation.

Rosario Catelli et al. [8] proposed a system lexicon-based and Bert system that were analyzed with machine learning techniques. Using an ad hoc dataset, this research suggests comparing these products in the Italian market, one of the biggest in the world. This study's small dataset reflects the reality that many languages besides English and Chinese must deal with. While evidence usually shows the importance of machine translation like BERT built on deep cognitive networks, it opens several questions about the efficacy and advancement of these alternatives, especially in comparison to those based on dictionaries.

Qik Lin et al. [9] propose two lexical-based methods, knowledge-based and machine learning-based, to automatically extract opinions from short reviews. To begin, evaluating classroom teaching needs methods for sentiment analysis at the sentence and paragraph levels. Second, appropriate teacher modelling may be obtained by amalgamating many forms of student assessment data such as multiple-choice answers, evaluative texts, academic history, intellectual, social network, and research abilities.

Irfan Ali Kandhro et al. [10] proposed a system of student feedback analysis using numerous natural language processing and machine learning techniques. Various classifiers, including ANN, SVM, KNN and multiple supervised machine learning classification algorithms, were utilized to create the model. The outcomes were examined using assessment criteria such as the Confusion Matrix, Precise, Recall, and F-score.

Michelangelo Misuraca et al. [11] proposed a system of opinion-mining techniques for mining educational data on student feedback. Here it provides a method for determining the emotional tone of student evaluations. After justifying the proposal's mathematical establishment and determining the polarity for positive/negative sentiment categorization. Numerous machine learning algorithms have been used for classification, such as ANN, SVM, NB and random forest. The SVM outperforms the highest accuracy at 96.50% on the heterogeneous dataset.

Metadata Mohammed Almosawi and Dr Salma A. Mahmood [12] proposed a lexical-based sentiment analysis approach to determine the polarity of students' feedback. Using a free-form online survey, they could compile a comprehensive dataset. This data collects a lexicon of 2,217 terms related to higher learning in both the Iraqi dialect spoken in southern Iraq and contemporary Arabic. The suggested method of sentiment analysis was 98% accurate.

Various ML methods were used, including Naive Bayes, Support Vector Machines, and the Kernelized Naive Neighbor Finder. Concerning the lecturer's performance, the gathered data shows that 60% of the students had a favorable impression, whereas 40% had a negative one.

HtarHtar Lwin et al. [13] proposed a system of education data mining using feedback analysis with machine learning techniques. In this article, we'll look at how to put into practice a method for analyzing feedback that uses both quantitative ratings and qualitative remarks. It utilized numerical ratings and free-form text comments as our datasets of choice. When grouping ratings, the K-means clustering technique is used. After obtaining a tagged dataset via the clustering process, classification models may be constructed utilizing various classification techniques. It employed the Naive Bayes classifier for training a model, then used 10-fold cross-validation to sort comments in the test dataset by positive and negative sentiment. Additionally, it tested the model by trying to categorize some free-form languages that had not been classified.

Rahil Nawaz et al. [14] AI and machine learning-based feedback analysis for educational data mining and student feedback data. Additionally, it proposes a unique automated analytic methodology for mining useful insights from students' open-ended replies to survey questions.

Giuseppe Varvara et al. [15] proposed a system challenge of dental education on student feedback data using NLP and machine learning techniques. The purpose of this research was to ascertain how undergraduates see the incorporation of these strategies into their coursework. All in all, 353 first- through sixth-year students at the students. The students were surveyed using an online survey created in Italian using "Google Forms" and sent via email. There were three sections to the questionnaire: the first asked for demographic information like age, sex, and course year; the second contained multiple-choice questions about how they felt the e-learning was going; and the third contained two free-form questions about the benefits and drawbacks of the new methods of instruction

3. Objectives

The feature extraction, feature selection, creation of training matrix, and assessment of system performance using test dataset are all done sequentially by the ABSA. Various feature selection methods are examined in the system, and a hybrid feature selection strategy is suggested.

4. Methods

The feature selection methods that have been investigated are as follows:

4.1 Term frequency (TF)

The term frequency count is used to pick features in this method. The term frequency of each feature is computed for each category. For feature selection, a threshold is specified. In each category, features with a word frequency higher than '2' are chosen. As a consequence of this,

The Ciência & Engenharia - Science & Engineering Journal
ISSN: 0103-944X
Volume 11 Issue 1, 2023
pp: 1939 –1951
a term frequency matrix for each category is produced. In addition, a compound matrix comprising words and their occurrence counts in all categories is created. A binary train matrix is produced from this matrix, with '1' representing non-zero term frequency.

4.2 Weighted Term Frequency

These techniques utilized the weight of each term is calculated using equation in this method (1). The conditional probability of a term is X (t,k), where X t is the overall occurrence count of a term "t" in all categories and X (t,k) is the occurrence count of term "t" in category "k." The weight of "t" rises if the percentage of occurrence of a word "t" in category "k" is higher than the other categories. For each category, a weight threshold is set. To create a binary train matrix, terms (features) with a weight higher than threshold are chosen. Kim Schouten et al. [14] perform a weighted term computation as well. This paper takes a similar approach to [14], proposing a hybrid method for feature selection that uses correlation to prevent feature duplication. Weights are utilized to identify the category of the test phrase in [8], and weights are employed for feature selection and creation of a binary train matrix in this method.

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

4.3 Term Frequency with Correlation Coefficient

To improve the classification accuracy of the classifier, features must be useful but not redundant. The term frequency matrix produced in I is utilized in this approach. The information gained from this matrix is useful, however it is also redundant. To prevent duplication, each feature's correlation is computed in relation to other characteristics in the same category. To calculate correlation, the Pearson correlation coefficient is employed.

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

The correlation of each term with regard to other terms is calculated using Equation (2), where "x[] and y[]" are vectors of term t_i and t_{i+1}, respectively, containing term frequency with respect to each category. The average of the correlation of a word "t" with other terms in that term category. To create a binary train matrix, terms with a correlation value of less than or equal to 0.85 are chosen.

4.4 Weighted Term Frequency with Correlation Coefficient

The weighted matrix produced is utilized to create a new matrix containing the weight of a word in relation to each term category in this method. The correlation of each word with regard to other terms is calculated using Equation (2), where "x[] and y[]" are vectors of term t_i and t_{i+1}, respectively, including the weight of term "t" in each category. Finally, as previously stated, a binary train matrix is produced (iii). This article makes a contribution by proposing a supervised method for term extraction that chooses important features while avoiding duplication by assessing feature correlation. The obtained findings indicate that the weighted word frequency with correlation method has a higher F-score than the other approaches. It is discovered in this study that characteristics chosen using weighted word frequency are more relevant, but also redundant. By measuring the correlation between characteristics in a term category, redundancy is eliminated.

4.5 Part of Speech Tagging and Chunking

Parts of speech tagging, often known as POS tagging, is a linguistic method that has been employed since 1960 and has lately attracted the interest of NLP researchers for the extraction of product features, since product characteristics are typically nouns or noun phrases. POS tagging gives a tag to each word in a text and categorizes it into morphological categories such as norm, verb, adjective, and so on. In terms of accuracy, POS taggers are effective for explicit feature extraction; nevertheless, a difficulty emerges when the review includes implicit characteristics. Hidden Markova Models are extensively utilized in the development of POS taggers because they are more accurate than other methods such as rule-based, statistical, and machine learning. NL Processor linguistic parser, Stanford POS tagger, Gate ANNIE POS Tagger, and Claws POS tagger are among of the English language POS taggers used for this [28].

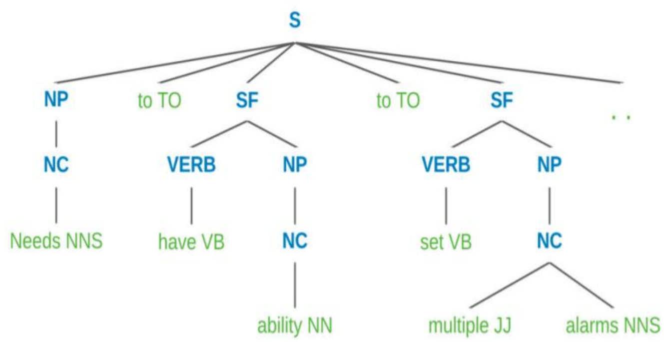


Figure 1 : POS tagging and chunking relationship

A phrase in a sentence retrieved using the POS tag pattern and a regular expression is known as POS chunking. To bring out the characteristics, the POS tag pattern comprising nouns, verbs, and descriptors is converted into a phrase. The phrases 'voice alerts' and satellite maps, are the example, denote to software features.

4.5.1 Bi-tagged

Unigrams and bi-grams are retrieved from text first, followed by bi-tagged phrase characteristics. The IG feature selection technique is used to extract the significant features. PromUni (prominent unigrams), PromBi (prominent bi-grams), and PromBiTa (prominent bi-tagged) characteristics are derived from unigrams, bigrams, and bi-tagged phrases, respectively. When unigram characteristics are coupled with bi-grams, their performance improves. Unigrams with bigrams and unigrams with bi-tagged features, referred to as ComUniBi and ComUniBiTa, respectively, are used to generate composite features. Finally, by merging prominent unigrams with prominent bi-tagged features, PromUniBiT is a feature set.

The Ciência & Engenharia - Science & Engineering Journal
ISSN: 0103-944X
Volume 11 Issue 1, 2023
pp: 1939 –1951
The phrases are tagged using a POS tagger, and bi-tagged characteristics are retrieved. The characteristics that have a significant connection between two successive words in a phrase are known as bi-tagged features. Table 1.1 lists the bi-tagged relationships that have been examined

Table 1.1: Bi-tagged relationships between two consecutive words [1]

First word	Second word
JJ	NN/NNS
RB/RBR/RBS	JJ
JJ	JJ
NN/NNS	JJ
RB/RBR/RBS	VB/VBD/VBG
VBN	NN/NNS
VB/VBG	JJ/JJR/JJS
JJ	VBN/VBG
RB/RBR/RBS	RB/RBR/RBS

A small snippet for some bi-tagged relationships is mentioned below.

```
for i in range(len(dataset)):
    sent = nltk.word_tokenize(dataset.iloc[i, 0].lower())
PoS_Tag_sent = nltk.pos_tag(sent)
for (w1, tag1), (w2, tag2) in nltk.bigrams(PoS_Tag_sent):
if tag1.startswith('JJ') and tag2.startswith('NN'):  # R1
corpora += w1 + ' ' + w2 + ';'
eliftag1.startswith('RB') and tag2.startswith('JJ'):  # R2
corpora += w1 + ' ' + w2 + ';'
eliftag1.startswith('JJ') and tag2.startswith('JJ'):  # R3
corpora += w1 + ' ' + w2 + ';'
eliftag1.startswith('NN') and tag2.startswith('JJ'):  # R4
corpora += w1 + ' ' + w2 + ';'
eliftag1.startswith('RB') and tag2.startswith('VB'):  # R5
corpora += w1 + ' ' + w2 + ';'
eliftag1.startswith('VB') and tag2.startswith('NN'):  # R6
corpora += w1 + ' ' + w2 + ';'
eliftag1.startswith('JJ') and tag2.startswith('VB'):  # R7
corpora += w1 + ' ' + w2 + ';'
eliftag1.startswith('RB') and tag2.startswith('RB'):  # R8
corpora += w1 + ' ' + w2 + ';'
eliftag1.startswith('RB') and tag2.startswith('VB'):  # R9
corpora += w1 + ' ' + w2 + ';'
S2_super_corpus.append(corpora)
    corpora = ''
bar.load(i, base=dataset, text='Stream 2')
```

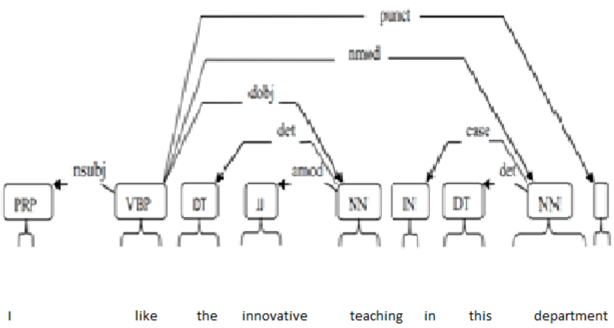
The Ciência & Engenharia - Science & Engineering Journal
ISSN: 0103-944X
Volume 11 Issue 1, 2023
pp: 1939 –1951

```
print('Stream 2: Processed')
stream2 = pd.Series (S2_super_corpus)
return stream2
```

Here, VB/VBN/VBG/VBD shows verb, JJ/JJR/JJS shows adjectives, RB/RBR/RBS shows adverb, and NN/NNS shows noun.

4.6 Dependency based Features

Features are chosen using this method based on grammatical connections between words. To extract grammatical connections between words, the Stanford NLP parser is utilized. When compared to single lemma characteristics, phrases or related words in a sentence may sometimes provide more information about an attribute. The arrows in Figure 2 show the connection between words in a phrase. The parser extracts characteristics so on. To extract rule-based characteristics in the initial testing, all dependency connections are utilized except the determinant (det) relation. The term frequency of each pair of features in an category is computed, and different features are chosen. Because many dependence connections do not occur on a regular basis, the word frequency is not used to choose features. A matrix including the review id, feature, word frequency in each category, and the actual category name is produced from these chosen characteristics. The training matrix is this matrix. The testing is similar to the lemma-based method, except instead of lemma characteristics, rule-based features are collected from each test phrase.



.

Figure 2: Dependency feature extraction using Sandford NLP

To extract features, dependency connections including any of the nouns, adjectives, or adverbs are utilised. The rules that were utilised are listed below.

conj: conjunct

A conjunct describes the relationship between two related and linked words by utilising co-ordinating conjunctions such as "and," "or," and so on. The conjuncts are handled unequally, with the first conjunct taking the lead and the remaining conjuncts relying on the first in some manner.

The Ciência & Engenharia - Science & Engineering Journal
ISSN: 0103-944X
Volume 11 Issue 1, 2023
pp: 1939 –1951

cop: copula

A copula is a linking word, especially a verb form that links the subject and complement. Copula is often seen to be reliant on its complement.

dobj: direct object

In a clause or sentence, a direct object is an adjective, noun phrase, or pronoun that denotes what or person gets the action of a transitive verb.

neg: negation modifier

The relationship between a negation word or phrase and the NP or VP it modifies is identified by a negation modifier.

nn: noun compound modifier

A noun compound modifier is a noun in a phrase that helps the reader comprehend the head noun and often modifies it.

nsubj: nominal subject

A nominal subject is a pro-agent of a clause in a sentence, which is sometimes referred to as a syntactic subject. Unlike conventional grammar principles, the verb does not regulate the connection in the phrase. The complement of the copular verb, which may be an adjective or a noun, is the root of the sentence when there is a copular verb.

nsubjpass: passive nominal subject

When a nominal subject (the syntactic subject) is employed in the passive voice, it is referred to as a passive nominal subject.

rcmod: relative clause modifier

A relative clause modifier is a relative clause that modifies a noun phrase. When an essential modifier is used, it is put next to the noun it changed, although it may also be separated. The link is established between the noun phrase and the relative clause, typically via the employment of a verb.

xcomp: open clausal complement

An open clausal complement is a predicative or clausal complement of a verb or adjective that does not have its own subject. In this instance, the topic is specified by a phrase or phrases that are not part of the xcomp. The clausal complements are always non-finite and are neither adjuncts nor modifiers.

nmod: Nominal modifier

The nmod connection describes the link between the noun / predicate modified by the prepositional supplement and the noun introduced by the preposition. The testing shows that

The Ciência & Engenharia - Science & Engineering Journal
ISSN: 0103-944X
Volume 11 Issue 1, 2023
pp: 1939 –1951
when compared to a system that utilises all dependency relations, the system that extracts features using selected dependency connections improves performance

5.Result

After weighing the pros and cons of different systems and alternatives, the decision to go with a more advanced system may have been made.

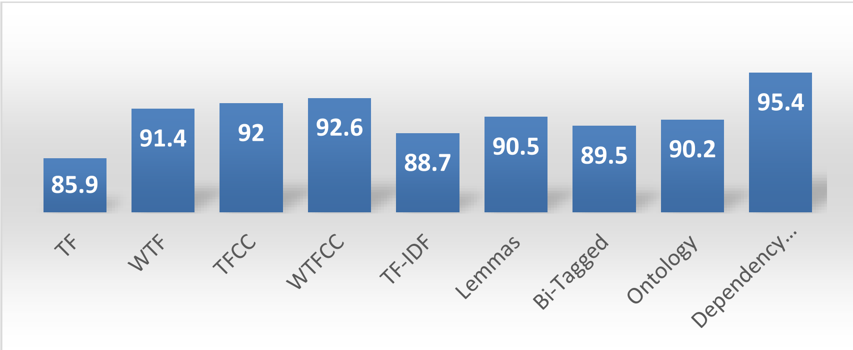


Figure 3 :- Classification Accuracy with various feature extraction tech. using Binary Classifier

6. Conclusion

This paper proposes various feature extraction selection techniques from large text data. These approaches are widely used for analysis and sentiment classification. Initially, we collected large text data and applied all feature extraction methods after pre-processing. The extracted features feed into a binary classifier and measure the classification accuracy. The NLP-based dependency feature extraction method provides higher accuracy with 95.4%, which is higher than other techniques. The application of the different machine learning and deep learning classifiers to the identification and classification of aspects will be the future work of this research. To validate the proposed model with hybrid machine learning and deep learning techniques will be future task of proposed system

References

1. Wook, Muslimah, et al. "Opinion mining technique for developing student feedback analysis system using lexicon-based approach (OMFeedback)." Education and Information Technologies 25.4 (2020): 2549-2560.
2. Chauhan, Ganpat Singh, Preksha Agrawal, and Yogesh Kumar Meena. "Aspect-based sentiment analysis of students' feedback to improve teaching–learning process." Information and communication technology for intelligent systems. Springer, Singapore, 2019. 259-266.
3. Singh, Laishram Kirtibas, and R. Renuga Devi. "Student feedback sentiment analysis: A review." Materials Today: Proceedings (2021).
4. Dsouza, DaneenaDeeksha, et al. "Sentimental analysis of student feedback using machine learning techniques." International Journal of Recent Technology and Engineering 8.14 (2019): 986-991.

5. Kastrati, Zenun, Ali Shariq Imran, and ArianitKurti. "Weakly supervised framework for aspect-based sentiment analysis on students' reviews of MOOCs." IEEE Access 8 (2020): 106799-106810.
6. Ulfa, Saida, et al. "Student feedback on online learning by using sentiment analysis: a literature review." 2020 6th International Conference on Education and Technology (ICET). IEEE, 2020.
7. Kastrati, Zenun, et al. "Sentiment analysis of students' feedback with NLP and deep learning: A systematic mapping study." Applied Sciences 11.9 (2021): 3986.
8. Catelli, Rosario, Serena Pelosi, and Massimo Esposito. "Lexicon-based vs. Bert-based sentiment analysis: A comparative study in Italian." Electronics 11.3 (2022): 374.
9. Lin, Qika, et al. "Lexical based automated teaching evaluation via students' short reviews." Computer Applications in Engineering Education 27.1 (2019): 194-205.
10. Kandhro, Irfan Ali, et al. "Student feedback sentiment analysis model using various machine learning schemes: a review." Indian Journal of Science and Technology 12.14 (2019): 1-9.
11. Misuraca, Michelangelo, GermanaScepi, and Maria Spano. "Using Opinion Mining as an educational analytic: An integrated strategy for the analysis of students' feedback." Studies in Educational Evaluation 68 (2021): 100979.
12. Almosawi, Mrtdaa Mohammed, and Salma A. Mahmood. "Lexicon-Based Approach For Sentiment Analysis To Student Feedback." Webology (ISSN: 1735-188X) 19.1 (2022).
13. Lwin, HtarHtar, et al. "Feedback analysis in outcome base education using machine learning." 2020 17th International conference on electrical engineering/electronics, computer, telecommunications and information technology (ECTI-CON). IEEE, 2020.
14. Nawaz, Reheel, et al. "Leveraging AI and Machine Learning for National Student Survey: Actionable Insights from Textual Feedback to Enhance Quality of Teaching and Learning in UK's Higher Education." Applied Sciences 12.1 (2022): 514.
15. Varvara, Giuseppe, et al. "Dental education challenges during the COVID-19 pandemic period in Italy: Undergraduate student feedback, future perspectives, and the needs of teaching strategies for professional development." Healthcare. Vol. 9. No. 4. MDPI, 2021.

Ms. Shital Abhimanyu Patil, completed B.E from Godavari college Of Engineering, Jalgaon. and ME Computer Science & Engineering from SSBT COET Bambhori. She is working as Assistant Professor in SSBT's College of Engineering and Technology since 2006.She is pursuing PhD in Computer Science & Engineering from Kavayitri Bahinabai Chaudhari North Maharashtra University, Jalgaon. Her areas of interest are Machine Learning, Data Analytics and Natural Language Processing.

The Ciência & Engenharia - Science & Engineering Journal
ISSN: 0103-944X
Volume 11 Issue 1, 2023
pp: 1939 –1951

Dr. Krishnakant Prabhudas Adhiya has completed PhD degree in Computer Science and Engineering from North Maharashtra University Jalgaon. Currently he is working as Professor, Department of Computer Engineering, SSBT's College of Engineering & Technology, Jalgaon and recognized PhD Guide in Kavayitri Bahinabai Chaudhari North Maharashtra University, Jalgaon. He has 44 research papers in reputed peer reviewed journals in addition to 14 papers in International Conferences to his credit. He is Professional Member in ACM, Life member of ISTE. His research interests are Wireless Sensor Networks and Security, Machine Learning, Block Chain and Natural Language Processing

Dr. Girishkumar K. Patnaik, is working as Professor & Principal of SSBT's College of Engineering & Technology, Jalgaon. He is recognized PhD Guide in Kavayitri Bahinabai Chaudhari North Maharashtra University, Jalgaon. Professional Member in ACM, Life member of ISTE and CSI. His research interests are Wireless Sensor Networks and Security, Machine Learning, Blockchain and Natural Language Processing. He has published 30 research papers in reputed peer reviewed journals in addition to 12 papers in International Conferences to his credit. He is Senior Member in IEEE.