IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

A SYNOPSIS OF THE ADVANCES AND CHALLENGES IN APPLYING DEEP LEARNING METHODS TO NATURAL LANGUAGE PROCESSING FOR FEEDBACK EVALUATION

Sujankumarreddy Vuyyuru (1) , Gogulamudi Sai Sandeep Reddy (2), Srinija Vuyyuru (3) , Pujitha Vajrala (4)

[1]Saispurthi institution of technology Affiliated to JNTU, HYDERABAD , India.

[2]Presidency university

[3 4]Sagi Rama KrishnamRaju Engineering College. Affiliated to JNTU, Kakinada, India.

Abstract: This research paper provides a comprehensive review of deep learning techniques applied to natural language processing (NLP) for the evaluation of feedback. In today's digital age, feedback is ubiquitous, and its analysis is crucial for understanding user sentiments, improving products and services, and making informed business decisions. Traditional methods of feedback analysis often fall short in handling the complexity and nuances of human language. Deep learning, a subset of machine learning, has shown promising results in capturing intricate patterns and semantics in natural language.

This paper begins by introducing the importance of feedback evaluation and the challenges associated with traditional approaches. Subsequently, it delves into an overview of deep learning techniques, including neural networks, recurrent neural networks (RNNs), long short-term memory (LSTM) networks, and transformer models, and their applications in NLP. The focus is on their ability to process and understand the context, sentiment, and semantics of textual feedback.

The main body of the paper reviews recent research studies and applications of deep learning techniques in feedback evaluation. This includes sentiment analysis, opinion mining, and emotion detection, highlighting the strengths and limitations of various models. Additionally, the paper explores the use of pre-trained language models, such as BERT (Bidirectional Encoder Representations from Transformers), GPT (Generative Pre-trained Transformer), and their variants, in feedback analysis.

The paper also discusses challenges and open research issues in applying deep learning to feedback evaluation, such as the need for labeled datasets, model interpretability, and ethical considerations. Furthermore, it provides insights into potential future directions for research in this domain, including advancements in model architectures, transfer learning, and the integration of multimodal data.

To validate the effectiveness of deep learning techniques in feedback evaluation, the paper presents a case study or experimental results using a specific dataset. It evaluates the performance of different models and compares them with traditional methods, showcasing the advantages of deep learning in capturing complex linguistic features.

In conclusion, this research paper consolidates the current state of the art in applying deep learning techniques to NLP for feedback evaluation. It provides a roadmap for researchers, practitioners, and industry professionals interested in leveraging advanced techniques to gain deeper insights from textual feedback in various applications.

Keywords: Deep learning, Natural Language Processing, Feedback Evaluation, Sentiment Analysis, Opinion Mining, Neural Networks, Transformer Models, BERT, GPT.

I. INTRODUCTION

In recent years, the intersection of Natural Language Processing (NLP) and deep learning has catalyzed transformative advancements in the field of feedback evaluation. The surge in digital communication channels has resulted in an unprecedented volume of textual feedback across diverse domains such as customer service, education, healthcare, and e-commerce. Extracting meaningful insights from this deluge of unstructured data has become a paramount challenge. Traditional NLP techniques, while effective to some extent, often struggle with the nuances and context inherent in human language.

Deep learning, a subfield of mechanism learning encouraged by the arrangement and function of the human brain, has emerged as an authoritative paradigm for tackling complex NLP tasks. Unlike conventional methods that rely on handcrafted features and rule-based systems, deep learning models autonomously learn hierarchical representations of data, allowing them to capture intricate patterns and relationships within the text. This paper explores the application of deep learning techniques to enhance the evaluation of feedback, shedding light on the methodologies, architectures, and challenges associated with this evolving research area.

The motivation behind delving into deep learning for feedback evaluation stems from the inherent need to discern sentiments, opinions, and valuable insights from the vast sea of textual feedback. Businesses strive to understand customer satisfaction, educators seek to improve teaching methodologies based on student feedback, and healthcare providers aim to gauge patient experiences. Conventional sentiment analysis techniques often fall short in capturing the subtleties and complexities of human language, motivating the exploration of more sophisticated models.

Deep learning models, with their ability to automatically learn intricate features and patterns, offer a promising avenue for overcoming the limitations of traditional approaches. By leveraging neural networks with multiple layers, these models can grasp the semantics and context of language, enabling more accurate and context-aware feedback evaluation.

1.1 Objectives of this manuscript:

1) Investigate and elucidate the various deep learning architectures employed in feedback evaluation, including Recurrent Neural Networks (RNNs), Long Short-Term Memory Networks (LSTMs), Convolutional Neural Networks (CNNs), and Transformer models.
2) Explore the preprocessing steps and feature extraction methods essential for preparing textual data for deep learning models, encompassing tokenization, word embeddings, and transfer learning.
3) Examine the diverse applications of deep learning in feedback evaluation across different domains, such as customer reviews in e-commerce, educational feedback, healthcare narratives, and social media sentiments.
4) Delve into the challenges inherent in applying deep learning to feedback analysis, including handling imbalanced datasets, ethical considerations, and the pursuit of more explainable models. Additionally, propose avenues for future research and development in this dynamic field.

Artificial Intelligence (AI) has revolutionized the educational landscape by introducing machine learning methodologies aimed at personalizing the student learning experience through various platforms, including learning management systems[1]. This transformative approach involves the integration of deep learning and transfer learning, utilizing pre-trained concepts to address novel, similar problems[2]. Additionally, natural language processing (NLP) methods have been instrumental in capturing student feedback, processing it, and generating predictive insights regarding their opinions on the learning infrastructure.

AI has the capacity to reshape traditional educational infrastructures[4], impacting online tutoring, learning management systems, curriculum development, and transitions in employment, teacher training, assessments, and research training. The institutional data collected spans diverse formats, including textual feedback from students and recordings in video and audio formats within classroom settings.

Chassignol et al.[5] defined AI as an activity dedicated to imbuing machines with intelligence, emphasizing the ability to function appropriately and foresee outcomes in a given environment. Educational institutions globally have embraced AI in various service delivery forms to students [6]. NLP, a widely employed AI methodology for mining student opinions [7], plays a pivotal role in interpreting end-users' feedback, providing insights that institutions invest resources and time to comprehend. NLP's versatility extends to analyzing textual data in multiple languages, offering valuable perceptions and opinions on services, products, or human experiences.

Eggert [9] highlighted the potential of AI in education, proposing methods to enhance teaching by collecting extensive data on students' prior knowledge, emotional states, and economic backgrounds. Intelligent tutoring systems (ITS) within adaptive learning platforms (ALP) are pivotal components, automating routine tasks and allowing teaching staff to focus on innovative approaches. The exposure of students to AI-driven tools is also emphasized to prepare them for a technology-dependent future, promoting lifelong learning through improved access to Massive Open Online Courses (MOOCs).

Holstein et al.[11] emphasized the necessity for personalized guidance in AI-enhanced classrooms, utilizing real-time support from AI systems to identify when students require human assistance for motivation. The challenges of involving non-technical stakeholders in the design of complex learning analytics systems were addressed by proposing tools like Konscia, a wearable and real-time awareness tool for teachers in AI-enhanced K-12 classrooms[12].

Analyzing student feedback in MOOCs is underscored by Alrajhi et al. [14] as crucial for understanding the need for instructor intervention. Chen et al.[6] conducted a comprehensive survey on the impact of AI on education, covering various technical aspects such as assessment, grading, smart schools, and personalized intelligent teaching.

The extensive research on AI's impact on education has led to the focus on building cognitive intelligent systems using AI, with the initial step being the collection of student opinions and feedback on existing educational infrastructure. Traditionally, educational institutions seek student feedback to gauge perceptions of the teaching team and the learning experience, which can be in quantitative or qualitative formats[18]. The manual monitoring and tracking of student feedback, however, are resource-intensive and time-consuming. NLP, with its annotation and summarization capabilities, offers a potential solution to this challenge.

This study delves into NLP methodologies applicable to the education domain, exploring existing methodologies, challenges, current trends, and the adoption of NLP methodologies from other disciplines. The contributions of this research include an enhanced understanding of AI's impact on education, the synthesis of NLP methodologies for annotating student user feedback, and an exploration of trends and challenges in NLP for adoption in the education domain. The subsequent sections provide a detailed exploration of feature extraction techniques, topic modeling, challenges in adopting NLP in education, and a discussion on the presented work. The paper concludes with limitations and outlines future directions for the study.

II. DEEP LEARNING ARCHITECTURES FOR FEEDBACK EVALUATION

Deep learning architectures have gained significant attention and success in various natural language processing (NLP) tasks, including feedback evaluation. These architectures leverage neural networks with multiple layers (deep neural networks) to automatically learn hierarchical representations of data, enabling them to capture intricate patterns and relationships in feedback text.

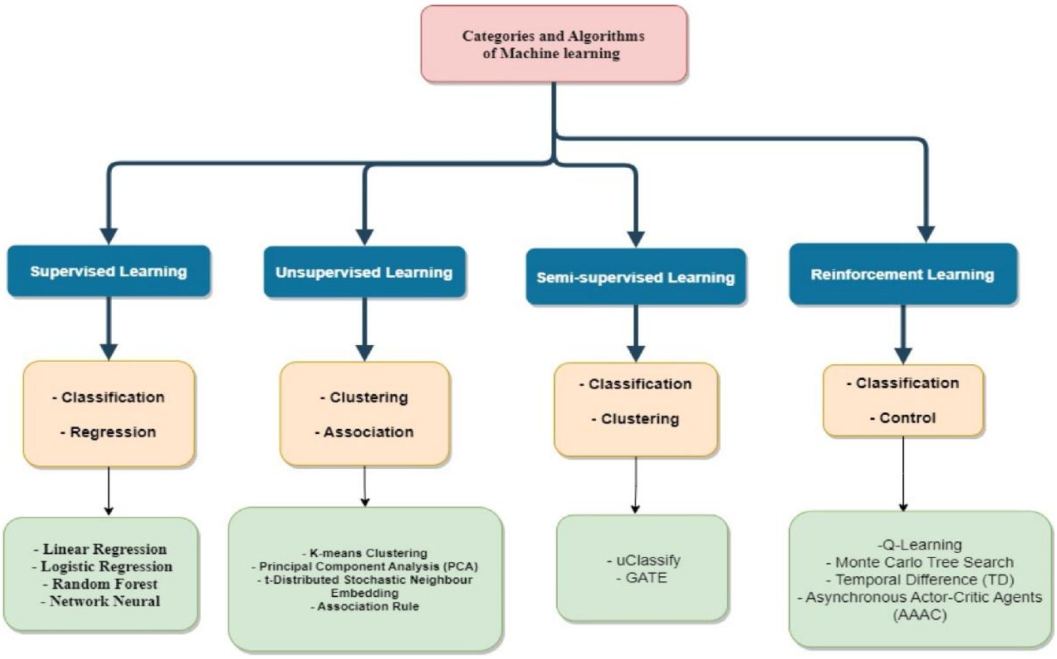

Figure 1 :Deep Learning Architectures for Feedback Evaluation

Here are some deep learning architectures commonly used for feedback evaluation:

1. Recurrent Neural Networks (RNNs):

Recurrent Neural Networks are a class of neural networks designed for sequential data processing[20]. In the context of feedback evaluation, RNNs can capture the temporal dependencies present in feedback texts. The ability to maintain a hidden state that considers previous information makes RNNs suitable for tasks where the order of words or context is crucial.

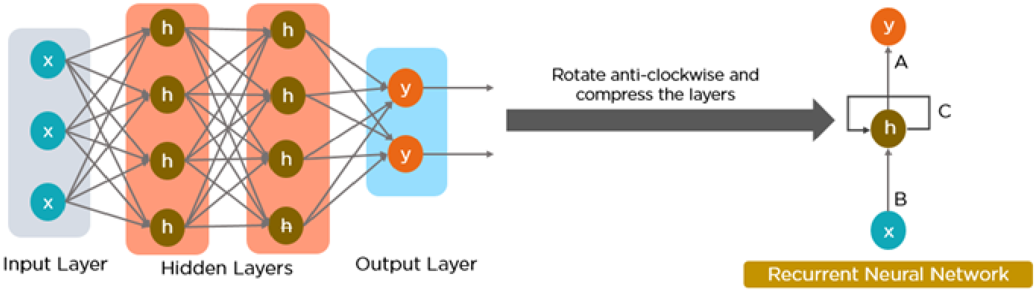

Figure 1 : Recurrent Neural Networks designed for sequential data processing

2. Long Short-Term Memory Networks (LSTMs):

LSTMs are a specialized form of RNNs designed to overcome the vanishing gradient problem. This problem often hinders the learning of long-term dependencies in sequential data[22]. LSTMs maintain a memory cell, allowing them to capture and store information for longer periods. In feedback evaluation, LSTMs can effectively handle the contextual nuances in longer feedback texts.

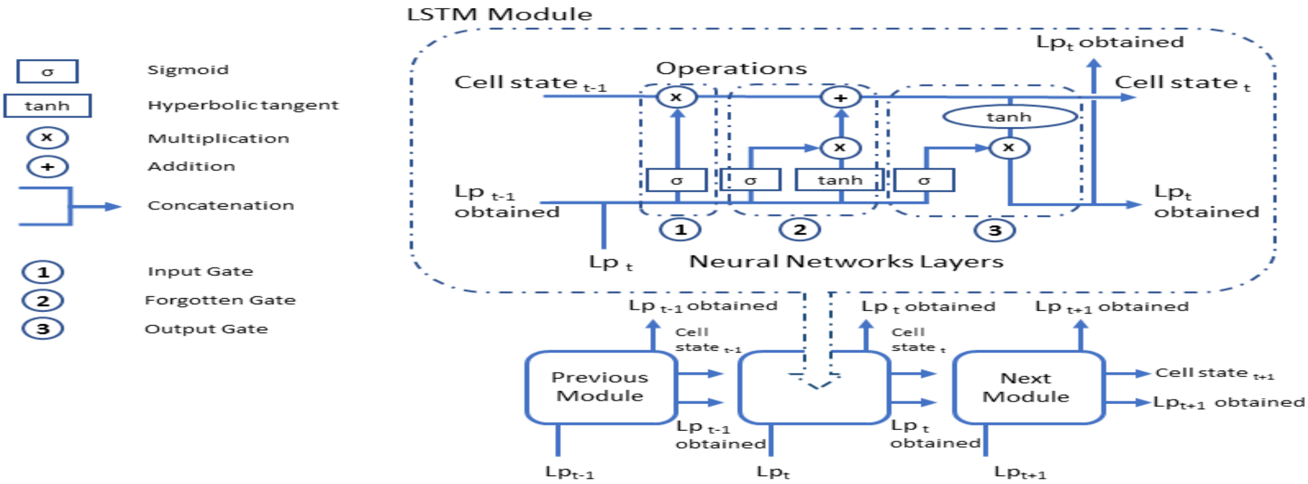

Figure 2 : General scheme of an Long Short-Term Memory neural networks (LSTM)

3. Gated Recurrent Units (GRUs):

GRUs are another variant of RNNs designed to address the limitations of traditional RNNs. They have a simplified structure compared to LSTMs, making them computationally more efficient[21]. GRUs are suitable for capturing dependencies in sequential data and can be applied to feedback evaluation tasks where a balance between performance and efficiency is essential.

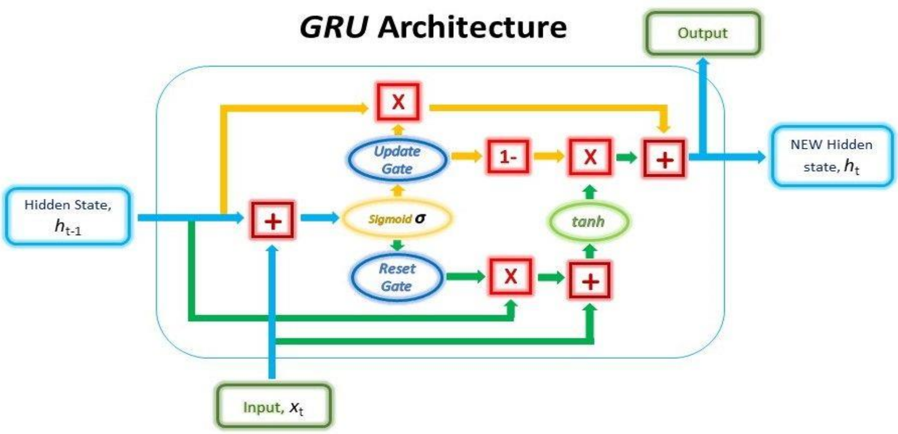

Figure 3 : Gated Recurrent Units (GRUs) Architecture

4. Convolutional Neural Networks (CNNs):

Originally developed for image processing, CNNs have been successfully adapted for natural language tasks, including feedback evaluation. In this context, CNNs can be applied to capture local patterns and features within feedback texts[16]. By employing convolutional layers, CNNs excel at identifying significant phrases or expressions that contribute to the overall sentiment or opinion.

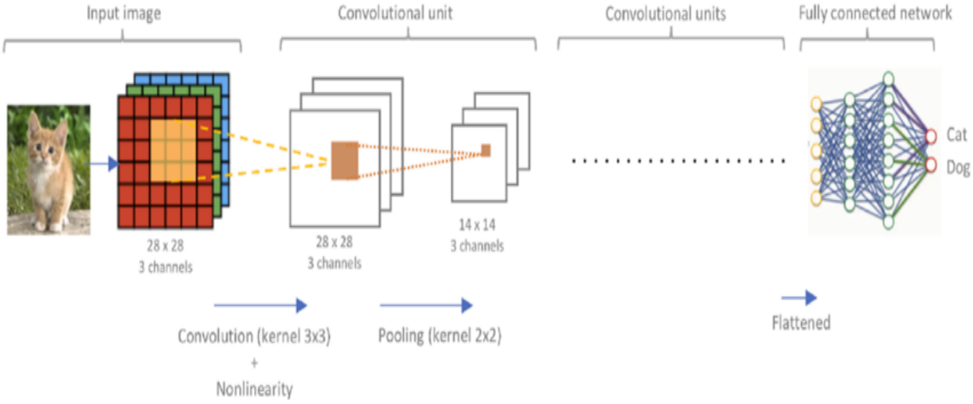

Figure 4 : Convolutional Neural Network (CNN) Architecture

5. Transformer Models:

Transformer models, such as BERT (Bidirectional Encoder Representations from Transformers) and GPT (Generative Pre-trained Transformer), have become prominent in natural language processing tasks[20]. These models use attention mechanisms to process input data in parallel, making them highly efficient. Transformer models excel in capturing contextual information and understanding the relationships between words, making them well-suited for feedback evaluation tasks.

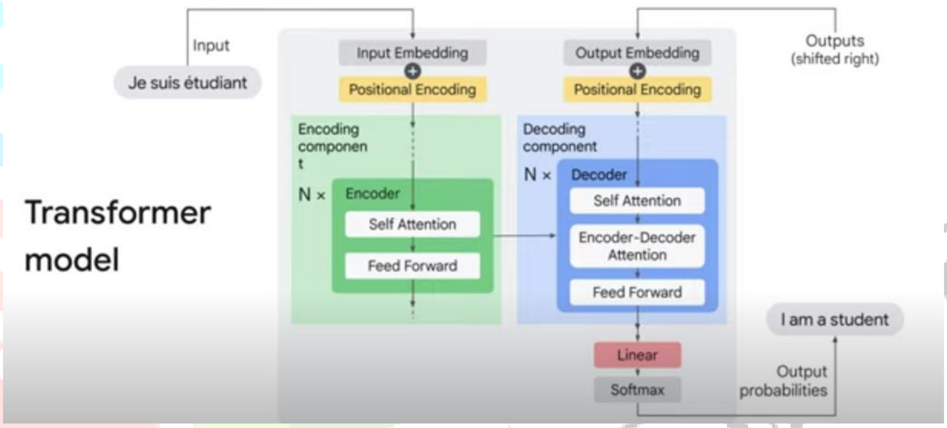

Figure 5 : Transformer Models and BERT Model: Overview

6. Attention Mechanisms:

Attention mechanisms have been integrated into various neural network architectures, including RNNs and Transformers. These mechanisms enable the model to focus more on certain parts of the input sequence, allowing it to assign different weights to different words or phrases[23]. Attention mechanisms are beneficial for feedback evaluation as they enable the model to give more significance to specific aspects of the feedback that contribute to the overall sentiment or opinion.

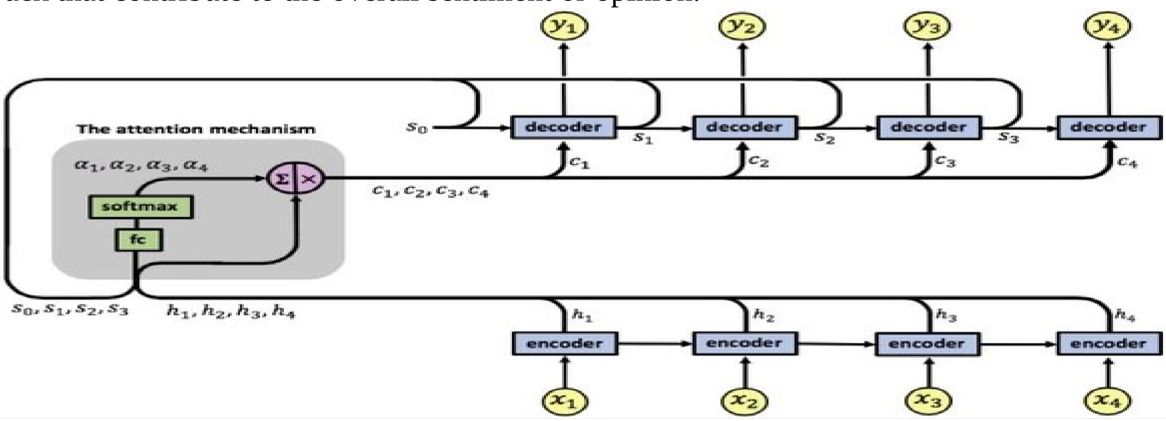

Figure 6 : Attention Mechanisms In RNNs

These deep learning architectures play a crucial role in advancing the field of feedback evaluation by providing models with the capacity to understand the contextual nuances, dependencies, and hierarchical structures present in natural language feedback. The choice of architecture depends on the specific characteristics of the feedback data and the goals of the evaluation task.

III. METHODOLOGY

To prepare textual data for traditional machine learning algorithms or techniques like topic modeling, it is imperative to undergo feature extraction and feature selection as crucial preprocessing steps. This transformation converts the students' feedback data into quantitative vector formats. This section delves into the existing methods employed in feature extraction, feature selection, and topic modeling.

3.1. Feature Extraction

Feature extraction techniques play a pivotal role in preparing students' feedback data for machine learning modeling, especially in the realm of Natural Language Processing (NLP). Within NLP, various feature extraction methods are utilized, including Bag of Words (BoW), Term Frequency-Inverse Document Frequency (TF-IDF), and Word Embedding[22].

1) **Bag of Words (BoW)[22]**It represents a prevalent feature extraction method involving a vocabulary of known words and a measurement of their presence in a document. BoW is solely concerned with the existence of known words, disregarding the structure or order of words in a document, thereby overlooking the context [24].
2) **TF-IDF[25]** Itestimates the importance of each word or term in a document based on their weights [26]. The Inverse Document Frequency (IDF) of a word indicates its commonality or rarity in a corpus. A lower IDF value signifies greater commonality in a corpus.
3) **Word Embedding** [27] It is a learned representation of text with similar meaning, contributing to generalization and dimensionality reduction. Common word embedding techniques include Word2Vec, GloVe, Doc2Vec [28], and Bidirectional encoder representations from transformers (BERT) [29]. The Word2Vec algorithm, for instance, builds on a neural networks model to discern word associations from a large text corpus. It can detect synonymous words and suggest additional words for incomplete sentences. Word2Vec assigns dimensions to each word based on its occurrence context in a sentence, grouping words with similar contexts in a vector space. GloVe combines matrix factorization and latent semantic analysis (LSA), and Doc2Vec generates numeric representations of documents. BERT, being a pre-trained deep bidirectional model, can dynamically produce word representations based on the surrounding words[30].

In the judgment conducted by Waykole et al.[31] different feature extraction techniques such as container of words, TF-IDF, and Word2Vec were compared using machine learning algorithms like logistic regression and random forest classifier, revealing Word2Vec as a superior feature extraction technique. Similar comparisons in other studies, such as those by Deepa et al.[32], demonstrated the effectiveness of feature extraction techniques like count vectorizer, Word2Vec, and TF-IDF, outperforming dictionary-based methods like valence-aware dictionary and sentiment reasoner (VADER) and SentiWordNet. Notably, count vectorizer achieved the highest classification accuracy of 81%. Twitter text analysis, akin to students' feedback for open-ended questions in educational institutions, benefits from these feature extraction techniques.

TF-IDF feature extraction generates high-dimensional feature vectors in a large text corpus[33]. In a study evaluating TF-IDF extraction, dimensionality reduction techniques such as latent semantic analysis (LSA) and linear discriminant analysis (LDA) were applied. The research concluded that TF-IDF outperformed the other two approaches (TF-IDF LSA and TF-IDF LDA) with larger datasets. In instances of smaller datasets, TF-IDF and TF-IDF LSA achieved comparable accuracy, while TF-IDF LDA faced challenges in accurate representation.

In the domain of deep learning, models such as LSTM, LSTM+ATT, multi-head ATT, and fusion have been compared for accuracy in text classification[35], showcasing varying levels of performance. Additionally, Zhang et al. [36] proposed a fine-tuned BERT model for sentiment analysis of student feedback to courses, integrating grammatical constraints and double attention layers for enhanced sentiment analysis. Masala et al. [37] utilized the BERT model to extract keywords from student feedback, reducing the feedback text by 59% with a marginal increase in mean average error. Wu et al.[38] introduced pre-trained word embedding to automatically create clusters for homogeneous and heterogeneous student groups based on their knowledge, facilitating collective feedback and collaborative learning, respectively.

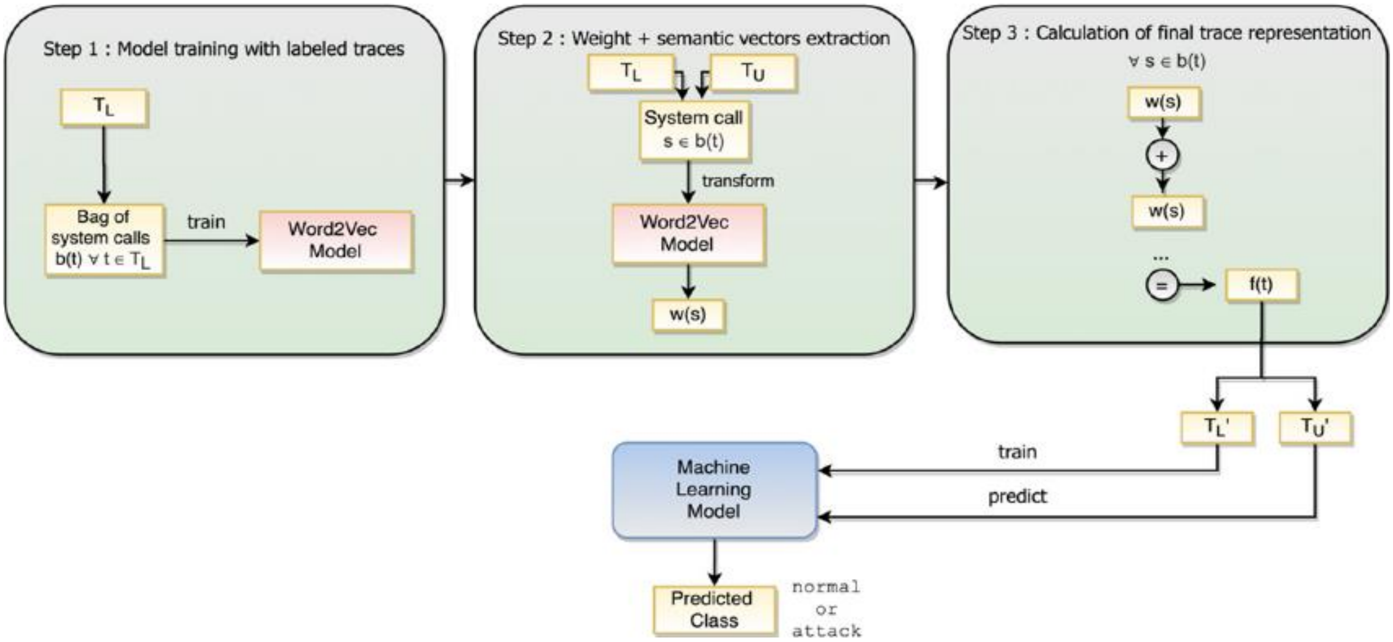

Figure 7 : Feature Extraction Machine Learning

Feature extraction methods typically disassemble students' feedback data into word tokens to facilitate semantic and grammatical analysis. Neural network-based Bag of Words (BoW), TF-IDF for word frequency, and word embedding techniques like Word2Vec, GloVe, Cove, and BERT aid in reducing the dimensionality of words while grouping those with similar contexts. The effectiveness of these feature extraction methods is often gauged through comparisons involving machine learning and deep learning methodologies[27][28][31][34][36][39].

3.2 Feature Selection

Feature selection is a crucial process aimed at reducing data dimensionality in terms of features, thereby preserving or enhancing the performance of a machine learning algorithm. The primary objective is to simplify a model's complexity while consistently maintaining accuracy. With n features in a dataset, the potential number of feature subsets becomes 2^n, and an increase in feature count could render modeling infeasible[40][41]. Evaluating the stability or robustness of feature subsets involves grouping similar features, considering all feature subsets, eliminating non-contributing features, and examining the size of feature subsets. Feature subset evaluation methods broadly fall into three categories: filters, wrappers, or embedded methods[42][43].

1) Filter Methods

Filter methods focus on ranking key features and selecting highly representative features by setting a threshold[44]. As illustrated in Figure 1, these methods rank and select features before actual modeling, filtering out low-importance features prior to training a model. The feature importance technique assesses two measures: predictive power and redundancy. Predictive power, measured through correlation criteria or dependence measures such as mutual information, χ^2 statistic, Markov blank, and minimal-redundancy-maximal-relevancy techniques, gauges a feature's correlation with the target variable(s). Redundancy, on the other hand, detects redundant features by evaluating relevant measures among independent variables. For instance, Wang[45] presented a redundant feature analysis method, identifying the most relevant features for predicting target variables and estimating redundancy in other features using these relevant features.

2) Wrapper Methods

Wrapper methods actively search for a subset of features using a predefined classifier, and the performance of this subset is then evaluated using predefined classifiers[44]. In wrapper methods, a machine learning algorithm enhances feature selection performance. As depicted in Figure 2, a subset of features is selected and trained by a classifier with the selected features. Subsequently, the classifier's performance is assessed. An example of a wrapper method is Sequential Forward Selection (SFS), a greedy search algorithm that iteratively extracts an optimal subset of features based on classifier performance. This iterative selection process involves sequentially choosing features from the pool of all features.

3) Embedded Methods

Embedded methods typically combine filter and wrapper methods[46], addressing the challenges of low accuracy in filter methods and slow computation speed in wrapper methods. These methods analyze optimum features contributing to the classifier's accuracy. As depicted in Figure 3, embedded methods estimate the performance of each subset of features. Regularization, one of the most common embedded methods, aims to reduce overfitting or variance by adding a penalty against model complexity, particularly for L1 regularization methods [47].

Parlar et al. [48]introduced a Query Expansion Ranking (QER) method for feature selection, comparing it with other methods such as information gain, Chi-square, document frequency difference, and optimal orthogonal centroid. They tested the model on English and Turkish review databases, demonstrating the superiority of the proposed feature selection method, particularly in the Naïve Bayes multinomial classifier. Similarly, feature selection was applied by Pong-Inwong et al. [49] in a teaching evaluation system using a filter method, reducing the number of attributes to 18 based on Chi-Square value. Machine learning algorithms, including ID3, J48, and Naïve Bayes, were then employed for student feedback classification, with the ensemble learning approach outperforming traditional algorithms with an accuracy of 87.16%.

Gutiérrez et al.[50] proposed a social mining model architecture for enhancing learning and e-learning quality based on students' feedback analysis. In their feature selection process, they used the random forest importance measure method to compute weights for each word, subsequently filtering them based on higher weights. The selected features were then fed into Support Vector Machine (SVM) classifiers with various kernels, and the SVM model with a radial kernel outperformed others with an accuracy of 85.17%. Additionally, Soukaina et al.[51] employed an information gain filter method for feature selection in an optimized sentiment analysis approach for students' feedback. Comparing SVM, random forest, and Naive Bayes classifiers before and after feature selection, random forest exhibited dominance before selection, while SVM outperformed with an accuracy of 85.9% after the feature selection.

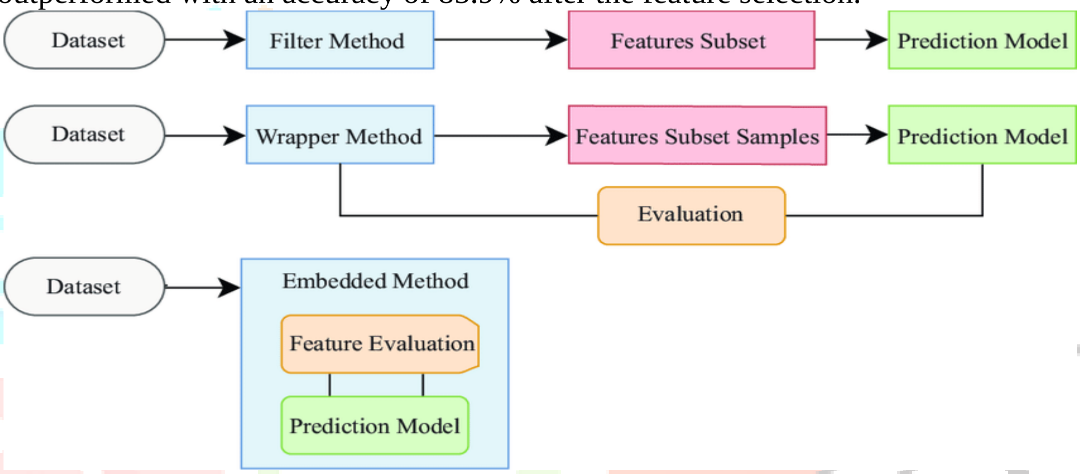

Figure 8: Feature Selection Methods for Machine Learning

Feature selection approaches exhibit noise resistance and aid in excluding irrelevant data, contributing to more effective data modeling. Modern feature selection methods have been proposed and compared with existing methods using machine learning methodologies[48][52].

3.3 Topic Modeling

Topic modeling refers to an automated procedure that employs machine learning techniques for text data to scrutinize a collection of documents, identifying grouped words [53][54]. In contrast to supervised methods, this technique doesn't necessitate training to categorize words within the corpus, operating as an unsupervised machine learning approach[55]. The principal objective of topic modeling is to partition a corpus of documents into clusters, revealing a list of encompassed topics. Subsequently, different sets of documents are organized based on the topics they cover. These techniques fall into two overarching categories: probabilistic and non-probabilistic models [56][57].

1) Non-Probabilistic Models

Non-probabilistic models involve algebraic approaches such as matrix factorization. Examples of these models include Latent Semantic Analysis (LSA) and Non-negative Matrix Factorization (NMF)[58]. Both LSA and NMF operate on Bag-of-Words (BoW) principles, where a corpus is transformed into a term-document matrix to capture term frequencies, disregarding the term order. LSA, an algebraic method, constructs a matrix representing words in a corpus, assuming that similar words are closely situated in the text[59]. It utilizes Single Value Decomposition (SVD) to reduce the number of words while preserving a similar structure. Text similarity is calculated using vector representation, and the texts are organized into semantic clusters. NMF transforms high-dimensional data into low-dimensional data without negative components, functioning as positive matrix factorization (PMF). This unsupervised technique extracts pertinent information without prior insights into the original data [60].

2) Probabilistic Models

Probabilistic models are entirely unsupervised approaches designed for latent Dirichlet Allocation (LDA) modeling and semi-supervised learning in probabilistic latent semantic analysis [61]. Probabilistic Latent Semantic Analysis (PLSA) is developed to identify the semantic co-occurrence of words or terms in a corpus[62]. Based on the initial statistical model revealing semantic co-occurrence in a document-term matrix, PLSA determines the number of topics, the probability of a topic, and the likelihood of a document containing the topic. It groups unknown topics in every existing document. LDA, a widely used technique in topic modeling, is constructed based on De Finetti's theorem, stating that positively correlated exchangeable observations are conditionally independent relative to some latent variable [63]. LDA captures inter and intra-document statistical structures, assuming a predefined number of topics in a corpus, with each document having a distinct proportion of the topics. It operates as a hidden variable model, revealing latent patterns in gathered data within a corpus.

To investigate the needs and perceptions of international students, Adriana et al.[70] Proposed a probabilistic topic model approach using LDA. Utilizing the machine learning for language toolkit (MALLET) [71], 20 topics were chosen based on 59,662 reviews. These topics encompassed language skills, convenient accommodation, weather, academic burdens, interesting courses, and more, ordered by their weight in the composition of the entire set of reviews. In the strategic planning for a university to enhance student enrollment, knowledge mining on online reviews was conducted using ensemble LDA (eLDA)[72]. The study employed multiple LDA models trained in parallel to extract the probabilistic score of words related to each generated topic. The held-out data were labeled using the trained LDA model and manually annotated with prior knowledge of identified topics in the database.

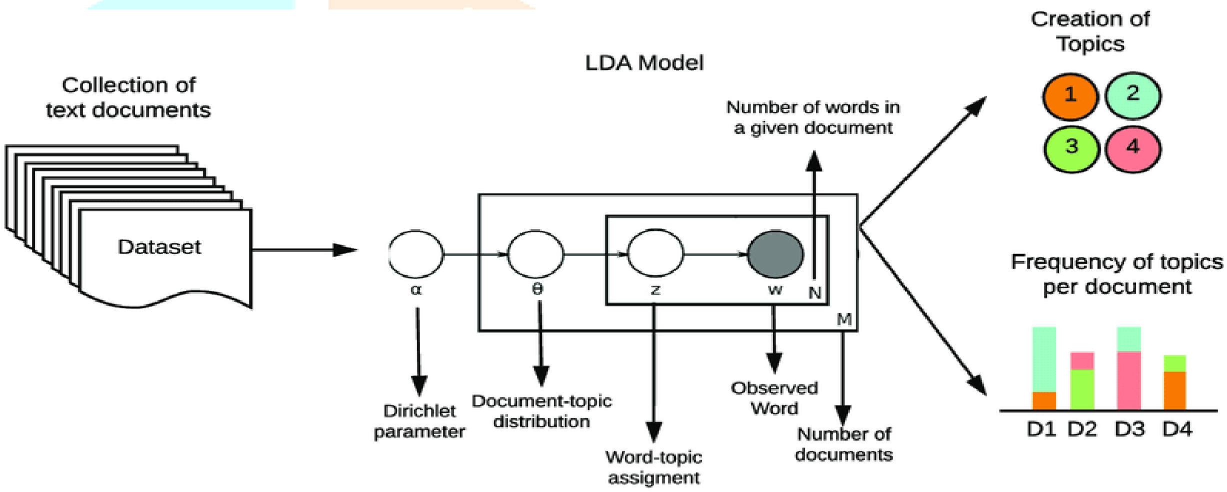

Figure 9 : Topic Modeling and Latent Dirichlet Allocation (LDA) using Gensim

3.4 Text Evaluation

In this subsection, we delve into NLP applications such as text summarization, document categorization, text annotation, and knowledge graphs.

1) Text Summarization

The exponential increase in student feedback collection in educational institutions has made content consolidation and resource extraction a laborious task. Text summarization techniques address this challenge by providing concise summaries of feedback, categorized into extractive, abstractive, and hybrid approaches.

Recent studies have explored different models for extractive summarization. Madhuri et al [100]. proposed a statistical method that assigns weights to each token based on its frequency, achieving high accuracy when compared to human summaries [99]. Fan et al. introduced Course MIRROR, employing automatic text summarization to aggregate students' feedback, outperforming existing techniques [100].

For abstractive summarization, Song et al. proposed a deep learning-based framework using LSTM-CNN, outperforming existing models in terms of evaluation metrics [101].

2) Document Categorization

Document categorization, or text classification, involves classifying the content, intent, and sentiment within a document into predefined labels. Li et al. integrated LSTM and CNN models for Chinese text classification, achieving high accuracy on benchmark datasets[19].

3) Entity Extraction

Entity annotation involves identifying named entities, parts of speech, and key phrases within a text. Dess et al. proposed an architecture for extracting entities and relations among entities, employing a deep learning model and entity detection module [102].

4) Knowledge Graphs

Knowledge graphs represent information abstractly and integrate data from various sources. Shi et al. proposed a learning path recommendation model based on a multidimensional knowledge graph framework, facilitating a well-organized learning path for students [103].

5) Sentiment Annotation

Sentiment annotation, a trending area in NLP, involves labeling emotion, opinion, and sentiment in a text. Kandhro et al. proposed an LSTM model for sentiment analysis, achieving high accuracy for positive and negative sentiment classification [104]. Hien et al. used NLP techniques to extract context and intention in student queries, achieving notable F1-scores for intent identification and context extraction [105].

Text annotation was utilized in a study on students' opinions in higher education, implementing the MATTER methodology for annotation, evaluation, and inter-rater agreement assessment [106].

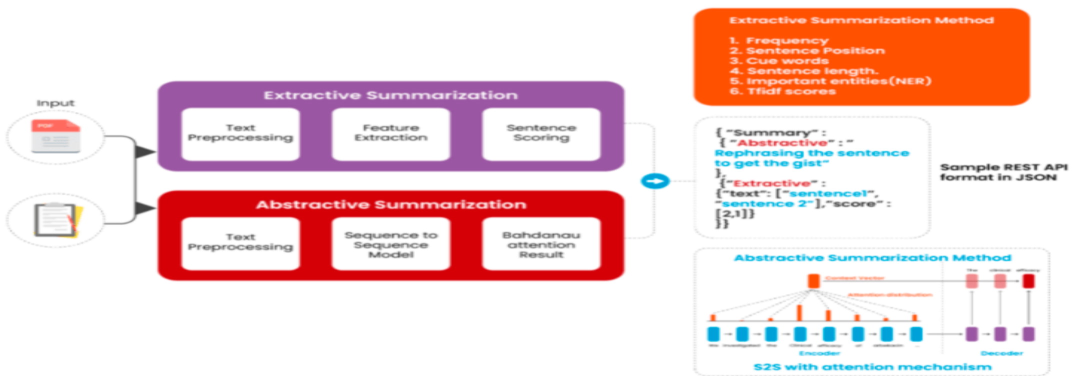

Figure 10 : Text Summarization in Natural Language Processing

IV. ANALYSIS OF SENTIMENT AND OPINION MINING

4.1 The Importance of Sentiment Analysis in Evaluating Feedback

Sentiment Analysis plays a crucial role in the assessment of feedback. Utilizing Natural Language Processing (NLP) techniques, it examines the emotional tone and attitude conveyed in textual data. In the realm of feedback evaluation, it facilitates the automated identification and categorization of sentiments, be they positive, negative, or neutral. This approach empowers institutions and organizations to gain valuable insights into the overall sentiment of feedback, fostering an understanding of stakeholders' perceptions and responses.

Effective Sentiment Analysis in feedback evaluation presents several benefits:

1) **Automated Processing:** Sentiment Analysis automates the laborious task of manually reviewing and categorizing each piece of feedback, allowing swift and efficient analysis of extensive datasets.
2) **Insight Generation:** It provides a deeper comprehension of the sentiments expressed by individuals, enabling organizations to pinpoint specific areas of satisfaction or concern. This insight proves invaluable for making data-driven decisions.
3) **Trend Identification:** By scrutinizing sentiments over time, organizations can discern trends and patterns in feedback. Recognizing evolving sentiments aids in proactive decision-making and the development of targeted interventions.
4) **Enhanced Customer Experience:** In customer-centric domains, Sentiment Analysis assists in gauging customer satisfaction and dissatisfaction levels. This information can be utilized to refine products, services, or communication strategies, ultimately enhancing the overall customer experience.

4.2 Techniques in Opinion Mining

Opinion Mining, also referred to as sentiment mining or sentiment analysis, involves the application of diverse techniques to extract subjective information from textual data.

Commonly used techniques in Opinion Mining include:

1) **Lexicon-Based Approaches:** These methods rely on predefined dictionaries or lexicons containing words associated with positive or negative sentiments. The sentiment of a text is determined by analyzing the presence of these words.
2) **Machine Learning Models:** Supervised machine learning models like Support Vector Machines (SVM) and Naive Bayes can be trained on labeled datasets to classify text into different sentiment categories. Unsupervised learning methods, such as clustering, are also applicable in Opinion Mining.
3) **Deep Learning Techniques:** Neural networks, especially Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks, are increasingly employed for sentiment analysis tasks due to their ability to capture contextual information in sequential data.
4) **Aspect-Based Opinion Mining:** This technique goes beyond overall sentiment analysis to identify sentiments associated with specific aspects or features of a product, service, or experience, providing a more detailed understanding of opinions.

4.3 Aspect-Oriented Sentiment Evaluation

Aspect-Based Sentiment Analysis (ABSA) is a specialized form of sentiment analysis that focuses on extracting sentiments related to specific aspects or features within a text. ABSA proves particularly valuable in feedback evaluation, allowing for a more detailed examination of sentiments associated with different elements of a product, service, or experience.

Key components of Aspect-Based Sentiment Analysis include:

1) **Aspect Extraction:** Identifying and extracting the aspects or features mentioned in the text associated with sentiments.
2) **Sentiment categorization:** Determining the response polarity (positive, negative, or neutral) articulated for each acknowledged portion.
3) **Contextual Understanding:** Considering the context in which aspects and sentiments are expressed, as the same aspect may elicit different sentiments based on the surrounding content.
4) **Aspect-Based Sentiment:** Analysis provides organizations with a nuanced view of stakeholder opinions, facilitating targeted improvements and strategic decision-making based on specific aspects of interest.

V. CASE STUDIES AND APPLICATIONS

5.1 Customer Feedback in E-commerce

In the realm of e-commerce, the analysis of customer feedback is crucial for understanding customer satisfaction, improving products, and enhancing overall user experience. Sentiment analysis and opinion mining techniques are employed to extract valuable insights from customer reviews, ratings, and comments. By categorizing sentiments expressed in feedback, e-commerce platforms can identify popular products, address customer concerns, and tailor marketing strategies to meet consumer expectations.

Case Study Example:

A major e-commerce platform utilized sentiment analysis to categorize customer reviews. Positive sentiments were associated with features like fast delivery, product quality, and user-friendly interfaces, while negative sentiments highlighted issues such as shipping delays, damaged products, or difficulties in the purchasing process. This analysis not only informed the platform about customer satisfaction but also guided improvements in logistics, packaging, and website functionality.

5.2 Educational Feedback Analysis

In the educational sector, feedback analysis is employed to assess the effectiveness of teaching methods, course content, and overall learning experiences. Natural Language processing techniques, including sentiment analysis and topic modeling, are utilized to extract meaningful information from student feedback. This aids educators and institutions in identifying areas of improvement, optimizing course structures, and enhancing the learning environment.

Case Study Example:

A university implemented sentiment analysis on student feedback to gain insights into the overall satisfaction of courses. Positive sentiments were associated with engaging lectures, helpful resources, and supportive faculty, while negative sentiments pinpointed areas requiring attention, such as challenging assessments or unclear instructions. This feedback analysis enabled the university to make targeted adjustments to curriculum design and teaching methodologies.

5.3 Healthcare and Patient Feedback

Patient feedback in the healthcare sector is invaluable for improving medical services, ensuring patient satisfaction, and enhancing overall healthcare experiences. Sentiment analysis and opinion mining are applied to patient reviews, survey responses, and online comments to discern sentiments related to healthcare facilities, medical staff, and the effectiveness of treatments.

Case Study Example:

A hospital employed sentiment analysis on patient feedback to understand the sentiments associated with different departments and services. Positive sentiments were linked to caring and competent staff, efficient processes, and successful treatments. Negative sentiments highlighted issues such as long wait times, communication challenges, or concerns about facility cleanliness. This analysis guided the hospital in implementing changes to enhance patient care and satisfaction.

5.4 Social Media Comments and Reviews

Social media platforms serve as vast sources of opinions and sentiments, making sentiment analysis crucial for understanding public perceptions. Brands and organizations leverage sentiment analysis to gauge reactions to their products, campaigns, or events. This enables them to adapt marketing strategies, address public concerns, and maintain a positive brand image.

Case Study Example:

A company monitored social media comments and reviews to assess the reception of a new product launch. Sentiment analysis helped categorize responses into positive, negative, or neutral sentiments. Positive sentiments were associated with product features, while negative sentiments highlighted issues such as pricing concerns or product functionality. This real-time feedback guided the company in making quick adjustments to marketing strategies and product offerings.

VI. EVALUATION METRICS

6.1 Accuracy, Precision, Recall

Evaluation metrics play a crucial role in assessing the performance of models and systems used for feedback evaluation. The following metrics are commonly employed:

1) **Accuracy:** This metric measures the overall correctness of the system by calculating the ratio of correctly predicted instances to the total instances. It is represented as:
$$\text{Accuracy} = \frac{\text{Number of Correct Predictions}}{\text{Total Number of Predictions}}$$
2) **Precision:** Precision quantifies the accuracy of positive predictions made by the system. It is calculated as the ratio of true positive predictions to the total positive predictions (both true positives and false positives):
$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives + False Positives}}$$
3) **Recall (Sensitivity):** Recall evaluates the ability of the system to capture all relevant instances, measuring the ratio of true positive predictions to the total actual positives (true positives and false negatives):
$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives + False Negatives}}$$

6.2 F1 Score

The F1 score is the harmonic mean of precision and recall. It provides a balanced assessment of a system's performance by considering both false positives and false negatives. The Following formula for F1:
$$F1 \, \text{Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision + Recall}}$$

6.3 Area under the Receiver Operating Characteristic (ROC-AUC)

In cases involving binary classification, where systems predict positive or negative outcomes, the ROC-AUC metric is employed. The Receiver Operating Characteristic (ROC) curve visualizes the trade-off between true positive rate (sensitivity) and false positive rate. The Area under the Curve (AUC) quantifies the overall performance of the model.

6.4 Challenges in Defining Metrics for Feedback Evaluation

Defining appropriate metrics for feedback evaluation encounters challenges due to the subjective nature of feedback and the diversity of domains.

Challenges include:

1) **Subjectivity:** Feedback often contains subjective elements, making it challenging to establish a universal metric that accurately captures its nuances.
2) **Multi-dimensional Evaluation**: Feedback may encompass various dimensions such as sentiment, relevance, and specificity, requiring a combination of metrics to provide a comprehensive evaluation.
3) **Context Dependency:** Metrics need to consider the context of feedback and the goals of evaluation. What is considered a successful prediction may vary based on the application.
4) **Imbalance in Data:** Class imbalance, where one class significantly outnumbers the other, can affect the effectiveness of metrics. In such cases, alternative metrics like precision-recall curves may be more informative than accuracy.
5) **Dynamic Nature of Feedback**: Feedback can evolve over time, and metrics need to adapt to changing patterns and sentiments.

VII. CHALLENGES AND FUTURE DIRECTIONS

This section delves into the challenges associated with implementing NLP techniques in the education domain.

1) Domain-Specific Language

To effectively classify academic datasets or students' feedback, a profound understanding of the core factors within the teaching context is essential [107]. This poses a significant challenge in implementing NLP in the education domain. Given the abundance of student feedback from diverse surveys, questionnaires, and educational feedback portals related to course teaching or learning management systems, NLP methodologies face the hurdle of comprehending or being trained on specific domains. Nhi et al. [108] addressed this challenge by introducing a domain-specific NLP tailored for students, faculty members, and universities in computer science or information technology within higher education. Their approach involved extracting tech-related skills through named entity recognition (NER) and constructing a personalized multi-level course recommendation system. This domain-specific NER was designed to gather data, such as job postings, course descriptions, and information from Massive Open Online Courses (MOOCs), from various websites. The system was enriched using an annotated corpus from StackOverflow and GitHub [109]. Another strategy proposed by Pashev et al. [110] involved extracting entities and relations using MeaningCloud API and Google Translate API, calculating grades based on relevance to teacher-created topics or auto-generated text from the subject area.

2) Sarcasm

Decoding sarcasm is a crucial aspect of NLP tasks like sentiment annotation and opinion analysis, particularly for understanding student opinions on course structure and educational infrastructure. A survey by [111] explicitly studied automatic sarcasm detection, highlighting the challenges and research gaps in this area. Sarcasm detection involves three main approaches: rule-based, statistical, and deep learning. Rule-based approaches identify sarcasm based on key indicators captured as evidence, while statistical approaches consider features such as punctuations, sentiment-lexicon-based features, unigrams, word embedding similarity, frequency of rare words, and sentiment flips. Traditional machine learning algorithms and deep learning algorithms, including RNN and LSTM methods individually or in combination with CNNs, are employed for automatic sarcasm detection. This survey provided a comprehensive understanding of sarcasm detection.

3) Ambiguity

Ambiguity is inherent in natural languages, and its complexity increases in machine learning language processing due to challenges in decoding context. Ambiguity can arise structurally, syntactically, or lexically within a sentence [112]. Addressing ambiguity is crucial in feedback analysis. In a study[85], word sense disambiguation was tackled by customizing BERT, a language representation model, and selecting the best context-gloss pairs from a group of related pairs. The context-gloss pairs were classified into positive and negative sentiment, and the proposed BERT model outperformed existing state-of-the-art models.

4) Emoticons And Special Characters

Emoticons and special characters, frequently used in students' feedback to express emotions, present a challenge for NLP in opinion mining. Processing these emoticons and appropriately labeling them with emotion tags is a challenging phase in NLP. In a 2020 study [113], cross-cultural reactions to the novel coronavirus were analyzed from tweets, and sentiment polarity and emotion were detected and validated using emoticons. A deep learning model based on LSTM, combined with feature extraction methods like GloVe and word embeddings, was used for this analysis. Another study by Cappallo et al. [114] proposed a dataset with real-world emojis, addressing challenges in emoji processing, anticipation, and query-by-emoji. They utilized a Bi-LSTM model for text-to-emoji baseline results and a CNN model for image-to-emoji, combining both for a multi-modal approach to emoticons processing.

5) Aspect-Based Sentiment Analysis

Chauhan et al. [115] highlighted that sentiment analysis tools, often underused in education, struggle to identify opinions on different aspects. Most research focuses on classifying positive or negative sentiment at the document level, overlooking opinions on specific aspects in student comments or feedback. Nazir et al. [116] conducted a survey addressing issues and challenges related to the extraction of different aspects in sentiment analysis, including explicit aspect extraction, implicit aspect extraction, aspect-level sentiment analysis, entity-level sentiment analysis, multi-word sentiment analysis, recognition of factors in sentiment evolution, and predicting sentiment evolution.

6) Data Imbalance

One prevalent challenge in AI [117] is data imbalance, where the number of samples in one class surpasses that in other classes. In the realm of NLP, a subset of AI, this challenge persists. Particularly in the education domain, obtaining massive labeled data is arduous due to the need for manual annotation by domain experts. Even when labeled data is obtained and utilized by deep learning algorithms, classification performance tends to be biased due to discrepancies in data distribution [118]. An effective approach to tackle this challenge involves leveraging transfer learning [119]. This entails training a deep learning model on an extensive corpus of student feedback to perform similar tasks on another data source. Additional techniques to address data imbalance include sampling methods [120], such as under-sampling majority classes or over-sampling minority classes, which may necessitate tasks related to text augmentation [121].

VIII. DISCUSSION

In Figure 8 [122], a Gartner diagram illustrates that decision intelligence, deep learning, and knowledge graphs have peaked, signaling their adaptability to the education domain for constructing decision support systems. These domains efficiently analyze existing data and streamline data storage processes. Deep learning methods, requiring minimal expertise in the application domain, can construct semantic networks to store interlinked entity data within a domain. The projected shift of 70% of organizations from big to small and wide data by 2025 underscores the need for diverse, structured, and unstructured data sources [123].

NLP, as an integral component of AI, facilitates the comprehension of human language, enabling the understanding of opinions and feedback. Educational institutions stand to gain significantly from adopting NLP methods to enhance the student learning experience, personalized learning management systems [9], and teacher training. This transformation exposes students to AI-driven tools. The purpose of AI in learning extends to an assortment of areas, including well-dressed classrooms with video and audio data annotation [124][125], NLP for textual data annotation, classification, summarization, and image processing for gesture detection [126][127] . While this study primarily focuses on NLP methodologies, it is crucial to note that even research articles indirectly related to AI in education can be adapted for educational purposes.

Holmes et al. [128] addressed the challenges and future implications of AI in education, emphasizing two key questions: "What we teach" and "How we teach it." "What we teach" pertains to the learning goals, emphasizing versatility, relevance, and transferability. Strategies to achieve these goals include selective emphasis on traditional knowledge areas, incorporation of modern knowledge, interdisciplinary concepts, embedded skills, and meta-learning. The question of "How we teach it" involves the transformation and enhancement of education through AI. The authors differentiated between educational technology, aimed at amending the field's taxonomy and ontology, and AI in education, which involves a layered framework encompassing substitution, augmentation, modification, and redefinition. This framework seeks to enhance and transform education, moving beyond the mere enhancement of teaching practices. Intelligent tutoring systems, incorporating domain models, pedagogy models, and learner models, were proposed to enhance individual student learning.

NLP techniques can be applied to feedback data using various programming languages. While different languages offer pre-built packages for NLP tasks, Python, Java, and R programming languages are widely employed [129]. Considerations in selecting a programming language include expertise, the availability of libraries or package tools for NLP tasks [130]. Python [131], known for its versatility and human-like syntax, provides a vast array of NLP packages for tasks like topic modeling, word embeddings, document classification, and sentiment annotations. Java [132], as a platform-independent language, supports robust architecture for comprehensive text analysis tasks such as clustering, tagging, and information extraction. R programming language, renowned for statistical learning, is extensively used in NLP tasks, especially in handling computationally intensive data analytics and investigating big data applications. Table 1 outlines these three programming languages, their respective NLP packages, features, and documentation sources.

In this study, four investigate questions pertaining to NLP in learning were explored. The first question delved into the existing methodologies used in NLP, covering data preprocessing methods, feature extraction, feature selection, and a comprehensive discussion on machine learning and deep learning models. Additionally, various approaches to topic modeling and student feedback topic extraction were explored. Text evaluation techniques, such as text summarization, document categorization, text annotation, and knowledge graphs, were extensively explained.

The second research question listening cautiously on the challenges of NLP in the teaching domain. Generic NLP challenges, including domain-specific language, sarcasm, ambiguity, and data imbalance, pose difficulties in presentation the latent semantic meaning of undergraduate feedback. The research community has addressed these challenges using approaches such as Named Entity Recognition (NER), rule-based methods, statistical techniques, deep learning, and BERT modeling. Emoticons and special characters, used by students to express sentiment in feedback, are processed through multimodal approaches, converting emojis to corresponding unicodes or employing image processing to determine sentiment. Aspect-based feeling psychotherapy, also known as fine-grained emotion analysis, emerges as a trending confront in NLP within the teaching sphere.

The third research question sought to identify trends in NLP methodologies applicable to the education domain. Various language processing methods were discussed in detail, emphasizing the need for a quantitative approach to train AI models. This involves preprocessing textual data into vectors using feature extraction and feature selection techniques. Topic modeling techniques, encompassing both probabilistic and non-probabilistic models, were explored, with Latent Dirichlet Allocation (LDA) emerging as a commonly used technique for unsupervised topic extraction from a corpus.

The fourth research question explored the community's work from other industry applications related to short text analysis, aiming to understand and adapt these findings to the education domain. Studies on Twitter text analysis using VADER and SentiWordNet techniques [32][56], as well as query expansion ranking [48], present valuable insights applicable to student feedback analysis. Approaches to analyzing Amazon product reviews [54] and processing emojis in student comments[114] offer further perspectives for application in higher education feedback analysis.

The discussion also highlights the distribution of references over the years (Figure 9), providing insights into the evolving landscape of NLP in education. The references distribution (Figure 10) showcases the diverse sources contributing to the study, reflecting a comprehensive exploration of NLP methodologies and their applications in the education domain.

IX. CONCLUSION

This study aimed to explore existing NLP methodologies applicable or adoptable in the education domain, providing insights into the impact of AI on education and open opportunities. The goal was to synthesize methods for processing student feedback and annotating their views. The search results from Google Scholar were meticulously examined to identify relevant NLP techniques. The primary objective of this investigation was to examine the existing methodologies in Natural Language Processing (NLP) that are applicable or adaptable within the education domain. The study aimed to shed light on the influence of Artificial Intelligence (AI) on education and the potential opportunities it opens up. The focus was on synthesizing techniques for processing student feedback and annotating their perspectives. The scrutiny of search results from Google Scholar was meticulous, aiming to pinpoint relevant NLP techniques that could be applied to student feedback or educational applications for analysis.

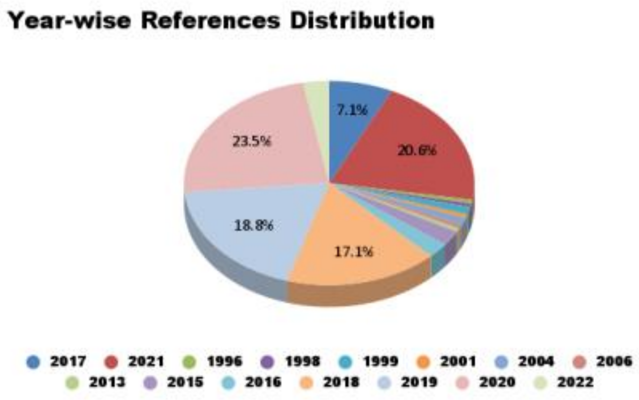

FIGURE 9. Year-wise References Distribution

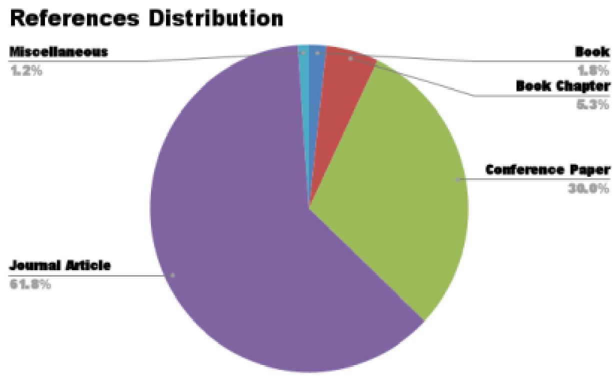

FIGURE 10. References Distribution

Figure 9 illustrates that a significant proportion of references in this study originate from the last five years, indicating recent advancements in NLP methodologies. Additionally, over 90% of the citations included in this study are drawn from journal articles and conference papers. Table 2 provides a comprehensive overview of the NLP techniques investigated in this study, along with citations from the research community.

The review article delves into the impact of AI on education, delineating the potential for introducing AI into educational institutions based on available opportunities. While the primary focus is on introducing NLP methodologies for feedback analysis in education, the article explores the existing landscape of NLP methodologies. It furnishes explanations and definitions for feature extraction, feature selection, and topic modeling methodologies. Moreover, text evaluation techniques, including text summarization, annotation, and knowledge graphs, are thoroughly examined, with each application defined and existing approaches discussed. The study also addresses challenges in the adoption of NLP methodologies in the education domain.

However, it is crucial to acknowledge certain limitations in this research. The study confines itself to AI implementation methodologies with less emphasis on pedagogical concepts. Specific challenges related to data, such as data scarcity and class imbalance, were not extensively discussed, despite their potential impact on deep learning algorithms, which depend on substantial data. Furthermore, strategies for interpreting deep learning models, often regarded as "black boxes," were not explored in depth. A promising avenue for future research could involve an exploration of data challenges associated with extracting feedback or opinions without compromising privacy.

REFERENCES:

[1] M. A. Peters, ‘‘Deep learning, education and the final stage of automation,’’Educ. Philosophy Theory, vol. 50, nos. 6–7, pp. 549–553, Jul. 2017, doi:10.1080/00131857.2017.1348928.
[2] X. J. Hunt, I. K. Kabul, and J. Silva, ‘‘Trasfer learning for education data,’’in Proc. ACM SIGKDD Conf., vol. 1. Halifax, NS, Canada, 2017, pp. 1–6.
[3] Z. Kastrati, F. Dalipi, A. S. Imran, K. P. Nuci, and M. A. Wani, ‘‘Sentimentanalysis of students’ feedback with NLP and deep learning: A systematic mapping study,’’ Appl. Sci., vol. 11, no. 9, p. 3986, Apr. 2021, doi:10.3390/app11093986. OECD Education Working Papers, Org. Econ. Co-Oper. Develop., Paris,France, 2021, doi: 10.1787/19939019.
[4] M. Chassignol, A. Khoroshavin, A. Klimova, and A. Bilyatdinova, ‘‘Artificial intelligence trends in education: A narrative overview,’’ Proc. Comput.Sci., vol. 136, pp. 16–24, Jan. 2018, doi: 10.1016/j.procs.2018.08.233.
[5] L. Chen, P. Chen, and Z. Lin, ‘‘Artificial intelligence in education:A review,’’ IEEE Access, vol. 8, pp. 75264–75278, 2020.
[6] M. L. B. Estrada, R. Z. Cabada, R. O. Bustillos, and M. Graff,‘‘Opinion mining and emotion recognition applied to learning environments,’’ Expert Syst. Appl., vol. 150, Jul. 2020, Art. no. 113265, doi:10.1016/j.eswa.2020.113265.
[7] A. Yadav and D. K. Vishwakarma, ‘‘Sentiment analysis using deep learningarchitectures: A review,’’ Artif. Intell. Rev., vol. 53, no. 6, pp. 4335–4385,Dec. 2019, doi: 10.1007/s10462-019-09794-5.

[8] K. Eggert, ''How artificial intelligence will shape universities of tomorrow,'' in Proc. Int. Conf., 2021, p. 50.
[9] K. Holstein, B. M. McLaren, and V. Aleven, ''Designing for complementarity: Teacher and student needs for orchestration support in AI-enhancedclassrooms,'' in Artificial Intelligence in Education (Lecture Notes inComputer Science). Cham, Switzerland: Springer, 2019, pp. 157–171, doi:10.1007/978-3-030-23204-7_14.
[10] A. Bhimdiwala, R. C. Neri, and L. M. Gomez, ''Advancing the design andimplementation of artificial intelligence in education through continuousimprovement,'' Int. J. Artif. Intell. Educ., vol. 11625, pp. 1–27, Oct. 2021,doi: 10.1007/s40593-021-00278-8.
[11] K. Holstein, B. M. McLaren, and V. Aleven, ''Co-designing a realtime classroom orchestration tool to support teacher–AI complementarity,'' J. Learn. Anal., vol. 6, no. 2, pp. 27–52, Jul. 2019, doi:10.18608/jla.2019.62.3.
[12] H. Labarthe, V. Luengo, and F. Bouchet, ''Analyzing the relationships between learning analytics, educational data mining and ai for education,''in Proc. 14th Int. Conf. Intell. Tutoring Syst. (ITS), Workshop Learn. Anal.,2018, pp. 10–19.
[13] L. Alrajhi, A. Alamri, F. D. Pereira, and A. I. Cristea, ''Urgency analysisof learners' comments: An automated intervention priority model forMOOC,'' in Proc. Int. Conf. Intell. Tutoring Syst. Cham, Switzerland:Springer, 2021, pp. 148–160.
[14] R. C. Sharma, P. Kawachi, and A. Bozkurt, ''The landscape of artificialintelligence in open, online and distance education: Promises and concerns,'' Asian J. Distance Educ., vol. 14, no. 2, pp. 1–2, 2019.
[15] K. Gulson, A. Murphie, S. Taylor, and S. Sellar, ''Education, work andAustralian society in an AI world,'' Gonski Inst. Educ., Univ. New SouthWales Sydney, Sydney, NSW, Australia, Tech. Rep., 2018.
[16] F. Pedro, M. Subosa, A. Rivas, and P. Valverde, ''Artificial intelligencein education: Challenges and opportunities for sustainable development,''Nat. Inst. Educ. Develop., Okahandja, Namibia, Tech. Rep., 2019.
[17] S. Gottipati, V. Shankararaman, and S. Gan, ''A conceptual framework foranalyzing students' feedback,'' in Proc. IEEE Frontiers Educ. Conf. (FIE),Oct. 2017, pp. 1–8.
[18] Y. Li, X. Wang, and P. Xu, ''Chinese text classification model based ondeep learning,'' Future Internet, vol. 10, no. 11, p. 113, Nov. 2018, doi:10.3390/fi10110113.
[19] S. Ramaswamy and N. DeClerck, ''Customer perception analysis usingdeep learning and NLP,'' Proc. Comput. Sci., vol. 140, pp. 170–178,Jan. 2018, doi: 10.1016/j.procs.2018.10.326.
[20] S. Prokhorov and V. Safronov, ''AI for AI: What NLP techniques helpresearchers find the right articles on NLP,'' in Proc. Int. Conf. Artif.Intell., Appl. Innov. (IC-AIAI), Sep. 2019, pp. 765–776, doi: 10.1109/icaiai48757.2019.00023.56734 VOLUME 10, 2022
[21] T. Shaik et al.: Review of Trends and Challenges in Adopting NLP Methods for Education Feedback Analysis
[22] R. Ahuja, A. Chug, S. Kohli, S. Gupta, and P. Ahuja, ''The impact offeatures extraction on the sentiment analysis,'' Proc. Comput. Sci., vol. 152,pp. 341–348, Jan. 2019, doi: 10.1016/j.procs.2019.05.008.
[23] A. Balamurali and B. Ananthanarayanan, ''Develop a neural model toscore bigram of words using bag-of-words model for sentiment analysis,''in Neural Networks for Natural Language Processing. Pennsylvania, PA,USA: IGI Global, 2020, pp. 122–142, doi: 10.4018/978-1-7998-1159-6.ch008.
[24] Y. Goldberg and G. Hirst, Neural Network Methods in Natural LanguageProcessing. San Rafael, CA, USA: Morgan & Claypool, 2017.
[25] I. Arroyo-Fernández, C.-F. Méndez-Cruz, G. Sierra, J.-M. Torres-Moreno,and G. Sidorov, ''Unsupervised sentence representations as word information series: Revisiting TF–IDF,'' Comput. Speech Lang., vol. 56,pp. 107–129, Jul. 2019, doi: 10.1016/j.csl.2019.01.005.
[26] F. M. Shah, F. Haque, R. U. Nur, S. A. Jahan, and Z. Mamud,''A hybridized feature extraction approach to suicidal ideation detectionfrom social media post,'' in Proc. IEEE Region 10 Symp. (TENSYMP),Jun. 2020, pp. 985–988, doi: 10.1109/TENSYMP50017.2020.9230733.
[27] A. Onan, ''Sentiment analysis on product reviews based on weighted wordembeddings and deep neural networks,'' Concurrency Comput., Pract.Exper., vol. 33, no. 23, p. e5909, Jun. 2020, doi: 10.1002/cpe.5909.

[28] S. Amin, M. I. Uddin, M. A. Zeb, A. A. Alarood, M. Mahmoud, andM. H. Alkinani, ‘‘Detecting dengue/flu infections based on tweets usingLSTM and word embedding,’’ IEEE Access, vol. 8, pp. 189054–189068,2020, doi: 10.1109/ACCESS.2020.3031174.
[29] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, ‘‘BERT: Pre-trainingof deep bidirectional transformers for language understanding,’’ 2018,arXiv:1810.04805.
[30] J. Hou, X. Li, H. Yao, H. Sun, T. Mai, and R. J. I. A. Zhu, ‘‘Bert-basedChinese relation extraction for public security,’’ IEEE Access, vol. 8,pp. 132367–132375, 2020.
[31] R. N. Waykole and A. D. Thakare, ‘‘A review of feature extraction methodsfor text classification,’’ Int. J. Adv. Eng. Res. Develop., vol. 5, no. 4,pp. 351–354, 2018.
[32] D. Deepa and A. Tamilarasi, ‘‘Sentiment analysis using feature extractionand dictionary-based approaches,’’ in Proc. 3rd Int. Conf. I-SMAC (IoTSocial, Mobile, Anal. Cloud) (I-SMAC), Dec. 2019, pp. 786–790, doi:10.1109/i-smac47947.2019.9032456.
[33] R. Dzisevic and D. Sesok, ‘‘Text classification using different featureextraction approaches,’’ in Proc. Open Conf. Electr., Electron. Inf. Sci.(eStream), Apr. 2019, pp. 1–4, doi: 10.1109/estream.2019.8732167.
[34] D. Goularas and S. Kamis, ‘‘Evaluation of deep learning techniques insentiment analysis from Twitter data,’’ in Proc. Int. Conf. Deep Learn.Mach. Learn. Emerg. Appl. (Deep-ML), Aug. 2019, pp. 12–17, doi:10.1109/deep-ml.2019.00011.
[35] K. Sangeetha and D. Prabha, ‘‘Sentiment analysis of student feedbackusing multi-head attention fusion model of word and context embedding for LSTM,’’ J. Ambient Intell. Humanized Comput., vol. 12, no. 3,pp. 4117–4126, Mar. 2020, doi: 10.1007/s12652-020-01791-9.
[36] Zhang, J. Dong, L. Min, and P. Bi, ‘‘A BERT fine-tuning model fortargeted sentiment analysis of Chinese online course reviews,’’ Int. J. Artif.Intell. Tools, vol. 29, no. 7, Dec. 2020, Art. no. 2040018.
[37] M. Masala, S. Ruseti, M. Dascalu, and C. Dobre, ‘‘Extracting and clustering main ideas from student feedback using language models,’’ in Proc. Int.Conf. Artif. Intell. Educ. Cham, Switzerland: Springer, 2021, pp. 282–292.
[38] Y. Wu, J. Nouri, X. Li, R. Weegar, M. Afzaal, and A. Zia, ‘‘A wordembeddings based clustering approach for collaborative learning groupformation,’’ in Proc. Int. Conf. Artif. Intell. Educ. Cham, Switzerland:Springer, 2021, pp. 395–400.
[39] D. Wang, J. Su, and H. Yu, ‘‘Feature extraction and analysis of naturallanguage processing for deep learning English language,’’ IEEE Access,vol. 8, pp. 46335–46345, 2020.
[40] B. Ghojogh, M. N. Samad, S. A. Mashhadi, T. Kapoor, W. Ali, F. Karray,and M. Crowley, ‘‘Feature selection and feature extraction in pattern analysis: A literature review,’’ 2019, arXiv:1905.02845.
[41] J. Cai, J. Luo, S. Wang, and S. Yang, ‘‘Feature selection in machine learning: A new perspective,’’ Neurocomputing, vol. 300, pp. 70–79, Jul. 2018,doi: 10.1016/j.neucom.2017.11.077.
[42] G. Kou, P. Yang, Y. Peng, F. Xiao, Y. Chen, and F. E. Alsaadi, ‘‘Evaluationof feature selection methods for text classification with small datasets usingmultiple criteria decision-making methods,’’ Appl. Soft Comput., vol. 86,Jan. 2020, Art. no. 105836, doi: 10.1016/j.asoc.2019.105836.
[43] N. Nikolić, O. Grljević, and A. Kovačević, ‘‘Aspect-based sentiment analysis of reviews in the domain of higher education,’’ Electron. Library,vol. 38, no. 1, pp. 44–64, Feb. 2020, doi: 10.1108/el-06-2019-0140.
[44] A. Bommert, X. Sun, B. Bischl, J. Rahnenführer, and M. Lang, ‘‘Benchmark for filter methods for feature selection in high-dimensional classification data,’’ Comput. Statist. Data Anal., vol. 143, Mar. 2020, Art. no. 106839, doi: 10.1016/j.csda.2019.106839.
[45] M. Wang, X. Tao, and F. Han, ‘‘A new method for redundancy analysis infeature selection,’’ in Proc. 3rd Int. Conf. Algorithms, Comput. Artif. Intell.,Dec. 2020, pp. 1–5, doi: 10.1145/3446132.3446153.
[46] P. Dhal and C. Azad, ‘‘A comprehensive survey on feature selection in thevarious fields of machine learning,’’ Int. J. Speech Technol., vol. 52, no. 4,pp. 4543–4581, Jul. 2021, doi: 10.1007/s10489-021-02550-9.
[47] H. Osman, M. Ghafari, and O. Nierstrasz, ‘‘Automatic feature selection byregularization to improve bug prediction accuracy,’’ in Proc. IEEE Workshop Mach. Learn. Techn. Softw. Quality Eval. (MaLTeSQuE), Feb. 2017,pp. 27–32, doi: 10.1109/MALTESQUE.2017.7882013.
[48] [48] T. Parlar, S. A. Özel, and F. Song, ‘‘QER: A new feature selection methodfor sentiment analysis,’’ Hum.-Centric Comput. Inf. Sci., vol. 8, no. 1,pp. 1–19, Dec. 2018.

[49] C. Pong-Inwong and K. Kaewmak, ‘‘Improved sentiment analysis forteaching evaluation using feature selection and voting ensemble learningintegration,’’ in Proc. 2nd IEEE Int. Conf. Comput. Commun. (ICCC),Oct. 2016, pp. 1222–1225.
[50] G. Gutiérrez, J. Canul-Reich, A. O. Zezzatti, L. Margain, and J. Ponce,‘‘Mining: Students comments about teacher performance assessment usingmachine learning algorithms,’’ Int. J. Combinat. Optim. Problems Informat., vol. 9, no. 3, p. 26, 2018.
[51] S. Mihi, B. A. B. Ali, I. E. Bazi, S. Arezki, and N. Laachfoubi, ‘‘Howdigitalization is perceived from Moroccan students: A sentiment analysisstudy,’’ in Proc. 4th Int. Conf. Intell. Comput. Data Sci. (ICDS), Oct. 2020,pp. 1–7.
[52] W. Tian, J. Li, and H. Li, ‘‘A method of feature selection based onWord2Vec in text categorization,’’ in Proc. 37th Chin. Control Conf.(CCC), Jul. 2018, pp. 9452–9455.
[53] H. M. Wallach, ‘‘Topic modeling,’’ in Proc. 23rd Int. Conf. Mach. Learn.(ICML), 2006, pp. 977–984, doi: 10.1145/1143844.1143967.
[54] P. P. Patil, S. Phansalkar, and V. V. Kryssanov, ‘‘Topic modelling for aspectlevel sentiment analysis,’’ in Proc. 2nd Int. Conf. Data Eng. Commun.Technol. Singapore: Springer, Oct. 2018, pp. 221–229, doi: 10.1007/978-981-13-1610-4_23.
[55] S. I. Nikolenko, S. Koltcov, and O. Koltsova, ‘‘Topic modelling for qualitative studies,’’ J. Inf. Sci., vol. 43, no. 1, pp. 88–102, Jul. 2016, doi:10.1177/0165551515617393.
[56] S. A. Curiskis, B. Drake, T. R. Osborn, and P. J. Kennedy, ‘‘An evaluationof document clustering and topic modelling in two online social networks:Twitter and Reddit,’’ Inf. Process. Manage., vol. 57, no. 2, Mar. 2020,Art. no. 102034, doi: 10.1016/j.ipm.2019.04.002.
[57] C. B. Asmussen and C. Møller, ‘‘Smart literature review: A practical topicmodelling approach to exploratory literature review,’’ J. Big Data, vol. 6,no. 1, pp. 1–18, Oct. 2019, doi: 10.1186/s40537-019-0255-7.
[58] B. V. Barde and A. M. Bainwad, ‘‘An overview of topic modeling methodsand tools,’’ in Proc. Int. Conf. Intell. Comput. Control Syst. (ICICCS),Jun. 2017, pp. 745–750.
[59] T. K. Landauer, P. W. Foltz, and D. Laham, ‘‘An introduction to latentsemantic analysis,’’ Discourse Process., vol. 25, nos. 2–3, pp. 259–284,Jan. 1998, doi: 10.1080/01638539809545028.
[60] D. D. Lee and H. S. Seung, ‘‘Learning the parts of objects by non-negativematrix factorization,’’ Nature, vol. 401, no. 6755, pp. 788–791, Oct. 1999,doi: 10.1038/44565.
[61] S. F. Chen, Building Probabilistic Models for Natural Language Cambridge, MA, USA: Harvard Univ., 1996.
[62] T. Hofmann, ‘‘Unsupervised learning by probabilistic latent semanticanalysis,’’ Mach. Learn., vol. 42, no. 1, pp. 177–196, Jan. 2001, doi:10.1023/a:1007617005950.
[63] B. De Finetti, Theory of Probability: A Critical Introductory Treatment,vol. 6. Hoboken, NJ, USA: Wiley, 2017.
[64] A. Hamzah, A. F. Hidayatullah, and A. G. Persada, ‘‘Discovering trends ofmobile learning research using topic modelling approach,’’ Int. J. Interact.Mobile Technol., vol. 14, no. 9, pp. 1–4, Jun. 2020.
[65] D. F. Onah and E. L. Pang, ‘‘MOOC design principles: Topic modellingpyLDAvis visualization &summarisation of learners’ engagement,’’ inProc. 13th Annu. Int. Conf. Educ. New Learn. Technol., 2021, pp. 1–10.VOLUME 10, 2022 56735
[66] T. Shaik et al.: Review of Trends and Challenges in Adopting NLP Methods for Education Feedback Analysis
[67] I. Marçal, R. E. Garcia, D. Eler, and R. C. M. Correia, ‘‘A strategyto enhance computer science teaching material using topic modelling:Towards overcoming the gap between college and workplace skills,’’ inProc. 51st ACM Tech. Symp. Comput. Sci. Educ., Feb. 2020, pp. 366–371.
[68] S. Unankard and W. Nadee, ‘‘Topic detection for online course feedbackusing LDA,’’ in Emerging Technologies for Education. Cham, Switzerland:Springer, 2020, pp. 133–142, doi: 10.1007/978-3-030-38778-5_16.
[69] S. Cunningham-Nelson, M. Baktashmotlagh, and W. Boles, ‘‘Visualizingstudent opinion through text analysis,’’ IEEE Trans. Educ., vol. 62, no. 4,pp. 305–311, Nov. 2019.
[70] J. K. N. Singh, ‘‘Academic resilience among international students: Lived experiences of postgraduate international students in Malaysia,’’ Asia Pacific Educ. Rev., vol. 22, no. 1, pp. 129–138, Nov. 2020, doi: 10.1007/s12564-020-09657-7.
[71] A. Perez-Encinas and J. Rodriguez-Pomeda, ‘‘International students’ perceptions of their needs when going abroad: Services on demand,’’ J. Stud. Int. Educ., vol. 22, no. 1, pp. 20–36, Feb. 2018.

[72] V. Liermann, ‘‘Overview machine learning and deep learning frameworks,’’ in The Digital Journey of Banking and Insurance, vol. 3. Cham, Switzerland: Springer, 2021, pp. 187–224, doi: 10.1007/978-3-030-78821- 6_12.
[73] S. Srinivas and R. Rajendran, ‘‘Topic-based knowledge mining of online student reviews for strategic planning in universities,’’ Comput. Ind. Eng., vol. 128, pp. 974–984, Feb. 2019.
[74] S. Pyasi, S. Gottipati, and V. Shankararaman, ‘‘SUFAT—An analytics tool for gaining insights from student feedback comments,’’ in Proc. IEEE Frontiers Educ. Conf. (FIE), Oct. 2018, pp. 1–9.
[75] S. Loria, Textblob Documentation, Release 0.15, vol. 2, 2018, p. 269.
[76] M. M. Tadesse, H. Lin, B. Xu, and L. Yang, ‘‘Detection of depressionrelated posts in Reddit social media forum,’’ IEEE Access, vol. 7, pp. 44883–44893, 2019.
[77] T. Ruan, Q. Kong, S. K. McBride, A. Sethjiwala, and Q. Lv, ‘‘Crossplatform analysis of public responses to the 2019 Ridgecrest earthquake sequence on Twitter and Reddit,’’ Sci. Rep., vol. 12, no. 1, pp. 1–14, Jan. 2022, doi: 10.1038/s41598-022-05359-9.
[78] L. N. Rani, S. Defit, and L. J. Muhammad, ‘‘Determination of student subjects in higher education using hybrid data mining method with the K-means algorithm and FP growth,’’ Int. J. Artif. Intell. Res., vol. 5, no. 1, pp. 91–101, Dec. 2021, doi: 10.29099/ijair.v5i1.223.
[79] R. K. Dinata, S. Retno, and N. Hasdyna, ‘‘Minimization of the number of iterations in k-medoids clustering with purity algorithm,’’ Revue d’Intell. Artificielle, vol. 35, no. 3, pp. 193–199, Jun. 2021, doi: 10.18280/ria.350302.
[80] T.-C. Wang, B. N. Phan, and T. T. T. Nguyen, ‘‘Evaluating operation performance in higher education: The case of Vietnam public universities,’’ Sustainability, vol. 13, no. 7, p. 4082, Apr. 2021, doi: 10.3390/su13074082.
[81] M. Rahmanian and E. G. Mansoori, ‘‘An unsupervised gene selection method based on multivariate normalized mutual information of genes,’’ ChemometricIntell. Lab. Syst., vol. 222, Mar. 2022, Art. no. 104512, doi: 10.1016/j.chemolab.2022.104512.
[82] Y. Hu and K. Dai, ‘‘Foreign-born Chinese students learning in China: (Re)shaping intercultural identity in higher education institution,’’ Int. J. Intercultural Relations, vol. 80, pp. 89–98, Jan. 2021, doi: 10.1016/j.ijintrel.2020.11.010.
[83] A. D’Ambrosio, S. Amodio, C. Iorio, G. Pandolfo, and R. Siciliano, ‘‘Adjusted concordance index: An extensionl of the adjusted Rand index to fuzzy partitions,’’ J. Classification, vol. 38, no. 1, pp. 112–128, Jun. 2020, doi: 10.1007/s00357-020-09367-0.
[84] S. Vanaja and M. Belwal, ‘‘Aspect-level sentiment analysis on E-commerce data,’’ in Proc. Int. Conf. Inventive Res. Comput. Appl. (ICIRCA), Jul. 2018, pp. 1275–1279.
[85] S. Wassan, X. Chen, T. Shen, M. Waqar, and N. Jhanjhi, ‘‘Amazon product sentiment analysis using machine learning techniques,’’ Revista Argentina de ClínicaPsicológica, vol. 30, no. 1, p. 695, 2021.
[86] S. Mystakidis, A. Christopoulos, and N. Pellas, ‘‘A systematic mapping review of augmented reality applications to support STEM learning in higher education,’’ Educ. Inf. Technol., vol. 27, no. 2, pp. 1883–1927, Aug. 2021, doi: 10.1007/s10639-021-10682-1.
[87] C. A. Palacios, J. A. Reyes-Suárez, L. A. Bearzotti, V. Leiva, and C. Marchant, ‘‘Knowledge discovery for higher education student retention based on data mining: Machine learning algorithms and case study in Chile,’’ Entropy, vol. 23, no. 4, p. 485, Apr. 2021, doi: 10.3390/e23040485.
[88] P. D. Gil, S. da Cruz Martins, S. Moro, and J. M. Costa, ‘‘A data-driven approach to predict first-year students’ academic success in higher education institutions,’’ Educ. Inf. Technol., vol. 26, no. 2, pp. 2165–2190, Oct. 2020, doi: 10.1007/s10639-020-10346-6.
[89] J. G. Perez, P. Bulacan, and E. S. Perez, ‘‘Predicting student program completion using Naïve Bayes classification algorithm,’’ Int. J. Mod. Educ. Comput. Sci., vol. 13, no. 3, pp. 57–67, Jun. 2021, doi: 10.5815/ijmecs.2021.03.05.
[90] Z. Kastrati, A. S. Imran, and A. Kurti, ‘‘Weakly supervised framework for aspect-based sentiment analysis on students’ reviews of MOOCs,’’ IEEE Access, vol. 8, pp. 106799–106810, 2020
[91] S. Gottipati, V. Shankararaman, and J. R. Lin, ‘‘Text analytics approach to extract course improvement suggestions from students’ feedback,’’ Res. Pract. Technol. Enhanced Learn., vol. 13, no. 1, pp. 1–19, Jun. 2018, doi: 10.1186/s41039-018-0073-0.
[92] K. S. Krishnaveni, R. R. Pai, and V. Iyer, ‘‘Faculty rating system based on student feedbacks using sentimental analysis,’’ in Proc. Int. Conf. Adv. Comput.,Commun. Informat. (ICACCI), Sep. 2017, pp. 1648–1653.

[93] W. S. El-Kassas, C. R. Salama, A. A. Rafea, and H. K. Mohamed, ‘‘Automatic text summarization: A comprehensive survey,’’ Expert Syst. Appl., vol. 165, Mar. 2021, Art. no. 113679, doi: 10.1016/j.eswa.2020.113679.
[94] M. Allahyari, S. Pouriyeh, M. Assefi, S. Safaei, E. D. Trippe, J. B. Gutierrez, and K. Kochut, ‘‘Text summarization techniques: A brief survey,’’ 2017, arXiv:1707.02268.
[95] L. Abualigah, M. Q. Bashabsheh, H. Alabool, and M. Shehab, ‘‘Text summarization: A brief review,’’ in Recent Advances in NLP: The Case of Arabic Language. Cham, Switzerland: Springer, Nov. 2019, pp. 1–15, doi: 10.1007/978-3-030-34614-0_1.
[96] B. Mutlu, E. A. Sezer, and M. A. Akcayol, ‘‘Candidate sentence selection for extractive text summarization,’’ Inf. Process. Manage., vol. 57, no. 6, Nov. 2020, Art. no. 102359, doi: 10.1016/j.ipm.2020.102359.
[97] A. Sefid, J. Wu, P. Mitra, and L. Giles, ‘‘Extractive research slide generation using windowed labeling ranking,’’ 2021, arXiv:2106.03246.
[98] M. P. Dhaliwal, R. Kumar, M. Rungta, H. Tiwari, and V. Vala, ‘‘On-device extractive text summarization,’’ in Proc. IEEE 15th Int. Conf. Semantic Comput. (ICSC), Jan. 2021, pp. 347–354.
[99] Y. Wu and B. Hu, ‘‘Learning to extract coherent summary via deep reinforcement learning,’’ in Proc. 32nd AAAI Conf. Artif. Intell., 2018, pp. 1–8.
[100]J. N. Madhuri and R. G. Kumar, ‘‘Extractive text summarization using sentence ranking,’’ in Proc. Int. Conf. Data Sci. Commun. (IconDSC), Mar. 2019, pp. 1–3, doi: 10.1109/IconDSC.2019.8817040.
[101]S. Song, H. Huang, and T. Ruan, ``Abstractive text summarization usingLSTM-CNN based deep learning,'' Multimedia Tools Appl., vol. 78, no. 1, pp. 857875, Jan. 2019.
[102]D. Dessì, F. Osborne, D. R. Recupero, D. Buscaldi, and E. Motta, ``Generating knowledge graphs by employing natural language processing and machine learning techniques within the scholarly domain,'' Future Gener. Comput. Syst., vol. 116, pp. 253264, Mar. 2021, doi: 10.1016/j.future.2020.10.026.
[103]D. Shi, T. Wang, H. Xing, and H. Xu, ``A learning path recommendation model based on a multidimensional knowledge graph framework for E-learning,'' Knowl.-Based Syst., vol. 195, May 2020, Art. no. 105618, doi: 10.1016/j.knosys.2020.105618.
[104]I. A. Kandhro, S. Wasi, K. Kumar, M. Rind, and M. Ameen, ``Sentiment analysis of students' comment using long-short term model,'' Indian J. Sci. Technol., vol. 12, no. 8, pp. 116, 2019.
[105]H. T. Hien, P.-N. Cuong, L. N. H. Nam, H. L. T. K. Nhung, and L. D. Thang, ``Intelligent assistants in higher-education environments: The FIT-EBot, a chatbot for administrative and learning support,'' in Proc. 9th Int. Symp. Inf. Commun. Technol. (SoICT), 2018, pp. 6976.
[106]O. Grljevi¢, Z. Bo²njak, and A. Kova£evi¢, ``Opinion mining in higher education: A corpus-based approach,'' Enterprise Inf. Syst., vol. 16, no. 5, pp. 126, May 2022.
[107]I. Sindhu, S. M. Daudpota, K. Badar, M. Bakhtyar, J. Baber, and M. Nurunnabi, ``Aspect-based opinion mining on student's feedback for faculty teaching performance evaluation,'' IEEE Access, vol. 7, pp. 108729108741, 2019, doi: 10.1109/ACCESS.2019.2928872.
[108]N. N. Y. Vo, Q. T. Vu, N. H. Vu, T. A. Vu, B. D. Mach, and G. Xu, ``Domain-specic NLP system to support learning path and curriculum design at tech universities,'' Comput. Educ., Artif. Intell., vol. 3, Jan. 2022, Art. no. 100042, doi: 10.1016/j.caeai.2021.100042.
[109]J. Tabassum, M. Maddela,W. Xu, and A. Ritter, ``Code and named entity recognition in StackOverow,'' 2020, arXiv:2005.01634.
[110]G. Pashev, S. Gaftandzhieva, and Y. Hopterieva, ``Domain specic automated essay scoring using cloud based NLP API,'' Int. J. Comput. Sci. Mobile Comput., vol. 10, no. 10, pp. 3339, Oct. 2021, doi: 10.47760/ijcsmc. 2021.v10i10.006.
[111]A. Joshi, P. Bhattacharyya, and M. J. Carman, ``Automatic sarcasm detection: A survey,'' ACM Comput. Surv., vol. 50, no. 5, pp. 122, 2017.
[112]S. Jusoh, ``A study on NLP applications and ambiguity problems,'' J. Theor. Appl. Inf. Technol., vol. 96, no. 6, pp. 114, 2018.
[113]A. S. Imran, S. M. Daudpota, Z. Kastrati, and R. Batra, ``Cross-cultural polarity and emotion detection using sentiment analysis and deep learning on COVID-19 related tweets,'' IEEE Access, vol. 8, pp. 181074181090, 2020.
[114]S. Cappallo, S. Svetlichnaya, P. Garrigues, T. Mensink, and C. G. M. Snoek, ``New modality: Emoji challenges in prediction, anticipation, and retrieval,'' IEEE Trans. Multimedia, vol. 21, no. 2, pp. 402415, Feb. 2019.

[115]G. S. Chauhan, P. Agrawal, and Y. K. Meena, ``Aspect-based sentiment analysis of students' feedback to improve teachingprocess," in Information and Communication Technology for Intelligent Systems. Singapore: Springer, Dec. 2018, pp. 259266, doi: 10.1007/978-981-13-1747-7_25.

[116]A. Nazir, Y. Rao, L. Wu, and L. Sun, ``Issues and challenges of aspectbased sentiment analysis: A comprehensive survey," IEEE Trans. Affect. Comput., early access, Jan. 30, 2020, doi: 10.1109/TAFFC.2020.2970399.

[117]F. Thabtah, S. Hammoud, F. Kamalov, and A. Gonsalves, ``Data imbalance in classication: Experimental evaluation," Inf. Sci., vol. 513, pp. 429441, Mar. 2020, doi: 10.1016/j.ins.2019.11.004.

[118]L. Guo, Y. Lei, S. Xing, T. Yan, and N. Li, ``Deep convolutional transfer learning network: A new method for intelligent fault diagnosis of machines with unlabeled data," IEEE Trans. Ind. Electron., vol. 66, no. 9, pp. 73167325, Sep. 2019.

[119]S. Ruder, M. Peters, S. Swayamdipta, and T. Wolf, ``Transfer learning in natural language processing," in Proc. Conf. North Amer. Chapter Assoc. Comput. Linguistics, Tuts., Minneapolis, MN, USA: Assoc. Comput. Linguistics, Jun. 2019, pp. 1518. [Online]. Available: https://aclanthology.org/N19-5004

[120]S. Shaikh, S. M. Daudpota, A. S. Imran, and Z. Kastrati, ``Towards improved classication accuracy on highly imbalanced text dataset using deep neural language models," Appl. Sci., vol. 11, no. 2, p. 869, Jan. 2021, doi: 10.3390/app11020869.

[121]C. Shorten, T. M. Khoshgoftaar, and B. Furht, ``Text data augmentation for deep learning," J. Big Data, vol. 8, no. 1, pp. 134, Jul. 2021, doi: 10.1186/s40537-021-00492-0.

[122]Gartner Identies Four Trends Driving Near-Term Articial Intelligence Innovation. Accessed: Mar. 2022. [Online]. Available: https://www.gartner.com/en/newsroom/press-releases/ 2021-09-07-gartner-identiesfour- trends-driving-near-term-articial-intelligence-innovation

[123]N. Almazmomi, A. Ilmudeen, and A. A. Qaffas, ``The impact of business analytics capability on data-driven culture and exploration: Achieving a competitive advantage," Benchmarking, Int. J., vol. 29, no. 4, pp. 12641283, Aug. 2021, doi: 10.1108/bij-01-2021-0021.

[124]G. Steinbauer, M. Kandlhofer, T. Chklovski, F. Heintz, and S. Koenig, ``A differentiated discussion about AI education K-12," KI-Künstliche Intelligenz, vol. 35, no. 2, pp. 131137, May 2021, doi: 10.1007/s13218- 021-00724-8.

[125]M. J. Timms, ``Letting articial intelligence in education out of the box: Educational cobots and smart classrooms," Int. J. Artif. Intell. Educ., vol. 26, no. 2, pp. 701712, Jan. 2016, doi: 10.1007/s40593-016-0095-y.

[126]L. Chen and D. Gerritsen, ``Building interpretable descriptors for student posture analysis in a physical classroom," in Proc. 22nd Int. Conf. Artif. Intell. Educ. (AIED), 2021, pp. 15.

[127]W. Holmes, M. Bialik, and C. Fadel, Articial Intelligence in Education. Boston, MA, USA: Center Curriculum Redesign, 2019, pp. 135.

[128]R. Egger, ``Software and tools," in Applied Data Science in Tourism. Cham, Switzerland: Springer, 2022, pp. 547588, doi: 10.1007/978-3-030- 88389-8_26.

[129]J. J. Thomas, V. Suresh, M. Anas, S. Sajeev, and K. S. Sunil, ``Programming with natural languages: A survey," in Computer Networks and Inventive Communication Technologies. Singapore: Springer, Sep. 2021, pp. 767779, doi: 10.1007/978-981-16-3728-5_57.

[130]D. Sarkar, Text Analytics With Python: A Practitioner's Guide to Natural Language Processing. Bangalore, India: Springer, 2019.

[131]R. M. Reese and A. Bhatia, Natural Language Processing With Java: Techniques for Building Machine Learning and Neural Network Models for NLP. Birmingham, U.K.: Packt Publishing Ltd, 2018.

[132]M. L. Jockers and R. Thalken, Text Analysis with R. Cham, Switzerland: Springer, 2020.

BIOGRAPHIES

	**Sujankumarreddy Vuyyuru** is an undergraduate student currently in his final year of Bachelor of Technology in Computer Science and Engineering with an honors degree in Artificial Intelligence and Machine Learning (AIML) from Sai Spurthi Institute of Technology.
	**Srinija Vuyyuru** is a graduate student in Bachelor of Technology in Computer Science and Engineering from Sagi Rama Krishnam Raju Engineering College, Affiliated to JNTU, Kakinada, India. Currently working DBS Technology India Pvt.Ltd as Software Engineer - Full Stack, India
	**Gogulamudi Sai Sandeep Reddy** is an undergraduate student currently in his final year of Bachelor of Technology in Computer Science and Engineering with an honors degree in Data Analyst from Presidency University, Bengaluru.
	**Pujitha Vajrala** is an undergraduate student currently in her final year of Bachelor of Technology in Computer Science and Engineering with an honors degree in Artificial Intelligence and Machine Learning (AIML) from Sai Spurthi Institute of Technology.