Faculty 4: Computer Science

Institute for Web Science
and Technologies

Exploring Academic Perspectives: Sentiments and Discourse on ChatGPT Adoption in Higher Education

Master's Thesis

in partial fulfillment of the requirements for
the degree of Master of Science (M.Sc.)
in Informatik

submitted by
Yashwanth Gopal

First supervisor: Prof. Dr. Frank Hopfgartner
Institute for Web Science and Technologies

Second supervisor: Dr.-Ing. Stefania Zourlidou
Institute for Web Science and Technologies

Koblenz, August 2024

Statement

I hereby certify that this thesis has been composed by me and is based on my own work, that I did not use any further resources than specified – in particular no references unmentioned in the reference section – and that I did not submit this thesis to another examination before. The paper submission is identical to the submitted electronic version.

	Yes	No
I agree to have this thesis published in the library.	☑	☐
I agree to have this thesis published on the Web.	☑	☐
The thesis text is available under a Creative Commons License (CC BY-SA 4.0).	☑	☐
The source code is available under a GNU General Public License (GPLv3).	☑	☐
The collected data is available under a Creative Commons License (CC BY-SA 4.0).	☑	☐

Koblenz, 8-8-2024

(Place, Date) (Signature)

Note

- If you would like us to contact you for the graduation ceremony, please provide your personal E-mail address: yashwanth.0968@gmail.com
- If you would like us to send you an invite to join the WeST Alumni and Members group on LinkedIn, please provide your LinkedIn ID : ....

Zusammenfassung

Künstliche Intelligenz (KI) wird in zahlreichen Branchen zunehmend eingesetzt, einschließlich im Bildungsbereich. Anwendungen der künstlichen Intelligenz (KI) werden für Schulen und Universitäten immer wichtiger, sei es für die automatisierte Bewertung, intelligente Bildungssysteme, individuelles Lernen oder die Unterstützung des Personals. ChatGPT, ein KI-basierter Chatbot, bietet kohärente und hilfreiche Antworten basierend auf der Analyse großer Datenmengen. Die Integration von ChatGPT, einem fortschrittlichen Natural Language Processing (NLP)-Tool, das von OpenAI entwickelt wurde, in die Hochschulbildung hat großes Interesse und Diskussionen ausgelöst. Da die Technologie bereits von vielen Studierenden und Lehrkräften adaptiert wurde, untersucht diese Studie die auf Universitätswebsites geäußerten Meinungen zur Integration von ChatGPT in die Bildung durch die Erstellung eines umfassenden Sentiment-Analyse-Rahmens unter Verwendung des Hierarchical Residual RSigELU Attention Network (HR-RAN). Der vorgeschlagene Rahmen adressiert mehrere Herausforderungen der Sentiment-Analyse, wie die Erfassung feinkörniger Sentiment-Nuancen, die Einbeziehung kontextueller Informationen und die Bewältigung komplexer sprachlicher Ausdrücke in universitären Bewertungsdaten.

Die Methodik umfasst mehrere Schritte, darunter die Datensammlung von verschiedenen Bildungswebsites, Blogs und Nachrichtenplattformen. Die Daten werden vorverarbeitet, um Emoticons, URLs und Tags zu behandeln und dann sarkastische Texte mit dem eXtreme Learning Hyperband Network (XLHN) zu erkennen und zu entfernen. Sätze werden basierend auf Ähnlichkeiten gruppiert und Themen werden mithilfe des Non-negative Term Document Matrix Factorization (NTDMF)-Ansatzes modelliert. Es werden Merkmale wie lexiko semantische, lexiko strukturelle und numerische Merkmale extrahiert. Eine Abhängigkeitsanalyse und Koreferenzauflösung werden durchgeführt, um grammatikalische Strukturen zu analysieren und semantische Beziehungen zu verstehen. Word Embeddings verwenden das Word2Vec Modell, um semantische Beziehungen zwischen Wörtern zu erfassen. Der vorverarbeitete Text und die extrahierten Merkmale werden in den HR-RAN-Klassifikator eingegeben, um die Stimmungen als positiv, negativ oder neutral zu kategorisieren.

Die Ergebnisse der Sentiment-Analyse zeigen, dass 74,8 % der Stimmungen gegenüber ChatGPT in der Hochschulbildung neutral sind, 21,5 % positiv und nur 3,7 % negativ. Dies deutet auf eine überwiegende Neutralität unter den Nutzern hin, wobei ein signifikanter Anteil positive Ansichten äußert und ein sehr kleiner Prozentsatz negative Meinungen hat. Darüber hinaus zeigt die Analyse regionale Unterschiede, wobei Kanada die höchste Anzahl von Stimmungen aufweist, überwiegend neutral, gefolgt von Deutschland, dem Vereinigten Königreich und den USA. Die Ergebnisse der Sentiment-Analyse werden anhand verschiedener Metriken wie Genauigkeit, Präzision, Recall, F-Maß und Spezifität bewertet. Die Ergebnisse zei-

gen, dass der vorgeschlagene Rahmen herkömmliche Sentiment-Analyse-Modelle übertrifft. Die HR-RAN-Technik erreichte eine Präzision von 98,98 %, einen Recall von 99,23 %, ein F-Maß von 99,10 %, eine Genauigkeit von 98,88 % und eine Spezifität von 98,31 %. Zusätzlich werden Wortwolken erstellt, um die häufigsten Begriffe innerhalb positiver, neutraler und negativer Stimmungen visuell darzustellen und ein klares und unmittelbares Verständnis der Schwerpunktthemen in den Daten zu vermitteln. Diese Erkenntnisse können Pädagogen, Administratoren und Entwicklern Informationen über die Vorteile und Herausforderungen bei der Integration von ChatGPT in Bildungseinrichtungen bieten, um Verbesserungen in der Bildungspraxis und der Entwicklung von KI-Tools zu leiten.

Abstract

Artificial intelligence (AI) is becoming more widely used in a number of industries, including in the field of education. Applications of artificial intelligence (AI) are becoming crucial for schools and universities, whether for automated evaluation, smart educational systems, individualized learning, or staff support. ChatGPT, an AI-based chatbot, offers coherent and helpful replies based on analyzing large volumes of data. Integrating ChatGPT, a sophisticated Natural Language Processing (NLP) tool developed by OpenAI, into higher education has sparked significant interest and debate. Since the technology is already adapted by many students and teachers, this study delves into analyzing the sentiments expressed on university websites regarding ChatGPT integration into education by creating a comprehensive sentiment analysis framework using Hierarchical Residual RSigELU Attention Network (HR-RAN). The proposed framework addresses several challenges in sentiment analysis, such as capturing fine-grained sentiment nuances, including contextual information, and handling complex language expressions in university review data.

The methodology involves several steps, including data collection from various educational websites, blogs, and news platforms. The data is preprocessed to handle emoticons, URLs, and tags and then, detect and remove sarcastic text using the eXtreme Learning Hyperband Network (XLHN). Sentences are then grouped based on similarity and topics are modeled using the Non-negative Term-Document Matrix Factorization (NTDMF) approach. Features, such as lexico-semantic, lexico-structural, and numerical features are extracted. Dependency parsing and coreference resolution are performed to analyze grammatical structures and understand semantic relationships. Word embedding uses the Word2Vec model to capture semantic relationships between words. The preprocessed text and extracted features are inputted into the HR-RAN classifier to categorize sentiments as positive, negative, or neutral.

The sentiment analysis results indicate that 74.8% of the sentiments towards ChatGPT in higher education are neutral, 21.5% are positive, and only 3.7% are negative. This suggests a predominant neutrality among users, with a significant portion expressing positive views and a very small percentage holding negative opinions. Additionally, the analysis reveals regional variations, with Canada showing the highest number of sentiments, predominantly neutral, followed by Germany, the UK, and the USA. The sentiment analysis results are evaluated based on various metrics, such as accuracy, precision, recall, F-measure, and specificity. Results indicate that the proposed framework outperforms conventional sentiment analysis models. The HR-RAN technique achieved a precision of 98.98%, recall of 99.23%, F-measure of 99.10%, accuracy of 98.88%, and specificity of 98.31%. Additionally, word clouds are generated to visually represent the most common terms within positive, neutral, and negative sentiments, providing a clear and immediate understanding of the key

themes in the data. These findings can inform educators, administrators, and developers about the benefits and challenges of integrating ChatGPT into educational settings, guiding improvements in educational practices and AI tool development.

List of Acronyms

Symbol	**Description**
D	Collected data
D_{clean}	Emoticons, tags, and URLs handled data
N_{clean}	Total number of handled data
T_{in}	Input text data
$D_{handled}$	Handled data
T_{split}	Sentence split text
$D_{collected}$	Collection of data from input review text
$T_{sc\_removed}$	Special character removed text or data
T_{abbrev}	Abbreviations and contractions handled data
T_{case}	Case converted text
T_{spell}	Spell checked text
T_{norm}	Normalized text
$T_{preproc}$	Pre-processed data
$\mathbf{v}$	Vector of each element in the pre-processed data
$\mathbf{I}$	Input layer
$\mathbf{H}$	Hidden layer
$\mathbf{W}$	Weights
$\mathbf{b}$	Bias
$\mathbf{W}_{init}$	Random initialization
F	Number of features in the input review text
$H_{neurons}$	Number of neurons in the hidden layer
ϕ	Activation function
H_{max}	Maximum number of hidden neurons
H_{min}	Minimum number of hidden neurons
B_{comp}	Computation budget
R_{max}	Maximum resources
R_{factor}	Reduction factor
$\mathbf{\Theta}$	Hyper-parameter configurations
N_{Θ}	Total number of hyper-parameter configurations
R_{Θ}	Allocated resources to configurations
$\mathbf{H}_{tuned}$	Hidden layer after hyper-parameter tuning
$\mathbf{W}_{out}$	Output weights
$\mathbf{H}^{\dagger}$	Invertible form of hidden layer's output
$\hat{y}$	Output predictions for sentiment analysis
$N_{no\_sarcasm}$	Total number of texts without sarcastic texts
$T_{no\_sarcasm}$	Sarcastic text removed data
d_{LeHe}	LeHe based distance calculation equation
C_{new}	New cluster after merging nearest distance
L	The linkage methods
S_{group}	Grouped sentences

N_{group}	Total number of grouped sentences
$\mathbf{TDM}$	Term-document matrix
$\mathbf{TTM}$	Topic-term matrix
$\mathbf{DTM}$	Document-topic matrix
$\mathbf{TTM}_{upd}$	Updated topic-term matrix
$\mathbf{DTM}_{upd}$	Updated document-topic matrix
$\mathbf{M}^T$	Transpose matrix
$\odot$	Element-wise multiplication operator
$\mathbf{Topics}$	Modeled topics
$C_{relevant}$	Relevant content
$N_{relevant}$	Total number of relevant contents
$\mathbf{F}_{ext}$	Extracted features
N_{ext}	Total number of extracted features
D_{dep}	Analyzed data for dependency parsing
$\mathbf{Tree}$	Dependency tree
T_{dep}	Data after dependency parsing
E_{ling}	Linguistic expressions
N_{ling}	Total number of linguistic expressions
$E_{cluster}$	Clustered linguistic expressions
$N_{clusters}$	Total number of clusters
b	Binary variable
J	Objective function
$\mathbf{Score}$	Score value
T_{coref}	Text after coreference resolution
N_{coref}	Total number of texts after coreference resolution
$T_{context}$	Input context texts
$\mathbf{W}_{indices}$	Word indices
W_{size}	Context window size
$\mathbf{v}_{context}$	Single context vector
y_{target}	Output of the target words
$\mathbf{W}^T$	Transpose of the weight matrix
T_{target}	Input target words
P_{output}	Final output probability
T_{embed}	Word embedded data or text
N_{embed}	Total number of word embedded texts
$\mathbf{F}_{extract}$	Feature extraction
T_{data}	Text data
$\mathbf{v}_{rep}$	Vector representation
$\mathbf{h}_{fwd}$	Forward direction of the hidden state in GRU
$\mathbf{h}_{bwd}$	Backward direction of the hidden state in GRU
$\mathbf{h}_{concat}$	Concatenation of both directions
$\mathbf{h}_{hid}$	Hidden representation of concatenation of both directions
c^w	Trainable parameters

$\mathbf{w}_{norm}$	Normalized weights
K^w	Word-level context vector
Φ^{word}	Word vector
$\mathbf{c}_{sent}$	Sentence-level context vector
$\mathbf{s}_{vector}$	Sentence vector
$\hat{y}_{class}$	Classification of sentiments
λ	Hyper-parameter
σ	Sigmoid function
μ_{Tripo}	Tripo membership function
$\mathbf{F}_{set}$	Fuzzy set
k	Scaling constant
μ_{upper}, μ_{lower}	Upper limit peak value and lower limit of membership function
μ_{match}	Matching degree of the fuzzy set
$f_{strength}$	Firing strength
$R_{combined}$	Combined rule
$C_{actions}$	Overall control actions
$\mathbf{C}_{crisp}$	Crisp value for analyzing sentiments
$\mathbf{c}_{centroid}$	Centroid of the fuzzy set
$\mathbf{S}_{degree}$	Analyzed sentiments in degree
$N_{polarity}$	Maximum polarity word count
P_{pos}	Percentage of positive sentiment analysis
P_{neg}	Percentage of negative sentiment analysis
P_{neut}	Percentage of neutral sentiment analysis
TP	True positives
TN	True negatives
FP	False positives
FN	False negatives

Contents

List of Figures

List of Algorithms

List of Tables

1 Introduction

Artificial Intelligence (AI) has the power to transform the way we learn and teach, thus making it more effective, engaging, and individualized. Artificial intelligence (AI) in education makes use of technology like natural language processing and machine learning to improve the educational process [37]. It involves applying algorithms to data analysis, pattern recognition, and prediction making, allowing teachers to tailor instruction to the needs of individual students. There are numerous potential benefits to using AI in education [3]. Personalized learning, which enables students to learn at their own pace and in a way that best suits their learning preferences, is one of the most significant advantages of AI in education. This may improve student performance [72]. Chatbots, automated grading and assessment, and intelligent tutoring systems can boost productivity, and saveteachers' time, and deliver more consistent and accurate feedback [38]. One such organization which has worked towards development of such an intelligence system is OpenAI. The organization made a substantial contribution to the advancement of natural language processing and artificial intelligence. In order to guarantee that Artificial General Intelligence (AGI) serves humanity as a whole, OpenAI was established in December 2015. Early initiatives from OpenAI, such OpenAI Universe and Gym, were centered on reinforcement learning and developing environments in which AI agents might interact and learn. Consequently, the groundwork for OpenAI's future efforts to create increasingly complex AI models was established by these projects. The creation of the Generative Pre-trained Transformer (GPT) marked the first significant advancement in natural language processing [60]. n language generation tasks, GPT has represented a major advancements [56].Better data analysis can be achieved by AI, enabling educators to make data-driven decisions [4]. Additionally, it can raise student engagement by offering dynamic and captivating educational opportunities [33]. AI can help make education more inclusive and accessible, allowing learners from all backgrounds to access high-quality education [15, 80].

Within this landscape of Artificial Intelligence in education, ChatGPT has got significant attention for its unique capabilities. Launched in November 2022, ChatGPT swiftly gained prominence, gaining over 100 million users within two months [85]. Unlike conventional question-answering models, ChatGPT's applications extend to AI art prompts, coding assistance, and essay generation [85]. Its versatility has led to widespread adoption in various educational scenarios, offering assistance in assignments, explanations, and personalized learning experiences [55]. The decision to focus this research on ChatGPT arises from its rapid adoption and transformative potential. Comparative analyses have highlighted ChatGPT's superiority over other conversational AI options [6]. For instance, the research paper [68] compared ChatGPT with LaMDA and BlenderBot, which demonstrated ChatGPT's ability to handle a larger range of topics and create different opinions within particular situations. When compared to Google's feature snippet, ChatGPT demonstrated an exceptional capacity to generate understandable and easily interpretable replies [34]. Comparative analyses conducted by conducted by [**?**, **?**, 73], showed that ChatGPT outperformed to other models like Bing Chat and Bard in a variety of tasks. The success of ChatGPT, surpassing major social media platforms in user acquisition, underscores its unique position in reshaping educational practices [22].

The integration of ChatGPT into educational systems has also sparked diverse opinions and debates. On one hand, its ability to facilitate learning and assist educators is widely acknowledged [55]. On the other hand, concerns regarding its impact on academic integrity, dependence, and the potential for misuse raise essential questions [85]. The University of British Columbia, as highlighted on their news page [96], acknowledges the potential of generative AI tools like ChatGPT in academics. Yet, the university emphasizes addressing the limitations of such tools, including potential issues with generating incorrect responses and references, as well as privacy concerns. Therefore, due to the privacy concerns the university recommends against using ChatGPT in its courses as highlighted in [95]. A similar approach is observed at the University of Texas Austin (UT Austin), where the Information Security Office [91] delves into acceptable usage guidelines and potential risks associated with integrating AI tools like ChatGPT. The issue is that UT Austin and OpenAI do not currently have an agreement regarding privacy and security, and that all content submitted into or generated by ChatGPT is accessible to ChatGPT, OpenAI, and their employees. Hence, the university emphasize on restricting the use of sensitive information due to privacy and security concerns. These perspectives underscore the complex terrain universities navigate when dealing with AI tools. These diverse opinions and restrictions necessitates a comprehensive analysis of user sentiments towards ChatGPT, particularly within the context of higher education.

Sentiment analysis is a technique in NLP which enables extracting and analyzing opinions, emotions, and attitudes from textual data [10]. By applying sentiment analysis to university review data, this study aims to evaluate the perceptions and experiences of students and educators using ChatGPT. The goal is to understand the positive and negative sentiments, providing insights that can inform the development and implementation of ChatGPT in educational settings.

1.1 Role of AI in Education

AI is incorporated into learner-centered education systems as part of Education 4.0 [81]. To detail the history of Artificial Intelligence (AI) in education, we must first explore its evolution from the 1960's when early intelligent tutoring systems like

PLATO and the "Automatic Grader" were developed to provide personalized instruction and automate grading tasks [5]. These early systems opened the way for advances in the 1970's with the development of systems like TICCIT (Time-shared, InteractiveComputer-Controlled Instructional Television) that utilized interactive multimedia to deliver educational content. The emergence of AI-based education from the 1980's saw a shift towards computer-based instruction enhanced by the advent of micro-computers and the World-Wide-Web (WWW) in the 1990's, leading to the development of more intelligent and adaptive learning services [5]. The 21st century has witnessed significant breakthroughs in AI applications in education, particularly with the advancements in Language Models (LLMs) and generative AI models. The role of AI in education can be categorized into several key areas:

- **Personalized Learning** AI makes individualized learning possible by adjusting course materials to meet the unique requirements of every student. By analyzing students' learning styles, progress, and difficulties, AI systems can customize instructional materials and learning paths. Applications like Knewton and Cerego use machine learning algorithms to make real-time recommendations for students, adjusting the content to optimize their learning outcomes [36].

- **Adaptive Learning** Adaptive learning technologies utilize AI to provide a customized learning experience by dynamically adjusting the difficulty and nature of tasks based on student performance. Platforms such as ALEKS and BYJU'S are examples of adaptive learning systems that assess students' abilities and provide personalized learning paths to improve learning outcomes [36] These systems collect data on student interactions and use it to optimize their learning journeys, enhancing engagement and effectiveness.

- **Support for Teachers** Artificial Intelligence (AI) assists educators by automating tasks like grading and offering insights into student performance. Tools like automated essay scoring systems and intelligent grading platforms help reduce the workload on teachers, allowing them to focus more on instruction and student engagement [36]. AI can also help in creating curriculum content, planning lessons, and identifying areas where students may need additional support.

- **Enhanced Accessibility** AI enhances accessibility in education by providing tools and resources that support diverse learning needs [64]. For instance, AI tools such as speech recognition and text-to-speech (TTS) technologies significantly aid students with disabilities. These tools can transform spoken language into written text, facilitating communication for those with speech impairments, and convert text into speech, aiding students with visual impairments or reading difficulties. These technologies help students engage with educational content more effectively and inclusively [64].

1.2 Overview of ChatGPT and its Capabilities

OpenAI developed ChatGPT, or Chat Generative Pre-trained Transformer, an advanced conversational AI model [77]. Leveraging the sophisticated transformer architecture, it can generate human-like responses to various inputs, thus making it highly effective in diverse applications. It was developed by OpenAI, which was founded in 2015 by Elon Musk, Sam Altman, Greg Brockman, Ilya Sutskever, and Wojciech Zaremba. While OpenAI has developed various programs, ChatGPT was launched on November 30, 2022. ChatGPT has evolved from earlier models like GPT-2 and GPT-3, thereby improving accuracy and contextual understanding with each iteration [75]. It is optimized for dialogue using Reinforcement Learning with Human Feedback (RLHF), which employs preference comparisons and human examples to guide the model towards desired behavior. ChatGPT has gained significant attention for its ability to engage in coherent and contextually relevant conversations, highlighting its potential in various fields, especially education and research.

Figure 1.1 shows the preview of homescreen interface of ChatGPT, where queries can be posted to get replies.

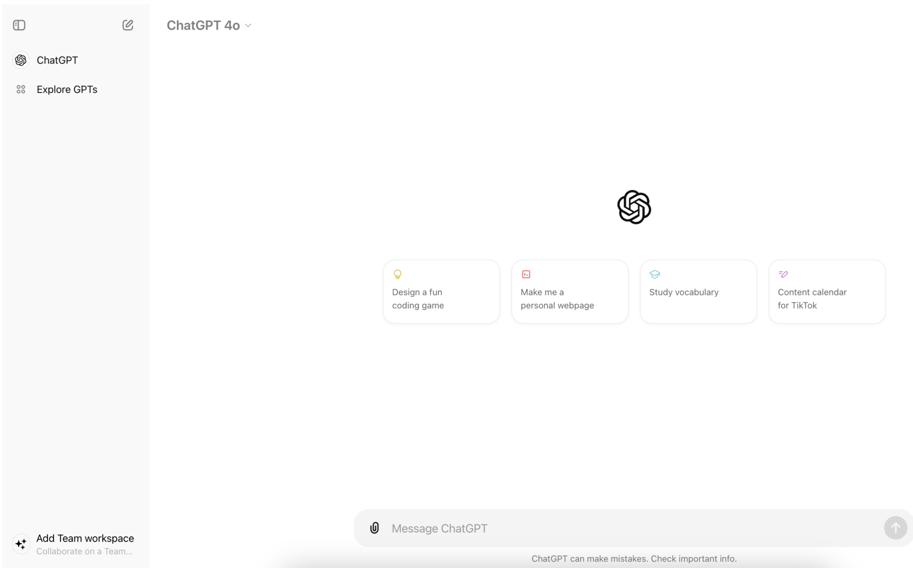


Figure 1.1: ChatGPT interface [69]

1.2.1 Technological Framework

ChatGPT integrates Reinforcement Learning from Human Feedback (RLHF) methods with the generative pre-trained transformer to align large language models with human intent [5]. A variation of the Transformer architecture, a well-liked deep-learning model for problems involving natural language processing, is used in the construction of ChatGPT [11]. The Transformer design comprises an encoder-decoder structure integrating self-attention mechanisms. Using a large dataset of paired input-output samples, the model is trained in a supervised manner. The model gains the ability to predict the most likely token in a sequence based on prior context during training. This is achieved through the minimization of a loss function, usually cross-entropy loss, which measures the disparity between the predicted and actual target distributions [11].

The training dataset used for training ChatGPT included a vast amount of text data from books, websites, and other sources [99]. ChatGPT's training dataset consists of billions of tokens of text used to create training examples for the model. The pretraining objective is next-word prediction, involving examples where the model predicts the next word in a sentence based on the input sequence. This training approach helps ChatGPT generate coherent and contextually relevant responses [28]. The learning architecture of ChatGPT 3.5 involves fine-tuning the pre-trained GPT models using reinforcement learning techniques with human feedback. This is learning happens in 3 main steps [28] as shown in Figure 1.2, Supervised Fine-Tuning (SFT) Model, Reward Model Training and Proximal Policy Optimization (PPO) RL Algorithm.

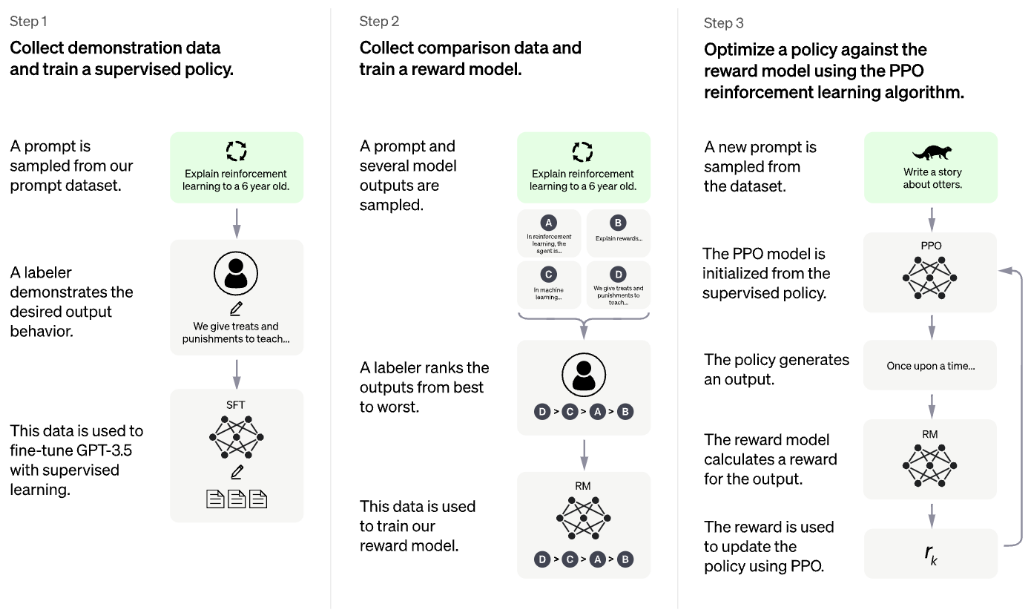


Figure 1.2: Reinforcement learning from human feedback [99]

1.2.2 ChatGPT Relevance to Education

In the research paper [6], the researcher explored the ways in which educators and students have improved teaching and learning strategies by using ChatGPT in a variety of educational contexts. It highlights the advantages, challenges, and impacts of integrating ChatGPT into education, shedding light on the potential benefits. Figure 1.3 illustrates the main findings from the reviewed studies regarding the utilization of ChatGPT in education. It categorizes these findings into two domains: "ChatGPT in Learning" and "ChatGPT in Teaching".The findings illustrate that ChatGPT enhances learning by serving as a virtual assistant, providing on-demand answers, personalized learning support, and aiding in writing and language proficiency. It enhances students' competencies and thereby helping to achieve overall academic success. For teachers, ChatGPT increases productivity and efficiency by quickly generating educational content and offering new teaching methodologies. [6].

The integration of ChatGPT in education holds significant promise for personalized learning experiences and enhanced academic efficiency. By assisting in designing tailored learning routes for students, serving as online research assistants, overcoming language barriers, automating grading, and offering AI-powered tutoring support, ChatGPT can optimize time management, foster collaboration, and improve student outcomes [79]. Additionally, its potential to imitate laboratory experiments, assist in curriculum design, provide data analysis support, and ensure ethical considerations make it a versatile tool with the capacity to revolutionize teaching, learning, and research methodologies in the realm of education [79].

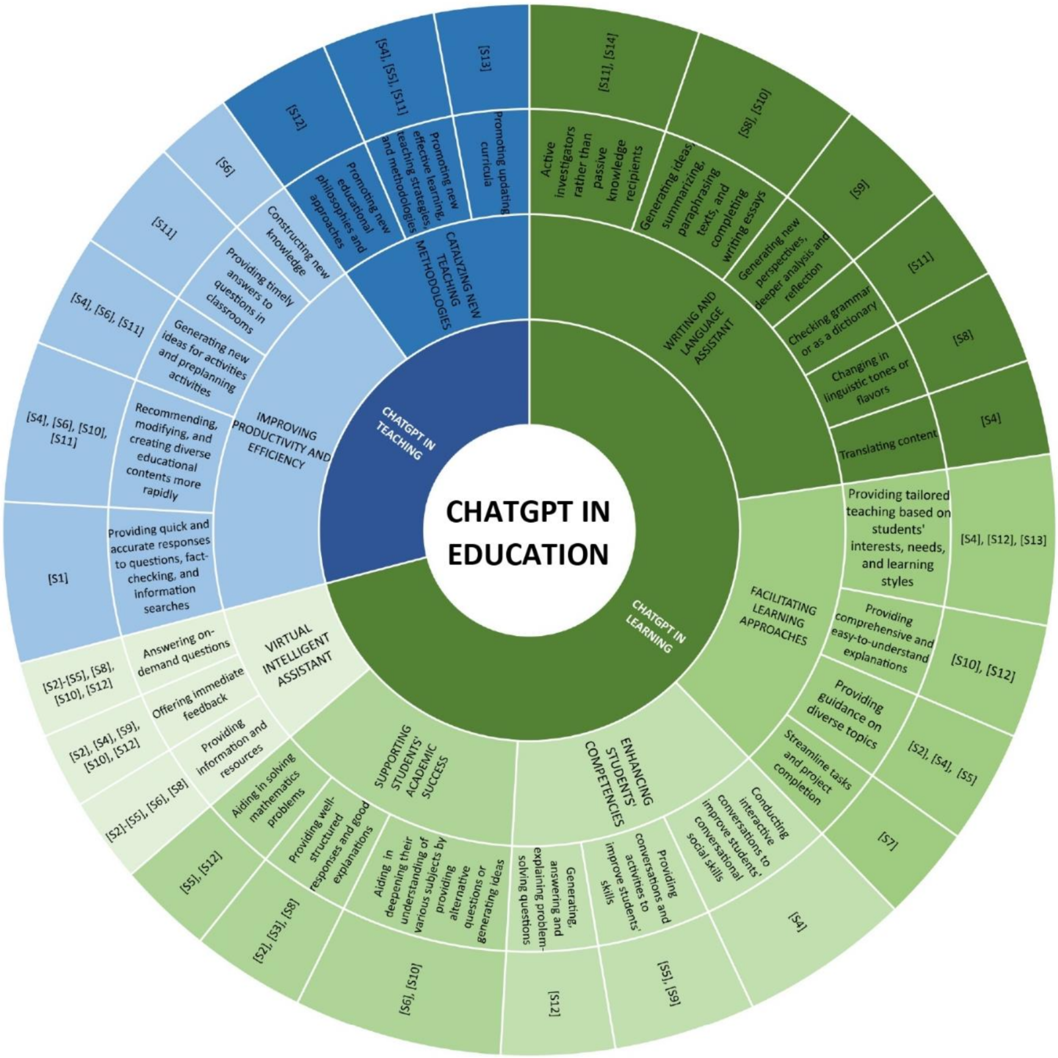


Figure 1.3: An illustration of Utilization of ChatGPT in education [6]

The paper [79] explains briefly about ChatGPT's utility which extends to diverse academic domains within higher education, such as basic sciences, engineering, health sciences, agriculture, management, and social sciences. In basic sciences, it serves as an interactive tutor, offering clarification on complex topics and promoting self-paced learning. In engineering, it provides concept explanations, aids in problem-solving, and supports students in understanding challenging engineering principles, enhancing their learning experiences. Moreover, in social sciences education, ChatGPT aids in concept clarification, guides research methodology, promotes critical thinking exercises, and engages with students on ethical and societal issues. In health sciences education, ChatGPT aids in understanding medical terminology, simulating clinical scenarios and patient interactions, therefore facilitating case-based learning. Within management sciences education, ChatGPT assists in analyzing case studies, simulating business scenarios, offering entrepreneurship guidance, aiding in leadership development, promoting creative thinking and can help in strategic decision making.

The relevance of ChatGPT in education is highlighted through its potential to enhance learning performance and efficiency for students [14]. ChatGPT provides benefits such as breaking spatial and temporal constraints, offering comprehensive information, assisting with academic tasks, simplifying complex concepts, and providing interactive question-and-answer interactions similar to human conversations. These applications of ChatGPT in education aim to improve student learning experiences, increase engagement, and provide personalized feedback to support adaptive learning, ultimately contributing to enhanced academic performance [14].

1.2.3 Impact of ChatGPT in the Field of Education

As illustrated in Figure 1.4, the influence of ChatGPT in education is extensive, leveraging its world-leading Natural Language Processing (NLP) capabilities. Its word vector model's performance allows for excellent context conversation ability, which is well-suited to natural language processing (NLP) tasks like named entity recognition, text classification, and part-of-speech tagging.

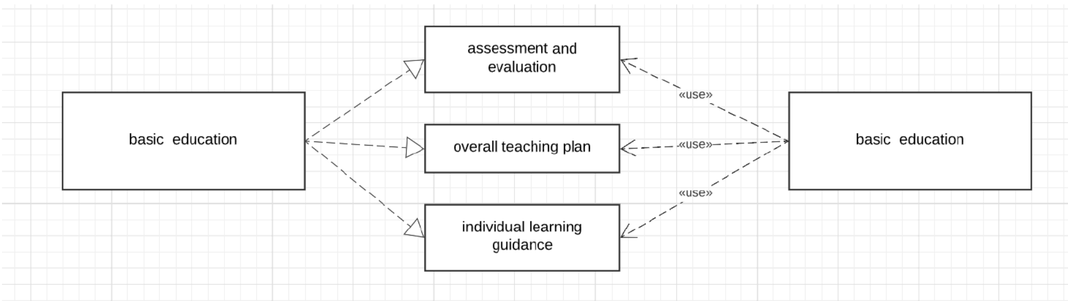


Figure 1.4: Impact of ChatGPT on education [90]

- **Influence on Education** ChatGPT can assist in knowledge cultivation in primary education, which is fundamental for primary and secondary schools. Knowledge serves as a personal core literacy and a resource essential for modern society, driving innovation, and entrepreneurship. While traditional education focuses on simple knowledge impartation, ChatGPT can significantly enhance this process [97]. In higher education, the emphasis shifts to teaching students how to use ChatGPT effectively to improve learning and research efficiency and enhance the quality of education. Educators are encouraged to move beyond traditional teaching methods, fostering students' creative thinking beyond AI capabilities.

- **Educational Assessment and Evaluation** ChatGPT's capabilities also influence educational assessment and evaluation methods. Schools might need to innovate and update their evaluation techniques to focus on creative and speculative thinking, thereby moving away from traditional exam-based knowledge assessments [97].
- **Teaching Plans and Individual Learning Guidance** ChatGPT has a profound impact on overall teaching plans and individual learning guidance. Its powerful information search and organization abilities allow it to answer students' questions quickly and facilitate natural language-based information retrieval [97]. This capability can revolutionize internet-based information search methods. Additionally, ChatGPT can serve as an AI teaching assistant, thus aiding teachers and students in developing course plans, generating question lists, and assisting with academic papers, programs, and tests.

1.2.4 Strengths of ChatGPT

One of the significant strengths of ChatGPT is its versatility as it can handle a wide range of topics and conversational styles. This adaptability makes it suitable for various educational contexts, thereby offering dynamic and engaging learning experiences. Moreover, ChatGPT enhances accessibility by providing educational support to students who may lack access to traditional educational resources, thereby bridging gaps in learning opportunities. Its ability to generate contextually relevant and coherent responses further underscores its potential as a valuable educational tool [24].

1.2.5 Limitations and Challenges

Despite its strengths, ChatGPT also faces several limitations and challenges. Ethical concerns, such as data privacy, potential bias in responses, and the risk of misuse are significant issues that must be addressed [59]. There is also the concern that students might become overly reliant on AI tools, thus impacting their critical thinking and problem-solving skills [59]. Ensuring the accuracy of responses and preventing misinterpretation of nuanced questions remain ongoing challenges in deploying ChatGPT in educational settings.Moreover, introducing AI technologies such as ChatGPT into the classroom necessitates substantial investments in computing power and infrastructure. Unfortunately, not all educational institutions have equal access to these resources, which could widen the existing gap between well-funded and under-resourced schools.

1.2.6 Future Prospects

The prospects of ChatGPT in education are promising with continuous advancements in AI expected to enhance its capabilities further [16]. Integration with other technologies, such as Virtual Reality (VR) and Augmented Reality (AR), could create immersive learning experiences, thus transforming the educational landscape. Ongoing efforts to improve ChatGPT through user feedback and continuous learning will ensure that it remains a valuable and effective tool for education. These advancements will likely address current limitations, making ChatGPT even more integral to the educational ecosystem.

1.3 Perspectives and Sentiments on ChatGPT

Sentiment analysis, also referred as opinion mining, is a field within Natural Language Processing (NLP) that focuses on determining the sentiment conveyed in text [30]. This involves classifying text as positive, negative, or neutral and can be extended to more nuanced categories, such as very positive or negative. Sentiment analysis aims to understand textual data's underlying emotions, opinions, and attitudes. It leverages various techniques from NLP, machine learning, and computational linguistics to process and analyze large volumes of text data. Sentiment analysis typically operates at different levels:

1. Document Level: Evaluates the overall sentiment of an entire document or review.

2. Sentence Level: Analyzes sentiment expressed in individual sentences within a document.

3. Entity Level: Combines entity recognition with sentiment analysis to determine sentiments about specific entities like products, services, or individuals.

4. Aspect-Based Level: A fine-grained analysis that identifies sentiment related to specific aspects or features within a document, such as customer service, product quality, or delivery time.

The process of sentiment analysis generally involves the following:

1. Text Preprocessing: Cleaning and preparing text for analysis by removing noise, such as punctuation, stop words, and stemming words to their root forms.

2. Feature Extraction: Identifying relevant features or attributes in the text that contribute to sentiment.

3. Sentiment Classification: Using machine learning algorithms or lexicon-based approaches to classify text into sentiment categories.

Several commentators argue that academics must adapt technologies like ChatGPT into higher education [88]. They emphasize that while there are concerns regarding academic integrity, these AI tools have the potential to enhance student learning.

Since there is very little academic literature related to this argument, the authors in [88] focus on exploring key themes in news articles, especially about ChatGPT in a higher education context, authors perform content analysis of 100 media articles. The authors estimated the positive and negative valence of publications on the subject using Nvivo's Sentiment Analysis tool. Table 1.1 presents a summary of the content analysis results. The themes included academic integrity concerns, ways to encourage students to avoid using ChatGPT, discussions on policy decisions regarding ChatGPT, considerations on incorporating ChatGPT into teaching practices, and the voices represented in the articles, including university staff, students, and ChatGPT itself.

Table 1.1: Content analysis of news articles [88]

Code	**Definition**	**Article count**
Academic Integrity		
Catching	Discussion of tools that can be used for detecting the use of ChatGPT	87
Concern	General concerns about cheating/contract cheating/unfair admissions	51
Educate	Addressing concerns by educating students or referring to a Code of Conduct	54
Example	Specific stories and examples about failing or penalising students for using ChatGPT	25
Subject	Some disciplines or types of assignments that might be more at risk than others	20
Avoidance		
Adaptation	Plans to restructure assignments or courses to minimise use of ChatGPT, including examples of specific assignments or tasks that ChatGPT cannot do	87
Errors	General criticism of errors made by ChatGPT or mentioning false referencing (outside of a specific context, such as learning or adapting assignments)	62
Learning	Specific concerns about negative impacts on learning outcomes	50
Policy		
Undecided	University is considering their policy on ChatGPT	41
No Use	University has banned or discouraged ChatGPT	22

Code	Definition	Article count
Allowed	University has encouraged or not banned ChatGPT	18
Embrace		
Teaching	Ideas for how ChatGPT can be usefully incorporated into teaching (e.g., using as a class activity, producing teaching resources)	58
Too hard	It is too hard to ban, for practical or other reasons	45
Workplace	Justifying the use of ChatGPT in universities by linking to real-world/workplace practice	25
Equity	ChatGPT can be used to improve/enhance/address concerns with equity or help struggling students. This does not have to be a specific equity group (e.g., reducing student stress or anxiety)	24
Voice		
Academic	Story, quote or example of a university academic or other university staff member	86
Student	Story, quote or example of a university student	79
ChatGPT	It is acknowledged that ChatGPT wrote part of the article, or ChatGPT responses are quoted as examples in text, or a ChatGPT spokesperson	30

The findings of the analysis in [88] shed light on significant academic integrity concerns within universities, prompting discussions on how to address the potential risks associated with AI tools like ChatGPT. The results showed a mixed response from universities and the general public, with a primary focus on academic integrity concerns and innovative assessment design opportunities [88]. Specifically, the analysis highlighted that some universities were considering banning ChatGPT, while others were in the process of reviewing and updating their policies regarding the use of AI tools in academia [88]. The authors also identified discussions on how universities might adapt their teaching practices to integrate AI tools ethically and effectively, emphasizing the importance of preparing students for a digital world where AI technologies are becoming increasingly prevalent.

1.4 Identification of Key Issues and Challenges

While the usage of ChatGPT in higher education presents several benefits, it also comes with a set of disadvantages, risks, and challenges. These include the potential for misuse leading to academic misconduct such as fabrication and spread of false information [14]. Additionally, concerns exist regarding the impact of ChatGPT on traditional educational norms and practices, the possibility of students becoming overly dependent on AI technologies, and the ethical considerations surrounding the use of AI in educational settings. Striking a balance between leveraging the advantages of ChatGPT while mitigating these risks is crucial for successful implementation in higher education [14].

Generative AI models like ChatGPT present both opportunities and challenges in educational settings. While they offer valuable assistance, they also exhibit limitations that need to be addressed. One significant drawback is the lack of human interaction, which can be detrimental for students who thrive on personal connections with educators [12]. Additionally, these models may have a limited understanding of concepts, as they rely on statistical patterns rather than true comprehension, affecting their ability to tailor explanations to individual learning needs. Moreover, biases present in the training data can lead to inaccuracies in assessing or grading student work, potentially perpetuating disparities [12]. Creativity is another area where generative models fall short, as their responses are constrained by existing data patterns, impacting the diversity and originality of output. Furthermore, the effectiveness of these models is heavily reliant on the quality and relevance of the data they are trained on, which can pose challenges if the data is insufficient or irrelevant. Contextual understanding is also a concern, as generative models may struggle to grasp situational nuances, resulting in inappropriate or irrelevant responses. Personalisation of instruction is another challenge, as these models may offer general assistance but struggle to tailor instruction to individual student needs, limiting personalized learning experiences. Lastly, there are privacy and data security concerns while using generative AI tools in educational settings, highlighting the importance of making sure the data is handled confidentiality [12].

ChatGPT's utilization in generating text and explanations brings forth various concerns that necessitate careful consideration in educational and informational contexts which is explained in detail in [84]. Firstly, the model's reliance on training data introduces biases into its responses, encompassing gender, racial, political, and data incompleteness biases, thereby compromising the accuracy and fairness of generated content. Additionally, the potential for misinformation dissemination emerges, as ChatGPT may inadvertently provide incorrect citations or sources, eroding the credibility and reliability of the information disseminated. Moreover, distinguishing between texts generated by ChatGPT and those authored by humans poses a challenge, impeding efforts to evaluate the authenticity and trustworthiness of generated content, consequently hindering misinformation detection and correction. Furthermore, ChatGPT's database may lack the currency and inclusion of pertinent scientific sources, leading to informational gaps and diminished relevance. Lastly, the ambiguity surrounding the authorship and citation of ChatGPT-generated text raises concerns regarding intellectual property rights and academic integrity, necessitating clarity in citation guidelines to ensure proper sourcing and crediting practices. These considerations underscore the importance of critical evaluation when employing ChatGPT in various domains [84].

1.5 Objective of the Study

This study aims to evaluate opinions on integrating ChatGPT in higher education by developing a sentiment analysis framework. This framework aims to capture detailed sentiments, incorporate context, and handle complex language expressions in the collected data. The understanding gained will help improve educational practices and guide developers in enhancing ChatGPT to serve educational needs better.

- Employ webscraping techniques to extract discourse information from official university websites discussing about ChatGPT. The resulting dataset would capture academic discussions concerning ChatGPT in education, addressing a significant gap in current research.
- To develop and implement the proposed HR-RAN classifier to classify sentiments as positive, negative, or neutral.
- To utilize dependency parsing to analyze grammatical structures and understand semantic relationships between words, thus ensuring effective incorporation of contextual information in sentiment analysis.
- To integrate lexico-semantic features, lexico-structural features, and numerical features into the sentiment analysis process to improve the accuracy and depth of sentiment detection.
- To develop the eXtreme Learning Hyperband Network (XLHN) for effectively detecting and interpreting sarcasm in the collected data.
- To develop the topic modelling logic using Non-negative based Term-document matrix Factorization (NTDMF) for effective topic modelling.

By achieving these objectives, the research aims to provide a comprehensive sentiment analysis framework that shows attitudes of universities regarding ChatGPT integration in higher education, thereby aiding educators and developers in making informed decisions.

1.6 Significance of the Study

The study's significance is to understanding sentiments concerning the integration of ChatGPT in higher education.

- **Enhancement of Educational Practices** By identifying positive and negative attitudes expressed in the blogs, articles, news etc., from students, educators or official university policies, the study can inform improvements in teaching methods and learning content, thus leading to a more effective learning environment.
- **Informed Decision-Making** Educational administrators and policymakers can benefit from the insights generated through this study. The detailed analysis of sentiments provides valuable feedback on the impact of ChatGPT, enabling informed decision-making regarding its integration into the educational system.
- **Improvement of ChatGPT's Educational Applications** For developers and researchers working on ChatGPT, this study offers crucial feedback that can guide further development and enhancement of the technology. By understanding the specific areas, where users express positive or negative sentiments, developers can focus on improving those aspects, making ChatGPT a more effective tool for education.
- **Addressing Sentiment Analysis Challenges** The study contributes to Natural Language Processing (NLP) and sentiment analysis by addressing significant challenges, such as capturing fine-grained sentiment nuances, incorporating contextual information, and handling complex language expressions. The proposed framework using HR-RAN can serve as a model for future research in sentiment analysis across various domains.
- **Contribution to Educational Research** This study contributes to educational research by comprehensively analyzing sentiments toward using advanced AI tools like ChatGPT in education. The findings can be used as a basis for further research into the impacts of AI on teaching and learning processes, thereby contributing to the broader academic discourse on educational technology.
- **Guidance for Future AI Integration** The study's outcomes can guide the integration of AI technologies in educational settings beyond ChatGPT. By highlighting the benefits and challenges of using AI for education, the research provides a roadmap for future AI applications, ensuring that they are implemented to maximize positive outcomes for students and educators.

In summary, this study is significant for its potential to enhance educational practices, inform decision-making, improve AI tools for education, address sentiment analysis challenges, contribute to educational research, and guide future AI integration. These contributions highlight the importance of utilizing sentiment analysis to optimize the use of AI in higher education.

1.7 Scope and Delimitations

Scope

- Focus on Higher Education: The study specifically evaluates the adoption of ChatGPT in higher education, analyzing its impact on teaching and learning processes.
- Sentiment Analysis Framework: Development of a sentiment analysis framework using HR-RAN to assess user opinions and feedback on ChatGPT.
- Data Sources: Utilizes data collected from official university websites, blogs, news articles, etc., to gather diverse opinions on ChatGPT adoption.
- Evaluation Metrics: Measures the performance of the sentiment analysis model through accuracy, precision, recall, and other relevant metrics to ensure effectiveness.
- Technological Integration: It integrates advanced NLP techniques, such as dependency parsing and coreference resolution, to enhance sentiment analysis accuracy.

Delimitations

- The data collected is only upto a sample size chosen from each university.
- The study relies on publicly available data from online platforms and does not include direct surveys or user interviews.

1.7.1 Definition of Terms

- **Web scraping**: Web Scraping is a technology that allows us to extract structured data from text such as HTML. When data is not provided in a machine-readable format, like JSON or XML, web scraping is quite helpful [48].
- **Sentiment analysis**: Sentiment analysis is a subfield of text analysis that may be used to extract and assess people's opinions on a given topic, highlighting its benefits and drawbacks by grouping user opinions into three categories: positive, negative, and neutral [94].
- **Topic modeling**:In machine learning and natural language processing, topic modeling is a technique used to find abstract subjects within a collection of texts. This unsupervised learning technique aims to automatically recognize the themes or subjects discussed in a collection of documents [27].

- **Dependency Parsing**: An NLP technique to for analyzing a sentence's grammatical structure and identify related words and the nature of their relationships. It helps to understand the semantic relationships and contextual dependencies in sentences [42].
- **Coreference Resolution**: A process in NLP that determines which words or phrases in a text refer to the same entity. It improves context understanding by linking pronouns to corresponding entities [98].
- **Kaggle** A platform for data science competitions and datasets. It provides datasets for training and validating machine learning models, thus facilitating data sharing and analysis for research.

1.8 Thesis Structure

Chapter 1: Introduction

This chapter introduces AI in education, ChatGPT, its significance and challenges in education, and the relevance of sentiment analysis in understanding its integration in higher education. It outlines the study's motivation, research objectives, questions, and the significance of the study. The chapter also defines key terms and sets the Scope and delimitations, providing a comprehensive foundation for the research.

Chapter 2: Literature Survey

This chapter reviews existing literature to provide a comprehensive overview of current state of knowledge regarding sentiments and attitudes of ChatGPT integration in education institutions. It also covers previous studies, technological advancements, techniques and identified challenges in sentiment analysis. The research gap analysis highlights areas lacking in current research, thereby setting the stage for the proposed study to address these gaps and contribute to the academic discourse. This chapter also identifies the problem statement and research questions.

Chapter 3: Theoretical Background

This chapter explains technical details such as classifiers, libraries, evaluation metrics, and other technical details implemented in this research.

Chapter 4: Proposed Methodology

The methodology chapter details the proposed sentiment analysis framework using HR-RAN. It describes the data collection process, preprocessing techniques, model development, and feature extraction methods. The classification and analysis steps and the evaluation metrics used to assess the model's performance are outlined. This chapter provides a step-by-step research approach.

Chapter 5: Results and Discussion

This chapter presents the sentiment analysis results, thus evaluating the model's performance using various metrics. It compares the findings of ChatGPT's adoption in higher education with those of previous studies.

Chapter 6: Conclusion and Future Scope

The final chapter summarizes the research findings. It discusses the practical implications for educators, administrators, and developers and provides recommendations for improving ChatGPT's integration into higher education. The chapter concludes with suggestions for future research directions, building on the study's findings to further explore the potential of AI in education.

2 Literature Review

The objective of this literature review is as follows. Firstly, it aims to provide a comprehensive overview of current state of knowledge regarding sentiments and attitudes of using ChatGPT for education. This involves reviewing existing research papers, identifying key themes, identifying sentiments expressed, and summarizing findings to offer insights into the broader discourse surrounding its integration. Additionally, this review aims critically to analyze the methodologies employed in previous studies, evaluating their strengths, limitations, and implications for future research. This serves as a guiding framework for this thesis research implementation regarding sentiment analysis. This literature review will not only provide knowledge to the ongoing discourse surrounding ChatGPT adoption in higher education but also contribute to the broader discourse on AI adoption in higher education. Additionally, this review identifies the existing research gap, thereby identifying the problem statement and research questions.

2.1 Overview of Sentiments and Opinions Regarding ChatGPT in Education

The achievement of student-centered learning will be enhanced by the use of AI-Chatbots in the education sector. The research [82] shows that students are using chatbots and that there have been significant benefits. Using a stratified random sampling technique, 47 students were selected, completing questionnaires on demographic information, Chatbot usage, communication preferences, potential use cases, and barriers. Data analysis, conducted using SPSS (Statistical Package for the Social Sciences), involved descriptive statistics to outline sample characteristics and inferential statistics to explore relationships between Chatbot adoption and respondents' demographics. This methodology aimed to uncover factors influencing Chatbot adoption in Indian higher education. the study revealed that the majority of students indicated they would use Chatbots to get help with educational issues and that they were less likely to use other forms of communication if they were chatting with Chatbots. This implies that students perceive AI-Chatbots as a valuable tool for improving their learning experience [82]. The following were the most likely anticipated application cases for chatbots in educational institutions: tutoring, learning feedback, resolving problems, and providing quick answers. However, concerns were raised regarding the potential barriers to using Chatbots, such as limited intelligence, privacy issues, and the risk of receiving incorrect advice.

The paper [85] explains that there is a research gap with respect to how ChatGPT is perceived and utilized by university students. Understanding the student's perception is important because it helps in analyzing how to efficiently integrate ChatGPT into the education system. The authors bridge this research gap by applying the "Unified Theory of Acceptance and Use of Technology 2" (UTAUT2) model to study university students' adoption and usage of ChatGPT. The study examined data from 503 Polish public university students and extended the Personal Innovativeness model. The model was tested using the PLS-SEM approach, which revealed significant impacts on behavioral intention and use behavior. The findings showed that Hedonic Motivation, Performance Expectancy, and Habit had the greatest effects on Behavioral Intention. Habit and Facilitating Conditions were the next most important factors influencing Use Behavior after Behavioral Intention. A total of 54.7% of the variance in Use Behavior and 72.8% of the variance in Behavioral Intention were explained by the study's model.According to [85], students shown a high degree of acceptability and readiness to utilize ChatGPT in their academic studies. Their perspectives on ChatGPT usage in education were influenced by their habits, performance expectations, and behavioral intentions.

2.2 Perceptions of Teachers and Students on ChatGPT Integration to Education

2.2.1 Perceptions of teachers

A systematic review conducted by [57] shows the perceptions of teachers towards ChatGPT in teaching and learning contexts. Employing a SWOT analysis approach, the study explored the strengths, weaknesses, opportunities, and threats associated with ChatGPT integration. Their research revealed that 78.6% of surveyed teachers expressed positive attitudes towards ChatGPT, citing its potential to revolutionize education. Despite concerns about potential decreases in critical thinking skills, teachers acknowledged the tool's ability to provide diverse learning materials and foster deeper understanding.

The authors in [22] focused on investigating the perceptions of scholars and PhD students regarding the implications of ChatGPT, an AI-powered language model, in the context of education. Through thematic content analysis, the study identified nine key themes including the changing role of educators, impact on assessment and evaluation, and personalized learning. It also highlights that AI, such as ChatGPT, has the potential to revolutionize traditional learning methods, emphasizing skills and competencies while redefining the roles of educators. The participants expressed optimism for the future of AI in education, acknowledging the transformative impact on the learning process while also recognizing challenges related to assessment, digital literacy, and ethical considerations [22]. Overall, the findings suggest that integrating AI in education, particularly ChatGPT, offers opportunities for enhancing learning experiences and outcomes for students and universities, although with careful consideration of potential risks and ethical implications. The study emphasizes the need for continuous dialogue, collaboration among stakeholders, and the development of best practices to ensure responsible and equitable use of AI technologies in education.

The study [66] focused on investigating university's English teachers' perceptions of ChatGPT utilization in language teaching and assessment. Employing qualitative methods, the study aimed to assess teachers' knowledge, perceived usefulness, challenges, and concerns regarding ChatGPT integration. Results indicated that while 60% of teachers recognized the potential of ChatGPT as a valuable support tool, concerns regarding students' overdependence and reliability were common. The study involved 30 university English teachers with varying levels of experience. The results and findings include:

- Teachers' Knowledge of ChatGPT: The study revealed that the participating teachers had varying levels of understanding about ChatGPT. They often confused it with other applications and did not have a comprehensive grasp of its functionalities.

- Usefulness of ChatGPT: While many teachers saw the potential of ChatGPT as a valuable support tool in education, there were concerns about proper guidance and training with regard to using the tool effectively.

- Challenges with ChatGPT: Teachers expressed worries about potential issues such as students cheating, overdependence on ChatGPT, and doubts about the reliability of the information provided by the tool. There were concerns about the changing role of teachers when students become familiar with ChatGPT.

Another study [39] utilized the Technology Acceptance Model (TAM) to explore university faculty members' attitudes towards ChatGPT technology in educational settings. Through interviews with 20 faculty members, the study revealed a predominant negative sentiment towards ChatGPT integration. Concerns about cheating, disruption of traditional learning environments, and lack of value addition to the learning experience were highlighted. However, some teachers acknowledged potential benefits such as increased student engagement.

The investigation on teachers' perceptions of ChatGPT as a supporting tool for teaching and learning through a mixed-methods design was made in [21]. The study involved in-service teachers from Gr. 6–12 in Dubai and Abu Dhabi, UAE. Results indicated that 85% of teachers perceived ChatGPT positively, emphasizing its impact on student learning outcomes, engagement, and motivation. Despite challenges in lesson planning and assessment, teachers recognized ChatGPT's ability to enhance learning experiences and provide personalized instruction.

2.2.2 Perceptions of students

The integration of ChatGPT into education in research has varied responses from students, as revealed by research [13]. Initially, students demonstrate limited awareness of cutting-edge AI technologies like ChatGPT. However, they display openness to incorporating AI to augment productivity and creativity. Analyzing students' readiness to embrace AI, the Technology Acceptance Model (TAM) examines students' preparedness to embrace AI, emphasizes perceived utility and ease of use as critical elements impacting acceptance intentions. For instance, in a survey involving 230 third-year interior architecture students in China, findings indicated a considerable percentage(62%) of students inclined towards integrating AI into their workflow to enhance productivity and creativity [13]. Despite this receptiveness, concerns about AI's potential impact on job opportunities and career prospects are evident among students. Notably, the study underscores the importance of educational institutions cultivating emerging technology competence among students to navigate the design industry's transformations driven by AI. Such insights can inform the design of AI-related curricula and strategies to align with future job market needs, ultimately enhancing students' preparedness for AI-driven industry shifts.

Similarly, the research [85] revealed varied responses among students. The study employs a modified version of the Unified Theory of Acceptance and Use of Technology (UTAUT2) model for identifying the students' perceptions of ChatGPT usage in education. It includes components such as hedonic motivation, performance Expectancy, Effort Expectancy, Social Influence, Habit, Personal Innovativeness, and Behavioral Intention [85]. The research findings demonstrate that students are comfortable adopting new technologies like ChatGPT, and their frequency of use aids to developing habitual behavior. Specifically, "Performance expectancy" is identified as a significant predictor of students' behavioral intention to use ChatGPT [85]. Furthermore, "Habit" is highlighted as a key factor influencing both behavioral intention and actual use behavior. The study reveals that students are more inclined to adopt functional technologies like ChatGPT when they have high levels of performance expectancy. Additionally, the positive link between "Hedonic motivation" and behavioral intention implies that students view AI chat as enjoyable.

Another study [65] delves into the students' responses to the integration of ChatGPT into education, offering insights into their perceptions, challenges, and suggestions. Methodologically, the research engaged 200 Vietnamese university students with prior experience using ChatGPT for academic purposes, utilizing random sampling techniques. Further, 30 students were selected for semi-structured interviews to delve deeper into the benefits, challenges, and potential solutions associated with ChatGPT in higher education. A tailored questionnaire, featuring Likert scale responses, assessed students' perceptions, while interviews provided qualitative insights. Students exhibited a positive attitude towards ChatGPT, citing its ease of use, utility as a search engine, and support for multiple languages. They valued it for saving time, offering diverse information, and aiding in learning and retention. Nonetheless, students recognized barriers such as assessing source reliability, accurate citation, and handling complex concepts. Despite these challenges, they proposed solutions including source verification and ethical usage guidelines.

The survey [65] revealed a mean score of 3.58, indicating agreement on ChatGPT's benefits, with specific scores highlighting its time-saving nature and knowledge diversity. Conversely, concerns about unreliable information and citation difficulties received a mean score of 3.64, indicating awareness of challenges. The qualitative data from interviews further enriched the understanding of students' perspectives.

2.3 Sentiment Analysis: Reviewing Opinions and Sentiments on ChatGPT Discourse

The author in [52] analyzes worries and concerns expressed on social media regarding ChatGPT adoption in education. Using the academic version of "Twitter Search API", relevant tweets were collected from December 1, 2022, to March 31, 2023. It resulted in 247,484 tweets, with 84,828 being original tweets. The research framework included various analytical techniques, including sentiment analysis using the RoBERTa architecture to classify sentiment in tweets, BERTopic modeling to cluster negative tweets into distinct topics, and social network analysis to identify propagating concerns. By leveraging NLP tools and methodologies, the study aimed to analyze discourse on Twitter related to ChatGPT in education as represented in Figure 2.1.

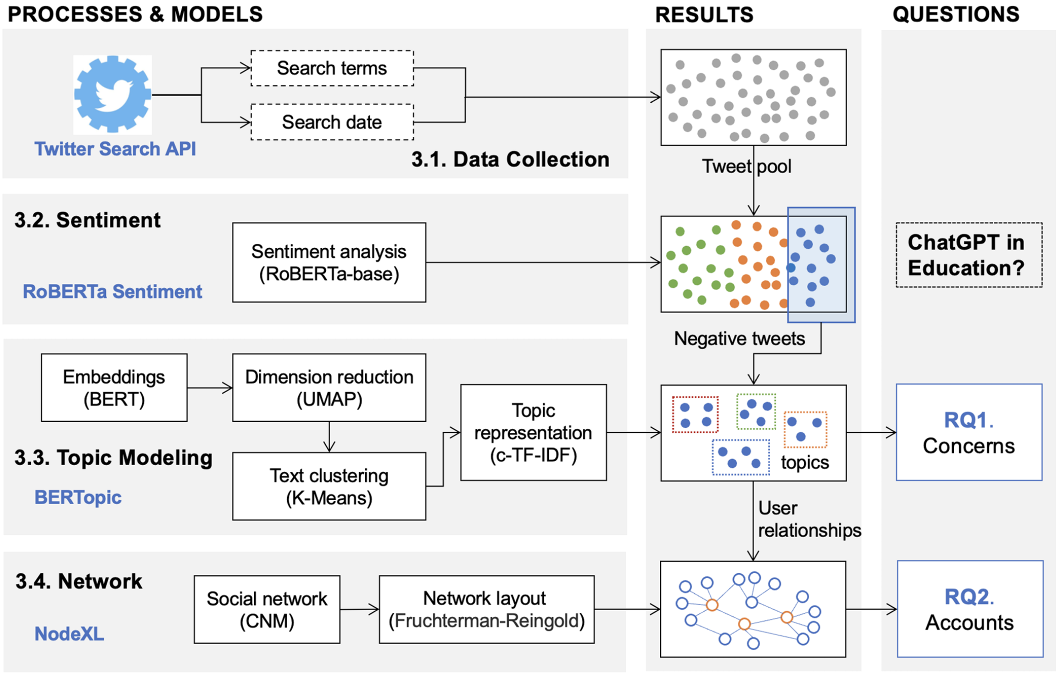


Figure 2.1: An illustration of the research framework used for analyzing Twitter data regarding ChatGPT discourses in education [52]

The sentiment analysis exposed 16,011 were negative, 42,495 were neutral, and 26,322 were positive with respect to ChatGPT in education. The orange line constantly positioned above the blue line in Figure 2.2 implies that twitter users typically expressed a positive attitude towards the deployment of ChatGPT in education. Events, such as the release of the GPT-4 model and ChatGPT surpassing 100 million users, triggered positive sentiment across Twitter during the study period.

Moreover, the topic modeling using BERTopic in the study clustered negative sentiment tweets into distinct topics, allowing for the identification of concerns expressed by Twitter users regarding the use of ChatGPT in education. Key concerns were identified through the analysis of 200 topics from 16,011 original negative-sentiment tweets. These topics fell into categories such as workforce challenges, policy and social concerns, impact on learning outcomes and skill development, limitation of capabilities, and academic integrity [52].

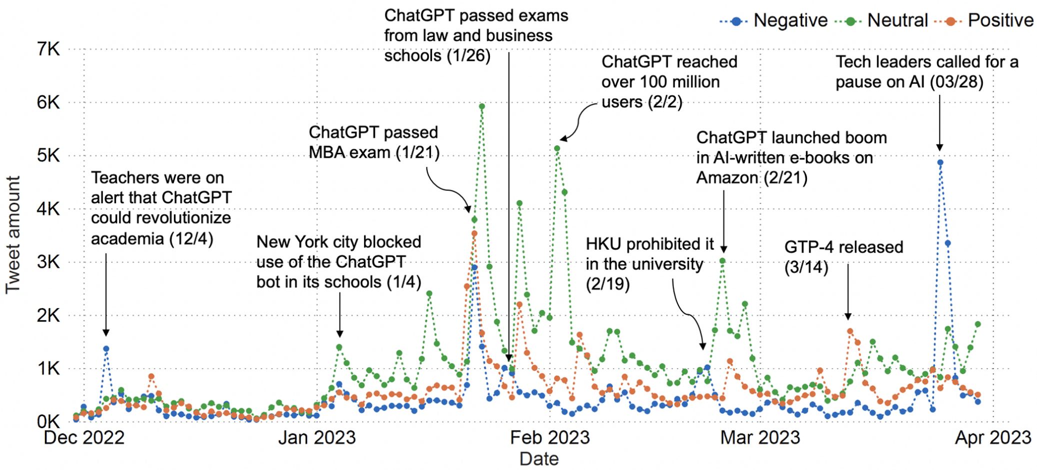


Figure 2.2: Twitter sentiment trend and significant events [52]

Another study [93] intended to uncover common feelings, subjects, and viewpoints that are expressed towards ChatGPT in the education field based on the data collected from Twitter. Data collection for investigating public sentiment on ChatGPT's educational use involved sourcing data from Twitter due to its real-time stream, vast data volume, and accessibility. Spanning from February 1, 2023, to February 12, 2023, tweets were gathered using specific keywords pertinent to ChatGPT and education. These keywords included phrases like 'ChatGPT AND education,' 'Teaching AND ChatGPT,' 'ChatGPT AND Students,' 'ChatGPT AND Exams,' and 'ChatGPT AND Learning.' Only English tweets were considered, resulting in a dataset of 11,830 unique tweets post-duplicate removal. Data preprocessing steps were implemented to ensure data quality and uniformity. Removal of retweets and duplicates streamlined the dataset to 11,830 tweets, setting the stage for subsequent analysis. Textual standardization was achieved by converting all text to lowercase, while unwanted noise like URLs, hashtags, emojis, and numbers were eliminated. Tokenization further refined the dataset by breaking down sentences into individual words (tokens), facilitating analysis. Additionally, stemming techniques were applied to ensure consistency by converting words into their root forms.

Figure 2.3: Word cloud based on top frequent positive opinion words [93]

Sentiment analysis was performed by annotating each tweet with its sentiment polarity, utilizing the TextBlob library's sentiment analysis capabilities. The sentiment polarity ranged from -1 to 1, categorizing tweets as positive (>0), negative (<0), or neutral (=0). Among the 11,830 tweets analyzed, the sentiment breakdown revealed 6,179 tweets expressing positive sentiments, 1,688 expressing negative sentiments, and 3,963 categorized as neutral. These findings provide insights into public sentiment regarding ChatGPT's role in education and serve as a basis for further analysis and interpretation. By using TF-IDF methodology, the most frequent used words in the tweets were analyzed. Word-cloud was created as per Figure 2.3 and Figure 2.4 which represents the Top frequent positive words such as free, creative, available, intelligent, useful etc., and negative opinion words such as bad, critical, wrong, worried, complex, false etc.

Figure 2.4: Word cloud based on top frequent negative opinion words [93]

The paper [93] highlights that most of the analyzed tweets expressed positive sentiments towards the use of ChatGPT in education. This positive sentiment indicates the potential for ChatGPT to play a significant role in education, with growing support from both educators and students. The conclusion suggests that ChatGPT has the capacity to transform the education sector, making learning more engaging and accessible to a wider range of learners. Additionally, the paper [93] compares the performance of different classifiers for tweet sentiment analysis, with the Support Vector Machine (SVM) classifier achieving the highest accuracy of 81.4%.

The researchers [48] utilized a dataset from Kaggle, comprising tweets with the ChatGPT hashtag (ChatGPT), which was continuously updated starting from December 5, 2022, until the last update on March 17, 2023. This dataset contained a substantial volume of tweets, totaling 274,581 entries by the final update. Various data cleaning techniques were applied to enhance the dataset's quality, including the removal of duplicates, near-duplicates, tweets from bots, and filtering out irrelevant content such as marketing impressions. These meticulous steps ensured that the dataset used for analysis was robust and representative of genuine user interactions.

For sentiment analysis, the researchers [48] preprocessed the data and evaluated three sentiment analysis models "VADER", "Twitter-roBERTa", and "XLM-T". Following a comparative analysis of their performance on a manually labeled subset of 1,000 tweets, the XLM-T model was selected for sentiment analysis on the full dataset, because it achieved a sensitivity metric of 67.6% compared to 50.9% and 65.9% for VADER and Twitter-roBERTa models respectively. The result of sentiment analysis were 35.28% of tweets being positive, 45.79% being neutral and 8.92% of negative attitude towards common ChatGPT usage. Additionally, the researchers employed the BERTopic model to extract topics from the tweets dataset, identifying major discussion topics related to ChatGPT on Twitter as shown in Figure 2.5. The sentiment scores calculated for each topic revealed varying sentiments among different topics, with most displaying neutral to positive sentiment, while the 'Job Replacement' topic leaned slightly towards negativity. Overall, the study indicated a generally positive public attitude towards ChatGPT on Twitter, emphasizing the importance of considering social context in technology adoption studies.

Topic	Count	Example
Artificial Intelligence	10,437	AI could make more work for us, instead of simplifying our lives.
Search Engines	7,118	I find myself Googling less and ChatGPT-ing more and more.
Education	5,565	Teachers are already using ChatGPT in Australian classrooms...
Writing	5,352	ChatGPT This is an edge for writers
Question Answering	2,893	Someone ask ChatGPT who the NBA MVP is.
Coding	2,799	wow, it can help me write code, great!!!
Job Replacement	2,761	Are Corporate Americas workers afraid of ChatGPT?
Google's Bard	2,572	AI Battle: ChatGPT Vs Google's Bard. Which one is better?
Hype	1,955	ChatGPT is a game changer. It'll be interesting to see how it evolves.
ChatGPT Plus	1,664	Assuming you're to paid before using ChatGPT will you ?? Me:I'll
Poem	1,659	Been using ChatGPT to write poems. Will speak in prose the entire day.
Music	1,601	I told ChatGPT to write me a song with Taylor Swift's lyrics and it did
Twitter	1,593	Can you imagine chatGPT having a twitter account and tweeting with us?
Healthcare	1,389	Don't start challenging your doctor with ChatGPT just yet!
Language Models	1,374	Don't mistake a language model (like chatGPT) for a verification engine.

Figure 2.5: Top 15 Topics with word count and examples [48]

The study [76] examined public sentiment regarding chatgpt's influence on education. The methodology implemented in this document involved the utilization of web mining to gather data for the study. A sample of 2003 internet articles was collected from varied sources including news websites, educational technology blogs, scientific forums, and popular science magazines. These articles were selected based on their relevance to the influence of ChatGPT on education. Additionally, text analysis techniques such as tokenization, word frequency analysis, text classification, sentiment analysis, topic modeling, named entity identification, co-occurrence analysis, and clustering were applied to extract, categorize, and analyze information from the gathered data. Figure 2.6 illustrates the word classification based on Principal Component Analysis (PCA). The PCA model categorizes the data into three significant clusters:

- Education technology and AI - This category reflects the impact of AI technology like ChatGPT on various educational stakeholders and contemporary pedagogical practices. Terms such as "education", "learning", "tool", and "technology" emphasize the growing integration of ChatGPT in educational ecosystems.

- Writing skills development - This cluster focuses on the particular effects of ChatGPT on writing-related activities, like "writing" and "essay". It demonstrates how ChatGPT supports language, structure, and content development while helping students create essays and other written projects.

- The technical and data-driven aspects of using ChatGPT in the classroom - This category explores ChatGPT's effects on education from a technological and linguistic perspective. Terms like "data," "model," "language," and "text" emphasize the AI-driven mechanisms of ChatGPT, its role in text processing, and aiding in language comprehension within educational contexts.

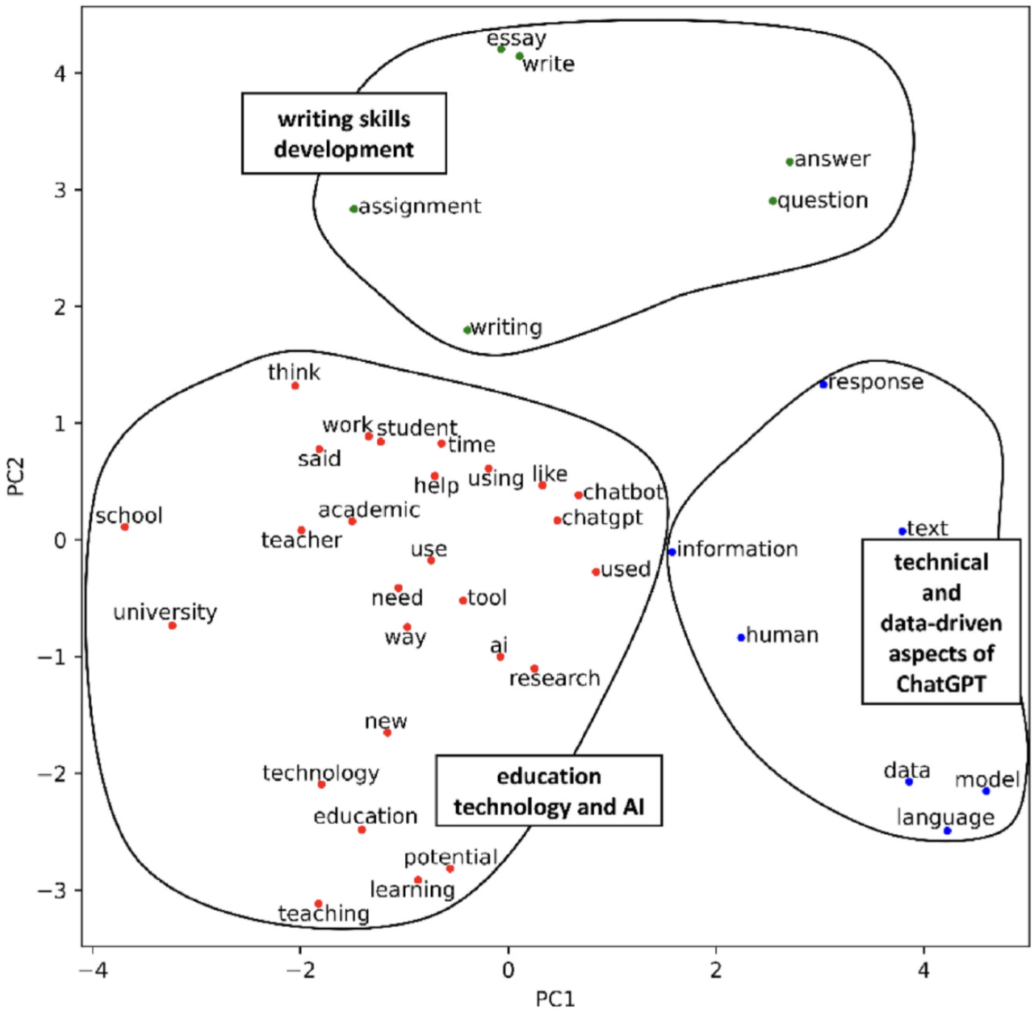


Figure 2.6: Word classification by Principal Component Analysis [76]

The findings of the study revealed significant insights into ChatGPT's influence on education. The research not only underscores the transformative potential of ChatGPT for both students and educators but also lays the groundwork for future investigations into its diverse pedagogical applications. Moreover, the study confronts the challenges and opportunities associated with ChatGPT's integration in educational settings, aiming to provide policymakers, educators, and institutions with insights to make informed decisions about its implementation [76].

Another study [92] emphasizes the need to address concerns regarding the use of chatbots, particularly ChatGPT, in education. It explains many evidences such as the swift bans imposed by educational networks in major cities like New York City and Los Angeles Unified schools due to fears of potential cheating in assignments. This highlights the necessity for a thorough investigation into the risks associated with ChatGPT's utilization in educational settings to ensure its safe deployment. In response to this important need for research, the study seeks to answer the fundamental research question: What are the specific concerns surrounding the integration of chatbots, particularly ChatGPT, in education? To answer the research question, the authors used a three-stage instrumental case study approach

- Social Network Analysis of Tweets: This method implemented includes a cross-sectional examination of tweets using Social Network examination (SNA). From December 23, 2022, to January 6, 2023, 2330 tweets from 1530 Twitter users were gathered and evaluated. The analysis intended to analyze public discourse regarding the use of ChatGPT in education. sentiment analysis of these tweets found that positive attitudes were roughly twice as frequent as negative sentiments. Positive sentiments showed a frequency of 5%, while negative sentiments were at 2.5%. The majority of sentiments, 92.5%, were non-categorized, suggesting indecision or neutrality towards ChatGPT in education.

- The Content Analysis of Interviews: it involved interviewing 19 participants who had experience using ChatGPT in education. These participants, including educators, developers, students, and AI freelancers. One of the findings from the Content Analysis of Interviews was that while some participants recognized ChatGPT's potential to enhance educational experiences, others expressed concerns about potential drawbacks. Specifically, a few participants highlighted that the misuse of ChatGPT by learners could potentially hinder their innovative capabilities and critical thinking skills.

- Investigation of user experiences: three experienced educators engaged with ChatGPT for a week to test various teaching and learning scenarios. Daily meetings were held throughout the week to discuss and summarize the obtained results.

The study concludes that while ChatGPT holds great potential to revolutionize education by providing innovative tools for both teaching and learning, it must be approached with caution. Key concerns include the risk of cheating, the accuracy and fairness of provided content, ethical issues such as the potential for reduced critical thinking, and privacy risks.

2.4 Sentiment analysis methodologies for evaluating ChatGPT in education

The Table 2.1 shows comprehensive overview of the sentiment analysis methods and their limitations.

Table 2.1: Comprehensive review of sentiment analysis methods and their limitations

Paper	Objective	Technique	Result	Limitation
[23]	Determinants of intention to use ChatGPT for educational purposes	Unified Theory of Acceptance and Use of Technology 2 (UTAUT2)	Consistency of 0.909	Due to the focus on general factors like missing trust and privacy concerns, AI adoption was limited by this model.
[83]	Sentiment analysis of student feedback using multi-head attention fusion model of word and context embedding	Ensemble-Long Short Term Memory (LSTM)	Specificity of 98.6%	This framework's performance was affected owing to the lack of capturing the nuances of multi-class classification tasks.
[86]	Public perception of ChatGPT and transfer learning for tweets sentiment analysis	GloVeLSTM	Accuracy of 81.1%	Owing to the need for high computation, this model took more time for execution.
[70]	Stability analysis of ChatGPT-based sentiment	Large Language Model (LLM)	Success rate of 0.0645	This framework had a limited sample size for analysis in AI Quality assurance testing, thus affecting the reliability of the system.
[67]	Sentiment analysis of student engagement with lecture recording	Microsoft Azure Cognitive Services text analytics	Prediction rate of 0.98	This methodology struggled to control the response rate owing to the missing data.

The paper [19] demonstrated the sentiment analysis for evaluating the qualitative responses from students. This model started by gathering the data from a qualitative feedback text after a semester-based course session at the University. Then, the text pre-processing stage was performed to find and remove the missing values and unidentified terms, respectively. Four classifiers were used for sentiment analysis: Naïve Bayes (NB), Support Vector Machine (SVM), J48 Decision Tree (DT), and Random Forest (RF), in which SVM outperformed other models with 63.79% accuracy. However, this model was not generalized for unseen data owing to the overfitting issues.

[104] depicted the framework for exploring the public response to ChatGPT with sentiment analysis and knowledge mapping. Initially, the Latent Dirichlet Allocation (LDA) topic modeling was applied to analyze the web tweets and comments from China regarding ChatGPT. The analysis using LDA revealed the main themes, such as technological impact, social impact, and educational development. Finally, knowledge mapping was employed to analyze the publication time and research hotspots that analyzed negative sentiment of 52.2%. Due to the reliance on publicly available data from social platforms, this model did not fully represent the diverse societal perspectives.

[71] depicted the methodology for exploring the use of artificial intelligence in undergraduate exams, especially focusing on the ability of Graduate Teaching Assistants (GTAs) to detect AI-generated assessments and the effectiveness of ChatGPT. The main findings were AI-generated assessments generally received higher marks than student submissions, often achieving A and A+ grades. To perform sentiment analysis of the feedback data, various ML models were used: Support Vector Machine (SVM), J48 Decision Tree (J48 DT), Naive Bayes (NB), Random Forest (RF). Among these, SVM demonstrated the highest accuracy in predicting sentiment with a 10-fold cross-validation accuracy of 63.79%. Yet, the generalization of this framework was affected owing to the lack of available sample size.

According to [74], sentiment analysis with a gender component and the diffusion of innovation theory were used to examine how ChatGPT was adopted by university students. Initially, five characteristics were investigated that shaped students' behavioral intentions toward ChatGPT: relative advantage, compatibility, ease of use, observability, and trialability. Then, the adoption and societal implications were examined. Finally, the mixed-method framework was employed for examining technology adoption in higher educational settings. Owing to the focus on a specific demographic, the study's findings are limited by the small sample size which may not be representative of broader educational contexts.

A discourse analysis of worries and concerns on social media based on ChatGPT in education was demonstrated by [53]. To determine the main issues surrounding the use of ChatGPT in education, Twitter data was examined. The study used the RoBERTa-based sentiment model for sentiment analysis and the BERT-based topic modeling for discourse analysis. In order to find influential users in the conversation, the social network was finally examined; this approach performed better when it came to the study of positive sentiments. The study's primary conclusions demonstrate that, on the whole, Twitter users had a favorable opinion of using ChatGPT in the classroom. Five specific problems were found in the investigation, though: policy and societal concerns, workforce challenges, impact on learning outcomes, limitations of AI capabilities, and academic integrity. The RoBERTa-based model that is employed determines the accuracy of sentiment analysis; it may not fully capture the complex emotions that are represented in tweets. Misclassification is a possibility, especially for complex or ambiguous tweets.

[58] examined higher education faculty perceptions of ChatGPT and the influencing factors to analyze the sentiment. Primarily, the comprehensive corpus of tweets was collected. Then, the collected data were analyzed by triangulating the Valence Aware Dictionary and sEntiment Reasoner (VADER), National Research Council (NRC) emotion lexicon, and ground coding. The findings show that 40% of sentiments expressed were positive, 51% were neutral, and 9% were negative. Due to the lack of generalization, this framework provided only a snapshot of information.

[49] described how natural language processing methods including sentiment analysis and topic modeling were used to assess public opinions regarding ChatGPT. Initially data from Twitter was gathered between December 5, 2022 and June 10, 2023. Then, the most popular topics such as education, bard, search engines, OpenAI, marketing, and cybersecurity were discussed. Education remained a consistently discussed topic, with tweets exploring both the benefits and challenges of integrating ChatGPT into educational settings. Out of the 3 sentiment analysis mdoels used: VADER (Valence Aware Dictionary and sEntiment Reasoner), Twitter-roBERTa, and XLM-T (Multilingual Transformer-based model), the XLM-T, outperformed the other three mdoels with appropriate preprocessing. It resulted in the closest match to human-labeled data. However, due to the reliance on user-provided descriptions, this model excluded some individuals and oversimplified career identities.

[87] evaluated the system for analyzing the public's response to ChatGPT through data obtained from Twitter. Firstly, the 10,722 tweets data was collected from Twitter and the collected data were selected for accurate sentiment analysis. Then, the preprocessing approach was employed, and the Term Frequency-Inverse Document Frequency (TF-IDF) method was applied for dataset division and vectorization of text data. Finally, the K-Nearest Neighbors (KNN) was applied for data mining that attained 88% accuracy when the value of k is 5. The analysis indicates a 74.3% of positive response to ChatGPT. Howvere, owing to the random selection of responses, this model did not represent the entire spectrum of user sentiments.

The document [76] primarily explores the integration and impact of ChatGPT in education. It focuses on sentiment analysis of higher education faculty perceptions, analyzing their attitudes towards ChatGPT's potential and challenges. Initially, the data was collected by employing web mining, and then the data were pre-processed based on corpus generation, tokenization, and stop word removal. Then, the information was extracted by using Word2Vec. Finally, these embeddings were then analyzed using Principal Component Analysis (PCA) and t-distributed Stochastic Neighbor Embedding (t-SNE) to classify the data and extract sentiment insights. The findings of sentiment analysis revealed a generally positive reception of ChatGPT in education, highlighting its role in enhancing student engagement and personalized learning experiences. Major concerns identified were related to academic integrity issues, such as plagiarism and cheating, emphasizing the need for ethical AI use and robust guidelines. The limitation of this model is that it might not capture nuanced opinions accurately due to the limitations of the Word2Vec model in understanding context-specific language variations and subtleties in sentiment.

[89] illustrated the sentiment analysis classification system using hybrid BERT models. The aim of this research was to enhance the accuracy of sentiment analysis by developing a hybrid BERT-based text classification model. The research proposed eight hybrid models combining BERT (RoBERTa and DistilBERT) with BiLSTM and BiGRU layers. The models included variations like DistilBERT-3G, DistilBERT-3L, DistilBERT-GLG, DistilBERT-LGL, RoBERTa-3G, RoBERTa-3L, RoBERTa-GLG, and RoBERTa-LGL. The effect of hybridizing the layer of this model using a combination of Bi-directional Encoder Representations from Transformers (BERT) with Bidirectional Long Short Term Memory (BiLSTM) and Bidirectional Gated Recurrent Unit (BiGRU) algorithms attained 90.49% accuracy in classifying sentiments. The study emphasized the need for robust preprocessing steps, including handling emojis, to improve model performance.

The moth flame optimization with hybrid deep learning-based sentiment categorization toward chatGPT on Twitter was illustrated by [7]. Researchers compared the performance of Hybrid Deep Learning Model (HDL) and Moth Flame Optimization (MFO) in sentiment analysis. Here, the Moth Flame Optimization with Hybrid Deep Learning-based Sentiment Analysis (MFOHDL-SA) outperformed other existing techniques with 95.09% accuracy in sentiment classification. Also, the findings on sentiments regarding chatgpt shows that the users appreciated ChatGPT's ability to generate human-like text and its potential to enhance user experiences in conversational AI settings. There were also negative sentiments showing concerns regarding potential for misuse, ethical concerns, and the accuracy of generated responses. There was also a limitation of the model owing to the small batch size, this model exhibited low stable training results.

2.5 Summary and Research Gap

The literature review provides a comprehensive overview of the integration of ChatGPT into educational settings, focusing on sentiments and attitudes towards its use. It examines various studies that highlight both positive and negative perceptions of ChatGPT in education. The positive aspects include enhanced student engagement and personalized learning experiences, while concerns center around academic integrity, privacy issues, and the accuracy of generated content. Different sentiment analysis methodologies are reviewed, comparing their effectiveness and limitations in capturing user sentiments. This section identifies gaps in current research, particularly in nuanced sentiment analysis and the broader implications of AI integration in education, setting the stage for the proposed study's objectives and methodologies.

- A noticeable gap exists in identifying specific sentiments and discourse topics related to ChatGPT adoption in Higher Education where students and academic staff actively incorporate such technologies into their daily learning, research, and teaching tasks.
- Existing research examines sentiments across platforms like Twitter and Reddit, there remains an unexplored perspective concerning how academic institutions perceive and adopt this technology in higher education.
- Sentiment analysis often requires understanding the context in which the text was written. Capturing and incorporating contextual information effectively remains a significant challenge, especially in cases, where context could drastically alter sentiment.
- One of the biggest challenges in sentiment analysis is identifying and accurately understanding nuances in sentiment expression, such as sarcasm, irony, and other figurative language [7].
- The presence of complex language and diverse expressions in the sentence remains a challenge in effectively analyzing the sentiments [49].
- Several research papers have highlighted the challenge of handling long texts or large paragraphs in sentiment analysis, necessitating the identification of attention at different levels. Sentiment analysis on long texts is challenging due to the presence of multiple sentiments and topics within the same text [63, 100, 40].

2.5.1 Research questions

1. What are the overall sentiments of universities towards the integration of ChatGPT in higher education, what factors influence these sentiments and what are the main topics or concerns discussed?

2. How can fine-grained sentiment nuances be effectively captured from the collected data about ChatGPT integration?

3 Theoretical background

This chapter delves into the machine learning algorithms, evaluation metrics, web scraping techniques and data processing libraries used in this thesis. It highlights the integration of these techniques within the sentiment analysis framework, providing a comprehensive overview of their application in assessing sentiments towards ChatGPT in education.

3.1 Sentiment Analysis

Sentiment analysis is a computational method that systematically identifies, extracts, quantifies, and studies subjective information from textual data provided as input to the classification step. It does this by combining natural language processing (NLP), text analysis, and machine learning[18]. It aims to determine the sentiment polarity (positive, negative, or neutral) expressed in a piece of text towards a particular topic, product, service, or entity [20].
Sentiment analysis can be implemented using various approaches:

- Rule-based methods: These use predefined lexicons, which are lists of words annotated with additional information such as their meanings and sentiment polarity [2]. Linguistic rules are then applied to identify sentiment in text. Commonly known approaches are SentiWordNet, VADER, and TextBlob.

- Machine learning techniques: Machine learning algorithms such as Support Vector Machines (SVM), Naive Bayes, or deep learning models are trained on labeled datasets to classify sentiment.

- Hybrid approaches: This approach is a combination of rule-based and machine-learning methods for improved accuracy. A system might use rule-based methods to pre-process data and then apply machine learning for classification.

While rule-based methods provide a straightforward approach to sentiment analysis by using predefined lexicons and linguistic rules, they often fall short in handling the complexity and variations of human language. Machine learning techniques, on the other hand, offer significant advantages by learning from data and adapting to various linguistic patterns. These methods can capture context, deal with ambiguities, and improve accuracy through continuous learning.

Deep learning is a subset of machine learninguses multi-layered artificial neural networks to extract complex patterns from large datasets. Conventional Machine learning typically uses traditional algorithms like Naive Bayes, Logistic Regression, and Random Forest. Deep learning employs neural network architectures such as Recurrent Neural Networks (RNNs), Deep Neural Networks (DNNs), and Convolutional Neural Networks (CNNs). While conventional Machine learning techniques often requires manual feature engineering where relevant features are extracted from the text data, Deep learning models can automatically learn useful feature representations from raw text data, eliminating the need for manual feature engineering [43].

3.2 Deep Learning Techniques for Sentiment Analysis

3.2.1 Deep Neural Networks (DNN)

A Deep Neural Network (DNN) is a complex architecture composed of multiple layers of artificial neurons designed to model intricate patterns in data. These networks excel in tasks involving unstructured data such as sequences, texts, and trees by learning representations through layered processing. A DNN's layers are made up of a collection of neurons, each of which takes in inputs, processes them using a weighted sum, and then runs the outcome via an activation function to add non-linearity to the model. The core components of DNNs include input layers, hidden layers, and an output layer which is shown in figure 3.1. The input layer receives raw data, which is then transformed by several hidden layers, and then the output layer produces the prediction or classification result [50].

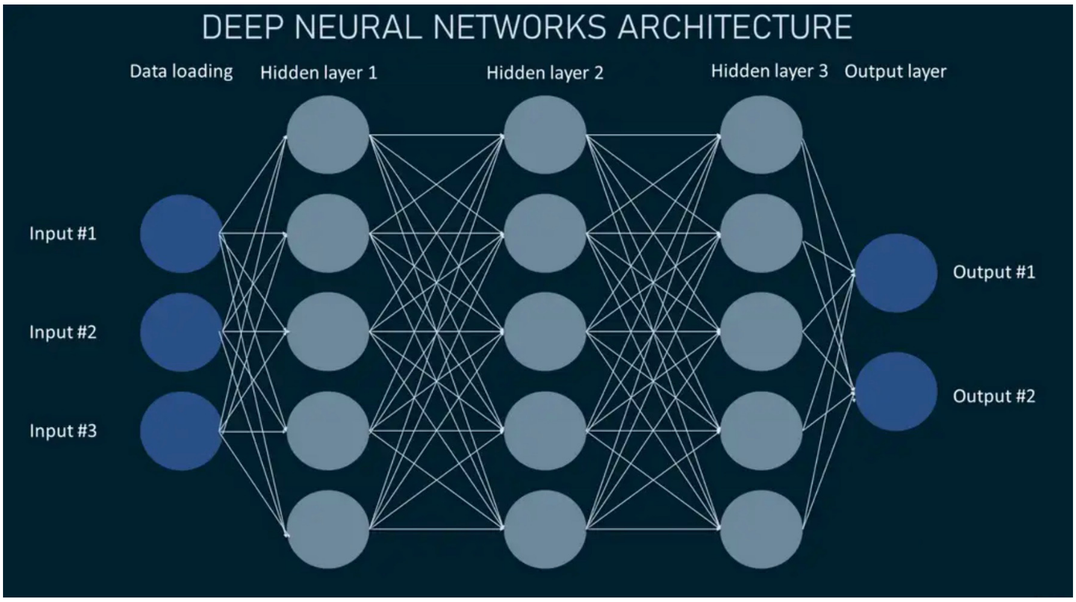


Figure 3.1: Deep Neural Network Architecture [8]

The working of a DNN involves two primary phases: forward propagation and backpropagation. Each layer receives the input data during forward propagation, with each neuron computing a weighted sum of its inputs and applying an activation function to produce an output. Mathematically, the output of a neuron in layer l can be represented as

$$h_i^l = f(W^l h^{l-1} + b^l)[50]$$

where W^l and b^l are the weight matrix and bias vector for layer l, respectively, h^{l-1} is the output from the previous layer, and f is the activation function [50]. Backpropagation involves computing the gradient of the loss function with respect to each weight using the chain rule and updating the weights to minimize the loss. The gradient of the loss with respect to the weights in layer l is given by

$$\frac{\partial L}{\partial W^l} = \delta^l (h^{l-1})^T [50]$$

where δ^l is the error term for layer l and L is the loss function. This iterative process continues until the network's performance is optimized. DNNs have shown remarkable success in various applications, such as sentiment analysis, by effectively capturing complex structures and relationships within the data [50].

3.2.2 Recurrent Neural Networks (RNN)

Recurrent Neural Networks (RNNs) are tailored for sequential data processing. They maintain a hidden state that captures information about the preceeding elements in the sequence, making them particularly suitable for tasks like text classification where context and order matters [54]. RNNs have connections that create directed cycles, in contrast to conventional feedforward neural networks, allowing them to maintain a 'memory' of previous inputs by persisting state across time steps. This makes them particularly suited for tasks where the order and context of inputs are crucial, such as language modeling, speech recognition, and time-series prediction.

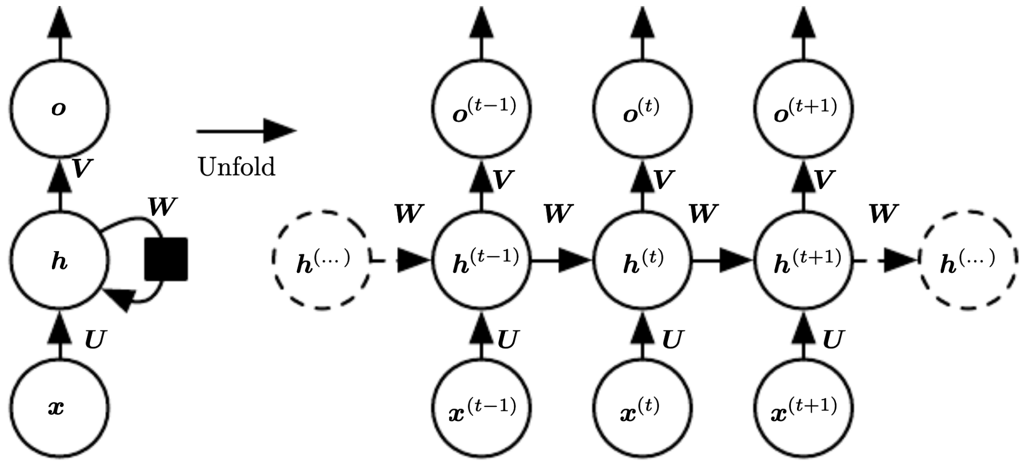


Figure 3.2: Recurrent Neural Network Architecture [26]

As per Figure 3.2, the architecture of an RNN can be visualized through its unfolded representation, which shows how the network processes a sequence of inputs over time. In the unfolded RNN architecture, each time step t involves three main components: the input x, the hidden state h, and the output o. The hidden state h is updated at each time step based on the current input x and the previous hidden state h, using the equation

$$h^{(t)} = \sigma(W_h h^{(t-1)} + W_x x^{(t)} + b)[26]$$

where W_h and W_x are weight matrices, b is a bias vector, and σ is a nonlinear activation function such as tanh or ReLU [26]. The output o at each time step is then computed using

$$o^{(t)} = V h^{(t)[26]}$$

[26] where V is another weight matrix. This parameter-sharing mechanism across different time steps allows the RNN to generalize well to sequences of varying lengths and capture long-term dependencies effectively [26].

3.2.3 Convolutional Neural Networks (CNN)

A Convolutional Neural Network (CNN) is a type of deep learning model which can be used for effective text classification tasks because of its ability to capture local dependencies and hierarchical structures in text [41].

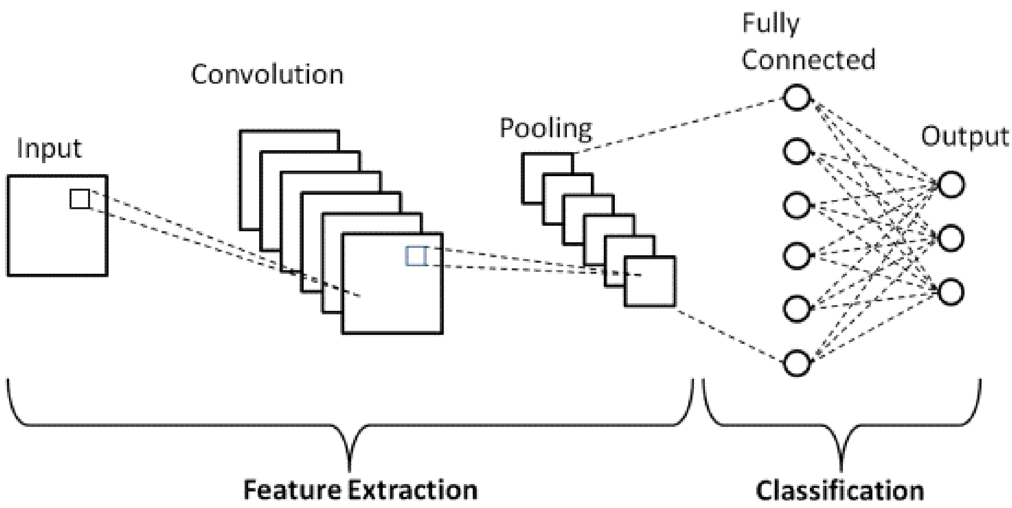


Figure 3.3: Convolutional Neural Network Architecture [29]

Figure 3.3 shows the basic CNN architecture which typically consists of the following layers: Input, Convolution, Pooling, Fully Connected, and Output.

1. **Input Layer:** Each word in a sentence is represented as a k-dimensional word vector using embeddings like Word2Vec or GloVe. For a sentence of length n, this results in an embedding matrix $X \in \mathbb{R}^{n \times k}$.

2. **Convolutional Layer:**
   - **Filters/Kernels:** Convolutional filters $w \in \mathbb{R}^{h \times k}$ of height h (number of words) and width k (dimension of word vectors) slide over the input matrix to extract local features.
   - **Convolution operation:** For each position i in the text, a filter generates a feature c_i by applying the convolution operation:

     $$c_i = f\left(w \cdot X_{i:i+h-1} + b\right) [47]$$

     where $X_{i:i+h-1}$ is the segment of the embedding matrix, b is a bias term, and f is a non-linear activation function (e.g., ReLU) [47].

3. **Feature map:**
   - The result of the convolution operation across all positions i forms a feature map:

     $$\mathbf{c} = [c_1, c_2, \ldots, c_{n-h+1}][47]$$

     where $\mathbf{c} \in \mathbb{R}^{n-h+1}$.

4. **Pooling layer:**

- **Max-pooling:** This operation reduces the dimensionality of the feature map while retaining the most important features by taking the maximum value:

$$\hat{c} = \max(\mathbf{c})[47]$$

This process helps to capture the most salient feature detected by each filter, creating a fixed-size output regardless of the input length [47] [41].

5. **Fully connected layer:**
   - The pooled features from different filters are concatenated into a single vector and passed to a fully connected layer. This layer combines the features to make a final classification:

$$\mathbf{z} = \mathbf{W} \cdot \hat{\mathbf{c}} + \mathbf{b}[47]$$

where $\mathbf{W}$ is the weight matrix, $\hat{\mathbf{c}}$ is the concatenated vector of pooled features, and $\mathbf{b}$ is the bias vector [47].

6. **Output layer:**
   - **Softmax Activation:** The fully connected layer's output is passed through a softmax function to produce a probability distribution over the possible classes:

$$\hat{y} = \text{softmax}(\mathbf{z})[47]$$

where $\hat{y}$ represents the predicted probabilities for each class [47, 41].

A Convolutional neural network (CNN) for text classification processes input text through several key layers. First, the text is converted into an embedding matrix, where each word is indicated by a vector. This matrix serves as the input layer. Convolutional layers then apply filters to this matrix to detect local patterns like n-grams. The output of these filters forms feature maps, which highlight important features in the text. These feature maps are then reduced in size by max-pooling layers, which retain the most significant features while discarding less important ones. The pooled features are flattened and passed through fully connected layers, which combine the information and produce the final classification output through a softmax activation function. This series of operations allows the CNN to effectively capture and utilize patterns in the text to classify it accurately. Advanced neural network architectures, like the Hierarchical Attention Network (HAN), are particularly effective for handling complex textual data. This architecture, which is the focus of our study, is discussed in the next section.

3.3 Hierarchical Attention Networks (HAN) for Sentiment Analysis

Hierarchical Attention Networks (HAN) is an advanced neural architecture designed to address the complexities of document classification. It is primarily used for document classification tasks. It mirrors the hierarchical structure of documents by incorporating two levels of attention mechanisms: word-level and sentence-level attention [102]. This allows the model to focus on the most relevant words and sentences, improving interpretability and classification accuracy. HAN effectively captures the context and importance of different parts of the text / document, making it particularly powerful for tasks involving long and complex documents.

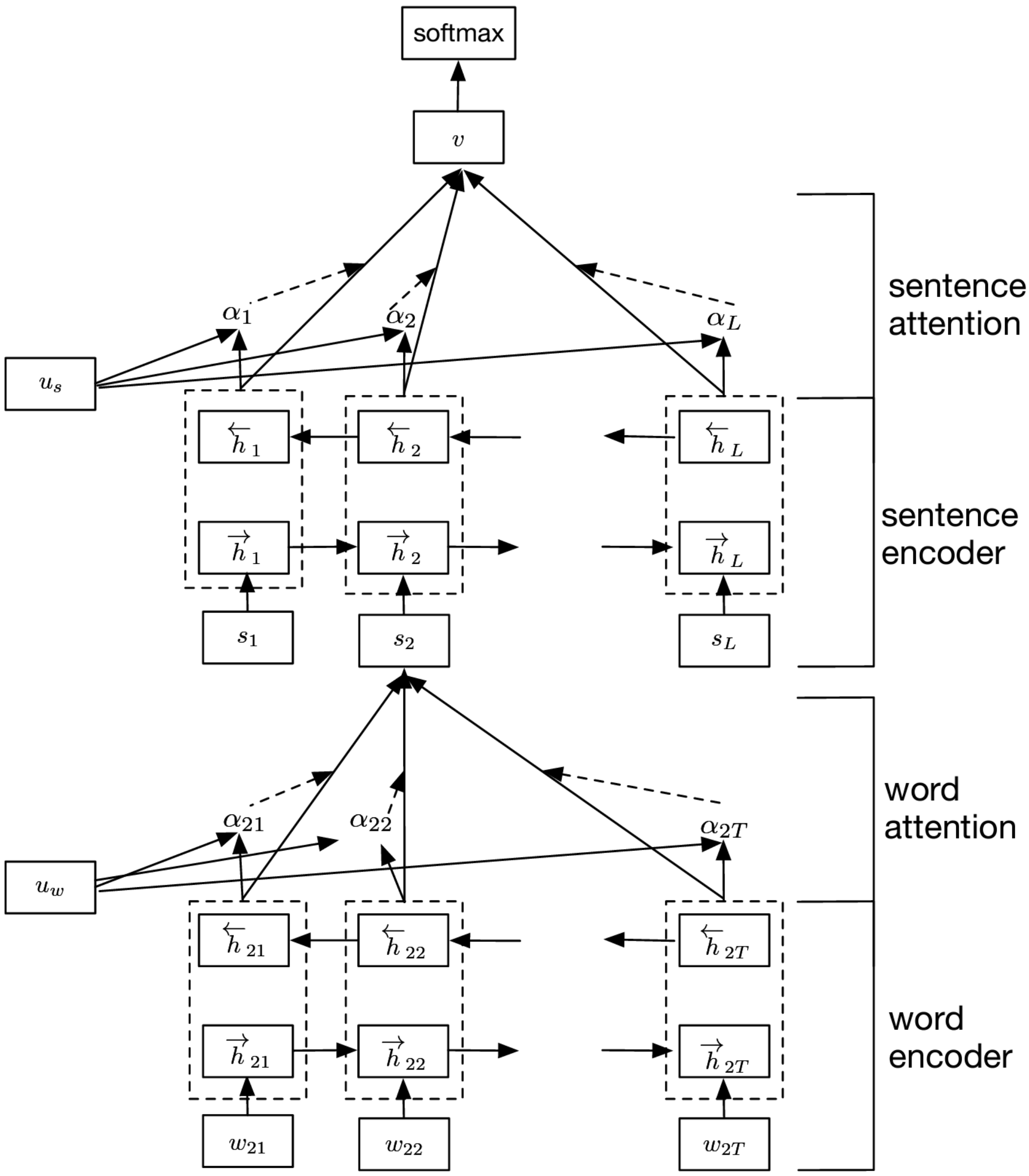


Figure 3.4: Hierarchical Attention Network Architecture [102]

U_s: sentence context vector
α: importance weight
U_w: word context vector (what is the informative word?)
$h^{\rightarrow}$: forward hidden state
$h^{\leftarrow}$: backward hidden state
$word_{\text{sen2,1}}$: The first word in sentence 2

3.3.1 Components of HAN

- **Word sequence encoder:** It uses a bidirectional GRU (Gated Recurrent Unit) to encode the sequence of words within each sentence. This captures the context from both directions (forward and backward) and produces annotations for each word.
- **Word-Level attention layer:** Applies an attention mechanism to the word annotations to identify the most important words within a sentence. This layer produces a sentence vector as a weighted sum of word annotations, where the weights reflect the importance of each word.
- **Sentence encoder:** Like the word sequence encoder, this component uses a bidirectional GRU to encode the sequence of sentence vectors, producing annotations for each sentence.
- **Sentence-Level attention layer:** This layer applies an attention mechanism to the sentence annotations to identify the most important sentences within a document. It produces a document vector as a weighted sum of sentence annotations, where the weights reflect the importance of each sentence.
- **Document classification layer:** The final document vector is used as input to a softmax classifier to assign labels to the document.

3.3.2 Working of HAN

The HAN's working mechanism capitalizes on the hierarchical structure of text data. By first focusing on encoding individual words into vectors, it captures the context-specific importance of each word within a sentence. The word-level attention mechanism assigns higher weights to more informative words, contributing significantly to the sentence representation. This sentence representation is further processed through a sentence-level encoder and attention mechanism, which aggregates the sentences' vectors to form a final document representation [102].

1. **Word sequence encoder:**
   - Each word w_{it} in a sentence s_i is embedded into a vector space: $x_{it} = W_e w_{it}$.

- A bidirectional GRU processes these word vectors to generate annotations:

$$\overrightarrow{h_{it}} = \text{GRU}(x_{it}), \quad \overleftarrow{h_{it}} = \text{GRU}(x_{it})$$

$$h_{it} = [\overrightarrow{h_{it}}, \overleftarrow{h_{it}}][102]$$

2. **Word-Level attention layer:**
   - Compute a hidden representation u_{it} for each word annotation:

$$u_{it} = \tanh(W_w h_{it} + b_w)[102]$$

   - Measure the importance of each word:

$$\alpha_{it} = \frac{\exp(u_{it}^\top u_w)}{\sum_t \exp(u_{it}^\top u_w)}[102]$$

   - Form the sentence vector as a weighted sum of word annotations:

$$s_i = \sum_t \alpha_{it} h_{it}[102]$$

3. **Sentence encoder:**
   - Process the sentence vectors with another bidirectional GRU:

$$\overrightarrow{h_i} = \text{GRU}(s_i), \quad \overleftarrow{h_i} = \text{GRU}(s_i)$$

$$h_i = [\overrightarrow{h_i}, \overleftarrow{h_i}][102]$$

4. **Sentence-Level attention layer:**
   - Compute a hidden representation u_i for each sentence annotation:

$$u_i = \tanh(W_s h_i + b_s)[102]$$

   - Measure the importance of each sentence:

$$\alpha_i = \frac{\exp(u_i^\top u_s)}{\sum_i \exp(u_i^\top u_s)}[102]$$

   - Form the document vector as a weighted sum of sentence annotations:

$$v = \sum_i \alpha_i h_i[102]$$

5. **Document classification:**

- The document vector v is used as input to a softmax classifier:

$$p = \text{softmax}(W_c v + b_c)[102]$$

- The training loss is the negative log likelihood of the correct labels:

$$L = -\sum_d \log p_{dj}[102]$$

where j is the correct label for document d.

3.3.3 Importance of HAN

The hierarchical attention network (HAN) improves document classification accuracy in several key ways:

- **Hierarchical structure:** The HAN mirrors the hierarchical structure of documents (word -> sentence -> document), allowing it to better capture the natural organization of text.
- Two levels of attention: It employs attention mechanisms at both the word and sentence levels, enabling the model to focus on the most relevant words within sentences and the most important sentences within documents.
- **Differential attention:** The model can attend differentially to more and less important content when constructing document representations, helping it identify the most salient information.
- **Improved representation:** Important sentence vectors that are similar to the word sequence represented are created by combining important words into sentence vectors and then into document vectors, the HAN builds a more informative document representation.
- **Context awareness:** The hierarchical structure allows the model to consider words and sentences in context, rather than treating all text elements equally.
- **Visualization and interpretability:** The attention weights can be visualized to show which words and sentences were most influential in the classification decision, providing insights into the model's reasoning

3.4 Neural Network Components of HAN

3.4.1 Residual Connections

Residual connections, also known as skip connections, are a technique introduced to address the degradation problem in deep neural networks. They bypass one or more layers by creating a shortcut for the input data. Instead of having layers directly fit a desired underlying function $H(x)$, residual connections let these layers fit a residual function $F(x) = H(x) - x$. The original function is then reformed as $H(x) = F(x) + x$ [31]. This reformulation simplifies the learning process of HAN by focusing on learning the residuals between layers rather than the entire transformation.

3.4.2 How Residual Connections Work

Residual connections are implemented by adding shortcut paths that skip one or more layers in advanced neural network architecture such as HAN. These shortcuts typically perform identity mapping, and their outputs are added to the outputs of the stacked layers. Formally, a building block in a residual network can be described as:

$$y = F(x, \{W_i\}) + x[31]$$

where x and y are the input and output vectors, respectively, and $F(x, \{W_i\})$ represents the residual function to be learned, typically consisting of a few layers with weights $\{W_i\}$ [31]. This approach allows for the use of deep networks without the common problems associated with training very deep models, such as vanishing gradients.

3.4.3 Why Use Residual Connections

1. **Ease of optimization**: The degradation issue, in which deeper models have greater training errors than their shallower equivalents, is lessened via residual connections. By reformulating the learning objective to focus on residuals, it becomes easier to optimize deep networks. If the optimal function is closer to an identity mapping, the residual connections help the solver find this optimal point by reducing the residuals towards zero [31].

2. **Addressing the vanishing gradient problem**: Deep networks often suffer from vanishing gradients, where gradients become exceedingly small during backpropagation, making training difficult. Residual connections allow gradients to flow more easily through the network, thus maintaining the effectiveness of backpropagation even in very deep networks.

3. **Improved training and generalization**: Residual networks (ResNets) have been shown to be easier to train and often achieve better accuracy compared to their plain counterparts. This has been validated across various datasets where deep residual networks significantly outperformed plain networks with similar or greater depths.

4. **Scalability**: Residual connections enable the construction of extremely deep networks, sometimes exceeding 1000 layers, without encountering severe train-

ing issues. This scalability opens up new possibilities for building more powerful models that can learn more complex patterns.

By leveraging these advantages, residual connections have become a foundational component in the design of deep learning architectures, contributing to the success of many state-of-the-art models in various fields of machine learning and computer vision [31].

3.4.4 RsigELU activation function

The RSigELU activation function is a novel activation function proposed to address the vanishing gradient problem and the negative region problem encountered in traditional activation functions like Sigmoid, Tanh, and ReLU. The RSigELU combines the features of ReLU, Sigmoid, and ELU (Exponential Linear Unit) functions to provide a more robust activation mechanism in neural networks [46].

RSigELUS (Single-parameter RSigELU):

- **Positive Region**: In the range $1 < x < \infty$, the function behaves as $x\left(\frac{1}{1+e^{-x}}\right)^{\alpha} + x$, where α is a slope parameter.
- **Linear Region**: In the range $0 \leq x \leq 1$, the function behaves linearly as x.
- **Negative Region**: For $-\infty < x < 0$, the function behaves as $\alpha(e^{x} - 1)$.

The RSigELUS function is designed to have three active regions: positive, linear, and negative. The parameter α controls the behavior in the positive and negative regions. When $\alpha = 0$, the RSigELUS function reduces to the ReLU function [46].

Equation:

$$f(x) = \begin{cases} x\left(\frac{1}{1+e^{-x}}\right)^{\alpha} + x & \text{if } 1 < x < \infty \\ x & \text{if } 0 \leq x \leq 1 \\ \alpha(e^{x} - 1) & \text{if } -\infty < x < 0 \end{cases}$$

RSigELUD (Double-parameter RSigELU):

- This function extends RSigELUS by introducing a second parameter β to provide additional control over the activation function's behavior in the negative region [46].
- **Positive Region**: Same as RSigELUS.
- **Linear Region**: Same as RSigELUS.
- **Negative Region**: For $-\infty < x < 0$, the function behaves as $\beta(e^{x} - 1)$.

Equation:

$$f(x) = \begin{cases} x\left(\frac{1}{1+e^{-x}}\right)^{\alpha} + x & \text{if } 1 < x < \infty \\ x & \text{if } 0 \leq x \leq 1 \\ \beta(e^{x} - 1) & \text{if } -\infty < x < 0 \end{cases}$$

Overcoming Vanishing Gradient Problem: By incorporating elements of the sigmoid function in the positive region and ELU in the negative region, RSigELU maintains gradient flow, thereby avoiding the vanishing gradient problem commonly faced in deep networks with Sigmoid and Tanh activations [46].

Handling Negative Region: Unlike ReLU, which outputs zero for all negative inputs, RSigELU leverages the ELU function in the negative region, ensuring that negative values contribute to the learning process [46].

3.5 Evaluation metrics

Evaluation metrics or performance metrics are used for evaluating how well a model performs. These metrics help in assessing the model's effectiveness and suitability for the intended task. They provide a quantitative basis for comparing and optimizing different models, guiding the selection of the most effective algorithms for specific tasks.

Different machine learning tasks require different evaluation metrics. For example, classification tasks, where the goal is to assign labels to instances, use metrics such as accuracy, precision, recall, and F1 score. On the other hand, regression tasks, which aim to predict continuous values, often utilize metrics like Mean Squared Error (MSE) and Mean Absolute Error (MAE) [61, 35]. There are also general evaluation metrics such as bias-variance trade-off and and overfitting/underfitting which helps to ensure that the model is neither too complex nor too simple for the data.

3.5.1 Classification Metrics

Classification metrics are used to evaluate the performance of classification algorithms, which predict categorical labels.

- **Confusion matrix**
  A table that provides an overview of a classification algorithm's performance is called a confusion matrix. It shows the true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN), providing a detailed breakdown of correct and incorrect predictions [61, 35].

		Predicted Class	
		Positive	Negative
Actual Class	Positive	**TP** The model correctly predicts a positive case	**FN** The model incorrectly predicts a negative case when it's actually positive
	Negative	**FP** The model incorrectly predicts a positive case when it's actually negative	**TN** The model correctly predicts a negative case

- **Accuracy** Accuracy represents the ratio of correctly predicted instances to the total instances [35].

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}[35]$$

- **Precision** Precision measures the proportion of true positive predictions among all positive predictions [35].

$$\text{Precision} = \frac{TP}{TP + FP}[35]$$

- **Recall** Recall, or sensitivity, calculates the proportion of true positives out of the actual positives [35].

$$\text{Recall} = \frac{TP}{TP + FN}[35]$$

- **F1 Score** The F1 score is the harmonic mean of precision and recall [35].

$$\text{F1 Score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}[35]$$

- **Specificity** Specificity measures the proportion of true negatives among all actual negatives, which is essential for tasks where identifying the negative class correctly is critical [35].

$$\text{Specificity} = \frac{TN}{TN + FP}[35]$$

3.6 Libraries and Tools

3.6.1 Web Scraping

Web scraping is the process of automating the information extraction from websites instead of manually copying data. Web scraping can be performed using many programming languages. Among these, Python stands out as the most powerful. This is due to its huge built-in libraries specifically for web scraping, extensive support for third-party open-source libraries, and generally higher execution speeds compared to other languages [78].

The web scraping process involves several key steps mentioned in [103]:

- **HTTP Request:** Send a request to the targeted website using a URL (GET) or form data (POST).
- **Response:** Receive the response from the server, which can be in HTML, JSON, XML, or other formats.
- **Parsing:** Use tools like Beautiful Soup or Pyquery to parse and navigate the response data, extracting the required information.
- **Data Storage:** Store the extracted data in a structured format such as CSV, JSON, or a database for further use.

All webpages are built using HTML as a primary backbone. Data can be extracted from both static and dynamic webpages and there are various techniques to extract depending on the type of webpage and the specific task. Some of these techniques described in [78, 44] include:

- **HTML parsing:** Extracting data by parsing the HTML structure of a webpage.
- **DOM (Document Object Model) parsing:** Navigating and manipulating the document structure of a webpage to extract data.
- **API usage:** Interacting with a webpage's API to retrieve structured data directly.
- **Headless browser scraping:** Using headless browsers like Selenium to render and interact with webpages programmatically.
- **Regular expressions:** Using pattern matching to find and extract data within the webpage's content.
- **Computer vision-based webpage analyzers:** Employing computer vision techniques to analyze and extract data from webpage images or layouts.

In addition to these techniques, there are many web scraping tools, or libraries available, such as Beautiful Soup and Selenium [78].

3.6.2 Beautiful Soup

Beautiful Soup is a Python library designed for parsing HTML and XML documents, making it a highly effective tool for web scraping to extract data from web pages. It constructs a parse tree from the given document, allowing for easy navigation and data extraction [78]. The library provides several methods, such as find(), find_all(), and select(), to locate elements based on tags, attributes, text content, and more. However, Beautiful Soup has limitations when dealing with dynamic web pages that rely on JavaScript to load content after the initial HTML is fetched, as it can only process static HTML content. To handle dynamic web pages, Beautiful Soup is often used in conjunction with tools like Selenium, which can render JavaScript and provide the fully loaded HTML for Beautiful Soup to parse.

3.6.3 Selenium

Selenium can interact with web pages in real-time, making it capable of handling dynamic content that is loaded through JavaScript. Selenium operates by controlling a web browser via scripts written in various programming languages, including Python, Java, and C [25]. It works by automating a web browser through its Web-Driver component, which can be controlled using scripts written in various programming languages. It initiates a browser instance, navigates to web pages, and interacts with elements on the page such as buttons, forms, and links, allowing for dynamic content to be loaded and data to be extracted. This interaction mimics real user behavior, making Selenium effective for scraping content that requires user actions. Selenium requires more computing resources and time compared to other scraping tools. Despite these demands, Selenium's ability to handle complex web interactions and its flexibility across different browsers make it an invaluable tool for web scraping projects that involve dynamic content.

In my implementation, I have utilized a combination of web scraping techniques, including HTML parsing with Beautiful Soup, real-time interaction with dynamic content using Selenium. This approach ensured that I could effectively extract data from both static and dynamic web pages across various university websites.

3.7 Topic Modelling

Topic Modeling is a statistical technique used to discover the underlying thematic structure within large collections of documents. It identifies hidden patterns by grouping words into topics, where a topic is a collection of words that frequently co-occur [45]. The technique is beneficial for organizing and summarizing large datasets of textual information. Word Clouds are a visualization where the size of each word reflects its importance or frequency in a given context, making them particularly useful for presenting the key words in each topic

3.7.1 Representation using Wordcloud

Word clouds are a popular method for visualizing text data, allowing users to quickly capture the most frequently occurring words in a dataset by representing them in varying sizes based on their frequency or importance. In Python, the matplotlib library is often used in conjunction with the wordcloud module to generate these visualizations. The process typically involves text preprocessing to remove stop words and irrelevant terms, followed by tokenizing the text to identify significant words and their frequencies. These words are then fed into the WordCloud function from the wordcloud library, which generates a word cloud image that can be displayed using matplotlib. The size and color of each word in the cloud reflect its prominence, offering an intuitive means of identifying key themes within a corpus. According to [32], word clouds can be a useful tool for exploratory data analysis, enabling researchers to uncover patterns and insights in textual data quickly. Furthermore, [17] highlight the importance of integrating visualization techniques, like word clouds, with computational methods for topic modeling, as they enhance the interpretability of complex models and facilitate the communication of results to broader audiences.

3.8 GUI Development

3.8.1 Tkinter

Tkinter is Python's standard GUI (Graphical User Interface) library, providing an interface to the Tcl/Tk toolkit. It allows for the creation of desktop applications with a range of widgets, such as buttons, labels, and text fields [1]. Tkinter integrates with the Tcl interpreter, enabling developers to design interactive applications. It supports themed widgets through the 'tkinter.ttk' module and offers various dialog boxes and layout management options. Tkinter is widely used for its ease of use and simplicity, making it a popular choice for building Python-based GUI applications To get started with Tkinter, it can be imported as "from tkinter import *" [1]. In my implementation, I used Tkinter to create a user-friendly interface for my sentiment analysis framework. This GUI enables easy interaction and visualization of the sentiment analysis results at each step, facilitating better understanding and usability of the framework

4 Methodology

In this proposed framework, Hierarchical Attention Networks with Residual connections and RSigELU activation function (HR-RAN) is developed. It is an advanced neural network architecture for performing sentiment analysis of collected data regarding ChatGPT integration in higher education. This approach offers an efficient and effective analysis of sentiments by focusing on different level of granularity in text and also capturing the fine-grained sentiment nuances. The proposed framework starts with Data collection and the following steps in sequence: handling Emoticons, Uniform Resource Locators (URLs), tags handling, pre-processing, sarcastic text detection, sentence grouping, topic modeling, content filtering, dependency parsing, coreference resolution, word embedding, feature extraction, classification, and calculating the degree of sentiments.

Based on the above steps, this methodology starts by collecting the text data regarding ChatGPT adoption for education from official university websites. Then, the emoticons, URLs, and tags are handled for effective sentiment classification. Meanwhile, the collected data and the handled data are pre-processed based on sentence splitting, special character removal, abbreviation and contraction handling, case conversion, spell checking, and text normalization. Next, the sarcastic texts are detected from the pre-processed output by using the eXtreme Learning Network (XLNet) with hyperband parameter. After the detection and removal of sarcastic texts, the sentences are grouped in terms of their similarity by using hierarchical clustering technique. Subsequently, the Non-negative-based Term-Document Matrix Factorization (NTDMF) is employed to model the topics for effective processing. Further, dependency parsing and coreference resolutions are identified from the modeled topics. Then, the Word2Vec approach is utilized to perform word embedding by capturing the semantic relationships between words. Now, the output from the pre-processing, word-embedding, and feature extraction are given to the classifier called the HR-RAN, which classifies the input texts into positive, negative, and neutral sentiments. The structural diagram for the proposed framework is illustrated in Figure 4.1.

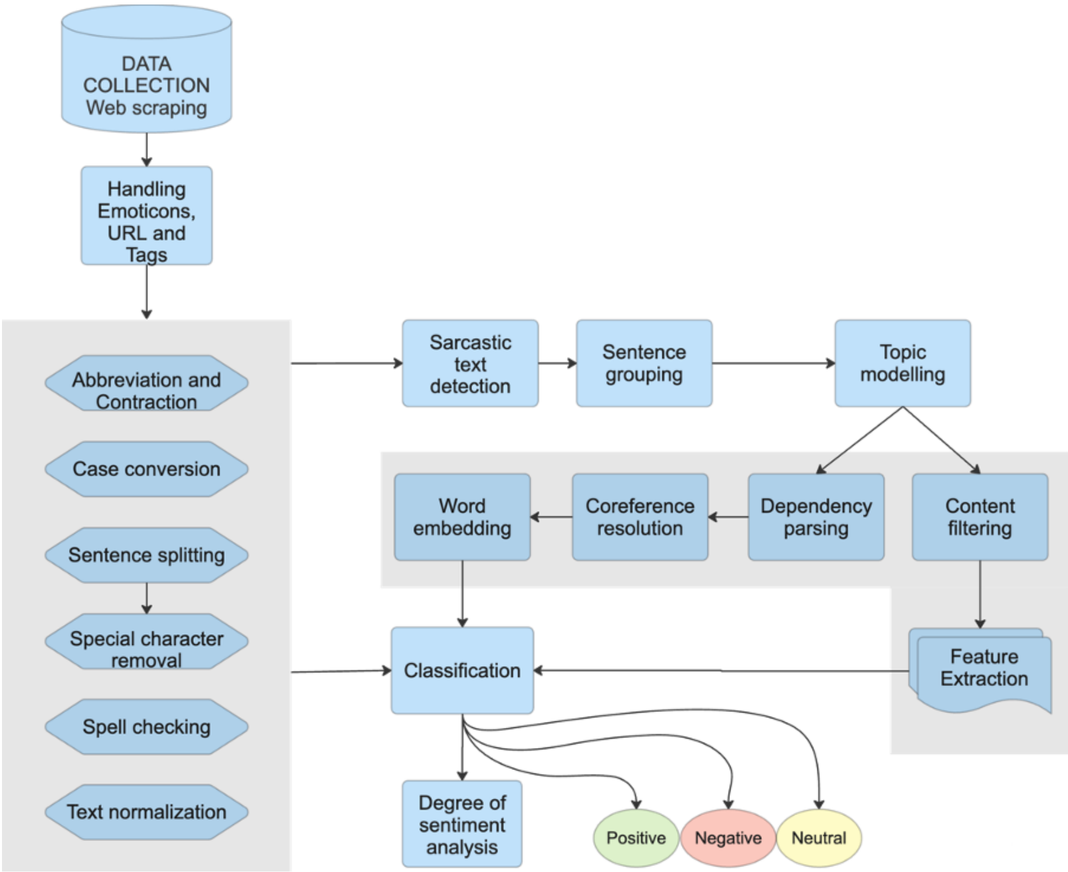


Figure 4.1: Structure of the proposed framework

4.1 Data Collection

The primary data source in this study consists of discussions about ChatGPT collected from official university websites. The data collection process employs two web scraping techniques: Beautiful Soup and Selenium. These tools effectively handle both static and dynamic web pages and facilitate real-time interaction with dynamic content. The data is gathered from universities in four countries: Germany, the United States, the United Kingdom, and Canada. The universities selected for this study not only have strong reputations in research and innovation but also represent diverse educational philosophies and environments, focusing on higher education. Additionally, these countries offer a broad spectrum of cultural and educational contexts, contributing to a comprehensive analysis of how universities globally communicate about ChatGPT. Table 4.1 shows the universities from which the data samples are extracted.

To identify relevant content, a systematic search was conducted using key terms such as: ‘Artificial Intelligence,’ ‘Machine Learning,’ ‘ChatGPT,’ ‘OpenAI,’ ‘GPT,’ ‘Conversational AI,’ ‘AI Integration,’ and ‘Integration into Academics.’ This search aimed to capture diverse perspectives and discussions on ChatGPT across different universities. The websites were evaluated for suitability based on the following inclusion and exclusion criteria:

Inclusion criteria:

- Discussion of ChatGPT about academic integrity issues.
- Discussion of ChatGPT in relation to its adoption in academics.
- University policies regarding the usefulness of AI tools such as ChatGPT.
- Discussion regarding the positive and negative aspects of ChatGPT concerning education.

Exclusion criteria:

- General discussions where the overall context is not relevant for my research.
- Discussions focusing on other conversational AIs (except ChatGPT).
- Content focused on primary and high schools.

The data collection process is conducted with a strong commitment to ethical web scraping standards as mentioned in [51] by respecting the terms of use and policies of the targeted websites. Sensitive or personally identifiable information is explicitly avoided to protect user privacy. To address potential problems, such as language variations in the extracted data, language translation tools like Google Translate, leveraging machine learning and neural machine translation, were utilized. The gathered data for analysis consists of 1,080 entries representing diverse opinions from official university websites. This data is used to understand the discourse surrounding ChatGPT in higher education. To train the sentiment analysis model, a separate dataset from Kaggle [9] is employed, comprising 98,760 entries of ChatGPT-related tweets from january to march 2023. This pre-labeled dataset includes positive, negative, and neutral sentiment labels about users’ opinions regarding ChatGPT. By leveraging both datasets, the study aims to provide a thorough analysis of sentiments towards ChatGPT in academic contexts.

The collected data using web scraping can be represented by the following mathematical expression:

$$D = \{d_1, d_2, d_3, \ldots, d_n\}[9]$$

Where $D = \{d_i \mid i = 1 \ldots n\}$, d_i indicates the collected review data, and n represents the total number of collected data. After collection, the data is processed in the next steps.

Table 4.1: List of universities from which the data was collected

Country	**University**
Canada	Fairleigh Dickinson University The King's University University of Alberta University of British Columbia University of Toronto University of Waterloo
Germany	Frankfurt School of Finance & Management Free University of Berlin Technical University of Munich (TUM) University of Hamburg University of Heidelberg
United Kingdom	Cranfield University Imperial College London University College London University of Cambridge University of Edinburgh University of Oxford
United States	Harvard University Massachusetts Institute of Technology Stanford University University of California, Berkeley University of Michigan

4.2 Handling Emoticons, URLs, and Tags

This phase processes the collected data by handling emoticons, URLs, and tags present in the input review text for effective sentiment classification. Emoticons express emotions directly, influencing the sentiment of the text and providing exact emotional context in sentiment analysis. Tags like hashtags provide extra context and focus on key subjects, improving the analysis of sentiments, whereas URLs are often unrelated to the sentiment and are removed to reduce noise and focus on the actual text data.

Example:

- Original: "I love ChatGPT! Check out this link: `http://example.com` #ChatGPT #Education"
- Processed: "I love ChatGPT!"

Thus, the handled review data can be represented as:

$$D_{clean} = \{d_{clean,1}, d_{clean,2}, \ldots, d_{clean,H}\}$$

Here, $h = 1 \to H$, $d_{clean,h}$ represents the emoticons, tags, and URLs handled data from the original review data, and H depicts the total number of handled data. Then, these data are input to the next step, which is pre-processing.

4.3 Pre-processing

In this step, the input text data T_{in} and the handled data D_{clean} are inputted into the pre-processing phase for enhancing the quality and consistency of sentiment analysis. Pre-processing includes sentence splitting, special character removal, abbreviation and contraction handling, case conversion, spell checking, and text normalization.

4.3.1 Sentence splitting

Initially, the data from the input review data T_{in} and handled data D_{clean} are given as input to the sentence splitting, which splits the individual sentence from the paragraph of the text reviews. This process is also important for sentiment analysis to analyze the sentence separately. Also, the sentiment of the specific parts of the text is accurately captured to provide more detailed information by splitting the text into sentences.

Example:

- Original: "ChatGPT is great. It helps a lot."
- Split:
  - "ChatGPT is great."
  - "It helps a lot."

This can be explained as,

$$T_{split} = \text{sentence-splitting}(T_{in}, D_{clean}) = T_{split}$$

Here, T_{split} indicates the sentence split text and $T_{split}(T_{in}, D_{clean})$ depicts the collection of data from the input review text T_{in} and handled data D_{clean}. Thus, the split sentences T_{split} are then processed for a further approach to accurate sentiment analysis.

4.3.2 Special character removal

After splitting the text into individual sentences T_{split}, the presence of special characters is analyzed and removed from each individual sentence. This step identifies and eliminates special characters, such as non-alphanumeric characters, punctuation marks, and other symbols.

Example:

- Original: "ChatGPT is great! @everyone #AI"
- Processed: "ChatGPT is great everyone AI"

Thus, the special characters are analyzed and removed to enhance the accuracy of sentiment analysis, which is explained as,

$$T_{sc\_removed} = T_{split} - SC(T_{split})$$

Where, $T_{sc\_removed}$ illustrates the special character removed text or data and $SC(T_{split})$ illustrates the special characters, which are present in T_{split}. Thus, the special character removal process improves the accuracy of sentiment analysis by removing the irrelevant characters. Also, the text after removing the special characters $T_{sc\_removed}$ is processed for further pre-processing steps.

4.3.3 Abbreviations and contraction handling

The special character-removed texts $T_{sc\_removed}$ are given as input to the abbreviations and contraction handling phase. In this phase of text pre-processing for sentiment analysis, the standardization of textual representations is focused on ensuring consistency and accuracy in sentiment interpretations. Abbreviations, which are shortened forms of words or phrases (e.g., “Dr.” for “Doctor”), are expanded, and contractions, which combine words by omitting certain letters (e.g., “can’t” for “cannot”), are normalized to maintain consistency. Owing to the utilization of these abbreviations and contractions in the sentence or text, there are several issues based on analyzing the sentiments accurately. Here, abbreviations need to be expanded to ensure accurate interpretations, whereas the contractions are normalized to maintain consistency.

Example:

- Original: "Prof. Johnson can’t imagine a class without ChatGPT."
- Processed: "Professor Johnson cannot imagine a class without ChatGPT."

These standardizations helped to improve the accuracy of sentiment classification, which is illustrated as,

$$T_{abbrev} = \Gamma_{fun}^{ACH}(T_{sc\_removed})$$

Here, T_{abbrev} indicates the abbreviations and contractions handled data or text and Γ_{fun}^{ACH} represents the function that handles the expansion and contractions. Thus, the handled data are applied to further processes for accurately analyzing the sentiments.

4.3.4 Case conversion

After expanding the abbreviations and normalizing the contractions from the review data T_{abbrev}, the case conversion process is performed. This phase converts all text data to lowercase to facilitate consistent processing for accurate sentiment analysis.

Example:

- Original: "ChatGPT is Great."
- Processed: "chatgpt is great."

This can be illustrated as,

$$T_{abbrev} \Rightarrow \text{lowercase}(T_{abbrev}) = T_{case}$$

Where, T_{case} indicates the case converted text. This conversion removes the variations that are caused by the inconsistent capitalization by standardizing the text. After the case conversion, the case-converted texts are fed to the next step for further analysis.

4.3.5 Spell checking

Now, the spellings of each word are checked by analyzing the case-converted text T_{case} from the previous stage. Spell check is the process of identifying and correcting spelling errors in textual data by comparing the text words with the dictionary or language model. Hence, the identification and correction of misspelled words help to improve the quality by ensuring the correct spelled words.

Example:

- Original: "ChatGPT is greaat."
- Processed: "ChatGPT is great."

This can be depicted as,

$$T_{spell} \Rightarrow T_{case}\,\Gamma_{fun}^{Spell}$$

Where, T_{spell} depicts the spell-checked text and review data and Γ_{fun}^{Spell} demonstrates the function that performs spell-checking. Thus, this step is important in sentiment analysis for preventing misinterpretations, which are caused by typing errors or misspellings. Also, integrating the spell check with the text pre-processing ensures the standardization of the input data. Then, the spell-checked sentence or text is provided to the next and final stage of pre-processing called text normalization.

4.3.6 Text normalization

Now, the spell-checked texts T_{spell} are converted into a similar format by using text normalization. In this phase, the spell-checked texts are transformed into a standard format for enhancing the consistency and accuracy of sentiment analysis. It reduces the complexity and variability of the text data. This function replaces multiple spaces with single space, and strips the leading and trailing spaces.

Example:

- Original: "ChatGPT is great. "
- Processed: "ChatGPT is great."

Text normalization is important in sentiment analysis for providing accurate sentiment classification based on approaching similar text expressions equally, which is examined by the following expression.

$$T_{spell} \xrightarrow{\text{Text Normalization}} T_{norm}$$

Where, T_{norm} represents the normalized text. Then, the normalized text is also known as the pre-processed data by performing several stages, such as sentence splitting, special character removal, abbreviations and contraction handling, case conversion, spell checking, and text normalization. By applying the above six steps, this proposed methodology receives a review text or data that has been pre-processed to guarantee the sentiment analysis's accuracy, which is shown as,

$$T_{preproc} = \{T_{preproc,1}, T_{preproc,2}, T_{preproc,3}, \ldots, T_{preproc,pp}\}$$

Here, $P = 1 \rightarrow pp$, $T_{preproc}$ illustrates the pre-processed data, and pp indicates the total number of pre-processed data. Then, the pre-processed data or text $T_{preproc}$ is provided as input to detect the sarcastic text in the present text.

4.4 Sarcastic Text Detection

After pre-processing each and every individual text or data for accurate analysis of sentiments, the sarcastic text is detected and removed. Finding and detecting text that uses sarcasm is known as sarcastic text detection. Here, by utilizing the XL-Net approach with hyper-parameter, the sarcastic texts are detected and then the detected sarcastic texts are removed. The eXtreme Learning Network (XLNet) uses

a bidirectional context modeling approach and optimizes both left and right context during training. This approach also allows the model to capture the complex dependencies and contextual nuances in the text, which are important for detecting sarcasm. The Hyper-band parameter tuning is also employed in the prevailing XLNet, which strikes a balance between exploration and exploitation by dynamically allocating resources to configurations based on their performance. Also, it efficiently identifies and focuses computational resources on the most promising configurations while exploring a diverse set of hyper-parameter combinations. The dataset used for training the sarcasm detection model is taken from kaggle [62].
It is a high-quality dataset specifically designed for the tasks of sarcasm and fake news detection. It comprises news headlines labeled as sarcastic or non-sarcastic, providing a robust foundation for training and evaluating models aimed at detecting sarcasm in textual data.

Example:

- Original: "Oh great, another useless AI tool."
- Detected and Removed: "Oh great, another useless AI tool." - (Identified as sarcastic)
- Original: "Yeah, ChatGPT was a complete waste of time, said no one ever."
- Detected and Removed: "Yeah, ChatGPT was a complete waste of time, said no one ever." - (Identified as sarcastic)

4.4.1 Input layer

Initially, the review text from the pre-processed data $T_{preproc}$ is accepted by the input layer, and this layer represents the input review text as a vector, which contains features extracted from the pre-processed text $T_{preproc}$. Also, each element in the input review text corresponds to a specific characteristic of the input review data. This can be illustrated by the following representation.

$$T_{preproc} = v[P]$$

Where, $v[P]$ indicates the vector of each element in the pre-processed data $T_{preproc}$. Then, the input layer passes this vector $v[P]$ to the hidden layer for further processing, which is depicted as,

$$I = v[P]$$

Here, I represents the input layer.

4.4.2 Hidden layer

This layer accepts the input layer's output I and performs it with a single hidden layer. Here, the weights and biases are randomly initialized by determining the number of hidden neurons to avoid iterative adjustments during training. Then, the random initialization of weights and biases are depicted as follows,

$$W = Winit[F, Hneurons]$$

$$b = Winit[Hneurons, 1]$$

Where, W depicts the weights, b represents the bias, $Winit$ indicates the random initialization, F illustrates the number of features in the input review text, and $Hneurons$ portrays the number of neurons in the hidden layer. Moreover, each neuron F in this layer is related to weights and biases for enabling the transformation of input features through an activation function to capture the intricate relationships within the review input data for accurate sentiment analysis. This can be elaborated as,

$$H = \phi^{act}[W \cdot v + b]$$

Here, H illustrates the hidden layer, and ϕ^{act} indicates the activation function, which introduces non-linearity for capturing the intricate relationship among the input text data. Then, the output of the hidden layer H is provided to the next layer, which is called the output layer to generate the output matrix for accurate sentiment analysis. Now, the hyperband hyper-parameter tuning technique is applied to prevent premature convergence for suboptimal solutions in analysis sentiments accurately. Then, the step-by-step process for the hyperband hyper-parameter tuning technique is illustrated as follows:

- Initially, the search space, such as the number of hidden neurons and the type of activation function, for the range of hyper-parameters tuning is defined and illustrated as,

  $$N = \{Hmax, Hmin\}$$

  $$\phi^{act} = \{ReLU\}$$

  Where, $Hmax$ and $Hmin$ indicate the maximum and minimum number of hidden neurons present in the hidden layer's output, respectively.

- Then, the total computation budget and the maximum resources are examined to initialize the hyperband parameters, which are depicted as follows,

  $$Bcomp = Rmax \log(\ell)$$

  Here, $Bcomp$ depicts the computation budget, $Rmax$ illustrates the maximum resources, and ℓ represents the reduction factor.

- Now, from the defined search space, the set of hyper-parameter configurations is sampled randomly for initiating the optimization process by the following representation.

  $$\Theta^i_{HC} = \{\Theta^{HC}_1, \Theta^{HC}_2, \Theta^{HC}_3, \ldots, \Theta^{HC}_i\}$$

  Where, Θ^i_{HC} represents the hyper-parameter configurations, and i indicates the total number of hyper-parameter configurations.

- Then, the resources are allocated to the configuration by using the successive halving approach, which is illustrated as,

$$R_\Theta = \frac{Bcomp}{\ell}\left(\frac{1}{i}\right)$$

  Here, R_Θ depicts the allocated resources to the configurations. Next, each configuration is evaluated to compute the loss. Then, based on the loss values, the top configurations are selected. Finally, more resources are allocated for the selection of top configurations to tune the hyper-parameters. Thus, the hidden layer after hyper-parameter tuning is expressed by

$$H_{tuned}$$

4.4.3 Output layer

In this layer, the output layer weights are calculated, which is done by using the Moore-Penrose generalized inverse of the hidden layer output matrix H_{tuned}, which is depicted as,

$$W_{out} = H_{tuned}^{\dagger} \cdot O$$

Where, W_{out} depicts the output weights, $H_{tuned}^{\dagger}$ represents the invertible form of the hidden layer's output, and O indicates the output predictions for sentiment analysis. Then, the new output predictions are evaluated by the multiplication of hidden layer output and the output weights, which are elaborated as follows,

$$O = H_{tuned} \cdot W_{out}$$

Based on new output predictions, the sarcastic texts are identified and deleted, thus the final texts after the removal of sarcastic text by using XLHN are depicted as follows,

$$T_{no\_sarcasm} = \{\hat{\lambda}_{a1}^{a}, \hat{\lambda}_{a2}^{a}, \hat{\lambda}_{a3}^{a}, \ldots, \hat{\lambda}_{aN_{no\_sarcasm}}^{a}\}$$

Here, $N_{no\_sarcasm}$ depicts the total number of texts without sarcastic texts, and $T_{no\_sarcasm}$ represents the sarcastic text removed data. Thus, the removed sarcastic texts are given as input to the next process, which is known as the sentence grouping.

4.5 Sentence Grouping

From the sarcastic text removed data $Tno\_sarcasm$, the sentences are grouped using the hierarchical clustering technique based on the similarity of the data. This process involves organizing and categorizing the sentences into comprehensible groups based on their semantic meaning, context, or relationship to one another.

In hierarchical clustering, there is no need to specify the number of clusters in advance. Each sentence from the sarcastic text removed data $Tno\_sarcasm$ is iteratively merged into larger clusters based on their similarity, creating a dendrogram that visually represents the hierarchy of clusters. This dendrogram can be cut at various levels to form different numbers of clusters, ensuring that each sentence is grouped with others that have similar semantic content. This clustering process helps in effective sentence grouping by categorizing the sentences into hierarchical clusters, where each cluster represents a group of sentences with similar characteristics.

Finally, the clustered sentences are represented as the grouped sentence, which is illustrated as $Sgroup$. Here, $G = 1 \to g$, $Sgroup$ depicts the grouped sentences or data, and $Ngroup$ represents the total number of grouped sentences. Thus, the grouped sentences enhance accuracy by clustering similar sentences for the identification of refined sentiments within the groups. These grouped sentences are then provided to the next stage of effective sentiment analysis.

4.6 Topic Modeling

Now, topics are modeled by accepting the grouped sentences S_{group} as input, which is done by using the NTDMF (Non-negative-based Term-Document Matrix Factorization) technique. Topic modeling is one of the text mining techniques, which is utilized for discovering the summary topics by identifying the patterns in the text to cluster words into topics. Here, the methodology utilized the NTDMF approach for modeling the topics from the grouped sentences for effective sentiment analysis. Here, the rows correspond to the term and the columns correspond to the documents for effective topic modeling. Moreover, each cell in the matrix represents the frequency of occurrence of a particular term in a particular document. Thus, the steps involved in the proposed NTDMF approach based on topic modeling are depicted as follows:

- Initially, the grouped sentences S_{group} are represented as the term-document matrix, where the row corresponds to terms and columns correspond to the document. Here, terms indicate the words from the grouped sentences and the document represents the grouped sentences, which is elaborated as,

$$TDM \Rightarrow TDM_{xy} = \begin{bmatrix} TDM_{11} & TDM_{12} & \cdots & TDM_{1D} \\ TDM_{21} & TDM_{22} & \cdots & TDM_{2D} \\ \vdots & \vdots & \ddots & \vdots \\ TDM_{x1} & TDM_{x2} & \cdots & TDM_{xD} \end{bmatrix}$$

  Where, TDM indicates the term-document matrix, and TDM_{xy} illustrates the frequency of the term x in the document y.

- Next, the two non-negative matrices, such as the topic-term matrix and the document-topic matrix, are initialized based on the term-document matrix factorization technique, which is represented by the following expressions,

$$TTM \in TTM_{xy} = \begin{bmatrix} TTM_{11} & TTM_{12} & \cdots & TTM_{1K} \\ TTM_{21} & TTM_{22} & \cdots & TTM_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ TTM_{x1} & TTM_{x2} & \cdots & TTM_{xK} \end{bmatrix}$$

$$DTM \in DTM_{xy} = \begin{bmatrix} DTM_{11} & DTM_{12} & \cdots & DTM_{1Y} \\ DTM_{21} & DTM_{22} & \cdots & DTM_{2Y} \\ \vdots & \vdots & \ddots & \vdots \\ DTM_{x1} & DTM_{x2} & \cdots & DTM_{xY} \end{bmatrix}$$

  Here, TTM indicates the topic-term matrix, DTM represents the document-topic matrix.

- Now, matrix factorization is performed, in which the topic-term matrix TTM and the document-term matrix DTM are multiplied to get the approximate of the term-document matrix TDM, which is explained as follows,

$$TDM \approx TTM \cdot DTM$$

- Then, the element-wise multiplication is applied to update the topic-term matrix TTM and the document-term matrix DTM. By applying this multiplication method, the reconstruction errors are minimized, and the following representation illustrates the element-wise multiplication for updating the topic-term matrix and the document-topic matrix.

$$\widetilde{TTM} = \left(\frac{(DTM)^T TDM}{(DTM)^T DTM \cdot TTM} \right) \odot TTM$$

$$\widetilde{DTM} = \left(\frac{(TTM) TDM}{(TTM) TTM^T \cdot DTM} \right) \odot DTM$$

  Where, $\widetilde{TTM}$ and $\widetilde{DTM}$ indicates the updated topic-term matrix and document-topic matrix, respectively, T^T illustrates the transpose matrix and $\odot$ represents the element-wise multiplication operator. Then, the steps are repeated until they converge, and the modeled topics are obtained by updating the topic-term matrix and document-term matrix. Thus, the modeled topics are illustrated as follows,

$$T^{topics}_{modeled} = \{T^{topics}_{modeled,1}, T^{topics}_{modeled,2}, T^{topics}_{modeled,3}, \ldots, T^{topics}_{modeled,n}\}$$

Here, $m = 1 \to mm$, $T^{topics}_{modeled}$ illustrated the topics, which are modeled by using the proposed NTDMF, and mm depicts the total number of modeled topics. Then, the modeled topics are provided to the two distinct processes, such as content filtering and dependency parsing, for the sentiment analysis based on the ChatGPT adoption for higher education.

4.7 Data Transformations

The pre-processed review text $T_{preproc}$ undergoes several preprocessing steps to enhance the data quality for sentiment classification. Firstly, dependency parsing is employed to analyze the grammatical structure of sentences, which helps in understanding the relationships between words. Co-reference resolution is then applied to identify and link pronouns and other referring expressions to the appropriate entities within the text. Following this, word embedding T_{embed} is used to transform words into continuous vector representations that capture semantic meanings. Content filtering is performed to remove irrelevant or redundant information, ensuring that the data remains focused on the essential aspects. Subsequently, feature extraction $F_{extract}$ is conducted to derive meaningful features from the text, capturing both syntactic and semantic properties. The output of feature extraction, denoted as $F_{extract}$, is a crucial input for the classification stage, where it will be utilized to classify sentiments accurately.

Example

- Original: "ChatGPT helps students understand complex topics."
- Dependency Parsed: "[ChatGPT] (subject) [helps] (verb) [students] (object) [understand complex topics] (object complement)"
- Original: "Dr. Smith said ChatGPT is useful. He recommends it to all his students."
- Coreference Resolved: "Dr. Smith said ChatGPT is useful. Dr. Smith recommends ChatGPT to all Dr. Smith's students."

4.8 Classification

In this step, the output from the pre-processing $T_{preproc}$, word embedding T_{embed}, and feature extraction $F_{extract}$ is given as a source to the classification stage for the classification of sentiments in terms of analyzing positive, negative, and neutral contents. Here, the HR-RAN approach is applied to classify positive, negative, and neutral sentiments. The prevailing Hierarchical Attention Network (HAN) has the ability to capture the hierarchical structures in the text data to focus on different levels of granularity, thus capturing both local and global context in the text. Also, it has the capability to weigh the importance of each word within a sentence dynamically by allowing it to focus on the most relevant information for sentiment classification. But, this model is struggling with several limitations, such as vanishing gradients or difficulty in capturing the long-range dependencies mainly in deep hierarchical structures. Hence, to conquer these limitations, the Residual connection with the Rectified Sigmoid Exponential Linear Unit (RSigELU) activation function is applied. Here, residual connections are employed to learn residual mappings rather than trying to learn the desired mapping directly. Thus, this helps to overcome the vanishing gradient problem by allowing the gradients to flow directly through the identity mappings. Moreover, the RSigELU function is smooth everywhere, including the point where it transitions from linear to exponential behavior, thus potentially leading to faster convergence. The classifier diagram for the proposed HR-RAN is depicted in Figure 5.7.

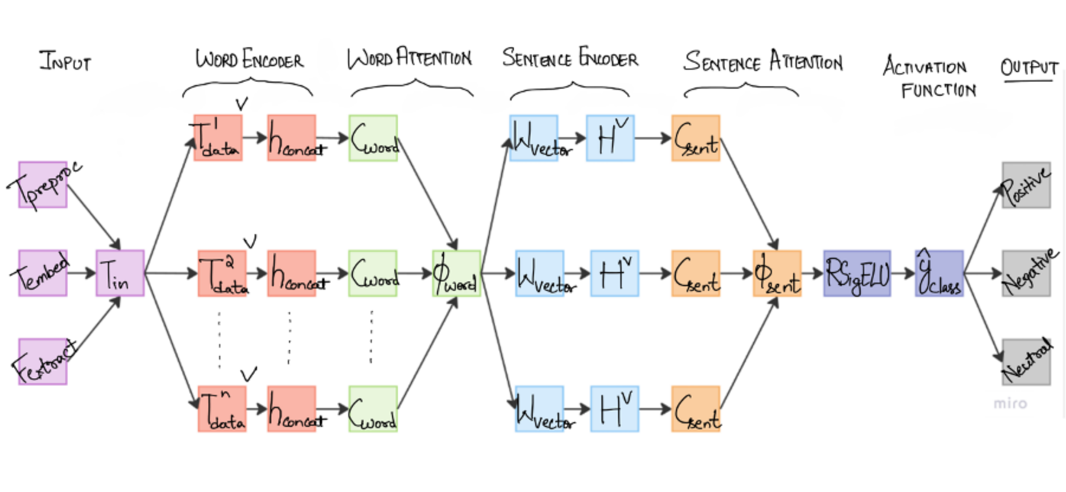


Figure 4.2: Classifier diagram of HR-RAN

Initially, the proposed classifier accepts the input from the pre-processing $T_{preproc}$, word embedding T_{embed}, and feature extraction $F_{extract}$ for the accurate classification of positive, negative, and neutral sentiments, and is represented by the following equation,

$$\begin{pmatrix} T_{preproc} \\ T_{embed} \\ F_{extract} \end{pmatrix} = T_{in}$$

Where, T_{in} depicts the text data, which are collected from pre-processing, word embedding, and feature extraction. Then, the step involved in the proposed HR-RAN algorithm is elaborated as follows.

4.8.1 Word level processing

Step 1: Firstly, the word in the input sentence T_{in} is represented as tokens, which are expressed as T^w_{split} where w indicates the words in the sentence T_{in}.

Step 2: Then, the embedding layer is applied to each token T^w_{split} for representation as a vector, which is depicted by the following representation.

$$v = E[T^w_{split}]$$

Here, v indicates the vector representation.

Step 3: Next, the Gated Recurrent Unit (GRU) is applied for encoding the word sequences bi-directionally. Here, the bidirectional GRU is employed for getting the annotations of words based on summarizing the information from both directions like forward and backward, and these directions are computed at time step η by the following equations,

$$h^{\eta}_{\text{fwd}} = \text{GRU}\left(v, h^{\eta-1}_{\text{fwd}}\right)$$

$$h^{\eta}_{\text{bwd}} = \text{GRU}\left(v, h^{\eta+1}_{\text{bwd}}\right)$$

Where, h^{η}_{fwd} and h^{η}_{bwd} indicate the forward and backward direction of the hidden state in GRU, respectively. Then, both forward and backward states are combined to form the concatenation of word representation, which is depicted as follows,

$$h_{\text{concat}} = \left[h^{\eta}_{\text{fwd}}, h^{\eta}_{\text{bwd}}\right]$$

Here, h_{concat} represents the concatenation of both directions.

Step 4: Now, the attention mechanism is employed for extracting the words that are essential for the meaning of the sentence and collecting the representation of those informative words to provide a sentence vector. This can be illustrated by the following representations,

$$c_{\text{word}} = \tanh\left(h^{\eta}_{\text{concat}} + c^{w}\right)$$

Where, c_{word} indicates the hidden representation of h_{concat}, c^{w} illustrates the trainable parameters. Based on the similarity of the hidden representation, the importance of the word is measured with a word-level context vector for getting the normalized importance weight, which is demonstrated as,

$$\alpha^{w} = \frac{\exp\left(\left(c_{\text{word}}\right)^{T} K^{w}\right)}{\sum_{\eta} \exp\left(\left(c_{\text{word}}\right)^{T} K^{w}\right)}$$

Here, α^{w} illustrates the normalized weights, and K^{w} indicates the word-level context vector. Next, the word vector is computed based on the weighted sum of the word annotations, which is depicted as,

$$\Phi^{\text{word}} = \sum_{\eta} \alpha^{w} h^{\eta}_{\text{concat}}$$

Where, Φ^{word} represents the word vector

4.8.2 Sentence level processing

Step 1: Similarly, the bi-directional GRU is applied to represent the sentence vectors by accepting the input from the previous level processing called word vector Φ_{word} which is depicted by the following representation,

$$H^{\eta}_{\text{fwd}} = \left[\text{GRU}_{\text{fwd}}\left(\Phi_{\text{word}}, v, \text{GRU}_{\text{fwd}}\left(H^{\eta-1}_{\text{fwd}}\right)\right)\right]$$

$$H^{\eta}_{\text{bwd}} = \left[\text{GRU}_{\text{bwd}}\left(\Phi_{\text{word}}, v, \text{GRU}_{\text{bwd}}\left(H^{\eta+1}_{\text{bwd}}\right)\right)\right]$$

Step 2: Now, the importance of weight for each sentence is represented by computing the attention mechanism, which is elaborated as follows,

$$\alpha^{s} = \tanh\left(H_{\text{concat}} + c^{s}\right)$$

Here, H_{concat} depicts the concatenation of the sentence representation, K^{s} illustrates the hidden representation of H_{concat}, α^{s} and c^{s} indicates the trainable parameters in sentence representation. Next, the sentence vector is evaluated based on the weighted sum of the sentence representation, which is illustrated by the following equation,

$$\Phi_{\text{sent}} = \sum_{\eta} \alpha^s H$$

Where, α^s indicates the normalized weight calculation in sentence level processing, and Φ_{sent} depicts the sentence vector.

Step 3: Finally, the RSigELU activation function is employed for each sentence vector to represent the features for the classification of sentiments, which is depicted by the following mathematical equation,

$$\hat{y} = \text{RSigELU}\left(w \cdot \Phi_{\text{sent}} + \lambda\right)$$

$$\text{RSigELU} = \begin{cases} \Phi_{\text{sent}} & \text{if } \Phi_{\text{sent}} > 0 \\ \sigma\left(\Phi_{\text{sent}}\right) - 1 & \text{if } \Phi_{\text{sent}} \leq 0 \end{cases}$$

Here, $\hat{y}$ depicts the predicted probability distribution for sentiment classification, λ indicates the hyper-parameter, and σ depicts the sigmoid function.

Step 4: Based on the represented features, which are present in the predicted probability distribution $\hat{y}$, this proposed method classifies the positive, negative, and neutral sentiments. It is depicted by the following mathematical equation, as

$$\hat{y} = \begin{cases} \hat{y}^{POS} & \text{positive sentiments} \\ \hat{y}^{NEG} & \text{negative sentiments} \\ \hat{y}^{NEU} & \text{neutral sentiments} \end{cases}$$

Here, $\hat{y}^{POS}$, $\hat{y}^{NEG}$, and $\hat{y}^{NEU}$ illustrate the positive, negative, and neutral sentiments, respectively. The pseudo-code for the classification task using HR-RAN approach is shown below

Algorithm 1: Pseudo-code for the proposed HR-RAN

Input: T_{in}, (W, b, σ, λ)
Output: Classified sentiments $\hat{y}_{class}$
Initialize T_{in}, v, h_{fwd}, h_{bwd}, h_{concat}, h_{hid}, c_{word}, s_{vector}, λ and maximum iteration ITR_{max}
Set $ITR = 1$
while $ITR \leq ITR_{max}$ **do**
  foreach T_{in} **do**
    `// Word Level Processing`
    Tokenize input sentence T_{in} into T^{w}_{split}
    Estimate $v = E[T^{w}_{split}]$
    Apply bi-directional GRU
    $h^{\eta}_{\text{fwd}} = \text{GRU}(v, h^{\eta-1}_{\text{fwd}})$
    $h^{\eta}_{\text{bwd}} = \text{GRU}(v, h^{\eta+1}_{\text{bwd}})$
    Concatenate forward and backward states
    $h_{\text{concat}} = [h^{\eta}_{\text{fwd}}, h^{\eta}_{\text{bwd}}]$
    Apply attention mechanism
    $c_{\text{word}} = \tanh(h^{\eta}_{\text{concat}} + c^{w})$
    Calculate normalized weights
    $\alpha^{w} = \frac{\exp((c_{\text{word}})^{T} K^{w})}{\sum_{\eta} \exp((c_{\text{word}})^{T} K^{w})}$
    Compute word vector
    $\Phi^{\text{word}} = \sum_{\eta} \alpha^{w} h^{\eta}_{\text{concat}}$
    `// Sentence Level Processing`
    Apply bi-directional GRU on Φ_{word}
    $H^{\eta}_{\text{fwd}} = \text{GRU}_{\text{fwd}}(\Phi_{\text{word}}, v, \text{GRU}_{\text{fwd}}(H^{\eta-1}_{\text{fwd}}))$
    $H^{\eta}_{\text{bwd}} = \text{GRU}_{\text{bwd}}(\Phi_{\text{word}}, v, \text{GRU}_{\text{bwd}}(H^{\eta+1}_{\text{bwd}}))$
    Compute sentence vector using attention mechanism
    $\alpha^{s} = \tanh(H_{\text{concat}} + c^{s})$
    Calculate sentence vector
    $\Phi_{\text{sent}} = \sum_{\eta} \alpha^{s} H$
    `// Classification`
    Apply RSigELU activation function
    if $\Phi_{sent} > 0$ **then**
      $\text{RSigELU} = \lambda - s_{\text{vector}}$
    else
      $\text{RSigELU} = \sigma(\Phi_{\text{sent}}) - 1$
    Predict the probability distribution
    $\hat{y}_{class} = \text{RSigELU}(w \cdot \Phi_{\text{sent}} + \lambda)$
return $\hat{y}^{POS}_{class}$: positive sentiments
return $\hat{y}^{NEG}_{class}$: negative sentiments
return $\hat{y}^{NEU}_{class}$: neutral sentiments

4.9 Degree of Sentiment Analysis

From the classification of sentiments , the fuzzy logic is applied for the identification of the degree of sentiments for the effective analysis of sentiments on ChatGPT adoption in higher education. Fuzzy logic enables the classification of data into multiple linguistic terms that provide a more nuanced understanding of uncertainty. Fuzzy rule-based classifiers can capture the complex relationships among the input features and class labels. The triangular polynomial (Tripo) membership function is applied in the fuzzy rule. Here, the degree of sentiment is analyzed based on the extracted features from the classified sentiments and the polarity score of the sentences. Thus, the steps involved in the fuzzy logic are explained in detail as follows:

- Initially, the classified sentiment features $\hat{y}^{class}$ are converted into the fuzzy set for the accurate analysis of the degree of sentiments by using the Triangular-Poly membership function, which is illustrated as follows,

  $$\hat{y}^{class} \xrightarrow{\mu_{Tripo}} F_{set}$$

  $$\mu_{Tripo}(\hat{y}^{class})_{k,a,b,c} = \begin{cases} 0 & \text{if } \hat{y}^{class} \leq a \\ k(\hat{y}^{class} - a)(b - \hat{y}^{class}) & \text{if } a < \hat{y}^{class} \leq b \\ k(c - \hat{y}^{class})(c - b) & \text{if } b < \hat{y}^{class} < c \\ 0 & \text{if } \hat{y}^{class} \geq c \end{cases}$$

  Where, μ_{Tripo} illustrates the Tripo membership function, F_{set} depicts the fuzzy set, k indicates the scaling constant, and a, b, and c represent the upper limit, peak value, and lower limit of the membership function, respectively.

- Based on the input field with respect to each rule, the matching degree of the fuzzy set F_{set} is to be fired, and it is demonstrated by the following representation,

  $$[\mu_{match}] = m(F_{set} \cdot f_{strength})$$

  Here, $[\mu_{match}]$ indicates the matching degree of the fuzzy set along with the rule ∇ and $f_{strength}$ represents the firing strength.

- Then, the firing strengths are combined to form the control actions, and the overall control actions are calculated by the following equation,

  $$\nabla^{\Delta} \rightarrow R_{combined}$$

  Where, ∇^{Δ} depicts the combined rule and $R_{combined}$ illustrates the overall control actions.

- Finally, the overall control actions $R_{combined}$ are converted into crisp values for analyzing the degree of sentiments by using the following representation,

$$C_{crisp} = \frac{\sum_{w=1}^{n} [\mu_{match,w}] \cdot c_{centroid}}{\sum_{w=1}^{n} [\mu_{match,w}]}$$

  Here, C_{crisp} depicts the crisp value for analyzing the sentiments in degree and $c_{centroid}$ illustrates the centroid of the fuzzy set. Then, the final output for the degree of sentiment analysis is done based on the crisp value C_{crisp}, which also indicates the polarity value. This can be depicted as follows,

$$S_{degree} = \begin{cases} \text{if } C_{crisp} == \text{positive}\& \max(C_{crisp})^{\text{new}} & : \text{then, } P_{pos} \\ \text{if } C_{crisp} == \text{negative}\& \max(C_{crisp})^{\text{new}} & : \text{then, } P_{neg} \\ \text{if } C_{crisp} == \text{neutral}\& \max(C_{crisp})^{\text{new}} & : \text{then, } P_{neut} \end{cases}$$

  Where, S_{degree} illustrates the final output, which represents the degree of sentiments, $\max(C_{crisp})^{\text{new}}$ indicates the maximum polarity word count, P_{pos}, P_{neg} and P_{neut} depict the percentage of the degree of positive, negative, and neutral sentiment analysis, respectively.

Thus, the proposed framework for the sentiment analysis regarding ChatGPT adoption for higher education from various university review data using HR-RAN, which accurately classifies the positive, negative, and neutral sentiments and effectively predicts the degree of those classified sentiments.

5 Results and Discussions

This chapter highlights a comprehensive analysis of the study, highlighting the key findings from various analytical methods applied. The results are structured to provide a clear overview of the sentiments of universities regarding ChatGPT integration into education, modeling efforts, and performance of the proposed sentiment analysis model. Each subsection delves into specific aspects of the analysis, offering detailed insights supported by visualizations and statistical metrics. The importance of these results lies in their contribution to understanding the sentiment and opinions expressed in the data and the effectiveness of the methods used to process and analyze this information. By examining the overall sentiment distribution which is then classified into positive, negative, and neutral, country-wise sentiment analysis, and sentiment scores by universities, this research aims to uncover trends that can inform decision-making and strategy development in related fields.

5.1 Interactive User Interface

The Graphical User Interface (GUI) was developed to streamline and simplify the sentiment analysis process. The GUI window is illustrated in the Figure 5.1. It depicts a software interface designed for conducting sentiment analysis and it outlines a series of steps derived from a proposed framework to ensure an efficient implementation process. It also details a series of interactive buttons that each trigger specific steps in the analysis process. Clicking on these buttons in sequence allows users to effectively manage the workflow of sentiment analysis from data input to result interpretation.

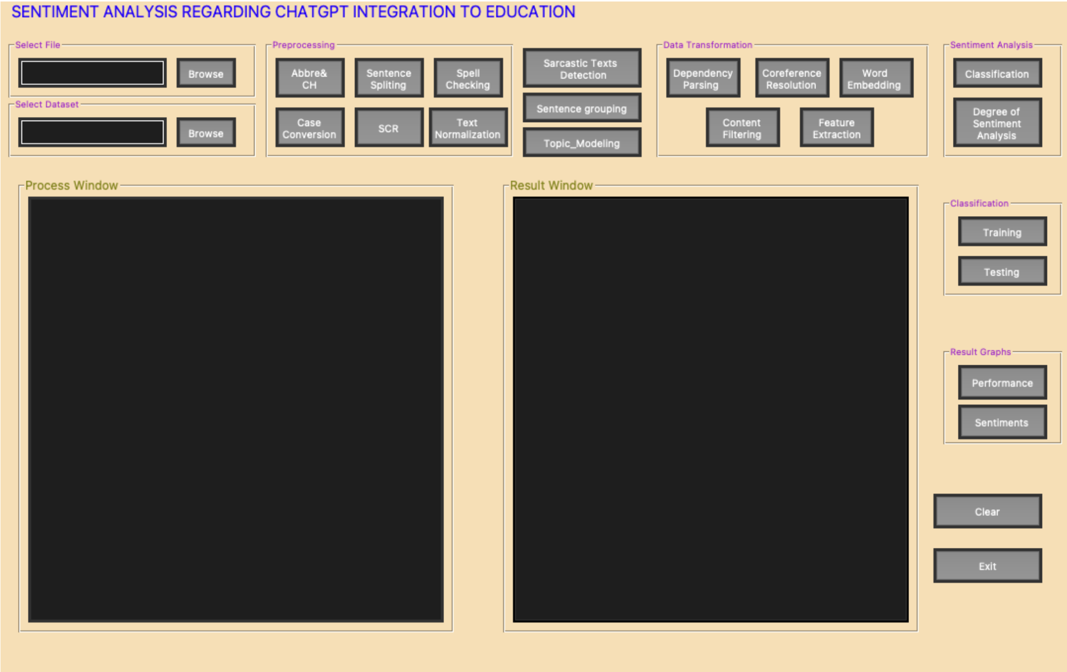


Figure 5.1: Graphical User Interface for Sentiment Analysis

This interface is used to both train and test the proposed HR-RAN model. The results of series of the steps are as follows

1. **File and Dataset Selection**: This stage involves selecting files or datasets for analysis. The select file button is used to select the collected data for analysis and the select dataset button is used to select the training data. The sample testing data is illustrated in Figure 5.2.

```
Original Text
I love ChatGPT! :-) says supervisors of university. Check out this link:  #ChatGPT #Education.
 Prof. Johnson says imagine a class without ChatGPT because he feels ChatGPT is   Grrrreeeat!!! .
Prof. johnson comments that for both EDUCATORS and         STUDENTS if used in   right way     it would be a wonderful tool.
"If someone questions against this opinion of the prof., he usually replies - Yeah, ChatGPT was a complete waste of time, said no one ever. "
"now, another supervior Dr. Smith said ChatGPT is useful. He recommends it to all his students.  @@"

Data was selected successfully...
```

Figure 5.2: Sample data before processing

2. **Pre-processing**: The pre-processing steps are essential for preparing the data for both training and testing phases. After text normalization, as illustrated in Figure 5.3, the resulting data is free from spelling errors, extra spaces, special characters, and has consistent case formatting.

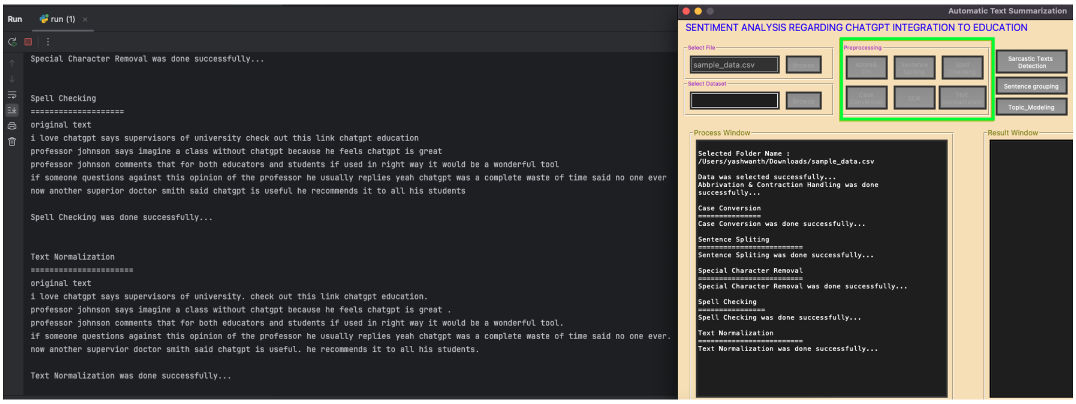


Figure 5.3: Output of data pre-processing

3. **Sarcastic text detection**: When the Sarcastic Text Detection button is clicked, the system analyzes the uploaded sample data to identify instances of sarcasm. For instance, as shown in Figure 5.4, a sarcastic comment like "yeah chatgpt was a complete waste of time said no one ever" in a review about a ChatGPT implementation might be incorrectly classified as negative without this detection step, which accurately labels it as sarcasm.

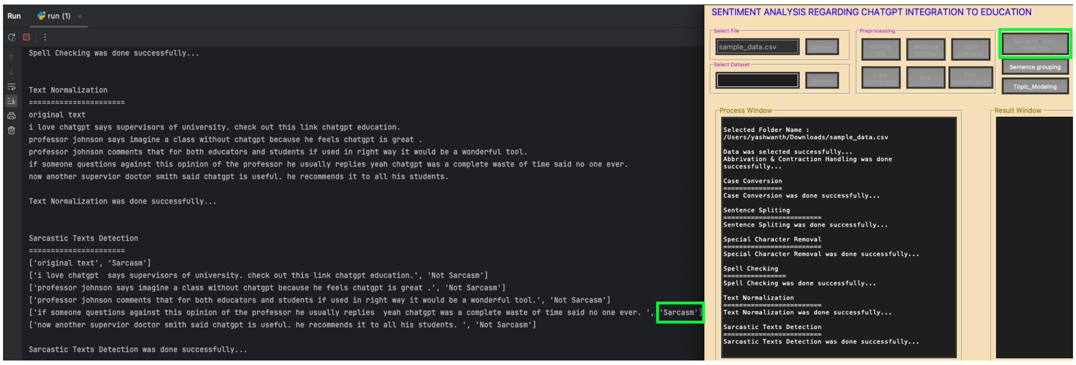


Figure 5.4: Output of sarcastic text detection

4. **Topic modelling**: Clicking the Topic Modeling button activates an algorithm that analyzes the output from grouped sentences to identify similar themes or topics within the sample data. For example, as shown in Figure 5.5, the topic modeling process reveals clusters of discussions and themes such as "students," "professor," or "ChatGPT."

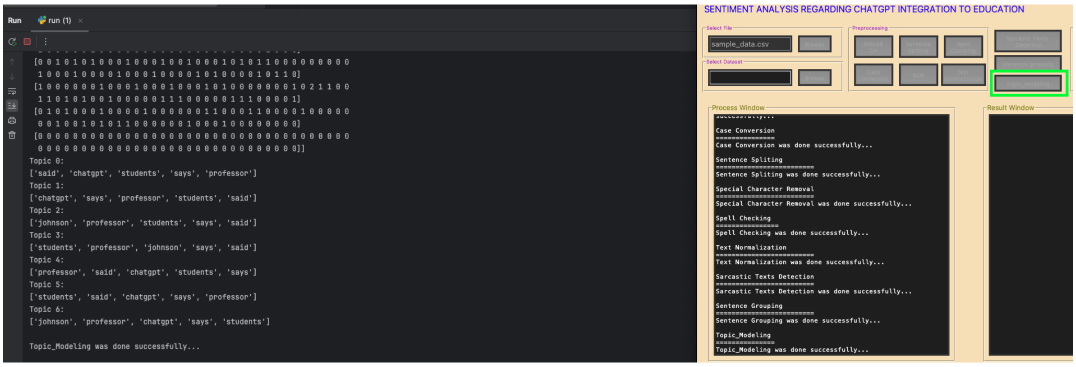


Figure 5.5: Topic modelling

5. **Data transformation**: This final step is essential for preparing the data for training and testing. After performing dependency parsing and coreference resolution, relationships within the text data are identified. The subsequent steps, such as word embedding and feature extraction, convert the data into a numerical format that captures semantic relationships between words. This transformation enhances the model's ability to identify and categorize sentiments. Figure 5.6 illustrates the numerical representation of embedded words and features used for training and testing the model.

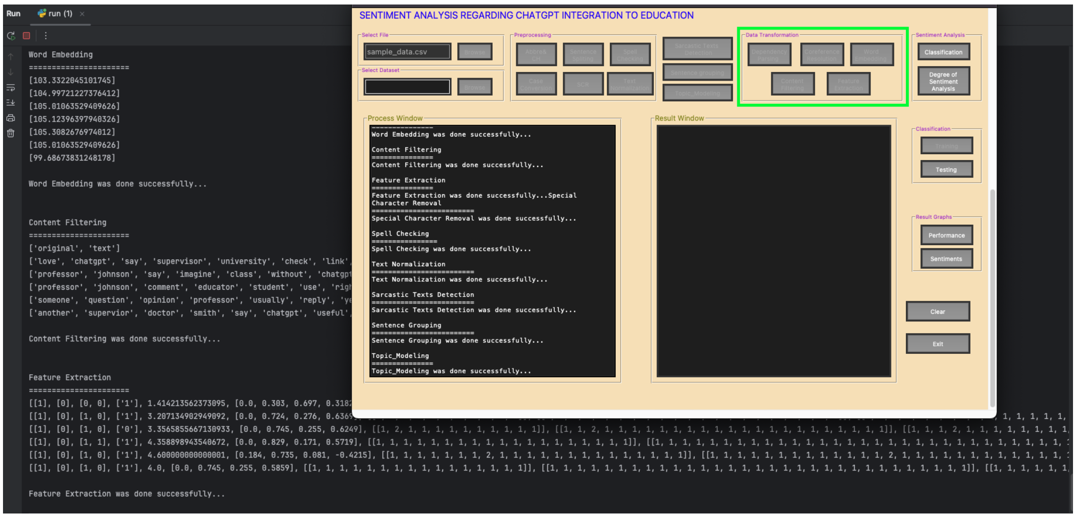


Figure 5.6: Data transformation technique

6. **Classification**:

During the training phase, the classification button is disabled because the models—HR-RAN, Hierarchical Attention Network (HAN), Convolutional Neural Network (CNN), Deep Neural Network (DNN), and Recurrent Neural

Network (RNN)—are trained on labeled data. In the testing phase, clicking the classify button performs classification, resulting in labels such as positive, negative, and neutral sentiments. Finally, the degree of sentiment analysis is calculated. The output is shown in Figure 5.7.

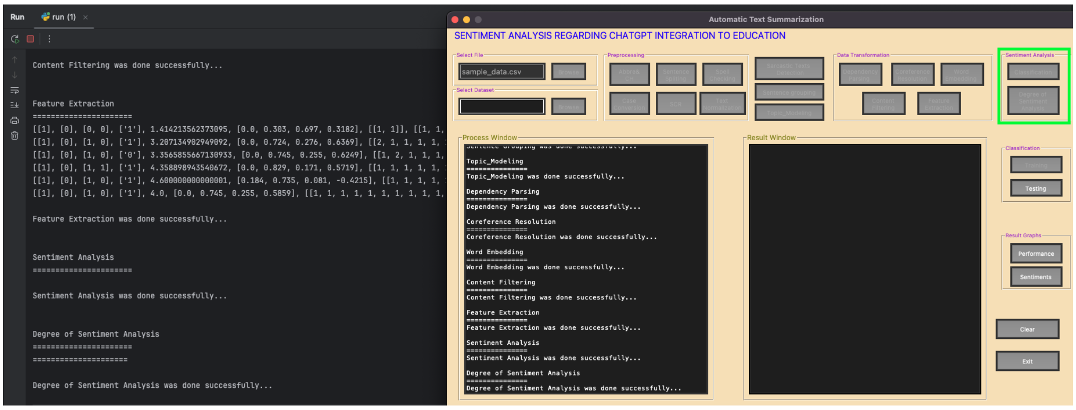


Figure 5.7: Classification and degree of sentiment analysis

7. **Training and Testing**: The training button is used to train the sentiment analysis models using the transformed data, providing outputs such as training time and parameter details. The testing phase evaluates the performance of sentiment analysis models like DNN, CNN, RNN, and HR-RAN using evaluation metrics such as accuracy, precision, recall, F-measure, and sensitivity.

8. **Performance Metrics**: Clicking this button generates graphs and plots for evaluation metrics like accuracy, precision, recall, and F-measure, offering enhanced visualization and comparison of results. This provides deeper insights into the models' effectiveness in sentiment classification.

9. **Sentiments Visualization**: Visual representations, such as graphs or charts, display the distribution and general trends of sentiments within the data, facilitating easier interpretation and presentation of results.

10. **Utility Functions**:
    - **Clear**: Resets all current selections and data entries in the interface, allowing users to start again with new data for analysis.
    - **Exit**: Closes the application, typically used after completing all required analyses.

5.2 Sentiment Analysis Results

The sentiment analysis conducted on the adoption of ChatGPT within higher education settings reveals a mostly neutral stance among the academic communities discussed. As illustrated in Figure 5.8, a substantial majority, accounting for 74.8% of the sentiments, are categorized as neutral. This suggests a cautious or observational approach of the universities towards the integration of ChatGPT into educational frameworks. The dominance of neutral sentiment indicates that while there is a recognition of the potential and presence of ChatGPT, there is still hesitation or insufficient experience to form a definitive positive or negative judgment.

In contrast, a smaller yet significant portion of the sentiments, about 21.5%, are positive. This positive sentiment underscores the benefits and useful applications of ChatGPT as perceived by some educators and students. It reflects acknowledgment of the tool's capability to enhance educational practices through support in tutoring, personalized learning, and administrative efficiency. However, the analysis also identifies a minimal portion of the responses, precisely 3.7%, expressing negative sentiments. This minority indicates concerns or unsatisfactory experiences with ChatGPT, which could be attributed to challenges such as the potential for over dependency, issues with content accuracy, or the ethical implications of AI in educational settings.

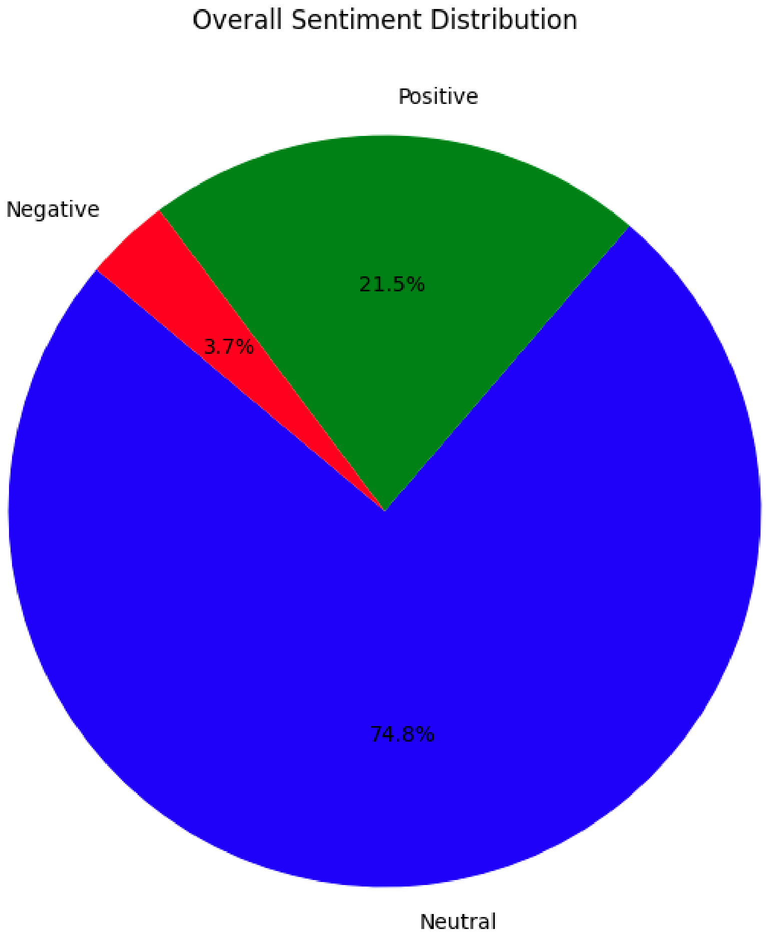


Figure 5.8: Overall sentiment distribution

5.2.1 Word-clouds for Sentiment Terms

To better understand the main themes and concerns in the collected data, word clouds were used to show how often certain terms appear in each sentiment category. These visuals are particularly effective in identifying the most common words associated with positive, neutral, and negative sentiments about integrating ChatGPT into higher education. The word clouds provide a clear visual summary of the key issues and positive aspects surrounding ChatGPT's adoption, helping to gain a more nuanced understanding of the various perspectives held by stakeholders in the academic community

Positive Sentiment Word Cloud

Overview: The positive sentiment word cloud in Figure 5.9 highlights the words frequently associated with positive feedback on ChatGPT's use in education. Words like "support," "potential," "learning," and "innovation" are used the most, reflecting a strong appreciation for the tool's benefits and innovative features.

Figure 5.9: Positive Sentiment Word Cloud for ChatGPT in Higher Education

Details:

- Support: The discussion in "Rethinking Educational Paradigms: Jason Jay's Innovative Use of ChatGPT in the Classroom" illustrates how ChatGPT supports educational dynamics by providing crucial assistance that enhances both teaching and student learning experiences.
- Potential: "What will the future of education look like in a world with generative AI?" examines the transformative potential of ChatGPT, emphasizing its significant impact on future educational methodologies.
- Learning: The article "Teaching & Learning with ChatGPT: Opportunity or Quagmire? Part I" focuses on the role of ChatGPT in enriching the educational process through tailored and interactive learning options.
- Innovation: In "MIT faculty, instructors, students experiment with generative AI in teaching and learning," the narrative explores how ChatGPT is spearheading innovations in educational practices.

Neutral Sentiment Word Cloud

Overview: Figure 5.10 displays a word cloud representing neutral attitude, which conveys a well-rounded viewpoint of ChatGPT's role in education. Words such as "data," "technology," "information," and "research" indicate a focus on the technical and informational aspects of ChatGPT. These does not indicate strong positive or negative connotations.

Figure 5.10: Negtaive Sentiment Word Cloud for ChatGPT in Higher Education

Details:

- Data: Data is used in titles like "How to Use ChatGPT's Advanced Data Analysis Feature" discussing how data is influencing ChatGPT's integration in the field of education.
- Technology: "Using ChatGPT to Perform a Skills Gap Analysis" demonstrates how ChatGPT is used to integrate technology into modern education.
- Information: "Advice and responses from faculty on ChatGPT and A.I.-assisted writing" talks about how ChatGPT serves as a key resource in managing content related to educational.
- Research: The exploration in "Can generative AI unlock technology-enabled tutoring, for everyone?" indicates academic focus on ChatGPT's integration.

Negative Sentiment Word Cloud

Overview: Words related to worries about ChatGPT in education are featured in the negative emotion word cloud shown in Figure 5.11. Noticeably, words like "challenges", "concern", "ethical", and "privacy" highlight areas where users see potential risks or drawbacks.

Figure 5.11: Negative Sentiment Word Cloud for ChatGPT in Higher Education

Details:

- **Challenges:** "Teaching & Learning with ChatGPT: Opportunity or Quagmire? Part III" talks about the challenges in integrating ChatGPT into existing educational frameworks.

- **Concern:** The study "Public Opinions on ChatGPT: An Analysis of Reddit Discussions by Using Sentiment Analysis, Topic Modeling, and SWOT Analysis" details various concerns about the use of ChatGPT.

- **Ethical:** "The Limitations and Ethical Considerations of ChatGPT" talks about the ethical dilemmas presented by the use of AI technologies like ChatGPT in educational settings.

- **Privacy:** "From ChatGPT to HackGPT: Meeting the Cybersecurity Threat of Generative AI" focuses on the privacy and security risks associated with using ChatGPT, stressing the need for improved data protection measures.

Even though there is a lot of enthusiasm for using ChatGPT in the classroom, there are some important things to keep in mind and potential risks that need to be taken into account. These observations align with the overall results of the sentiment analysis. These insights will help educators, lawmakers, and developers deploy ChatGPT in educational contexts correctly and successfully, maximising its advantages and minimising its drawbacks.

5.2.2 Country-wise sentiment analysis

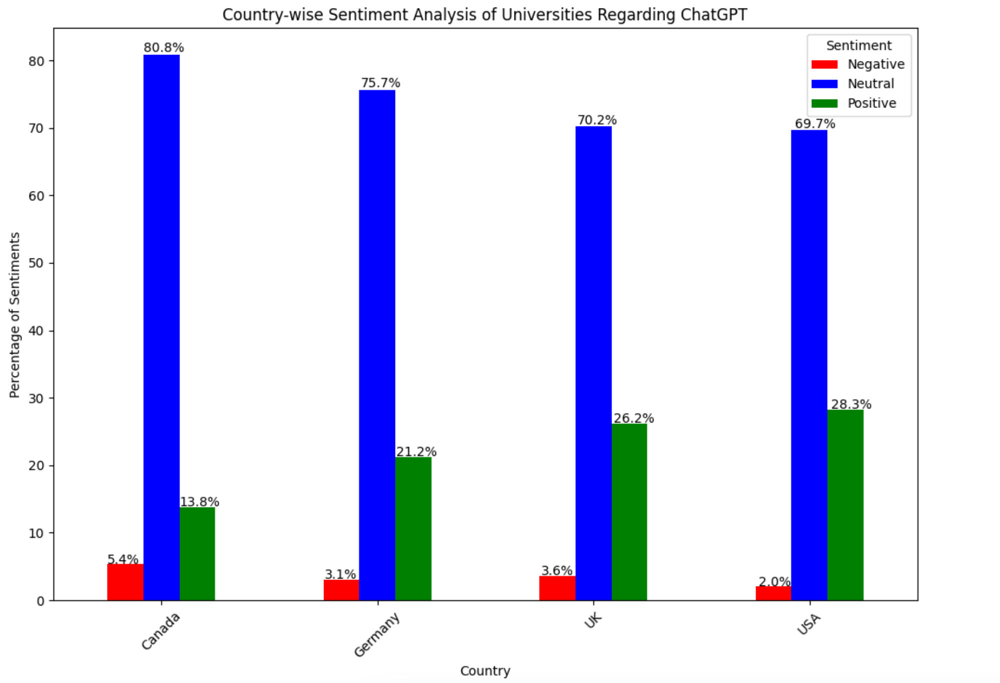


Figure 5.12: Countrywise sentiment analysis regarding chatgpt

Figure 5.12 shows the sentiment study with respect to countries illustrating diverse opinions across different regions. The country-wise sentiment analysis reveals varying levels of acceptance and concerns regarding ChatGPT among universities across different countries. In Canadian institutions, emotions are mostly neutral (80.8%), with a small percentage of positive (13.8%) and negative (5.4%) sentiments. German universities maintain a high degree of neutrality (75.7%) and a larger percentage of positive (21.2%) as compared to Canadian universities. On the other hand, opinions at US and UK universities are rather evenly distributed, with the US showing a substantial positive feeling (28.3%) and the UK showing a considerable positive sentiment (26.2%). This analysis of different nations reveals a generally positive but cautious approach towards integrating ChatGPT into higher education. Globally, opinion is largely neutral, indicating that academic institutions are weighing the advantages and possible drawbacks of ChatGPT. This careful viewpoint suggests a methodical and cautious approach, suggesting that although educational institutions recognise the benefits of AI technology, they also remain fully aware of the challenges and possible problems associated with their integration.

This uniform pattern of optimism across different regions underscores a common international perspective on the adoption of AI in education. It suggests that while there is a global sense of interest in using ChatGPT in education, there is also a need for thorough evaluation and strategic implementation as per the opinions taken from these regions. These insights can guide educators and policymakers in developing tailored AI adoption strategies that address both the opportunities and the challenges specific to their regional and institutional contexts.

5.2.3 Average sentiment scores by university

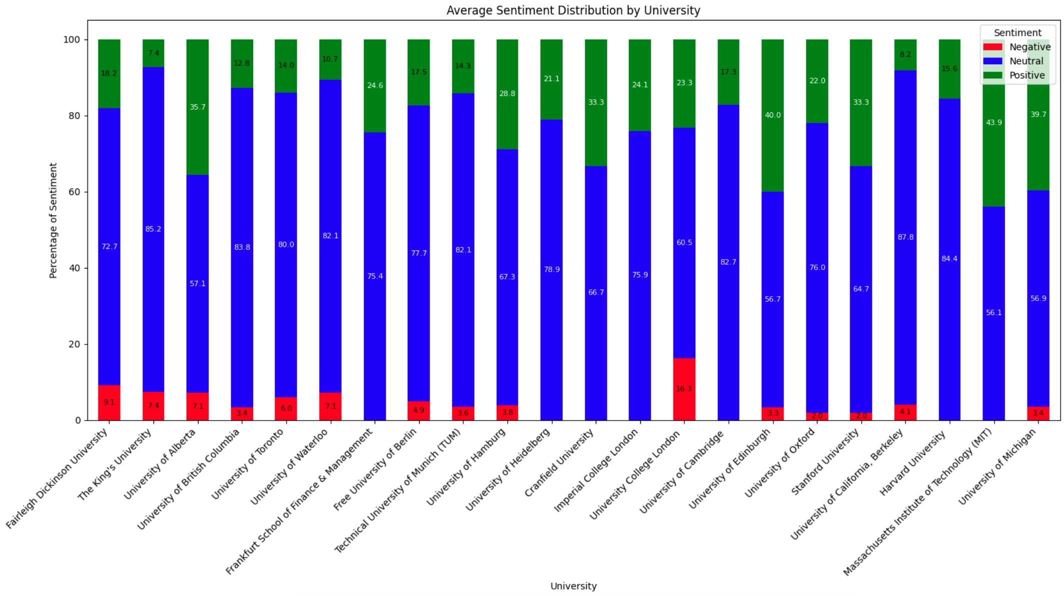


Figure 5.13: Average sentiment scores by university

The Figure 5.13 representing the average sentiment scores by university demonstrates the distribution of sentiments across different institutions. The data indicates:

Universities such as University of California and The King's University shows almost neutral emotions. More than 80% of the sentiments identified in these universities have a neutral opinions towards ChatGPT. This high rate of neutrality might indicate that universities are neither rejecting the advanced technology nor they are embracing it completely. It shows a cautious attitude towards fully integrating ChatGPT into educational practices.

Massachusetts Institute of Technology (MIT) is a university in the USA. It shows 43.9% of positive sentiments. MIT's is known for its leading role in technology and innovation sector which explains the strong positive sentiments identified. This also suggests a correlation between a university's technological focus and its openness to adopting advanced AI tools like ChatGPT. The university of alberta and stanford university is well known for its achievements in science, technology, engineering, and mathematics (STEM). These universities also show higher positive sentiments. This could also indicate that institutions with robust technological infrastructures and research cultures are more likely to perceive ChatGPT as beneficial. University College London (UCL) stands out with a significant negative sentiment of 16.3%. UCL is renowned for its traditional educational values. The presence of a higher negative sentiment at this institution might indicate concerns about the impact of ChatGPT on academic integrity and the quality of student learning.

The Figure 5.13 indicates that canadian universities have almost same percentages of positive and negative opinions and on the other hand, european univerisites show very less or no negative sentiments. This could be reflective of different educational policies, cultural attitudes towards technology, and the existing level of AI integration within the curriculum

The examination of positive sentiments towards ChatGPT's use in higher education reveals a strong focus on its transformative potential. For instance, at the University of Hamburg, an article called "We Became What We Once Admired" talks about how ChatGPT has helped students and professionals do better in their studies and careers by making it easier for them to succeed. Similarly, Imperial College London has articles like "Celebrating Advanced Creative Writing," which discuss how AI is used in creative writing classes. These articles show how AI can boost creativity and open up more learning opportunities, helping to push education and the arts forward. At University College London, another article titled "Launching Our Annual Report 2023" talks about how ChatGPT and other AI tools have greatly impacted healthcare engineering by improving research and development. Another example from Imperial College London, "DoC Lights Up Imperial Lates with Innovative AI," describes a successful event that explains AI advancements and highlighted community engagement. Lastly, Stanford University's article "ChatGPT: Revolutionizing Learning at Stanford" explains how ChatGPT is changing how students learn by making classes more engaging and improving learning efficiency. Overall, these articles highlight the positive ways ChatGPT is enhancing education, encouraging creativity and innovation and supporting research and development. They provide a positive view of how ChatGPT can benefit higher education, contrasting with the challenges AI might bring.

Several universities have raised concerns about using ChatGPT in higher education. For instance, an article from the Free University of Berlin, "ChatGPT Takes Errors and Biases from Its Data Sources," discusses worries about ChatGPT inheriting biases and mistakes from its training data. This raises questions about fairness and reliability. It important for understanding the bias in AI and the challenges universities face when using these technologies. Similarly, the University of California, Berkeley has a post titled "ChatGPT Teaches Students About the World They're Entering," which talks about how ChatGPT sometimes struggles with complex real-world scenarios. This limitation could mislead students and impact its use in education. Stanford University also looks at the pros and cons of AI in their article "AI in Education: A Double-Edged Sword?" They point out that while AI has benefits, it also risks undermining traditional teaching methods and critical thinking skills, which is important for addressing fears that AI could harm the quality of learning. At the Massachusetts Institute of Technology (MIT), the article "Evaluating the Risks of Generative AI" focuses on the rapid spread of ChatGPT usage. The article also explain that the rapid adoption can lead to over dependency, ethical concerns, and overshadowing human expertise, which suggests a cautious approach to AI use. The Technical University of Munich also discusses similar issues in "Balancing Innovation and Ethics in AI Adoption," highlighting the ethical dilemmas of AI in education and the need for careful implementation. Common themes across these universities include ethical and bias concerns, the risk of over-dependence on AI, and questions about the reliability of AI-generated content.

5.2.4 Sentiment scores by type of university

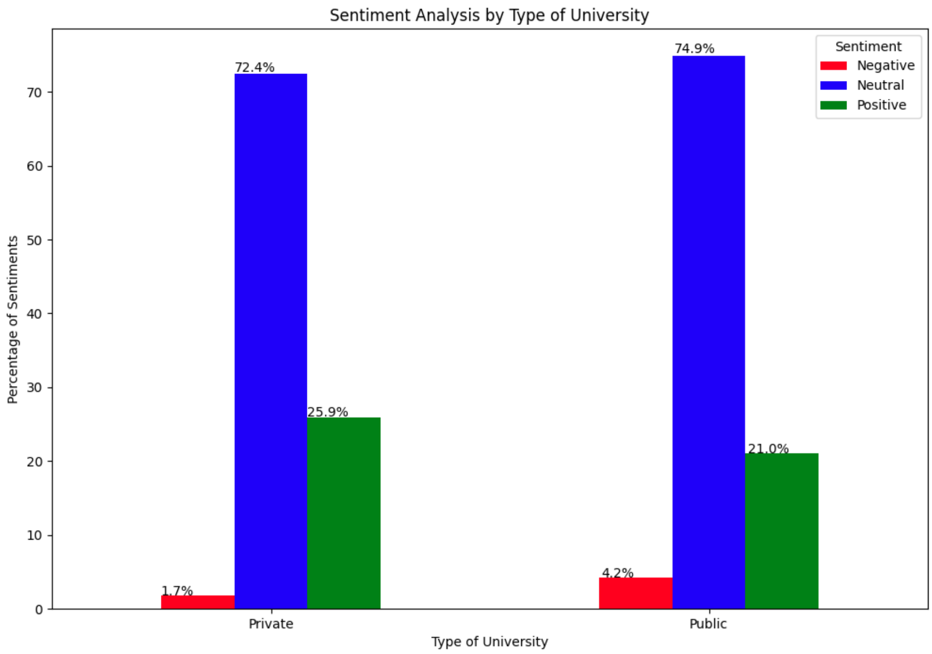


Figure 5.14: Sentiment scores by type of university

The university-type-specific sentiment analysis, as shown in the Figure 5.14, indicates how ChatGPT is perceived in public and private institutions. The sentiments expressed by private and public universities show that private institutions have a slightly higher positive sentiment (25.9%) compared to public universities (21.0%). Both types of institutions exhibit predominantly neutral sentiments, with private universities at 72.4% and public at 74.9% respectively.

These findings suggest that while there is a general cautiousness towards ChatGPT across both types of universities, private institutions might be slightly more open to or optimistic about integrating such technologies. This could be due to the flexibility and possibly easier applications of policies and rules at private universities to invest in and experiment with new technologies. Public universities, often larger and subject to more regulatory constraints, shows more conservative opinions. The greater negativity in public universities could reflect concerns about the scalability of ChatGPT and it's potential disruptive impacts on traditional educational models.

5.3 Modeling and Analysis Results

5.3.1 Topic modeling

The Pointwise Mutual Information (PMI) is used to measure how the words in data are related to each other. It is used to evaluate the quality of topics generated by a topic modeling algorithm. Higher PMI values indicate that the words within the topics are more closely associated with each other, suggesting that the results have more coherent and meaningful topics. The Table 5.1 shows the evaluation of the PMI for various numbers of sentences to analyze the performance of NTDMF model. The formula for PMI is given by:

$$\mathrm{PMI}(x;y) = \log\left(\frac{p(x,y)}{p(x)\cdot p(y)}\right) \text{ [101]}$$

where: $p(x, y)$ is the joint probability of events x and y occurring together. $p(x)$ and $p(y)$ are the individual probabilities of events x and y occurring independently [101].

No. of Sentences	PMI of NTDMF model
100	3.52
200	4.26
300	4.88
400	5.6
500	6.1

Table 5.1: Pointwise Mutual Information for NTDMF

From the Table 5.1, it is evident that the PMI values rise as the number of sentences increases. This trend indicates that the proposed NTDMF model performs better with larger datasets, leading to more coherent topics. Specifically, the PMI value increased from 3.52 with 100 sentences to 6.10 with 500 sentences, showing a significant improvement in topic coherence with more data.

5.3.2 Performance evaluation of sentiment analysis models

This section illustrates the effectiveness of the proposed framework in terms of analyzing the performance of the proposed HR-RAN techniques. The performance of the proposed HR-RAN is validated based on accuracy, precision, recall, f-measure, sensitivity, specificity. The training time required for the model is also calculated and compared it with other approaches like Convolutional Neural Network(CNN), Recurrent Neural Network(RNN), and Deep Neural Network(DNN).

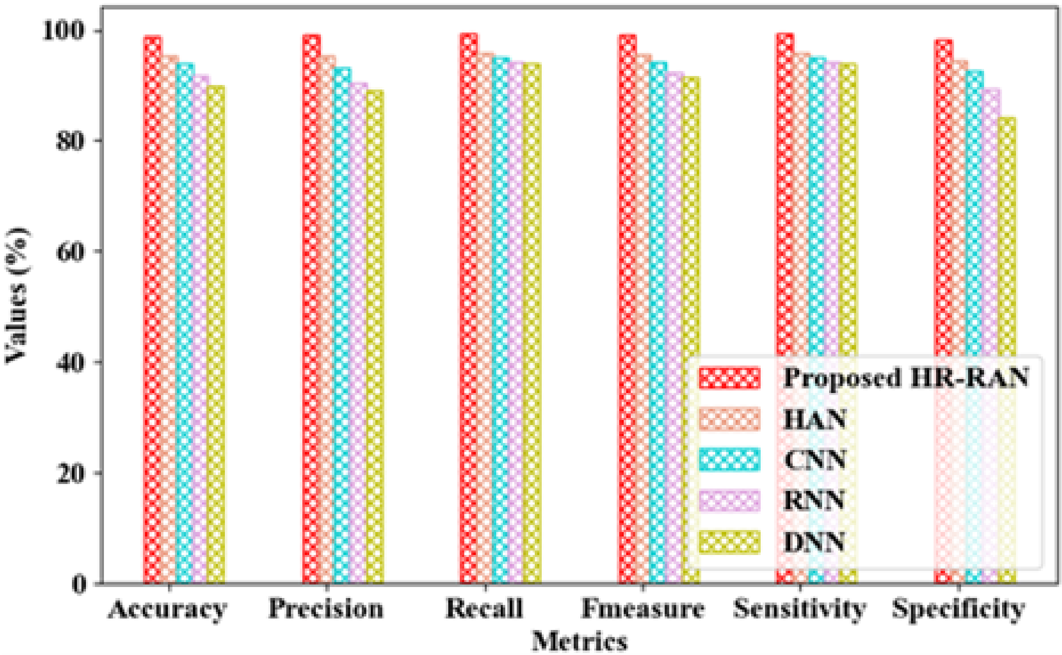
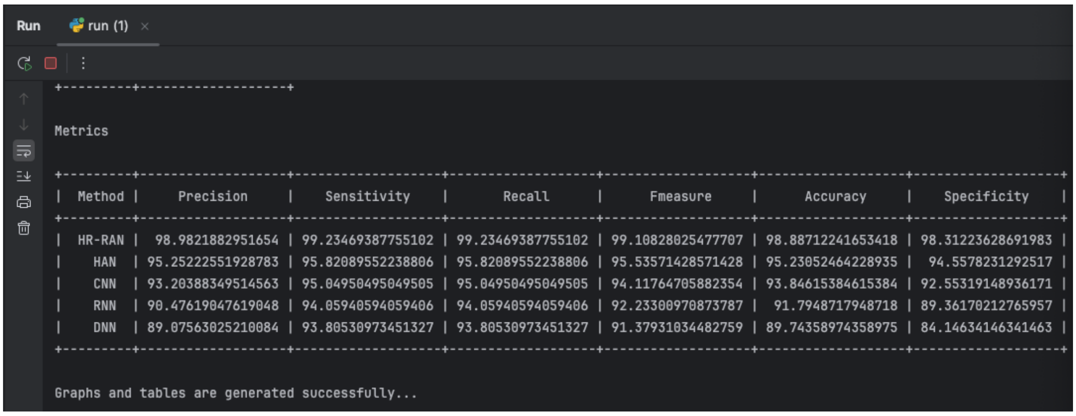


Figure 5.15: Performance evaluation of the HR-RAN approach

Figure 5.15 provides a comparative performance evaluation of HR-RAN, HAN, CNN, RNN, and DNN. This comparision is made using the key metrics such as accuracy, precision, recall, F-measure, and specificity. The HR-RAN model demonstrates superior performance, achieving high scores in the evaluation metrices. It achieved 98.98% in precision, 99.23% in recall, 99.10% in F-measure, 98.88% in accuracy, and 98.31% in specificity. These results underscore the effectiveness of the HR-RAN's architectural innovations, particularly its use of residual connections and the RSigELU activation function. This configuration not only enhances learning efficiency by promoting better gradient flow but also significantly improves the model's ability to classify sentences accurately.



Method	Precision	Sensitivity	Recall	Fmeasure	Accuracy	Specificity
HR-RAN	98.9821882951654	99.23469387755102	99.23469387755102	99.10828025477707	98.88712241653418	98.31223628691983
HAN	95.25222551928783	95.82089552238806	95.82089552238806	95.53571428571428	95.23052464228935	94.5578231292517
CNN	93.20388349514563	95.04950495049505	95.04950495049505	94.11764705882354	93.84615384615384	92.55319148936171
RNN	90.47619047619048	94.05940594059406	94.05940594059406	92.23300970873787	91.7948717948718	89.36170212765957
DNN	89.07563025210084	93.80530973451327	93.80530973451327	91.37931034482759	89.74358974358975	84.14634146341463

Figure 5.16: Quantitative Performance Metrics

The Figure 5.16 shows the analysis of various machine learning models revealing detailed differences in their performance across several key metrics. The Hierarchical Attention Network (HAN) and Convolutional Neural Network (CNN) show good performance, with HAN achieving a precision of 95.25% and a sensitivity of 95.82%, resulting in a robust F-measure of 95.54%. This indicates its excellent ability to identify relevant instances with high accuracy while minimizing false negatives. The CNN follows closely with a precision of 93.20% and sensitivity of 95.04%, indicating its effectiveness in spatial data interpretation, essential for tasks like image recognition. Both models show high specificity, 94.58% for HAN and 93.86% for CNN, highlighting their capacity to accurately identify negative instances.

On the other hand, the Recurrent Neural Network (RNN) and Deep Neural Network (DNN) display varying strengths, with RNN showcasing a higher sensitivity of 94.06% compared to DNN's 93.80%. However, both models have comparatively low specificity, where RNN and DNN record 89.36% and 84.16%, respectively. The DNN's precision at 89.08% and F-measure of 91.38% suggest that while it is broadly applicable, refinements are needed for complex applications. These comparative insights shows the advanced capabilities of the HR-RAN approach in handling sentence classification tasks more effectively.

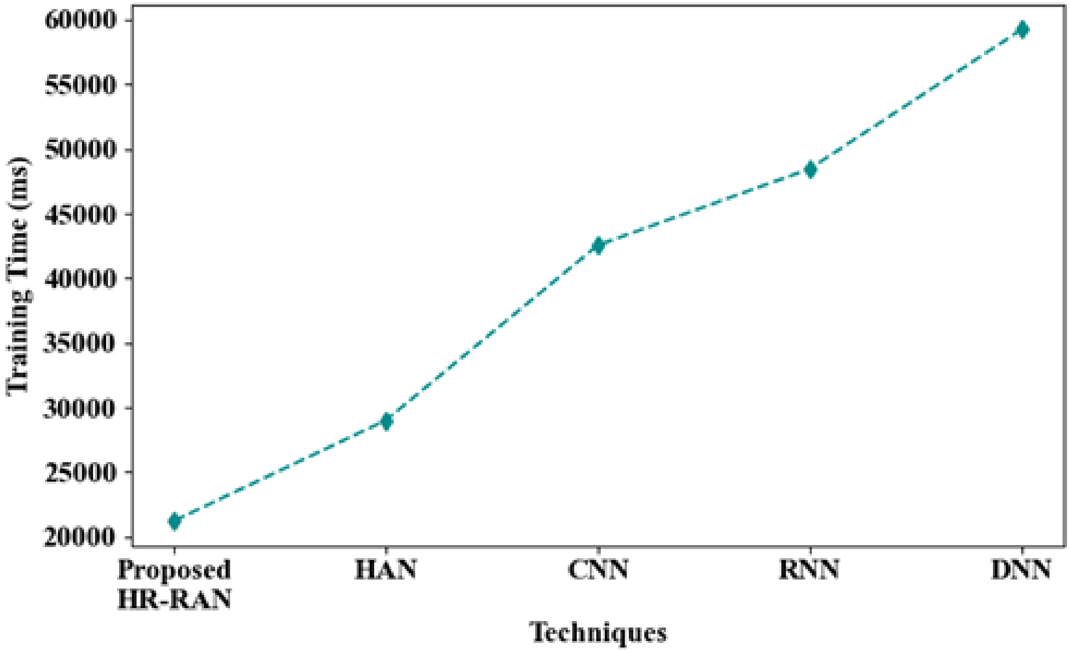


Figure 5.17: Training time evaluation of the HR-RAN

Figure 5.17 depicts the training time of the proposed model HR-RAN. The training time of the proposed HR-RAN is very fast when compared to the other existing sentiment analysis approaches like CNN, RNN, and DNN. The proposed HR-RAN took 21232ms of time for training, whereas the existing HAN, CNN, RNN, and DNN took 29016ms, 42559ms, 48472ms, and 59252ms of time for training. These quantitative details demonstrates that the proposed framework had good performance in how quickly the model can learn which is evident because of the less time required for training. The reason behind this fast learning rate is due to the incorporation of the RSigELU activation function, which avoids the vanishing gradient problem, thus leading to faster convergence.

5.4 Findings Summary

The results presents a comprehensive analysis of sentiments surrounding the integration of ChatGPT into higher education, leveraging advanced sentiment analysis models like HR-RAN, HAN, CNN, DNN, and RNN. The development of an interactive Graphical User Interface (GUI) in this research simplifies the sentiment analysis process. This GUI is used for providing a streamlined workflow, allowing users to manage sentiment analysis from data input to interpretation with ease. Various analytical steps, including sarcastic text detection, topic modeling, and data transformation ensures an accurate and nuanced understanding of sentiment dynamics within academic communities.

The sentiment analysis results reveal that a significant majority of sentiments, approximately 74.8%, are neutral, indicating a cautious stance among universities regarding ChatGPT's integration into educational frameworks. This suggests that while there is recognition of ChatGPT's potential, there is also hesitation or insufficient experience to form strong positive or negative opinions. The study makes use of visualization techniques such word clouds to present the recurring themes and issues related to ChatGPT's integration. These visualizations effectively identify common terms within each sentiment category, aiding in the interpretation of trends and key issues. Positive sentiments, highlighted by terms such as "support," "potential," "impact," and "learning," reflect an appreciation for ChatGPT's ability to enhance educational practices. This includes its potential to improve student engagement and provide personalized learning experiences. However, concerns regarding issues like bias, misuse, privacy, and misinformation are also prevalent, emphasizing the need for careful consideration of these challenges.

The advanced sentiment analysis model used in the study proved effective. It performed better than traditional methods by accurately capturing subtle sentiment nuances and contextual information. The model's high precision and recall rates illustrate its proficiency in sentiment classification, highlighting the value of employing advanced machine learning techniques to understand complex educational sentiments. However there are scenarios where other models might still be preferable depending on the specific requirements of the task or data characteristics.

6 Conclusion and future work

This study analyzed the sentiments and discourse expressed in university websites regarding the integration of ChatGPT in higher education. The incorporation of ChatGPT into the field of education has attracted significant attention, as evidenced by the comprehensive sentiment analysis conducted in this study. With the increase in the adoption of AI technologies, especially the ones developed by OpenAI, ChatGPT is a tool that can transform educational practices. The research revealed that a significant portion of sentiments expressed in the universities were neutral, indicating a cautious open stance toward ChatGPT's potential benefits and challenges. Positive sentiments highlighted the promise of enhanced student engagement and personalized learning, while negative sentiments underscored concerns about bias, privacy, and misinformation. Privacy concerns arise from how the sensitive data is handled, necessitating robust frameworks to ensure data protection and confidentiality. The concern regarding misinformation highlights the importance of improving ChatGPT's accuracy and reliability, ensuring that responses are based on verified information and do not mislead users. Addressing these challenges requires ongoing research and development efforts to refine ChatGPT models and enhance their capabilities.

By leveraging advanced sentiment analysis models, such as Hierarchical Residual Recurrent Attention Network (HR-RAN), the study effectively captured the complex opinions and sentiments expressed by academic communities regarding the use of ChatGPT in educational settings. The HR-RAN model demonstrated superior performance in classifying sentiments, outperforming other machine learning models like Convolutional Neural Network (CNN), Recurrent Neural Networks (RNN), and Deep Neural Network (DNN). Its high precision and recall rates validate its effectiveness in understanding the nuances of educational sentiments. This advanced approach allowed for a nuanced understanding of the sentiments, capturing fine-grained detailed and contextual information, leading to a more accurate classification of sentiments providing a robust framework for future sentiment analysis endeavors.

Building on the results of this study, there are a number of directions that future research can go. Firstly, expanding the dataset to include a broader range of institutions and geographical regions could provide a more comprehensive understanding of global sentiments towards ChatGPT. To record a wider range of emotions and interactions, future studies could potentially look into creating more complex models that use multimodal data, such audio and video. Another potential and promising future work could be collaborating with educational institutions to conduct pilot studies and real-world implementations of ChatGPT can provide valuable feedback and inform best practices for its integration.

In conclusion, this study highlights that universities are either neutral or show significant positive attitudes toward the adoption of ChatGPT in educational settings. This indicates a promising future for ChatGPT in education, provided that the challenges are effectively managed. By leveraging the strengths of ChatGPT and addressing its limitations, educational institutions can create more dynamic and effective learning environments, ultimately benefiting both students and educators.

Bibliography

[1] Python interface to tcl/tk. https://docs.python.org/3/library/tkinter.html.

[2] Sentiment analysis guide. https://monkeylearn.com/sentiment-analysis/.

[3] T. Adiguzel, M. H. Kaya, and F. K. Cansu. Revolutionizing education with ai: Exploring the transformative potential of chatgpt. *Contemporary Educational Technology*, 15(3):1–13, 2023.

[4] K. Ahmad, W. Iqbal, A. El-Hassan, J. Qadir, D. Benhaddou, M. Ayyash, and A. Al-Fuqaha. Data-driven artificial intelligence in education: A comprehensive review. *IEEE Transactions on Learning Technologies*, 17:12–31, 2024.

[5] M. AL-Smadi. Chatgpt and beyond: The generative ai revolution in education. (arXiv:2311.15198), Nov. 2023. arXiv:2311.15198 [cs].

[6] Y. AlBadarin, M. Tukiainen, M. Saqr, and N. Pope. A systematic literature review of empirical research on chatgpt in education. *SSRN Electronic Journal*, 2023.

[7] M. Aljebreen, B. Alabduallah, M. M. Asiri, A. S. Salama, M. Assiri, and S. S. Ibrahim. Moth flame optimization with hybrid deep learning based sentiment classification toward chatgpt on twitter. *IEEE Access*, 11:104984–104991, 2023.

[8] AltexSoft. Deep learning and the future of machine learning. https://www.altexsoft.com/blog/deep-learning/, oct 2021.

[9] K. Ansari. 500k chatgpt-related tweets jan-mar 2023. https://www.kaggle.com/datasets/khalidryder777/500k-chatgpt-tweets-jan-mar-2023, 2023. AI Conversations: Discover ChatGPT's Impact Through Tweets.

[10] A. A. Q. Aqlan, B. Manjula, and R. Lakshman Naik. A study of sentiment analysis: Concepts, techniques, and challenges. In *Lecture Notes on Data Engineering and Communications Technologies*, volume 28. Springer Singapore, 2019.

[11] Aydin and E. Karaarslan. Is chatgpt leading generative ai? what is beyond expectations? *Academic Platform Journal of Engineering and Smart Systems*, 11(3):118–134, Sept. 2023.

[12] D. Baidoo-Anu and L. O. Ansah. Education in the era of generative artificial intelligence (ai): Understanding the potential benefits of chatgpt in promoting teaching and learning.

[13] Y. Cao, A. A. Aziz, and W. N. R. M. Arshard. University students' perspectives on artificial intelligence: A survey of attitudes and awareness among interior architecture students. *IJERI: International Journal of Educational Research and Innovation*, (2020):1–21, Dec. 2023.

[14] J. Chen, Z. Zhuo, and J. Lin. Does chatgpt play a double-edged sword role in the field of higher education? an in-depth exploration of the factors affecting student performance. *Sustainability*, 15(2424):16928, Jan. 2023.

[15] L. Chen, P. Chen, and Z. Lin. Artificial intelligence in education: A review. *IEEE Access*, 8:75264–75278, 2020.

[16] O. E. Chinonso, A. M.-E. Theresa, and T. C. Aduke. Chatgpt for teaching, learning and research: Prospects and challenges. *Global Academic Journal of Humanities and Social Sciences*, 5(02):33–40, 2023.

[17] J. Chuang, C. Manning, and J. Heer. Termite: Visualization techniques for assessing textual topic models. may 2012.

[18] W. Contributors. Sentiment analysis: Overview, applications, and future directions. `https://en.wikipedia.org/w/index.php?title=Sentiment_analysis&oldid=1230364084`, June 2024. Page Version ID: 1230364084.

[19] D. K. Dake and E. Gyimah. Using sentiment analysis to evaluate qualitative students' responses. *Education and Information Technologies*, 28(4):4629–4647, 2023.

[20] DataEQ. What is sentiment analysis. `https://dataeq.com/resources/news/what-is-sentiment-analysis`, October 2021. DataEQ Resources.

[21] A. ElSayary. An investigation of teachers' perceptions of using chatgpt as a supporting tool for teaching and learning in the digital era. *Journal of Computer Assisted Learning*, n/a(n/a).

[22] M. Firat. What chatgpt means for universities: Perceptions of scholars and students. *Journal of Applied Learning and Teaching*, 6:1–22, Apr. 2023.

[23] B. Foroughi, M. G. Senali, M. Iranmanesh, A. Khanfar, M. Ghobakhloo, N. Annamalai, and B. Naghmeh-Abbaspour. Determinants of intention to use chatgpt for educational purposes: Findings from pls-sem and fsqca. *International Journal of Human-Computer Interaction*, pages 1–20, 2023.

[24] K. A. A. Gamage, S. C. P. Dehideniya, Z. Xu, and X. Tang. Chatgpt and higher education assessments: More opportunities than concerns? *Journal of Applied Learning and Teaching*, 6(2):358–369, 2023.

[25] M. Gheorghe, F.-C. Mihai, and M. Dârdală. Modern techniques of web scraping for data scientists. *International Journal of User-System Interaction*, 11(1):63–75, 2018.

[26] I. Goodfellow, Y. Bengio, and A. Courville. *Deep Learning*. MIT Press, 2016.

[27] M. Grootendorst. Bertopic: Neural topic modeling with a class-based tf-idf procedure. *arXiv preprint arXiv:2203.05794*, March 2022.

[28] Y. Gupta. Chat gpt and gpt 3 detailed architecture study-deep nlp horse. `https://medium.com/nerd-for-tech/gpt3-and-chat-gpt-detailed-architecture-study-deep-nlp-horse-db3af9de8a5d`, March 2023.

[29] M. Gurucharan. Basic cnn architecture. `https://www.upgrad.com/blog/basic-cnn-architecture/`, 2024.

[30] J. Hartmann, M. Heitmann, C. Siebert, and C. Schamp. More than a feeling: Accuracy and application of sentiment analysis. *International Journal of Research in Marketing*, 40(1):75–87, 2023.

[31] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. *arXiv*, (arXiv:1512.03385), Dec. 2015. arXiv:1512.03385 [cs].

[32] F. Heimerl, S. Lohmann, S. Lange, and T. Ertl. *Word Cloud Explorer: Text Analytics Based on Word Clouds*. Proceedings of the Annual Hawaii International Conference on System Sciences, jan 2014. journalAbbreviation: Proceedings of the Annual Hawaii International Conference on System Sciences.

[33] M. Hooda, C. Rana, O. Dahiya, A. Rizwan, and M. S. Hossain. Artificial intelligence for assessment and feedback to enhance student success in higher education. *Mathematical Problems in Engineering*, 2022:1–19, 2022.

[34] A. M. Hopkins, J. M. Logan, G. Kichenadasse, and M. J. Sorich. Artificial intelligence chatbots will revolutionize how cancer patients access information: Chatgpt represents a paradigm-shift. *JNCI cancer spectrum*, 7(2):pkad010, Mar. 2023.

[35] M. Hossin and M. Sulaiman. A review on evaluation metrics for data classification evaluations. *International Journal of Data Mining & Knowledge Management Process*, 5(2):01–11, mar 2015.

[36] J. Huang, S. Saleh, and Y. Liu. A review on artificial intelligence in education. *Academic Journal of Interdisciplinary Studies*, 10(3):206, May 2021.

[37] X. Huang, D. Zou, G. Cheng, X. Chen, and H. Xie. Trends, research issues and applications of artificial intelligence in education. *Educational Technology and Society*, 26(1):112–131, 2023.

[38] J. Hutson and D. Plate. Enhancing institutional assessment and reporting through conversational technologies: Exploring the potential of ai-powered tools and natural language processing. *DS Journal of Artificial Intelligence and Robotics*, 1(1):11–22, 2023.

[39] N. Iqbal, H. Ahmed, and K. Azhar. Exploring teachers' attitudes towards using chat gpt. *Global Journal for Management and Administrative Sciences*, 3, Feb. 2023.

[40] M. e. a. Islam. Challenges and future in deep learning for sentiment analysis: a comprehensive review and a proposed novel hybrid approach. *Artificial Intelligence Review*, 57(3):62, 2024.

[41] A. Jacovi, O. Sar Shalom, and Y. Goldberg. Understanding convolutional neural networks for text classification. In *Proceedings of the 2018 EMNLP Workshop BlackboxNLP: Analyzing and Interpreting Neural Networks for NLP*, pages 56–65, Brussels, Belgium, 2018. Association for Computational Linguistics.

[42] S. Jaiswal. Natural language processing — dependency parsing: Different ways for dependency parsing using spacy, nltk with stanford corenlp and stanza. `https://towardsdatascience.com/natural-language-processing-dependency-parsing-cf094bbbe3f7`, Aug 2021.

[43] D. Kansara and V. Sawant. Comparison of traditional machine learning and deep learning approaches for sentiment analysis. In H. Vasudevan, A. Michalas, N. Shekokar, and M. Narvekar, editors, *Advanced Computing Technologies and Applications*, pages 365–377, Singapore, 2020. Springer.

[44] M. Khder. Web scraping or web crawling: State of art, techniques, approaches and application. *International Journal of Advances in Soft Computing and its Applications*, 13(3):145–168, Dec. 2021.

[45] P. Kherwa and P. Bansal. Topic modeling: A comprehensive review. *ICST Transactions on Scalable Information Systems*, 7, jul 2018.

[46] S. Kiliçarslan and M. Celik. Rsigelu: A nonlinear activation function for deep neural networks. *Expert Systems with Applications*, 174:114805, jul 2021.

[47] Y. Kim. Convolutional neural networks for sentence classification. *arXiv*, (arXiv:1408.5882), Sept. 2014. arXiv:1408.5882 [cs].

[48] R. Koonchanok, Y. Pan, and H. Jang. Tracking public attitudes toward chatgpt on twitter using sentiment analysis and topic modeling. *arXiv preprint arXiv:2306.12951*, June 2023.

[49] R. Koonchanok, Y. Pan, and H. Jang. Public attitudes toward chatgpt on twitter: sentiments, topics, and occupations. *Social Network Analysis and Mining*, 14(1):1–21, 2024.

[50] M. Kraus and S. Feuerriegel. Sentiment analysis based on rhetorical structure theory: Learning deep neural networks from discourse trees. *Expert Systems with Applications*, 118:65–79, mar 2019.

[51] V. Krotov, L. Johnson, and L. Silva. Tutorial: Legality and ethics of web scraping. *Communications of the Association for Information Systems*, 47, 2020.

[52] L. Li, Z. Ma, L. Fan, S. Lee, H. Yu, and L. Hemphill. Chatgpt in education: a discourse analysis of worries and concerns on social media. Oct. 2023.

[53] L. Li, Z. Ma, L. Fan, S. Lee, H. Yu, and L. Hemphill. Chatgpt in education: a discourse analysis of worries and concerns on social media. *Education and Information Technologies*, pages 1–35, 2023.

[54] P. Liu, X. Qiu, and X. Huang. Recurrent neural network for text classification with multi-task learning. *arXiv*, (arXiv:1605.05101), May 2016. arXiv:1605.05101 [cs].

[55] C. K. Lo. What is the impact of chatgpt on education? a rapid review of the literature. *Education Sciences*, 13(4):1–15, 2023.

[56] M. Mager, R. F. Astudillo, T. Naseem, M. A. Sultan, Y. S. Lee, R. Florian, and S. Roukos. Gpt-too: A language-model-first approach for amr-to-text generation. In *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, volume 1, pages 1846–1852, 2020.

[57] D. T. T. Mai, C. V. Da, and N. V. Hanh. The use of chatgpt in teaching and learning: a systematic review through swot analysis approach. *Frontiers in Education*, 9, 2024.

[58] Y. Mamo, H. Crompton, D. Burke, and C. Nickel. Higher education faculty perceptions of chatgpt and the influencing factors: A sentiment analysis of x. *TechTrends*, 68(3):520–534, 2024.

[59] B. Memarian and T. Doleck. Chatgpt in education: Methods, potentials, and limitations. *Computers in Human Behavior: Artificial Humans*, 1(2):1–11, 2023.

[60] B. Min, H. Ross, E. Sulem, A. P. B. Veyseh, T. H. Nguyen, O. Sainz, E. Agirre, I. Heintz, and D. Roth. Recent advances in natural language processing via large pre-trained language models: A survey. *ACM Computing Surveys*, 56(2):1–49, 2023.

[61] A. Mishra. Metrics to evaluate your machine learning algorithm. `https://towardsdatascience.com/metrics-to-evaluate-your-machine-learning-algorithm-f10ba6e38234`, feb 2018.

[62] R. Misra and P. Arora. Sarcasm detection using news headlines dataset. *AI Open*, 4:13–18, 2023.

[63] S. Mohammad. Challenges in sentiment analysis. In *Sentiment Analysis and Opinion Mining*, pages 61–83. Springer, Cham, 2017.

[64] K. Neha, R. Kumar, and M. Sankat. Ai wizards: Pioneering assistive technologies for higher education inclusion of students with learning disabilities. In *Applied Assistive Technologies and Informatics for Students with Disabilities*, pages 55–72. Springer, Singapore, 2024.

[65] T. Ngo. The perception by university students of the use of chatgpt in education. *International Journal of Emerging Technologies in Learning (iJET)*, 18:4–19, Sept. 2023.

[66] C. Nguyen. University teachers' perceptions of using chatgpt in language teaching and assessment. *Proceedings of the AsiaCALL International Conference*, 4:116–128, Jan. 2024.

[67] L. M. Nkomo and B. K. Daniel. Sentiment analysis of student engagement with lecture recording. *TechTrends*, 65(2):213–224, 2021.

[68] D. E. O'Leary. An analysis of three chatbots: Blenderbot, chatgpt and lamda. Research Paper Sponsored by iORB Forthcoming, USC Marshall School of Business, feb 2023.

[69] OpenAI. Chatgpt for windows. `https://www.openai.com/chatgpt`, 2023. Desktop application.

[70] T. Ouyang, A. MaungMaung, K. Konishi, Y. Seo, and I. Echizen. Stability analysis of chatgpt-based sentiment analysis in ai quality assurance. *ArXiv*, pages 1–8, 2024.

[71] L. Parker, C. Carter, A. Karakas, A. J. Loper, and A. Sokkar. Graduate instructors navigating the ai frontier: The role of chatgpt in higher education. *Computers and Education Open*, 6:1–13, 2024.

[72] F. Pedro, M. Subosa, A. Rivas, and P. Valverde. Artificial intelligence in education: Challenges and opportunities for sustainable development, 2019.

[73] V. Plevris, G. Papazafeiropoulos, and A. J. Rios. Chatbots put to the test in math and logic problems: A preliminary comparison and assessment of chatgpt-3.5, chatgpt-4, and google bard. `https://arxiv.org/abs/2305.18618`, 2023.

[74] R. Raman, S. Mandal, P. Das, T. Kaur, J. P. Sanjanasri, and P. Nedungadi. Exploring university students' adoption of chatgpt using the diffusion of innovation theory and sentiment analysis with gender dimension. *Human Behavior and Emerging Technologies*, pages 1–21, 2024.

[75] P. P. Ray. Chatgpt: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. *Internet of Things and Cyber-Physical Systems*, 3:121–154, 2023.

[76] A. Rejeb, K. Rejeb, A. Appolloni, H. Treiblmaier, and M. Iranmanesh. Exploring the impact of chatgpt on education: A web mining and machine learning approach. *International Journal of Management Education*, 22(1):1–14, 2024.

[77] K. I. Roumeliotis and N. D. Tselikas. Chatgpt and open-ai models: A preliminary review. *Future Internet*, 15(6):1–24, 2023.

[78] V. M. Ruchitaa Raj, Nandhakumar Raj. Web scraping tools and techniques: A brief survey. In *2023 4th International Conference on Innovative Trends in Information Technology (ICITIIT)*, pages 1–4, Feb. 2023.

[79] A. P. S and S. Aithal. Application of chatgpt in higher education and research – a futuristic analysis. *International Journal of Applied Engineering and Management Letters (IJAEML)*, 7(33):168–194, Sept. 2023.

[80] S. Z. Salas-Pilco, K. Xiao, and J. Oshima. Artificial intelligence and new technologies in inclusive education for minority students: A systematic review. *Sustainability (Switzerland)*, 14(20):1–17, 2022.

[81] N. Sandu and E. Gide. Adoption of ai-chatbots to enhance student learning experience in higher education in india. In *2019 18th International Conference on Information Technology Based Higher Education and Training (ITHET)*, pages 1–5, 2019.

[82] N. Sandu and E. Gide. Adoption of ai-chatbots to enhance student learning experience in higher education in india. In *2019 18th International Conference on Information Technology Based Higher Education and Training (ITHET)*, page 1–5, Magdeburg, Germany, Sept. 2019. IEEE.

[83] K. Sangeetha and D. Prabha. Sentiment analysis of student feedback using multi-head attention fusion model of word and context embedding for lstm. *Journal of Ambient Intelligence and Humanized Computing*, 12(3):1–10, 2021.

[84] M. Schönberger. *ChatGPT in Higher Education: The Good, The Bad, and The University*. June 2023.

[85] A. Strzelecki. Students' acceptance of chatgpt in higher education: An extended unified theory of acceptance and use of technology. *Innovative Higher Education*, page 1–23, Nov. 2023.

[86] Y. Su and Z. J. Kabala. Public perception of chatgpt and transfer learning for tweets sentiment analysis using wolfram mathematica. *Data*, 8(12):1–27, 2023.

[87] A. A. Sulaeman, M. Danny, S. Butsianto, and S. Pratama. Sentiment analysis on social media x (twitter) against chatgbt using the k-nearest neighbors algorithm. *Brilliance: Research of Artificial Intelligence*, 4(1):265–275, 2024.

[88] M. Sullivan, A. Kelly, and P. McLaughlan. Chatgpt in higher education: Considerations for academic integrity and student learning. *Journal of Applied Learning and Teaching*, 6(1):31–40, March 2023.

[89] A. S. Talaat. Sentiment analysis classification system using hybrid bert models. *Journal of Big Data*, 10(1):1–18, 2023.

[90] X. Tan. The impact of chatgpt on education and future prospects. *Highlights in Science, Engineering and Technology*, 61:138–143, 2023.

[91] The University of Texas at Austin Information Security Office. AI Tools and Resources. `https://security.utexas.edu/ai-tools`, 2024. Accessed: 2024-07-28.

[92] A. Tlili, B. Shehata, M. A. Adarkwah, A. Bozkurt, D. T. Hickey, R. Huang, and B. Agyemang. What if the devil is my guardian angel: Chatgpt as a case study of using chatbots in education. *Smart Learning Environments*, 10(1):15, Feb. 2023.

[93] M. Tubishat, F. Al-Obeidat, and A. Shuhaiber. Sentiment analysis of using chatgpt in education. In *2023 International Conference on Smart Applications, Communications and Networking (SmartNets)*, page 1–7, July 2023.

[94] M. Tubishat, F. Al-Obeidat, and A. Shuhaiber. Sentiment analysis of using chatgpt in education. *WSEAS Transactions on Advances in Engineering Education*, 20:60–66, 2023.

[95] UBC Centre for Teaching, Learning and Technology. Privacy Impact Assessments for Generative AI Instructional Use at UBC. `https://ai.ctlt.ubc.ca/privacy-impact-assessments-for-generativeai-instructional-use-at-ubc/`, 2024. Accessed: 2024-07-28.

[96] UBC News. UBC's Approach to Academic Integrity in the Age of AI. `https://news.ubc.ca/2023/03/ubcs-approach-to-academic-integrity-in-the-age-of-ai/`, 2023. Accessed: 2024-07-28.

[97] G. van den Berg and E. du Plessis. Chatgpt and generative ai: Possibilities for its contribution to lesson planning, critical thinking and openness in teacher education. *Education Sciences*, 13(10):1–12, 2023.

[98] N. Van Otten. Coreference resolution in natural language processing (nlp) simplified [8 powerful techniques & 2 models].

https://spotintelligence.com/2024/01/17/coreference-resolution-nlp/, Jan 2024.

[99] A. Wagh. Open ai – understand foundational concepts of chatgpt and cool stuff you can explore! https://medium.com/@amol-wagh/open-ai-understand-foundational-concepts-of-chatgpt-and-cool-stuff-you-can-explore-a7a77baf0ee3, July 2023.

[100] M. Wankhade, A. Rao, and C. Kulkarni. A survey on sentiment analysis methods, applications, and challenges. *Artificial Intelligence Review*, 55(7):5731–5780, 2022.

[101] Wikipedia contributors. Pointwise mutual information — Wikipedia, the free encyclopedia. https://en.wikipedia.org/w/index.php?title=Pointwise_mutual_information&oldid=XXXXXXX, 2024. [Online; accessed August 8, 2024].

[102] Z. Yang, D. Yang, C. Dyer, X. He, A. Smola, and E. Hovy. Hierarchical attention networks for document classification. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 1480–1489, San Diego, California, 2016. Association for Computational Linguistics.

[103] B. Zhao. *Web Scraping*, pages 1–3. Springer International Publishing, Cham, 2017.

[104] J. Zhou, Z. Liang, Y. Fang, and Z. Zhou. Exploring public response to chatgpt with sentiment analysis and knowledge mapping. *IEEE Access*, 12:50504–50516, 2024.