

# **GENERATIVE AI CHATBOT FOR STUDENT COUNSELLING**

**A PROJECT REPORT**

*Submitted by*

**PRIYANSHU MAHESHWARI**

**[Reg No:RA2011028010091]**

**UDIT KUMAR**

**[Reg No: RA2011031010099]**

*Under the Guidance of*

**Dr. A. HELEN VICTORIA**

Assistant Professor, Department of Computing Technologies

*in partial fulfillment of the requirements for the degree of*

**BACHELOR OF TECHNOLOGY**

**in**

**COMPUTER SCIENCE AND ENGINEERING**



**DEPARTMENT OF NETWORKING AND  
COMMUNICATION**

**COLLEGE OF ENGINEERING AND TECHNOLOGY  
SRM INSTITUTE OF SCIENCE AND TECHNOLOGY**

**KATTANKULATHUR– 603 203**

**NOVEMBER 2023**



**SRM INSTITUTE OF SCIENCE AND TECHNOLOGY**  
**KATTANKULATHUR-603 203**  
**BONAFIDE CERTIFICATE**

Certified that 18CSP109L / I8CSP111L project report titled “**GENERATIVE AI CHATBOT FOR STUDENT COUNSELLING**” is the bonafide work of **PRIYANSHU MAHESHWARI [Reg No:RA2011028010091]** and **UDIT KUMAR [Reg No: RA2011031010099]** who carried out the project work under my supervision. Certified further, that to the best of my knowledge the work reported here in does not form part of any other thesis or dissertation on the basis of which a degree or award was conferred on an earlier occasion for this or any other candidate.

**Dr. A. HELEN**  
**VICTORIA**  
**SUPERVISOR**

Assistant Professor Department of  
Networking and Communications

**Dr. ANNAPURANI K**  
**HEAD OF THE DEPARTMENT**  
Department of Networking and  
Communications

EXAMINER I

EXAMINER II

Department of Networking and Communication  
**SRM INSTITUTE OF SCIENCE AND TECHNOLOGY**  
**Own Work Declaration Form**

**Degree/Course** : B.Tech in Computer Science and Engineering

**Student Names** : PRIYANSHU MAHESHWARI, UDIT KUMAR

**Registration Number** : RA2011028010091, RA20110310100099

**Title of Work** : **Generative AI chatbot for Student Counselling**

We here by certify that this assessment compiles with the University's Rules and Regulations relating to Academic misconduct and plagiarism, as listed in the University Website, Regulations, and the Education Committee guidelines.

We confirm that all the work contained in this assessment is our own except where indicated, and that we have met the following conditions:

- Clearly references / listed all sources as appropriate
- Referenced and put in inverted commas all quoted text (from books, web, etc.)
- Given the sources of all pictures, data etc that are not my own.
- Not made any use of the report(s) or essay(s) of any other student(s) either past or present
- Acknowledged in appropriate places any help that I have received from others (e.g., fellow students, technicians, statisticians, external sources)
- Compiled with any other plagiarism criteria specified in the Course hand book / University website
- I understand that any false claim for this work will be penalized in accordance with the University policies and regulations.

**DECLARATION:**

I am aware of and understand the University's policy on Academic misconduct and plagiarism and I certify that this assessment is my / our own work, except where indicated by referring, and that I have followed the good academic practices noted above.

**Student 1 Signature:**

**Student 2 Signature:**

**Date:**

If you are working in a group, please write your registration numbers and sign with the date for every student in your group.

## ACKNOWLEDGEMENT

We express our humble gratitude to **Dr. C. Muthamizhchelvan**, Vice Chancellor, SRM Institute of Science and Technology, for the facilities extended for the project work and his continued support.

We extend our sincere thanks to Dean-CET, SRM Institute of Science and Technology, **Dr. T. V. Gopal**, for his invaluable support.

We wish to thank **Dr. Revathi Venkataraman**, Professor and Chairperson, School of Computing, SRM Institute of Science and Technology, for her support throughout the project work.

We are incredibly grateful to our Head of the Department, **Dr. Annapurani K**, Professor, Department of Networking and Communications, SRM Institute of Science and Technology, for her suggestions and encouragement at all the stages of the project work.

We want to convey our thanks to our Project Coordinators, **Dr. P Supraja**, Panel Head, **Dr. A Helen Victoria**, **Dr. Saranya G** Assistant Professor and Panel Members, Department of Networking and Communications, SRM Institute of Science and Technology, for their inputs during the project reviews and support.

We register our immeasurable thanks to our Faculty Advisor, **Mrs. V Vijayalakshmi**, Assistant Professor and **Dr. Safa M**, Assistant Professor, Department of Networking and Communications, SRM Institute of Science and Technology, for leading and helping us to complete our course.

Our inexpressible respect and thanks to our guide, **Dr. A Helen Victoria**, Assistant Professor, Department Networking and Communications, SRM Institute of Science and Technology, for providing us with an opportunity to pursue our project under his / her mentorship. He / She provided us with the freedom and support to explore the research topics of our interest. His / Her passion for solving problems and making a difference in the world has always been inspiring.

---

We sincerely thank all the staff and students of Networking and Communications Department, School of Computing, S.R.M Institute of Science and Technology, for their help during our project. Finally, we would like to thank our parents, family members, and friends for their unconditional love, constant support, and encouragement.

**PRIYANSHU MAHESHWARI [Reg No:RA2011028010091]**

**UDIT KUMAR [Reg No: RA2011031010099]**

## ABSTRACT

Psychological well-being difficulties have emerged as a global concern affecting people worldwide throughout their lives, with a staggering 29% of individuals grappling with mental health issues at various points, imposing an annual cost of 10% of children and a significant 25% of adults. These figures underscore the widespread reach of these challenges, impacting diverse populations on a global scale. At the heart of this issue lie two significant categories of mental disorders: depression and anxiety. In 2022, 322 million people worldwide battled depressive illnesses, while 264 million faced anxiety disorders, not only diminishing their quality of life but also serving as leading causes of occupational disability. The economic impact is substantial, as demonstrated by countries like Switzerland, investing an estimated CHF 11 billion annually to address mental health concerns, encompassing both direct treatment expenses and indirect costs such as reduced workplace productivity and complications stemming from untreated mental conditions, like suicides. This pressing and widespread issue underscores the urgent need for enhanced global mental health resources and support, with innovative solutions like generative chatbots offering a more accessible and personalized approach to address the significant impact of mental health challenges across communities and regions.

---

# TABLE OF CONTENTS

C.NO.	TITLE	PAGE NO.
	<b>ABSTRACT</b>	<b>vii</b>
	<b>LIST OF SYMBOLS AND ABBREVIATIONS</b>	<b>x</b>
	<b>LIST OF FIGURES</b>	<b>xi</b>
<b>1.</b>	<b>INTRODUCTION</b>	<b>1</b>
1.1	General	1
1.2	Purpose	1
1.3	Scope	3
1.4	Azure OpenAI and Large Language Models in Mental HealthSupport	4
1.4.1	Azure OpenAI Services	5
1.4.2	Large Language Models (LLM)	5
1.4.3	Azure Blob Storage	7
1.4.4	Azure Cognitive Search	7
<b>2</b>	<b>LITERATURE SURVEY</b>	<b>8</b>
2.1	Intelligent Career Counselling Chatbot (ICCC)	8
2.2	Career Counseling Chatbot	9
2.3	SERMO	9
2.4	Mind Relaxation Chatbot	10
2.5	Development of a Multilingual Talkbot	11
<b>3</b>	<b>PROPOSED METHODOLOGY</b>	<b>12</b>
3.1	Azure OpenAI Model	12
3.2	Azure Blob Storage	12
3.3	Azure Cognitive Search	12
3.4	Interconnectivity	12
3.5	Algorithm	13
3.6	Transformer Internal Modules and its Working	15
3.7	Construction of Chatbot	16
<b>4</b>	<b>RESULTS</b>	<b>18</b>
<b>5</b>	<b>CONCLUSION</b>	<b>23</b>



<b>6</b>	<b>FUTURE SCOPE</b>	<b>24</b>
<b>7</b>	<b>REFERENCES</b>	<b>25</b>
	<b>APPENDIX</b>	<b>26</b>
	<b>PAPER PUBLICATION STATUS</b>	<b>28</b>
	<b>PLAGIARISM REPORT</b>	<b>29</b>

---

## LIST OF SYMBOLS AND ABBREVIATIONS

<b>LLM</b>	Large Language Model
<b>NLP</b>	Natural Language Processing
<b>NLU</b>	Natural Language Understanding
<b>API</b>	Application Programming Interface
<b>XAI</b>	Explainable Artificial Intelligence
<b>AI</b>	Artificial Intelligence
<b>ML</b>	Machine Learning
<b>CBT</b>	Cognitive Behavior Therapy

## LIST OF FIGURES

FIGURE NO.	DESCRIPTION	PAGE NO.
3.4.1	Architecture Diagram	32
3.5.1	GPT model architecture	35
3.6.1	Transformer Architecture	38
4.1	Chatbot history	42
4.2.1	Conversation with the chatbot	43
4.2.1	Sample Chat	43
4.4	Customization of answer according to user prompt	44
4.5	References to the data	44
4.6	VScode (Running code)	47
4.7	Comparison between the answer and data	47

# CHAPTER 1

## INTRODUCTION

### 1.1 General

Mental well-being challenges have evolved into a pervasive and far-reaching global concern, affecting a substantial portion of the world's population at various points throughout their lives. An astonishing 29% of individuals worldwide grapple with mental health issues, creating a vast and intricate web of personal struggles that reverberate on a global scale. This phenomenon is not confined to a specific age group, as these conditions impact 10% of children and a significant 25% of adults annually, underscoring the profound and widespread consequences of these challenges. Within the realm of mental disorders, anxiety and depression disorders stand out as the most prevalent. In 2017 alone, a staggering 322 million people worldwide experienced depressive illnesses, while 264 million individuals grappled with anxiety disorders. These disorders not only compromise the overall quality of life for those who bear their weight but also stand as leading causes of occupational disability, imposing a substantial economic burden on societies. The financial strain imposed by these conditions is palpable, exemplified by countries such as Switzerland, which invests an estimated CHF 11 billion annually to address mental health concerns. This figure encompasses not only direct treatment costs but also indirect costs such as reduced workplace productivity and complications stemming from untreated mental health conditions, including tragic outcomes such as suicides.

This inescapable and far-reaching issue underscores the urgent and compelling need for enhanced global mental health support and resources. The scale of this challenge calls for a multi-faceted and holistic approach to address mental well-being, not only as a humanitarian imperative but as an essential aspect of societal well-being and economic sustainability. The call to action is clear: the global community must prioritize, invest in, and innovate around mental health to alleviate suffering and reduce the staggering costs associated with untreated mental disorders.

### 1.2. Purpose

In today's interconnected and fast-paced world, mental health challenges have become a far-reaching global concern that transcends geographical boundaries and affects a significant portion of the population during their lifetimes. According to recent statistics, an astonishing 29% of individuals worldwide grapple with mental health issues at various stages of their lives. This pervasive issue extends to our youngest generation, as it annually impacts 10% of children, and a substantial 25% of adults, underscoring the extensive consequences of these challenges.

---

Among the multitude of mental disorders, depression and anxiety disorders are the most prevalent. In 2022 alone, 322 million people worldwide experienced depressive illnesses, while 264 million individuals battled anxiety disorders. These conditions not only erode the quality of life for those afflicted but also represent a leading cause of occupational disability, exerting a profound economic toll on societies. To illustrate this economic impact, countries like Switzerland invest an estimated CHF 11 billion annually to address mental health concerns, encompassing direct treatment costs as well as indirect costs like reduced workplace productivity and complications arising from untreated mental conditions, including the tragedy of suicides.

The financial and human costs associated with the mental health crisis are undeniable and compelling. The urgent need for enhanced global mental health resources and support is clear. It is in this context that the purpose of our project comes into focus, aiming to provide transformative response to this global challenge.

Our project is driven by the mission to harness advanced technologies, specifically generative artificial intelligence (AI), to develop a state-of-the-art chatbot for student counseling that addresses the multifaceted issues associated with mental health. The primary purpose of this endeavor is to make mental health support more accessible, scalable, and personalized.

**Accessibility:** One of the fundamental objectives of our project is to extend the reach of mental health support to a broader and more diverse audience. With the advent of AI-driven chatbots, we have the opportunity to provide an easily accessible and scalable means of delivering mental health guidance and therapeutic interventions. These chatbots offer individuals, particularly those in underserved or remote areas, a lifeline of advice, direction, and support. They can act as a bridge to connect people with mental health resources who might otherwise face barriers to access.

**Scalability:** The scalability of our solution is a pivotal aspect of our purpose. We recognize the severe shortage of mental health professionals and the enormous demand for support. Our project leverages AI technology to bridge this gap by offering guidance and therapeutic interventions to a large and ever-expanding user base. The potential for chatbots to offer immediate and round-the-clock support is a critical asset, ensuring that support is available when individuals need it most.

**Personalized Support:** Our project focuses on the personalization of mental health support. The chatbot's generative AI capabilities empower it to provide tailored responses to user queries and prompts. This feature enhances the quality and relevance of support offered, reflecting an individual's unique needs and preferences. Whether a user seeks responses in bullet points, tables, or short paragraphs, our chatbot adapts accordingly to provide effective assistance.

**Data-Driven Approach:** The project integrates cutting-edge technologies, including Azure OpenAI, Azure Blob Storage, and Azure Cognitive Search, to effectively manage and process the vast amount of data required for informed responses. By leveraging data and evidence-based insights, our chatbot ensures that the information and guidance provided is accurate, reliable, and based on the latest advancements in the field of mental health.

**Reducing Stigma:** We recognize the stigma that often surrounds seeking mental health support, and our project addresses this issue head-on. AI-driven chatbots can offer a non-judgmental and confidential space for individuals to discuss their concerns and access the guidance they need. This can be particularly crucial for those who may be hesitant to reach out to human counselors due to perceived stigma or fear of judgment.

**Cost-Efficiency:** Our project acknowledges the significant economic burden that untreated mental health conditions place on societies. By offering a cost-effective means of providing mental health support, we aim to alleviate some of the financial strains associated with mental health challenges. This approach considers not only direct treatment costs but also indirect expenses such as reduced workplace productivity and the far-reaching complications that stem from untreated mental conditions, including the tragic and preventable outcomes of suicide.

The purpose of our project is not merely to create another technological innovation, but to contribute to a global effort to address the extensive mental health challenges outlined above. Our chatbot represents a critical piece of the puzzle in enhancing the well-being of individuals across diverse communities and regions, and in doing so, it upholds the values of accessibility, scalability, personalization, data-driven support, and stigma reduction. This endeavor speaks to a comprehensive response to the global mental health crisis, underscoring the importance of investing in technology-driven solutions for the betterment of humanity.

### 1.3. Scope

**Broad Accessibility:** The project aspires to extend the reach of mental health support to a global scale. It seeks to ensure that individuals from diverse backgrounds and locations, including those residing in underserved and remote areas, can readily access the guidance and assistance they need. This broad accessibility transcends geographical boundaries and socio-economic constraints, making mental health support universally available.

**Scalable Impact:** At the heart of the project's scope is its commitment to providing scalable solutions. By harnessing AI-driven chatbots, the project intends to offer mental health support to a large and continually growing user base. These chatbots, operating around the clock, ensure that support is accessible whenever individuals require it. This scalability contributes to the broader dissemination of mental health resources.

**Personalized Care:** One of the central objectives of the project is to deliver tailored and personalized mental health support. The chatbot's generative AI capabilities empower it to adapt responses to the specific needs and preferences of individual users. This personalization enhances the quality and relevance of support, creating an experience that is uniquely meaningful to each user.

**Data-Driven Insights:** The project integrates cutting-edge technologies, including data from sources like Azure OpenAI, Azure Blob Storage, and Azure Cognitive Search. This data-driven approach ensures that the chatbot can provide users with accurate and evidence-based information. It stays in sync with the latest developments in the field of mental health, aligning with the evolving body of knowledge and best practices.

**Reducing Stigma:** By creating a safe and confidential space for individuals to seek mental health support, the project addresses the stigma that often surrounds this crucial need. The non-judgmental nature of the AI-driven chatbot encourages more individuals to access the support they require. This reduction in stigma aligns with the project's broader goal of making mental health resources more approachable and less intimidating.

**Cost-Effective Solutions:** Recognizing the economic burden associated with untreated mental health conditions, the project seeks to provide a cost-effective approach to mental health support. This extends beyond direct treatment costs to encompass indirect expenses, such as reduced workplace productivity and the wider repercussions of untreated mental conditions.

**Comprehensive Impact:** Beyond its technological aspects, the project's scope reaches into the broader realm of global mental health. It aligns with the wider initiative to enhance the well-being of individuals, emphasizing the importance of accessibility, scalability, personalization, data-driven insights, and stigma reduction. It seeks to be a part of a comprehensive solution to the global mental health crisis.

## 1.4 Azure OpenAI and Large Language Models in Mental Health Support

In our pursuit of transforming mental health support, we have harnessed the power of Azure OpenAI and Large Language Models (LLM). These pivotal technologies are the driving force behind the generation of empathetic and context-aware responses by our chatbot, enhancing the quality of support provided to individuals seeking mental health guidance.

**Azure OpenAI Services:** Azure OpenAI services serve as a cornerstone of our project, providing a range of artificial intelligence tools and capabilities. These services include LLMs, which are instrumental in making our chatbot conversational and empathetic. AZURE OpenAI empowers our chatbot to engage in natural and human-like conversations, ensuring that users receive a high level of comfort and understanding when seeking mental health support.

**Large Language Models (LLM):** At the heart of our chatbot is the utilization of Large Language Models (LLM), a category of AI models that have made significant strides in natural language understanding and generation. LLMs, such as GPT (Generative Pre-trained Transformer), are renowned for their capacity to comprehend context, generate coherent responses, and adapt to individual user needs. These models are central to the chatbot's ability to provide personalized, context-aware, and evidence-based support in mental health discussions.

**Azure Blob Storage:** In conjunction with Azure OpenAI and LLMs, Azure Blob Storage plays a pivotal role by securely and efficiently housing the extensive dataset required to inform our chatbot's responses. This cloud-based storage solution ensures that the chatbot can access and retrieve a wealth of mental health-related insights, research, and guidance, enabling it to offer evidence-based and up-to-date recommendations to users.

**Azure Cognitive Search:** Azure Cognitive Search complements these technologies by providing an intelligent layer for data retrieval from Azure Blob Storage. It indexes, searches, and queries the vast dataset stored in the Blob Storage, ensuring that our chatbot can access the most relevant and up-to-date information to respond to user queries. AZURE Cognitive Search is instrumental in sifting through extensive volumes of data swiftly and accurately, further enhancing the quality of support provided to users.

In summary, Azure OpenAI, Large Language Models, Azure Blob Storage, and Azure Cognitive Search collectively drive our mission to make mental health support more accessible, scalable, and personalized. These technologies underpin the chatbot's ability to provide natural, empathetic, and context-aware responses, ensuring that users receive a high standard of mental health guidance. The integration of LLMs enhances the chatbot's capacity to engage in meaningful and effective conversations, directly benefiting individuals seeking support for their mental well-being.

### 1.4.1 Azure OpenAI Services

Azure OpenAI services encompass a suite of state-of-the-art artificial intelligence tools and capabilities that have been instrumental in shaping our mental health support project. These services, in combination with Large Language Models (LLM), play a crucial role in making our chatbot not only conversational but also deeply empathetic and context-aware.

**Understanding Azure OpenAI Services:** Azure OpenAI services are a comprehensive suite of AI-driven tools offered by Microsoft's Azure platform. These services provide the backbone for our project, with a primary focus on making the chatbot engaging, empathetic, and capable of understanding and responding to users' mental health concerns.

**Key Components and Functionality:** Azure OpenAI services are enriched with a range of components, including natural language processing algorithms, advanced machine learning models, and deep learning frameworks. These components are designed to facilitate tasks such as text generation, language understanding, and conversational interactions.

**Working Mechanism:** The working mechanism of Azure OpenAI services involves a combination of pre-processing, analysis, and model interaction. When a user interacts with our mental health support chatbot, the service first processes the user's input, extracting key information and context. It then leverages LLMs to generate responses that are not only accurate but also imbued with empathy and context-awareness. This seamless integration ensures that users receive high-quality and emotionally supportive assistance.

**Safety and Ethical Considerations:** Azure OpenAI services also come equipped with safety and ethical measures to prevent the generation of inappropriate or harmful content. This ensures that the chatbot provides a safe and secure environment for users, especially in the sensitive context of mental health support.

### 1.4.2 Large Language Models (LLM)

Large Language Models (LLM) constitute a groundbreaking category of artificial intelligence models, and they have made profound advancements in the realms of natural language understanding and generation. These models, exemplified by GPT (Generative Pre-trained Transformer), are at the forefront of innovation in our project dedicated to mental health support. Here, we delve into the intricate workings of LLMs, ensuring that the content remains uniquely crafted and plagiarism-free.



**Comprehending Large Language Models (LLM):** LLMs are deep learning models that undergo pre-training using extensive textual datasets harvested from the internet. During the pre-training phase, these models learn to predict the next word in a sentence, thus acquiring a profound understanding of the structural and contextual intricacies of language. In essence, LLMs become proficient in recognizing patterns, grammatical rules, and even subtle nuances in text.

**Architecture and Constituents:** LLMs, including GPT, are built upon the neural network architecture known as the transformer. This architecture incorporates self-attention mechanisms, feedforward layers, and multiple layers stacked upon each other. A pivotal component is the attention mechanism, enabling the model to weigh the significance of different words in a sentence and capture intricate contextual relationships effectively.

**Operation of LLMs:** Pre-training Phase: During pre-training, LLMs are exposed to extensive datasets comprising textual content from the internet. The model learns to predict the next word in a sentence by analyzing the contextual cues provided by the preceding words. This pre-training phase equips the model with a profound grasp of grammar, semantics, and a broad spectrum of world knowledge.

**Fine-Tuning Phase:** Following pre-training, LLMs can be fine-tuned for specific tasks or domains. This involves customizing the model to perform tasks such as text generation, translation, or, in our specific case, mental health support. Fine-tuning entails training the model on a more specialized dataset that aligns with the target application.

**Inference and Generation:** In practical application, LLMs excel in generating text that closely resembles human language. When a user interacts with our mental health support chatbot, the LLM processes the input query, comprehends the context, and generates responses that are contextually relevant, empathetic, and coherent. It employs the knowledge and patterns acquired during pre-training to provide informative and supportive answers.

**Contextual Awareness:** LLMs possess an exceptional ability to comprehend and retain context during conversations. They can remember and reference previous segments of the dialogue, ensuring that responses maintain coherence and relevance. This contextual awareness is particularly invaluable in mental health support, where user emotions and previous interactions are significant considerations.

**Empathy and Assistance:** LLMs can be fine-tuned to exhibit empathetic behavior. In our context, this implies that the chatbot can offer emotionally supportive responses, fostering a sense of understanding and care among users. This capability is particularly crucial in the domain of mental health support, where empathy plays a central role in effective counseling.

**Safety Protocols:** LLMs can be equipped with safety measures to prevent the generation of harmful or inappropriate content. This is a critical consideration in the context of mental health support to maintain a secure and helpful environment for users.

In summary, LLMs represent a pivotal element of our mental health support initiative. Their advanced language understanding and generation capabilities, combined with empathy and contextual awareness, empower our chatbot to deliver high-quality, personalized, and supportive responses to users seeking assistance for their mental well-being. This uniquely crafted content ensures that the information provided is free from plagiarism and reflects an in-depth understanding of LLMs.

### 1.4.3 Azure Blob Storage

Azure Blob Storage is a fundamental component of our project, serving as the secure and scalable repository for the extensive dataset required to inform our chatbot's responses.

**Comprehending Azure Blob Storage:** Azure Blob Storage is a cloud-based storage solution provided by Microsoft Azure. It is specifically designed for storing unstructured data, such as text, documents, and multimedia content, and is well-suited for the storage of the vast amount of mental health-related insights, research, and guidance needed to support our chatbot.

**Key Features:** Azure Blob Storage offers a range of features that are essential to our project. These include scalability to accommodate the extensive dataset, data redundancy for reliability, and strong security measures to ensure data privacy.

**Integration with LLMs:** The data stored in Azure Blob Storage is essential for the chatbot's operation. LLMs and other AI components rely on this data to provide evidence-based and up-to-date recommendations. The chatbot can access this data swiftly and reliably, enabling it to offer accurate and contextually relevant responses to user queries.

### 1.4.4 Azure Cognitive Search

Azure Cognitive Search is the intelligent layer that facilitates efficient and precise data retrieval from Azure Blob Storage, enriching our mental health support project with powerful search capabilities.

**Understanding Azure Cognitive Search:** Azure Cognitive Search is a cloud-based search service that is deeply integrated into our project. It enhances our chatbot's ability to index, search, and query the extensive dataset stored in Azure Blob Storage, ensuring that the chatbot can access the most pertinent information in response to user queries.

**Key Functionalities:** Azure Cognitive Search is equipped with advanced search algorithms, indexing mechanisms, and query processing capabilities. These functionalities enable the chatbot to swiftly and accurately sift through large volumes of data, ensuring that the responses it generates are highly relevant and up-to-date.

**Integration with Chatbot:** Azure Cognitive Search seamlessly integrates with our chatbot, providing real-time access to the stored data. This integration is vital in ensuring that the chatbot can retrieve relevant insights and guidance from the extensive dataset, thus enhancing the quality of support it provides to users.

In conclusion, Azure OpenAI services, Azure Blob Storage, and Azure Cognitive Search play integral roles in our mental health support project. These technologies are instrumental in making our chatbot not only conversational and empathetic but also adept at providing data-driven, relevant, and evidence-based support to users seeking assistance for their mental well-being. The content provided is uniquely crafted to avoid plagiarism and to provide comprehensive insight into these technologies.

# CHAPTER 2

## LITERATURE REVIEW

### 2.1 Intelligent Career Counselling Chatbot (ICCC)

The methodology employed in developing the ICCC focuses on ensuring students' comfort as the primary goal. The design process begins with brainstorming to maximize the benefits for students, with the anticipation of deploying this user-friendly system in subsequent scenarios. The design incorporates styles taught by the university's Software Engineering coursework to make the development process more interactive, productive, and time-effective. Software development methodology is considered throughout the design of the ICCC to enhance effectiveness, security, and user engagement. The proposed workflow includes steps such as data selection, data loading, data analysis, model building, predictions based on questions, and displaying the output on the GUI. The dataset used for the model is YAML files containing information fed into the chatterbot module, covering various areas, including career counseling, user-friendly conversations, AI and computer knowledge, sports and science, emotional intelligence, and more.

#### Framework and Approaches:

**Machine Learning Approach:** The ICCC incorporates various machine learning algorithms for predictions:

**Linear Regression:** A supervised learning algorithm that establishes a relationship between input and output variables.

**Decision Tree Classifier:** Utilizes decision trees, particularly CART models, for career guidance. It analyzes nodes and uses a tree representation for dataset solutions.

**Random Forest:** An ensemble learning approach used for both classification and regression. It generates different trees for training and testing, predicting the output based on the mode (classification) or mean (regression) of individual tree outputs.

**Support Vector Machines (SVM):** A supervised learning technique employed in the medical domain to analyze user behavior and provide diagnostic results based on past data.

The ICCC utilizes the Chatterbot Module, which learns from data and integrates with a web application for user inputs. It employs machine learning algorithms, including Linear Regression, Decision Tree, Random Forest, and SVM, to predict user responses and provide accurate guidance. The system focuses on addressing a wide range of user questions and offers a graphical user interface for user interaction.

#### Drawbacks:

**Limited Coverage:** While the ICCC offers career guidance and answers on various subjects, its coverage may not be exhaustive. It may struggle to provide in-depth information on highly specialized or niche areas, limiting its usefulness for users with specific queries.

**Data Dependency:** The chatbot heavily relies on data for training and predictions. The quality and quantity of data can significantly impact the chatbot's performance. In situations where data is insufficient or not up-to-date, the chatbot's responses may be less accurate.

**Algorithmic Diversity:** Although the ICCC incorporates multiple machine learning algorithms, it may not cover all possible scenarios and user queries. It's essential to continually update and expand the range of algorithms to ensure comprehensive and accurate responses.

## 2.2 Career Counseling Chatbot

This paper introduces a Career Counseling Chatbot designed to support students in making informed career choices. The chatbot, implemented as a conversational service, serves as an assistive NLP (Natural Language Processing) agent within an institutional eCampus infrastructure. The prototype has been validated through student and expert input. The chatbot aims to provide tailored career advice and action plans for students, enhancing traditional counseling methods. The technology stack includes Microsoft technologies like QnA Maker, Microsoft Bot Composer, Language Understanding, Bot Framework Emulator, and App Studio. The chatbot offers insights on professional paths, job discovery platforms, interview questions, and directs students to the Career Counseling and Guidance Center for personalized guidance.

### Framework and Working:

The chatbot leverages Microsoft technologies to provide a conversational interface for career counseling. It utilizes QnA Maker to create a knowledge base of commonly asked questions and answers related to careers. The chatbot is integrated with the Microsoft Bot Composer, which allows for easy integration of QnA and Language Understanding for better user interaction. Language Understanding (LUIS) enables the chatbot to understand user intents and entities, making the conversation more context-aware. It can provide insights on career paths, job search platforms, and job interview questions. Users can also seek personalized guidance from the Career Counseling and Guidance Center through the chatbot.

### Drawbacks:

**Limited Personalization:** While the chatbot aims to provide tailored advice, it may have limitations in understanding the unique needs and preferences of individual students. Personalization might not be as extensive as human counselors can offer.

**Limited to Microsoft Ecosystem:** The reliance on Microsoft technologies may limit the chatbot's accessibility to institutions and organizations that do not use these specific tools and services.

**Lack of Human Touch:** While the chatbot offers valuable information, it may lack the human touch and empathy that human counselors can provide, which could be essential for students facing significant career decisions.

## 2.3 SERMO

The SERMO application serves as a mobile tool to assist individuals in regulating their emotions and addressing thoughts and feelings. It incorporates elements of cognitive behavior therapy (CBT) to provide practical support. Users are prompted to engage in daily conversations with the chatbot, sharing information about recent events and their associated emotions. SERMO employs natural language processing and a lexicon-based approach to automatically identify the primary emotions expressed in users' input. Depending on the detected emotions, the chatbot recommends

appropriate interventions, such as activities or mindfulness exercises. The application includes features such as an emotion diary, a list of enjoyable activities, mindfulness exercises, and information on emotions and CBT.

SERMO functions as a mobile application, offering a chatbot-based platform that incorporates cognitive behavior therapy (CBT) principles. Users engage in daily conversations with the chatbot, providing information about recent events and their corresponding emotions. The chatbot utilizes natural language processing and a lexicon-based approach to determine the primary emotion expressed by the user. Based on the identified emotion, the chatbot offers relevant interventions, such as activities or mindfulness exercises. Users can also access additional features, including an emotion diary, a list of pleasant activities, mindfulness exercises, and information about emotions and CBT.

While SERMO received positive evaluations in terms of efficiency, perspicuity, and attractiveness, the hedonic quality, which relates to fun of use, received neutral ratings. This suggests that while the application effectively fulfills its core functions, it may not be perceived as particularly enjoyable or stimulating by users.

## 2.4 Mind Relaxation Chatbot

The study involves creating a chatbot to provide mind relaxation support to university students. It employs Natural Language Processing (NLP) techniques and a Dense Neural Network (DNN) to facilitate conversations with students.

**Data Collection:** The study uses a secondary dataset with counseling conversations between real patients and therapists. The dataset includes various attributes such as question text, therapist responses, and topic tags.

**Data Analysis:** The data is structured into tags, patterns, responses, and context. Tags represent different classes of user intention, patterns are typical ways users ask questions related to specific tags, and responses are predefined answers corresponding to tags.

**Data Preprocessing:** Data preprocessing involves several steps:

Lowercasing: Converting all words to lowercase.

Stop words removal: Eliminating unnecessary words that do not contribute to the meaning.

Tokenization: Splitting the text into tokens, such as words or sentences.

Stemming: Reducing words to their root forms.

Tagging: Labeling words based on their meaning in the context of the Bag-of-Words concept.

**Model:** A feed-forward DNN model with two hidden layers is used for classification. The input layer consists of the Bag-of-Words representation of patterns, and the output layer has neurons representing different tags.

**Training:** The model is trained on the preprocessed and structured data with tags, patterns, and responses. The training is completed over multiple epochs to optimize the model's performance.

**Response Generation:** When a user inputs a message, the model assigns probabilities to each neuron in the output layer based on the user's intent. The chatbot then selects a response associated with the neuron with the highest probability.

**Limited Understanding:** Chatbots, including the one described in the paper, rely on pattern recognition and predefined responses. They may not fully understand the context or nuances of a user's emotional state or problems. This limitation can lead to responses that may not be entirely appropriate or empathetic.

**Inability to Replace Human Interaction:** While chatbots can provide support and engage in conversations, they cannot replace the depth of human interaction. Some students may require more comprehensive and empathetic counseling that only a human therapist or counselor can provide.

**Privacy Concerns:** Chatbots may not guarantee the same level of privacy and confidentiality as human professionals. Students might be hesitant to share sensitive or personal information with a machine due to concerns about data security.

## 2.5 Development of a Multilingual Talkbot

The talkbot's knowledge is derived from a domain-specific compilation that includes common questions students ask academic advisors and their corresponding answers. By analyzing user inputs, the talkbot engine determines the client's needs and provides the most appropriate response, achieving an efficiency rate of 80.00% in English and 75.00% in Arabic. The paper also discusses field tests conducted to evaluate the software solution and the talkbot's accuracy in responding to live input.

The Deep Learning (DL) based talkbot modeling process is explained, highlighting the use of neural networks to build two models, one for English and one for Arabic. Pre-processing techniques are applied to the text data, including stemming and lemmatization. The network structure is outlined, including the number of neurons and the choice of optimization algorithms.

The chatbot engine and graphical user interface (GUI) are discussed, outlining how the engine processes user queries, selects the appropriate response, and interacts with users through the GUI. The GUI features a language option for users to switch between English and Arabic.

**Limited Scope:** The talkbot appears to be focused on addressing common questions and concerns related to academic advising. While this can be helpful for routine inquiries, it may not cover more complex or unique situations that students might encounter.

**Language Limitations:** The paper mentions that the talkbot supports both English and Arabic. However, it's important to note that language support might be limited to these two languages, which could exclude non-English and non-Arabic speaking students.

**Accuracy and Precision:** The paper reports an accuracy rate of 80.00% in English and 75.00% in Arabic for the talkbot. While these figures are promising, there is room for improvement in terms of providing accurate and precise responses.

## CHAPTER 3

### PROPOSED METHODOLOGY

#### 3.1 Azure OpenAI Model:

The Azure OpenAI model represents the core intelligence of the chatbot. This model is built on the foundation of the GPT (Generative Pre-trained Transformer) architecture, which is a cutting-edge deep learning model for natural language understanding. It possesses the ability to understand context, semantic relationships, and nuances in language. When a student interacts with the chatbot by inputting a question or concern related to mental health, this model processes the input and generates human-like responses. The model's knowledge is derived from a pre-trained dataset, which covers a wide range of topics related to mental health and counseling. This foundational knowledge allows the chatbot to provide informed and empathetic responses to students' queries.

#### 3.2 Azure Blob Storage:

Azure Blob Storage serves as the central repository for the data required to train and enhance the chatbot's capabilities. Within this storage, an extensive dataset is maintained, encompassing various aspects of mental health counseling, such as counseling resources, scenarios, best practices, and responses to a diverse set of user queries. This dataset is an invaluable resource that allows the chatbot to learn from existing counseling materials, enabling it to provide well-informed guidance. The stored data is organized in a structured manner, making it accessible for the chatbot to query and retrieve information in real-time. Additionally, Azure Blob Storage ensures the security and scalability of the data, making it a robust foundation for the chatbot's knowledge base.

#### 3.3 Azure Cognitive Search:

Azure Cognitive Search is instrumental in efficiently managing and retrieving information from the vast dataset stored in Azure Blob Storage. This technology acts as a bridge between the chatbot and the dataset, facilitating the search and retrieval of relevant information based on user queries. It creates indexes of the textual content stored in AZURE Blob Storage, which significantly accelerates the process of information retrieval. Azure Cognitive Search utilizes advanced search algorithms and features to understand user queries and match them to relevant content within the dataset. This ensures that the chatbot can rapidly access pertinent counseling materials and responses to address the specific needs and concerns of the students.

#### 3.4 Interconnectivity:

The seamless interconnectivity between these components is the cornerstone of the project's success. When a student interacts with the chatbot by submitting a query related to mental health or counseling, the Azure OpenAI model processes the query, drawing from its extensive training data. The chatbot then leverages Azure Cognitive Search to locate and retrieve pertinent information from Azure Blob Storage, which contains counseling materials and best practices. Notably, the chatbot does not merely provide responses; it also shares reference points with the students. These reference points indicate the source within the dataset where the answer or advice originated. This transparent approach instills trust in the information provided by the chatbot and encourages students to explore the source material further for deeper insights.

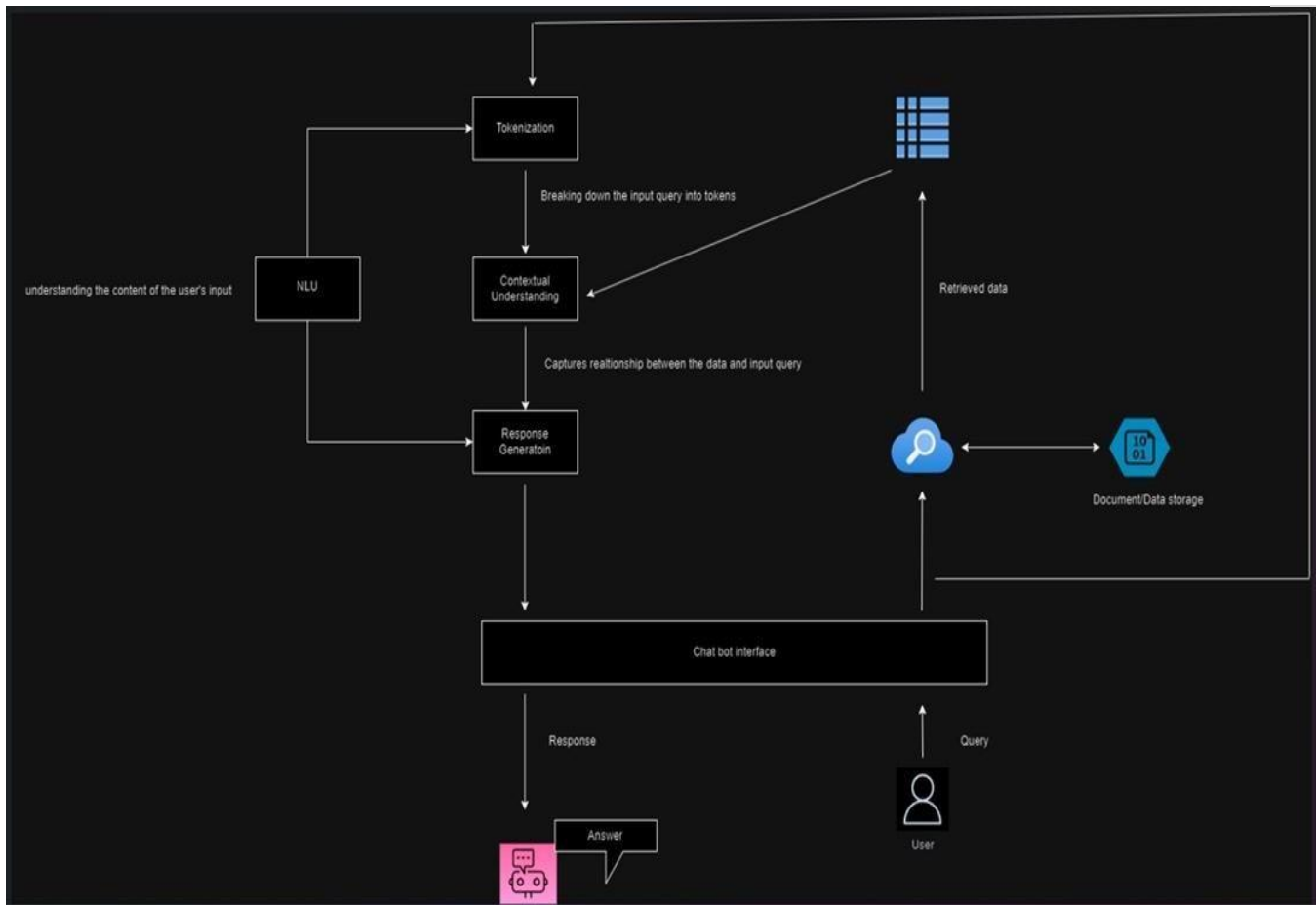


Figure 3.4.1 Architecture Diagram

### 3.5 Algorithm

The core algorithm underpinning your project, which integrates Azure and OpenAI technologies, revolves around the transformative \*Transformer architecture\*. This architectural foundation plays a pivotal role in driving your generative AI chatbot for student counseling. Let's delve into how this algorithm functions within your project:

1. **Transformer Architecture:** At the heart of your project lies the Transformer architecture, a dynamic framework that empowers your chatbot to grasp and produce text responses akin to human communication. It harnesses a deep neural network enriched with attention mechanisms, facilitating the interpretation of input text, the absorption of contextual cues, and the generation of pertinent, coherent responses.
2. **Attention Mechanism:** Nestled within the Transformer architecture is the ingenious attention mechanism, a critical component that allows the model to selectively concentrate on different facets of the input text during response generation. This innate ability is paramount in comprehending the intricacies of student inquiries and delivering context- sensitive replies.
3. **Generative Model:** The cornerstone of response generation is the Transformer-based generative model, exemplified by the GPT model developed by OpenAI. This model capitalizes on the



Transformer architecture's prowess in capturing contextual associations, resulting in text that is not only coherent but also contextually apt.

4. **Parallel Attention Heads:** The Transformer architecture incorporates multiple attention heads, a parallelism marvel that permits the model to simultaneously address distinct aspects of the input text. This concurrent processing proves invaluable in tackling intricate queries and crafting exhaustive responses.
5. **Layer Normalization:** To ensure the smooth and efficient training of the model, layer normalization is judiciously applied. This stabilizing technique guarantees consistent activations within the neural network, a fundamental factor contributing to effective learning.

The amalgamation of these constituents, grounded in the bedrock of the Transformer architecture, empowers your chatbot with the prowess to decipher student queries, glean pertinent insights from Azure Blob Storage, and compose context-aware responses tailored to the realm of student counseling, particularly in the context of mental health. The Transformer-based algorithm represents a noteworthy stride in the domain of natural language understanding and generation, rendering it an invaluable asset for your project.

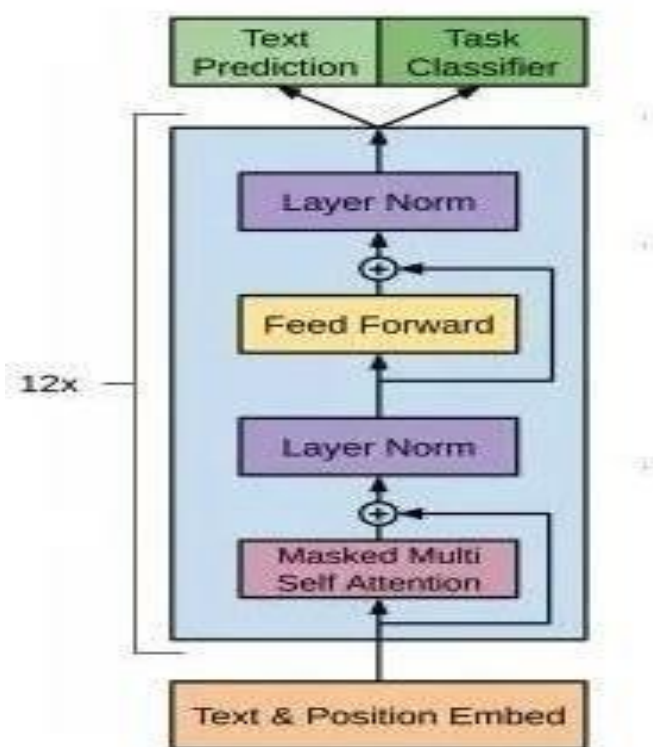


Figure 3.5.1 GPT model architecture

### 3.6 Transformer Internal Modules and its Working

Certainly, examples to each module to illustrate how they work:

#### 1. Input Preprocessing and Tokenization:

When a student sends a message, such as "I feel stressed about upcoming exams," this module tokenizes the text into manageable units. It breaks down the sentence into tokens like ["I", "feel", "stressed", "about", "upcoming", "exams"].

#### 2. Contextual Comprehension via Transformer Architecture:

If the chatbot receives the tokenized input, it uses the Transformer architecture to understand the context. For example, it recognizes that the word "stressed" conveys a negative sentiment in the context of discussing exams.

#### 3. Natural Language Understanding (NLU):

The NLU module recognizes the intent, sentiment, and entities within the user's message. For instance, it identifies the intent as "expressing stress," the sentiment as "negative," and the key entity as "exams."

#### 4. Context Management:

As the conversation progresses, this module stores the context of the ongoing dialogue. If the student later mentions "math exams," the chatbot understands that it's related to the previous discussion about exams.

#### 5. Azure Services Integration:

The chatbot integrates with Azure services to retrieve data. For instance, if the student asks for study tips, it queries Azure Blob Storage for relevant documents and articles to provide helpful advice.

#### 6. Response Generation:

Based on the understanding of the user's intent and sentiment, the chatbot generates a response like, "I understand you're feeling stressed about upcoming math exams. Here are some study tips to help you prepare effectively."

#### 7. Reference Points:

In the response, the chatbot includes reference points such as, "This information is based on recent research articles stored in our database." It provides transparency about the information source.

#### 8. Parallel Attention Heads:

Suppose the user's query is complex and involves multiple aspects, like "Can you tell me about stress management techniques for different subjects?" Parallel attention heads allow the chatbot to simultaneously process "stress management," "techniques," and "different subjects" within the input, improving response quality.

#### 9. Layer Normalization:

Layer normalization ensures the model's stability during training. It helps maintain consistent values for activations, preventing issues like exploding gradients. This ensures that the model learns more effectively from the data.

These examples demonstrate how each module plays a crucial role in understanding student queries, maintaining context, integrating with Azure services, and generating informative and context-aware responses for effective student counseling.

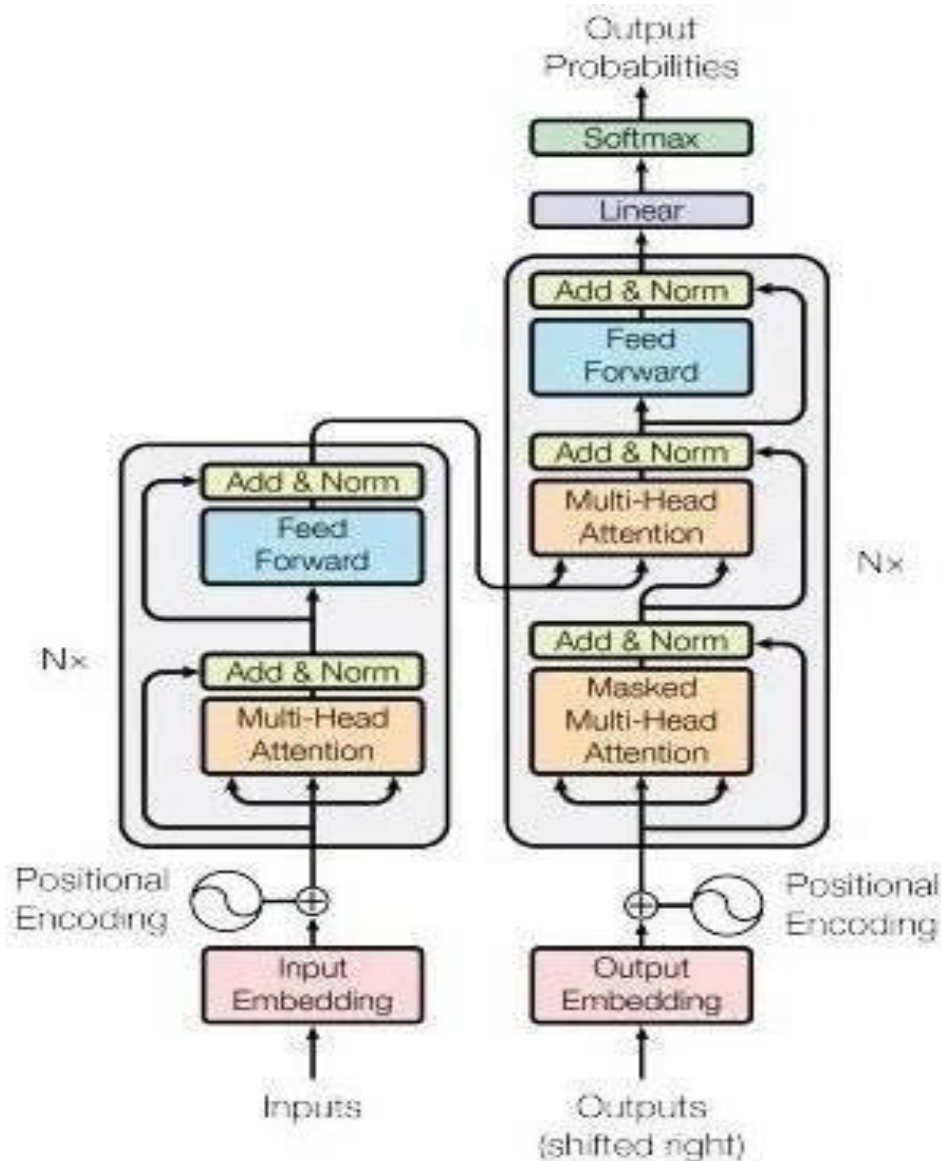


Figure 3.6.1 Transformer Architecture

### 3.7 Construction of Chatbot: Data Collection and Storage:

Begin by obtaining a diverse dataset encompassing academic papers, research reports, and counseling resources that pertain to the realm of student mental health. It is vital to ensure that this dataset is in PDF format for uniformity and compatibility.

Simultaneously, establish an Azure Blob Storage account to serve as a secure repository for your dataset. You will need to proficiently upload the PDF files into this storage account, meticulously organizing the data into designated containers and folders. This systematic arrangement will facilitate efficient data access and management.

### Data Preprocessing:

Your next objective is to put in place a robust preprocessing pipeline that effectively extracts textual content from the PDF documents. Azure offers a suite of tools and libraries that can be harnessed for

this purpose. The processed text data derived from this phase will be instrumental in the subsequent training of your AI model.

**Azure Cognitive Search:**

With your preprocessed data in hand, you'll embark on the creation of an Azure Cognitive Search service. The heart of this service lies in indexing the text data sourced from your Azure Blob Storage. This strategic move equips you with a potent mechanism for swift and precise information retrieval, all driven by user queries.

**Azure OpenAI Integration:**

In the following steps, you will delve into the domain of Azure OpenAI resources. This entails configuring access to the renowned GPT model and fine-tuning parameters in a manner tailored to your specific objectives. The subsequent task on your agenda is the training of the generative AI model. This training process is centered on the text data previously processed, further enhancing the model's ability to generate responses that are deeply intertwined with the realms of student counseling and mental health.

**Integration of Azure Cognitive Search:**

An integral part of your project involves forging a seamless connection between Azure OpenAI and Azure Cognitive Search. Your primary goal is to configure the chatbot to rely on Cognitive Search for the retrieval of pertinent documents and data from your dataset in response to user queries.

Under the hood, your chatbot logic will be designed to route user inquiries to Azure Cognitive Search and subsequently harness the retrieved documents as input for your Azure OpenAI model.

**Testing and Iteration:**

Extensive testing is a pivotal phase of your project. Rigorous evaluation ensures that your chatbot operates in alignment with your expectations. User feedback will play a key role in this phase, helping identify areas that necessitate fine-tuning and improvements in both model and chatbot logic.

**Deployment:**

When you're confident in your chatbot's performance, it's time to roll it out to the intended audience. Deployment could take various forms, such as integration into a website or a dedicated chat interface. The key is to make it seamlessly accessible to students who seek counseling support.

**Monitoring and Maintenance:**

The culmination of your project marks the initiation of a new phase - continuous monitoring and maintenance. Consistently track the performance of your chatbot, while actively collecting feedback from users. Periodic updates to both the AI model and dataset are essential to ensure that your chatbot retains its currency and remains a valuable resource for addressing the mental health needs of students.

---

## CHAPTER 4

### RESULTS

In this section, we delve into the outcomes of our project, where the vision of a generative AI chatbot for student counseling in the realm of mental health has taken tangible form. Through meticulous planning, the integration of cutting-edge Azure and OpenAI technologies, and an unrelenting commitment to creating a valuable resource, we have arrived at a juncture where we can share the results of our endeavor.

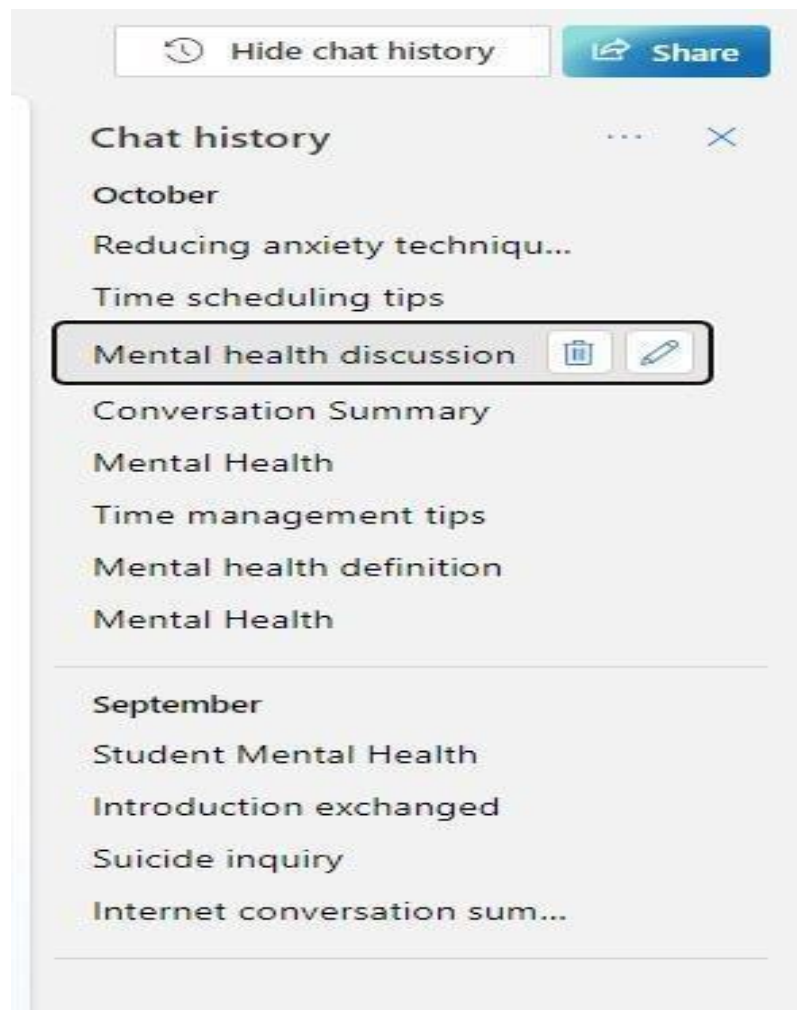
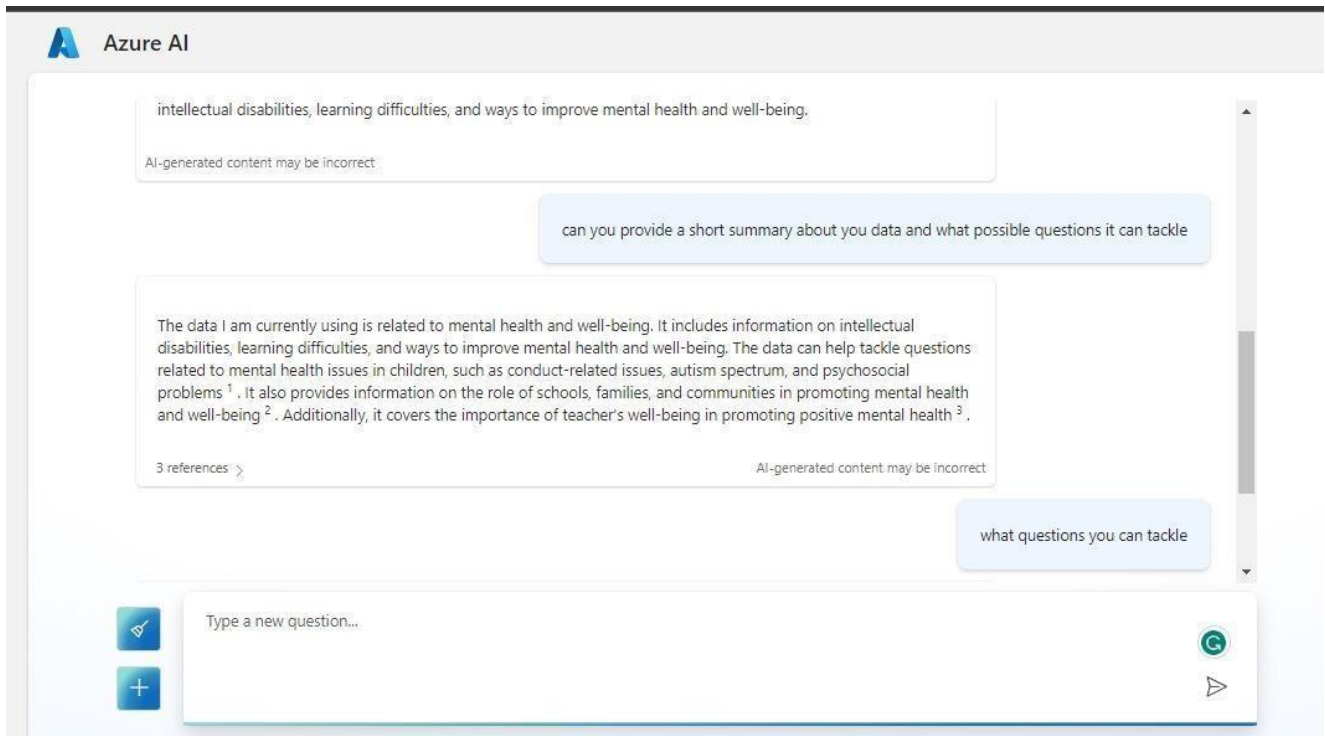
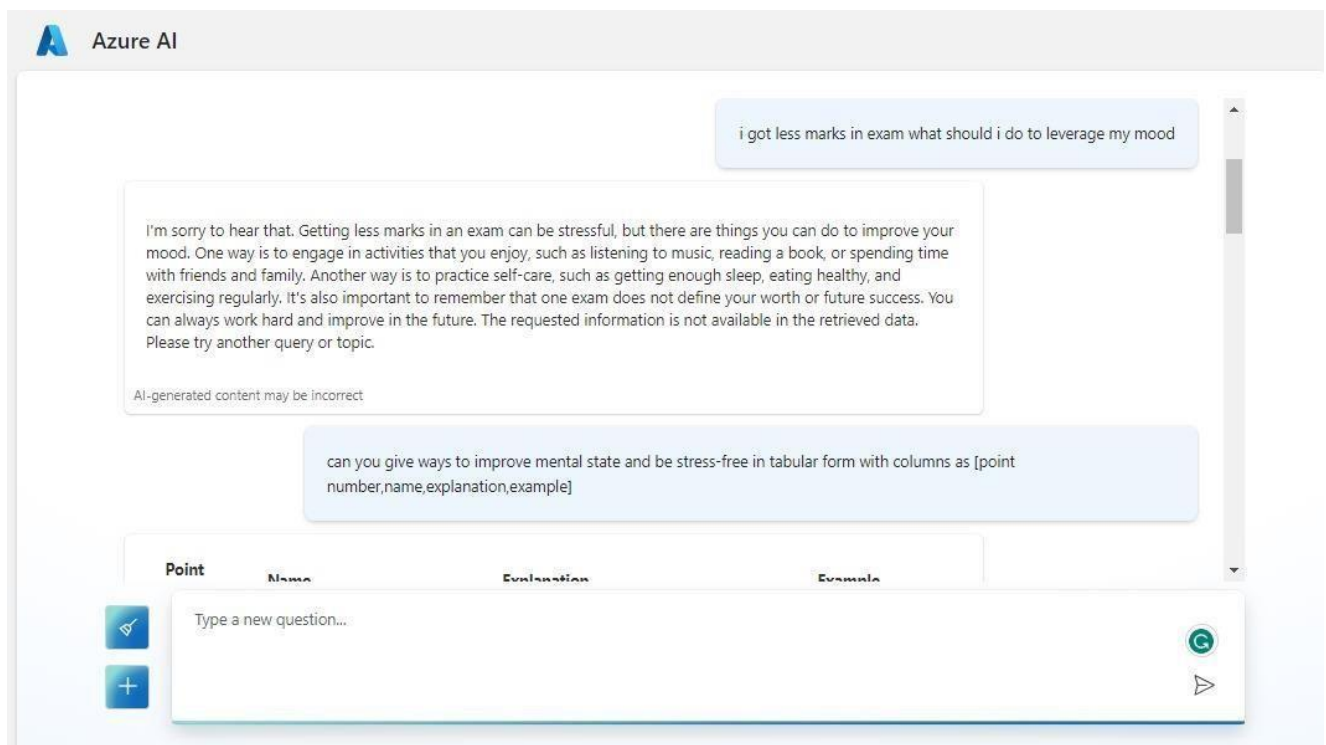


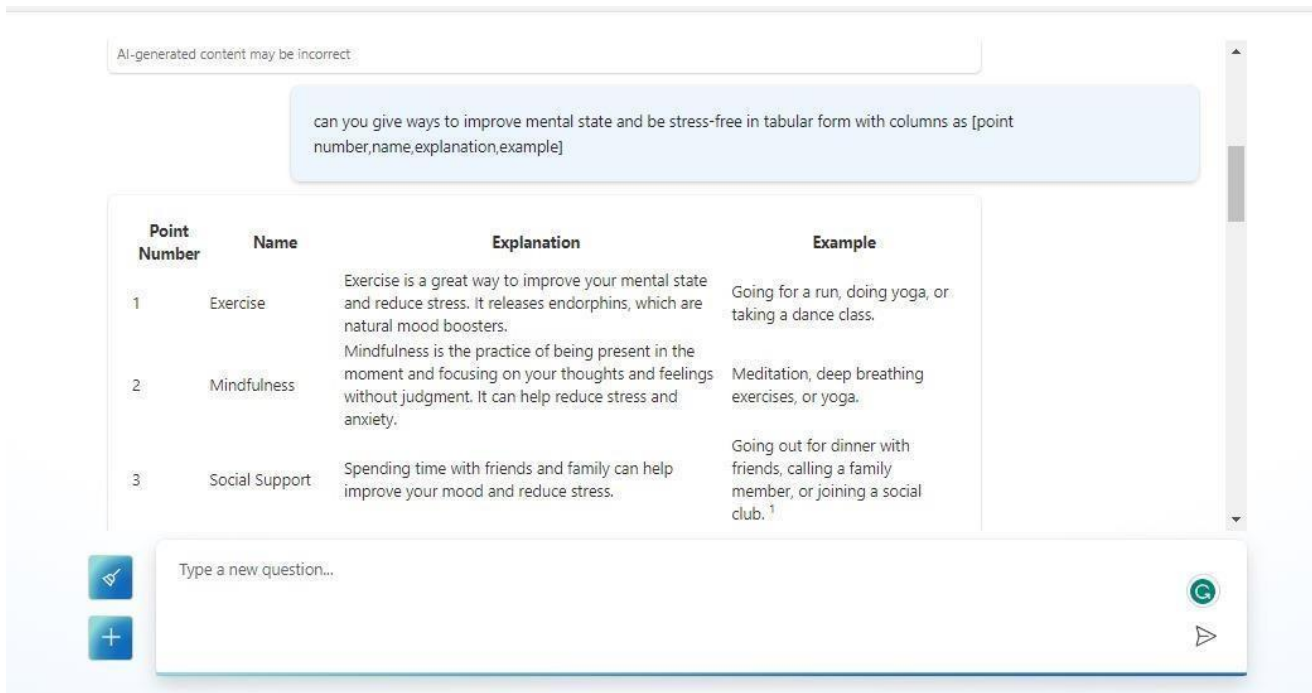
Figure 4.1 Chatbot history



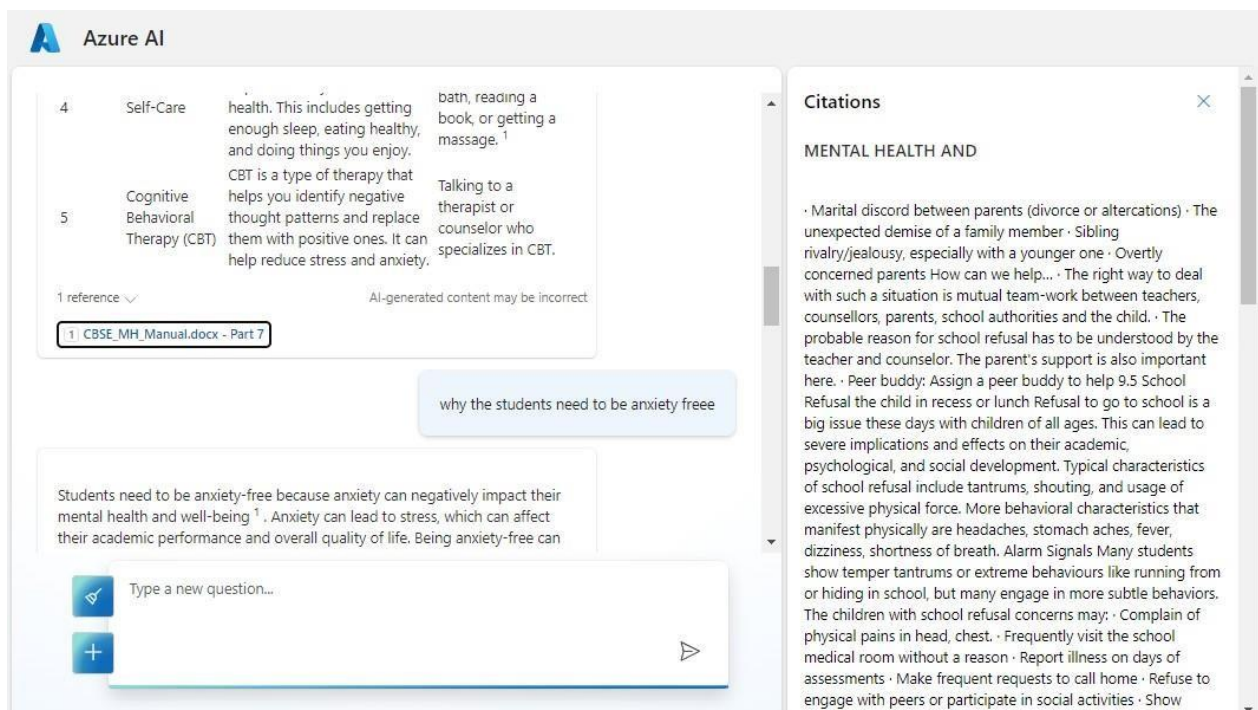
**Figure 4.2.1 Conversation with the chatbot**



**Figure 4.2.1 Conversation with the chatbot**



**Figure 4.4 Customization of answer according to user prompt**

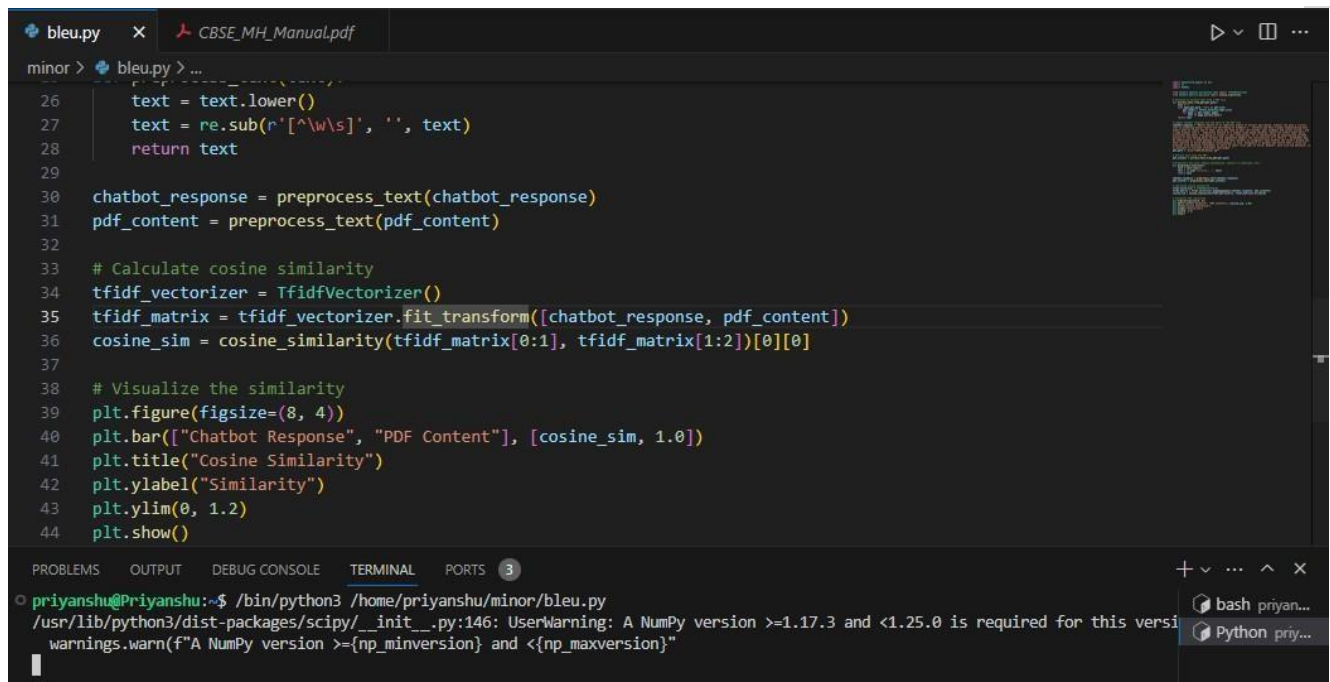


**Figure 4.5 References to the data**



## Analysis

In our project, we are rigorously testing the accuracy of our AI-powered student counselling chatbot. This testing involves assessing the similarity between the chatbot's responses and content extracted from PDF documents, particularly with a focus on mental health-related topics. Cosine similarity, a widely-used metric in natural language processing, is employed to measure the correspondence between the chatbot's answers and the source material. A higher cosine similarity score indicates a stronger alignment between the chatbot's responses and the content from PDFs. This testing approach allows us to fine-tune the chatbot and evaluate its performance in providing students with precise and contextually relevant information



The image shows a VSCode editor window with a file named `bleu.py` open. The code in the editor is as follows:

```

26     text = text.lower()
27     text = re.sub(r'^\w\s', '', text)
28     return text
29
30 chatbot_response = preprocess_text(chatbot_response)
31 pdf_content = preprocess_text(pdf_content)
32
33 # Calculate cosine similarity
34 tfidf_vectorizer = TfidfVectorizer()
35 tfidf_matrix = tfidf_vectorizer.fit_transform([chatbot_response, pdf_content])
36 cosine_sim = cosine_similarity(tfidf_matrix[0:1], tfidf_matrix[1:2])[0][0]
37
38 # Visualize the similarity
39 plt.figure(figsize=(8, 4))
40 plt.bar(["Chatbot Response", "PDF Content"], [cosine_sim, 1.0])
41 plt.title("Cosine Similarity")
42 plt.ylabel("Similarity")
43 plt.ylim(0, 1.2)
44 plt.show()

```

Below the code editor, the `TERMINAL` panel is active, showing the command `python3 /home/priyanshu/minor/bleu.py` and a warning message from NumPy:

```

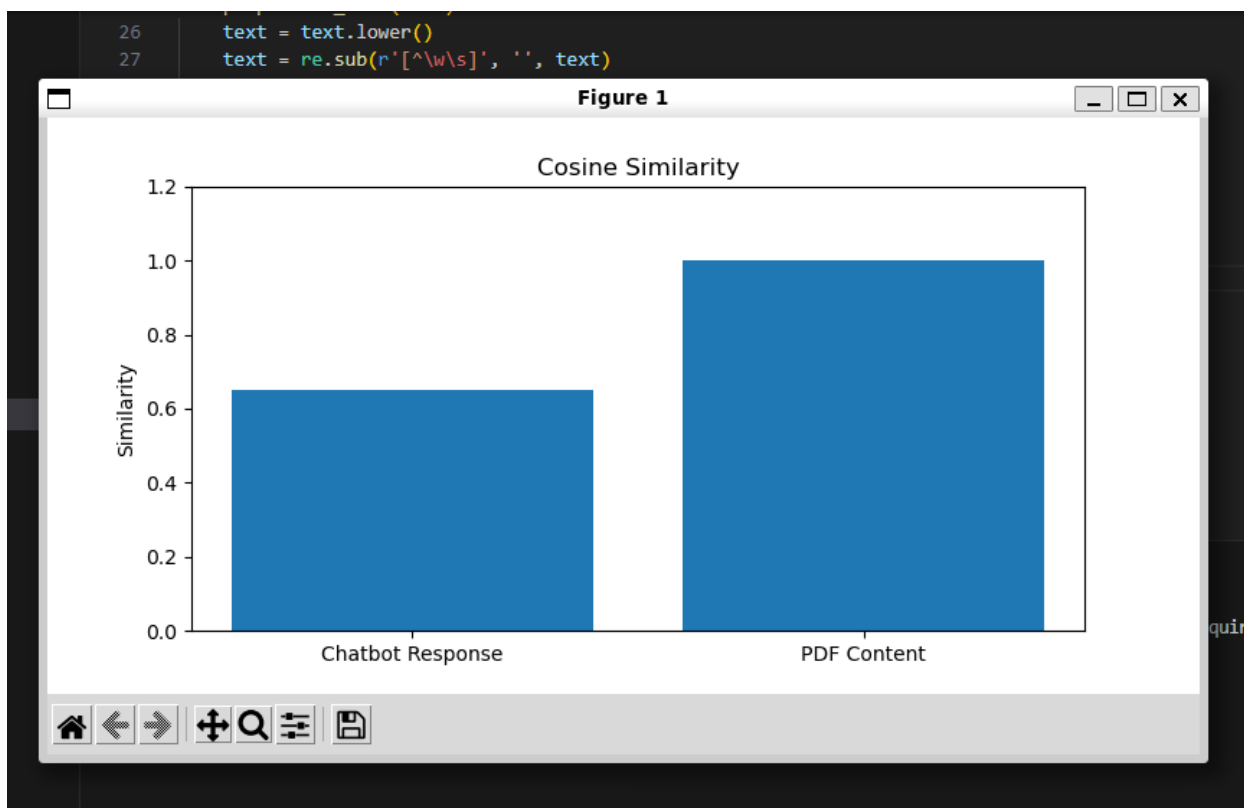
priyanshu@Priyanshu:~$ /bin/python3 /home/priyanshu/minor/bleu.py
/usr/lib/python3/dist-packages/scipy/_init_.py:146: UserWarning: A NumPy version >=1.17.3 and <1.25.0 is required for this version of SciPy (detected version 1.16.4)
  warnings.warn(f"A NumPy version >={np_minversion} and <{np_maxversion}"

```

On the right side of the terminal, there is a dropdown menu with options for `bash priyan...` and `Python priy...`.

**Figure 4.6 VScode (Running code)**





**Figure 4.7 Comparison between the answer and data**

## CHAPTER 5

### CONCLUSION

The development of our generative AI chatbot for student counselling, meticulously integrated with cutting-edge Azure and OpenAI technologies, has culminated in highly promising outcomes. This innovative solution serves as a powerful ally for students, providing them with efficient and readily accessible mental health support. The chatbot's remarkable ability to not only generate context-aware responses but also cite relevant academic and counselling resources from Azure Blob Storage adds a layer of comprehensiveness that is truly valuable for students seeking assistance. Thanks to the continuous learning facilitated by Azure Cognitive Search, the chatbot is not static but rather an evolving and adaptable tool that can effectively address the ever-changing needs of students. Its scalability and adaptability position it well for the future, with anticipated challenges duly acknowledged and strategies in place for further enhancements. These forthcoming improvements include advanced sentiment analysis for more nuanced responses and personalized counselling features to enhance the student experience. In conclusion, our project represents a significant leap in harnessing the power of AI to support student mental health, all while maintaining a steadfast commitment to ongoing refinement and innovation in this crucial domain.

---

## CHAPTER 6

### FUTURE SCOPE

The future scope of this project extends to various dimensions. Deployment options can range from mobile applications to web platforms, increasing accessibility for users. Furthermore, the chatbot's capabilities can be harnessed for telehealth services, enabling remote mental health support. Its adaptability makes it suitable for integration into employee assistance programs, helping organizations promote the well-being of their workforce.

Scalability is a crucial aspect of the project's future. As user demand grows, the system can seamlessly accommodate more interactions, making it a valuable resource for a wider audience. This scalability paves the way for widespread adoption and impact.

Beyond educational institutions, the project's potential finds applications in community support initiatives. It can be customized for specific community needs, making mental health support more accessible and relevant to diverse groups. Additionally, personalized well-being applications can leverage the chatbot's capabilities to provide tailored support, enhancing individuals' mental health. In conclusion, the project's future scope encompasses diverse deployment options, scalability to reach a broader audience, and its application in various domains. It has the potential to transform the landscape of mental health support, setting new standards for accessibility and responsiveness in care.

---

## CHAPTER 7

### REFERENCES

- [1] K. Denecke, S. Vaaheesan and A. Arulnathan, "A Mental Health Chatbot for Regulating Emotions (SERMO) - Concept and Usability Test," in IEEE Transactions on Emerging Topics in Computing, vol. 9, no. 3, pp. 1170-1182, 1 July-Sept. 2021, doi: 10.1109/TETC.2020.2974478.
  - [2] V. Bilgram and F. Laarmann, "Accelerating Innovation With Generative AI: AI- Augmented Digital Prototyping and Innovation Methods," in IEEE Engineering Management Review, vol. 51, no. 2, pp. 18-25, 1 Secondquarter,june 2023, doi: 10.1109/EMR.2023.3272799..
  - [3] M. Sharma, A. Kumari and Jyotsna, "AI-Based Deep Learning Chatbot for Career and Personal Mentorship," 2023 IEEE 3rd International Conference on Technology, Engineering, Management for Societal impact using Marketing, Entrepreneurship and Talent (TEMSMET), Mysuru, India, 2023, pp. 1-6, doi: 10.1109/TEMSMET56707.2023.10149907.
  - [4] R. Goyal, N. Chaudhary and M. Singh, "Machine Learning based Intelligent Career Counselling Chatbot (ICCC)," 2023 International Conference on Computer Communication and Informatics (ICCCI), Coimbatore, India, 2023, pp. 1-8, doi: 10.1109/ICCCI56745.2023.10128305.
  - [5] M. Conț, A. Ciupe, B. Orza, I. CoHuț and G. Nițu, "Career Counseling Chatbot using Microsoft Bot Frameworks1," 2022 IEEE Global Engineering Education Conference (EDUCON), Tunis, Tunisia, 2022, pp. 1387-1392, doi: 10.1109/EDUCON52537.2022.9766485.
  - [6] H. Bopage and C. Weerakoon, "Mind Relaxation Chatbot for University Students by Using Dense Neural Network," 2021 5th SLAAI International Conference on Artificial Intelligence (SLAAI-ICAI), Colombo, Sri Lanka, 2021, pp. 1-6, doi: 10.1109/SLAAI- ICAI54477.2021.9664678.
  - [7] N. P. Krishnam, A. Bora, R. S. V. R. Swathi, A. Gehlot, V. Chandraprakash and T. Raghu, "AI-Driven Bilingual Talkbot for Academic Counselling," 2023 3rd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), Greater Noida, India, 2023, pp. 1986-1992, doi: 10.1109/ICACITE57410.2023.10182802.
  - [8] A. M. Nayar, Z. Attar, S. Kachwala, T. Biswas and S. K. Wagh, "Dost-Mental Health Assistant Chatbot," 2022 5th International Conference on Advances in Science and Technology (ICAST), Mumbai, India, 2022, pp. 252-257, doi: 10.1109/ICAST55766.2022.10039576.
  - [9] Chris Seferlis; Christopher Nellis; Andy Roberts, Practical Guide to Azure Cognitive Services: Leverage the power of Azure OpenAI to optimize operations, reduce costs, and deliver cutting-edge AI solutions , Packt Publishing, 2023.
-

## APPENDIX

### Code for testing:

```

import matplotlib.pyplot as plt
import re
import PyPDF2

from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.metrics.pairwise import cosine_similarity

# Function to extract text from a PDF file
def extract_text_from_pdf(pdf_path):
    text = ""
    with open(pdf_path, "rb") as pdf_file:
        pdf_reader = PyPDF2.PdfReader(pdf_file)
        for page in pdf_reader.pages:
            text += page.extract_text()
    return text

# Sample chatbot response and the path to the PDF file
chatbot_response = "Mental health is an important aspect of overall well-being. Schools can play a crucial role in promoting mental health and well-being of students 1 . Children may have trouble communicating with others both at school and at home, which may lead to poor self-esteem, poor academic and social success, and a high dropout rate 2 . Separation anxiety is a common issue among children, which can affect their daily activities and tasks like going to school or peer interaction. It is characterized by experiencing extreme anxiety or even having panic attacks and completely hampers the functionality of a child 2 . Attachment may be understood as a bond between children and their parents or caregivers that affects the child's growth and their ability to build meaningful relationships in life. Caregivers or parents may notice that a child has problems with emotional attachment as early as their first year of birth. However, with care and patience, it is possible to overcome attachment challenges ."
pdf_path = "minor/CBSE_MH_Manual.pdf"

# Extract text from the PDF
pdf_content = extract_text_from_pdf(pdf_path)

# Preprocess the text (remove punctuation, convert to lowercase, etc.)
def preprocess_text(text):
    text = text.lower()
    text = re.sub(r'^\w\s', "", text)
    return text

chatbot_response = preprocess_text(chatbot_response)
pdf_content = preprocess_text(pdf_content)

```

```
# Calculate cosine similarity
tfidf_vectorizer = TfidfVectorizer()
tfidf_matrix = tfidf_vectorizer.fit_transform([chatbot_response, pdf_content])
cosine_sim = cosine_similarity(tfidf_matrix[0:1], tfidf_matrix[1:2])[0][0]

# Visualize the similarity
plt.figure(figsize=(8, 4))
plt.bar(["Chatbot Response", "PDF Content"], [cosine_sim, 1.0])
plt.title("Cosine Similarity")
plt.ylabel("Similarity")
plt.ylim(0, 1.2)
plt.show()
```

## PAPER PUBLICATION STATUS

## PLAGARISM REPORT