

Analyzing Emotional Trends in Time-Stamped Chats

A PROJECT REPORT

Submitted by

AADAV SRINIVAS (2116210701001)

ADARSH R (2116210701012)

in partial fulfillment for the award of the

degree of

BACHELOR OF ENGINEERING

in

COMPUTER SCIENCE AND ENGINEERING



RAJALAKSHMI ENGINEERING COLLEGE

ANNA UNIVERSITY, CHENNAI

MAY 2024

RAJALAKSHMI ENGINEERING COLLEGE, CHENNAI

BONAFIDE CERTIFICATE

Certified that this Thesis titled “**Analyzing Emotional Trends in Time-Stamped Chats**” is the bonafide work of “**AADAV SRINIVAS (2116210701001), ADARSH R (2116210701012)**” who carried out the work under my supervision. Certified further that to the best of my knowledge the work reported herein 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 on this or any other candidate.

SIGNATURE

Dr. S.Vinod Kumar

PROJECT COORDINATOR

Professor

Department of Computer Science and Engineering

Rajalakshmi Engineering College

Chennai - 602 105

Submitted to Project Viva-Voce Examination held on_____

Internal Examiner

External Examiner

ACKNOWLEDGMENT

First, we thank the almighty god for the successful completion of the project. Our sincere thanks to our chairman **Mr. S. Meganathan B.E., F.I.E.**, for his sincere endeavor in educating us in his premier institution. We would like to express our deep gratitude to our beloved Chairperson **Dr. Thangam Meganathan Ph.D.**, for her enthusiastic motivation which inspired us a lot in completing this project and Vice Chairman **Mr. Abhay Shankar Meganathan B.E., M.S.**, for providing us with the requisite infrastructure.

We also express our sincere gratitude to our college Principal, **Dr. S. N. Murugesan M.E., PhD.**, and **Dr. P. KUMAR M.E., PhD, Director computing and information science , and Head Of Department of Computer Science and Engineering** and our project coordinator **Dr. S.Vinod Kumar**, for her encouragement and guiding us throughout the project towards successful completion of this project and to our parents, friends, all faculty members and supporting staffs for their direct and indirect involvement in successful completion of the project for their encouragement and support.

AADAV SRINIVAS

ADARSH R

ABSTRACT

In the dynamic landscape of digital communication, understanding emotional expressions in online dialogues is increasingly crucial. This project explores the complex dynamics of emotions across various digital conversations, ranging from casual chats among friends to professional collaborations. With a specific focus on the interconnected relationship between chatbots, Large Language Models (LLMs), and Natural Language Processing (NLP), our goal is to create a comprehensive tool for analyzing emotions within digital conversations, particularly those occurring on messaging apps and online platforms. The motivation for such a tool is from the ever-changing nature of online communication, where text-based interactions are prevalent. By uncovering patterns and trends in emotional dynamics over time, our project aims to offer valuable insights into how emotions influence conversations. This understanding holds significance for individuals, researchers, and companies. Additionally, the project explores tracking historical data to observe changes in emotions over time for individuals. This includes chat logs, scripts, and public reviews. Users can upload these conversations, pose questions, and receive comprehensive analyses of emotional expressions, sentiments, and key insights. Ultimately, the overarching goal of this project is to contribute to mental health awareness by providing timely support to individuals expressing distress

TABLE OF CONTENTS

CHAPTER NO.	TITLE	
	ABSTRACT	iii
	LIST OF TABLES	v
	LIST OF FIGURES	vii
1.	INTRODUCTION	1
	1.1.PROBLEM STATEMENT	
	1.2.OBJECTIVES	
	1.3.SCOPE OF THE PROJECT	
	1.1. NEED FOR PROJECT	
	1.1.RESOURCES	
1.	LITERATURE SURVEY	4
	1.1.SURVEY	
	1.1.LIMITATIONS	

1.	SYSTEM REQUIREMENTS	6
	1.1.HARDWARE REQUIREMENTS	
	1.2.SOFTWARE REQUIREMENTS	
1.	SYSTEM DESIGN	11
	1.1.EXISTING SYSTEM	
	1.1.PROPOSED SYSTEM	
	1.1.PROCESS FLOW OF PROPOSED METHODOLOGY	12
1.	SYSTEM IMPLEMENTATION	25
	1.1.INTRODUCTION	
	1.2.EMOTION DETECTION	
	1.3.USER INTERFACE	
	1.4.CHATBOT INTEGRATION DATA UPLOAD AND STORE	
6.	RESULTS AND DISCUSSIONS	
	6.1.USER INTERFACE DESIGN	
	6.2.SUMMARY	
7.	CONCLUSION	29
	REFERENCES	31

	LIST OF FIGURES	
FIGURE NO	TITLE	PAGE NO
4 . 4	PROCESS FLOW OF THE PROPOSED SYSTEM	5
4.4.1	PDF DOC INTO PROMPTING	6
5.3	USER INTERFACE	8
5.4	HTML CODE FOR TEMPLATES	11
5.5	CODE FOR DATA UPLOAD	25
6.2.1	USER INTERFACE WINDOW	
6.2.2	PDF UPLOAD WNDOW	
6.2.3	PDF PROCESSING WINDOW	
6.2.4	SEARCH BY VOICE	
6.2.5	RESULTS GENERATED	

LIST OF ABBREVIATIONS

NLP	Natural language processing
BTM	Biterm topic model
WLT M	Weighted labeled topic model
XETM	X-term emotion-topic model
NLU	Natural language understanding
SENN	Semantic emotion neural network
CNN	Convolution neural network
LLM	Large language models
AI	Artificial intelligence
UX	User experience
HTML	Hypertext markup language
CSS	Cascading style sheets
API	Application programming interface

CHAPTER 1

INTRODUCTION

In today's digital age, communication has transcended traditional boundaries, evolving into a rapid and multifaceted exchange of thoughts and emotions. Whether it's friends catching up through instant messaging, students collaborating on academic projects, or colleagues brainstorming in virtual workspaces, the digital realm is the new frontier for human interaction. As we navigate this vast and dynamic landscape, understanding how people convey and perceive emotions in these digital dialogues becomes increasingly vital.

Our project aims to delve deep into the heart of these interactions, exploring the patterns and trends in emotional expressions as they unfold over time. We seek to illuminate the subtle and overt ways emotions shape our digital conversations, shedding light on how different contexts and relationships influence these exchanges. By examining various scenarios—be it casual chats between friends, serious academic discussions, or professional communications—we aim to uncover the emotional undercurrents that define our digital communication era.

Central to our exploration is the role of advanced technologies like chatbots, Large Language Models (LLMs), and Natural Language Processing (NLP). These intelligent conversational agents, powered by sophisticated AI algorithms, are not just tools but pivotal participants in our daily digital dialogues. They facilitate seamless interactions across multiple domains, enhancing user experience and bridging gaps between human and machine communication.

Artificial Intelligence, particularly through the development of LLMs and NLP, has revolutionized the way we interact with digital platforms. LLMs, such as OpenAI's GPT-4, are capable of understanding and generating human-like text, making them indispensable in various applications—from customer service chatbots to personal virtual assistants. Their ability to process and interpret large volumes of text data allows them to respond with remarkable accuracy and relevance, creating more natural and engaging interactions..

1. PROBLEM STATEMENT

The prevalent expression of stress in interpersonal chats poses a challenge, intensified by the extensive textual data. Identifying emotions and understanding context become formidable tasks due to the sheer volume of text. The project addresses the necessity for a tool that navigates through extensive textual datasets, facilitating the identification of emotions in interpersonal chats. Additionally, it aims to provide a nuanced understanding of context specifically in product reviews or comments. The tool further explores tracking historical data to observe changes in emotions over time for an individual.

1. OBJECTIVES

Develop a comprehensive tool that analyzes emotions within digital conversations, specifically those occurring on messaging apps and online platforms. Our primary aim is to understand the intricacies of how individuals express their feelings in the fast-paced realm of instant communication. We try to unveil patterns and 2 trends in emotional dynamics over time, enhancing our comprehension of the evolving nature of emotions in various conversational settings. This can be helpful for individuals and researchers in predicting how people feel. Ultimately, our goal is to contribute to a more nuanced understanding of emotional expressions in the digital age.

4. NEED FOR PROJECT

In the age of digital communication, we often chat online instead of talking. This way of communicating is always changing. Understanding how emotions shift in these chats over time gives us important insights into how feelings impact our conversations. People use chats for learning, working together, and making decisions. Looking at the emotions in these conversations helps us really get what users feel, making our interactions more meaningful. This can be helpful for individuals and researchers in predicting how people feel. Companies can also use this to understand the emotions of their employees by checking their chat messages. The primary goal is to contribute to mental health awareness and provide timely support to individuals expressing distress.

5. SCOPE OF THE PROJECT

Our project not only addresses challenges in emotion detection, NLP, and AI but also strives to enhance user empowerment. By introducing innovative models like WLTm and XETm, the project contributes to overcoming emotion detection challenges, fostering a more empathetic understanding. Emphasizing advancements in speech processing, especially for the Polish language, it aligns with the convergence of computer science and human languages. Additionally, the project evaluates widely-used NLUs, providing valuable insights for the software engineering community. Beyond these advancements, the tool's unique feature includes the ability to predict suicide thoughts in people, adding a crucial dimension to its societal impact and reinforcing its commitment to mental health awareness.

6. RESOURCES

This project has been developed through widespread secondary research of accredited manuscripts, standard papers, business journals, white papers, analysts' information, and conference reviews. Significant resources are required to achieve an efficacious completion of this project.

The following prospectus details a list of resources that will play a primary role in the successful execution of our project:

- A properly functioning workstation (PC, laptop, net-books etc.) to carry out desired research and collect relevant content.
- Unlimited internet access.
- Unrestricted access to the university lab in order to gather a variety of literature including academic resources (for e.g. Prolog tutorials, online programming examples, bulletins, publications, e-books, journals etc.), technical manuscripts, etc. Prolog development kit in order to program the desired system and other related software that will be required to perform our research.

CHAPTER 2

LITRETURE SURVEY

The literature survey deals with various aspects of emotion detection and natural language processing (NLP) to provide a comprehensive understanding of the challenges and advancements in these fields. Firstly, there are challenges in emotion detection within short messages on social media platforms, primarily due to feature sparsity and limitations in traditional word-level algorithms. The importance of topic extraction is highlighted, with a critique of the existing biterm topic model (BTM) for its time-consuming nature and unsuitability for emotion prediction. In response, two models, the weighted labeled topic model (WLTM) and the X-term emotion- topic model (XETM), are introduced to address these issues, incorporating mapping between emotions and topics, as well as leveraging user scores for emotion-topic probabilities.

Moving to the interdisciplinary domain of natural language processing (NLP) and deep learning, the convergence of computer science, artificial intelligence, and human languages is explored. NLP challenges, including speech recognition, natural language understanding, and language generation, are discussed as focal points within this domain. The literature then shifts focus to a specific application, namely elevating speech processing and recognition capabilities for the Polish language. The research involves an in-depth exploration of Polish phonemes, statistical language modeling methods, and applications such as speech and speaker recognition and speech translation.

The exploration extends to the realm of software engineering, where the use of chatbots is increasingly prevalent. The paper emphasizes the crucial role of Natural Language Understanding (NLU) platforms in this domain. A gap is identified, highlighting the lack of studies on NLU performance in the software engineering domain. The research aims to address this gap by evaluating four widely-used

NLUs, namely IBM Watson, Google Dialogflow, Rasa, and Microsoft LUIS, in software engineering tasks related to repository data and Stack Overflow posts. IBM Watson emerges as the top performer, excelling in intents classification, and the study provides an in-depth analysis of NLU features and confidence score thresholds.

Lastly, the growing importance of emotion recognition in artificial intelligence and human-computer interaction is addressed. The focus is on detecting six basic emotions from text, with an acknowledgment of the challenges posed by real-world data and the limitations of manual feature design. The proposed solution introduces a novel neural network architecture, Semantic-Emotion Neural Network (SENN), which leverages pre-trained word embeddings, Bidirectional Long Short-Term Memory (BiLSTM), and Convolutional Neural Network (CNN) to enhance emotion recognition. Experimental results demonstrate the superior performance of SENN compared to baseline methods.

In the context of this comprehensive literature survey, the project, "EmotionAnalysisPDF," aims to revolutionize user interaction with textual content in PDF documents. The project focuses on extracting and analyzing emotions embedded in chat conversations within PDFs, with the primary goal of empowering users, including students, professionals, researchers, and general users. The transformative tool seamlessly uploads PDF documents and performs emotion analysis on chat-based content, aligning with the advancements and challenges highlighted in the literature survey across emotion detection, NLP, and software engineering domains.

Innovative models like WLTM and XETM integrate labeled data, mapping emotions to topics and refining emotion-topic probabilities for accurate analyses. In the realm of NLP and deep learning, our project aligns with advancements in speech

processing, emphasizing the efficacy of statistical language modeling in precise speech recognition for the Polish language. Acknowledging the surge in chatbot use in software engineering, our research evaluates NLUs, highlighting IBM Watson as a standout performer, especially in intents classification. The project navigates literature challenges, addressing the demand for substantial training corpora, pre-processing inconsistencies, and the risk of overfitting in deep learning models. Introducing SENN, a Semantic-Emotion Neural Network, the project leverages pre-trained word embeddings, BiLSTM, and CNN architectures, outperforming baseline methods in emotion recognition. Despite challenges, our project aims to push the boundaries, contributing to advancements in emotion.

2.2 LIMITATIONS

- Feature Sparsity and Algorithm Limitations: Emotion detection faces challenges in short messages due to feature sparsity and the limitations of traditional wordlevel algorithms.
- Speech Processing for the Polish Language
- Large Corpus Requirement: The need for a large corpus of texts for generating and training exact vectors may be a limitation. 6
- Pre-processing Inconsistencies: Potential impact on performance due to inconsistencies in pre-processing.
- Lack of Human Empathy: Recognition process may lack human empathy and understanding.
- Pre-processing Consistency Impact: Inconsistencies in pre-processing may impact performance.
- Adaptability to Trends: Potential difficulty in adapting to evolving trends in emotional expressions.

CHAPTER 3

SYSTEM SPECIFICATIONS

1. HARDWARE REQUIREMENTS

Processor	: Intel Core i3 or above
Speed	: 1.8 GHz or above
System Type	: 32bit/64 bit
Hard Disk	: 5 GB or more
Key Board	: Standard Windows Keyboard
Memory (RAM)	: 512 MB or above
Monitor	: SVGA or higher resolution

2. SOFTWARE REQUIREMENTS

Operating System	: Windows, Linux, or MacOS
Programming Language	: Python 3.6 or above, HTML, CSS, JavaScript
Web Framework	: React.js, Node.js, Three.js, Web3.js.
Database	: MongoDB
Web Server	: Apache or Nginx

CHAPTER 4

PROJECT DESCRIPTION

1. EXISTING SYSTEM

The tools we currently have for understanding emotions in online chats usually use basic methods to guess feelings. But these methods might miss the subtle ways people express emotions. Also, these tools can't analyze emotions quickly, which makes them less helpful in fast conversations. Plus, the AI and NLP in these tools might not be very good at detecting emotions accurately. Consequently, there is a need for a more sophisticated and comprehensive tool that can effectively analyze emotions in digital conversations, taking into account the evolving nature of communication channels and the complexities of human emotion.

2. PROPOSED SYSTEM

The proposed system will leverage advanced AI and NLP techniques to provide a comprehensive analysis of emotions within digital conversations. By employing state-of-the-art sentiment analysis algorithms, coupled with machine learning models trained on vast datasets of conversational data, the system will offer enhanced 9 accuracy in emotion detection. Real-time analysis capabilities will ensure timely insights into evolving emotional dynamics during conversations. Furthermore, the system will incorporate adaptive learning mechanisms to continuously improve its emotion detection capabilities based on user feedback and evolving linguistic patterns.

3. PROCESS FLOW OF THE PROPOSED METHODOLOGY

This section explores the structural layout, data flow, and integration points, outlining the backbone of our innovative system. By meticulously detailing how different components interact and communicate, it ensures a robust, scalable, and efficient architecture, essential for delivering seamless user experiences.

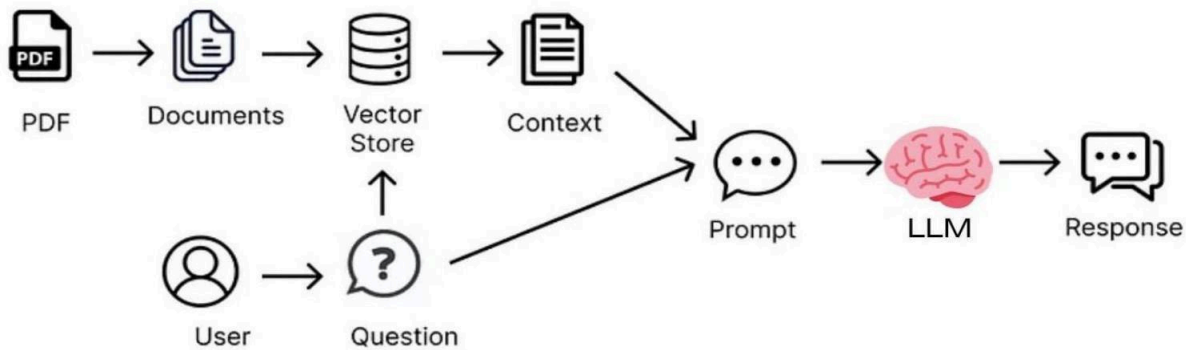


Figure 4.4 PROCESS FLOW OF THE PROPOSED SYSTEM

PDF Uploading: Users initiate the interaction by uploading PDF documents containing relevant textual information.

Text Processing and Segmentation: The uploaded PDFs are processed to extract text, which is then segmented into manageable chunks for efficient handling.

Vector Store: Processed text chunks are stored in a vector store, optimizing storage and retrieval processes for large-scale text data

User Query or Prompt: Users input questions or prompts, which trigger the context- building process within the system.

Context Analysis: The system analyzes user queries and comprehends the context from the uploaded documents, forming a clear understanding of the user's information needs.

Prompt Refinement: Based on the context, a refined prompt is generated, capturing the essence of the user's query and the document context.

Language Model (LLM) Processing: The refined prompt is processed by the Language Model (LLM), a sophisticated tool capable of understanding complex human language patterns.

Response Generation: The LLM generates a tailored response based on the refined prompt and the context, ensuring the response is relevant and accurate

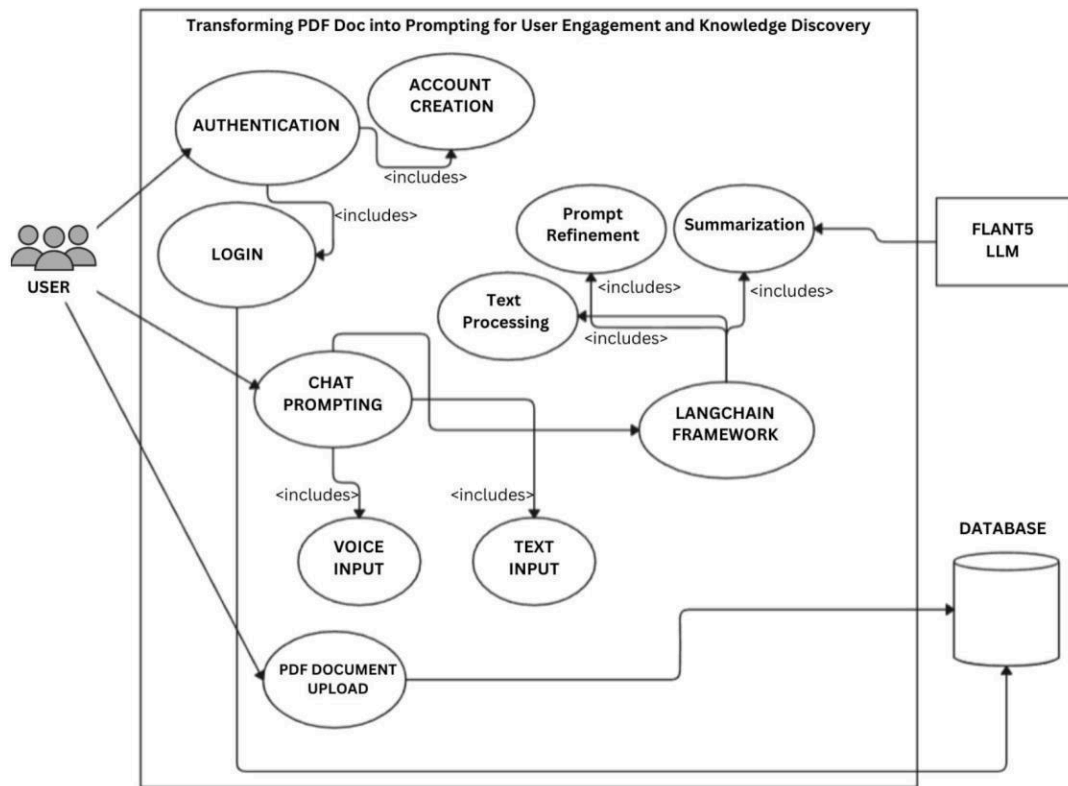


Figure 4.4.1 PDF DOC INTO PROMPTING

CHAPTER 5

IMPLEMENTATION

1. INTRODUCTION

In the ever-changing world of online communication, human emotions play a big role in how we interact. This project aims to explore digital conversations to find out how emotions show up and change. Today, online communication is everywhere in our personal lives, studies, and work. It's crucial to understand how emotions work in digital chats, whether it's chatting with friends, discussing ideas at work, or seeking help online. In this fast-paced digital era, where conversations happen in the blink of an eye, grasping the intricacies of emotions online is more important than ever. From the excitement of sharing good news to the comfort of seeking advice during tough times, emotions color every message we send and receive. By analysing these digital dialogues, we can gain valuable insights into human behavior, fostering deeper connections and enhancing the quality of our online interactions

2. EMOTION DETECTION

In our approach to emotion detection, we employ a comprehensive methodology designed to transform text data into numerical vectors for efficient storage and retrieval. Central to this methodology is the utilization of embeddings and the FAISS library, which establish a robust vector store capable of handling large volumes of textual data effectively. By leveraging embeddings, we encode textual information into numerical representations, enabling semantic similarity comparisons and facilitating rapid retrieval of relevant information. To enhance natural language understanding and improve chatbot interactions, we implement the Google/Flan-t5 Large Language Model (LLM). This powerful model provides contextually relevant responses by processing and generating text based on the input provided. By incorporating the torch library, particularly PyTorch, an open-source machine learning framework built on the Torch library, we ensure flexibility and scalability in our implementation, allowing for efficient training and deployment of machine

learning models. Furthermore, our methodology encompasses essential libraries such as `dotenv`, `PyPDF2`, `embeddings`, `ConversationalRetrievalChain`, `ConversationBufferMemory`, `HuggingFaceHub`, `CharacterTextSplitter`, and `FAISS`. These libraries play pivotal roles in various stages of our emotion detection pipeline, from data, they form the backbone of our system, enabling us to achieve accurate and reliable emotion detection capabilities in digital conversations.

Step 1: Preprocessing

1. Clean and tokenize the text.
2. Remove stop words and punctuation.
3. Normalize text (e.g., convert to lowercase).

Step 2: Feature Extraction

1. Extract features like word frequencies and sentiment scores.

Step 3: Model Training

1. Train a model on labeled data.
1. Use machine learning or deep learning techniques to classify emotions based on extracted features.

Step 4: Model Evaluation

Evaluate model performance using appropriate metrics such as accuracy, precision, recall, and F1-score

Step 5: Model Adjustment

1. Adjust parameters or features if necessary based on evaluation results.

Step 6: Application

1. Apply the trained model to new data for emotion analysis.

End Algorithm.

3. USER INTERFACE

Incorporating Streamlit for the user interface in our project has significantly enhanced the accessibility and usability of our applications. With its intuitive Python syntax and seamless integration with data science and machine learning libraries, Streamlit has enabled us to quickly develop interactive and visually appealing web apps. Its simplicity allows for rapid prototyping and iteration, empowering us to efficiently experiment with different features and functionalities. Overall, leveraging Streamlit for our user interface has been instrumental in delivering a user-friendly and engaging experience

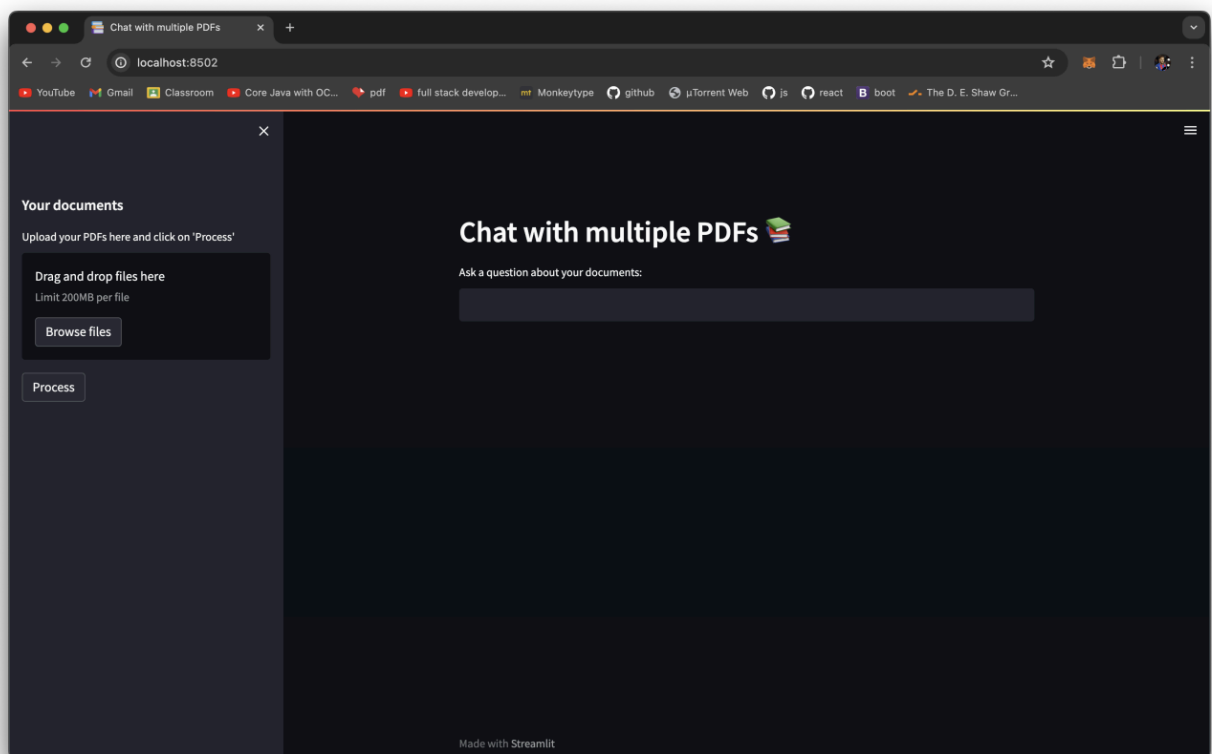
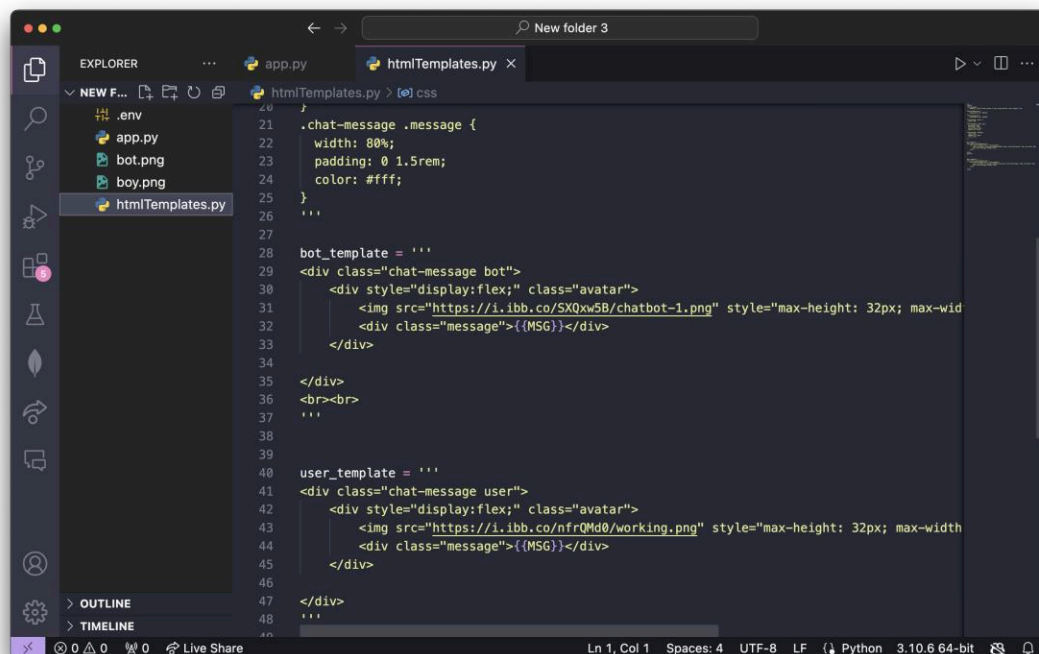


Figure 5.3 USER INTERFACE

User experience (UX) is a cornerstone of the project, as it directly impacts the efficiency and user-friendliness for the user. Users are central to the system, and their experience is a priority. Here, the assessing various aspects of user experience: The system is been designed an intuitive user interface that simplifies the interaction. It guides users through the document upload, query submission, and report retrieval processes. The user interface is user-friendly, reducing the learning curve for users.

4. CHATBOT INTEGRATION

In integrating the chatbot, we prioritize user engagement and emotional awareness. Users can query specific emotions and receive instant feedback, promoting understanding. The chatbot is compatible across platforms for flexible engagement. It offers informative content on emotional well-being and may integrate external emotional intelligence APIs for enhanced responses. Overall, our goal is to provide users with valuable insights and support for managing emotions effectively.

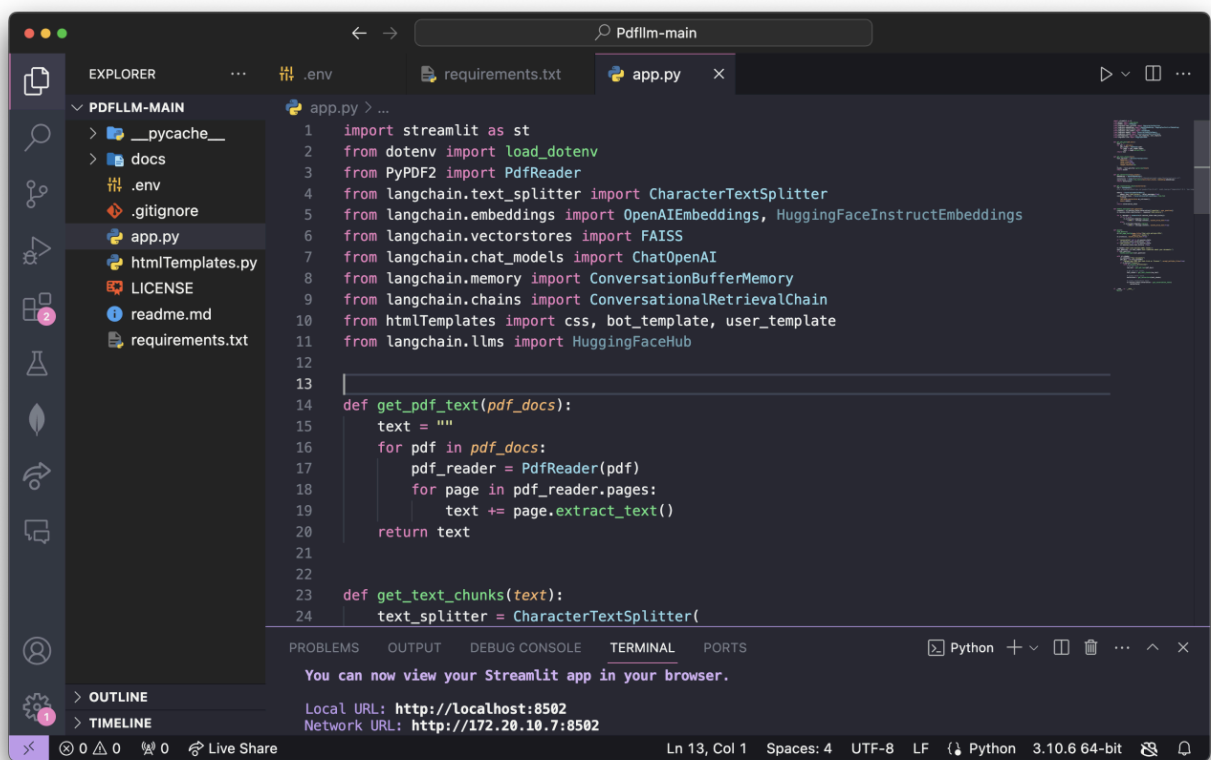


The image shows a code editor window with a dark theme. The Explorer panel on the left shows a file structure with files like .env, app.py, bot.png, boy.png, and htmlTemplates.py. The main editor area displays the content of htmlTemplates.py, which contains CSS and HTML code for chat messages. The code defines a .chat-message class and two templates: bot_template and user_template. The bot_template includes an avatar image and a message div. The user_template includes an avatar image and a message div. The status bar at the bottom indicates the current line and column (Ln 1, Col 1), the number of spaces (4), the encoding (UTF-8), the line feed (LF), the interpreter (Python 3.10.6 64-bit), and a Live Share icon.

```
21 .chat-message .message {
22   width: 80%;
23   padding: 0 1.5rem;
24   color: #fff;
25 }
26 '''
27
28 bot_template = '''
29 <div class="chat-message bot">
30   <div style="display:flex;" class="avatar">
31     {{MSG}}</div>
33   </div>
34 </div>
35 <br><br>
36 '''
37
38
39 user_template = '''
40 <div class="chat-message user">
41   <div style="display:flex;" class="avatar">
42     {{MSG}}</div>
44   </div>
45 </div>
46
47 </div>
48 '''
```

5. DATA UPLOAD AND STORAGE

Our data storage strategy ensures security, efficiency, and user convenience. We implement user authentication for data upload, use MongoDB for structured storage, and support multiple formats like PDFs. Users can select and download chat messages as PDFs for offline reference, enhancing accessibility and productivity.



The screenshot shows a VS Code editor window titled 'Pdfllm-main'. The Explorer sidebar on the left displays the project structure, including files like `__pycache__`, `docs`, `.env`, `.gitignore`, `app.py`, `htmlTemplates.py`, `LICENSE`, `readme.md`, and `requirements.txt`. The main editor area shows the `app.py` file with the following Python code:

```
1 import streamlit as st
2 from dotenv import load_dotenv
3 from PyPDF2 import PdfReader
4 from langchain.text_splitter import CharacterTextSplitter
5 from langchain.embeddings import OpenAIEmbeddings, HuggingFaceInstructEmbeddings
6 from langchain.vectorstores import FAISS
7 from langchain.chat_models import ChatOpenAI
8 from langchain.memory import ConversationBufferMemory
9 from langchain.chains import ConversationalRetrievalChain
10 from htmlTemplates import css, bot_template, user_template
11 from langchain.llms import HuggingFaceHub
12
13
14 def get_pdf_text(pdf_docs):
15     text = ""
16     for pdf in pdf_docs:
17         pdf_reader = PdfReader(pdf)
18         for page in pdf_reader.pages:
19             text += page.extract_text()
20     return text
21
22
23 def get_text_chunks(text):
24     text_splitter = CharacterTextSplitter(
```

At the bottom of the editor, the TERMINAL panel shows the message: "You can now view your Streamlit app in your browser." with the following URLs:

```
Local URL: http://localhost:8502
Network URL: http://172.20.10.7:8502
```

The status bar at the bottom indicates the current file is at line 13, column 1, with 4 spaces, UTF-8 encoding, LF line endings, Python 3.10.6 64-bit interpreter, and a Live Share session.

algorithm to evolve neural networks for an AI player. Through iterative training and evolutionary processes, the AI learned to control paddles and respond dynamically to the game's dynamics, showcasing adaptive gameplay strategies. The results unveiled a progressive evolution in the AI's gameplay proficiency, from initial random movements to refined, strategic gameplay over multiple generations. This endeavor provided valuable insights into the potential of evolutionary algorithms for training AI agents in gaming environments, highlighting the adaptability and learning capabilities of NEAT-based neural networks. Furthermore, it underscored the importance of continuous refinement and

CHAPTER 6

RESULTS AND DISCUSSION

6.1 USER INTERFACE DESIGN

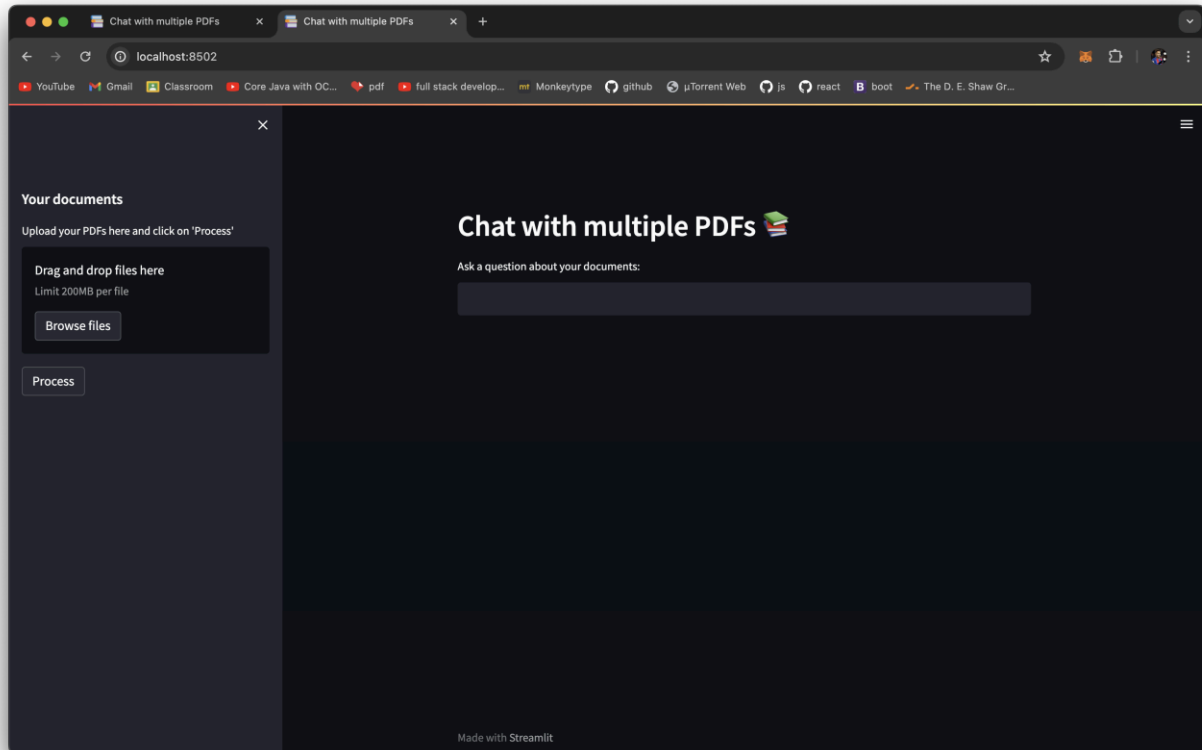


Figure 6.2.1 USER INTERFACE WINDOW

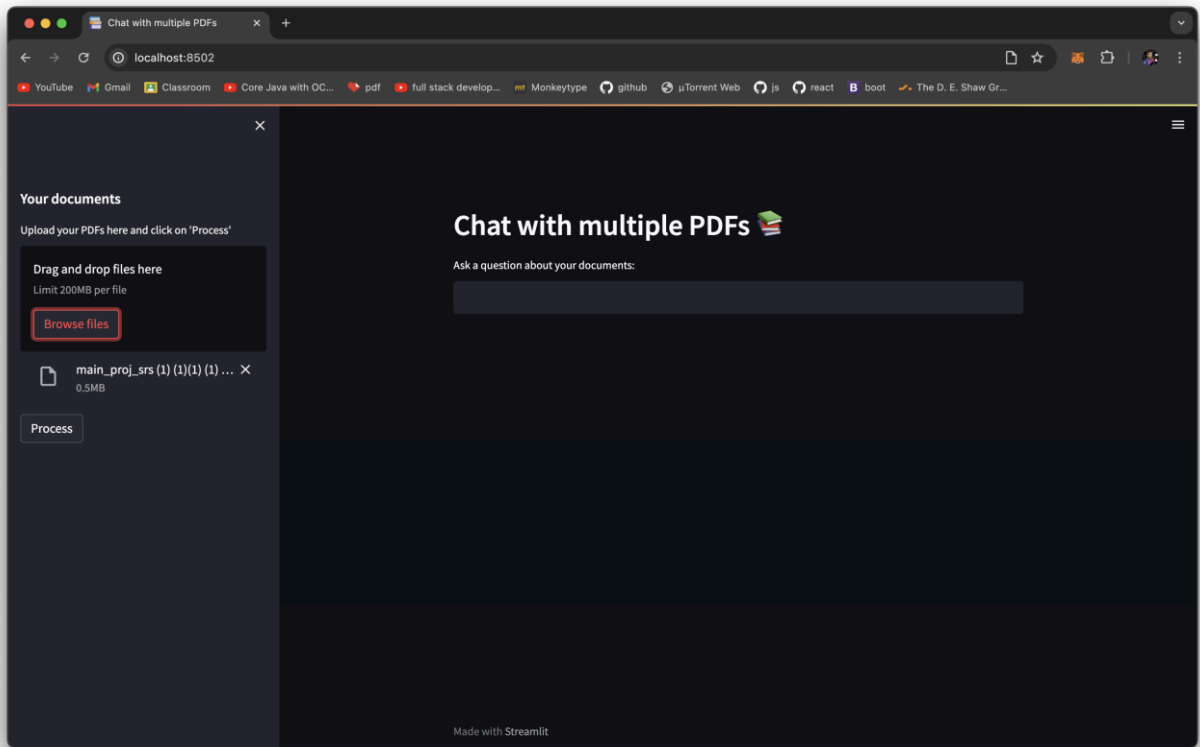


Figure 6.2.2 PDF UPLOAD WINDOW

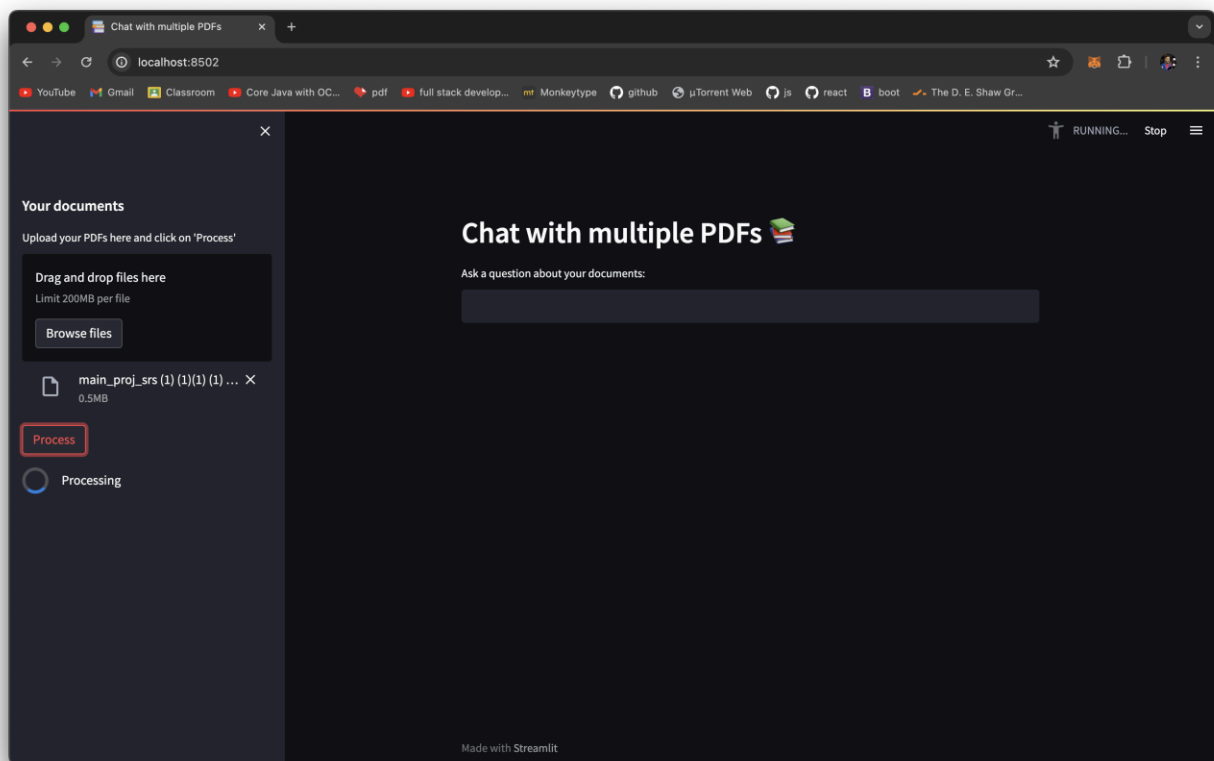
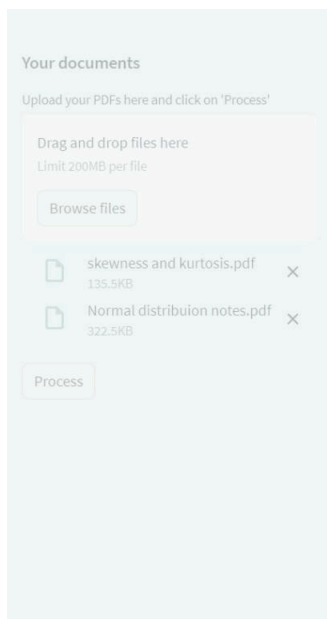



Figure 6.2.3 PDF PROCESSING WINDOW



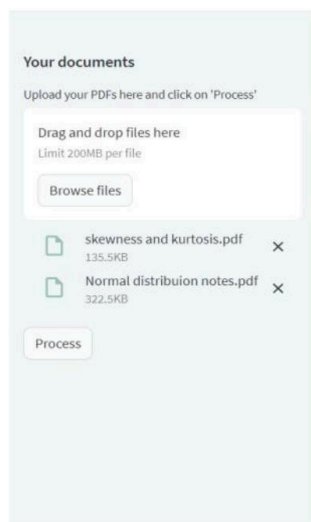
Chatbot 

Ask a question about your documents

 Search by Voice

Listening...

Figure 6.2.4 SEARCH BY VOICE



Chatbot 

Ask a question about your documents

 Search by Voice

Ask a question about your documents:

What is skewness and kurtosis?



What is skewness and kurtosis?



Select message 1 for download



Skewness is a measure of the symmetry of a distribution. It tells us whether the distribution is symmetric or skewed to the left or right. A distribution is symmetric if the values are uniformly distributed around the mean. Skewness can be positive (right-skewed), negative (left-skewed), or zero (symmetric).

Kurtosis is a measure of the peakness or flatness of a distribution. It tells us about the shape of the distribution's tails compared to a normal distribution. A distribution with high kurtosis has heavier tails and a sharper peak, while a distribution with low kurtosis has lighter tails and a flatter peak. Kurtosis is often measured using Pearson's coefficient, with a value of 3 for a normal distribution.

Figure 6.2.5 RESULTS GENERATED

3. SUMMARY

A user-friendly UI/UX has been developed, ensuring that even inexperienced users can comfortably navigate the system without technical assistance. Separate interfaces have been created to cater to the unique needs of each actor involved in the system, providing comprehensive functionality to all users. To compare the effectiveness of different techniques, a comparative analysis was conducted, measuring the accuracy of each approach. The application heavily relies on Artificial Intelligence, which plays a vital role in addressing various issues and vulnerabilities present in the current system.

APPENDIX

SOURCE CODE:

```
import streamlit as st

from dotenv import load_dotenv

from PyPDF2 import PdfReader

from langchain.text_splitter import CharacterTextSplitter

from langchain.embeddings import OpenAIEmbeddings,
HuggingFaceInstructEmbeddings

from langchain.vectorstores import FAISS

from langchain.chat_models import ChatOpenAI

from langchain.memory import ConversationBufferMemory

from langchain.chains import ConversationalRetrievalChain

from htmlTemplates import css, bot_template, user_template

from langchain.llms import HuggingFaceHub

import torch

device = torch.device("cuda:0" if torch.cuda.is_available() else "cpu")

def get_pdf_text(pdf_docs):

    text = ""

    for pdf in pdf_docs:

        pdf_reader = PdfReader(pdf)

        for page in pdf_reader.pages:
```

```

        text += page.extract_text()

    return text

def get_text_chunks(text):
    text_splitter = CharacterTextSplitter(
        separator="\n",
        chunk_size=1000,
        chunk_overlap=200,
        length_function=len
    )
    chunks = text_splitter.split_text(text)

    return chunks

@st.cache_resource()
def get_vectorstore(text_chunks):
    embeddings = OpenAIEmbeddings()

    # embeddings =
    HuggingFaceInstructEmbeddings(model_name="hkunlp/instructor-xl")

    vectorstore = FAISS.from_texts(texts=text_chunks, embedding=embeddings)

    return vectorstore

@st.cache_resource()
def get_conversation_chain(_vectorstore):
    llm = ChatOpenAI()

    #llm=HuggingFaceHub(repo_id="tiiuae/falcon-40b",
    model_kwargs={"temperature":0.5, "max_length":512})

    memory = ConversationBufferMemory(

```

```

memory_key='chat_history', return_messages=True)

conversation_chain = ConversationalRetrievalChain.from_llm(

    llm=llm,

    retriever=_vectorstore.as_retriever(),

    memory=memory

)

return conversation_chain

```

```

def handle_userinput(user_question):

    print(st.session_state)

    response = st.session_state.conversation({"question": user_question})

    st.session_state.chat_history = response['chat_history']

    for i, message in enumerate(st.session_state.chat_history):

        if i % 2 == 0:

            st.write(user_template.replace(

                "{{MSG}}", message.content), unsafe_allow_html=True)

        else:

            st.write(bot_template.replace(

                "{{MSG}}", message.content), unsafe_allow_html=True)

def main():

    load_dotenv()

```

```

st.set_page_config(page_title="Chat with multiple PDFs",
                    page_icon=":books:")

# st.write(css, unsafe_allow_html=True)

if "conversation" not in st.session_state:
    st.session_state.conversation = None

if "chat_history" not in st.session_state:
    st.session_state.chat_history = None

st.header("Chatbot:books:")

st.header("Ask a question about your documents")

user_question = st.text_input("Ask a question about your documents:")

if user_question:
    print(user_question)

    handle_userinput(user_question)

with st.sidebar:
    st.subheader("Your documents")

    pdf_docs = st.file_uploader(
        "Upload your PDFs here and click on 'Process'", accept_multiple_files=True)

    if st.button("Process"):
        with st.spinner("Processing"):
            # get pdf text
            raw_text = get_pdf_text(pdf_docs)

            # get the text chunks
            text_chunks = get_text_chunks(raw_text)

```

```
# create vector store

vectorstore = get_vectorstore(text_chunks)

# create conversation chain

st.session_state.conversation = get_conversation_chain(
    vectorstore)

if __name__ == '__main__':
    main()
```


CHAPTER 7

CONCLUSION

In conclusion, the project on analyzing emotions within digital conversations represents a significant endeavor with far-reaching implications for both individuals and society at large. Through the integration of advanced technologies such as artificial intelligence (AI) and natural language processing (NLP), coupled with a commitment to Corporate Social Responsibility (CSR) principles, this project has the potential to yield valuable insights and contribute positively to various facets of digital communication and mental well-being. By systematically analyzing the emotional dynamics inherent in digital conversations, the project aims to uncover patterns, trends, and nuances in how individuals express and perceive emotions online. This understanding can inform a wide range of applications, from enhancing the effectiveness of chatbots and virtual assistants to providing valuable insights for researchers studying human behavior and emotion. Furthermore, the project's adherence to CSR principles ensures that ethical considerations, fairness, transparency, and accountability are prioritized throughout the project lifecycle. By incorporating measures to protect user privacy, mitigate biases, combat online harassment, and promote mental health awareness, the project seeks to foster a safer, more inclusive, and empathetic digital environment. Through community engagement, collaboration with stakeholders, and a commitment to education and awareness, the project aims to empower individuals to navigate digital communication spaces responsibly and ethically. By raising awareness about the ethical implications of emotion analysis and promoting digital literacy, the project strives to equip users with the knowledge and tools needed to engage in online interactions with empathy and respect.

REFERENCES

1. A Jianhui Pang;Yanghui Rao;Haoran Xie;Xizhao Wang;Fu Lee Wang;Tak-Lam Wong;Qing Li | Fast Supervised Topic Models for Short Text Emotion Detection | Year: 2021 | IEEE Transactions on Cybernetics
1. Piotr Kłosowski | Deep Learning for Natural Language Processing and Language Modelling | 2018 Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA) | Year: 2018 | Conference Paper | Publisher: IEEE The Institute Of Electrical and Electronics Engineers Inc. signal processing
2. Ahmad Abdellatif;Khaled Badran;Diego Elias Costa;Emad Shihab | IEEE Transactions on Software Engineering |A Comparison of Natural Language Understanding Platforms for Chatbots in Software Engineering | Year: 2022 | Volume: 48, Issue: 8 | Journal Article | Publisher: IEEE IEEE Transactions on Software Engineering
3. Erdenebileg Batbaatar;Meijing Li;Keun Ho Ryu | Semantic-Emotion Neural Network for Emotion Recognition From Text | IEEE Access | Year: 2019 | Volume: 7 | Journal Article | Publisher: IEEE
4. A Leeja Mathew;V R Bindu |A Review of Natural Language Processing Techniques for Sentiment Analysis using Pre-trained Models | 2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC) | Year: 2020 | Conference Paper | Publisher: IEEE
5. Design development and execution of Smart Contract : An Overview Pa Martina; R Dhanvardini; R Vijay R Amirtharajan; Padmapriya Pravinkumar 2023 International Conference on Computer Communication and Informatics (ICCCI) Year: 2023 | Conference Paper | Publisher: IEEE
6. Blockchain-Enabled Smart Contracts: Architecture, Applications, and Future Trends Shuai Wang; Liwei Ouyang; Yong Yuan; Xiaochun Ni; Xuan Han; Fei-Yue Wang IEEE Transactions on Systems, Man, and Cybernetics: Systems Year: 2019 | Volume: 49, Issue: 11 | Journal Article | Publisher: IEEE