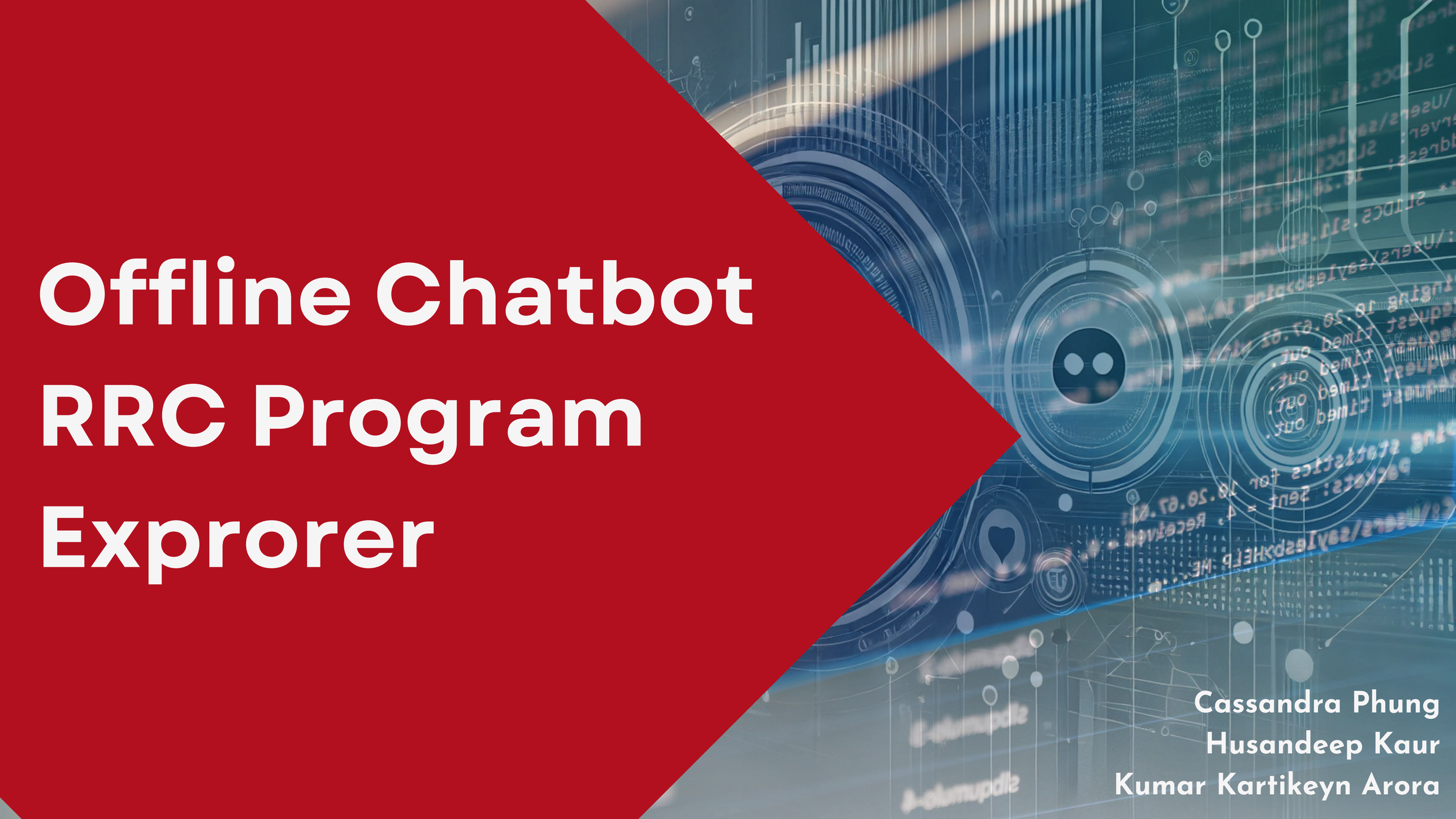
Offline Chatbot RRC Program Exprorer



Cassandra Phung
Husandeep Kaur
Kumar Kartikeyn Arora

Meet the Team

Cassandra Phung

Kumar Kartikeyn Arora

Husandeep Kaur

Introduction

Objective

This project developed a chatbot to operate locally and offline.

Key Features

- Processes and interacts with PDF files using a language model (LLM) within a Jupyter Notebook.
- Efficiently retrieves information from large documents, saving time and effort.
- By operating offline, the chatbot avoids relying on online platforms where user-uploaded data could be utilized for training generative AI models.
- Due to this being a local system, it is suitable for handling sensitive files such as financial, medical or general personal information.

Introduction

Use Case

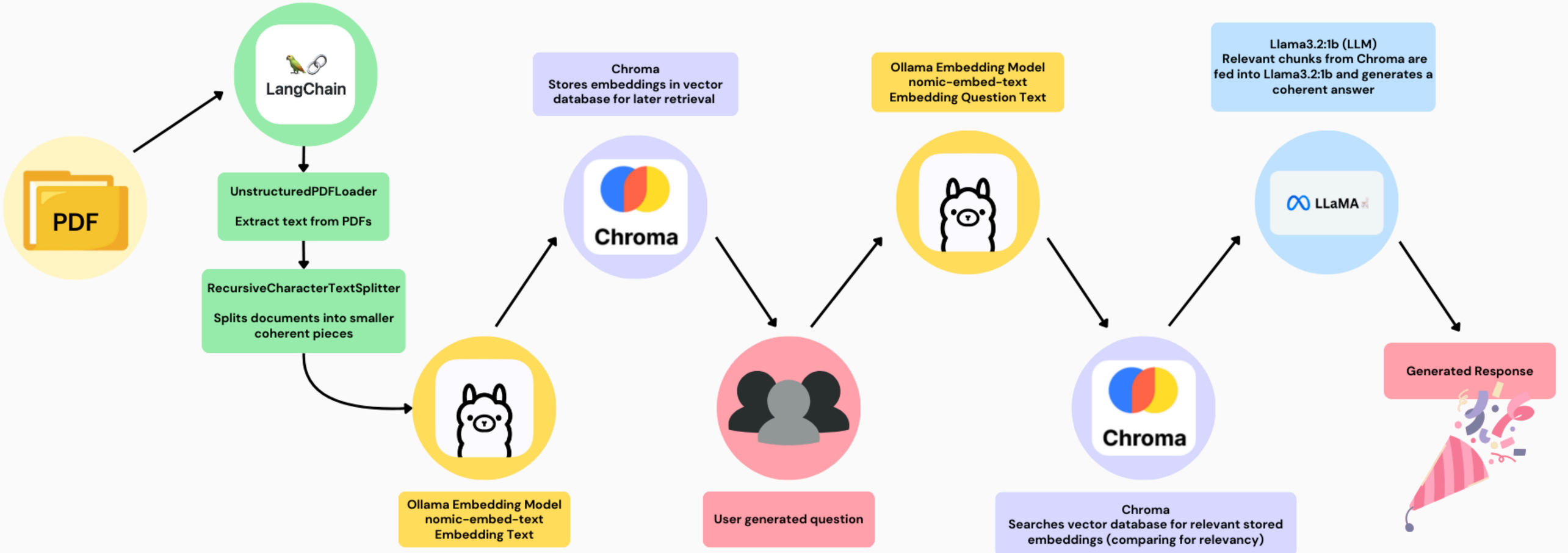
Demonstration includes processing information from three Information Technology Programs at Red River College, sourced directly from the official RRC website.

Using these programs as input data, our chatbot analyzed the PDFs to generate responses to user queries.

Chatbot Project Overview

Neural Networks & Deep Learning

Cassandra Phung
Kumar Kartikeyn Arora
Husandeep Kaur



Chatbot Responses

Utilizes RRC Program Explorer

Identify
User Query

Leveraging LLM to create coherent responses with speed and accuracy in mind

Chatbots are simple, but what can we do to optimize for better results?

Methodology

- Created separate program PDFs.
- Scripted multiple versions of the chatbot, leveraging different libraries and techniques.
- Switching from a LLM with 7 billion parameters to 1 billion to improve generated response time
- Hyperparameter testing for splitting and chunking text.
- Implemented metadata and tested retrievers for better results.

Prototype & Results

```python
# Run the chatbot
if __name__ == "__main__":
    chatbot_interactive()
```

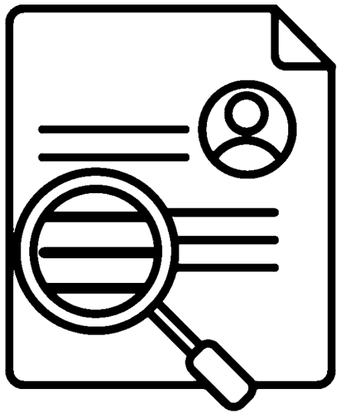
```
Welcome to the Academic Advisor Chatbot!
Type 'exit' to quit the chatbot.
Enter your question:  tell me about data science and machine learning program
Specify the program (e.g., Data_Science_Machine_Learning):  data science and machine learning

Chatbot Response:
As an academic advisor, I'm happy to provide you with general information about data science and machine learning programs!

Data science is an interdisciplinary field that combines elements of computer science, statistics, and domain-specific knowledge to extract insights and knowledge from data. The primary goal of data science is to uncover hidden
in large datasets, which can inform decision-making in various industries such as healthcare, finance, marketing, and more.

Machine learning is a subfield of artificial intelligence that involves training algorithms to make predictions or decisions based on data. It's a key component of data science, as it enables data scientists to build models tha
erformance over time. Machine learning has applications in various areas, such as image recognition, natural language processing, recommendation systems, and predictive modeling.

Data science and machine learning programs typically cover a range of topics, including:

1. Programming languages: Python, R, or SQL are commonly used programming languages in data science.
2. Data structures and algorithms: Understanding data structures like arrays, lists, and dictionaries, as well as algorithms like sorting, searching, and graph traversal, is crucial for working with large datasets.
3. Statistics and probability: Statistical analysis and probability theory provide the foundation for understanding how to interpret and model data.
4. Machine learning: This includes supervised and unsupervised learning techniques, such as linear regression, logistic regression, decision trees, clustering, and neural networks.
```

```
  * Undergraduate Data Science programs:
        + Public universities: $10,000 - $25,000 per year
        + Private universities: $20,000 - $40,000 per year
  * Graduate Data Science programs (MS or MA):
        + Public universities: $15,000 - $30,000 per year
        + Private universities: $25,000 - $50,000 per year
  * Online Data Science programs:
        + Certificate courses: $1,000 - $5,000
        + Master's degree programs: $20,000 - $40,000 per year

  **Tuition Fees for Machine Learning Programs**

  Machine Learning is often offered as a specialization within Data Science programs. Tuition fees for Machine Learning programs are typically similar to those mentioned above.

  * Undergraduate Machine Learning courses:
        + Public universities: $5,000 - $15,000
```

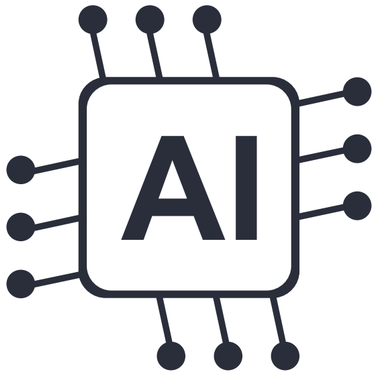
RRC Chatbot

Key Takeaways

User Query



Match query embeddings to file embeddings



Responses

Entirely local system for privacy and security

THANK YOU