

**Name : Hetal Gada**

**NUID : 002114112**

**Title: Development of an AI-Powered PDF Recommendation System**  
**Introduction**

In today's data-rich environment, extracting meaningful insights from documents such as PDFs is crucial for informed decision-making. Traditional search methods often fall short when dealing with large volumes of unstructured data. To address this challenge, an AI-powered PDF recommendation system has been developed, leveraging advanced natural language processing (NLP) and machine learning techniques.

**System Overview**

The developed system utilizes several key components:

**Document Extraction:** PDF documents are processed using PyMuPDF to extract text content efficiently.

**Embeddings and Vectorization:** Textual content from PDFs is converted into high-dimensional vector representations using OpenAI's embeddings model. These embeddings capture semantic similarities between documents.

**Pinecone Indexing:** These embeddings are indexed using Pinecone, a vector database service, to enable fast and scalable retrieval based on cosine similarity.

**Chatbot Interface:** The system features a Streamlit-based chat interface where users can upload PDFs and query information. The chatbot leverages a combination of retrieval-based and generative models (GPT-3.5-turbo) to interact with users and provide relevant responses.

**Functionality**

**PDF Upload and Processing:** Users can upload multiple PDFs, which are asynchronously processed and chunked for efficient handling.

**Recommendation Engine:** Based on the vectorized embeddings, the system recommends relevant PDFs to users based on their queries and the content similarity stored in the Pinecone index.

**Interactive Chat Interface:** Users interact with the chatbot to ask questions about uploaded PDFs. The system maintains a chat history and uses it to refine future recommendations and responses.

## Learning and Adaptation

The system continually learns from user interactions and updates its knowledge base:

**Dynamic Learning:** Through user queries and feedback, the system adapts its recommendations and responses to improve relevance over time.

**Personalization:** Recommendations become increasingly personalized as the system learns user preferences and the context of their queries.

## Use Cases

**Research and Analysis:** Researchers can quickly access relevant literature and related studies.

**Corporate Knowledge Management:** Organizations can leverage the system for efficient document retrieval and knowledge sharing among employees.

**Educational Support:** Students and educators can benefit from personalized learning resources and study materials.

## Conclusion

The AI-powered PDF recommendation system represents a significant advancement in document management and information retrieval. By integrating AI technologies such as NLP, embeddings, and machine learning models, the system offers scalable, personalized solutions for accessing and utilizing vast amounts of textual data.

## Future Enhancements

Future enhancements could focus on:

**Enhanced Natural Language Understanding:** Improving the chatbot's ability to understand complex queries and nuances in language.

**Integration with External APIs:** Incorporating external data sources and APIs to enrich document recommendations further.

## Acknowledgments

Acknowledging the tools and libraries used, such as PyMuPDF, OpenAI's models, Pinecone for vector indexing, and Streamlit for the user interface.

This report outlines the development and capabilities of the AI-powered PDF recommendation system, highlighting its potential applications and future directions for improvement.

## Recommended system from the drop downs:

Powered by OpenAI and Pinecone

Deploy

Drag and drop files here  
Limit 200MB per file • PDF

Browse files

AI.pdf 12.6MB

Hetal\_Gada\_Resume.pdf 167.8KB

Select a PDF to query

/tmp/Hetal\_Gada\_Resume.pdf

Ask your question here:

is there masters degree required for software engineer job?

Send

Assistant

In the field of software engineering, a master's degree is not always required for a job. Many software engineering positions may only require a bachelor's degree in computer science, software engineering, or a related field. However, having a master's degree can be beneficial for advancing your career, especially if you are interested in roles that require more specialized knowledge or research-oriented positions. Ultimately, the educational requirements for software engineering jobs can vary depending on the company, the specific role, and the level of experience required.

Powered by OpenAI and Pinecone

Deploy

Drag and drop files here  
Limit 200MB per file • PDF

Browse files

AI.pdf 12.6MB

Hetal\_Gada\_Resume.pdf 167.8KB

Select a PDF to query

/tmp/AI.pdf

Ask your question here:

give me the best AI used for education

Send


Assistant

One of the significant AI applications in education is the use of AI-powered course recommendation platforms like Skill Lake. These platforms analyze employee performance and learning patterns to recommend courses and training materials that are most suitable for each individual, enhancing the overall effectiveness of training programs. Additionally, AI enables immersive learning experiences through technologies like virtual and augmented reality training, providing realistic simulations for employees to practice and apply their skills in a safe environment. It's essential to implement these AI technologies responsibly and ethically, ensuring that the data used to power AI algorithms is diverse and inclusive, with human oversight in place to maintain ethical standards.


Answers from my pdf:

Powered by OpenAI and Pinecone


Deploy


 **PDF Information Chatbot**

Upload PDFs

 Drag and drop files here  
Limit 200MB per file • PDF

Browse files

 HetaL\_Gada\_Resume.pdf 167.8KB

 AI.pdf 12.6MB

Select a PDF to query

/tmp/AI.pdf

Ask your question here:

when did I work in deloitte?

Send

Assistant


You worked in Deloitte as a Senior Analyst from June 2021 to July 2022.

Clear


If I ask any irrelevant question:

Powered by OpenAI and Pinecone


Deploy


 **PDF Information Chatbot**

Upload PDFs

 Drag and drop files here  
Limit 200MB per file • PDF

Browse files

 AI.pdf 12.6MB

 HetaL\_Gada\_Resume.pdf 167.8KB

Select a PDF to query

/tmp/AI.pdf

Ask your question here:

blue

Send

Assistant

It seems like you have just entered the word "blue" without any context. Can you please provide more information or specify a question related to AI or any of the individuals mentioned in the database? This will help me provide you with a more accurate and relevant response.

Powered by OpenAI and Pinecone

# PDF Information Chatbot

Deploy

Upload PDFs



Drag and drop files here

Limit 200MB per file • PDF

Browse files



AI.pdf 12.6MB



HetaL\_Gada\_Resume.pdf 167.8KB



Select a PDF to query

/tmp/AI.pdf



Ask your question here:

do you know anything about Formula 1?

Send

Assistant

I can provide information related to AI experts and professionals. If you have any questions about individuals in the AI field or related topics, feel free to ask!

Clear