STATEMENT OF PURPOSE

VISION CHATBOT

Swaksh Patwari(B22AI065)

Why to Working on this Project?

I am very much interested to learn Deep Learning and Computer Vision model and their implementation. It will going enhance my skills. Working on this project will let me learn a lot from my mentor. It will also going to help me for my CV.

Technical Understanding

- Python, pandas, matplotlib, numpy, sklearn, MongoDB, Basics Of ML, HTML, CSS, Basics of CNN.

Work Flow Of The Project

1. Data Collection:

Gather a diverse dataset containing images and corresponding text descriptions.

2. Data Preprocessing:

Normalize and preprocess images (resize, normalization).

Tokenize and clean textual data.

Prepare data for training GPT-4.

3. Model Architecture:

Design a multimodal architecture that integrates GPT-4 with computer vision models.

Implement Convolutional Neural Networks (CNN) for image processing.

Utilize Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM), or Transformer models for text processing.

4. Training:

Train the vision model on the image dataset using CNN.

Train the text model using GPT-4 or a similar language model.

Fine-tune the multimodal model on the combined dataset.

5. Integration of GPT-4:

Integrate GPT-4 for handling text-based interactions.

Develop mechanisms for seamless communication between vision and language models.

6. Real-time Contextual Understanding:

Implement mechanisms to understand and adapt to the dynamic nature of conversations.

Use attention mechanisms or context-aware models for better understanding.

7. User Interface Design:

Design an intuitive and visually appealing interface for user interactions.

Implement features for uploading images, typing text, and receiving responses.

8. Testing and Evaluation:

Conduct thorough testing with various inputs (images, text, mixed).

Evaluate the performance of the Vision ChatBot in terms of accuracy and response time.

9. Optimization:

Optimize the model if the performance of the model is not as the required.

10. Documentation:

Create comprehensive documentation covering the project's architecture, code structure, and usage guidelines.

11. Deployment:

Deploy the Vision ChatBot in a suitable environment (cloud, server).

Monitor for performance and scalability.

12.Provide Updates Based on user reviews and fixing of some bugs.

Contact No - 9407190501
email id - b22ai065@iitj.ac.in