Model Development Phase Template

Date	9th July 2024
Team ID	SWTID1719935665
Project Title	GeminiDecode: Multilanguage Document Extraction by Gemini Pro
Maximum Marks	5 Marks

Feature Selection Report Template

In the forthcoming update, each feature will be accompanied by a brief description. Users will indicate whether it's selected or not, providing reasoning for their decision. This process will streamline decision-making and enhance transparency in feature selection.

Feature	Description	Selected (Yes/No)	Reasoning
PDF Reader	Reads and extracts text from PDF documents.	Yes	Essential for user interaction and input handling.
Text Search	Brief description	Yes	Enhances user experience by enabling quick access to specific information.
Text Summarization	Provides a summary of the PDF content.	Yes	Helps users quickly understand the main points of lengthy documents.
Named Entity Recognition	Identifies and highlights entities (names, places, dates).	Yes	Useful for extracting key information and enhancing user queries.

Keyword Extraction	Extracts and highlights key terms in the document.	Yes	Improves user understanding of important concepts and terms in the document.
Contextual Chat	Engages in a conversation with the user about the PDF content.	Yes	Central feature for interactive and meaningful user engagement.
Annotation	Allows users to add notes and comments to the PDF.	No	Adds complexity and may not be necessary for basic information retrieval.
Multiple Language Support	Supports PDFs in various languages.	Yes	Increases accessibility and usability for a diverse user base.
Document Navigation	Enables easy navigation through different sections of the PDF.	Yes	Facilitates smooth and efficient user experience.
User Authentication	Requires users to log in before accessing the PDF chatbot.	No	Depends on use case; may not be needed for general information retrieval.
Export Chat History	Allows users to export the chat history as a text or PDF file.	Yes	Useful for record-keeping and future reference.
Voice Interaction	Enables voice-based interaction with the chatbot.	No	Adds complexity; text-based interaction may suffice for initial implementation.
PDF Annotation Detection	Detects existing annotations and comments in the PDF.	No	May be redundant with text extraction and search capabilities.
Real-time Collaboration	Allows multiple users to interact with the PDF simultaneously.	No	Complex feature that might not be necessary for the primary function of information retrieval.