Srikanth B

Final Project

PROJECT TITLE

"Image Descriptor:
Enhancing Image Understanding with AI"

AGENDA

- Introduction
- Problem Statement
- Project Overview
- Solution and Value Proposition
- Modelling Approach
- Results and Impact
- Conclusion

PROBLEM STATEMENT

- Difficulty in comprehending images without human intervention.
- Lack of efficient tools for extracting meaningful information from images.
- Need for an automated solution to describe images accurately.

PROJECT OVERVIEW

- Introduce the Image Descriptor project.
- Briefly explain the technologies and frameworks used (e.g., Google Generative AI, Express.js, Multer).
- Highlight the aim of the project: To develop a system capable of generating textual descriptions for uploaded images.

WHO ARE THE END USERS?

- Content creators
- Social media managers
- E-commerce businesses
- Researchers
- Anyone requiring image understanding and analysis

YOUR SOLUTION AND ITS VALUE PROPOSITION

- **Describe the Image Descriptor system:**
    1. User-friendly web interface for uploading images and entering questions.
    2. Integration with Google Generative AI for image description generation.

- **Value proposition:**
    1. Provides accurate textual descriptions for images.
    2. Saves time and effort in manually annotating images.
    3. Enhances accessibility for visually impaired individuals.
    4. Improves search engine optimization (SEO) for images.

THE WOW IN YOUR SOLUTION

- **Showcase the unique features and advantages of the Image Descriptor system:**

1. Real-time image description generation.
2. Seamless integration with existing applications.
3. Scalability to handle large volumes of image data.
4. Ability to customize prompts for generating descriptions.

MODELLING

- **Explain the modelling approach:**

1. Utilization of Google Generative AI's Gem-Pro Vision model.
2. Processing of image and text inputs for content generation.
3. Handling of multi-modal data for accurate descriptions.

RESULTS

- **Present the outcomes and performance metrics:**

1. Accuracy of generated descriptions compared to ground truth.
2. Speed of processing and response time.
3. User feedback and satisfaction ratings.
4. Potential applications and future improvements.

Demo Link