**Project Initialization and Planning Phase**

|  |  |
| --- | --- |
| Date | 7 November 2024 |
| Team ID | 739889 |
| Project Title | Image Caption Generator |
| Maximum Marks | 3 Marks |

**Project Proposal (Proposed Solution) template**

This project proposal outlines a solution to address the challenge of generating descriptive captions for images.With a clear objective, defined scope, and a concise problem statement,  
the proposed solution details the use of deep learning techniques for feature extraction and text generation, including hardware (GPU), software libraries (TensorFlow/Keras), and skilled personnel in AI.

|  |  |
| --- | --- |
| **Project Overview** | |
| Objective | To develop an AI-based system capable of generating accurate and descriptive captions for images using deep learning techniques. |
| Scope | The project focuses on automating image captioning using neural networks, targeting applications in accessibility, digital content management, and multimedia platforms. It covers model development, dataset preparation, training, and deployment. |
| **Problem Statement** | |
| Description | Manually writing captions for large volumes of images is time-consuming, inconsistent, and not scalable. Additionally, lack of captions can make content inaccessible to visually impaired users. |
| Impact | Solving this problem enables efficient content management, improves accessibility, and saves significant human effort while maintaining consistency and scalability. |
| **Proposed Solution** | |
| Approach | The system uses a combination of Convolutional Neural Networks (CNNs) for feature extraction and Recurrent Neural Networks (RNNs) or Transformer models for generating natural language descriptions of the images. |
| Key Features | * Automatic, real-time caption generation * Support for diverse image types * Improved accessibility for visually impaired users * Scalable for large datasets and platforms |

**Resource Requirements**

|  |  |  |
| --- | --- | --- |
| **Resource Type** | **Description** | **Specification/Allocation** |
| **Hardware** | | |
| Computing Resources | CPU/GPU specifications, number of cores | e.g., 2 x NVIDIA V100 GPUs |
| Memory | RAM specifications | e.g., 8 GB |
| Storage | Disk space for data, models, and logs | e.g., 1 TB SSD |
| **Software** | | |
| Frameworks | Python frameworks | e.g., Flask |
| Libraries | Additional libraries | e.g., tensorflow |
| Development Environment | IDE, version control | e.g., Google Colab, Anaconda prompt, VS code |
| **Data** | | |
| Data | Source, size, format | e.g., Kaggle dataset, 2,000 images |