Project Design Phase-I
Proposed Solution Template

Date	19 September 2022
Team ID	598189
Project Name	Project - Image Caption Generation
Maximum Marks	2 Marks

Proposed Solution:

PROJECT TITLE: **IMAGE CAPTION GENERATION**

S.No	Parameters	Description
1.	Problem Statement (Problem to be solved)	The problem at hand is to automatically generate human-readable textual descriptions (captions) for images, particularly photographs. This task is challenging for computers as it requires understanding the image content and translating it into natural language
2.	Idea / Solution description	Our solution involves building an Image Caption Generator using the Flickr 8K dataset, which consists of 8,000 images , each paired with five different captions. We will use a CNN, specifically the Inception V3 model, to extract image features and encode the input image. These features will serve as the initial input to the RNN-based LSTM layers, which will generate image captions. This approach distinguishes our model from others, as it provides the image

		embedding as the first input to the RNN only once.
3.	Novelty / Uniqueness	The novelty of our solution lies in the use of a hybrid model combining CNNs and RNNs with LSTM layers to generate image captions. Additionally, we use the Inception V3 model for feature extraction. The solution aims to provide accurate and contextually relevant image captions.
4.	Social Impact / Customer Satisfaction	The social impact of our solution is significant, as it can benefit individuals with visual impairments by providing them with detailed image descriptions. It can also be used in various applications, such as content indexing, image retrieval, and improving the accessibility of image-based content. Customer satisfaction is expected to increase as the system can automatically generate descriptive captions for images,reducing the need for manual captioning.
5.	Business Model (Revenue Model)	The business model for this project can include licensing the Image Caption Generator to various industries that rely on image data. Potential revenue sources may come from selling licenses to media companies, e-commerce platforms, or content management systems. Alternatively, a subscription-based service for individuals and businesses looking to automatically caption their

		images can be explored.
6.	Scalability of the Solution	The solution is designed to be scalable. As more data becomes available, the model can be retrained to improve caption quality. The use of pre-trained CNN models and deep learning techniques makes it adaptable to larger datasets and evolving technology. Furthermore, the solution can be deployed on cloud infrastructure to handle a higher volume of image caption requests, ensuring scalability.