Project Initialization and Planning Phase

Date	16 June 2025
Team Lead Name	Jayanth Srinivas Bommisetty
Project Title	Sloan Digital Sky Survey (SDSS) galaxy classification using machine learning
Maximum Marks	3 Marks

Project Proposal (Proposed Solution) template

We propose a CNN-based image classification solution that takes galaxy images as input and returns one of the five defined morphological types with associated confidence. This approach leverages the strengths of deep learning in visual pattern recognition and ensures efficient, scalable, and objective classification of galaxy morphologies.

Project Overview	
Objective	To develop a Convolutional Neural Network (CNN)-based model capable of classifying galaxy images into five morphological categories to automate and accelerate astronomical data analysis.
Scope	Automate classification of SDSS galaxy images into five types with confidence scores, covering preprocessing, model training, evaluation, and web deployment—excluding real-time telescope data and spectral analysis.
Problem Statement	
Description	Manual galaxy classification is slow, labor-intensive, and prone to inconsistency. With the scale of data from astronomical surveys like SDSS, manual methods cannot keep up with classification demands.
Impact	Automating galaxy classification will enable large-scale processing of astronomical data, improve accuracy and consistency, and assist researchers in studying the universe more efficiently.
Proposed Solution	
Approach	Build a CNN with Keras/TensorFlow using augmented Galaxy Classification data and deploy it via a Flask web app for image-based galaxy classification.
Key Features	Supports 5 galaxy classes, provides softmax confidence scores, and features a clean, interactive web interface.

Resource Requirements

Resource Type	Description	Specification/Allocation
Hardware		
Computing Resources	CPU/GPU	1 x NVIDIA RTX 3060 (or similar)
Memory	RAM specifications	8 GB
Storage	Disk space for data, models, and logs	1 TB SSD
Software		
Frameworks	Python frameworks	Flask
Libraries	Additional libraries	scikit-learn, pandas, numpy, matplotlib
Development Environment	IDE, version control	Jupyter Notebook, Git/GitHub
Data		
Data	Source, size, format	Kaggle(Galaxy classification dataset), 25,000 images