

Project Initialization and Planning Phase

| | |
|---------------|---|
| Date | 15 July 2024 |
| Team ID | 740051 |
| Project Title | SDSS galaxy classification using Machine Learning |
| Maximum Marks | 3 Marks |

Project Proposal (Proposed Solution) template

The Sloan Digital Sky Survey(SDSS) has amassed a vast repository of galaxy images aand spectra, offering a rich dataset for astronomical research. Traditional methods of classifying galaxies based on visual inspection are time-consuming and subjective. This project purposes a leaverage machine learning techniques to automate the classification process, aiming to enhance accuracy and efficiency in identifying galaxy types within the SDSS database.

| Project Overview | |
|-------------------|---|
| Objective | Improving efficiency and accuracy over manual methods, facilitating faster analysis and deeper insights into astronomical data |
| Scope | Classify galaxies by analyzing their spectra, identifying types like spirals or ellipticals, aiding in understanding cosmic structure and evolution. |
| Problem Statement | |
| Description | SDSS uses machine learning to classify galaxies by analyzing their spectra, distinguishing types like spirals or ellipticals, enhancing our understanding of cosmic structure and evolution. |
| Impact | SDSS's machine learning-driven galaxy classification revolutionizes astronomy by automating and refining categorization, enabling large-scale studies of galaxy populations, evolution, and the broader universe. |
| Proposed Solution | |
| Approach | SDSS employs supervised machine learning models trained on galaxy spectra to classify types such as spirals or ellipticals, enhancing accuracy and scalability in large-scale astronomical surveys. |
| Key Features | SDSS uses machine learning to classify galaxies based on spectral features, employing algorithms like random forests and neural networks for accurate identification of galaxy types and properties. |

Resource Requirements

| Resource Type | Description | Specification/Allocation |
|---------------------|---|--------------------------|
| Hardware | | |
| Computing Resources | CPU/GPU specifications, number of cores | NVIDIA RTX 3090 |
| Memory | RAM specifications | 8 GB |
| Storage | Disk space for data, models, | 1 TB SSD |

| | | |
|-------------------------|----------------------|--------------------------|
| | and logs | |
| Software | | |
| Frameworks | Python frameworks | TensorFlow, PyTorch |
| Libraries | Additional libraries | Matplotlib, Seaborn |
| Development Environment | IDE, version control | Jupyter Notebook, Python |
| Data | | |
| Data | Source, size, format | Kaggle dataset, csv |