

Abstract

- The classification of galaxies plays a crucial role in astronomical research and understanding the structure of the universe. Manual classification of galaxies is time-consuming and prone to human error.
- This project proposes an automated galaxy classification system using machine learning techniques and a web-based interface. The model is trained on features derived from the Sloan Digital Sky Survey (SDSS) dataset and is integrated with a Flask-based web application that allows users to input galaxy features and obtain predictions.

Problem Statement

Galaxy classification is an essential task in astrophysics for studying galaxy formation and evolution. Traditional manual methods are inefficient when dealing with large datasets such as SDSS.

Hence, there is a need for an automated, efficient, and scalable system that can classify galaxies using machine learning models.

Objective

The objectives of this project are:

- To build a machine learning model for galaxy classification
- To integrate the trained model with a Flask web application
- To provide a user-friendly interface for prediction
- To demonstrate end-to-end ML workflow from data preprocessing to deployment

Technology Stack

- **Programming Language:** Python
- **Web Framework:** Flask
- **Frontend:** HTML, CSS
- **Machine Learning Libraries:** Scikit-learn
- **Data Processing:** Pandas, NumPy
- **Model Serialization:** Joblib
- **Version Control:** GitHub

Dataset Description

The dataset used in this project is derived from the **Sloan Digital Sky Survey (SDSS)**. It contains various numerical features representing physical properties of galaxies. These

features are used as inputs to the machine learning model for classification.

System Architecture

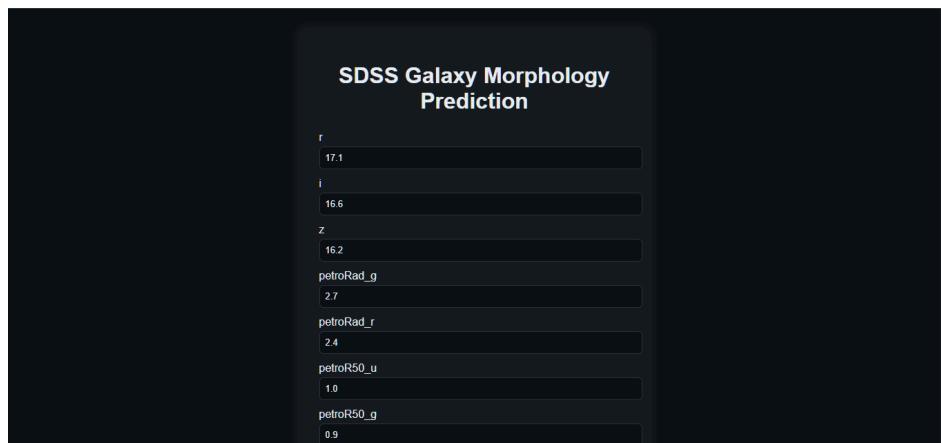
The system architecture consists of the following components:

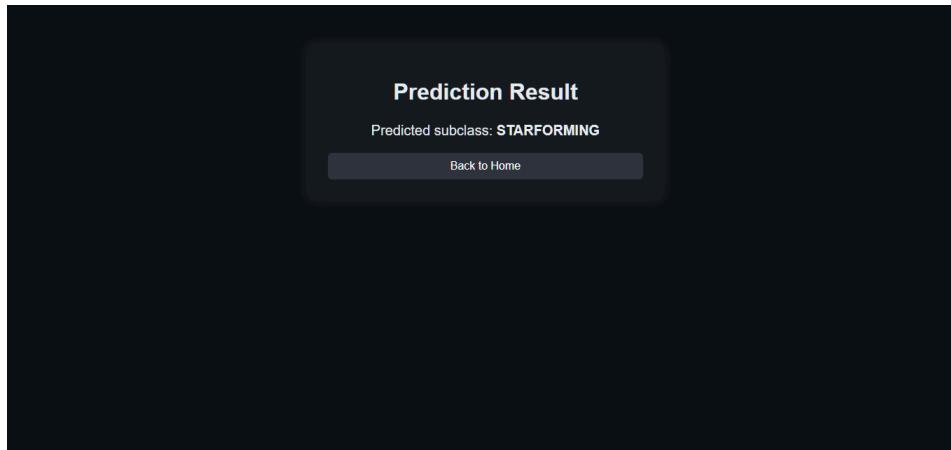
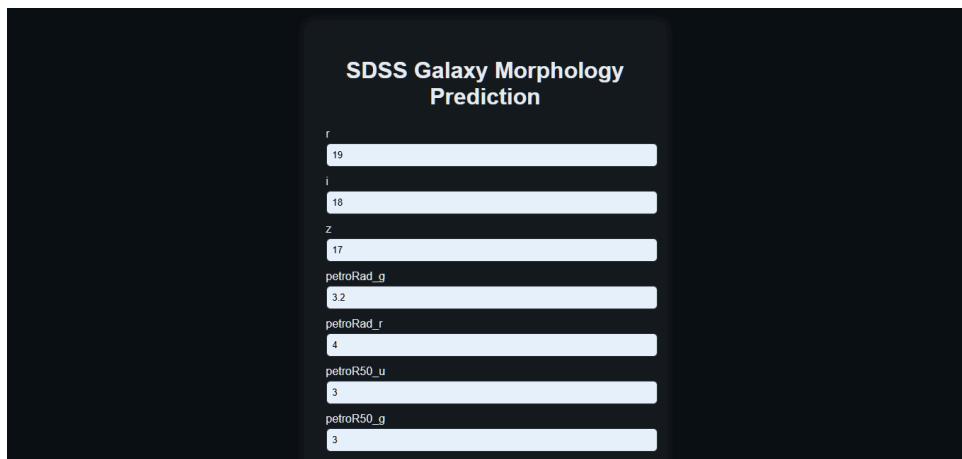
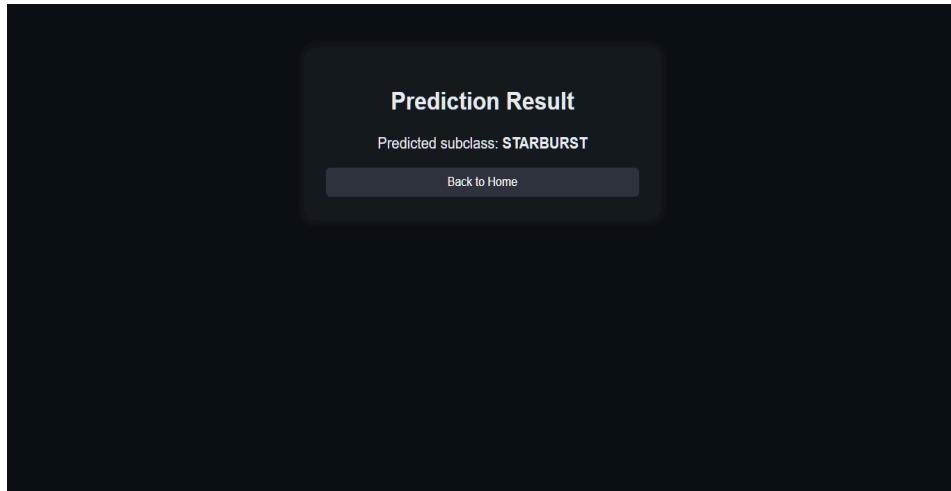
1. User Interface (HTML/CSS)
2. Flask Backend
3. Preprocessing Module (Scaler)
4. Machine Learning Model
5. Prediction Output

Architecture Flow:

User → Web Interface → Flask Backend → ML Model → Prediction Result

Results and Output





Project Structure

```
SDSS-Galaxy-Classification/
|__ app.py
|__ templates/
|   |__ index.html
|   |__ inner-page.html
|__ static/
|__ scaler.pkl
|__ selected_features.json
|__ output_screenshots/
|__ requirements.txt
|__ README.md
```

Implementation Details

- The machine learning model is trained using scikit-learn.
- Feature scaling is applied using a pre-trained scaler.
- The Flask backend loads the model and scaler using Joblib.
- The web interface collects user inputs and sends them to the backend.
- The backend processes inputs and returns the predicted galaxy class

Conclusion

This project demonstrates how machine learning can be effectively used to automate galaxy classification. The integration of the ML model with a Flask-based web application provides a simple and interactive way to obtain predictions.

The system can be further enhanced by using deep learning models, adding more galaxy classes, and deploying the application on cloud platforms.

Future Scope

- Integration of deep learning models
- Deployment on cloud platforms
- Support for image-based galaxy classification
- Improved UI and user experience