Model Development Phase

Date	19 June 2025
Team ID	xxxxxx
Project Title	sloan digital sky survey (sdss) galaxy classification using machine learning
Maximum Marks	5 Marks

Feature Selection Report

In the forthcoming update, each feature will be accompanied by a brief description. Users will indicate whether it's selected or not, providing reasoning for their decision. This process will streamline decision-making and enhance transparency in feature selection.

Feature	Description	Selected (Yes/No)	Reasoning
u	Ultraviolet magnitude	Yes	Used in regression to estimate redshift as it contributes to photometric patterns
g	Green magnitude	Yes	Important for both classification and redshift estimation
r	Red magnitude	Yes	Strong correlation with both redshift and star formation class

i	Infrared magnitude	Yes	Helps in regression and classification of galaxies
z	Near-infrared magnitude	Yes	Crucial for redshift shift patterns
modelFlux_u	Flux in u-band model	No	High variance and not strongly correlated with target variables
modelFlux_g	Flux in g-band model	No	Similar to modelFlux_u; excluded due to noise and multicollinearity
modelFlux_r	Flux in r-band model	No	Lower predictive power in preliminary testing
modelFlux_i	Flux in i-band model	No	Removed due to high correlation and redundancy
modelFlux_z	Flux in z-band model	No	Did not improve performance in validation
petroR50_u	Petrosian 50% light radius in u-band	No	Not selected due to inconsistent values and noise
petroR50_g	Petrosian 50% light radius in g-band	Yes	Used in classification to differentiate galaxy structure
petroR50_r	etrosian 50% light radius in r-band	Yes	Selected for redshift estimation

petroR50_i	Petrosian 50% light radius in i-band	Yes	Correlated well with redshift; improves regression performance
psfMag_u	PSF magnitude in u-band	Yes	Used in regression model for better accuracy
psfMag_g	PSF magnitude in g-band	Yes	Included for regression due to its moderate influence on redshift
psfMag_r	PSF magnitude in r-band	Yes	Helps improve redshift regression results
psfMag_i	PSF magnitude in i-band	Yes	Chosen for classification of galaxy morphology
psfMag_z	PSF magnitude in z-band	Yes	Enhances classification accuracy when combined with psfMag_i
class	General classification label (e.g., GALAXY)	No	Redundant, not used as input
subclass	Star formation classification (STARFORMING/STAR BURST)	Yes	Target variable for classification task
redshift	Measure of cosmic distance via spectral shift	Yes	Target variable for regression task

redshift_err	Error in redshift measurement	No	Not used in training, used only for evaluation reference