A Comparison of Classification Algorithms:

Classification of celestial objects: stars, galaxies and quasars

Classifiers	VALIDATION
1. Logistic Regression **cross validation logistic regression** **with regularization**	Accuracy_Score = 0.97 Accuracy_Score = 0.968 Cs=[1000, 10000] Accuracy_Score = 0.9693333 Cs=[1, 10] important featrures: psfMag_u, psfMag_g , petromag_g , gr, ri, ug Decision boundary not separate the classes very well. I tested for all the features.
2. SVM classifier **Using "OneVsRestClassifier"**	Accuracy_Score = 0.954 AUC curve = 0.98 I also calculated micro and macro roc' s Macro-average roc curve = 0.98 Micro average roc curve = 0.99 Roc curve of class 0(QSO =0.96 Roc curve of class 1(GALAXY) =0.98 Roc curve of class 2(STAR) =0.98 Micro average precision-recall = 0.98 precision-recall for class 0(QSO =0.77 precision-recall for class 1(GALAXY) =0.99 precision-recall for class 2(STAR) =0.87
3. KMeans	Predicted classes = 3
4. KNN classifier	Accuracy_Score = 0.9306(n_neighbors=8)(KNeighborsClassifier) Tree boundary is not very good.
5. Ensemble Classifications • **Random Forest Classifier**	No bagging:

• **Random Forest Classifier** **(by bootstrap sampling)** • **XGB classifier**	Accuracy_Score = 0.973 (rfclf=RandomForestClassifier(n_estimators=100)) Accuracy_Score = 0.9746 (rfclf=RandomForestClassifier(n_estimators=100 , max_features='auto', random_state=1))) Accuracy_Score = 0.974666666667 (rfclf = RandomForestClassifier(n_estimators=100, max_features='auto', oob_score=True, random_state=1)) rfclf.oob_score_= 0.9804706106506621 Accuracy_Score = 0.9727 Important Features: Ug, iz, ri, psfFlux_u, psfMag_u, psfMag_g, petromag_g, cModelFlux_u, gr,spectroFlux_u, petromag_u
7. Decision Tree classifier max_depth=6	Accuracy_Score = 0.928 Roc curve =0.77 Macro-average roc curve = 0.86 Micro average roc curve = 0.95 Roc curve of class 0(QSO =0.92 Roc curve of class 1(GALAXY) =0.87 Roc curve of class 2(STAR) =0.77 AUC for decision tree with max depth = 3: 0.8816449225486803 AUC for decision tree with max depth = 6: 0.9128212766939313 AUC for decision tree max depth = 10: 0.8787117944772443 the scores of the grid search = 6