Results:

Class Imbalance: In all three task there is class imbalance issues which must be taken care. In these three task Positive class is very small as compared to negate class which makes difficult for the learning algorithm for learning minority class. Also in class imbalance problem accuracy is not the true validation for model performance, same is the case in our task in most of the cases accuracy is above 90% but most of the minority class are misclassified so we have to take Precision and Recall for model selection

I cater the class imbalance problem by up sampling the minority class and down sampling the majority class.

Results for Bankruptcy with class imbalance

Precision (+ve)	Recall (+ve)	AUC	Accuracy
38.596	30.137	0.643	96.18%

Results for Bankruptcy with catering class imbalance

Precision (+ve)	Recall (+ve)	AUC	Accuracy
96.068	100.000	0.979	97.92%

Feature Selection: In our tasks we have multiple features for classification even 90+ in case of Bankruptcy dataset so we must select the features which are contributing the most in classification. Different techniques are employed for feature selection like feature selection based on ANOVA F score and Random forest Feature selection. Optimizing the no of features make model generalize resulting in increased accuracy and precision.

Results for Bankruptcy without feature selection

Precision (+ve)	Recall (+ve)	AUC	Accuracy
96.068	100.000	0.979	97.92%

Results for Bankruptcy with feature selection

Precision (+ve)	Recall (+ve)	AUC	Accuracy
97.256	100.000	0.098	98.59%

As visible from above statistics results improve.

Cross Validation: Different cross validation techniques are employed for training the algorithms include KFold, Stratified Kfold, and repeated Kfold.

ML Classification Algorithm: Different types of model are used for training on the given datasets and different algorithm perform better in different parameter setting and datasets. In general Random Forest perform better on two on the datasets in all different conditions.

Details analysis is done can be found in Notebook file.