Insights:

from data distribution: Male 1380 Female 956

Churn: distribution: 1: 1616 (67.87%) and 0: 765 (32.12%)

- Probability of Churn is higher in case of education level 0 and 1 than 2.
- In case of joining designation 1, probability of churn is higher.
- In case of quarterly rating is 1, probability of churn is significantly higher and also same pattern is observed in case of when driver's quarterly rating has increased throughout tenure.
- due to some reason, for drivers who joined in 2018 and 2019 , probability of churn is very high compare to 2020 and before 2018.

Random Forest:

- train and test score: (0.8697478991596639, 0.8679245283018868)
- feature importance: highest is: joining year , followed by No of records available in data, and total business value.
- recall: 0.866
- precision: 0.928
- f1-score: 0.89

on Grid Search CV: RF:

- best params: ccp_alpha=0.001, max_depth=10, max_features=7, n_estimators=300
- Gridsearch RF best score: 0.8881417819617973

Bagging Classfier: with Decision Tree:

- with 50 DTs. when max_depth=7, class_weight="balanced"
- f1 score: 0.9064039408866995
- precision: 0.9387755102040817
- recall_score: 0.8761904761904762
- accuracy: 0.880

XGBoost Classifier: (Grid SEARCH CV:) 'max_depth': 2, 'n_estimators': 100
- Accuracy: 0.87
- f1 score: 0.90
- recall: 0.923
- precision: 0.884
- feature importance: highest is : joining year , followed by No of records available in data, and total business value.

GradientBoostingClassifier: GBDC:

- Train Score: 0.914390756302521
- Test Score: 0.8909853249475891
- Accuracy Score: 0.8909853249475891
- ROC-AUC score test dataset: 0.9447855910621867
- precision score test dataset: 0.9287925696594427
- Recall score test dataset: 0.9118541033434651
- f1 score test dataset: 0.9202453987730062

Recommendations:

- We can see that probability of churn is higher in the case of education level 0 and 1. So the drivers who are less educated have high probability of churn. So ola can aim at these drivers and provide more training and awareness to increase their income so the churning can be reduced.

- In case of joining designation 1, probability of churn is higher. So, ola can give more incentives to the drivers who are joining at designation 1 to reduce the churning of these drivers.

- In case of quarterly rating is 1, probability of churn is significantly higher. Ola can provide training programs and awareness to drivers to increase their ratings so there are less drivers who get the lowest rating