Report of training

# Introduction

Report of search and training made on March 24, 2021 at 08:17:56.

## Training data

There are 690 training samples. The distribution of the labels is the following:

* Class 1: 307 instances.
* Class 0: 383 instances.

## Optimizing procedure

The parameters for the bayesian search are:

* Nested Cross Validation using 10 outer folds and 10 inner folds.
* Some of the folds will be skipped. In particular, [0, 2, 4, 6, 8] outer folds and [0, 2, 4, 6, 8, 9] inner folds will be skipped.
* For each outer fold search, a model will be fitted. In order to search for the best hyperparameters, 10 initial points will be evaluated, and 10 additional calls will be made.
* Models will be calibrated using their inner validation set.
* The optimizing metric for the bayesian search is average\_precision.
* The function used for the bayesian search is gbrt\_minimize.
* Additionally, 20 instances will be left out for assessing the variance of all models.

The search spaces for the optimization is the following:

* Search space for xgboost model.
* model: XGBClassifier(base\_score=None, booster=None, colsample\_bylevel=None,  
   colsample\_bynode=None, colsample\_bytree=None, gamma=None,  
   gpu\_id=None, importance\_type='gain', interaction\_constraints=None,  
   learning\_rate=None, max\_delta\_step=None, max\_depth=None,  
   min\_child\_weight=None, missing=nan, monotone\_constraints=None,  
   n\_estimators=100, n\_jobs=None, num\_parallel\_tree=None,  
   random\_state=None, reg\_alpha=None, reg\_lambda=None,  
   scale\_pos\_weight=None, subsample=None, tree\_method=None,  
   validate\_parameters=None, verbosity=None)
* pipeline\_post\_process: Pipeline(steps=[('post\_process',  
   <nestedcvtraining.api.OptionedPostProcessTransformer object at 0x0000025B8D09DA60>),  
   ('resample', SMOTE())])
* Search space:
* undersampling\_majority\_class: Categorical(categories=(True, False), prior=None)
* max\_k\_undersampling: Integer(low=5, high=6, prior='uniform', transform='identity')
* resample\_\_sampling\_strategy: Categorical(categories=('minority', 'all'), prior=None)
* post\_process\_\_option: Categorical(categories=('option\_1', 'option\_2', 'option\_3'), prior=None)
* model\_\_max\_depth: Integer(low=5, high=15, prior='uniform', transform='identity')
* model\_\_learning\_rate: Real(low=0.05, high=0.31, prior='log-uniform', transform='identity')
* model\_\_min\_child\_weight: Integer(low=1, high=10, prior='uniform', transform='identity')
* model\_\_subsample: Real(low=0.8, high=1, prior='log-uniform', transform='identity')
* model\_\_colsample\_bytree: Real(low=0.13, high=0.8, prior='log-uniform', transform='identity')
* model\_\_scale\_pos\_weight: Real(low=0.1, high=10, prior='log-uniform', transform='identity')
* model\_\_objective: Categorical(categories=('binary:logistic',), prior=None)
* Search space for random\_forest model.
* model: RandomForestClassifier()
* pipeline\_post\_process: None
* Search space:
* undersampling\_majority\_class: Categorical(categories=(True, False), prior=None)
* model\_\_bootstrap: Integer(low=0, high=1, prior='uniform', transform='identity')
* model\_\_n\_estimators: Integer(low=10, high=100, prior='uniform', transform='identity')
* model\_\_max\_depth: Integer(low=2, high=10, prior='uniform', transform='identity')
* model\_\_min\_samples\_split: Integer(low=5, high=20, prior='uniform', transform='identity')
* model\_\_min\_samples\_leaf: Integer(low=1, high=4, prior='uniform', transform='identity')
* model\_\_max\_features: Categorical(categories=('auto', 'sqrt'), prior=None)
* model\_\_class\_weight: Categorical(categories=('balanced', 'balanced\_subsample'), prior=None)

# Report of validation of the model in the outer Cross Validation

## Winner models of each fold and main metrics

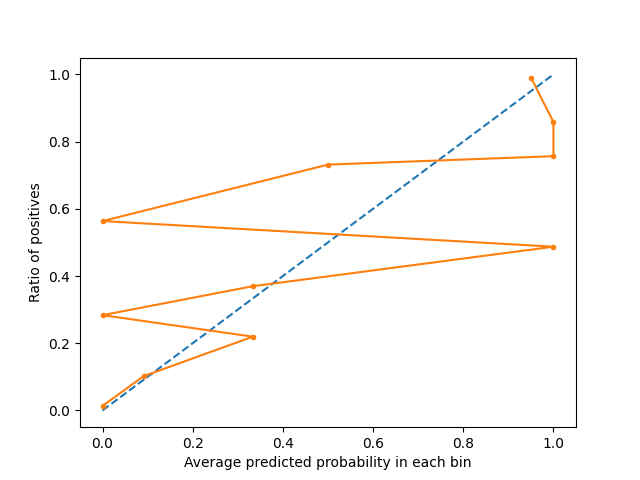
|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Fold** | **Model** | **Params** | **Comments** | **roc\_auc** | **neg\_log\_loss** | **average\_precision** | **neg\_brier\_score** |
| 1 | RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 1 * model\_\_n\_estimators: 70 * model\_\_max\_depth: 6 * model\_\_min\_samples\_split: 9 * model\_\_min\_samples\_leaf: 4 * model\_\_max\_features: sqrt * model\_\_class\_weight: balanced | * option: build model without resampling * number of folds: 4 * average size of training set: 542.75 * average prop of minority class: 0.446 | 0.955 | 0.719 | 0.916 | 0.077 |
| 3 | XGBClassifier | * undersampling\_majority\_class: False * max\_k\_undersampling: 6 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_3 * model\_\_max\_depth: 7 * model\_\_learning\_rate: 0.084 * model\_\_min\_child\_weight: 10 * model\_\_subsample: 0.806 * model\_\_colsample\_bytree: 0.246 * model\_\_scale\_pos\_weight: 1.374 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 542.75 * average prop of minority class before resampling: 0.444 * average size of training set after resampling: 603.0 * average prop of minority class after resampling: 0.5 | 0.912 | 0.395 | 0.915 | 0.129 |
| 5 | XGBClassifier | * undersampling\_majority\_class: True * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_3 * model\_\_max\_depth: 12 * model\_\_learning\_rate: 0.085 * model\_\_min\_child\_weight: 4 * model\_\_subsample: 0.89 * model\_\_colsample\_bytree: 0.139 * model\_\_scale\_pos\_weight: 0.318 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 392.0 * average prop of minority class before resampling: 0.385 * average size of training set after resampling: 482.5 * average prop of minority class after resampling: 0.5 | 0.947 | 0.294 | 0.934 | 0.097 |
| 7 | RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 0 * model\_\_n\_estimators: 62 * model\_\_max\_depth: 6 * model\_\_min\_samples\_split: 6 * model\_\_min\_samples\_leaf: 4 * model\_\_max\_features: sqrt * model\_\_class\_weight: balanced | * option: build model without resampling * number of folds: 4 * average size of training set: 542.75 * average prop of minority class: 0.444 | 0.964 | 0.264 | 0.953 | 0.079 |
| 9 | XGBClassifier | * undersampling\_majority\_class: False * max\_k\_undersampling: 6 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_3 * model\_\_max\_depth: 7 * model\_\_learning\_rate: 0.063 * model\_\_min\_child\_weight: 10 * model\_\_subsample: 0.974 * model\_\_colsample\_bytree: 0.399 * model\_\_scale\_pos\_weight: 0.437 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 542.75 * average prop of minority class before resampling: 0.444 * average size of training set after resampling: 603.0 * average prop of minority class after resampling: 0.5 | 0.961 | 0.26 | 0.954 | 0.079 |

For the selected optimization metric average\_precision the average score is 0.934, and the standard deviation is 0.017.

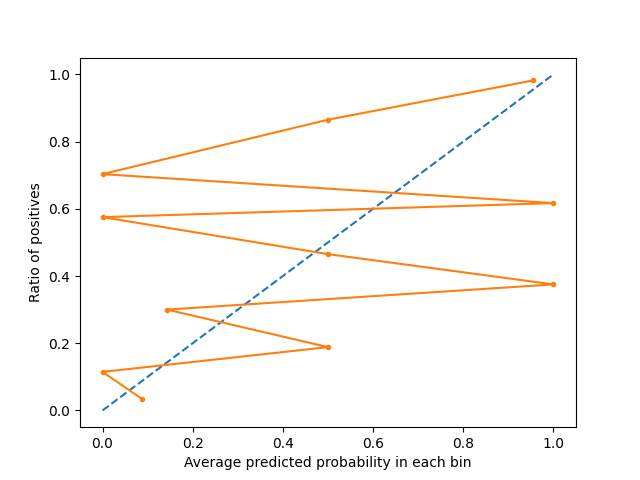
## Main plots

### Calibration plots

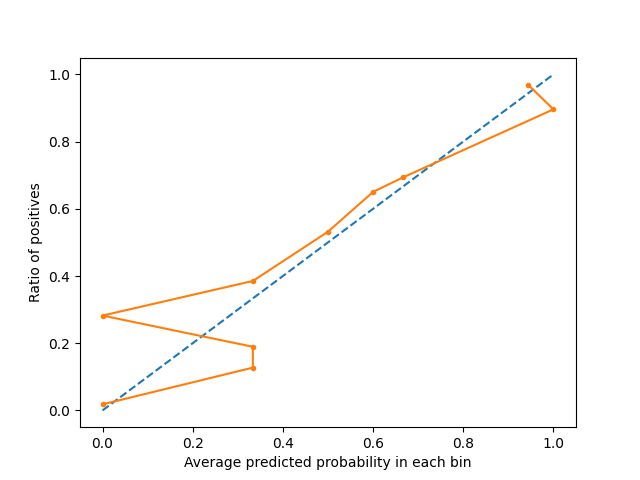
#### Calibration plot of fold 1



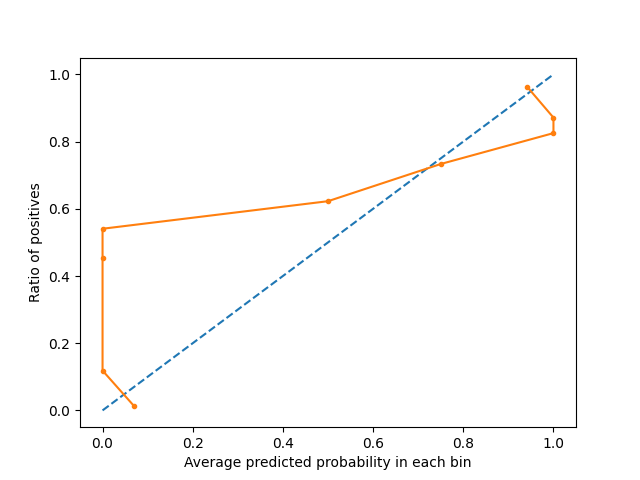
#### Calibration plot of fold 3



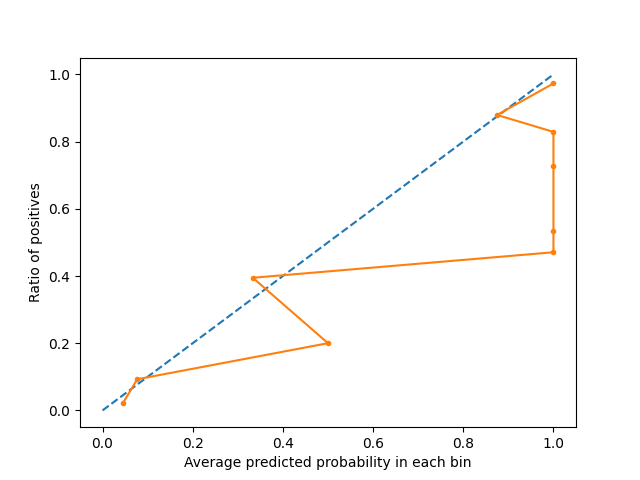
#### Calibration plot of fold 5



#### Calibration plot of fold 7

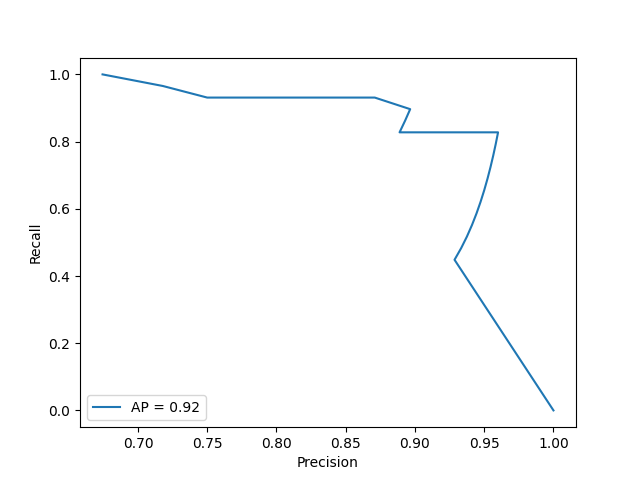


#### Calibration plot of fold 9

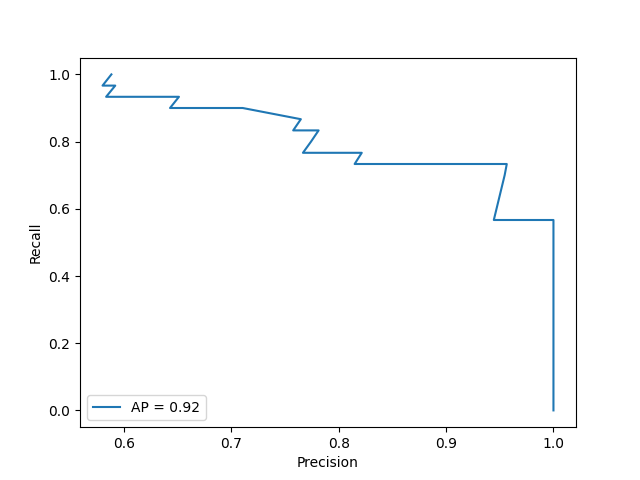


### Precision-recall curve plots

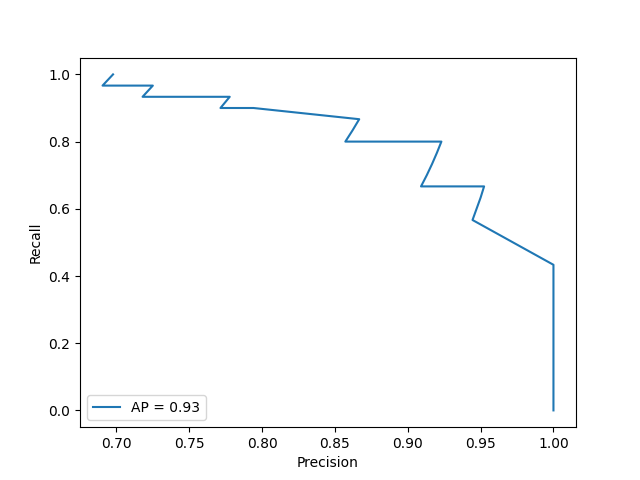
#### Precision-recall curve plot of fold 1



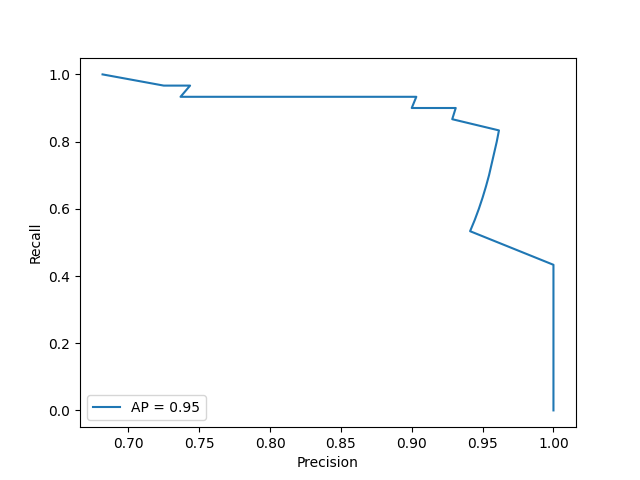
#### Precision-recall curve plot of fold 3



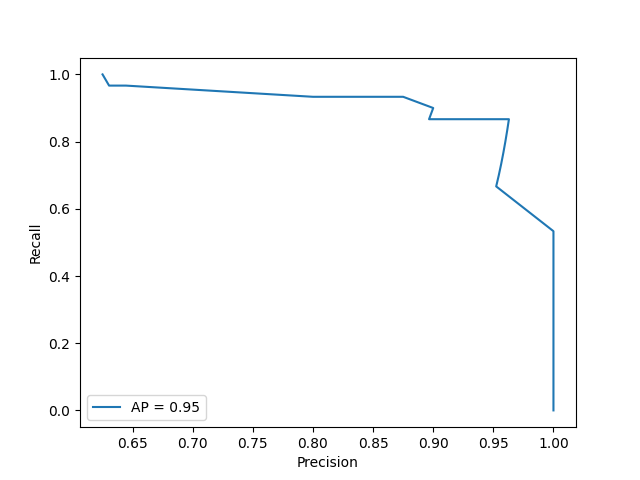
#### Precision-recall curve plot of fold 5



#### Precision-recall curve plot of fold 7

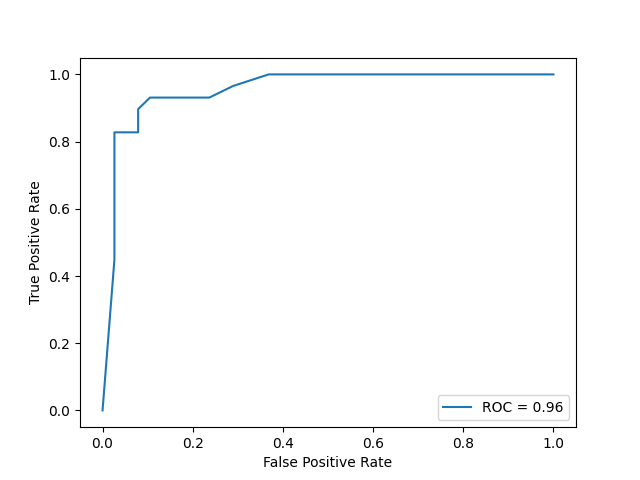


#### Precision-recall curve plot of fold 9

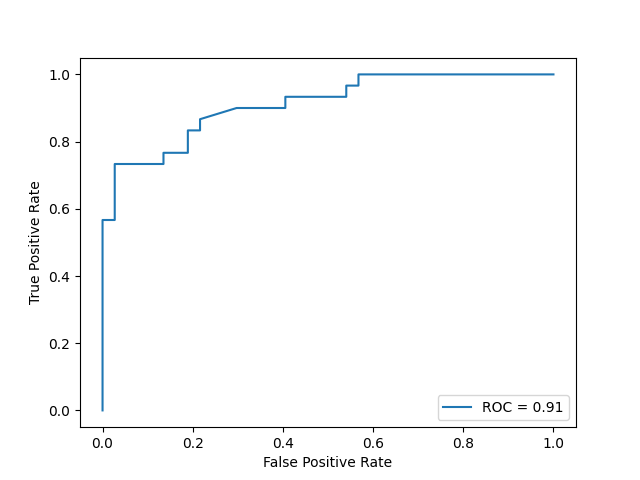


### ROC curve plots

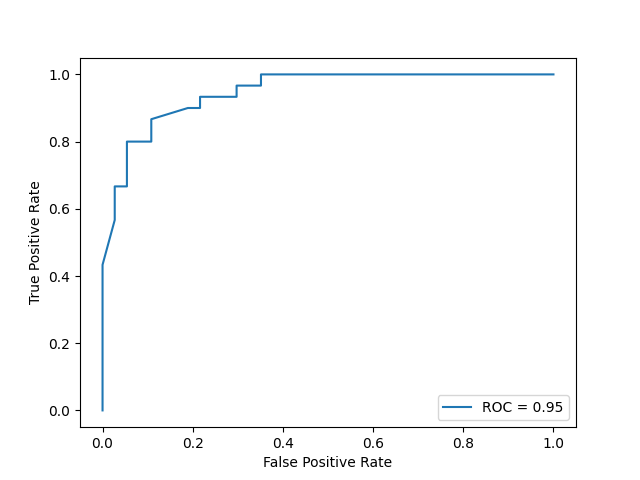
#### ROC curve plot of fold 1



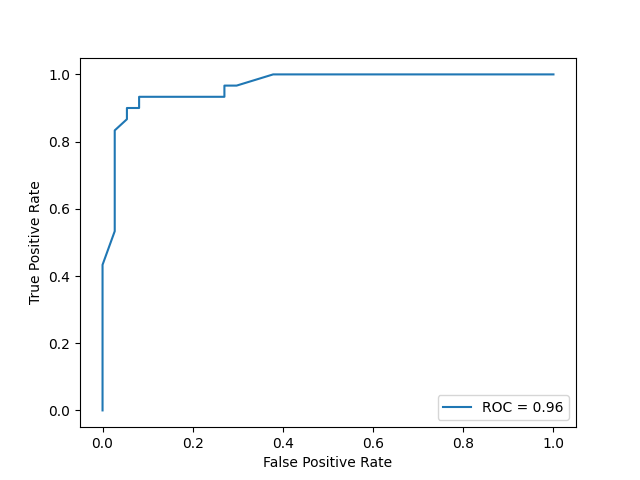
#### ROC curve plot of fold 3



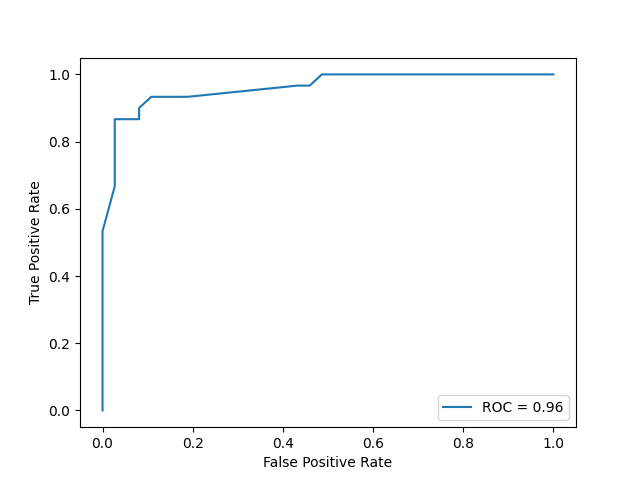
#### ROC curve plot of fold 5



#### ROC curve plot of fold 7

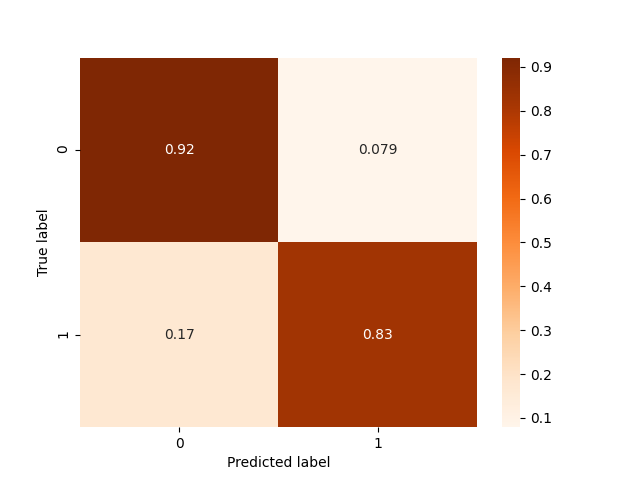


#### ROC curve plot of fold 9

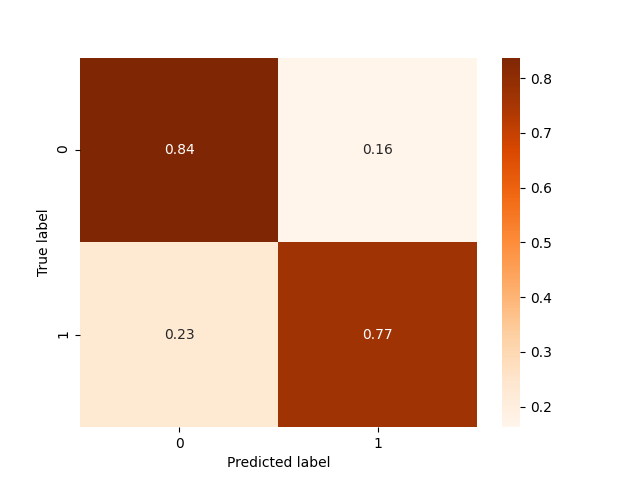


### Confusion matrix

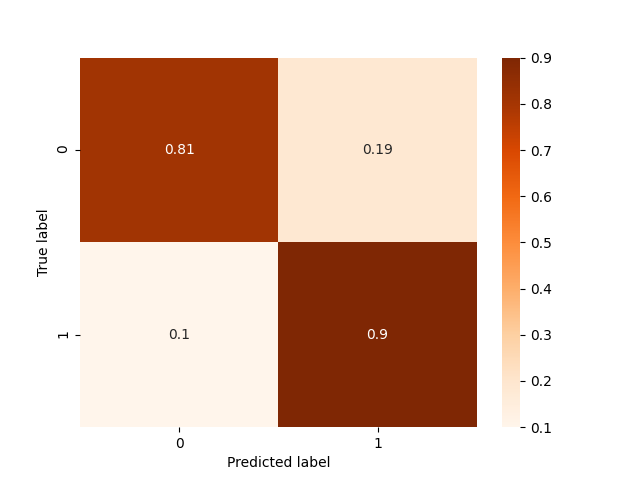
#### Confusion matrix of fold 1



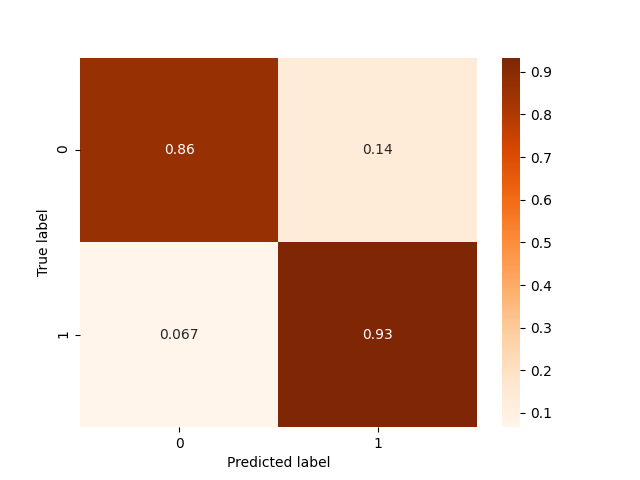
#### Confusion matrix of fold 3



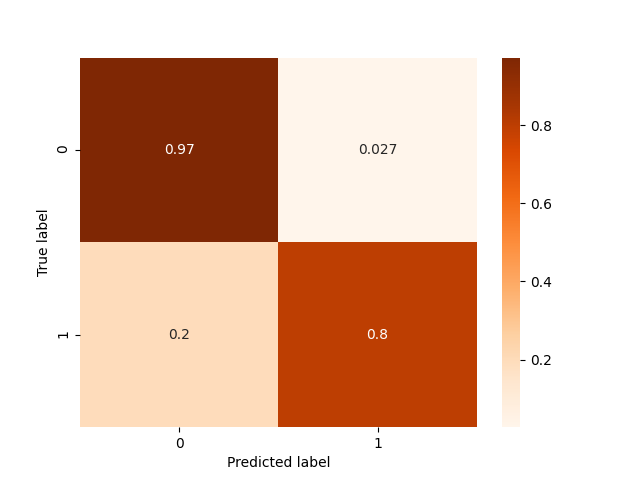
#### Confusion matrix of fold 5



#### Confusion matrix of fold 7

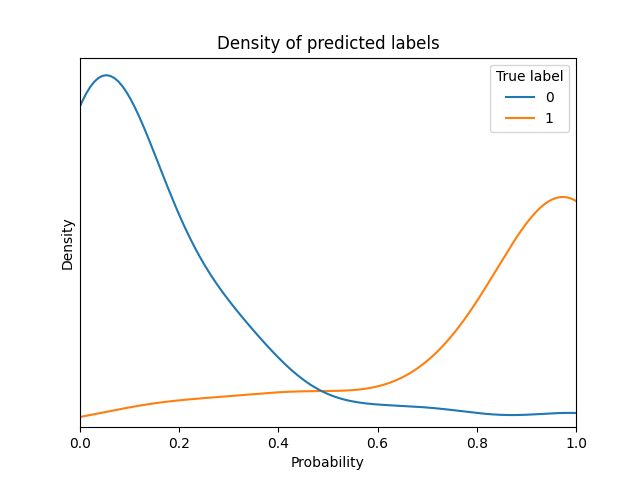


#### Confusion matrix of fold 9

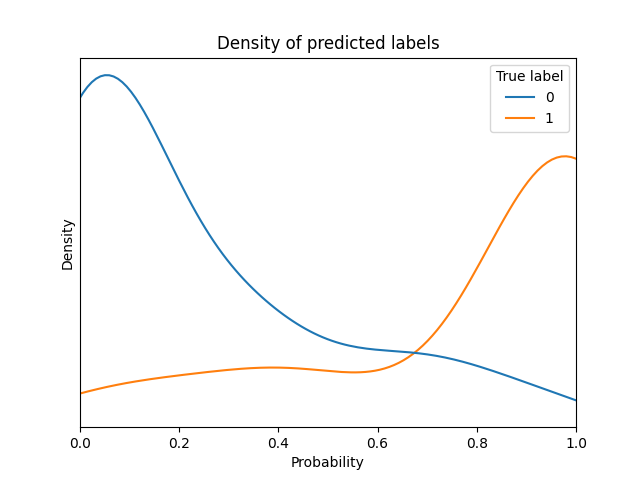


### Histograms

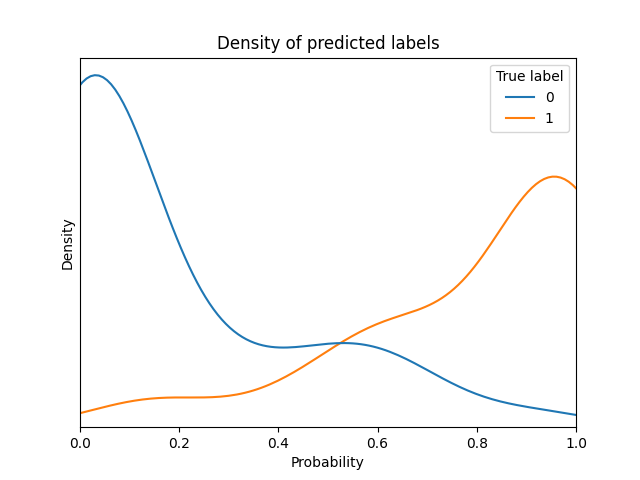
#### Histogram of fold 1



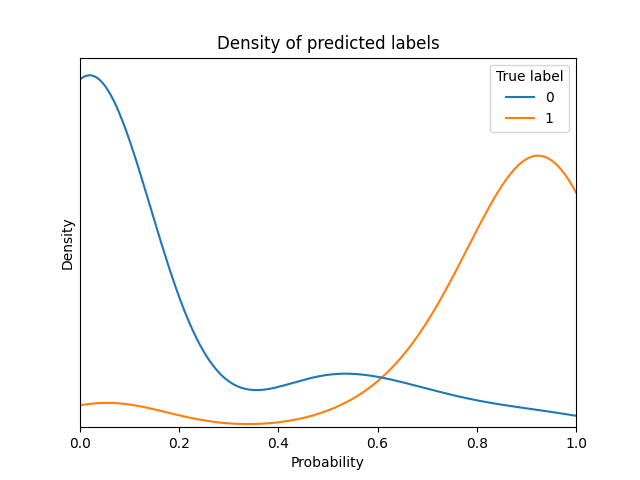
#### Histogram of fold 3



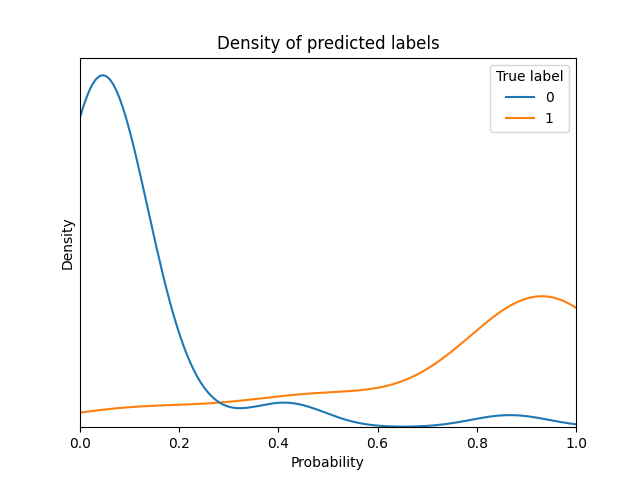
#### Histogram of fold 5



#### Histogram of fold 7



#### Histogram of fold 9



## Comparison of several predictions to assess variance

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Instance** | **Real label** | **Prediction by model of fold 1** | **Prediction by model of fold 3** | **Prediction by model of fold 5** | **Prediction by model of fold 7** | **Prediction by model of fold 9** | **Standard deviation in predictions of this instance** |
| 1 | 0 | 0.0 | 0.052 | 0.009 | 0.028 | 0.028 | 0.018 |
| 2 | 1 | 1.0 | 1.0 | 0.918 | 0.987 | 0.99 | 0.031 |
| 3 | 1 | 0.143 | 0.052 | 0.009 | 0.039 | 0.0 | 0.051 |
| 4 | 0 | 0.219 | 0.052 | 0.009 | 0.113 | 0.079 | 0.071 |
| 5 | 0 | 0.287 | 0.281 | 0.345 | 0.071 | 0.46 | 0.126 |
| 6 | 0 | 0.285 | 0.125 | 0.009 | 0.155 | 0.174 | 0.089 |
| 7 | 1 | 1.0 | 1.0 | 1.0 | 0.925 | 0.936 | 0.034 |
| 8 | 1 | 1.0 | 1.0 | 1.0 | 1.0 | 0.99 | 0.004 |
| 9 | 1 | 1.0 | 1.0 | 1.0 | 0.987 | 0.936 | 0.025 |
| 10 | 0 | 0.119 | 0.177 | 0.385 | 0.028 | 0.103 | 0.121 |
| 11 | 0 | 0.026 | 0.052 | 0.009 | 0.028 | 0.077 | 0.024 |
| 12 | 1 | 1.0 | 1.0 | 1.0 | 0.987 | 0.99 | 0.006 |
| 13 | 1 | 1.0 | 1.0 | 1.0 | 0.987 | 0.99 | 0.006 |
| 14 | 0 | 0.285 | 0.26 | 0.624 | 0.113 | 0.084 | 0.192 |
| 15 | 0 | 0.391 | 0.458 | 0.624 | 0.427 | 0.535 | 0.083 |
| 16 | 0 | 0.363 | 0.172 | 0.009 | 0.155 | 0.107 | 0.116 |
| 17 | 0 | 0.286 | 0.228 | 0.12 | 0.155 | 0.2 | 0.058 |
| 18 | 1 | 0.533 | 0.323 | 0.524 | 0.626 | 0.711 | 0.13 |
| 19 | 0 | 0.839 | 0.896 | 0.918 | 0.842 | 0.731 | 0.065 |
| 20 | 1 | 1.0 | 1.0 | 1.0 | 0.987 | 0.99 | 0.006 |

The average standard deviation is 0.063

# Report of inner trainings

## Report of inner training in fold 1 of outer Cross Validation

## Report of training in this outer fold

Best model with respect to selected metric is RandomForestClassifier with the following params:

* undersampling\_majority\_class: False
* model\_\_bootstrap: 1
* model\_\_n\_estimators: 70
* model\_\_max\_depth: 6
* model\_\_min\_samples\_split: 9
* model\_\_min\_samples\_leaf: 4
* model\_\_max\_features: sqrt
* model\_\_class\_weight: balanced

### Comparison of all models trained in this outer fold

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Model** | **Params** | **Comments** | **roc\_auc on inner fold** | **neg\_log\_loss on inner fold** | **average\_precision on inner fold** | **neg\_brier\_score on inner fold** | **roc\_auc on outer fold** | **neg\_log\_loss on outer fold** | **average\_precision on outer fold** | **neg\_brier\_score on outer fold** |
| XGBClassifier | * undersampling\_majority\_class: True * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_3 * model\_\_max\_depth: 9 * model\_\_learning\_rate: 0.107 * model\_\_min\_child\_weight: 1 * model\_\_subsample: 0.869 * model\_\_colsample\_bytree: 0.185 * model\_\_scale\_pos\_weight: 0.193 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 392.5 * average prop of minority class before resampling: 0.383 * average size of training set after resampling: 484.5 * average prop of minority class after resampling: 0.5 | 0.956 | 0.222 | 0.936 | 0.07 | 0.95 | 0.723 | 0.911 | 0.074 |
| XGBClassifier | * undersampling\_majority\_class: True * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_3 * model\_\_max\_depth: 11 * model\_\_learning\_rate: 0.064 * model\_\_min\_child\_weight: 8 * model\_\_subsample: 0.979 * model\_\_colsample\_bytree: 0.539 * model\_\_scale\_pos\_weight: 0.168 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 392.5 * average prop of minority class before resampling: 0.383 * average size of training set after resampling: 484.5 * average prop of minority class after resampling: 0.5 | 0.943 | 0.252 | 0.923 | 0.081 | 0.95 | 0.741 | 0.914 | 0.08 |
| XGBClassifier | * undersampling\_majority\_class: False * max\_k\_undersampling: 6 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_1 * model\_\_max\_depth: 8 * model\_\_learning\_rate: 0.149 * model\_\_min\_child\_weight: 10 * model\_\_subsample: 0.982 * model\_\_colsample\_bytree: 0.535 * model\_\_scale\_pos\_weight: 0.13 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 542.75 * average prop of minority class before resampling: 0.446 * average size of training set after resampling: 601.0 * average prop of minority class after resampling: 0.5 | 0.908 | 0.329 | 0.877 | 0.104 | 0.916 | 0.353 | 0.91 | 0.106 |
| XGBClassifier | * undersampling\_majority\_class: False * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_2 * model\_\_max\_depth: 5 * model\_\_learning\_rate: 0.067 * model\_\_min\_child\_weight: 9 * model\_\_subsample: 0.812 * model\_\_colsample\_bytree: 0.226 * model\_\_scale\_pos\_weight: 3.374 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 542.75 * average prop of minority class before resampling: 0.446 * average size of training set after resampling: 601.0 * average prop of minority class after resampling: 0.5 | 0.945 | 0.239 | 0.93 | 0.076 | 0.954 | 0.274 | 0.95 | 0.081 |
| XGBClassifier | * undersampling\_majority\_class: True * max\_k\_undersampling: 6 * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_3 * model\_\_max\_depth: 12 * model\_\_learning\_rate: 0.166 * model\_\_min\_child\_weight: 6 * model\_\_subsample: 0.818 * model\_\_colsample\_bytree: 0.189 * model\_\_scale\_pos\_weight: 4.202 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 392.5 * average prop of minority class before resampling: 0.383 * average size of training set after resampling: 484.5 * average prop of minority class after resampling: 0.5 | 0.955 | 0.227 | 0.938 | 0.071 | 0.953 | 0.74 | 0.905 | 0.085 |
| XGBClassifier | * undersampling\_majority\_class: True * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_2 * model\_\_max\_depth: 6 * model\_\_learning\_rate: 0.059 * model\_\_min\_child\_weight: 3 * model\_\_subsample: 0.809 * model\_\_colsample\_bytree: 0.263 * model\_\_scale\_pos\_weight: 0.135 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 392.5 * average prop of minority class before resampling: 0.383 * average size of training set after resampling: 484.5 * average prop of minority class after resampling: 0.5 | 0.948 | 0.231 | 0.93 | 0.074 | 0.937 | 0.302 | 0.934 | 0.089 |
| XGBClassifier | * undersampling\_majority\_class: False * max\_k\_undersampling: 6 * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_1 * model\_\_max\_depth: 6 * model\_\_learning\_rate: 0.271 * model\_\_min\_child\_weight: 10 * model\_\_subsample: 0.927 * model\_\_colsample\_bytree: 0.341 * model\_\_scale\_pos\_weight: 1.275 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 542.75 * average prop of minority class before resampling: 0.446 * average size of training set after resampling: 601.0 * average prop of minority class after resampling: 0.5 | 0.921 | 0.316 | 0.879 | 0.102 | 0.893 | 0.404 | 0.89 | 0.121 |
| XGBClassifier | * undersampling\_majority\_class: False * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_2 * model\_\_max\_depth: 9 * model\_\_learning\_rate: 0.146 * model\_\_min\_child\_weight: 4 * model\_\_subsample: 0.84 * model\_\_colsample\_bytree: 0.268 * model\_\_scale\_pos\_weight: 4.461 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 542.75 * average prop of minority class before resampling: 0.446 * average size of training set after resampling: 601.0 * average prop of minority class after resampling: 0.5 | 0.943 | 0.26 | 0.921 | 0.085 | 0.95 | 0.294 | 0.948 | 0.089 |
| XGBClassifier | * undersampling\_majority\_class: True * max\_k\_undersampling: 6 * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_2 * model\_\_max\_depth: 8 * model\_\_learning\_rate: 0.136 * model\_\_min\_child\_weight: 4 * model\_\_subsample: 0.89 * model\_\_colsample\_bytree: 0.187 * model\_\_scale\_pos\_weight: 2.278 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 392.5 * average prop of minority class before resampling: 0.383 * average size of training set after resampling: 484.5 * average prop of minority class after resampling: 0.5 | 0.946 | 0.246 | 0.932 | 0.077 | 0.953 | 0.274 | 0.948 | 0.081 |
| XGBClassifier | * undersampling\_majority\_class: True * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_3 * model\_\_max\_depth: 12 * model\_\_learning\_rate: 0.072 * model\_\_min\_child\_weight: 6 * model\_\_subsample: 0.818 * model\_\_colsample\_bytree: 0.155 * model\_\_scale\_pos\_weight: 8.15 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 392.5 * average prop of minority class before resampling: 0.383 * average size of training set after resampling: 484.5 * average prop of minority class after resampling: 0.5 | 0.954 | 0.231 | 0.937 | 0.072 | 0.954 | 0.719 | 0.912 | 0.074 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 1 * model\_\_n\_estimators: 95 * model\_\_max\_depth: 7 * model\_\_min\_samples\_split: 15 * model\_\_min\_samples\_leaf: 2 * model\_\_max\_features: sqrt * model\_\_class\_weight: balanced\_subsample | * option: build model without resampling * number of folds: 4 * average size of training set: 392.5 * average prop of minority class: 0.383 | 0.96 | 0.208 | 0.939 | 0.068 | 0.957 | 0.727 | 0.926 | 0.079 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 1 * model\_\_n\_estimators: 19 * model\_\_max\_depth: 8 * model\_\_min\_samples\_split: 12 * model\_\_min\_samples\_leaf: 4 * model\_\_max\_features: sqrt * model\_\_class\_weight: balanced\_subsample | * option: build model without resampling * number of folds: 4 * average size of training set: 392.5 * average prop of minority class: 0.383 | 0.959 | 0.216 | 0.94 | 0.069 | 0.956 | 0.262 | 0.936 | 0.075 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 1 * model\_\_n\_estimators: 63 * model\_\_max\_depth: 10 * model\_\_min\_samples\_split: 13 * model\_\_min\_samples\_leaf: 1 * model\_\_max\_features: sqrt * model\_\_class\_weight: balanced\_subsample | * option: build model without resampling * number of folds: 4 * average size of training set: 542.75 * average prop of minority class: 0.446 | 0.957 | 0.228 | 0.935 | 0.076 | 0.958 | 0.268 | 0.937 | 0.078 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 1 * model\_\_n\_estimators: 75 * model\_\_max\_depth: 8 * model\_\_min\_samples\_split: 7 * model\_\_min\_samples\_leaf: 4 * model\_\_max\_features: auto * model\_\_class\_weight: balanced | * option: build model without resampling * number of folds: 4 * average size of training set: 392.5 * average prop of minority class: 0.383 | 0.957 | 0.228 | 0.939 | 0.074 | 0.955 | 0.708 | 0.911 | 0.071 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 1 * model\_\_n\_estimators: 38 * model\_\_max\_depth: 7 * model\_\_min\_samples\_split: 10 * model\_\_min\_samples\_leaf: 3 * model\_\_max\_features: sqrt * model\_\_class\_weight: balanced\_subsample | * option: build model without resampling * number of folds: 4 * average size of training set: 542.75 * average prop of minority class: 0.446 | 0.961 | 0.216 | 0.941 | 0.07 | 0.957 | 0.721 | 0.915 | 0.077 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 1 * model\_\_n\_estimators: 67 * model\_\_max\_depth: 6 * model\_\_min\_samples\_split: 16 * model\_\_min\_samples\_leaf: 3 * model\_\_max\_features: auto * model\_\_class\_weight: balanced | * option: build model without resampling * number of folds: 4 * average size of training set: 392.5 * average prop of minority class: 0.383 | 0.959 | 0.215 | 0.94 | 0.07 | 0.953 | 0.717 | 0.915 | 0.074 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 1 * model\_\_n\_estimators: 37 * model\_\_max\_depth: 5 * model\_\_min\_samples\_split: 19 * model\_\_min\_samples\_leaf: 2 * model\_\_max\_features: sqrt * model\_\_class\_weight: balanced | * option: build model without resampling * number of folds: 4 * average size of training set: 542.75 * average prop of minority class: 0.446 | 0.959 | 0.214 | 0.939 | 0.071 | 0.948 | 0.725 | 0.914 | 0.077 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 1 * model\_\_n\_estimators: 70 * model\_\_max\_depth: 6 * model\_\_min\_samples\_split: 9 * model\_\_min\_samples\_leaf: 4 * model\_\_max\_features: sqrt * model\_\_class\_weight: balanced | * option: build model without resampling * number of folds: 4 * average size of training set: 542.75 * average prop of minority class: 0.446 | 0.962 | 0.208 | 0.944 | 0.069 | 0.955 | 0.719 | 0.916 | 0.077 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 0 * model\_\_n\_estimators: 36 * model\_\_max\_depth: 2 * model\_\_min\_samples\_split: 8 * model\_\_min\_samples\_leaf: 3 * model\_\_max\_features: auto * model\_\_class\_weight: balanced\_subsample | * option: build model without resampling * number of folds: 4 * average size of training set: 392.5 * average prop of minority class: 0.383 | 0.947 | 0.241 | 0.925 | 0.08 | 0.938 | 0.759 | 0.897 | 0.087 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 1 * model\_\_n\_estimators: 75 * model\_\_max\_depth: 4 * model\_\_min\_samples\_split: 5 * model\_\_min\_samples\_leaf: 2 * model\_\_max\_features: auto * model\_\_class\_weight: balanced | * option: build model without resampling * number of folds: 4 * average size of training set: 392.5 * average prop of minority class: 0.383 | 0.956 | 0.225 | 0.936 | 0.075 | 0.948 | 0.729 | 0.917 | 0.08 |

## Report of inner training in fold 3 of outer Cross Validation

## Report of training in this outer fold

Best model with respect to selected metric is XGBClassifier with the following params:

* undersampling\_majority\_class: False
* max\_k\_undersampling: 6
* resample\_\_sampling\_strategy: minority
* post\_process\_\_option: option\_3
* model\_\_max\_depth: 7
* model\_\_learning\_rate: 0.084
* model\_\_min\_child\_weight: 10
* model\_\_subsample: 0.806
* model\_\_colsample\_bytree: 0.246
* model\_\_scale\_pos\_weight: 1.374
* model\_\_objective: binary:logistic

### Comparison of all models trained in this outer fold

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Model** | **Params** | **Comments** | **roc\_auc on inner fold** | **neg\_log\_loss on inner fold** | **average\_precision on inner fold** | **neg\_brier\_score on inner fold** | **roc\_auc on outer fold** | **neg\_log\_loss on outer fold** | **average\_precision on outer fold** | **neg\_brier\_score on outer fold** |
| XGBClassifier | * undersampling\_majority\_class: True * max\_k\_undersampling: 6 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_3 * model\_\_max\_depth: 11 * model\_\_learning\_rate: 0.056 * model\_\_min\_child\_weight: 8 * model\_\_subsample: 0.884 * model\_\_colsample\_bytree: 0.147 * model\_\_scale\_pos\_weight: 0.146 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 392.0 * average prop of minority class before resampling: 0.385 * average size of training set after resampling: 482.5 * average prop of minority class after resampling: 0.5 | 0.95 | 0.247 | 0.929 | 0.077 | 0.904 | 0.389 | 0.909 | 0.12 |
| XGBClassifier | * undersampling\_majority\_class: True * max\_k\_undersampling: 6 * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_3 * model\_\_max\_depth: 13 * model\_\_learning\_rate: 0.112 * model\_\_min\_child\_weight: 2 * model\_\_subsample: 0.816 * model\_\_colsample\_bytree: 0.136 * model\_\_scale\_pos\_weight: 3.677 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 392.0 * average prop of minority class before resampling: 0.385 * average size of training set after resampling: 482.5 * average prop of minority class after resampling: 0.5 | 0.95 | 0.253 | 0.927 | 0.079 | 0.909 | 0.385 | 0.899 | 0.119 |
| XGBClassifier | * undersampling\_majority\_class: True * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_2 * model\_\_max\_depth: 12 * model\_\_learning\_rate: 0.286 * model\_\_min\_child\_weight: 7 * model\_\_subsample: 0.979 * model\_\_colsample\_bytree: 0.141 * model\_\_scale\_pos\_weight: 2.402 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 392.0 * average prop of minority class before resampling: 0.385 * average size of training set after resampling: 482.5 * average prop of minority class after resampling: 0.5 | 0.96 | 0.21 | 0.947 | 0.068 | 0.895 | 0.452 | 0.895 | 0.149 |
| XGBClassifier | * undersampling\_majority\_class: True * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_3 * model\_\_max\_depth: 13 * model\_\_learning\_rate: 0.3 * model\_\_min\_child\_weight: 6 * model\_\_subsample: 0.876 * model\_\_colsample\_bytree: 0.195 * model\_\_scale\_pos\_weight: 0.2 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 392.0 * average prop of minority class before resampling: 0.385 * average size of training set after resampling: 482.5 * average prop of minority class after resampling: 0.5 | 0.948 | 0.246 | 0.925 | 0.077 | 0.891 | 0.398 | 0.898 | 0.129 |
| XGBClassifier | * undersampling\_majority\_class: False * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_1 * model\_\_max\_depth: 7 * model\_\_learning\_rate: 0.066 * model\_\_min\_child\_weight: 6 * model\_\_subsample: 0.843 * model\_\_colsample\_bytree: 0.247 * model\_\_scale\_pos\_weight: 0.736 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 542.75 * average prop of minority class before resampling: 0.444 * average size of training set after resampling: 603.0 * average prop of minority class after resampling: 0.5 | 0.931 | 0.296 | 0.89 | 0.095 | 0.869 | 0.475 | 0.872 | 0.149 |
| XGBClassifier | * undersampling\_majority\_class: True * max\_k\_undersampling: 6 * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_1 * model\_\_max\_depth: 7 * model\_\_learning\_rate: 0.265 * model\_\_min\_child\_weight: 10 * model\_\_subsample: 0.936 * model\_\_colsample\_bytree: 0.473 * model\_\_scale\_pos\_weight: 0.185 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 392.0 * average prop of minority class before resampling: 0.385 * average size of training set after resampling: 482.5 * average prop of minority class after resampling: 0.5 | 0.932 | 0.279 | 0.896 | 0.091 | 0.849 | 0.975 | 0.78 | 0.162 |
| XGBClassifier | * undersampling\_majority\_class: True * max\_k\_undersampling: 6 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_3 * model\_\_max\_depth: 15 * model\_\_learning\_rate: 0.074 * model\_\_min\_child\_weight: 10 * model\_\_subsample: 0.826 * model\_\_colsample\_bytree: 0.548 * model\_\_scale\_pos\_weight: 1.048 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 392.0 * average prop of minority class before resampling: 0.385 * average size of training set after resampling: 482.5 * average prop of minority class after resampling: 0.5 | 0.96 | 0.217 | 0.94 | 0.069 | 0.905 | 0.842 | 0.912 | 0.124 |
| XGBClassifier | * undersampling\_majority\_class: True * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_2 * model\_\_max\_depth: 15 * model\_\_learning\_rate: 0.075 * model\_\_min\_child\_weight: 4 * model\_\_subsample: 0.897 * model\_\_colsample\_bytree: 0.436 * model\_\_scale\_pos\_weight: 1.027 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 392.0 * average prop of minority class before resampling: 0.385 * average size of training set after resampling: 482.5 * average prop of minority class after resampling: 0.5 | 0.959 | 0.22 | 0.944 | 0.07 | 0.892 | 0.44 | 0.896 | 0.147 |
| XGBClassifier | * undersampling\_majority\_class: False * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_3 * model\_\_max\_depth: 14 * model\_\_learning\_rate: 0.161 * model\_\_min\_child\_weight: 10 * model\_\_subsample: 0.922 * model\_\_colsample\_bytree: 0.503 * model\_\_scale\_pos\_weight: 3.639 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 542.75 * average prop of minority class before resampling: 0.444 * average size of training set after resampling: 603.0 * average prop of minority class after resampling: 0.5 | 0.962 | 0.202 | 0.941 | 0.066 | 0.91 | 0.367 | 0.917 | 0.114 |
| XGBClassifier | * undersampling\_majority\_class: False * max\_k\_undersampling: 6 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_3 * model\_\_max\_depth: 7 * model\_\_learning\_rate: 0.084 * model\_\_min\_child\_weight: 10 * model\_\_subsample: 0.806 * model\_\_colsample\_bytree: 0.246 * model\_\_scale\_pos\_weight: 1.374 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 542.75 * average prop of minority class before resampling: 0.444 * average size of training set after resampling: 603.0 * average prop of minority class after resampling: 0.5 | 0.966 | 0.202 | 0.948 | 0.065 | 0.912 | 0.395 | 0.915 | 0.129 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 0 * model\_\_n\_estimators: 70 * model\_\_max\_depth: 8 * model\_\_min\_samples\_split: 19 * model\_\_min\_samples\_leaf: 1 * model\_\_max\_features: sqrt * model\_\_class\_weight: balanced | * option: build model without resampling * number of folds: 4 * average size of training set: 392.0 * average prop of minority class: 0.385 | 0.966 | 0.199 | 0.942 | 0.064 | 0.91 | 0.398 | 0.904 | 0.132 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 1 * model\_\_n\_estimators: 65 * model\_\_max\_depth: 3 * model\_\_min\_samples\_split: 10 * model\_\_min\_samples\_leaf: 4 * model\_\_max\_features: sqrt * model\_\_class\_weight: balanced\_subsample | * option: build model without resampling * number of folds: 4 * average size of training set: 542.75 * average prop of minority class: 0.444 | 0.955 | 0.229 | 0.933 | 0.073 | 0.904 | 0.408 | 0.898 | 0.131 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 1 * model\_\_n\_estimators: 64 * model\_\_max\_depth: 8 * model\_\_min\_samples\_split: 5 * model\_\_min\_samples\_leaf: 3 * model\_\_max\_features: auto * model\_\_class\_weight: balanced\_subsample | * option: build model without resampling * number of folds: 4 * average size of training set: 542.75 * average prop of minority class: 0.444 | 0.965 | 0.203 | 0.945 | 0.066 | 0.914 | 0.383 | 0.911 | 0.127 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 0 * model\_\_n\_estimators: 17 * model\_\_max\_depth: 2 * model\_\_min\_samples\_split: 14 * model\_\_min\_samples\_leaf: 4 * model\_\_max\_features: auto * model\_\_class\_weight: balanced | * option: build model without resampling * number of folds: 4 * average size of training set: 542.75 * average prop of minority class: 0.444 | 0.95 | 0.249 | 0.925 | 0.082 | 0.891 | 0.442 | 0.893 | 0.144 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 0 * model\_\_n\_estimators: 95 * model\_\_max\_depth: 10 * model\_\_min\_samples\_split: 14 * model\_\_min\_samples\_leaf: 1 * model\_\_max\_features: sqrt * model\_\_class\_weight: balanced | * option: build model without resampling * number of folds: 4 * average size of training set: 392.0 * average prop of minority class: 0.385 | 0.964 | 0.203 | 0.942 | 0.065 | 0.92 | 0.363 | 0.921 | 0.125 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 0 * model\_\_n\_estimators: 84 * model\_\_max\_depth: 6 * model\_\_min\_samples\_split: 6 * model\_\_min\_samples\_leaf: 3 * model\_\_max\_features: sqrt * model\_\_class\_weight: balanced | * option: build model without resampling * number of folds: 4 * average size of training set: 392.0 * average prop of minority class: 0.385 | 0.96 | 0.223 | 0.941 | 0.072 | 0.922 | 0.364 | 0.922 | 0.123 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 1 * model\_\_n\_estimators: 17 * model\_\_max\_depth: 7 * model\_\_min\_samples\_split: 8 * model\_\_min\_samples\_leaf: 4 * model\_\_max\_features: auto * model\_\_class\_weight: balanced | * option: build model without resampling * number of folds: 4 * average size of training set: 542.75 * average prop of minority class: 0.444 | 0.958 | 0.229 | 0.938 | 0.072 | 0.901 | 0.387 | 0.909 | 0.122 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 1 * model\_\_n\_estimators: 95 * model\_\_max\_depth: 7 * model\_\_min\_samples\_split: 9 * model\_\_min\_samples\_leaf: 2 * model\_\_max\_features: sqrt * model\_\_class\_weight: balanced\_subsample | * option: build model without resampling * number of folds: 4 * average size of training set: 542.75 * average prop of minority class: 0.444 | 0.966 | 0.194 | 0.946 | 0.061 | 0.913 | 0.406 | 0.91 | 0.13 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 0 * model\_\_n\_estimators: 42 * model\_\_max\_depth: 9 * model\_\_min\_samples\_split: 13 * model\_\_min\_samples\_leaf: 3 * model\_\_max\_features: sqrt * model\_\_class\_weight: balanced\_subsample | * option: build model without resampling * number of folds: 4 * average size of training set: 392.0 * average prop of minority class: 0.385 | 0.96 | 0.217 | 0.939 | 0.069 | 0.919 | 0.381 | 0.918 | 0.127 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 0 * model\_\_n\_estimators: 47 * model\_\_max\_depth: 6 * model\_\_min\_samples\_split: 7 * model\_\_min\_samples\_leaf: 1 * model\_\_max\_features: sqrt * model\_\_class\_weight: balanced\_subsample | * option: build model without resampling * number of folds: 4 * average size of training set: 392.0 * average prop of minority class: 0.385 | 0.962 | 0.21 | 0.942 | 0.068 | 0.915 | 0.394 | 0.91 | 0.13 |

## Report of inner training in fold 5 of outer Cross Validation

## Report of training in this outer fold

Best model with respect to selected metric is XGBClassifier with the following params:

* undersampling\_majority\_class: True
* max\_k\_undersampling: 5
* resample\_\_sampling\_strategy: all
* post\_process\_\_option: option\_3
* model\_\_max\_depth: 12
* model\_\_learning\_rate: 0.085
* model\_\_min\_child\_weight: 4
* model\_\_subsample: 0.89
* model\_\_colsample\_bytree: 0.139
* model\_\_scale\_pos\_weight: 0.318
* model\_\_objective: binary:logistic

### Comparison of all models trained in this outer fold

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Model** | **Params** | **Comments** | **roc\_auc on inner fold** | **neg\_log\_loss on inner fold** | **average\_precision on inner fold** | **neg\_brier\_score on inner fold** | **roc\_auc on outer fold** | **neg\_log\_loss on outer fold** | **average\_precision on outer fold** | **neg\_brier\_score on outer fold** |
| XGBClassifier | * undersampling\_majority\_class: True * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_3 * model\_\_max\_depth: 12 * model\_\_learning\_rate: 0.085 * model\_\_min\_child\_weight: 4 * model\_\_subsample: 0.89 * model\_\_colsample\_bytree: 0.139 * model\_\_scale\_pos\_weight: 0.318 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 392.0 * average prop of minority class before resampling: 0.385 * average size of training set after resampling: 482.5 * average prop of minority class after resampling: 0.5 | 0.96 | 0.215 | 0.934 | 0.07 | 0.947 | 0.294 | 0.934 | 0.097 |
| XGBClassifier | * undersampling\_majority\_class: False * max\_k\_undersampling: 6 * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_2 * model\_\_max\_depth: 15 * model\_\_learning\_rate: 0.077 * model\_\_min\_child\_weight: 1 * model\_\_subsample: 0.842 * model\_\_colsample\_bytree: 0.514 * model\_\_scale\_pos\_weight: 2.086 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 542.75 * average prop of minority class before resampling: 0.444 * average size of training set after resampling: 603.0 * average prop of minority class after resampling: 0.5 | 0.942 | 0.246 | 0.909 | 0.079 | 0.947 | 0.297 | 0.935 | 0.092 |
| XGBClassifier | * undersampling\_majority\_class: True * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_1 * model\_\_max\_depth: 6 * model\_\_learning\_rate: 0.078 * model\_\_min\_child\_weight: 3 * model\_\_subsample: 0.926 * model\_\_colsample\_bytree: 0.757 * model\_\_scale\_pos\_weight: 9.152 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 392.0 * average prop of minority class before resampling: 0.385 * average size of training set after resampling: 482.5 * average prop of minority class after resampling: 0.5 | 0.924 | 0.315 | 0.88 | 0.101 | 0.909 | 0.373 | 0.899 | 0.113 |
| XGBClassifier | * undersampling\_majority\_class: False * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_2 * model\_\_max\_depth: 6 * model\_\_learning\_rate: 0.082 * model\_\_min\_child\_weight: 4 * model\_\_subsample: 0.871 * model\_\_colsample\_bytree: 0.23 * model\_\_scale\_pos\_weight: 1.05 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 542.75 * average prop of minority class before resampling: 0.444 * average size of training set after resampling: 603.0 * average prop of minority class after resampling: 0.5 | 0.947 | 0.246 | 0.909 | 0.08 | 0.949 | 0.31 | 0.945 | 0.089 |
| XGBClassifier | * undersampling\_majority\_class: False * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_1 * model\_\_max\_depth: 9 * model\_\_learning\_rate: 0.184 * model\_\_min\_child\_weight: 4 * model\_\_subsample: 0.96 * model\_\_colsample\_bytree: 0.514 * model\_\_scale\_pos\_weight: 1.384 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 542.75 * average prop of minority class before resampling: 0.444 * average size of training set after resampling: 603.0 * average prop of minority class after resampling: 0.5 | 0.916 | 0.335 | 0.865 | 0.109 | 0.912 | 0.377 | 0.895 | 0.121 |
| XGBClassifier | * undersampling\_majority\_class: False * max\_k\_undersampling: 6 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_2 * model\_\_max\_depth: 13 * model\_\_learning\_rate: 0.121 * model\_\_min\_child\_weight: 9 * model\_\_subsample: 0.897 * model\_\_colsample\_bytree: 0.221 * model\_\_scale\_pos\_weight: 0.274 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 542.75 * average prop of minority class before resampling: 0.444 * average size of training set after resampling: 603.0 * average prop of minority class after resampling: 0.5 | 0.954 | 0.217 | 0.923 | 0.07 | 0.951 | 0.284 | 0.926 | 0.09 |
| XGBClassifier | * undersampling\_majority\_class: False * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_2 * model\_\_max\_depth: 10 * model\_\_learning\_rate: 0.128 * model\_\_min\_child\_weight: 3 * model\_\_subsample: 0.891 * model\_\_colsample\_bytree: 0.178 * model\_\_scale\_pos\_weight: 0.276 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 542.75 * average prop of minority class before resampling: 0.444 * average size of training set after resampling: 603.0 * average prop of minority class after resampling: 0.5 | 0.948 | 0.236 | 0.909 | 0.076 | 0.951 | 0.295 | 0.941 | 0.092 |
| XGBClassifier | * undersampling\_majority\_class: False * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_3 * model\_\_max\_depth: 13 * model\_\_learning\_rate: 0.054 * model\_\_min\_child\_weight: 1 * model\_\_subsample: 0.856 * model\_\_colsample\_bytree: 0.37 * model\_\_scale\_pos\_weight: 4.761 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 542.75 * average prop of minority class before resampling: 0.444 * average size of training set after resampling: 603.0 * average prop of minority class after resampling: 0.5 | 0.959 | 0.215 | 0.933 | 0.07 | 0.945 | 0.314 | 0.918 | 0.096 |
| XGBClassifier | * undersampling\_majority\_class: False * max\_k\_undersampling: 6 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_2 * model\_\_max\_depth: 14 * model\_\_learning\_rate: 0.064 * model\_\_min\_child\_weight: 3 * model\_\_subsample: 0.96 * model\_\_colsample\_bytree: 0.192 * model\_\_scale\_pos\_weight: 8.429 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 542.75 * average prop of minority class before resampling: 0.444 * average size of training set after resampling: 603.0 * average prop of minority class after resampling: 0.5 | 0.943 | 0.24 | 0.905 | 0.08 | 0.945 | 0.312 | 0.931 | 0.098 |
| XGBClassifier | * undersampling\_majority\_class: False * max\_k\_undersampling: 6 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_3 * model\_\_max\_depth: 11 * model\_\_learning\_rate: 0.126 * model\_\_min\_child\_weight: 10 * model\_\_subsample: 0.925 * model\_\_colsample\_bytree: 0.651 * model\_\_scale\_pos\_weight: 0.962 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 542.75 * average prop of minority class before resampling: 0.444 * average size of training set after resampling: 603.0 * average prop of minority class after resampling: 0.5 | 0.956 | 0.208 | 0.924 | 0.065 | 0.961 | 0.268 | 0.943 | 0.083 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 1 * model\_\_n\_estimators: 13 * model\_\_max\_depth: 9 * model\_\_min\_samples\_split: 12 * model\_\_min\_samples\_leaf: 3 * model\_\_max\_features: auto * model\_\_class\_weight: balanced\_subsample | * option: build model without resampling * number of folds: 4 * average size of training set: 542.75 * average prop of minority class: 0.444 | 0.951 | 0.236 | 0.923 | 0.078 | 0.95 | 0.298 | 0.911 | 0.086 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 0 * model\_\_n\_estimators: 38 * model\_\_max\_depth: 4 * model\_\_min\_samples\_split: 9 * model\_\_min\_samples\_leaf: 1 * model\_\_max\_features: sqrt * model\_\_class\_weight: balanced | * option: build model without resampling * number of folds: 4 * average size of training set: 392.0 * average prop of minority class: 0.385 | 0.957 | 0.208 | 0.922 | 0.069 | 0.955 | 0.255 | 0.942 | 0.084 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 1 * model\_\_n\_estimators: 35 * model\_\_max\_depth: 10 * model\_\_min\_samples\_split: 16 * model\_\_min\_samples\_leaf: 4 * model\_\_max\_features: sqrt * model\_\_class\_weight: balanced | * option: build model without resampling * number of folds: 4 * average size of training set: 542.75 * average prop of minority class: 0.444 | 0.957 | 0.209 | 0.923 | 0.066 | 0.955 | 0.269 | 0.937 | 0.079 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 1 * model\_\_n\_estimators: 59 * model\_\_max\_depth: 5 * model\_\_min\_samples\_split: 16 * model\_\_min\_samples\_leaf: 3 * model\_\_max\_features: auto * model\_\_class\_weight: balanced | * option: build model without resampling * number of folds: 4 * average size of training set: 392.0 * average prop of minority class: 0.385 | 0.956 | 0.221 | 0.927 | 0.073 | 0.955 | 0.256 | 0.939 | 0.081 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 1 * model\_\_n\_estimators: 11 * model\_\_max\_depth: 10 * model\_\_min\_samples\_split: 8 * model\_\_min\_samples\_leaf: 1 * model\_\_max\_features: sqrt * model\_\_class\_weight: balanced\_subsample | * option: build model without resampling * number of folds: 4 * average size of training set: 392.0 * average prop of minority class: 0.385 | 0.954 | 0.232 | 0.923 | 0.076 | 0.953 | 0.731 | 0.912 | 0.084 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 1 * model\_\_n\_estimators: 81 * model\_\_max\_depth: 10 * model\_\_min\_samples\_split: 17 * model\_\_min\_samples\_leaf: 4 * model\_\_max\_features: sqrt * model\_\_class\_weight: balanced\_subsample | * option: build model without resampling * number of folds: 4 * average size of training set: 542.75 * average prop of minority class: 0.444 | 0.96 | 0.214 | 0.929 | 0.07 | 0.955 | 0.268 | 0.937 | 0.083 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 0 * model\_\_n\_estimators: 59 * model\_\_max\_depth: 2 * model\_\_min\_samples\_split: 14 * model\_\_min\_samples\_leaf: 2 * model\_\_max\_features: sqrt * model\_\_class\_weight: balanced | * option: build model without resampling * number of folds: 4 * average size of training set: 392.0 * average prop of minority class: 0.385 | 0.944 | 0.253 | 0.909 | 0.083 | 0.949 | 0.299 | 0.94 | 0.092 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 1 * model\_\_n\_estimators: 58 * model\_\_max\_depth: 4 * model\_\_min\_samples\_split: 11 * model\_\_min\_samples\_leaf: 2 * model\_\_max\_features: auto * model\_\_class\_weight: balanced\_subsample | * option: build model without resampling * number of folds: 4 * average size of training set: 542.75 * average prop of minority class: 0.444 | 0.955 | 0.222 | 0.922 | 0.073 | 0.952 | 0.258 | 0.932 | 0.082 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 1 * model\_\_n\_estimators: 54 * model\_\_max\_depth: 8 * model\_\_min\_samples\_split: 13 * model\_\_min\_samples\_leaf: 3 * model\_\_max\_features: auto * model\_\_class\_weight: balanced | * option: build model without resampling * number of folds: 4 * average size of training set: 542.75 * average prop of minority class: 0.444 | 0.956 | 0.23 | 0.925 | 0.075 | 0.955 | 0.265 | 0.943 | 0.082 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 0 * model\_\_n\_estimators: 92 * model\_\_max\_depth: 8 * model\_\_min\_samples\_split: 9 * model\_\_min\_samples\_leaf: 3 * model\_\_max\_features: sqrt * model\_\_class\_weight: balanced | * option: build model without resampling * number of folds: 4 * average size of training set: 542.75 * average prop of minority class: 0.444 | 0.959 | 0.216 | 0.931 | 0.072 | 0.96 | 0.275 | 0.935 | 0.084 |

## Report of inner training in fold 7 of outer Cross Validation

## Report of training in this outer fold

Best model with respect to selected metric is RandomForestClassifier with the following params:

* undersampling\_majority\_class: False
* model\_\_bootstrap: 0
* model\_\_n\_estimators: 62
* model\_\_max\_depth: 6
* model\_\_min\_samples\_split: 6
* model\_\_min\_samples\_leaf: 4
* model\_\_max\_features: sqrt
* model\_\_class\_weight: balanced

### Comparison of all models trained in this outer fold

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Model** | **Params** | **Comments** | **roc\_auc on inner fold** | **neg\_log\_loss on inner fold** | **average\_precision on inner fold** | **neg\_brier\_score on inner fold** | **roc\_auc on outer fold** | **neg\_log\_loss on outer fold** | **average\_precision on outer fold** | **neg\_brier\_score on outer fold** |
| XGBClassifier | * undersampling\_majority\_class: True * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_2 * model\_\_max\_depth: 13 * model\_\_learning\_rate: 0.083 * model\_\_min\_child\_weight: 7 * model\_\_subsample: 0.928 * model\_\_colsample\_bytree: 0.676 * model\_\_scale\_pos\_weight: 0.11 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 392.0 * average prop of minority class before resampling: 0.385 * average size of training set after resampling: 482.5 * average prop of minority class after resampling: 0.5 | 0.939 | 0.27 | 0.897 | 0.083 | 0.936 | 0.337 | 0.924 | 0.096 |
| XGBClassifier | * undersampling\_majority\_class: True * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_1 * model\_\_max\_depth: 9 * model\_\_learning\_rate: 0.226 * model\_\_min\_child\_weight: 8 * model\_\_subsample: 0.933 * model\_\_colsample\_bytree: 0.217 * model\_\_scale\_pos\_weight: 0.135 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 392.0 * average prop of minority class before resampling: 0.385 * average size of training set after resampling: 482.5 * average prop of minority class after resampling: 0.5 | 0.912 | 0.335 | 0.874 | 0.11 | 0.936 | 0.331 | 0.914 | 0.105 |
| XGBClassifier | * undersampling\_majority\_class: False * max\_k\_undersampling: 6 * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_3 * model\_\_max\_depth: 12 * model\_\_learning\_rate: 0.132 * model\_\_min\_child\_weight: 2 * model\_\_subsample: 0.845 * model\_\_colsample\_bytree: 0.311 * model\_\_scale\_pos\_weight: 2.026 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 542.75 * average prop of minority class before resampling: 0.444 * average size of training set after resampling: 603.0 * average prop of minority class after resampling: 0.5 | 0.947 | 0.242 | 0.91 | 0.079 | 0.963 | 0.269 | 0.963 | 0.077 |
| XGBClassifier | * undersampling\_majority\_class: True * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_2 * model\_\_max\_depth: 7 * model\_\_learning\_rate: 0.206 * model\_\_min\_child\_weight: 1 * model\_\_subsample: 0.941 * model\_\_colsample\_bytree: 0.132 * model\_\_scale\_pos\_weight: 2.124 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 392.0 * average prop of minority class before resampling: 0.385 * average size of training set after resampling: 482.5 * average prop of minority class after resampling: 0.5 | 0.932 | 0.285 | 0.886 | 0.093 | 0.937 | 0.753 | 0.942 | 0.088 |
| XGBClassifier | * undersampling\_majority\_class: True * max\_k\_undersampling: 6 * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_1 * model\_\_max\_depth: 12 * model\_\_learning\_rate: 0.066 * model\_\_min\_child\_weight: 7 * model\_\_subsample: 0.92 * model\_\_colsample\_bytree: 0.325 * model\_\_scale\_pos\_weight: 6.441 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 392.0 * average prop of minority class before resampling: 0.385 * average size of training set after resampling: 482.5 * average prop of minority class after resampling: 0.5 | 0.897 | 0.36 | 0.833 | 0.119 | 0.92 | 0.383 | 0.886 | 0.116 |
| XGBClassifier | * undersampling\_majority\_class: True * max\_k\_undersampling: 6 * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_3 * model\_\_max\_depth: 8 * model\_\_learning\_rate: 0.053 * model\_\_min\_child\_weight: 9 * model\_\_subsample: 0.829 * model\_\_colsample\_bytree: 0.293 * model\_\_scale\_pos\_weight: 0.677 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 392.0 * average prop of minority class before resampling: 0.385 * average size of training set after resampling: 482.5 * average prop of minority class after resampling: 0.5 | 0.952 | 0.228 | 0.917 | 0.073 | 0.955 | 0.7 | 0.947 | 0.07 |
| XGBClassifier | * undersampling\_majority\_class: True * max\_k\_undersampling: 6 * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_3 * model\_\_max\_depth: 14 * model\_\_learning\_rate: 0.127 * model\_\_min\_child\_weight: 8 * model\_\_subsample: 0.893 * model\_\_colsample\_bytree: 0.782 * model\_\_scale\_pos\_weight: 0.185 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 392.0 * average prop of minority class before resampling: 0.385 * average size of training set after resampling: 482.5 * average prop of minority class after resampling: 0.5 | 0.936 | 0.275 | 0.892 | 0.086 | 0.932 | 0.764 | 0.932 | 0.09 |
| XGBClassifier | * undersampling\_majority\_class: False * max\_k\_undersampling: 6 * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_1 * model\_\_max\_depth: 5 * model\_\_learning\_rate: 0.052 * model\_\_min\_child\_weight: 2 * model\_\_subsample: 0.802 * model\_\_colsample\_bytree: 0.399 * model\_\_scale\_pos\_weight: 0.236 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 542.75 * average prop of minority class before resampling: 0.444 * average size of training set after resampling: 603.0 * average prop of minority class after resampling: 0.5 | 0.888 | 0.38 | 0.839 | 0.125 | 0.926 | 0.359 | 0.897 | 0.113 |
| XGBClassifier | * undersampling\_majority\_class: False * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_2 * model\_\_max\_depth: 5 * model\_\_learning\_rate: 0.133 * model\_\_min\_child\_weight: 8 * model\_\_subsample: 0.836 * model\_\_colsample\_bytree: 0.265 * model\_\_scale\_pos\_weight: 0.167 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 542.75 * average prop of minority class before resampling: 0.444 * average size of training set after resampling: 603.0 * average prop of minority class after resampling: 0.5 | 0.939 | 0.246 | 0.904 | 0.081 | 0.947 | 0.731 | 0.935 | 0.079 |
| XGBClassifier | * undersampling\_majority\_class: False * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_1 * model\_\_max\_depth: 9 * model\_\_learning\_rate: 0.07 * model\_\_min\_child\_weight: 8 * model\_\_subsample: 0.869 * model\_\_colsample\_bytree: 0.178 * model\_\_scale\_pos\_weight: 4.598 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 542.75 * average prop of minority class before resampling: 0.444 * average size of training set after resampling: 603.0 * average prop of minority class after resampling: 0.5 | 0.884 | 0.38 | 0.814 | 0.128 | 0.918 | 0.385 | 0.885 | 0.111 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 0 * model\_\_n\_estimators: 88 * model\_\_max\_depth: 9 * model\_\_min\_samples\_split: 7 * model\_\_min\_samples\_leaf: 4 * model\_\_max\_features: auto * model\_\_class\_weight: balanced | * option: build model without resampling * number of folds: 4 * average size of training set: 392.0 * average prop of minority class: 0.385 | 0.95 | 0.236 | 0.917 | 0.077 | 0.967 | 0.257 | 0.962 | 0.076 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 0 * model\_\_n\_estimators: 62 * model\_\_max\_depth: 6 * model\_\_min\_samples\_split: 6 * model\_\_min\_samples\_leaf: 4 * model\_\_max\_features: sqrt * model\_\_class\_weight: balanced | * option: build model without resampling * number of folds: 4 * average size of training set: 542.75 * average prop of minority class: 0.444 | 0.951 | 0.234 | 0.919 | 0.075 | 0.964 | 0.264 | 0.953 | 0.079 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 1 * model\_\_n\_estimators: 17 * model\_\_max\_depth: 8 * model\_\_min\_samples\_split: 20 * model\_\_min\_samples\_leaf: 3 * model\_\_max\_features: auto * model\_\_class\_weight: balanced | * option: build model without resampling * number of folds: 4 * average size of training set: 392.0 * average prop of minority class: 0.385 | 0.949 | 0.236 | 0.911 | 0.076 | 0.964 | 0.255 | 0.958 | 0.078 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 1 * model\_\_n\_estimators: 77 * model\_\_max\_depth: 3 * model\_\_min\_samples\_split: 8 * model\_\_min\_samples\_leaf: 3 * model\_\_max\_features: sqrt * model\_\_class\_weight: balanced\_subsample | * option: build model without resampling * number of folds: 4 * average size of training set: 392.0 * average prop of minority class: 0.385 | 0.945 | 0.243 | 0.906 | 0.076 | 0.959 | 0.267 | 0.957 | 0.081 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 1 * model\_\_n\_estimators: 92 * model\_\_max\_depth: 10 * model\_\_min\_samples\_split: 10 * model\_\_min\_samples\_leaf: 2 * model\_\_max\_features: auto * model\_\_class\_weight: balanced | * option: build model without resampling * number of folds: 4 * average size of training set: 542.75 * average prop of minority class: 0.444 | 0.949 | 0.242 | 0.915 | 0.078 | 0.97 | 0.253 | 0.965 | 0.075 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 0 * model\_\_n\_estimators: 100 * model\_\_max\_depth: 10 * model\_\_min\_samples\_split: 6 * model\_\_min\_samples\_leaf: 2 * model\_\_max\_features: auto * model\_\_class\_weight: balanced | * option: build model without resampling * number of folds: 4 * average size of training set: 392.0 * average prop of minority class: 0.385 | 0.949 | 0.238 | 0.913 | 0.077 | 0.973 | 0.247 | 0.972 | 0.076 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 0 * model\_\_n\_estimators: 65 * model\_\_max\_depth: 4 * model\_\_min\_samples\_split: 8 * model\_\_min\_samples\_leaf: 1 * model\_\_max\_features: sqrt * model\_\_class\_weight: balanced\_subsample | * option: build model without resampling * number of folds: 4 * average size of training set: 392.0 * average prop of minority class: 0.385 | 0.95 | 0.225 | 0.91 | 0.073 | 0.959 | 0.269 | 0.95 | 0.083 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 1 * model\_\_n\_estimators: 51 * model\_\_max\_depth: 2 * model\_\_min\_samples\_split: 6 * model\_\_min\_samples\_leaf: 1 * model\_\_max\_features: auto * model\_\_class\_weight: balanced | * option: build model without resampling * number of folds: 4 * average size of training set: 392.0 * average prop of minority class: 0.385 | 0.945 | 0.247 | 0.906 | 0.077 | 0.942 | 0.738 | 0.944 | 0.084 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 1 * model\_\_n\_estimators: 65 * model\_\_max\_depth: 7 * model\_\_min\_samples\_split: 17 * model\_\_min\_samples\_leaf: 1 * model\_\_max\_features: sqrt * model\_\_class\_weight: balanced\_subsample | * option: build model without resampling * number of folds: 4 * average size of training set: 542.75 * average prop of minority class: 0.444 | 0.953 | 0.229 | 0.919 | 0.074 | 0.973 | 0.235 | 0.967 | 0.076 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 0 * model\_\_n\_estimators: 40 * model\_\_max\_depth: 9 * model\_\_min\_samples\_split: 9 * model\_\_min\_samples\_leaf: 3 * model\_\_max\_features: auto * model\_\_class\_weight: balanced | * option: build model without resampling * number of folds: 4 * average size of training set: 392.0 * average prop of minority class: 0.385 | 0.95 | 0.231 | 0.918 | 0.076 | 0.968 | 0.26 | 0.965 | 0.079 |

## Report of inner training in fold 9 of outer Cross Validation

## Report of training in this outer fold

Best model with respect to selected metric is XGBClassifier with the following params:

* undersampling\_majority\_class: False
* max\_k\_undersampling: 6
* resample\_\_sampling\_strategy: minority
* post\_process\_\_option: option\_3
* model\_\_max\_depth: 7
* model\_\_learning\_rate: 0.063
* model\_\_min\_child\_weight: 10
* model\_\_subsample: 0.974
* model\_\_colsample\_bytree: 0.399
* model\_\_scale\_pos\_weight: 0.437
* model\_\_objective: binary:logistic

### Comparison of all models trained in this outer fold

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Model** | **Params** | **Comments** | **roc\_auc on inner fold** | **neg\_log\_loss on inner fold** | **average\_precision on inner fold** | **neg\_brier\_score on inner fold** | **roc\_auc on outer fold** | **neg\_log\_loss on outer fold** | **average\_precision on outer fold** | **neg\_brier\_score on outer fold** |
| XGBClassifier | * undersampling\_majority\_class: False * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_2 * model\_\_max\_depth: 15 * model\_\_learning\_rate: 0.198 * model\_\_min\_child\_weight: 3 * model\_\_subsample: 0.988 * model\_\_colsample\_bytree: 0.188 * model\_\_scale\_pos\_weight: 7.234 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 542.75 * average prop of minority class before resampling: 0.444 * average size of training set after resampling: 603.0 * average prop of minority class after resampling: 0.5 | 0.934 | 0.281 | 0.893 | 0.094 | 0.934 | 0.33 | 0.926 | 0.107 |
| XGBClassifier | * undersampling\_majority\_class: False * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_2 * model\_\_max\_depth: 11 * model\_\_learning\_rate: 0.127 * model\_\_min\_child\_weight: 10 * model\_\_subsample: 0.896 * model\_\_colsample\_bytree: 0.5 * model\_\_scale\_pos\_weight: 0.174 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 542.75 * average prop of minority class before resampling: 0.444 * average size of training set after resampling: 603.0 * average prop of minority class after resampling: 0.5 | 0.941 | 0.257 | 0.905 | 0.085 | 0.932 | 0.316 | 0.919 | 0.096 |
| XGBClassifier | * undersampling\_majority\_class: True * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_1 * model\_\_max\_depth: 14 * model\_\_learning\_rate: 0.188 * model\_\_min\_child\_weight: 5 * model\_\_subsample: 0.869 * model\_\_colsample\_bytree: 0.368 * model\_\_scale\_pos\_weight: 0.234 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 392.0 * average prop of minority class before resampling: 0.385 * average size of training set after resampling: 482.5 * average prop of minority class after resampling: 0.5 | 0.938 | 0.281 | 0.906 | 0.091 | 0.936 | 0.331 | 0.933 | 0.102 |
| XGBClassifier | * undersampling\_majority\_class: False * max\_k\_undersampling: 6 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_3 * model\_\_max\_depth: 7 * model\_\_learning\_rate: 0.063 * model\_\_min\_child\_weight: 10 * model\_\_subsample: 0.974 * model\_\_colsample\_bytree: 0.399 * model\_\_scale\_pos\_weight: 0.437 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 542.75 * average prop of minority class before resampling: 0.444 * average size of training set after resampling: 603.0 * average prop of minority class after resampling: 0.5 | 0.954 | 0.229 | 0.919 | 0.073 | 0.961 | 0.26 | 0.954 | 0.079 |
| XGBClassifier | * undersampling\_majority\_class: False * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_3 * model\_\_max\_depth: 5 * model\_\_learning\_rate: 0.309 * model\_\_min\_child\_weight: 2 * model\_\_subsample: 0.894 * model\_\_colsample\_bytree: 0.175 * model\_\_scale\_pos\_weight: 0.16 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 542.75 * average prop of minority class before resampling: 0.444 * average size of training set after resampling: 603.0 * average prop of minority class after resampling: 0.5 | 0.949 | 0.242 | 0.912 | 0.078 | 0.963 | 0.26 | 0.963 | 0.079 |
| XGBClassifier | * undersampling\_majority\_class: True * max\_k\_undersampling: 6 * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_2 * model\_\_max\_depth: 5 * model\_\_learning\_rate: 0.28 * model\_\_min\_child\_weight: 1 * model\_\_subsample: 0.835 * model\_\_colsample\_bytree: 0.72 * model\_\_scale\_pos\_weight: 5.463 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 392.0 * average prop of minority class before resampling: 0.385 * average size of training set after resampling: 482.5 * average prop of minority class after resampling: 0.5 | 0.934 | 0.276 | 0.891 | 0.091 | 0.913 | 0.364 | 0.892 | 0.119 |
| XGBClassifier | * undersampling\_majority\_class: True * max\_k\_undersampling: 6 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_2 * model\_\_max\_depth: 12 * model\_\_learning\_rate: 0.075 * model\_\_min\_child\_weight: 10 * model\_\_subsample: 0.976 * model\_\_colsample\_bytree: 0.785 * model\_\_scale\_pos\_weight: 0.352 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 392.0 * average prop of minority class before resampling: 0.385 * average size of training set after resampling: 482.5 * average prop of minority class after resampling: 0.5 | 0.944 | 0.255 | 0.9 | 0.08 | 0.949 | 0.27 | 0.949 | 0.079 |
| XGBClassifier | * undersampling\_majority\_class: False * max\_k\_undersampling: 6 * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_3 * model\_\_max\_depth: 10 * model\_\_learning\_rate: 0.288 * model\_\_min\_child\_weight: 7 * model\_\_subsample: 0.913 * model\_\_colsample\_bytree: 0.316 * model\_\_scale\_pos\_weight: 8.795 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 542.75 * average prop of minority class before resampling: 0.444 * average size of training set after resampling: 603.0 * average prop of minority class after resampling: 0.5 | 0.946 | 0.234 | 0.917 | 0.075 | 0.964 | 0.242 | 0.967 | 0.066 |
| XGBClassifier | * undersampling\_majority\_class: True * max\_k\_undersampling: 6 * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_1 * model\_\_max\_depth: 8 * model\_\_learning\_rate: 0.152 * model\_\_min\_child\_weight: 5 * model\_\_subsample: 0.908 * model\_\_colsample\_bytree: 0.383 * model\_\_scale\_pos\_weight: 2.752 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 392.0 * average prop of minority class before resampling: 0.385 * average size of training set after resampling: 482.5 * average prop of minority class after resampling: 0.5 | 0.932 | 0.299 | 0.897 | 0.096 | 0.918 | 0.381 | 0.927 | 0.113 |
| XGBClassifier | * undersampling\_majority\_class: True * max\_k\_undersampling: 6 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_1 * model\_\_max\_depth: 5 * model\_\_learning\_rate: 0.184 * model\_\_min\_child\_weight: 7 * model\_\_subsample: 0.835 * model\_\_colsample\_bytree: 0.393 * model\_\_scale\_pos\_weight: 1.152 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 392.0 * average prop of minority class before resampling: 0.385 * average size of training set after resampling: 482.5 * average prop of minority class after resampling: 0.5 | 0.942 | 0.265 | 0.914 | 0.086 | 0.921 | 0.358 | 0.919 | 0.115 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 1 * model\_\_n\_estimators: 30 * model\_\_max\_depth: 9 * model\_\_min\_samples\_split: 5 * model\_\_min\_samples\_leaf: 1 * model\_\_max\_features: auto * model\_\_class\_weight: balanced\_subsample | * option: build model without resampling * number of folds: 4 * average size of training set: 542.75 * average prop of minority class: 0.444 | 0.953 | 0.236 | 0.915 | 0.074 | 0.96 | 0.251 | 0.956 | 0.065 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 0 * model\_\_n\_estimators: 99 * model\_\_max\_depth: 8 * model\_\_min\_samples\_split: 19 * model\_\_min\_samples\_leaf: 2 * model\_\_max\_features: sqrt * model\_\_class\_weight: balanced\_subsample | * option: build model without resampling * number of folds: 4 * average size of training set: 392.0 * average prop of minority class: 0.385 | 0.955 | 0.214 | 0.916 | 0.068 | 0.96 | 0.254 | 0.932 | 0.07 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 0 * model\_\_n\_estimators: 65 * model\_\_max\_depth: 2 * model\_\_min\_samples\_split: 15 * model\_\_min\_samples\_leaf: 2 * model\_\_max\_features: sqrt * model\_\_class\_weight: balanced | * option: build model without resampling * number of folds: 4 * average size of training set: 392.0 * average prop of minority class: 0.385 | 0.948 | 0.24 | 0.906 | 0.076 | 0.946 | 0.306 | 0.945 | 0.09 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 1 * model\_\_n\_estimators: 25 * model\_\_max\_depth: 4 * model\_\_min\_samples\_split: 18 * model\_\_min\_samples\_leaf: 2 * model\_\_max\_features: auto * model\_\_class\_weight: balanced\_subsample | * option: build model without resampling * number of folds: 4 * average size of training set: 392.0 * average prop of minority class: 0.385 | 0.951 | 0.228 | 0.915 | 0.072 | 0.965 | 0.254 | 0.945 | 0.075 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 1 * model\_\_n\_estimators: 18 * model\_\_max\_depth: 9 * model\_\_min\_samples\_split: 8 * model\_\_min\_samples\_leaf: 2 * model\_\_max\_features: auto * model\_\_class\_weight: balanced | * option: build model without resampling * number of folds: 4 * average size of training set: 542.75 * average prop of minority class: 0.444 | 0.947 | 0.255 | 0.912 | 0.083 | 0.965 | 0.252 | 0.96 | 0.07 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 1 * model\_\_n\_estimators: 80 * model\_\_max\_depth: 10 * model\_\_min\_samples\_split: 6 * model\_\_min\_samples\_leaf: 4 * model\_\_max\_features: sqrt * model\_\_class\_weight: balanced\_subsample | * option: build model without resampling * number of folds: 4 * average size of training set: 392.0 * average prop of minority class: 0.385 | 0.954 | 0.223 | 0.918 | 0.071 | 0.962 | 0.258 | 0.958 | 0.072 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 1 * model\_\_n\_estimators: 90 * model\_\_max\_depth: 4 * model\_\_min\_samples\_split: 5 * model\_\_min\_samples\_leaf: 2 * model\_\_max\_features: sqrt * model\_\_class\_weight: balanced\_subsample | * option: build model without resampling * number of folds: 4 * average size of training set: 392.0 * average prop of minority class: 0.385 | 0.952 | 0.22 | 0.915 | 0.07 | 0.969 | 0.241 | 0.958 | 0.072 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 1 * model\_\_n\_estimators: 88 * model\_\_max\_depth: 2 * model\_\_min\_samples\_split: 7 * model\_\_min\_samples\_leaf: 4 * model\_\_max\_features: auto * model\_\_class\_weight: balanced | * option: build model without resampling * number of folds: 4 * average size of training set: 392.0 * average prop of minority class: 0.385 | 0.946 | 0.249 | 0.908 | 0.078 | 0.945 | 0.284 | 0.944 | 0.084 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 0 * model\_\_n\_estimators: 69 * model\_\_max\_depth: 10 * model\_\_min\_samples\_split: 8 * model\_\_min\_samples\_leaf: 2 * model\_\_max\_features: auto * model\_\_class\_weight: balanced | * option: build model without resampling * number of folds: 4 * average size of training set: 392.0 * average prop of minority class: 0.385 | 0.952 | 0.226 | 0.913 | 0.073 | 0.964 | 0.253 | 0.953 | 0.069 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 0 * model\_\_n\_estimators: 11 * model\_\_max\_depth: 10 * model\_\_min\_samples\_split: 7 * model\_\_min\_samples\_leaf: 1 * model\_\_max\_features: sqrt * model\_\_class\_weight: balanced\_subsample | * option: build model without resampling * number of folds: 4 * average size of training set: 542.75 * average prop of minority class: 0.444 | 0.943 | 0.267 | 0.898 | 0.086 | 0.973 | 0.231 | 0.96 | 0.066 |