Report of training

# Introduction

## 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] 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 gp\_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.utils.pipes\_and\_transformers.OptionedPostProcessTransformer object at 0x0000019C81E0E7C0>),  
   ('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

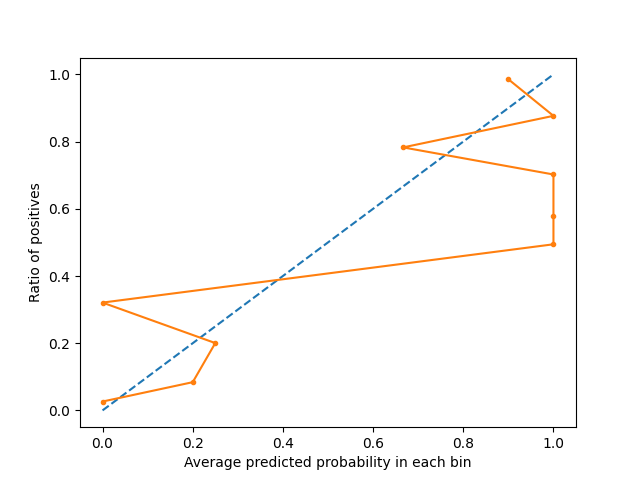
## Winner models of each fold and main metrics

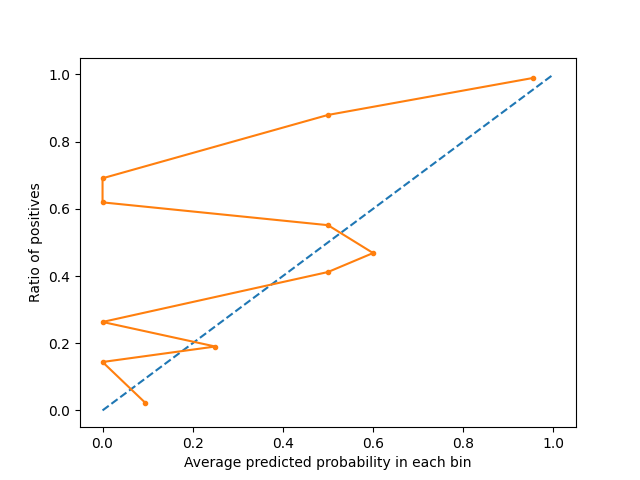
|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Fold** | **Model** | **Params** | **Comments** | **roc\_auc** | **neg\_log\_loss** | **average\_precision** | **neg\_brier\_score** |
| 1 | XGBClassifier | * undersampling\_majority\_class: True * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_3 * model\_\_max\_depth: 7 * model\_\_learning\_rate: 0.103 * model\_\_min\_child\_weight: 1 * model\_\_subsample: 0.986 * model\_\_colsample\_bytree: 0.557 * model\_\_scale\_pos\_weight: 0.596 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 5 * 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: 484.4 * average prop of minority class after resampling: 0.5 | 0.968 | 0.242 | 0.956 | 0.072 |
| 3 | XGBClassifier | * undersampling\_majority\_class: True * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_3 * model\_\_max\_depth: 10 * model\_\_learning\_rate: 0.058 * model\_\_min\_child\_weight: 4 * model\_\_subsample: 0.94 * model\_\_colsample\_bytree: 0.38 * model\_\_scale\_pos\_weight: 1.667 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 5 * 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: 482.4 * average prop of minority class after resampling: 0.5 | 0.917 | 0.373 | 0.921 | 0.12 |
| 5 | RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 0 * model\_\_n\_estimators: 91 * model\_\_max\_depth: 6 * model\_\_min\_samples\_split: 15 * model\_\_min\_samples\_leaf: 3 * model\_\_max\_features: sqrt * model\_\_class\_weight: balanced | * option: build model without resampling * number of folds: 5 * average size of training set: 542.8 * average prop of minority class: 0.444 | 0.948 | 0.745 | 0.904 | 0.086 |
| 7 | XGBClassifier | * undersampling\_majority\_class: True * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_3 * model\_\_max\_depth: 14 * model\_\_learning\_rate: 0.101 * model\_\_min\_child\_weight: 2 * model\_\_subsample: 0.874 * model\_\_colsample\_bytree: 0.534 * model\_\_scale\_pos\_weight: 6.33 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 5 * 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: 482.4 * average prop of minority class after resampling: 0.5 | 0.969 | 0.241 | 0.966 | 0.072 |
| 9 | RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 1 * model\_\_n\_estimators: 50 * model\_\_max\_depth: 7 * model\_\_min\_samples\_split: 11 * model\_\_min\_samples\_leaf: 3 * model\_\_max\_features: auto * model\_\_class\_weight: balanced | * option: build model without resampling * number of folds: 5 * average size of training set: 392.0 * average prop of minority class: 0.385 | 0.978 | 0.242 | 0.969 | 0.072 |

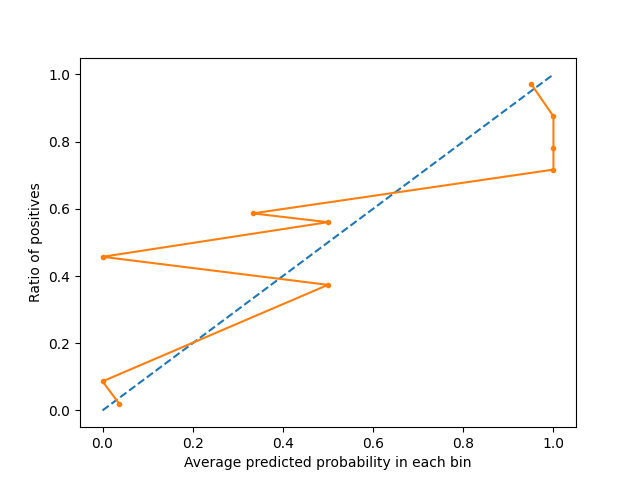
For the selected optimization metric average\_precision the average score is 0.943.

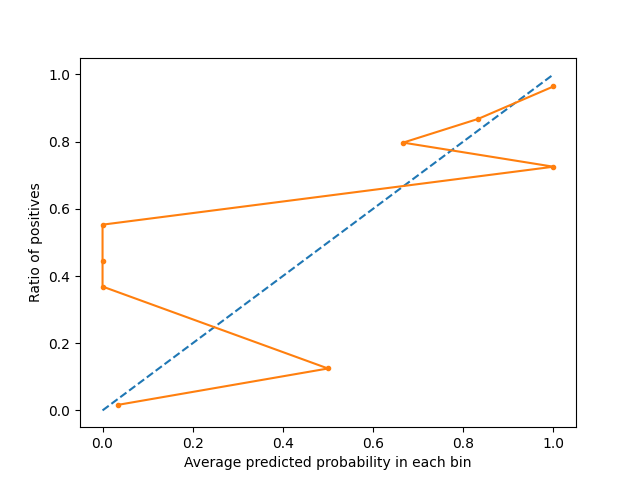
## Main plots

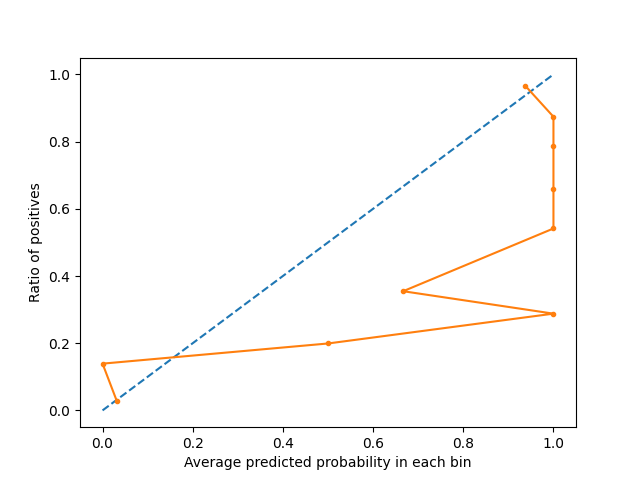
### Calibration plots



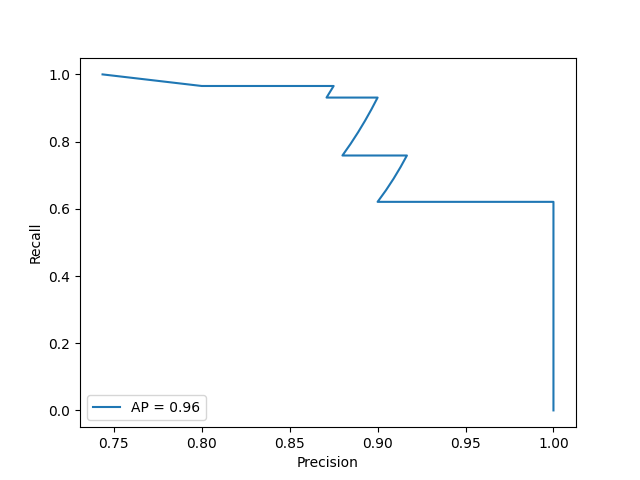


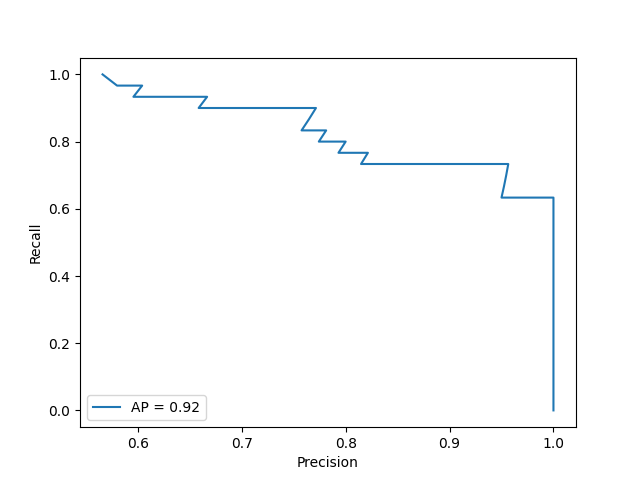


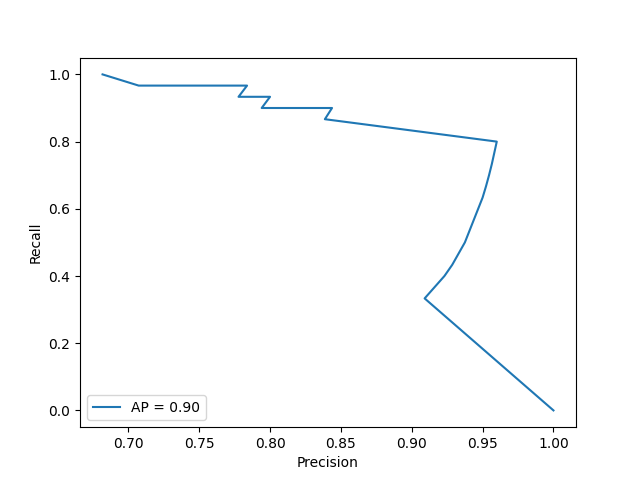


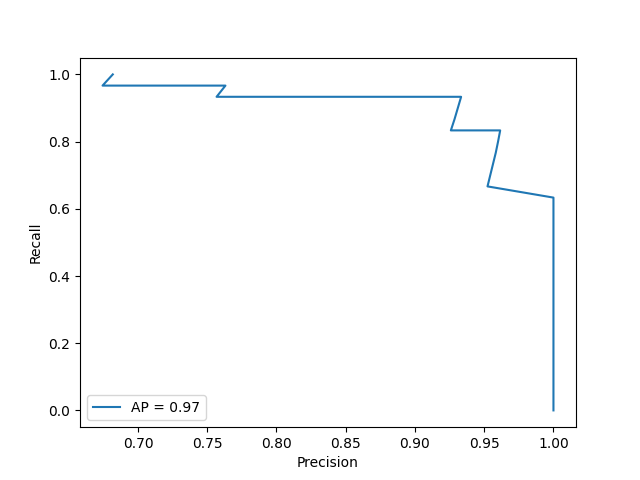


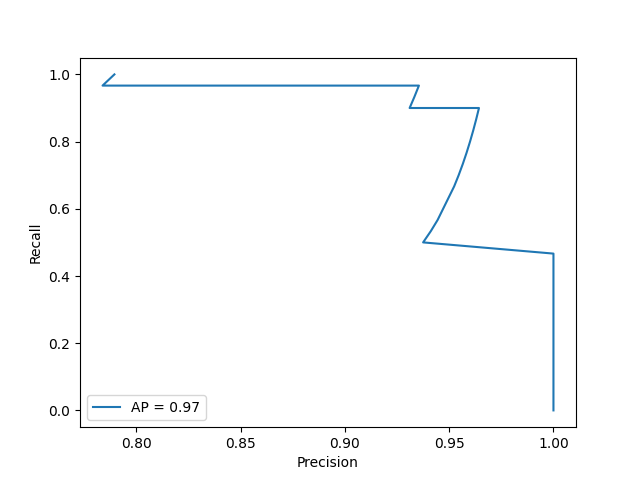
### Precision-recall curve plots



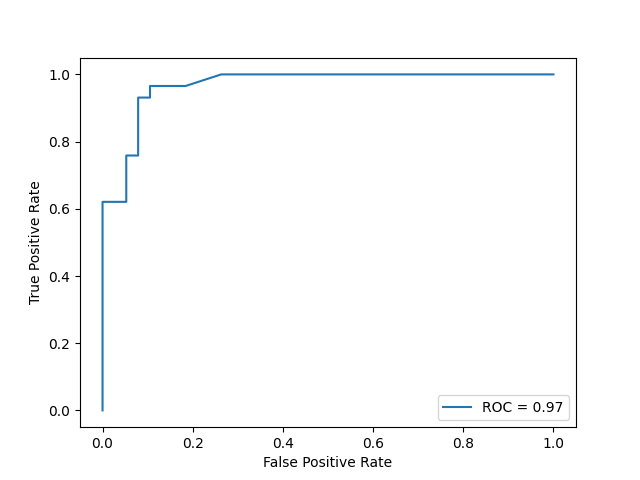


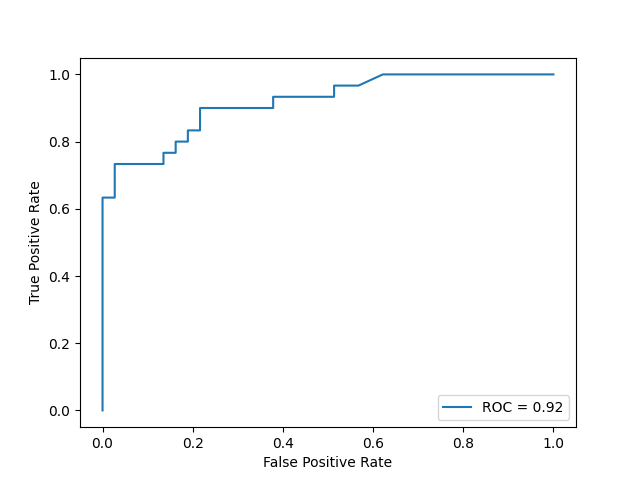


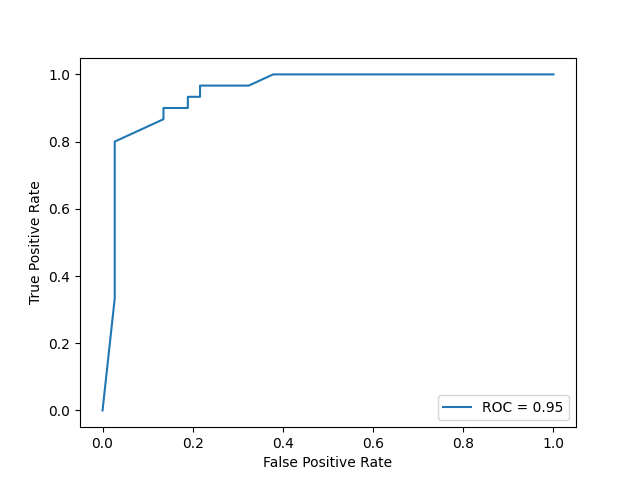


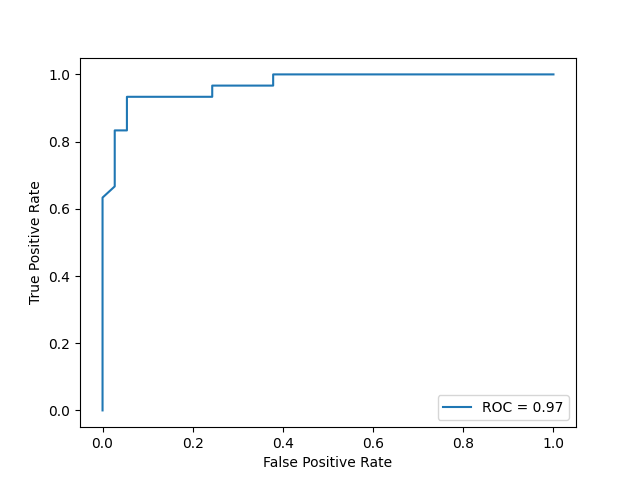


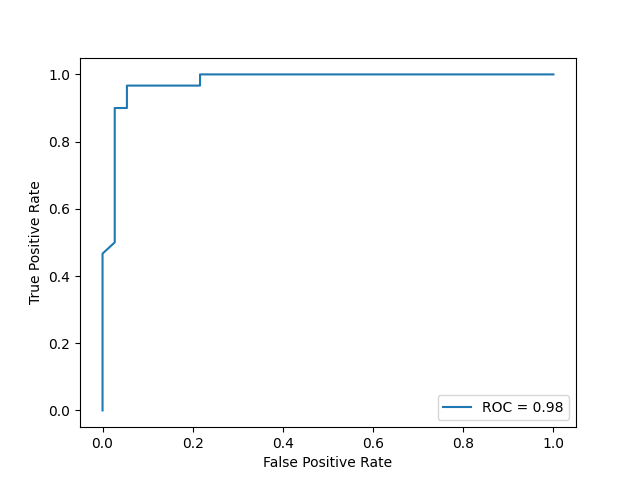
### ROC curve plots



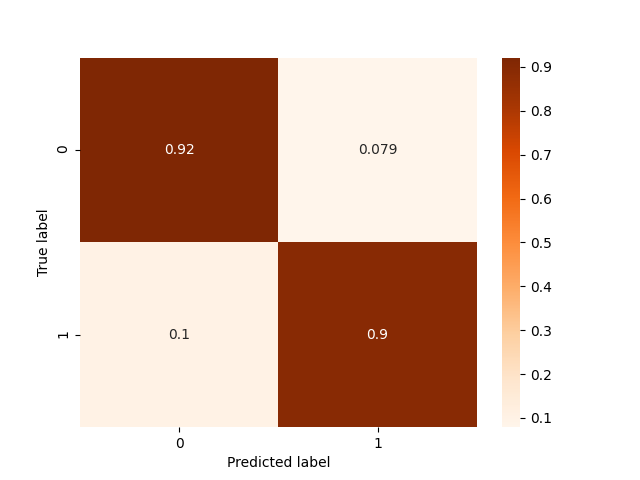


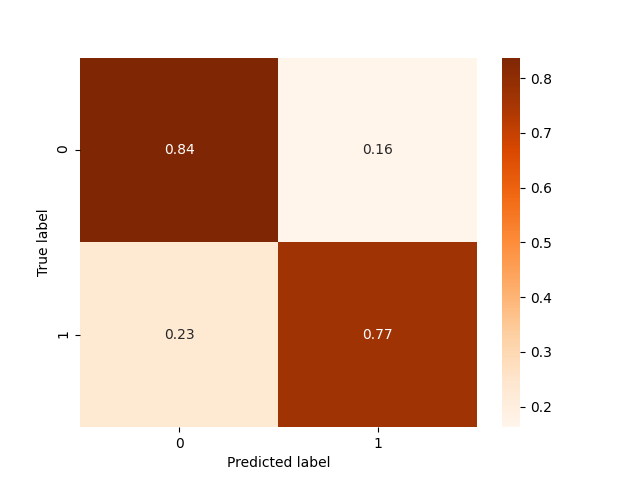


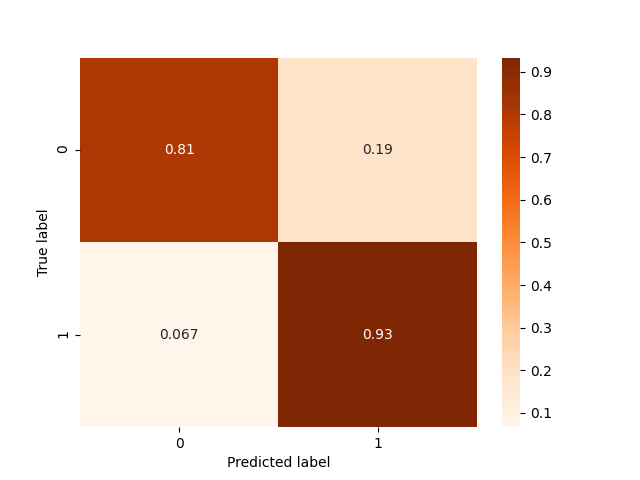


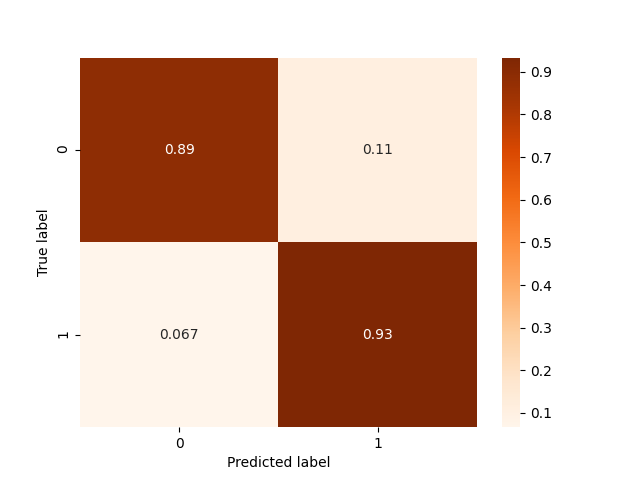


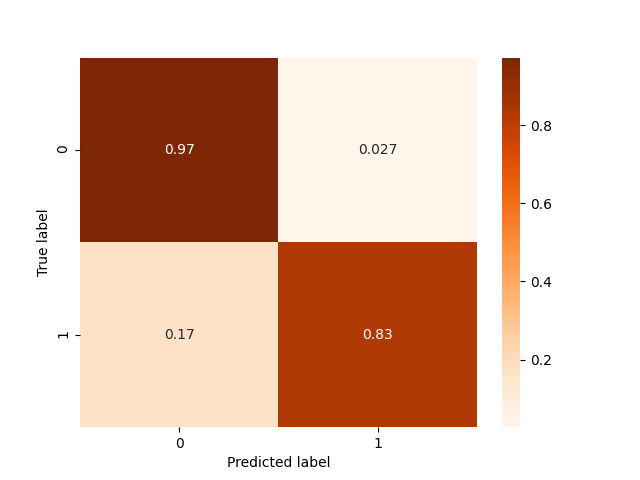
### Confusion matrix



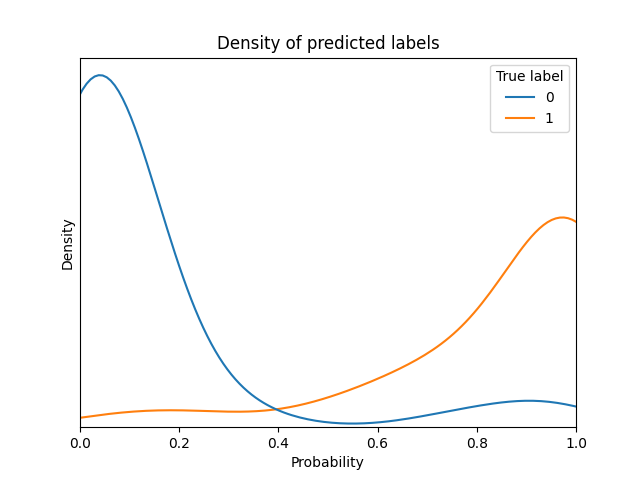


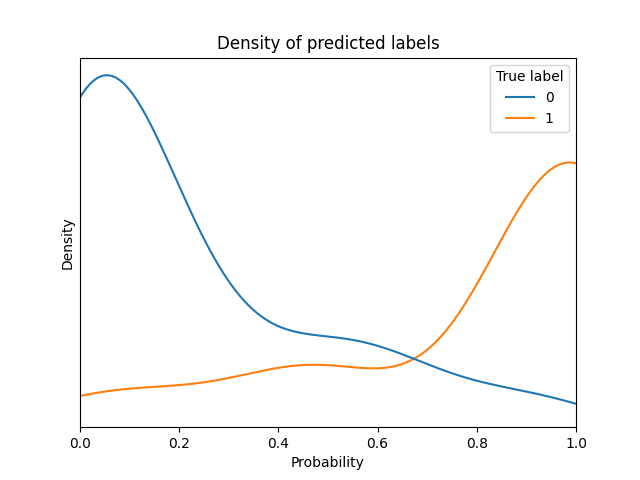


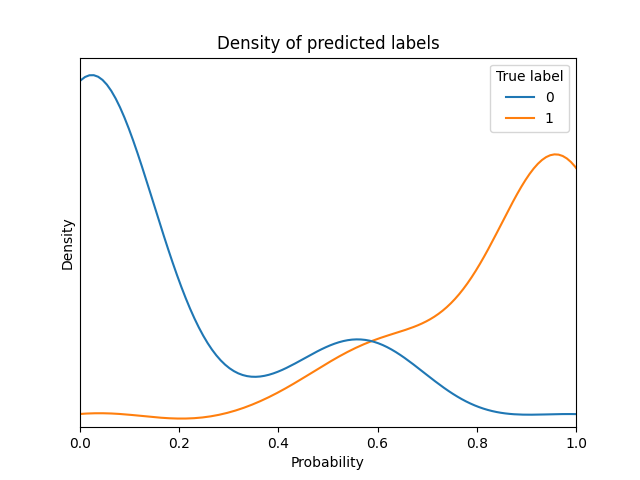


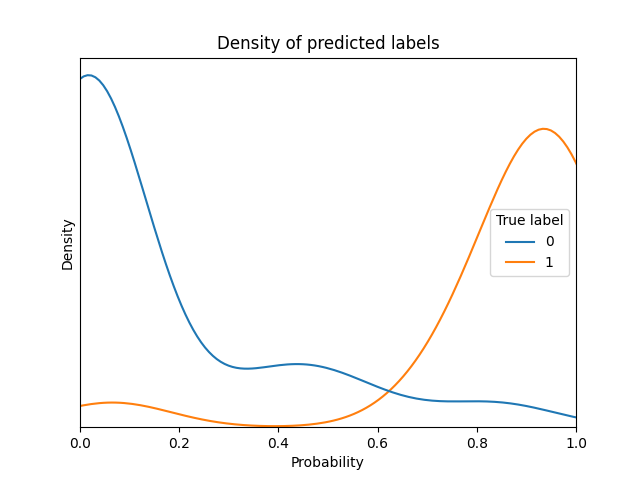


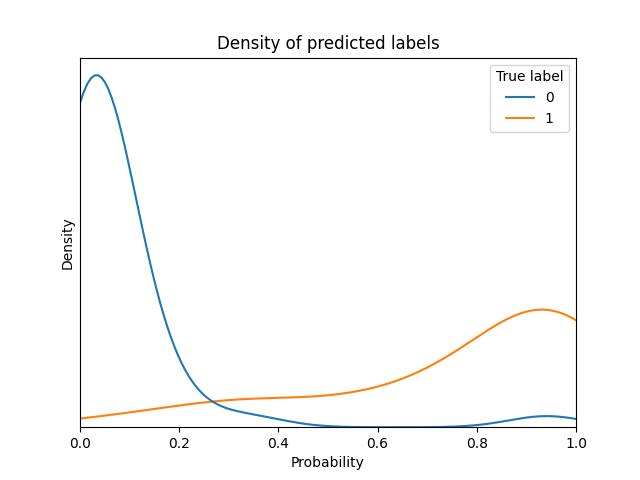
### Histograms











## 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 | 0.02 | 0.01 | 0.029 | 0.044 | 0.031 | 0.012 |
| 2 | 1 | 1.0 | 1.0 | 1.0 | 0.958 | 0.991 | 0.016 |
| 3 | 1 | 0.179 | 0.105 | 0.036 | 0.069 | 0.034 | 0.054 |
| 4 | 0 | 0.052 | 0.01 | 0.151 | 0.013 | 0.031 | 0.052 |
| 5 | 0 | 0.077 | 0.228 | 0.167 | 0.046 | 0.348 | 0.109 |
| 6 | 0 | 0.084 | 0.079 | 0.186 | 0.08 | 0.143 | 0.043 |
| 7 | 1 | 1.0 | 1.0 | 0.943 | 0.958 | 0.958 | 0.024 |
| 8 | 1 | 1.0 | 1.0 | 1.0 | 0.99 | 1.0 | 0.004 |
| 9 | 1 | 1.0 | 1.0 | 1.0 | 0.987 | 0.958 | 0.016 |
| 10 | 0 | 0.184 | 0.125 | 0.029 | 0.078 | 0.038 | 0.058 |
| 11 | 0 | 0.052 | 0.097 | 0.029 | 0.053 | 0.011 | 0.029 |
| 12 | 1 | 1.0 | 1.0 | 1.0 | 0.99 | 0.991 | 0.005 |
| 13 | 1 | 1.0 | 1.0 | 1.0 | 0.99 | 0.958 | 0.016 |
| 14 | 0 | 0.152 | 0.192 | 0.12 | 0.069 | 0.143 | 0.041 |
| 15 | 0 | 0.624 | 0.467 | 0.569 | 0.513 | 0.443 | 0.066 |
| 16 | 0 | 0.084 | 0.109 | 0.291 | 0.079 | 0.143 | 0.078 |
| 17 | 0 | 0.084 | 0.155 | 0.253 | 0.077 | 0.127 | 0.064 |
| 18 | 1 | 0.74 | 0.419 | 0.586 | 0.63 | 0.658 | 0.106 |
| 19 | 0 | 0.532 | 0.83 | 0.888 | 0.639 | 0.941 | 0.155 |
| 20 | 1 | 1.0 | 1.0 | 1.0 | 0.99 | 0.991 | 0.005 |

The average standard deviation is 0.048

# 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 XGBClassifier with the following params:

* undersampling\_majority\_class: True
* max\_k\_undersampling: 5
* resample\_\_sampling\_strategy: minority
* post\_process\_\_option: option\_3
* model\_\_max\_depth: 7
* model\_\_learning\_rate: 0.103
* model\_\_min\_child\_weight: 1
* model\_\_subsample: 0.986
* model\_\_colsample\_bytree: 0.557
* model\_\_scale\_pos\_weight: 0.596
* 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: 6 * model\_\_learning\_rate: 0.073 * model\_\_min\_child\_weight: 7 * model\_\_subsample: 0.996 * model\_\_colsample\_bytree: 0.485 * model\_\_scale\_pos\_weight: 0.275 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 5 * 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: 484.4 * average prop of minority class after resampling: 0.5 | 0.955 | 0.217 | 0.939 | 0.068 | 0.955 | 0.309 | 0.943 | 0.079 |
| XGBClassifier | * undersampling\_majority\_class: True * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_2 * model\_\_max\_depth: 14 * model\_\_learning\_rate: 0.207 * model\_\_min\_child\_weight: 8 * model\_\_subsample: 0.812 * model\_\_colsample\_bytree: 0.596 * model\_\_scale\_pos\_weight: 0.173 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 5 * 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: 484.4 * average prop of minority class after resampling: 0.5 | 0.946 | 0.253 | 0.93 | 0.082 | 0.942 | 0.296 | 0.939 | 0.089 |
| XGBClassifier | * undersampling\_majority\_class: True * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_1 * model\_\_max\_depth: 10 * model\_\_learning\_rate: 0.122 * model\_\_min\_child\_weight: 5 * model\_\_subsample: 0.96 * model\_\_colsample\_bytree: 0.228 * model\_\_scale\_pos\_weight: 1.211 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 5 * 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: 484.4 * average prop of minority class after resampling: 0.5 | 0.919 | 0.32 | 0.883 | 0.105 | 0.914 | 0.347 | 0.9 | 0.1 |
| XGBClassifier | * undersampling\_majority\_class: False * max\_k\_undersampling: 6 * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_1 * model\_\_max\_depth: 8 * model\_\_learning\_rate: 0.06 * model\_\_min\_child\_weight: 8 * model\_\_subsample: 0.934 * model\_\_colsample\_bytree: 0.216 * model\_\_scale\_pos\_weight: 4.197 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 5 * average size of training set before resampling: 542.8 * average prop of minority class before resampling: 0.446 * average size of training set after resampling: 601.2 * average prop of minority class after resampling: 0.5 | 0.92 | 0.319 | 0.877 | 0.105 | 0.89 | 0.861 | 0.891 | 0.12 |
| 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: 0.06 * model\_\_min\_child\_weight: 1 * model\_\_subsample: 0.871 * model\_\_colsample\_bytree: 0.138 * model\_\_scale\_pos\_weight: 0.968 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 5 * 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: 484.4 * average prop of minority class after resampling: 0.5 | 0.949 | 0.242 | 0.931 | 0.077 | 0.952 | 0.283 | 0.948 | 0.083 |
| XGBClassifier | * undersampling\_majority\_class: True * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_3 * model\_\_max\_depth: 11 * model\_\_learning\_rate: 0.118 * model\_\_min\_child\_weight: 7 * model\_\_subsample: 0.909 * model\_\_colsample\_bytree: 0.164 * model\_\_scale\_pos\_weight: 7.145 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 5 * 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: 484.4 * average prop of minority class after resampling: 0.5 | 0.963 | 0.203 | 0.948 | 0.064 | 0.958 | 0.723 | 0.918 | 0.078 |
| XGBClassifier | * undersampling\_majority\_class: True * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_3 * model\_\_max\_depth: 7 * model\_\_learning\_rate: 0.103 * model\_\_min\_child\_weight: 1 * model\_\_subsample: 0.986 * model\_\_colsample\_bytree: 0.557 * model\_\_scale\_pos\_weight: 0.596 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 5 * 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: 484.4 * average prop of minority class after resampling: 0.5 | 0.967 | 0.176 | 0.953 | 0.056 | 0.968 | 0.242 | 0.956 | 0.072 |
| XGBClassifier | * undersampling\_majority\_class: False * max\_k\_undersampling: 6 * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_2 * model\_\_max\_depth: 14 * model\_\_learning\_rate: 0.306 * model\_\_min\_child\_weight: 8 * model\_\_subsample: 0.823 * model\_\_colsample\_bytree: 0.273 * model\_\_scale\_pos\_weight: 0.376 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 5 * average size of training set before resampling: 542.8 * average prop of minority class before resampling: 0.446 * average size of training set after resampling: 601.2 * average prop of minority class after resampling: 0.5 | 0.949 | 0.234 | 0.93 | 0.075 | 0.946 | 0.293 | 0.944 | 0.09 |
| XGBClassifier | * undersampling\_majority\_class: True * max\_k\_undersampling: 6 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_2 * model\_\_max\_depth: 7 * model\_\_learning\_rate: 0.118 * model\_\_min\_child\_weight: 6 * model\_\_subsample: 0.838 * model\_\_colsample\_bytree: 0.517 * model\_\_scale\_pos\_weight: 0.691 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 5 * 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: 484.4 * average prop of minority class after resampling: 0.5 | 0.952 | 0.234 | 0.933 | 0.077 | 0.963 | 0.267 | 0.957 | 0.082 |
| XGBClassifier | * undersampling\_majority\_class: False * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_2 * model\_\_max\_depth: 6 * model\_\_learning\_rate: 0.055 * model\_\_min\_child\_weight: 3 * model\_\_subsample: 0.897 * model\_\_colsample\_bytree: 0.293 * model\_\_scale\_pos\_weight: 5.263 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 5 * average size of training set before resampling: 542.8 * average prop of minority class before resampling: 0.446 * average size of training set after resampling: 601.2 * average prop of minority class after resampling: 0.5 | 0.941 | 0.258 | 0.923 | 0.084 | 0.937 | 0.77 | 0.898 | 0.087 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 1 * model\_\_n\_estimators: 87 * model\_\_max\_depth: 8 * model\_\_min\_samples\_split: 19 * model\_\_min\_samples\_leaf: 2 * model\_\_max\_features: auto * model\_\_class\_weight: balanced\_subsample | * option: build model without resampling * number of folds: 5 * average size of training set: 392.5 * average prop of minority class: 0.383 | 0.963 | 0.196 | 0.944 | 0.064 | 0.953 | 0.715 | 0.913 | 0.075 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 0 * model\_\_n\_estimators: 88 * model\_\_max\_depth: 4 * model\_\_min\_samples\_split: 20 * model\_\_min\_samples\_leaf: 3 * model\_\_max\_features: auto * model\_\_class\_weight: balanced\_subsample | * option: build model without resampling * number of folds: 5 * average size of training set: 542.8 * average prop of minority class: 0.446 | 0.962 | 0.205 | 0.944 | 0.067 | 0.954 | 0.717 | 0.912 | 0.074 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 1 * model\_\_n\_estimators: 83 * model\_\_max\_depth: 5 * model\_\_min\_samples\_split: 18 * model\_\_min\_samples\_leaf: 3 * model\_\_max\_features: sqrt * model\_\_class\_weight: balanced\_subsample | * option: build model without resampling * number of folds: 5 * average size of training set: 542.8 * average prop of minority class: 0.446 | 0.965 | 0.194 | 0.946 | 0.063 | 0.958 | 0.707 | 0.915 | 0.073 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 0 * model\_\_n\_estimators: 61 * model\_\_max\_depth: 5 * model\_\_min\_samples\_split: 9 * model\_\_min\_samples\_leaf: 1 * model\_\_max\_features: auto * model\_\_class\_weight: balanced | * option: build model without resampling * number of folds: 5 * average size of training set: 542.8 * average prop of minority class: 0.446 | 0.964 | 0.194 | 0.943 | 0.064 | 0.954 | 0.724 | 0.915 | 0.078 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 0 * model\_\_n\_estimators: 18 * 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: 5 * average size of training set: 542.8 * average prop of minority class: 0.446 | 0.962 | 0.208 | 0.946 | 0.065 | 0.964 | 0.268 | 0.951 | 0.082 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 0 * model\_\_n\_estimators: 66 * model\_\_max\_depth: 2 * model\_\_min\_samples\_split: 11 * model\_\_min\_samples\_leaf: 2 * model\_\_max\_features: auto * model\_\_class\_weight: balanced | * option: build model without resampling * number of folds: 5 * average size of training set: 542.8 * average prop of minority class: 0.446 | 0.956 | 0.219 | 0.938 | 0.07 | 0.926 | 0.767 | 0.903 | 0.083 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 1 * model\_\_n\_estimators: 18 * model\_\_max\_depth: 5 * model\_\_min\_samples\_split: 11 * model\_\_min\_samples\_leaf: 1 * model\_\_max\_features: auto * model\_\_class\_weight: balanced\_subsample | * option: build model without resampling * number of folds: 5 * average size of training set: 542.8 * average prop of minority class: 0.446 | 0.958 | 0.221 | 0.935 | 0.071 | 0.952 | 0.733 | 0.902 | 0.079 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 1 * model\_\_n\_estimators: 13 * 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: 5 * average size of training set: 392.5 * average prop of minority class: 0.383 | 0.965 | 0.194 | 0.95 | 0.062 | 0.953 | 0.282 | 0.946 | 0.083 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 0 * model\_\_n\_estimators: 83 * model\_\_max\_depth: 9 * model\_\_min\_samples\_split: 7 * model\_\_min\_samples\_leaf: 3 * model\_\_max\_features: auto * model\_\_class\_weight: balanced | * option: build model without resampling * number of folds: 5 * average size of training set: 392.5 * average prop of minority class: 0.383 | 0.965 | 0.186 | 0.948 | 0.06 | 0.961 | 0.268 | 0.945 | 0.077 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 1 * model\_\_n\_estimators: 57 * model\_\_max\_depth: 2 * model\_\_min\_samples\_split: 16 * model\_\_min\_samples\_leaf: 3 * model\_\_max\_features: sqrt * model\_\_class\_weight: balanced\_subsample | * option: build model without resampling * number of folds: 5 * average size of training set: 542.8 * average prop of minority class: 0.446 | 0.952 | 0.238 | 0.931 | 0.077 | 0.931 | 0.754 | 0.893 | 0.084 |

## 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: True
* max\_k\_undersampling: 5
* resample\_\_sampling\_strategy: minority
* post\_process\_\_option: option\_3
* model\_\_max\_depth: 10
* model\_\_learning\_rate: 0.058
* model\_\_min\_child\_weight: 4
* model\_\_subsample: 0.94
* model\_\_colsample\_bytree: 0.38
* model\_\_scale\_pos\_weight: 1.667
* 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: 10 * model\_\_learning\_rate: 0.083 * model\_\_min\_child\_weight: 6 * model\_\_subsample: 0.981 * model\_\_colsample\_bytree: 0.282 * model\_\_scale\_pos\_weight: 8.69 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 5 * 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: 482.4 * average prop of minority class after resampling: 0.5 | 0.968 | 0.184 | 0.954 | 0.058 | 0.911 | 0.39 | 0.916 | 0.12 |
| XGBClassifier | * undersampling\_majority\_class: False * max\_k\_undersampling: 6 * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_2 * model\_\_max\_depth: 10 * model\_\_learning\_rate: 0.248 * model\_\_min\_child\_weight: 8 * model\_\_subsample: 0.876 * model\_\_colsample\_bytree: 0.138 * model\_\_scale\_pos\_weight: 0.188 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 5 * average size of training set before resampling: 542.8 * average prop of minority class before resampling: 0.444 * average size of training set after resampling: 603.2 * average prop of minority class after resampling: 0.5 | 0.958 | 0.223 | 0.944 | 0.07 | 0.907 | 0.412 | 0.908 | 0.135 |
| XGBClassifier | * undersampling\_majority\_class: False * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_1 * model\_\_max\_depth: 15 * model\_\_learning\_rate: 0.097 * model\_\_min\_child\_weight: 7 * model\_\_subsample: 0.94 * model\_\_colsample\_bytree: 0.647 * model\_\_scale\_pos\_weight: 2.363 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 5 * average size of training set before resampling: 542.8 * average prop of minority class before resampling: 0.444 * average size of training set after resampling: 603.2 * average prop of minority class after resampling: 0.5 | 0.933 | 0.287 | 0.904 | 0.091 | 0.876 | 0.447 | 0.878 | 0.144 |
| XGBClassifier | * undersampling\_majority\_class: False * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_1 * model\_\_max\_depth: 13 * model\_\_learning\_rate: 0.084 * model\_\_min\_child\_weight: 4 * model\_\_subsample: 0.963 * model\_\_colsample\_bytree: 0.456 * model\_\_scale\_pos\_weight: 0.188 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 5 * average size of training set before resampling: 542.8 * average prop of minority class before resampling: 0.444 * average size of training set after resampling: 603.2 * average prop of minority class after resampling: 0.5 | 0.932 | 0.283 | 0.905 | 0.089 | 0.873 | 0.456 | 0.884 | 0.143 |
| XGBClassifier | * undersampling\_majority\_class: False * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_2 * model\_\_max\_depth: 11 * model\_\_learning\_rate: 0.053 * model\_\_min\_child\_weight: 6 * model\_\_subsample: 0.909 * model\_\_colsample\_bytree: 0.332 * model\_\_scale\_pos\_weight: 1.42 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 5 * average size of training set before resampling: 542.8 * average prop of minority class before resampling: 0.444 * average size of training set after resampling: 603.2 * average prop of minority class after resampling: 0.5 | 0.951 | 0.234 | 0.931 | 0.074 | 0.909 | 0.408 | 0.909 | 0.138 |
| XGBClassifier | * undersampling\_majority\_class: True * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_1 * model\_\_max\_depth: 8 * model\_\_learning\_rate: 0.08 * model\_\_min\_child\_weight: 2 * model\_\_subsample: 0.961 * model\_\_colsample\_bytree: 0.436 * model\_\_scale\_pos\_weight: 0.185 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 5 * 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: 482.4 * average prop of minority class after resampling: 0.5 | 0.938 | 0.271 | 0.915 | 0.088 | 0.9 | 0.401 | 0.886 | 0.13 |
| XGBClassifier | * undersampling\_majority\_class: True * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_3 * model\_\_max\_depth: 10 * model\_\_learning\_rate: 0.058 * model\_\_min\_child\_weight: 4 * model\_\_subsample: 0.94 * model\_\_colsample\_bytree: 0.38 * model\_\_scale\_pos\_weight: 1.667 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 5 * 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: 482.4 * average prop of minority class after resampling: 0.5 | 0.972 | 0.178 | 0.958 | 0.057 | 0.917 | 0.373 | 0.921 | 0.12 |
| XGBClassifier | * undersampling\_majority\_class: False * max\_k\_undersampling: 6 * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_2 * model\_\_max\_depth: 6 * model\_\_learning\_rate: 0.232 * model\_\_min\_child\_weight: 7 * model\_\_subsample: 0.986 * model\_\_colsample\_bytree: 0.17 * model\_\_scale\_pos\_weight: 0.184 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 5 * average size of training set before resampling: 542.8 * average prop of minority class before resampling: 0.444 * average size of training set after resampling: 603.2 * average prop of minority class after resampling: 0.5 | 0.95 | 0.243 | 0.929 | 0.077 | 0.877 | 0.45 | 0.89 | 0.147 |
| XGBClassifier | * undersampling\_majority\_class: True * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_2 * model\_\_max\_depth: 5 * model\_\_learning\_rate: 0.284 * model\_\_min\_child\_weight: 2 * model\_\_subsample: 0.816 * model\_\_colsample\_bytree: 0.475 * model\_\_scale\_pos\_weight: 1.162 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 5 * 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: 482.4 * average prop of minority class after resampling: 0.5 | 0.946 | 0.26 | 0.92 | 0.083 | 0.878 | 0.467 | 0.876 | 0.157 |
| XGBClassifier | * undersampling\_majority\_class: False * max\_k\_undersampling: 6 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_2 * model\_\_max\_depth: 12 * model\_\_learning\_rate: 0.168 * model\_\_min\_child\_weight: 8 * model\_\_subsample: 0.816 * model\_\_colsample\_bytree: 0.444 * model\_\_scale\_pos\_weight: 1.043 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 5 * average size of training set before resampling: 542.8 * average prop of minority class before resampling: 0.444 * average size of training set after resampling: 603.2 * average prop of minority class after resampling: 0.5 | 0.959 | 0.208 | 0.944 | 0.067 | 0.897 | 0.419 | 0.904 | 0.141 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 1 * model\_\_n\_estimators: 32 * model\_\_max\_depth: 3 * model\_\_min\_samples\_split: 14 * model\_\_min\_samples\_leaf: 3 * model\_\_max\_features: auto * model\_\_class\_weight: balanced\_subsample | * option: build model without resampling * number of folds: 5 * average size of training set: 542.8 * average prop of minority class: 0.444 | 0.958 | 0.222 | 0.936 | 0.07 | 0.904 | 0.439 | 0.901 | 0.136 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 0 * model\_\_n\_estimators: 78 * model\_\_max\_depth: 7 * model\_\_min\_samples\_split: 7 * model\_\_min\_samples\_leaf: 2 * model\_\_max\_features: sqrt * model\_\_class\_weight: balanced | * option: build model without resampling * number of folds: 5 * average size of training set: 542.8 * average prop of minority class: 0.444 | 0.971 | 0.176 | 0.952 | 0.057 | 0.918 | 0.391 | 0.915 | 0.133 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 0 * model\_\_n\_estimators: 30 * model\_\_max\_depth: 6 * model\_\_min\_samples\_split: 12 * model\_\_min\_samples\_leaf: 2 * model\_\_max\_features: auto * model\_\_class\_weight: balanced | * option: build model without resampling * number of folds: 5 * average size of training set: 392.0 * average prop of minority class: 0.385 | 0.965 | 0.195 | 0.944 | 0.062 | 0.918 | 0.39 | 0.919 | 0.131 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 0 * model\_\_n\_estimators: 18 * model\_\_max\_depth: 7 * model\_\_min\_samples\_split: 16 * model\_\_min\_samples\_leaf: 2 * model\_\_max\_features: sqrt * model\_\_class\_weight: balanced | * option: build model without resampling * number of folds: 5 * average size of training set: 542.8 * average prop of minority class: 0.444 | 0.964 | 0.196 | 0.94 | 0.062 | 0.903 | 0.421 | 0.902 | 0.134 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 0 * model\_\_n\_estimators: 57 * 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: 5 * average size of training set: 392.0 * average prop of minority class: 0.385 | 0.963 | 0.207 | 0.941 | 0.066 | 0.915 | 0.381 | 0.91 | 0.127 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 0 * model\_\_n\_estimators: 69 * model\_\_max\_depth: 5 * model\_\_min\_samples\_split: 8 * model\_\_min\_samples\_leaf: 4 * model\_\_max\_features: sqrt * model\_\_class\_weight: balanced\_subsample | * option: build model without resampling * number of folds: 5 * average size of training set: 392.0 * average prop of minority class: 0.385 | 0.967 | 0.193 | 0.948 | 0.062 | 0.921 | 0.374 | 0.918 | 0.126 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 1 * model\_\_n\_estimators: 11 * 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: 5 * average size of training set: 542.8 * average prop of minority class: 0.444 | 0.963 | 0.209 | 0.94 | 0.067 | 0.922 | 0.367 | 0.92 | 0.123 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 0 * model\_\_n\_estimators: 93 * model\_\_max\_depth: 6 * model\_\_min\_samples\_split: 11 * model\_\_min\_samples\_leaf: 2 * model\_\_max\_features: sqrt * model\_\_class\_weight: balanced | * option: build model without resampling * number of folds: 5 * average size of training set: 542.8 * average prop of minority class: 0.444 | 0.969 | 0.184 | 0.952 | 0.059 | 0.919 | 0.38 | 0.916 | 0.127 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 1 * model\_\_n\_estimators: 48 * model\_\_max\_depth: 5 * model\_\_min\_samples\_split: 6 * model\_\_min\_samples\_leaf: 2 * model\_\_max\_features: sqrt * model\_\_class\_weight: balanced\_subsample | * option: build model without resampling * number of folds: 5 * average size of training set: 392.0 * average prop of minority class: 0.385 | 0.969 | 0.185 | 0.95 | 0.06 | 0.912 | 0.397 | 0.911 | 0.133 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 0 * model\_\_n\_estimators: 39 * model\_\_max\_depth: 5 * model\_\_min\_samples\_split: 12 * model\_\_min\_samples\_leaf: 4 * model\_\_max\_features: auto * model\_\_class\_weight: balanced\_subsample | * option: build model without resampling * number of folds: 5 * average size of training set: 542.8 * average prop of minority class: 0.444 | 0.965 | 0.193 | 0.945 | 0.062 | 0.915 | 0.394 | 0.911 | 0.129 |

## 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 RandomForestClassifier with the following params:

* undersampling\_majority\_class: False
* model\_\_bootstrap: 0
* model\_\_n\_estimators: 91
* model\_\_max\_depth: 6
* model\_\_min\_samples\_split: 15
* model\_\_min\_samples\_leaf: 3
* 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: 6 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_3 * model\_\_max\_depth: 13 * model\_\_learning\_rate: 0.052 * model\_\_min\_child\_weight: 3 * model\_\_subsample: 0.824 * model\_\_colsample\_bytree: 0.32 * model\_\_scale\_pos\_weight: 0.129 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 5 * 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: 482.4 * average prop of minority class after resampling: 0.5 | 0.958 | 0.223 | 0.932 | 0.07 | 0.96 | 0.253 | 0.947 | 0.08 |
| XGBClassifier | * undersampling\_majority\_class: True * max\_k\_undersampling: 6 * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_2 * model\_\_max\_depth: 6 * model\_\_learning\_rate: 0.054 * model\_\_min\_child\_weight: 4 * model\_\_subsample: 0.813 * model\_\_colsample\_bytree: 0.15 * model\_\_scale\_pos\_weight: 6.37 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 5 * 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: 482.4 * average prop of minority class after resampling: 0.5 | 0.952 | 0.232 | 0.921 | 0.075 | 0.947 | 0.286 | 0.93 | 0.091 |
| XGBClassifier | * undersampling\_majority\_class: False * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_2 * model\_\_max\_depth: 13 * model\_\_learning\_rate: 0.238 * model\_\_min\_child\_weight: 9 * model\_\_subsample: 0.97 * model\_\_colsample\_bytree: 0.209 * model\_\_scale\_pos\_weight: 0.952 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 5 * average size of training set before resampling: 542.8 * average prop of minority class before resampling: 0.444 * average size of training set after resampling: 603.2 * average prop of minority class after resampling: 0.5 | 0.952 | 0.228 | 0.916 | 0.073 | 0.954 | 0.273 | 0.944 | 0.085 |
| XGBClassifier | * undersampling\_majority\_class: True * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_2 * model\_\_max\_depth: 15 * model\_\_learning\_rate: 0.259 * model\_\_min\_child\_weight: 7 * model\_\_subsample: 0.945 * model\_\_colsample\_bytree: 0.188 * model\_\_scale\_pos\_weight: 0.142 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 5 * 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: 482.4 * average prop of minority class after resampling: 0.5 | 0.945 | 0.259 | 0.916 | 0.083 | 0.949 | 0.304 | 0.94 | 0.093 |
| XGBClassifier | * undersampling\_majority\_class: False * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_2 * model\_\_max\_depth: 12 * model\_\_learning\_rate: 0.088 * model\_\_min\_child\_weight: 5 * model\_\_subsample: 0.931 * model\_\_colsample\_bytree: 0.68 * model\_\_scale\_pos\_weight: 0.271 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 5 * average size of training set before resampling: 542.8 * average prop of minority class before resampling: 0.444 * average size of training set after resampling: 603.2 * average prop of minority class after resampling: 0.5 | 0.956 | 0.218 | 0.926 | 0.071 | 0.964 | 0.252 | 0.955 | 0.082 |
| XGBClassifier | * undersampling\_majority\_class: True * max\_k\_undersampling: 6 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_1 * model\_\_max\_depth: 8 * model\_\_learning\_rate: 0.135 * model\_\_min\_child\_weight: 9 * model\_\_subsample: 0.926 * model\_\_colsample\_bytree: 0.267 * model\_\_scale\_pos\_weight: 0.289 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 5 * 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: 482.4 * average prop of minority class after resampling: 0.5 | 0.93 | 0.304 | 0.897 | 0.098 | 0.9 | 0.393 | 0.904 | 0.11 |
| XGBClassifier | * undersampling\_majority\_class: True * max\_k\_undersampling: 6 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_2 * model\_\_max\_depth: 14 * model\_\_learning\_rate: 0.125 * model\_\_min\_child\_weight: 9 * model\_\_subsample: 0.91 * model\_\_colsample\_bytree: 0.177 * model\_\_scale\_pos\_weight: 0.181 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 5 * 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: 482.4 * average prop of minority class after resampling: 0.5 | 0.936 | 0.283 | 0.9 | 0.089 | 0.946 | 0.3 | 0.922 | 0.091 |
| XGBClassifier | * undersampling\_majority\_class: False * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_2 * model\_\_max\_depth: 14 * model\_\_learning\_rate: 0.071 * model\_\_min\_child\_weight: 3 * model\_\_subsample: 0.85 * model\_\_colsample\_bytree: 0.244 * model\_\_scale\_pos\_weight: 0.746 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 5 * average size of training set before resampling: 542.8 * average prop of minority class before resampling: 0.444 * average size of training set after resampling: 603.2 * average prop of minority class after resampling: 0.5 | 0.95 | 0.234 | 0.92 | 0.075 | 0.946 | 0.301 | 0.935 | 0.094 |
| XGBClassifier | * undersampling\_majority\_class: False * max\_k\_undersampling: 6 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_2 * model\_\_max\_depth: 7 * model\_\_learning\_rate: 0.152 * model\_\_min\_child\_weight: 6 * model\_\_subsample: 0.86 * model\_\_colsample\_bytree: 0.343 * model\_\_scale\_pos\_weight: 0.227 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 5 * average size of training set before resampling: 542.8 * average prop of minority class before resampling: 0.444 * average size of training set after resampling: 603.2 * average prop of minority class after resampling: 0.5 | 0.954 | 0.219 | 0.923 | 0.069 | 0.959 | 0.265 | 0.943 | 0.081 |
| XGBClassifier | * undersampling\_majority\_class: True * max\_k\_undersampling: 6 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_1 * model\_\_max\_depth: 8 * model\_\_learning\_rate: 0.139 * model\_\_min\_child\_weight: 4 * model\_\_subsample: 0.969 * model\_\_colsample\_bytree: 0.296 * model\_\_scale\_pos\_weight: 1.221 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 5 * 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: 482.4 * average prop of minority class after resampling: 0.5 | 0.917 | 0.326 | 0.883 | 0.108 | 0.922 | 0.329 | 0.934 | 0.098 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 1 * model\_\_n\_estimators: 27 * model\_\_max\_depth: 4 * model\_\_min\_samples\_split: 12 * model\_\_min\_samples\_leaf: 4 * model\_\_max\_features: auto * model\_\_class\_weight: balanced\_subsample | * option: build model without resampling * number of folds: 5 * average size of training set: 542.8 * average prop of minority class: 0.444 | 0.959 | 0.215 | 0.933 | 0.068 | 0.959 | 0.249 | 0.951 | 0.079 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 0 * model\_\_n\_estimators: 67 * model\_\_max\_depth: 6 * model\_\_min\_samples\_split: 6 * model\_\_min\_samples\_leaf: 2 * model\_\_max\_features: sqrt * model\_\_class\_weight: balanced | * option: build model without resampling * number of folds: 5 * average size of training set: 542.8 * average prop of minority class: 0.444 | 0.961 | 0.199 | 0.931 | 0.065 | 0.952 | 0.284 | 0.932 | 0.089 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 1 * model\_\_n\_estimators: 73 * model\_\_max\_depth: 6 * 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: 5 * average size of training set: 542.8 * average prop of minority class: 0.444 | 0.964 | 0.186 | 0.938 | 0.062 | 0.957 | 0.274 | 0.934 | 0.086 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 0 * model\_\_n\_estimators: 46 * model\_\_max\_depth: 10 * model\_\_min\_samples\_split: 8 * model\_\_min\_samples\_leaf: 1 * model\_\_max\_features: auto * model\_\_class\_weight: balanced\_subsample | * option: build model without resampling * number of folds: 5 * average size of training set: 392.0 * average prop of minority class: 0.385 | 0.966 | 0.183 | 0.939 | 0.061 | 0.954 | 0.725 | 0.909 | 0.084 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 1 * model\_\_n\_estimators: 13 * model\_\_max\_depth: 10 * model\_\_min\_samples\_split: 16 * model\_\_min\_samples\_leaf: 4 * model\_\_max\_features: auto * model\_\_class\_weight: balanced | * option: build model without resampling * number of folds: 5 * average size of training set: 542.8 * average prop of minority class: 0.444 | 0.956 | 0.227 | 0.926 | 0.075 | 0.951 | 0.282 | 0.936 | 0.091 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 1 * model\_\_n\_estimators: 42 * model\_\_max\_depth: 9 * model\_\_min\_samples\_split: 19 * model\_\_min\_samples\_leaf: 4 * model\_\_max\_features: auto * model\_\_class\_weight: balanced\_subsample | * option: build model without resampling * number of folds: 5 * average size of training set: 392.0 * average prop of minority class: 0.385 | 0.963 | 0.202 | 0.937 | 0.066 | 0.956 | 0.257 | 0.941 | 0.083 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 0 * model\_\_n\_estimators: 79 * model\_\_max\_depth: 3 * model\_\_min\_samples\_split: 6 * model\_\_min\_samples\_leaf: 2 * model\_\_max\_features: sqrt * model\_\_class\_weight: balanced\_subsample | * option: build model without resampling * number of folds: 5 * average size of training set: 542.8 * average prop of minority class: 0.444 | 0.955 | 0.22 | 0.924 | 0.072 | 0.95 | 0.284 | 0.94 | 0.088 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 1 * model\_\_n\_estimators: 75 * model\_\_max\_depth: 4 * model\_\_min\_samples\_split: 19 * model\_\_min\_samples\_leaf: 4 * model\_\_max\_features: auto * model\_\_class\_weight: balanced | * option: build model without resampling * number of folds: 5 * average size of training set: 392.0 * average prop of minority class: 0.385 | 0.957 | 0.219 | 0.927 | 0.071 | 0.954 | 0.275 | 0.945 | 0.089 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 0 * model\_\_n\_estimators: 77 * model\_\_max\_depth: 5 * model\_\_min\_samples\_split: 13 * model\_\_min\_samples\_leaf: 3 * model\_\_max\_features: sqrt * model\_\_class\_weight: balanced | * option: build model without resampling * number of folds: 5 * average size of training set: 542.8 * average prop of minority class: 0.444 | 0.961 | 0.202 | 0.934 | 0.065 | 0.954 | 0.276 | 0.941 | 0.086 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 0 * model\_\_n\_estimators: 91 * model\_\_max\_depth: 6 * model\_\_min\_samples\_split: 15 * model\_\_min\_samples\_leaf: 3 * model\_\_max\_features: sqrt * model\_\_class\_weight: balanced | * option: build model without resampling * number of folds: 5 * average size of training set: 542.8 * average prop of minority class: 0.444 | 0.968 | 0.182 | 0.942 | 0.06 | 0.948 | 0.745 | 0.904 | 0.086 |

## 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 XGBClassifier with the following params:

* undersampling\_majority\_class: True
* max\_k\_undersampling: 5
* resample\_\_sampling\_strategy: minority
* post\_process\_\_option: option\_3
* model\_\_max\_depth: 14
* model\_\_learning\_rate: 0.101
* model\_\_min\_child\_weight: 2
* model\_\_subsample: 0.874
* model\_\_colsample\_bytree: 0.534
* model\_\_scale\_pos\_weight: 6.33
* 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: 6 * model\_\_learning\_rate: 0.099 * model\_\_min\_child\_weight: 9 * model\_\_subsample: 0.807 * model\_\_colsample\_bytree: 0.544 * model\_\_scale\_pos\_weight: 1.256 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 5 * average size of training set before resampling: 542.8 * average prop of minority class before resampling: 0.444 * average size of training set after resampling: 603.2 * average prop of minority class after resampling: 0.5 | 0.949 | 0.239 | 0.911 | 0.077 | 0.949 | 0.318 | 0.948 | 0.091 |
| 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: 0.229 * model\_\_min\_child\_weight: 8 * model\_\_subsample: 0.959 * model\_\_colsample\_bytree: 0.681 * model\_\_scale\_pos\_weight: 0.914 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 5 * 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: 482.4 * average prop of minority class after resampling: 0.5 | 0.949 | 0.24 | 0.911 | 0.078 | 0.937 | 0.749 | 0.938 | 0.088 |
| XGBClassifier | * undersampling\_majority\_class: False * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_2 * model\_\_max\_depth: 6 * model\_\_learning\_rate: 0.104 * model\_\_min\_child\_weight: 8 * model\_\_subsample: 0.969 * model\_\_colsample\_bytree: 0.277 * model\_\_scale\_pos\_weight: 7.827 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 5 * average size of training set before resampling: 542.8 * average prop of minority class before resampling: 0.444 * average size of training set after resampling: 603.2 * average prop of minority class after resampling: 0.5 | 0.941 | 0.248 | 0.903 | 0.08 | 0.939 | 0.744 | 0.943 | 0.087 |
| XGBClassifier | * undersampling\_majority\_class: True * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_2 * model\_\_max\_depth: 11 * model\_\_learning\_rate: 0.147 * model\_\_min\_child\_weight: 1 * model\_\_subsample: 0.821 * model\_\_colsample\_bytree: 0.149 * model\_\_scale\_pos\_weight: 4.208 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 5 * 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: 482.4 * average prop of minority class after resampling: 0.5 | 0.937 | 0.265 | 0.898 | 0.084 | 0.95 | 0.305 | 0.957 | 0.078 |
| XGBClassifier | * undersampling\_majority\_class: True * max\_k\_undersampling: 6 * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_1 * model\_\_max\_depth: 10 * model\_\_learning\_rate: 0.081 * model\_\_min\_child\_weight: 5 * model\_\_subsample: 0.807 * model\_\_colsample\_bytree: 0.245 * model\_\_scale\_pos\_weight: 0.483 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 5 * 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: 482.4 * average prop of minority class after resampling: 0.5 | 0.92 | 0.326 | 0.877 | 0.106 | 0.943 | 0.315 | 0.93 | 0.1 |
| XGBClassifier | * undersampling\_majority\_class: True * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_2 * model\_\_max\_depth: 10 * model\_\_learning\_rate: 0.14 * model\_\_min\_child\_weight: 5 * model\_\_subsample: 0.928 * model\_\_colsample\_bytree: 0.147 * model\_\_scale\_pos\_weight: 0.464 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 5 * 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: 482.4 * average prop of minority class after resampling: 0.5 | 0.949 | 0.238 | 0.917 | 0.077 | 0.949 | 0.71 | 0.956 | 0.076 |
| XGBClassifier | * undersampling\_majority\_class: True * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_3 * model\_\_max\_depth: 14 * model\_\_learning\_rate: 0.101 * model\_\_min\_child\_weight: 2 * model\_\_subsample: 0.874 * model\_\_colsample\_bytree: 0.534 * model\_\_scale\_pos\_weight: 6.33 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 5 * 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: 482.4 * average prop of minority class after resampling: 0.5 | 0.959 | 0.205 | 0.932 | 0.065 | 0.969 | 0.241 | 0.966 | 0.072 |
| XGBClassifier | * undersampling\_majority\_class: False * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_2 * model\_\_max\_depth: 14 * model\_\_learning\_rate: 0.269 * model\_\_min\_child\_weight: 5 * model\_\_subsample: 0.807 * model\_\_colsample\_bytree: 0.569 * model\_\_scale\_pos\_weight: 5.054 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 5 * average size of training set before resampling: 542.8 * average prop of minority class before resampling: 0.444 * average size of training set after resampling: 603.2 * average prop of minority class after resampling: 0.5 | 0.943 | 0.243 | 0.91 | 0.081 | 0.928 | 0.772 | 0.92 | 0.089 |
| XGBClassifier | * undersampling\_majority\_class: False * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_3 * model\_\_max\_depth: 11 * model\_\_learning\_rate: 0.119 * model\_\_min\_child\_weight: 6 * model\_\_subsample: 0.953 * model\_\_colsample\_bytree: 0.347 * model\_\_scale\_pos\_weight: 1.796 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 5 * average size of training set before resampling: 542.8 * average prop of minority class before resampling: 0.444 * average size of training set after resampling: 603.2 * average prop of minority class after resampling: 0.5 | 0.958 | 0.201 | 0.93 | 0.064 | 0.967 | 0.24 | 0.967 | 0.067 |
| XGBClassifier | * undersampling\_majority\_class: False * max\_k\_undersampling: 6 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_2 * model\_\_max\_depth: 7 * model\_\_learning\_rate: 0.281 * model\_\_min\_child\_weight: 3 * model\_\_subsample: 0.838 * model\_\_colsample\_bytree: 0.239 * model\_\_scale\_pos\_weight: 3.355 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 5 * average size of training set before resampling: 542.8 * average prop of minority class before resampling: 0.444 * average size of training set after resampling: 603.2 * average prop of minority class after resampling: 0.5 | 0.934 | 0.275 | 0.895 | 0.09 | 0.913 | 0.79 | 0.918 | 0.096 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 1 * model\_\_n\_estimators: 70 * model\_\_max\_depth: 6 * model\_\_min\_samples\_split: 12 * model\_\_min\_samples\_leaf: 2 * model\_\_max\_features: sqrt * model\_\_class\_weight: balanced | * option: build model without resampling * number of folds: 5 * average size of training set: 392.0 * average prop of minority class: 0.385 | 0.957 | 0.211 | 0.926 | 0.067 | 0.974 | 0.24 | 0.969 | 0.072 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 0 * model\_\_n\_estimators: 80 * model\_\_max\_depth: 6 * model\_\_min\_samples\_split: 15 * model\_\_min\_samples\_leaf: 2 * model\_\_max\_features: auto * model\_\_class\_weight: balanced | * option: build model without resampling * number of folds: 5 * average size of training set: 542.8 * average prop of minority class: 0.444 | 0.959 | 0.204 | 0.929 | 0.066 | 0.963 | 0.271 | 0.961 | 0.083 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 1 * model\_\_n\_estimators: 27 * model\_\_max\_depth: 4 * model\_\_min\_samples\_split: 16 * model\_\_min\_samples\_leaf: 2 * model\_\_max\_features: sqrt * model\_\_class\_weight: balanced | * option: build model without resampling * number of folds: 5 * average size of training set: 392.0 * average prop of minority class: 0.385 | 0.957 | 0.211 | 0.928 | 0.067 | 0.97 | 0.253 | 0.964 | 0.079 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 0 * model\_\_n\_estimators: 51 * model\_\_max\_depth: 3 * model\_\_min\_samples\_split: 8 * model\_\_min\_samples\_leaf: 2 * model\_\_max\_features: sqrt * model\_\_class\_weight: balanced | * option: build model without resampling * number of folds: 5 * average size of training set: 542.8 * average prop of minority class: 0.444 | 0.952 | 0.222 | 0.918 | 0.071 | 0.946 | 0.737 | 0.943 | 0.086 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 0 * model\_\_n\_estimators: 21 * model\_\_max\_depth: 7 * model\_\_min\_samples\_split: 15 * model\_\_min\_samples\_leaf: 4 * model\_\_max\_features: auto * model\_\_class\_weight: balanced\_subsample | * option: build model without resampling * number of folds: 5 * average size of training set: 542.8 * average prop of minority class: 0.444 | 0.957 | 0.214 | 0.925 | 0.07 | 0.958 | 0.272 | 0.946 | 0.084 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 0 * model\_\_n\_estimators: 22 * model\_\_max\_depth: 5 * model\_\_min\_samples\_split: 16 * model\_\_min\_samples\_leaf: 3 * model\_\_max\_features: auto * model\_\_class\_weight: balanced\_subsample | * option: build model without resampling * number of folds: 5 * average size of training set: 392.0 * average prop of minority class: 0.385 | 0.958 | 0.211 | 0.926 | 0.068 | 0.96 | 0.267 | 0.95 | 0.081 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 0 * model\_\_n\_estimators: 91 * model\_\_max\_depth: 8 * model\_\_min\_samples\_split: 14 * model\_\_min\_samples\_leaf: 3 * model\_\_max\_features: auto * model\_\_class\_weight: balanced | * option: build model without resampling * number of folds: 5 * average size of training set: 542.8 * average prop of minority class: 0.444 | 0.957 | 0.212 | 0.925 | 0.067 | 0.96 | 0.283 | 0.952 | 0.085 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 0 * model\_\_n\_estimators: 24 * model\_\_max\_depth: 9 * model\_\_min\_samples\_split: 20 * model\_\_min\_samples\_leaf: 1 * model\_\_max\_features: sqrt * model\_\_class\_weight: balanced | * option: build model without resampling * number of folds: 5 * average size of training set: 542.8 * average prop of minority class: 0.444 | 0.956 | 0.221 | 0.92 | 0.071 | 0.949 | 0.325 | 0.948 | 0.094 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 0 * model\_\_n\_estimators: 99 * model\_\_max\_depth: 9 * model\_\_min\_samples\_split: 17 * model\_\_min\_samples\_leaf: 4 * model\_\_max\_features: auto * model\_\_class\_weight: balanced | * option: build model without resampling * number of folds: 5 * average size of training set: 392.0 * average prop of minority class: 0.385 | 0.958 | 0.206 | 0.927 | 0.066 | 0.96 | 0.289 | 0.955 | 0.086 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 0 * model\_\_n\_estimators: 85 * model\_\_max\_depth: 6 * model\_\_min\_samples\_split: 13 * model\_\_min\_samples\_leaf: 3 * model\_\_max\_features: sqrt * model\_\_class\_weight: balanced | * option: build model without resampling * number of folds: 5 * average size of training set: 392.0 * average prop of minority class: 0.385 | 0.958 | 0.208 | 0.927 | 0.067 | 0.969 | 0.245 | 0.964 | 0.077 |

## 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 RandomForestClassifier with the following params:

* undersampling\_majority\_class: True
* model\_\_bootstrap: 1
* model\_\_n\_estimators: 50
* model\_\_max\_depth: 7
* model\_\_min\_samples\_split: 11
* model\_\_min\_samples\_leaf: 3
* model\_\_max\_features: auto
* 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: 14 * model\_\_learning\_rate: 0.244 * model\_\_min\_child\_weight: 4 * model\_\_subsample: 0.821 * model\_\_colsample\_bytree: 0.469 * model\_\_scale\_pos\_weight: 0.19 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 5 * 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: 482.4 * average prop of minority class after resampling: 0.5 | 0.949 | 0.243 | 0.916 | 0.076 | 0.947 | 0.276 | 0.941 | 0.074 |
| XGBClassifier | * undersampling\_majority\_class: True * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_2 * model\_\_max\_depth: 11 * model\_\_learning\_rate: 0.101 * model\_\_min\_child\_weight: 9 * model\_\_subsample: 0.954 * model\_\_colsample\_bytree: 0.639 * model\_\_scale\_pos\_weight: 1.723 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 5 * 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: 482.4 * average prop of minority class after resampling: 0.5 | 0.952 | 0.232 | 0.916 | 0.076 | 0.962 | 0.281 | 0.943 | 0.084 |
| 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: 0.142 * model\_\_min\_child\_weight: 6 * model\_\_subsample: 0.93 * model\_\_colsample\_bytree: 0.376 * model\_\_scale\_pos\_weight: 0.329 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 5 * average size of training set before resampling: 542.8 * average prop of minority class before resampling: 0.444 * average size of training set after resampling: 603.2 * average prop of minority class after resampling: 0.5 | 0.95 | 0.238 | 0.916 | 0.077 | 0.956 | 0.26 | 0.957 | 0.074 |
| XGBClassifier | * undersampling\_majority\_class: False * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_2 * model\_\_max\_depth: 8 * model\_\_learning\_rate: 0.103 * model\_\_min\_child\_weight: 6 * model\_\_subsample: 0.803 * model\_\_colsample\_bytree: 0.201 * model\_\_scale\_pos\_weight: 0.156 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 5 * average size of training set before resampling: 542.8 * average prop of minority class before resampling: 0.444 * average size of training set after resampling: 603.2 * average prop of minority class after resampling: 0.5 | 0.94 | 0.26 | 0.903 | 0.086 | 0.946 | 0.298 | 0.946 | 0.089 |
| XGBClassifier | * undersampling\_majority\_class: True * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_1 * model\_\_max\_depth: 9 * model\_\_learning\_rate: 0.12 * model\_\_min\_child\_weight: 2 * model\_\_subsample: 0.891 * model\_\_colsample\_bytree: 0.705 * model\_\_scale\_pos\_weight: 1.566 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 5 * 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: 482.4 * average prop of minority class after resampling: 0.5 | 0.914 | 0.329 | 0.875 | 0.106 | 0.946 | 0.329 | 0.941 | 0.102 |
| 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: 0.102 * model\_\_min\_child\_weight: 7 * model\_\_subsample: 0.869 * model\_\_colsample\_bytree: 0.251 * model\_\_scale\_pos\_weight: 4.865 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 5 * 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: 482.4 * average prop of minority class after resampling: 0.5 | 0.955 | 0.224 | 0.921 | 0.072 | 0.962 | 0.272 | 0.954 | 0.085 |
| XGBClassifier | * undersampling\_majority\_class: True * max\_k\_undersampling: 6 * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_2 * model\_\_max\_depth: 12 * model\_\_learning\_rate: 0.162 * model\_\_min\_child\_weight: 6 * model\_\_subsample: 0.996 * model\_\_colsample\_bytree: 0.242 * model\_\_scale\_pos\_weight: 0.881 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 5 * 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: 482.4 * average prop of minority class after resampling: 0.5 | 0.956 | 0.218 | 0.923 | 0.071 | 0.947 | 0.295 | 0.942 | 0.09 |
| XGBClassifier | * undersampling\_majority\_class: False * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_2 * model\_\_max\_depth: 7 * model\_\_learning\_rate: 0.18 * model\_\_min\_child\_weight: 9 * model\_\_subsample: 0.866 * model\_\_colsample\_bytree: 0.204 * model\_\_scale\_pos\_weight: 0.917 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 5 * average size of training set before resampling: 542.8 * average prop of minority class before resampling: 0.444 * average size of training set after resampling: 603.2 * average prop of minority class after resampling: 0.5 | 0.944 | 0.244 | 0.904 | 0.079 | 0.95 | 0.28 | 0.946 | 0.084 |
| XGBClassifier | * undersampling\_majority\_class: False * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_2 * model\_\_max\_depth: 7 * model\_\_learning\_rate: 0.091 * model\_\_min\_child\_weight: 5 * model\_\_subsample: 0.972 * model\_\_colsample\_bytree: 0.146 * model\_\_scale\_pos\_weight: 1.12 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 5 * average size of training set before resampling: 542.8 * average prop of minority class before resampling: 0.444 * average size of training set after resampling: 603.2 * average prop of minority class after resampling: 0.5 | 0.951 | 0.237 | 0.918 | 0.077 | 0.954 | 0.288 | 0.953 | 0.088 |
| XGBClassifier | * undersampling\_majority\_class: True * max\_k\_undersampling: 6 * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_2 * model\_\_max\_depth: 15 * model\_\_learning\_rate: 0.154 * model\_\_min\_child\_weight: 6 * model\_\_subsample: 0.82 * model\_\_colsample\_bytree: 0.296 * model\_\_scale\_pos\_weight: 5.161 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 5 * 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: 482.4 * average prop of minority class after resampling: 0.5 | 0.953 | 0.231 | 0.917 | 0.075 | 0.939 | 0.315 | 0.937 | 0.096 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 0 * model\_\_n\_estimators: 96 * model\_\_max\_depth: 5 * model\_\_min\_samples\_split: 6 * model\_\_min\_samples\_leaf: 1 * model\_\_max\_features: sqrt * model\_\_class\_weight: balanced | * option: build model without resampling * number of folds: 5 * average size of training set: 392.0 * average prop of minority class: 0.385 | 0.954 | 0.219 | 0.92 | 0.07 | 0.973 | 0.231 | 0.959 | 0.067 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 0 * model\_\_n\_estimators: 98 * model\_\_max\_depth: 3 * model\_\_min\_samples\_split: 19 * model\_\_min\_samples\_leaf: 2 * model\_\_max\_features: auto * model\_\_class\_weight: balanced | * option: build model without resampling * number of folds: 5 * average size of training set: 392.0 * average prop of minority class: 0.385 | 0.949 | 0.238 | 0.914 | 0.075 | 0.956 | 0.27 | 0.951 | 0.08 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 1 * model\_\_n\_estimators: 50 * model\_\_max\_depth: 7 * model\_\_min\_samples\_split: 11 * model\_\_min\_samples\_leaf: 3 * model\_\_max\_features: auto * model\_\_class\_weight: balanced | * option: build model without resampling * number of folds: 5 * average size of training set: 392.0 * average prop of minority class: 0.385 | 0.955 | 0.219 | 0.926 | 0.069 | 0.978 | 0.242 | 0.969 | 0.072 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 0 * model\_\_n\_estimators: 99 * model\_\_max\_depth: 7 * model\_\_min\_samples\_split: 14 * model\_\_min\_samples\_leaf: 2 * model\_\_max\_features: sqrt * model\_\_class\_weight: balanced\_subsample | * option: build model without resampling * number of folds: 5 * average size of training set: 392.0 * average prop of minority class: 0.385 | 0.953 | 0.221 | 0.922 | 0.071 | 0.968 | 0.233 | 0.964 | 0.068 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 0 * model\_\_n\_estimators: 48 * model\_\_max\_depth: 9 * model\_\_min\_samples\_split: 10 * model\_\_min\_samples\_leaf: 2 * model\_\_max\_features: sqrt * model\_\_class\_weight: balanced | * option: build model without resampling * number of folds: 5 * average size of training set: 392.0 * average prop of minority class: 0.385 | 0.95 | 0.235 | 0.918 | 0.076 | 0.964 | 0.241 | 0.952 | 0.067 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 1 * model\_\_n\_estimators: 77 * model\_\_max\_depth: 5 * model\_\_min\_samples\_split: 10 * model\_\_min\_samples\_leaf: 2 * model\_\_max\_features: sqrt * model\_\_class\_weight: balanced\_subsample | * option: build model without resampling * number of folds: 5 * average size of training set: 392.0 * average prop of minority class: 0.385 | 0.953 | 0.23 | 0.923 | 0.073 | 0.972 | 0.235 | 0.956 | 0.067 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 1 * model\_\_n\_estimators: 84 * model\_\_max\_depth: 9 * model\_\_min\_samples\_split: 7 * model\_\_min\_samples\_leaf: 1 * model\_\_max\_features: auto * model\_\_class\_weight: balanced\_subsample | * option: build model without resampling * number of folds: 5 * average size of training set: 542.8 * average prop of minority class: 0.444 | 0.95 | 0.233 | 0.92 | 0.076 | 0.967 | 0.241 | 0.957 | 0.069 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 0 * model\_\_n\_estimators: 96 * model\_\_max\_depth: 10 * model\_\_min\_samples\_split: 13 * model\_\_min\_samples\_leaf: 4 * model\_\_max\_features: auto * model\_\_class\_weight: balanced\_subsample | * option: build model without resampling * number of folds: 5 * average size of training set: 392.0 * average prop of minority class: 0.385 | 0.949 | 0.238 | 0.916 | 0.075 | 0.964 | 0.247 | 0.952 | 0.07 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 0 * model\_\_n\_estimators: 70 * model\_\_max\_depth: 10 * model\_\_min\_samples\_split: 16 * model\_\_min\_samples\_leaf: 2 * model\_\_max\_features: sqrt * model\_\_class\_weight: balanced\_subsample | * option: build model without resampling * number of folds: 5 * average size of training set: 542.8 * average prop of minority class: 0.444 | 0.948 | 0.238 | 0.913 | 0.076 | 0.978 | 0.236 | 0.971 | 0.069 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 1 * model\_\_n\_estimators: 67 * model\_\_max\_depth: 6 * 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: 5 * average size of training set: 542.8 * average prop of minority class: 0.444 | 0.952 | 0.234 | 0.92 | 0.075 | 0.973 | 0.235 | 0.964 | 0.069 |