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 [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',  
   <utils.pipes\_and\_transformers.OptionedPostProcessTransformer object at 0x0000021101E9B6A0>),  
   ('resample', SMOTE())])
* Search space:
* undersampling\_majority\_class: Categorical(categories=(True, False), prior=None)
* 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: Pipeline(steps=[('scale', StandardScaler())])
* 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)

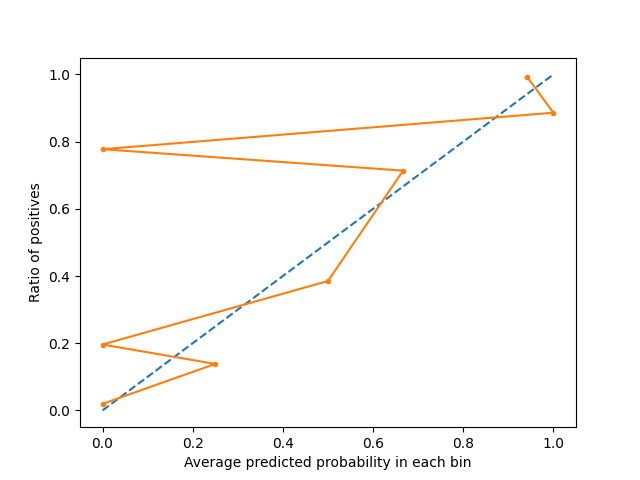
# Report of validation of the model in the outer Cross Validation

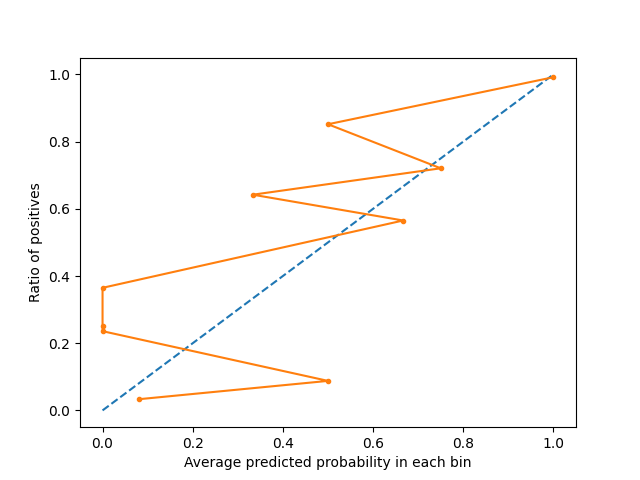
## Winner models of each fold and main metrics

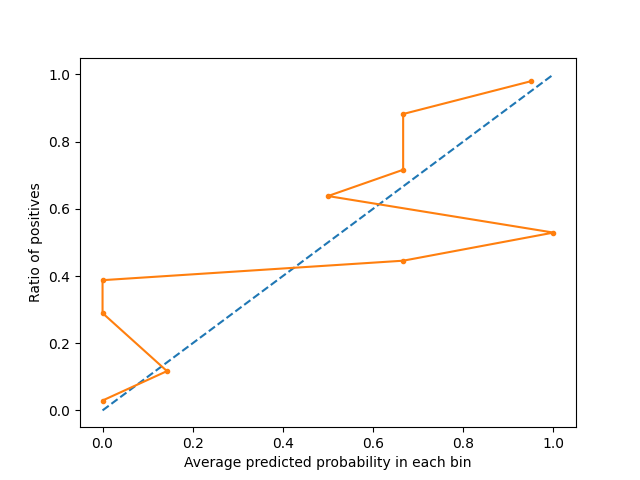
|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Fold** | **Model** | **Params** | **Comments** | **roc\_auc** | **neg\_log\_loss** | **average\_precision** | **neg\_brier\_score** |
| 1 | RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 0 * model\_\_n\_estimators: 72 * model\_\_max\_depth: 9 * model\_\_min\_samples\_split: 11 * 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 | 0.959 | 0.264 | 0.947 | 0.078 |
| 3 | XGBClassifier | * undersampling\_majority\_class: True * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_3 * model\_\_max\_depth: 14 * model\_\_learning\_rate: 0.113 * model\_\_min\_child\_weight: 7 * model\_\_subsample: 0.876 * model\_\_colsample\_bytree: 0.535 * model\_\_scale\_pos\_weight: 1.827 * 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 | 0.369 | 0.921 | 0.118 |
| 5 | XGBClassifier | * undersampling\_majority\_class: True * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_3 * model\_\_max\_depth: 7 * model\_\_learning\_rate: 0.234 * model\_\_min\_child\_weight: 7 * model\_\_subsample: 0.844 * model\_\_colsample\_bytree: 0.615 * model\_\_scale\_pos\_weight: 3.644 * 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 | 0.279 | 0.948 | 0.09 |
| 7 | RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 1 * model\_\_n\_estimators: 78 * model\_\_max\_depth: 6 * model\_\_min\_samples\_split: 9 * 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 | 0.968 | 0.239 | 0.967 | 0.072 |
| 9 | XGBClassifier | * undersampling\_majority\_class: False * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_3 * model\_\_max\_depth: 14 * model\_\_learning\_rate: 0.116 * model\_\_min\_child\_weight: 8 * model\_\_subsample: 0.817 * model\_\_colsample\_bytree: 0.178 * model\_\_scale\_pos\_weight: 2.216 * 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 | 0.973 | 0.231 | 0.972 | 0.069 |

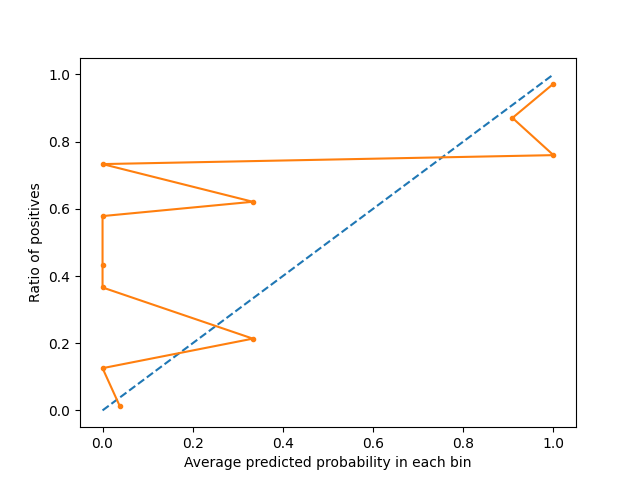
## Main plots

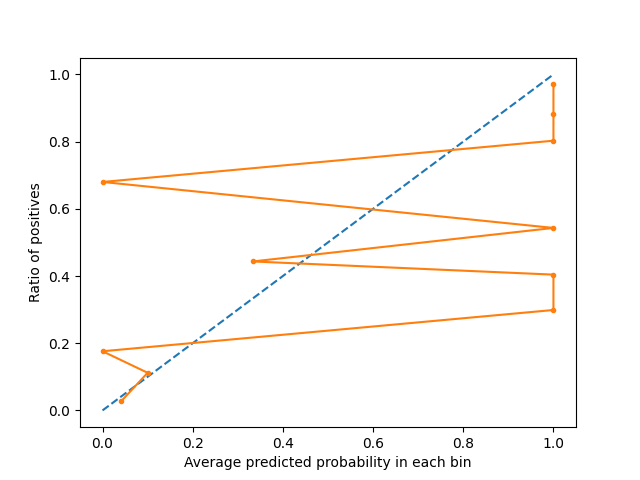
### Calibration plots



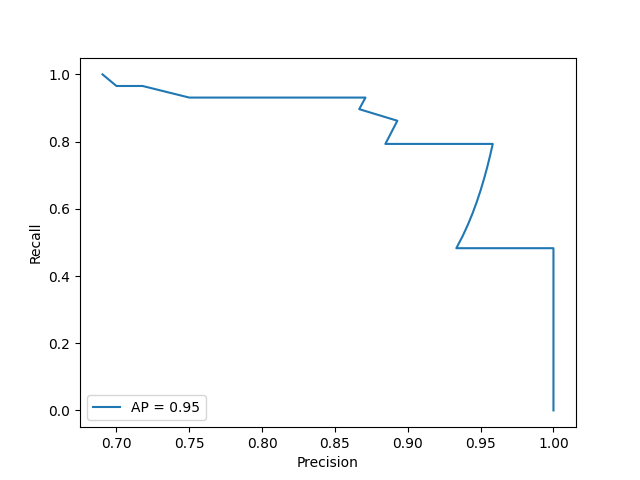


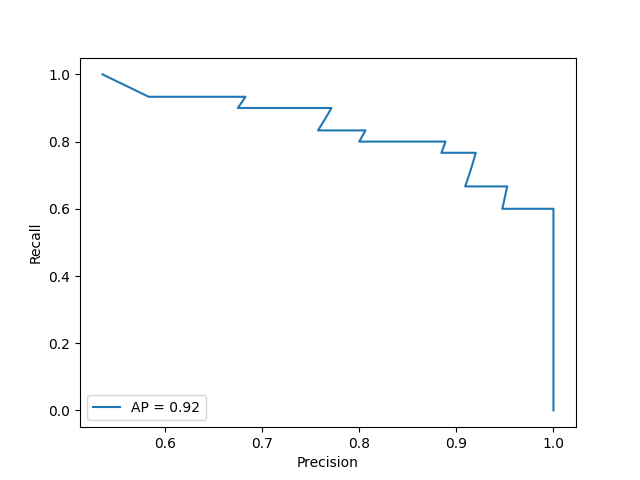


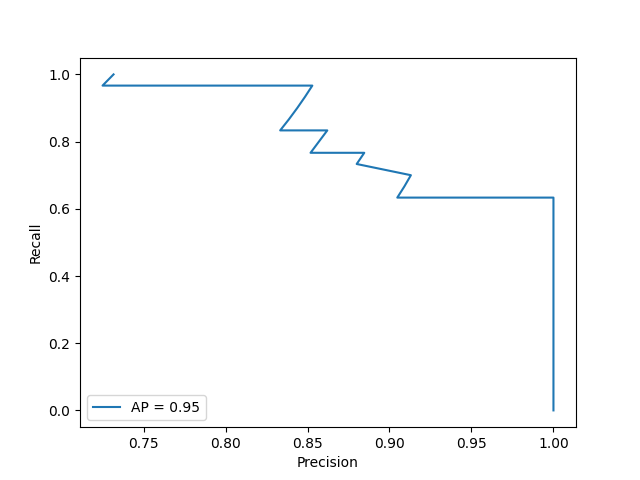


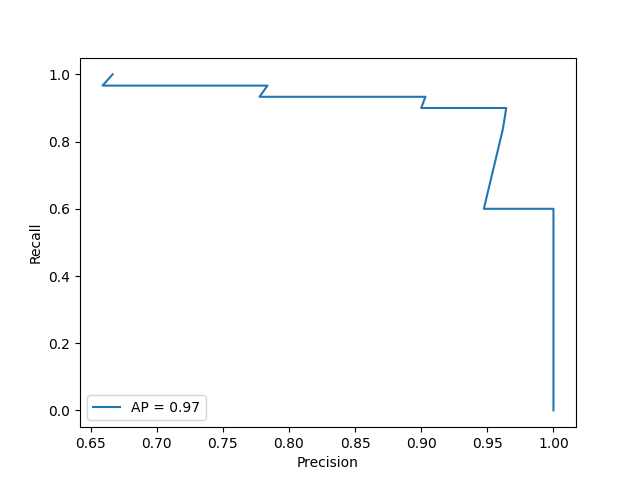


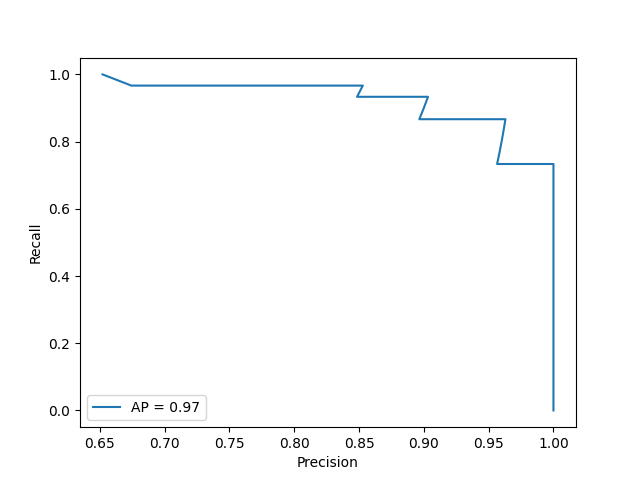
### Precision-recall curve plots



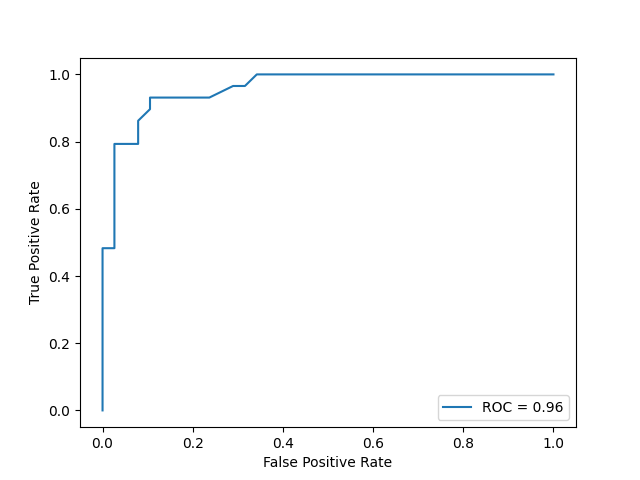


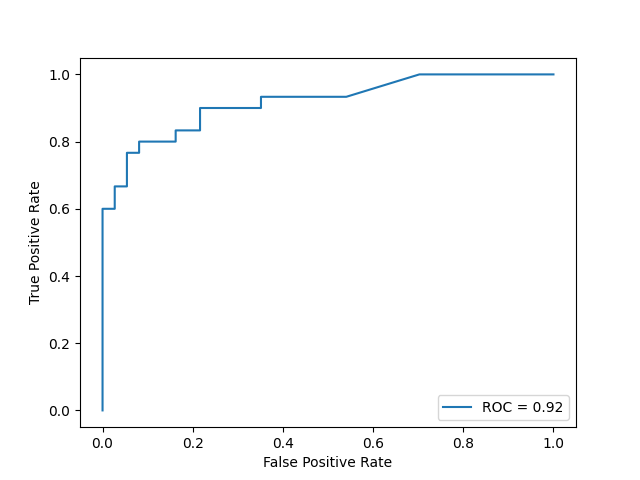


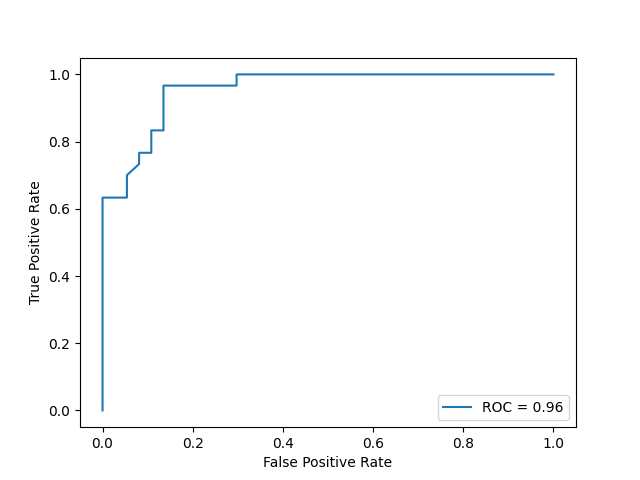


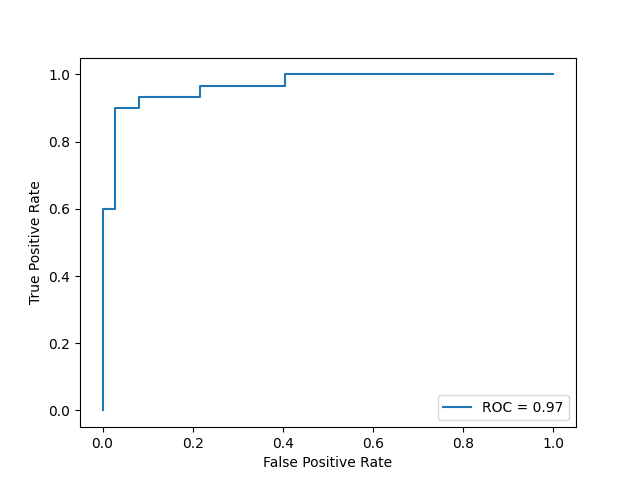


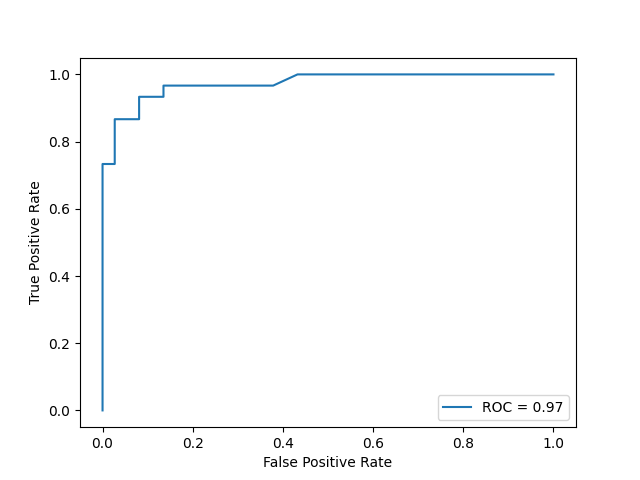
### ROC curve plots



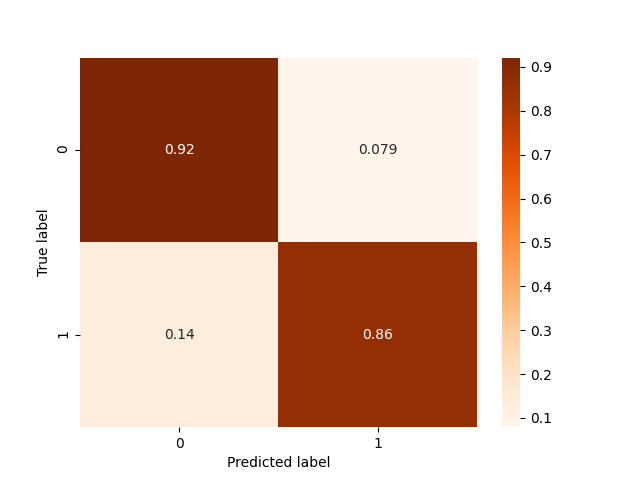


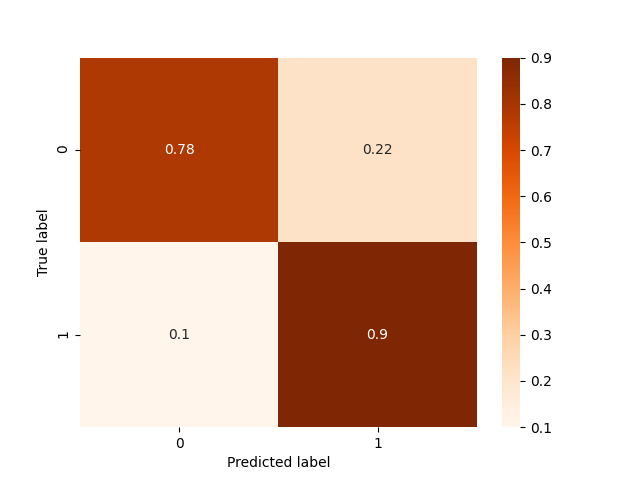


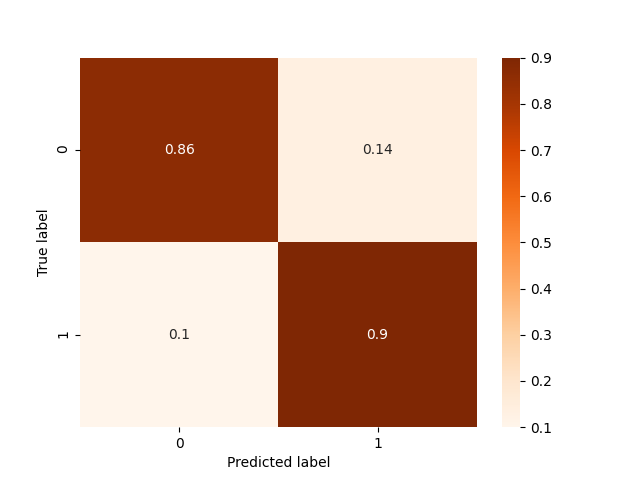


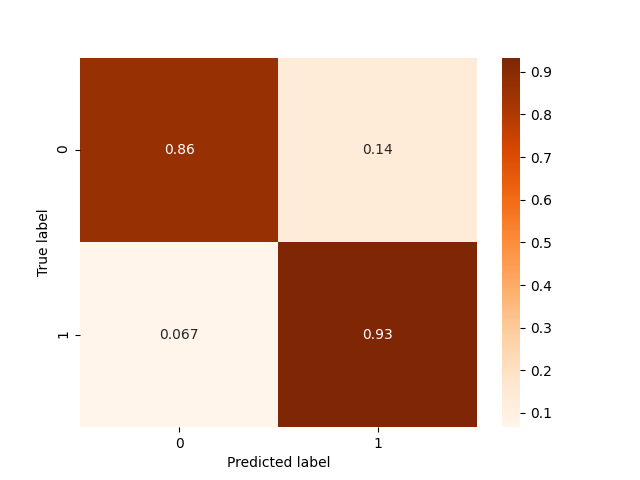


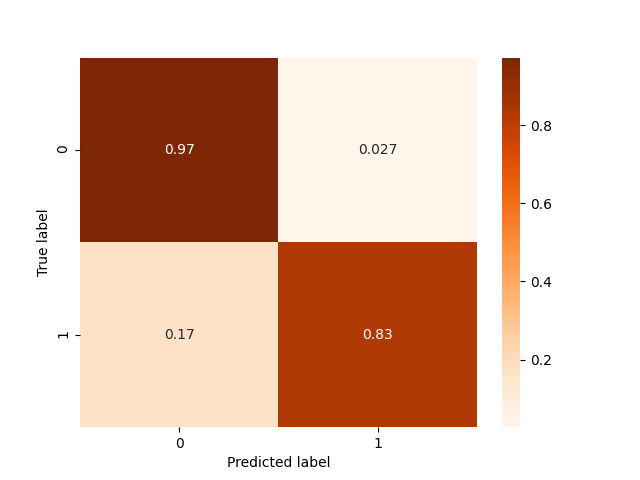
### Confusion matrix



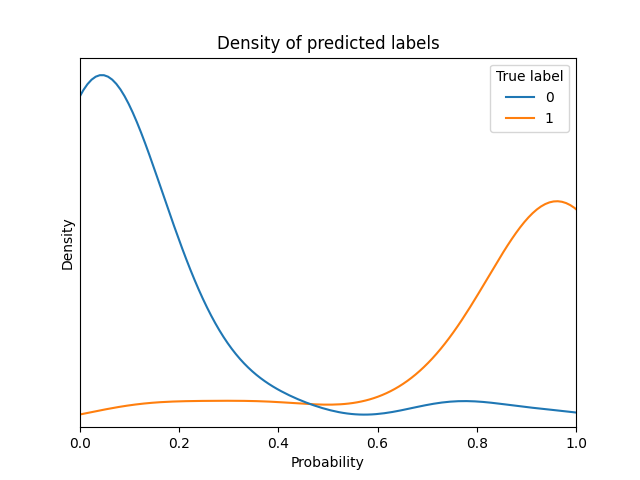


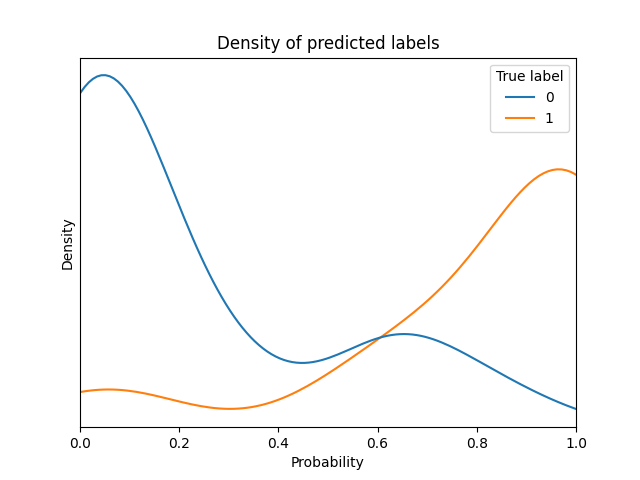


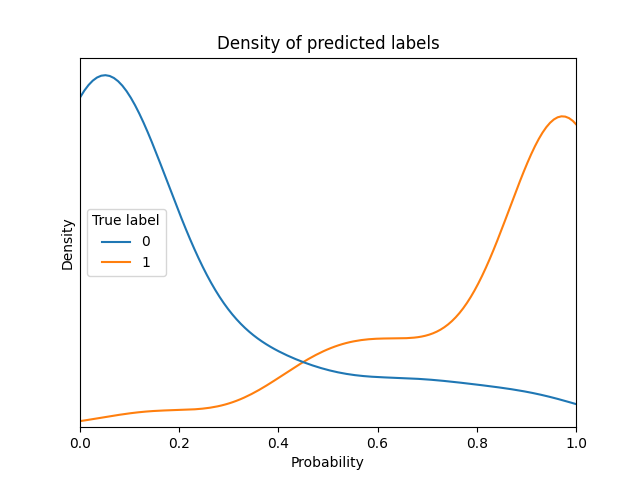


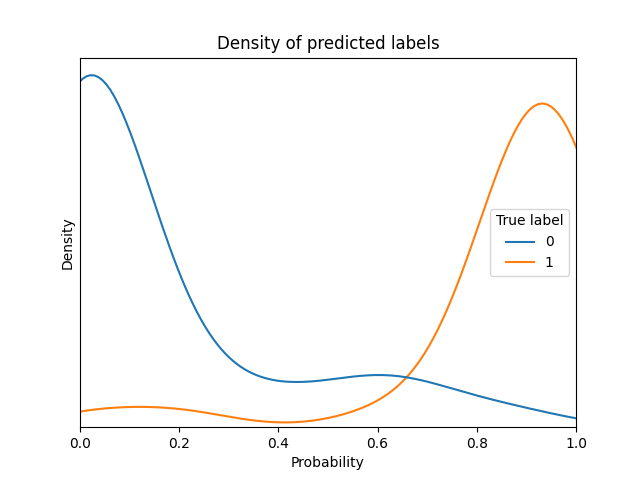


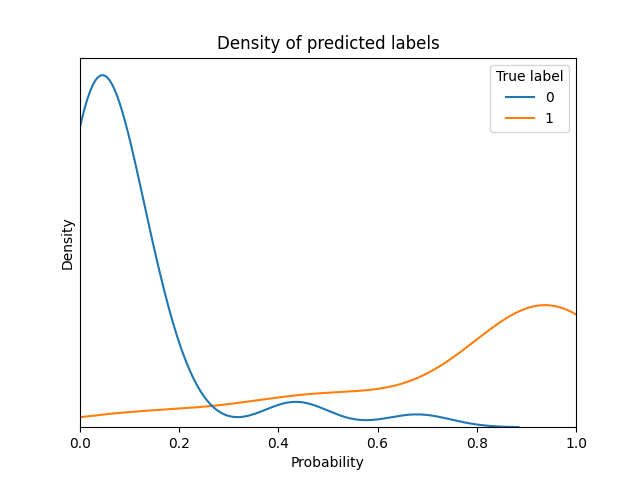
### Histograms











## 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 | 0.084 | 0.08 | 0.071 | 0.014 | 0.024 | 0.029 |
| 2 | 1 | 1.0 | 1.0 | 0.987 | 0.957 | 0.955 | 0.02 |
| 3 | 1 | 0.159 | 0.115 | 0.138 | 0.014 | 0.111 | 0.05 |
| 4 | 0 | 0.184 | 0.038 | 0.013 | 0.062 | 0.024 | 0.062 |
| 5 | 0 | 0.21 | 0.08 | 0.082 | 0.232 | 0.459 | 0.138 |
| 6 | 0 | 0.172 | 0.31 | 0.082 | 0.292 | 0.038 | 0.109 |
| 7 | 1 | 1.0 | 1.0 | 0.987 | 0.956 | 0.991 | 0.016 |
| 8 | 1 | 1.0 | 1.0 | 0.987 | 0.991 | 0.991 | 0.005 |
| 9 | 1 | 1.0 | 1.0 | 0.987 | 0.991 | 0.955 | 0.017 |
| 10 | 0 | 0.026 | 0.138 | 0.111 | 0.033 | 0.138 | 0.05 |
| 11 | 0 | 0.027 | 0.058 | 0.037 | 0.014 | 0.024 | 0.015 |
| 12 | 1 | 1.0 | 1.0 | 0.987 | 0.991 | 0.97 | 0.011 |
| 13 | 1 | 1.0 | 1.0 | 0.987 | 0.987 | 0.991 | 0.006 |
| 14 | 0 | 0.148 | 0.058 | 0.063 | 0.087 | 0.307 | 0.093 |
| 15 | 0 | 0.657 | 0.549 | 0.233 | 0.56 | 0.318 | 0.16 |
| 16 | 0 | 0.24 | 0.08 | 0.071 | 0.309 | 0.058 | 0.103 |
| 17 | 0 | 0.215 | 0.083 | 0.036 | 0.205 | 0.098 | 0.071 |
| 18 | 1 | 0.74 | 0.56 | 0.635 | 0.569 | 0.446 | 0.096 |
| 19 | 0 | 0.768 | 0.682 | 0.635 | 0.865 | 0.68 | 0.082 |
| 20 | 1 | 1.0 | 1.0 | 0.987 | 0.991 | 0.991 | 0.005 |

The average standard deviation is 0.057

# 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: 0
* model\_\_n\_estimators: 72
* model\_\_max\_depth: 9
* model\_\_min\_samples\_split: 11
* model\_\_min\_samples\_leaf: 3
* model\_\_max\_features: sqrt
* model\_\_class\_weight: balanced\_subsample

### 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 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_2 * model\_\_max\_depth: 10 * model\_\_learning\_rate: 0.119 * model\_\_min\_child\_weight: 6 * model\_\_subsample: 0.94 * model\_\_colsample\_bytree: 0.721 * model\_\_scale\_pos\_weight: 9.095 * 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 | 0.939 | 0.268 | 0.916 | 0.087 | 0.959 | 0.272 | 0.949 | 0.085 |
| XGBClassifier | * undersampling\_majority\_class: False * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_3 * model\_\_max\_depth: 8 * model\_\_learning\_rate: 0.081 * model\_\_min\_child\_weight: 8 * model\_\_subsample: 0.889 * model\_\_colsample\_bytree: 0.18 * model\_\_scale\_pos\_weight: 7.435 * 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 | 0.961 | 0.206 | 0.946 | 0.066 | 0.959 | 0.725 | 0.917 | 0.078 |
| XGBClassifier | * undersampling\_majority\_class: True * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_2 * model\_\_max\_depth: 11 * model\_\_learning\_rate: 0.119 * model\_\_min\_child\_weight: 9 * model\_\_subsample: 0.931 * model\_\_colsample\_bytree: 0.6 * model\_\_scale\_pos\_weight: 0.395 * 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 | 0.951 | 0.232 | 0.931 | 0.074 | 0.944 | 0.299 | 0.933 | 0.091 |
| XGBClassifier | * undersampling\_majority\_class: False * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_2 * model\_\_max\_depth: 10 * model\_\_learning\_rate: 0.099 * model\_\_min\_child\_weight: 8 * model\_\_subsample: 0.846 * model\_\_colsample\_bytree: 0.157 * model\_\_scale\_pos\_weight: 5.131 * 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 | 0.24 | 0.928 | 0.077 | 0.956 | 0.268 | 0.948 | 0.081 |
| XGBClassifier | * undersampling\_majority\_class: False * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_2 * model\_\_max\_depth: 10 * model\_\_learning\_rate: 0.147 * model\_\_min\_child\_weight: 9 * model\_\_subsample: 0.991 * model\_\_colsample\_bytree: 0.227 * model\_\_scale\_pos\_weight: 0.296 * 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 | 0.952 | 0.235 | 0.936 | 0.074 | 0.957 | 0.273 | 0.954 | 0.084 |
| XGBClassifier | * undersampling\_majority\_class: True * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_1 * model\_\_max\_depth: 11 * model\_\_learning\_rate: 0.179 * model\_\_min\_child\_weight: 8 * model\_\_subsample: 0.843 * model\_\_colsample\_bytree: 0.643 * model\_\_scale\_pos\_weight: 0.666 * 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 | 0.933 | 0.29 | 0.899 | 0.095 | 0.918 | 0.354 | 0.903 | 0.107 |
| XGBClassifier | * undersampling\_majority\_class: False * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_2 * model\_\_max\_depth: 13 * model\_\_learning\_rate: 0.247 * model\_\_min\_child\_weight: 10 * model\_\_subsample: 0.867 * model\_\_colsample\_bytree: 0.26 * model\_\_scale\_pos\_weight: 6.072 * 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 | 0.945 | 0.25 | 0.924 | 0.08 | 0.948 | 0.298 | 0.939 | 0.091 |
| XGBClassifier | * undersampling\_majority\_class: True * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_1 * model\_\_max\_depth: 9 * model\_\_learning\_rate: 0.117 * model\_\_min\_child\_weight: 3 * model\_\_subsample: 0.895 * model\_\_colsample\_bytree: 0.688 * model\_\_scale\_pos\_weight: 0.251 * 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 | 0.258 | 0.916 | 0.086 | 0.918 | 0.364 | 0.905 | 0.107 |
| XGBClassifier | * undersampling\_majority\_class: True * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_1 * model\_\_max\_depth: 13 * model\_\_learning\_rate: 0.223 * model\_\_min\_child\_weight: 7 * model\_\_subsample: 0.855 * model\_\_colsample\_bytree: 0.186 * model\_\_scale\_pos\_weight: 2.782 * 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 | 0.921 | 0.313 | 0.884 | 0.103 | 0.906 | 0.395 | 0.884 | 0.105 |
| XGBClassifier | * undersampling\_majority\_class: False * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_2 * model\_\_max\_depth: 12 * model\_\_learning\_rate: 0.195 * model\_\_min\_child\_weight: 4 * model\_\_subsample: 0.867 * model\_\_colsample\_bytree: 0.422 * model\_\_scale\_pos\_weight: 0.3 * 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 | 0.946 | 0.248 | 0.927 | 0.079 | 0.95 | 0.282 | 0.941 | 0.082 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 1 * model\_\_n\_estimators: 67 * model\_\_max\_depth: 7 * 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: 542.8 * average prop of minority class: 0.446 | 0.968 | 0.185 | 0.949 | 0.062 | 0.956 | 0.715 | 0.918 | 0.078 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 0 * model\_\_n\_estimators: 44 * model\_\_max\_depth: 3 * 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: 392.5 * average prop of minority class: 0.383 | 0.957 | 0.217 | 0.937 | 0.071 | 0.934 | 0.745 | 0.901 | 0.082 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 0 * model\_\_n\_estimators: 95 * model\_\_max\_depth: 6 * 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: 5 * average size of training set: 392.5 * average prop of minority class: 0.383 | 0.964 | 0.194 | 0.947 | 0.064 | 0.954 | 0.712 | 0.906 | 0.075 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 1 * model\_\_n\_estimators: 15 * 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: 5 * average size of training set: 542.8 * average prop of minority class: 0.446 | 0.966 | 0.197 | 0.946 | 0.064 | 0.954 | 0.264 | 0.936 | 0.073 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 1 * model\_\_n\_estimators: 15 * model\_\_max\_depth: 5 * model\_\_min\_samples\_split: 19 * 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: 0.446 | 0.963 | 0.202 | 0.945 | 0.066 | 0.941 | 0.754 | 0.898 | 0.083 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 0 * model\_\_n\_estimators: 47 * model\_\_max\_depth: 9 * model\_\_min\_samples\_split: 18 * 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: 0.446 | 0.966 | 0.187 | 0.949 | 0.061 | 0.955 | 0.718 | 0.915 | 0.076 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 0 * model\_\_n\_estimators: 72 * model\_\_max\_depth: 9 * model\_\_min\_samples\_split: 11 * 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 | 0.969 | 0.182 | 0.953 | 0.059 | 0.959 | 0.264 | 0.947 | 0.078 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 0 * model\_\_n\_estimators: 99 * model\_\_max\_depth: 8 * model\_\_min\_samples\_split: 7 * 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: 392.5 * average prop of minority class: 0.383 | 0.964 | 0.197 | 0.947 | 0.064 | 0.964 | 0.26 | 0.951 | 0.075 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 0 * model\_\_n\_estimators: 36 * model\_\_max\_depth: 4 * 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: 5 * average size of training set: 542.8 * average prop of minority class: 0.446 | 0.962 | 0.207 | 0.944 | 0.069 | 0.949 | 0.723 | 0.912 | 0.076 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 0 * model\_\_n\_estimators: 16 * model\_\_max\_depth: 7 * 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: 5 * average size of training set: 392.5 * average prop of minority class: 0.383 | 0.968 | 0.189 | 0.948 | 0.062 | 0.958 | 0.257 | 0.946 | 0.076 |

# 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
* resample\_\_sampling\_strategy: all
* post\_process\_\_option: option\_3
* model\_\_max\_depth: 14
* model\_\_learning\_rate: 0.113
* model\_\_min\_child\_weight: 7
* model\_\_subsample: 0.876
* model\_\_colsample\_bytree: 0.535
* model\_\_scale\_pos\_weight: 1.827
* 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 * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_2 * model\_\_max\_depth: 13 * model\_\_learning\_rate: 0.097 * model\_\_min\_child\_weight: 3 * model\_\_subsample: 0.826 * model\_\_colsample\_bytree: 0.271 * model\_\_scale\_pos\_weight: 0.589 * 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 | 0.233 | 0.932 | 0.072 | 0.898 | 0.439 | 0.898 | 0.147 |
| XGBClassifier | * undersampling\_majority\_class: False * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_2 * model\_\_max\_depth: 8 * model\_\_learning\_rate: 0.179 * model\_\_min\_child\_weight: 10 * model\_\_subsample: 0.979 * model\_\_colsample\_bytree: 0.538 * model\_\_scale\_pos\_weight: 0.363 * 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 | 0.949 | 0.241 | 0.929 | 0.076 | 0.9 | 0.395 | 0.911 | 0.131 |
| XGBClassifier | * undersampling\_majority\_class: True * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_3 * model\_\_max\_depth: 14 * model\_\_learning\_rate: 0.113 * model\_\_min\_child\_weight: 7 * model\_\_subsample: 0.876 * model\_\_colsample\_bytree: 0.535 * model\_\_scale\_pos\_weight: 1.827 * 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 | 0.187 | 0.954 | 0.06 | 0.917 | 0.369 | 0.921 | 0.118 |
| XGBClassifier | * undersampling\_majority\_class: False * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_2 * model\_\_max\_depth: 7 * model\_\_learning\_rate: 0.301 * model\_\_min\_child\_weight: 10 * model\_\_subsample: 0.859 * model\_\_colsample\_bytree: 0.387 * model\_\_scale\_pos\_weight: 0.175 * 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 | 0.947 | 0.25 | 0.923 | 0.076 | 0.897 | 0.406 | 0.9 | 0.135 |
| XGBClassifier | * undersampling\_majority\_class: True * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_2 * model\_\_max\_depth: 13 * model\_\_learning\_rate: 0.099 * model\_\_min\_child\_weight: 5 * model\_\_subsample: 0.973 * model\_\_colsample\_bytree: 0.432 * model\_\_scale\_pos\_weight: 0.299 * 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 | 0.957 | 0.224 | 0.941 | 0.069 | 0.907 | 0.418 | 0.904 | 0.139 |
| XGBClassifier | * undersampling\_majority\_class: False * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_2 * model\_\_max\_depth: 12 * model\_\_learning\_rate: 0.084 * model\_\_min\_child\_weight: 8 * model\_\_subsample: 0.863 * model\_\_colsample\_bytree: 0.24 * model\_\_scale\_pos\_weight: 0.38 * 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 | 0.233 | 0.933 | 0.073 | 0.901 | 0.414 | 0.904 | 0.139 |
| XGBClassifier | * undersampling\_majority\_class: True * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_1 * model\_\_max\_depth: 6 * model\_\_learning\_rate: 0.219 * model\_\_min\_child\_weight: 6 * model\_\_subsample: 0.904 * model\_\_colsample\_bytree: 0.379 * model\_\_scale\_pos\_weight: 3.139 * 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 | 0.923 | 0.315 | 0.891 | 0.1 | 0.899 | 0.438 | 0.871 | 0.142 |
| XGBClassifier | * undersampling\_majority\_class: True * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_2 * model\_\_max\_depth: 14 * model\_\_learning\_rate: 0.214 * model\_\_min\_child\_weight: 5 * model\_\_subsample: 0.873 * model\_\_colsample\_bytree: 0.142 * model\_\_scale\_pos\_weight: 7.906 * 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 | 0.951 | 0.242 | 0.931 | 0.077 | 0.886 | 0.442 | 0.888 | 0.152 |
| XGBClassifier | * undersampling\_majority\_class: True * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_3 * model\_\_max\_depth: 6 * model\_\_learning\_rate: 0.117 * model\_\_min\_child\_weight: 3 * model\_\_subsample: 0.967 * model\_\_colsample\_bytree: 0.727 * model\_\_scale\_pos\_weight: 0.504 * 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 | 0.964 | 0.192 | 0.948 | 0.06 | 0.919 | 0.368 | 0.919 | 0.117 |
| XGBClassifier | * undersampling\_majority\_class: True * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_1 * model\_\_max\_depth: 14 * model\_\_learning\_rate: 0.126 * model\_\_min\_child\_weight: 8 * model\_\_subsample: 0.806 * model\_\_colsample\_bytree: 0.2 * model\_\_scale\_pos\_weight: 6.335 * 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 | 0.927 | 0.306 | 0.892 | 0.099 | 0.891 | 0.435 | 0.884 | 0.146 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 1 * model\_\_n\_estimators: 92 * model\_\_max\_depth: 4 * model\_\_min\_samples\_split: 9 * 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 | 0.971 | 0.181 | 0.951 | 0.058 | 0.915 | 0.379 | 0.91 | 0.127 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 1 * model\_\_n\_estimators: 79 * model\_\_max\_depth: 5 * model\_\_min\_samples\_split: 15 * model\_\_min\_samples\_leaf: 4 * 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 | 0.968 | 0.191 | 0.947 | 0.061 | 0.918 | 0.384 | 0.918 | 0.129 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 0 * model\_\_n\_estimators: 90 * model\_\_max\_depth: 4 * model\_\_min\_samples\_split: 16 * 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: 542.8 * average prop of minority class: 0.444 | 0.963 | 0.21 | 0.944 | 0.067 | 0.904 | 0.415 | 0.905 | 0.137 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 1 * model\_\_n\_estimators: 31 * model\_\_max\_depth: 9 * model\_\_min\_samples\_split: 17 * 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: 0.444 | 0.965 | 0.196 | 0.946 | 0.063 | 0.921 | 0.363 | 0.918 | 0.121 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 1 * model\_\_n\_estimators: 14 * model\_\_max\_depth: 9 * model\_\_min\_samples\_split: 18 * 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 | 0.965 | 0.204 | 0.94 | 0.065 | 0.916 | 0.386 | 0.904 | 0.13 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 1 * model\_\_n\_estimators: 74 * model\_\_max\_depth: 3 * 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: 392.0 * average prop of minority class: 0.385 | 0.962 | 0.214 | 0.941 | 0.068 | 0.912 | 0.409 | 0.909 | 0.135 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 0 * model\_\_n\_estimators: 55 * model\_\_max\_depth: 8 * model\_\_min\_samples\_split: 12 * 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: 392.0 * average prop of minority class: 0.385 | 0.967 | 0.185 | 0.949 | 0.059 | 0.923 | 0.378 | 0.924 | 0.13 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 0 * model\_\_n\_estimators: 60 * model\_\_max\_depth: 7 * model\_\_min\_samples\_split: 6 * 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 | 0.966 | 0.195 | 0.946 | 0.063 | 0.923 | 0.348 | 0.924 | 0.119 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 0 * model\_\_n\_estimators: 90 * model\_\_max\_depth: 4 * 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 | 0.965 | 0.202 | 0.946 | 0.065 | 0.912 | 0.412 | 0.905 | 0.134 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 0 * model\_\_n\_estimators: 89 * model\_\_max\_depth: 8 * 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 | 0.969 | 0.186 | 0.951 | 0.06 | 0.91 | 0.407 | 0.909 | 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 XGBClassifier with the following params:

* undersampling\_majority\_class: True
* resample\_\_sampling\_strategy: all
* post\_process\_\_option: option\_3
* model\_\_max\_depth: 7
* model\_\_learning\_rate: 0.234
* model\_\_min\_child\_weight: 7
* model\_\_subsample: 0.844
* model\_\_colsample\_bytree: 0.615
* model\_\_scale\_pos\_weight: 3.644
* 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 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_3 * model\_\_max\_depth: 9 * model\_\_learning\_rate: 0.149 * model\_\_min\_child\_weight: 4 * model\_\_subsample: 0.905 * model\_\_colsample\_bytree: 0.441 * model\_\_scale\_pos\_weight: 1.699 * 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 | 0.968 | 0.172 | 0.944 | 0.055 | 0.941 | 0.313 | 0.919 | 0.095 |
| XGBClassifier | * undersampling\_majority\_class: True * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_3 * model\_\_max\_depth: 7 * model\_\_learning\_rate: 0.234 * model\_\_min\_child\_weight: 7 * model\_\_subsample: 0.844 * model\_\_colsample\_bytree: 0.615 * model\_\_scale\_pos\_weight: 3.644 * 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 | 0.177 | 0.948 | 0.056 | 0.956 | 0.279 | 0.948 | 0.09 |
| XGBClassifier | * undersampling\_majority\_class: True * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_2 * model\_\_max\_depth: 10 * model\_\_learning\_rate: 0.248 * model\_\_min\_child\_weight: 4 * model\_\_subsample: 0.816 * model\_\_colsample\_bytree: 0.582 * model\_\_scale\_pos\_weight: 0.635 * 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 | 0.961 | 0.207 | 0.934 | 0.066 | 0.95 | 0.269 | 0.921 | 0.083 |
| XGBClassifier | * undersampling\_majority\_class: True * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_1 * model\_\_max\_depth: 5 * model\_\_learning\_rate: 0.104 * model\_\_min\_child\_weight: 9 * model\_\_subsample: 0.955 * model\_\_colsample\_bytree: 0.763 * model\_\_scale\_pos\_weight: 0.407 * 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 | 0.929 | 0.3 | 0.893 | 0.098 | 0.92 | 0.361 | 0.918 | 0.109 |
| XGBClassifier | * undersampling\_majority\_class: True * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_1 * model\_\_max\_depth: 9 * model\_\_learning\_rate: 0.106 * model\_\_min\_child\_weight: 10 * model\_\_subsample: 0.833 * model\_\_colsample\_bytree: 0.177 * model\_\_scale\_pos\_weight: 2.575 * 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 | 0.925 | 0.305 | 0.887 | 0.1 | 0.925 | 0.338 | 0.926 | 0.103 |
| XGBClassifier | * undersampling\_majority\_class: False * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_3 * model\_\_max\_depth: 14 * model\_\_learning\_rate: 0.207 * model\_\_min\_child\_weight: 5 * model\_\_subsample: 0.824 * model\_\_colsample\_bytree: 0.137 * model\_\_scale\_pos\_weight: 1.988 * 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 | 0.963 | 0.214 | 0.937 | 0.069 | 0.945 | 0.299 | 0.942 | 0.098 |
| XGBClassifier | * undersampling\_majority\_class: False * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_1 * model\_\_max\_depth: 8 * model\_\_learning\_rate: 0.07 * model\_\_min\_child\_weight: 2 * model\_\_subsample: 0.849 * model\_\_colsample\_bytree: 0.611 * model\_\_scale\_pos\_weight: 6.902 * 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 | 0.915 | 0.327 | 0.878 | 0.106 | 0.912 | 0.372 | 0.9 | 0.114 |
| XGBClassifier | * undersampling\_majority\_class: True * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_2 * model\_\_max\_depth: 15 * model\_\_learning\_rate: 0.08 * model\_\_min\_child\_weight: 4 * model\_\_subsample: 0.817 * model\_\_colsample\_bytree: 0.251 * model\_\_scale\_pos\_weight: 0.574 * 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 | 0.95 | 0.235 | 0.919 | 0.077 | 0.959 | 0.258 | 0.95 | 0.082 |
| XGBClassifier | * undersampling\_majority\_class: False * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_1 * model\_\_max\_depth: 8 * model\_\_learning\_rate: 0.057 * model\_\_min\_child\_weight: 5 * model\_\_subsample: 0.991 * model\_\_colsample\_bytree: 0.235 * model\_\_scale\_pos\_weight: 0.124 * 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 | 0.928 | 0.3 | 0.893 | 0.096 | 0.89 | 0.864 | 0.903 | 0.122 |
| XGBClassifier | * undersampling\_majority\_class: True * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_2 * model\_\_max\_depth: 6 * model\_\_learning\_rate: 0.188 * model\_\_min\_child\_weight: 3 * model\_\_subsample: 0.869 * model\_\_colsample\_bytree: 0.17 * model\_\_scale\_pos\_weight: 0.313 * 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 | 0.239 | 0.925 | 0.078 | 0.95 | 0.279 | 0.934 | 0.087 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 0 * model\_\_n\_estimators: 93 * model\_\_max\_depth: 5 * model\_\_min\_samples\_split: 6 * 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: 542.8 * average prop of minority class: 0.444 | 0.963 | 0.19 | 0.936 | 0.061 | 0.951 | 0.28 | 0.937 | 0.084 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 1 * model\_\_n\_estimators: 39 * model\_\_max\_depth: 2 * 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: 5 * average size of training set: 542.8 * average prop of minority class: 0.444 | 0.948 | 0.239 | 0.914 | 0.076 | 0.954 | 0.286 | 0.94 | 0.088 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 1 * model\_\_n\_estimators: 98 * model\_\_max\_depth: 8 * model\_\_min\_samples\_split: 12 * 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: 542.8 * average prop of minority class: 0.444 | 0.964 | 0.193 | 0.935 | 0.063 | 0.953 | 0.284 | 0.933 | 0.09 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 0 * model\_\_n\_estimators: 83 * model\_\_max\_depth: 7 * model\_\_min\_samples\_split: 11 * model\_\_min\_samples\_leaf: 4 * 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 | 0.184 | 0.943 | 0.061 | 0.952 | 0.728 | 0.904 | 0.084 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 0 * model\_\_n\_estimators: 88 * model\_\_max\_depth: 9 * 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: 392.0 * average prop of minority class: 0.385 | 0.963 | 0.195 | 0.935 | 0.065 | 0.95 | 0.286 | 0.921 | 0.087 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 1 * model\_\_n\_estimators: 71 * model\_\_max\_depth: 6 * 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: 5 * average size of training set: 392.0 * average prop of minority class: 0.385 | 0.965 | 0.189 | 0.94 | 0.062 | 0.953 | 0.256 | 0.938 | 0.08 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 1 * model\_\_n\_estimators: 85 * model\_\_max\_depth: 10 * model\_\_min\_samples\_split: 18 * 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 | 0.964 | 0.194 | 0.939 | 0.065 | 0.955 | 0.27 | 0.939 | 0.085 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 1 * model\_\_n\_estimators: 14 * model\_\_max\_depth: 4 * 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: 5 * average size of training set: 392.0 * average prop of minority class: 0.385 | 0.964 | 0.195 | 0.936 | 0.063 | 0.949 | 0.292 | 0.94 | 0.093 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 1 * model\_\_n\_estimators: 65 * model\_\_max\_depth: 3 * 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: 5 * average size of training set: 542.8 * average prop of minority class: 0.444 | 0.956 | 0.223 | 0.926 | 0.072 | 0.95 | 0.295 | 0.936 | 0.094 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 1 * model\_\_n\_estimators: 57 * model\_\_max\_depth: 3 * 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: 542.8 * average prop of minority class: 0.444 | 0.956 | 0.222 | 0.927 | 0.072 | 0.949 | 0.286 | 0.938 | 0.09 |

# 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: True
* model\_\_bootstrap: 1
* model\_\_n\_estimators: 78
* model\_\_max\_depth: 6
* model\_\_min\_samples\_split: 9
* model\_\_min\_samples\_leaf: 3
* model\_\_max\_features: auto
* model\_\_class\_weight: balanced\_subsample

### 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 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_2 * model\_\_max\_depth: 13 * model\_\_learning\_rate: 0.23 * model\_\_min\_child\_weight: 8 * model\_\_subsample: 0.907 * model\_\_colsample\_bytree: 0.558 * model\_\_scale\_pos\_weight: 0.236 * 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 | 0.947 | 0.247 | 0.913 | 0.079 | 0.947 | 0.742 | 0.951 | 0.087 |
| XGBClassifier | * undersampling\_majority\_class: False * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_1 * model\_\_max\_depth: 8 * model\_\_learning\_rate: 0.235 * model\_\_min\_child\_weight: 6 * model\_\_subsample: 0.956 * model\_\_colsample\_bytree: 0.297 * model\_\_scale\_pos\_weight: 6.857 * 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 | 0.861 | 0.424 | 0.799 | 0.142 | 0.932 | 0.385 | 0.934 | 0.114 |
| XGBClassifier | * undersampling\_majority\_class: True * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_2 * model\_\_max\_depth: 15 * model\_\_learning\_rate: 0.146 * model\_\_min\_child\_weight: 2 * model\_\_subsample: 0.908 * model\_\_colsample\_bytree: 0.409 * model\_\_scale\_pos\_weight: 1.87 * 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 | 0.95 | 0.227 | 0.914 | 0.076 | 0.938 | 0.334 | 0.937 | 0.089 |
| XGBClassifier | * undersampling\_majority\_class: False * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_2 * model\_\_max\_depth: 12 * model\_\_learning\_rate: 0.134 * model\_\_min\_child\_weight: 2 * model\_\_subsample: 0.815 * model\_\_colsample\_bytree: 0.557 * model\_\_scale\_pos\_weight: 2.22 * 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 | 0.938 | 0.265 | 0.898 | 0.089 | 0.933 | 0.781 | 0.937 | 0.09 |
| XGBClassifier | * undersampling\_majority\_class: True * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_1 * model\_\_max\_depth: 14 * model\_\_learning\_rate: 0.087 * model\_\_min\_child\_weight: 7 * model\_\_subsample: 0.834 * model\_\_colsample\_bytree: 0.491 * model\_\_scale\_pos\_weight: 1.498 * 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 | 0.919 | 0.327 | 0.871 | 0.104 | 0.948 | 0.303 | 0.943 | 0.095 |
| XGBClassifier | * undersampling\_majority\_class: False * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_1 * model\_\_max\_depth: 11 * model\_\_learning\_rate: 0.296 * model\_\_min\_child\_weight: 6 * model\_\_subsample: 0.818 * model\_\_colsample\_bytree: 0.337 * model\_\_scale\_pos\_weight: 2.644 * 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 | 0.881 | 0.391 | 0.828 | 0.129 | 0.936 | 0.352 | 0.939 | 0.104 |
| XGBClassifier | * undersampling\_majority\_class: False * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_3 * model\_\_max\_depth: 6 * model\_\_learning\_rate: 0.056 * model\_\_min\_child\_weight: 8 * model\_\_subsample: 0.98 * model\_\_colsample\_bytree: 0.326 * model\_\_scale\_pos\_weight: 0.295 * 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 | 0.953 | 0.229 | 0.922 | 0.073 | 0.955 | 0.709 | 0.957 | 0.074 |
| XGBClassifier | * undersampling\_majority\_class: False * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_1 * model\_\_max\_depth: 15 * model\_\_learning\_rate: 0.294 * model\_\_min\_child\_weight: 5 * model\_\_subsample: 0.912 * model\_\_colsample\_bytree: 0.296 * model\_\_scale\_pos\_weight: 2.493 * 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 | 0.869 | 0.412 | 0.808 | 0.138 | 0.94 | 0.366 | 0.942 | 0.107 |
| XGBClassifier | * undersampling\_majority\_class: False * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_2 * model\_\_max\_depth: 8 * model\_\_learning\_rate: 0.188 * model\_\_min\_child\_weight: 8 * model\_\_subsample: 0.951 * model\_\_colsample\_bytree: 0.357 * model\_\_scale\_pos\_weight: 1.309 * 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 | 0.24 | 0.912 | 0.078 | 0.95 | 0.734 | 0.955 | 0.084 |
| XGBClassifier | * undersampling\_majority\_class: False * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_2 * model\_\_max\_depth: 15 * model\_\_learning\_rate: 0.16 * model\_\_min\_child\_weight: 2 * model\_\_subsample: 0.894 * model\_\_colsample\_bytree: 0.37 * model\_\_scale\_pos\_weight: 0.258 * 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 | 0.256 | 0.905 | 0.082 | 0.95 | 0.732 | 0.956 | 0.084 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 0 * model\_\_n\_estimators: 23 * model\_\_max\_depth: 8 * model\_\_min\_samples\_split: 19 * 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 | 0.956 | 0.216 | 0.926 | 0.07 | 0.965 | 0.269 | 0.965 | 0.082 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 0 * model\_\_n\_estimators: 11 * model\_\_max\_depth: 5 * model\_\_min\_samples\_split: 11 * 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 | 0.958 | 0.215 | 0.924 | 0.068 | 0.97 | 0.241 | 0.967 | 0.077 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 1 * model\_\_n\_estimators: 78 * model\_\_max\_depth: 6 * model\_\_min\_samples\_split: 9 * 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 | 0.96 | 0.212 | 0.932 | 0.068 | 0.968 | 0.239 | 0.967 | 0.072 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 0 * model\_\_n\_estimators: 16 * model\_\_max\_depth: 10 * 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 | 0.947 | 0.234 | 0.911 | 0.075 | 0.966 | 0.255 | 0.96 | 0.082 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 1 * model\_\_n\_estimators: 46 * model\_\_max\_depth: 9 * 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: 5 * average size of training set: 392.0 * average prop of minority class: 0.385 | 0.959 | 0.204 | 0.928 | 0.065 | 0.968 | 0.254 | 0.961 | 0.08 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 1 * model\_\_n\_estimators: 52 * model\_\_max\_depth: 9 * model\_\_min\_samples\_split: 18 * 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 | 0.959 | 0.203 | 0.928 | 0.064 | 0.968 | 0.261 | 0.959 | 0.08 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 1 * model\_\_n\_estimators: 16 * model\_\_max\_depth: 7 * 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: 542.8 * average prop of minority class: 0.444 | 0.956 | 0.222 | 0.925 | 0.071 | 0.97 | 0.245 | 0.966 | 0.078 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 1 * model\_\_n\_estimators: 74 * model\_\_max\_depth: 3 * model\_\_min\_samples\_split: 20 * 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 | 0.953 | 0.225 | 0.919 | 0.07 | 0.954 | 0.273 | 0.947 | 0.083 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 0 * model\_\_n\_estimators: 66 * model\_\_max\_depth: 2 * model\_\_min\_samples\_split: 14 * 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: 392.0 * average prop of minority class: 0.385 | 0.947 | 0.244 | 0.913 | 0.076 | 0.934 | 0.755 | 0.936 | 0.092 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 1 * model\_\_n\_estimators: 69 * model\_\_max\_depth: 8 * 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 | 0.956 | 0.217 | 0.926 | 0.07 | 0.972 | 0.231 | 0.969 | 0.073 |

# 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
* resample\_\_sampling\_strategy: all
* post\_process\_\_option: option\_3
* model\_\_max\_depth: 14
* model\_\_learning\_rate: 0.116
* model\_\_min\_child\_weight: 8
* model\_\_subsample: 0.817
* model\_\_colsample\_bytree: 0.178
* model\_\_scale\_pos\_weight: 2.216
* 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 * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_2 * model\_\_max\_depth: 9 * model\_\_learning\_rate: 0.102 * model\_\_min\_child\_weight: 8 * model\_\_subsample: 0.803 * model\_\_colsample\_bytree: 0.39 * model\_\_scale\_pos\_weight: 0.713 * 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 | 0.951 | 0.235 | 0.919 | 0.078 | 0.941 | 0.306 | 0.934 | 0.089 |
| XGBClassifier | * undersampling\_majority\_class: True * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_2 * model\_\_max\_depth: 6 * model\_\_learning\_rate: 0.273 * model\_\_min\_child\_weight: 6 * model\_\_subsample: 0.811 * model\_\_colsample\_bytree: 0.537 * model\_\_scale\_pos\_weight: 1.06 * 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 | 0.951 | 0.225 | 0.916 | 0.073 | 0.959 | 0.29 | 0.941 | 0.089 |
| XGBClassifier | * undersampling\_majority\_class: True * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_1 * model\_\_max\_depth: 11 * model\_\_learning\_rate: 0.055 * model\_\_min\_child\_weight: 4 * model\_\_subsample: 0.996 * model\_\_colsample\_bytree: 0.345 * model\_\_scale\_pos\_weight: 6.539 * 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 | 0.903 | 0.356 | 0.863 | 0.117 | 0.887 | 0.421 | 0.872 | 0.132 |
| XGBClassifier | * undersampling\_majority\_class: False * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_1 * model\_\_max\_depth: 14 * model\_\_learning\_rate: 0.164 * model\_\_min\_child\_weight: 1 * model\_\_subsample: 1.0 * model\_\_colsample\_bytree: 0.21 * model\_\_scale\_pos\_weight: 2.863 * 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 | 0.875 | 0.392 | 0.817 | 0.132 | 0.842 | 0.486 | 0.788 | 0.16 |
| XGBClassifier | * undersampling\_majority\_class: False * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_2 * model\_\_max\_depth: 6 * model\_\_learning\_rate: 0.184 * model\_\_min\_child\_weight: 8 * model\_\_subsample: 0.911 * model\_\_colsample\_bytree: 0.653 * model\_\_scale\_pos\_weight: 5.095 * 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 | 0.273 | 0.896 | 0.09 | 0.943 | 0.307 | 0.937 | 0.098 |
| XGBClassifier | * undersampling\_majority\_class: False * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_2 * model\_\_max\_depth: 14 * model\_\_learning\_rate: 0.059 * model\_\_min\_child\_weight: 7 * model\_\_subsample: 0.862 * model\_\_colsample\_bytree: 0.164 * model\_\_scale\_pos\_weight: 9.216 * 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 | 0.249 | 0.912 | 0.08 | 0.95 | 0.297 | 0.94 | 0.092 |
| XGBClassifier | * undersampling\_majority\_class: True * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_2 * model\_\_max\_depth: 9 * model\_\_learning\_rate: 0.128 * model\_\_min\_child\_weight: 3 * model\_\_subsample: 0.862 * model\_\_colsample\_bytree: 0.465 * model\_\_scale\_pos\_weight: 6.952 * 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 | 0.94 | 0.265 | 0.907 | 0.087 | 0.938 | 0.321 | 0.92 | 0.099 |
| XGBClassifier | * undersampling\_majority\_class: False * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_1 * model\_\_max\_depth: 9 * model\_\_learning\_rate: 0.062 * model\_\_min\_child\_weight: 6 * model\_\_subsample: 0.813 * model\_\_colsample\_bytree: 0.362 * model\_\_scale\_pos\_weight: 3.353 * 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 | 0.897 | 0.366 | 0.845 | 0.121 | 0.894 | 0.419 | 0.897 | 0.133 |
| XGBClassifier | * undersampling\_majority\_class: False * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_3 * model\_\_max\_depth: 8 * model\_\_learning\_rate: 0.117 * model\_\_min\_child\_weight: 5 * model\_\_subsample: 0.923 * model\_\_colsample\_bytree: 0.67 * model\_\_scale\_pos\_weight: 1.593 * 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 | 0.948 | 0.243 | 0.921 | 0.079 | 0.977 | 0.212 | 0.974 | 0.062 |
| XGBClassifier | * undersampling\_majority\_class: False * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_3 * model\_\_max\_depth: 14 * model\_\_learning\_rate: 0.116 * model\_\_min\_child\_weight: 8 * model\_\_subsample: 0.817 * model\_\_colsample\_bytree: 0.178 * model\_\_scale\_pos\_weight: 2.216 * 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 | 0.236 | 0.925 | 0.076 | 0.973 | 0.231 | 0.972 | 0.069 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 0 * model\_\_n\_estimators: 69 * model\_\_max\_depth: 4 * model\_\_min\_samples\_split: 8 * 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 | 0.951 | 0.231 | 0.918 | 0.073 | 0.969 | 0.244 | 0.963 | 0.073 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 0 * model\_\_n\_estimators: 88 * model\_\_max\_depth: 6 * model\_\_min\_samples\_split: 12 * 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 | 0.951 | 0.236 | 0.918 | 0.075 | 0.973 | 0.242 | 0.961 | 0.071 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 1 * model\_\_n\_estimators: 12 * model\_\_max\_depth: 2 * model\_\_min\_samples\_split: 5 * 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 | 0.939 | 0.272 | 0.904 | 0.087 | 0.94 | 0.313 | 0.936 | 0.093 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 0 * model\_\_n\_estimators: 59 * model\_\_max\_depth: 6 * model\_\_min\_samples\_split: 10 * 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 | 0.952 | 0.227 | 0.919 | 0.073 | 0.966 | 0.247 | 0.954 | 0.07 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 1 * model\_\_n\_estimators: 11 * model\_\_max\_depth: 8 * model\_\_min\_samples\_split: 19 * 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 | 0.95 | 0.234 | 0.916 | 0.075 | 0.971 | 0.235 | 0.967 | 0.07 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 1 * model\_\_n\_estimators: 56 * model\_\_max\_depth: 10 * 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: 5 * average size of training set: 392.0 * average prop of minority class: 0.385 | 0.953 | 0.229 | 0.922 | 0.073 | 0.976 | 0.236 | 0.959 | 0.067 |
| RandomForestClassifier | * undersampling\_majority\_class: True * model\_\_bootstrap: 0 * model\_\_n\_estimators: 24 * model\_\_max\_depth: 7 * 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: 392.0 * average prop of minority class: 0.385 | 0.949 | 0.237 | 0.914 | 0.076 | 0.963 | 0.249 | 0.956 | 0.071 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 0 * model\_\_n\_estimators: 60 * model\_\_max\_depth: 10 * model\_\_min\_samples\_split: 20 * 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 | 0.949 | 0.239 | 0.918 | 0.075 | 0.965 | 0.245 | 0.952 | 0.069 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 0 * model\_\_n\_estimators: 84 * model\_\_max\_depth: 5 * model\_\_min\_samples\_split: 5 * 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: 0.444 | 0.948 | 0.239 | 0.913 | 0.077 | 0.976 | 0.244 | 0.97 | 0.072 |
| RandomForestClassifier | * undersampling\_majority\_class: False * model\_\_bootstrap: 0 * model\_\_n\_estimators: 42 * model\_\_max\_depth: 5 * model\_\_min\_samples\_split: 12 * 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: 542.8 * average prop of minority class: 0.444 | 0.954 | 0.224 | 0.921 | 0.071 | 0.971 | 0.244 | 0.962 | 0.073 |