Report of training

# Introduction

Report of search and training made on March 25, 2021 at 00:18:35.

## 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 9 outer folds and 10 inner folds.
* Some of the folds will be skipped. In particular, [0, 2, 3, 4, 6, 8] outer folds and [0, 2, 4, 6, 8, 9] inner folds will be skipped.
* For each outer fold search, a model will be fitted. In order to search for the best hyperparameters, 10 initial points will be evaluated, and 10 additional calls will be made.
* Models will not be calibrated.
* A final model will be fitted using the whole inner dataset.
* The optimizing metric for the bayesian search is average\_precision.
* The function used for the bayesian search is gbrt\_minimize.
* Additionally, 20 instances will be left out for assessing the variance of all models.

The search spaces for the optimization is the following:

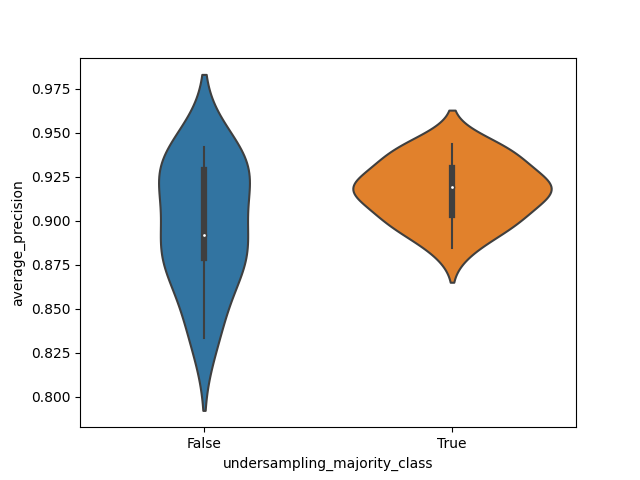
* Search space for xgboost model.
* model: XGBClassifier(base\_score=None, booster=None, colsample\_bylevel=None,  
   colsample\_bynode=None, colsample\_bytree=None, gamma=None,  
   gpu\_id=None, importance\_type='gain', interaction\_constraints=None,  
   learning\_rate=None, max\_delta\_step=None, max\_depth=None,  
   min\_child\_weight=None, missing=nan, monotone\_constraints=None,  
   n\_estimators=100, n\_jobs=None, num\_parallel\_tree=None,  
   random\_state=None, reg\_alpha=None, reg\_lambda=None,  
   scale\_pos\_weight=None, subsample=None, tree\_method=None,  
   validate\_parameters=None, verbosity=None)
* pipeline\_post\_process: Pipeline(steps=[('post\_process',  
   <nestedcvtraining.api.OptionedPostProcessTransformer object at 0x0000016C46351160>),  
   ('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

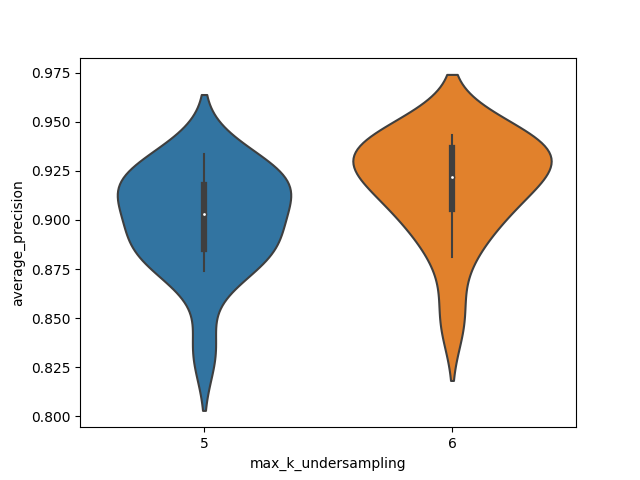
# Plots for all trained models and params with respect to loss metric on the inner folds

## Plots for xgboost model

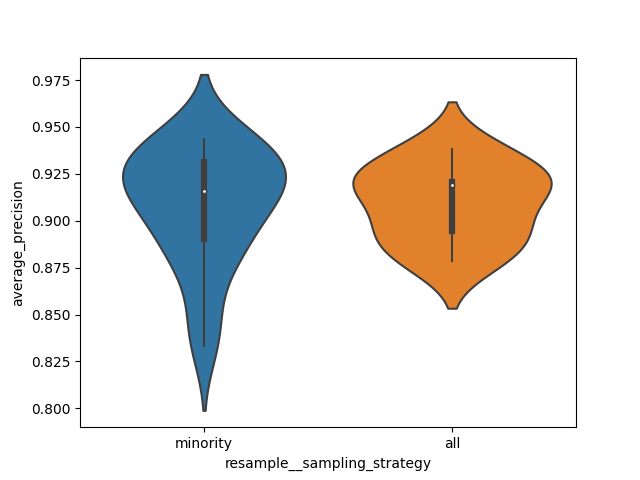
### Plots for undersampling\_majority\_class param



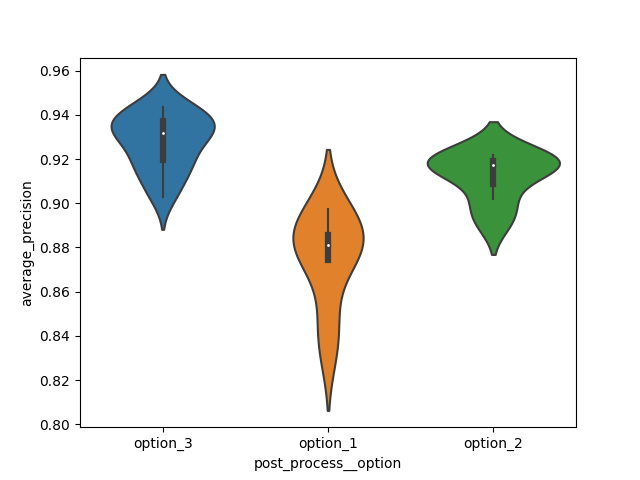
### Plots for max\_k\_undersampling param



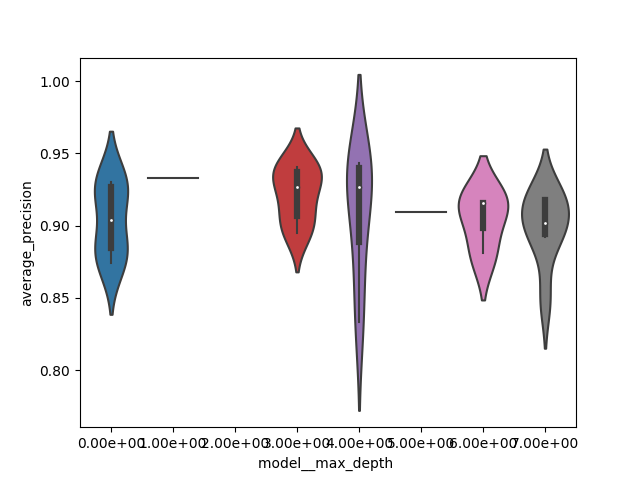
### Plots for resample\_\_sampling\_strategy param



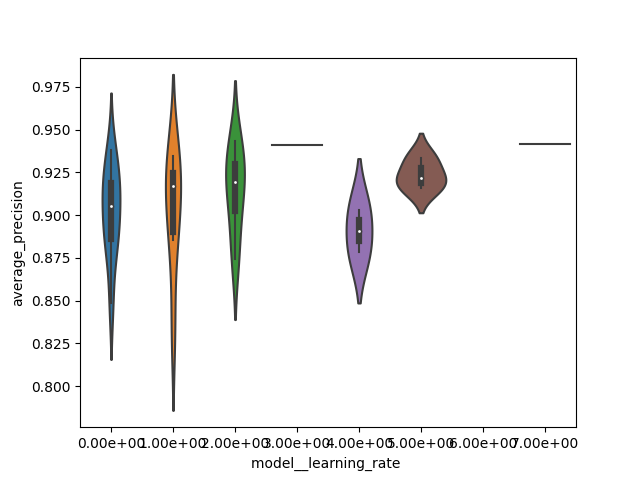
### Plots for post\_process\_\_option param



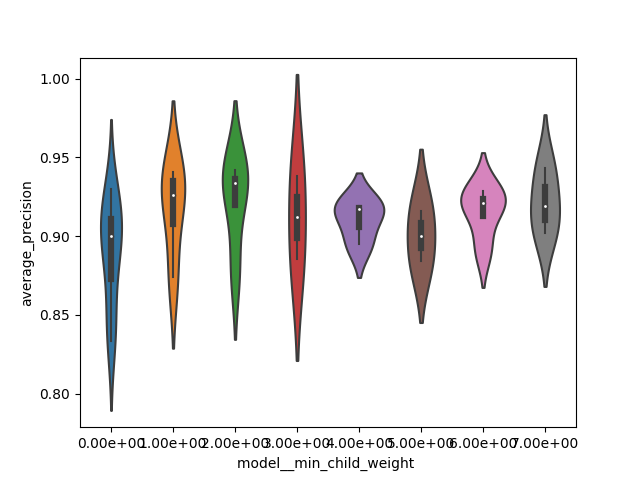
### Plots for model\_\_max\_depth param



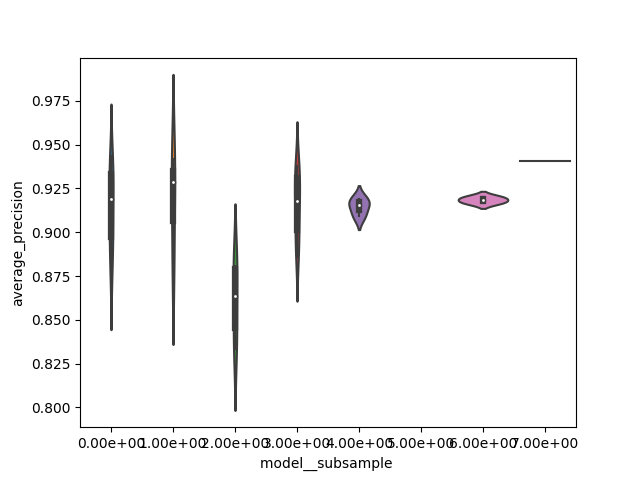
### Plots for model\_\_learning\_rate param



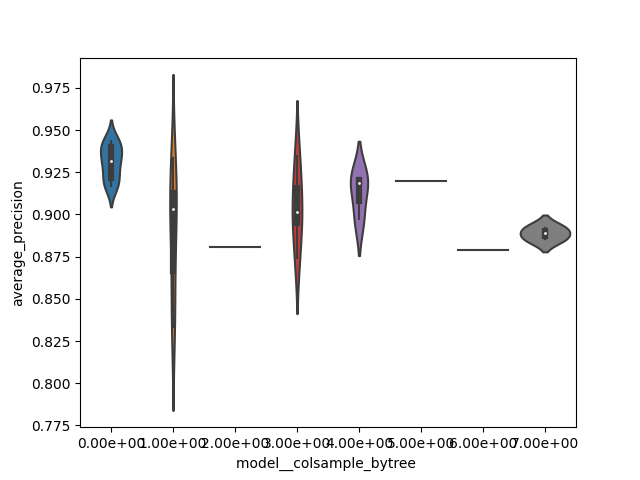
### Plots for model\_\_min\_child\_weight param



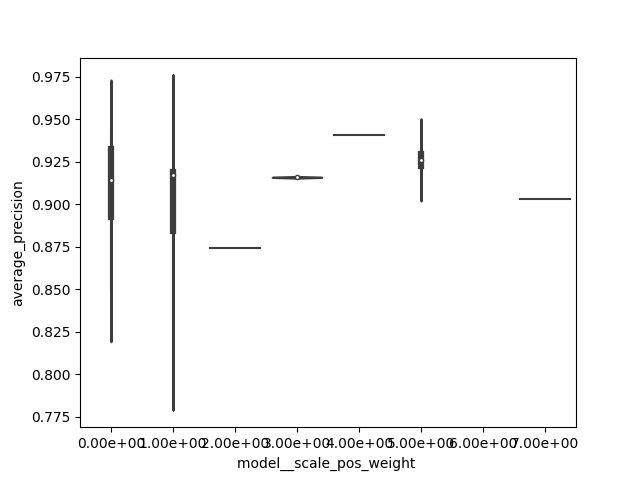
### Plots for model\_\_subsample param



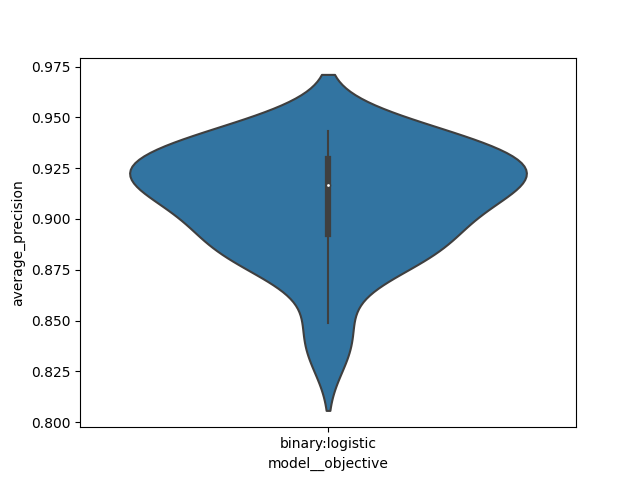
### Plots for model\_\_colsample\_bytree param



### Plots for model\_\_scale\_pos\_weight param

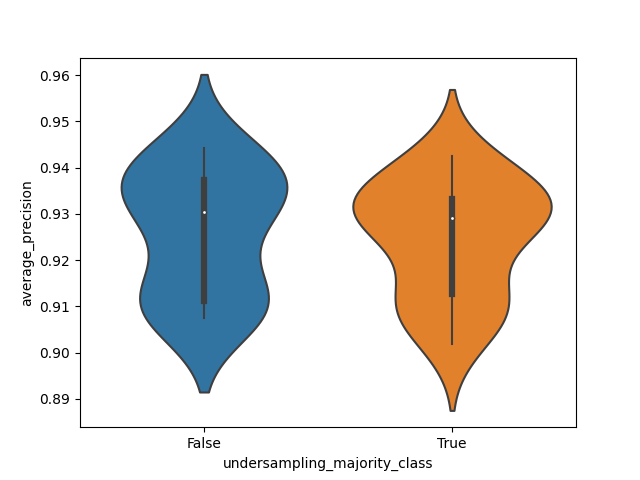


### Plots for model\_\_objective param

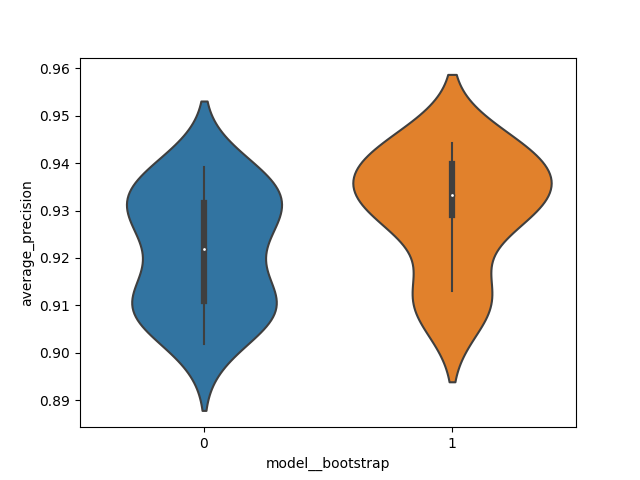


## Plots for random\_forest model

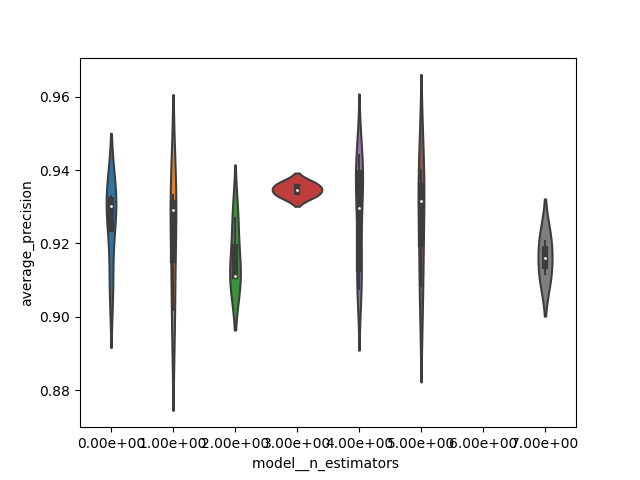
### Plots for undersampling\_majority\_class param



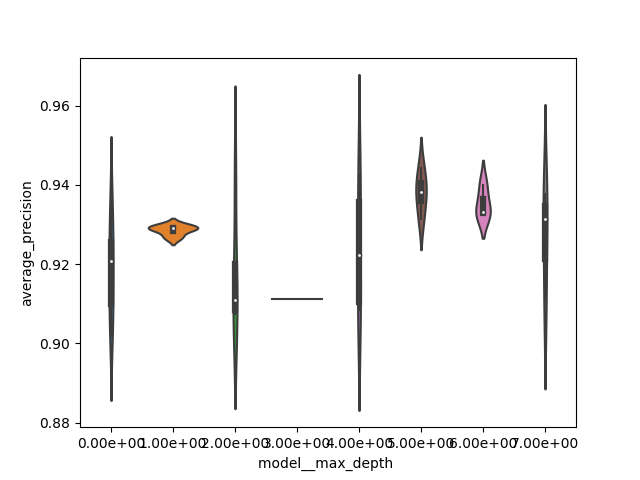
### Plots for model\_\_bootstrap param



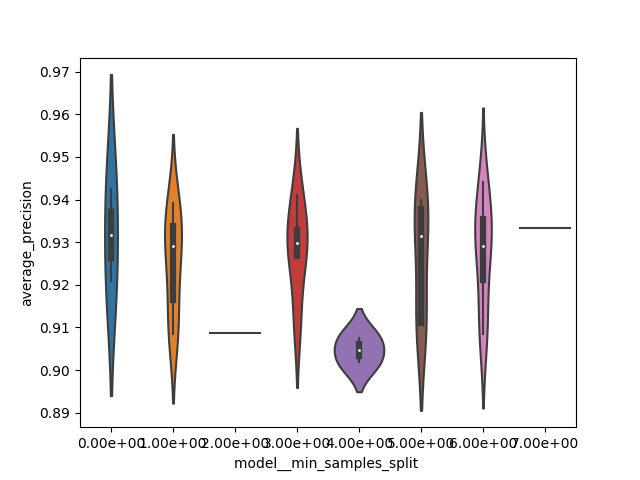
### Plots for model\_\_n\_estimators param



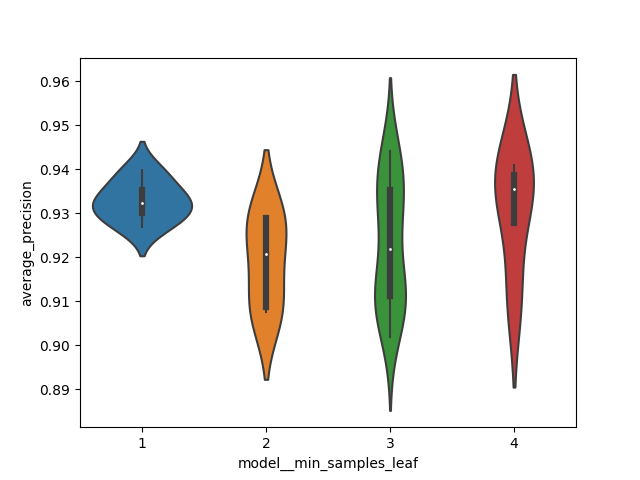
### Plots for model\_\_max\_depth param



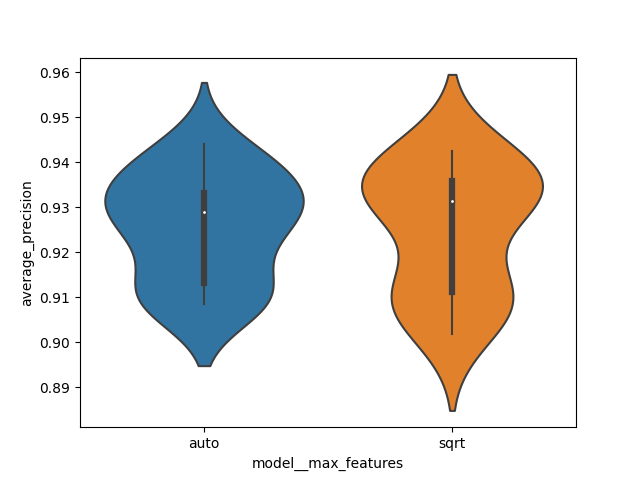
### Plots for model\_\_min\_samples\_split param



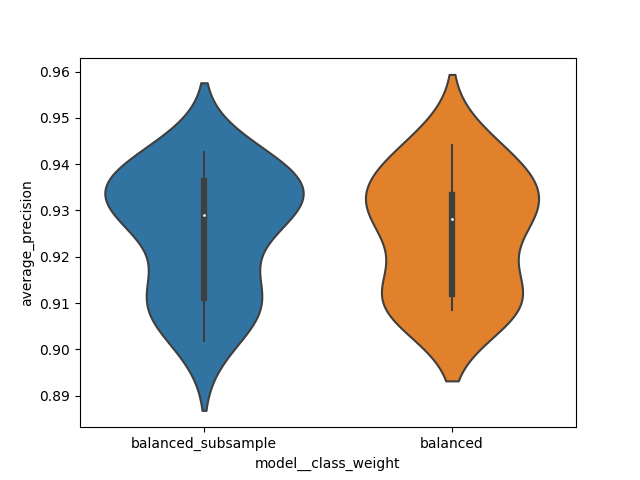
### Plots for model\_\_min\_samples\_leaf param



### Plots for model\_\_max\_features param



### Plots for model\_\_class\_weight param



## Winner models of each fold and main metrics

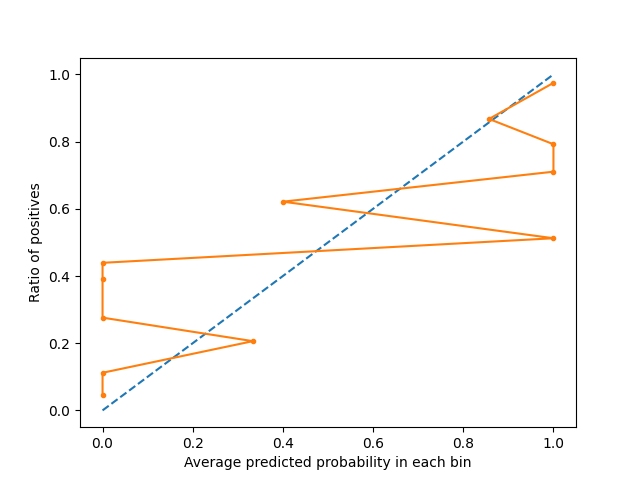
|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Fold** | **Model** | **Params** | **Comments** | **roc\_auc** | **neg\_log\_loss** | **average\_precision** | **neg\_brier\_score** |
| 1 | random\_forest | * undersampling\_majority\_class: False * model\_\_bootstrap: 1 * model\_\_n\_estimators: 57 * model\_\_max\_depth: 8 * model\_\_min\_samples\_split: 18 * model\_\_min\_samples\_leaf: 3 * model\_\_max\_features: auto * model\_\_class\_weight: balanced | * option: build model without resampling * number of folds: 4 * average size of training set: 535.5 * average prop of minority class: 0.445 | 0.954 | 0.304 | 0.941 | 0.088 |
| 5 | xgboost | * undersampling\_majority\_class: True * max\_k\_undersampling: 6 * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_3 * model\_\_max\_depth: 9 * model\_\_learning\_rate: 0.056 * model\_\_min\_child\_weight: 5 * model\_\_subsample: 0.882 * model\_\_colsample\_bytree: 0.149 * model\_\_scale\_pos\_weight: 0.835 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 387.25 * average prop of minority class before resampling: 0.385 * average size of training set after resampling: 476.5 * average prop of minority class after resampling: 0.5 | 0.956 | 0.315 | 0.962 | 0.088 |
| 7 | xgboost | * undersampling\_majority\_class: True * max\_k\_undersampling: 6 * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_2 * model\_\_max\_depth: 14 * model\_\_learning\_rate: 0.103 * model\_\_min\_child\_weight: 9 * model\_\_subsample: 0.809 * model\_\_colsample\_bytree: 0.152 * model\_\_scale\_pos\_weight: 5.624 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 387.25 * average prop of minority class before resampling: 0.385 * average size of training set after resampling: 476.5 * average prop of minority class after resampling: 0.5 | 0.933 | 0.536 | 0.938 | 0.177 |

For the selected optimization metric average\_precision the average score is 0.947, and the standard deviation is 0.011.

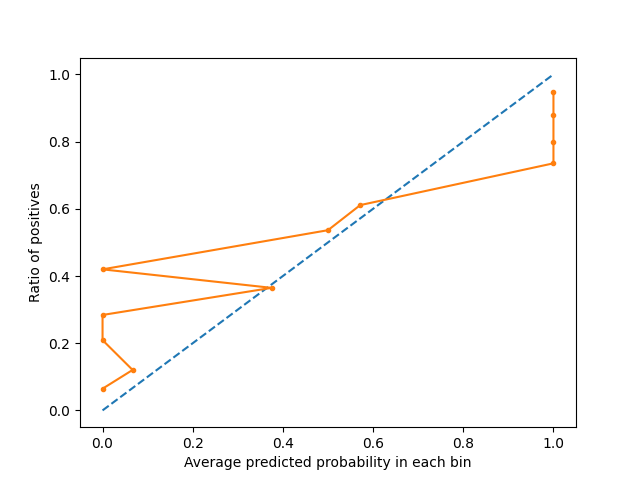
## Main plots

### Calibration plots

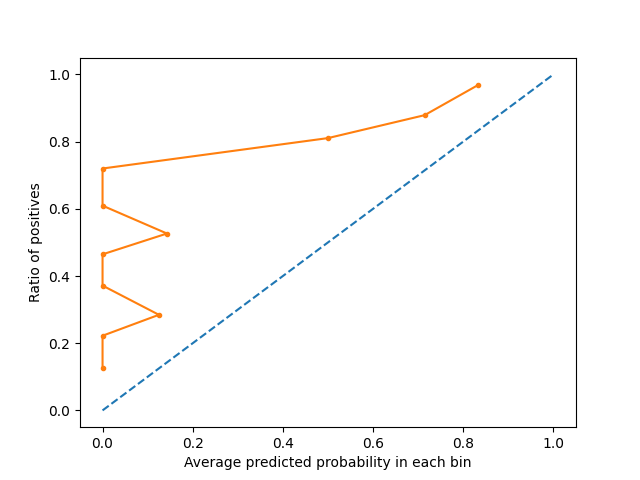
#### Calibration plot of fold 1



#### Calibration plot of fold 5

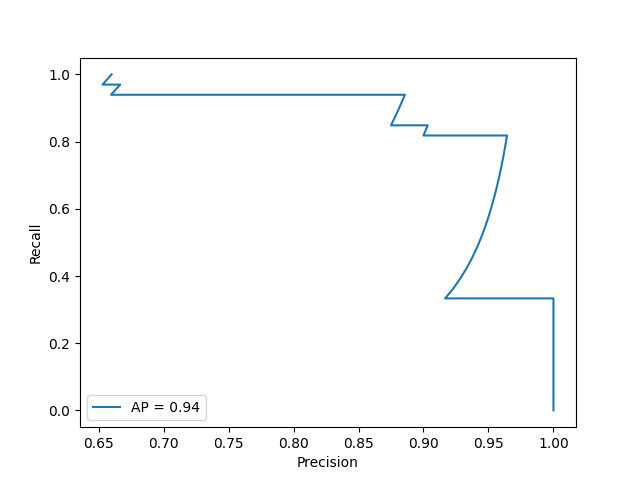


#### Calibration plot of fold 7

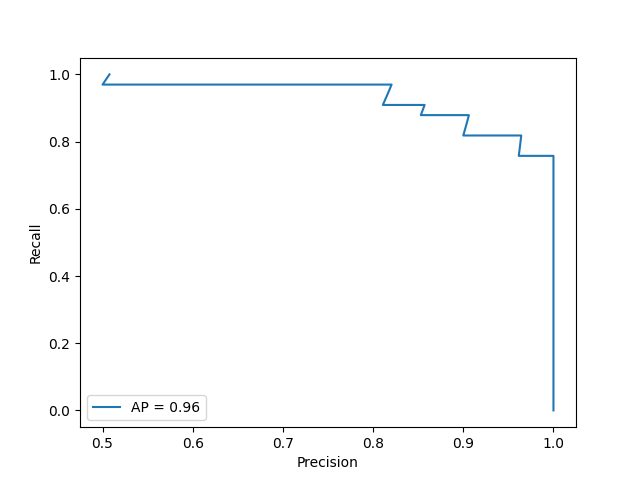


### Precision-recall curve plots

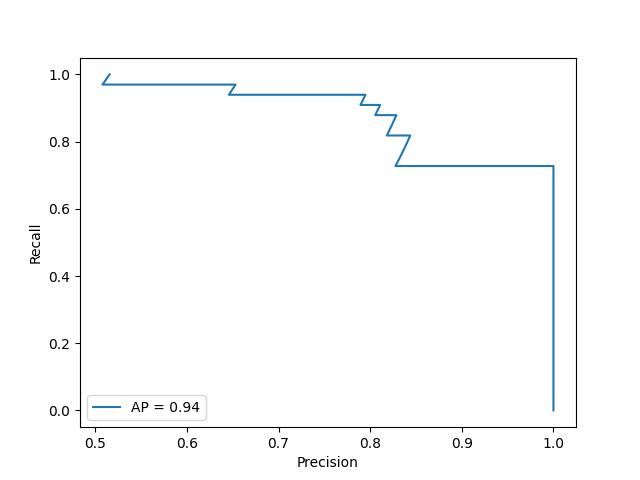
#### Precision-recall curve plot of fold 1



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

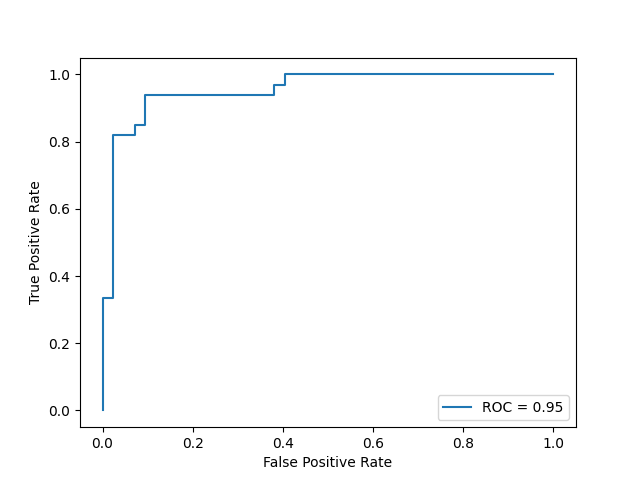


#### Precision-recall curve plot of fold 7

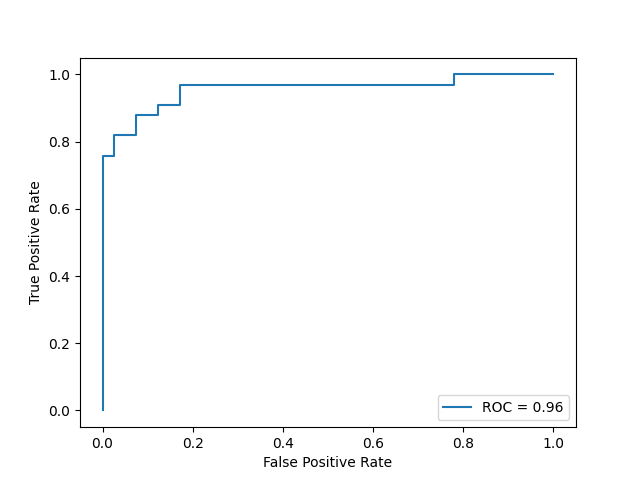


### ROC curve plots

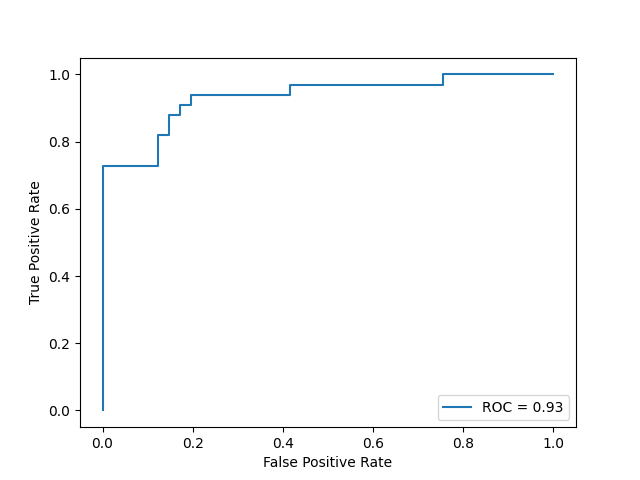
#### ROC curve plot of fold 1



#### ROC curve plot of fold 5

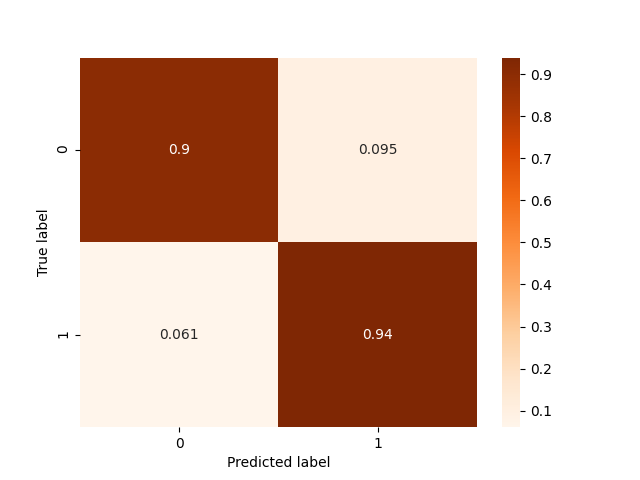


#### ROC curve plot of fold 7

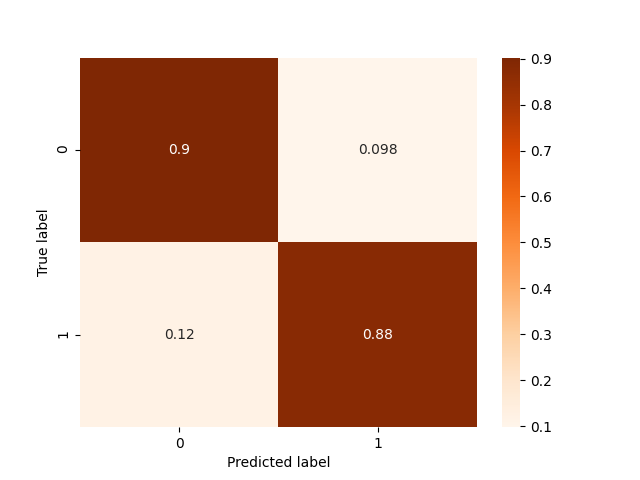


### Confusion matrix

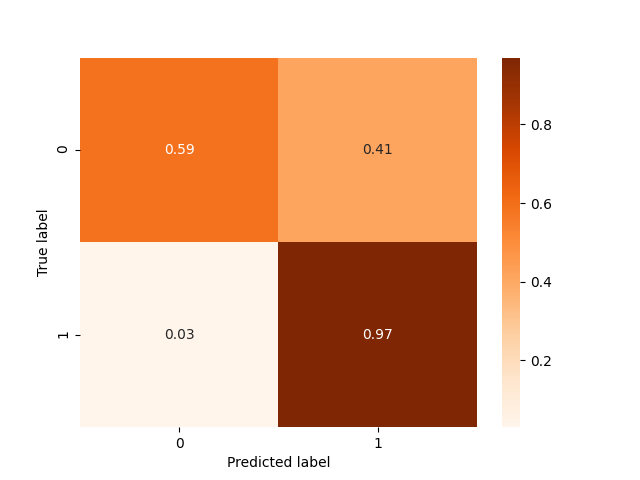
#### Confusion matrix of fold 1



#### Confusion matrix of fold 5

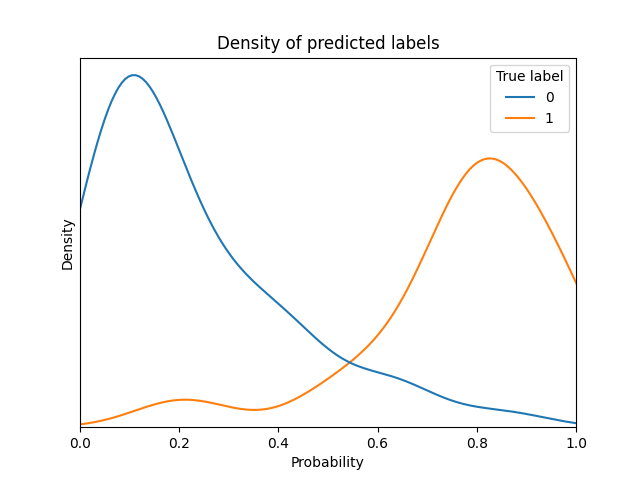


#### Confusion matrix of fold 7

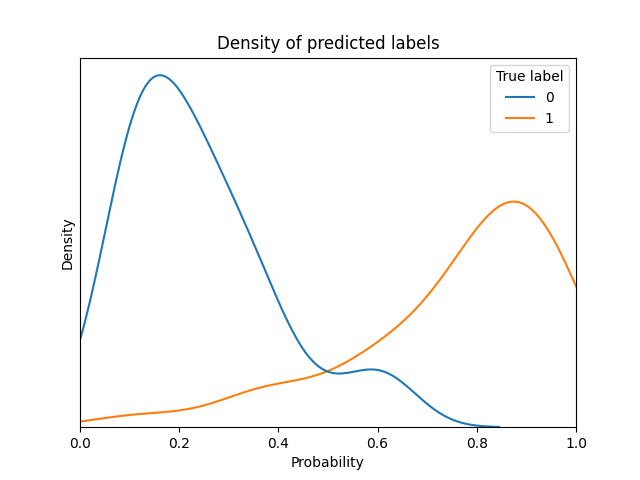


### Histograms

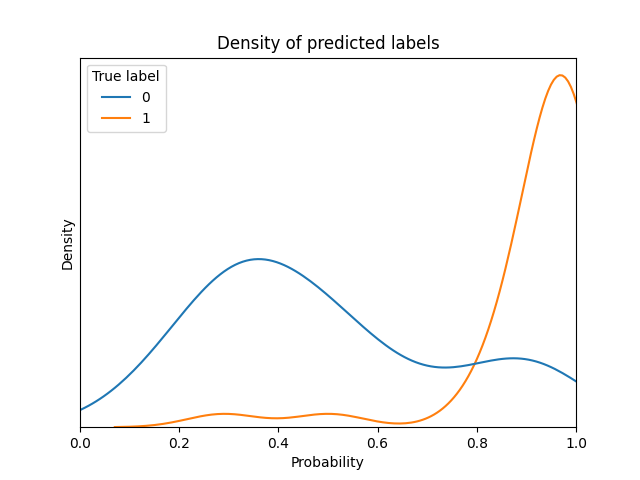
#### Histogram of fold 1



#### Histogram of fold 5



#### Histogram of fold 7



## Comparison of several predictions to assess variance

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Instance** | **Real label** | **Prediction by model of fold 1** | **Prediction by model of fold 5** | **Prediction by model of fold 7** | **Standard deviation in predictions of this instance** |
| 1 | 0 | 0.071 | 0.169 | 0.293 | 0.091 |
| 2 | 1 | 0.935 | 0.88 | 0.986 | 0.043 |
| 3 | 1 | 0.135 | 0.243 | 0.298 | 0.068 |
| 4 | 0 | 0.302 | 0.168 | 0.492 | 0.133 |
| 5 | 0 | 0.394 | 0.444 | 0.862 | 0.21 |
| 6 | 0 | 0.306 | 0.221 | 0.669 | 0.194 |
| 7 | 1 | 0.89 | 0.933 | 0.987 | 0.04 |
| 8 | 1 | 0.962 | 0.94 | 0.994 | 0.022 |
| 9 | 1 | 0.918 | 0.904 | 0.993 | 0.039 |
| 10 | 0 | 0.234 | 0.342 | 0.588 | 0.148 |
| 11 | 0 | 0.115 | 0.196 | 0.43 | 0.134 |
| 12 | 1 | 0.9 | 0.934 | 0.994 | 0.039 |
| 13 | 1 | 0.936 | 0.958 | 0.995 | 0.024 |
| 14 | 0 | 0.392 | 0.429 | 0.654 | 0.116 |
| 15 | 0 | 0.554 | 0.544 | 0.893 | 0.162 |
| 16 | 0 | 0.386 | 0.258 | 0.798 | 0.231 |
| 17 | 0 | 0.41 | 0.327 | 0.831 | 0.22 |
| 18 | 1 | 0.541 | 0.569 | 0.945 | 0.184 |
| 19 | 0 | 0.75 | 0.728 | 0.954 | 0.102 |
| 20 | 1 | 0.986 | 0.965 | 0.994 | 0.012 |

The average standard deviation is 0.111

# 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 random\_forest with the following params:

* undersampling\_majority\_class: False
* model\_\_bootstrap: 1
* model\_\_n\_estimators: 57
* model\_\_max\_depth: 8
* model\_\_min\_samples\_split: 18
* 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** |
| xgboost | * undersampling\_majority\_class: True * max\_k\_undersampling: 6 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_3 * model\_\_max\_depth: 6 * model\_\_learning\_rate: 0.09 * model\_\_min\_child\_weight: 1 * model\_\_subsample: 0.894 * model\_\_colsample\_bytree: 0.14 * model\_\_scale\_pos\_weight: 0.757 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 387.0 * average prop of minority class before resampling: 0.384 * average size of training set after resampling: 477.0 * average prop of minority class after resampling: 0.5 | 0.922 | 0.378 | 0.93 | 0.113 | 0.93 | 0.395 | 0.911 | 0.121 |
| xgboost | * undersampling\_majority\_class: True * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_1 * model\_\_max\_depth: 10 * model\_\_learning\_rate: 0.089 * model\_\_min\_child\_weight: 6 * model\_\_subsample: 0.89 * model\_\_colsample\_bytree: 0.379 * model\_\_scale\_pos\_weight: 0.345 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 387.0 * average prop of minority class before resampling: 0.384 * average size of training set after resampling: 477.0 * average prop of minority class after resampling: 0.5 | 0.906 | 0.436 | 0.895 | 0.137 | 0.901 | 0.425 | 0.893 | 0.132 |
| xgboost | * undersampling\_majority\_class: True * max\_k\_undersampling: 6 * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_2 * model\_\_max\_depth: 5 * model\_\_learning\_rate: 0.2 * model\_\_min\_child\_weight: 8 * model\_\_subsample: 0.816 * model\_\_colsample\_bytree: 0.511 * model\_\_scale\_pos\_weight: 1.588 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 387.0 * average prop of minority class before resampling: 0.384 * average size of training set after resampling: 477.0 * average prop of minority class after resampling: 0.5 | 0.916 | 0.39 | 0.922 | 0.113 | 0.965 | 0.304 | 0.962 | 0.095 |
| xgboost | * undersampling\_majority\_class: True * max\_k\_undersampling: 6 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_3 * model\_\_max\_depth: 10 * model\_\_learning\_rate: 0.143 * model\_\_min\_child\_weight: 3 * model\_\_subsample: 0.983 * model\_\_colsample\_bytree: 0.186 * model\_\_scale\_pos\_weight: 4.751 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 387.0 * average prop of minority class before resampling: 0.384 * average size of training set after resampling: 477.0 * average prop of minority class after resampling: 0.5 | 0.929 | 0.425 | 0.941 | 0.134 | 0.948 | 0.428 | 0.928 | 0.13 |
| xgboost | * undersampling\_majority\_class: False * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_1 * model\_\_max\_depth: 11 * model\_\_learning\_rate: 0.102 * model\_\_min\_child\_weight: 2 * model\_\_subsample: 0.872 * model\_\_colsample\_bytree: 0.27 * model\_\_scale\_pos\_weight: 1.885 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 535.5 * average prop of minority class before resampling: 0.445 * average size of training set after resampling: 594.0 * average prop of minority class after resampling: 0.5 | 0.852 | 0.483 | 0.833 | 0.153 | 0.872 | 0.475 | 0.864 | 0.15 |
| xgboost | * undersampling\_majority\_class: False * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_2 * model\_\_max\_depth: 14 * model\_\_learning\_rate: 0.087 * model\_\_min\_child\_weight: 6 * model\_\_subsample: 0.944 * model\_\_colsample\_bytree: 0.196 * model\_\_scale\_pos\_weight: 1.673 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 535.5 * average prop of minority class before resampling: 0.445 * average size of training set after resampling: 594.0 * average prop of minority class after resampling: 0.5 | 0.91 | 0.382 | 0.917 | 0.115 | 0.944 | 0.36 | 0.931 | 0.111 |
| xgboost | * undersampling\_majority\_class: True * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_2 * model\_\_max\_depth: 13 * model\_\_learning\_rate: 0.189 * model\_\_min\_child\_weight: 1 * model\_\_subsample: 0.913 * model\_\_colsample\_bytree: 0.402 * model\_\_scale\_pos\_weight: 3.543 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 387.0 * average prop of minority class before resampling: 0.384 * average size of training set after resampling: 477.0 * average prop of minority class after resampling: 0.5 | 0.908 | 0.497 | 0.916 | 0.141 | 0.957 | 0.347 | 0.95 | 0.112 |
| xgboost | * undersampling\_majority\_class: True * max\_k\_undersampling: 6 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_3 * model\_\_max\_depth: 11 * model\_\_learning\_rate: 0.119 * model\_\_min\_child\_weight: 9 * model\_\_subsample: 0.82 * model\_\_colsample\_bytree: 0.192 * model\_\_scale\_pos\_weight: 0.482 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 387.0 * average prop of minority class before resampling: 0.384 * average size of training set after resampling: 477.0 * average prop of minority class after resampling: 0.5 | 0.934 | 0.32 | 0.943 | 0.091 | 0.946 | 0.307 | 0.929 | 0.091 |
| xgboost | * undersampling\_majority\_class: False * max\_k\_undersampling: 6 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_3 * model\_\_max\_depth: 11 * model\_\_learning\_rate: 0.269 * model\_\_min\_child\_weight: 4 * model\_\_subsample: 0.848 * model\_\_colsample\_bytree: 0.137 * model\_\_scale\_pos\_weight: 0.404 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 535.5 * average prop of minority class before resampling: 0.445 * average size of training set after resampling: 594.0 * average prop of minority class after resampling: 0.5 | 0.933 | 0.347 | 0.942 | 0.099 | 0.947 | 0.334 | 0.93 | 0.108 |
| xgboost | * undersampling\_majority\_class: False * max\_k\_undersampling: 6 * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_3 * model\_\_max\_depth: 10 * model\_\_learning\_rate: 0.1 * model\_\_min\_child\_weight: 4 * model\_\_subsample: 0.819 * model\_\_colsample\_bytree: 0.388 * model\_\_scale\_pos\_weight: 0.114 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 535.5 * average prop of minority class before resampling: 0.445 * average size of training set after resampling: 594.0 * average prop of minority class after resampling: 0.5 | 0.923 | 0.512 | 0.935 | 0.152 | 0.949 | 0.45 | 0.941 | 0.146 |
| random\_forest | * undersampling\_majority\_class: False * model\_\_bootstrap: 1 * model\_\_n\_estimators: 61 * model\_\_max\_depth: 2 * model\_\_min\_samples\_split: 17 * model\_\_min\_samples\_leaf: 2 * model\_\_max\_features: auto * model\_\_class\_weight: balanced\_subsample | * option: build model without resampling * number of folds: 4 * average size of training set: 535.5 * average prop of minority class: 0.445 | 0.923 | 0.422 | 0.929 | 0.126 | 0.937 | 0.422 | 0.926 | 0.125 |
| random\_forest | * undersampling\_majority\_class: False * model\_\_bootstrap: 0 * model\_\_n\_estimators: 11 * model\_\_max\_depth: 10 * model\_\_min\_samples\_split: 16 * model\_\_min\_samples\_leaf: 1 * model\_\_max\_features: sqrt * model\_\_class\_weight: balanced | * option: build model without resampling * number of folds: 4 * average size of training set: 535.5 * average prop of minority class: 0.445 | 0.93 | 0.346 | 0.931 | 0.104 | 0.959 | 0.284 | 0.941 | 0.083 |
| random\_forest | * undersampling\_majority\_class: True * model\_\_bootstrap: 0 * model\_\_n\_estimators: 57 * model\_\_max\_depth: 8 * model\_\_min\_samples\_split: 17 * model\_\_min\_samples\_leaf: 3 * model\_\_max\_features: sqrt * model\_\_class\_weight: balanced\_subsample | * option: build model without resampling * number of folds: 4 * average size of training set: 387.0 * average prop of minority class: 0.384 | 0.934 | 0.332 | 0.937 | 0.099 | 0.958 | 0.303 | 0.944 | 0.086 |
| random\_forest | * undersampling\_majority\_class: True * model\_\_bootstrap: 1 * model\_\_n\_estimators: 55 * model\_\_max\_depth: 7 * 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: 4 * average size of training set: 387.0 * average prop of minority class: 0.384 | 0.938 | 0.333 | 0.943 | 0.099 | 0.957 | 0.313 | 0.937 | 0.089 |
| random\_forest | * undersampling\_majority\_class: False * model\_\_bootstrap: 1 * model\_\_n\_estimators: 57 * model\_\_max\_depth: 8 * model\_\_min\_samples\_split: 18 * model\_\_min\_samples\_leaf: 3 * model\_\_max\_features: auto * model\_\_class\_weight: balanced | * option: build model without resampling * number of folds: 4 * average size of training set: 535.5 * average prop of minority class: 0.445 | 0.94 | 0.325 | 0.944 | 0.097 | 0.957 | 0.308 | 0.939 | 0.088 |
| random\_forest | * undersampling\_majority\_class: True * model\_\_bootstrap: 0 * model\_\_n\_estimators: 56 * model\_\_max\_depth: 8 * 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: 4 * average size of training set: 387.0 * average prop of minority class: 0.384 | 0.939 | 0.328 | 0.939 | 0.098 | 0.958 | 0.299 | 0.947 | 0.086 |
| random\_forest | * undersampling\_majority\_class: False * model\_\_bootstrap: 1 * model\_\_n\_estimators: 55 * model\_\_max\_depth: 5 * 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: 4 * average size of training set: 535.5 * average prop of minority class: 0.445 | 0.937 | 0.339 | 0.941 | 0.1 | 0.951 | 0.319 | 0.936 | 0.091 |
| random\_forest | * undersampling\_majority\_class: False * model\_\_bootstrap: 1 * model\_\_n\_estimators: 59 * model\_\_max\_depth: 10 * 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: 4 * average size of training set: 535.5 * average prop of minority class: 0.445 | 0.935 | 0.335 | 0.938 | 0.099 | 0.959 | 0.3 | 0.945 | 0.086 |
| random\_forest | * undersampling\_majority\_class: False * model\_\_bootstrap: 0 * model\_\_n\_estimators: 46 * model\_\_max\_depth: 3 * model\_\_min\_samples\_split: 7 * model\_\_min\_samples\_leaf: 1 * model\_\_max\_features: sqrt * model\_\_class\_weight: balanced\_subsample | * option: build model without resampling * number of folds: 4 * average size of training set: 535.5 * average prop of minority class: 0.445 | 0.933 | 0.369 | 0.936 | 0.108 | 0.949 | 0.359 | 0.931 | 0.102 |
| random\_forest | * undersampling\_majority\_class: False * model\_\_bootstrap: 1 * model\_\_n\_estimators: 67 * model\_\_max\_depth: 9 * model\_\_min\_samples\_split: 15 * model\_\_min\_samples\_leaf: 1 * model\_\_max\_features: auto * model\_\_class\_weight: balanced\_subsample | * option: build model without resampling * number of folds: 4 * average size of training set: 535.5 * average prop of minority class: 0.445 | 0.935 | 0.327 | 0.94 | 0.095 | 0.965 | 0.288 | 0.95 | 0.082 |

## 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 xgboost with the following params:

* undersampling\_majority\_class: True
* max\_k\_undersampling: 6
* resample\_\_sampling\_strategy: all
* post\_process\_\_option: option\_3
* model\_\_max\_depth: 9
* model\_\_learning\_rate: 0.056
* model\_\_min\_child\_weight: 5
* model\_\_subsample: 0.882
* model\_\_colsample\_bytree: 0.149
* model\_\_scale\_pos\_weight: 0.835
* 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** |
| xgboost | * undersampling\_majority\_class: True * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_1 * model\_\_max\_depth: 6 * model\_\_learning\_rate: 0.082 * model\_\_min\_child\_weight: 7 * model\_\_subsample: 0.842 * model\_\_colsample\_bytree: 0.256 * model\_\_scale\_pos\_weight: 1.383 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 387.25 * average prop of minority class before resampling: 0.385 * average size of training set after resampling: 476.5 * average prop of minority class after resampling: 0.5 | 0.9 | 0.453 | 0.884 | 0.144 | 0.91 | 0.431 | 0.914 | 0.135 |
| xgboost | * undersampling\_majority\_class: True * max\_k\_undersampling: 6 * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_3 * model\_\_max\_depth: 9 * model\_\_learning\_rate: 0.056 * model\_\_min\_child\_weight: 5 * model\_\_subsample: 0.882 * model\_\_colsample\_bytree: 0.149 * model\_\_scale\_pos\_weight: 0.835 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 387.25 * average prop of minority class before resampling: 0.385 * average size of training set after resampling: 476.5 * average prop of minority class after resampling: 0.5 | 0.949 | 0.33 | 0.938 | 0.095 | 0.956 | 0.315 | 0.962 | 0.088 |
| xgboost | * undersampling\_majority\_class: False * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_3 * model\_\_max\_depth: 6 * model\_\_learning\_rate: 0.063 * model\_\_min\_child\_weight: 8 * model\_\_subsample: 0.846 * model\_\_colsample\_bytree: 0.161 * model\_\_scale\_pos\_weight: 0.189 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 536.25 * average prop of minority class before resampling: 0.444 * average size of training set after resampling: 596.0 * average prop of minority class after resampling: 0.5 | 0.94 | 0.445 | 0.929 | 0.149 | 0.963 | 0.428 | 0.968 | 0.139 |
| xgboost | * undersampling\_majority\_class: False * max\_k\_undersampling: 6 * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_1 * model\_\_max\_depth: 13 * model\_\_learning\_rate: 0.08 * model\_\_min\_child\_weight: 1 * model\_\_subsample: 0.856 * model\_\_colsample\_bytree: 0.319 * model\_\_scale\_pos\_weight: 0.105 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 536.25 * average prop of minority class before resampling: 0.444 * average size of training set after resampling: 596.0 * average prop of minority class after resampling: 0.5 | 0.896 | 0.626 | 0.881 | 0.211 | 0.916 | 0.598 | 0.924 | 0.197 |
| xgboost | * undersampling\_majority\_class: False * max\_k\_undersampling: 6 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_3 * model\_\_max\_depth: 7 * model\_\_learning\_rate: 0.133 * model\_\_min\_child\_weight: 4 * model\_\_subsample: 0.878 * model\_\_colsample\_bytree: 0.158 * model\_\_scale\_pos\_weight: 5.605 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 536.25 * average prop of minority class before resampling: 0.444 * average size of training set after resampling: 596.0 * average prop of minority class after resampling: 0.5 | 0.944 | 0.381 | 0.933 | 0.119 | 0.942 | 0.357 | 0.947 | 0.114 |
| xgboost | * undersampling\_majority\_class: True * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_3 * model\_\_max\_depth: 14 * model\_\_learning\_rate: 0.12 * model\_\_min\_child\_weight: 3 * model\_\_subsample: 0.876 * model\_\_colsample\_bytree: 0.135 * model\_\_scale\_pos\_weight: 0.15 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 387.25 * average prop of minority class before resampling: 0.385 * average size of training set after resampling: 476.5 * average prop of minority class after resampling: 0.5 | 0.93 | 0.428 | 0.919 | 0.14 | 0.949 | 0.426 | 0.956 | 0.137 |
| xgboost | * undersampling\_majority\_class: False * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_2 * model\_\_max\_depth: 10 * model\_\_learning\_rate: 0.113 * model\_\_min\_child\_weight: 6 * model\_\_subsample: 0.898 * model\_\_colsample\_bytree: 0.517 * model\_\_scale\_pos\_weight: 2.174 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 536.25 * average prop of minority class before resampling: 0.444 * average size of training set after resampling: 596.0 * average prop of minority class after resampling: 0.5 | 0.933 | 0.372 | 0.919 | 0.11 | 0.957 | 0.301 | 0.954 | 0.091 |
| xgboost | * undersampling\_majority\_class: False * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_1 * model\_\_max\_depth: 11 * model\_\_learning\_rate: 0.163 * model\_\_min\_child\_weight: 4 * model\_\_subsample: 0.853 * model\_\_colsample\_bytree: 0.624 * model\_\_scale\_pos\_weight: 0.493 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 536.25 * average prop of minority class before resampling: 0.444 * average size of training set after resampling: 596.0 * average prop of minority class after resampling: 0.5 | 0.894 | 0.425 | 0.879 | 0.129 | 0.931 | 0.356 | 0.925 | 0.115 |
| xgboost | * undersampling\_majority\_class: True * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_3 * model\_\_max\_depth: 11 * model\_\_learning\_rate: 0.191 * model\_\_min\_child\_weight: 3 * model\_\_subsample: 0.828 * model\_\_colsample\_bytree: 0.247 * model\_\_scale\_pos\_weight: 0.712 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 387.25 * average prop of minority class before resampling: 0.385 * average size of training set after resampling: 476.5 * average prop of minority class after resampling: 0.5 | 0.948 | 0.3 | 0.933 | 0.09 | 0.958 | 0.269 | 0.961 | 0.077 |
| xgboost | * undersampling\_majority\_class: True * max\_k\_undersampling: 6 * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_3 * model\_\_max\_depth: 11 * model\_\_learning\_rate: 0.068 * model\_\_min\_child\_weight: 8 * model\_\_subsample: 0.957 * model\_\_colsample\_bytree: 0.601 * model\_\_scale\_pos\_weight: 0.571 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 387.25 * average prop of minority class before resampling: 0.385 * average size of training set after resampling: 476.5 * average prop of minority class after resampling: 0.5 | 0.937 | 0.303 | 0.92 | 0.089 | 0.97 | 0.249 | 0.97 | 0.072 |
| random\_forest | * undersampling\_majority\_class: True * model\_\_bootstrap: 0 * model\_\_n\_estimators: 19 * model\_\_max\_depth: 4 * model\_\_min\_samples\_split: 8 * model\_\_min\_samples\_leaf: 3 * model\_\_max\_features: auto * model\_\_class\_weight: balanced\_subsample | * option: build model without resampling * number of folds: 4 * average size of training set: 387.25 * average prop of minority class: 0.385 | 0.944 | 0.346 | 0.929 | 0.102 | 0.959 | 0.309 | 0.963 | 0.086 |
| random\_forest | * undersampling\_majority\_class: False * model\_\_bootstrap: 0 * model\_\_n\_estimators: 56 * model\_\_max\_depth: 2 * model\_\_min\_samples\_split: 8 * model\_\_min\_samples\_leaf: 3 * model\_\_max\_features: auto * model\_\_class\_weight: balanced | * option: build model without resampling * number of folds: 4 * average size of training set: 536.25 * average prop of minority class: 0.444 | 0.933 | 0.416 | 0.922 | 0.124 | 0.951 | 0.39 | 0.954 | 0.112 |
| random\_forest | * undersampling\_majority\_class: True * model\_\_bootstrap: 1 * model\_\_n\_estimators: 46 * model\_\_max\_depth: 7 * model\_\_min\_samples\_split: 20 * model\_\_min\_samples\_leaf: 4 * model\_\_max\_features: sqrt * model\_\_class\_weight: balanced | * option: build model without resampling * number of folds: 4 * average size of training set: 387.25 * average prop of minority class: 0.385 | 0.948 | 0.333 | 0.933 | 0.098 | 0.961 | 0.301 | 0.965 | 0.085 |
| random\_forest | * undersampling\_majority\_class: True * model\_\_bootstrap: 0 * model\_\_n\_estimators: 93 * model\_\_max\_depth: 3 * model\_\_min\_samples\_split: 5 * model\_\_min\_samples\_leaf: 2 * model\_\_max\_features: auto * model\_\_class\_weight: balanced\_subsample | * option: build model without resampling * number of folds: 4 * average size of training set: 387.25 * average prop of minority class: 0.385 | 0.936 | 0.376 | 0.921 | 0.111 | 0.956 | 0.343 | 0.959 | 0.097 |
| random\_forest | * undersampling\_majority\_class: True * model\_\_bootstrap: 1 * model\_\_n\_estimators: 30 * model\_\_max\_depth: 9 * model\_\_min\_samples\_split: 17 * model\_\_min\_samples\_leaf: 1 * model\_\_max\_features: auto * model\_\_class\_weight: balanced\_subsample | * option: build model without resampling * number of folds: 4 * average size of training set: 387.25 * average prop of minority class: 0.385 | 0.948 | 0.333 | 0.933 | 0.1 | 0.956 | 0.299 | 0.964 | 0.084 |
| random\_forest | * undersampling\_majority\_class: True * model\_\_bootstrap: 1 * model\_\_n\_estimators: 24 * model\_\_max\_depth: 4 * 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: 4 * average size of training set: 387.25 * average prop of minority class: 0.385 | 0.946 | 0.355 | 0.929 | 0.104 | 0.956 | 0.328 | 0.961 | 0.093 |
| random\_forest | * undersampling\_majority\_class: True * model\_\_bootstrap: 0 * model\_\_n\_estimators: 66 * model\_\_max\_depth: 8 * model\_\_min\_samples\_split: 8 * model\_\_min\_samples\_leaf: 3 * model\_\_max\_features: sqrt * model\_\_class\_weight: balanced\_subsample | * option: build model without resampling * number of folds: 4 * average size of training set: 387.25 * average prop of minority class: 0.385 | 0.946 | 0.325 | 0.932 | 0.098 | 0.961 | 0.279 | 0.965 | 0.079 |
| random\_forest | * undersampling\_majority\_class: True * model\_\_bootstrap: 1 * model\_\_n\_estimators: 56 * model\_\_max\_depth: 4 * model\_\_min\_samples\_split: 11 * model\_\_min\_samples\_leaf: 1 * model\_\_max\_features: auto * model\_\_class\_weight: balanced | * option: build model without resampling * number of folds: 4 * average size of training set: 387.25 * average prop of minority class: 0.385 | 0.945 | 0.353 | 0.93 | 0.104 | 0.958 | 0.319 | 0.961 | 0.09 |
| random\_forest | * undersampling\_majority\_class: True * model\_\_bootstrap: 0 * model\_\_n\_estimators: 37 * model\_\_max\_depth: 4 * model\_\_min\_samples\_split: 11 * model\_\_min\_samples\_leaf: 1 * model\_\_max\_features: auto * model\_\_class\_weight: balanced | * option: build model without resampling * number of folds: 4 * average size of training set: 387.25 * average prop of minority class: 0.385 | 0.942 | 0.344 | 0.927 | 0.101 | 0.957 | 0.306 | 0.961 | 0.085 |
| random\_forest | * undersampling\_majority\_class: False * model\_\_bootstrap: 0 * model\_\_n\_estimators: 18 * model\_\_max\_depth: 9 * 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: 4 * average size of training set: 536.25 * average prop of minority class: 0.444 | 0.942 | 0.315 | 0.933 | 0.098 | 0.962 | 0.273 | 0.965 | 0.078 |

## 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 xgboost with the following params:

* undersampling\_majority\_class: True
* max\_k\_undersampling: 6
* resample\_\_sampling\_strategy: all
* post\_process\_\_option: option\_2
* model\_\_max\_depth: 14
* model\_\_learning\_rate: 0.103
* model\_\_min\_child\_weight: 9
* model\_\_subsample: 0.809
* model\_\_colsample\_bytree: 0.152
* model\_\_scale\_pos\_weight: 5.624
* 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** |
| xgboost | * undersampling\_majority\_class: True * max\_k\_undersampling: 6 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_3 * model\_\_max\_depth: 12 * model\_\_learning\_rate: 0.071 * model\_\_min\_child\_weight: 1 * model\_\_subsample: 0.902 * model\_\_colsample\_bytree: 0.27 * model\_\_scale\_pos\_weight: 0.405 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 387.25 * average prop of minority class before resampling: 0.385 * average size of training set after resampling: 476.5 * average prop of minority class after resampling: 0.5 | 0.941 | 0.322 | 0.909 | 0.096 | 0.931 | 0.328 | 0.939 | 0.099 |
| xgboost | * undersampling\_majority\_class: True * max\_k\_undersampling: 6 * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_2 * model\_\_max\_depth: 14 * model\_\_learning\_rate: 0.103 * model\_\_min\_child\_weight: 9 * model\_\_subsample: 0.809 * model\_\_colsample\_bytree: 0.152 * model\_\_scale\_pos\_weight: 5.624 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 387.25 * average prop of minority class before resampling: 0.385 * average size of training set after resampling: 476.5 * average prop of minority class after resampling: 0.5 | 0.942 | 0.512 | 0.919 | 0.158 | 0.933 | 0.536 | 0.938 | 0.177 |
| xgboost | * undersampling\_majority\_class: False * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_1 * model\_\_max\_depth: 5 * model\_\_learning\_rate: 0.119 * model\_\_min\_child\_weight: 3 * model\_\_subsample: 0.814 * model\_\_colsample\_bytree: 0.413 * model\_\_scale\_pos\_weight: 2.994 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 536.25 * average prop of minority class before resampling: 0.444 * average size of training set after resampling: 596.0 * average prop of minority class after resampling: 0.5 | 0.893 | 0.456 | 0.874 | 0.144 | 0.862 | 0.526 | 0.854 | 0.175 |
| xgboost | * undersampling\_majority\_class: False * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_1 * model\_\_max\_depth: 6 * model\_\_learning\_rate: 0.088 * model\_\_min\_child\_weight: 5 * model\_\_subsample: 0.895 * model\_\_colsample\_bytree: 0.781 * model\_\_scale\_pos\_weight: 0.238 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 536.25 * average prop of minority class before resampling: 0.444 * average size of training set after resampling: 596.0 * average prop of minority class after resampling: 0.5 | 0.916 | 0.429 | 0.886 | 0.135 | 0.861 | 0.53 | 0.84 | 0.163 |
| xgboost | * undersampling\_majority\_class: False * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_2 * model\_\_max\_depth: 14 * model\_\_learning\_rate: 0.058 * model\_\_min\_child\_weight: 8 * model\_\_subsample: 0.81 * model\_\_colsample\_bytree: 0.771 * model\_\_scale\_pos\_weight: 0.191 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 536.25 * average prop of minority class before resampling: 0.444 * average size of training set after resampling: 596.0 * average prop of minority class after resampling: 0.5 | 0.94 | 0.435 | 0.892 | 0.144 | 0.942 | 0.407 | 0.94 | 0.125 |
| xgboost | * undersampling\_majority\_class: True * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_3 * model\_\_max\_depth: 13 * model\_\_learning\_rate: 0.082 * model\_\_min\_child\_weight: 7 * model\_\_subsample: 0.888 * model\_\_colsample\_bytree: 0.272 * model\_\_scale\_pos\_weight: 3.358 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 387.25 * average prop of minority class before resampling: 0.385 * average size of training set after resampling: 476.5 * average prop of minority class after resampling: 0.5 | 0.945 | 0.413 | 0.916 | 0.127 | 0.938 | 0.434 | 0.944 | 0.147 |
| xgboost | * undersampling\_majority\_class: False * max\_k\_undersampling: 6 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_1 * model\_\_max\_depth: 14 * model\_\_learning\_rate: 0.066 * model\_\_min\_child\_weight: 1 * model\_\_subsample: 0.865 * model\_\_colsample\_bytree: 0.229 * model\_\_scale\_pos\_weight: 0.24 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 536.25 * average prop of minority class before resampling: 0.444 * average size of training set after resampling: 596.0 * average prop of minority class after resampling: 0.5 | 0.873 | 0.515 | 0.849 | 0.169 | 0.837 | 0.547 | 0.815 | 0.182 |
| xgboost | * undersampling\_majority\_class: True * max\_k\_undersampling: 6 * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_1 * model\_\_max\_depth: 15 * model\_\_learning\_rate: 0.121 * model\_\_min\_child\_weight: 2 * model\_\_subsample: 0.811 * model\_\_colsample\_bytree: 0.481 * model\_\_scale\_pos\_weight: 0.168 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 387.25 * average prop of minority class before resampling: 0.385 * average size of training set after resampling: 476.5 * average prop of minority class after resampling: 0.5 | 0.917 | 0.403 | 0.897 | 0.123 | 0.853 | 0.511 | 0.831 | 0.153 |
| xgboost | * undersampling\_majority\_class: True * max\_k\_undersampling: 5 * resample\_\_sampling\_strategy: minority * post\_process\_\_option: option\_3 * model\_\_max\_depth: 9 * model\_\_learning\_rate: 0.171 * model\_\_min\_child\_weight: 2 * model\_\_subsample: 0.878 * model\_\_colsample\_bytree: 0.246 * model\_\_scale\_pos\_weight: 8.557 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 387.25 * average prop of minority class before resampling: 0.385 * average size of training set after resampling: 476.5 * average prop of minority class after resampling: 0.5 | 0.934 | 0.462 | 0.903 | 0.131 | 0.936 | 0.476 | 0.941 | 0.153 |
| xgboost | * undersampling\_majority\_class: True * max\_k\_undersampling: 6 * resample\_\_sampling\_strategy: all * post\_process\_\_option: option\_2 * model\_\_max\_depth: 14 * model\_\_learning\_rate: 0.076 * model\_\_min\_child\_weight: 9 * model\_\_subsample: 0.823 * model\_\_colsample\_bytree: 0.389 * model\_\_scale\_pos\_weight: 0.208 * model\_\_objective: binary:logistic | * option: build model with resampling * number of folds: 4 * average size of training set before resampling: 387.25 * average prop of minority class before resampling: 0.385 * average size of training set after resampling: 476.5 * average prop of minority class after resampling: 0.5 | 0.934 | 0.437 | 0.902 | 0.143 | 0.941 | 0.417 | 0.941 | 0.129 |
| random\_forest | * undersampling\_majority\_class: True * model\_\_bootstrap: 1 * model\_\_n\_estimators: 58 * model\_\_max\_depth: 5 * model\_\_min\_samples\_split: 18 * model\_\_min\_samples\_leaf: 3 * model\_\_max\_features: auto * model\_\_class\_weight: balanced | * option: build model without resampling * number of folds: 4 * average size of training set: 387.25 * average prop of minority class: 0.385 | 0.945 | 0.346 | 0.913 | 0.099 | 0.931 | 0.362 | 0.937 | 0.11 |
| random\_forest | * undersampling\_majority\_class: False * model\_\_bootstrap: 0 * model\_\_n\_estimators: 41 * model\_\_max\_depth: 10 * model\_\_min\_samples\_split: 7 * model\_\_min\_samples\_leaf: 4 * model\_\_max\_features: sqrt * model\_\_class\_weight: balanced\_subsample | * option: build model without resampling * number of folds: 4 * average size of training set: 536.25 * average prop of minority class: 0.444 | 0.941 | 0.445 | 0.911 | 0.094 | 0.925 | 0.348 | 0.931 | 0.107 |
| random\_forest | * undersampling\_majority\_class: True * model\_\_bootstrap: 1 * model\_\_n\_estimators: 54 * model\_\_max\_depth: 7 * 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: 4 * average size of training set: 387.25 * average prop of minority class: 0.385 | 0.942 | 0.343 | 0.908 | 0.098 | 0.931 | 0.363 | 0.938 | 0.113 |
| random\_forest | * undersampling\_majority\_class: False * model\_\_bootstrap: 0 * model\_\_n\_estimators: 19 * model\_\_max\_depth: 5 * model\_\_min\_samples\_split: 9 * model\_\_min\_samples\_leaf: 2 * model\_\_max\_features: auto * model\_\_class\_weight: balanced | * option: build model without resampling * number of folds: 4 * average size of training set: 536.25 * average prop of minority class: 0.444 | 0.94 | 0.337 | 0.909 | 0.1 | 0.918 | 0.361 | 0.925 | 0.111 |
| random\_forest | * undersampling\_majority\_class: False * model\_\_bootstrap: 0 * model\_\_n\_estimators: 61 * model\_\_max\_depth: 5 * model\_\_min\_samples\_split: 13 * model\_\_min\_samples\_leaf: 2 * model\_\_max\_features: sqrt * model\_\_class\_weight: balanced\_subsample | * option: build model without resampling * number of folds: 4 * average size of training set: 536.25 * average prop of minority class: 0.444 | 0.939 | 0.336 | 0.907 | 0.098 | 0.927 | 0.346 | 0.933 | 0.105 |
| random\_forest | * undersampling\_majority\_class: True * model\_\_bootstrap: 1 * model\_\_n\_estimators: 33 * model\_\_max\_depth: 7 * model\_\_min\_samples\_split: 15 * model\_\_min\_samples\_leaf: 3 * model\_\_max\_features: sqrt * model\_\_class\_weight: balanced\_subsample | * option: build model without resampling * number of folds: 4 * average size of training set: 387.25 * average prop of minority class: 0.385 | 0.942 | 0.346 | 0.911 | 0.099 | 0.928 | 0.366 | 0.938 | 0.113 |
| random\_forest | * undersampling\_majority\_class: True * model\_\_bootstrap: 0 * model\_\_n\_estimators: 29 * model\_\_max\_depth: 2 * model\_\_min\_samples\_split: 13 * model\_\_min\_samples\_leaf: 3 * model\_\_max\_features: sqrt * model\_\_class\_weight: balanced\_subsample | * option: build model without resampling * number of folds: 4 * average size of training set: 387.25 * average prop of minority class: 0.385 | 0.935 | 0.414 | 0.902 | 0.122 | 0.927 | 0.422 | 0.931 | 0.127 |
| random\_forest | * undersampling\_majority\_class: False * model\_\_bootstrap: 0 * model\_\_n\_estimators: 84 * model\_\_max\_depth: 2 * model\_\_min\_samples\_split: 11 * model\_\_min\_samples\_leaf: 3 * model\_\_max\_features: sqrt * model\_\_class\_weight: balanced | * option: build model without resampling * number of folds: 4 * average size of training set: 536.25 * average prop of minority class: 0.444 | 0.937 | 0.426 | 0.912 | 0.127 | 0.922 | 0.428 | 0.925 | 0.13 |
| random\_forest | * undersampling\_majority\_class: False * model\_\_bootstrap: 0 * model\_\_n\_estimators: 59 * model\_\_max\_depth: 6 * model\_\_min\_samples\_split: 15 * model\_\_min\_samples\_leaf: 3 * model\_\_max\_features: auto * model\_\_class\_weight: balanced\_subsample | * option: build model without resampling * number of folds: 4 * average size of training set: 536.25 * average prop of minority class: 0.444 | 0.943 | 0.327 | 0.911 | 0.095 | 0.925 | 0.345 | 0.935 | 0.105 |
| random\_forest | * undersampling\_majority\_class: True * model\_\_bootstrap: 0 * model\_\_n\_estimators: 67 * model\_\_max\_depth: 3 * model\_\_min\_samples\_split: 18 * model\_\_min\_samples\_leaf: 3 * model\_\_max\_features: auto * model\_\_class\_weight: balanced | * option: build model without resampling * number of folds: 4 * average size of training set: 387.25 * average prop of minority class: 0.385 | 0.938 | 0.374 | 0.908 | 0.109 | 0.928 | 0.388 | 0.935 | 0.116 |