C:\Users\ayrisbud\AppData\Roaming\Python\Python39\site-packages\xgboost\compat.py:36: FutureWarning: pandas.Int64Index is deprecated and will be removed from pandas in a future version. Use pandas.Index with the appropriate dtype instead.

from pandas import MultiIndex, Int64Index

Before resampling

False 424647

True 1655

Name: target, dtype: int64

After resampling

False 424158

True 1655

Name: target, dtype: int64

Running randomForest for the file fall11Mod

[[181707 286]

[ 357 352]]

precision recall f1-score support

False 1.00 1.00 1.00 181993

True 0.55 0.50 0.52 709

accuracy 1.00 182702

macro avg 0.77 0.75 0.76 182702

weighted avg 1.00 1.00 1.00 182702

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Running xgBoost for the file fall11Mod

C:\Users\ayrisbud\AppData\Roaming\Python\Python39\site-packages\xgboost\sklearn.py:1224: UserWarning: The use of label encoder in XGBClassifier is deprecated and will be removed in a future release. To remove this warning, do the following: 1) Pass option use\_label\_encoder=False when constructing XGBClassifier object; and 2) Encode your labels (y) as integers starting with 0, i.e. 0, 1, 2, ..., [num\_class - 1].

warnings.warn(label\_encoder\_deprecation\_msg, UserWarning)

C:\Users\ayrisbud\AppData\Roaming\Python\Python39\site-packages\xgboost\data.py:262: FutureWarning: pandas.Int64Index is deprecated and will be removed from pandas in a future version. Use pandas.Index with the appropriate dtype instead.

elif isinstance(data.columns, (pd.Int64Index, pd.RangeIndex)):

[19:33:53] WARNING: C:/Users/Administrator/workspace/xgboost-win64\_release\_1.5.1/src/learner.cc:1115: Starting in XGBoost 1.3.0, the default evaluation metric used with the objective 'binary:logistic' was changed from 'error' to 'logloss'. Explicitly set eval\_metric if you'd like to restore the old behavior.

[[181953 40]

[ 425 284]]

precision recall f1-score support

False 1.00 1.00 1.00 181993

True 0.88 0.40 0.55 709

accuracy 1.00 182702

macro avg 0.94 0.70 0.77 182702

weighted avg 1.00 1.00 1.00 182702

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

For file fall11Mod the best model is {XGBClassifier(base\_score=0.5, booster='gbtree', colsample\_bylevel=1,

colsample\_bynode=1, colsample\_bytree=1, enable\_categorical=False,

gamma=0, gpu\_id=-1, importance\_type=None,

interaction\_constraints='', learning\_rate=0.300000012,

max\_delta\_step=0, max\_depth=6, min\_child\_weight=1, missing=nan,

monotone\_constraints='()', n\_estimators=100, n\_jobs=48,

num\_parallel\_tree=1, predictor='auto', random\_state=0,

reg\_alpha=0, reg\_lambda=1, scale\_pos\_weight=1, subsample=1,

tree\_method='exact', validate\_parameters=1, verbosity=None): [0.5498547918683446, 0.9974548718678504, 0.8765432098765432, 0.4005641748942172, 0.8531602477998727, array([[181953, 40],

[ 425, 284]], dtype=int64), ' precision recall f1-score support\n\n False 1.00 1.00 1.00 181993\n True 0.88 0.40 0.55 709\n\n accuracy 1.00 182702\n macro avg 0.94 0.70 0.77 182702\nweighted avg 1.00 1.00 1.00 182702\n'], RandomForestClassifier(class\_weight='balanced', n\_jobs=16, random\_state=42): [0.5226429101707498, 0.9964806077656512, 0.5517241379310345, 0.49647390691114246, 0.7689143718094122, array([[181707, 286],

[ 357, 352]], dtype=int64), ' precision recall f1-score support\n\n False 1.00 1.00 1.00 181993\n True 0.55 0.50 0.52 709\n\n accuracy 1.00 182702\n macro avg 0.77 0.75 0.76 182702\nweighted avg 1.00 1.00 1.00 182702\n']}

Before resampling

False 469216

True 3811

Name: target, dtype: int64

After resampling

False 468687

True 3811

Name: target, dtype: int64

Running randomForest for the file fall12Mod

[[200693 401]

[ 426 1207]]

precision recall f1-score support

False 1.00 1.00 1.00 201094

True 0.75 0.74 0.74 1633

accuracy 1.00 202727

macro avg 0.87 0.87 0.87 202727

weighted avg 1.00 1.00 1.00 202727

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Running xgBoost for the file fall12Mod

C:\Users\ayrisbud\AppData\Roaming\Python\Python39\site-packages\xgboost\sklearn.py:1224: UserWarning: The use of label encoder in XGBClassifier is deprecated and will be removed in a future release. To remove this warning, do the following: 1) Pass option use\_label\_encoder=False when constructing XGBClassifier object; and 2) Encode your labels (y) as integers starting with 0, i.e. 0, 1, 2, ..., [num\_class - 1].

warnings.warn(label\_encoder\_deprecation\_msg, UserWarning)

C:\Users\ayrisbud\AppData\Roaming\Python\Python39\site-packages\xgboost\data.py:262: FutureWarning: pandas.Int64Index is deprecated and will be removed from pandas in a future version. Use pandas.Index with the appropriate dtype instead.

elif isinstance(data.columns, (pd.Int64Index, pd.RangeIndex)):

[19:34:17] WARNING: C:/Users/Administrator/workspace/xgboost-win64\_release\_1.5.1/src/learner.cc:1115: Starting in XGBoost 1.3.0, the default evaluation metric used with the objective 'binary:logistic' was changed from 'error' to 'logloss'. Explicitly set eval\_metric if you'd like to restore the old behavior.

[[201010 84]

[ 499 1134]]

precision recall f1-score support

False 1.00 1.00 1.00 201094

True 0.93 0.69 0.80 1633

accuracy 1.00 202727

macro avg 0.96 0.85 0.90 202727

weighted avg 1.00 1.00 1.00 202727

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

For file fall12Mod the best model is {XGBClassifier(base\_score=0.5, booster='gbtree', colsample\_bylevel=1,

colsample\_bynode=1, colsample\_bytree=1, enable\_categorical=False,

gamma=0, gpu\_id=-1, importance\_type=None,

interaction\_constraints='', learning\_rate=0.300000012,

max\_delta\_step=0, max\_depth=6, min\_child\_weight=1, missing=nan,

monotone\_constraints='()', n\_estimators=100, n\_jobs=48,

num\_parallel\_tree=1, predictor='auto', random\_state=0,

reg\_alpha=0, reg\_lambda=1, scale\_pos\_weight=1, subsample=1,

tree\_method='exact', validate\_parameters=1, verbosity=None): [0.79551034724658, 0.9971242113778629, 0.9310344827586207, 0.6944274341702388, 0.9347859637800164, array([[201010, 84],

[ 499, 1134]], dtype=int64), ' precision recall f1-score support\n\n False 1.00 1.00 1.00 201094\n True 0.93 0.69 0.80 1633\n\n accuracy 1.00 202727\n macro avg 0.96 0.85 0.90 202727\nweighted avg 1.00 1.00 1.00 202727\n'], RandomForestClassifier(class\_weight='balanced', n\_jobs=16, random\_state=42): [0.7448318420240666, 0.9959206223147385, 0.7506218905472637, 0.7391304347826086, 0.8731008810990977, array([[200693, 401],

[ 426, 1207]], dtype=int64), ' precision recall f1-score support\n\n False 1.00 1.00 1.00 201094\n True 0.75 0.74 0.74 1633\n\n accuracy 1.00 202727\n macro avg 0.87 0.87 0.87 202727\nweighted avg 1.00 1.00 1.00 202727\n']}

Before resampling

False 1164587

True 47270

Name: target, dtype: int64

After resampling

False 1160713

True 47270

Name: target, dtype: int64

Running randomForest for the file spring1Mod

[[497196 1914]

[ 1303 18955]]

precision recall f1-score support

False 1.00 1.00 1.00 499110

True 0.91 0.94 0.92 20258

accuracy 0.99 519368

macro avg 0.95 0.97 0.96 519368

weighted avg 0.99 0.99 0.99 519368

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Running xgBoost for the file spring1Mod

C:\Users\ayrisbud\AppData\Roaming\Python\Python39\site-packages\xgboost\sklearn.py:1224: UserWarning: The use of label encoder in XGBClassifier is deprecated and will be removed in a future release. To remove this warning, do the following: 1) Pass option use\_label\_encoder=False when constructing XGBClassifier object; and 2) Encode your labels (y) as integers starting with 0, i.e. 0, 1, 2, ..., [num\_class - 1].

warnings.warn(label\_encoder\_deprecation\_msg, UserWarning)

C:\Users\ayrisbud\AppData\Roaming\Python\Python39\site-packages\xgboost\data.py:262: FutureWarning: pandas.Int64Index is deprecated and will be removed from pandas in a future version. Use pandas.Index with the appropriate dtype instead.

elif isinstance(data.columns, (pd.Int64Index, pd.RangeIndex)):

[19:35:18] WARNING: C:/Users/Administrator/workspace/xgboost-win64\_release\_1.5.1/src/learner.cc:1115: Starting in XGBoost 1.3.0, the default evaluation metric used with the objective 'binary:logistic' was changed from 'error' to 'logloss'. Explicitly set eval\_metric if you'd like to restore the old behavior.

[[497694 1416]

[ 1664 18594]]

precision recall f1-score support

False 1.00 1.00 1.00 499110

True 0.93 0.92 0.92 20258

accuracy 0.99 519368

macro avg 0.96 0.96 0.96 519368

weighted avg 0.99 0.99 0.99 519368

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

For file spring1Mod the best model is {XGBClassifier(base\_score=0.5, booster='gbtree', colsample\_bylevel=1,

colsample\_bynode=1, colsample\_bytree=1, enable\_categorical=False,

gamma=0, gpu\_id=-1, importance\_type=None,

interaction\_constraints='', learning\_rate=0.300000012,

max\_delta\_step=0, max\_depth=6, min\_child\_weight=1, missing=nan,

monotone\_constraints='()', n\_estimators=100, n\_jobs=48,

num\_parallel\_tree=1, predictor='auto', random\_state=0,

reg\_alpha=0, reg\_lambda=1, scale\_pos\_weight=1, subsample=1,

tree\_method='exact', validate\_parameters=1, verbosity=None): [0.92351246647462, 0.9940697155003774, 0.9292353823088456, 0.9178596110178695, 0.9618522266577925, array([[497694, 1416],

[ 1664, 18594]], dtype=int64), ' precision recall f1-score support\n\n False 1.00 1.00 1.00 499110\n True 0.93 0.92 0.92 20258\n\n accuracy 0.99 519368\n macro avg 0.96 0.96 0.96 519368\nweighted avg 0.99 0.99 0.99 519368\n'], RandomForestClassifier(class\_weight='balanced', n\_jobs=16, random\_state=42): [0.9217788800544655, 0.9938059333651669, 0.9082850160525181, 0.935679731464113, 0.9553882959067133, array([[497196, 1914],

[ 1303, 18955]], dtype=int64), ' precision recall f1-score support\n\n False 1.00 1.00 1.00 499110\n True 0.91 0.94 0.92 20258\n\n accuracy 0.99 519368\n macro avg 0.95 0.97 0.96 519368\nweighted avg 0.99 0.99 0.99 519368\n']}

Before resampling

False 614405

True 8458

Name: target, dtype: int64

After resampling

False 612735

True 8458

Name: target, dtype: int64

Running randomForest for the file summer11Mod

[[260584 2733]

[ 1250 2375]]

precision recall f1-score support

False 1.00 0.99 0.99 263317

True 0.46 0.66 0.54 3625

accuracy 0.99 266942

macro avg 0.73 0.82 0.77 266942

weighted avg 0.99 0.99 0.99 266942

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Running xgBoost for the file summer11Mod

C:\Users\ayrisbud\AppData\Roaming\Python\Python39\site-packages\xgboost\sklearn.py:1224: UserWarning: The use of label encoder in XGBClassifier is deprecated and will be removed in a future release. To remove this warning, do the following: 1) Pass option use\_label\_encoder=False when constructing XGBClassifier object; and 2) Encode your labels (y) as integers starting with 0, i.e. 0, 1, 2, ..., [num\_class - 1].

warnings.warn(label\_encoder\_deprecation\_msg, UserWarning)

C:\Users\ayrisbud\AppData\Roaming\Python\Python39\site-packages\xgboost\data.py:262: FutureWarning: pandas.Int64Index is deprecated and will be removed from pandas in a future version. Use pandas.Index with the appropriate dtype instead.

elif isinstance(data.columns, (pd.Int64Index, pd.RangeIndex)):

[19:36:07] WARNING: C:/Users/Administrator/workspace/xgboost-win64\_release\_1.5.1/src/learner.cc:1115: Starting in XGBoost 1.3.0, the default evaluation metric used with the objective 'binary:logistic' was changed from 'error' to 'logloss'. Explicitly set eval\_metric if you'd like to restore the old behavior.

[[262880 437]

[ 1982 1643]]

precision recall f1-score support

False 0.99 1.00 1.00 263317

True 0.79 0.45 0.58 3625

accuracy 0.99 266942

macro avg 0.89 0.73 0.79 266942

weighted avg 0.99 0.99 0.99 266942

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

For file summer11Mod the best model is {XGBClassifier(base\_score=0.5, booster='gbtree', colsample\_bylevel=1,

colsample\_bynode=1, colsample\_bytree=1, enable\_categorical=False,

gamma=0, gpu\_id=-1, importance\_type=None,

interaction\_constraints='', learning\_rate=0.300000012,

max\_delta\_step=0, max\_depth=6, min\_child\_weight=1, missing=nan,

monotone\_constraints='()', n\_estimators=100, n\_jobs=48,

num\_parallel\_tree=1, predictor='auto', random\_state=0,

reg\_alpha=0, reg\_lambda=1, scale\_pos\_weight=1, subsample=1,

tree\_method='exact', validate\_parameters=1, verbosity=None): [0.575985977212971, 0.9909381064051367, 0.7899038461538461, 0.4532413793103448, 0.840705789776893, array([[262880, 437],

[ 1982, 1643]], dtype=int64), ' precision recall f1-score support\n\n False 0.99 1.00 1.00 263317\n True 0.79 0.45 0.58 3625\n\n accuracy 0.99 266942\n macro avg 0.89 0.73 0.79 266942\nweighted avg 0.99 0.99 0.99 266942\n'], RandomForestClassifier(class\_weight='balanced', n\_jobs=16, random\_state=42): [0.5439138898431237, 0.9850791557716657, 0.4649569303054033, 0.6551724137931034, 0.7438596041965442, array([[260584, 2733],

[ 1250, 2375]], dtype=int64), ' precision recall f1-score support\n\n False 1.00 0.99 0.99 263317\n True 0.46 0.66 0.54 3625\n\n accuracy 0.99 266942\n macro avg 0.73 0.82 0.77 266942\nweighted avg 0.99 0.99 0.99 266942\n']}

Before resampling

False 695140

True 7660

Name: target, dtype: int64

After resampling

False 692657

True 7660

Name: target, dtype: int64

Running randomForest for the file summer12Mod

[[296020 1897]

[ 1275 2008]]

precision recall f1-score support

False 1.00 0.99 0.99 297917

True 0.51 0.61 0.56 3283

accuracy 0.99 301200

macro avg 0.75 0.80 0.78 301200

weighted avg 0.99 0.99 0.99 301200

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Running xgBoost for the file summer12Mod

C:\Users\ayrisbud\AppData\Roaming\Python\Python39\site-packages\xgboost\sklearn.py:1224: UserWarning: The use of label encoder in XGBClassifier is deprecated and will be removed in a future release. To remove this warning, do the following: 1) Pass option use\_label\_encoder=False when constructing XGBClassifier object; and 2) Encode your labels (y) as integers starting with 0, i.e. 0, 1, 2, ..., [num\_class - 1].

warnings.warn(label\_encoder\_deprecation\_msg, UserWarning)

C:\Users\ayrisbud\AppData\Roaming\Python\Python39\site-packages\xgboost\data.py:262: FutureWarning: pandas.Int64Index is deprecated and will be removed from pandas in a future version. Use pandas.Index with the appropriate dtype instead.

elif isinstance(data.columns, (pd.Int64Index, pd.RangeIndex)):

[19:36:50] WARNING: C:/Users/Administrator/workspace/xgboost-win64\_release\_1.5.1/src/learner.cc:1115: Starting in XGBoost 1.3.0, the default evaluation metric used with the objective 'binary:logistic' was changed from 'error' to 'logloss'. Explicitly set eval\_metric if you'd like to restore the old behavior.

[[297297 620]

[ 1812 1471]]

precision recall f1-score support

False 0.99 1.00 1.00 297917

True 0.70 0.45 0.55 3283

accuracy 0.99 301200

macro avg 0.85 0.72 0.77 301200

weighted avg 0.99 0.99 0.99 301200

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

For file summer12Mod the best model is {RandomForestClassifier(class\_weight='balanced', n\_jobs=16, random\_state=42): [0.5587089593767389, 0.989468791500664, 0.5142125480153649, 0.6116356990557417, 0.7632137420838805, array([[296020, 1897],

[ 1275, 2008]], dtype=int64), ' precision recall f1-score support\n\n False 1.00 0.99 0.99 297917\n True 0.51 0.61 0.56 3283\n\n accuracy 0.99 301200\n macro avg 0.75 0.80 0.78 301200\nweighted avg 0.99 0.99 0.99 301200\n'], XGBClassifier(base\_score=0.5, booster='gbtree', colsample\_bylevel=1,

colsample\_bynode=1, colsample\_bytree=1, enable\_categorical=False,

gamma=0, gpu\_id=-1, importance\_type=None,

interaction\_constraints='', learning\_rate=0.300000012,

max\_delta\_step=0, max\_depth=6, min\_child\_weight=1, missing=nan,

monotone\_constraints='()', n\_estimators=100, n\_jobs=48,

num\_parallel\_tree=1, predictor='auto', random\_state=0,

reg\_alpha=0, reg\_lambda=1, scale\_pos\_weight=1, subsample=1,

tree\_method='exact', validate\_parameters=1, verbosity=None): [0.547450688500186, 0.991925630810093, 0.7034911525585844, 0.4480657934815717, 0.813113966731251, array([[297297, 620],

[ 1812, 1471]], dtype=int64), ' precision recall f1-score support\n\n False 0.99 1.00 1.00 297917\n True 0.70 0.45 0.55 3283\n\n accuracy 0.99 301200\n macro avg 0.85 0.72 0.77 301200\nweighted avg 0.99 0.99 0.99 301200\n']}

Before resampling

False 1035541

True 94205

Name: target, dtype: int64

After resampling

False 1030003

True 94205

Name: target, dtype: int64

Running randomForest for the file winter1Mod

[[437500 6304]

[ 2357 38016]]

precision recall f1-score support

False 0.99 0.99 0.99 443804

True 0.86 0.94 0.90 40373

accuracy 0.98 484177

macro avg 0.93 0.96 0.94 484177

weighted avg 0.98 0.98 0.98 484177

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Running xgBoost for the file winter1Mod

C:\Users\ayrisbud\AppData\Roaming\Python\Python39\site-packages\xgboost\sklearn.py:1224: UserWarning: The use of label encoder in XGBClassifier is deprecated and will be removed in a future release. To remove this warning, do the following: 1) Pass option use\_label\_encoder=False when constructing XGBClassifier object; and 2) Encode your labels (y) as integers starting with 0, i.e. 0, 1, 2, ..., [num\_class - 1].

warnings.warn(label\_encoder\_deprecation\_msg, UserWarning)

C:\Users\ayrisbud\AppData\Roaming\Python\Python39\site-packages\xgboost\data.py:262: FutureWarning: pandas.Int64Index is deprecated and will be removed from pandas in a future version. Use pandas.Index with the appropriate dtype instead.

elif isinstance(data.columns, (pd.Int64Index, pd.RangeIndex)):

[19:37:55] WARNING: C:/Users/Administrator/workspace/xgboost-win64\_release\_1.5.1/src/learner.cc:1115: Starting in XGBoost 1.3.0, the default evaluation metric used with the objective 'binary:logistic' was changed from 'error' to 'logloss'. Explicitly set eval\_metric if you'd like to restore the old behavior.

[[441675 2129]

[ 3379 36994]]

precision recall f1-score support

False 0.99 1.00 0.99 443804

True 0.95 0.92 0.93 40373

accuracy 0.99 484177

macro avg 0.97 0.96 0.96 484177

weighted avg 0.99 0.99 0.99 484177

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

For file winter1Mod the best model is {XGBClassifier(base\_score=0.5, booster='gbtree', colsample\_bylevel=1,

colsample\_bynode=1, colsample\_bytree=1, enable\_categorical=False,

gamma=0, gpu\_id=-1, importance\_type=None,

interaction\_constraints='', learning\_rate=0.300000012,

max\_delta\_step=0, max\_depth=6, min\_child\_weight=1, missing=nan,

monotone\_constraints='()', n\_estimators=100, n\_jobs=48,

num\_parallel\_tree=1, predictor='auto', random\_state=0,

reg\_alpha=0, reg\_lambda=1, scale\_pos\_weight=1, subsample=1,

tree\_method='exact', validate\_parameters=1, verbosity=None): [0.9307134950186172, 0.9886239949439978, 0.9455818827799504, 0.9163054516632403, 0.9662716621396612, array([[441675, 2129],

[ 3379, 36994]], dtype=int64), ' precision recall f1-score support\n\n False 0.99 1.00 0.99 443804\n True 0.95 0.92 0.93 40373\n\n accuracy 0.99 484177\n macro avg 0.97 0.96 0.96 484177\nweighted avg 0.99 0.99 0.99 484177\n'], RandomForestClassifier(class\_weight='balanced', n\_jobs=16, random\_state=42): [0.8977365307640537, 0.9821119136183668, 0.8577617328519855, 0.9416193991033611, 0.9330881410944418, array([[437500, 6304],

[ 2357, 38016]], dtype=int64), ' precision recall f1-score support\n\n False 0.99 0.99 0.99 443804\n True 0.86 0.94 0.90 40373\n\n accuracy 0.98 484177\n macro avg 0.93 0.96 0.94 484177\nweighted avg 0.98 0.98 0.98 484177\n']}

C:\Users\ayrisbud\Downloads\aldaPipeline\Econet>





