

In [1]:

```
#Installing all necessary packages
!pip install snowflake-connector-python[pandas]
!pip install holidays
!pip install --upgrade snowflake-sqlalchemy
!pip install imblearn
!pip install catboost
!pip install probatus
from probatus.feature_elimination import ShapRFEVCV
!pip install tensorflow
!pip install ipywidgets
!pip install pandas_profiling
!pip install optuna

import optuna
from sklearn.metrics import mean_squared_error
from pandas_profiling import ProfileReport
import snowflake.connector
import numpy as np
import pandas as pd
import datetime as dt
import boto3
import base64
from numpy import dtype, isnan, sqrt
from sqlalchemy import create_engine
from sqlalchemy import pool
from sqlalchemy.dialects import registry
from cryptography.hazmat.backends import default_backend
from cryptography.hazmat.primitives.asymmetric import rsa
from cryptography.hazmat.primitives.asymmetric import dsa
from cryptography.hazmat.primitives import serialization
from botocore.exceptions import ClientError
from snowflake.connector.pandas_tools import write_pandas, pd_writer
from dateutil.easter import *
from pathlib import Path
from imblearn.over_sampling import SMOTE
import gc
import dask.dataframe as dd
import dask
from sklearn.linear_model import LogisticRegression
from sklearn.svm import SVC
from sklearn.model_selection import KFold, StratifiedKFold
from sklearn.model_selection import cross_val_score
from sklearn.preprocessing import StandardScaler
from sklearn.preprocessing import LabelEncoder, OneHotEncoder
from datetime import datetime
from sklearn.ensemble import RandomForestClassifier
from sklearn.naive_bayes import GaussianNB
from sklearn.model_selection import KFold, StratifiedKFold
from sklearn.metrics import roc_auc_score
from sklearn.model_selection import train_test_split
from catboost import Pool, CatBoostClassifier
import catboost
import time
```

```
from sklearn.ensemble import RandomForestClassifier
from sklearn.metrics import f1_score, log_loss
from imblearn.over_sampling import SMOTE
from imblearn.under_sampling import RandomUnderSampler
from imblearn.pipeline import Pipeline

import matplotlib.pyplot as plt
from sklearn.datasets import make_classification
from sklearn.metrics import plot_confusion_matrix
from sklearn.svm import SVC

import seaborn as sns
import shap
from catboost import CatBoostClassifier

from sklearn.metrics import accuracy_score
from sklearn.metrics import classification_report
from optuna.samplers import TPESampler
import catboost
import pickle

!jupyter nbextension enable --py widgetsnbextension

%matplotlib inline
```

Looking in indexes: <https://pypi.org/simple>, <https://pip.repos.neuron.amazonaws.com>
 Requirement already satisfied: snowflake-connector-python[pandas] in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (2.8.0)
 Requirement already satisfied: setuptools>34.0.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from snowflake-connector-python[pandas]) (59.4.0)
 Requirement already satisfied: pytz in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from snowflake-connector-python[pandas]) (2021.3)
 Requirement already satisfied: idna<4,>=2.5 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from snowflake-connector-python[pandas]) (3.1)
 Requirement already satisfied: urllib3<1.27,>=1.21.1 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from snowflake-connector-python[pandas]) (1.26.8)
 Requirement already satisfied: filelock<4,>=3.5 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from snowflake-connector-python[pandas]) (3.8.0)
 Requirement already satisfied: pyjwt<3.0.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from snowflake-connector-python[pandas]) (2.5.0)
 Requirement already satisfied: requests<3.0.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from snowflake-connector-python[pandas]) (2.26.0)
 Requirement already satisfied: typing-extensions<5,>=4.3 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from snowflake-connector-python[pandas]) (4.3.0)
 Requirement already satisfied: pyOpenSSL<23.0.0,>=16.2.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from snowflake-connector-python[pandas]) (21.0.0)
 Requirement already satisfied: certifi>=2017.4.17 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from snowflake-connector-python[pandas]) (2021.10.8)
 Requirement already satisfied: asn1crypto<2.0.0,>0.24.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from snowflake-connector-python[pandas]) (1.5.1)
 Requirement already satisfied: charset-normalizer<3,>=2 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from snowflake-connector-python[pandas]) (2.0.8)
 Requirement already satisfied: cffi<2.0.0,>=1.9 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from snowflake-connector-python[pandas]) (1.15.0)
 Requirement already satisfied: cryptography<37.0.0,>=3.1.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from snowflake-connector-python[pandas]) (36.0.0)
 Requirement already satisfied: oscrypto<2.0.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from snowflake-connector-python[pandas]) (1.3.0)
 Requirement already satisfied: pycryptodome!=3.5.0,<4.0.0,>=3.2 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from snowflake-connector-python[pandas]) (3.15.0)
 Requirement already satisfied: pandas<1.5.0,>=1.0.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from snowflake-connector-python[pandas]) (1.3.4)
 Requirement already satisfied: pyarrow<8.1.0,>=8.0.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from snowflake-connector-python[pandas]) (8.0.0)
 Requirement already satisfied: pycparser in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from cffi<2.0.0,>=1.9->snowflake-connector-python[pandas]) (2.21)
 Requirement already satisfied: python-dateutil>=2.7.3 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from pandas<1.5.0,>=1.0.0->snowflake-connector-python[pandas]) (2.8.2)
 Requirement already satisfied: numpy>=1.17.3 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from pandas<1.5.0,>=1.0.0->snowflake-connector-python[pandas]) (1.20.3)
 Requirement already satisfied: six>=1.5.2 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from pyOpenSSL<23.0.0,>=16.2.0->snowflake-connector-python[pandas]) (1.16.0)

[notice] A new release of pip available: 22.2.2 -> 22.3.1

[notice] To update, run: `pip install --upgrade pip`

Looking in indexes: <https://pypi.org/simple>, <https://pip.repos.neuron.amazonaws.com>

Requirement already satisfied: holidays in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (0.16)

Requirement already satisfied: python-dateutil in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from holidays) (2.8.2)

Requirement already satisfied: hijri-converter in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from holidays) (2.2.4)

Requirement already satisfied: convertdate>=2.3.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from holidays) (2.4.0)

Requirement already satisfied: korean-lunar-calendar in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from holidays) (0.3.1)
Requirement already satisfied: pymeeus<=1,>=0.3.13 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from convertdate>=2.3.0->holidays) (0.5.11)
Requirement already satisfied: six>=1.5 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from python-dateutil->holidays) (1.16.0)

[notice] A new release of pip available: 22.2.2 -> 22.3.1

[notice] To update, run: `pip install --upgrade pip`

Looking in indexes: <https://pypi.org/simple>, <https://pip.repos.neuron.amazonaws.com>

Requirement already satisfied: snowflake-sqlalchemy in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (1.4.4)

Requirement already satisfied: sqlalchemy<2.0.0,>=1.4.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from snowflake-sqlalchemy) (1.4.27)

Requirement already satisfied: snowflake-connector-python<3.0.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from snowflake-sqlalchemy) (2.8.0)

Requirement already satisfied: requests<3.0.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from snowflake-connector-python<3.0.0->snowflake-sqlalchemy) (2.26.0)

Requirement already satisfied: asn1crypto<2.0.0,>0.24.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from snowflake-connector-python<3.0.0->snowflake-sqlalchemy) (1.5.1)

Requirement already satisfied: pyjwt<3.0.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from snowflake-connector-python<3.0.0->snowflake-sqlalchemy) (2.5.0)

Requirement already satisfied: pycryptodomex!=3.5.0,<4.0.0,>=3.2 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from snowflake-connector-python<3.0.0->snowflake-sqlalchemy) (3.15.0)

Requirement already satisfied: pytz in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from snowflake-connector-python<3.0.0->snowflake-sqlalchemy) (2021.3)

Requirement already satisfied: urllib3<1.27,>=1.21.1 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from snowflake-connector-python<3.0.0->snowflake-sqlalchemy) (1.26.8)

Requirement already satisfied: charset-normalizer<3,>=2 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from snowflake-connector-python<3.0.0->snowflake-sqlalchemy) (2.0.8)

Requirement already satisfied: filelock<4,>=3.5 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from snowflake-connector-python<3.0.0->snowflake-sqlalchemy) (3.8.0)

Requirement already satisfied: typing-extensions<5,>=4.3 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from snowflake-connector-python<3.0.0->snowflake-sqlalchemy) (4.3.0)

Requirement already satisfied: cffi<2.0.0,>=1.9 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from snowflake-connector-python<3.0.0->snowflake-sqlalchemy) (1.15.0)

Requirement already satisfied: oscrypto<2.0.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from snowflake-connector-python<3.0.0->snowflake-sqlalchemy) (1.3.0)

Requirement already satisfied: idna<4,>=2.5 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from snowflake-connector-python<3.0.0->snowflake-sqlalchemy) (3.1)

Requirement already satisfied: cryptography<37.0.0,>=3.1.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from snowflake-connector-python<3.0.0->snowflake-sqlalchemy) (36.0.0)

Requirement already satisfied: setuptools>34.0.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from snowflake-connector-python<3.0.0->snowflake-sqlalchemy) (59.4.0)

Requirement already satisfied: pyOpenSSL<23.0.0,>=16.2.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from snowflake-connector-python<3.0.0->snowflake-sqlalchemy) (21.0.0)

Requirement already satisfied: certifi>=2017.4.17 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from snowflake-connector-python<3.0.0->snowflake-sqlalchemy) (2021.10.8)

Requirement already satisfied: greenlet!=0.4.17 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from sqlalchemy<2.0.0,>=1.4.0->snowflake-sqlalchemy) (1.1.2)

Requirement already satisfied: pycparser in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from cffi<2.0.0,>=1.9->snowflake-connector-python<3.0.0->snowflake-sqlalchemy) (2.21)

Requirement already satisfied: six>=1.5.2 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from pyOpenSSL<23.0.0,>=16.2.0->snowflake-connector-python<3.0.0->snowflake-sqlalchemy) (1.16.0)

[notice] A new release of pip available: 22.2.2 -> 22.3.1

[notice] To update, run: `pip install --upgrade pip`

Looking in indexes: <https://pypi.org/simple>, <https://pip.repos.neuron.amazonaws.com>

Requirement already satisfied: imblearn in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (0.0)
 Requirement already satisfied: imbalanced-learn in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from imblearn) (0.9.1)
 Requirement already satisfied: joblib>=1.0.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from imbalanced-learn->imblearn) (1.1.0)
 Requirement already satisfied: scikit-learn>=1.1.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from imbalanced-learn->imblearn) (1.1.2)
 Requirement already satisfied: threadpoolctl>=2.0.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from imbalanced-learn->imblearn) (3.0.0)
 Requirement already satisfied: numpy>=1.17.3 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from imbalanced-learn->imblearn) (1.20.3)
 Requirement already satisfied: scipy>=1.3.2 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from imbalanced-learn->imblearn) (1.8.1)

[notice] A new release of pip available: 22.2.2 -> 22.3.1

[notice] To update, run: `pip install --upgrade pip`

Looking in indexes: <https://pypi.org/simple>, <https://pip.repos.neuron.amazonaws.com>

Requirement already satisfied: catboost in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (1.1)
 Requirement already satisfied: matplotlib in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from catboost) (3.5.0)
 Requirement already satisfied: six in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from catboost) (1.16.0)
 Requirement already satisfied: pandas>=0.24.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from catboost) (1.3.4)
 Requirement already satisfied: graphviz in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from catboost) (0.20.1)
 Requirement already satisfied: scipy in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from catboost) (1.8.1)
 Requirement already satisfied: plotly in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from catboost) (5.6.0)
 Requirement already satisfied: numpy>=1.16.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from catboost) (1.20.3)
 Requirement already satisfied: python-dateutil>=2.7.3 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from pandas>=0.24.0->catboost) (2.8.2)
 Requirement already satisfied: pytz>=2017.3 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from pandas>=0.24.0->catboost) (2021.3)
 Requirement already satisfied: fonttools>=4.22.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from matplotlib->catboost) (4.28.2)
 Requirement already satisfied: kiwisolver>=1.0.1 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from matplotlib->catboost) (1.3.2)
 Requirement already satisfied: packaging>=20.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from matplotlib->catboost) (21.3)
 Requirement already satisfied: cycycler>=0.10 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from matplotlib->catboost) (0.11.0)
 Requirement already satisfied: pyparsing>=2.2.1 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from matplotlib->catboost) (3.0.6)
 Requirement already satisfied: pillow>=6.2.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from matplotlib->catboost) (9.0.1)
 Requirement already satisfied: tenacity>=6.2.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from plotly->catboost) (8.0.1)

[notice] A new release of pip available: 22.2.2 -> 22.3.1

[notice] To update, run: `pip install --upgrade pip`

Looking in indexes: <https://pypi.org/simple>, <https://pip.repos.neuron.amazonaws.com>

Requirement already satisfied: probatus in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (1.8.9)
 Requirement already satisfied: shap<0.39.0, >=0.38.1 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from probatus) (0.38.1)
 Requirement already satisfied: numpy>=1.19.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from probatus) (1.20.3)
 Requirement already satisfied: matplotlib>=3.1.1 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from probatus) (3.5.0)
 Requirement already satisfied: scikit-learn>=0.22.2 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from probatus) (1.1.2)
 Requirement already satisfied: scipy>=1.4.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from probatus) (1.8.1)
 Requirement already satisfied: joblib>=0.13.2 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from probatus) (1.1.0)
 Requirement already satisfied: pandas>=1.0.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from probatus) (1.3.4)
 Requirement already satisfied: tqdm>=4.41.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from probatus) (4.62.3)

Requirement already satisfied: pyparsing>=2.2.1 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from matplotlib>=3.1.1->probatus) (3.0.6)

Requirement already satisfied: kiwisolver>=1.0.1 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from matplotlib>=3.1.1->probatus) (1.3.2)

Requirement already satisfied: pillow>=6.2.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from matplotlib>=3.1.1->probatus) (9.0.1)

Requirement already satisfied: fonttools>=4.22.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from matplotlib>=3.1.1->probatus) (4.28.2)

Requirement already satisfied: cyclor>=0.10 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from matplotlib>=3.1.1->probatus) (0.11.0)

Requirement already satisfied: python-dateutil>=2.7 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from matplotlib>=3.1.1->probatus) (2.8.2)

Requirement already satisfied: packaging>=20.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from matplotlib>=3.1.1->probatus) (21.3)

Requirement already satisfied: pytz>=2017.3 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from pandas>=1.0.0->probatus) (2021.3)

Requirement already satisfied: threadpoolctl>=2.0.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from scikit-learn>=0.22.2->probatus) (3.0.0)

Requirement already satisfied: cloudpickle in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from shap<0.39.0,>=0.38.1->probatus) (2.0.0)

Requirement already satisfied: numba in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from shap<0.39.0,>=0.38.1->probatus) (0.54.1)

Requirement already satisfied: slicer==0.0.7 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from shap<0.39.0,>=0.38.1->probatus) (0.0.7)

Requirement already satisfied: six>=1.5 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from python-dateutil>=2.7->matplotlib>=3.1.1->probatus) (1.16.0)

Requirement already satisfied: setuptools in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from numba->shap<0.39.0,>=0.38.1->probatus) (59.4.0)

Requirement already satisfied: llvmlite<0.38,>=0.37.0rc1 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from numba->shap<0.39.0,>=0.38.1->probatus) (0.37.0)

[notice] A new release of pip available: 22.2.2 -> 22.3.1

[notice] To update, run: `pip install --upgrade pip`

2022-11-23 16:11:10.670518: I tensorflow/core/platform/cpu_feature_guard.cc:193] This TensorFlow binary is optimized with oneAPI Deep Neural Network Library (oneDNN) to use the following CPU instructions in performance-critical operations: AVX2 AVX512F FMA To enable them in other operations, rebuild TensorFlow with the appropriate compiler flags.

2022-11-23 16:11:11.537706: E tensorflow/stream_executor/cuda/cuda_blas.cc:2981] Unable to register cuBLAS factory: Attempting to register factory for plugin cuBLAS when one has already been registered

```
-----
RuntimeError                                Traceback (most recent call last)
RuntimeError: module compiled against API version 0xe but this version of numpy is 0xd
```

```
-----
RuntimeError                                Traceback (most recent call last)
RuntimeError: module compiled against API version 0xe but this version of numpy is 0xd
```

```
-----
RuntimeError                                Traceback (most recent call last)
RuntimeError: module compiled against API version 0xe but this version of numpy is 0xd
```

Looking in indexes: <https://pypi.org/simple>, <https://pip.repos.neuron.amazonaws.com>
Requirement already satisfied: tensorflow in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (2.10.0)
Requirement already satisfied: grpcio<2.0,>=1.24.3 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from tensorflow) (1.49.1)
Requirement already satisfied: typing-extensions>=3.6.6 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from tensorflow) (4.3.0)
Requirement already satisfied: tensorboard<2.11,>=2.10 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from tensorflow) (2.10.1)
Requirement already satisfied: opt-einsum>=2.3.2 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from tensorflow) (3.3.0)
Requirement already satisfied: termcolor>=1.1.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from tensorflow) (2.0.1)
Requirement already satisfied: keras<2.11,>=2.10.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from tensorflow) (2.10.0)
Requirement already satisfied: astunparse>=1.6.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from tensorflow) (1.6.3)
Requirement already satisfied: numpy>=1.20 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from tensorflow) (1.20.3)
Requirement already satisfied: gast<=0.4.0,>=0.2.1 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from tensorflow) (0.4.0)
Requirement already satisfied: tensorflow-estimator<2.11,>=2.10.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from tensorflow) (2.10.0)
Requirement already satisfied: flatbuffers>=2.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from tensorflow) (22.9.24)
Requirement already satisfied: packaging in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from tensorflow) (21.3)
Requirement already satisfied: tensorflow-io-gcs-filesystem>=0.23.1 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from tensorflow) (0.27.0)
Requirement already satisfied: libclang>=13.0.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from tensorflow) (14.0.6)
Requirement already satisfied: protobuf<3.20,>=3.9.2 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from tensorflow) (3.19.1)
Requirement already satisfied: six>=1.12.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from tensorflow) (1.16.0)
Requirement already satisfied: keras-preprocessing>=1.1.1 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from tensorflow) (1.1.2)
Requirement already satisfied: setuptools in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from tensorflow) (59.4.0)
Requirement already satisfied: wrapt>=1.11.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from tensorflow) (1.13.3)
Requirement already satisfied: google-pasta>=0.1.1 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from tensorflow) (0.2.0)
Requirement already satisfied: h5py>=2.9.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from tensorflow) (3.3.0)
Requirement already satisfied: absl-py>=1.0.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from tensorflow) (1.2.0)
Requirement already satisfied: wheel<1.0,>=0.23.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from tensorflow) (0.37.0)
Requirement already satisfied: requests<3,>=2.21.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from tensorflow) (2.26.0)
Requirement already satisfied: tensorboard-data-server<0.7.0,>=0.6.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from tensorflow) (0.6.1)
Requirement already satisfied: google-auth-oauthlib<0.5,>=0.4.1 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from tensorflow) (0.4.6)
Requirement already satisfied: google-auth<3,>=1.6.3 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from tensorflow) (2.12.0)
Requirement already satisfied: werkzeug>=1.0.1 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from tensorflow) (2.0.3)
Requirement already satisfied: tensorboard-plugin-wit>=1.6.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from tensorflow) (1.8.1)
Requirement already satisfied: markdown>=2.6.8 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from tensorflow) (3.4.1)
Requirement already satisfied: pyparsing!=3.0.5,>=2.0.2 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from tensorflow) (3.0.9)


```
>tensorflow) (3.0.6)
Requirement already satisfied: pyasn1-modules<=0.2.1 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from google-auth<3,>=1.6.3->tensorboard<2.11,>=2.10->tensorflow) (0.2.8)
Requirement already satisfied: rsa<5,>=3.1.4 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from google-auth<3,>=1.6.3->tensorboard<2.11,>=2.10->tensorflow) (4.7.2)
Requirement already satisfied: cachetools<6.0,>=2.0.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from google-auth<3,>=1.6.3->tensorboard<2.11,>=2.10->tensorflow) (5.2.0)
Requirement already satisfied: requests-oauthlib<=0.7.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from google-auth-oauthlib<0.5,>=0.4.1->tensorboard<2.11,>=2.10->tensorflow) (1.3.1)
Requirement already satisfied: importlib-metadata<=4.4 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from markdown<=2.6.8->tensorboard<2.11,>=2.10->tensorflow) (4.8.2)
Requirement already satisfied: certifi<=2017.4.17 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from requests<3,>=2.2.1.0->tensorboard<2.11,>=2.10->tensorflow) (2021.10.8)
Requirement already satisfied: charset-normalizer<=2.0.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from requests<3,>=2.2.1.0->tensorboard<2.11,>=2.10->tensorflow) (2.0.8)
Requirement already satisfied: idna<4,>=2.5 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from requests<3,>=2.2.1.0->tensorboard<2.11,>=2.10->tensorflow) (3.1)
Requirement already satisfied: urllib3<1.27,>=1.21.1 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from requests<3,>=2.2.1.0->tensorboard<2.11,>=2.10->tensorflow) (1.26.8)
Requirement already satisfied: zipp<=0.5 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from importlib-metadata<=4.4->markdown<=2.6.8->tensorboard<2.11,>=2.10->tensorflow) (3.6.0)
Requirement already satisfied: pyasn1<0.5.0,>=0.4.6 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from pyasn1-modules<=0.2.1->google-auth<3,>=1.6.3->tensorboard<2.11,>=2.10->tensorflow) (0.4.8)
Requirement already satisfied: oauthlib<=3.0.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from requests-oauthlib<=0.7.0->google-auth-oauthlib<0.5,>=0.4.1->tensorboard<2.11,>=2.10->tensorflow) (3.2.1)
```

[notice] A new release of pip available: 22.2.2 -> 22.3.1

[notice] To update, run: `pip install --upgrade pip`

Looking in indexes: <https://pypi.org/simple>, <https://pip.repos.neuron.amazonaws.com>

```
Requirement already satisfied: ipywidgets in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (7.6.5)
Requirement already satisfied: widgetsnbextension<=3.5.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from ipywidgets) (3.5.2)
Requirement already satisfied: jupyterlab-widgets<=1.0.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from ipywidgets) (1.0.2)
Requirement already satisfied: nbformat<=4.2.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from ipywidgets) (5.1.3)
Requirement already satisfied: ipython<=4.0.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from ipywidgets) (7.32.0)
Requirement already satisfied: ipykernel<=4.5.1 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from ipywidgets) (6.5.1)
Requirement already satisfied: ipython-genutils<=0.2.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from ipywidgets) (0.2.0)
Requirement already satisfied: traitlets<=4.3.1 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from ipywidgets) (5.1.1)
Requirement already satisfied: tornado<7.0,>=4.2 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from ipykernel<=4.5.1->ipywidgets) (6.1)
Requirement already satisfied: jupyter-client<8.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from ipykernel<=4.5.1->ipywidgets) (7.1.0)
Requirement already satisfied: debugpy<2.0,>=1.0.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from ipykernel<=4.5.1->ipywidgets) (1.5.1)
Requirement already satisfied: matplotlib-inline<0.2.0,>=0.1.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from ipykernel<=4.5.1->ipywidgets) (0.1.3)
Requirement already satisfied: decorator in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from ipython<=4.0.0->ipywidgets) (5.1.0)
Requirement already satisfied: pexpect<4.3 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from ipython<=4.0.0->ipywidgets) (4.8.0)
Requirement already satisfied: pickleshare in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from ipython<=4.0.0->ipywidgets) (0.7.5)
Requirement already satisfied: jedi<=0.16 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from ipython<=4.0.0->ipywidgets)
```

ts) (0.18.1)
Requirement already satisfied: backcall in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from ipython>=4.0.0->ipywidget
s) (0.2.0)
Requirement already satisfied: prompt-toolkit!=3.0.0,!<3.0.1,<3.1.0,>=2.0.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packa
ges (from ipython>=4.0.0->ipywidgets) (3.0.22)
Requirement already satisfied: pygments in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from ipython>=4.0.0->ipywidget
s) (2.10.0)
Requirement already satisfied: setuptools>=18.5 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from ipython>=4.0.0->ip
ywidgets) (59.4.0)
Requirement already satisfied: jupyter-core in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from nbformat>=4.2.0->ipywi
dgets) (4.9.1)
Requirement already satisfied: jsonschema!=2.5.0,>=2.4 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from nbformat>=
4.2.0->ipywidgets) (3.2.0)
Requirement already satisfied: notebook>=4.4.1 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from widgetsnbextension~
=3.5.0->ipywidgets) (6.4.6)
Requirement already satisfied: parso<0.9.0,>=0.8.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from jedi>=0.16->ipy
thon>=4.0.0->ipywidgets) (0.8.2)
Requirement already satisfied: six>=1.11.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from jsonschema!=2.5.0,>=2.4
->nbformat>=4.2.0->ipywidgets) (1.16.0)
Requirement already satisfied: pyrsistent>=0.14.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from jsonschema!=2.5.
0,>=2.4->nbformat>=4.2.0->ipywidgets) (0.18.0)
Requirement already satisfied: attrs>=17.4.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from jsonschema!=2.5.0,>=
2.4->nbformat>=4.2.0->ipywidgets) (21.2.0)
Requirement already satisfied: python-dateutil>=2.1 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from jupyter-client
<8.0->ipykernel>=4.5.1->ipywidgets) (2.8.2)
Requirement already satisfied: pyzmq>=13 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from jupyter-client<8.0->ipyke
rnel>=4.5.1->ipywidgets) (22.3.0)
Requirement already satisfied: nest-asyncio>=1.5 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from jupyter-client<8.
0->ipykernel>=4.5.1->ipywidgets) (1.5.1)
Requirement already satisfied: entrypoints in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from jupyter-client<8.0->ipy
kernel>=4.5.1->ipywidgets) (0.3)
Requirement already satisfied: argon2-cffi in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from notebook>=4.4.1->widget
snbextension~=3.5.0->ipywidgets) (21.1.0)
Requirement already satisfied: Send2Trash>=1.8.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from notebook>=4.4.1->
widgetsnbextension~=3.5.0->ipywidgets) (1.8.0)
Requirement already satisfied: prometheus-client in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from notebook>=4.4.1->
widgetsnbextension~=3.5.0->ipywidgets) (0.12.0)
Requirement already satisfied: Jinja2 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from notebook>=4.4.1->widgetsnbex
tension~=3.5.0->ipywidgets) (3.0.3)
Requirement already satisfied: nbconvert in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from notebook>=4.4.1->widgetsn
bextension~=3.5.0->ipywidgets) (6.3.0)
Requirement already satisfied: terminado>=0.8.3 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from notebook>=4.4.1->w
idgetsnbextension~=3.5.0->ipywidgets) (0.12.1)
Requirement already satisfied: ptyprocess>=0.5 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from pexpect>4.3->ipytho
n>=4.0.0->ipywidgets) (0.7.0)
Requirement already satisfied: wcwidth in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from prompt-toolkit!=3.0.0,!<3.
0.1,<3.1.0,>=2.0.0->ipython>=4.0.0->ipywidgets) (0.2.5)
Requirement already satisfied: cffi>=1.0.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from argon2-cffi->notebook>=
4.4.1->widgetsnbextension~=3.5.0->ipywidgets) (1.15.0)
Requirement already satisfied: MarkupSafe>=2.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from Jinja2->notebook>=
4.4.1->widgetsnbextension~=3.5.0->ipywidgets) (2.0.1)
Requirement already satisfied: bleach in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from nbconvert->notebook>=4.4.1->
widgetsnbextension~=3.5.0->ipywidgets) (4.1.0)
Requirement already satisfied: nbclient<0.6.0,>=0.5.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from nbconvert->n
otebook>=4.4.1->widgetsnbextension~=3.5.0->ipywidgets) (0.5.9)
Requirement already satisfied: defusedxml in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from nbconvert->notebook>=4.
4.1->widgetsnbextension~=3.5.0->ipywidgets) (0.7.1)

Requirement already satisfied: mistune<2,>=0.8.1 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from nbconvert->notebook>=4.4.1->widgetsnbextension~=3.5.0->ipywidgets) (0.8.4)
Requirement already satisfied: jupyterlab-pygments in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from nbconvert->notebook>=4.4.1->widgetsnbextension~=3.5.0->ipywidgets) (0.1.2)
Requirement already satisfied: testpath in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from nbconvert->notebook>=4.4.1->widgetsnbextension~=3.5.0->ipywidgets) (0.5.0)
Requirement already satisfied: pandocfilters>=1.4.1 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from nbconvert->notebook>=4.4.1->widgetsnbextension~=3.5.0->ipywidgets) (1.5.0)
Requirement already satisfied: pycparser in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from cffi>=1.0.0->argon2-cffi->notebook>=4.4.1->widgetsnbextension~=3.5.0->ipywidgets) (2.21)
Requirement already satisfied: packaging in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from bleach->nbconvert->notebook>=4.4.1->widgetsnbextension~=3.5.0->ipywidgets) (21.3)
Requirement already satisfied: webencodings in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from bleach->nbconvert->notebook>=4.4.1->widgetsnbextension~=3.5.0->ipywidgets) (0.5.1)
Requirement already satisfied: pyparsing!=3.0.5,>=2.0.2 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from packaging->bleach->nbconvert->notebook>=4.4.1->widgetsnbextension~=3.5.0->ipywidgets) (3.0.6)

[notice] A new release of pip available: 22.2.2 -> 22.3.1

[notice] To update, run: pip install --upgrade pip

Looking in indexes: <https://pypi.org/simple>, <https://pip.repos.neuron.amazonaws.com>

Requirement already satisfied: pandas_profiling in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (3.3.0)

Requirement already satisfied: matplotlib<3.6,>=3.2 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from pandas_profiling) (3.5.0)

Requirement already satisfied: tqdm<4.65,>=4.48.2 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from pandas_profiling) (4.62.3)

Requirement already satisfied: statsmodels<0.14,>=0.13.2 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from pandas_profiling) (0.13.2)

Requirement already satisfied: tangled-up-in-unicode==0.2.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from pandas_profiling) (0.2.0)

Requirement already satisfied: seaborn<0.12,>=0.10.1 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from pandas_profiling) (0.11.2)

Requirement already satisfied: htmlmin==0.1.12 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from pandas_profiling) (0.1.12)

Requirement already satisfied: missingno<0.6,>=0.4.2 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from pandas_profiling) (0.5.1)

Requirement already satisfied: Jinja2<3.2,>=2.11.1 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from pandas_profiling) (3.0.3)

Requirement already satisfied: pydantic<1.10,>=1.8.1 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from pandas_profiling) (1.9.2)

Requirement already satisfied: multimethod<1.9,>=1.4 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from pandas_profiling) (1.8)

Requirement already satisfied: scipy<1.10,>=1.4.1 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from pandas_profiling) (1.8.1)

Requirement already satisfied: phik<0.13,>=0.11.1 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from pandas_profiling) (0.12.2)

Requirement already satisfied: PyYAML<6.1,>=5.0.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from pandas_profiling) (5.4.1)

Requirement already satisfied: requests<2.29,>=2.24.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from pandas_profiling) (2.26.0)

Requirement already satisfied: joblib~=1.1.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from pandas_profiling) (1.1.0)

Requirement already satisfied: numpy<1.24,>=1.16.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from pandas_profiling) (1.20.3)

Requirement already satisfied: visions[type_image_path]==0.7.5 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from pandas_profiling) (0.7.5)

Requirement already satisfied: pandas!=1.4.0,<1.5,>1.1 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from pandas_profiling) (1.3.4)

Requirement already satisfied: networkx>=2.4 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from visions[type_image_path]==0.7.5->pandas_profiling) (2.6.3)

Requirement already satisfied: attrs>=19.3.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from visions[type_image_path]==0.7.5->pandas_profiling) (21.2.0)

Requirement already satisfied: imagehash in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from visions[type_image_path]=0.7.5->pandas_profiling) (4.3.1)

Requirement already satisfied: Pillow in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from visions[type_image_path]==0.7.5->pandas_profiling) (9.0.1)

Requirement already satisfied: MarkupSafe>=2.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from jinja2<3.2,>=2.11.1->pandas_profiling) (2.0.1)

Requirement already satisfied: kiwisolver>=1.0.1 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from matplotlib<3.6,>=3.2->pandas_profiling) (1.3.2)

Requirement already satisfied: cycycler>=0.10 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from matplotlib<3.6,>=3.2->pandas_profiling) (0.11.0)

Requirement already satisfied: pyparsing>=2.2.1 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from matplotlib<3.6,>=3.2->pandas_profiling) (3.0.6)

Requirement already satisfied: fonttools>=4.22.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from matplotlib<3.6,>=3.2->pandas_profiling) (4.28.2)

Requirement already satisfied: packaging>=20.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from matplotlib<3.6,>=3.2->pandas_profiling) (21.3)

Requirement already satisfied: python-dateutil>=2.7 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from matplotlib<3.6,>=3.2->pandas_profiling) (2.8.2)

Requirement already satisfied: pytz>=2017.3 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from pandas!=1.4.0,<1.5,>1.1->pandas_profiling) (2021.3)

Requirement already satisfied: typing-extensions>=3.7.4.3 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from pydantic<1.10,>=1.8.1->pandas_profiling) (4.3.0)

Requirement already satisfied: certifi>=2017.4.17 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from requests<2.29,>=2.24.0->pandas_profiling) (2021.10.8)

Requirement already satisfied: charset-normalizer~=2.0.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from requests<2.29,>=2.24.0->pandas_profiling) (2.0.8)

Requirement already satisfied: idna<4,>=2.5 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from requests<2.29,>=2.24.0->pandas_profiling) (3.1)

Requirement already satisfied: urllib3<1.27,>=1.21.1 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from requests<2.29,>=2.24.0->pandas_profiling) (1.26.8)

Requirement already satisfied: patsy>=0.5.2 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from statsmodels<0.14,>=0.13.2->pandas_profiling) (0.5.2)

Requirement already satisfied: six in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from patsy>=0.5.2->statsmodels<0.14,>=0.13.2->pandas_profiling) (1.16.0)

Requirement already satisfied: PyWavelets in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from imagehash->visions[type_image_path]==0.7.5->pandas_profiling) (1.2.0)

[notice] A new release of pip available: 22.2.2 -> 22.3.1

[notice] To update, run: `pip install --upgrade pip`

Looking in indexes: <https://pypi.org/simple>, <https://pip.repos.neuron.amazonaws.com>

Requirement already satisfied: optuna in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (3.0.2)

Requirement already satisfied: cmaes>=0.8.2 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from optuna) (0.8.2)

Requirement already satisfied: scipy<1.9.0,>=1.7.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from optuna) (1.8.1)

Requirement already satisfied: cliff in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from optuna) (4.0.0)

Requirement already satisfied: colorlog in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from optuna) (6.7.0)

Requirement already satisfied: sqlalchemy>=1.3.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from optuna) (1.4.27)

Requirement already satisfied: alembic>=1.5.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from optuna) (1.8.1)

Requirement already satisfied: PyYAML in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from optuna) (5.4.1)

Requirement already satisfied: tqdm in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from optuna) (4.62.3)

Requirement already satisfied: numpy in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from optuna) (1.20.3)

Requirement already satisfied: packaging>=20.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from optuna) (21.3)

Requirement already satisfied: Mako in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from alembic>=1.5.0->optuna) (1.2.3)

```

Requirement already satisfied: importlib-resources in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from alembic>=1.5.0->optuna) (5.4.0)
Requirement already satisfied: importlib-metadata in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from alembic>=1.5.0->optuna) (4.8.2)
Requirement already satisfied: pyparsing!=3.0.5,>=2.0.2 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from packaging>=20.0->optuna) (3.0.6)
Requirement already satisfied: greenlet!=0.4.17 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from sqlalchemy>=1.3.0->optuna) (1.1.2)
Requirement already satisfied: cmd2>=1.0.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from cliff->optuna) (2.4.2)
Requirement already satisfied: stevedore>=2.0.1 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from cliff->optuna) (4.0.0)
Requirement already satisfied: autopage>=0.4.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from cliff->optuna) (0.5.1)
Requirement already satisfied: PrettyTable>=0.7.2 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from cliff->optuna) (3.4.1)
Requirement already satisfied: wcwidth>=0.1.7 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from cmd2>=1.0.0->cliff->optuna) (0.2.5)
Requirement already satisfied: pyperclip>=1.6 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from cmd2>=1.0.0->cliff->optuna) (1.8.2)
Requirement already satisfied: attrs>=16.3.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from cmd2>=1.0.0->cliff->optuna) (21.2.0)
Requirement already satisfied: zipp>=0.5 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from importlib-metadata->alembic>=1.5.0->optuna) (3.6.0)
Requirement already satisfied: pbr!=2.1.0,>=2.0.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from stevedore>=2.0.1->cliff->optuna) (5.10.0)
Requirement already satisfied: MarkupSafe>=0.9.2 in /home/ec2-user/anaconda3/envs/python3/lib/python3.8/site-packages (from Mako->alembic>=1.5.0->optuna) (2.0.1)

```

[notice] A new release of pip available: 22.2.2 -> 22.3.1

[notice] To update, run: `pip install --upgrade pip`

Config option `kernel_spec_manager_class` not recognized by `EnableNBExtensionApp`.

Enabling notebook extension jupyter-js-widgets/extension...

- Validating: OK

In [2]:

```

# # Connect to snowflake
# ctx = snowflake.connector.connect(
#     user='TWINKLE',
#     password='@Luiwatson7',
#     account='mja29153.us-east-1',
#     warehouse='PROGRESSIVE_LARGE',
#     database='CLIENT',
#     schema='AUDIENCE_GROUP',
# )
# cur = ctx.cursor()

```

In []:

In [3]:

```
# sql = """
# SELECT IMPRESSION_ID ,
# PARTNER_ID , CONVERSION_TYPE,
# CAMPAIGN_ID,
# AD_GROUP_ID,
# AUDIENCE_ID,
# CREATIVE_ID,
# AD_FORMAT, MATCHED_FOLD_POSITION, REGION,CITY,
# ZIP_POSTAL_CODE, SUPPLY_VENDOR, SITE, OS_FAMILY,
# BROWSER,USER_HOUR_OF_WEEK, DEVICE_MODEL from CLIENT.AUDIENCE_GROUP.REDS_IMPRESSIONS i
# LEFT JOIN CLIENT.AUDIENCE_GROUP.REDS_CONVERSIONS c
# ON c.TDID = i.TDID
# AND i.TDID <> '00000000-0000-0000-0000-000000000000'
# WHERE i.CAMPAIGN_ID in ('q5vvxul', 'nov2ir3')

# limit 1000000
# """

# print('Executing impressions query')
# cur.execute(sql)
# # Fetch the result set from the cursor and deliver it as the Pandas DataFrame.
# df = cur.fetch_pandas_all()

# cur.close()
# ctx.close()
```

In [4]:

```
# df.to_pickle('df')
```

In [5]:

```
# #
# Primary Tag
# Tag Name: AG_EL_Ticket_Purchase_Conversion
# Tag ID: 80ignqj
# Secondary Tag
# Tag Name: Cart
# Tag ID: uiifnc9

df = pd.read_pickle('df')
# df

#df = df.sample(frac=0.3)
```

In []:

In [6]:

```
df['Conversion_Binarized'] = np.where(df['CONVERSION_TYPE'] == 'other:80ignqj', '1', '0')
```

In [7]:

```
df['Conversion_Binarized'] = df['Conversion_Binarized'].astype(float)
df['Conversion_Binarized'].value_counts()
```

Out[7]:

```
0.0    982053
1.0    17947
Name: Conversion_Binarized, dtype: int64
```

In [8]:

```
(df.CONVERSION_TYPE == 'other:80ignqj').sum()
```

Out[8]:

```
17947
```

In [9]:

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1000000 entries, 0 to 999999
Data columns (total 19 columns):
#   Column                Non-Null Count  Dtype
---  -
0   IMPRESSION_ID          1000000 non-null object
1   PARTNER_ID             1000000 non-null object
2   CONVERSION_TYPE        1000000 non-null object
3   CAMPAIGN_ID            1000000 non-null object
4   AD_GROUP_ID            1000000 non-null object
5   AUDIENCE_ID            1000000 non-null object
6   CREATIVE_ID            1000000 non-null object
7   AD_FORMAT              1000000 non-null object
8   MATCHED_FOLD_POSITION  1000000 non-null object
9   REGION                 1000000 non-null object
10  CITY                   1000000 non-null object
11  ZIP_POSTAL_CODE        1000000 non-null object
12  SUPPLY_VENDOR          1000000 non-null object
13  SITE                   1000000 non-null object
14  OS_FAMILY              999985 non-null float64
15  BROWSER                 999902 non-null float64
16  USER_HOUR_OF_WEEK      999125 non-null float64
17  DEVICE_MODEL            1000000 non-null object
18  Conversion_Binarized    1000000 non-null float64
dtypes: float64(4), object(15)
memory usage: 145.0+ MB
```

In []:

In [10]:

```
null_vals = df.isnull().sum().apply(lambda x: x * 100 / df.shape[0]).sort_values(ascending = False).loc[lambda x: x > 0]
null_vals
```

Out[10]:

```
USER_HOUR_OF_WEEK    0.0875
BROWSER              0.0098
OS_FAMILY            0.0015
dtype: float64
```

In [11]:

```
#profile = ProfileReport(df, title='Detailed Report Endeavour')
```

In [12]:

```
#profile.to_notebook_iframe()
#profile.to_file("Endeavour.html")
```

In [13]:

```
df = df.dropna()
```

In [14]:

```
##### Data without SMOTE #####
```

In [15]:

```
X = df.drop(['IMPRESSION_ID', 'PARTNER_ID', \
            'CONVERSION_TYPE', 'CAMPAIGN_ID', 'AD_GROUP_ID', 'AUDIENCE_ID', 'CREATIVE_ID', 'CITY', 'ZIP_POSTAL_CODE', 'Conversion_Binarized'
            ], axis=1)
Y = df['Conversion_Binarized']
```

In [16]:

```
Xdata_train, Xdata_val, Ydata_train, Ydata_val = train_test_split(X, Y, test_size=0.3, random_state=0)

categorical_features_indices = np.where(X.dtypes != np.float)[0]
print(categorical_features_indices)

[0 1 2 3 4 8]
```


In [17]:

```

from imblearn.over_sampling import ADASYN, SMOTE, RandomOverSampler
from imblearn.under_sampling import NearMiss, RandomUnderSampler
from imblearn.over_sampling import SMOTENC

#Instantiate SMOTENC algorithm with 60% oversampling
#smt = SMOTENC(categorical_features=categorical_features_indices, random_state=123, sampling_strategy=.6)
#X_ROS, Y_ROS = smt.fit_resample(X, Y)
#smt = SMOTE(random_state=777)
rus = RandomUnderSampler(random_state=777)
#ros = RandomOverSampler(random_state=777)
X_ROS, Y_ROS = rus.fit_resample(X, Y)

```

In [18]:

```

Xdata_train, Xdata_val, Ydata_train, Ydata_val = train_test_split(X_ROS, Y_ROS, test_size=0.3, random_state=0)

categorical_features_indices = np.where(X.dtypes != np.float)[0]
print(categorical_features_indices)

```

[0 1 2 3 4 8]

In [19]:

```

model = catboost.CatBoostClassifier(verbose=False)
model.fit(Xdata_train, Ydata_train, cat_features=categorical_features_indices)

```

Out[19]:

<catboost.core.CatBoostClassifier at 0x7ff06375f700>

In []:

In [20]:

```
y_pred = model.predict(Xdata_val)
```

In [21]:

```
print(classification_report(Ydata_val, y_pred))
```

	precision	recall	f1-score	support
0.0	0.91	0.87	0.89	5360
1.0	0.88	0.92	0.90	5401
accuracy			0.90	10761
macro avg	0.90	0.90	0.90	10761
weighted avg	0.90	0.90	0.90	10761

In [22]:

```
print(f1_score(Ydata_val, y_pred))
```

0.8982989504162141

In [23]:

```
def objective(trial):
    model = catboost.CatBoostClassifier(
        iterations=trial.suggest_int("iterations", 100, 200),
        learning_rate=trial.suggest_float("learning_rate", 1e-3, 1e-1, log=True),
        depth=trial.suggest_int("depth", 4, 10),
        l2_leaf_reg=trial.suggest_float("l2_leaf_reg", 1e-8, 100.0, log=True),
        bootstrap_type=trial.suggest_categorical("bootstrap_type", ["Bayesian"]),
        random_strength=trial.suggest_float("random_strength", 1e-8, 10.0, log=True),
        bagging_temperature=trial.suggest_float("bagging_temperature", 0.0, 10.0),
        od_type=trial.suggest_categorical("od_type", ["IncToDec", "Iter"]),
        od_wait=trial.suggest_int("od_wait", 10, 50),
        loss_function='Logloss',
        eval_metric='AUC',
        nan_mode='Min',
        thread_count=5,
        task_type='CPU',
        custom_metric=['Accuracy', 'F1', 'Precision', 'Recall'],
        class_weights = [0.4, 0.6],
        verbose=True
    )
    model.fit(Xdata_train, Ydata_train, eval_set=(Xdata_val, Ydata_val),
              cat_features=categorical_features_indices,
              verbose_eval=50,
              early_stopping_rounds=30,
              use_best_model=True,
              plot=True)
    y_pred = model.predict(Xdata_val)
    return accuracy_score(Ydata_val, y_pred)
```

In [24]:

```
optuna.logging.set_verbosity(optuna.logging.WARNING)

sampler = TPESampler(seed=1)
study = optuna.create_study(study_name="catboost", direction="maximize", sampler=sampler)
study.optimize(objective, n_trials=10)
```

0:	test: 0.7916182	best: 0.7916182 (0)	total: 6.73ms	remaining: 948ms
50:	test: 0.8815387	best: 0.8815387 (50)	total: 205ms	remaining: 367ms
100:	test: 0.8876019	best: 0.8876019 (100)	total: 395ms	remaining: 160ms
141:	test: 0.8893162	best: 0.8893162 (141)	total: 550ms	remaining: 0us

bestTest = 0.8893162405
bestIteration = 141

0:	test: 0.8204637	best: 0.8204637 (0)	total: 8.55ms	remaining: 1.31s
50:	test: 0.8433739	best: 0.8434852 (49)	total: 301ms	remaining: 608ms
100:	test: 0.8466507	best: 0.8466507 (100)	total: 593ms	remaining: 311ms
150:	test: 0.8790070	best: 0.8790070 (150)	total: 882ms	remaining: 17.5ms
153:	test: 0.8804120	best: 0.8804120 (153)	total: 900ms	remaining: 0us

bestTest = 0.8804119677
bestIteration = 153

0:	test: 0.7310909	best: 0.7310909 (0)	total: 13.3ms	remaining: 1.5s
50:	test: 0.8508389	best: 0.8508389 (50)	total: 382ms	remaining: 471ms
100:	test: 0.8693411	best: 0.8693411 (100)	total: 732ms	remaining: 94.2ms
113:	test: 0.8706025	best: 0.8706025 (113)	total: 823ms	remaining: 0us

bestTest = 0.8706024935
bestIteration = 113

0:	test: 0.7351810	best: 0.7351810 (0)	total: 13.1ms	remaining: 1.33s
50:	test: 0.8366935	best: 0.8366935 (50)	total: 433ms	remaining: 442ms
100:	test: 0.8394009	best: 0.8396747 (89)	total: 846ms	remaining: 16.8ms
102:	test: 0.8395273	best: 0.8396747 (89)	total: 862ms	remaining: 0us

bestTest = 0.8396746767
bestIteration = 89

Shrink model to first 90 iterations.

0:	test: 0.7082770	best: 0.7082770 (0)	total: 5.98ms	remaining: 1s
50:	test: 0.8730356	best: 0.8731459 (49)	total: 203ms	remaining: 470ms
100:	test: 0.8766958	best: 0.8767086 (99)	total: 392ms	remaining: 264ms
150:	test: 0.8813057	best: 0.8813057 (150)	total: 581ms	remaining: 69.2ms
168:	test: 0.8821060	best: 0.8821060 (168)	total: 649ms	remaining: 0us

bestTest = 0.8821060466
bestIteration = 168

0:	test: 0.8157439	best: 0.8157439 (0)	total: 6.45ms	remaining: 928ms
50:	test: 0.8945177	best: 0.8945345 (49)	total: 241ms	remaining: 444ms
100:	test: 0.9010570	best: 0.9010570 (100)	total: 474ms	remaining: 207ms
144:	test: 0.9053701	best: 0.9053701 (144)	total: 679ms	remaining: 0us

bestTest = 0.9053700842
bestIteration = 144

0:	test: 0.7345624	best: 0.7345624 (0)	total: 10.2ms	remaining: 1.5s
50:	test: 0.8732136	best: 0.8732136 (50)	total: 316ms	remaining: 608ms
100:	test: 0.8803731	best: 0.8804085 (99)	total: 612ms	remaining: 291ms
148:	test: 0.8819594	best: 0.8820650 (143)	total: 898ms	remaining: 0us

bestTest = 0.8820649748
bestIteration = 143

Shrink model to first 144 iterations.

0:	test: 0.7236267	best: 0.7236267 (0)	total: 7.21ms	remaining: 1.01s
50:	test: 0.8649040	best: 0.8649040 (50)	total: 273ms	remaining: 481ms
100:	test: 0.8704759	best: 0.8704759 (100)	total: 533ms	remaining: 211ms
140:	test: 0.8727164	best: 0.8727164 (140)	total: 742ms	remaining: 0us

bestTest = 0.8727164089
bestIteration = 140

0:	test: 0.8005908	best: 0.8005908 (0)	total: 7ms	remaining: 791ms
50:	test: 0.8872777	best: 0.8872777 (50)	total: 245ms	remaining: 303ms
100:	test: 0.8910785	best: 0.8910785 (100)	total: 481ms	remaining: 61.8ms
113:	test: 0.8919725	best: 0.8919725 (113)	total: 542ms	remaining: 0us

bestTest = 0.8919725168
bestIteration = 113

0:	test: 0.8081118	best: 0.8081118 (0)	total: 8.99ms	remaining: 1.45s
50:	test: 0.8844944	best: 0.8846061 (49)	total: 255ms	remaining: 556ms
100:	test: 0.8882748	best: 0.8882748 (100)	total: 491ms	remaining: 296ms
150:	test: 0.8908088	best: 0.8908088 (150)	total: 726ms	remaining: 52.9ms
161:	test: 0.8912609	best: 0.8912609 (161)	total: 778ms	remaining: 0us

bestTest = 0.8912609467
bestIteration = 161

In [25]:

```

print("Number of finished trials: ", len(study.trials))
print("Best trial:")
trial = study.best_trial
print("  Value: ", trial.value)
print("  Params: ")
for key, value in trial.params.items():
    print("    {}: {}".format(key, value))

```

Number of finished trials: 10

Best trial:

Value: 0.7979741659697054

Params:

```

iterations: 145
learning_rate: 0.06564339077069614
depth: 6
l2_leaf_reg: 7.546635702360232e-06
bootstrap_type: Bayesian
random_strength: 1.4799844388224288e-07
bagging_temperature: 0.19366957870297075
od_type: IncToDec
od_wait: 20

```

In [26]:

```

model = catboost.CatBoostClassifier(**trial.params,
                                     loss_function='Logloss',
                                     eval_metric='AUC',
                                     nan_mode='Min',
                                     thread_count=5,
                                     task_type='CPU',
                                     custom_metric=['Accuracy', 'F1', 'Precision', 'Recall'],
                                     #per_float_feature_quantization='2:border_count=1024',
                                     class_weights = [0.4, 0.6],
                                     verbose=True)
model.fit(Xdata_train, Ydata_train, eval_set=(Xdata_val, Ydata_val),
          cat_features=categorical_features_indices,
          verbose_eval=50,
          early_stopping_rounds=30,
          use_best_model=True,
          plot=True)
y_pred = model.predict(Xdata_val)

```

```

0:      test: 0.8157439 best: 0.8157439 (0)      total: 7.64ms   remaining: 1.1s
50:      test: 0.8945177 best: 0.8945345 (49)     total: 245ms   remaining: 451ms
100:     test: 0.9010570 best: 0.9010570 (100)    total: 477ms   remaining: 208ms
144:     test: 0.9053701 best: 0.9053701 (144)    total: 681ms   remaining: 0us

```

bestTest = 0.9053700842

bestIteration = 144

In [27]:

```
print('CatBoost model is fitted: ' + str(model.is_fitted()))
print('CatBoost model parameters:')
print(model.get_params())

print('Best Score:')
print(model.get_best_score())

print('Best Iteration:')
print(model.get_best_iteration())
```

CatBoost model is fitted: True

CatBoost model parameters:

```
{'iterations': 145, 'learning_rate': 0.06564339077069614, 'depth': 6, 'l2_leaf_reg': 7.546635702360232e-06, 'loss_function': 'Logloss', 'od_w
ait': 20, 'od_type': 'IncToDec', 'nan_mode': 'Min', 'thread_count': 5, 'verbose': True, 'class_weights': [0.4, 0.6], 'random_strength': 1.479
9844388224288e-07, 'custom_metric': ['Accuracy', 'F1', 'Precision', 'Recall'], 'eval_metric': 'AUC', 'bagging_temperature': 0.193669578702970
75, 'task_type': 'CPU', 'bootstrap_type': 'Bayesian'}
```

Best Score:

```
{'learn': {'Accuracy:use_weights=false': 0.7964312741466524, 'Recall:use_weights=false': 0.905449613021623, 'Precision:use_weights=false': 0.
7486932046642542, 'Accuracy:use_weights=true': 0.8154015342721116, 'Logloss': 0.3845140993159121, 'Precision:use_weights=true': 0.81714453267
51762, 'Recall:use_weights=true': 0.905449613021623, 'F1:use_weights=true': 0.8526617078015725, 'F1:use_weights=false': 0.813840830449827},
'validation': {'Accuracy:use_weights=false': 0.7980670941362327, 'Recall:use_weights=false': 0.9239029809294576, 'Precision:use_weights=fals
e': 0.7446163178647255, 'Accuracy:use_weights=true': 0.8205994887691196, 'AUC': 0.9053700841745724, 'Logloss': 0.3722219744080487, 'Precisio
n:use_weights=true': 0.8139020923902635, 'Recall:use_weights=true': 0.9239029809294576, 'F1:use_weights=true': 0.859314926951269, 'F1:use_w
eights=false': 0.8190222370284}}
```

Best Iteration:

144

In [28]:

```
print(classification_report(Ydata_val, y_pred))
```

	precision	recall	f1-score	support
0.0	0.88	0.69	0.77	5360
1.0	0.74	0.91	0.82	5401
accuracy			0.80	10761
macro avg	0.81	0.80	0.80	10761
weighted avg	0.81	0.80	0.80	10761

In [29]:

```
from sklearn.metrics import accuracy_score, confusion_matrix, f1_score, roc_auc_score
preds_class = model.predict(Xdata_val)
```

```
display(F'Learning Rate set to: {model.get_all_params()["learning_rate"]}')
display(F'Accuracy Score: {accuracy_score(Ydata_val, preds_class)}')
display(F"F1-Score: {f1_score(Ydata_val, preds_class,average='weighted')}")
```

```
matrix = confusion_matrix(Ydata_val, preds_class)
print('-'*50)
print('Confusion Matrix: ')
print(matrix)
print('-'*50)
```

```
'Learning Rate set to: 0.06564339250326157'
```

```
'Accuracy Score: 0.7979741659697054'
```

```
'F1-Score: 0.7953585164955788'
```

```
-----
Confusion Matrix:
```

```
[[3675 1685]
 [ 489 4912]]
-----
```

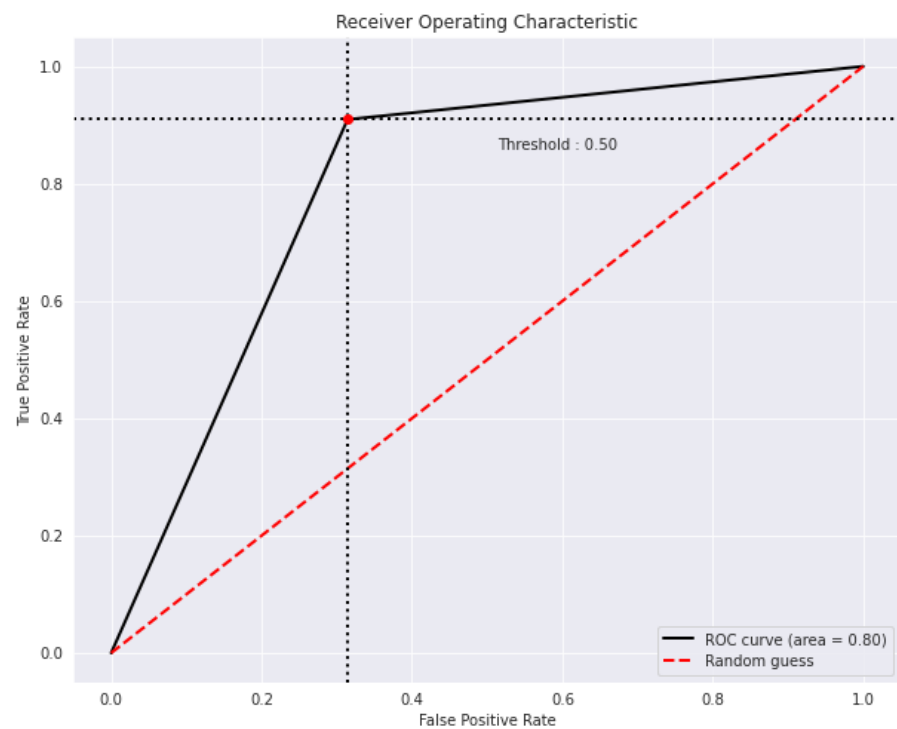
In [30]:

```
from sklearn.metrics import roc_curve, auc
from itertools import cycle
from sklearn.preprocessing import label_binarize
from sklearn import metrics
```


In [31]:

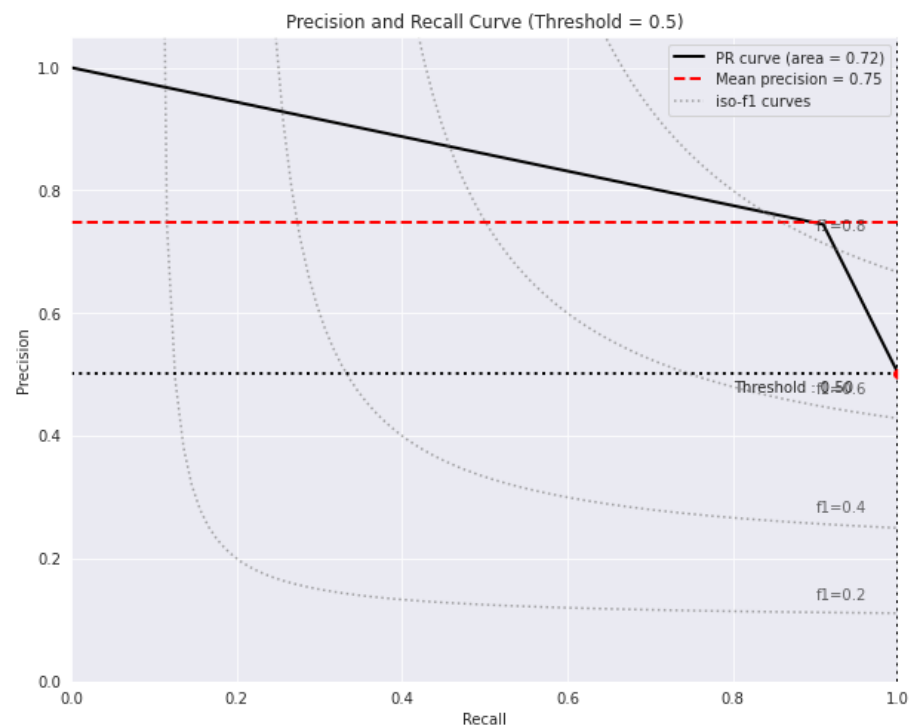
```
from plot_metric.functions import BinaryClassification
# Visualisation with plot_metric
bc = BinaryClassification(Ydata_val, preds_class, labels=["Class 0", "Class 1"], threshold=0.5)

# Figures
plt.figure(figsize=(10,8))
bc.plot_roc_curve()
plt.show()
```



In [32]:

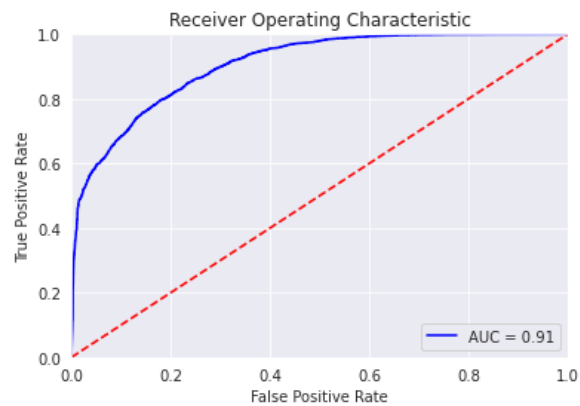
```
# Visualisation with plot_metric  
  
plt.figure(figsize=(10,8))  
  
bc.plot_precision_recall_curve()  
plt.show()
```



In [33]:

```
import sklearn.metrics as metrics
# calculate the fpr and tpr for all thresholds of the classification
probs = model.predict_proba(Xdata_val)
preds = probs[:,1]
fpr, tpr, threshold = metrics.roc_curve(Ydata_val, preds)
roc_auc = metrics.auc(fpr, tpr)

# method 1: plt
import matplotlib.pyplot as plt
plt.title('Receiver Operating Characteristic')
plt.plot(fpr, tpr, 'b', label = 'AUC = %0.2f' % roc_auc)
plt.legend(loc = 'lower right')
plt.plot([0, 1], [0, 1], 'r--')
plt.xlim([0, 1])
plt.ylim([0, 1])
plt.ylabel('True Positive Rate')
plt.xlabel('False Positive Rate')
plt.show()
```



In [34]:

```
import bctools as bc

conf_matrix, metrics_fixed_thresh_df = bc.utilities.get_confusion_matrix_and_metrics_df(
    true_y = Ydata_val,
    predicted_proba = preds_class,
    threshold = 0.5 # default = 0.5
)

display(conf_matrix, metrics_fixed_thresh_df)

array([[3675, 1685],
       [ 489, 4912]])
```

	threshold_dependent_metric	value
0	accuracy	0.7980
1	balanced_accuracy	0.7975
2	f1_score	0.8188
3	precision	0.7446
4	recall	0.9095
5	cohens_kappa	0.5956
6	matthews_corr_coef	0.6109

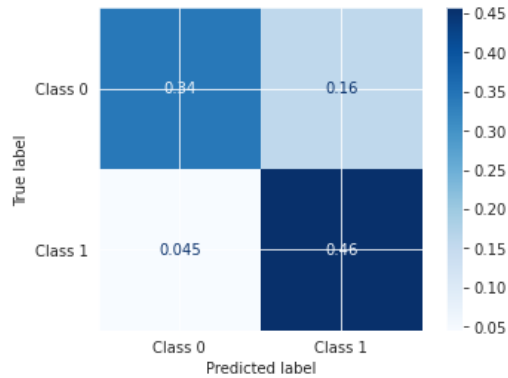
In [35]:

```
class_names = ['Class 0', 'Class 1']
```

In [36]:

```
disp = plot_confusion_matrix(model, Xdata_val, Ydata_val, display_labels=class_names, cmap=plt.cm.Blues, xticks_rotation='horizontal', normalize = 'all'
)
```

Function plot_confusion_matrix is deprecated; Function `plot_confusion_matrix` is deprecated in 1.0 and will be removed in 1.2. Use one of the class methods: ConfusionMatrixDisplay.from_predictions or ConfusionMatrixDisplay.from_estimator.



In [37]:

```
import shap
# shap_values = model.get_feature_importance(Pool(Xdata_val, label=Ydata_val, cat_features=categorical_features_indices),
#                                           type="ShapValues")
# expected_value = shap_values[0,-1]
# shap_values = shap_values[:, :-1]

# shap.initjs()
# shap.force_plot(expected_value, shap_values[3,:], Xdata_val.iloc[3,:])

# shap.summary_plot(shap_values, Xdata_val)
# shap.summary_plot(shap_values, Xdata_val)
```

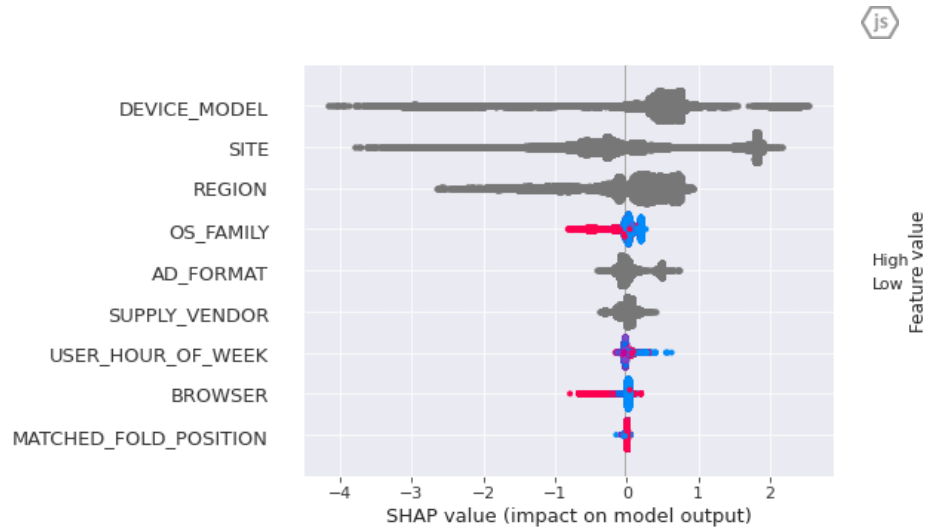
```
In [38]:  
explainer = shap.TreeExplainer(model)  
shap_values = explainer.shap_values(Xdata_train)
```

```
In [39]:  
model.get_feature_importance(prettified=True)
```

Out[39]:

	Feature Id	Importances
0	SITE	32.876189
1	DEVICE_MODEL	29.689352
2	REGION	18.364656
3	AD_FORMAT	5.644158
4	OS_FAMILY	4.731725
5	SUPPLY_VENDOR	4.368040
6	BROWSER	1.937947
7	USER_HOUR_OF_WEEK	1.838675
8	MATCHED_FOLD_POSITION	0.549259

```
In [40]:  
shap.initjs()  
##Class 1  
shap.summary_plot(shap_values, Xdata_train.values, feature_names = Xdata_train.columns)
```



In [41]:

```
##Training set
train = Xdata_train
train['id_target'] = Ydata_train

#Testing Set
mydf = Xdata_val
mydf['id_target'] = Ydata_val

#Specify the column names that you will use for the bidlist (These are just example names)
col_list = ['SITE', 'DEVICE_MODEL', 'REGION', 'AD_FORMAT', 'SUPPLY_VENDOR', 'USER_HOUR_OF_WEEK']
```

In [42]:

```

### Define Bidlist Functions

import pandas as pd
from ttd_api_authenticate import *
from TradedeskCredentials import *
from bid_lists_generator import *

def bid_col (mydf, col, target, num):
    df1 = mydf.loc[mydf[target] == 1, [col, target]]
    df2 = df1.groupby(col).agg(np.count_nonzero).sort_values(by = target, ascending=False)
    df3 = df2.iloc[0:num,]
    df3 = df3.reset_index()
    df3['key'] = 0
    return df3

def df_list_make(train, col_list, targeted_id):
    df_list = []
    for i in col_list:
        j = bid_col (train, col = i, target = targeted_id, num = 100000)
        df_list.append(j)
    return df_list
df_list = df_list_make(train, col_list, targeted_id = 'id_target')

def cross_join (df_list, col_list):
    lenth = len(df_list)-1
    i = 0
    mydf = df_list[i]
    while i < lenth:
        mydf = pd.merge(mydf, df_list[i+1], on= 'key')
        i = i+1
    mydf = mydf[col_list]
    return mydf

def apply_bid_lists_to_adgroups(token, bid_list_ids, adgroupids, enable_on_apply=True,
                                on_error='continue', verbose=False):
    #on_error in ['continue', 'raise']
    fails = []
    results = []
    for i, adgroup in enumerate(adgroupids):
        if verbose:
            printstr = f"Adding {len(bid_list_ids)} to adgroup {adgroup}"
            print(printstr, end='\r')
            print(' '*len(printstr), end='\r')
        try:
            current_state = generic_get(f'adgroup/{adgroup}', token=token).json()
        except Exception as e:
            if on_error == 'continue':
                fails.append(adgroup)
                continue
            elif on_error == 'raise':
                raise ValueError(f"Error getting adgroup id '{adgroup}' (index {i})")

        current_bid_lists = current_state['AssociatedBidLists']
        current_bid_lists.extend([{'BidListId':x,
                                   'IsEnabled':enable_on_apply} for x in bid_list_ids])

```



```

payload={'AdGroupId':adgroup, 'AssociatedBidLists':current_bid_lists}
try:
    result = generic_put('adgroup', token=token, payload=payload).json()
    results.append({adgroup:result['AssociatedBidLists']})
except Exception as e:
    if on_error == 'continue':
        fails.append(adgroup)
        continue
    elif on_error == 'raise':
        raise ValueError(f"Error updating adgroup id '{adgroup}' (index {i})")
print('Uploaded Successfully')

```

In []:

```

#Call the bidlist functions
df_list = df_list_make(train, col_list, targeted_id = 'id_target')
df_cross = cross_join (df_list, col_list)

categorical_features = col_list

categorical_features_indices = [df_cross.columns.get_loc(col) for col in categorical_features]

#Use your trained catboost model. If it isn't called clf, change the name here.
probs = clf.predict_proba(df_cross)
df_prob = pd.DataFrame({'Negative': probs[:, 0], 'Positive': probs[:, 1]})
bid_model = df_cross
bid_model['probs'] = df_prob['Positive']

#This outputs every permutation of the columns selected for the bidlist and assigned them a probability
bid_model

```

Passing 'suffixes' which cause duplicate columns {'id_target_x'} in the result is deprecated and will raise a MergeError in a future version.

In [1]:

```

#This is a sample scaling function. This is how we scale the probabilities to the correct price multipliers.
#Mark is running this campaign so we'll be consulting with him on this range

min = bid_model['probs'].min()
max = bid_model['probs'].max()

a, b = 0, 2
x, y = min, max
bid_model['scaled'] = (bid_model.probs - x) / (y - x) * (b - a) + a
bid_model['value'] = (bid_model.scaled + 0.5)
#Just cleaning the dataframe up to keep only the value
bid_model = bid_model.drop(['probs', 'scaled', 'Unnamed: 0'],1)

```

```

-----
NameError                                Traceback (most recent call last)
/tmp/ipykernel_28361/192485853.py in <cell line: 4>()
      2 #Mark is running this campaign so we'll be consulting with him on this range
      3
----> 4 min = bid_model['probs'].min()
      5 max = bid_model['probs'].max()
      6

NameError: name 'bid_model' is not defined

```

In []:

```

#Rename the columns to match trade desk specified names
#You may also need to change the type. The excel sheet specifies these
bid_model = bid_model.rename(columns={'AD_FORMAT': 'AdFormatId'})
bid_model.AdFormatId = bid_model.AdFormatId.astype('string')

```

In []:

```

#Api calls
#Some trade desk variables require an API call to get the mapping for each variable.
#This is our sandbox api so you can test out the functions

#These are all ids associated with the sandbox
#Campaign ID: kwh8cvf
#Ad Group ID: z6z3tsc
#Advertiser ID: ca5g5oz
#Partner ID: hregxvd

#Authentication api call
#Make sure the associated python scripts are located in your current directory
sandboxuser = "tttd_api_hregxvd@chalice.com"
token = authenticate(sandboxuser, verbose=True)

```

In []:

```

#If you need to get supply vendor ids from the api, this an example of the call you would make
payload={
    "AdvertiserId":"ca5g5oz",
    #Here, you would insert the name of the supply vendor and it will return the associated id.
    #Probably will have to loop through a list
    "SearchTerms":["adcolony"],
    "PageStartIndex":0,
    "PageSize":100
}
resjson = generic_post("supplyvendor/query/advertiser", token, payload=payload).json()

print(resjson)

```

```

#Here is the trade desk documentation for various api calls
#https://api.thetradedesk.com/v3/portal/api/doc/ApiOverview

```

In []:

```

#Once everything is in Trade Desk Format, this is how you push a bid list to trade desk

#First, save the bidlist as a csv split the bidlist into csv chunks
#We can only up load ~40k lines at a time
#Make sure you save without an index in the csv

rows = pd.read_csv("mytest.csv", chunksize=40000)
for i, chunk in enumerate(rows):
    chunk.to_csv('mytest_{}.csv'.format(i), index=False)

```

In []:

```

#Next, we get old bid list ids and remove them from the ad group we want to apply the new list to
#This is not necessary at the start of the campaign
adgroupids = ['z6z3tsc']

#If we remove anything, it goes into oldbidlist ids
oldbidlistids = []

for adgroup in lst:
    info = generic_get(f"/adgroup/{adgroup}", token).json()
    bidlists = info['AssociatedBidLists']
    bidlists = [x for x in bidlists if x['BidListId'] not in oldbidlistids]
    payload={'AdGroupId':adgroup, 'AssociatedBidLists':bidlists}
    result = generic_put('adgroup', token=token, payload=payload).json()

bidlists = resjson['AssociatedBidLists']
print(bidlists)

```

In []:

```
#Next we post the bidlist to the campaign level. It lives here and can get applied to any ad groups

#Name needs to be specified, as well as owner_id, which is the campaign id.
#Everything else stays

bidids = []

for i in range('Range of csv files'):
    bidlistlines = bid_list_json(bidlist_id=1,
                                name="Test_bids_date",
                                bidlist_source="User",
                                bidlist_adjustment_type="Optimized",
                                resolution_type="ApplyMultiplyAdjustment",
                                bidlines_csv=f"mytest_{i}.csv",
                                bidlist_owner="Campaign",
                                bidlist_owner_id="kwh8cvf",
                                is_available_for_library_use=True,
                                bidlist_type='BidAdjustment',
                                csv_value_col='value',
                                csv_sep=',',
                                show=False)
    bidpost = post_bid_list(token=token, payload=bidlistlines).json()['BidListId']
    bidids.append(bidpost)
print(bidids)
```

In []:

```
#Next we apply these bidids to the selected adgroup
newbidlines = bidids
apply_bid_lists_to_adgroups(token=token, bid_list_ids=newbidlines, adgroupids=adgroupids, enable_on_apply=True,
                             on_error='continue', verbose=True)
```