```
!pip uninstall -y numpy
!pip install numpy==1.23.5
#!pip install catboost
#!pip install optuna
```

```
Found existing installation: numpy 2.0.2
Uninstalling numpy-2.0.2:
  Successfully uninstalled numpy-2.0.2
Collecting numpy==1.23.5
  Downloading numpy-1.23.5-cp311-cp311-manylinux_2_17_x86_64.manylinux2014_x86_64.whl.metadata (2.3 kB)
Downloading numpy-1.23.5-cp311-cp311-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (17.1 MB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 17.1/17.1 MB 62.4 MB/s eta 0:00:00
Installing collected packages: numpy
ERROR: pip's dependency resolver does not currently take into account all the packages that are installed. This beha
chex 0.1.89 requires numpy>=1.24.1, but you have numpy 1.23.5 which is incompatible.
pymc 5.21.1 requires numpy>=1.25.0, but you have numpy 1.23.5 which is incompatible.
tensorflow 2.18.0 requires numpy<2.1.0,>=1.26.0, but you have numpy 1.23.5 which is incompatible.
xarray 2025.1.2 requires numpy>=1.24, but you have numpy 1.23.5 which is incompatible.
blosc2 3.2.0 requires numpy>=1.26, but you have numpy 1.23.5 which is incompatible.
albucore 0.0.23 requires numpy>=1.24.4, but you have numpy 1.23.5 which is incompatible.
jax 0.5.2 requires numpy>=1.25, but you have numpy 1.23.5 which is incompatible.
treescope 0.1.9 requires numpy>=1.25.2, but you have numpy 1.23.5 which is incompatible.
jaxlib 0.5.1 requires numpy>=1.25, but you have numpy 1.23.5 which is incompatible.
imbalanced-learn 0.13.0 requires numpy<3,>=1.24.3, but you have numpy 1.23.5 which is incompatible.
scikit-image 0.25.2 requires numpy>=1.24, but you have numpy 1.23.5 which is incompatible.
bigframes 1.41.0 requires numpy>=1.24.0, but you have numpy 1.23.5 which is incompatible.
albumentations 2.0.5 requires numpy>=1.24.4, but you have numpy 1.23.5 which is incompatible.
Successfully installed numpy-1.23.5
WARNING: The following packages were previously imported in this runtime:
  [numpy]
You must restart the runtime in order to use newly installed versions.
```

RESTART SESSION

```
!pip install catboost
!pip install optuna
```

```
!pip install optuna
```

```
Collecting catboost
  Downloading catboost-1.2.7-cp311-cp311-manylinux2014_x86_64.whl.metadata (1.2 kB)
Requirement already satisfied: graphviz in /usr/local/lib/python3.11/dist-packages (from catboost) (0.20.3)
Requirement already satisfied: matplotlib in /usr/local/lib/python3.11/dist-packages (from catboost) (3.10.0)
Requirement already satisfied: numpy<2.0,>=1.16.0 in /usr/local/lib/python3.11/dist-packages (from catboost) (1.23.5
Requirement already satisfied: pandas>=0.24 in /usr/local/lib/python3.11/dist-packages (from catboost) (2.2.2)
Requirement already satisfied: scipy in /usr/local/lib/python3.11/dist-packages (from catboost) (1.14.1)
Requirement already satisfied: plotly in /usr/local/lib/python3.11/dist-packages (from catboost) (5.24.1)
Requirement already satisfied: six in /usr/local/lib/python3.11/dist-packages (from catboost) (1.17.0)
Requirement already satisfied: python-dateutil>=2.8.2 in /usr/local/lib/python3.11/dist-packages (from pandas>=0.24-
Requirement already satisfied: pytz>=2020.1 in /usr/local/lib/python3.11/dist-packages (from pandas>=0.24->catboost)
Requirement already satisfied: tzdata>=2022.7 in /usr/local/lib/python3.11/dist-packages (from pandas>=0.24->catboos
Requirement already satisfied: contourpy>=1.0.1 in /usr/local/lib/python3.11/dist-packages (from matplotlib->catboos
Requirement already satisfied: cycler>=0.10 in /usr/local/lib/python3.11/dist-packages (from matplotlib->catboost) (
Requirement already satisfied: fonttools>=4.22.0 in /usr/local/lib/python3.11/dist-packages (from matplotlib->catboo
Requirement already satisfied: kiwisolver>=1.3.1 in /usr/local/lib/python3.11/dist-packages (from matplotlib->catboo
Requirement already satisfied: packaging>=20.0 in /usr/local/lib/python3.11/dist-packages (from matplotlib->catboost
Requirement already satisfied: pillow>=8 in /usr/local/lib/python3.11/dist-packages (from matplotlib->catboost) (11.
Requirement already satisfied: pyparsing>=2.3.1 in /usr/local/lib/python3.11/dist-packages (from matplotlib->catboos
Requirement already satisfied: tenacity>=6.2.0 in /usr/local/lib/python3.11/dist-packages (from plotly->catboost) (9
Downloading catboost-1.2.7-cp311-cp311-manylinux2014_x86_64.whl (98.7 MB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 98.7/98.7 MB 9.8 MB/s eta 0:00:00
Installing collected packages: catboost
Successfully installed catboost-1.2.7
Collecting optuna
  Downloading optuna-4.2.1-py3-none-any.whl.metadata (17 kB)
Collecting alembic>=1.5.0 (from optuna)
  Downloading alembic-1.15.1-py3-none-any.whl.metadata (7.2 kB)
Collecting colorlog (from optuna)
  Downloading colorlog-6.9.0-py3-none-any.whl.metadata (10 kB)
Requirement already satisfied: numpy in /usr/local/lib/python3.11/dist-packages (from optuna) (1.23.5)
Requirement already satisfied: packaging>=20.0 in /usr/local/lib/python3.11/dist-packages (from optuna) (24.2)
Requirement already satisfied: sqlalchemy>=1.4.2 in /usr/local/lib/python3.11/dist-packages (from optuna) (2.0.39)
Requirement already satisfied: tqdm in /usr/local/lib/python3.11/dist-packages (from optuna) (4.67.1)
Requirement already satisfied: PyYAML in /usr/local/lib/python3.11/dist-packages (from optuna) (6.0.2)
Requirement already satisfied: Mako in /usr/lib/python3/dist-packages (from alembic>=1.5.0->optuna) (1.1.3)
Requirement already satisfied: typing-extensions>=4.12 in /usr/local/lib/python3.11/dist-packages (from alembic>=1.5
Requirement already satisfied: greenlet!=0.4.17 in /usr/local/lib/python3.11/dist-packages (from sqlalchemy>=1.4.2->
Downloading optuna-4.2.1-py3-none-any.whl (383 kB)
```

```
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 383.6/383.6 kB 30.0 MB/s eta 0:00:00
Downloading alembic-1.15.1-py3-none-any.whl (231 kB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 231.8/231.8 kB 22.8 MB/s eta 0:00:00
Downloading colorlog-6.9.0-py3-none-any.whl (11 kB)
Installing collected packages: colorlog, alembic, optuna
Successfully installed alembic-1.15.1 colorlog-6.9.0 optuna-4.2.1
```

```
import pandas as pd
from catboost import CatBoostClassifier, Pool
from sklearn.model_selection import train_test_split
from sklearn.metrics import accuracy_score
import optuna

# Step 1: Load CSV file
file_path = 'global_fund_transactions_dataset.csv'
data = pd.read_csv(file_path)

# Step 2: Define target and features
target_column = 'company_name'  # Update this if needed
X = data.drop(columns=[target_column])
y = data[target_column]

# Step 3: Identify categorical features
cat_features = X.select_dtypes(include=['object', 'category']).columns.tolist()

# Step 4: Train/test split
X_train, X_valid, y_train, y_valid = train_test_split(X, y, test_size=0.2, random_state=42)

# Step 5: Define Optuna objective function
def objective(trial):
    params = {
        'iterations': trial.suggest_int('iterations', 10, 50),
        'depth': trial.suggest_int('depth', 4, 6),
        'learning_rate': trial.suggest_float('learning_rate', 0.01, 0.3),
        'l2_leaf_reg': trial.suggest_float('l2_leaf_reg', 1.0, 10.0),
        'bagging_temperature': trial.suggest_float('bagging_temperature', 0.0, 1.0),
        'random_strength': trial.suggest_float('random_strength', 1e-9, 10.0),
        'border_count': trial.suggest_int('border_count', 32, 255),
```

```
        'task_type': 'GPU',
        'devices': '0',
        'verbose': 0
    }

    model = CatBoostClassifier(**params, cat_features=cat_features)
    model.fit(X_train, y_train, eval_set=(X_valid, y_valid), early_stopping_rounds=30, verbose=0)
    preds = model.predict(X_valid)
    return accuracy_score(y_valid, preds)

# Step 6: Run Optuna study
study = optuna.create_study(direction='maximize')
study.optimize(objective, n_trials=10)

# Step 7: Train final model using best parameters
best_params = study.best_params
best_params['task_type'] = 'GPU'
best_params['devices'] = '0'
print("\n✅ Best Hyperparameters Found by Optuna:")
print(best_params)

final_model = CatBoostClassifier(**best_params, cat_features=cat_features, verbose=100)
train_pool = Pool(X_train, y_train, cat_features=cat_features)
test_pool = Pool(X_valid, y_valid, cat_features=cat_features)

final_model.fit(train_pool, eval_set=test_pool)

# Step 8: Feature importances
feature_importances = final_model.get_feature_importance(prettified=True)
print("\n🔍 Top Features Contributing to the Prediction:\n")
print(feature_importances.sort_values(by='Importances', ascending=False).head(15))
```

```
[I 2025-03-25 17:24:57,728] A new study created in memory with name: no-name-e974e1ef-5ef6-4745-ab8b-3bc3470ae594
[I 2025-03-25 17:28:57,685] Trial 0 finished with value: 0.9738 and parameters: {'iterations': 27, 'depth': 5, 'lear
[I 2025-03-25 17:33:12,768] Trial 1 finished with value: 0.98605 and parameters: {'iterations': 38, 'depth': 4, 'lea
[I 2025-03-25 17:38:00,333] Trial 2 finished with value: 0.996 and parameters: {'iterations': 43, 'depth': 4, 'learn
[I 2025-03-25 17:44:52,529] Trial 3 finished with value: 1.0 and parameters: {'iterations': 50, 'depth': 5, 'learnin
[I 2025-03-25 17:48:40,883] Trial 4 finished with value: 1.0 and parameters: {'iterations': 35, 'depth': 4, 'learnin
```

```
[I 2025-03-25 17:50:50,356] Trial 5 finished with value: 0.975 and parameters: {'iterations': 16, 'depth': 5, 'learn
[I 2025-03-25 17:52:19,836] Trial 6 finished with value: 0.9525 and parameters: {'iterations': 13, 'depth': 4, 'lear
[I 2025-03-25 17:53:57,412] Trial 7 finished with value: 0.2962 and parameters: {'iterations': 11, 'depth': 5, 'lear
[I 2025-03-25 17:56:40,827] Trial 8 finished with value: 0.98985 and parameters: {'iterations': 14, 'depth': 6, 'lea
[I 2025-03-25 18:01:56,053] Trial 9 finished with value: 0.99425 and parameters: {'iterations': 47, 'depth': 4, 'lea

✅ Best Hyperparameters Found by Optuna:
{'iterations': 50, 'depth': 5, 'learning_rate': 0.23065733902990812, 'l2_leaf_reg': 7.044413140737836, 'bagging_temp
0:      learn: 4.6265492        test: 4.6179508 best: 4.6179508 (0)     total: 8.37s    remaining: 6m 50s
49:     learn: 0.1599087        test: 0.1399547 best: 0.1399547 (49)    total: 6m 48s   remaining: 0us
bestTest = 0.1399547241
bestIteration = 49

🔍 Top Features Contributing to the Prediction:

             Feature Id   Importances
0               address  4.833297e+01
1          jurisdiction  2.651347e+01
2        company_number  1.834897e+01
3         is_sanctioned  6.786221e+00
4      transaction_date  1.044872e-02
5           amount_usd  4.007281e-03
6      transaction_type  3.916008e-03
7   transaction_comment  2.920572e-07
8        transaction_id  0.000000e+00
9     jurisdiction_code  0.000000e+00
10             currency  0.000000e+00
```