

```
[1]: # pip install xgboost
```

```
[3]: # Lift Booster Notebook: Improving Rare Event Detection
```

```
import os
import glob
import joblib
import pandas as pd
import numpy as np
from sklearn.metrics import average_precision_score
from xgboost import XGBClassifier
from sklearn.model_selection import train_test_split

# Paths
champion_stack_folder = "champion_stacks"
data_folder = "market_shock_synthetic_datasets"
base_model_folder = "champion_packages"
xgb_output_folder = "xgboost_stacks"

# Create output folder for xgboost models
os.makedirs(xgb_output_folder, exist_ok=True)

# Load Champion Logistic Stack Models
model_paths = sorted(glob.glob(os.path.join(champion_stack_folder, "*.pkl")))
models = {os.path.basename(p).replace("_stacked.pkl", ""): joblib.load(p) for p in model_paths}

# Load All Datasets
csv_paths = sorted(glob.glob(os.path.join(data_folder, "*.csv")))
datasets = {os.path.basename(p).replace(".csv", ""): pd.read_csv(p) for p in csv_paths}

# Retrain with XGBoost Stacker
results = []
for name in models:
    if name in datasets:
        df = datasets[name]
        X = df.drop("rare_event", axis=1)
        y = df["rare_event"]

        # Rebuild meta features from base champions
        meta_features = []
        for base_path in glob.glob(os.path.join(base_model_folder, "*.pkl")):
            if "_meta" in base_path: continue
            model = joblib.load(base_path)
            try:
                meta_features.append(model.predict_proba(X)[:, 1])
            except AttributeError:
                meta_features.append(model.decision_function(X))

        meta_X = np.vstack(meta_features).T
        X_train, X_test, y_train, y_test = train_test_split(meta_X, y, stratify=y, test_size=0.3, random_state=42)

        xgb = XGBClassifier(use_label_encoder=False, eval_metric='logloss')
        xgb.fit(X_train, y_train)

        # Save XGBoost stacker
        joblib.dump(xgb, os.path.join(xgb_output_folder, f"{name}_xgb_stacked.pkl"))

        preds = xgb.predict_proba(X_test)[:, 1]
        pr_auc = average_precision_score(y_test, preds)
        baseline = y.mean()

        results.append({
            "Dataset": name,
            "Model": "XGBoost Stacker",
            "PR AUC": round(pr_auc, 3),
            "Baseline": round(baseline, 3),
            "Lift": round(pr_auc - baseline, 3),
            "Liftoff": "🚀" if pr_auc >= 0.6 else "🐢"
        })

# Save Results
results_df = pd.DataFrame(results)
results_df.to_csv("xgboost_lift_results.csv", index=False)
print(results_df)
```

```
/Users/sophia-boettcher/anaconda3/lib/python3.11/site-packages/xgboost/training.py:183: UserWarning: [01:18:17] WARNING: /Users/runner/work/xgboost/xgboost/src/learner.cc:738:
Parameters: { "use_label_encoder" } are not used.

bst.update(dtrain, iteration=1, fobj=obj)
/Users/sophia-boettcher/anaconda3/lib/python3.11/site-packages/xgboost/training.py:183: UserWarning: [01:18:17] WARNING: /Users/runner/work/xgboost/xgboost/src/learner.cc:738:
Parameters: { "use_label_encoder" } are not used.

bst.update(dtrain, iteration=1, fobj=obj)
/Users/sophia-boettcher/anaconda3/lib/python3.11/site-packages/xgboost/training.py:183: UserWarning: [01:18:18] WARNING: /Users/runner/work/xgboost/xgboost/src/learner.cc:738:
Parameters: { "use_label_encoder" } are not used.

bst.update(dtrain, iteration=1, fobj=obj)
/Users/sophia-boettcher/anaconda3/lib/python3.11/site-packages/xgboost/training.py:183: UserWarning: [01:18:18] WARNING: /Users/runner/work/xgboost/xgboost/src/learner.cc:738:
Parameters: { "use_label_encoder" } are not used.

bst.update(dtrain, iteration=1, fobj=obj)
/Users/sophia-boettcher/anaconda3/lib/python3.11/site-packages/xgboost/training.py:183: UserWarning: [01:18:19] WARNING: /Users/runner/work/xgboost/xgboost/src/learner.cc:738:
Parameters: { "use_label_encoder" } are not used.

bst.update(dtrain, iteration=1, fobj=obj)
/Users/sophia-boettcher/anaconda3/lib/python3.11/site-packages/xgboost/training.py:183: UserWarning: [01:18:19] WARNING: /Users/runner/work/xgboost/xgboost/src/learner.cc:738:
Parameters: { "use_label_encoder" } are not used.

bst.update(dtrain, iteration=1, fobj=obj)
/Users/sophia-boettcher/anaconda3/lib/python3.11/site-packages/xgboost/training.py:183: UserWarning: [01:18:19] WARNING: /Users/runner/work/xgboost/xgboost/src/learner.cc:738:
Parameters: { "use_label_encoder" } are not used.

bst.update(dtrain, iteration=1, fobj=obj)
/Users/sophia-boettcher/anaconda3/lib/python3.11/site-packages/xgboost/training.py:183: UserWarning: [01:18:20] WARNING: /Users/runner/work/xgboost/xgboost/src/learner.cc:738:
Parameters: { "use_label_encoder" } are not used.

bst.update(dtrain, iteration=1, fobj=obj)
/Users/sophia-boettcher/anaconda3/lib/python3.11/site-packages/xgboost/training.py:183: UserWarning: [01:18:20] WARNING: /Users/runner/work/xgboost/xgboost/src/learner.cc:738:
Parameters: { "use_label_encoder" } are not used.

bst.update(dtrain, iteration=1, fobj=obj)
/Users/sophia-boettcher/anaconda3/lib/python3.11/site-packages/xgboost/training.py:183: UserWarning: [01:18:20] WARNING: /Users/runner/work/xgboost/xgboost/src/learner.cc:738:
Parameters: { "use_label_encoder" } are not used.

bst.update(dtrain, iteration=1, fobj=obj)
/Users/sophia-boettcher/anaconda3/lib/python3.11/site-packages/xgboost/training.py:183: UserWarning: [01:18:21] WARNING: /Users/runner/work/xgboost/xgboost/src/learner.cc:738:
Parameters: { "use_label_encoder" } are not used.

bst.update(dtrain, iteration=1, fobj=obj)
/Users/sophia-boettcher/anaconda3/lib/python3.11/site-packages/xgboost/training.py:183: UserWarning: [01:18:21] WARNING: /Users/runner/work/xgboost/xgboost/src/learner.cc:738:
Parameters: { "use_label_encoder" } are not used.

bst.update(dtrain, iteration=1, fobj=obj)
/Users/sophia-boettcher/anaconda3/lib/python3.11/site-packages/xgboost/training.py:183: UserWarning: [01:18:21] WARNING: /Users/runner/work/xgboost/xgboost/src/learner.cc:738:
Parameters: { "use_label_encoder" } are not used.

bst.update(dtrain, iteration=1, fobj=obj)
/Users/sophia-boettcher/anaconda3/lib/python3.11/site-packages/xgboost/training.py:183: UserWarning: [01:18:22] WARNING: /Users/runner/work/xgboost/xgboost/src/learner.cc:738:
Parameters: { "use_label_encoder" } are not used.
```

bst.update(dtrain, iteration=i, fobj=obj)						
	Dataset	Model	PR AUC	Baseline	Lift	Liftoff
0	market_shock_sim	XGBoost Stacker	0.181	0.165	0.016	🔴
1	marketshock_easy_s1	XGBoost Stacker	0.581	0.141	0.440	🔴
2	marketshock_easy_s2	XGBoost Stacker	0.659	0.136	0.524	🔴
3	marketshock_easy_s3	XGBoost Stacker	0.695	0.133	0.562	🔴
4	marketshock_extreme_s1	XGBoost Stacker	0.258	0.261	-0.003	🔴
5	marketshock_extreme_s2	XGBoost Stacker	0.268	0.255	0.013	🔴
6	marketshock_extreme_s3	XGBoost Stacker	0.281	0.264	0.018	🔴
7	marketshock_hard_s1	XGBoost Stacker	0.185	0.177	0.008	🔴
8	marketshock_hard_s2	XGBoost Stacker	0.156	0.164	-0.007	🔴
9	marketshock_hard_s3	XGBoost Stacker	0.144	0.155	-0.011	🔴
10	marketshock_medium_s1	XGBoost Stacker	0.277	0.141	0.136	🔴
11	marketshock_medium_s2	XGBoost Stacker	0.256	0.144	0.112	🔴
12	marketshock_medium_s3	XGBoost Stacker	0.210	0.140	0.070	🔴
13	stock_prediction_clean	XGBoost Stacker	0.921	0.502	0.418	🔴

[]: