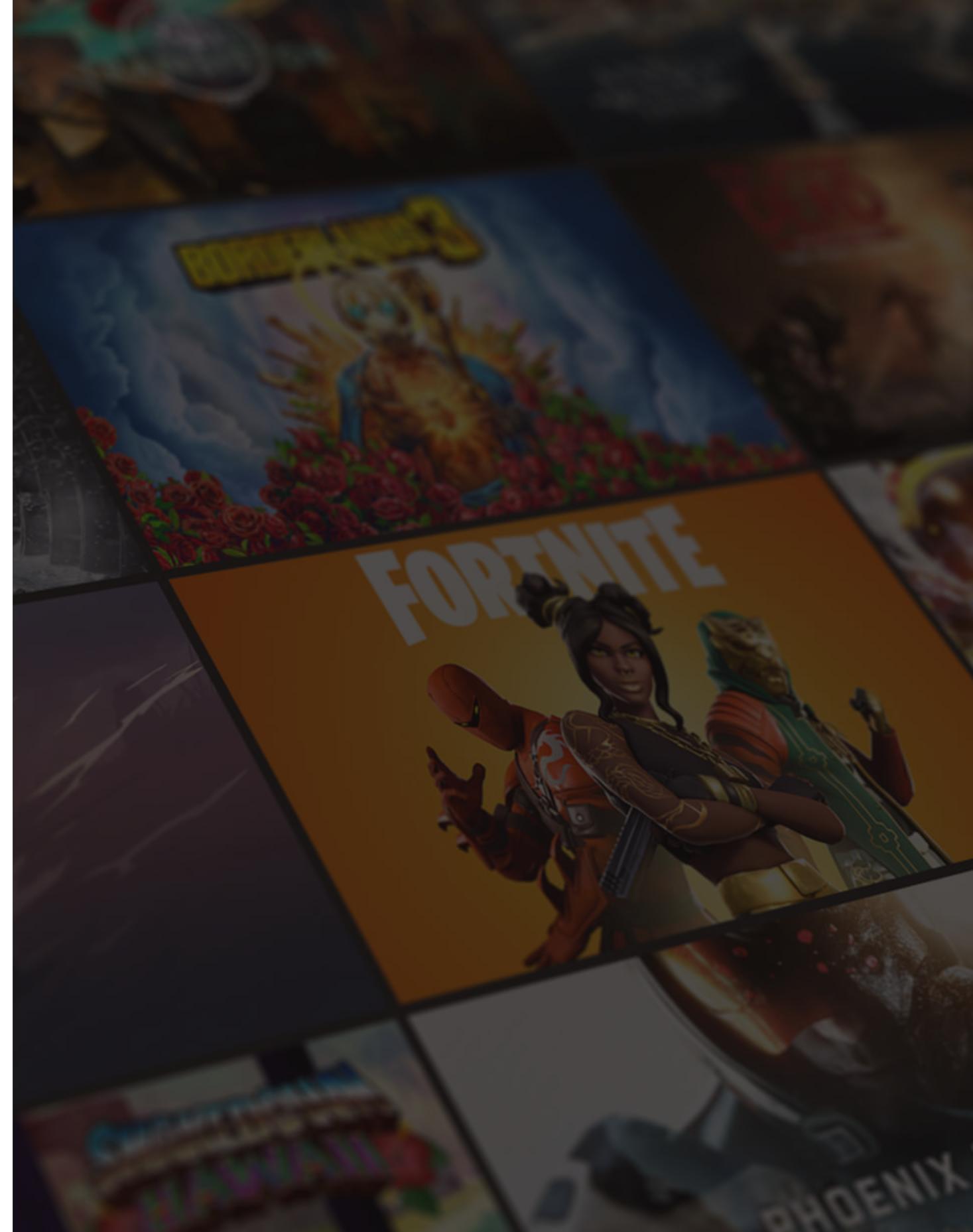


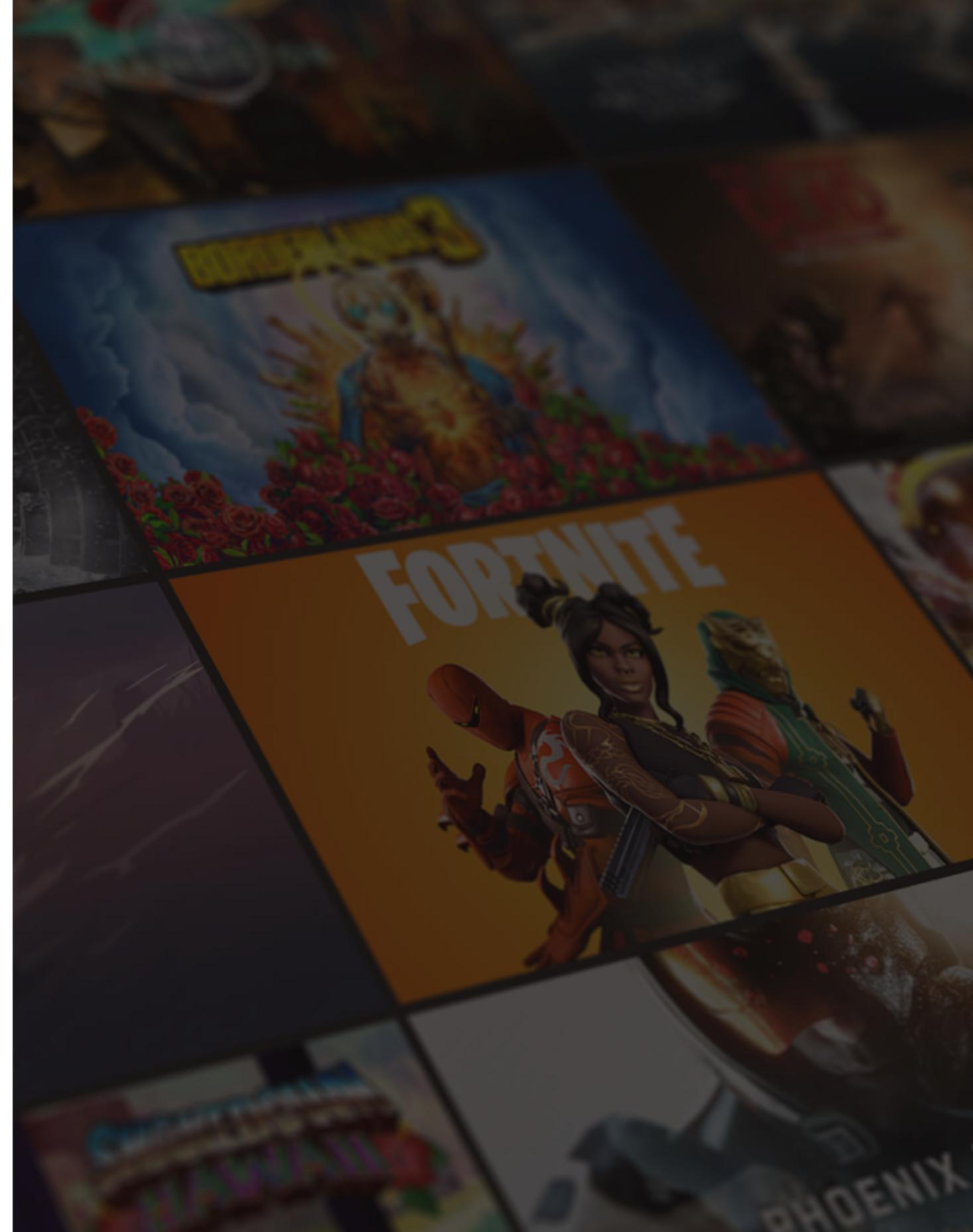
Epic Games Store

# Exclusive Games Selection Model



Epic Games Store

# Exclusive Games Selection Model





Motivation

**Aggressive Pursuit is  
Unsustainable**

# Methodology

## Data Ingestion

- Steam Store API
- Steam Spy API

# Data

## Target Variable:

Games played for  $\geq 1$  hour a day per player, for everyday the game has been available.

~1000 positive, 40000 negative

## 29 Feature Variables

# Methodology

## Data Ingestion

- Steam Store API
- Steam Spy API

## Data Cleaning

- EDA
- Remove Anomalies

# Methodology

## Data Ingestion

- Steam Store API
- Steam Spy API

## Data Cleaning

- EDA
- Remove Anomalies

## Baselining

- Metric Selection
- Random Forest

# Baseline Model

**Classifier:**

Random Forest (200 estimators)

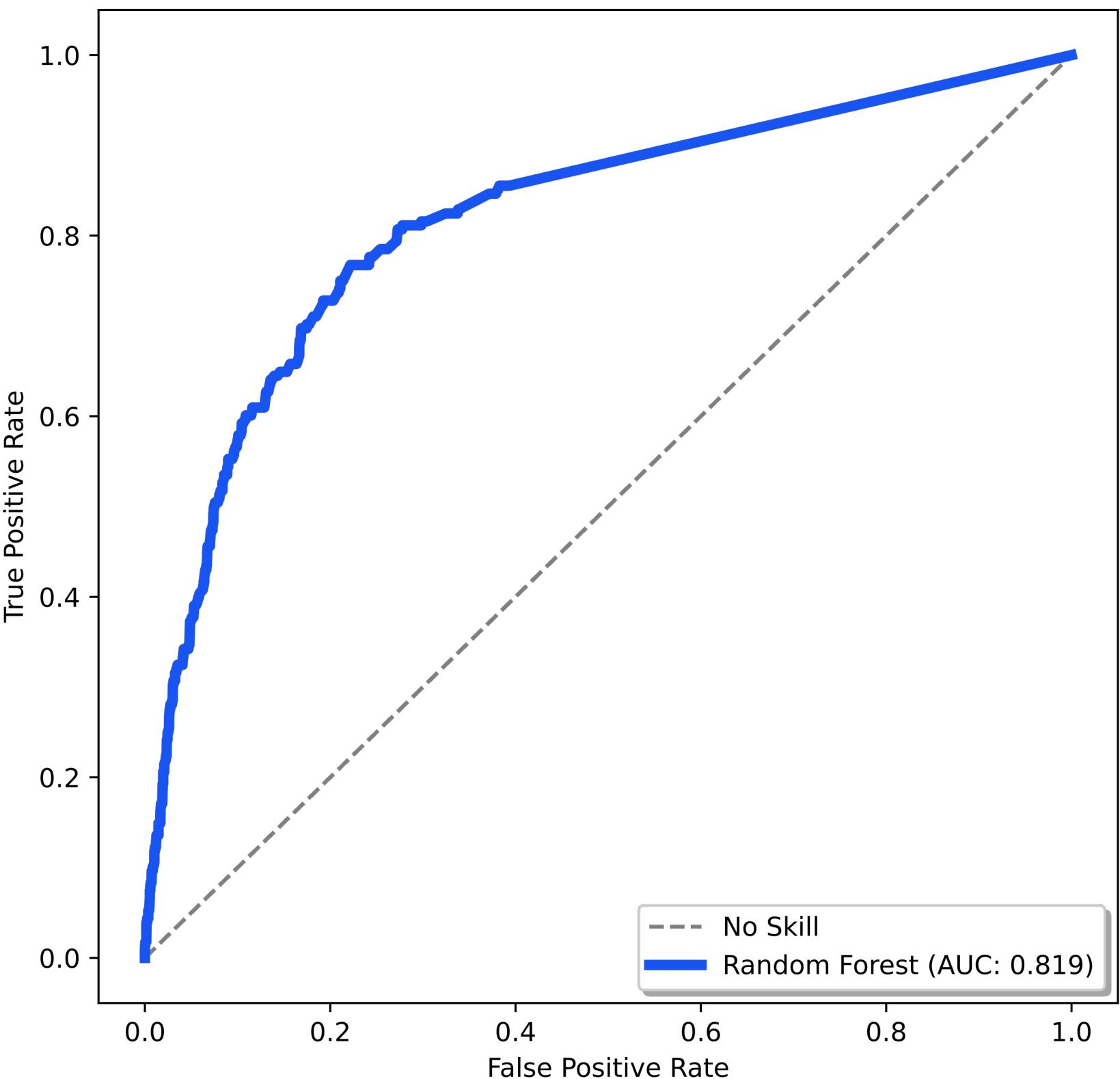
**Features:**

Genres, Categories, Platforms

**Precision:** 0.219

**Recall:** 0.135

**F(0.6):** 0.177



# Baseline Model

**Classifier:**

Random Forest (200 estimators)

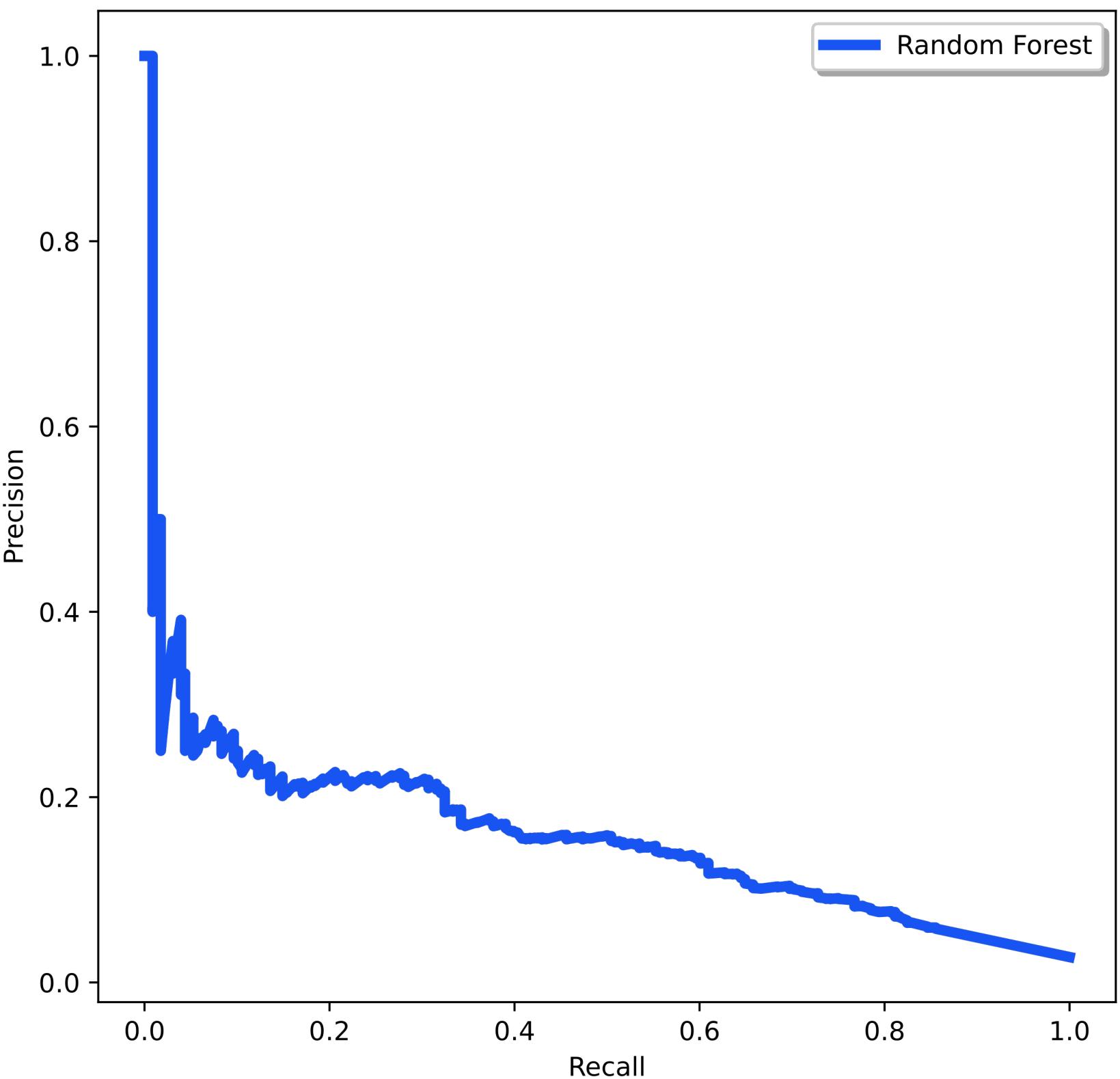
**Features:**

Genres, Categories, Platforms

**Precision:** 0.219

**Recall:** 0.135

**F(0.6):** 0.177



# Modelling Pipeline

**Train-Test Split**  
StratifiedSplit

**Cross Validator**  
StratifiedKFold  
GridSearchCV

**Pre-processor**  
ColumnTransformer  
SimpleImputer  
MultiHotEncoder

**Classifier**  
RandomForest  
XGBoost

# Modelling Pipeline

- Explore other features
  - Tags

**Train-Test Split**  
StratifiedSplit

**Cross Validator**  
StratifiedKFold  
GridSearchCV

**Pre-processor**  
ColumnTransformer  
SimpleImputer  
MultiHotEncoder

**Classifier**  
RandomForest  
XGBoost

# Modelling Pipeline

- Explore other features
  - Tags
- Feature engineering
  - Historical Records Only
  - Producer Owners, Hours...

**Train-Test Split**  
StratifiedSplit

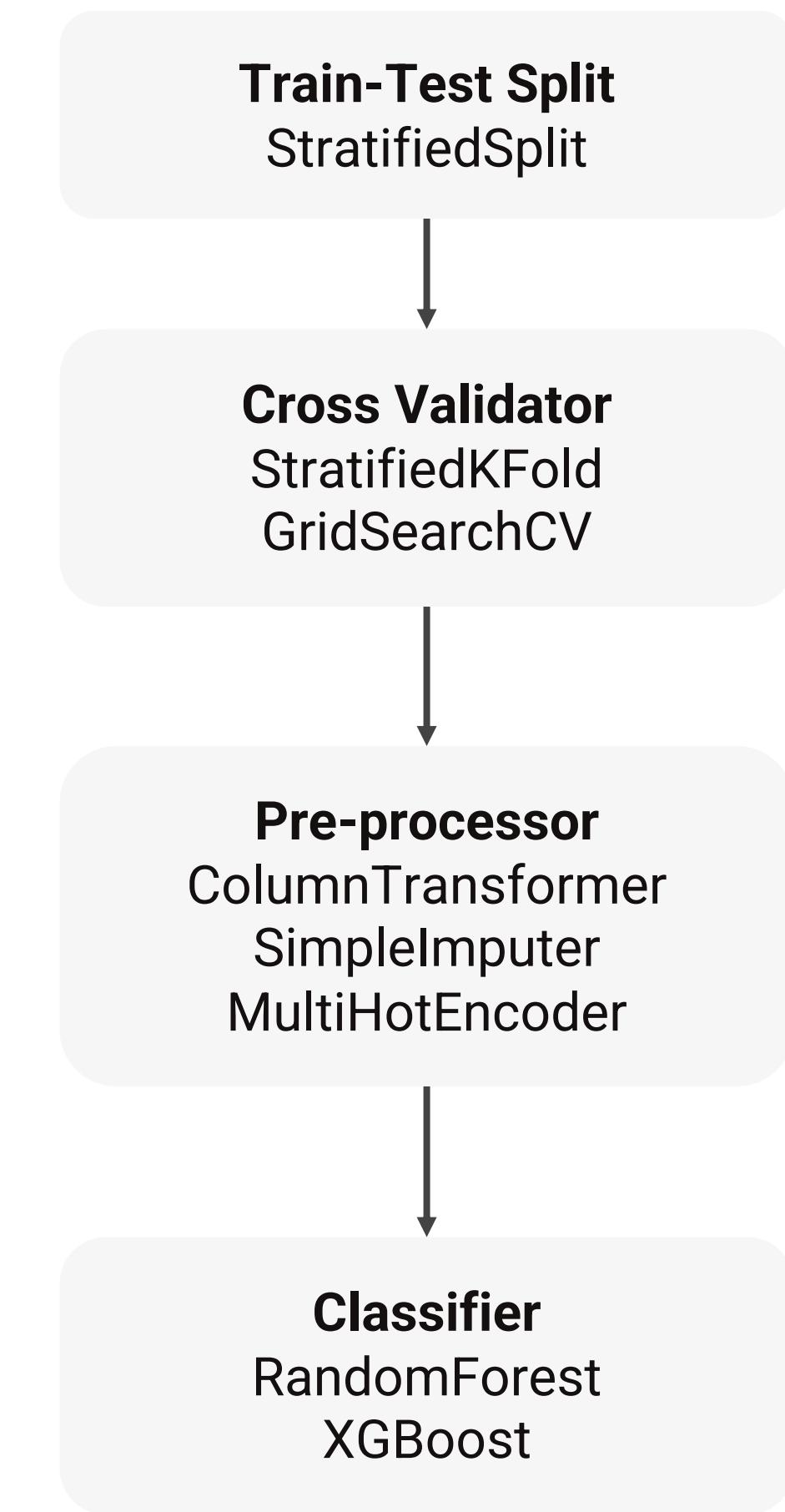
**Cross Validator**  
StratifiedKFold  
GridSearchCV

**Pre-processor**  
ColumnTransformer  
SimpleImputer  
MultiHotEncoder

**Classifier**  
RandomForest  
XGBoost

# Modelling Pipeline

- Explore other features
  - Tags
- Feature engineering
  - Historical Records Only
  - Producer Owners, Hours...
- Explore other models
  - XGBoost
- Parameter tuning
  - Class weights
  - GridSearchCV



# Final Model

## Classifier:

XGBoost

500 estimators, class weight=10,  
evaluation metric='auc'

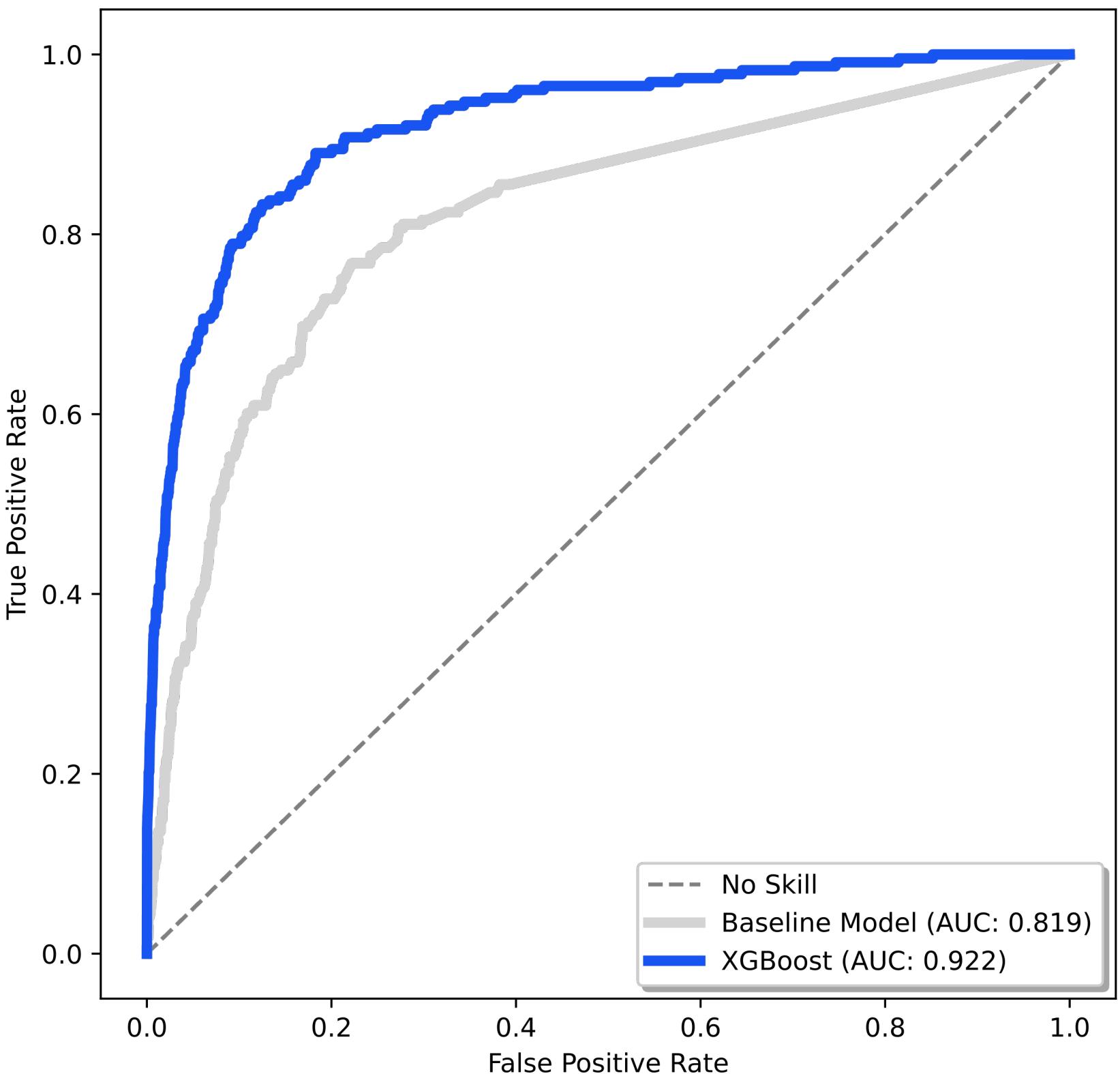
## Features:

Tags, Genres, Categories, Platforms,  
Historic Publisher Owners, Historic  
Publisher Hours, Supported Languages ...

Precision: 0.400

Recall: 0.478

F(0.6): 0.436



# Final Model

**Classifier:**

XGBoost

500 estimators, class weight=10,  
evaluation metric='auc'

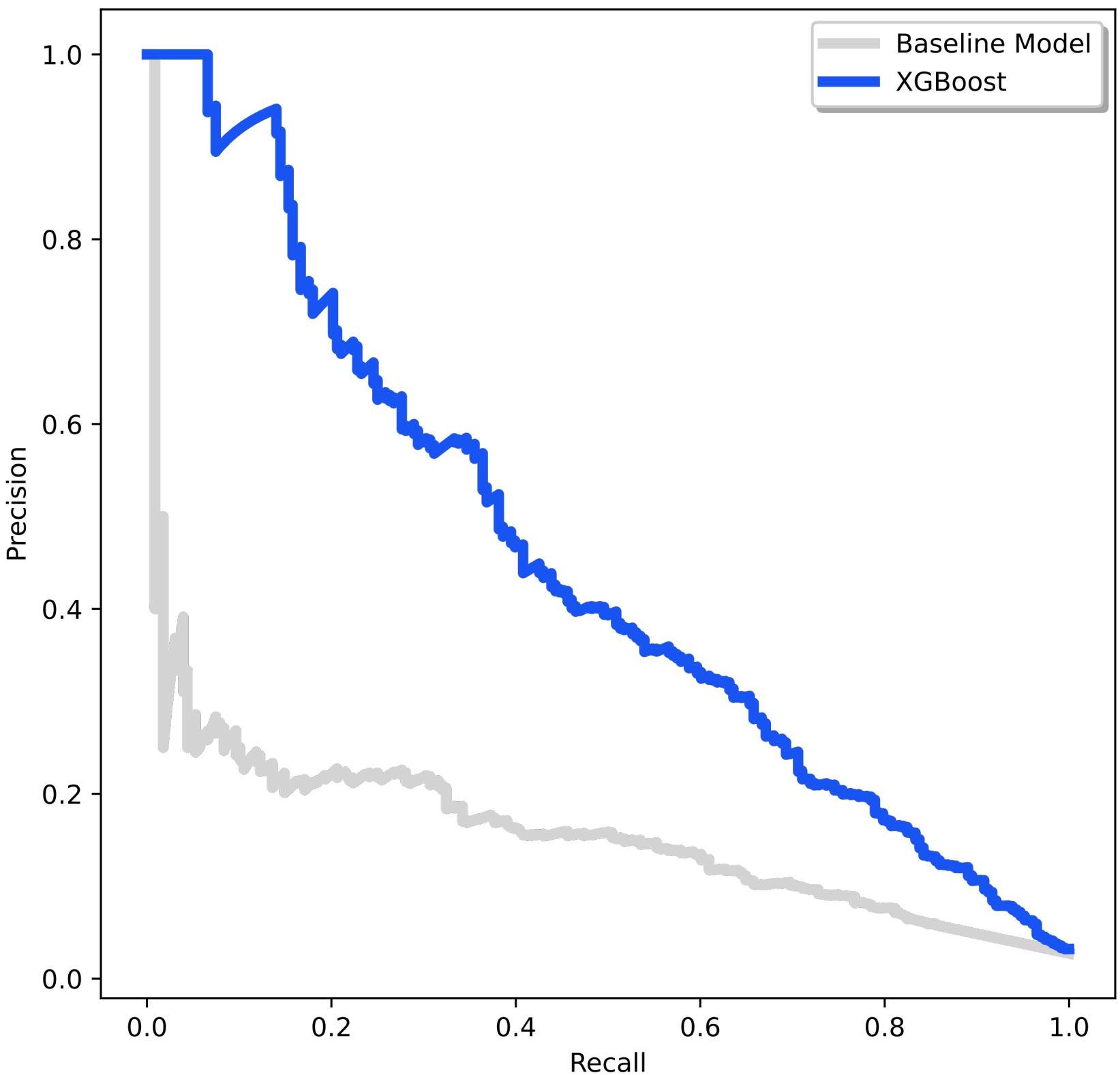
**Features:**

Tags, Genres, Categories, Platforms,  
Historic Publisher Owners, Historic  
Publisher Hours, Supported Languages ...

**Precision:** 0.400

**Recall:** 0.478

**F(0.6):** 0.436



# Conclusions and Future Work

- There is potential for a classification model to select games
- The limiting factor in this exploration was the data available
- Collect time-series data
- Explore other datasets/ data sources

A blue-tinted landscape from the video game Fortnite. In the foreground, three characters are standing in a grassy field. From left to right: a female character with short dark hair holding a pink shotgun; a male character with short brown hair holding a black rifle; and another male character with short brown hair holding a black shotgun. Behind them is a small campsite with a blue tent and some equipment. To the left, there's a wooden bridge over a stream. The background consists of rolling green hills covered in trees under a cloudy sky.

Thank you!