BOARD GAME TIME!
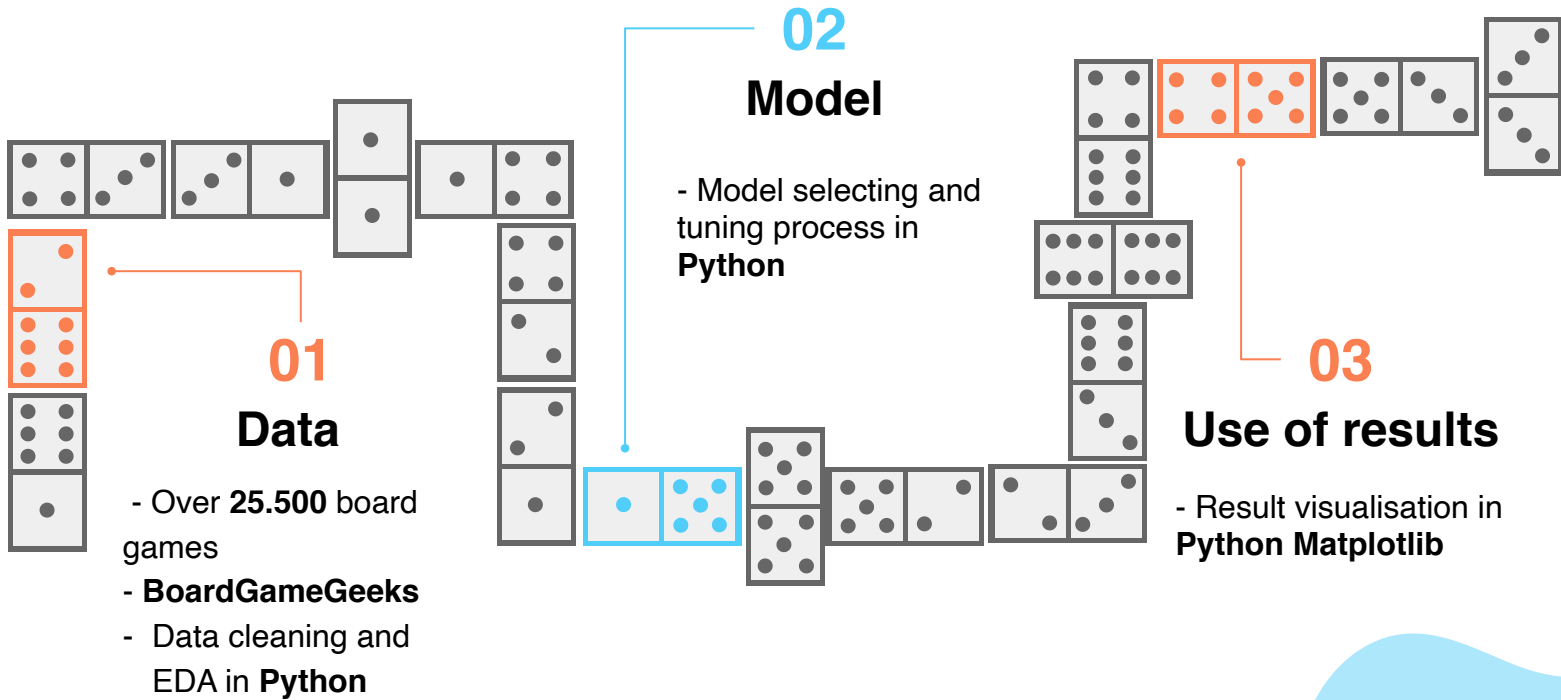
Classification Project

METIS - Data Science Bootcamp 2021
Gabriella Nemeth

- Games and Puzzle Market
  - $10.4 bn 2019
  - $11.3 bn 2020
  - **- $12.1 bn 2021 (predicted)**

- Lockdown free time

- **Crowdfunding** - Kickstarter
  - game projects $236.6 m in 2020
  - **33% increase to 2019**

- **BoardGameGeeks**

02

Model

- Model selecting and tuning process in **Python**

01

Data

 - Over **25.500** board games
- **BoardGameGeeks**
- Data cleaning and EDA in **Python**

03

Use of results

- Result visualisation in **Python Matplotlib**

Features

- players (min, max),
- game time (avg, min, max),
- owned, num of votes on the page,
- year of release,
- age limit,
- weight (complexity),
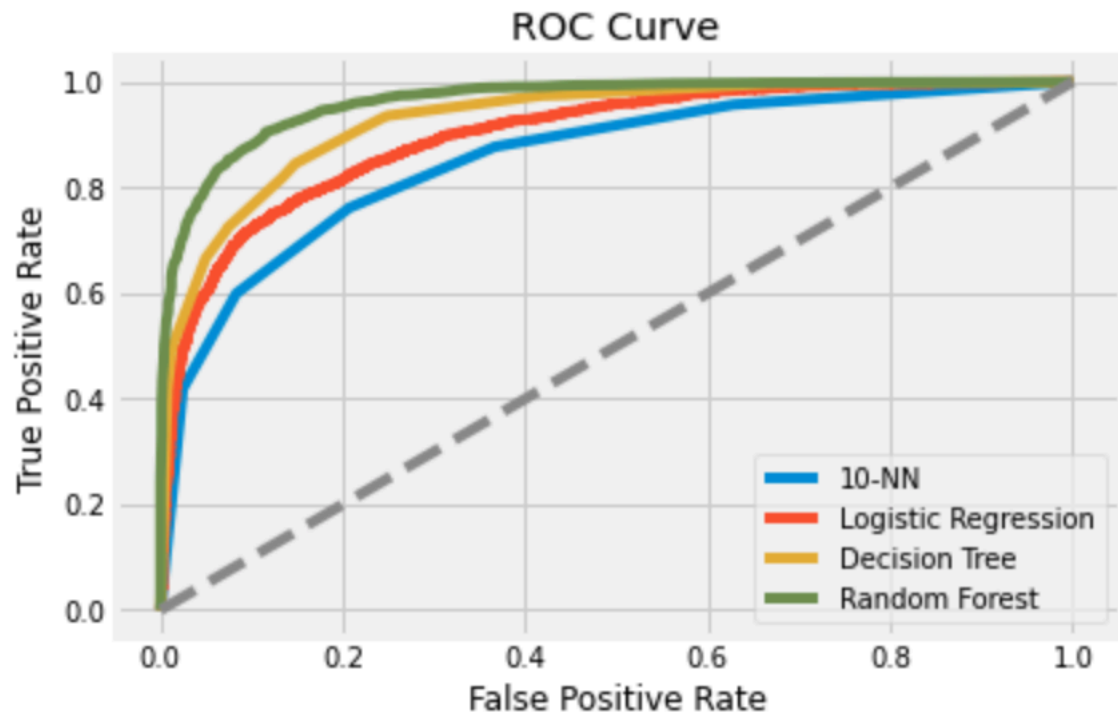- expands,
- reimplements

Target

Geek rating
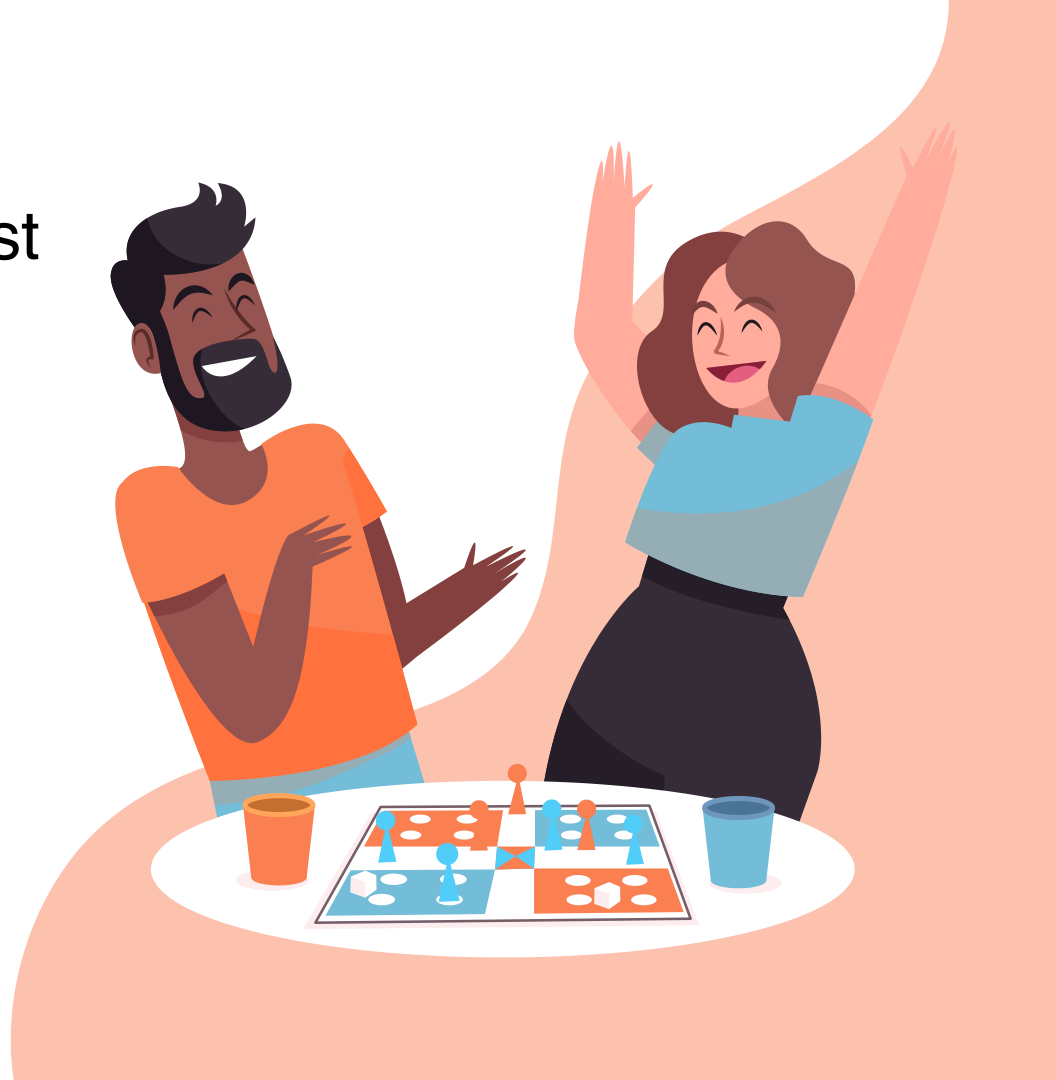Threshold at median - balanced dataset

Metric

Accuracy
Positive and negative event equally important

ROC Curve

Model Choice: Random Forest

- Automatic feature interaction

- Many types of data

- High performance important

- Interpretability not as much
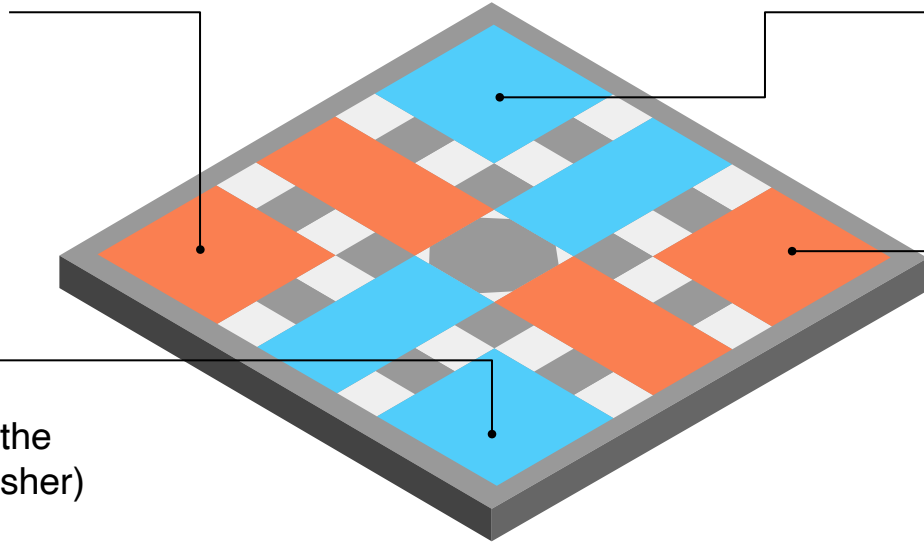
- **Baseline model's accuracy ~ 0.8**

Treat missing and Outlier values

Fill or drop lines, data errors

Feature engineering

Creating binary (dummy) variables

Adding data

More features added to the model (categories, publisher)

Tuning Hyperparameters

Finding best n-estimators and max_depth

Confusion Matrix
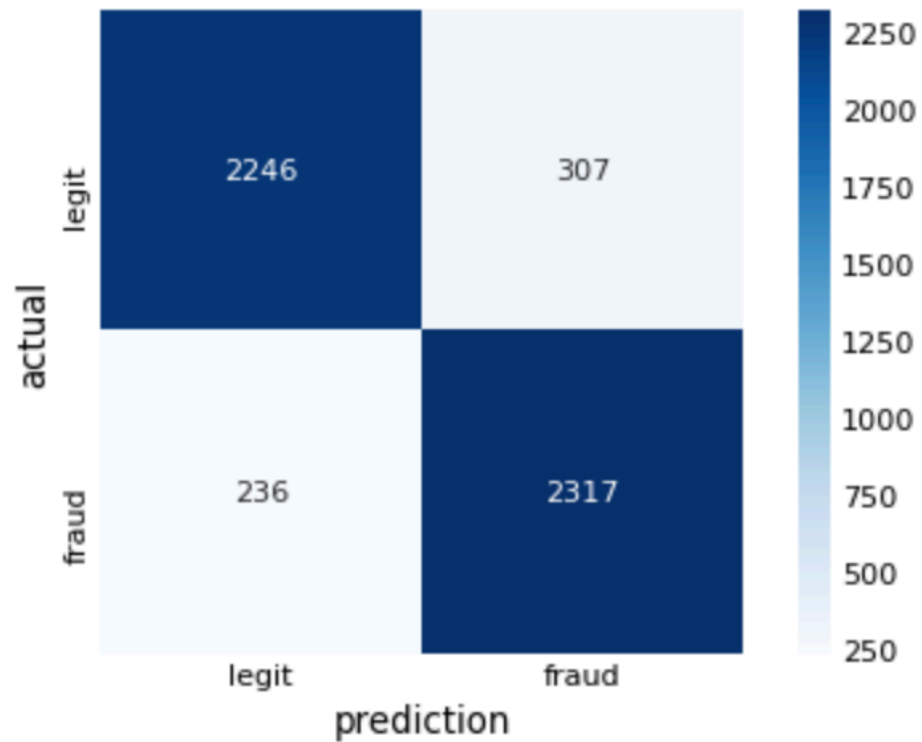
Features with the most

importance:

 - owned

 - number of votes
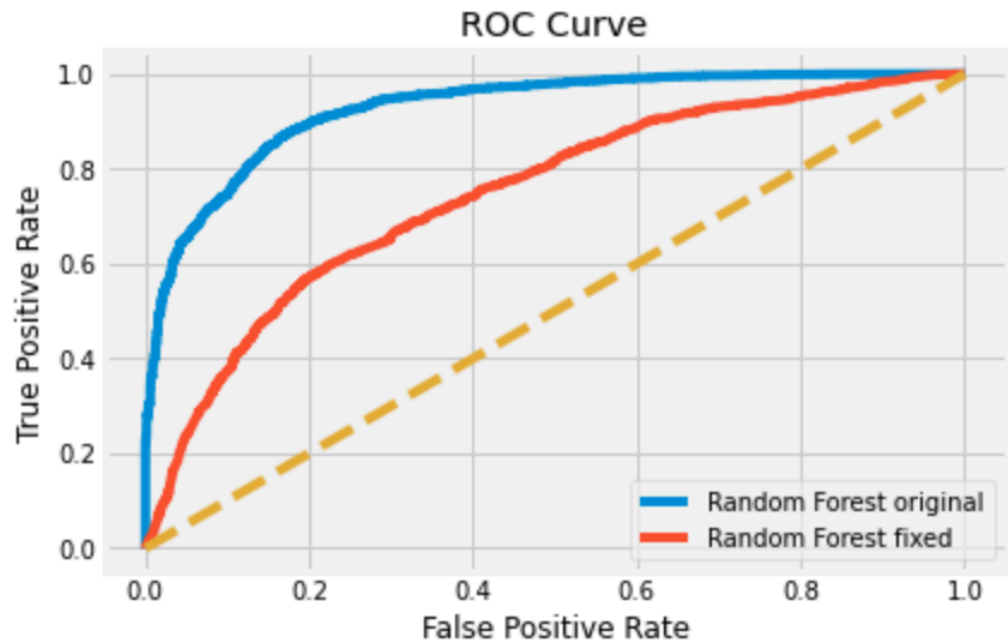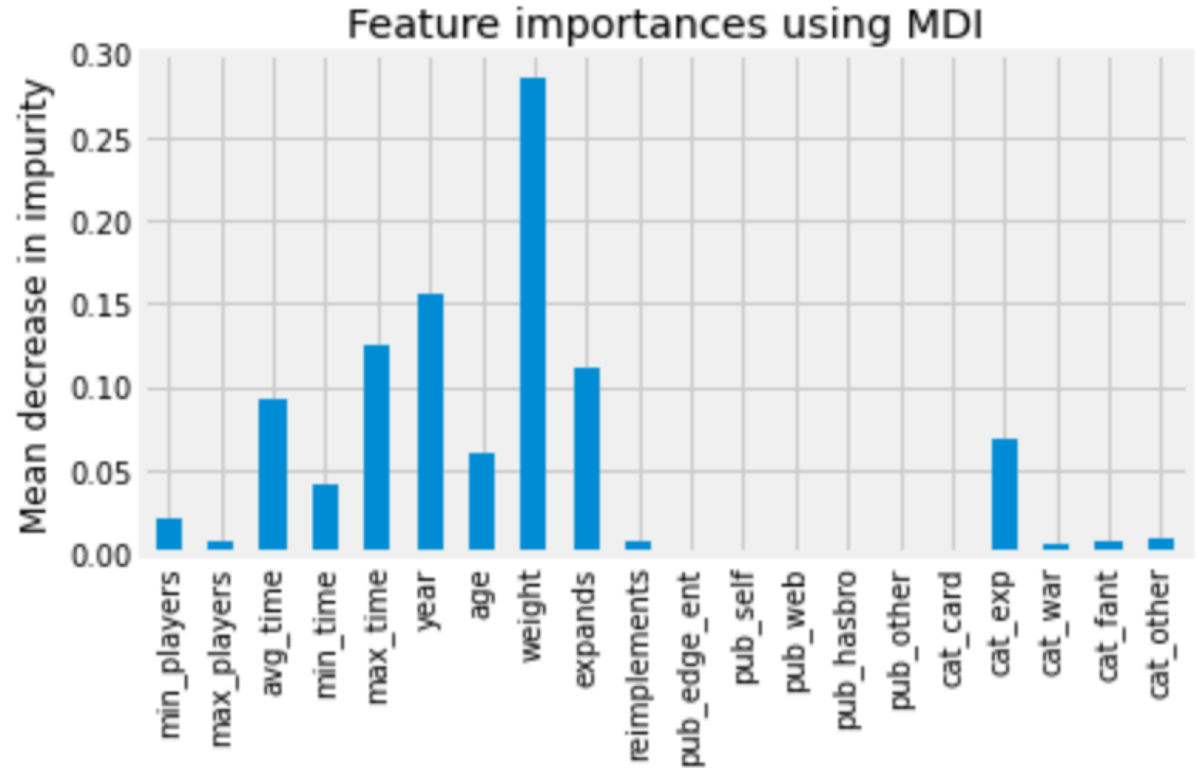
IS THIS RIGHT?

No..

Feature importances using MDI
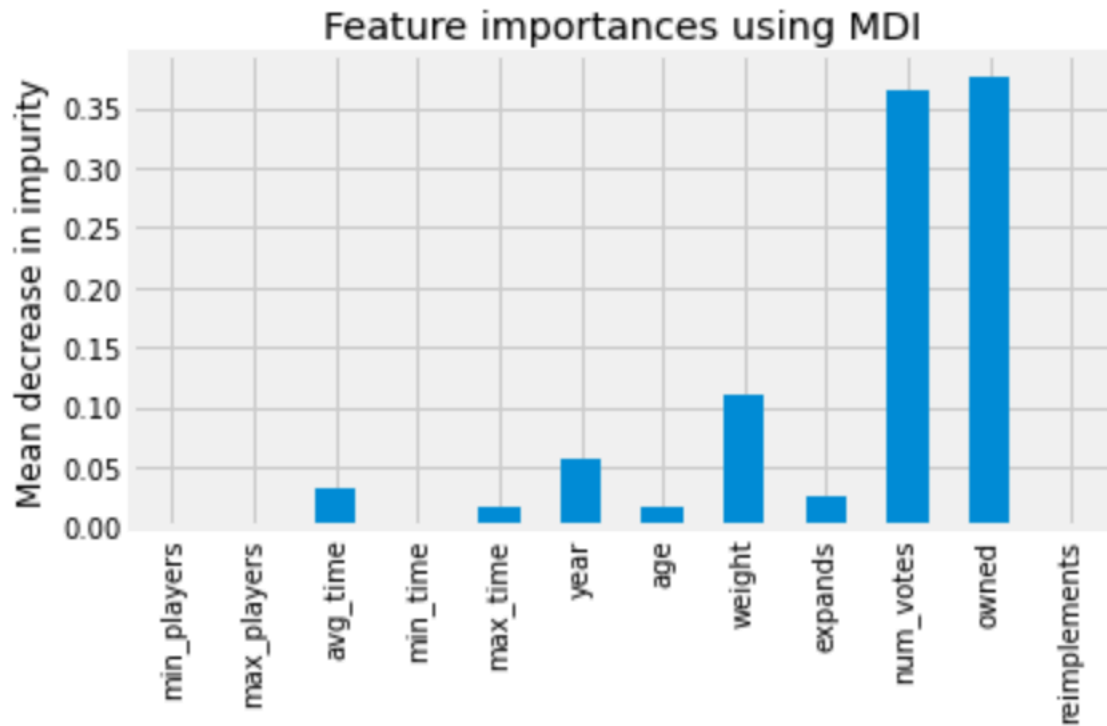
- Random Forest is the chosen

  model

- Accuracy the metric

- Accuracy dropped to ~ 0.6

- ROC Curve still not too bad

ROC Curve

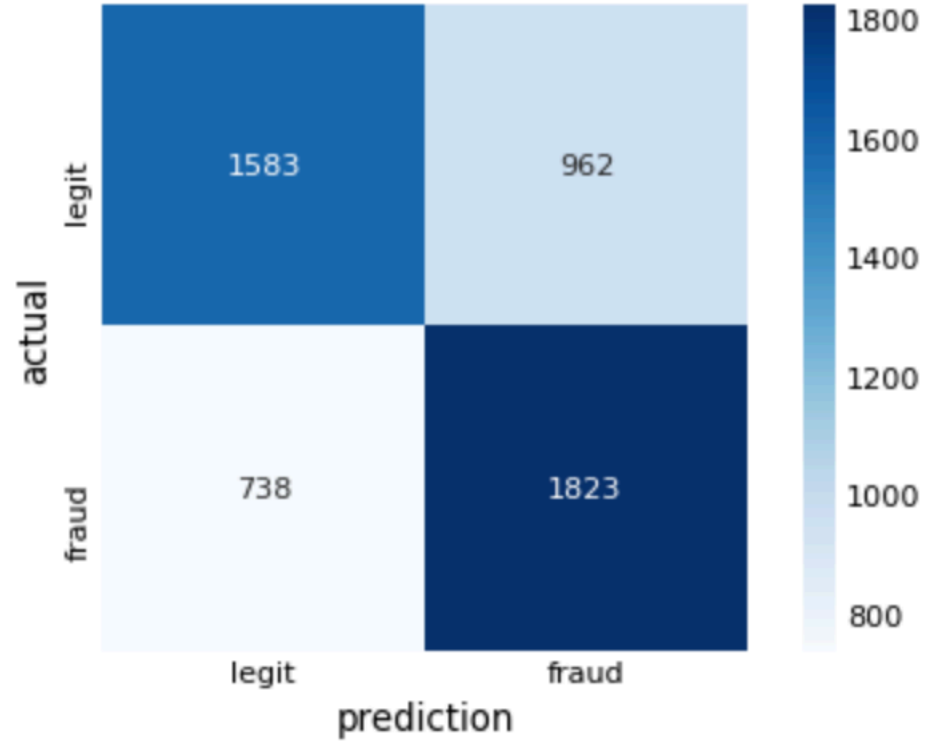# Corrected feature importance chart

Features with the
most importance:
- Weight
- Year
- Avg time
- Max_time
- Expands

Feature importances using MDI
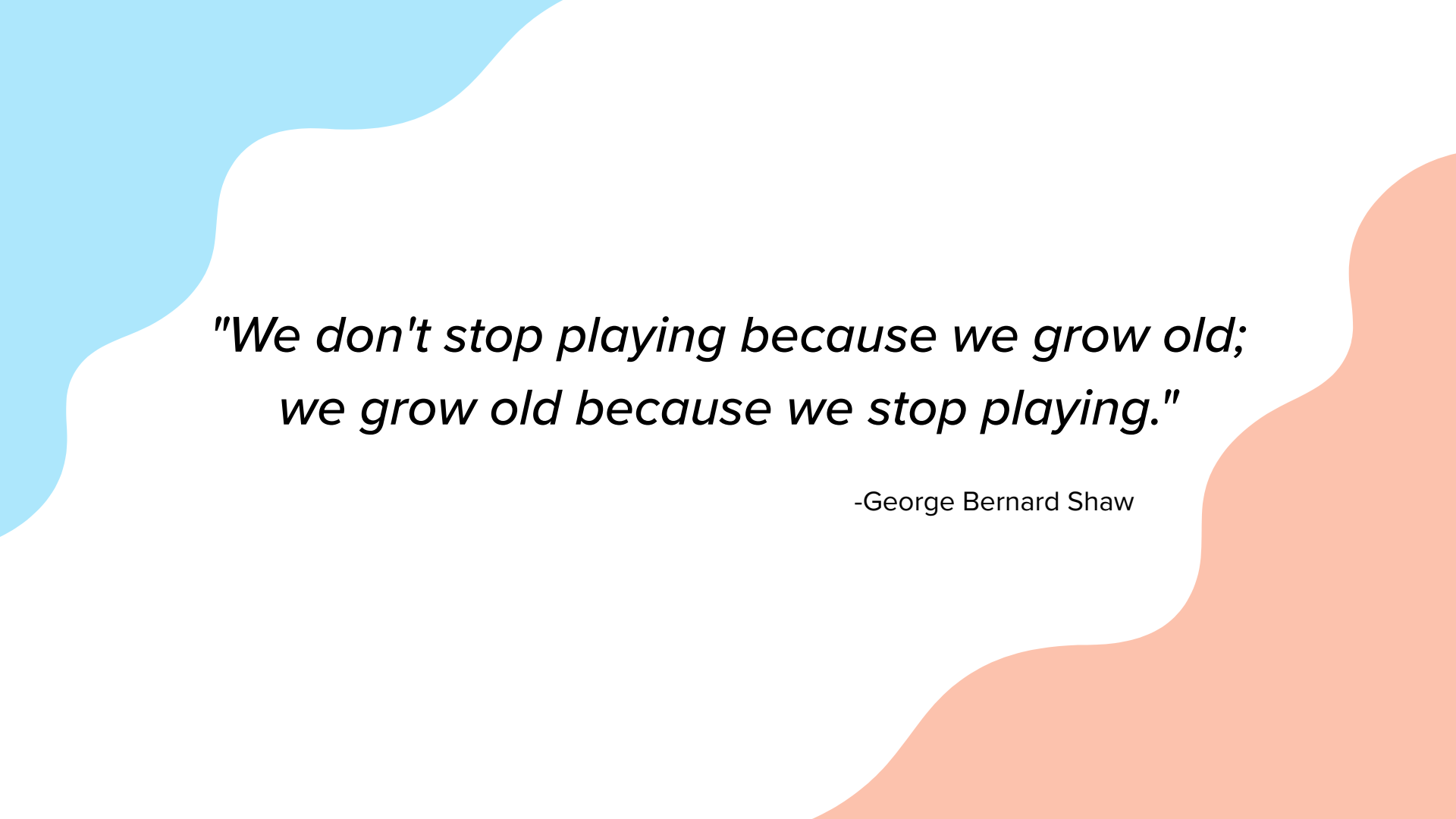
Confusion Matrix on corrected model

Future work

Improve the model further with the new feature set

Build **application** from model - help people planning to crowdfunding

Monitor demand for game projects post pandemic

*"We don't stop playing because we grow old;
we grow old because we stop playing."*

-George Bernard Shaw