Twitch Channel Partnership Classification

Griffin Hundley

Overview

Twitch.tv

- Video game, music, art related livestreams
- Heavy community focus
- Chatroom

Congresswoman AOC playing Among Us
with other top streamers to ~300k viewers

Custom Settings
Map: Polus
Impostors: 2
Confirm Ejects: Off
Emergency Meetings: 1
Anonymous Votes: Off
Emergency Cooldown: 20s
Discussion Time: 15s
Voting Time: 90s
Player Speed: 1.25x
Crewmate Vision: 0.5x
Impostor Vision: 1.25x
Kill Cooldown: 22.5s
Kill Distance: Short
Task Bar Updates: Meetings
Visual Tasks: Off
Common Tasks: 2
Long Tasks: 1
Short Tasks: 5

Ping: 50 ms

Among Us with Ilhan Omar, Hasan, and Pokimane, etc.
19 days ago

Among Us

👁 2,126,636   ⬆ Share

AOC ✔
last live 19 days ago

♡ Follow

Advantages of being a partner

- Subscriber button
- Ad revenue
- Bounties
- Can apply to be shown on the front page

Business Problem

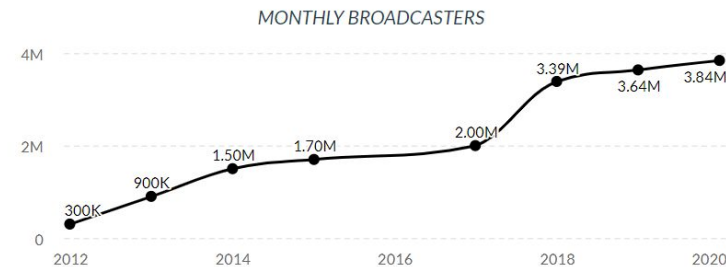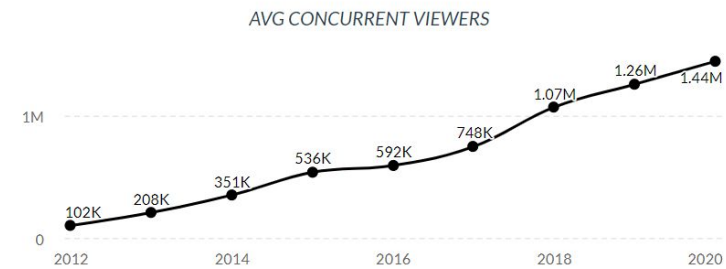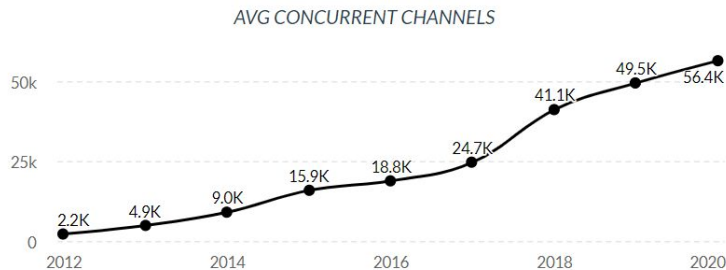Partner applications are manually reviewed

- Not scalable - stats double every few years
- Personal bias can influence who is approved

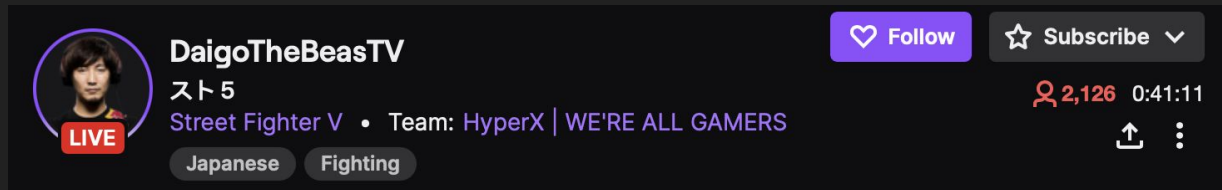Solution: Model that can approve partners based on continuous data

Source:
Mansoor Iqbal
https://www.businessofapps.com/data/twitch-statistics/

AVG CONCURRENT CHANNELS

2.2K (2012), 4.9K, 9.0K, 15.9K, 18.8K, 24.7K, 41.1K (2018), 49.5K, 56.4K (2020)

AVG CONCURRENT VIEWERS

102K (2012), 208K, 351K, 536K, 592K, 748K, 1.07M (2018), 1.26M, 1.44M (2020)

MONTHLY BROADCASTERS

300K (2012), 900K, 1.50M, 1.70M, 2.00M, 3.39M (2018), 3.64M, 3.84M (2020)

Data

Twitch API and Twitchtracker

- ~30,000 channels

Metrics such as

- Views and followers
- How often they stream
- How many games they play, how long do they stream
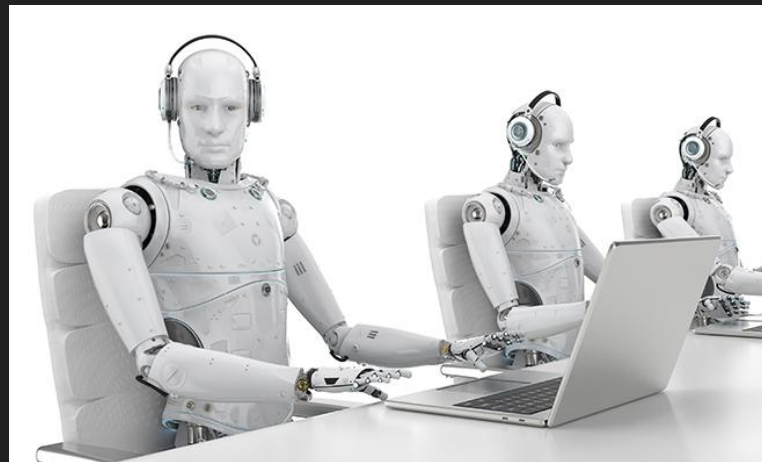- How long do viewers stay and watch

Assumptions

Bias due to only having current data
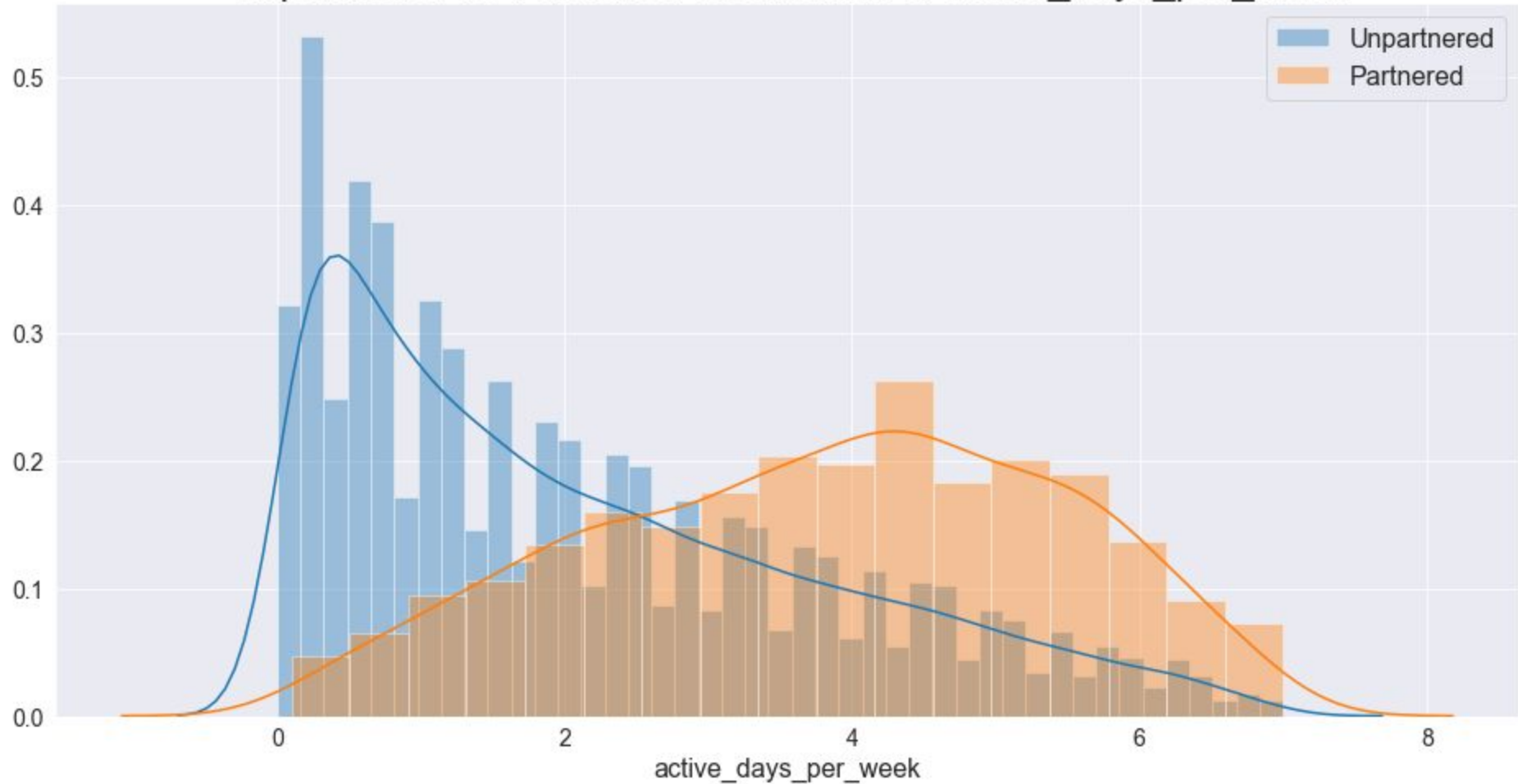
Unable to distinguish 'artificial views'

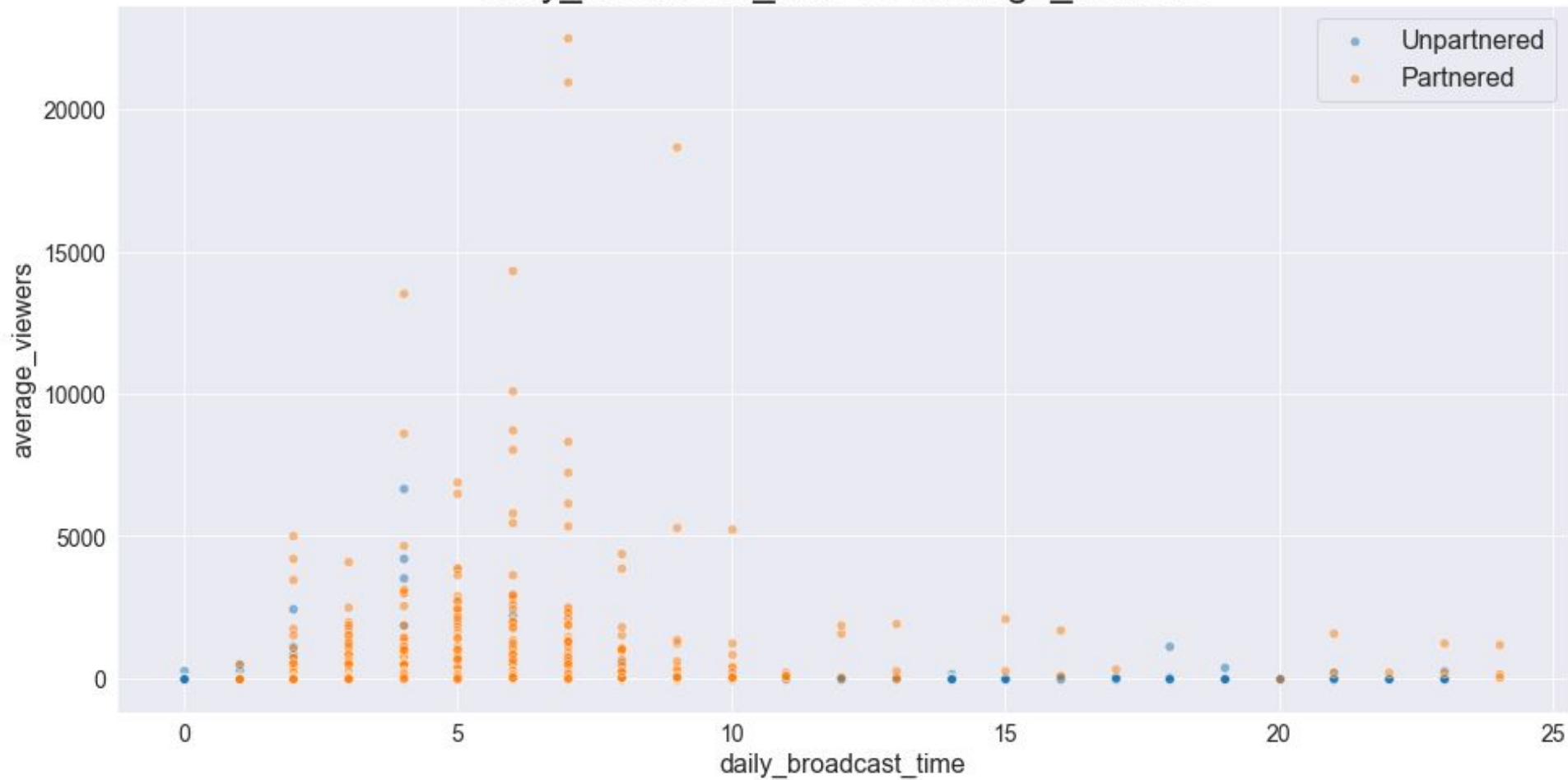Data collection period - late fall

Excluding chat interaction

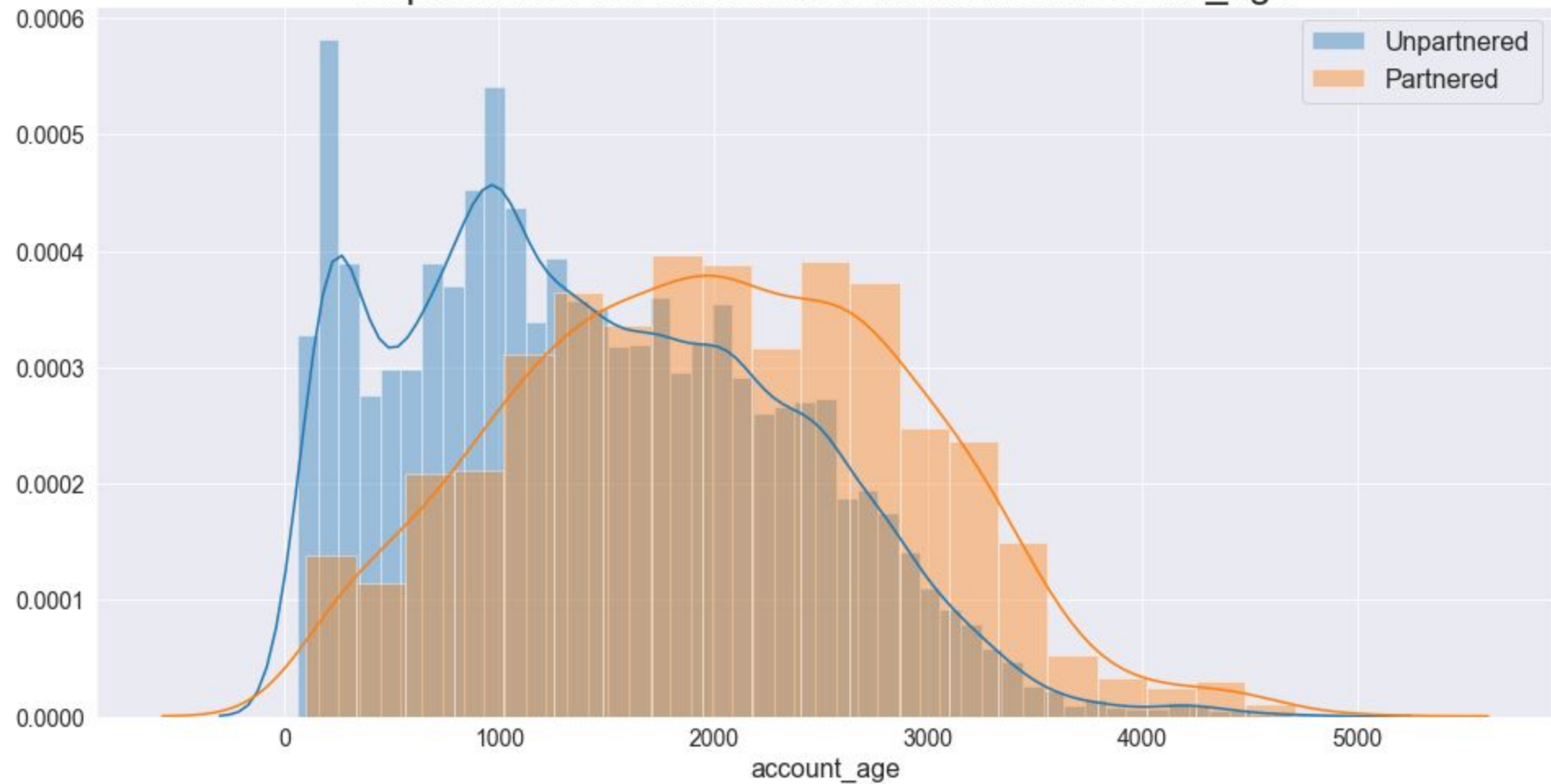Source: R. Massicotte @ sfmagazine.com

Unpartnered vs Partnered Distribution of active_days_per_week
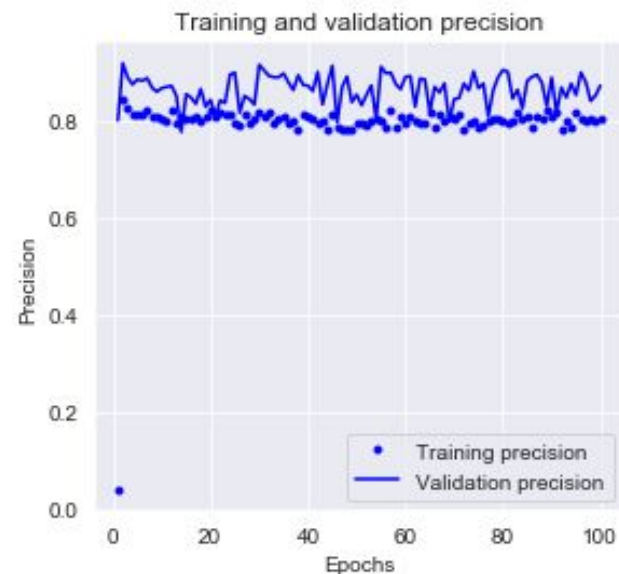
daily_broadcast_time vs average_viewers

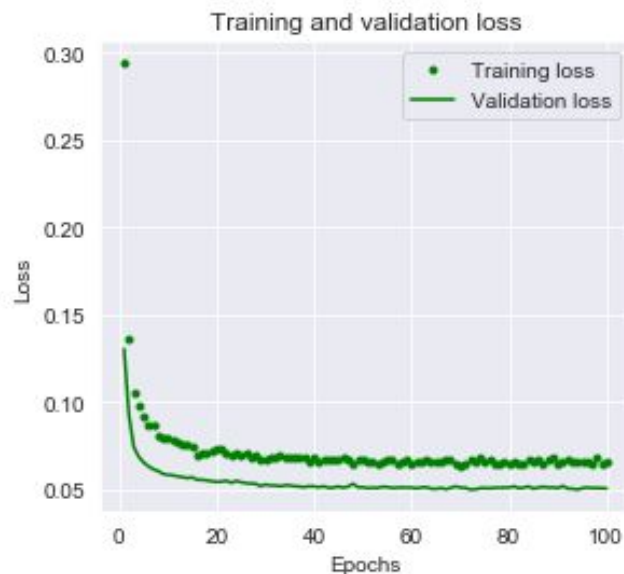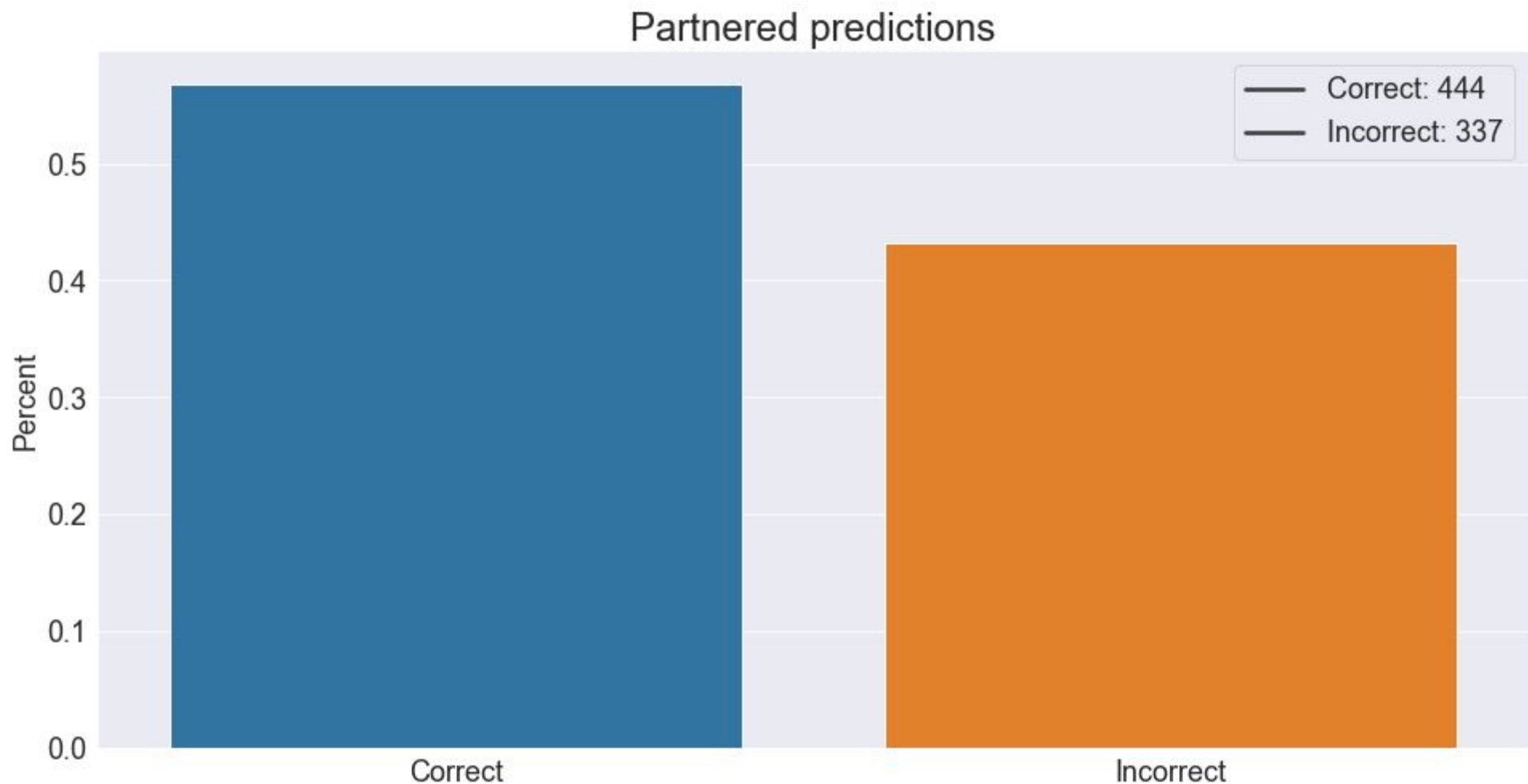Unpartnered vs Partnered Distribution of account_age

Methods

- Highly selective, minimize false positives

Results

Partnered predictions

Correct: 444
Incorrect: 337

Results

Unpartnered predictions

Correct: 16573
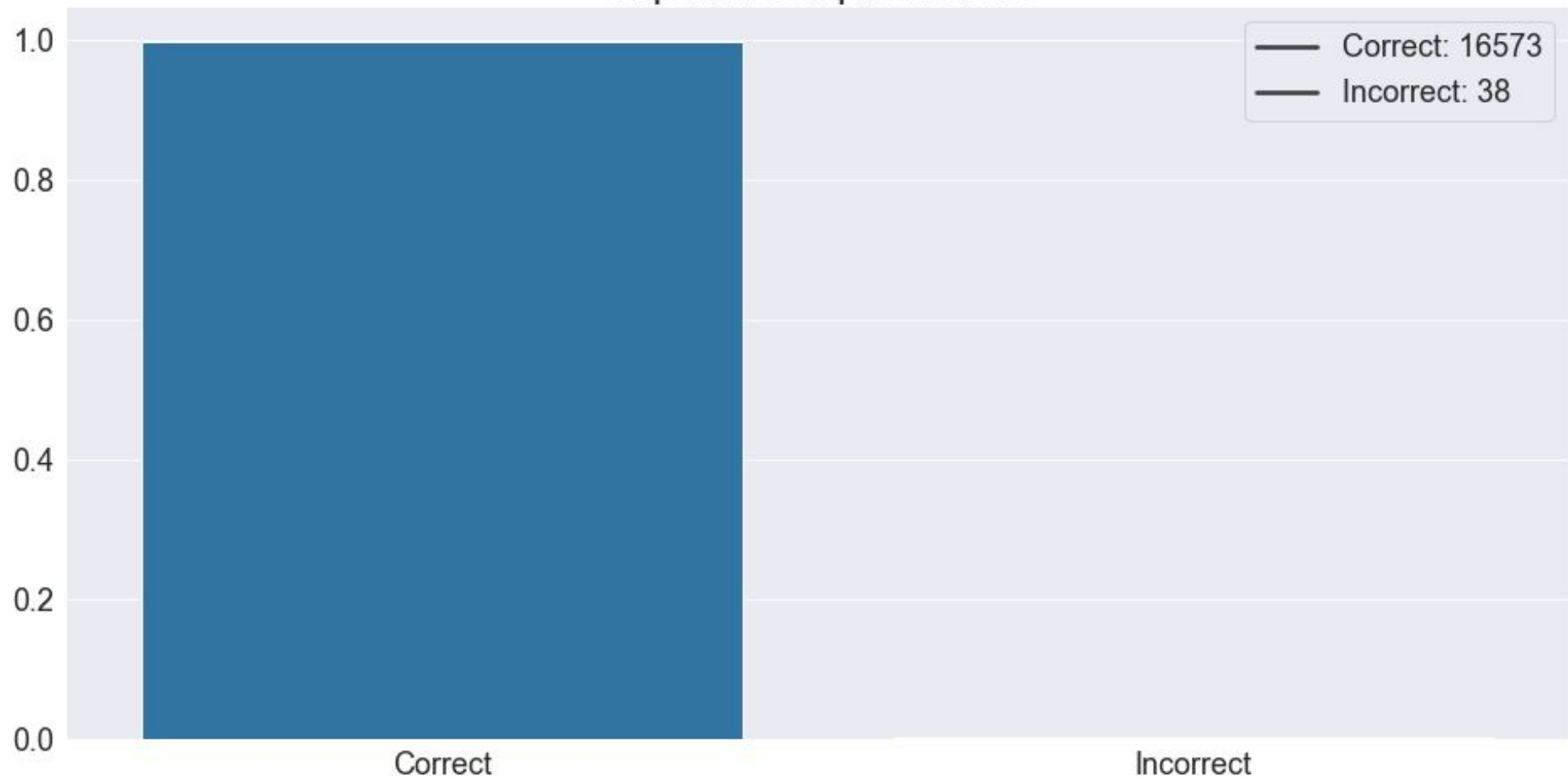Incorrect: 38

Recommendations

First step:

- Start testing on a subset of applications
- High probability channels are still manually reviewed
- Incorporate chat activity/sentiment/brand risk into model

Second step:

- Transition to full automated review

Next Steps

- More data for training (only 4% of partners were in the dataset)
- Tackle assumptions - natural views and historical data
- Get chat activity data
- Do the same analysis but on only affiliates, excluding non-partners
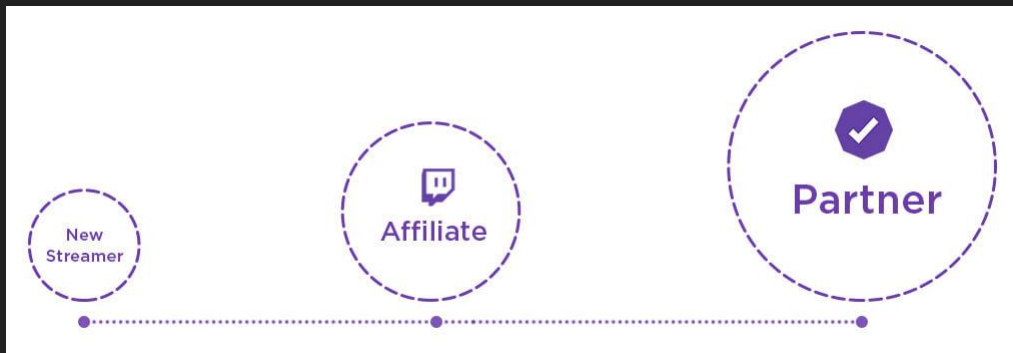
Contact

Griffin Hundley

Email: hundlegq@dukes.jmu.edu

Github: griffinhundley

Linkedin: griffin-hundley-61b020118