Twitch Channel Partnership Classification

Griffin Hundley

Overview

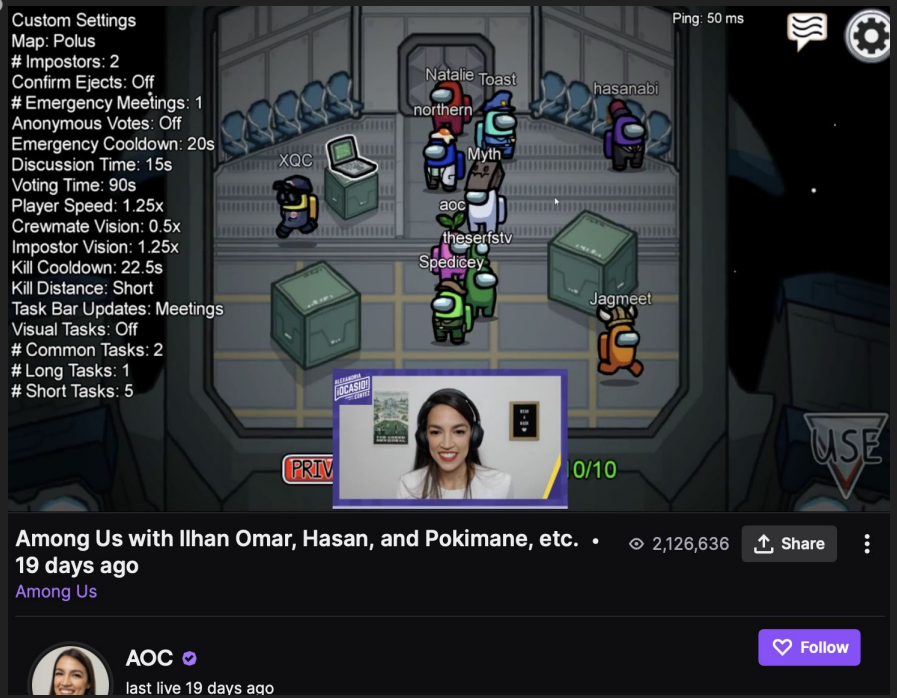
Twitch.tv

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

Partners get:

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

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

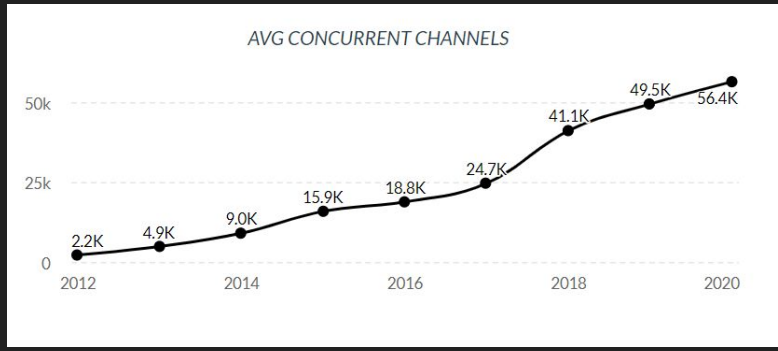


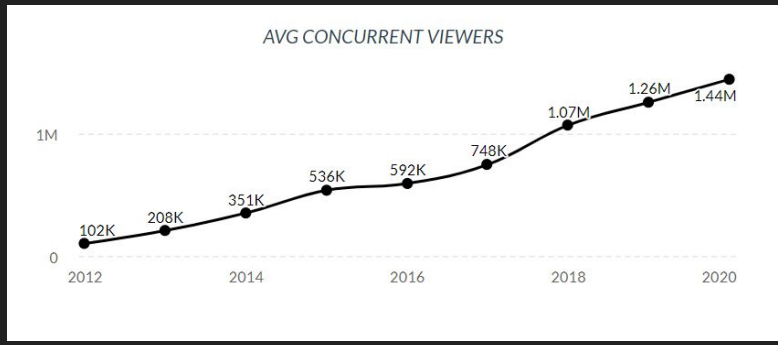
Business Problem

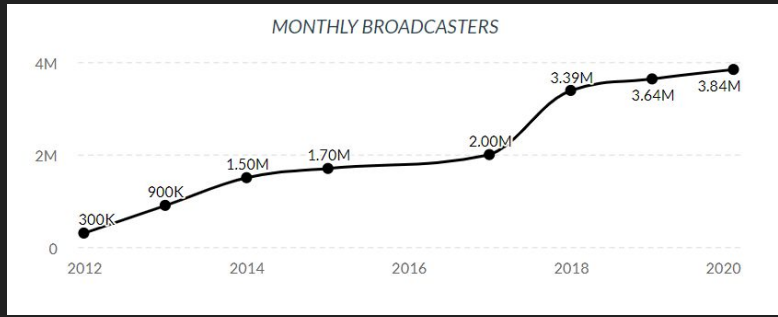
Partner applications are manually reviewed

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

Solution: Model that can approve partners based on continuous data

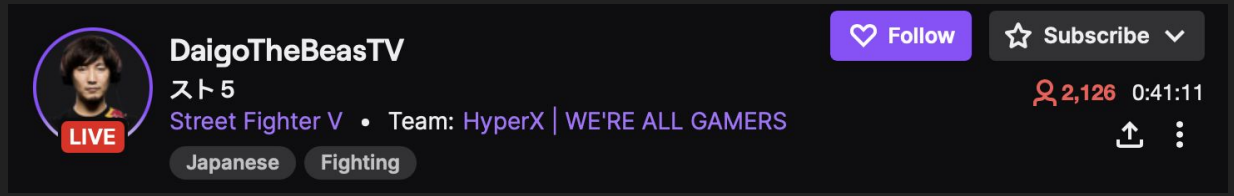






Source:
Mansoor Iqbal
https://www.businessofapps.com/data/twitch-statistics/

Data



Twitch API and Twitchtracker

- ~30,000 channels
- Class imbalance with ~28,000 unpartnered and ~1,500 partners

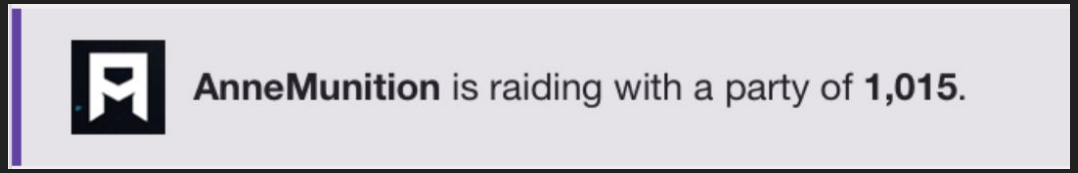
Metrics such as

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

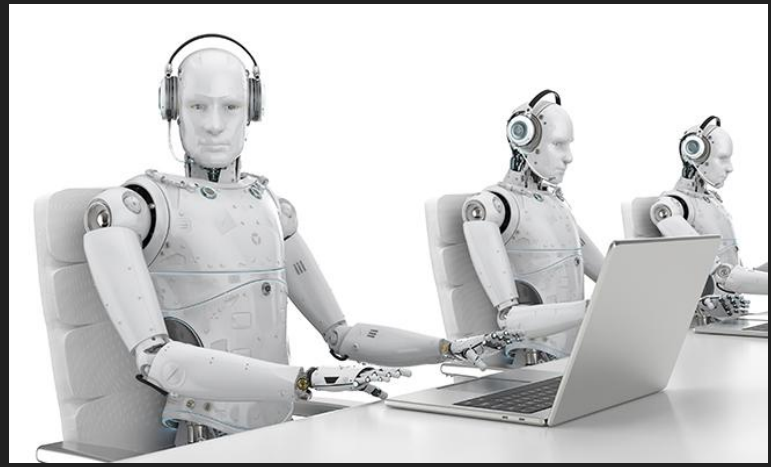
Assumptions

Bias due to only having current data

Unable to distinguish ‘artificial views’ from raids, hosts, viewbots

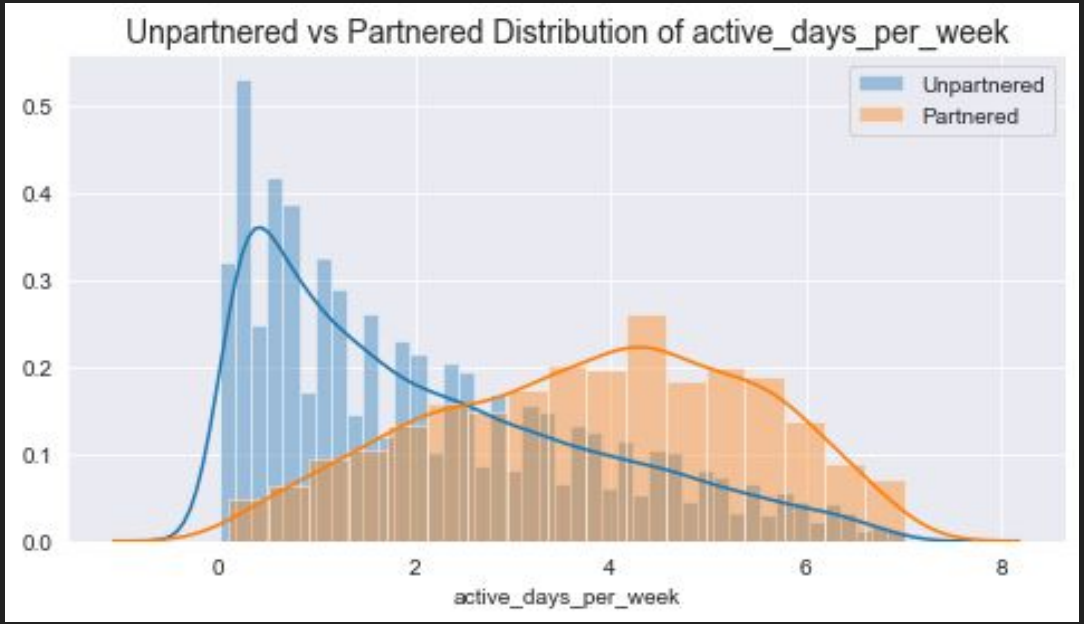


Source: Twitch

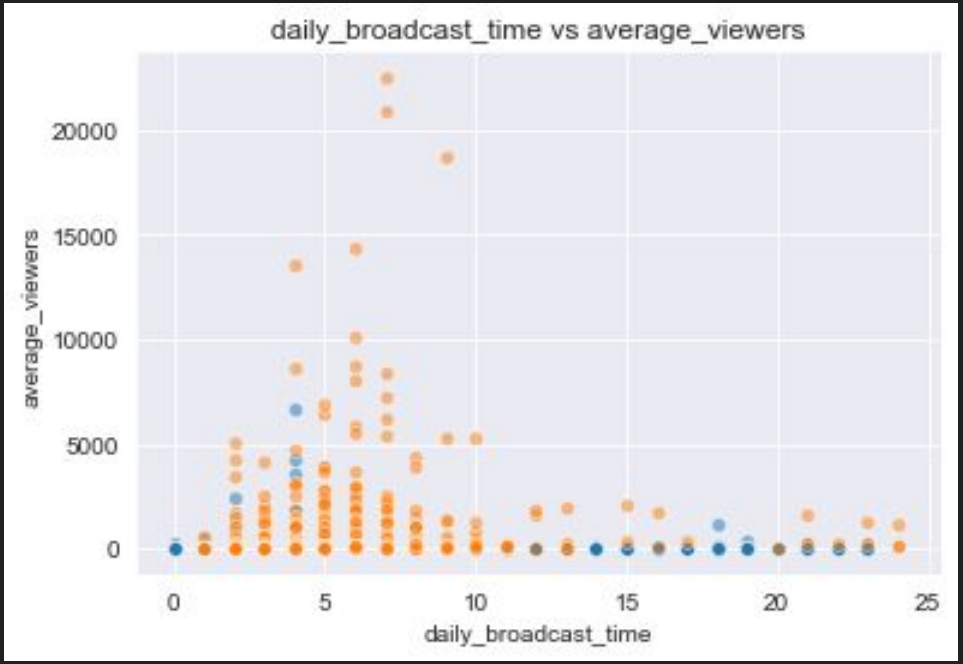


Source: R. Massicotte @ sfmagazine.com

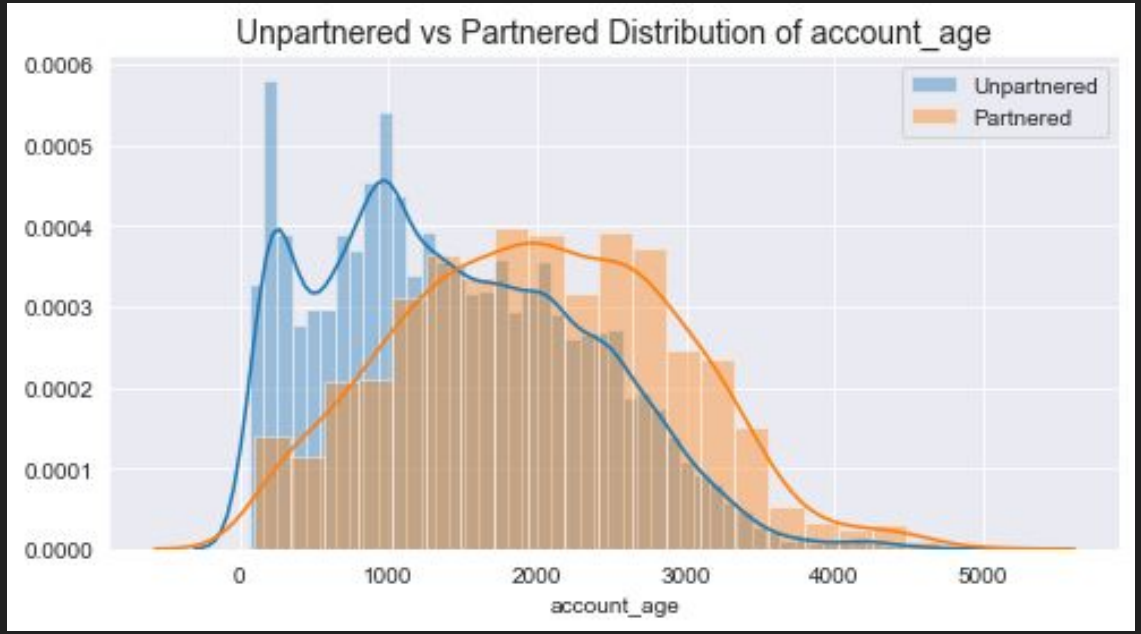
Data


Unpartnered vs Partnered Distribution of active_days_per_week
Unpartnered
Partnered
0.5
0.4
0.3
0.2
0.1
0.0
0
2
4
6
8
active_days_per_week

Data

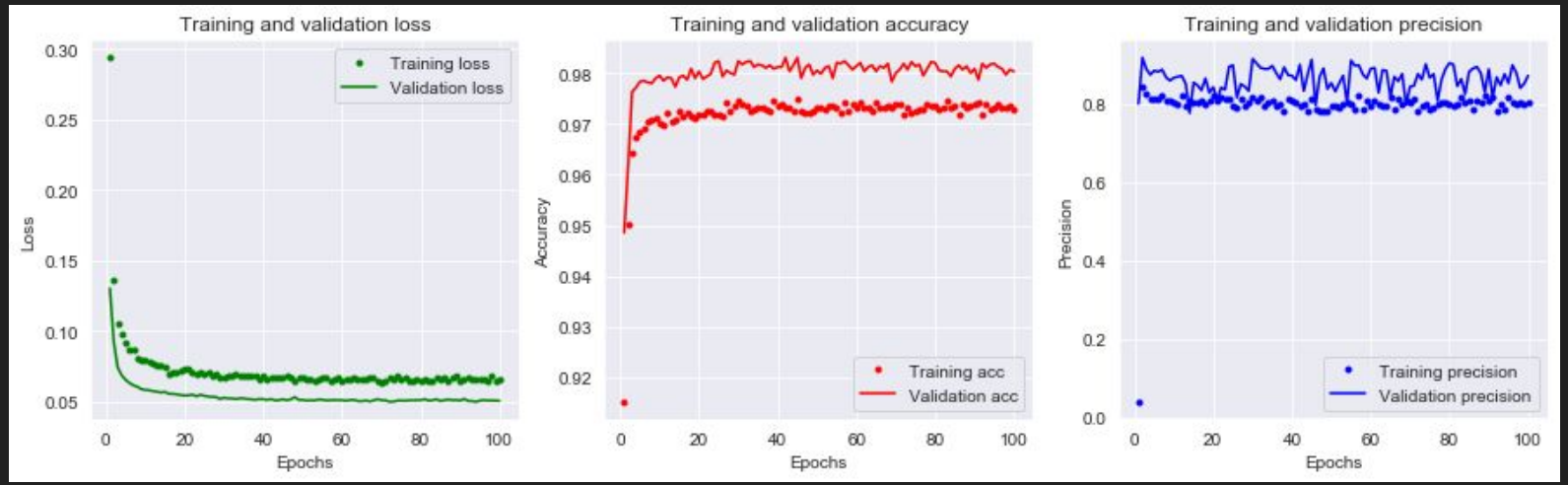

daily_broadcast_time vs average_viewers
average_viewers
20000
15000
10000
5000
0
0
5
10
15
20
25
daily_broadcast_time

Data


Unpartnered vs Partnered Distribution of account_age
Unpartnered
Partnered
0.0006
0.0005
0.0004
0.0003
0.0002
0.0001
0.0000
0
1000
2000
3000
4000
5000
account_age

Methods

- Keras neural net - binary classification
- ADAM optimization, binary crossentropy loss
- Tuned to reduce overfitting
- 100 epochs, batch size of 32
- Precision and specificity



Results

Test set

- 5,591 out of 5,601 non-partners correctly classified as non-partner
- 185 out of 307 partners correctly classified as partner

Scores

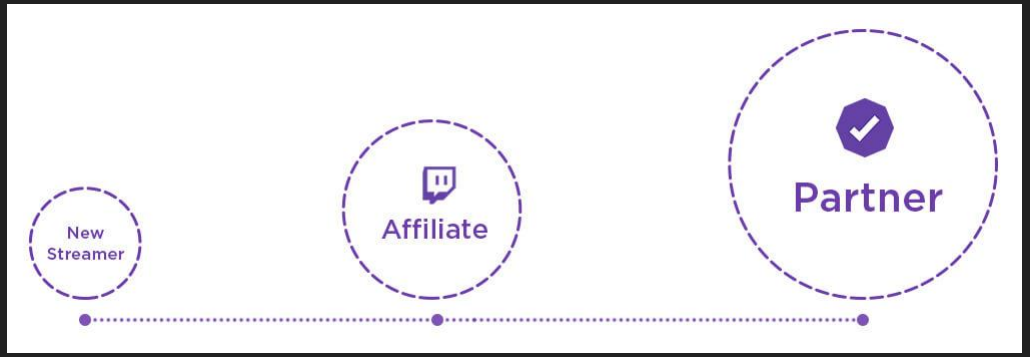
- Precision: 95%
- Specificity: 99.8%

99% effective at selectiveness against non-partners

60% effective at selecting partners

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