



Console-iation



How Xbox Can Increase Market Share
at Playstation's Expense



Stakes

Microsoft's "Xbox Series X" vs. Playstation's "PS5":

System	Xbox Series X	PS5
Release Date	11/10/2020	11/12/2020
Price	\$500	\$500
CPU	8-core 3.8 GHz AMD Zen 2	8-core 3.5 GHz AMD Zen 2
Ram	16 GB	16 GB
Resolution	Up to 8k	Up to 8k

Problem

- Hardware is no longer an important differentiator
 - Games will all load fast and look great!
- Differentiating ourselves:
 - Reliability and longevity
 - Public perception - marketing choices, charity, sponsorships
 - Community - how do we connect our gamers?
 - SOCIAL MEDIA
 - What can we learn about our competitor's customer base by performing Natural Language Processing on Reddit posts?
 - Self-styled "front page of the internet"
 - Community for sharing and ranking content via upvote/downvote

Operation

Scraping 10,000 posts from each subreddit

- r/Xbox - created November 13, 2008 - 212,000+ members
- r/Playstation - created January 25, 2008 - 401,000+ members

Filtering:

- Text only
- Not “removed” (deleted by moderator)
- Chronological

+++Random time delay to prevent server overload

Cleaning and Concatenating

- Load saved .csv files of scraped data into single dataframe
- Combine 10,005 Xbox posts and 10,026 Playstation posts
 - Set Xbox = 0 and Playstation = 1 as the target of my classification analysis
- Replace null values with empty strings, remove duplicates
- Combine 'title' and 'selftext' (subtitle) values into single column

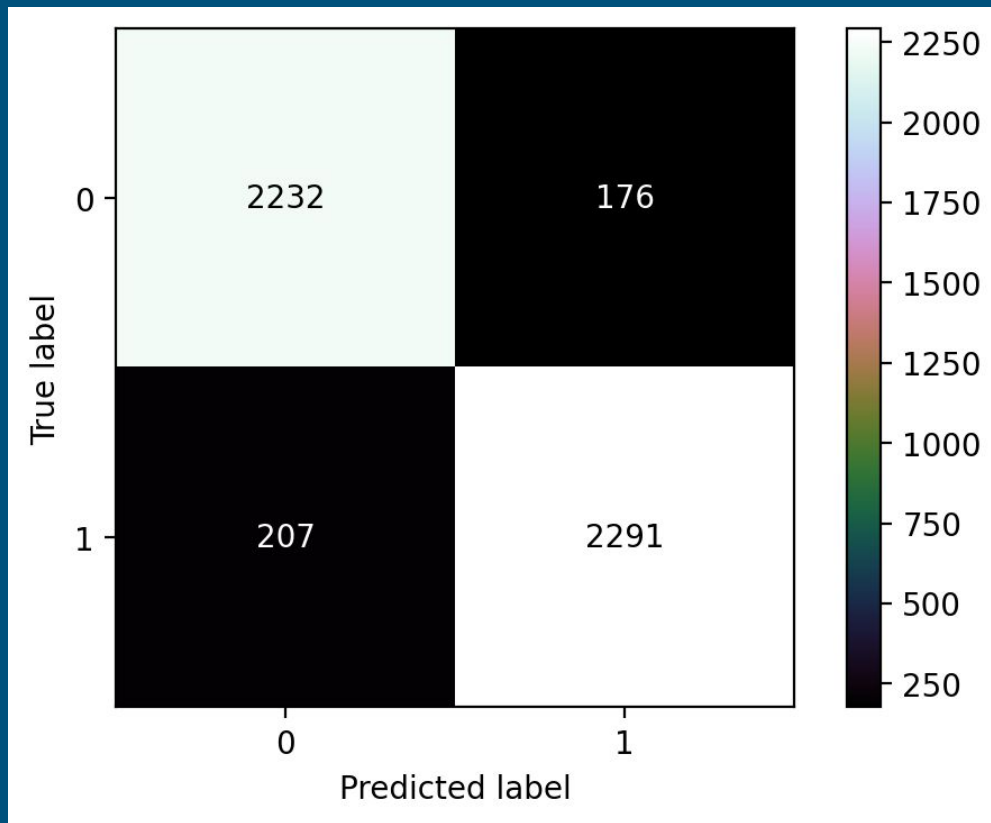
EDA

Preliminary findings:

- Average sentiment across both subreddits is 0.27 (slightly positive)
 - Xbox subreddit is slightly more positive (0.29) compared to Playstation (0.26)
- There are 17,993 unique authors represented in the data set
 - 8,135 from r/Xbox and 9,858 from r/Playstation
- There are 183 cross-posters who appear in both subreddits
- Average post length overall was ~ 71 words
 - Xbox posts averaged 68 words, Playstation posts averaged 75 words

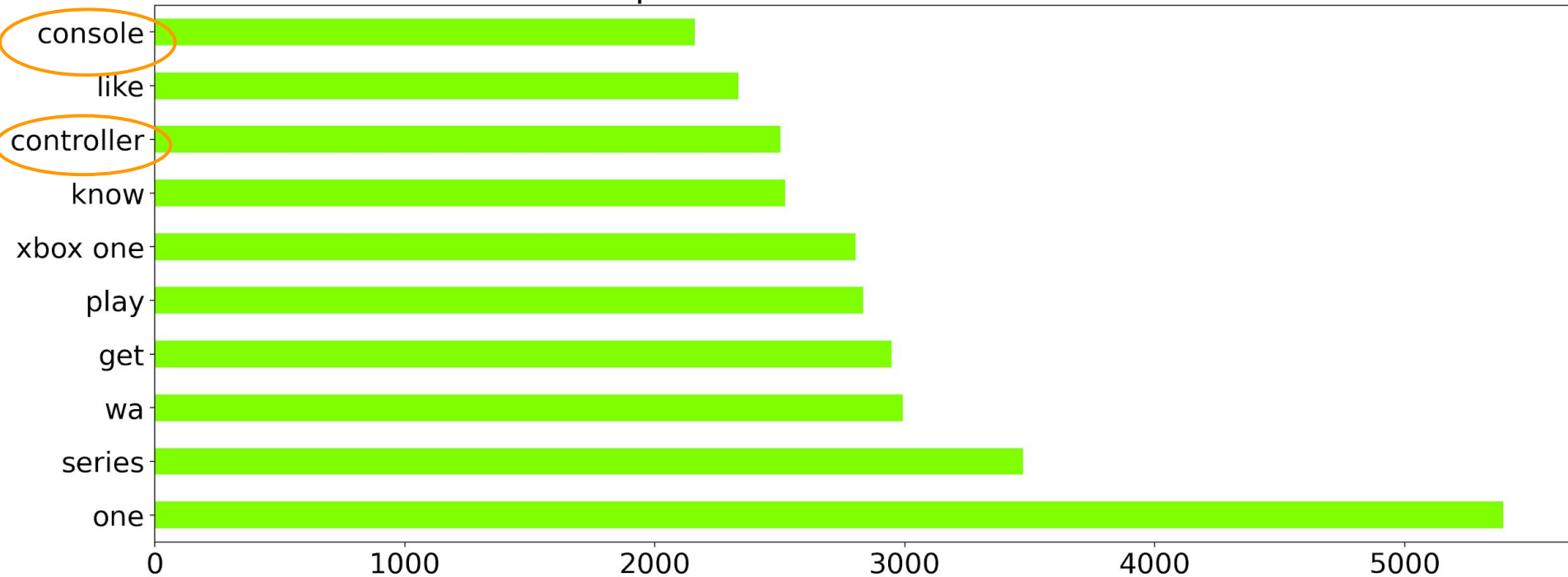
Modeling

- Tested 2 different Vectorizers and 3 Classification methods
- Highest Accuracy with Logistic Regression
 - Cross-validated training data score of 92% accuracy
 - Unseen test data score also 92% accurate
 - Sensitivity of 92%
 - Precision of 93%



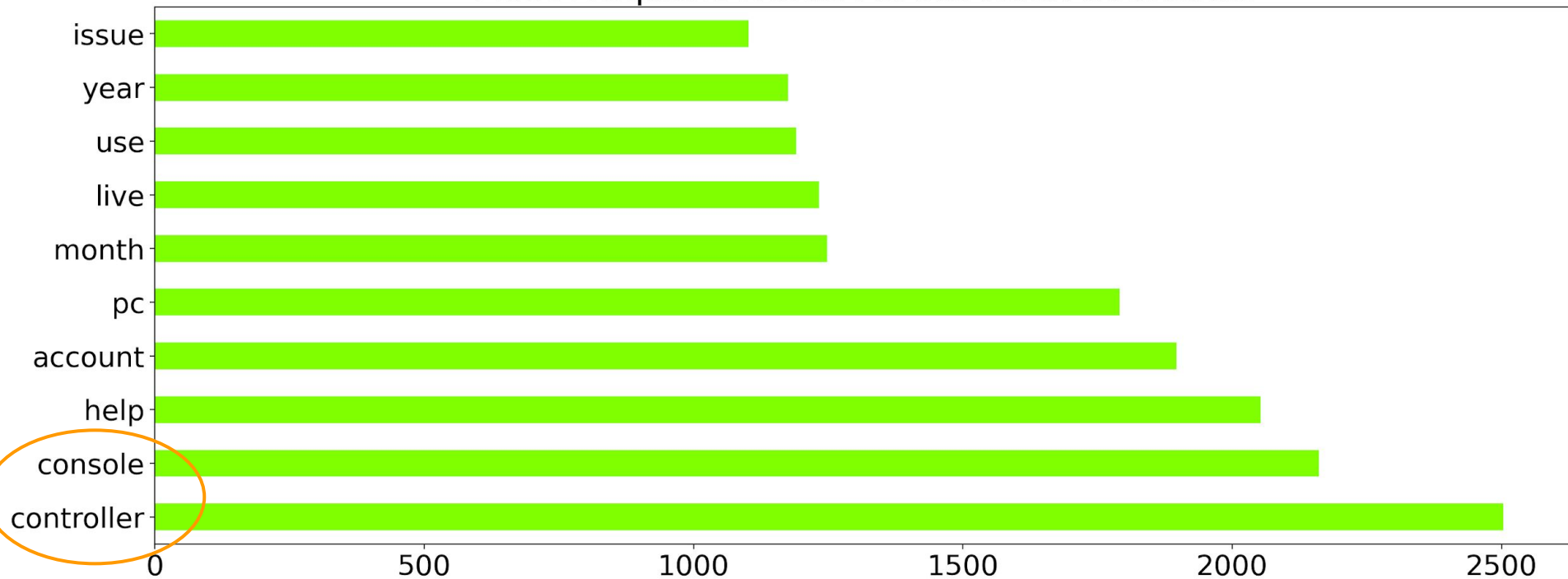
Unfiltered Frequency

Most Frequent Words Across Xbox Subreddit



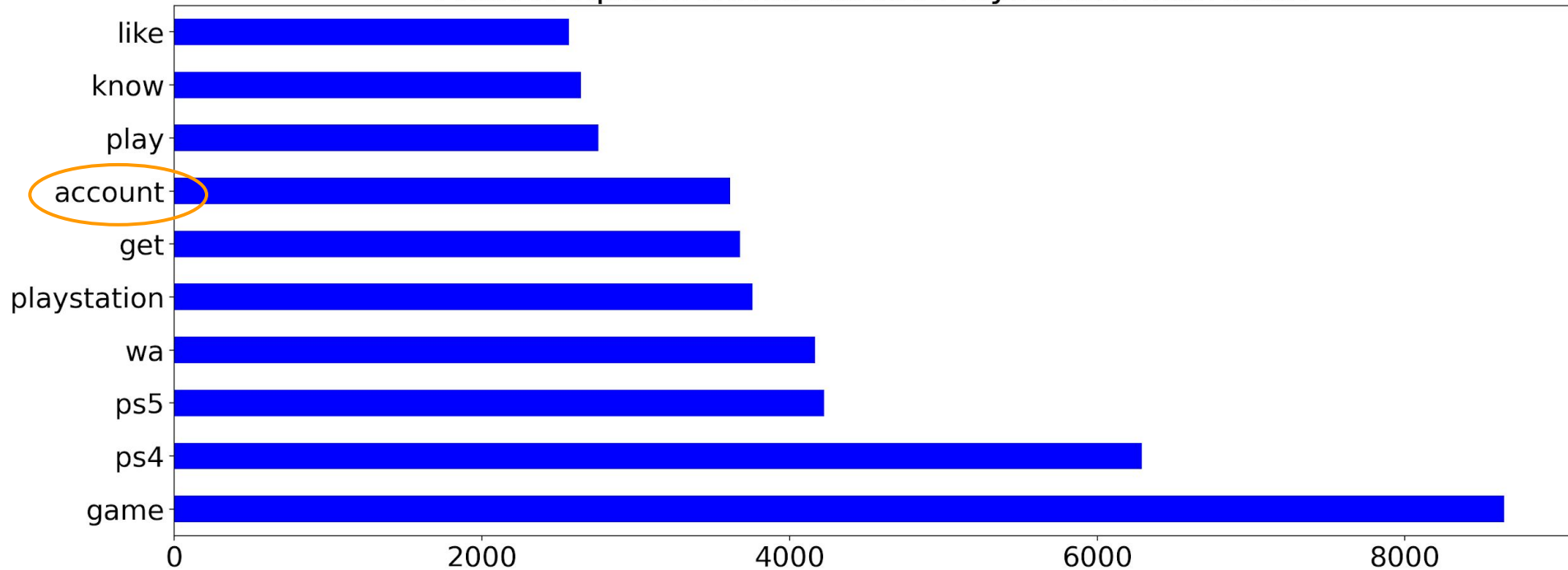
Removing Custom Stop Words

Most Frequent Words Across Xbox Subreddit



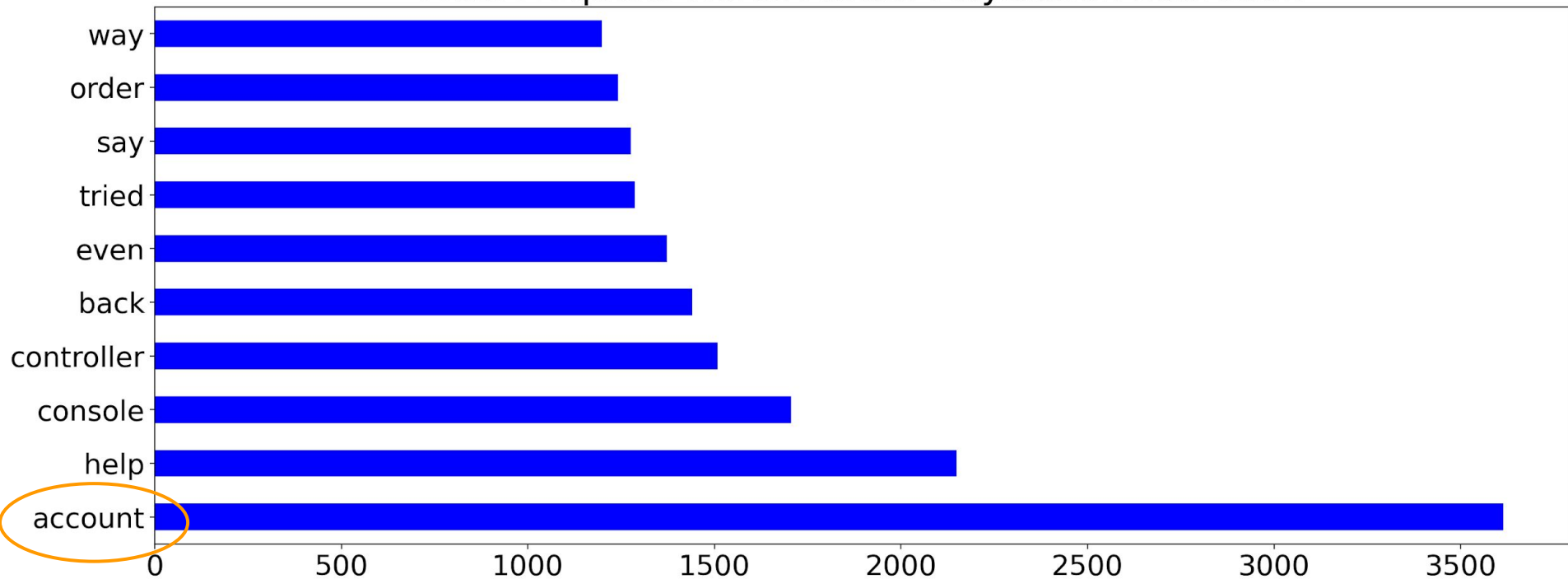
Playstation Unfiltered

Most Frequent Words Across Playstation Subreddit



Playstation Filtered

Most Frequent Words Across Playstation Subreddit



A Closer Look

Removing unique identifiers like
'xbox', 'psn', 'microsoft', 'sony'

Exclusive Games

Account & Security

Advanced features

Word or Pair	Exponentiated Coefficient
Last (of) us	2.43
Remote play	2.40
banned	2.00
hacked	1.98
Fall guys	1.96
wallet	1.95
Safe mode	1.92
database	1.85
3d	1.80
network	1.80

Recommendations

- Highlight games exclusive to Xbox, “Special Edition”, etc.
 - Build and maintain exclusive relationships with top developers
- Advertise ease of Microsoft account creation / maintenance / security and superior customer service
 - Respond to complaints (especially those visible on social media) quickly and comprehensively
- Promote new ways for gamers to connect and new features like VR and remote play

ME AND MY FRIENDS

IN 60 YEARS