

Project 3: Reddit Post Classification

Jason Lee

WC-S5

7-24-20

Building an Community



Gauging player interest through machine learning

A Dime a Dozen



Trading Card Games (TCGs) have been a popular form of gaming for over 25 years

Often based on popular and well-established franchises in order to tap into an existing fandom

Digital Games for a Digital Age



Eternal has gathered a strong fan base for its strong gameplay and design

Being designed for digital play provides advantages over traditional paper games

Growing the Game

In order to succeed, any game needs to:

1. Retain enfranchised players
2. Draw in new players

How do we determine the topics of interest for these two groups in order to best allocate time and resources?



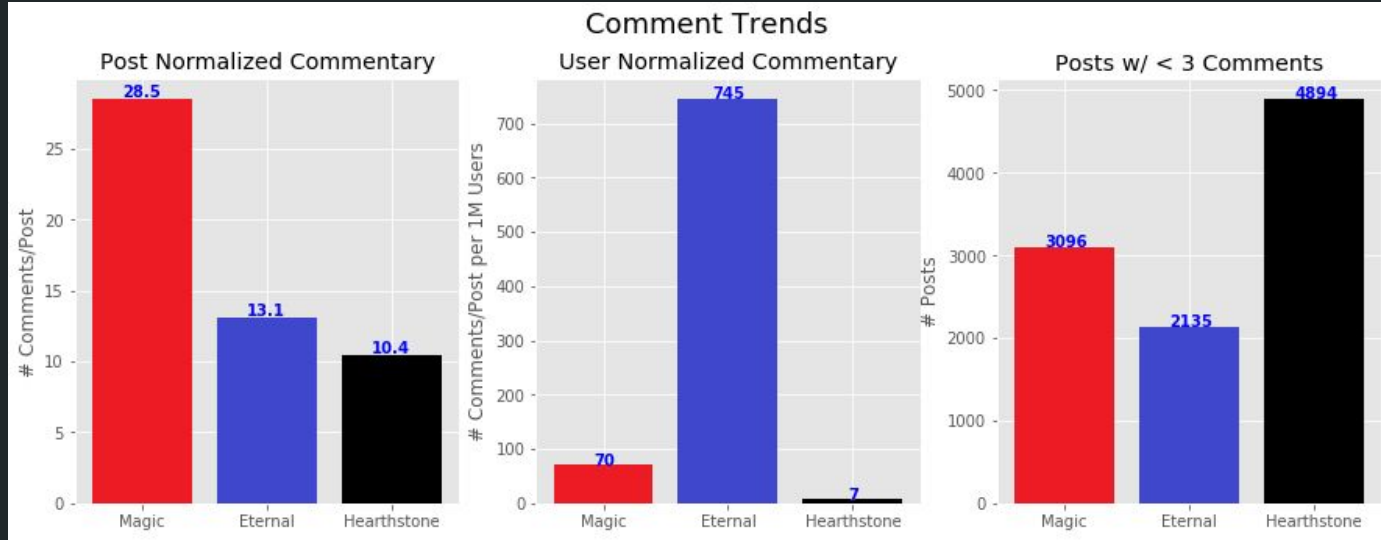
Reading the Players

To gauge the topics of interest for the Eternal community, I utilized machine learning to analyze Reddit posts from the Eternal, Magic, and Hearthstone subs

Over 30,000 posts from the past two years were analyzed to measure member input and create models for inference

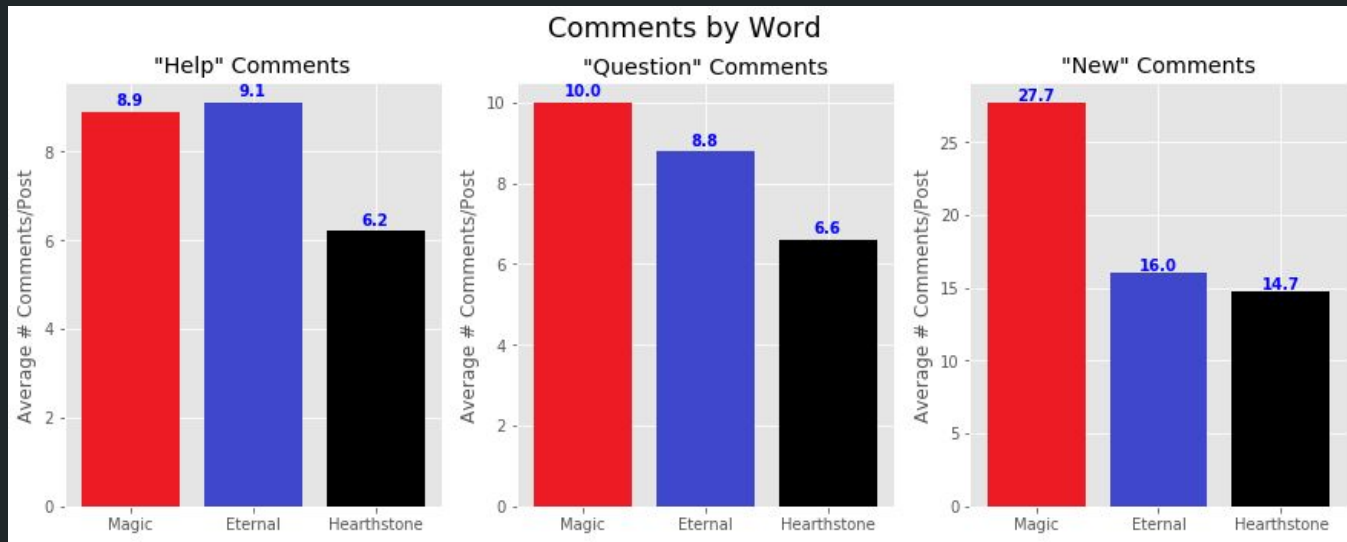


Overall Sub Activity



I measured activity by quantifying the number of comments generated by each sub. Normalized to the number of members, Eternal players are much more involved, which is a positive sign for level of investment

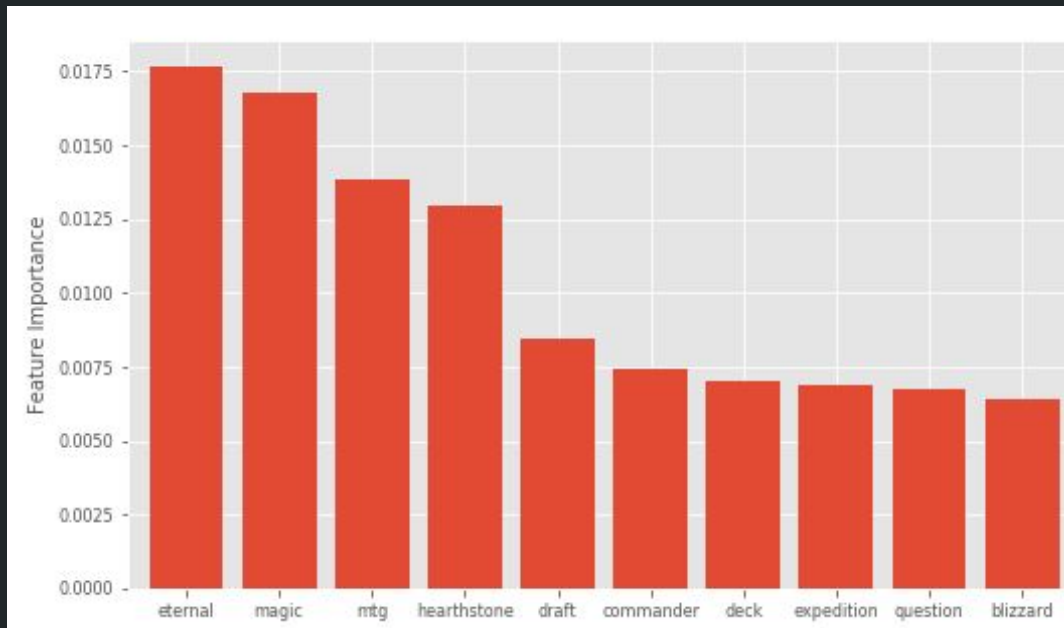
Stay Awhile, and Listen



Creating a welcoming community is vital to drawing in new blood. Counting responses to “help” words shows Eternal members are good about responding to these players



Creating a Clear Identity



Using classification models, I was able to create a predictor with over 80% accuracy between each of the three games based solely on the title of a post. Specificity for Eternal was approximately 78%

Digging Deeper



However, if we use this model, the distinguishing words is not particularly useful. The top four features were:

1. Eternal
2. Magic
3. MTG
4. Hearthstone

To gain better insight, I created models that removed game specific words

The Power of Many

As would be expected, model accuracy dropped significantly without these “easy” words.

To combat this, I used five different models and aggregated feature importances from all the results

	Logistic Regression	Linear SVM	Random Forest	Gradient Boost	Ada Boost
0	scarlatch	highlight	twitch	twitch	sunday
1	chronicle	scarlatch	spoiler	thread	aggro
2	highlight	chronicle	top	list	chronicle
3	fjs	twitch	highlight	top	art
4	xbox	switch	list	highlight	help
5	exalted	achievement	help	podcast	twitch
6	leaderboard	fjs	change	spoiler	list
7	achievement	exalted	update	booster	monday
8	farming	overwhelm	week	scarlatch	check
9	gunslinger	wednesday	podcast	monday	month

Putting it Together

	Magic	Eternal	Hearthstone
0	booster	scarlatch	class
1	prerelease	twitch	bob
2	lgs	highlight	lethal
3	ruling	thread	highlander
4	reprint	xbox	nerfs
5	myb	achievement	otk
6	edition	chronicle	hp
7	sleeve	monday	odd
8	maro	sunday	solo
9	paper	farming	til

Looking at the aggregate results, several topics for Eternal stand out

- Scarlatch, a developer with in-game interactions with players
- Twitch, the popular streaming platform
- Xbox and achievement, which are tied to console play
- Chronicle - Eternal: Chronicles of the Throne is a tie-in board game

Conclusions and Recommendations

From my results, I believe focusing on these areas will help grow Eternal's playerbase:

1. Foster human relationships, whether it is through developer/fan interaction, or promoting the growth of leaders in the community itself, such as prominent streamers (for example, Team Rankstar ranked 34 on the aggregate list). This keeps enfranchised players involved even if gameplay quality drops
2. Pursue alternate methods of presenting the game to players. Whether this is through video game consoles (where games like Magic and Hearthstone are not available), board games (to reach people who like games in general but may not have played a TCG) or other forms of media



AMA

Thanks for Listening!

Questions?