



# Steam Game Review Analysis

Kai Zhang, Fanyu Wang, Fu Yang, Shuai Wang

**Professor Rong Liu**

BIA-660

# Content

- Background & Motivation
- Data Description
- Features Extraction
- Review Classification
- Trend Prediction
- Conclusions & Further Work



# Background & Motivation

- Game industry is booming in recent years. 2017 US gaming revenue had reached \$ 36 billion.
- Steam is the most effective digital distribution platform for video games.
- This project will focus on the review data to analyze players' feedback.
- The goal of project is to help developer improve game quality





## Related Work

In Kyle Speaker's work, his team analyzed the names of 5820 steam games released in 2018. They have discussed:

1. Title size. The average title length is between 1 and 5 words. The top 10% games have a slightly longer title on average.
2. Unique Word Usage. 27.93% of titles were found to include a unique word within their title. 48.91% of the top 10% games title utilized at least one unique word.
3. Multi-Part titles. 15.4% of the total dataset was found to include multiple part. 19.17% of top 10% games had multi-part titles.

Kyle Speaker is a data scientist. Head of Analytics at Aces Design.

## What's new

In this project, we focus on the reviews of some specific games, which is helpful for developers and players to further understand games.



# Web Scrapping and Data Preprocessing



# Web Scrapping


## Original Sin I

<https://steamcommunity.com/app/230230/reviews>

## Original Sin II

<https://steamcommunity.com/app/435150/reviews>

490 people found this review helpful  
9 people found this review funny

 **Recommended**  
165.6 hrs on record

Posted: June 6, 2017

EARLY ACCESS REVIEW


You will love this game if...

- You like games that don't hold your hand
- You like games that allow you freedom to customise your characters and experiences
- You like games that TRULY reward exploration and investigation

\*Legit, if you don't sit there scrolling over the map looking for hidden chests, hidden dungeons, etc. you're doing it wrong!


- You like high strategy, turn based combat that rewards forward thinking and planning
- You like finding, using, selling or collecting loot
- You have 1 or more friends that will play with you. This game shines in multiplayer

The handful of little annoying bugs and glitches are completely and utterly outweighed by how captivation and enthralled this

 **ProfessorButts**  
112 products in account

11


145 people found this review helpful  
39 people found this review funny

 **Recommended**  
136.9 hrs on record

Posted: March 2

Divinity: Original Sin 2 is like playing a very enthusiastic Dungeon Master's homebrew campaign. Said DM really fell in love with the cosmology and deities section of the Dungeon Master's Guide as well as the idea of creating an epic, REALLY LONG arc where his players eventually achieve godhood. Mr. DM also put an awful lot of time into fleshing out his world and making sure you know that HIS elves are DIFFERENT (they EAT people and get their MEMORIES, see), though for some reason he decided to keep his dwarves Scottish.

This DM is not perfect and has his faults. For one, he really loves throwing environmental effects all over the battlefield to the point of it becoming annoying. Fire spreads 50 feet in an arc! Now ice shoots out over the fire! Now smoke and lighting and tentacles spread atop the fire and ice! Yeah! Near the end of the campaign, his planning starts to slip, both in combat design and in lore. For

 **PixelGrotto**  
152 products in account

6



# Data Preprocessing

5500+ Reviews from 'Original Sin II' and 1300+ Reviews from 'Original Sin I' are scraped in the following format:

	user_name	user_product	helpful	funny	post_date	recommend_or_not	game_time	review
0	facciola007	304	4	0	2016/09/15	Recommended	235.2	amaze happy seem good far alpha release pre da...
1	Dixie Normous	257	2	0	2016/09/15	Recommended	156.3	enjoy i've play far graphically look great bar...
2	InquisitorTea	364	31	1	2016/09/15	Recommended	8.5	run well early access original sin good every ...
3	Longmire	115	2	0	2016/09/15	Recommended	16.6	play hour far seem like great road far multipl...
4	Fanest	0	6	2	2016/09/15	Recommended	212.8	play arena pvp tell source point op option pla...

The data below is removed:

- (1). Duplicates and reviews written in other language except English.
- (2). Useless and unreadable information such as symbols and emoji.



# Data Preprocessing

## Tokenization

Data formatting and sorting.

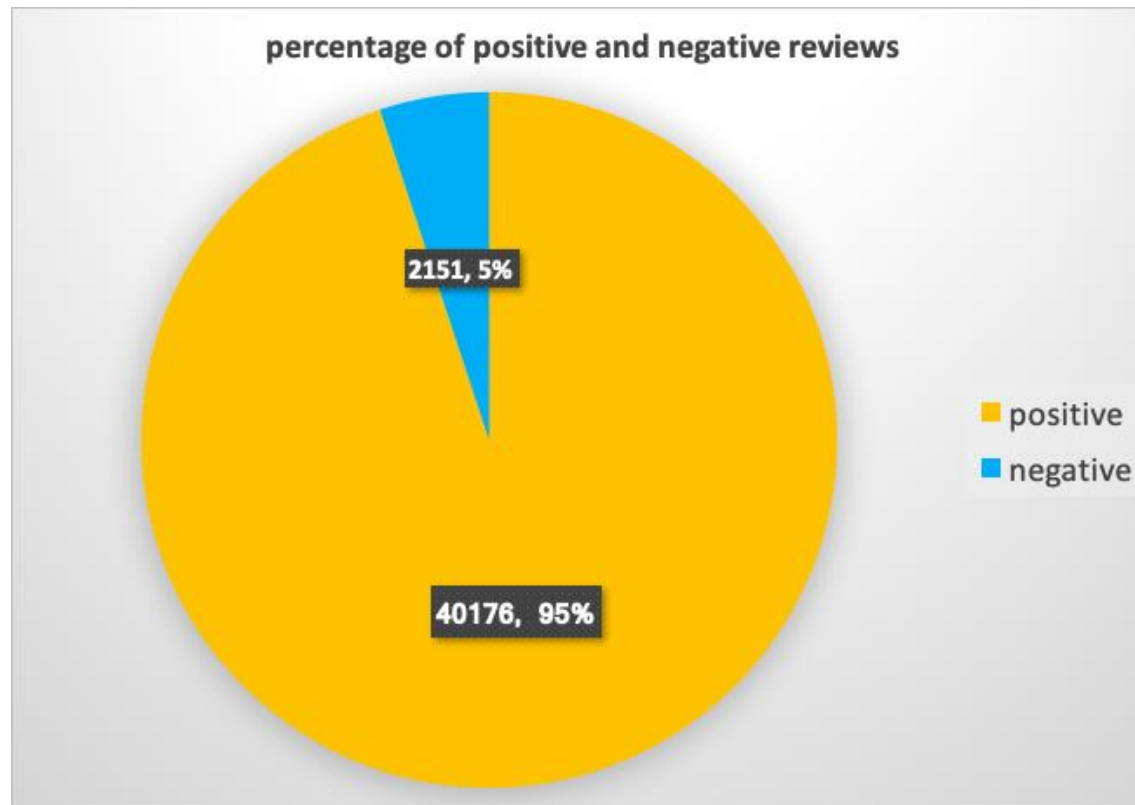
user_name	user_product	helpful	funny	post_date	recommend_or_not	game_time	review							
facciola007	304	4	0	2016/9/15	Recommended	235.2	amaze happy seem good far alpha release pre day review post release still a							
Dixie Norm	257	2	0	2016/9/15	Recommended	156.3	enjoy i've play far graphically look great bar minor visual glitch much expect t							
InquisitorTe	364	31	1	2016/9/15	Recommended	8.5	run well early access original sin good every way love first reason get							
Longmire	115	2	0	2016/9/15	Recommended	16.6	play hour far seem like great road far multiplayer pretty addictive visuals reall							
Fanest	0	6	2	2016/9/15	Recommended	212.8	play arena pvp tell source point op option play without would really nice cam							
newpain	38	2	0	2016/9/15	Recommended	104.3	probably one best ui character customisation ive see long time 60 minute alre							
Krazyhawai	582	0	0	2016/9/15	Recommended	71	amaze mechanic look great 10 10							
FastCrab	239	0	0	2016/9/15	Recommended	155.7	excellent successor already can't wait see next real complaint arena ai easy							
Biwya	101	0	0	2016/9/15	Recommended	24	see clear improvement form do amaze still challenge even alpha version							
JakeeL	174	0	0	2016/9/15	Recommended	599.6	back kickstarter say money well-spent o amazing sequel already awesome o							
Elethiomel	487	0	0	2016/9/15	Recommended	226.2	baldur's gate reborn steroid							
[TAW] Kaka	394	0	0	2016/9/15	Recommended	104.7	good stuff							
Gargall aml	96	0	0	2016/9/15	Recommended	179.9	discuss suicide big red lizzard first couple minute 10 10 would recommend su							
mbpopolar	420	1	2	2016/9/15	Recommended	1,000.10	early fully playable least 6-7 month let tell level refinement already unbelieval							
Podcast Int	196	0	0	2016/9/15	Recommended	50.7	kickstarer backer larian agian i'm impress alpha far great buy i'll add soon							
J Hawk	156	5	0	2016/9/15	Recommended	131.3	love							
Liqui	530	1	0	2016/9/15	Recommended	179.2	original sin improve every way							
KelseyGG	53	1	0	2016/9/15	Recommended	4.2	edit 10 sorry post help review board thought problem shoulnt problem graph							
ArenMook	174	16	0	2016/9/15	Recommended	390.6	enjoy first divinity original sin owe pick one larian take everything great first n							
Dirty Dan	252	1	0	2016/9/15	Recommended	17.6	really good early access still bit rough around edge dialogue bit wonky voice							
1Mr.Guru	252	1	4	2016/9/15	Recommended	35.1	run run run hole walk foot another hole oh look hole need hole hole wait car							
peaces4efn	42	8	5	2016/9/15	Recommended	1,382.60	like first one like beta though finish like first one taste game							
Azalus Dark	368	1	0	2016/9/15	Recommended	101.7	awesome sequel awesome wait forward full release							
Scominggre	204	1	0	2016/9/15	Recommended	102.9	ok early alpha extremely concern magic physical armor block cc damage stat							
Grenaderac	343	3	0	2016/9/15	Recommended	190.6	love far							
Azbolg	167	38	2	2016/9/15	Recommended	128.6	super basic atm voice dialogue like deluxe addition cool thing basically force							
Banjo Kabo	385	11	26	2016/9/15	Recommended	78.8	want kill thing want fire want lizzard want environment death frozen slow pois							



# Exploratory Data Analysis (EDA)

Analysis of reviews:

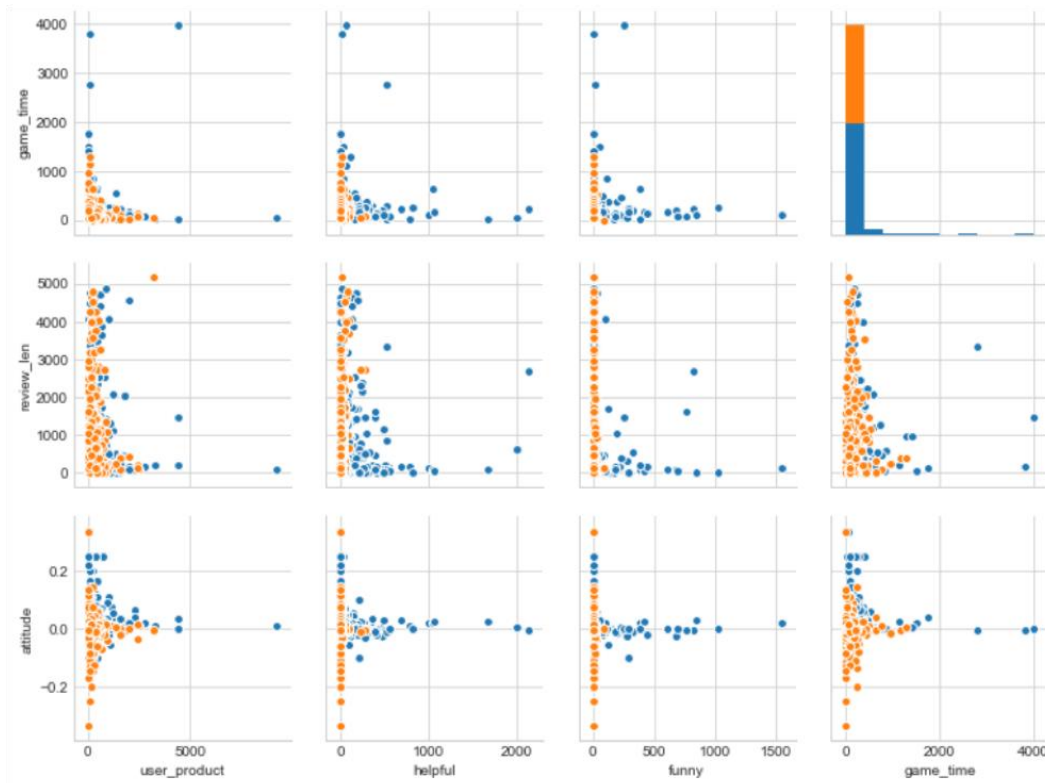
Almost 95% players gave 'recommend'. However, in the 'recommend' reviews, players gave not only praise, but also meaningful suggestions.



# Exploratory Data Analysis (EDA)

Try to find variable relations by scatter plot

- Reviews with neutral attitude get more 'funny' from other players
- Long reviews are not easy to get more 'funny'





# Natural Language Processing

# Natural Language Processing

## POS (Part of Speech) Tagging

For each token, we give it a Part of Speech tag.

```
[('initial', 'JJ'), ('introduction', 'NN'), ('introduction', 'NN'), ('lack', 'NN'), ('lack', 'NN'), ('goal', 'NN'), ('goal', 'NN'), ('unclear', 'JJ'), ('unclear', 'JJ'), ('unclear', 'JJ'), ('jump', 'NN'), ('dialogue', 'NN'), ('dialogue', 'NN'), ('compelling', 'NN'), ('compelling', 'NN'), ('plain', 'NN'), ('plain', 'NN'), ('text', 'NN'), ('text', 'NN'), ('be', 'NN'), ('narration', 'NN'), ('narration', 'NN'), ('tutorial', 'JJ'), ('tutorial', 'JJ'), ('really', 'RB'), ('really', 'RB'), ('explain', 'VB'), ('explain', 'VB'), ('combat', 'NN'), ('there', 'EX'), ('there', 'EX'), ('s', 'VBZ'), ('s', 'VBZ'), ('weak', 'JJ'), ('enemy', 'JJ'), ('enemy', 'JJ'), ('start', 'NN'), ('start', 'NN'), ('practice', 'NN'), ('combat', 'NN'), ('combat', 'NN'), ('new', 'JJ'), ('find', 'VBP'), ('find', 'VBP'), ('sufficient', 'JJ'), ('sufficient', 'JJ'), ('sca', 'NN'), ('help', 'NN'), ('learn', 'VB'), ('learn', 'VB'), ('experience', 'NN'), ('experience', 'JJ'), ('enough', 'JJ'), ('enough', 'JJ'), ('make', 'NN'), ('make', 'NN'), ('want', 'NN'), ('play', 'VB'), ('play', 'VB'), ('didnt', 'NN'), ('didnt', 'NN'), ('prst', 'RB'), ('interested', 'JJ'), ('interested', 'JJ'), ('second', 'JJ'), ('section', 'NN'), ('button', 'NN'), ('work', 'NN'), ('work', 'NN'), ('well', 'RB'), ('story', 'NN'), ('story', 'NN'), ('boring', 'JJ'), ('boring', 'JJ'), ('play', 'NN'), ('slow', 'JJ'), ('slow', 'JJ'), ('interact', 'NN'), ('interact', 'NN'), ('n', 'POS'), ('opinion', 'NN'), ('opinion', 'NN'), ('dont', 'NN'), ('dont', 'NN'), ('move', 'NN'), ('early', 'JJ'), ('early', 'JJ'), ('access', 'NN'), ('access',
```



# Natural Language Processing

## Bigrams

We use tagged tokens to get bigrams and we choose the bigrams that combined with an adjective and a noun.

[ 'initial introduction', 'suppose jump', 'tutorial battle', 'enemy start', 'new player', 'sufficient scaffold', 'und', 'boring play', 'slow interact', 'early access', 'implement i', 'current state', 'many bug', 'first divi', 'i', 'kill thing', 'good someday', 'euros thing', 'advanced beta', 'right spoil', 'final version', 'bad traile', 'ovie', 'early access', 'update version', 'reviewer participate', 'full review', 'com divinity', 'html note', 'te', 'clear detail', 'best see', 'armor quality', 'dull way', 'great attack', 'usage particle', 'pyrotechnic', 't bit', 'voice-acting ambient', 'decent nothing', 'special get', 'lush jungle', 'possible battle', 'choose we', 'avoid battle', 'wide selection', 'avoidable everything', 'replenish turn', 'certain ability', 'loot require', 'environmental status', 'frozen stun', 'multiple enemy', 'simple idea', 'give surprising', 'classic mode', 'd', 'y', 'certain enemy', 'enough power', 'allow farm', 'tough opponent', 'final bos', 'first act', 'classic hour', 'skill', 'enemy distance', 'final bos', 'nail plenty', 'instant death', 'certain dialog', 'new story', 'defini', 'ertain people', 'lizard others', 'specific tag', 'due history', 'provide case', 'good dialog', 'apparent deal', 'rfare', 'stat opponent', 'multiple way', 'playthrough time', 'reliant combat', 'dead voidwoken', 'due use', 'unknown protagonist', 'questionable authority', 'open npcs', 'unknown consequence', 'several side', 'quest r', 'avoid', 'overall struggle', 'cellar climb', 'lower someone', 'current character', 'current character', 'secon', 'obvious crash', 'early access', 'annoy inventory', 'auto-sort feature', 'stingy try', 'real direction', 'poi', 'ue inability', 'little time', 'weak impression', 'main story', 'ominous vibe', 'enough play', 'kiss stranger', 'om line', 'thin support', 'hefty price', 'original sin', 'larian studio', 'successful history', 'early access', 'iew', 'full release', 'great combat', 'decent macabre', 'great side', 'quest con', 'lite crash', 'lore music', 'cant figure', 'mechanic plenty', 'end pick', 'first place', 'certain defeat', 'mechanic build', 'able use',





# LDA

Split dataset into train (80%) and test sets (20%).

Set topic number to 4.

Print out top 20 words per topic.

```
tf_vectorizer = CountVectorizer(max_df=0.90, min_df=50, stop_words = stop_words)
tf = tf_vectorizer.fit_transform(text)
tf_feature_names = text

X_train, X_test = train_test_split(tf, test_size=0.2, random_state=0)

num_topics = 4
lda = LatentDirichletAllocation(n_components=num_topics, \
                                learning_method='online', \
                                max_iter=10, verbose=0, n_jobs=1, \
                                random_state=0).fit(X_train)

num_top_words = 20
for topic_idx, topic in enumerate(lda.components_):
    print("Topic %d:" % (topic_idx))
    words={tf_feature_names[i]:topic[i] for i in topic.argsort()[::-1][0:num_top_words]}
    print(words)
    print("\n")
```



# LDA

## Results:

Topic 0:

[('plenty side', 132.2828678816968), ('tough opponent', 126.82672750479561), ('reviewer participate', 106.001695727162), ('advanced beta', 100.95591763788345), ('much magic', 97.46107958322085), ('certain enemy', 69.97405748608946), ('dull i', 66.12593878881988), ('certain ability', 59.22440362815999), ('special get', 58.044298751392006), ('kill thing', 56.57469414989552), ('great attack', 51.723872230660255), ('initial introduction', 47.47876079402863), ('replenish turn', 46.367363769384134), ('particular skill', 45.879466331037634), ('first act', 45.708073974109645), ('good someday', 45.6195465080732), ('enough make', 45.02096915497249), ('environmental status', 41.69890757781668), ('classical mode', 41.34467258397585), ('multiple enemy', 41.020385102991106)]

Topic 1:

[('html note', 239.26711195393622), ('simple idea', 181.47224415462733), ('bad trailer', 136.42634777700994), ('allow farm', 131.8975289825392), ('decent nothing', 117.3196675230409), ('euros thing', 109.43647424118134), ('usage particle', 83.96747998270111), ('first minute', 71.50604605293807), ('due limitation', 71.01342487015154), ('fight part', 70.4033624581616), ('many bug', 65.69975771676764), ('sufficient scaffold', 64.98393131366339), ('armor quality', 60.942813153168025), ('frozen stun', 56.50884453796401), ('early access', 54.56961508213819), ('clear detail', 52.203472026170004), ('single quest', 51.20734761017036), ('avoid battle', 50.13230405102245), ('possible battle', 40.63340966179505), ('boring play', 0.25378092637572575)]

Topic 2:

[('final bos', 203.96747349110572), ('choose word', 136.09790932613055), ('new player', 115.78644515156189), ('com divinity', 71.92638351164469), ('suppose jump', 71.07544558533834), ('give surprising', 65.35396600361818), ('full review', 64.33016712064662), ('second refund', 63.69556250022164), ('dull way', 61.38347623729287), ('update version', 61.02563401992151), ('loot require', 60.634582170247256), ('enough power', 58.53594726429162), ('slow interact', 48.88189348107163), ('tutorial battle', 47.616229013881906), ('enemy distance', 44.92810877065587), ('quick note', 43.730077955049275), ('first divinity', 41.29239449300832), ('pyrotechnic surface', 0.25656165111520146), ('decent nothing', 0.25604814953990923), ('awesome movie', 0.2548147505725157)]

Topic 3:

[('early access', 162.99018073045903), ('awesome movie', 124.09591629129724), ('boring play', 92.62992693225662), ('implement i', 84.69463920380292), ('avoidable everything', 82.45492864512563), ('enemy start', 80.69932840145631), ('starter ability', 63.489347134788574), ('wide selection', 57.843071461046584), ('classic hour', 55.8330795122074), ('initiate possibility', 55.42475588257037), ('voice-acting ambient', 53.67893554764839), ('lush jungle', 52.489537987117004), ('ap future', 52.207395912233004), ('best see', 48.03225844338926), ('whole movie', 47.63022672195256), ('current state', 46.189269822971184), ('final version', 43.85414831639133), ('great bit', 42.476891935144366), ('right spoil', 41.209381481613214), ('pyrotechnic surface', 36.64505031532709)]

# Feature Extraction

## Original Sin I

### Positive Features

beautifully-designed world	interesting system	excellent experience	high-quality voice
developed system	useful npc	playable character	awesome experience

### Negative Features

endless dialogue	obnoxious partner	stupid game	tedious dialogue
big bug			

# Feature Extraction

## Original Sin II

### Positive Features

nice camera	real joy	huge improvement	good soundtrack
best rpg	wonderful world	awesome sequel	gorgeous art
worth investment	tactical combat	addictive visuals	gorgeous soundtrack
wonderful adventure	good game	awesome movie	

### Negative Features

tough opponent	slow interact	many bug	boring play
bad trailer			

# Topic Modeling

Different from previous part, here positive and negative data will be processed together, because both of them can express the same topic.

Next, I will compare reviews from the periods before the first 4 big patches after release. The time info of review sub-data sets are as following.

	Released date
Release	09/14/17
1st Patch	09/21/17
2nd Patch	10/06/17
3rd Patch	12/05/17
4nd Patch	02/01/18

Review	Starting Date	Ending Date
dataset 1	09/14/17	09/20/17
dataset 2	09/21/17	10/05/17
dataset 3	10/06/17	12/04/17
dataset 4	12/05/17	01/31/18



# Topic Modeling

Extract gameplay related keywords:

Total topic number = 10

It's common to see some players express same general ideas, like 'I like rpg very much', 'This is a really a good studio.'

This kind of general ideas are excluded through two ways:

- (1). Add them in stop words.
- (2). Ignore after generating word cloud

Word Cloud Sample



# Topic Modeling

## Evaluate the extracted features:

Keywords related gameplay in Graph 1 are 'story', 'content', 'issue' and 'quest', which are all updated in the next big patch.

	Bug/issue	Story	Combat System	Armor & Damage	Journal	Quest	Co-op Mode	Character
1st Patch	1	1				1	1	
2nd Patch	1	1	1	1			1	

(Note: '1' means the keywords mentioned in players reviews were updated in next patch.)



Graph 1: Keywords from data set 1



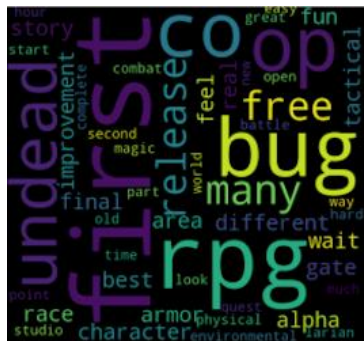
Graph 2: Keywords from data set 2



# Topic Modeling

In the graph 3 and 4, a new keyword is 'coop' was found, which meant many players hope new play system 'cooperation mode' can be released as soon as possible. In the next big patch, the new system 'cooperation mode' was released.

	Bug/issue	Story	Combat System	Armor & Damage	Journal	Quest	Co-op Mode	Character
3rd Patch	1	1	1	1	1	1		
4th Patch		1	1				1	



Graph 3: Keywords from data set 3



Graph 4: Keywords from data set 4



# Topic Modeling

## Insights from developer's perspective:

The result showed that developers are really caring about the feedback from players. If they take advantage of 'topic modeling', it would be more efficient.

After release, many players reviewed that they needed chinese version. However, it took developer half a year to update the Chinese version. During this period, many players who needed chinese version already gave negative review and left game.





# Topic Modeling

## Insights from player's perspective:

From the perspective of players, if you use 'Topic modeling' to check others' feedback, they can find that many players review the damage was not balance, so it's not their issue if they didn't win in battle, that's caused by unbalanced design, which should be fixed immediately.

In the next big patch, tons of balancing changes were updated by developer.



### Balancing changes

- Rebalanced trader prices in the Nameless Isle and Arx regions
- Break the Shackles can be cast while Silenced-- removes Silence and Plague in addition to other statuses
- Maddening Song now targets only enemy characters
- Demonic Stare damage and restoration increased by 50%
- Taunt now pierces Physical Armour, costs 2AP, has a cooldown of 5 turns
- Incarnate no longer has Taunt on spawn
- Breathing Bubble now provides Silence immunity
- Improved scaling of Attribute buff from Encourage. E.g. +2 to Primary attributes at L7.
- Oily Carapace now removes Slowed
- Turn to Oil radius increased to 5m
- Gag Order now costs 2AP, damage increased by 35%
- Smoke Cover cost decreased to 1AP, increased cloud radius to 3m
- Black Shroud and Blessed Smoke Cloud now have 4m radius
- Bleed Fire now decreases Fire Resistance by 20%
- Wind Up Toy's Suicide now costs 1 AP, deals 15% more damage
- Dimensional Bolt damage increased by 10%
- Mend Metal and Soothing Cold armour restored increased by a 35%
- Cryotherapy now clears Burning and Necrofire
- Shocking Touch damage increased by 10%
- Petrifying Touch now costs 1AP
- Meteor Shower damage reduced by 10%
- Corrupted Blade damage increased by 30%
- Corrosive Touch damage increased by 10%
- Vacuum Touch damage increased by 100%
- Ground Smash damage reduced by 35%
- Death Wish no longer deals damage to the target
- Rain Of Arrows damage reduced by 50% to match other 3SP skills
- Sabotage range increased to 13m. Only characters with grenades and arrows are highlighted by Sabotage





# Review Classification Models



## **Independent Variables:**

TF-IDF matrix created by the text of each review

Number of positive words and negative words in each review

Number of players think this review helpful

## **Label:**

Positive review or negative review

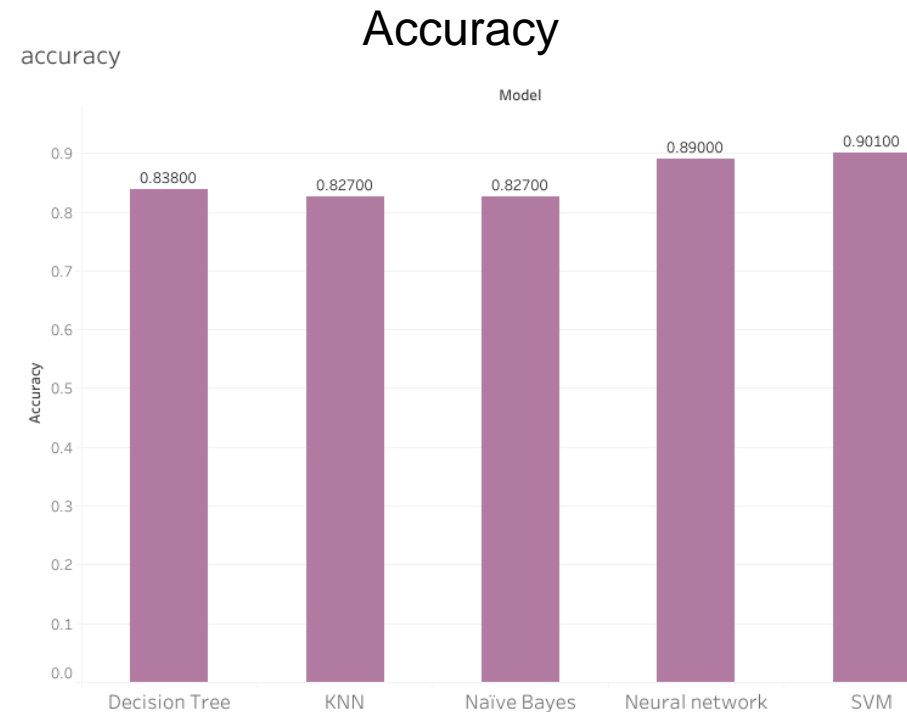
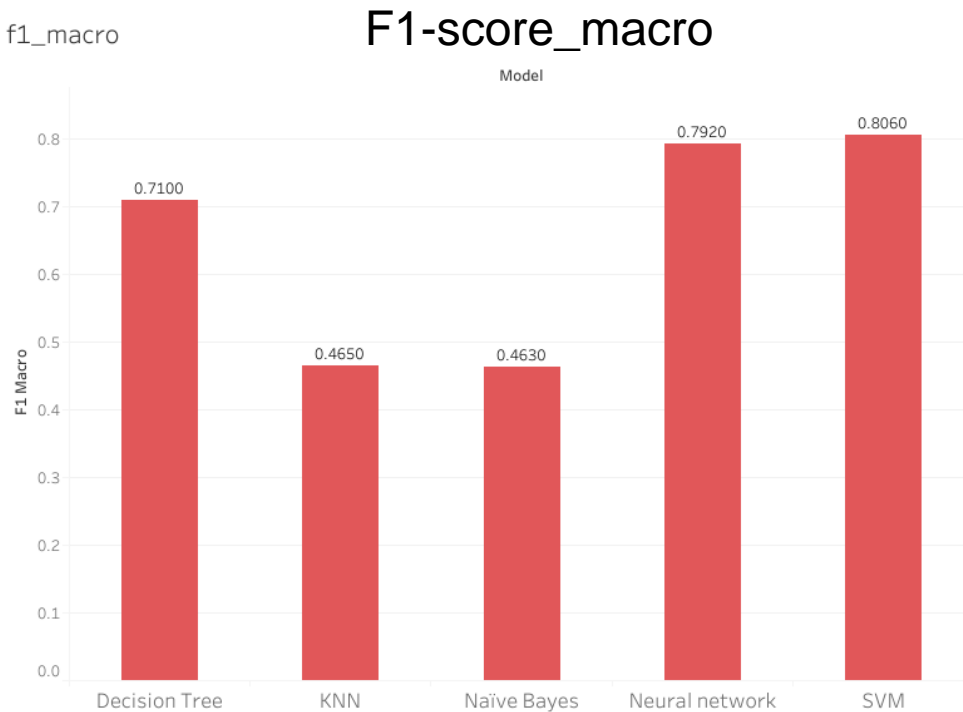
## **Method:**

Naive Bayes, SVM, KNN, Decision Tree, Neural Network

## **Evaluation:**

Macro F1 Score (precision, recall) and Accuracy

# Classification

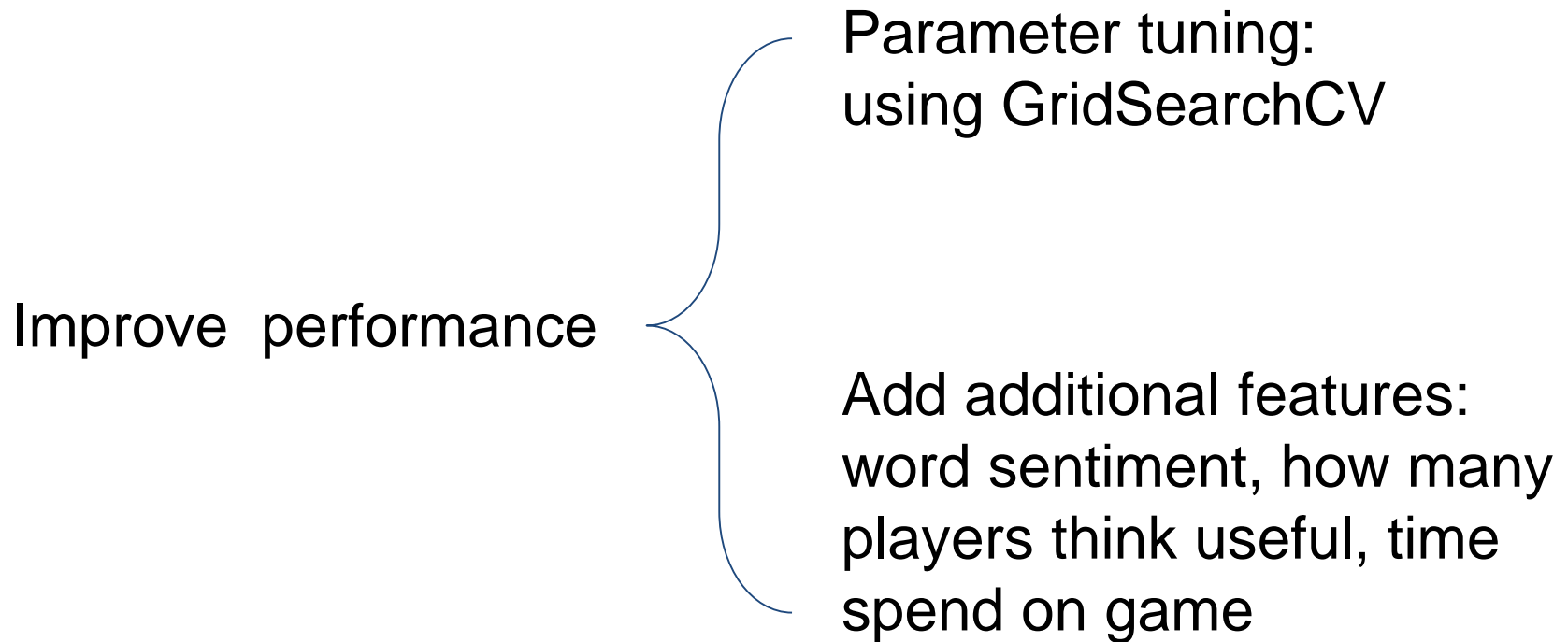


Best Performance: SVM, Neural Network

Worst Performance: Naive Bayes, KNN

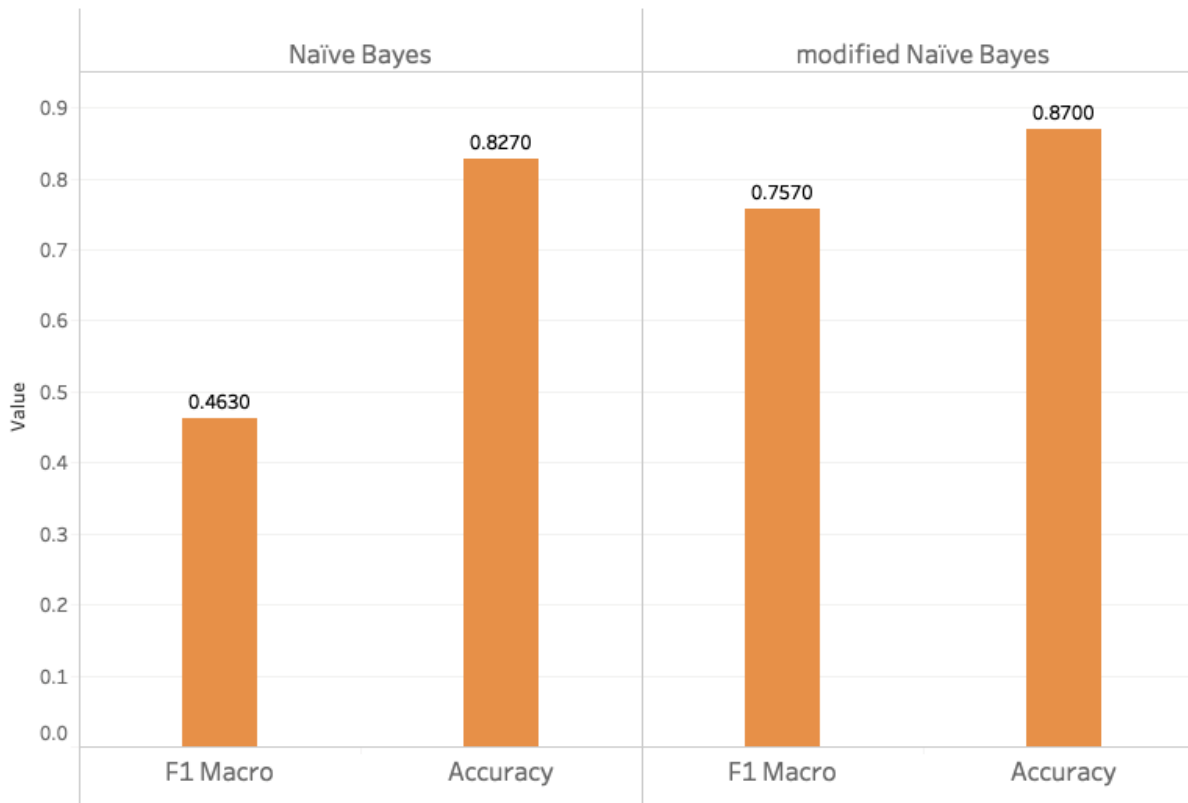


# Classification



# Modified Naive Bayes:

clf__alpha	tf-idf__min_df	tf-idf__stopwords	best f1_macro
0.5	18	English	0.749



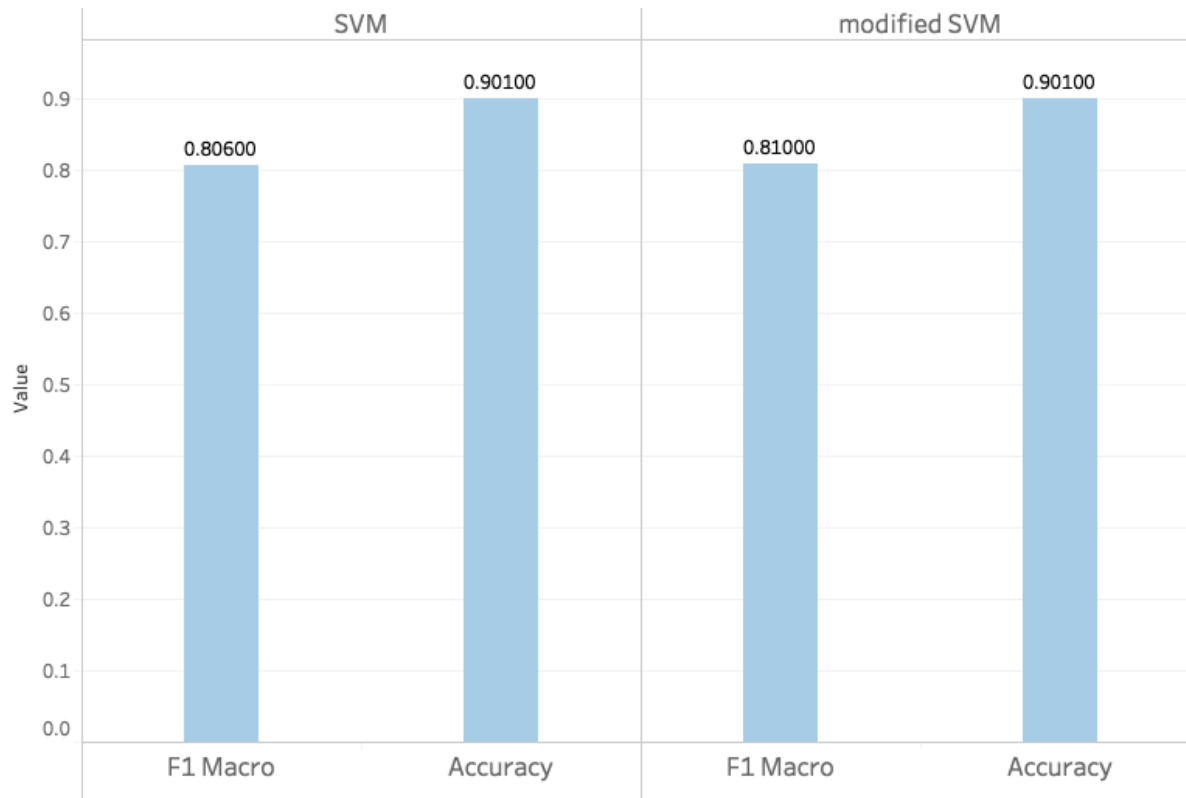
	cleaned_text	pos	neg
0	amaze happy seem good far alpha release pre da...	4	0
1	enjoy i have play far graphically look great b...	2	3
2	run well early access original sin good every ...	3	1
3	play hour far seem like great road far multipl...	7	0
4	play arena pvp tell source point op option pla...	2	2
5	probably one good ui character customisation i...	3	0
6	amaze mechanic look great 10 10	2	0
7	excellent successor already can not wait see n...	2	1
8	see clear improvement form do amaze still chal...	3	0
9	back kickstarter say money well spend o amazin...	23	10

Improved significantly



# Modified SVM:

kernel	clf__C	tf-idf__min_df	tf-idf_stopwords	best f1_macro
liner	1.5	2	English	0.812



	cleaned_text	pos	neg
0	amaze happy seem good far alpha release pre da...	4	0
1	enjoy i have play far graphically look great b...	2	3
2	run well early access original sin good every ...	3	1
3	play hour far seem like great road far multipl...	7	0
4	play arena pvp tell source point op option pla...	2	2
5	probably one good ui character customisation i...	3	0
6	amaze mechanic look great 10 10	2	0
7	excellent successor already can not wait see n...	2	1
8	see clear improvement form do amaze still chal...	3	0
9	back kickstarter say money well spend o amazin...	23	10

Not improved significantly while the original model is good enough

# Review Exceptions

	cleaned_text	pos	neg	label
1	enjoy i have play far graphically look great b...	2	3	Recommended
18	enjoy first divinity original sin owe pick one...	6	7	Recommended

"Enjoyed what I've played so far, graphically looks great (barring some minor visual glitches) and is much of what I expected in terms of combat and roleplaying elements.

Had some issues with the camera and some textures being a bit wonky on player models, besides that very few complaints. Can't wait for the full release!"

positive word: ['enjoy', 'great']

negative word: ['glitch', 'issue', 'complaint']

***Number of negative word is larger than positive word while its label is 'positive'***

"If you've enjoyed the first Divinity: Original Sin, you owe it to yourself to pick up this one.

Larian took everything that was great about the first game and made it even better. Literally a minute into the game you can either recruit or kill your first companion, simply depending on how you talk to them.

Stealing mechanics changed -- both guards and whoever owns the item will actively seek out the thief (you), and should they find you red-handed, throw you in prison, attack or simply take the stolen merchandise, again depending on how you talk your way out of the situation.

Fortunately even if you choose the violent option, the game is still perfectly playable.

Merchants and civilians don't participate in the fight.

I'd write more, but I think I'll just go ahead and play some more instead."

positive word: ['enjoy', 'great', 'well', 'easy', 'fortunately', 'perfectly']

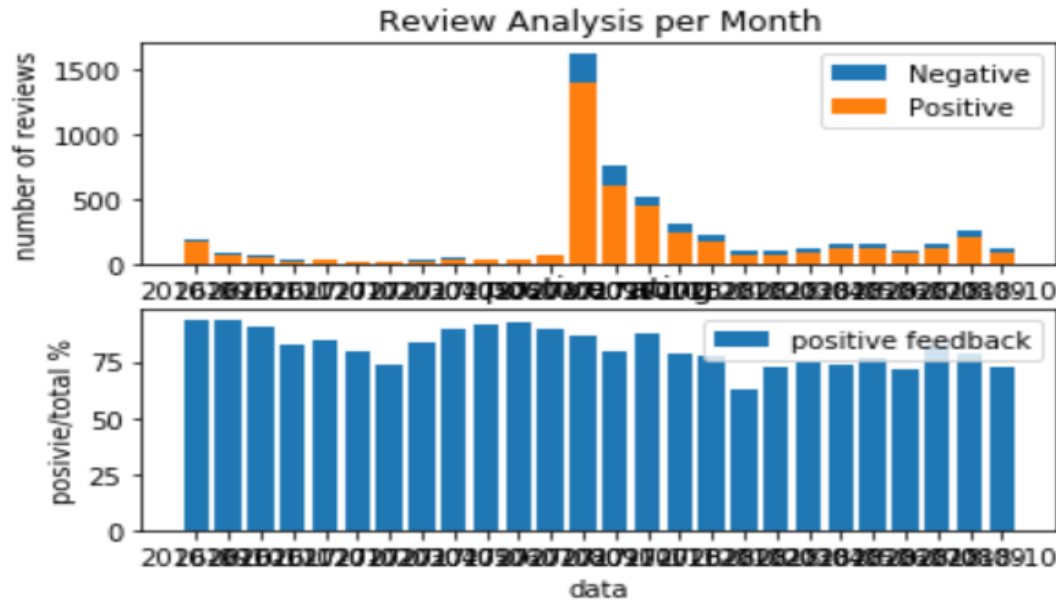
negative word: ['sin', 'kill', 'steal', 'prison', 'attack', 'steal', 'violent']

***Some negative word may not have negative meaning in RPG game.***

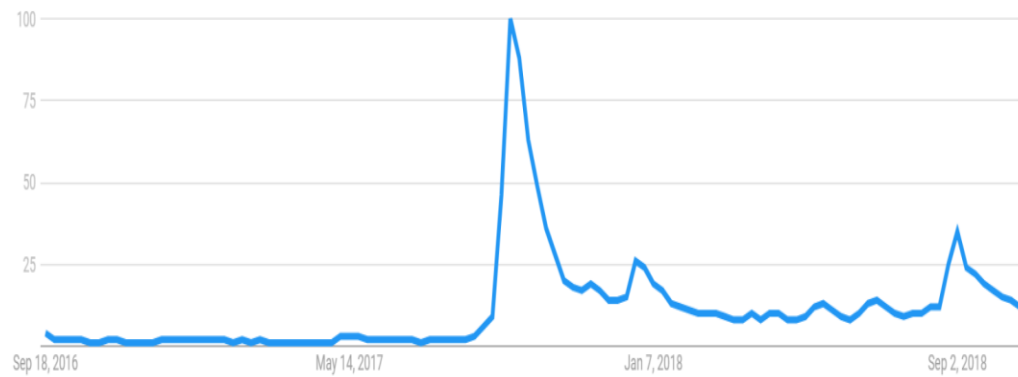


# Prevailing Trend Forecasting

## General Prevailing Trend



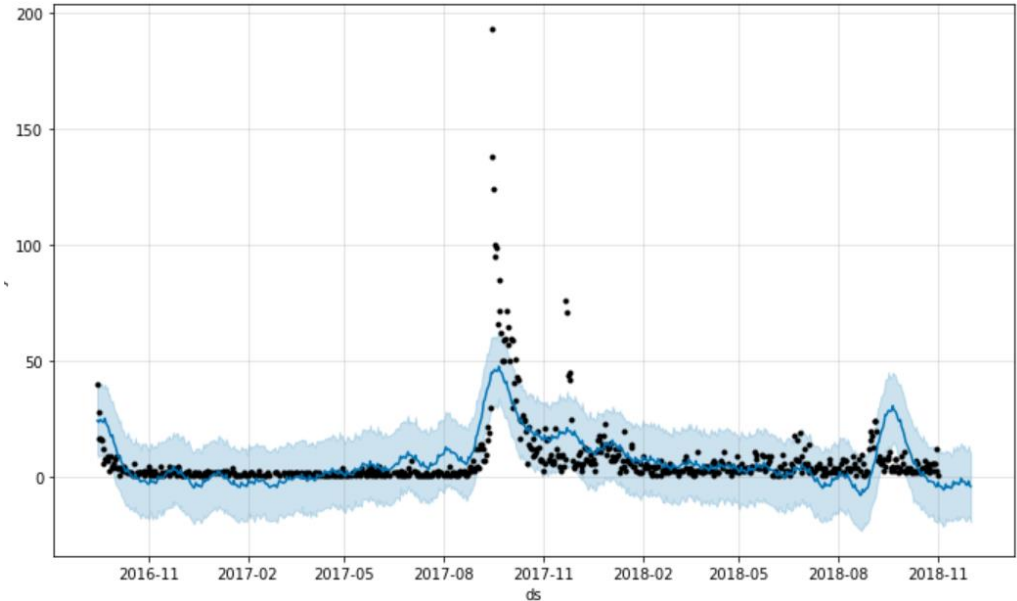
Average Positive Rating:  
**81.69031102811851%**



# Google Web Search Trend

# Trend Prediction

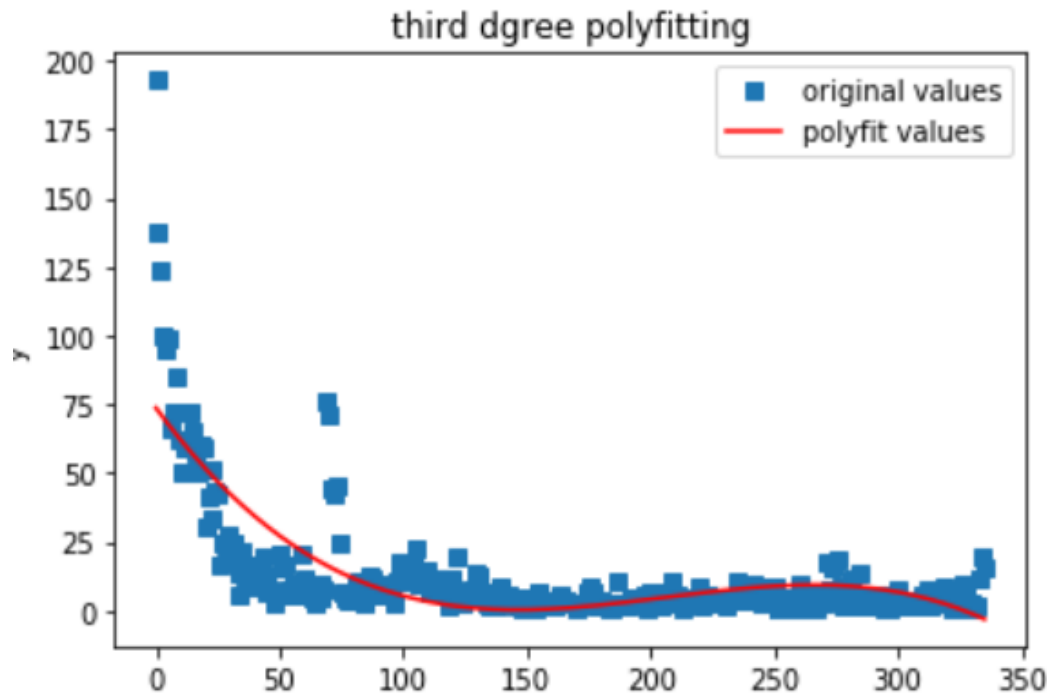
- Polynomial Fitting
- Exponential Curve
- Time-series Model



We took reviews from 09/14/2017 (release date) to 08/31/2018 to fit the forecasting model.

The reviews from 09/01/2018 to 10/31/2018 is for verification on prediction

# ● Polynomial Fitting



similarity score is :

163.62900364951395

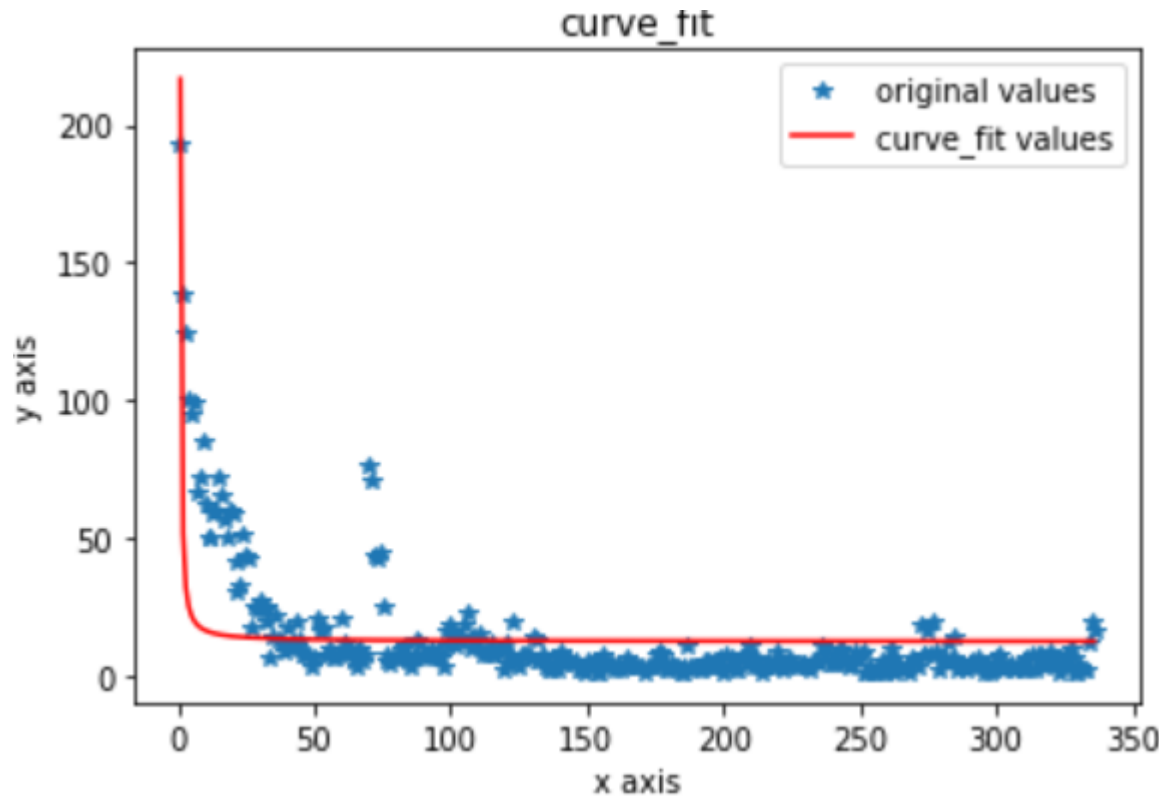
Polynomial Curve Fitting

$x^3 + x^2$

$-1.054e-05 x + 0.006519 x - 1.23 x + 73.64$

$$similarity.score = \frac{\sum ||actual.value - predicted.value||}{sample.size}$$

# ● Exponential Curve



similarity score is :

7.792709825757792

parameter a: 12.30452657404054

parameter b: 2.869100023130686

$$similarity.score = \frac{\sum ||actual.value - predicted.value||}{sample.size}$$





# ● Time-series Forecasting

## What makes Time Series Special?

As the name suggests, TS is a collection of data points collected at **constant time intervals**. These are analyzed to determine the long term trend so as to forecast the future or perform some other form of analysis. But what makes a TS different from say a regular regression problem? There are 2 things:

1. It is **time dependent**. So the basic assumption of a linear regression model that the observations are independent doesn't hold in this case.
2. Along with an increasing or decreasing trend, most TS have some form of **seasonality trends**, i.e. variations specific to a particular time frame. For example, if you see the sales of a woolen jacket over time, you will invariably find higher sales in winter seasons.

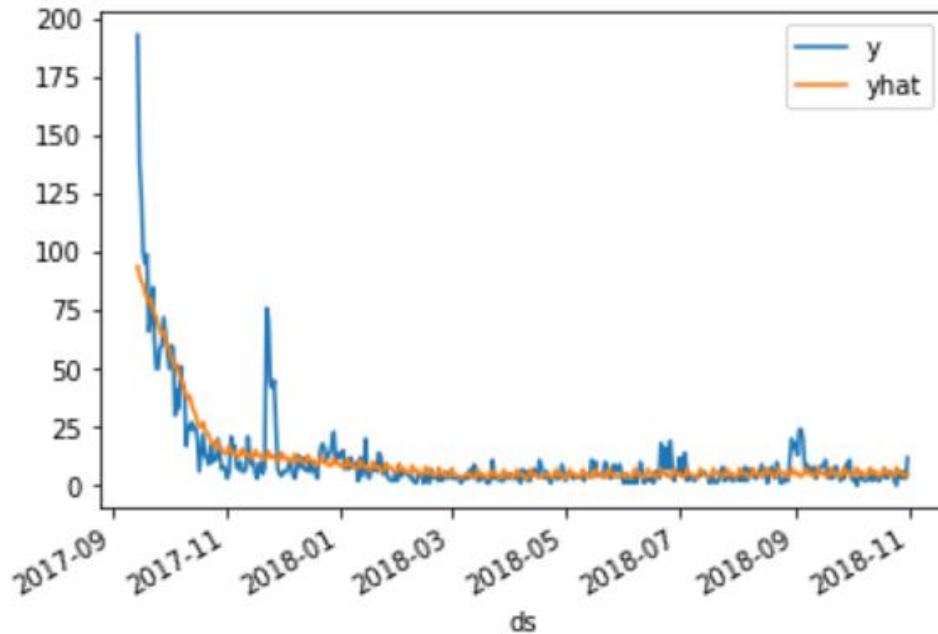
# ● fbprophet

1. Prophet follows the **sklearn** model API. We create an instance of the **Prophet** class and then call its **fit** and **predict** methods.
2. The input to Prophet is always a dataframe with two columns: **ds** and **y**. The **ds** (datestamp) column should be of a format expected by Pandas, ideally YYYY-MM-DD for a date or YYYY-MM-DD HH:MM:SS for a timestamp. The **y** column must be numeric, and represents the measurement we wish to forecast.

	DS	Y
<b>0</b>	2007-12-10	9.590761
<b>1</b>	2007-12-11	8.519590
<b>2</b>	2007-12-12	8.183677
<b>3</b>	2007-12-13	8.072467
<b>4</b>	2007-12-14	7.893572

	DS	YHAT	YHAT_LOWER	YHAT_UPPER
<b>3265</b>	2017-01-15	8.199274	7.489884	8.969065
<b>3266</b>	2017-01-16	8.524244	7.790682	9.266504
<b>3267</b>	2017-01-17	8.311615	7.553025	9.049803
<b>3268</b>	2017-01-18	8.144232	7.428174	8.864747
<b>3269</b>	2017-01-19	8.156091	7.395160	8.883232

# ● Time-series Forecasting - fbprophet



Time series analysis is a **statistical technique** that deals with time series data, or trend analysis.

similarity score is :

4.87978148756082

ds	y	yhat	yhat_lower	yhat_upper
2017-09-14	193.0	93.818660	81.106089	107.866684
2017-09-15	138.0	89.916886	75.988997	102.689786
2017-09-16	124.0	87.409611	75.444152	99.631395
2017-09-17	100.0	86.075322	74.123834	98.353564
2017-09-18	95.0	82.512991	68.904490	96.118781



# Conclusion & Future Work



# Conclusions

- The game **Original Sin** are really caring about players' feedback and the developer update new content based on players' reviews.
- Game company can update the game according to reviews' features, like bug fixes and balancing changes.



# Further Work

- Scrape more data for comprehensive analysis
- Extract more meaningful keywords and features
- Increase f1 score and accuracy of our model
- Increase accuracy of Time Series Analysis