SENTIMENT ANALYSIS REPORT

Duration of the project: 6 Months

Project Mentors: Shalaka, Sadhana

Team Members: Tarush Shankar, Amogh Umesh, Hari Chetan Krapati

Introduction:

The aim of Sentiment Analysis is to analyze the subjective information in the text and mine the general opinion. Information is extracted from the opinions, appraisals, and emotions of people in regards to entities, events, and their attributes. It is the process of classifying text as either positive, negative, or neutral. Machine learning techniques are used to evaluate a piece of text and determine the sentiment behind it.

For the project, we have focused on 3 games - Gwent, The Witcher 3 and Cyberpunk 2077, developed by CD Projekt RED. We have chosen a timeline of around 2 months from the date of release of each game to analyze tweets and determine the general sentiment, whether the game has been spoken about favorably or unfavorably.

Method:

PHASE 1 :

Learning the basics:

We first start with phase 1, where we spent time learning all the topics that will be needed for making the model and pre-processing and handling the data. We learn NLP, NumPy, pandas, NLTK etc. to be able to handle the programs we have to write in the future. We also learn the basics of ML and the way an ML model is built, and what steps are needed to ensure a proper model.

PHASE 2 :

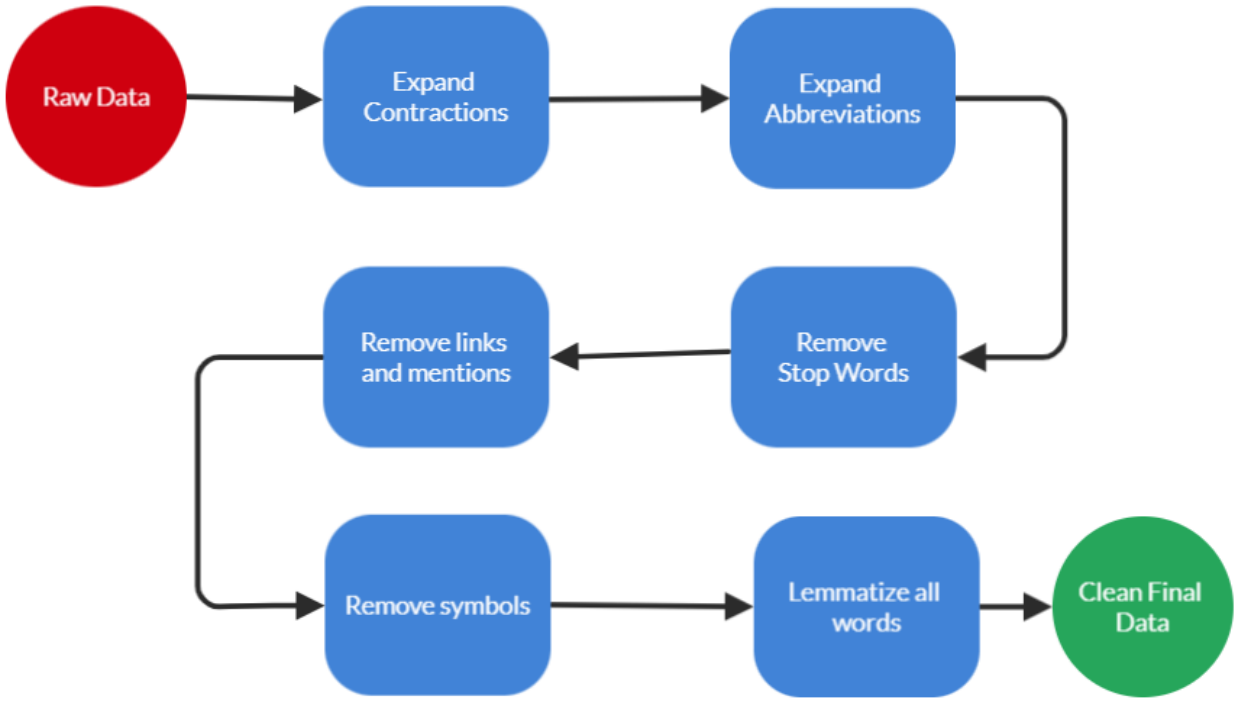
Data:

In the next phase, we first needed to decide the subject of our sentiment analysis, and we have decided on the company CD Projekt Red, and primarily focus on 3 games, that is , The Witcher 3, Gwent, and Cyberpunk 2077, and we would analyse the impact on the company by the sentiment of these three games. We now needed to get the tweets for all three of these games. We decided to measure the sentiment from 10 days before the release to 2 months after the release of each game. This would give us a period when the game is a hot topic, hence giving us more data and a

good impact point for the company as well. We used a twitter API to fetch the tweets based on our search. We obtained thousands of tweets for each game, and now we had to prepare this data.

Pre-processing:

We used multiple steps to pre-process and clean the data to clean it up and make it ready for the predictions. Also, we knew the pre-processing program would be useful to clean and prepare the training data.

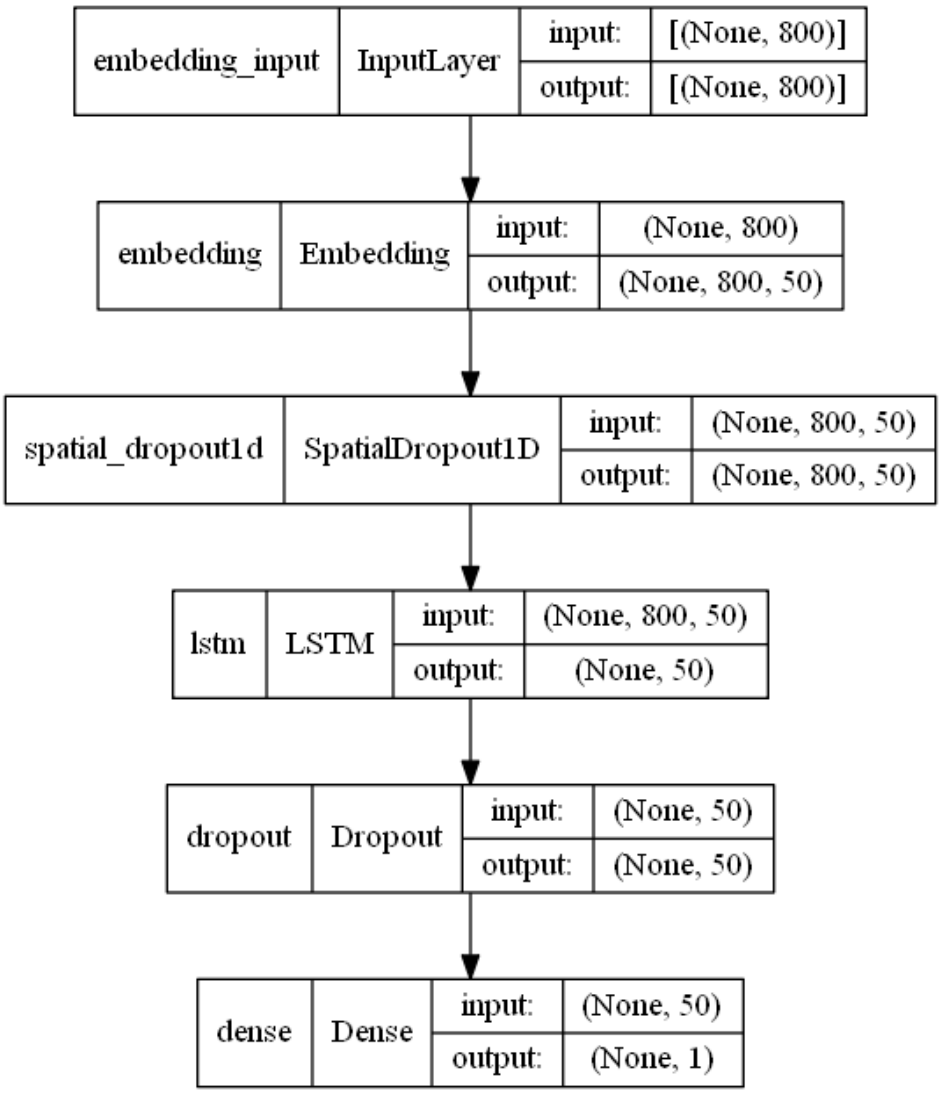


Hence, using all these steps, we not only prepare the tweets for the final sentiment prediction, but we also have a program that can pre-process the training data when we get to that stage.

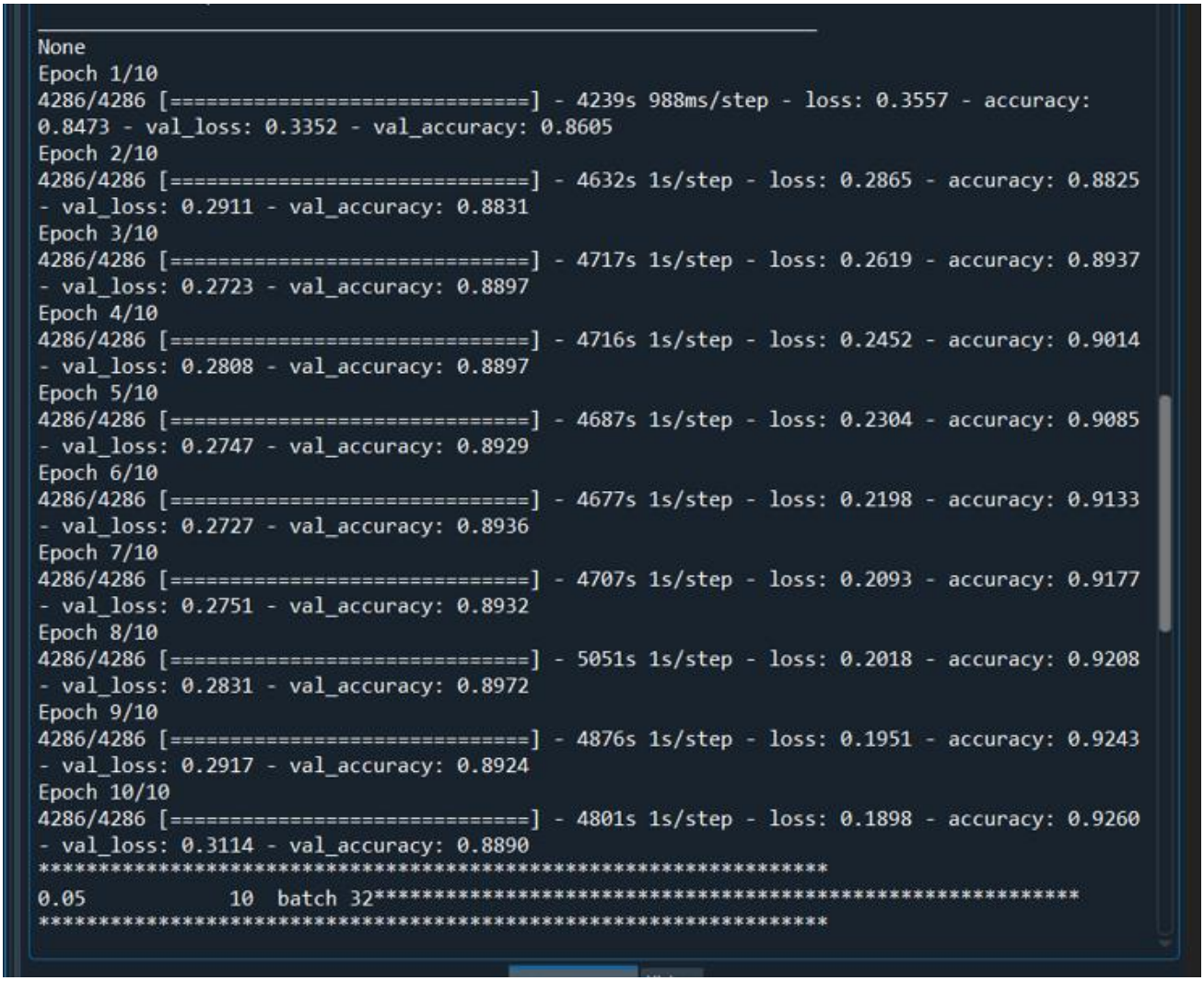
PHASE 3:

Making the model:

Arguably the most important step of the sentiment analysis project, we needed a high accuracy model for NLP to be able to get a reliable sentiment prediction on the tweets. We first tested out various methods to check their accuracies. We took a lot of guidance from the mentors and also the internet to be able to come up with sensible models. At the end of the day, we realised soon that a neural network model was the only one that gave us reliable outcomes, and we used the keras library from TensorFlow in python to get a good NLP model. We also came to the conclusion that LSTMs are good for sentiment analysis, as they use context and remember key points and may use them to predict the sentiment more accurately. We hence tried a couple of configurations, again taking the guidance of the mentors and the internet, of an LSTM model and finally settled for this architecture.

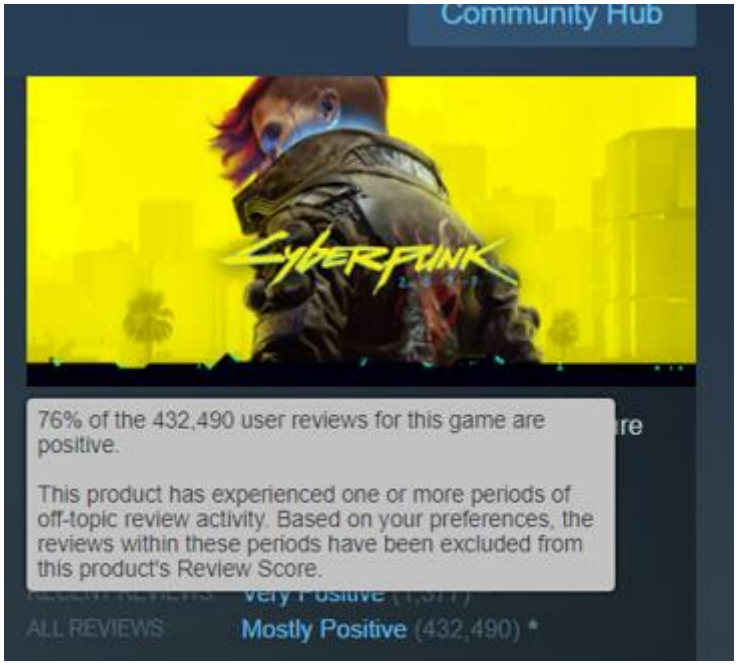


The model has an embedding layer first, to reduce the dimensionality of the tokenized sentences, then we use a dropout to add some variance, and we have the core LSTM that does most of the heavy lifting by memorizing the context and by trying to understand the context better, and finally we have a dropout again to prevent overfitting and a dense layer to compile all of this to a final result. We trained the model for 10 epochs on an example of steam game data we had, and we got around 89-90% accuracy from this model, which seemed satisfactory.



Getting the training data:

We needed good training data to be able to get a good trained model and proper context. Therefore, we looked to steam for the training data, as it had many reviews, and all of them were classified as positive or negative based on the reviewer's choice on whether they liked the game or disliked it. This was perfect, because we were using data created by reviewers, i.e, people to predict the sentiment of the tweets. We used a basic internet scraping program to fetch the reviews, and cyberpunk 2077, which had 25% of the reviews by players who disliked the game, and also had a large number of reviews, was perfect, as we would get a large amount of training data for both positive and negative reviews, making the model more reliable to predict both positive and negative sentiment.



We used the pre-processing program that we already had to clean up the reviews, and trained the model based on this data, which gave us over 90% accuracy. We used this model to predict the sentiment on all three games, and hence our work on getting the twitter sentiment data is done.

We also get a confusion matrix, with the difference in accuracy between the positive and negative sentiments seemed to be only 5%, which was tolerable.

```
0.9977899
Confusion matrix :
 [[94606  5391]
 [ 4607 39742]]
Outcome values :
 94606 5391 4607 39742
Classification report :
              precision    recall  f1-score   support

           1       0.95      0.95      0.95     99997
           0       0.88      0.90      0.89     44349

    accuracy                           0.93    144346
   macro avg       0.92      0.92      0.92    144346
weighted avg       0.93      0.93      0.93    144346


In [7]:
```

PHASE 4:

Product and sentiment analysis:

Here, we go through multiple articles to understand each product, i.e, each game better. We have documented the game's functioning, cost, budget to make, sales, aesthetics, etc. and we also add the analysis of the twitter sentiment, to see why there was more positive or negative discussion of the game, and also, we see the impact of the release of the game on the stock of the company, and interrelate that with the positive or negative sentiment of the game. All of this gives us a good understanding of the relative sentiment of each individual game, and how it has impacted the company, CD Projekt Red.

We also did cluster analysis for cyberpunk 2077, where, we extracted clusters using KMeans clustering for positive and negative sentiment tweets individually. This allowed us to get the main positive points that people talked more about, and the main negative points that people were most discontent with. This allowed us to find exact information on the top reasons for the positive or negative sentiment, and hence, this information can be used in a product to be able to work and fix the negative points that people were discontent with, and also advertise the positive points to more people, as it seems to be the most appealing points of the product.

Results:

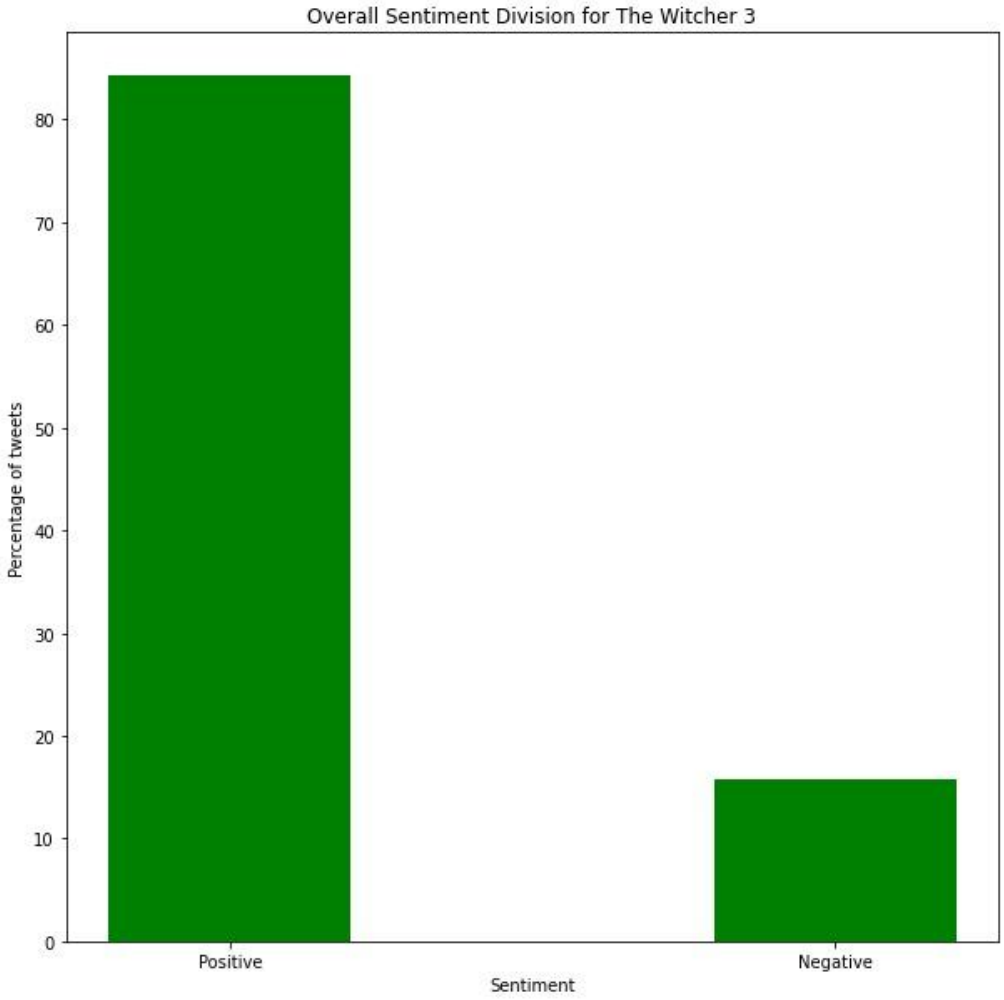
We now take a closer look at the three games and the impact that they had on the company.

The Witcher 3: Wild Hunt

The Witcher 3: Wild Hunt is an action role-playing game developed by Polish developer CD Projekt Red. It is played in an open world with a third-person perspective. The game was localized in 15 languages, with a total of 500 voice actors and was created with the REDengine 3, CD Projekt Red's proprietary game engine designed for nonlinear role-playing video games set in open world environments

Public Reception and Sentiment on release:

The Witcher 3 had a largely positive response from the community upon its release. Based on our model's predictions and the tweets gathered for the first 2 months after its release, the game had an around 85% positive sentiment.

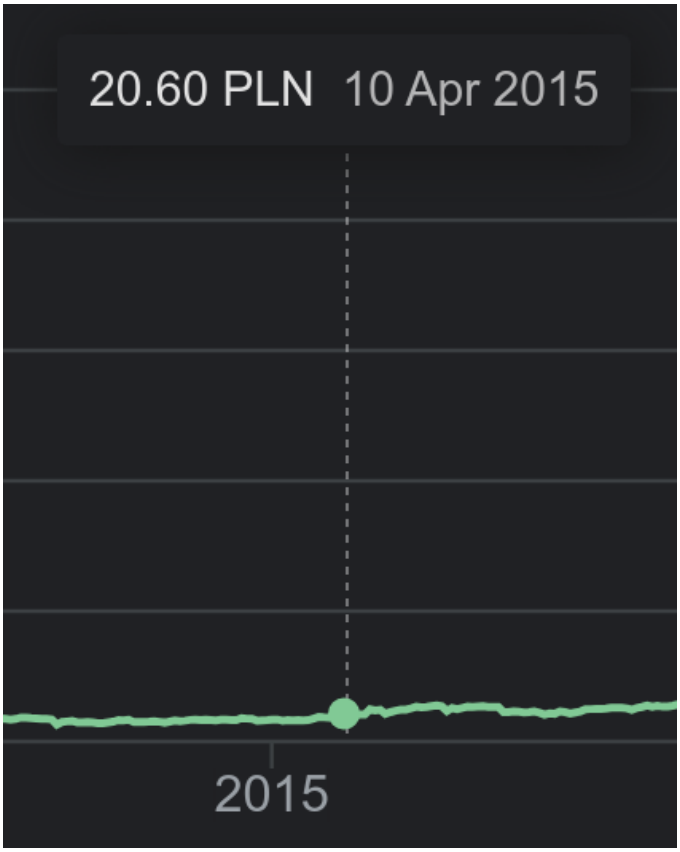


As seen, Witcher 3 has been a massive hit because it has an extraordinary story-line along with beautiful graphics which makes it so unique and interesting but at the same time it requires a lot of cpu power. Also, it has a lot of interesting side quests which keeps you entangled with the game even when you are not completing the main objectives.

Impact of Witcher 3 Release on CD Projekt RED stock prices:

Upon release, CD Projekt RED stock prices did not witness huge fluctuations, though it steadily kept on an upward trend. Between 1 month before the game launch and 2 months after, it's stock prices increased by around 20%.

Note PLN is Polish currency, Poland Zloty.



Another interesting point that can be noted is the drop in stock prices during the latter half of 2015, when CD Projekt RED released a huge 15 GB update with fixes, improvements and its first ever expansion, Hearts of Stone on October 8th.

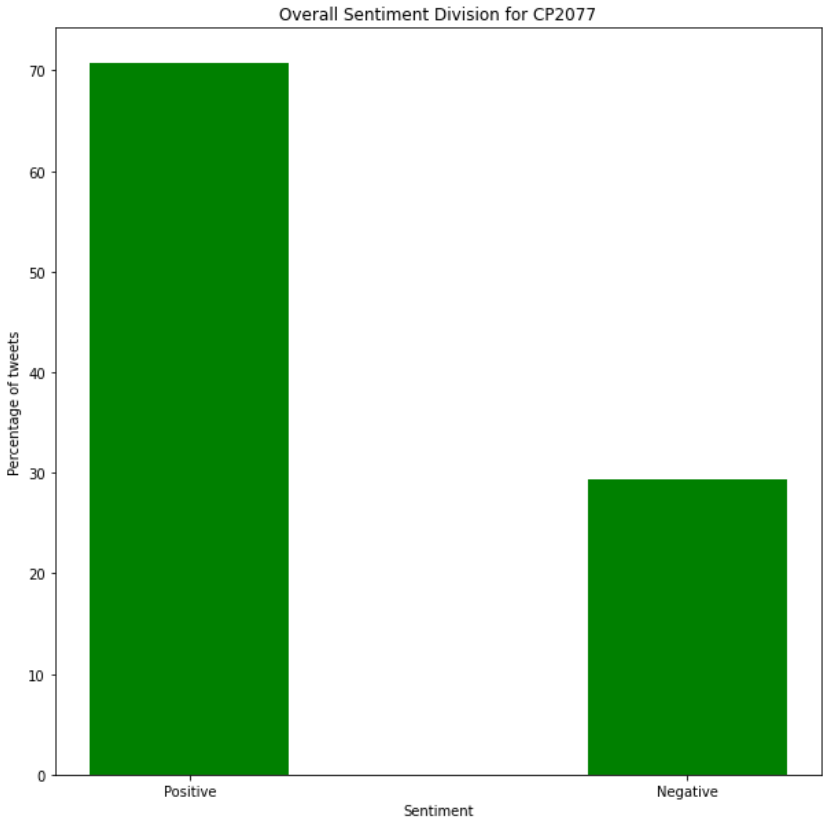
Hearts of Stone was a disappointment for most of the community since there wasn't anything particularly different from the enormous amount of stuff to do that already existed in the main game, aside from a new artisan and crafting options.

Cyberpunk 2077

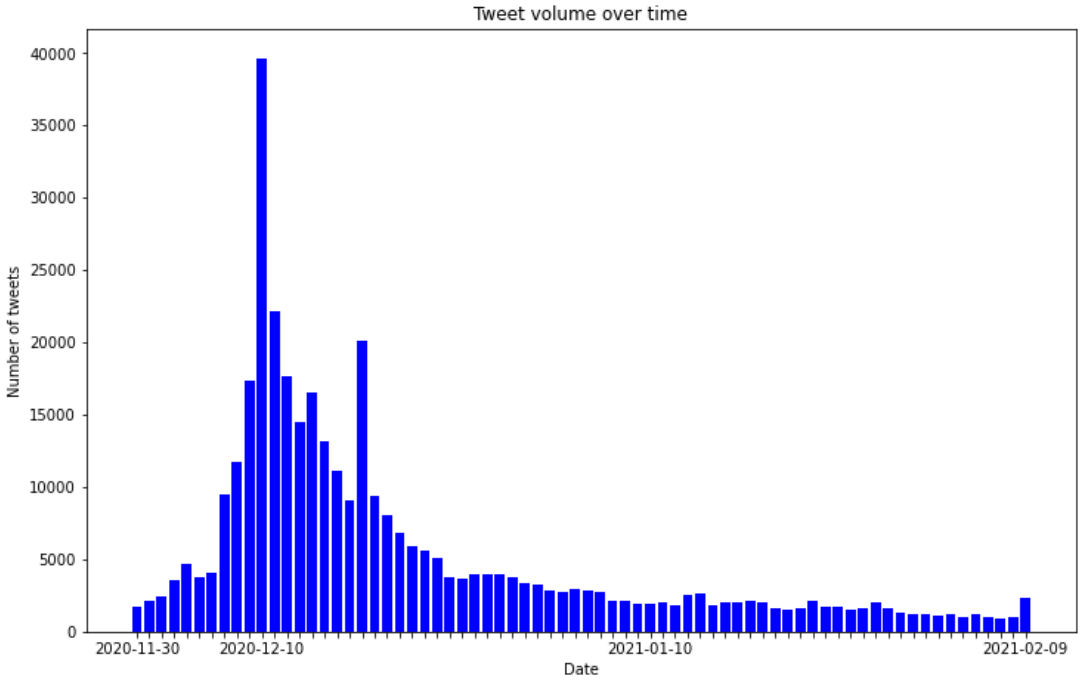
Cyberpunk 2077 is an action role-playing video game developed and published by CD Projekt. The story takes place in Night City, an open world set in the Cyberpunk universe. Players assume the first-person perspective of a customizable mercenary known as V, who can acquire skills in hacking and machinery with options for melee and ranged combat.

Sentiment analysis:

We use our final model to predict the sentiment on the tweets about the three games. Here, we take a look at the results obtained for the sentiment analysis of the tweets related to Cyberpunk 2077 over a period of 10 days before release to 2 months after release.



Here, the overall sentiment division gives us over 70% of tweets are positive, and around 30% being negative. This coincides with the reviews present on steam as well, where we have 75% positive reviews. Even though the percentage of positive tweets is pretty high, we can clearly see there is noticeable discontent regarding the game, based on the fact that on steam, a decent doing game can have anywhere between 85-95% positive reviews. We have already seen the cause of this noticeable discontent being the game's bad and buggy launch, and many players' disappointment over the incomplete state of the game not being what was promised and marketed by CDPR. The 70% positive tweets also tell us that the players are discontent, but do not hate the game.

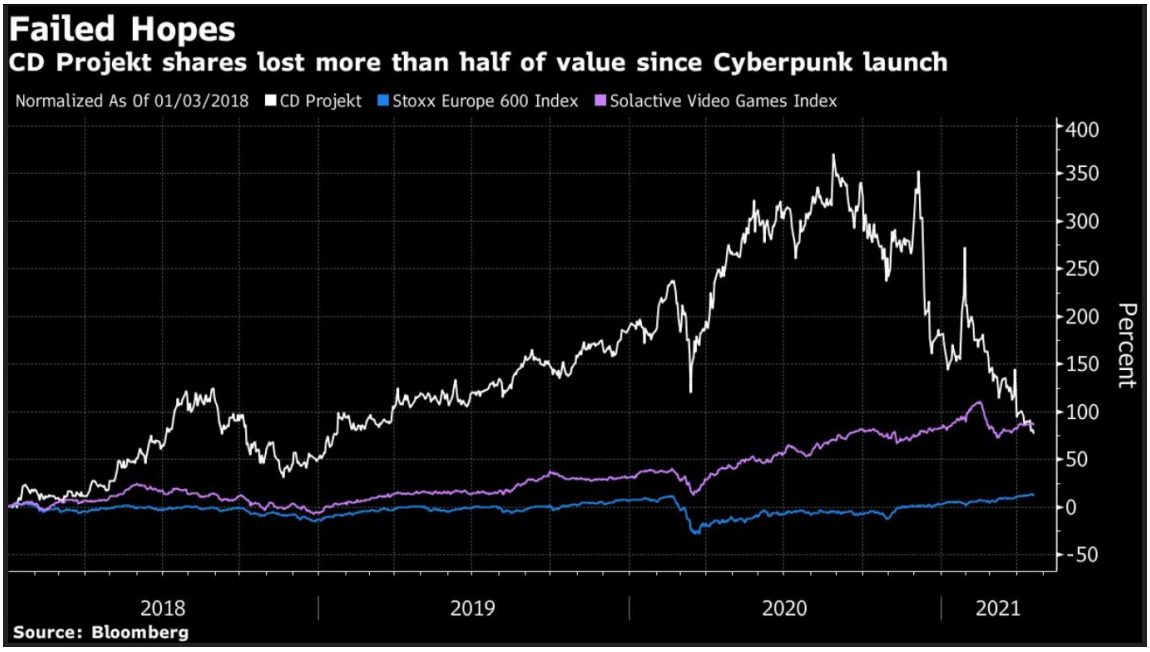


The tweet volume over time shows us the massive spike on release day, with nearly 40,000 tweets on December 10th about the game. Even a few days before release, the game was very actively discussed, showing the excitement of players over the release of the game. The two months after release do not paint the same picture. As with the case with any game after its new release, the discussions about it will die down, but the impact of the messy launch had reduced the active players heavily. We can see a similar trend even in the monthly active steam players we had analyzed in the "How long will cyberpunk be in the market" section of the product analysis.

Impact of Cyberpunk 2077 on the stock price of CDPR:

A few months before the release of Cyberpunk 2077, CD Projekt SA was among Europe's hottest stocks, riding the wave of pandemic gaming demand and on the way to a record high ahead of the release of the studio's much-anticipated Cyberpunk 2077 game. Fast forward 12 months, the stock is the region's worst performer of 2021, down 63% from its August record, with the company facing questions about how it plans to fix problems with the botched release of the bug-ridden game.

Cyberpunk 2077 is a complex futuristic role-playing game that had been predicted to take the gaming world by storm when it was launched on Dec. 10. But it wasn't long before players were reporting multiple glitches, causing negative reviews and triggering Sony Corp.'s unprecedented decision to remove the title from its PlayStation store. A strategy update at the end of march 2021 failed to address concern over how the Polish company will affect a turnaround, showing only that Cyberpunk wasn't yet ready for the full multi-player version that could restore its popularity. It also left questions unanswered over when CD Projekt will be ready with an updated version for next-generation consoles.



The discontent we noticed during the sentiment analysis and also the massive fall of the tweet volume shows that the release not only disappointed many fans, but also, the topic of discussion quickly moved on. Since before the release, the game was very heavily hyped and discussed about, the stocks soared to CDR's all time highest, but the discontent and also the company losing the status of being the "Hot topic" had heavily impacted the stock, which had a downward fall until 4 months after the release, where it settled down.

From after April 15th, stock price has remained more or less the same, with not much change. CDPR has since given multiple patches and updates to fix the bugs and have tried to redeem themselves, but the impact of the bug-ridden launch still remains. We can also notice another small dip before the release of the game, which was when the game release had been postponed.

Clustering and analysis of key points for the game:

After we do the sentiment analysis on the tweets we can cluster these tweets to understand the main reasons for people's discontent. This gives us a better insight at the main causes of the hate, allowing us to focus on the main points.

For example, when we divide the negative tweets into 10 main clusters, we get the top features from each cluster as:

Top terms per cluster:

Cluster 0:

Months, game, cyberpunk, wait, delay, ago, play, like, come, release

Analysis: game got delayed by a few months after years of waiting for the release

Cluster 1:

Bad, cyberpunk, game, really, feel, play, like, want, know, think

Analysis: Generally saying the game is bad

Cluster 2:

Refund, cyberpunk, request, Microsoft, offer, playstation, store, purchase, process, sony

Analysis: people asking for refunds from Microsoft and playstation store because the console release had it the worst

Cluster 3:

Play, stream, cyberpunk, game, twitch, want, crash, minutes, time, like

Analysis: People experienced and also watched streams of the game crashing every few minutes

Cluster 4:

Store, playstation, sony, offer, remove, refund, customers, pull, cyberpunk, game

Analysis: sony removed the game from the playstation store because too many players were asking for refunds

Cluster 5:

Release, delay, cyberpunk, game, time, buggy, years, cd, project, like

Analysis: the game is buggy even after being worked on for years and getting delayed as well

Cluster 6:

Sell, copy, million, overpromise, underdeliver, rig, cyberpunk, promise, despite, refund

Analysis: the marketing hyped the game too much and the game underdelivered on all the promises made in the years before release

Cluster 7:

Cyberpunk, cd, project, game, like, red, make, fuck, buy, save

Analysis: could be saying don't buy the game, save the money instead, or it is just a miscellaneous cluster

Cluster 8:

Game, cyberpunk, video, biggest, year, bug, happen, suppose, break, crash

Analysis: disappointed about the game breaking bugs and crashes in the biggest release in years

Cluster 9:

Ps, refund, cyberpunk, version, unplayable, xbox, sony, game, textures, inexcusable

Analysis: complaining that the ps and xbox versions of the game are at an unplayable state and the textures were bad on the consoles to make it run at a decent framerate

If we have a lower number of clusters, we get a basic overview of the reasons for discontent. As we keep increasing the number of clusters, we get more specific reasons for the negative sentiment. When we increase the number of clusters to 20, we get a few more specific points such as:

Cluster 4:

lawsuit, mislead, investors, class, action, publisher, allege, beware, free, download

Analysis: Here, we see people talking about the class action lawsuits that a few investors put on CDPR for misleading and lying to the investors by giving them false promises about what the game is going to be at release, hence, the market price of CDPR had been artificially and falsely inflated by the Company's misleading statements.

Cluster 9:

ransomware, hit, attack, projekt, cd, developer, makers, hack, mobile, cyberpunk

Analysis: This talks about the time a ransomware attack had happened at CDPR HQ just a few months after the release of CP2077

Cluster 14:

save, corrupt, file, mb, big, cyberpunk, permanently, items, larger, size

analysis: This talks about a bug where the save file for the game got corrupted and players lost a lot of items and progress due to this bug. This frustrated players as they worked hard and even after ignoring other bugs and crashes, they ended up with severe and game breaking bugs like this one.

We can also use the elbow method to get the ideal k value and analyse all the clusters at that point, to try to get the maximum main points of discontent noticed in the tweets.

We also take a look at the clusters for the positive tweets, where we get:

Cluster 0:

judy, panam, romance, cyberpunk, alvarez, palmer, love, guide, best, character

Analysis: We see that players really liked the characters, mainly the main romantic pursuit characters like judy Alvarez, and panam palmer. This shows that people enjoyed time with these two main characters a lot.

Cluster 4:

cyberpunk, like, pc, good, love, look, character, new, best, time

The pc players had relatively fewer bugs, therefore it was obvious they were the ones giving it more praise. Also, pc had very few game breaking bugs, so people who played the game on pc had lesser complaints. Here we also see that they loved the look(aesthetic) and characters of the game, making them the strong points of CP2077.

Cluster 6:

stream, live, come, twitch, cyberpunk, hang, pm, join, play, twitchstreamer

Analysis: We also see that a lot of twitch streamers played the game on stream, which made more people watch and talk about the game. Due to the game's hype, a lot of streamers were hoping to csh in on the hype train, and this positively impacted the number of people talking about the game, and kept eyes on CP2077 for longer.

Cluster 8:

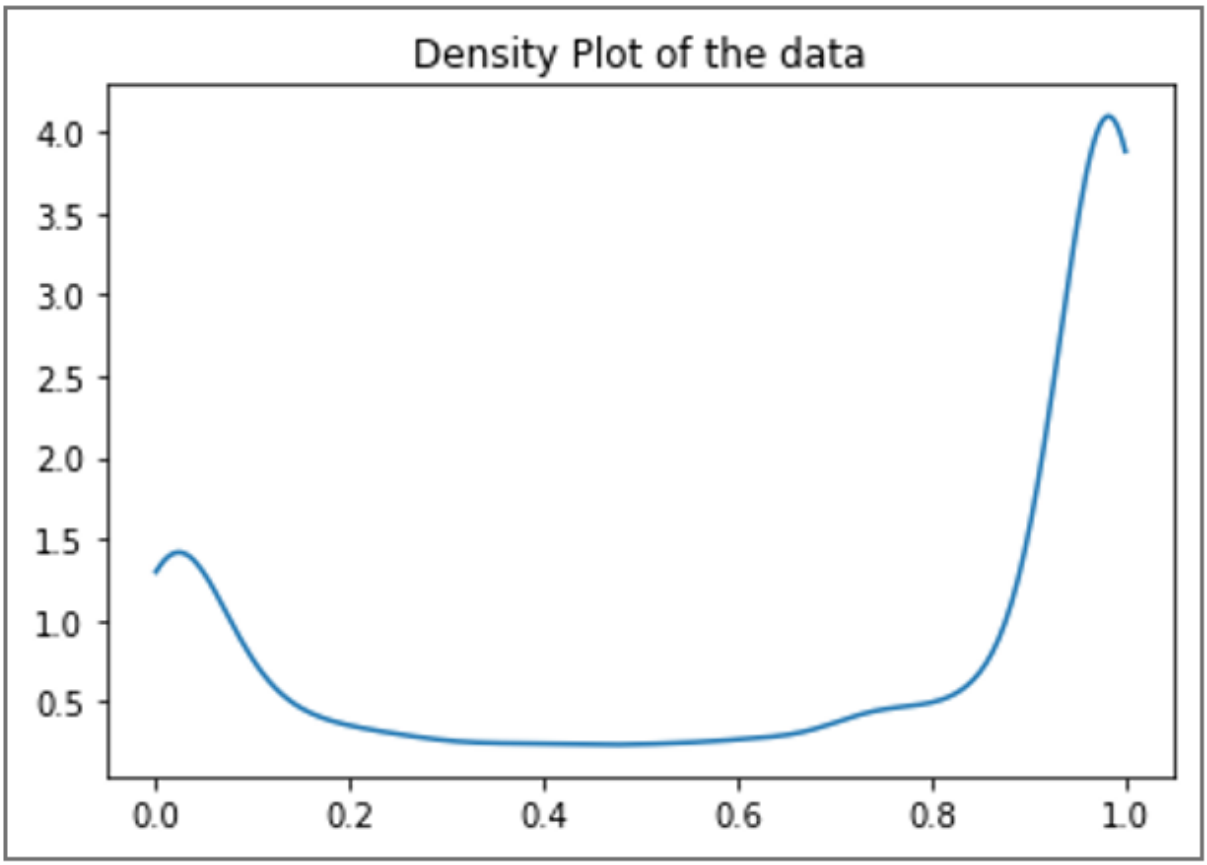
photomode, virtualphotography, cyberpunk, vgpunite, cdprojektred, thecapturedcollective, shutterpunk, gamergram, nightcity, thephotomode

Analysis: Here, we again see that the players loved the aesthetic of the game, and the positive tweets shared pictures in-game and enjoyed the graphics to the fullest.

Hence, clustering analysis of the game based on our sentiment analysis allows the company to look at the main strengths and weaknesses of the game, allowing them to work on the main points of discontent, and also advertise on the most positively talked about points in the game.

Gwent

-Sentiment:



The above graph is a density distribution of the sentiment of tweets related to Gwent during the months of may, june and july of the year 2017 when gwent was released. As we can see a huge majority of the tweets are positive whereas a few of them are negative.

By considering 0.5 as the cutoff point for positive and negative tweets, we get

Positive tweets: 70% of ~36000 tweets.

Note: Gwent was in beta, i.e it was a part of the DLC Wild Hunt to Witcher 3 for about a year before release. The released game however had differences from its beta version. The sentiment during the time of release also takes into account the community disliking some of the new additions as compared to the beta.

-Stock Market:

Gwent being a small release did not impact CD projekt's stock price by a huge margin.

This graph shows the stock price of CD Projekt for the year 2017.

As we can see from above a small bump is noticeable around May when Gwent was released as an independent game. The constant steady increase of CD Projekt's stock price can however be attributed to various other factors including the Cyberpunk hype and the Witcher 3 effect which became one of the most loved and played game of all time.

Obstacles faced:

Error while predicting sentiment:

We forgot to fit the training data onto the tokenizer while predicting, which gave us the same sentiment score for all the data. We recognized our mistake pretty quickly and fixed it

Finding training data:

It was hard looking for good training data. At first, we thought of using an example dataset from online for training, but later we realised steam is a better source, and so we fetched the training data from steam.

Web Scraping data:

Getting a twitter API with a high enough limit to get tens or even hundreds of thousands of tweets was hard at first, but we found one soon enough. Also, the steam web scraper had a problem where it fetched incomplete data, so we fixed that as well.

Problems with plotting sentiment:

The gwent data displayed some indexing error while we were trying to plot its sentiment and the tweet volume, this is a problem we are yet to fix

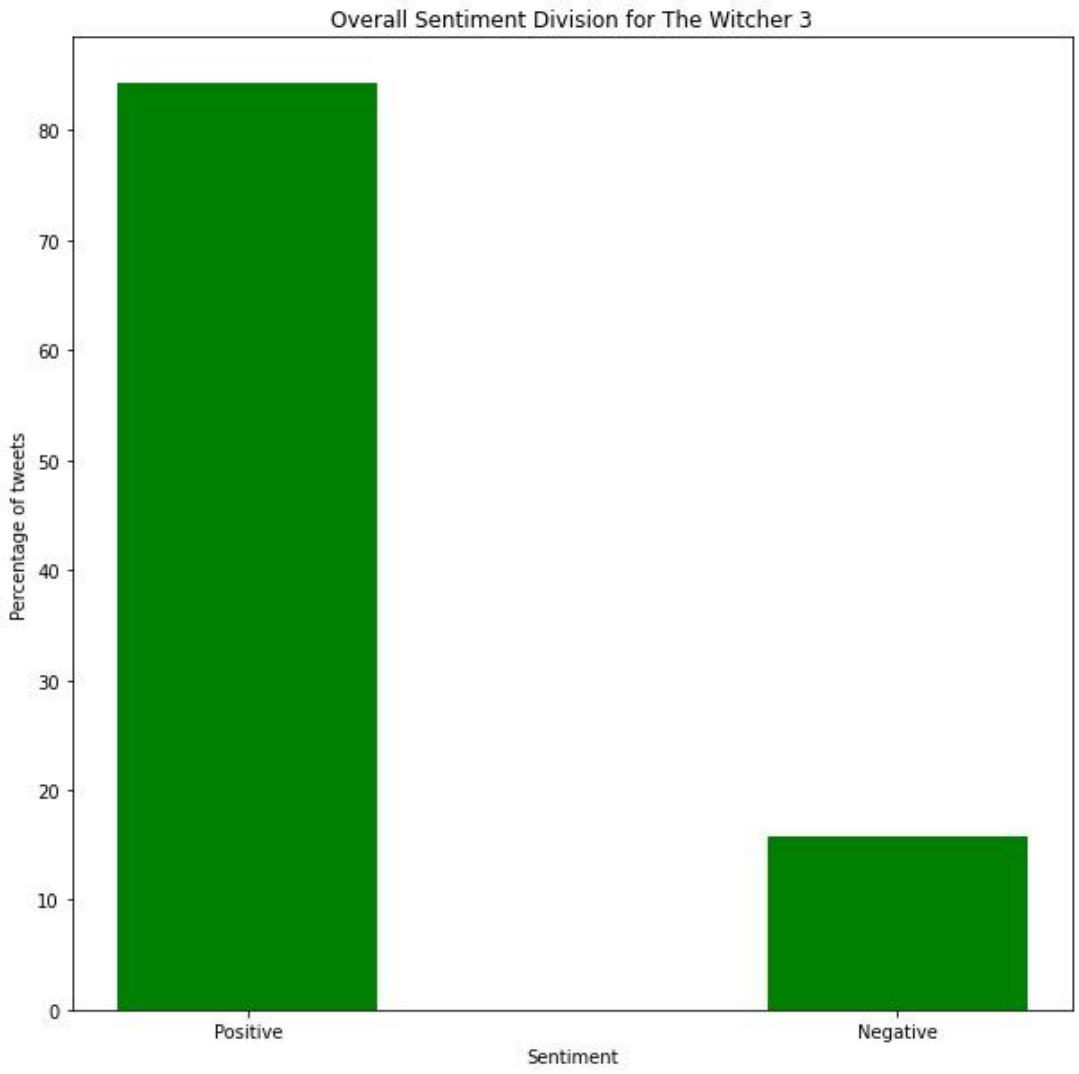
Conclusion:

The main outcomes of the Sentiment Analysis and Stock Market analysis performed on each of the games are as listed:

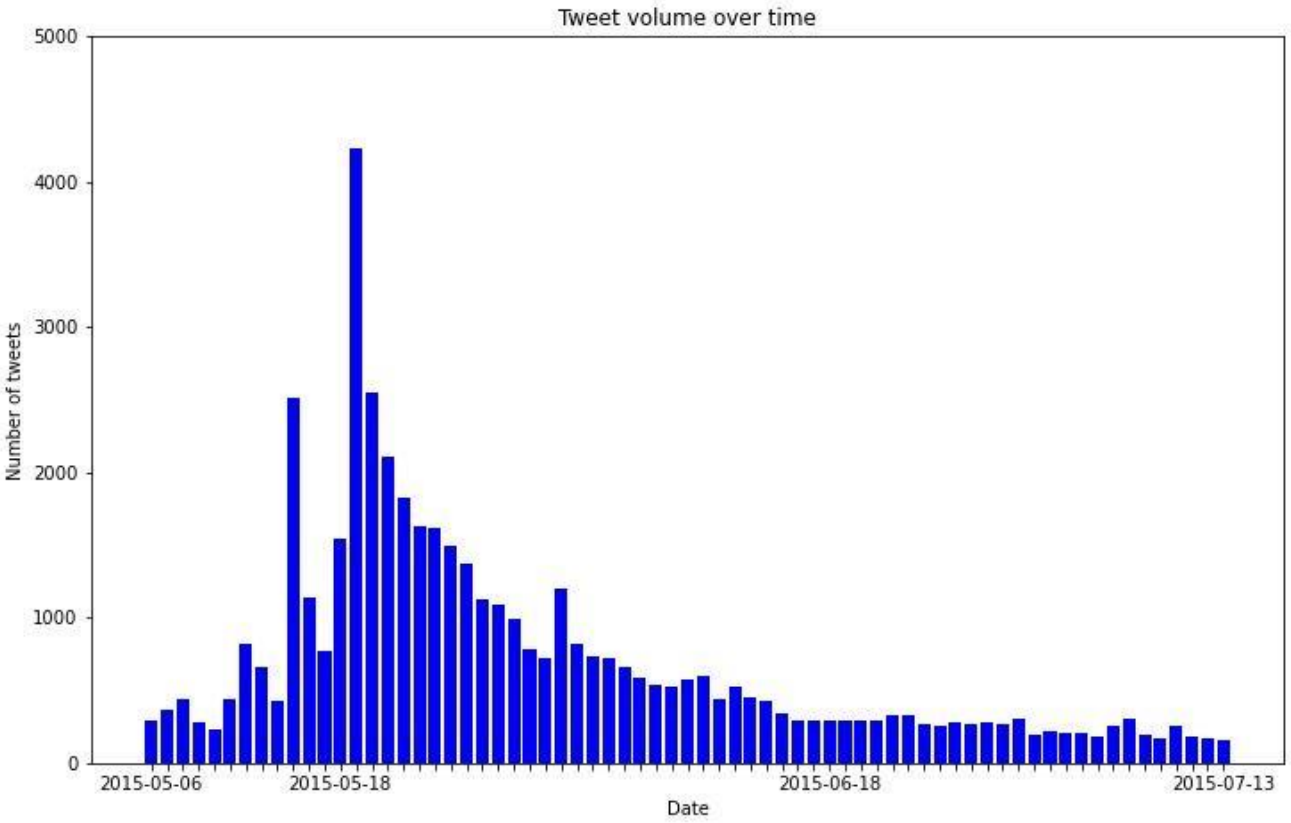
Witcher 3

-Sentiment:

The Witcher 3 had a largely positive response from the community upon its release. Based on our model's predictions and the tweets gathered for the first 2 months after its release, the game had an around 85% positive sentiment.



As seen, Witcher 3 has been a massive hit because it has an extraordinary story-line along with mind beautiful graphics which makes it so unique and interesting but at the same time it requires a lot of cpu power. Also, it has a lot of interesting side quests which keeps you entangled with the game even when you are not completing the main objectives.

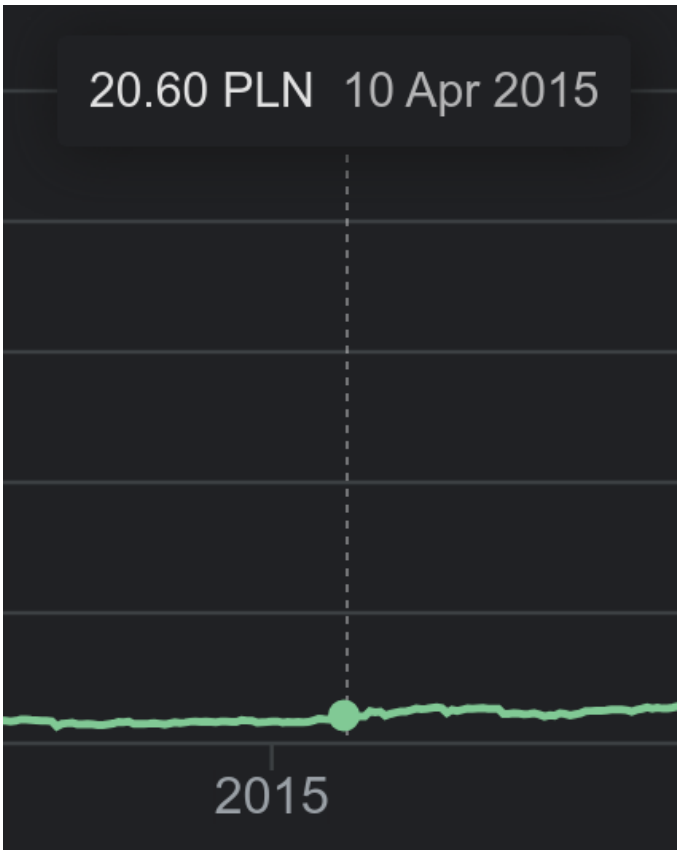


As is the trend whenever a new product releases, we can see the build up and a big jump in the number of tweets on the day of release, close to around 4200 tweets on 18th May, then eventually dying down slowly over the next 2 months.

-Stock Market

Upon release, CD Projekt RED stock prices did not witness huge fluctuations, though it steadily kept on an upward trend. Between 1 month before the game launch and 2 months after, it's stock prices increased by around 20%.

Note PLN is Polish currency, Poland Zloty.

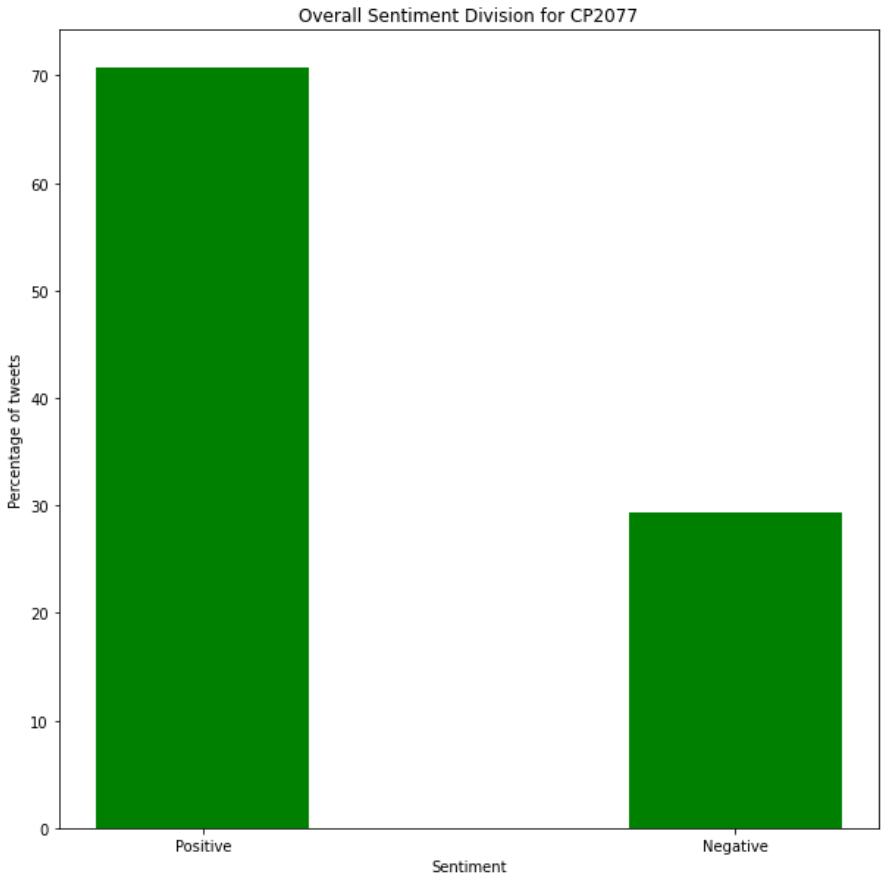


Another interesting point that can be noted is the drop in stock prices during the latter half of 2015, when CD Projekt RED released a huge 15 GB update with fixes, improvements and its first ever expansion, Hearts of Stone on October 8th.

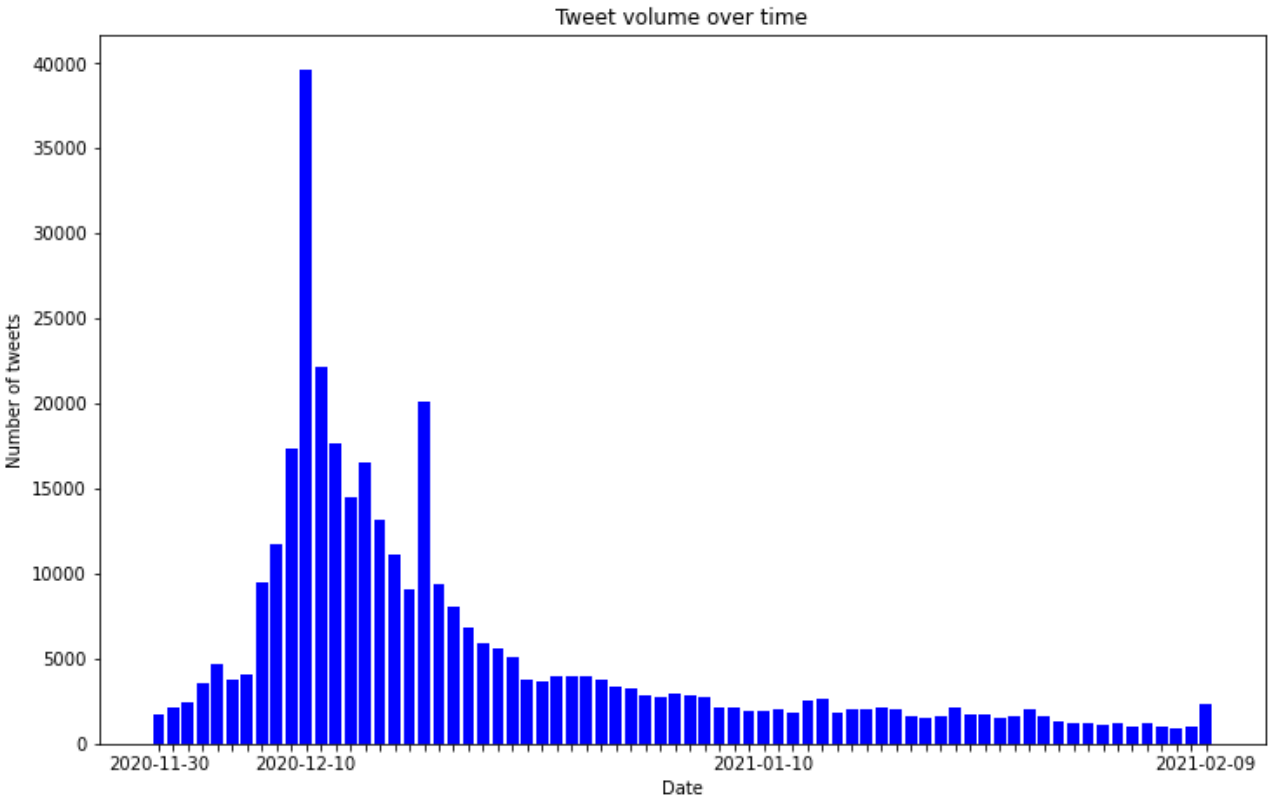
Hearts of Stone was a disappointment for most of the community since there wasn't anything particularly different from the enormous amount of stuff to do that already existed in the main game, aside from a new artisan and crafting options.

Cyberpunk 2077

-Sentiment:

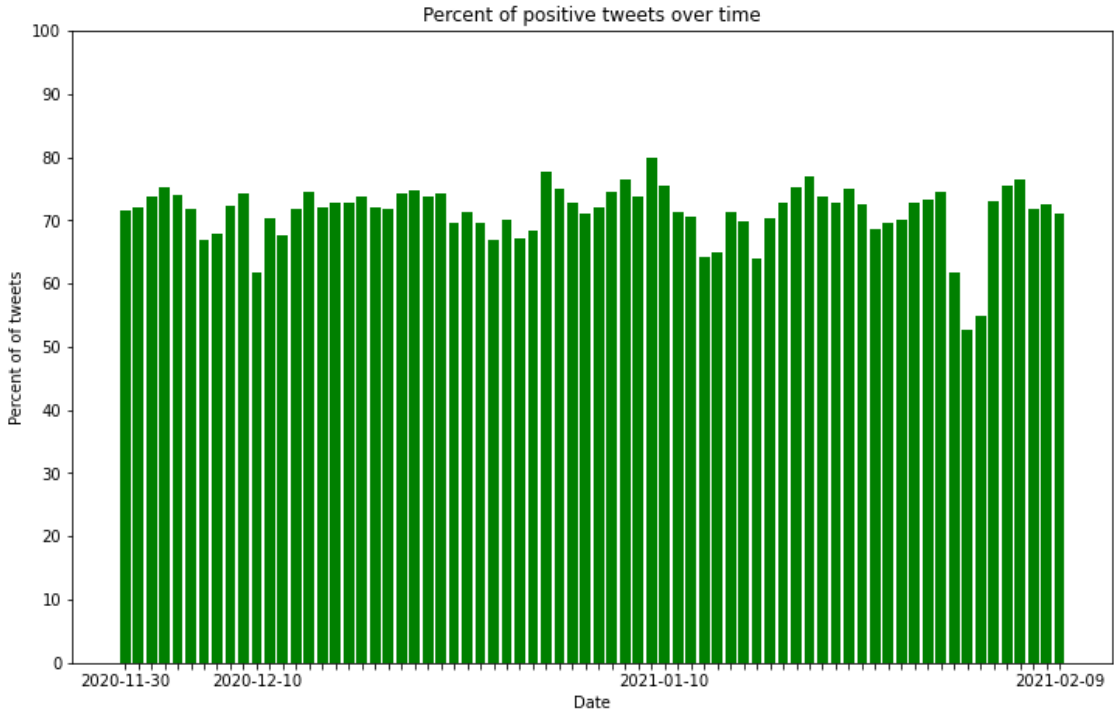


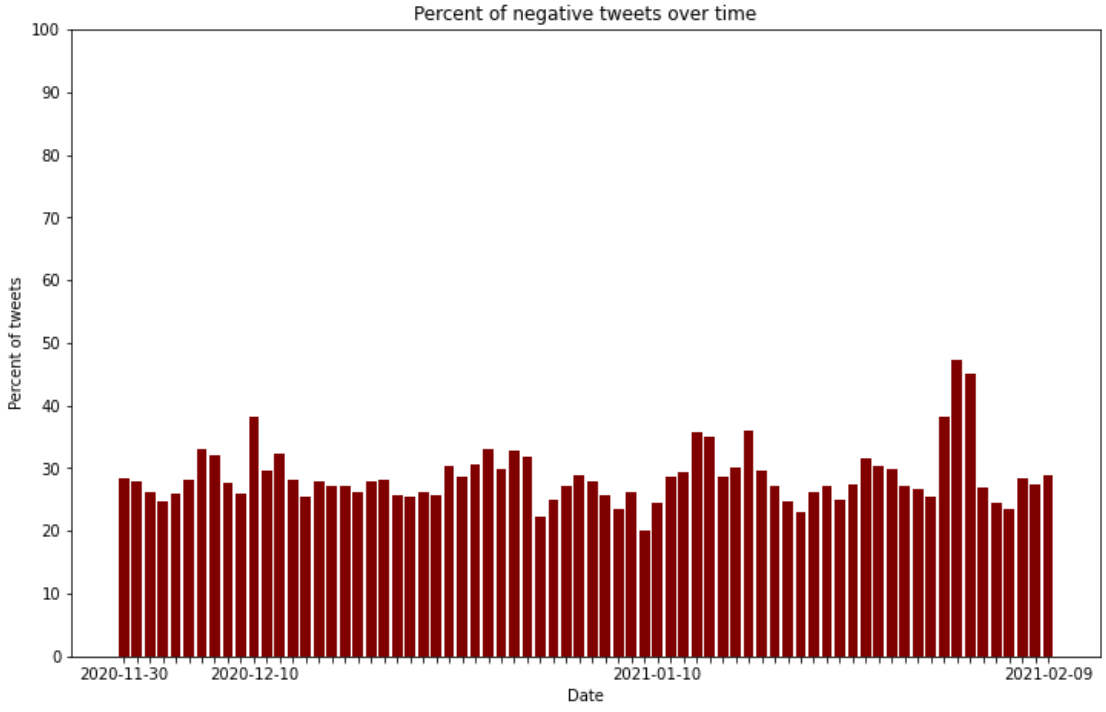
Here, the overall sentiment division gives us over 70% of tweets are positive, and around 30% being negative. This coincides with the reviews present on steam as well, where we have 75% positive reviews. Even though the percentage of positive tweets is pretty high, we can clearly see there is noticeable discontent regarding the game, based on the fact that on steam, a decent doing game can have anywhere between 85-95% positive reviews. We have already seen the cause of this noticeable discontent being the game's bad and buggy launch, and many players' disappointment over the incomplete state of the game not being what was promised and marketed by CDPR. The 70% positive tweets also tell us that the players are discontent, but do not hate the game.



The tweet volume over time shows us the massive spike on release day, with nearly 40,000 tweets on December 10th about the game. Even a few days before release, the game was very actively discussed, showing the excitement of players over the release of the game. The two months after release do not paint the same picture. As with the case with any game after its new release, the discussions about it will die down, but the impact of the messy launch had reduced the active players heavily. We can see a similar trend even in the monthly active steam players we had analyzed in the "How long will cyberpunk be in the market" section of the product analysis.

We can take a look at the sentiment over time for the tweets.

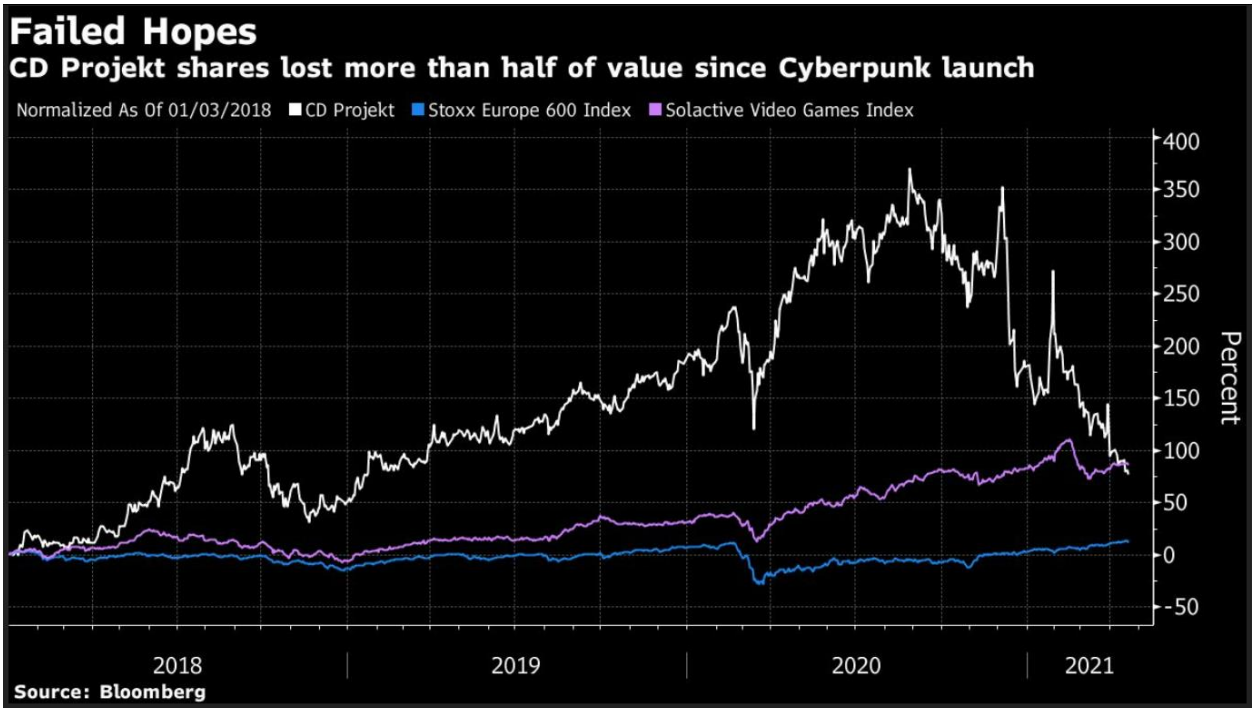




The percent of negative and positive tweets remain pretty constant over time, as for the 2-month period after the release of the game, there was only 1 major patch and multiple minor hotfixes, so the problems still continued and that's we can see the discontent among players remaining more or less constant, with very few deviations from the overall average.

-Stock Market

A few months before the release of Cyberpunk 2077, CD Projekt SA was among Europe's hottest stocks, riding the wave of pandemic gaming demand and on the way to a record high ahead of the release of the studio's much-anticipated Cyberpunk 2077 game. Fast forward 12 months, the stock is the region's worst performer of 2021, down 63% from its August record, with the company facing questions about how it plans to fix problems with the botched release of the bug-ridden game.



Cyberpunk 2077 is a complex futuristic role-playing game that had been predicted to take the gaming world by storm when it was launched on Dec. 10. But it wasn't long before players were reporting multiple glitches, causing negative reviews and triggering Sony Corp.'s unprecedented decision to remove the title from its PlayStation store. A strategy update at the end of march 2021 failed to address concern over how the Polish company will affect a turnaround, showing only that Cyberpunk wasn't yet ready for the full multi-player version that could restore its popularity. It also left questions unanswered over when CD Projekt will be ready with an updated version for next-generation consoles.

Here's a look at the flashpoints in CD Projekt's wild year:

• April 3, 2020: CD Projekt becomes the most valuable company listed in Warsaw as stock became a pandemic winner

• Aug. 27: Shares soar to a record high ahead of Cyberpunk release

• Dec. 8: Stock falls 7.1% as initial press-reviews fail to match high hopes for title

• Dec. 10: Cyberpunk launches with 8 million copies sold in pre-orders, but shares fall a further 8.5% as players report numerous glitches

• Dec. 18: Sony pulls Cyberpunk from its PlayStation Store, sending stock plunging to its lowest since March

• March 29: Studio releases large patch to fix almost 500 bugs

• March 30: Strategy update provides no clear path to fix franchise, pushing stock to a 2-year low

• April 15 2021: Shares drop to 13-month low after company unexpectedly publishes preliminary FY2020 data, delaying full report by 4 days

From after April 15th, stock price has remained more or less the same, with not much change. CDPR has since given multiple patches and updates to fix the bugs and have tried to redeem themselves, but the impact of the bug-ridden launch still remains.

-Clustering analysis:

We have taken a look at the main talking points of both the positive and the negative sentiment tweets to get us a few main topics that people discussed heavily by clustering the data and extracting the top features of each cluster. This mainly gave us

For the negative sentiment, people were mainly talking about:

- the game release being delayed
- people asking for refunds of the game
- the game crashing every few minutes
- sony pulling the game from the ps store due to the console version being unplayable for many
- A lot of graphical, AI and open world related bugs in the game
- The game being overhyped by the marketing team and the game itself underdelivering
- The game being unplayable for console
- lawsuit on CDPR for misleading the investors with false promises that it couldn't deliver
- The ransomware attack on CDPR HQ just a few months after release
- The bug that corrupts save files, making players lose their progress and items

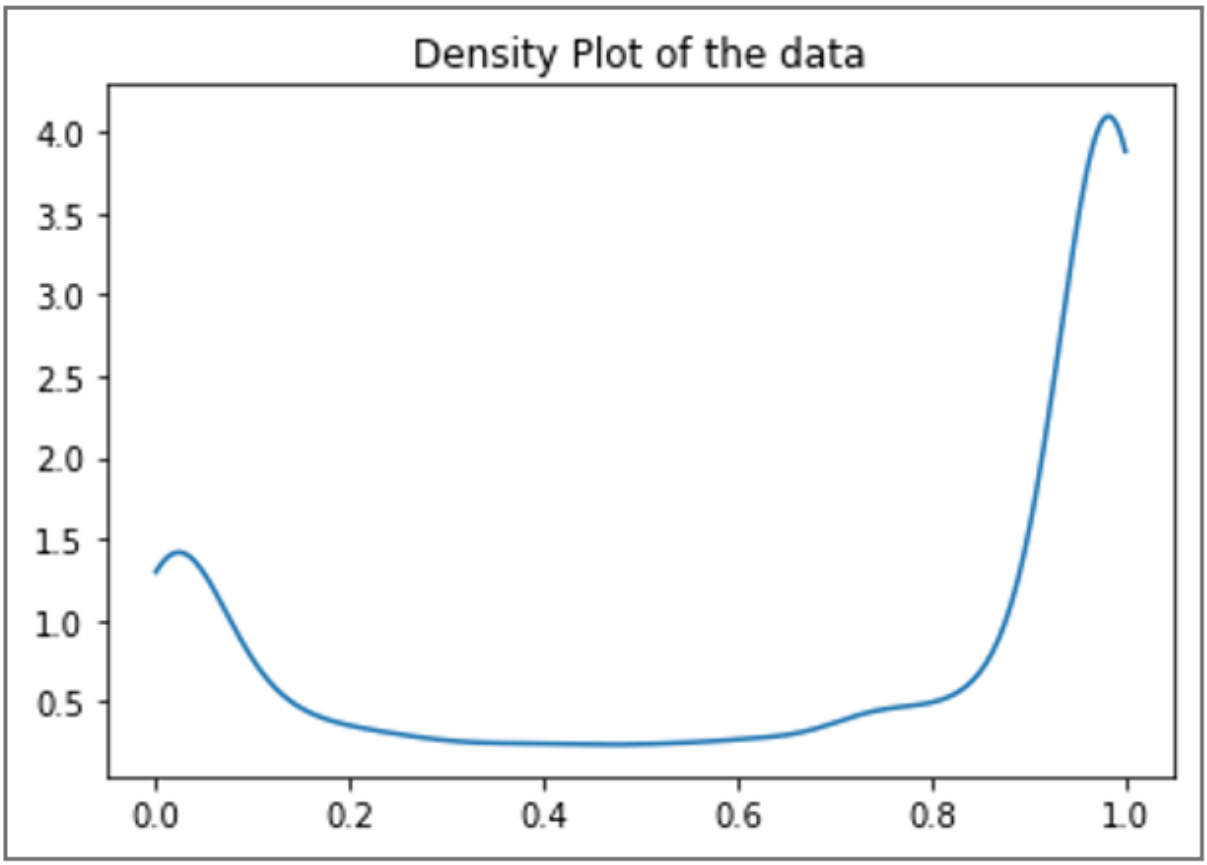
For the positive sentiment, people were mainly talking about:

- People loved the main romantic interest characters judy Alvarez and panam palmer
- The pc players had relatively fewer bugs, therefore it was obvious they were the ones giving it more praise. We also see that they loved the look(aesthetic) and characters of the game, making them the strong points of CP2077.
- A lot of twitch streamers played the game on stream, which made more people watch and talk about the game.
- The players loved the aesthetic of the game, and the positive tweets shared pictures in-game and enjoyed the graphics to the fullest.

Hence we were able to extract the key points for both the sentiments by using KMeans clustering. Clustering analysis of the game based on our sentiment analysis allows the company to look at the main strengths and weaknesses of the game, allowing them to work on the main points of discontent, and also advertise on the most positively talked about points in the game.

Gwent

-Sentiment:



The above graph is a density distribution of the sentiment of tweets related to Gwent during the months of may, june and july of the year 2017 when gwent was released. As we can see a huge majority of the tweets are positive whereas a few of them are negative.

By considering 0.5 as the cutoff point for positive and negative tweets, we get

Positive tweets: 70% of ~36000 tweets.

Note: Gwent was in beta, i.e it was a part of the DLC Wild Hunt to Witcher 3 for about a year before release. The released game however had differences from its beta version. The sentiment during the time of release also takes into account the community disliking some of the new additions as compared to the beta.

-Stock Market:

Gwent being a small release did not impact CD projekt's stock price by a huge margin.

This graph shows the stock price of CD Projekt for the year 2017.

As we can see from above a small bump is noticeable around May when Gwent was released as an independent game. The constant steady increase of CD Projekt's stock price can however be attributed to various other factors including the Cyberpunk hype and the Witcher 3 effect which became one of the most loved and played game of all time.

Applications of the Results:

Sentiment Analysis is a good feedback mechanism for companies to learn how their product has been received by the customers. An even more in-depth sentiment analysis could focus on specific features, updates etc for more concentrated results.

Stock Market analysis is also an accurate method for companies to understand investors and stakeholders reactions to their products, in this case how certain big updates have been received by investors as discussed above.

Also, clustering analysis of the product based on our sentiment analysis allows the company to look at the main strengths and weaknesses of the product, allowing them to work on the main points of discontent, and also advertise on the most positively talked about points in the product. This allows the company to look into the exact topics that people are talking the most positively or negatively about, giving the company complete insight about people's sentiment on the product.

Future Work:

We can try to improve the model accuracy by trying for a better, but more computationally heavy model.

We can increase the amount of data used for the sentiment analysis to get a larger overall view major updates or fixes, and also DLCs on the company, and on the sentiment of the game itself

We can try to find a more accurate training dataset, as the positive reviews on steam can have both positive and neutral reviews, and so having a better dataset like metacritic would allow us to filter out only the polarized reviews and not the neutral reviews to get a better training dataset for better accuracy.

We need to fix the error in plotting the gwent sentiment data that might have happened due to incomplete data.

We can use the elbow method to determine the ideal number of clusters, giving us the exact details of each product with the ideal clustering data