***Cyberpunk 2077***

**Topic Modeling and Sentiment Analysis of YouTube Comments**

Hao Ma, Yufei Cai

**1. Introduction**

*Cyberpunk 2077*, to be released in mid-December, is an action role-playing video game developed by *CD Projekt Red*.It has continuously gained popularity across the globe since its title announcement in 2012, especially during E3 2018 and E3 2019. As more and more information about the game made available to the public, such as videos posted by the developer on YouTube, data containing players’ sentiment, engagement, expectations, and interests for the game are generated and have great potential for data mining.

Yet there are no studies related to video game among us. As video game lovers, players, and fans of the developer, we propose the study of topic modeling and sentiment analysis for *Cyberpunk 2077* via mining video comments from the game’s official YouTube channel (https://www.youtube.com/user/CyberPunkGame). Also, this would be an important and useful project for the game company to learn about their customers. We use machine learning tools to find the latent topics from the comments and identify their themes. In addition, we analyze the sentiments from the comments - how they feel about the game.

**2. Related Work**

Jia Xue, Junxiang Chen, Chen Chen, Chengda Zheng, Sijia Li, Tingshao Zhu. Public discourse and sentiment during the COVID 19 pandemic: using Latent Dirichlet Allocation for topic modeling on Twitter. arXiv: 2005.08817

George Pipis (2020). How to Run Sentiment Analysis in Python using VADER

https://python-bloggers.com/2020/10/how-to-run-sentiment-analysis-in-python-using-vader/

**3. Methodology**

**3.1 Design**

This study follows the procedures of (1) data collection; (2) data pre-processing; (3) data analysis, including latent topic detection, identification, and sentiment analysis. The programming language used in this study was Python.

**3.2 Data Collection**

The dataset was obtained from Kaggle, created by Andrew Dadario. It comprises 4 csv files with the following content: statistics of videos, more than 230,000 comments & replies with like counts and date published, and Google Trend data for the key word *Cyberpunk 2077*. For

the purpose of this study, we only used the one containing 171,781 comments (as of May 26, 2020) from 19 videos published in 2020, 2019, 2018, 2013, and 2012.

**3.3 Data Pre-processing**

As mentioned above, we had a total of 171,781 YouTube comments. To ensure the quality of our analysis, the data was pre-processed. Firstly, we removed 2 missing values (comments). Secondly, we removed all non-ASCII characters such as comments in Korean, Chinese, Russian and all kinds of emojis, since they were not of interest to us. Due to the nature of Pandas “read\_csv()” and “tolist()” functions, the removal of “\n” characters was also included. Then we removed punctuations in this list: !()-[]{};:'",<>./?@#$%^&\*\_~. Lastly, we removed repeated patterns such as “I want this I want this I want this”, etc. The number of remaining comments was 162,850.

**3.4 Methods**

Since we only had plain text without any labels, unsupervised learning method would be appropriate. For the topic modeling task, we chose the Latent Dirichlet Allocation (LDA) method. “LDA is a generative probabilistic model of a corpus. The basic idea is that documents are represented as random mixtures over latent topics, where each topic is characterized by a distribution over words.” “LDA then tries to backtrack from the documents to find a set of topics that are likely to have generated the collection.” We could gain knowledge of the topics people discussed in their comments using this LDA model. The specific package used in this study was Gensim.

We then analyzed the key words of the learned topics and their top 20 representative comments. We came up with a final version of themes of all the topics by manually reading the text.

Then we moved on to sentiment analysis. Sentiment analysis is a text analysis method that detects polarity (e.g. a positive or negative opinion) within the text, whether a whole document, paragraph, sentence, or clause. Sentiment analysis aims to measure the attitude, sentiments, evaluations, attitudes, and emotions of a speaker/writer based on the computational treatment of subjectivity in a text.

VADER (Valence Aware Dictionary for Sentiment Reasoning) is a model used for text sentiment analysis that is sensitive to both polarity (positive/negative) and intensity (strength) of emotion. It is available in the NLTK package and can be applied directly to unlabeled text data.

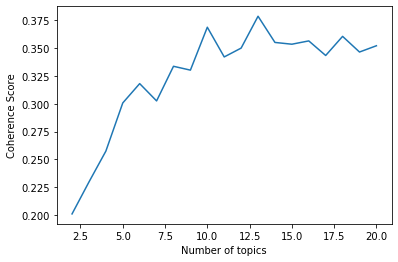
VADER sentimental analysis relies on a dictionary that maps lexical features to emotion intensities known as sentiment scores. The sentiment score of a text can be obtained by summing up the intensity of each word in the text. For example, words like ‘love’, ‘enjoy’, ‘happy’, ‘like’ all convey a positive sentiment. Also, VADER is intelligent enough to understand the basic context of these words, such as “did not love” as a negative statement. It also understands the emphasis of capitalization and punctuation, such as “ENJOY”. Then we could study people’s sentiments from the comments and each of the topics.

**4. Experiment & Results**

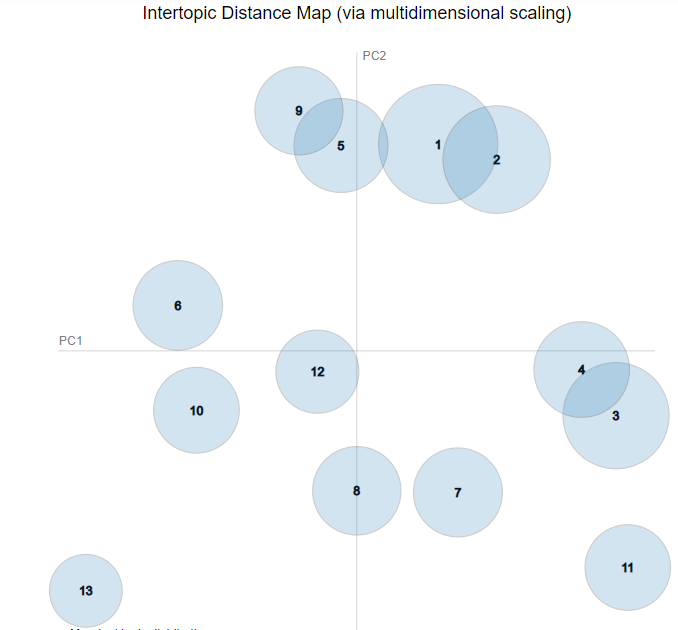
**4.1 Topic Modeling**

With the 162,850 comments obtained from the above pre-processing steps, we further processed the data so that it could be fed into the Gensim LDA model. The processing steps include: (1) converting documents into a list of lowercase tokens; (2) filtering out stop words, such as *the*, *but*, *at*, etc.; (3) removing commonly used words in casual occasions (*wow*, *wtf*, *shit*, *fuck*, *fucking*, *crap*, *lol*, *plz*, *yo*, *xd*, *hey*, *omg*, *im*, *ive*, *oh*). Note that these words were collected from pre-experiments (i.e. we ran the model several times and took notes of those meaningless words with high probability of belonging to one of the topics). Then bigrams and trigrams were generated and added to our collection of documents.

Next, we trained 19 LDA models using different number of topics from 2 to 20. Coherence score was computed for each trial (see Figure 1). Based on the results, we chose the number of topics to be 13. After training the model, an intertopic distance map was generated to evaluate performance (see Figure 2). In the figure, we had big and small circles with acceptable overlapping area and they were scattered in all four quadrants. It showed that we have obtained a good model.

**Figure 1**

**Figure 2**



We extracted the top 15 keywords for each topic based on probabilities, as shown in Table 1. It might not be easy to get the theme of each topic only by looking at the words, so the top 20 representative comments of each topic were also generated (see Table 2). Note that these representative comments did not go through the processing steps discussed at the beginning of 4.1. Then we could analyze those representatives (together with our knowledge of the game) to get the themes (see Table 2). The number of comments for each topic (based on the model output) was also included. There will be remarks for the 10th topic in the next paragraph.

**Table 1**

|  |  |
| --- | --- |
| Topic | Keywords |
| 1 | time, game, damn, yes, god, holy, archive, hyped, teaser, nice, finally, life, trailer,  teaser\_trailer, thank |
| 2 | person, bullets, cool, character, love, like, 3rd\_person, guy, girl, archive\_bullets, 3rd, point, use, car, hope  *Note: “3” was added in front of “rd”* |
| 3 | years, release, release\_date, trailer, cyberpunk, date, waiting, game, old, views,  called, team, stop, title, day |
| 4 | like, looks, looks\_like, keanu\_reeves, keanu, man, game, good, reeves, cgi, dude,  police, shot, super, killed |
| 5 | cd, red, cd\_projekt, project, awesome, ready, coming\_ready, cd\_projekt\_red,  cd\_project, project, yeah, coming, message, cd\_project\_red, dead |
| 6 | blade\_runner, amazing, new, gameplay, game, video, blade, runner, world, watch,  looks, looks\_amazing, hype, open\_world, year |
| 7 | game, people, coming, didnt, trailer, time, long, ciri, ps, watching, came, cyberpunk,  theyre, know, remember |
| 8 | song, music, trailer, like, end, read, years\_ago, video, ago, thought, years, know,  epic, perfect, movie |
| 9 | like, deus\_ex, look, great, ex, deus, game, ghost\_shell, gta, looks, cyberpunk,  looking, shell, ghost, money |
| 10 | pretty, city, night, ok, theres, night\_city, lets, bullet, screen, understand, big, footage, video, daft\_punk, shadowrun |
| 11 | game, wait, good, witcher, games, best, going, cdpr, play, come, gonna, trailer,  excite, hope, cyberpunk |
| 12 | future, real, comment, whats, john\_wick, wonder, secret\_message, john, wick, stupid, cyberpunk, check, secret, sick, kill |
| 13 | don’t, game, want, pc, know, need, buy, dont\_know, like, consoles, right, play, think,console, games |

**Table 2**

|  |  |
| --- | --- |
| Topic, Theme  &  Number of Comments | Comments  *(Note: periods were added to separate each sentence)* |
| 1  **The teaser trailer (hyped)**  **28,030** | BOYIIIIIIIIIIIIIIIII the hype is real first game of my life that im gonna preorder IDGAF fuck the gas  money.  Thats a TEASER The REAL trailer is yet to be released teaser is not trailer.  4th time watching this and still cant get enough This game is released So hyped.  FUCKING ACTUALLY FUCKING HYPED AS FUCK FOR THIS FUCKING GAME GOD  FUCKING DAMN FUCK.  Between this and Star Citizen I think I might as well resign from the next two decades of my life right  now.  No one Me Time to rewatch the teaser for the 100th time.  Its the first real teaser trailer of a conceptually promising game.  Happy 5 Year Anniversary Cyberpunk 2077 Teaser Trailer 3.  Damn half life alyx is 2 days away.  god damn thats a hell of a game.  215 Thank God thats not real water.  Thank God for giving us CDPR.  I was 13 I will be 20 when this game comes out Time flies.  GOD DAMN IT I JUST GOT MY MONEY.  Oh my this game will ruin my social life.  You can preorder 1 year from now The hype is real lol.  Thats how a game is made god damn it.  God damn they actually released it.  Which one will drop first Star Citizen or CyberPunk 2077.  It would be great if we could get a trailer on the 5 year anniversary of the teaser Which was released on January 10th 2013. |
| 2  **1st/3rd person perspective**  **(3rd person preferred)**  **11,779** | Im hoping they add the options to play the game in 3rd person view Always been like that when I play  rpg.  Can we change between 1st and 3rd Person I Wanna Play but i cant play 1st Person Shooter.  I still prefer 3RD person camera so other people with FPS motion sickness in FPS give me a  headache.  i hope there was an option to either play as 1st person or 3rd person play mode.  i like 3rd person view more than 1st person.  Wish they had the option to play in 3rd person I like to have that option.  1st person game give me motion sickness fuck.  id like to think that the main character is Markus and thats Detroit after some decades D.  PLEASE PLEASE PLEASE put Garelt of Rivia in this game somehow Can you imagine seeing him as a side quest character being a bounty hunter with all of those cool augments on him And he has two  pistols that say Silver and Steel On the side of them Little girl squeal engage.  Does the camera can turn into third person just like skyrim i think it would be better can be in FPS andTPS.  I like the game a lot but i personally think it should also be on 3rd person.  I hope they also add an option for 3rd person camera view.  well i hope the male voice is more badass cause im not playing female for sure.  Wish we could switch to on foot 3rd person that wouldve been cool.  111 female main character Im assuming same jacket as the on the guy is wearing.  1ST PERSON TO 3RD PERSON VIEW SWITCH.  is ther any way to view in 3rd person i would like to see my character since i can dress up etc.  No 3rd person rpgs It is sad and no way I will play it now.  This versions not bad but I like the first version of chippin in a lot better.  A little disappointed that its not a 3rd person view was optional and not locked. |
| 3  **Release date (been waiting for a long time)**  **9,532** | Imagine waiting 8 years for a game until its release since you graduated from elementary school to  graduating high school Crazy.  4146 Jesus Christ head got blow off mmy legs.  After 5 years of waiting I realized that 2077 isnt the title of game its the release date.  Seven years SEVEN FUCKING YEARS IVE WAITED FOR ONE GAME.  Lol this trailer came out my first year in high school and Im graduating next week lol.  Ser que ele vai ter dublado para PT BR se tiver vai ser + Top ainda \o Que venha 2020 logo rsrss.  I was 10 years old when this trailer came out when the game releases i will be 17.  The game is not called Cyberpunk 2077 the 2077 part is the release date of Cyberpunk.  The game is actually just called cyberpunk the 2077 part is the release date.  2077 is release date oh i will be 78 years old when this game release.  maybe its just called Cyberpunk and the release date is 2077.  Well guys the game finally has a release date over 7 years after this trailer was released.  Sonystada j ta aparecendo pra encher o saco pq esse jogo que vai chutar o rabo de the last of lacrao no ano que vem.  i was 10 years old when this trailer came out jesus.  The original trailer had 12 million views in 5 years This one has 11 million in less than a week.  Oh I get it The game is called Cyberpunk and the release date is 2077.  4145 Jesus christ head blownoff mmy legs.  Olha q esse trailer do comeo de 2013 ou seja o jogo est sendo feito desde aquela poca ento vai ser um  jogo foda demais.  JESUS CHRIST Loses his head MMY LEGS.  7 years later Release date. |
| 4  **Keanu Reeves**  **9,168** | The guy in the picture looks like keanu reeves OHHH IT IS HIM.  Keanu Reeves in cyberpunk This game just got even better than it already was.  Looks like max Payne 3 with cyberpunk coat.  Omg keanu reeves for reaaal this game just got a whole lot better than it already is.  Only thing that would make this game better would be idk Keanu Reeves or something.  not gonna lie still really good but i miss the e3 2018 trailer.  Game just got 10x more better when Keanu reeves appears.  KEANU REEVES IS GONNA MAKE THIS GAME 100 BETTER.  This looks like Keanu Reeves.  Not gonna lie it looks like apex.  Literally i got goosbumps when keanu Reeves showed up.  Keanu Reeves better have a cyborg dog.  Keanu Reeves has entred the Game Now thats what i am talking about.  Keanu Reeves its just getting better and better.  Cyberpunk 2077 2019 Trailer I got Keanu Reeves.  Seeing Keanu Reeves in this made the game EVEN BETTER.  Looks like someone killed his dog again.  My one goal in this game is getting Keanu Reeves to like me.  Is that fucking Keanu Reeves This game is gonna be wild.  Imagine if you killed a dog in this game. |
| 5  **CD Projekt Red (has best games, is best company)**  **8,706** | EASingle player games are dead CD PROJEKT RED EA is dead.  This Game Is BREATHTAKING Thank You CD RED PROJEKT RED.  CD PROJECT is the most fucking awesome company of games in the world I love all  games of you.  CD Projekt Red we really love you guys DO IT FOR LOVE Its from gamers to gamers.  Cyberpunk 2077 is shaping up amazingly All it just needs is official modding support  from CD Projekt RED.  Coming When its ready this is why we love you CD Projekt Red.  Cyberpunk 2077 Red dead redemption 2 The elder scrolls 6 AAAAAAAAAAAAAAAAAAAAA.  EASingle player games are dead CD Project Take My Keenu.  I will Support everything from CD Projekt Red because of The Witcher 3 They can do  great Games.  CD Project Red produces some the best games in the industry Love it.  COMING WHEN ITS READY Why CD Projekt Red WHY you trolls.  Dear CD Projekt Red youre all breathtaking.  CD PROJEKT RED YOURE BREATHTAKING.  Youre breathtaking CD Project RED.  Well it would be the First shooter from CD projekt RED as they are more rpg focused  game developers.  HEY CD PROJEKT RED YOURE BREATHTAKING.  My body is ready CD projekt Red.  Youre BREATHTAKING CD PROJEKT RED.  CD PROJEKT RED youre breathtaking.  Coming When its ready Hell yeah. |
| 6  **Gameplay, graphics**  **(many comparisons with Blade Runner &Watch Dogs)**  **10,069** | Im just blown away by how amazing this game looks This is the standard that video  games should have.  The open world aspect is just amazing Well all of it looks amazing.  Dawno nie widziaem takiego stosunku apek do wywietle jak i apek w gr do tych w d  Ludzie ju wiedz e redzi dotrzymuj sowa ukazanego na zwiastunie i nie jest to tylko pusty obraz a zapowied czego wielkiego Nic tylko czeka.  watch dogs2 + blade runner = cyberpunk 2077.  So is cyberpunk 2077 an rpg I hope its something with an open world and the graphics  will be high performance.  Is it just me or the graphics of the game looks worse than the first gameplay reveal trailer.  Kurcz naczekam si troch na Cyberpunka 2077 Lubi kupowagra w gry w wersji ostatecznej tzn wydanej z wszystkimi dodatkami wpeni poatan Zawsze te poluj na jak promocj Jeeli ma on wyj w kwietniu 2020 roku to najprdzej odpal j w 2021 roku Moja cierpliwo zostanie poddana prbie jeli j przetrzymam bdzie mona o mnie powiedzie e jestem cnotliwym czowiekiem.  Watch dogs 2 + blade runner= ciberpunk 2077.  blade runner and snacher cyber punk.  If Blade Runner was an open world RPG.  basically the Blade Runner game Ive always wanted.  I think this game is a better watch dogs 1.  This is the trailer we get No actual gameplay no nothing Just an edited series of clips  showing nothing off.  I hope this game will be better then watch dogs.  Imagine watch dogs legion concept in cyberpunk 2077 world.  Japierdole jakie to jest dobre miodne super wspaniae wymienite dobrane fantastyczne  wyjebane w kosmos na szybko nie jestem wstanie znale wicej sw na to arcydzieo  miodnoci Ju zaczynam oszczdzac na now kart graficzn D.  I do hope they include some elements from Blade Runner.  this will be the blade runner of video games.  Zalatuje Saint rowem ale tylko odrobink xd nie jest te to wiedmin ale 100 bdzie mojej  kolekcji xd.  The Graphics Looks INSANE And Its OPEN WORLD |
| 7  **Surroundings, characters in the game trailer**  **10,273** | I think about getting the super duper collectors edition of this game I will also have to  write my Bachelor work then Fml.  I didnt know that a game with master yi would come D.  This trailer went through ps3 ps4 ps4 pro and about to go through the ps5.  I appreciate CDPR showing the surroundings during the day because some people thought that there would be no actual daynight cycle in this game.  Didnt know rick ross was going to be in the game.  Guys the hype train is leaving soon Get your tickets now and preorder the shit out of it No regrets.  3355 Master Yi has come a long way.  Didnt know Rick Ross was in Cyberpunk.  Imagine if the Hack on Xbox wasnt Cyberpunk it was Cyberpunk when Xbox is being  Robbed Lol.  I havent said wow for a game in a long time WOW.  Maybe this Game Coming to PS5 and Xbox Two.  Fucking 2013 i didnt know that this game had been in development for so long.  I said models Do you know what that means in the 3d models can come very close to this Cant anyone reply to someone without insulting them.  My first time buying a collectors edition anything Its hype 3.  When is the collectors edition going to be available for preorder.  I cant be the only one finding it hard to believe that the game is coming out soon.  1425 Didnt know Rick Ross was in the game.  This is the game Im buying a PS4 pro for I guess.  Cocksuckin attention is the cringiest thing Ive heard in a long time.  I didnt know Rick Ross was in this game 116. |
| 8  **Background music**  **(the one sounds like Deewani Mastani)**  **10,600** | At 207 background music sounds similar to a bollywood Movie Padmavat songs Deewani Mastani.  Why I hear voices from the end of Deewani Mastani song OST Bajirao Mastani in that  melody.  During the interview the background music made me leave as it was too annoying having to force myself to listen to what was being said.  this sounds like the song dewani mastani from bajirao mastani movie.  From 105 this song bgm is copied from Bajirao mastani movie song mai deevani ho  gayi its an Indian song lol.  is it just me or 106 sounds like deewani mastani from bajirao mastani.  This sounds very similar to that song in the bajirao mastani movie where mastani danced  for bajirao.  Listen Deewani Mastani song from Bajirao Mastani movie.  A part of this theme song is copied or same as a song in the Indian movie Bajirao Mastani.  Sounds like deewani mastani from Bajirao mastani.  Sounds like the ending Deewani Mastani.  Copied Listen to bajirao mastani Bollywood song Mastani.  From 204 music copied from bajirao mastani song deewani mastani.  103 Its copied from Deewani mastani movie name=bajirao mastania Bollywood movie  song.  The music that starts 207 I am sure is from the Bollywood movie Bajirao Mastani.  there is a part of the song is taken from Diwani Mastani song from Bajirao Mastani.  There was something familiar about this song I can hear Deewani Mastani from Bajirao  Mastani movie with Deepika Padukone.  Hey thats deewani Mastani bollywood song in the background.  105 Do you hear dewani mastani song from bajirao mastani film or is it just me.  Bgm copied from Indian movie song bajirao mastani. |
| 9  **Comparison with other games &films**  **(GTA5,** **Deus Ex, Ghost in the Shell)**  **14,058** | GTA5 + DEUS EX = CYBERPUNK Just the type of game ive been looking for since I  played GTA series and the Deux Ex series.  Great game It reminds me of like a mix between Fallout GTA KOTR in a world like  Ghost In The Shell.  1020 Whatever caused this ghosting If its not a gameplay element hope it would not be in the final game I love the Deus Ex style but not so much for the smearingghosting artifacts that games antialiasing have.  This game looks so good Although its got look and style of a GTA game hoping it isnt  tho.  I like this game The metropolis is Deus Ex inspired but that is it The story of the game  maybe very good.  Reminds me a lot of Ghost in the Shell and the manga Eden Its an Endless World Looking forward to playing this.  This is like if Deus ex GTA 5 and ghost in the shell without the philosophy had a  threesome.  This gameplay feels like a mix Deus Ex Fallout and Killzone.  look like deus ex btw like a good game.  Honestly I think this game graphically is between FEAR and GTA5 it doesnt look like an improvement over the latter.  Looking forward to this hope it has some Ghost in the Shell inspiration.  Feels like i will be playing ghost in the shell Looks awsm.  Doesnt look interesting at all hope its just a shitty demo of a good game.  Its Like Dredd Deus Ex and Ghost In The Shell had a threesome.  Cyberpunk world looks a lot like the world of Ghost in a shell.  This gives me Deus ex vibes but way better emersion and gameplay with different  Mechanics.  this reminds me a lot of Ghost in The Shell not the shitty movie one of course.  Fuuuuuuck Its like gta + witcher + deus ex + an action movie O.  I hope this games story is a bit like Judge Dredd Gritty and raw.  Dexter DeShawn voiced by Mike Pondsmith himself. |
| 10  **DISCARDED\*** | Literalmente estoy entrando en una fase de depresion esperando el juego por que ya  ninguno me entretiene y no hago mas que pensar en el dia que pueda experimentar CyberPunk2077.  Isnt Jackie Welles the recurring sidekick character in the trailers suppose to be a big  character so this trailer is a spoiler in that case Bit annoyed by that.  Crear el juego en primera persona y obvio no lo comprare Deben saber que asi como yo  hay una comunidad de miles de personas Ojo hay que ser mas inteligentes.  Crear el juego en primera persona y obvio no lo comprare Deben saber que as como yo  hay una comunidad de miles de personas Ojo hay que ser mas inteligentes.  Este juego junta todo lo detestable acerca del neocyberpunk Tiene tan poco del gnero que para hacertelo creer lo dice en el nombre.  para que van a personalizar el personaje si va a estar en primera persona osea es al  pedo.  es una pena que el juego sea en primera persona y no en 3ra.  An sigo deseando que implementen el modo 3ra persona As podramos apreciar diferentes maneras de ver el mundo cyberpunk y poder disfrutar de la moda en nuestro personaje Pnks Nt ead .  Does this mean theres no character creation I wanted to make my own character.  Soy la nica que no puede parar de ver este gameplay Imagnate cuando salga la viciada que le vamos a pegar.  Por mas que lo intento no me gusta la vista en primera persona.  Soy el unico al que no le llama la atencion este juego.  This thing ruins entire trailer.  Si el juego viene en espaol latino sera perfecto El espaol de Espaa y sus modismos en las traducciones me la sudan.  Im throwing my money at my phone screen but nothing is happening.  Que suerte que no supe nunca de este juego hasta ahora me cuesta esperar 1 ao por un  juego as que imagnate 6.  SI AGREGAN LA OPCIN DE VISTA EN TERCERA PERSONA ES PERFECTO SI LO HACEN NO TIENEN COMPETENCIA.  I keep throwing money at the screen but nothing happens Edit Now my screen is cracked.  por que no lo sacaron el 21 abril que es mi cumpleaos jajjajaja Esta muy bueno juego  me encanta mucho la ambientacion futuristica y el estilo de mundo libre amo eso enserio  estoy con hypee por este juego desde julio a comenzar ahorrar para comprar este grandisimo juegazo 3.  No entiendo para qu tanta personalizacin de mi personaje si al final no se puede ver en  pantalla Es lo nico que no me gusta Por lo dems es una compra segura. |
| 11  **Better single player games**  **(against EA)**  **21,491** | I think Witcher 3 Wild Hunt is best game ever made but I gotta feeling this will be even  better Story seems insanely good as well as graphicsJust typical CDPR game D.  This game will transcend its market many non gamers will enjoy it and maybe will  become gamers afterwards Seriously its being a long while since I am that excited about a game cant wait.  EA Single player games are dead CDPR hold my beer.  EA Single player games are dead CDPR Hold my beer.  EA single player games are dead CDPR hold my beer.  EA Single Player games are Dead CDPR Hold my beer.  EA Single Player games are Dead CDPR Hold my beer InCDPRweTrust.  CDPR EA is dead games are dead.  EA No more Single player games CDPR Hold my beer.  EA single player games are dead CDPR EA is dead.  Cyberpunk 2077 will be game of the decade Im gonna have to play this game a couple  times to see the different outcomes.  EA Single player games are dead CDPR pffft.  EA Single player games are dead Metro Exodus Hold my vodka.  This is going to be the best game for a really really fucking good single player game.  A good single player RPG ahhh just cant wait to play it.  EA single player games are dead CDPR LOL.  Hands down one of the best teams out there you have so much on all other devs right now the quality is just outsanding Cant wait to play the game.  CDPR put Elon Musk Tesla Cybertruck in Cyberpunk 2077 PLEASE I am getting the  truck and the game.  CDPR put Elon Musk Tesla Cybertruck in Cyberpunk 2077 PLEASE I am getting the  truck and the game.  First The Witcher 3 now Cyberpunk 2077 CDPR I love you guys so much This game is  going to be released. |
| 12  **John Wick**  **6,152** | 2014john wick 2017 john wick cyberpunk.  Plot twist Cyberpunk 2077 is actually John Wick Chapter 77.  Can believe they put John Wick Fortnite in Cyberpunk.  cant believe they got john wick from fortnite.  I cant believe John Wick from Fortnite is in Cyberpunk.  All here scream John Wick Jonh Wick Matrix is all and Neo.  Fortnite we got REAL JOHN WICK.  We got John wick from fortnite in cyberpunk 2077.  Nobody Fortnite kids hey that John wick from fortnite.  John Wick in cyberpunk is better than from fortnite lol.  When John Wick mashed with Neo in future 2077.  John wick w cyberpunk haha.  So thats what John Wick has been up to lately after his stint in the Matrix.  When you accidentally kill a dog and its from one John Wick 0.  Isnt that John wick from fortnite.  Omg guys its john wick from fortnite.  I bet someone is gonna dress as John Wick.  Fortnite kids We have John Wick.  John Wick pff is NEO of Matrix.  John Wick 4 Cyber Pencil. |
| 13  **Hardware, performance, &platforms**  **15,432** | wanna pre order on steam but Im not sure if it will work on i79750H gtx 1660ti 24GB ram8 + 16 flex at 60fps medium settings.  dont buy from steam guys buy for GOG they deserve all the money dont let this happend like the witcher 3.  intel i15 114 ghz gtx 7080ti OC 32gb ram 6000 mhz will run it at 40 fps in low sarcasm i want this game so bad.  I truly want this game but i cant buy The game and i have no xbox one or ps4 and thats  really fuck me i need to play this think.  minimum requirements to run i9 9900k rtx 2080ti 32gb Ram.  Isnt your PC customizable though How are they gonna handle giving a customizable PC a personality Its not like the Witcher series.  Here i5 6600k 46 GHz GTX 1060 Xtreme Gaming 6G and 16 GB RAM 2x 8GB 3200  MHz CL15 I hope to play at maximum settings 60 FPS in FullHD.  Ive installed 2 GTX 1080 ti in SLI just for this particular game Hope the game will be enough to play it in 4K on ultra settings.  I hope it releases on Xbox one but I dont care I want to play this game.  I dont even have a pc that can run the game but still probably buying this just to give  money to cdpr.  Ill upgrade my pc just to play this game at ultra settings with rtx on.  Who cares that theyre making less money Money isnt all there is you know The game will be better thanks to this.  I was gonna waste my money on this steam sale but i wont anymore Ill save for  cyberpunk 2077.  going to buy this from GOG want to support these guys fully.  This game I will pre order I dont care.  What fps should I expect 1440 with i7 8700 and gtx 1660 Id like to run high graphics  Fps.  Will realese before ps5 So dont know if i buy it for ps4 and then ps5 when i will buy one.  I have a Ryzen 2600 rx 580 and 16 gb of ram would my pc be able to play this.  There were only a few games made me upgrade my PC Cyberpunk will be for consoles as well I will play on the PC.  I dont have a ps4 anymore but I still plan to buy this game just to support this people they deserve it. |

\*: We decided to disregard topic 10 because it seemed like a collection of Spanish comments not detected in the pre-processing stage (which was one of the limitations of this study) and did not contain enough English comments with substantial contribution. This would not affect the validity of the model and the other topics.

**4.2 Sentiment Analysis**

The VADER library returns 4 values:

* **pos:** The probability of the sentiment to be positive
* **neu:** The probability of the sentiment to be neutral
* **neg:** The probability of the sentiment to be negative
* **compound:** The normalized compound score which calculates the sum of all lexicon ratings and takes values from -1 to 1

Notice that the pos, neu and neg probabilities add up to 1. Also, the compound score is a very useful metric in case we want a single measure of sentiment. Typical threshold values are the following:

* **positive:** compound score >= 0.05
* **neutral:** compound score between -0.05 and 0.05
* **negative:** compound score <= -0.05

Since we have cleaned the data in our previous steps, we could analyze the comments directly. Let us analyze some random comments through our sentimental analyzer (Figure 3).

**Figure 3**

**图形用户界面, 应用程序

描述已自动生成**

The output is 31.4% positive and 68.6% neutral. The compound score is 0.4936.

Instead of using “SentimentIntensityAnalyzer()” and “polarity\_scores()”, we also wrote a function named “sentiment\_scores()” to get the sentiment score of any sentences. We randomly tested three comments (Figure 4).

**Figure 4**

**文本, 信件

描述已自动生成**

Now, we would add 4 new columns: **neg**, **neu**, **pos** and **compound** using the lambda function (Figure 5).

**Figure 5**

图形用户界面, 文本, 应用程序, 电子邮件

描述已自动生成

Since we have cleaned the data and have gathered the required information in a structured format, we could perform any kind of analysis. For instance, let us have a look at the compound score for each of the topics (Figure 6).

**Figure 6**

表格

描述已自动生成

We sorted the data by the mean of each topic, and we can see that topic “CD Projekt Red (has best games, is best company)” have the highest average sentiment score, topic “John Wick” have the lowest average sentiment score. Moreover, there is no topic that has the average score of negative sentiment. 4 of them are neutral and the other 9 are positive.

Let us also have a look at the boxplot (Figure 7).

图表, 箱线图

描述已自动生成**Figure 7**

Finally, we studied the distribution of all the comments (not grouped by topics, see Figure 8 & 9).

**图表, 直方图

描述已自动生成Figure 8**

**Figure 9图表, 条形图

描述已自动生成**

**5. Conclusion & Discussion**

# This study shows the topics people discussed and their sentiments to the game *Cyberpunk 2077* after watching several videos from the game’s official YouTube channel, which is a major source for people to learn about the game. The results will be very interesting to players and fans as well as the developer. This is also a study that makes use of video comments from the public instead of just looking at the titles and the number of views.

# Firstly, there are 12 (English) latent topics in the comments and we identified them as: (1) the teaser video; (2) 1st or 3rd person perspective; (3) release date; (4) Keanu Reeves; (5) *CD Projekt Red*; (6) gameplay and graphics; (7) surroundings and characters in the trailer; (8) background music; (9) comparison with other games and films; (10) quality of single player games; (11) John Wick; (12) hardware, performance, and platforms. By digging deeper into the representative comments, details of each topic could be found (e.g. people are very excited about this new game and have given a lot of compliments; some players want 3rd person perspective to be added to avoid headache; people are very, very interested in the character played by Keanu Reeves; game performance on different machines and platforms; finally, people have been waiting for many years (same for the authors of this report).

# Secondly, as we have mentioned earlier, topic “CD Projekt Red (has best games, is best company)” have the highest average sentiment score, topic “John Wick” have the lowest average sentiment score. There is no topic that has the average score of negative sentiment. 4 of them are neutral and the other 9 are positive. We also found that for most of the topics, the median have the same sentiment score (0) with the 25th percentile, which means most of the comments for these topics have positive and neutral sentiment. For topic “The teaser trailer (hyped)”, “Release date (been waiting for a long time)” and “John Wick”, they all have many outliers both above the max. value and below the min. value, which shows that people have extremely opposite opinions over these three topics. Finally, it is obvious that nearly half of the comments (about 70,000) have a sentiment score between 0 and 0.2. Most of the comments have positive or neutral sentiment on this game, whereas there are still a small number of comments (about 30,000) showing negative sentiment. All in all, people are holding relatively positive attitude of the game.

As mentioned above,the natural language processing tasks may require more research since we failed to remove Spanish comments written in a-z/A-Z, but the validity of the model and our analysis was not affected. Moreover, we did not take the replies to some of the comments into consideration to avoid bias. It might be useful to give them some weight and include them in our analysis. It would be interesting to find out why some hot comments got so many replies and their content. Another limitation or challenge of this study is that YouTube comments are written in a more casual fashion, which is harder to deal with.

**6. References**

Andrew Dadario. (2020). Cyberpunk 2077 Youtube Reception.

https://www.kaggle.com/andrewmvd/cyberpunk-2077

vRehruvrek, R. & Sojka, P. (2010). Software Framework for Topic Modelling with Large Corpora. *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks* (p./pp. 45--50), May, Valletta, Malta: ELRA.

Blei, D., Ng, A., & Jordan, M.I. (2003). Latent Dirichlet Allocation. J. Mach. Learn. Res., 3, 993-1022.

# Edwin Chen. Introduction to Latent Dirichlet Allocation. http://blog.echen.me/2011/08/22/introduction-to-latent-dirichlet-allocation/

# Susan Li. Topic Modeling and Latent Dirichlet Allocation (LDA) in Python.

# https://towardsdatascience.com/topic-modeling-and-latent-dirichlet-allocation-in-python-9bf156893c24

Selva Prabhakaran. Topic Modeling with Gensim (Python).

https://www.machinelearningplus.com/nlp/topic-modeling-gensim- python/#13viewthetopicsinldamodel

Jia Xue, Junxiang Chen, Chen Chen, Chengda Zheng, Sijia Li, Tingshao Zhu.

Public discourse and sentiment during the COVID 19 pandemic: using Latent Dirichlet Allocation for topic modeling on Twitter arXiv: 2005.08817

Hutto, C. J. & Gilbert, E. (2014). VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text.. In E. Adar, P. Resnick, M. D. Choudhury, B. Hogan & A. H. Oh (eds.), *ICWSM*, : The AAAI Press. ISBN: 978-1-57735-659-2

George Pipis (2020). How to Run Sentiment Analysis in Python using VADER

https://python-bloggers.com/2020/10/how-to-run-sentiment-analysis-in-python-using-vader/

Aditya Beri (2020). SENTIMENTAL ANALYSIS USING VADER

https://towardsdatascience.com/sentimental-analysis-using-vader-a3415fef7664

Python | Sentiment Analysis using VADER

https://www.geeksforgeeks.org/python-sentiment-analysis-using-vader/

Natassha Selvaraj (2020). A Beginner’s Guide to Sentiment Analysis with Python

https://towardsdatascience.com/a-beginners-guide-to-sentiment-analysis-in-python-95e354ea84f6