

Università degli Studi di Milano Bicocca

Dipartimento di Informatica

Corso di Laurea Magistrale



Analisi della crisi del mercato del Gaming tramite tecniche
NLP e Graph

Lorenzo Megna - 868929

Massimo Trippetta - 869286

ANNO ACCADEMICO 2023/2024

Indice

Introduzione	1
1 Raccolta e analisi dei dati	3
1.1 Raccolta dei Dati dalla Community di Reddit	3
1.2 Analisi e pulizia del dataset	4
1.3 Preprocessing dei testi	6
2 Sentiment Analysis e NER	8
2.1 Sentiment Analysis	8
2.2 Named Entity Recognition	13
3 Graph Analysis	17
3.1 Introduzione	17
3.2 Primo post	17
3.2.1 Visualizzazione del grafo	17
3.2.2 Analisi temporale	18
3.2.3 Grado dei Nodi	19
3.2.4 Analisi discussioni	21
3.2.5 Analisi dei voti	24
3.3 Secondo post	25
3.3.1 Visualizzazione del grafo	26
3.3.2 Analisi temporale	26
3.3.3 Grado dei Nodi	27
3.3.4 Community detection	29
Conclusioni	34

Elenco delle figure

1.1	Numero di risposte per utente	5
1.2	Numero di upvote per utente	5
1.3	Commento dell'utente con più upvote	6
2.1	Distribuzione delle label	9
2.2	Distribuzione del sentiment per post	10
2.3	Variazione del sentiment del mese	11
2.4	Commenti negativi di Gennaio 2024	12
2.5	Appartenenza alla label	12
2.6	Sentiment delle aziende	15
2.7	Sentiment delli giochi più nominati	16
3.1	Grafo primo post	18
3.2	Frequenza vs ore	19
3.3	Indegree vs Outdegree	19
3.4	Frequenza dei messaggi per ogni coppia di utenti	20
3.5	Coppie con più messaggi scambiati	20
3.6	Discussione tra Manjorno316 ed eeeeeeeeEeeEEeeeE6	21
3.7	discussione tra Pickles_68 e eeeeeeeeEeeEEeeeE6	22
3.8	Discussione tra Randi_Butternubs_3 e deleted	23
3.9	Discussione tra deleted e SuperSeriousStud	24
3.10	Commenti con più upVotes nel primo post	25
3.11	Grafo relativo al secondo post	26
3.12	Frequenza di messaggi oraria del secondo post	27
3.13	Indegree vs Outdegree	28
3.14	Commenti scritti da RareCodeMonkey	28
3.15	Risposte al commento di RareCodeMonkey	29
3.16	Ripartizione dei nodi dopo la community detection con clustering.	30

Introduzione

A partire dal 2023 e continuando nel 2024, l'industria dei videogiochi ha subito massicci licenziamenti. Nel 2023 sono stati persi oltre 10.000 posti di lavoro e ulteriori 9.500 licenziamenti sono avvenuti solo nel periodo da gennaio a maggio 2024 [7]. Questi licenziamenti hanno avuto effetti risonanti sia sugli studi di sviluppo di videogiochi già affermati che sulle aziende emergenti, influenzando dipendenti, progetti e il panorama complessivo dell'industria videoludica. I licenziamenti hanno causato la cancellazione di diversi videogiochi, la chiusura di studi di sviluppo o la cessione di questi dalle loro società madri, e migliaia di dipendenti hanno perso il loro lavoro.

Negli ultimi anni, purtroppo, i licenziamenti nell'industria dei videogiochi sono diventati sempre più comuni. Nel 2023, abbiamo assistito a licenziamenti quasi settimanali in tutto il settore e, purtroppo, sembra che il 2024 supererà questa cifra.

L'industria dei videogiochi è più grande e genera più denaro di film e musica combinati, portando 180 miliardi di dollari solo nel 2021 [9]. È anche un settore che diventa ogni anno più rischioso e costoso poiché i giochi AAA richiedono più tempo e denaro per essere realizzati (circa 300 milioni di dollari), portando a una situazione in cui anche un singolo flop può affondare uno studio o un editore. Inoltre, l'intera industria ha disperatamente bisogno di sindacati per aiutare a proteggere i suoi milioni di lavoratori quando le cose non vanno come previsto.

Cause principali:

- **Aumento dei costi di sviluppo:** I costi di sviluppo dei giochi sono in costante crescita, rendendo più difficile per le aziende mantenere la redditività.
- **Cambiamento dei consumatori:** I gusti e le abitudini dei consumatori stanno cambiando, il che può rendere difficile prevedere quali giochi avranno successo.
- **Rallentamento post-pandemia:** La fine della pandemia ha portato a una diminuzione dell'interesse per i videogiochi, che avevano visto un boom durante i periodi di lockdown.

- **Altri fattori:** Altri fattori, come le fluttuazioni economiche globali e le innovazioni tecnologiche, contribuiscono alla precarietà del settore.

Dopo aver osservato l'impatto significativo dei licenziamenti nel settore dei videogiochi nel 2023 e 2024, ci siamo posti delle domande fondamentali, immaginando che un'azienda ci abbia chiesto di effettuare un'analisi per rispondere alle seguenti domande:

1. Quali sono le opinioni e i sentimenti prevalenti degli utenti riguardo alla situazione nel mercato del gaming nel 2024?
2. In che modo il feedback degli utenti sui social media può essere correlato ai licenziamenti avvenuti presso le aziende di videogiochi?
3. Quali sono le principali preoccupazioni e criticità emerse dall'analisi dei commenti e dei post online?

Per rispondere a queste domande, abbiamo deciso di analizzare il sentiment degli utenti espressi sui social media come Reddit, passando poi a una parte di Named Entity Recognition e infine unendo i vari collegamenti dei commenti tramite grafi, applicando le relative tecniche.

1. Raccolta e analisi dei dati

Per comprendere a fondo le opinioni e i sentimenti prevalenti degli utenti riguardo alla situazione del mercato del gaming nel 2024, abbiamo deciso di raccogliere dati dalla community di Reddit, in particolare dalla subreddit r/gaming, il quale contiene più di 40 milioni di utenti. Reddit è noto per essere un luogo dove gli appassionati di videogiochi discutono vivacemente di una vasta gamma di argomenti, rendendolo una fonte preziosa di informazioni dirette e non filtrate.

1.1 Raccolta dei Dati dalla Community di Reddit

Per raccogliere i dati abbiamo utilizzato il sito Apify, una piattaforma di web scraping, la quale, attraverso l'url di un post, riesce a raccogliere tutti i commenti e le informazioni di essi. Il nostro obiettivo era ottenere un campione rappresentativo delle discussioni riguardanti i licenziamenti nell'industria dei videogiochi. A tal fine, abbiamo scelto sette post specifici all'interno della subreddit r/gaming, ognuno dei quali trattava temi rilevanti come i recenti licenziamenti, l'impatto sulle aziende e le preoccupazioni degli utenti.

I titoli dei sette post selezionati coprono una varietà di aspetti relativi ai licenziamenti, tramite i seguenti titoli ai quali segue una breve descrizione:

1. Is the gaming industry on decline or is it getting better?[5]
2. The Video Game Industry is NOT Collapsing. A Lawyer Explains. [8]
3. What's a good video game that made the industry worse? [4]
4. Riot Games Announces Layoff of 530 Positions, 11% of Workforce [1]
5. Why are so many gaming industry employees getting laid off? [6]
6. Abusive privacy policies in video games a grim sign for industry [2]
7. Layoff in the Gaming Industry [3]

Attraverso Apify, siamo riusciti a raccogliere circa 3000 commenti dai sette post selezionati. Questo processo ci ha permesso di ottenere una grande quantità di dati in modo efficiente e sistematico.

1.2 Analisi e pulizia del dataset

A questo punto ci troviamo con 7 dataset diversi, uno per post, quindi li uniamo in un unico dataframe che comprenderà 3306 righe e 24 variabili; molte variabili risultano inutili ai fini della nostra ricerca e quindi andiamo a togliere le variabili con l'url e l'html del post, una variabile che nomina da che community è stato preso il commento(nel nostro caso sono stati presi tutti dalla stessa) ed altre che non portano nessuna informazione, ritrovandoci, infine, con le seguenti 9 variabili:

- body: testo del commento
- createdAt: momento esatto della creazione del commento/post (in formato yy-mm-dd-h-m-s)
- dataType: indica se si tratta del post o di un commento
- comment_id: id del commento
- numberOfreplies: numero di risposte
- parent_id: id del commento/post a cui si è risposto (gli utenti possono rispondere anche ad altri commenti, non solo ai post)
- postId: id del posto sotto al quale si trova il commento
- upVotes: numero di like al commento/post
- username: username dell'utente

A questo punto vogliamo capire quali siano gli utenti più influenti all'interno del dataset, ciò possiamo farlo verificando gli utenti con più Upvote e più risposte ai loro commenti:

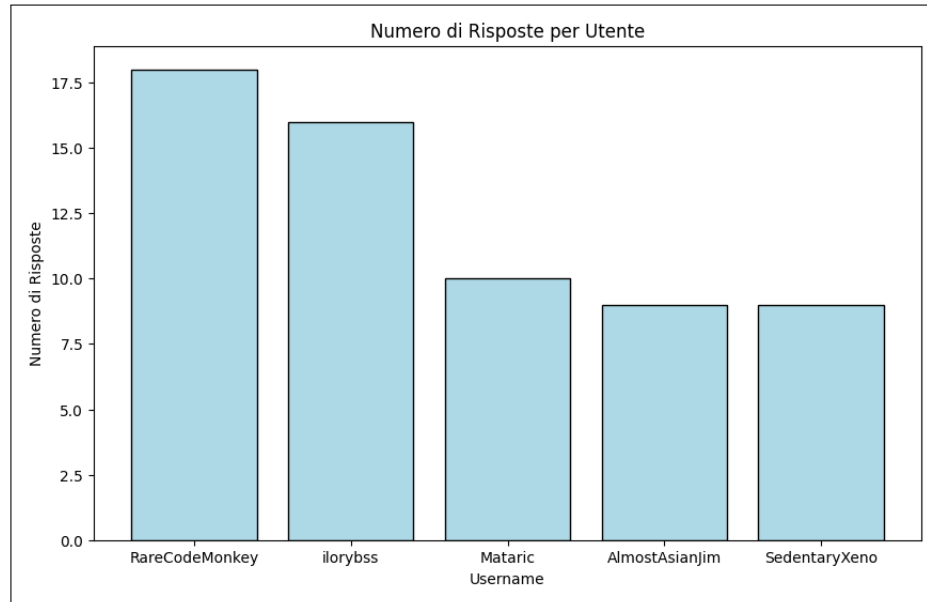


Figura 1.1: Numero di risposte per utente

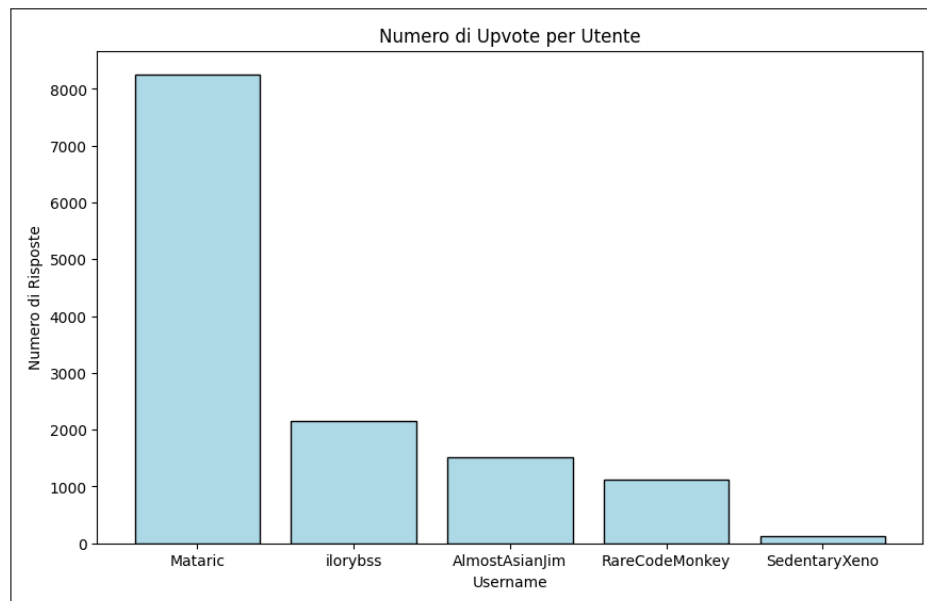


Figura 1.2: Numero di upvote per utente

Possiamo subito notare che i 5 utenti si ripetono in entrambe le immagini, verificando quindi di avere un alta influenza all'interno dei post; rimane sicuramente impresso l'utente Mataric il quale ha ricevuto 10 risposte e, soprattutto, 8000 mi piace al suo commento,

andiamo a vedere il commento da più vicino:

Il commento risponde alla terzo post, il quale chiede quali sono i giochi più belli che però hanno reso l'industria videoludica peggiore.

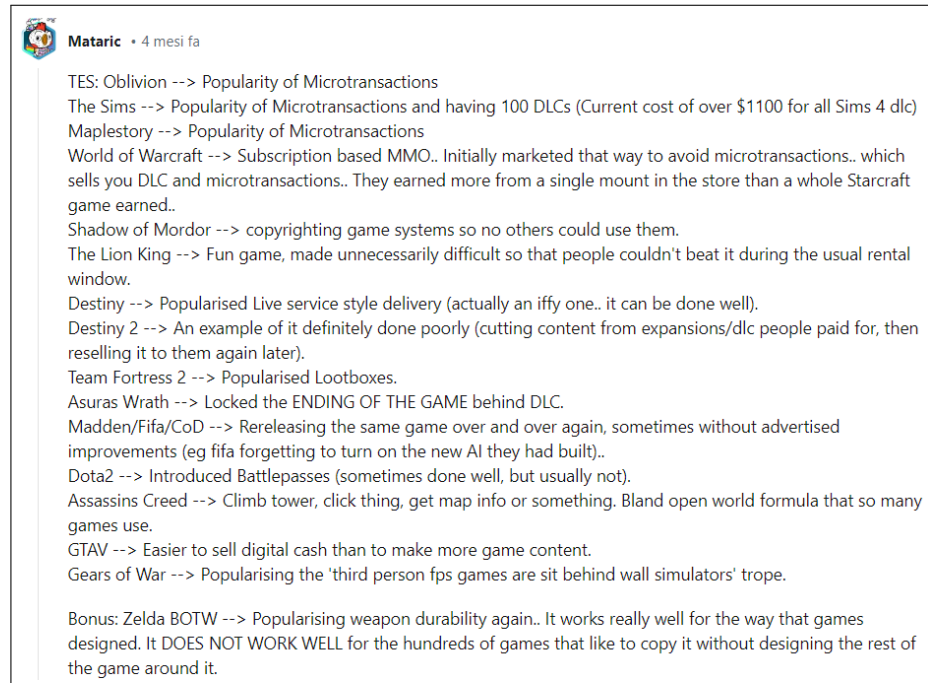


Figura 1.3: Commento dell'utente con più upvote

L'utente in questione offre un elenco dei giochi che, dal suo punto di vista, hanno peggiorato l'industria, tra i giochi che nomina la maggior parte di essi offre microtransazioni nel gioco, ovvero degli acquisti che, ad un basso prezzo, consentono di ottenere dei vantaggi, come dei potenziamenti o qualche miglioramento.

1.3 Preprocessing dei testi

In seguito, siamo passati alla pulizia dei testi per prepararci ad affrontare i task successivi, ovvero sentiment e NER.

Inizialmente sono state rimosse tutte le righe che presentavano commenti cancellati o ri-

mossi dall'autore del post, successivamente è stata controllata la presenza di valori mancanti. Poi si è passati alla rimozione di:

- HTML tags
- @-mentions
- # hash-tags
- numbers

In seguito sono stati resi tutti i caratteri minuscoli e abbiamo eliminato tutte le stopwords, ovvero parole che gli algoritmi considerano poco significative perché possono essere usate spesso all'interno delle frasi e, prese singolarmente, non hanno significato; ad esse abbiamo tolto le negazioni, le quali ci servono per effettuare una buona sentiment analysis.

A questo punto abbiamo eliminato i segni di punteggiatura ed abbiamo verificato se tra le 100 parole più presenti all'interno del dataframe ci fosse qualcosa di sbagliato o "strano".

Adesso siamo pronti per passare alle analisi dei sentimenti e named entity recognition.

2. Sentiment Analysis e NER

2.1 Sentiment Analysis

La sentiment analysis, o analisi del sentiment, è una tecnica di elaborazione del linguaggio naturale (NLP) che consiste nel determinare l'orientamento emotivo di un testo. Questa tecnica può essere utilizzata per analizzare opinioni, valutazioni, emozioni o atteggiamenti espressi in un testo.

La sentiment analysis è il processo di identificazione e categorizzazione delle opinioni espresse in un testo, con l'obiettivo di determinare se l'atteggiamento dell'autore verso un argomento specifico è positivo, negativo o neutro.

Viene applicata in numerosi ambiti come marketing e analisi dei social media, assistenza clienti, politica e ricerca di mercato.

Nel contesto della presente ricerca, abbiamo impiegato l'algoritmo "cardiffnlp/twitter-roberta-base" per condurre un'analisi del sentiment sui commenti raccolti. Questo modello, basato su una versione pre-addestrata di RoBERTa specificamente addestrata su dati provenienti da Twitter, è stato scelto per la sua capacità di gestire testi informali e brevi, tipici dei social media.

L'obiettivo principale di questa analisi era comprendere l'atteggiamento generale degli utenti nei confronti del tema in esame, identificando se i commenti espressi fossero principalmente positivi, negativi o neutri. Questa valutazione del sentiment è stata essenziale per trarre conclusioni significative sulle opinioni degli utenti e per identificare eventuali tendenze o pattern nelle risposte raccolte.

Per condurre l'analisi del sentiment, abbiamo utilizzato il modello "cardiffnlp/twitter-roberta-base" insieme alla libreria Hugging Face per l'elaborazione del linguaggio naturale (NLP). In particolare, abbiamo utilizzato la funzione softmax per ottenere le probabilità predette per ciascuna classe di sentiment (positivo, negativo, neutro) e abbiamo selezionato la classe con la probabilità più alta come etichetta di sentiment per ciascun commento. Successivamente, abbiamo integrato le etichette di sentiment ottenute nel nostro processo

di analisi dei dati, consentendoci di esplorare e comprendere più approfonditamente l'opinione degli utenti rispetto al tema trattato. Questa metodologia ci ha fornito preziose informazioni sull'atteggiamento prevalente degli utenti, consentendoci di contestualizzare e interpretare meglio i risultati della nostra ricerca.

In conclusione, l'utilizzo dell'algoritmo "cardiffnlp/twitter-roberta-base" per analizzare il sentiment dei commenti è stato cruciale per il nostro studio, consentendoci di ottenere una panoramica completa delle opinioni degli utenti e di arricchire la nostra comprensione del tema in esame.

Nell'output abbiamo ricevuto l'etichetta del commento ed un valore che va da 0 ad 1 che ci indica la sicurezza con cui il commento viene associato al sentiment.

A questo punto, attraverso il seguente grafico a torta possiamo verificare la distribuzione delle label:

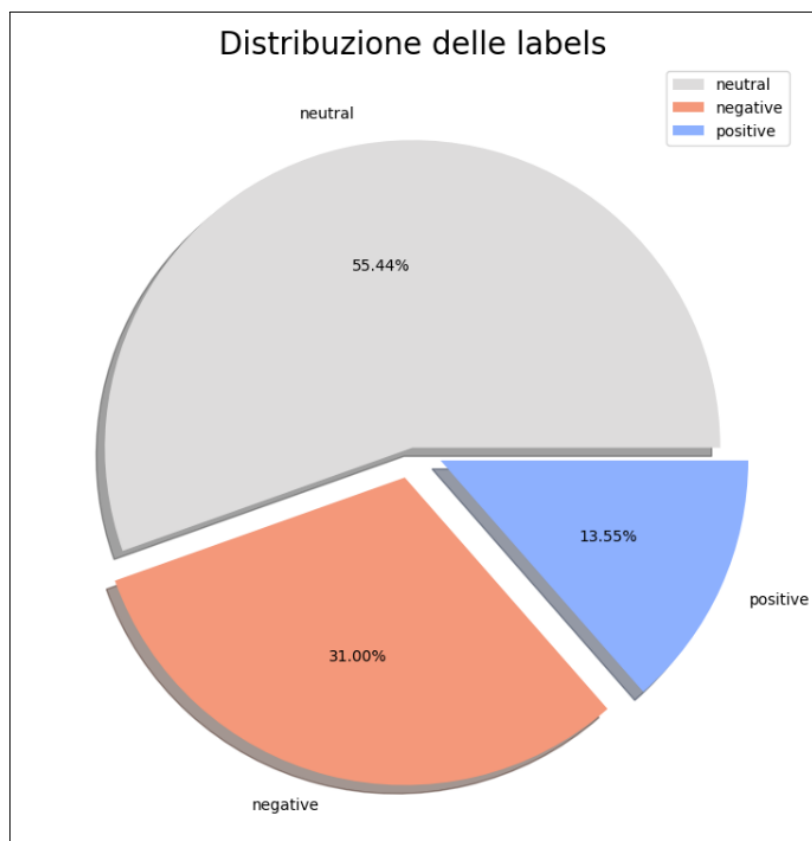


Figura 2.1: Distribuzione delle label

Già da questo grafico possiamo capire come ci siano molti più commenti negativi

rispetto a quelli positivi, implicando quindi un feedback negativo riguardo alla situazione attuale del mondo del gaming, in totale, sono stati raccolti 1833 commenti neutrali, 1025 negativi e 448 positivi.

Passando ad analisi più complesse abbiamo studiato la distribuzione dei sentiment per ogni post, la distribuzione dell'appartenenza dei commenti alla label ed abbiamo anche verificato il cambiamento dei sentiment nel tempo.

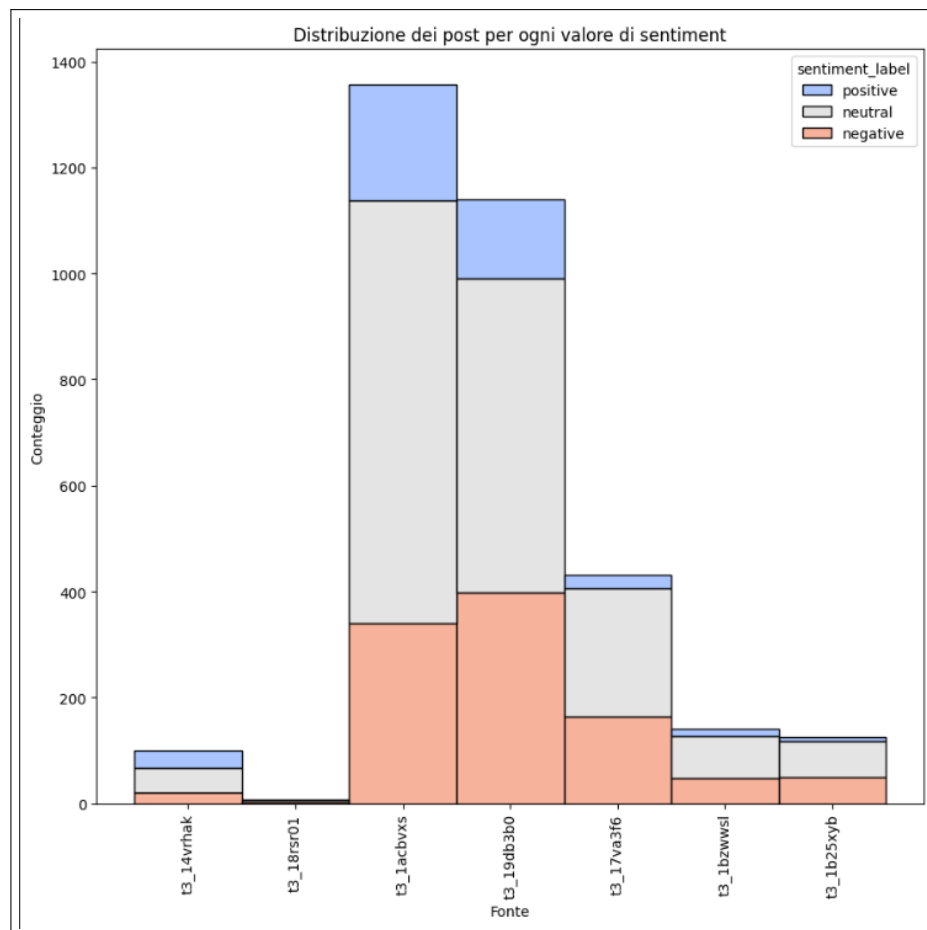


Figura 2.2: Distribuzione del sentiment per post

Da questo grafico possiamo vedere la quantità di sentiment positive o negative relative ad ogni post, il terzo ed il quarto post rappresentano la maggior parte dei nostri commenti, in particolare il terzo post è quello con più commenti negativi, più di 400; in generale vediamo che ogni post ha una quantità piccola di commenti positivi rispetto ai neutrali ed ai negativi (tranne il primo) e si può notare specialmente negli ultimi tre post.

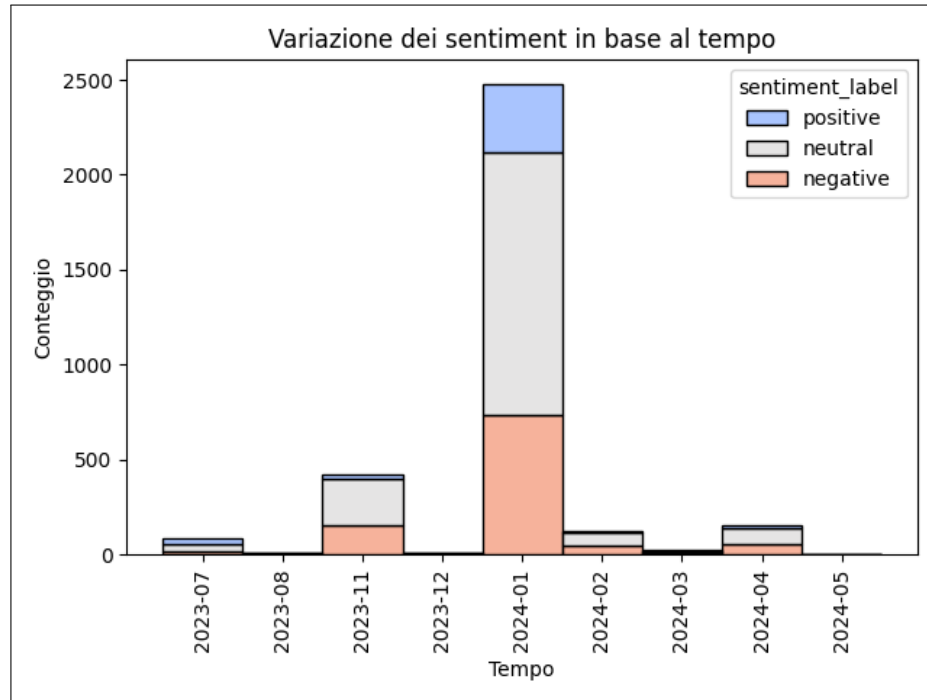


Figura 2.3: Variazione del sentiment del mese

Successivamente, è molto interessante vedere la variazione del numero di commenti e dei valori di sentiment per ogni mese. L'unico mese in cui risultano esserci commenti più positivi è il primo (19 negativi contro 33 positivi), da lì in poi i commenti negativi sono sempre maggiori rispetto ai positivi, questo può esser spiegato dal fatto che, secondo wikipedia [7], il mese di Luglio 2023 è stato il mese con meno licenziamenti dal 2023 fino ad oggi insieme al mese di Dicembre, il quale, difatti, risulta avere solo 8 commenti. Il mese in cui gli utenti sono stati più attivi è il mese di Gennaio 2024, nel quale, come detto in precedenza, sono stati annunciati annunciati circa sei mila licenziamenti per il mese stesso e circa altri quattro mila da effettuare entro fine anno. Questo ha portato quindi gli utenti a parlare dell'argomento come mai prima.

Vedendo questi dati, siamo incuriositi dai commenti che, secondo il nostro algoritmo, risultano con più appartenenza al sentiment negativo durante il primo mese del 2024:

index	body	createdAt	dataType	numberOfReplies	upVotes	username	sentiment_label	sentiment_value
1912	riot sucks ass clown company	2024-01	comment	0.0	-1	hope_they_died	negative	0.979
200	reddit certainly got worse soon acquiring ft2 fucking periwinkles ruined everything	2024-01	comment	1.0	56	redpanda eater	negative	0.9789
613	fortnite battle pass shit ruined everything	2024-01	comment	0.0	1	bootlegportalfluid	negative	0.9781
1016	fuckin shit ow2 day long many horrible things blizzard done understanding ow genuinely die hill clone fundamentally different ways	2024-01	comment	0.0	3	Poopeefighter2001	negative	0.9761
141	fortnite battle passes suck	2024-01	comment	0.0	2	Fingey	negative	0.975

Figura 2.4: Commenti negativi di Gennaio 2024

Da questi commenti possiamo notare come gli utenti siano infuriati con il gioco Fortnite, il quale contiene un "battlepass" che, attraverso un piccolo pagamento, consente di ottenere personaggi che saranno disponibili solo per un breve periodo di tempo. A questo punto vogliamo vedere quanto è sicuro il nostro algoritmo associare un commento ad una determinata label, lo vediamo attraverso il seguente grafico:

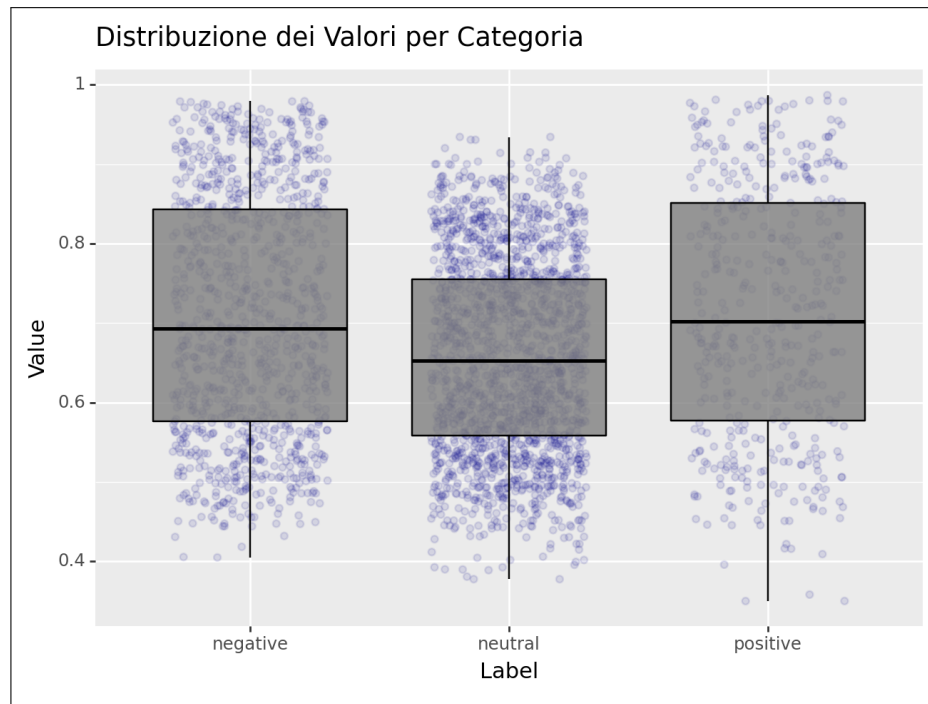


Figura 2.5: Appartenenza alla label

Ci fornisce una chiara visualizzazione delle distribuzioni dei valori per ciascuna categoria di sentiment, mostrando che le differenze tra le categorie non sono drastiche e che la maggior parte dei dati è concentrata in un intervallo ristretto. I dati positivi e negativi risultano avere qualche commento in più vicino al valore massimo.

Infine andiamo a vedere, in generale, il commento che con più sicurezza viene classificato

come positivo ed anche quello negativo:

Il commento più negativo risulta essere "profiteering corporate greed fuckin sucks need like" con un valore di appartenenza pari a 0,9792.

Il più positivo è invece "really loved br mode fh4 favorite part game", con 0,9869 di appartenenza.

Adesso possiamo passare alla seconda task del nostro studio, la Named Entity Recognition.

2.2 Named Entity Recognition

Named entity recognition (NER) è il task che mira a identificare e successivamente classificare entità (sequenze di token) nel testo in categorie predefinite come:

- i nomi di persone
- organizzazioni
- luoghi
- espressioni temporali
- quantità
- valori monetari
- percentuali
- ecc...

In questo progetto, l'uso della Named Entity Recognition (NER) ha diversi scopi specifici che possono arricchire e approfondire l'analisi del feedback degli utenti.

Ad esempio, un utilizzo potrebbe essere quello di identificare le aziende coinvolte, andando a determinare quali aziende di videogiochi sono menzionate frequentemente nei commenti e nei post degli utenti. Questo ci aiuta a capire quali aziende sono maggiormente sotto il riflettore degli utenti e possono essere coinvolte nei licenziamenti o percepite

negativamente. Inoltre, possiamo verificare quali siano i giochi più nominati e coinvolti, in modo da comprendere quali giochi o prodotti sono associati ai licenziamenti o a feedback negativi, aiutando a identificare possibili cause legate a specifici prodotti.

Abbiamo utilizzato un modello prealllenato di Hugging Face. `dbmdz/bert-large-cased-finetuned-conll03-english` è una versione pre-addestrata di BERT (Bidirectional Encoder Representations from Transformers), che è stata ulteriormente addestrata specificamente per il compito di Named Entity Recognition (NER) utilizzando il dataset CoNLL-2003. Il modello è pre-addestrato in maniera bidirezionale, cioè considera il contesto di una parola guardando sia a sinistra che a destra, migliorando così la comprensione del significato delle parole nel loro contesto.

”Large” si riferisce alla dimensione del modello BERT, che in questo caso è la versione più grande con 24 strati (layers), 16 teste di attenzione (attention heads) e 340 milioni di parametri.

”Cased” significa che il modello distingue tra maiuscole e minuscole. Ad esempio, ”Apple” (l’azienda) e ”apple” (il frutto) sono trattati come entità diverse.

Nell’output del modello di Named Entity Recognition (NER) fornito da Hugging Face, ogni entità riconosciuta è rappresentata da un dizionario che contiene diversi parametri.

Ogni dizionario è composto in questo modo:

- **Entity:** Indica il tipo di entità riconosciuta e la posizione della parola all’interno dell’entità.
Valori Possibili: I-ORG (parte interna di un’organizzazione), B-ORG (inizio di un’entità organizzazione), I-LOC (parte interna di un luogo), B-LOC (inizio di un’entità luogo), I-PER (parte interna di un nome di persona), B-PER (inizio di un’entità persona), ecc.
- **Score:** Il punteggio di confidenza del modello sulla classificazione dell’entità.
Valori Possibili: Un numero tra 0 e 1, dove valori più vicini a 1 indicano una maggiore confidenza.
- **Index:** La posizione del token all’interno della sequenza di token del testo di input.
Valori Possibili: Numeri interi positivi.

- **Start:** La posizione iniziale (offset) del primo carattere dell'entità riconosciuta nel testo originale.

Valori Possibili: Numeri interi non negativi che indicano la posizione del carattere.

- **End:** La posizione finale dell'ultimo carattere dell'entità riconosciuta nel testo originale.

Valori Possibili: Numeri interi non negativi che indicano la posizione del carattere.

- **Word:** Il testo esatto della parola o sequenza di parole che sono state riconosciute come entità.

Valori Possibili: Qualsiasi stringa di caratteri che rappresenta una parola o sequenza di parole.

A questo punto proseguiamo con le due analisi che ci eravamo posti all'inizio del task; per prima cosa, estraiamo tutte le Org trovate e andiamo a verificarne il sentiment per quelle più volte citate all'interno dei commenti:

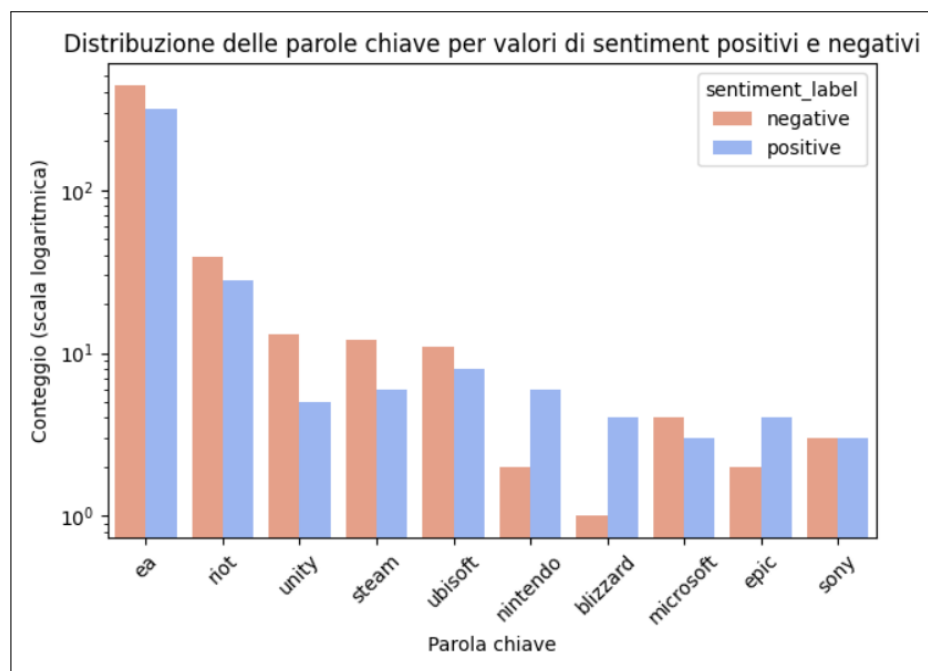


Figura 2.6: Sentiment delle aziende

Il grafico mostra la distribuzione delle parole chiave per valori di sentiment positivi e negativi. Le parole chiave rappresentano diverse aziende di videogiochi e sono posiziona-

te sull'asse delle ascisse (x), mentre l'asse delle ordinate (y) rappresenta il conteggio delle occorrenze di queste parole chiave in una scala logaritmica. I valori positivi sono indicati in blu, mentre i valori negativi sono indicati in arancione. L'uso di una scala logaritmica suggerisce che ci sono grandi variazioni nel numero di menzioni tra le diverse aziende, con alcune aziende che ricevono molte più menzioni di altre.

Per molte aziende, il sentiment negativo è leggermente superiore a quello positivo, indicando una generale insoddisfazione degli utenti nelle discussioni online.

Nintendo e Blizzard sono le uniche due aziende con una netta prevalenza di sentiment positivo, suggerendo una percezione più favorevole da parte degli utenti.

Adesso, realizziamo il medesimo grafico per i giochi più nominati:

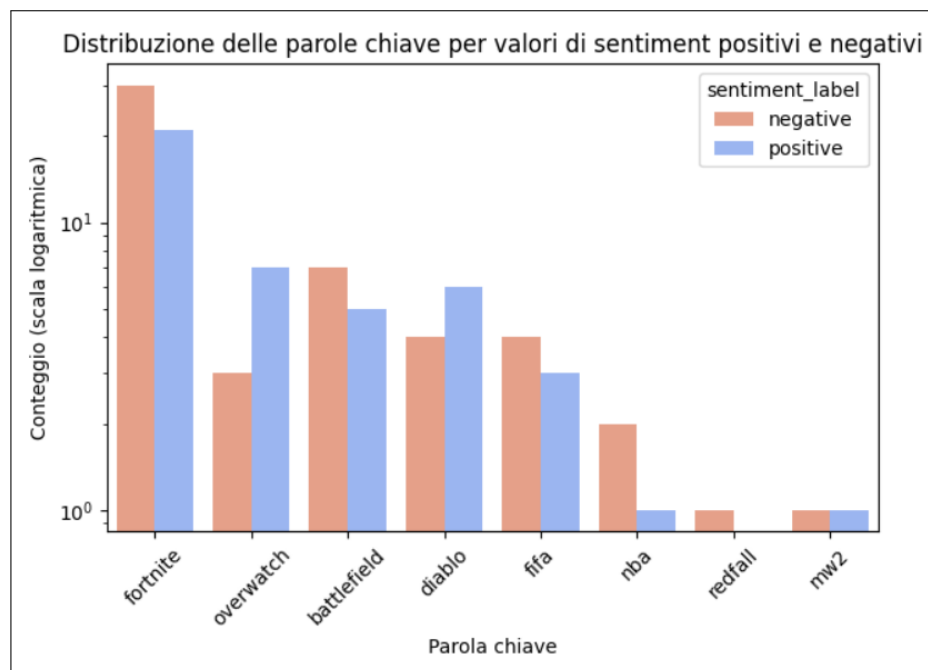


Figura 2.7: Sentiment delli giochi più nominati

Per molti giochi, il sentiment negativo è superiore a quello positivo, suggerendo una generale insoddisfazione o critiche da parte degli utenti.

Diablo ed Overwatch sono gli unici giochi che mostrano una netta prevalenza di sentiment positivo, suggerendo una percezione più favorevole da parte degli utenti.

3. Graph Analysis

3.1 Introduzione

Questa sezione finale si concentrerà sull'analisi delle connessioni tra gli utenti, basata sui commenti raccolti da tre post, ognuno dei quali contiene circa cento commenti. L'obiettivo è esaminare le interazioni tra gli utenti per identificare pattern di connessione e dinamiche di comunicazione.

In oltre, verrà utilizzata la sentiment analysis effettuata precedentemente, per provare a capire cosa ne pensa la community di **Reddit** riguardo il primo post e il sesto post; quindi gli utenti che dicono la propria sui seguenti argomenti: *"Is the gaming industry on decline or is it getting better?"* e *"Abusive privacy policies in video games a grim sign for industry"*. Per comodità verranno definiti come primo e secondo post.

3.2 Primo post

Verranno analizzati, ora, tutti i commenti riferiti al primo post, quindi vedremo cosa ne pensano le persone riguardo all'andamento dell'industria dei videogames.

Nel grafo, il nodo in **verde acqua** sarà quello relativo allo scrittore del post, mentre tutti gli altri in **blu** appartengono agli altri. Per quanto riguarda gli archi, invece, quelli colorati di **rosso chiaro** corrispondono a tutti i commenti riferiti al post, mentre tutti gli altri, quelli in **grigio** si riferiscono ai messaggi tra utenti.

3.2.1 Visualizzazione del grafo

In questo primo paragrafo, verrà visualizzato il grafo per intero e verranno presentate delle idee preliminari.

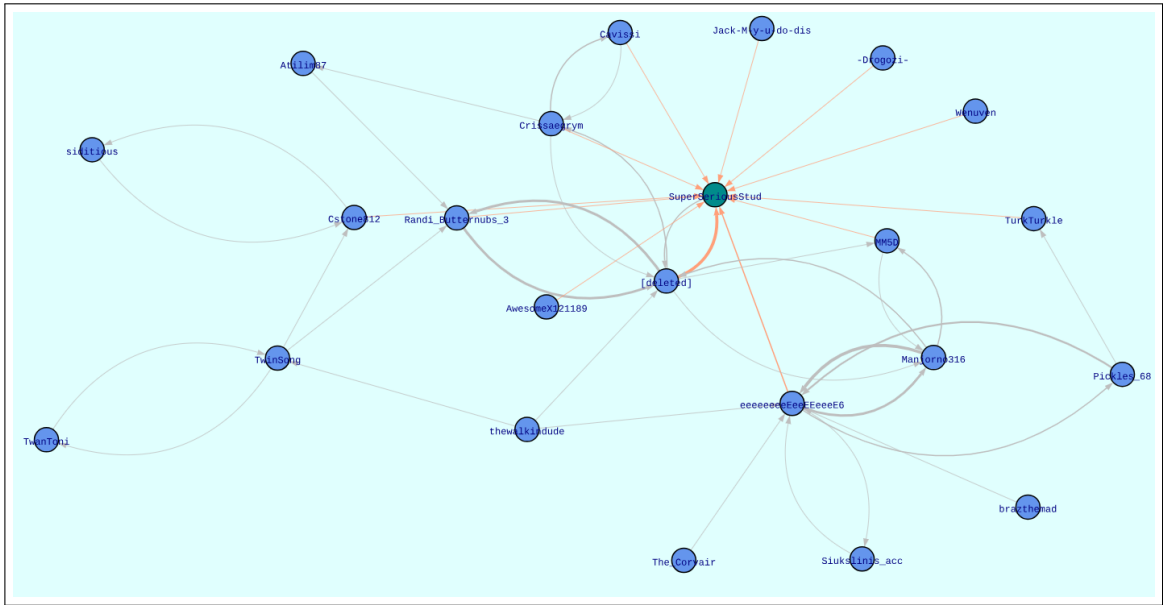


Figura 3.1: Grafo primo post

Dalla figura 3.1 si possono già vedere alcune interazioni particolari. Per esempio tra gli utenti **eeeeeeeeEeeEeeeeE6**, **Manjorno316** e **Pickles_68**. Anche tra gli utenti **Ran-di_Butternubs_3** e **deleted** (questo in particolare si chiama così perchè il suo account non esiste più; può essersi eliminato lui stesso, oppure essere stato bannato). Per quanto riguarda il post, ci sono 12 archi diretti esclusivamente verso **SuperSeriusStud**, circa il 16,9%.

3.2.2 Analisi temporale

In questa sezione ci occuperemo di osservare il comportamento degli utenti da un punto di vista temporale; quindi, vedere, come col passare del tempo è variata la frequenza di commenti.

Nel seguente grafico, viene mostrato l'istogramma che mostra la frequenza di messaggi di ora in ora:

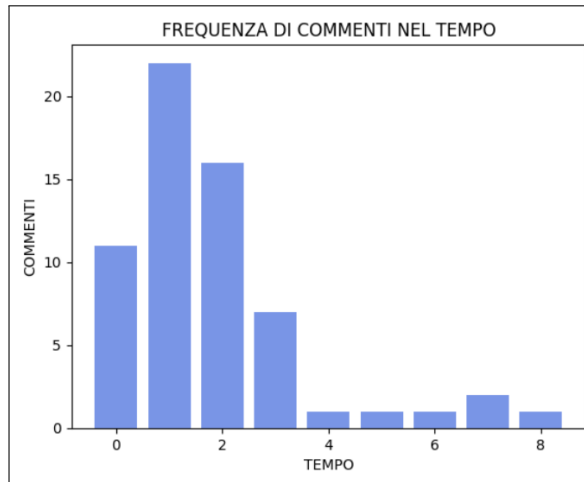


Figura 3.2: Frequenza vs ore

Come ci si può facilmente aspettare, si ha un picco di messaggi nelle prime ore. Dopo di che la coda destra¹ si abbassa velocemente.

3.2.3 Grado dei Nodi

Andiamo a vedere chi sono, anche se prima è stato accenato, gli utenti che hanno avuto più interazioni.

Col seguente grafico andremo a mostrare il *numero di archi entranti* (commenti ricevuti) con il *numero di archi uscenti* (commenti scritti).

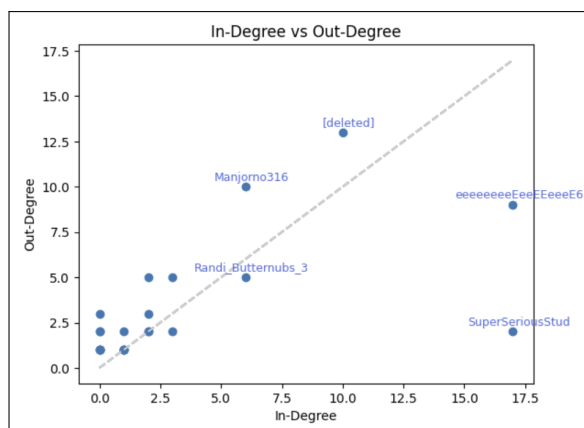


Figura 3.3: Indegree vs Outdegree

¹**NB:** la coda, per rendere leggibile il grafico, è stata tagliata. Nel grafico originale si sarebbe protratta per qualche altra ora ma con frequenze basse.

Dalla figura 3.3, vediamo che gli utenti citati precedentemente sono effettivamente curiosi da studiare.

Gli utenti che hanno ricevuto più messaggi sono `eeeeeeeeEeeEEeeeE6` e `SuperSeriousStud` che sono entrambi a quota 18. Per quanto riguarda l'outdegree, colui che ha il numero più alto, è `deleted`.

Verrà, ora, mostrato un istogramma contenente il numero di messaggi scambiati per ogni coppia di utenti:

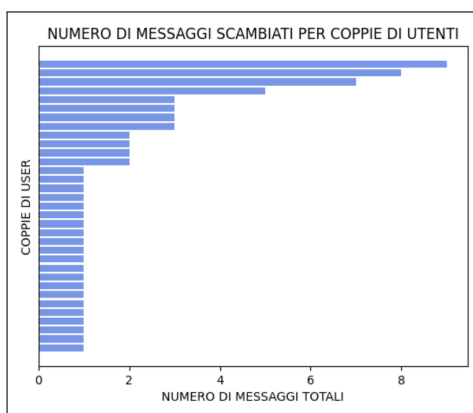


Figura 3.4: Frequenza dei messaggi per ogni coppia di utenti

Come si può notare, solo 4 coppie di user si sono scambiate più di 4 (quindi almeno un "botta e risposta") messaggi totali; vediamo chi sono.

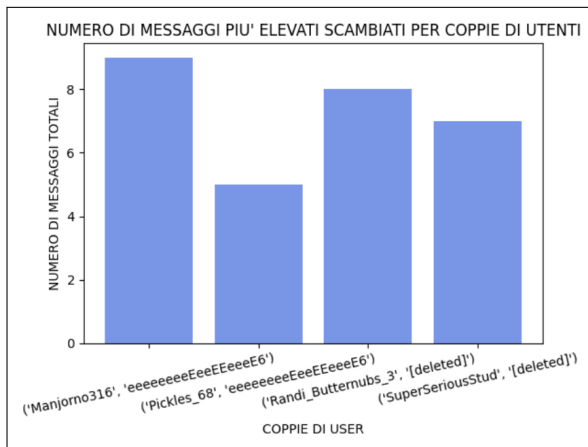


Figura 3.5: Coppie con più messaggi scambiati

Si nota che le coppie contenenti l'utente **eeeeeeeeEeeEEeeeE6** sono quelle con un maggior numero di messaggi scambiati. Potrebbe significare che lui abbia discusso animatamente con loro. Nella prossima sezione verranno analizzate proprio le discussioni di queste 4 coppie di utenti.

3.2.4 Analisi discussioni

Analizziamo la discussione più lunga che abbiamo, ovvero quella tra **Manjorno316** e **eeeeeeeeEeeEEeeeE6**:

Indice	Username	Sentiment	UpVote	Testo
0	Manjorno316	positive	1	for every tears of the kingdom there are 4-5 battlefield 2042s. Could you name me 4-5 battlefield 2042s from the past year? It hasn't been close to that many per good
1	eeeeeeeeEeeEEeeeE6	neutral	1	gollem, redfall, forsoken, overwatch 2. battlefield 2042. Diablo immortal. MW2. of late 2022-2023. so far.
2	Manjorno316	neutral	1	Yeah those are the ones I can think of as well. Which doesn't put us close to 4-5 bad games per good game.
3	eeeeeeeeEeeEEeeeE6	neutral	1	you asked for 4-5. you want me, to categorically list, every single good game, and every single bad game, at that ratio. that would take hours. and I'm not that pettv. take
4	Manjorno316	negative	0	I never asked you to do that. Don't even know what part of what I wrote in my second comment made you think I asked you to do that? You got me all confused here.

Figura 3.6: Discussione tra Manjorno316 ed eeeeeeeeEeeEEeeeE6

Possiamo vedere dai commenti come non ci sia un vero e proprio litigio in questo caso, il primo commento risponde alla domanda dell'autore del post in cui veniva chiesto se l'industria del gaming è in una situazione di declino o meno. L'utente risponde che ad ogni uscita di un gioco carino ne corrispondono almeno 4-5 pessimi e chiede, agli altri utenti, di fornire un elenco di questi giochi pessimi che escono annualmente; il secondo utente risponde con una lista di giochi che ritiene non adatti. L'unico accenno di litigio lo troviamo nell'ultimo commento, in cui Manjorno316 risponde che il commento dell'utente precedente lo ha confuso e non corrispondeva alla sua richiesta, a questo commento

non seguono risposte.

Ora passiamo alla discussione tra **Pickles_68** e **eeeeeeeeEeeEeeE6**:

Indice	Username	Sentiment	UpVote	Testo
0	Pickles_68	neutral	1	Didn't know art could release broken
1	eeeeeeeeEeeEeeE6	neutral	1	does if you break the frame.
2	Pickles_68	neutral	1	How does one break a digital frame
3	eeeeeeeeEeeEeeE6	neutral	1	digitally
4	Pickles_68	neutral	1	With water

Figura 3.7: discussione tra Pickles_68 e eeeeeeeeEeeEeeE6

In questo caso la discussione nasce per rispondere al commento di **Siukslinis_acc**, il quale scrive che le edizioni speciali non offrono nulla in più rispetto alle edizioni classiche, la discussione che ne segue è ironia verso il fatto che i giochi vengano pubblicati rotti.

Passiamo a quella tra **Randi_Butternubs_3** e **deleted**:

Indice	Username	Sentiment	UpVote	Testo
0	[deleted]	positive	6	A better year than we've had in a long time, but "one of the best"?
1	Randi_Butternubs_3	neutral	1	Totally, and I've been gaming longer than most redditors have been alive.
2	[deleted]	positive	2	Well, that's your opinion. It's a good year.
3	Randi_Butternubs_3	positive	0	Everything about gaming is subjective. But glad it makes you feel better to say the obvious.
4	[deleted]	negative	1	Grumpy aren't we. All you did was say "i'm old and I think this". You didn't give any actual reasoning behind your statement. But I'm not going to sit here and say "give me a well written out essay documenting your findings on reddit snaarf." Cranky old man.

Figura 3.8: Discussione tra Randi_Butternubs_3 e deleted

La discussione parte a seguito di un commento di Randi_Butternubs_3 verso il post principale in cui spiega che secondo lui il 2023 è stato un anno tra i migliori di sempre per il gaming. Viene seguito una risposta in cui viene scritto che si tratta di un buon anno, ma assolutamente non uno tra i migliori; Randi risponde che ne è certo dato che è un giocatore da più tempo della maggior parte degli utenti di reddit. Infine, l'utente deleted risponde in modo più cattivo, confermato dalla sentiment con label negative, dando dello scorbutico e del vecchio all'altro utente, rimproverandolo di non aver dato motivazioni alla sua tesi.

Passiamo a quella tra **deleted** e **SuperSeriousStud**

Indice	Username	Sentiment	UpVote	Testo
0	[deleted]	neutral	0	I dont think its in decline, but the quality of games certainly is
1	SuperSeriousStud	neutral	1	The average is. But the best games of 2023 are worse than those of 2022? of 2013?
2	[deleted]	positive	-2	I have legit not played a game released after 2013 that would truly captivate me and have me spend tons of hours with and want to return, graphics are better sure, but its usually one playthrough and im like "eeh it was okay" but nfsu, dragon age origins, kkend2, empire earth, silent hill 1-3, resident evil 1-3 plus veronica and outbreaks, diablo 2, the longest journey series, and many more are games i keep replaying sinking hundreds if not thousand of hours into. Newest game i can think of that was somewhat fun was AC odyssey, think i spent about 120h with it, but again, didnt even finish the story and i dont feel like ever revisiting it again.
3	SuperSeriousStud	positive	2	Last time I checked, Elden ring released after 2013 :) so maybe is not "not any game" but "few games"? Have you played: Cuphead, Hollow knight, Dont starve, Neon White? I specifically mention this ones because it seems to me that you like challenging games.
4	[deleted]	positive	1	Well as i said Elden Ring is an exception, literally the only game after that day that i see myself playing years later. Apart from Neon White, i know the games, Cuphead was okay, Hollow Knight is quite good, but another one of those once is enough, Dont Starve is not my kind of game. After 2013 not all games are bad, there are some good ones, but nothing really that great to replay or give a second thought after being done with it.

Figura 3.9: Discussione tra deleted e SuperSeriousStud

In quest'ultima discussione, non troviamo un litigio, ma uno scambio di opinioni riguardo al declino del mondo del gaming, l'utente deleted crede che la qualità dei giochi sia scesa a picco dal 2013 in poi, non avendo più trovato un gioco che lo abbia attratto abbastanza da spendere troppe ore; ovviamente è anche conscio che la qualità della grafica sia migliorata molto. Alla fine dello scambio i due si trovano d'accordo sul fatto che l'unica eccezione sia Elden Ring.

3.2.5 Analisi dei voti

Dopo aver analizzato le discussioni presenti sotto al primo post, andiamo a vedere quali sono i commenti che hanno riscosso maggior successo all'interno della community e a valutarne la loro posizione in merito. Per farlo useremo il sistema più intuitivo: ci baseremo sugli **upVotes** ovvero i *like*.

Username	Sentiment	UpVote	Testo
TurkTurkle	neutral	19	Neither. I've been gaming since the late 80s and can confidently say Its always been like this. A few islands of quality in a sea of sewage
eeeeeeeeeEeeEeeEe	positive	16	it's weird. gaming is evolving as an industry, and an artform. which is both good, and bad. there are lots of positives and lots of negatives. for every tears of the kingdom there are 4-5 battlefield 2042s. on a super high note. Iarian studios. have decided that every single person who purchased Baldur's gate 3 in early access, gets the full game, and an upgrade to the deluxe edition which includes a decent amount of things, free of charge. not bad for a game I've technically owned for 3 years already. they work so closely with the community to develop something truly special.
[deleted]	positive	6	A better year than we've had in a long time, but "one of the best"?
Jack-M-y-u-do-dis	neutral	4	It's definitely a weird moment. I no longer see massive game studios as trustworthy, quite the opposite infact. Consistently the best made, maintained games I've played either come from indie devs, small studios or the occasional fan service game from a studio that doesn't focus on monetizing every aspect of a game at all times at the cost of gameplay. I am however burnt out on a large part of the market, as the moment I see microtransactions, basic content locked behind DLC, Season Passes and "ultimate editions" of games that cost upwards of \$100, I instantly get the desire to throw my computer put the window and go back to playing on Xbox 360. Another pet peeve of mine is paying for online on console, but there's literally no alternative, since for the same performance PC gaming still requires a higher initial cost of entry, so... Yeah I'm both insanely bummed out that everuthing seems to be i creasongly expensive, but also happy to see a lot of cheap, quickly made but still well polished and FINISHED games be available, even playable on a 10 year old office computer or PS4. Weird times.

Figura 3.10: Commenti con più upVotes nel primo post

3.3 Secondo post

Questa seconda analisi, andrà ad indagare il parere degli utenti sulla questione **micro transazioni** all'interno dei giochi. Il post, infatti, è solo una serie di link che fanno riferimento ad un sito che approfondisce l'argomento.

3.3.1 Visualizzazione del grafo

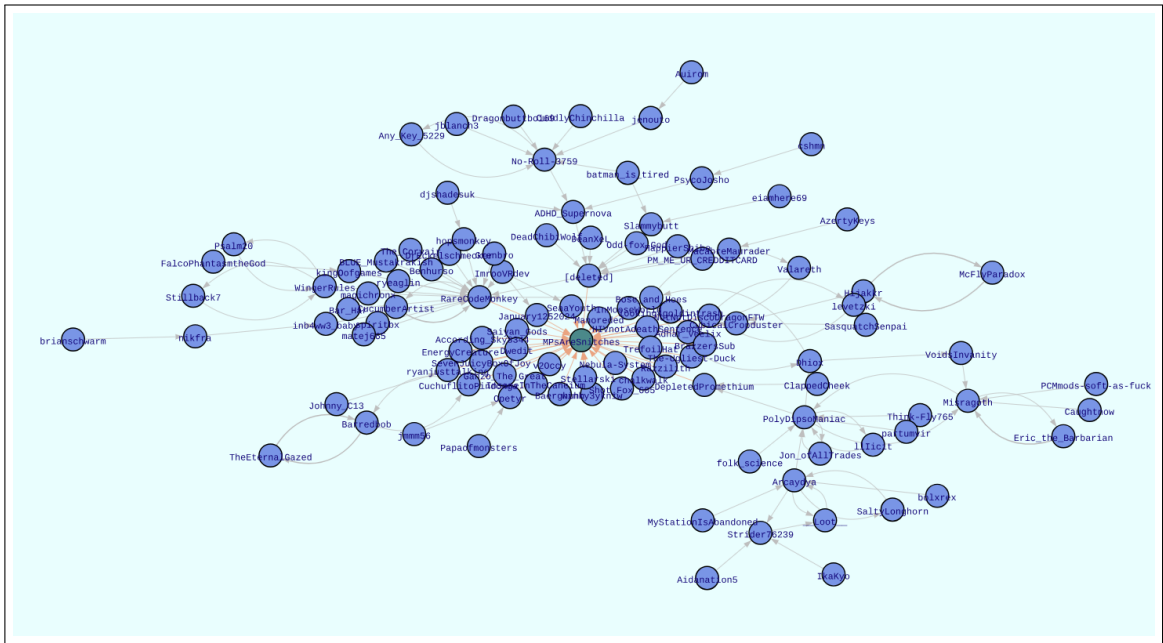


Figura 3.11: Grafo relativo al secondo post

Come possiamo notare dalla figura 3.11, sembra che gli utenti siano sparsi in più *sotto-comunità*. Come analisi preliminare si può dire ben poco, bisognerà approfondire tramite le successive analisi.

3.3.2 Analisi temporale

Iniziamo anche in questo caso con il guardare il post da un punto di vista temporale, quindi vedere come aumenta o diminuisce il numero di messaggi scritti di ora in ora.

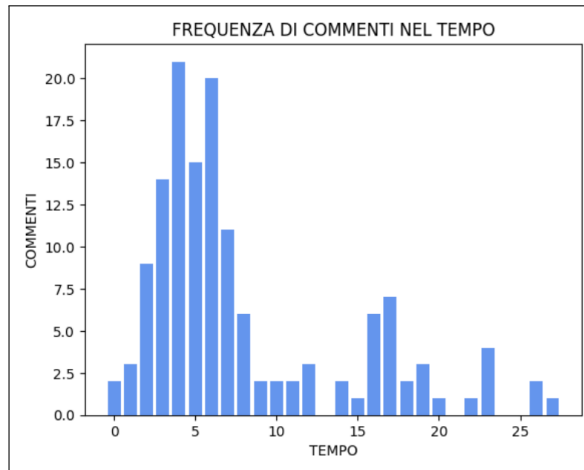


Figura 3.12: Frequenza di messaggi oraria del secondo post

Ci sono due aspetti da analizzare che lo caratterizzano e lo differenziano da quello precedente. In primis si nota subito che la distribuzione delle frequenze è molto diversa e particolare; infatti, inizialmente cresce e scende velocemente (come da normale aspettativa), ma poi sembra avere ancora una piccola affluenza nelle ore successive. In secondo luogo, è da notare che il tempo (in ore) si protrae a lungo.²

3.3.3 Grado dei Nodi

Dalla figura 3.11 si nota che ci sono moltissimi nodi (utenti) che, però, scrivono pochi commenti. Andiamo ad analizzare il grado di ogni nodo tramite il seguente grafico:

²Si consideri anche in questo caso che la coda è stata tagliata. Ovviamente il taglio è stato effettuato in punti dai quali la distribuzione andava sempre più a scendere.

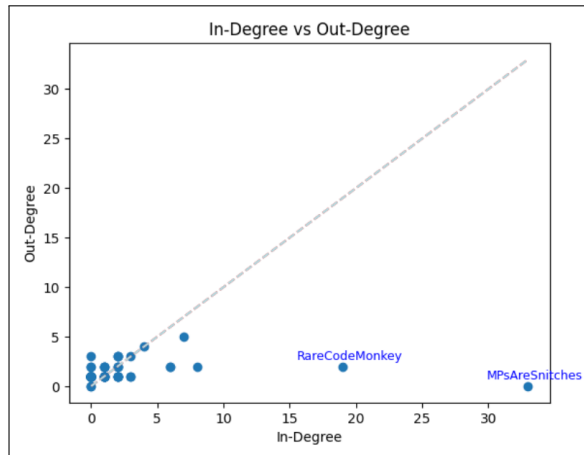


Figura 3.13: Indegree vs Outdegree

Possiamo notare che due utenti si sono distinti dagli altri. **RareCodeMonkey** e **MPsAreSnitches**, infatti, presentano un numero elevatissimo di *indegree*. Tra i due, quello veramente anomalo è il primo; infatti, **MPsAreSnitches** è l'autore del post e quindi ha senso che abbia ricevuto molte risposte.

Per andare a capire meglio cosa ha scatenato la serie di commenti in risposta a **RareCodeMonkey** andiamo a vedere cosa ha scritto.

Indice	UpVotes	Sentiment	numberOfreplies	Testo
0	1113	Negative	18	It is just impossible for the average person to read all the licenses, policies, etc. that come with on-line services, software and games. You could not do that even if it was your full-time job. You "clicked accept to play the game, so, we can do whatever we want" it's out of hand. It should be way more limited what this kind of licenses can include, and make illegal the most outrageous. Even children are accepting these policies. All of that for what? To extract as much money as possible from users. Pricing for virtual items that are carefully crafted to maximize profit and minimize value.
1	3	Negative	1	Yes. This is something that needs to be changed by law. Companies will not change it even if there is people in that companies that care for the issue.

Figura 3.14: Commenti scritti da **RareCodeMonkey**.

Il maggior numero delle risposte le ha ricevute il primo commento. Come possiamo vedere, la sentiment riferita al testo ci indica che è negativo, ed effettivamente si sta lamentando del fatto che leggere tutte le polize da accettare è quasi impossibile e che per disperazione e voglia di giocare si clicchi su **accetta** senza darci peso. In oltre si riferisce anche giocatori più piccoli che molto prbabilmente non andranno mai a leggere le polize poichè solo interessati al gioco.

Questo commento ha riscosso molto successo; ben 1113 "like". Ci fa capire che i "gamers", non sono affatto felici della direzione in cui si sta andando e che sono d'accordo sul fatto che le case produttrici si concentrino più sul guadagno che sul prodotto stesso. A questo proposito, vediamo cosa ne pensano gli altri utenti. Nella seguente tabella verranno mostrati i commenti più votati e più inerenti con l'argomento che si sta trattando.

Indice	UpVotes	Sentiment	numberOfreplies	UserName	Testo
0	403	Negative	18	[deleted]	And you're going to accept. It's not as if declining is a real option because almost every product comes with an equally terrible privacy policy. In order for people to decline there needs to be an alternative
1	49	Neutral	1	hopsmonkey	All of that for what? And also to allow them to shut off access whenever they feel like you should no longer get to play the game you bought.
2	23	Neutral	2	spiritbx	Even children are accepting these policies. Imagine that instead of fixing the problems they just change the law so that children can consent... I wouldn't surprised, the shareholders must be appeased after all.
3	64	Neutral	3	WingerRules	2 things. 1. Its basically the Republicans on the Supreme Court that thinks companies should have no limits of what they can write in contracts. Mandatory individual arbitration and class action waivers was illegal about 10 years ago, then the Republicans on the Supreme Court voted 5-4 on party lines to allow companies to do it. Every democrat justice voted against it. 2. If you're in a multi person residence or you have friends over frequently, theres basically no way for them to verify it was actually you that signed the agreement for a game unless they required something like making an account that requires verification.
4	12	Positive	4	Bose_and_Hoes	Totally agree with all your points. I am obviously an outlier because I write those policies, but I can and do read all ToS and PP on basically all services I use. After years, I can read them super fast.

Figura 3.15: Risposte al commento di **RareCodeMonkey**.

Come possiamo intuire, la community è abbastanza d'accordo e la pensa più o meno allo stesso modo. Nonostante risultino valutazioni positive di sentiment, nessuno vede, quello che stanno facendo le *grandi case di videogames*, in modo positivo.

3.3.4 Community detection

Come accennato in precedenza, il grafo completo sembra possedere molte comunità. Si è quindi deciso di effettuare una community detection tramite algoritmo di clustering per andare a dividere il grafo in *sotto-gruppi*.

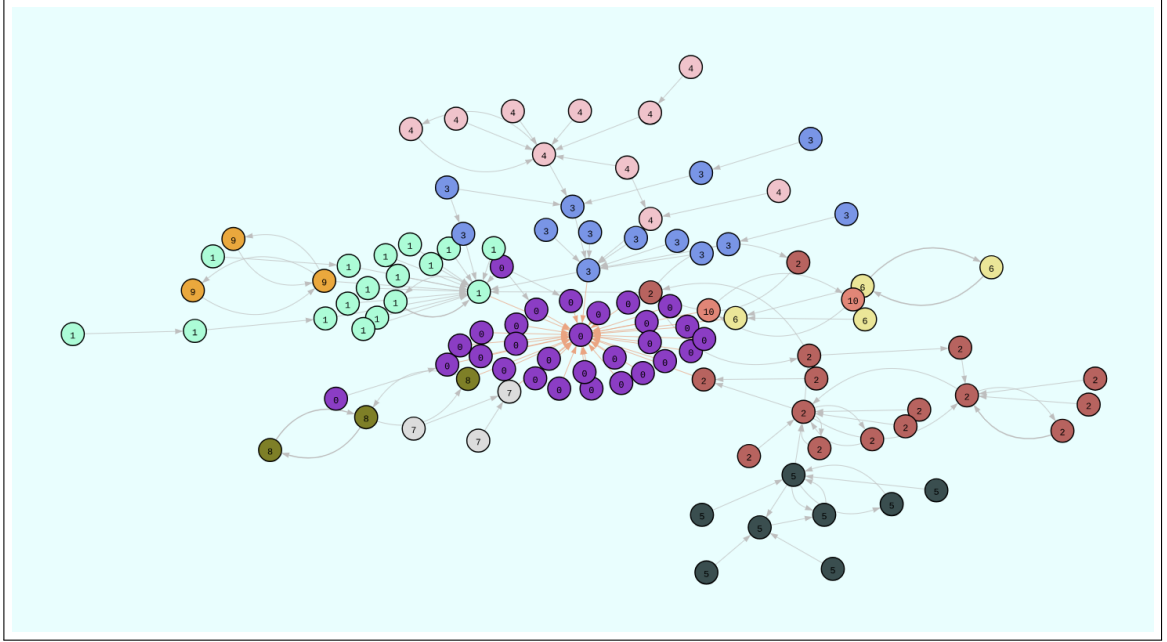


Figura 3.16: Ripartizione dei nodi dopo la community detection con clustering.

Il grafo è stato suddiviso in 11 parti tramite l'algoritmo **Leiden**. Questo algoritmo cerca di massimizzare la **modularità** ovvero quanto le connessioni all'interno delle comunità siano più dense rispetto a quelle tra le comunità. Il calcolo della modularità si effettua con la seguente formula:

$$Q = \frac{1}{2m} \sum_{ij} \left(A_{ij} - \frac{k_i k_j}{2m} \right) \delta(c_i, c_j) \quad (3.1)$$

dove:

- Q è la modularità della partizione.
- A_{ij} è l'elemento i,j -esimo della matrice di adiacenza.
- k_i e k_j sono rispettivamente i gradi dei nodi i e j .
- m è il numero totale di archi nel grafo.
- $\delta(c_i, c_j)$ è un valore che è pari a 1 se i e j sono nella stessa comunità e pari a 0 altrimenti.

Andiamo ora ad indagare i commenti scritti dagli utenti appartenenti ad ogni cluster. Per fare ciò si è deciso di guardare il commento con più risposte di ogni utente presente in ogni cluster per provare a dare un *orientamento* alla community. L'idea è la seguente: *"il commento con più risposte sarà quello che, nel cluster, fa partire una discussione e di conseguenza rappresenta il tema della community"*.

Ecco i risultati:

- **CLUSTER 0:** il post, che non verrà riportato in quanto si limita ad essere una serie di link. Riportiamo solo il titolo: "Abusive privacy policies in video games a grim sign for industry". **MPsAreSnitches**
- **CLUSTER 1:** "it is just impossible for the average person to read all the licenses, policies, etc. that come with on-line services, software and games. You could not do that even if it was your full-time job. You "clicked accept to play the game, so, we can do whatever we want" it's out of hand. It should be way more limited what this kind of licenses can include, and make illegal the most outrageous. Even children are accepting these policies. All of that for what? To extract as much money as possible from users. Pricing for virtual items that are carefully crafted to maximize profit and minimize value." **RareCodeMonkey**
- **CLUSTER 2:** if they don't show ads, what are we calling the pop up that shows current sells and new releases. **Misragoth**
- **CLUSTER 3:** And you're going to accept. It's not as if declining is a real option because almost every product comes with an equally terrible privacy policy. In order for people to decline there needs to be an alternative. **deleted**
- **CLUSTER 4:** i'm an old person, but i got my first fancy graphics card in 2020. pandemic boredom. at this point i'm so disappointed with what the AAA gaming space looks like i've crawled back in to my indie cubbyhole and i don't expect to emerge. **No-Roll-3759**

- **CLUSTER 5:** Yeah I'm scared when Newell dies though. I hope he has a successor ready who he trusts. **Arcaydya**
- **CLUSTER 6:** It is not just games that does this though. It is literally any and all forms of digital service delivery. Went down this rabbit hole for a university project a while back. I dug into the policies of a few common services i used every day (email and phone user ID, online grocery service that allows me to order from my phone and pick up goods ready at store, entertainment streaming site and lastly an IM messaging app). The policies are generally huge and link to other policies. They use generalized jargon and terms that you need to look up unless you have a legal background. One policy generally also allows for data gathering and use by the primary service provider plus also all of its 3rd party affiliates. It took me a long time, but just focusing on those 4 services above i found that some of them were actually related to each other as affiliates (you would never expect it because they as so vastly different from each other). At the end of the day, they had access and record of literally everything that determines my place in this world. Full name, surname, middle name, address, contact details in full, banking details in full, list of bank branches i frequently visit in person, list of stores (and even other stores not affiliated with the primary store) that i frequently visit in person, general descriptions of products i buy, foods i buy, foods i avoid possibly due to allergies, brands i buy frequently, brands i avoid, location where i work, details of my insurance providers, details of my close contacts and their contact info, details of local and even some foreign tourism spots i have visited in the last number of years, details of any medications i have used and may use in future (i think this is speculative based on other bio data). Of course they also had sensitive data like credit card, ID number, drivers license number but those are apparently masked to a greater extent. Most interestingly they also recorded some form of internet browsing history, in addition to device hardware and even software data related to other programs on my devices. It gets weirder when you look into app specific data being captured (distance traveled per day, phone usage per day broken down into subcategories for entertainment/news/-commerce, even data and internet usage statistics). The list goes on. My project got

a really good score. It also made me start using a vpn on my phone and pc, which doesnt help against 90% of the stuff above, but to an extent at least obscures some location data (it will never be 100% because of your phone being on the cell network anyway). There's a ton of data being gathered on all of us every day. I cannot think that half of it has any actual useful use case, at least not directly to the provider gathering it... but that is where 3rd party sale of that data comes in. A whole network of companies doing different things, targeting their products/markets based on trends that come out of that data. I will tell you one thing for sure. If someone with ill intent got a hold of all of that data on me, they can most certainly ruin my life in multiple different ways (identity theft to gain cash, commit crimes, cancel my existing policies/services without me knowing, impersonation with intent to ruin reputation, or go after and harm the people I care about). Now throw generative AI in that mix, and things get even nastier. **CubicalCropduster**

- **CLUSTER 7:** We need to make a EULA for an app that politicians will use. Put in the most egregious things in there. Then use that against them all until they actually put in laws that protect the customers. **Opetyr**
- **CLUSTER 8:** 2 things: 1. Its basically the Republicans on the Supreme Court that thinks companies should have no limits of what they can write in contracts. Mandatory individual arbitration and class action waivers was illegal about 10 years ago, then the Republicans on the Supreme Court voted 5-4 on party lines to allow companies to do it. Every democrat justice voted against it. 2. If you're in a multi person residence or you have friends over frequently, theres basically no way for them to verify it was actually you that signed the agreement for a game unless they required something like making an account that requires verification.” **WingerRules**
- **CLUSTER 9:** Yeah man just don't do the thing this subreddit is made specifically for. **Barredbob**
- **CLUSTER 10:** If it's free, you are the product. **NotNotDiscoDragonFTW**

Conclusioni

Inizialmente, ci siamo proposti di rispondere a tre domande principali per seguire un percorso di analisi, ciò ci ha portato a concludere che la situazione del mondo del gaming è sicuramente critica; tramite la sentiment siamo riusciti ad estrapolare che il 31% degli utenti prova emozioni negative al riguardo, contro il 13% di positivi.

Inoltre, il mese che ha creato più scompiglio negli utenti è Gennaio, ciò è sicuramente correlato al fatto che è stato il mese con più licenziamenti degli ultimi due anni, portando gli utenti a lamentarsi in modo maggiore con più di 500 commenti negativi nel lasso di tempo in questione.

Attraverso l'utilizzo della Named Entity Recognition volevamo estrarre le aziende e i giochi sotto il riflettore degli utenti. La maggior parte delle aziende viene vista in modo negativo dagli utenti per il fatto che quasi tutte hanno effettuato dei licenziamenti in questi ultimi due anni. Questo anche per quanto riguarda i giochi singoli, anche se non tutte le critiche sui giochi riguardano il nostro obiettivo.

Per quanto riguarda l'analisi coi grafi, possiamo provare a rispondere alle domande proposte nei due post analizzati.

1. Is the gaming industry on decline or is it getting better?

Per rispondere con precisione a questa domanda bisognerebbe leggersi tutti i commenti. Noi abbiamo deciso di utilizzare come "*campioni*" gli utenti che avevano scritto e discusso di più; tra l'altro, alcuni di essi, si sono anche rivelati essere gli utenti con maggior numero di "*like*" e quindi con molta approvazione da parte della community.

Detto ciò, leggendo le loro discussioni sembra che nessuno voglia dare una vera risposta *secca*, ma che preferiscano lamentarsi di ciò che non gli piace, limitandosi ad elencare e parlare male di giochi **flop**.

Nel complesso, però, non si può nemmeno rispondere positivamente, poichè la com-

munity non è affatto fiduciosa, ma anzi, è severissima con ogni problematica emergente.

2. **Abusive privacy policies in video games a grim sign for industry.**

Su questo tema invece, possiamo dare una risposta molto chiara e molto più *robusta* della precedente. Gli utenti non sono affatto d'accordo con le strategie per far accettare le clausole più nascoste ai videogiocatori. Lo capiamo utilizzando una logica simile a quella precedente: utilizziamo il numero di risposte ad un commento per vedere se quel commento è stato approvato dagli altri. Risulta che l'utente **RareCodeMonkey** scrive un commento che riceve più di 1000 *like* e 18 risposte. Ne sono state riportate solo poche, però assicuriamo che ad ora, nessuno dei 18 commenti provava a smentirlo o, quanto meno, a andare in contro alle grandi case produttrici di videogiochi.

Per quanto riguarda la community detection, si può dire che ha portato a risultati pronosticabili. La logica dietro a tutto è stata: dal grafo risultava evidente una frammentazione in piccoli gruppetti; quindi, per trovarli, si è deciso di trovare, anzitutto, con un metodo di clustering, tutte le community e poi provare a dargli un nome. Il nome di ogni cluster viene assegnato cercando il commento dell'utente con più risposte. Anche qui la logica dietro è chiara, un commento con tante risposte sarà quello che ha dato il via ad una discussione (cluster) che potrebbe essere diversa.

Come previsto, tenendo conto dei cluster più grandi, dal commento di **Misragoth** (cluster 2), si parla di pubblicità all'interno dei giochi; questo è un argomento che è comunque inerente al tema "*speculazione sul gaming*", ma che non c'entra con la domanda fatta nel post.

Una cosa simile la si può riscontrare nel cluster 4, dove **No-Roll-3759** parla di grafica.

Si ha quindi il cluster 0 dove ci sono i commenti inerenti al post, mentre più ci si allontana, in termini di distanza di nodi, più si tende a parlare di altro.

Bibliografia

- [1] MissionHairyPosition. *Riot Games Announces Layoff of 530 Positions, 11% of Workforce*. 2024. URL: https://www.reddit.com/r/gaming/comments/19db3b0/riot_games_announces_layoff_of_530_positions_11/.
- [2] MPsAreSnitches. *Abusive privacy policies in video games a grim sign for industry*. 2024. URL: https://www.reddit.com/r/gaming/comments/1bzwwsl/abusive_privacy_policies_in_video_games_a_grim/.
- [3] Onlyadeej. *Layoff in the Gaming Industry*. 2023. URL: https://www.reddit.com/r/gaming/comments/1b25xyb/layoff_in_the_gaming_industry/.
- [4] RobIson240YT. *What's a good video game that made the industry worse?* 2024. URL: https://www.reddit.com/r/gaming/comments/1acbvx/whats_a_good_video_game_that_made_the_industry/.
- [5] SuperSeriousStud. *Is the gaming industry on decline or is it getting better?* 2023. URL: https://www.reddit.com/r/gaming/comments/14vrhak/is_the_gaming_industry_on_decline_or_is_it/.
- [6] *Why are so many gaming industry employees getting laid off?* 2023. URL: https://www.reddit.com/r/gaming/comments/17va3f6/why_are_so_many_gaming_industry_employees_getting/.
- [7] Wikipedia contributors. *2023–2024 video game industry layoffs — Wikipedia, The Free Encyclopedia*. [Online; accessed 21-May-2024]. 2024. URL: https://en.wikipedia.org/w/index.php?title=2023%E2%80%932024_video_game_industry_layoffs&oldid=1224858736.
- [8] XsStreamMonsterX. *The Video Game Industry is NOT Collapsing. A Lawyer Explains*. 2024. URL: https://www.reddit.com/r/gaming/comments/18rsr01/the_video_game_industry_is_not_collapsing_a/.

- [9] Zack Zwiezen, Kenneth Shepard. *89 Days Into 2024 And 8,800+ Video Game Layoffs Have Been Announced*. 2024. URL: <https://kotaku.com/game-industry-layoffs-how-many-2024-unity-twitch-1851155818>.