



VALORANT

# Agenda

1. Introduction
2. Problem Statement
3. Data Collection
4. Data Pre-Processing
5. Sentiment Analysis
6. Insights from Analysis
7. Data Modelling
8. Conclusion
9. Future Scope

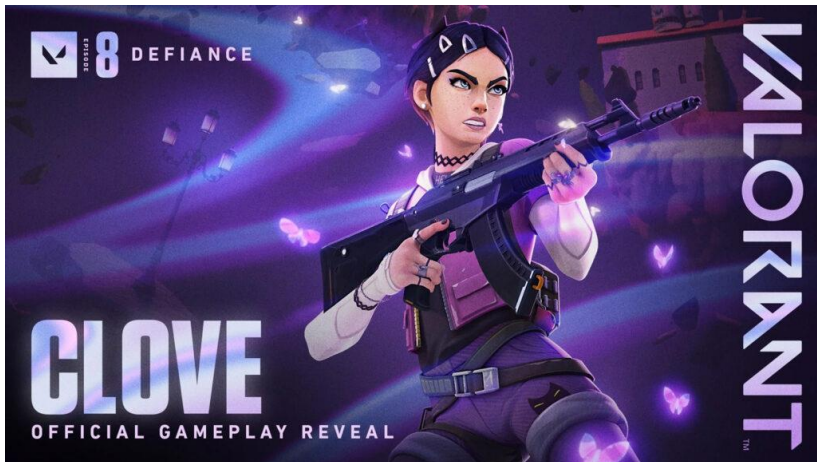


# Introduction

- **Valorant**, a free-to-play first-person tactical hero shooter developed and published by Riot Games
- Released on June 2, 2020
- Estimated 19M total players as of 2024 with a daily average of nearly 5M players
- Players play as one of a set of Agents, characters based on several countries and cultures around the world



# Problem Statement



- The 25<sup>th</sup> agent, **Clove** debuted in the game alongside patch 8.05 on 26<sup>th</sup> March 2024
- What are people's reactions to the new agent before and after its release?
- We conducted sentiment analysis on people's discussions on Reddit



# Data Collection

Collected the Reddit's comments using PRAW from Reddit's below five subreddit

Subreddit	Subscribed Users	# of posts per day	Collected count
r/Valorant	2 million	~9	11489
r/ValorantCompetitive	222k	~2	9380
r/ValorantPBE	40k	~1	1600
r/ValorantClips	22k	~2	632
r/Eggwick	25k	~3	826

To optimize storage and retrieval efficiency, we implemented storing data from CSV to MySQL database due to the large volume of data from these subreddits.



# Data Pre-Processing/Challenges

Pre-processing the collected data involved removing filler and stop-words to enhance its relevance for further analysis.

## Challenges

- The character name "clove" inconsistently appeared as "cleave," potentially causing confusion in preprocessing.
- Some gamer verbal expressions were accidentally removed, risking loss of valuable context.

## Solution

- Developed logic to handle "clove" variations, maintaining accuracy.
- Refined preprocessing to retain crucial gamer expressions for comprehensive data processing.



# Sentiment Analysis

Implemented sentiment analysis on pre & post-release dataset

## **Google Bert:**

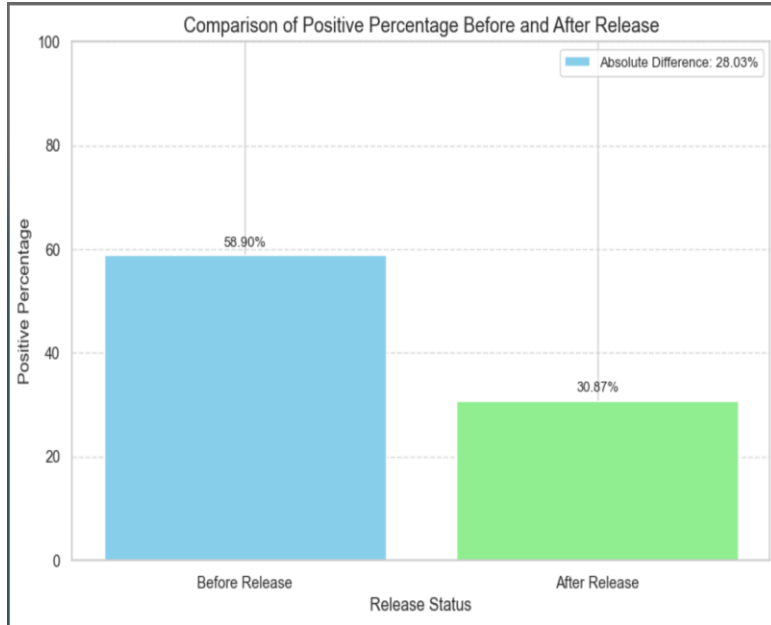
- Training Data Set - IMDB Data set
- Challenges – Longer execution time to train model, incorrect sentiment predictions due to context mismatch in IMDB and reddit comments

## **Llama 2:**

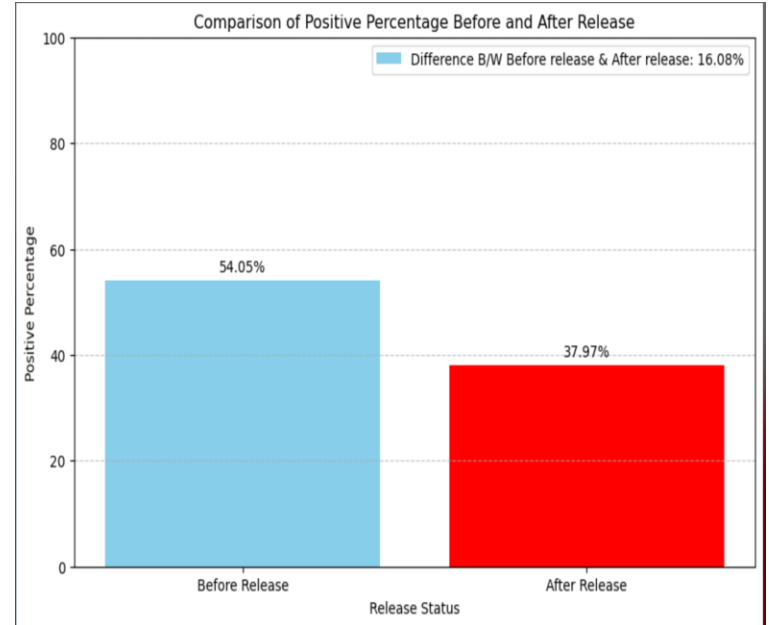
- Ollama - used to run Llama 2 LLM locally
- LangChain - used to interact with Ollama Llama 2
- Challenges - prompted to download models from paid sources and pay for collab pro due to extensive CPU utilization, giving csv files as input



# Insights from Analysis (BERT)



**BERT's Sentiment Analysis for Raw Data**

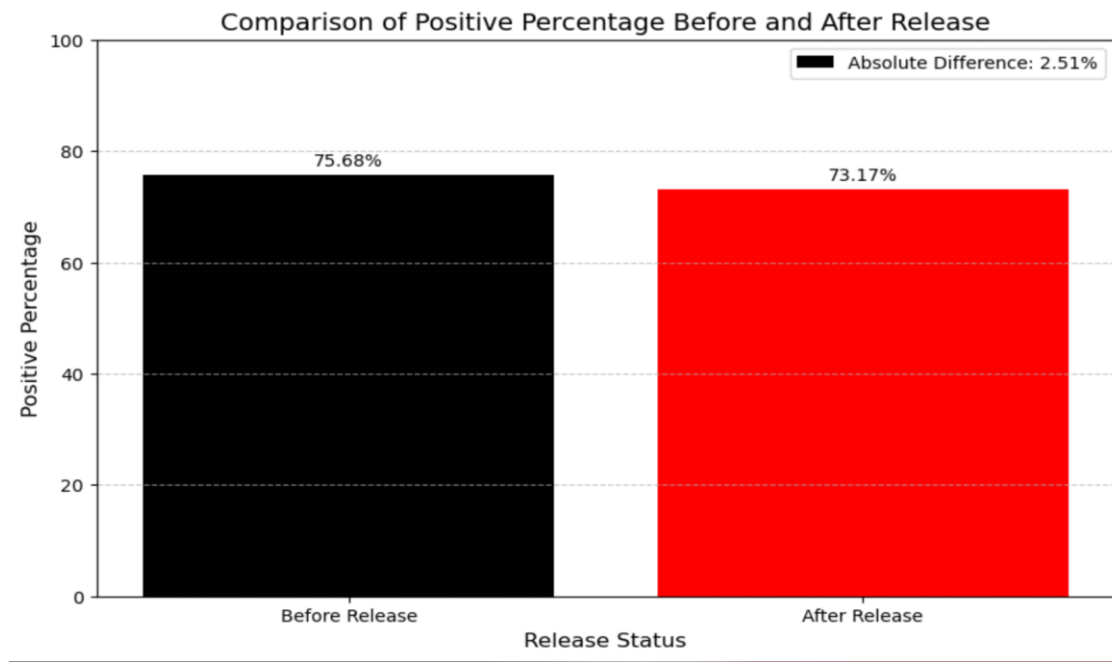


**BERT's Sentiment Analysis for Filtered Data**





# Insights from Analysis (Llama)



**Llama 2's Sentiment Analysis for Filtered Data**



# Data Modelling

- Data modelling is crucial for extracting meaningful insights from data.
- LDA (Latent Dirichlet Allocation): Probabilistic topic modeling technique used to discover hidden topics within a collection of documents or texts.
- Topic modelling done on 4 distinct groups of data formed by sentiment analysis:
  - Before Release Positive Comments
  - Before Release Negative Comments
  - After Release Positive Comments
  - After Release Negative Comments



# Data Modelling (WordCloud)

Top 10 topic in Negative comments of Before Clove



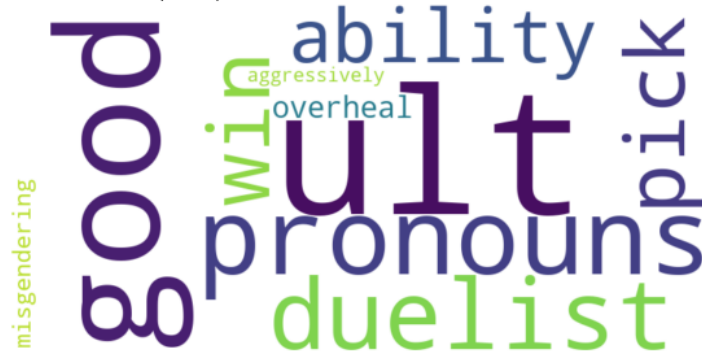
Top 10 topic in Positive comments of Before Clove



Top 10 topic in Negative comments After release of Clove



Top 10 topic in Positive comments After release of Clove



# Valorant Community's Reception of Agent Clove

- Consistent Themes:
  - "Smoke" remains a critical point of discussion pre and post-release
  - Persistent mention of "pronouns" signals ongoing gender identity conversations
- Evolving Discussions:
  - Transition from speculative "hype" to actual gameplay focus on "ability" and "ult."
  - Pre-release concerns about competitive fit lessen post-launch
- Sentiment Improvements:
  - Positive reception of Clove's "abilities" post-launch suggests exceeded expectations
  - Clove is embraced as an "aggressive controller," fitting the desired playstyle
- Areas for Attention:
  - Recurrent misgendering highlights the need for community education
  - Negative comments on "team" and "lurk" suggest potential issues with team dynamics

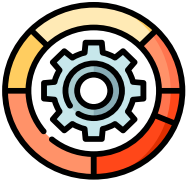


# Future Scope



## Enhanced Data Collection

- Collecting additional pre-release data for improved sentiment analysis
- Exploring data from other social media platforms



## Advanced Data Processing:

- Refining keyword filtering for precise analysis
- Implementing advanced algorithms for optimized insights



## Optimized Data Management:

- Considering NoSQL for scalability with unstructured data



## Automated Data Retrieval:

- Deploying API Wrappers on remote VMs for continuous data retrieval





**Thank You!**

