

PLAYER 1

HIGHSCORE 2500

PLAYER 2

UNCOVERING PLAYER SENTIMENTS: A SENTIMENT ANALYSIS ON THE FAULLOUT SERIES STEAM REVIEWS FOR BETHESDA GAME STUDIOS

START

MENU

SIGN IN

PRESENTED BY: JOCEL ARCELONA

MENU

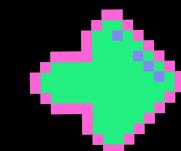
⚡ 01

♦ 07

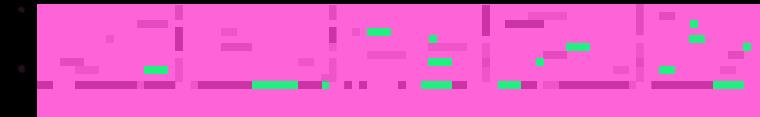
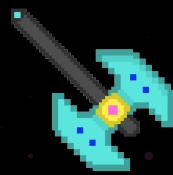
★ 12



BUSINESS PROBLEM:



BETHESDA GAME STUDIOS JUST GOT BOUGHT BY MICROSOFT AND THEY CURRENTLY HAVE A HUGE INFLUX OF BUDGET TO PRODUCE THEIR NEXT FALLOUT TITLE. HOWEVER, THEY LACK SPECIFIC UNDERSTANDING OF WHAT FACTORS ARE DRIVING POSITIVE OR NEGATIVE SENTIMENT AMONG THEIR PLAYERS.



OBJECTIVES:

- To understand the overall player sentiment towards the Fallout Series using Natural Language Processing
- To identify common themes and topics that emerge from the player reviews using Topic Modeling
- To provide insights and recommendations to Bethesda Game Studios on how to improve player experience based on player reviews
- To develop a sentiment analysis model that can accurately predict a players sentiment towards the game



BACK TO
AGENDA PAGE

SIGN IN



BACK TO AGENDA PAGE

DATA

- Web scraped 320,000 reviews from the steam app with 22 features
- Concatenated 6 dataframes from 6 fallout series titles
- Ended up using only 192,000 reviews after dropping non-english reviews

MENU



♦ WHAT IS STEAM?

Steam is a digital platform created by Valve Corporation to serve as a distributor of PC games



♦ WHAT IS THE FALLOUT SERIES ABOUT?

The Fallout Game Series is a post-apocalyptic action role-playing game set in a retro-futuristic world.



♦ WHAT IS BETHESDA GAME STUDIOS?

Bethesda Game Studios is known for its work on the Fallout and Elder Scrolls series. These games are large, open world RPG games

[BACK TO AGENDA PAGE](#)



PREPROCESSING STEPS:

- ❖ LOWERCASE TEXTUAL DATA, REMOVE STOP WORDS, REMOVE NON-ENGLISH WORDS, TOKENIZE AND LEMMATIZE USING NLTK
- ❖ USED GENSIM'S LDA FOR TOPIC MODELING AND GENSIM'S PHRASES OBJECT FOR EDA TO IDENTIFY AND VISUALIZE WORDS THAT SHOW UP THE MOST IN EACH DOMINANT TOPIC FOR POSITIVE AND NEGATIVE REVIEWS

[BACK TO AGENDA PAGE](#)

01 07 12

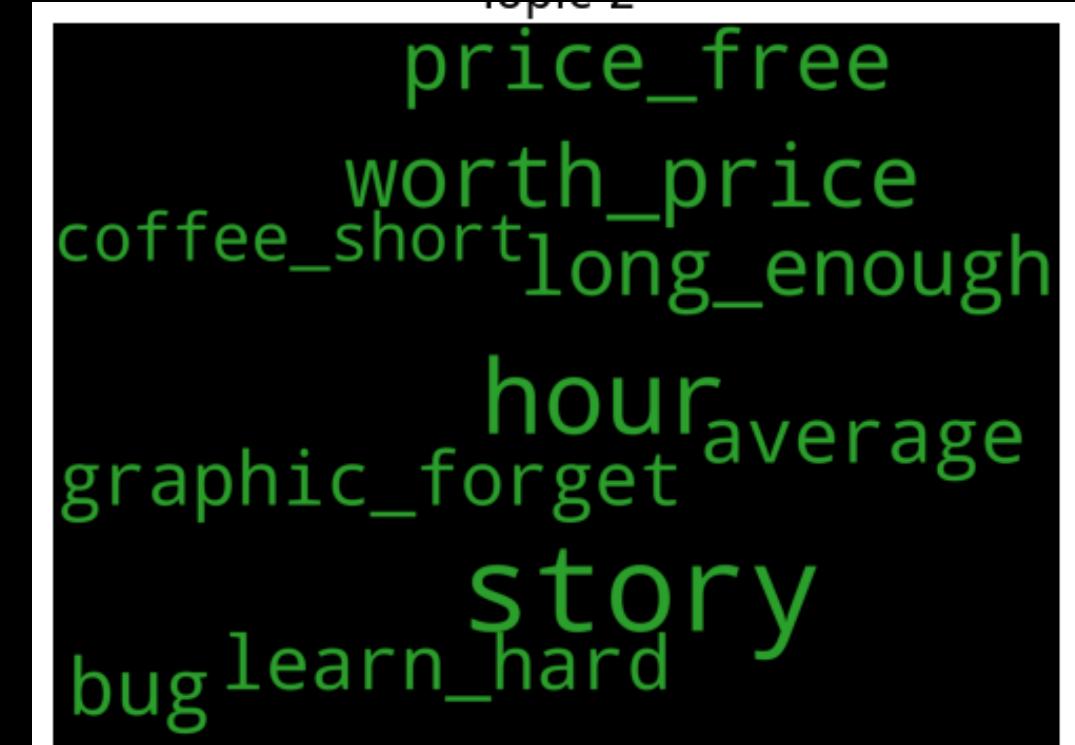


MODS/GREAT
STORYLINE

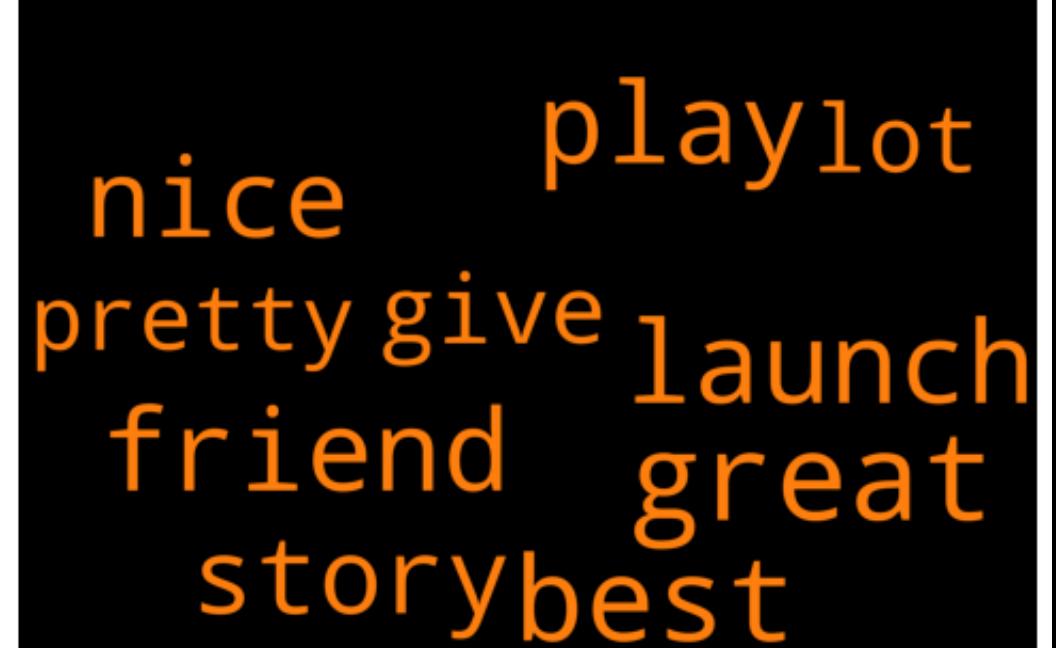


WORD
CLOUD
OF TOPICS FROM
POSITIVE
REVIEWS

PRICE



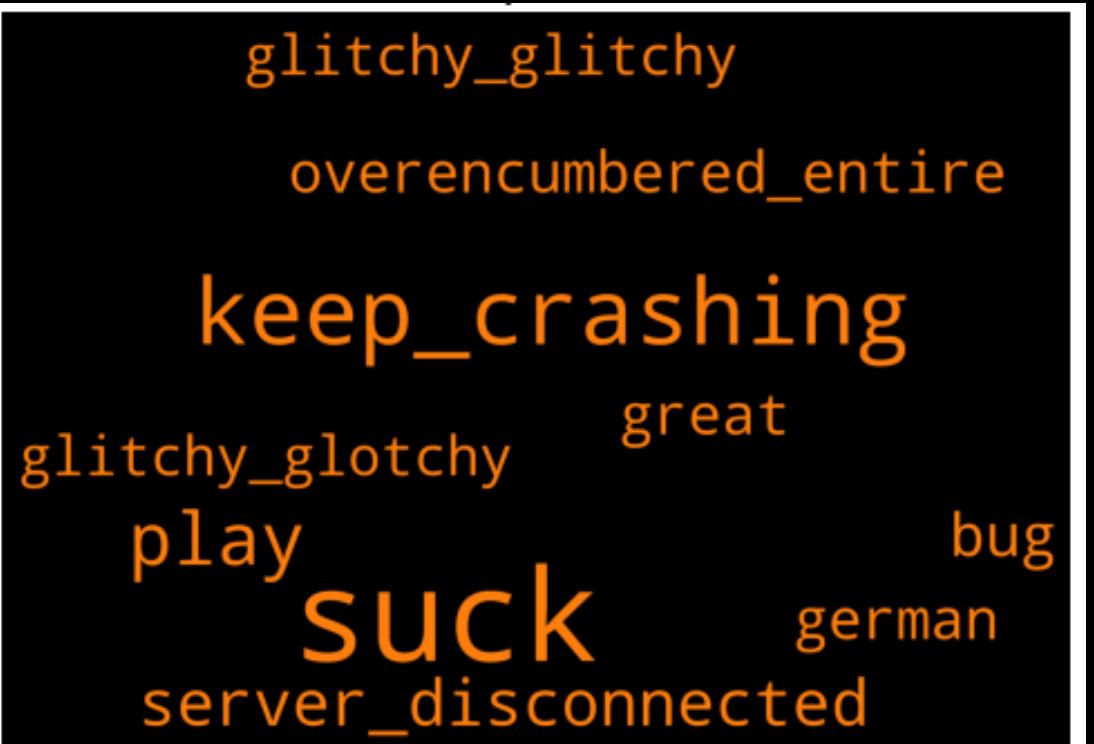
MULTIPLAYER



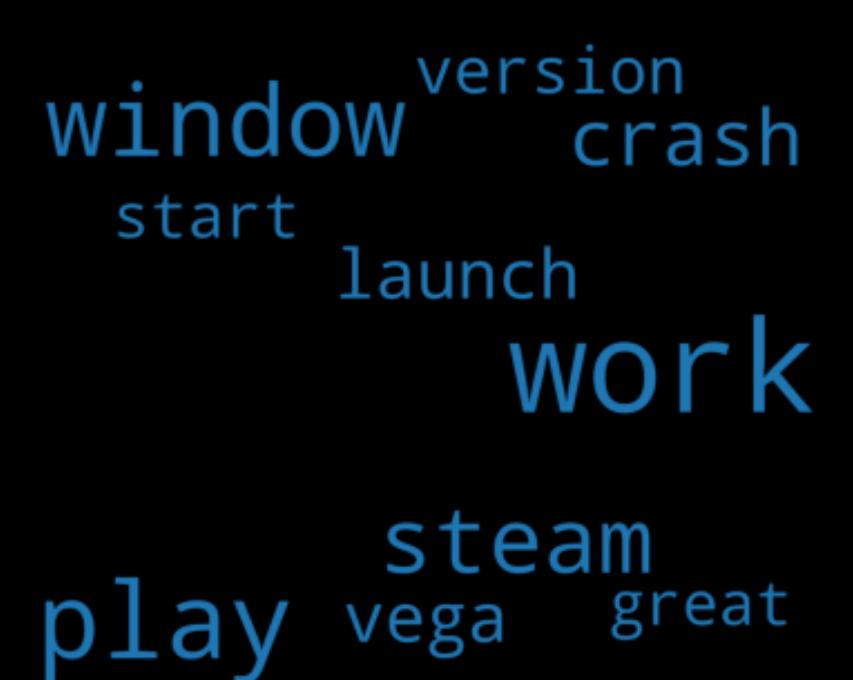
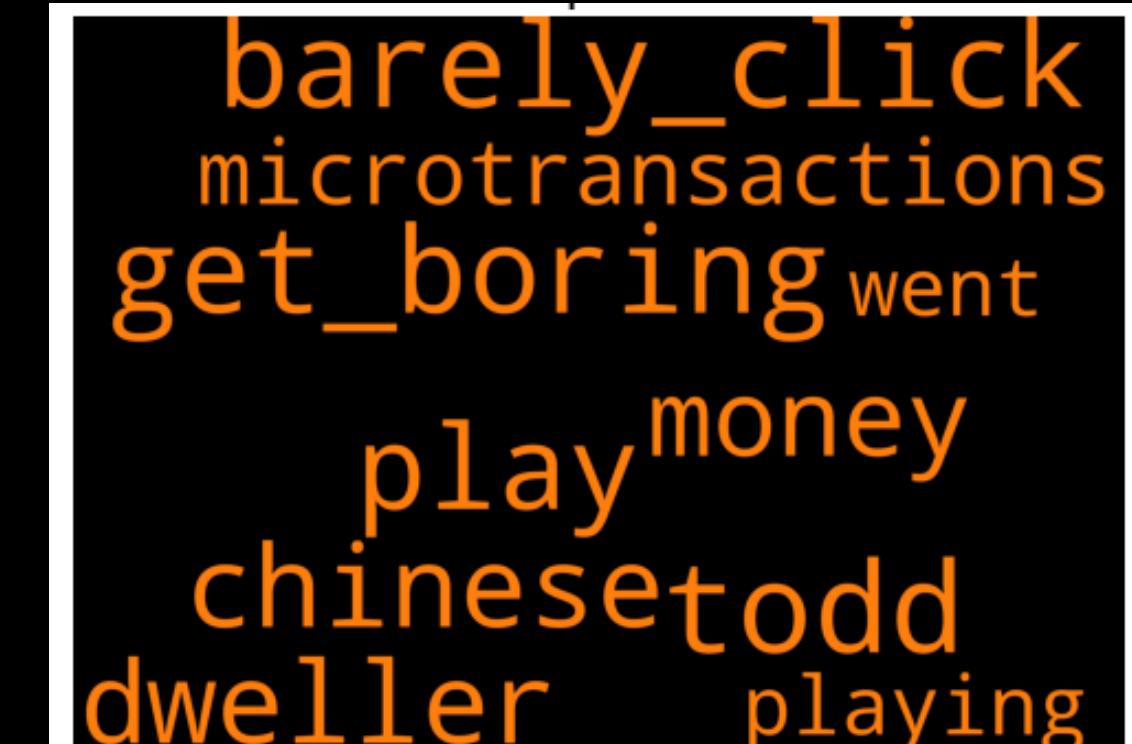
Coherence score: 55%

NON-PLAYABLE
CHARACTERS





WORD
CLOUD OF
TOPICS
FROM NEGATIVE
REVIEWS



Coherence score: 37%



PLAYER 1



PREDICTIVE MODELING

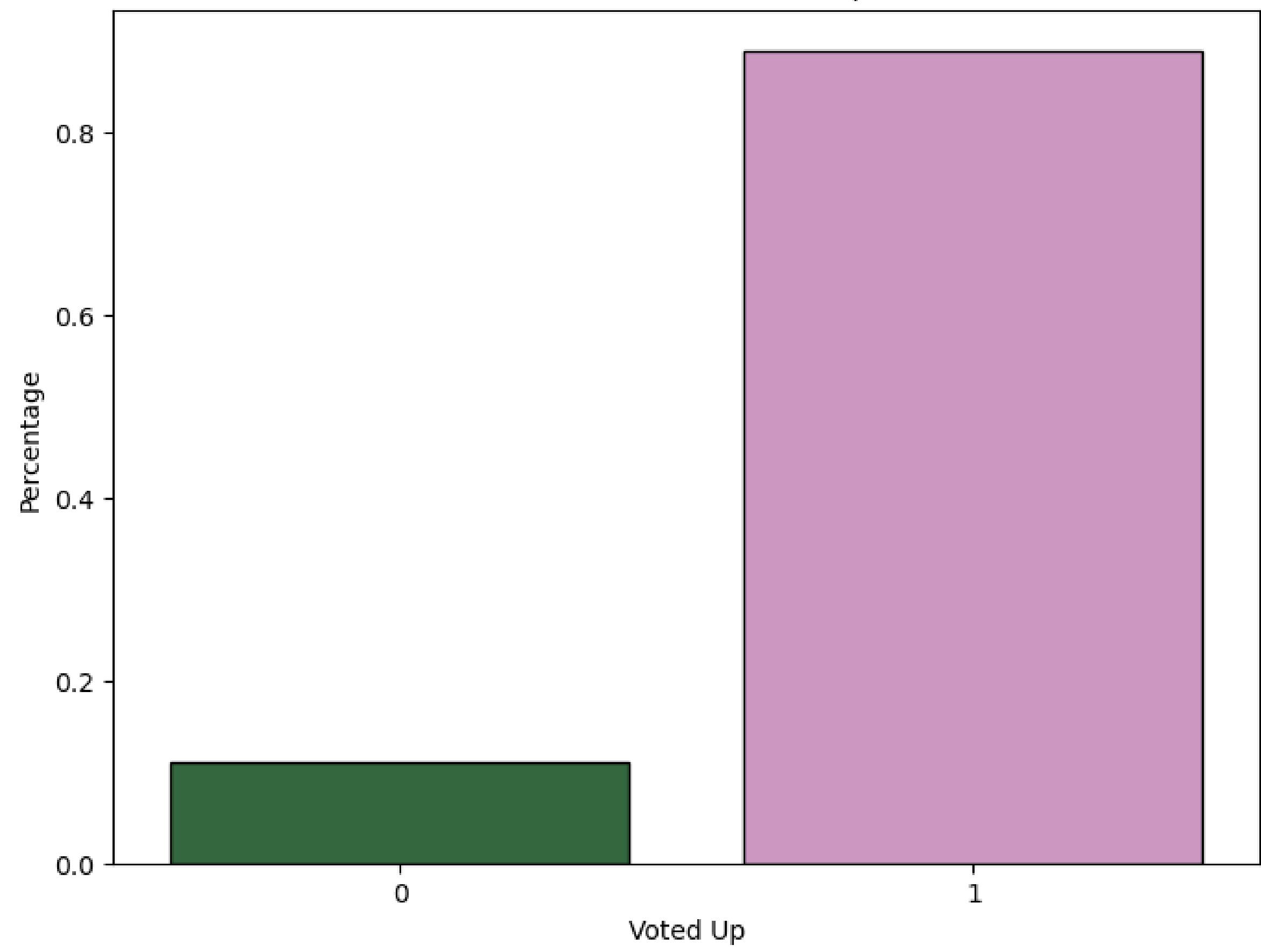


CLASS IMBALANCE

CLASS 0: 11%

CLASS 1: 89%

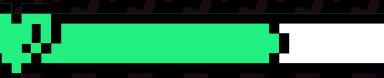
Distribution of Voted Up





★★★★★

HIGHSCORE 2500



PLAYER 2

MODELS



MULTINOMIAL
NAIVE BAYES



DECISION TREES



RANDOM FOREST



XGBOOST

NLTK (UNIGRAMS)

ACCURACY: 0.77

ACCURACY: 0.89

ACCURACY: 0.87

ACCURACY: 0.93

GENSIM (BIGRAMS)

ACCURACY: 0.77

ACCURACY: 0.89

ACCURACY: 0.86

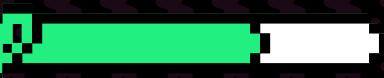
ACCURACY: 0.92





★★★★★

HIGHSCORE 2500



PLAYER 2

BEST MODEL

STACKED MULTINOMIAL
NAIVE BAYES & XGBOOST:

ACCURACY: 0.93

RECALL:

- CLASS 0: 0.70
- CLASS 1: 0.95

PRECISION:

- CLASS 0: 0.65
- CLASS 1: 0.96

F1 SCORE:

- CLASS 0: 0.67
- CLASS 1: 0.96

WHAT IS A STACKING CLASSIFIER?

A STACKING CLASSIFIER IS A TYPE OF ENSEMBLE MODEL IN MACHINE LEARNING THAT COMBINES MULTIPLE BASE CLASSIFIERS TO MAKE A FINAL PREDICTION.

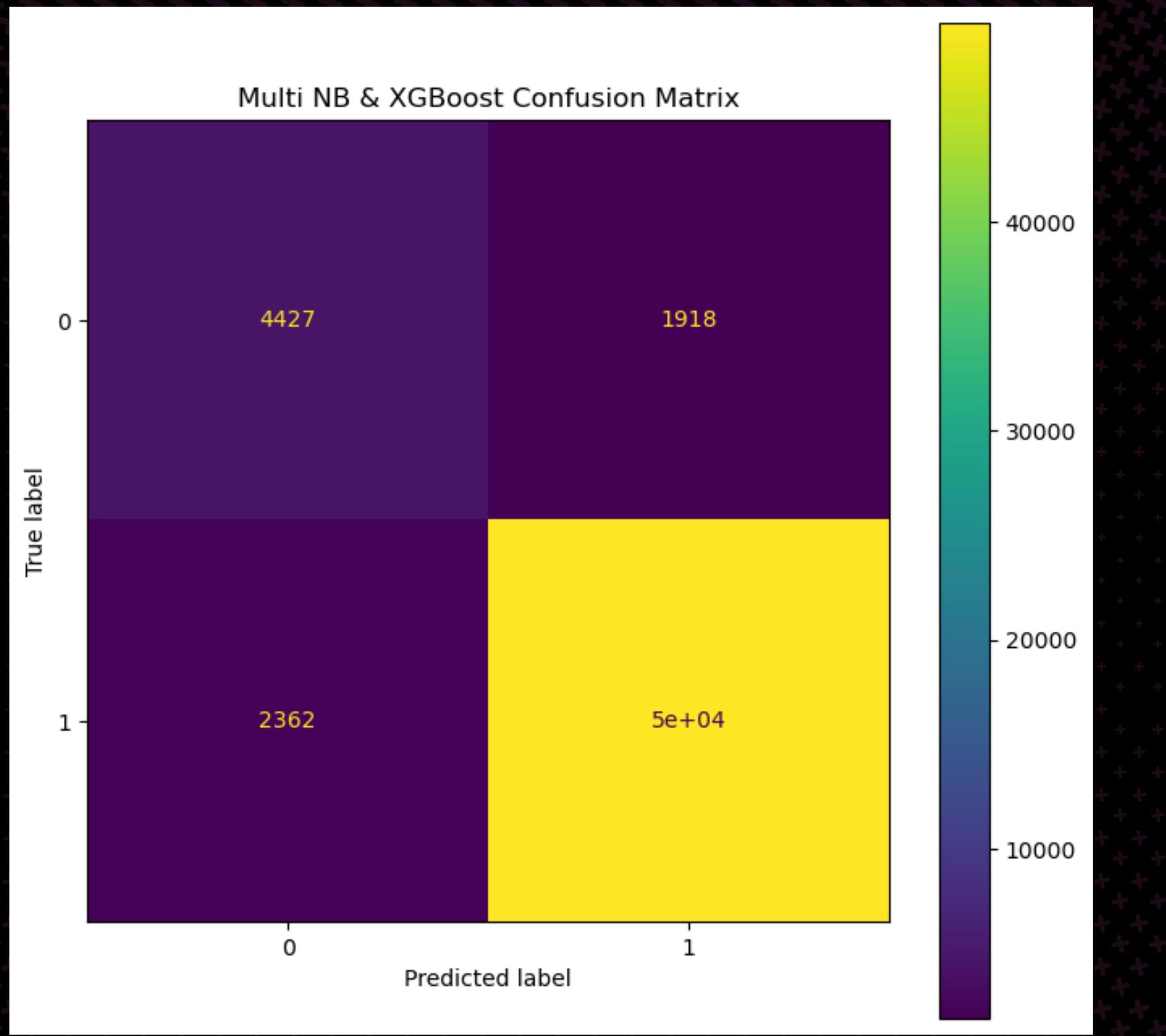


HIGHSCORE 2500

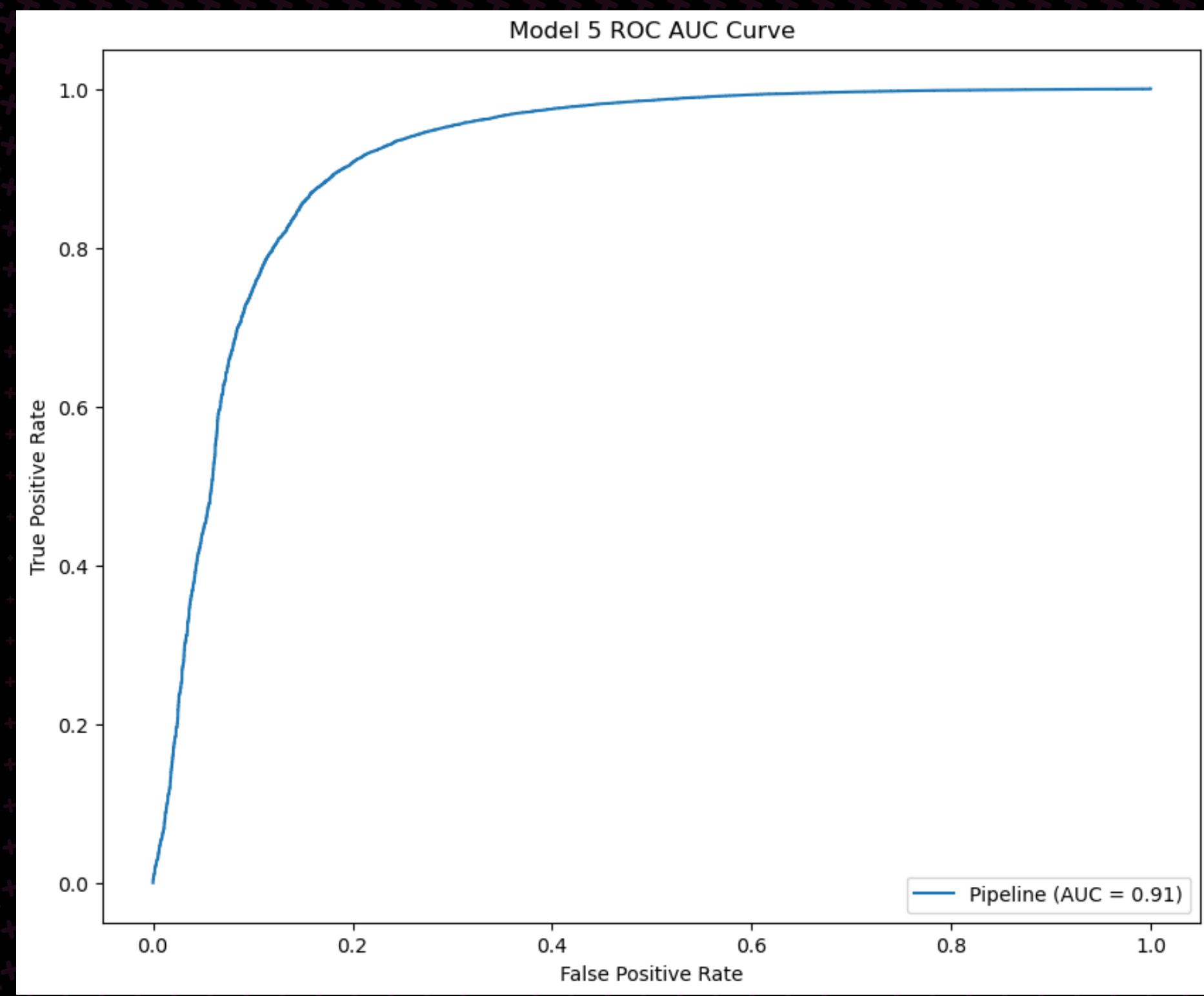


PLAYER 2

Multi NB & XGBoost Confusion Matrix



Model 5 ROC AUC Curve



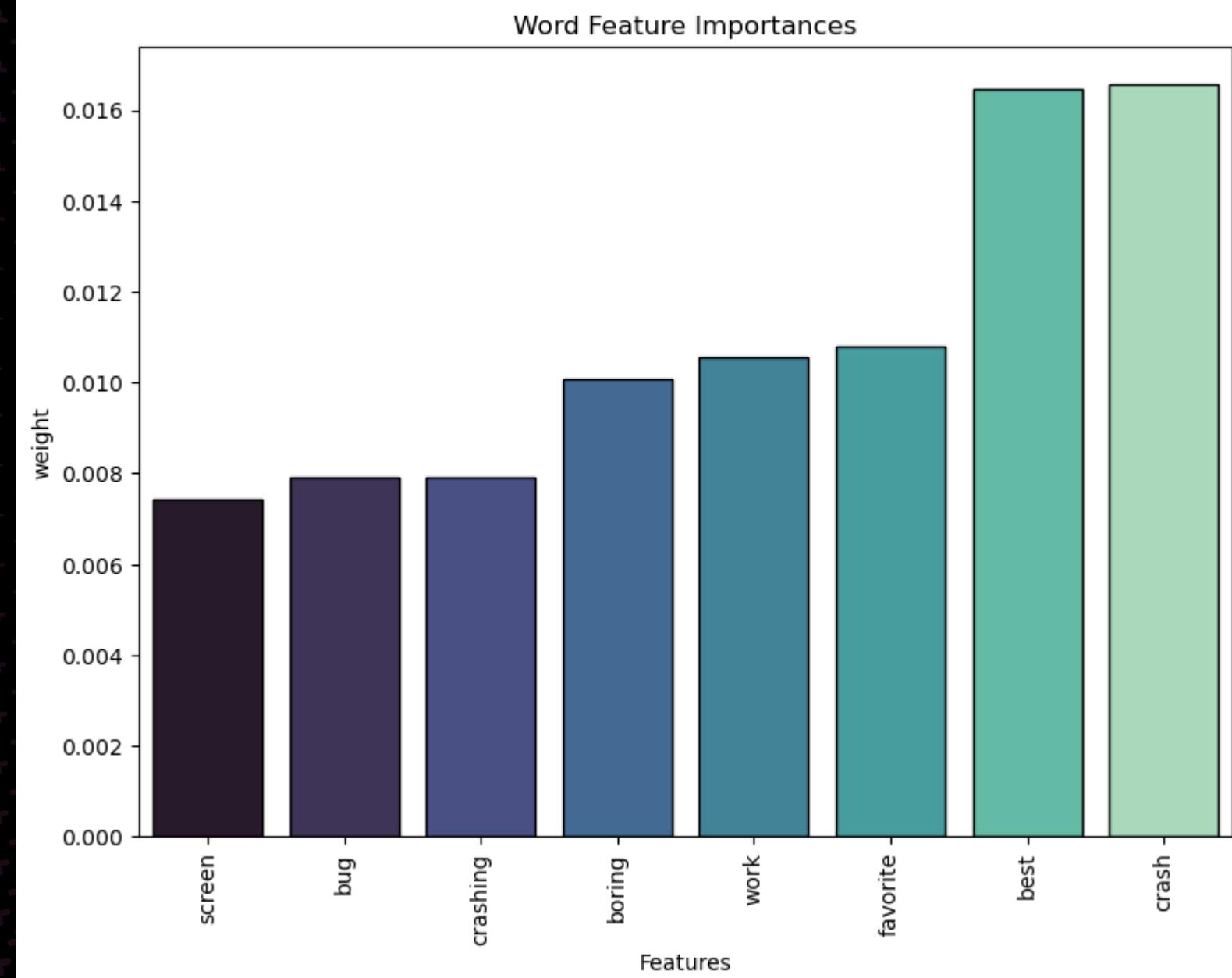
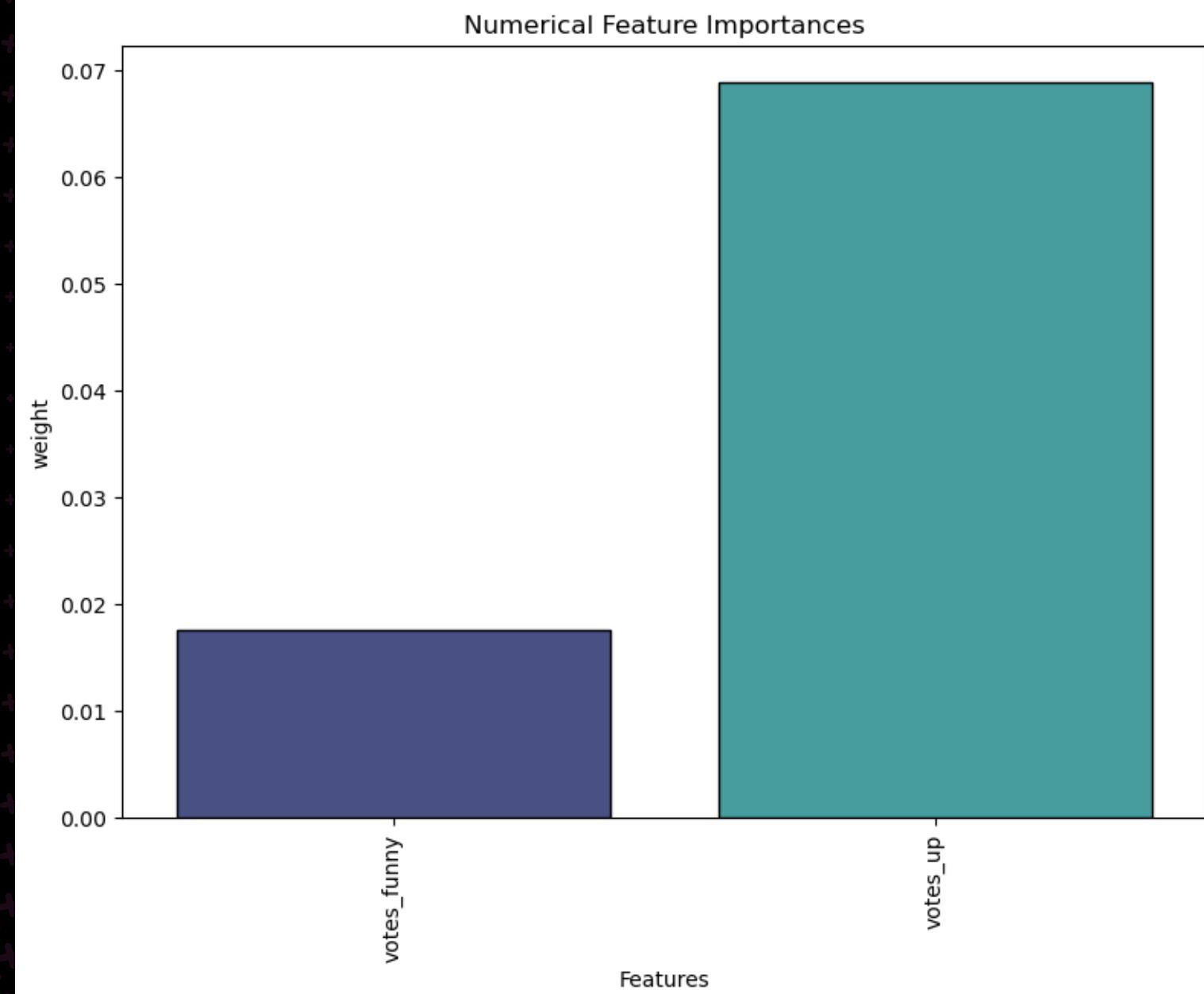


PLAYER 2

FEATURE IMPORTANCE

"VOTES UP" FEATURE
HELPS OTHER PLAYERS TO
QUICKLY IDENTIFY
REVIEWS THAT HAVE BEEN
FOUND USEFUL OR
INFORMATIVE BY OTHER
PLAYERS.

"VOTES FUNNY" FEATURE
IS INTENDED TO PROVIDE
A WAY FOR PLAYERS TO
SHARE THEIR SENSE OF
HUMOR AND APPRECIATE
REVIEWS THAT ARE
WRITTEN IN A
LIGHTEARTED OR
HUMOROUS TONE.



RECOMMENDATIONS

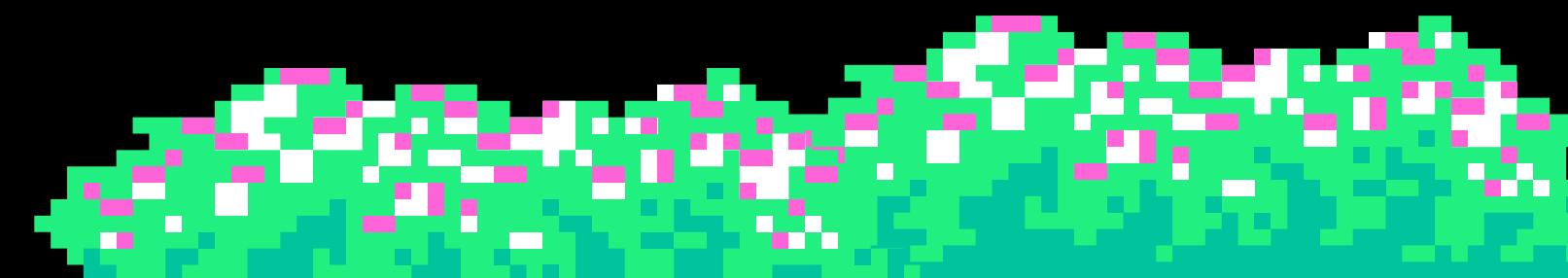
- ★ + Fix technical issues like glitches, frequent crashes and launching issues
- ★ + Reduce number of microtransactions
- ★ + Improve the storyline
- ★ + Remove "always online" mechanism to reduce reliance on centralized servers, if possible

❖ FUTURE STEPS ❖

DATA COLLECTION:
GATHER MORE REVIEWS
FROM OTHER GAMING
PLATFORMS

USE PYSPARK FOR BIGGER DATA

DATA TRANSLATION: USE A
TRANSLATOR TO
TRANSLATE NON-ENGLISH
REVIEWS FOR BETTER
GENERALIZABILITY



MENU



THANK YOU!



jocelarcelona30@gmail.com



<https://github.com/JocelArcelona>



<https://www.linkedin.com/in/jocel-arcelona/>