Natural Language Processing

In Steam Games

Mariam Kandilari
Eleni Chaitidou
Dionysios Vaggelatos

Overview

- What kind of games are more popular?

Overview

- What kind of games are more popular?

- How can we predict from a review whether a user recommends a game title?

Overview

- What kind of games are more popular?
- How can we predict from a review whether a user recommends a game title?
- We used some of the most popular video-games to make sentiment analysis on user reviews to answer these questions.

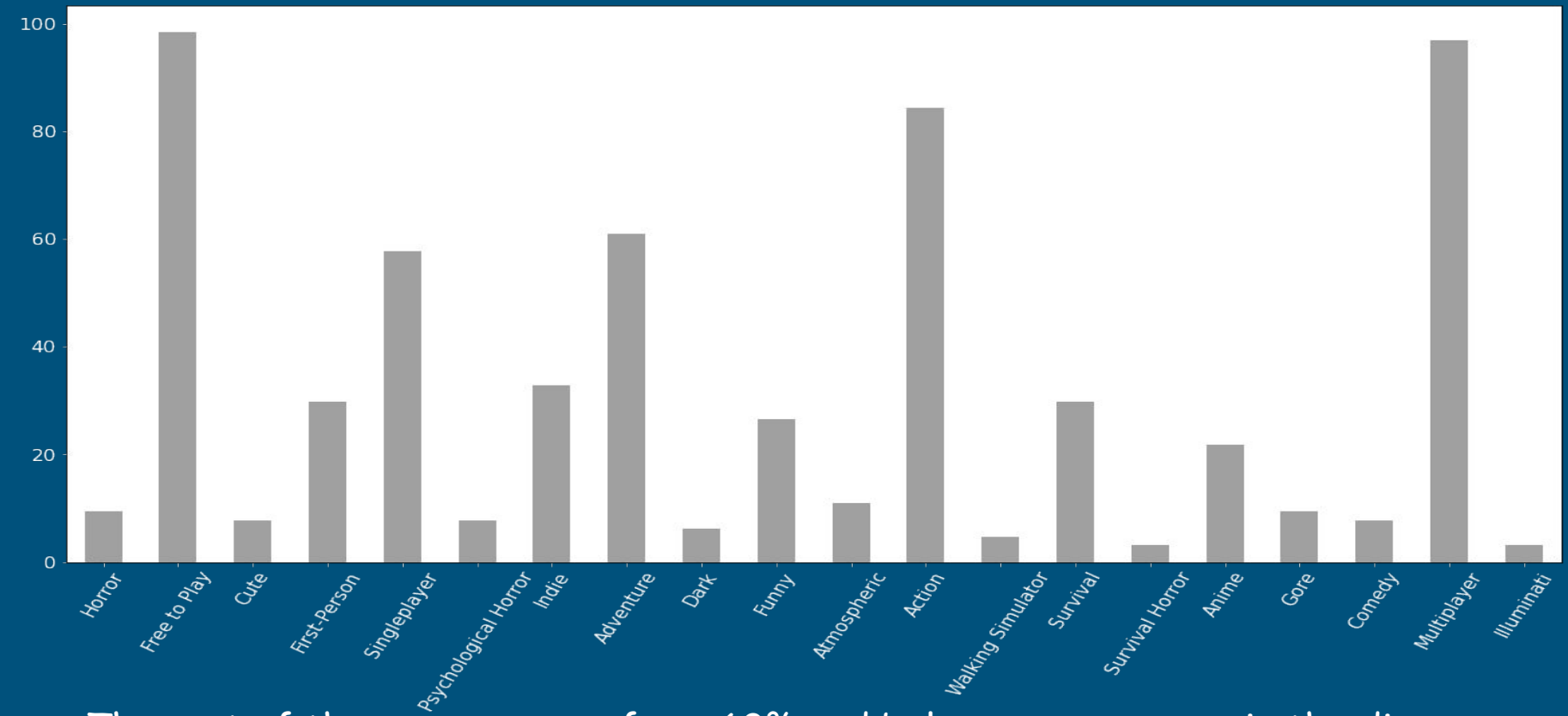
The three most popular games according to their genre (in %)

- Free to play with 98.5%

- Multiplayer with 96.8%

- Action with 84.3%

Game genres and their corresponding percentages



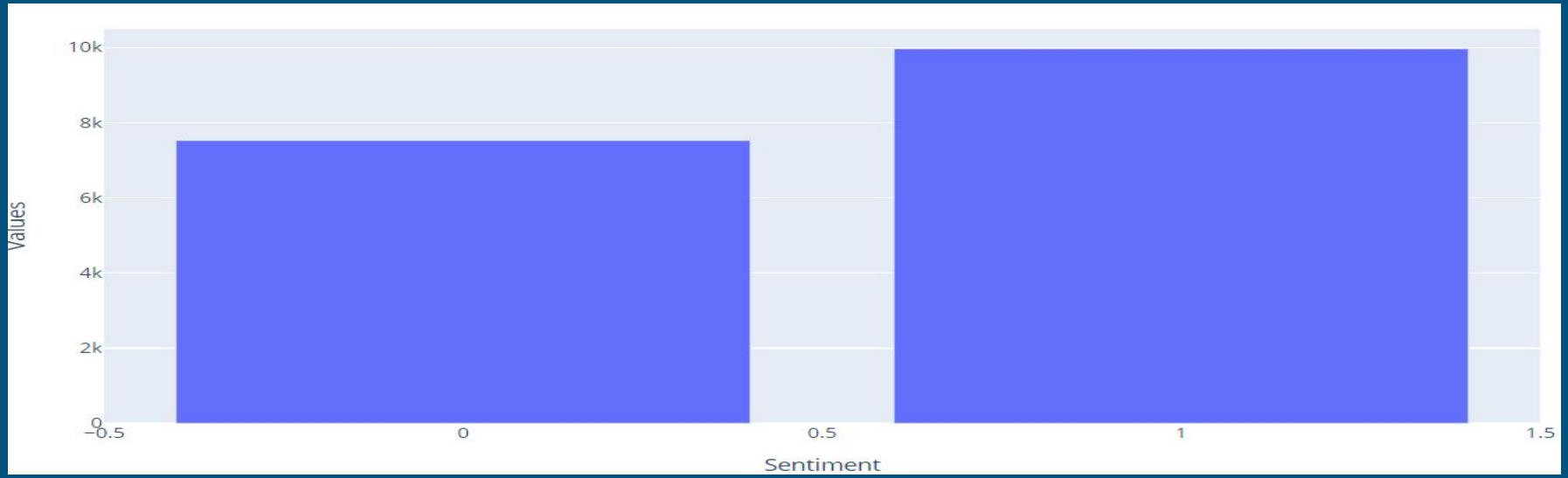
The rest of the genres range from 60% and below as we can see in the diagram above

Reviewers can be angry… but also very pleasant

The key to success in Video-game business is to avoid the first reaction!

Negative vs Positive User Input

Counter to the sterotype we see our dataset leans on positive reviews. Might be because the games in it are generally popular. We can of course balance our set to be more precise!



Model Training

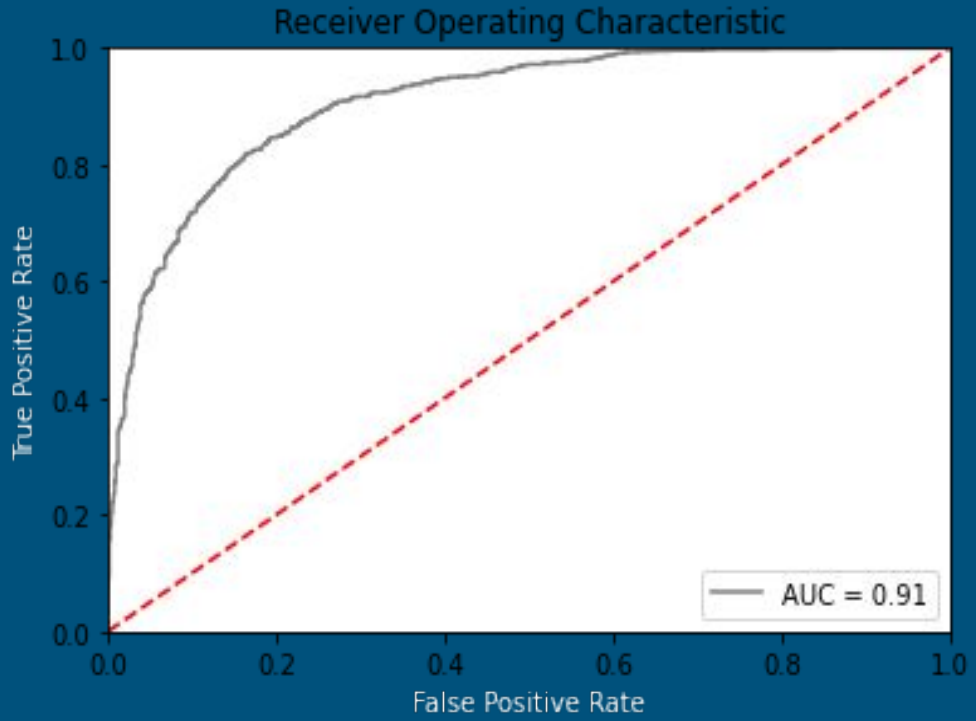
We trained a model to predict whether a review will be negative or not with users inputs: with an accuracy of 83%(with logistic regression)!

The AUC score of our model

Our model works very well

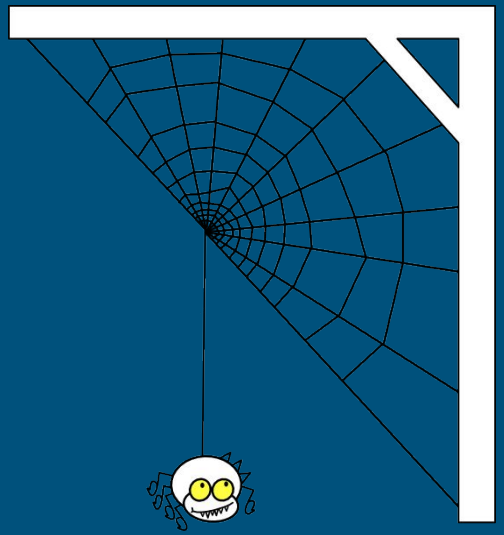
with identifying 'Relevant' words

As shown with its AUC score.



How does it work?

- We use the key-words from the reviews to detect sentiment of the reviewer!
- When these key-words are associated with a game, it is more likely that the game will be received positively!



A web of associations between these words will provide us insights whether the game will be appreciated by the users!

Examples!

- Words like ‘fun’, ‘great’, ‘enjoyable’, ‘cheerful’, ‘yeah’ appear usually in reviews and descriptions of games that are positively rated!
- On the other hand ‘hate’, ‘shame’, ‘disappointed’ or ‘boring’, ‘repetitive’ appear in negative reviews.

To close, our Sources:

1. Kaggle.com (dataset with game data)
2. Google.com (images)

Thank you all!