

Project Proposal

Web Data Processing Systems Assignment 2

Group 27

- Simei Li, 2738043
- Yiran Li, 2730767
- Kairui Wang, 2731737
- Summer Xia, 2703936

Intro & Motivation

The game industry has become one of the most profitable markets in the entertainment industry. Knowing what customers and users think and how they feel about a game (or a game event) is a central piece to drive the decision-making process, of any game developer or game studio, towards user satisfaction.

Our topic is to analyze some user comments toward a worldwide game "**Genshin Impact**" from popular social media such as Youtube(English), bilibili(Chinese), etc., finding the events, topics or characters that players are most interested in during this period, as well as the emotional tendencies of the comments under each video. If possible, we will also compare the interests and emotional tendencies of people in different languages on the same event to see whether there are differences.

Goal

Here we plan to use game version trailer videos.

These videos will be released one week before the version update. As a brief overview of the content of the new version, it shows what new gameplay, new characters, new events and other information will be available in the next version.

On a certain timeline (up to 8 month), for each video (up to 5), we will present these results:

- **Text Mining:**
 - **Keywords:** Including the events, characters and slangs, etc.
 - **Relationships** between some keywords.
- **Entity Linking:** For some typical keywords, especially character and item names, identify them with game wiki page.
- **Sentiment Analysis:**
 - General emotional tendency of the video
 - Emotional tendency of keywords

Roadmap

Task 1: Crawler

xx - 11/12/2021

Obtain comments of the game's official account on social platforms such as youtube and bilibili. (Youtube or GooglePlay/Appstore?)

Task 2.1: Text Mining

12/12/2021 - 18/12/2021

- Keywords extraction
- Named entity recognition
- Relation extraction

Task 2.2: Entity Linking

17/12/2021 - 19/12/2021

Task 3.1: Sentiment analysis on comments

12/12/2021 - 19/12/2021

Task 3.2: Sentiment analysis on keywords

16/12/2021 - 19/12/2021

Task 4: Report and Presentation

19/12/2021 - 21/12/2021

Work split

No matter how it's listed, we make sure that everyone's workload is almost the same.

Task 1: Simei Li, Kairui Wang

Task 2: Simei Li, Summer Xia, Kairui Wang

Task 3: Summer Xia, Simei Li, Yiran Li

Task 4: Yiran Li, Kairui Wang

Our concerns (questions)

- What kind of comments are better for this topic? Comments from video platform or comments from Game App Store?
- Is it a feasible way to do sentiment analysis that we use comments from App Store (which are with ratings) to train the model and then predict the video comments? Though we know some general sentiment analysis models work.