ANALYZING PLAYER FEEDBACK IN STEAM REVIEW ACROSS GAME GENRES

SAFIRA NURUL IZZA

UNIVERSITI TEKNOLOGI MALAYSIA

Table of Content

CHAPTER 3 METHODOLOGY

CHAPTER 3

METHODOLOGY

3.1 Introduction

This chapter provides information on the research method adopted in studying the Steam reviews of the players based on game genre. This outlines the phases of the study, data acquisition by web scraping and Application Programming Interfaces, sentiment analysis by different methods, and mapping out feedback patterns to the various genres. The application of the methodology is to make the process of solving the research questions consistent, following patterns and standards that would help developers understand how player sentiment differs by game genre and how to enhance it.

3.2 Research Framework

The research framework for this study consists of four main phases: preparation and planning, data collection, sentiment analysis and interpretation, and conclusion. Figure 3.1 represents the research framework phases, followed by details of each phase.
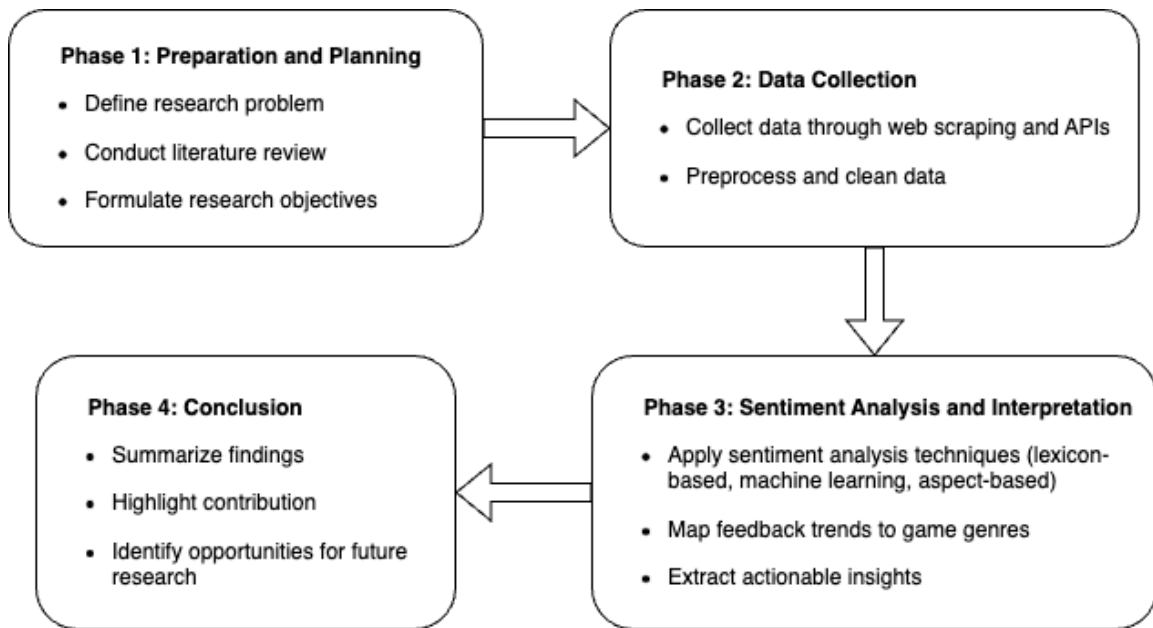
Figure 3-1Reseach Framework

The first phase is preparation and planning. it is an assessment of learning resources involving examining a set of materials associated with the issue, as well as staking out the research paradigm and goals. This step helps to set a background for the research to evaluate the limitations of other methods of analysing player input using the Steam review platform as well as to define the parameters of the study.

The second phase is data collection, and the major focus is made on the collection of the review data from the Steam platform. They include using web scraping together with an API to obtain reviews of choice games in five games categories. Furthermore, collected data is processed on an analytical nature for its quality and relevance of the research work.

The third phase is the sentiment analysis and interpretation phase, which its objective is to further analyse the collected data using sentimental analysis tools such as Lexicon base, Machine leaning sentimental analysis and Aspect base sentimental analysis. This

phase also involves linking this feedback patterns with the Genre, evaluating player attitudes as well as evaluating the information which is relevant to the game developers.

The last phase is conclusion, where the specific findings, implications to the theories and various suggestions for further routes are presented. This phase aims at achieving the study objectives and provide suggestion about improvement of game development.

3.3 Planning and Preparation

The planning and preparation phase focuses on investigating the research problems and conducting a thorough literature review. The research problems are introduced by examining the background of the study, followed by the research questions and objectives. Additionally, the scope and significance of the study are defined to ensure clarity and relevance. These aspects are presented in detail in Chapter 1.

The literature review phase aims to establish a comprehensive understanding of the current methods and concepts relevant to analysing player feedback in Steam reviews. This involves comparing existing sentiment analysis techniques, identifying their strengths and limitations, and highlighting areas for improvement. Key concepts such as sentiment analysis methods (e.g., lexicon-based, machine learning, deep learning, and aspect-based approaches), challenges in sentiment analysis (e.g., handling sarcasm, ambiguity, and multilingual data), and data collection methods (e.g., web scraping and APIs) are explored in Chapter 2.

Furthermore, this phase includes an investigation of gaming genres and their impact on player feedback, as well as the role of sentiment analysis in understanding user reviews. The findings from the literature review provide a foundation for the methodological choices and establish the relevance of the proposed approach.

3.4 Development of the Analytical Framework

The development phase involves creating and implementing components required for analysing player feedback from Steam reviews across various game genres. Figure 3.2 illustrates the operational framework, including data collection, data preprocessing, sentiment analysis, and insights generation. Detailed descriptions of each layer are provided in the following subsections.
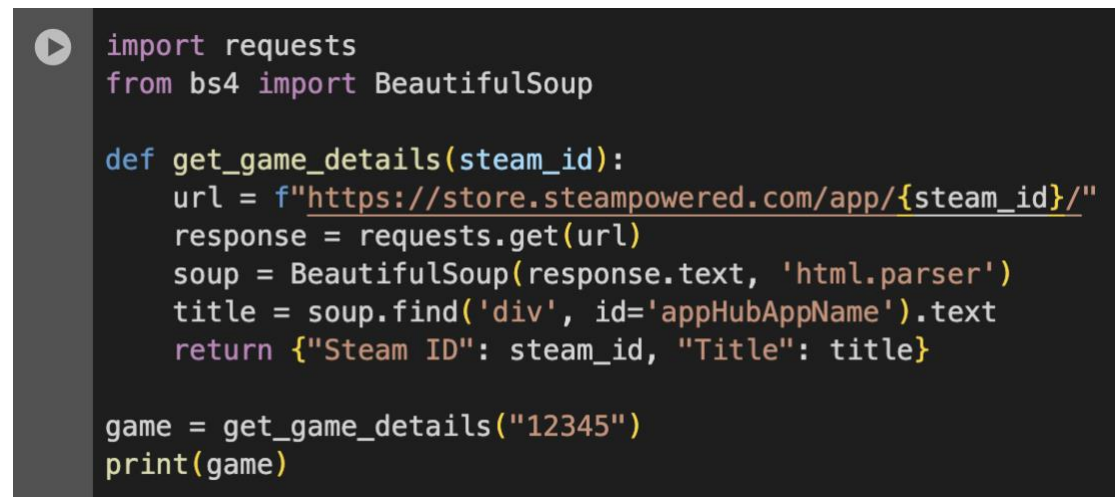
3.4.1 Data Collection Layer

The data collection layer gathers datasets from the Steam platform, focusing on reviews for selected games across five genres: Action, RPG, FPS, Strategy, and Indie. The games selected for each genre include:

- **Action**: Red Dead Redemption 2, Sekiro™: Shadows Die Twice - GOTY Edition, Lies of P, FINAL FANTASY VII REBIRTH.

- **RPG**: Fear & Hunger, Undertale, Baldur's Gate 3.

- **FPS**: Call of Duty®: Black Ops 6, Ready or Not, Metro Exodus, PUBG: BATTLEGROUNDS, S.T.A.L.K.E.R. 2: Heart of Chornobyl.

- **Strategy**:

- **Indie**: Nine Sols, Mouthwashing, Unpacking, Miside, Stardew Valley.

Data is collected using web scraping techniques and Steam APIs to extract relevant information such as game title, genre, developer, publisher, and release date. The information is validated and stored for further processing.

3.4.1.1 Data Collection

Data collection involves extracting game details from the Steam platform using web scraping techniques. The Python libraries requests and BeautifulSoup were used to retrieve and parse game information, such as title, genre, developer, publisher, and release date. The following snippet demonstrates the process:

```python
import requests
from bs4 import BeautifulSoup

def get_game_details(steam_id):
    url = f"https://store.steampowered.com/app/{steam_id}/"
    response = requests.get(url)
    soup = BeautifulSoup(response.text, 'html.parser')
    title = soup.find('div', id='appHubAppName').text
    return {"Steam ID": steam_id, "Title": title}

game = get_game_details("12345")
print(game)
```

This code fetches and parses the game details from Steam's store page using the specified steam_id. The data collected is then stored for further analysis. For this study, a total of five games from each genre were analysed.

3.4.2 Data Preprocessing Layer

In this layer, raw data is cleaned and structured for sentiment analysis. Key steps include:

- **Field Separation**: Splitting combined fields, such as user data and activity logs, into distinct fields for ease of analysis.

- **Filtering**: Removing irrelevant or redundant data entries.

- **Normalization**: Standardizing formats (e.g., date, text) to ensure consistency across the dataset.