Video Game Recommender System

Part 1: Search

Website: http://discorn.pythonanywhere.com/

1. Hosting, Framework, Language

Hosting: PythonAnywhere

Framework: Flask

Languages: Python 3.6, SQLite, HTML, CSS

2. Functionality

The website is to take the user input and return a list of games in according to the relevancy of the input. Information returned would include the title, release date, metacritic score, price, and description.

To save space and reduce load time several unused data columns in the dataset were left out. Data such as the publisher count, achievement count, and linux requirement.

3. Algorithm

The search function will use TF-IDF (Term Frequency – Inverse Document Frequency) by taking in every word on the input and calculates the frequency of the term among the documents then multiplied by the inverted frequency of term occurrences in each document. This repeats for every word and return a score to be outputted in the results page.

Challenges:

- Initial website made using JavaScript but reverted to using Python instead.
- Crashed when using Flask-tables, resolved by using HTML table.

Improvements to be Made:

- CSS visual aesthetics
- Expand TF-IDF to include the title with weighed scores so not to overshadow the description score
- Add stemming function
- Exclude stop words

Source Code: https://github.com/KRT0335/RecSystem

Dataset: https://data.world/craigkelly/steam-game-data

References:

Basic flask tutorial

http://www.blog.pythonlibrary.org/2017/12/12/flask-101-getting-started/

Video tutorial of Python Flask by Corey Schafer

https://www.youtube.com/watch?v=MwZwr5Tvyxo

Blog describing TF-IDF integration into Python

https://towardsdatascience.com/tf-idf-for-document-ranking-from-scratch-in-python-on-real-world-dataset-796d339a4089