Video Games Aspects Mining & Recommendation
(Master Project 30 ECTS)

Description:

Metacritic[1] is website aggregating reviews in order to create a normalized score. Additionally, it allows to compare various critics from professional medium to users. The different categories are movies, TV series, music and video games. In this work, we are interested in the latter. An example could be recent games such as God of War[2] with critic reviews[3] and user reviews[4].

The idea of this master thesis is 1) finding aspects in reviews and expressing the similarity between two reviews; 2) applying recommendation systems for video games. More specifically, we would like to explore various ways of doing collaborative filtering using textual similarity as a metric to define user's neighborhood, similarly to [1], [2].

Currently, our dataset contains approximately:
- 75'000 games
- 270'000 professional reviews in English
- 535'000 user reviews in English
- 217'000 users wrote at least one review

The master project includes three distinct parts:
1) Develop several methods to represent the semantics of reviews. It could be conditioned on game's genre. A first model would determine the aspects (with the clusters) of a professional and/or user review. Comparing reviews is important and thus, a similarity metric based on the aspects has to be defined. Other neural models such as auto-encoders and siamse networks could be used in this task. At the end, we would like to answer this kind of questions:
   a. Do two reviews contain the same (sub)set of aspects?
   b. Are user reviews aligned somehow with professional reviews?
   c. If we condition reviews by genre, do we have similar results as before?

2) Media are known to be biased. We would like to study whether there exist clusters of news medium. For this task, we can use the representations developed in 1) and measure the similarity among professional reviews with various distance metrics. A complementary method could be studying the correlation in terms of ratings. Finally, we could study their influence on users. Example of questions we could answer:
   a. Does IGN[5] attach the same importance to the aspects than Polygon[6] ?

---

[1] http://www.metacritic.com/
[2] http://www.metacritic.com/game/playstation-4/god-of-war
[3] http://www.metacritic.com/game/playstation-4/god-of-war/critic-reviews
[4] http://www.metacritic.com/game/playstation-4/god-of-war/user-reviews
[5] https://www.ign.com/reviews/games
[6] https://www.polygon.com/reviews

b. What are the most important aspects of racing games for IGN?
c. Are (some) users influenced by these mediums. Is there a correlation between their scores and some medias?

3) Personalized recommendation is an important task for online platforms, especially when the data is sparse. For examples, in hotels, it is rare that two users share many hotels. We believe this is also the case for video games (to be validated). We would like to explore collaborative filtering based on reviews similarity instead of item preferences, similarly to [1], [2]. Moreover, we could also experiment if using in addition professional reviews would be beneficial.

Prerequisites:
- Machine Learning
- (Deep) Neural networks
- Python
- Knowledge of PyTorch is a plus

Supervisor: Diego Antognini (diego.antognini@epfl.ch) & Boi Faltings

References:

[1] https://www.ijcai.org/Proceedings/13/Papers/395.pdf
[2] https://www.ijcai.org/Proceedings/15/Papers/122.pdf