Tên đề tài: Xây dựng trang web tìm kiếm và gợi ý phim

Giáo Viên hướng dẫn:

TS. Huỳnh Xuân Phụng

Nhóm thực hiện:

Võ Hà Nhật Tân - 18133047

Phạm Đình Nhiên – 18133038

Danh sách các yêu cầu của đề tài:

Các chức năng tìm kiếm và gợi ý

Gợi ý các phim phù hợp với phim đã tìm

Hiển thị chi tiết thông tin bộ phim, diễn viên, đánh giá

Công cụ sử dụng: Python, Jupyter notebook, API, Front-end:Java Script, CSS, AJAX.

Thuật toán sử dụng: Content-based filtering, Cosine Similarity, KNN

Milestones:

1. Tìm hiểu về hệ thống Recommendation (22/08/2021 – 29/08/2021)

* Tìm hiểu Content based filtering và Collaborative Filtering
* Tìm hiểu các thuật toán có liên quan

1. Tìm hiểu các thuật toán:

- Euclidean algorithm: https://en.wikipedia.org/wiki/Euclidean\_algorithm

- Pearson correlation: https://en.wikipedia.org/wiki/Pearson\_correlation\_coefficient

- Cosine similarity: https://en.wikipedia.org/wiki/Cosine\_similarity

- Similarity score:

https://en.wikipedia.org/wiki/Similarity\_score

- Cosine Similarity:

https://www.machinelearningplus.com/nlp/cosine-similarity/

- Naive Bayes:

<https://viblo.asia/p/thuat-toan-phan-lop-naive-bayes-924lJWPm5PM>

1. Tìm hiểu công cụ sử dụng trong đề tài (30/08/2021-05/09/2021)

Flask: https://flask.palletsprojects.com/en/2.0.x/

AJAX: https://developer.mozilla.org/en-US/docs/Web/Guide/AJAX/Getting\_Started

Npl: https://realpython.com/nltk-nlp-python/

Thư viện sử dụng:

gunicorn

MarkupSafe

Werkzeug

numpy

scipy

nltk

scikit-learn

pandas

beautifulsoup4

jsonschema

tmdbv3api

lxml

urllib3

requests

pickleshare

1. Xây dựng thiết kế database (06/09/2021-18/09/2021)

* Dữ liệu phim MovieLens: https://grouplens.org/datasets/movielens/
* Tiền xử lý dữ liệu

1. Thiết kế hệ thống Recommendation (19/09/2021-02/10/2021)

* Lấy API key: <https://www.themoviedb.org/>
* Lấy API để load poster thông tin của phim lên trang web
* Xây dựng hệ thống gợi ý theo Content based
* Sử dụng thuật toán Cosine similarity, Naïve Bayes sử dụng Accuracy score để đánh giá mô hình

1. Xây dựng các chức năng cho trang web (03/10/2021-10/10/2021)

* Xây dựng chức năng search, gợi ý search
* Xây dựng chức năng load thông tin của phim
* Xây dựng chức năng gợi ý các phim phù hợp với phim đã tìm

1. Triển khai ứng dụng trên Herokuapp (10/10/2021-12/10/2021)
2. Kiểm tra lỗi và hoàn thiện hệ thống (13/10/2021-30/10/2021)