# SENTIMENT ANALYSIS FOR MOVIE RECOMMENDATIONS

HILARIO JUNIOR NENGARE(174682)

Professor Zafer Erenel | EUROPEAN UNIVERSITY OF LEFKE | TRNC

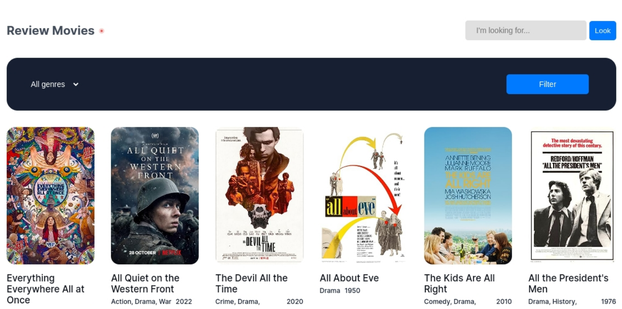
## INTRODUCTION

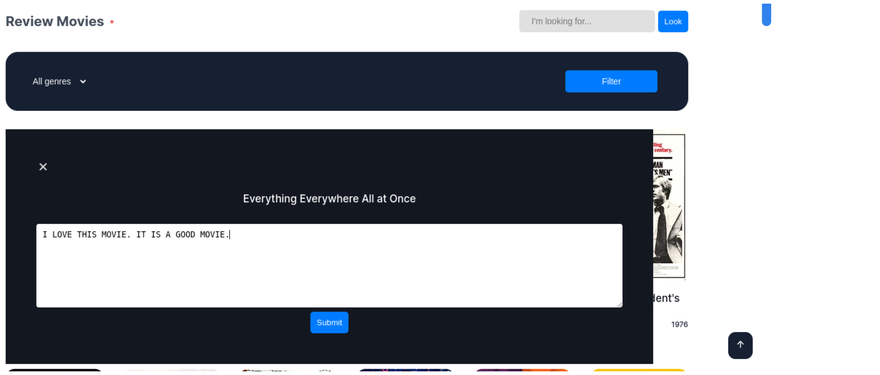
- With Sentiment Analysis we're able to mine for user opinions, felts and sentiments towards a certain movie and then use that input to create a very competent movie recommendation system.
- Sentiments in this regard are like a thumbs up or thumbs down in older movie rating systems.
- In this project the main aim is to achieve a movie-suggestion system which is optimal and personalized.
- Different users can share different sentiments towards a movie, and it is these reviews that we look at and then classify in accordance with positive or negative polarity.
- In order, to provide movie suggestions that align with user sentiments.

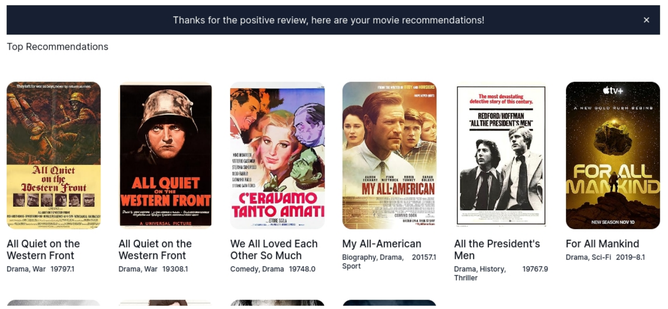
## GOALS

- **Viewer engagement** – with personalized movie recommendations, we achieve viewer traffic and manage to retain viewers.
- **Context-based movie recommendations** – each movie pitched to the viewer should be domain-compliant and contextualized to user sentiments.
- **Improved content Discovery –** this is what gears this project as we want to bring movies the user would love but would not have found out otherwise.
- **Adaptive movie recommender** system for users which enhances user retention, by automatically knowing what the user prefers and in what case they prefer it.
- **Increased satisfaction** when users receive recommendations that align with their preferences and sentiments, they are more likely to enjoy the recommended movies.

## RESULTS




## METHOD

- **Data collection** – A dataset with movie reviews from the IMDB website is scraped and used.
- On the server-end an OMDB API is used to recommend movies to users using content- based filtering.
- **Data cleaning** - An obtained review dataset is decoded into aspects during training, using the techniques unigram – if it's one word, bigram-if it's two words and trigram if it's three words.
- In data cleaning notebooks we extract useless features and symbols, known as stopwords .
- **Polarity Detection -** In this module classification of reviews into negative, positive or neutral is done.
- Machine learning models from scikit-learn library for polarity detection were selected. These included Naive Bayes, SVM, logistic regression and neural networks.
- SVM model outperformed the others and was trained on the labelled data that was attained. I used the extracted features and other corresponding labelled data to train the model.
- The modelwas saved as a pipeline to later be used in a flask server.

## CONCLUSION

- Sentiment analysis is a very relevant field of NLP and hence there loads of previous findings and research pertaining to this field.
- Its popularity ensures that there are a lot of resources such as libraries and datasets to build upon.
- It can be applied to other fields other than movie sites that use recommender systems and dependence is upon user feedback.
- Below **ethics** ought to be **enforced**:
- It is imperative to protect user data, ensure secure means of interaction, openness and inclusive usage for all people