MovieEdge – An Interactive Recommender System

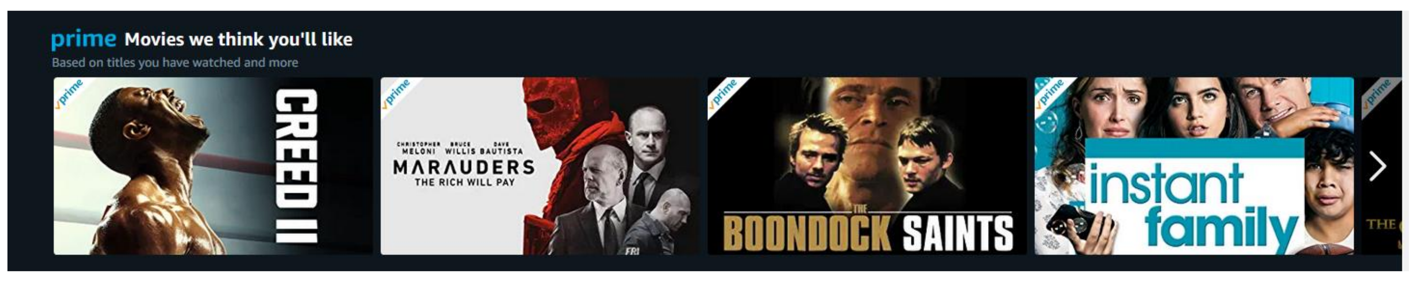
Georgia Tech College of Computing

Rocko Graziano, Daniel Klass, Yi Sun, Jonathan Tay

Data & Visual Analytics, Fall 2019

Motivation

Recommender Systems are a ubiquitous part of modern Digital Life



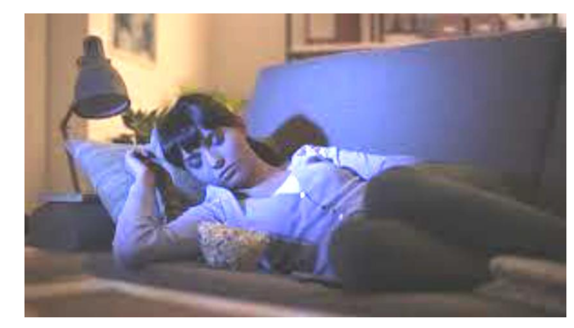
But they

- Lack interactivity [1]
- Offer recommendations which are hard to comprehend [1]
- Often lead to disappointing results and frustrated consumers [1,2,3]

MovieEdge

We present MovieEdge, a movie Recommender System which:

- Leverages accurate and efficient Word2vec modeling of user ratings to generate real-time movie recommendations
- Allows the user to interactively navigate through a movie "Taste Space"



Our Goal

Less of This

More of This

How We Did It

Gather & Preprocess Movie Rating Data

20 million ratings for 27,278 movies were downloaded from MovieLens. Ratings for each user were converted to a simple binary like/dislike from the original 5-star scale.

movielens

OMDb API
The Open Movie Database

To enhance UI visualization, we downloaded metadata about each movie using the Open Movie Database API. Movie thumbnails are dynamically fetched during interaction with MovieEdge.

Train Recommender System

We evaluated two best-of-breed models: *matrix factorization* and a novel *Word2vec* encoding.

Matrix Factorization, the current state-of-the-art approach, gained popularity during the Netflix Prize Competition. It relies on a sparse ***users X movies*** matrix of ratings and one low rank decomposition which can be used to predict a user's probable rating for unseen movies. While highly effective, it is compute-intensive and requires significant training time. We evaluated 4 different Matrix Factorization approaches, confirming that SVD++ remains the most effective [6].

Word2vec, a popular technique in Natural Language Processing, is built on the idea that words are represented by the context of words surrounding them. Our model treats a users' ratings history as a *sentence* where each movie is a *word*. Sequencing and windowing based on the history of ratings reduced training time and improved results by accounting for changing user tastes.

The Word2vec embedding of movies, essentially a "taste-space", provides the input for training a Ridge Regression model on the embedding vectors with the binarized rating labels as the prediction target. We call the combined approach **Word2vec+Ridge Regression** (W2v+RR) and found it produced performance approaching state-of-the-art accuracy.

Visualize the Movie Taste Space

Our recommender engine relies on a high dimensional model space. To effectively visualize this space for users, we applied dimensionality reduction using t-Stochasitc Neighbor Embedding (t-SNE), creating a two-dimensional model of the movie "taste space". To address rendering and navigational challenges, we implemented zoom techniques which clustered our 27,278 movies using Agglomerative clustering. This worked well on our large dataset and provided smooth zoom in/zoom out functionality for interactive graph exploration. We also deferred rendering of elements for as long as possible, and then rendered only elements currently visible to the user.

Experiments – Model Accuracy

Data was partitioned into 70% training, 15% validation, and 15% test sets, stratified by user. After training candidate models and hyper-parameter tuning using the validation test data we evaluated both the In-Sample and Out-of-Sample Receiver Operating Characteristic Area Under the Curve (AUC).

Method	In-Sample AUC	Out-of-Sample AUC
Baseline (user&movie mean ratings)	0.7644	0.7540
NMF	0.7381	0.7293
NMTF	0.6342	0.6300
SVD	0.8254	0.7822
SVD++	0.8308	0.7992
Word2vec, naive approach	0.6755	0.6588
Word2vec + Ridge Regression	**0.8801**	**0.7906**

Table 1: AUC Scores for Evaluated Models

W2v+RR nearly matched the accuracy of **SVD++** while providing significant speedup in training time (from 12+ hours for SVD to 1 hour for Word2vec). In the end, we selected the novel **W2v+RR** approach given its demonstrated useful properties in other contexts [7].

Our ad hoc testing also showed that the "taste space" model gave too much weight to a movie's year of release. Though this correlates to the generational taste of a user, a broader range of years may be captured while learning similarity by increasing the window range of the Word2vec learner.

Experiments – UI Effectiveness

A survey to gauge user reaction was distributed to classmates and colleagues. After a brief demo video, respondents were asked to review two versions of MovieEdge: one with taste space visualization, and one limited to just movie thumbnails. Response indicated recognition of the innovative nature of our approach (88.89%), but also suggested that our UI needs further refinement before adoption – only 18% were ready to recommend MovieEdge to others.

Question	With Visualization	Just Thumbnails
Preferred version	44.44%	55.56%
Positive Reaction	72.22%	77.78%
Innovative Product	88.89%	72.22%

Discussion & Conclusions

Our work acknowledges and builds upon two recent efforts:

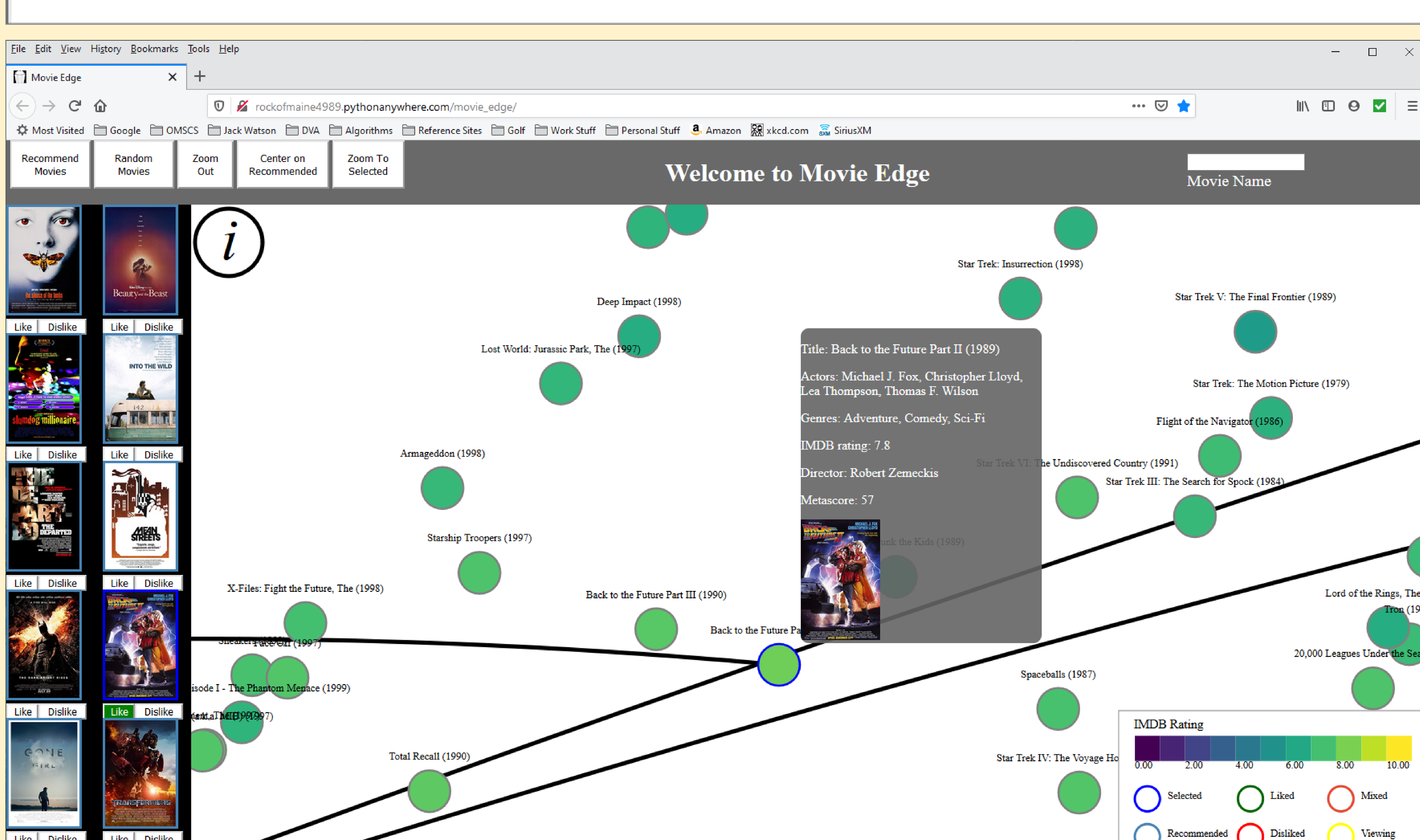
- MovieExplorer, which provides an exploration of ratings taste space [4]
- Embedding Projector, which visualizes learnt embeddings [5]

We have successfully demonstrated two objectives:

- A novel Word2vec recommender system can achieve state-of-the-art performance
- That movie taste space can be presented and interactively navigated

Our open research items include:

- Improving both performance and intuitive navigation within the taste space
- Refining the model & t-SNE clustering to deprioritize the year a film is created



References

[1] Herlocker, Konstan, Terveen, and Riedl. 2004. Evaluating collaborative filtering recommender systems. ACM Transactions on Information Systems (TOIS) 22, 1 (2004), 5–53.

[2] Smith and Linden. 2017. Two decades of recommender systems at Amazon. com. IEEE internet computing 21, 3 (2017), 12–18.

[3] Molnar. 2019. Interpretable Machine Learning -A Guide for Making Black Box Models Explainable. https://christophm.github.io/interpretable-mlbook/, Chapter Interpretability.

[4] Taijala, Willemsen, and Konstan. 2018. Movieexplorer: building an interactive exploration tool from ratings and latent taste spaces. In Proceedings of the 33rd Annual ACM Symposium on Applied Computing. ACM, ACM, 1383–1392.

[5] Smilkov, Thorat, Nicholson, Reif, Viégas, and Wattenberg. 2016. Embedding Projector: Interactive Visualization and Interpretation of Embeddings. arXiv:stat.ML/1611.05469

[6] Koren. 2008. Factorization meets the neighborhood: a multifaceted collaborative filtering model. In Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 426–434.

[7] Mikolov, Sutskever, Chen, Corrado, and Dean. 2013. Distributed representations of words and phrases and their compositionality. In Advances in neural information processing systems.