MOVIE RECOMMENDATION*

NATHAN GRANT , RUDRA GUIN , TYLER HOM , AND STEVEN QIU

Abstract. In this project we explored various techniques, both used in and out of class to try to accurately predict movie ratings based off of the data set used in lab. Using both techniques used in and out of class, we used algorithms which preformed very well on the test data.

Key words. Movie Recommendation, Decision Trees, SVD

AMS subject classifications. 15A18, 62P20

1. Introduction.

Recommendation systems are some of the most important applications in modern day machine learning. Companies such as Amazon must use purchase history and search trends to predict which items you might want to buy. Netflix faces a similar dilemma in that users will watch and rate movies and they must recommend new movies which the users will like. In this project we explored many approaches to movie recommendation to see what offered the best results.

The biggest challenge to recommendation systems is the sparsity to the data. In the data set any given user only rates a small fraction of the total of movies in the database. Such sparse information makes it hard to make very accurate predictions about what movies the users would enjoy. In our projects we used methods such as Decision Trees, XGBoost, Neural Networks, SVD, and a Weighted SVD Nearest Neighbor approach. We evaluate our results based off of two metrics, accuracy of the classification of a users movie rating, and the mean absolute deviation from the correct rating.

The paper is organized as follows. Our main results are in section 2, our new algorithm is in section 3, experimental results are in section 4, and the conclusions follow in section 6.

2. Main results.

The movie recommendation system was evaluated in one of two ways. One way is through human evaluation of the results. Given a user and their rated movies, the algorithm could output movies that they might like. However, this gets tricky because there is no metric to evaluate how good these recommendations actually are. The second option is to take out examples from the training data to use them as test data which was the most promising option. First, we approached the problem using classification models where a users data would be input along with the movie data in order to predict what the user would rate the movie. These techniques resulted in accuracies which were not much better than random.

?? is the first line, and **??** is the last line.

3. Algorithm.

Sed gravida lectus ut purus. Morbi laoreet magna. Pellentesque eu wisi. Proin turpis. Integer sollicitudin augue nec dui. Fusce lectus. Vivamus faucibus nulla nec lacus. Integer diam. Pellentesque sodales, enim feugiat cursus volutpat, sem mauris dignissim mauris, quis consequat sem est fermentum ligula. Nullam justo lectus, condimentum sit amet, posuere a, fringilla mollis, felis. Morbi nulla nibh, pellentesque at, nonummy eu, sollicitudin nec, ipsum. Cras neque. Nunc augue. Nullam vitae quam id quam pulvinar blandit. Nunc sit amet orci. Aliquam erat elit, pharetra nec, aliquet a, gravida in, mi. Quisque urna enim, viverra quis,

*Submitted to the editors DATE.

Funding: This work was funded by the Fog Research Institute under contract no. FRI-454.

TABLE 1
MAD of Methods Used

Neural Network	0.8799
Regression Tree	0.8938
Random Forest	0.8936
Gradient Boosted Tree	0.8938
SVD with Varying KNN	0.6326
SVD with Weighted Varying KNN	.4838

suscipit quis, tincidunt ut, sapien. Cras placerat consequat sem. Curabitur ac diam. Curabitur diam tortor, mollis et, viverra ac, tempus vel, metus.

Our analysis leads to the algorithm in Algorithm 3.1.

Algorithm 3.1 SVD Nearest Weighted Neighbors

Define Epochs: $1 - N$
while $n < Epochs$ **do**
  Find $\nabla E = (\frac{\delta E}{\delta W_i}, \cdots \frac{\delta E}{\delta w_{comp}})$
  Update $W := -\gamma * \nabla E$
end while
return W

Curabitur ac lorem. Vivamus non justo in dui mattis posuere. Etiam accumsan ligula id pede. Maecenas tincidunt diam nec velit. Praesent convallis sapien ac est. Aliquam ullamcorper euismod nulla. Integer mollis enim vel tortor. Nulla sodales placerat nunc. Sed tempus rutrum wisi. Duis accumsan gravida purus. Nunc nunc. Etiam facilisis dui eu sem. Vestibulum semper. Praesent eu eros. Vestibulum tellus nisl, dapibus id, vestibulum sit amet, placerat ac, mauris. Maecenas et elit ut erat placerat dictum. Nam feugiat, turpis et sodales volutpat, wisi quam rhoncus neque, vitae aliquam ipsum sapien vel enim. Maecenas suscipit cursus mi.

4. Experimental results. Quisque facilisis auctor sapien. Pellentesque gravida hendrerit lectus. Mauris rutrum sodales sapien. Fusce hendrerit sem vel lorem. Integer pellentesque massa vel augue. Integer elit tortor, feugiat quis, sagittis et, ornare non, lacus. Vestibulum posuere pellentesque eros. Quisque venenatis ipsum dictum nulla. Aliquam quis quam non metus eleifend interdum. Nam eget sapien ac mauris malesuada adipiscing. Etiam eleifend neque sed quam. Nulla facilisi. Proin a ligula. Sed id dui eu nibh egestas tincidunt. Suspendisse arcu.

Figure 1 shows some example results. Additional results are available in the supplement in **??**.

Maecenas dui. Aliquam volutpat auctor lorem. Cras placerat est vitae lectus. Curabitur massa lectus, rutrum euismod, dignissim ut, dapibus a, odio. Ut eros erat, vulputate ut, interdum non, porta eu, erat. Cras fermentum, felis in porta congue, velit leo facilisis odio, vitae consectetuer lorem quam vitae orci. Sed ultrices, pede eu placerat auctor, ante ligula rutrum tellus, vel posuere nibh lacus nec nibh. Maecenas laoreet dolor at enim. Donec molestie dolor nec metus. Vestibulum libero. Sed quis erat. Sed tristique. Duis pede leo, fermentum quis, consectetuer eget, vulputate sit amet, erat.

5. Discussion of $Z = X \cup Y$. Curabitur nunc magna, posuere eget, vene-

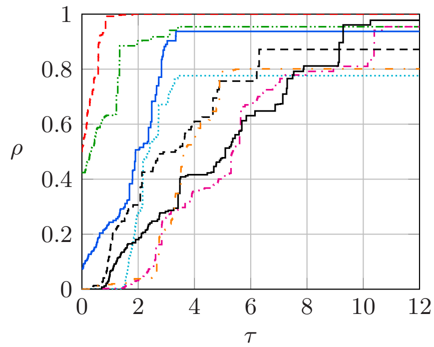

Fig. 1. *Example figure using external image files.*

natis eu, vehicula ac, velit. Aenean ornare, massa a accumsan pulvinar, quam lorem laoreet purus, eu sodales magna risus molestie lorem. Nunc erat velit, hendrerit quis, malesuada ut, aliquam vitae, wisi. Sed posuere. Suspendisse ipsum arcu, scelerisque nec, aliquam eu, molestie tincidunt, justo. Phasellus iaculis. Sed posuere lorem non ipsum. Pellentesque dapibus. Suspendisse quam libero, laoreet a, tincidunt eget, consequat at, est. Nullam ut lectus non enim consequat facilisis. Mauris leo. Quisque pede ligula, auctor vel, pellentesque vel, posuere id, turpis. Cras ipsum sem, cursus et, facilisis ut, tempus euismod, quam. Suspendisse tristique dolor eu orci. Mauris mattis. Aenean semper. Vivamus tortor magna, facilisis id, varius mattis, hendrerit in, justo. Integer purus.

6. Conclusions.

Some conclusions here.

Appendix A. An example appendix.

Aenean tincidunt laoreet dui. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Integer ipsum lectus, fermentum ac, malesuada in, eleifend ut, lorem. Vivamus ipsum turpis, elementum vel, hendrerit ut, semper at, metus. Vivamus sapien tortor, eleifend id, dapibus in, egestas et, pede. Pellentesque faucibus. Praesent lorem neque, dignissim in, facilisis nec, hendrerit vel, odio. Nam at diam ac neque aliquet viverra. Morbi dapibus ligula sagittis magna. In lobortis. Donec aliquet ultricies libero. Nunc dictum vulputate purus. Morbi varius. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. In tempor. Phasellus commodo porttitor magna. Curabitur vehicula odio vel dolor.

Lemma A.1. *Test Lemma.*

Acknowledgments. We would like to acknowledge the assistance of volunteers in putting together this example manuscript and supplement.

REFERENCES

[1] G. H. Golub and C. F. Van Loan, *Matrix Computations*, The Johns Hopkins University Press, Baltimore, 4th ed., 2013.