Recommender System

Hugo Veríssimo
Foundations of Machine Learning 24/25
University of Aveiro
Aveiro, Portugal
hugoverissimo@ua.pt

João Cardoso
Foundations of Machine Learning 24/25
University of Aveiro
Aveiro, Portugal
joaopcardoso@ua.pt

***Abstract*—**

Keywords: MovieLens, GroupLens, Recommender System, Collaborative Filtering

I. Introduction

Since the early 90's that the production rate of multimedia content has increased dramatically (pun intended). Initially, users relied mostly on video store owners, film critics on newspapers, friends. With increasing volumes of content, recommender systems have appeared as data-driven methods to reliably and quickly recommend movies based on the users' own appreciations.

Recommender systems are information tools that provide users with recommended items (ideally) based on their list of preferences [1], [2]. These can be divided in different types, depending on the algorithm and approach to the data. Three categories can be considered: non-personalized, content-based, and collaborative filtering algorithms. Regardless of the algorithm of choice, these face crucial challenges: cold start problem (where it is difficult to tailor recommendations to a user without known preferences, or recommend an item with no reviews); data sparsity (given that most users review few items in the universe of possible items, leaving most of the user/item matrix empty); and scalability (as data grows exponentially, processing becomes evermore expensive and troublesome). These systems have been widely used in many different areas (online shopping, music, books, movie recommendation), and significant investment has gone into developing evermore personalized algorithms. A notable case for this was the Netflix prize competition in 2006, a moment where the research in the field skyrocketed. As a result, it was needed to develop better frameworks for comparison, considering not only the metrics, but data preprocessing, preparation, and routine, in order to ensure reproducibility across models and authors [3].

For that reason, the developed recommender systems is based on one of the most widely used movie databases, MovieLens dataset, for education and development [4]. The choice of this dataset allowed for a based comparison with algorithms from the literature, and facilitate the analysis and interpretation of the results here presented.

II. State of the Art

The field of recommender systems is wide and covers many different algorithms and techniques. In the present work we focus on collaborative filtering using matrix factorization, as it is the main focus of the work here developed.

III. Methodology

A. dados engenharia? tratamento?

os dados contém

100836 ratings, to 9724 movies by 610 users

contudo de modo a tornar o dataset um pouco menor e tp por nao querermos utilizadores novos, com poucas classificacoes de filmes, visto q qnd um utilizador tem poucas classificacoes ha uma metodologia diferente (vet ppts / literatura ?), foram removidos utilizadores com menos de 50 filmes avaliados e foram removidos filmes sem avaliacoes

pelo q ficamos com (9633 filmes, 385 users) e 93812 ratings no total

os dados foram separados em teste e treino, pelo que os ratings levaram split de 20%, resultando em 75049 ratings para teste e 18763 para treino sendo o objetivo ver como a previsao dos ratings fica com os de treino

B. dados análise

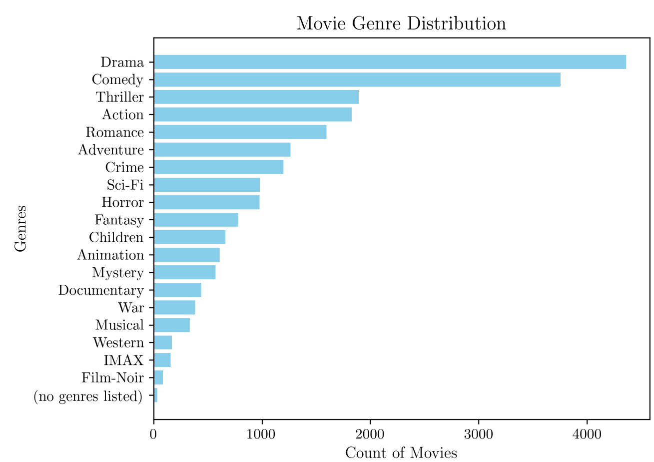

Fig. 1: CAPTIO CAPTION CAPTION

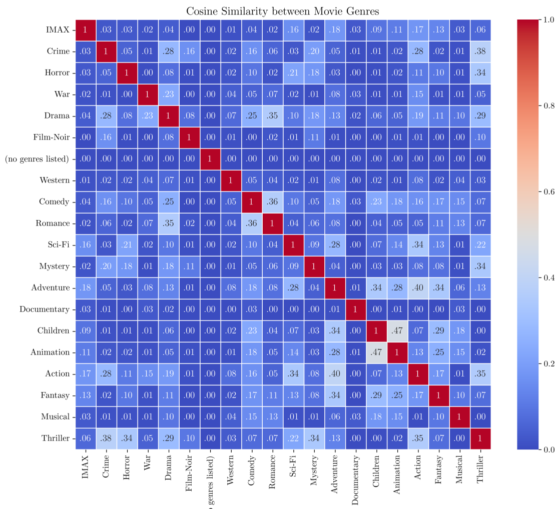

Fig. 2: CAPTIO CAPTION CAPTION

na fig 2 ve se relacoes como

TABLE I: CAPTIO CAPTION CAPTION

genre	**genre**	**sim**
Animation	Children	.47
Action	Adventure	.40
Crime	Thriller	.38
Romance	Comedy	.36
Romance	Drama	.35

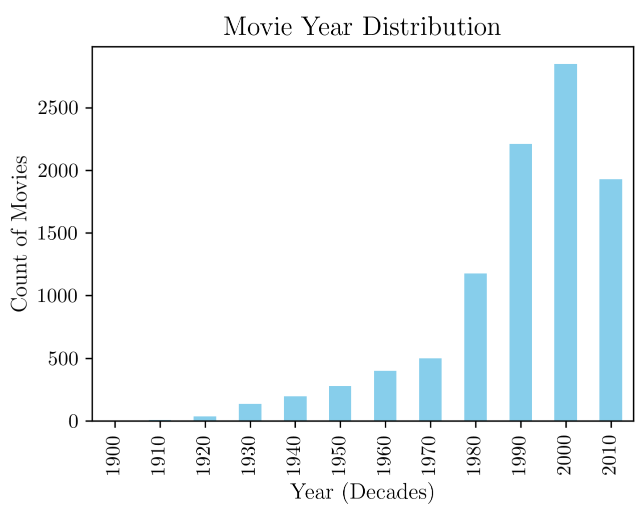

Fig. 3: CAPTIO CAPTION CAPTION

IV. Classification Models

V. Logistic Regression

VI. Discussion

A. Performance Metrics

B. Decision Boundaries

C. Literature Benchmark

VII. Conclusion

Work Load

Both authors contributed equally to the project.

References

[1] J. A. Konstan and J. Riedl, "Recommender systems: From algorithms to user experience," *User Modeling and User-Adapted Interaction*, vol. 22, no. 1-2, pp. 101–123, 2012. [Online]. Available: https://doi.org/10.1007/s11257-011-9112-x

[2] R. Katarya and O. P. Verma, "An effective collaborative movie recommender system with cuckoo search," *Egyptian Informatics Journal*, vol. 18, no. 2, pp. 105–112, 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1110866516300470

[3] A. Said and A. Bellogín, "Comparative recommender system evaluation: benchmarking recommendation frameworks," in *Proceedings of the 8th ACM Conference on Recommender Systems*, ser. RecSys '14. New York, NY, USA: Association for Computing Machinery, 2014, p. 129–136. [Online]. Available: https://doi.org/10.1145/2645710.2645746

[4] F. M. Harper and J. A. Konstan, "The movielens datasets: History and context," *ACM Transactions on Interactive Intelligent Systems (TiiS)*, vol. 5, no. 4, pp. 19:1–19:19, 2015. [Online]. Available: https://doi.org/10.1145/2827872