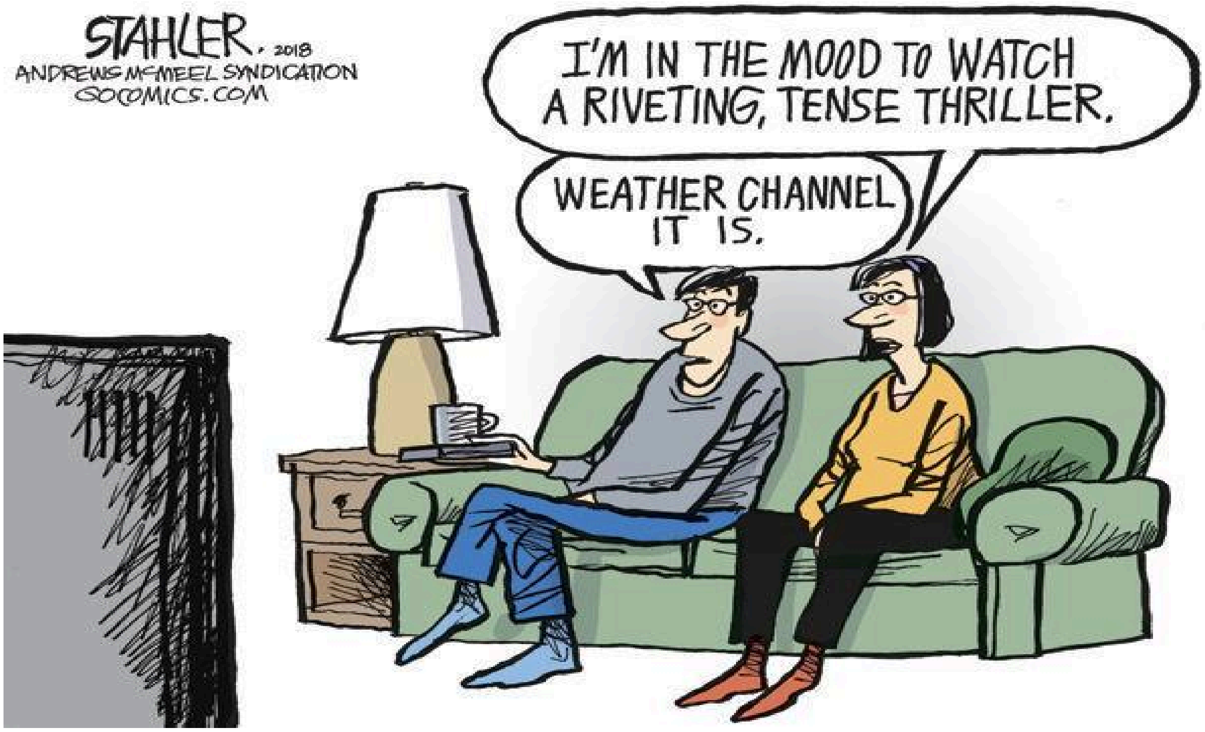
CMPT 419 Project - Spring 2024

FilmFinder



[1]

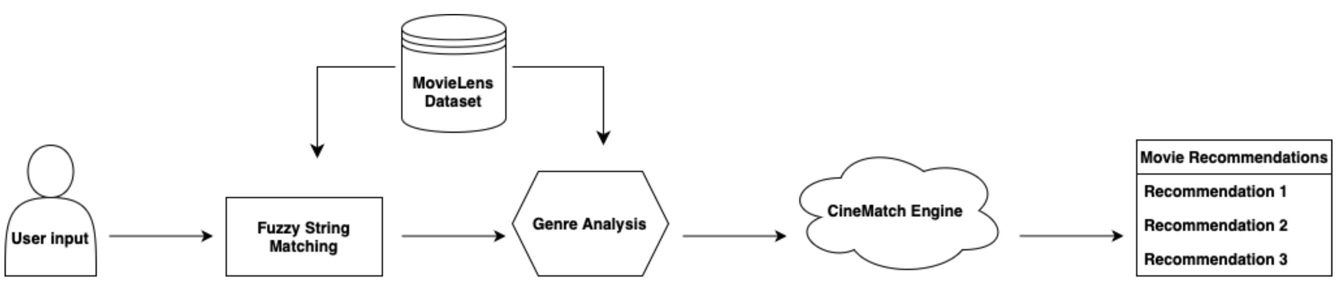
Group Members:
Ali Nikan
Kimia Naghavi

1. ABSTRACT

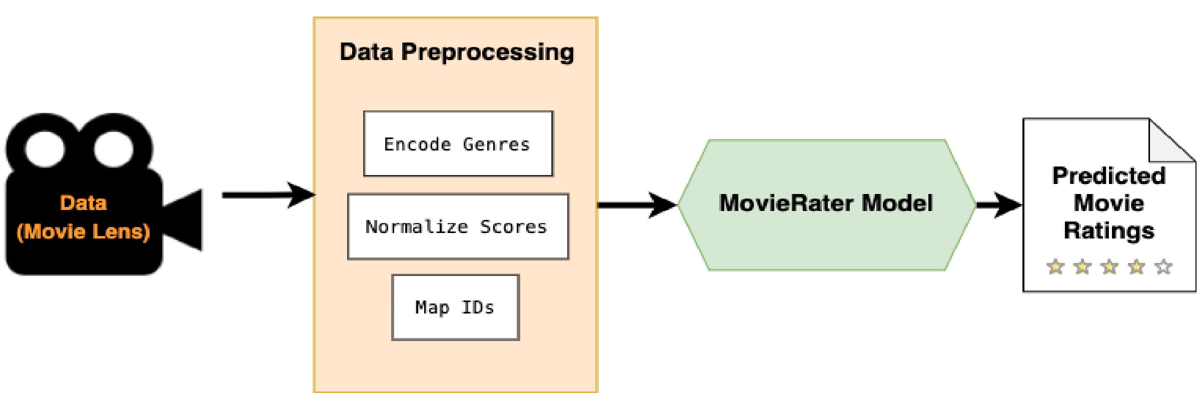
Our project aims to revolutionize movie recommendations by leveraging advanced techniques in natural language processing and machine learning. We have developed a system that analyzes movie genres using sophisticated algorithms, allowing users to input a movie title and receive personalized recommendations based on genre similarity. Unlike traditional recommendation systems that rely solely on user ratings or viewing history, our approach considers the inherent characteristics of movies themselves, providing more accurate and diverse suggestions. By combining TF-IDF vectorization for genre analysis with fuzzy string matching for user input, our system ensures robustness and flexibility in handling various input queries. This innovation not only enhances the user experience by offering tailored recommendations but also opens up new possibilities for applying similar techniques to other domains beyond movie recommendations, such as book recommendations, music suggestions, or even product recommendations in e-commerce platforms. In essence, our initiative is a pioneering effort to use machine learning and natural language processing to provide tailored suggestions and improve consumer satisfaction in the field of entertainment and beyond.

2. VISUAL ABSTRACT

2.1 CineMatch - Movie Recommender



2.2 MovieRater



3. INTRODUCTION

In today's digital world, when there are so many options, recommendation algorithms can help consumers choose relevant and engaging information. Personalized recommendations, whether for a movie, a book, or a product, may dramatically improve customer experience and happiness. In this regard, our study aims to address the difficulty of making effective and tailored movie recommendations by combining cutting-edge natural language processing and machine learning approaches. Our project draws inspiration from *Deep Dive into Content-Based Recommender Systems* article. This article highlights the importance of considering movie attributes beyond user ratings in recommendation algorithms. By incorporating movie characteristics such as genre, cast, director, and plot, the study demonstrates improved recommendation accuracy and diversity [2]. However, existing movie recommendation systems still face challenges in effectively leveraging these attributes to provide personalized recommendations tailored to individual user preferences.

3.1 PROBLEM STATEMENT

Our project seeks to improve upon existing movie recommendation systems, particularly those that rely entirely on user ratings or watching history. While these systems are effective to a certain extent, they often rely heavily on user ratings and viewing history, potentially overlooking key characteristics of movies themselves. As a result, recommendations may not fully align with users' preferences or fail to introduce them to new and diverse content. Additionally, these systems may struggle with handling ambiguous or incomplete user input, leading to less-than-optimal suggestions. To address these limitations, our project aims to develop a recommendation system that not only considers user preferences but also incorporates important movie attributes. By doing so, we aspire to enhance the overall movie-watching experience and empower users to discover content tailored to their individual preferences.

4. DATA GATHERING AND DATA CLEANING

The data utilized for this analysis was sourced from the MovieLens 25M dataset [3], obtained from the GroupLens website. With 25 million movie ratings and one million tag applications applied to a catalog of 62,000 movies by approximately 162,000 users, the dataset offers valuable insights into user preferences and movie characteristics. Additionally, it includes tag genome data, featuring 15 million relevance scores across 1,129 distinct tags.

The data cleaning process started with quality assurance measures to ensure the integrity and reliability of the dataset. Movies were filtered based on a minimum rating threshold, prioritizing those with significant user engagement. This step aimed to boost dataset relevance.

Afterward, the filtered ratings were merged with movie metadata, incorporating information such as titles, genres, and release years. This merge facilitated a comprehensive understanding of the dataset and enabled more meaningful analysis. Additionally, user tags were merged with ratings data, providing insights into user preferences and behaviors. Furthermore, to enrich the dataset with additional context, genome scores and tags were integrated, associating tag names with their respective relevance scores. This enhanced the dataset's utility by providing deeper insights into movie characteristics and attributes. Finally, the cleaned dataset (final_dataset.csv and movies.csv) was exported for further exploration and modeling, which will be discussed in more detail in the methodology section.

5. RELATED WORK

Our goal when building FilmFinder was to use the most recent developments in machine learning and natural language processing (NLP) to create a sophisticated movie recommendation engine. With an emphasis on neural network models, genre-based filtering, and user input processing, this section highlights important academic publications and technical frameworks that have shaped the methodology of our research.

Neural network-based recommendation systems mark a significant advancement in the direction of more precise and customized suggestions. The design of our model is based on a neural collaborative filtering framework that was proposed by He et al. (2017). Our study follows their method of capturing detailed user-item interactions without the need for explicit matrix factorization by employing embedding layers to represent users and movies [4].

Our focus on using NLP techniques—specifically, TF-IDF vectorization—to analyze movie genres is derived from classic works on the subject. Work on the mechanics of text analysis, including the TF-IDF weighting method, has been done by Rajaraman and Ullman (2011), which has proven to be a helpful resource [5]. Their talk on using similar techniques to increase the relevancy of search queries has influenced our approach to genre-based recommendation.

Fuzzy string matching was developed to handle unclear or partial user inputs. It was motivated by real-world recommendation systems' practical challenges. We used Winkler's (1990) discussion of the ways of doing string matching, and its uses and effectiveness in enhancing data quality to improve our system's user experience by making it easier to use and understandable [6]. Therefore, we decided to use Fuzzy to improve our system's user experience.

Moreover, the incorporation of collaborative components with content-based filtering is a hybrid method that Aggarwal (2016) discusses in detail. This paper emphasizes the power of hybrid models in resolving the weaknesses of collaborative or content-based systems, offering a comprehensive overview of several recommender system methodologies [7].

In addition, the incorporation of sentiment analysis into recommendation systems, as looked into by Xiong and Zhang (2023), is a promising method for increasing the level of customization of movie suggestions. Their research highlights the importance of using natural language processing to extract emotional tendencies from user evaluations. This method provides a more comprehensive insight of user preferences than standard ratings. Although the original goal of our project was to use sentiment analysis to enhance our recommendations, we ran into issues with dataset availability that slowed its complete execution. However, the methods and results outlined by Xiong and Zhang (2023) are a crucial point of reference for our research, showing how sentiment analysis may be used to develop more sophisticated and user-focused recommendation systems [8]. Their results provide guidance for FilmFinder's next steps, especially in the area of using emotional analysis to produce more individualized recommendations.

The goal of our project is to promote recommendation system development by placing FilmFinder in the context of these noteworthy papers and contributions. Using deep learning and strong input processing, our approach offers personalized movie recommendations by enhancing predictions with genre analysis.

6. METHODS

In this project, we addressed two primary challenges. Firstly, we developed a recommendation system tasked with suggesting similar movies to users based on their input movie. To achieve this objective, our recommender system identifies movies with the most similar genres to the user's selected film. Secondly, we constructed a rating prediction model designed to forecast the rating of a movie based on its features. To train this model effectively, we utilized diverse information about each movie, including its genre, user ratings, tags assigned by users, and the respective relevance of each of the tags to the movies. We will provide comprehensive descriptions of these challenges and our solutions in more detail.

6.1 CineMatch - Movie Recommender

Our method for creating personalized movie recommendations based on genre similarities is called the CineMatch model. With the use of fuzzy logic and natural language processing, this approach tries to show users movie recommendations that align closely with their interests and preferences based on their own inputs.

6.1.1 DATA PREPROCESSING

Two essential stages are included in our data preprocessing phase to get the movie dataset ready for analysis. First, we load the movie data, which consists of each movie's genre and title, from the MovieLens dataset. Second, we use the vectorization approach known as TF-IDF (Term Frequency-Inverse Document Frequency) to transform the genre information into a numerical representation that can be interpreted by our model. This transformation allows us to do a quantitative analysis of the genre data, finding the genres' significance throughout the movie dataset.

6.1.2 MODEL ARCHITECTURE

CineMatch's architecture is centred on matching user input with our movie database in order to identify matches based on similarities in genre. We base our comparison on the TF-IDF matrix that is produced after preprocessing. Genre-based recommendations are made possible by identifying films whose genres are most comparable to the user-specified film by applying the cosine similarity metric to this matrix.

6.1.3 MODEL COMPILATION AND TRAINING

Our compilation and training method mainly focuses on creating the TF-IDF matrix and similarity calculations since our model depends on precomputed similarity measures instead of a trainable neural network. This preparation guarantees that, when given user input, our model can effectively search and find recommendations.

6.1.4 MODEL EVALUATION

Through user input and participation, we assess CineMatch's performance. Recommendation accuracy and relevance are evaluated by how well the algorithm matches user choices with related movies' genres. Through continuous tests and making sure the movies were actually recommended correctly by either asking the users or actually finding it out by ourselves, we made sure we did a good job in evaluating our system.

6.1.5 VISUALIZATION AND ANALYSIS

CineMatch's visualization features include showing suggested films next to their ratings for genre similarity. By classifying these ratings into descriptive categories (e.g., "Very Similar," "Highly Similar"), we provide users a clear idea of the rationale behind the recommendation of a specific movie. Users find it easier to navigate the suggestions and have more faith in the system as a result of this transparency.

6.1.6 ERROR METRICS CALCULATION

The output of CineMatch is qualitative, therefore standard error measurements like MAE or MSE are not easily applicable. Instead, we focus on metrics like precision and recall, considering the user's acceptance of recommended movies as true positives. To further improve the recommendation's accuracy, the model could go through further quantitative analysis in future iterations based on user rating predictions.

6.2 MovieRater

In the MovieRater Model, we've developed a deep-learning framework aimed to forecast movie ratings. This model relies on multiple input features, including the movie's title, the user's identity, the movie's genre, user-assigned tags, and the relevance of those tags to the movie.

6.2.1 DATA PREPROCESSING

The initial step in constructing the movie rating system involved preprocessing the dataset to facilitate effective model training. Genres were encoded into a binary matrix to capture the categorical nature of movie genres. Additionally, relevance scores were normalized using Min-Max scaling to standardize the range of values, ensuring consistent feature representation across the dataset. Furthermore, user, movie, and tag IDs were mapped to integer indices to enable seamless integration into the embedding layers of the model.

6.2.2 MODEL ARCHITECTURE

The model architecture comprised several key components tailored to capture intricate relationships within the dataset. Embedding layers were employed to learn dense representations of categorical variables such as users, movies, and tags. These embeddings were concatenated with other pertinent features, including relevance scores and genre information, to form a comprehensive feature vector. Batch normalization was applied to stabilize the learning process and accelerate convergence. Dense layers with rectified linear unit (ReLU) activation and dropout regularization were utilized to capture complex patterns in the data while mitigating overfitting risks. Finally, a linear activation function was employed in the output layer to facilitate the regression task of predicting movie ratings.

6.2.3 MODEL COMPILATION AND TRAINING

The model was compiled using the Adam optimizer and Mean Squared Error (MSE) loss function, with Mean Absolute Error (MAE) and MSE serving as evaluation metrics. Training was conducted over 50 epochs with a batch size of 4096 instances. To prevent overfitting and enhance generalization performance, early stopping and learning rate reduction techniques were implemented as callback mechanisms during training.

6.2.4 MODEL EVALUATION

Following model training, evaluation was performed using a separate test dataset. The trained model's predictive performance was assessed through metrics such as test loss (MSE) and test MAE. Predictions were post-processed by rounding to discrete rating values for comparison with actual ratings.

6.2.5 VISUALIZATION AND ANALYSIS

To gain insights into the model's behavior and performance, various visualization techniques were employed. Training and validation loss curves were plotted to monitor model convergence and identify potential overfitting. Histogram plots of actual versus predicted ratings provided a visual assessment of similarity of prediction and real ratings distributions. Furthermore, histogram of residuals (true ratings - predicted ratings) indicated the distribution of prediction errors. We have provided these figures in Section 6.2 (Results - MovieRater).

6.2.6 ERROR METRICS CALCULATION

Mean Absolute Error (MAE), Mean Squared Error (MSE), and Root Mean Squared Error (RMSE) were computed to quantitatively evaluate the model's prediction accuracy, providing a comprehensive assessment of its performance.

7. RESULTS

7.1 CineMatch - Movie Recommender

A close-knit circle of friends and family served as the primary user base for the CineMatch model, validating its success. Their interactions with the system provided valuable insights into its real-world performance and user satisfaction.

7.1.1 USER EXPERIENCE AND SATISFACTION

When we first introduced CineMatch to our group, we encouraged people to try out different inputs, which represented a variety of preferences in movies. The majority of the participants gave extremely positive feedback, saying that the titles that were recommended to them were either new or closely fit their interests. This user satisfaction shows the model's ability to correctly understand and predict user preferences based on genre similarity.

7.1.2 VALIDATION OF RECOMMENDATIONS

We performed an in-depth validation procedure to guarantee the validity of our recommendations. This involved inputting a variety of film titles into CineMatch and then doing independent web research to confirm how similar the suggested films were. The results were positive: the majority of the suggested films had many similarities with the input titles, indicating the model's accuracy in detecting connections based on genre.

7.1.3 REFLECTIONS ON DATABASE AND FUTURE IMPROVEMENTS

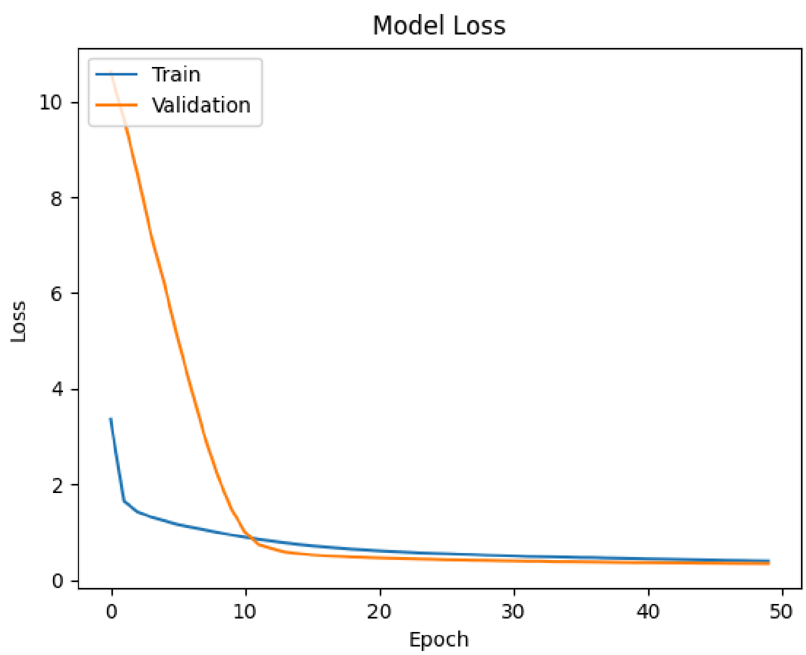
Although our users and ourselves found the recommendations to be satisfactory, we recognize that the existing database, which is mostly genre-focused, is not very comprehensive. This realization has let us realize how much more detailed datasets can do to improve recommendation accuracy. We could significantly improve our system's prediction capability by adding more movie attributes, such user tags, director and cast information, and plot summaries.

7.1.4 CONCLUSION

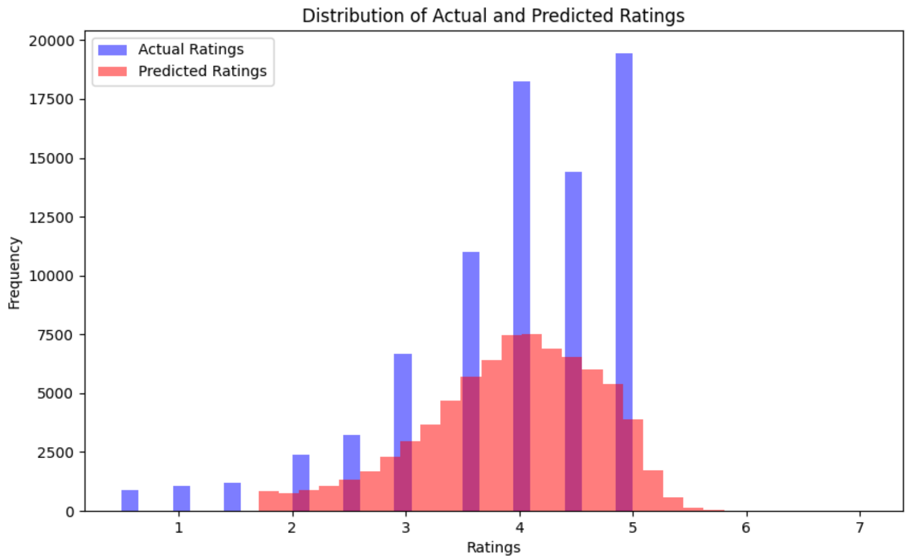
To sum up, the CineMatch model has shown to be an effective tool for finding and suggesting films that suit users' preferences. The core method of the system, which uses genre similarity for customized recommendations, has been confirmed by the positive feedback from our test group. In order to provide even more accurate and specific movie recommendations, we are motivated to grow our database and include more detailed information in the future.

7.2 MovieRater

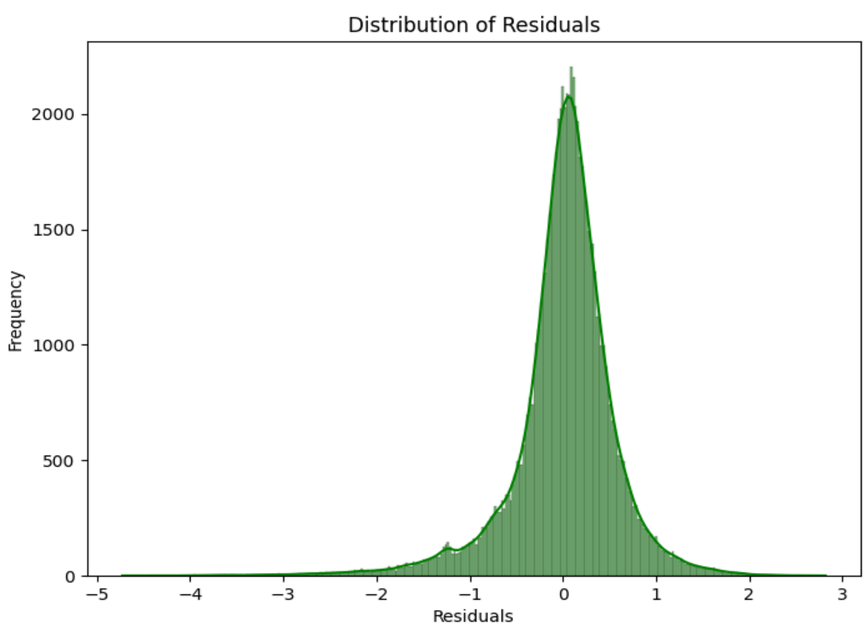
As mentioned in Section 5.2.5 (Visualization and Analysis), we have plotted the training and validation loss curves to monitor model convergence and identify potential overfitting. As shown in the figure below, our model was able to achieve acceptable validation loss after around 20 epochs, and continue to perform better with more epochs. By comparing train vs. validation curves, we can see no overfitting happening as validation loss is continuously decreasing as well as the training loss.



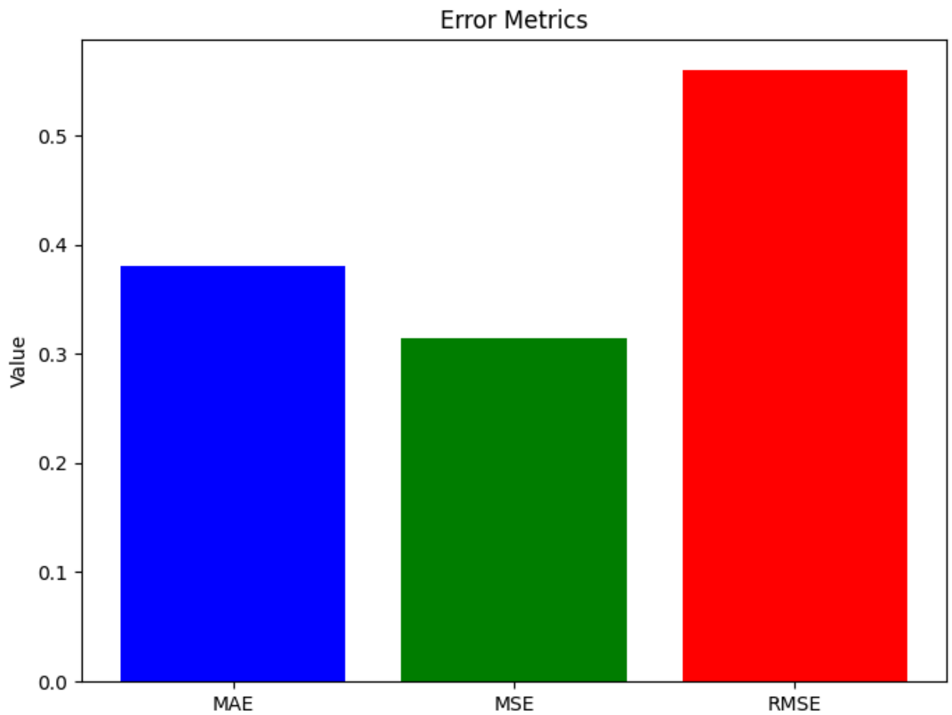
In order to compare our model's prediction with true ratings, we have plotted the distribution of our predictions as well as distribution of true ratings in the figure below. As shown in the figure, our model was not able to predict the low scores (below 1.5) really well.



In order to get a better understanding of our predictions, we have calculated the residuals which is the difference between predictions and true ratings. We have plotted the distribution of these residuals in the figure below. Closer residual values to 0 demonstrate correct predictions. As shown in the figure belos, many of our predictions have residuals near zero, and our residuals distributions are close to a Gaussian distribution with a mean of 0 and standard deviation of ~ 0.2.

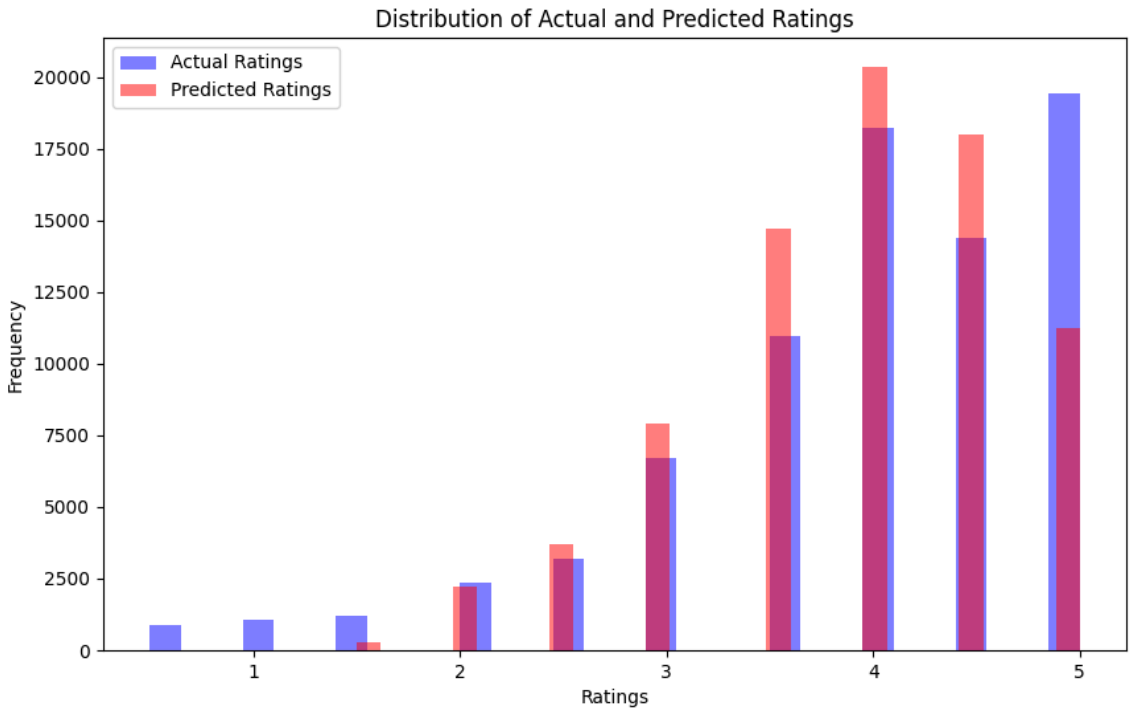


At the end, we also visualized the test set error of our model, which shows how well our model performed on our test dataset. As we can see in the figure below, our test MAE is less than 0.4, which means that on average we are predicting our ratings with ±0.38 value.

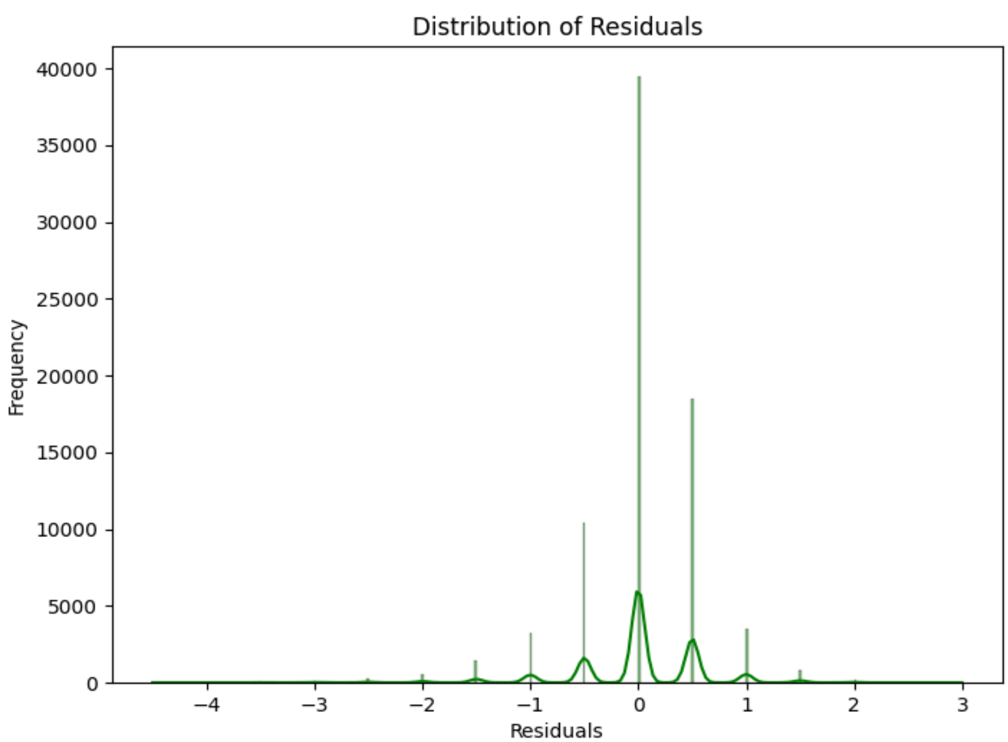


One important point about our model is that it is a regression model, so our output predictions are continues numbers. However, we know that ratings are discrete numbers like: 1, 1.5, 2, …, 4.5, 5. Therefore, we decided to perform a post-processing step and descretize our model's predictions to these defined bins. In the following we have provided same figures as above, with these new descretize predictions.

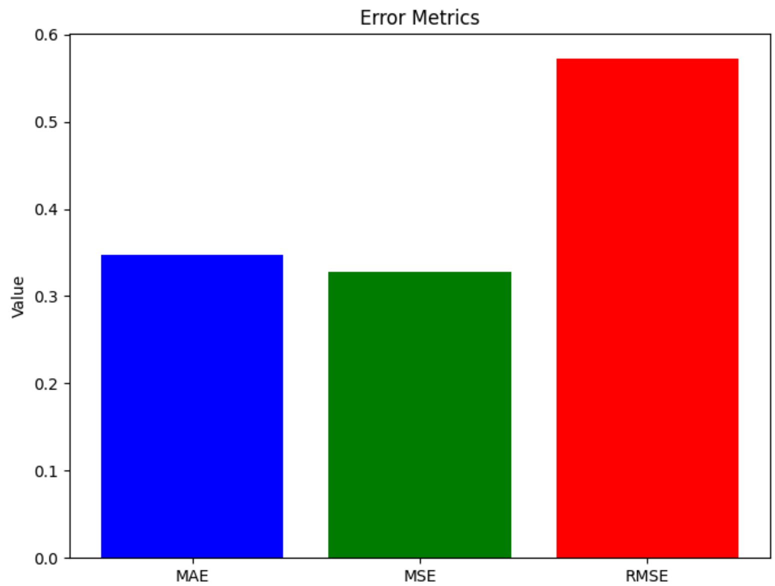
Now, in the distribution of actual vs. predicted ratings figure, we can clearly see the ratings that our model has been overpredicted (2.5, 3, 3.5, 4, 4.5), and those ratings that our model has been underpredicted (0.5, 1, 1.5, 2, 5). So one conclusion is that our model is not performing good on the lower and higher end of ratings.



As we can see in the figure below, most of our residuals are on 0, which is show perfect prediction of rating. The second and third most predictions, are those with extra 0.5 or less 0.5 ratings. We also had some predictions with more errors (1, 1.5, and often 2).



Finally, we have shown test errors after this post-processing step. As shown below, our test MAE, decreased from 0.38 to around 0.35, which means our post-processing step helped our model predict more accurate ratings.



8. DISCUSSION

As a prime example of the idea that good data selection is just as important as algorithmic innovation, FilmFinder's goal to redefine movie recommendations completely comprehends the potential of a data-centric strategy. Consistent with the argument stated by Brown (2020), our project puts great attention to data refinement. This can be seen by the extensive preprocessing steps we take, where data integrity is crucial to guaranteeing that recommendations are both correct and relevant to the user [9]. We used genre data as an example, and performed TF-IDF vectorization, which demonstrates a 'data-centric' commitment.

Although user feedback loops are not presently incorporated in our project, our approach follows the principles of human-centered machine learning as mentioned by Chancellor (2020). We try to understand and satisfy our users' varied preferences even in the lack of explicit feedback options [10].

In keeping with FairML's principles, FilmFinder works for transparent and fair practices [11]. Our tool tries to explain each recommendation by simply showing the similar genres to users so they can understand the rationale behind our genre-based selections. Our commitment to respecting the principles of trustworthy AI as expressed by Schneiderman (2020) ensures that, although not being as advanced as Netflix or other large platforms, our system remains as fair and unbiased as our available dataset allows [12].

8.1 LIMITATIONS AND CHALLENGES

While the proposed movie recommendation system exhibits promising capabilities, several limitations and challenges warrant acknowledgment.

First, there are drawbacks to emphasizing genre similarities. It's a simple method of making movie recommendations, but it doesn't necessarily cover the whole spectrum of what people would find engaging. This approach may ignore users' nuanced preferences and stress the need of taking users' wider values into account while designing [13]. It would be beneficial for future iterations of FilmFinder to comprehend and provide recommendations based on a larger range of movie qualities, rather than just genres.

Our dataset's diversity and quality are also very important. We acknowledge the significance of a representative dataset for fair and inclusive suggestions, taking inspiration from Jernite et al. (2022). Including a larger range of films in our dataset might contribute to the diversity and fairness of our recommendations [14].

Another problem is scalability. We know that as our data grows, so does the need for computational power. As we add more movies and user data, we will need to plan for this growth to maintain the performance of the system [15].

Moreover, the model's interpretability and explainability remain areas of concern. While the model can generate accurate recommendations based on learned patterns in the data, the underlying decision-making process may not be readily interpretable to end-users. This lack of transparency can hinder user trust and acceptance of the recommendations provided.

Lastly, the dynamic nature of user preferences presents an ongoing challenge for recommendation systems. User preferences and viewing habits may evolve over time, necessitating continuous model updates and adaptation to ensure relevance and accuracy in recommendation generation.

Despite these limitations and challenges, the proposed movie recommendation system represents a significant step towards personalized and contextually relevant movie suggestions. Addressing these limitations and overcoming challenges will be crucial for enhancing the system's effectiveness and ensuring user satisfaction in real-world deployment scenarios.

It is not just FilmFinder that faces these challenges. It is part of a larger ecosystem of innovation that is both human and data-centric, where each challenge offers a different chance to progress.

With the understanding of the course readings, we proceeded to develop a recommendation system that is not only technologically proficient but also ethically and humanely attuned.

8.2 FUTURE POSSIBLE WORK

There are many potential development opportunities to improve FilmFinder's user experience. We understand the need to change in order to create a platform that is more inclusive, transparent, and responsive, while also acknowledging the weaknesses of our existing model.

Contextual personalization integration is a critical area for improvement. We can greatly improve our recommendations by taking into account the particular environment in which a user is looking for them, such as the time of day, their mood, or even the weather at the time. With this method, we might adjust suggestions based on dynamic elements that affect viewing choices in addition to static preferences. For example, it could be more appropriate to advise a lighthearted comedy on a rainy day rather than a dark drama. Although this goal was part of our initial project proposal, achieving it became difficult due to practical limitations. Contextual customization, which offers a more complex and fulfilling user experience, is still a top goal for development in the future.

One potential line of future work is to explore the integration of real-time user feedback into recommendation algorithms. By continuously gathering and analyzing user interactions and feedback, recommendation systems can adapt and refine their suggestions in real-time, ensuring that they remain relevant and effective. This approach aligns with the principles of data-centric exploration, as it emphasizes the importance of using dynamic data to inform recommendation decisions. Additionally, incorporating user feedback into recommendation algorithms enhances the human-centric aspect of the system by prioritizing user preferences and satisfaction.

Another promising avenue for future research is to investigate the ethical implications of recommendation systems, particularly in terms of fairness and transparency. As recommendation algorithms play an increasingly influential role in shaping user experiences and decision-making processes, it is essential to ensure that they are fair, unbiased, and transparent. This involves not only addressing issues of algorithmic bias and discrimination but also providing users with clear explanations of how recommendations are generated. By adopting a data-centric and human-centric lens, future work can focus on developing ethical guidelines and best practices for recommendation systems, promoting fairness, accountability, and trustworthiness. Our goal is to establish Fairness and Transparency in our recommendation system, so that FilmFinder becomes an ethical model. As our system develops, we will work to make sure that it supports diversity and inclusivity while also avoiding the ongoing growth of biases. The CARE principles for indigenous data governance [16] will serve as our model for adopting a system that values the variety of our user base and the films we suggest. This involves considering the cultural and societal implications of our suggestions in addition to computational correctness.

Furthermore, future research could explore the integration of diverse datasets and perspectives into recommendation algorithms to enhance their inclusivity and relevance. By incorporating a wide range of cultural, linguistic, and demographic factors, recommendation systems can better cater to the diverse needs and preferences of users from different backgrounds. This aligns with the human-centric aspect of recommendation systems, as it recognizes and respects the unique perspectives and experiences of individual users. By embracing diversity and inclusivity, future recommendation systems can foster a more equitable and enriching user experience.

9. CONNECTION TO CLASS THEME

FilmFinder is deeply rooted in a data-centric approach, where every aspect of the project revolves around leveraging data to enhance movie recommendation systems. From the initial stages of data collection and preprocessing, where meticulous attention is given to ensuring the quality and fairness of the dataset, to the extensive data exploration component, which aims to identify key characteristics that impact recommendation quality, the project exemplifies a commitment to understanding and harnessing the inherent value of data. By employing data valuation approaches and experimenting with different dataset types, FilmFinder seeks to uncover insights into how data selection decisions influence the effectiveness of recommendation systems. Furthermore, the project's emphasis on reporting and documentation underscores the importance of transparently documenting methodologies and findings, ensuring that the insights gained from the data-centric exploration can be shared and utilized to advance the field of recommendation systems. In essence, FilmFinder's holistic approach to leveraging data exemplifies the core tenets of a data-centric philosophy, where decisions and insights are driven by a deep understanding of the underlying data, ultimately leading to more accurate and personalized movie recommendations.

10. OVERALL ARTIFACT QUALITY

Ali’s Accomplishments:

- Implemented data processing techniques on a large dataset sourced from MovieLens.
- Initiated implementation of a recommendation system to generate personalized movie recommendations based on user’s preferences.
- Gathered related articles and papers from SFU’s library database and found the related works.
- Contributed in writing technical documentation for the project.

Kimia’s Accomplishments:

- Gathered the data needed for analysis.

- Ensured data cleanliness and relevance for subsequent analysis.
- Leveraged methodologies to implement the MovieRater model which include user ratings, tags assigned by users, and the respective relevance of each of the tags to the movies as well as genre to predict the movie's rating.
- Contributed in writing technical documentation for the project.

GitHub Repository Lik: https://github.com/kimianaghavi/FilmFinder

10. REFERENCES

[1] Editorial cartoon: *Weather thriller*. (n.d.). https://santamariatimes.com/opinion/editorial/editorial-cartoon-weather-thriller/article_964ea2bc-998e-55d8-86b3-0f12acbf1295.html

[2] Torabi, N. (2023, August 26). *Deep dive into content-based recommender systems:* Unveiling the power of attribute-based... Medium. https://neemz.medium.com/deep-dive-into-content-based-recommender-systems-unveiling-the-power-of-attribute-based-48d27fc98812

[3] *Movielens 25M dataset*. GroupLens. (2021, March 2). https://grouplens.org/datasets/movielens/25m/

[4] He, X., Liao, L., Zhang, H., Nie, L., Hu, X., & Chua, T. (2017). *Neural Collaborative Filtering*. In Proceedings of the 26th International Conference on World Wide Web (WWW '17).

[5] Rajaraman, A., & Ullman, J. D. (2011). *Mining of Massive Datasets*. Cambridge University Press.

[6] Winkler, W. E. (1990). *String Comparator Metrics and Enhanced Decision Rules in the Fellegi-Sunter Model of Record Linkage*. In Proceedings of the Section on Survey Research Methods (pp. 354-359). American Statistical Association.

[7] Aggarwal, C. C. (2016). *Recommender Systems: The Textbook*. Springer.

[8] Xiong, W., & Zhang, Y. (2023). *An intelligent film recommender system based on emotional analysis*. PeerJ Computer Science, 9, e1243. https://doi.org/10.7717/peerj-cs.1243

[9] Brown, S. (2020). Why it's time for 'data-centric artificial intelligence'. MIT Sloan. Retrieved from https://mitsloan.mit.edu/ideas-made-to-matter/why-its-time-data-centric-artificial-intelligence

[10] Chancellor, S. (2020). Toward Practices for Human-Centered Machine Learning. Communications of the ACM. Retrieved from https://cacm.acm.org/research/toward-practices-for-human-centered-machine-learning/

[11] FairML: Principles of Fairness in Machine Learning. (2020). Retrieved from https://fairmlbook.org/introduction.html

[12] Schneiderman, B. (2020). Human-Centered Artificial Intelligence: Reliable, Safe & Trustworthy. International Journal of Human–Computer Interaction. Retrieved from https://www.tandfonline.com/doi/full/10.1080/10447318.2020.1741118

[13] Zhu, Q., Azaria, A., & Riedl, M. O. (2018). Value-Sensitive Algorithm Design: Method, Case Study, and Lessons. Proceedings of the ACM on Human-Computer Interaction, 2(CSCW), Article 194. https://dl.acm.org/doi/10.1145/3274463

[14] Jernite, Y., Nguyen, H., Biderman, S., Rogers, A., Masoud, M., Danchev, V., ... & Mitchell, M. (2022). Data Governance in the Age of Large-Scale Data-Driven Language Technology. In Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency (FAccT '22). Association for Computing Machinery, New York, NY, USA, 2206–2222. https://dl.acm.org/doi/abs/10.1145/3531146.3534637

[15] Kaplan, J., McCandlish, S., Henighan, T., Brown, T. B., Chess, B., Child, R., ... & Amodei, D. (2020). Scaling Laws for Neural Language Models. arXiv. Retrieved from https://arxiv.org/abs/2001.08361

[16] Carroll, S. R., Garba, I., Figueroa-Rodríguez, O. L., Holbrook, J., Lovett, R., Materechera, S., ... & Hudson, M. (2020). The CARE principles for indigenous data governance. Data Science Journal, 19(43). Retrieved from https://datascience.codata.org/articles/10.5334/dsj-2020-043/