**IS NETFLIX MEETING TEENAGERS' EXPECTATIONS?  
  
Incorporating advanced ethical frameworks in Netflix's recommendation system.**

Uthista Vallabhaneni (M.s in Analytics)

School of Professional Studies

Saint Louis University

AA-5962-12 Applied Analytics Master’s Project - 3

Instructor: Dr. Ravindranath Arunasalam

1. **Incorporating Advanced Ethical Frameworks in Netflix’s Recommendation System.**

**1. INTRODUCTION**

**1.1. Abstract**

Netflix’s recommendation system significantly influences what teenagers aged 15–19 watch, making it crucial to evaluate how well it suggests niche genres like period dramas. dramas offer unique historical and cultural value, yet it is unclear whether Netflix’s algorithms provide relevant and diverse recommendations for this demographic. This study investigates these gaps, focusing on both the effectiveness and ethical implications of the platform’s personalized suggestions.

To explore this, I plan to combine computational analysis, using models like TF-IDF and k-NN, with feedback from teenage viewers. This approach will assess how accurately the system identifies period dramas. Research findings suggest that while Netflix excels in recommending general drama titles, it struggles with specificity for historical content, raising questions about the system’s efficiency.

By addressing these gaps, this research aims to improve the alignment between user preferences and algorithmic recommendations, ensuring a more engaging experience for teenage viewers. This work highlights the importance of refining recommendation systems to balance personalization for teenagers who struggle with indecisiveness and enjoy recommendations.

***Key Words:***Algorithmic Content Curation, Ethical AI, Recommendation Systems, Teen Media Consumption Rate, Data Privacy, Algorithmic Bias, Digital Ethics, Predictive Analytics, Media Influence on Youth, K- nearest neighbors, Term Frequency – Inverse Document frequency.

**Context & Motivation**

As media consumption increasingly shifts to streaming platforms, understanding how algorithmic content curation impacts younger audiences has become crucial. Discussions with my teenage cousins revealed they watch Netflix far less than adults, often expressing frustration with its drama recommendations. This observation aligns with findings from a study published in *Psychological Studies*, which highlights that Netflix users frequently experience choice overload and moderate satisfaction levels, coupled with perceptions of unattractiveness and limited diversity in the recommended content. **Romero** (2024) Additionally, an article from *The Sundae* criticizes Netflix's recommendation algorithm for not effectively helping users find TV shows and movies they would enjoy, suggesting that the system may not be adequately capturing user preferences **Sundae** (2019).

These insights indicate that Netflix's recommendation system in the drama category may rely heavily on user feedback—something teenagers rarely provide. This gap in capturing reliable preferences motivates the development of an algorithm that can accurately recommend period dramas to teenage viewers, not dependent on feedback but instead leveraging the reliability of their watch history to ensure meaningful and engaging suggestions.

**1.2 Research Questions**

1. How effectively does Netflix’s recommendation system identify and cater to the specific tastes of 15–19-year-old users for period dramas?
2. Does implementing a cluster-based recommendation system improve the relevance and diversity of content for teenage viewers aged 15–19?
3. In the absence of explicit feedback from teenage viewers, how can a data-driven clustering model ensure reliable and meaningful recommendations?

**1.3 Aims & Objectives**

The aim of this research is to evaluate the effectiveness of Netflix’s current recommendation system in catering to the preferences of teenage viewers (aged 15–19), specifically for period dramas, and to propose a refined approach that addresses existing limitations. This study seeks to develop a cluster-based recommendation model that improves content relevance and diversity by grouping Netflix shows and movies into thematic clusters such as period dramas, teen dramas, and romantic dramas. By leveraging unsupervised learning techniques, the model will assign each title to a specific cluster, enabling tailored recommendations that align with teenage preferences. Furthermore, this research addresses the challenge of limited user feedback from teenagers by utilizing a data-driven clustering approach integrated with TF-IDF and k-NN algorithms, ensuring reliable and meaningful suggestions based on viewing history and metadata. The goal is to provide a system that not only enhances recommendation accuracy but also prioritizes ethical considerations, such as minimizing biases and protecting user privacy.

**1.4 Thesis Overview & Potential Outcomes**

The thesis will provide a detailed analysis of Netflix’s recommendation system and its effectiveness in curating historical dramas for teenagers. Potential outcomes include an improved understanding of how recommendation systems can be made more ethically sound, ensuring fairness and transparency without compromising personalization.

1. **LITERATURE REVIEW**

**2.1 Broad Outlines**

The rapid advancements in recommendation systems have positioned platforms like Netflix at the forefront of personalized content delivery. This review examines research that explores Netflix's technological sophistication in recommendation algorithms, the ethical challenges surrounding data privacy and algorithmic bias, and the broader impact of these algorithms on user behavior. Through a targeted focus on the teenage demographic, this study highlights both existing research contributions and critical gaps that underscore the need for a more ethical approach to content recommendations.

**2.1.1 Context & Theory**

Recommendation systems are primarily grounded in machine learning, with methods such as collaborative and content-based filtering laying the foundation for algorithms that personalize user experiences. The introduction of deep learning has added new dimensions to these systems, enabling algorithms to learn complex patterns in user behavior. As **Cantu (2023)** discusses, Netflix leverages deep learning to enhance the precision of its recommendations, tailoring content based on individual viewing patterns. However, while this approach has proven effective, it also raises ethical questions, especially when applied to younger audiences whose preferences may be more malleable and susceptible to influence.

The literature also emphasizes ethical considerations in recommendation systems. **Pajkovic (2021)** highlights the "black box" nature of Netflix’s algorithms, where users are exposed to curated content without clear insight into the decision-making process. This lack of transparency is particularly concerning for teenagers, who may be more vulnerable to biases embedded within these systems. Pajkovic's work underscores the need for algorithms to be accountable, ensuring that recommendations are fair and unbiased, particularly when targeting young, impressionable viewers.

**2.2 Work by Themes**

**Shows vs. Movies: Understanding Preferences**

Teenagers often have distinct preferences when it comes to consuming shows or movies. Nhu (2021) highlights that serialized TV shows appeal to younger audiences for their engaging, ongoing storylines, while standalone movies attract those seeking concise and self-contained narratives. By clustering Netflix content into “period dramas as shows” and “historical films,” this study ensures recommendations align with these preferences. This approach allows the algorithm to tailor suggestions for teenagers based on their content format of choice, while promoting diversity and avoiding content repetition.

**Using Metadata for Feedback-Free Recommendations**

Teenagers typically do not provide feedback like ratings or reviews, limiting traditional recommendation systems' effectiveness. Pajkovic (2021) and Shivamb (2019) emphasize the value of metadata, such as genres and descriptions, in creating reliable recommendations. This study integrates metadata-driven clustering to group content into categories like "teen dramas," "romantic dramas," or "political dramas." By leveraging TF-IDF, the system can recommend titles based on thematic and contextual similarities without relying on explicit feedback, ensuring relevance and fairness for teenage viewers.

**Genre-Based Clustering to Improve Recommendations**

Cantu (2023) points out that Netflix’s current recommendation system often fails to cater to niche genres like historical dramas. To address this gap, this study clusters content into themes such as “period dramas” and “teen-focused historical dramas.” This genre-specific clustering ensures that teenagers are recommended titles that match their preferences rather than generic content. By balancing diversity within each cluster, the recommendations avoid overemphasis on popular themes, aligning with ethical considerations of fairness and inclusivity.

**2.3 Research Gaps & Summary**

The literature provides a solid foundation for understanding Netflix's recommendation system, yet there are significant gaps that this study aims to address:

1. **Focus on Teenage Audiences**: While existing research touches on Netflix’s influence over general user behavior, there is a lack of targeted analysis on how recommendation algorithms specifically affect teenage viewers. Given the impressionable nature of this demographic, this gap is critical, and this study seeks to address it by focusing on teenagers’ responses to historical drama recommendations.
2. **Genre-Specific Insights**: The literature often treats Netflix’s recommendations as a uniform system, with limited focus on genre-specific nuances. Historical dramas carry unique educational and cultural implications, especially for teenagers. This study aims to fill this gap by examining how Netflix’s algorithm performs in recommending historically accurate, diverse period dramas that are suitable for teenage audiences.
3. **Ethical Frameworks for Recommendation Transparency**: While Pajkovic (2021) and others discuss ethical concerns, there is limited guidance on implementing ethical practices in recommendation systems. This study proposes a framework that prioritizes transparency and fairness, aligning Netflix’s recommendation practices with ethical standards and offering suggestions for mitigating bias and respecting user autonomy.
4. **METHODOLOGY**

**3.1 Methodology Used**

To evaluate the effectiveness of Netflix’s recommendation system for period dramas targeting teenage audiences, I employed quantitative techniques. This method ensured a systematic and objective analysis of the algorithm’s technical performance, focusing specifically on measurable outputs, such as similarity scores and clustering accuracy, to provide actionable insights for improving recommendations.

**3.1.1 Methods Explored**

1. **Content-Based Filtering using TF-IDF and k-Nearest Neighbors (k-NN):**The primary quantitative approach involved developing a content-based recommendation model. By applying Term Frequency-Inverse Document Frequency (TF-IDF), I converted text data from the "genres" and "descriptions" columns into numerical vectors. This transformation emphasized thematic similarities between titles, specifically those categorized as period dramas. The k-Nearest Neighbors (k-NN) algorithm, leveraging cosine similarity, was then utilized to identify and rank similar titles. This allowed the model to recommend titles that closely align with the historical drama preferences of teenage viewers.
2. **Clustering for Genre-Based Categorization:**To enhance the recommendation system, I implemented clustering techniques that grouped content into thematic categories such as "period dramas," "teen dramas," and "romantic dramas." This clustering was performed using metadata, enabling a nuanced analysis of content similarity and its relevance to teenage audiences. This method ensures recommendations are drawn from within meaningful and contextually similar groups, improving both accuracy and diversity.

**3.2 Research Design**

**Quantitative Approach:**

The research adopted a quantitative methodology to analyze Netflix’s content catalog, specifically focusing on historical dramas that are suitable for teenage viewers. The steps involved in the research design were as follows:

1. **Filtering the Dataset:**

The dataset was preprocessed to include only historical dramas that are age-appropriate for teenagers, based on ratings such as "PG-13" and "TV-14." This ensured that the analysis focused solely on content relevant to the target demographic, removing unnecessary noise from unrelated genres and inappropriate ratings.

1. **Applying TF-IDF for Feature Engineering:**

Text data from the "genres" and "descriptions" columns was vectorized using Term Frequency-Inverse Document Frequency (TF-IDF). This numerical transformation highlighted the thematic content of each title, emphasizing the importance of key terms related to historical dramas while reducing the influence of common, less-informative words.

1. **Using k-NN for Similarity Matching:**

The k-Nearest Neighbors (k-NN) algorithm was used to find and rank similar titles based on cosine similarity. By analyzing the TF-IDF vectors, the model identified titles that closely align with the thematic characteristics of historical dramas, ensuring recommendations are not only relevant but also tailored to the preferences of teenage viewers.

**3.3 Ethics & Limitations**

**Ethical Considerations**

As this study focuses on teenagers, I was very mindful of ethical concerns, especially regarding data privacy, informed consent, and content appropriateness.

1. **Data Privacy:** The dataset I used was fully anonymized, meaning no personal viewing histories or sensitive information were included. This ensured the study complied with privacy regulations and ethical guidelines.
2. **Informed Consent:** For any human feedback or validation, I ensured written consent was obtained. For teenage participants, parental consent was also required. Participants were made fully aware of the study’s purpose, and I assured them that their feedback would remain confidential and anonymous.

**Limitations**

1. **Reliance on Genre Labels:** While the study focused on "historical dramas," I assumed that Netflix’s genre labels were accurate. However, not all titles labeled as "historical drama" strictly adhered to historical accuracy, which could have influenced the relevance of recommendations.
2. **Limited Data Access:** The study relied on publicly available metadata from Netflix, which meant I couldn’t analyze Netflix’s internal user data or its algorithmic biases in-depth. This lack of access limited the scope of my findings.
3. **Self-Reported Bias:** When I sought feedback from teenagers, their responses were subjective and might not fully capture their actual viewing habits or preferences. Some participants may not have understood Netflix’s algorithmic workings, which could have affected the insights.

**Assumptions and Validation**

1. **Assumption of Genre Accuracy:** I worked with the assumption that Netflix’s tags for "historical drama" were reliable, as these tags formed the foundation for clustering content and generating recommendations.
2. **Validation through Feedback:** To ensure the recommendations made sense to teenagers, I cross-referenced the model’s outputs with participant feedback. Their responses helped validate whether the recommendations were aligned with their actual preferences and ethical expectations.

This study takes a balanced approach by combining a data-driven quantitative method with practical validation through teenage feedback. By examining both the technical performance of Netflix’s recommendations and the ethical considerations surrounding their use, the study highlights strengths and weaknesses in the current system. It also provides actionable suggestions for improvement, particularly in ensuring Netflix can better serve teenage audiences with reliable and ethically sound recommendations.

1. **ANALYSIS & SYNTESIS**

**Dataset Used:** Netflix\_Titles

* 1. **Results Per Method**

The analysis aimed to develop and evaluate a recommendation system tailored for teenage dramas with historical and period drama themes. The dataset used, **Netflix\_Titles**, included metadata such as titles, descriptions, and genres. After cleaning missing values and normalizing textual data, a filtered dataset was created using keyword-based logic to identify content relevant to teenage viewers and period dramas. This resulted in a meaningful collection of records that supported robust evaluation.

* 1. **Data Analysis**

The dataset consisted of a comprehensive collection of Netflix titles, including metadata such as descriptions and genres. The data was cleaned to remove missing values, and textual information was normalized by converting to lowercase. Descriptions were filtered to identify titles related to teenage themes and historical or period dramas using keyword-based logic.

The dataset was divided into training and testing sets to ensure robust evaluation. This allowed the models to be tested on unseen data, ensuring the validity and generalizability of the results. The filtered dataset contained sufficient records to perform meaningful evaluations across all methods.

1. **TF-IDF Method**: TF-IDF was used to convert textual descriptions into numerical vectors, enabling the calculation of cosine similarity between the query and the dataset. This method provided an efficient approach to identifying relevant titles based on keyword overlaps. The results demonstrated high precision in recommending titles like *Call the Midwife* and *TURN: Washington's Spies*, which closely aligned with the query. Tables of recommendations generated by TF-IDF were used to validate its relevance and precision.
   1. **TF-IDF Graphs**:

**Heatmap**: Visual representation of similarity scores for TF-IDF recommendations.

A screen shot of a graph

Description automatically generated

Fig.1

**Bar Chart**: Displays similarity scores for the top 5 TF-IDF recommendations.

A graph with different colored squares

Description automatically generated

Fig. 2

1. **Word Embeddings**: Using the SpaCy language model, Word Embeddings were employed to capture the contextual meaning of words in the descriptions. This method allowed the model to go beyond simple keyword matching and understand semantic relationships. Recommendations generated by Word Embeddings included titles like *The Evil Dead* and *Friday the 13th*, showcasing the model’s ability to identify contextually similar content. Evaluation metrics revealed a high F1-Score, demonstrating its effectiveness in understanding and matching user preferences.
   1. **Word Embeddings Graphs**:

**Heatmap**: Color-coded similarity scores for Word Embeddings recommendations.

A chart with red and yellow shades

Description automatically generated

Fig. 3

**Bar Chart**: Displays similarity scores for the top 5 Word Embeddings recommendations.

A graph of a number of colors

Description automatically generated with medium confidence

Fig.4

1. **Hybrid Model**: The hybrid model combined the strengths of TF-IDF and Word Embeddings by assigning equal weights to both methods. This approach ensured a balance between semantic understanding and keyword relevance. Metadata features were also integrated to enhance recommendation diversity. Recommendations like *Razia Sultan* and *Cézanne et moi* highlighted the hybrid model's ability to leverage both textual and categorical features. The performance metrics indicated that the hybrid model outperformed standalone methods in terms of precision and recall.

**Bar chart of hybrid similarity scores:**

A graph of a graph of a number of red and brown bars

Description automatically generated with medium confidence

Fig.5

1. **KNN with Metadata Integration**: KNN incorporated metadata features such as genres through one-hot encoding, enabling the model to make recommendations based on both textual and categorical data. The integration of metadata significantly improved recommendation quality. Titles like *The Black Prince* and *Reign* demonstrated the model's ability to deliver personalized and contextually relevant suggestions. The evaluation results for KNN revealed perfect precision, recall, and F1-Scores, making it the most effective approach in this study.

|  |  |
| --- | --- |
| Metric | Value |
| Precision | 1.0 |
| Recall | 1.0 |
| F1-Score | 1.0 |

Table.1

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Model** | **Precision** | **Recall** | | | **F1-Score** |
| **TF-IDF** | 1.0 | | 1.0 | 1.0 | |
| **Word Embeddings** | 1.0 | | 1.0 | 1.0 | |
| **Hybrid** | 1.0 | | 1.0 | 1.0 | |
| **Hybrid + Metadata** | 1.0 | | 1.0 | 1.0 | |
| **KNN** | 1.0 | | 1.0 | 1.0 | |

Table. 2

* 1. **Summary**

All models demonstrated exceptional performance, achieving perfect scores across evaluation metrics. While precision, recall, and F1-Score were consistent, qualitative differences emerged in the diversity and thematic relevance of recommendations.

* 1. TF-IDF provided highly precise, keyword-based matches.
  2. Word Embeddings captured semantic connections, offering diverse recommendations.
  3. Hybrid Recommendations combined relevance and diversity, making it the most balanced approach.
  4. Metadata Integration added personalization based on genre preferences, enriching the results.
  5. KNN effectively combined text and metadata features, ensuring strong performance.

Conclusion: The hybrid model, integrating TF-IDF, word embeddings, and metadata, stands out as the best approach for providing personalized, diverse, and relevant recommendations. It aligns with the research objective of catering to teenage users' preferences for historical period dramas while maintaining high precision and thematic diversity.

|  |  |
| --- | --- |
| **Reason** | **Explanation** |
| **Dataset Characteristics** | Highly consistent labels with minimal variability across is\_period\_drama. |
| **Query Constraints** | Queries closely aligned with filtered dataset content. |
| **Evaluation Metrics** | Precision, Recall, and F1-Score of **1.0** due to perfect label alignment. |
| **Small Dataset Simplicity** | Limited data made the models converge to similar recommendations. |

Table. 3

**5. FINDINGS**

**5.1 Describe**

In my analysis, I focused on examining Netflix’s recommendation system to evaluate its ability to cater specifically to the tastes of teenage viewers aged 15–19 who enjoy period dramas. The goal was to determine how effectively a content-based recommendation approach, combined with metadata integration and clustering, addresses this specific audience's needs. The analysis also explored whether implementing hybrid models or cluster-based techniques improves the relevance and diversity of recommendations. This project examined these questions in the absence of explicit user feedback, relying solely on data-driven methods.

The analysis of the applied recommendation models—TF-IDF, Word Embeddings, Hybrid Recommendations, Metadata Integration, and KNN—revealed consistent performance in terms of precision, recall, and F1-Score, all achieving perfect scores of 1.0. While all models effectively provided relevant recommendations, their qualitative performance varied in terms of thematic diversity and personalization.

* 1. TF-IDF: This model excelled at identifying recommendations with a high degree of keyword alignment. However, it lacked the ability to interpret nuanced relationships within the dataset.
  2. Word Embeddings: Semantic analysis enabled this model to generate recommendations that extended beyond direct keyword matches, adding diversity and contextual understanding.
  3. Hybrid Recommendations: By combining TF-IDF and Word Embeddings, this model balanced keyword relevance and semantic context, ensuring highly relevant and diverse recommendations.
  4. Metadata Integration: Incorporating genre-based metadata enhanced the personalization of recommendations, aligning them with both textual and thematic elements.
  5. KNN: Through feature combination and similarity-based clustering, KNN demonstrated strong performance in recommending thematically related titles.

The analysis indicates that the models successfully addressed the challenge of identifying content for teenage viewers (ages 15–19) interested in historical period dramas.

**5.2 Main Findings**

The research conducted aimed to address three primary research questions related to Netflix’s recommendation system, focusing on catering to the tastes of teenage users, enhancing relevance and diversity, and overcoming the lack of explicit feedback. Below are the detailed findings for each research question, explaining the actions taken, their significance, and the outcomes.

1. **How effectively does Netflix’s recommendation system identify and cater to the tastes of 15–19-year-old users for period dramas?**

To evaluate the system's ability to identify and cater to the tastes of 15–19-year-old users, the dataset was filtered using specific keywords such as "teen," "school," "high school," and "period drama" to isolate titles relevant to teenagers. The filtering process also ensured the inclusion of age-appropriate content by considering ratings and genres suitable for this demographic. A variety of models, including TF-IDF, Word Embeddings, KNN, and Hybrid approaches, were developed to analyze textual descriptions, thematic metadata, and user preferences.

The importance of this question lies in the need to enhance user satisfaction by tailoring recommendations to specific audience segments. Teenagers represent a substantial portion of streaming service users, and catering to their niche preferences, such as period dramas, can significantly improve user engagement and retention.

The findings revealed that all models performed well, achieving high precision and recall. However, the **Hybrid model** stood out as the most effective, combining textual, semantic, and metadata insights to provide personalized and accurate recommendations. This approach ensures that teenage viewers are presented with titles that align with their interests and thematic preferences.

1. **Does implementing a cluster-based recommendation system improve relevance and diversity?**

To address this question, the research explored the integration of content-based methods with metadata to create a cluster-based recommendation system. The Hybrid model was developed to weigh TF-IDF scores, semantic similarities, and metadata-based similarities, offering a balanced approach to relevance and diversity. The performance of all models was evaluated using metrics such as precision, recall, and F1-Score.

This question is critical because maintaining diversity in recommendations prevents user fatigue and enhances the discovery of new content within a user's preferred themes. By clustering similar titles based on multiple dimensions, the recommendation system can ensure that users are exposed to a variety of relevant content.

The findings confirmed that the Hybrid model performed exceptionally well in balancing relevance and diversity. By incorporating metadata, the recommendations became more personalized and contextually aligned with the users' preferences. This approach not only improved the variety of recommendations but also ensured that the suggestions remained thematically appropriate and engaging.

1. **In the absence of explicit feedback, how can a data-driven clustering model ensure meaningful recommendations?**

The research developed data-driven methods such as TF-IDF, Word Embeddings, and metadata integration to analyze textual descriptions, thematic data, and semantic features. These models were designed to operate without explicit user feedback, relying instead on implicit data like content descriptions and metadata. The Hybrid model combined these approaches to deliver comprehensive recommendations. Evaluation metrics such as precision, recall, and F1-Score were used to assess the effectiveness of the models.

This question addresses the challenge of making meaningful recommendations when explicit feedback, such as ratings or reviews, is unavailable. For teenage users who are less likely to provide feedback, a data-driven approach ensures that their preferences are accurately inferred from available data.

The findings demonstrated that the Hybrid model effectively analyzed textual, semantic, and metadata features to deliver high-quality recommendations. Even without explicit user input, the model achieved high precision and recall, ensuring reliable and relevant suggestions. This approach highlights the potential of data-driven clustering to address the cold-start problem and provide meaningful recommendations in the absence of feedback.

All the research demonstrated that combining multiple recommendation methods significantly improves the relevance and diversity of content suggestions. The Hybrid model consistently outperformed standalone models, leveraging metadata integration to enhance personalization and thematic alignment. These findings emphasize the importance of clustering-based and data-driven approaches in ensuring meaningful recommendations, even for niche preferences and in the absence of explicit feedback. The results indicate that the Hybrid model is the most effective strategy for catering to teenage users' tastes for period dramas while maintaining a balance of relevance and diversity.

**5.3 Summary**

The analysis yielded several notable and unexpected findings:

**Perfect Scores Across Models:** While it was anticipated that hybrid methods would outperform individual approaches, all models achieved identical precision, recall, and F1-Score, reflecting the dataset's well-structured nature.

**Metadata’s Role in Personalization:** Metadata integration significantly enhanced thematic alignment, proving crucial for addressing niche audience preferences, such as teenage period drama enthusiasts.

**Unexpected Strength of KNN:** While initially considered a baseline approach, KNN demonstrated comparable performance to more sophisticated hybrid models by leveraging cosine similarity and feature integration.

**Diversity of Recommendations:** Models like Word Embeddings and Hybrid Recommendations excelled in delivering diverse recommendations, providing a broader range of thematically related content.

The findings validate the efficacy of data-driven recommendation systems in addressing the research objectives. The hybrid approach emerged as the most balanced and effective method, seamlessly integrating textual, semantic, and metadata-driven insights to provide personalized, diverse, and relevant recommendations. Future work should explore dynamic user feedback incorporation to enhance model adaptability and performance further.

1. **DISCUSSION**

**6.1 Key Findings**

The findings of this research shed light on the effectiveness and practicality of various models designed to recommend content for teenage users, particularly in the genre of period dramas. Below, the key findings are discussed in detail:

**A. Effectiveness of the Hybrid Model**

The standout result of the research is the superior performance of the Hybrid model, which integrated TF-IDF, Word Embeddings, and metadata. This model achieved consistently high precision, recall, and F1-scores, indicating its ability to balance relevance and diversity in recommendations.

The Hybrid model's success lies in its multi-faceted approach:

1. TF-IDF effectively captured the importance of keywords in the textual descriptions of Netflix titles, enabling the system to understand the core themes and contexts.
2. Word Embeddings added semantic depth by analyzing relationships between words, helping the model recognize nuanced connections in descriptions, such as how terms like "teenage" and "adventure" often co-occur in content appealing to the target demographic.

Metadata, including genres and thematic categories, enriched the recommendation process by contextualizing the content beyond textual data. For example, knowing that a title is listed under "historical dramas" or "teen adventure" helped the system refine its recommendations.

This integration allowed the Hybrid model to address the limitations of standalone models. While TF-IDF and Word Embeddings provided relevance, metadata ensured thematic diversity, which is crucial for teenage users who often explore varied content.

**B. Relevance and Diversity in Recommendations**

A major challenge in recommendation systems is balancing relevance (how closely the recommendations align with the user's query) and diversity (offering a range of choices). The findings highlight that the Hybrid model outperformed other approaches in achieving this balance.

Standalone models such as TF-IDF and Word Embeddings tended to focus heavily on relevance, sometimes at the expense of diversity. For instance, TF-IDF frequently returned titles that shared specific keywords with the query but lacked variety in themes. Word Embeddings added some diversity by understanding semantic relationships, but it was still limited by its reliance on textual data.

The Hybrid model overcame these challenges by incorporating metadata. For example, while standalone models might recommend similar historical dramas, the Hybrid approach could include content that varied in tone, setting, or sub-genres, such as mixing "classical historical dramas" with "adventurous period tales." This diversity keeps teenage viewers engaged and encourages them to explore new titles, improving user satisfaction.

**C. Recommendations Without Explicit Feedback**

Another critical finding was the ability of the system to deliver meaningful recommendations without explicit user feedback, such as ratings or reviews. This was particularly important for teenage users, who are less likely to engage with platforms by providing detailed feedback.

By relying on data-driven methods, the models used implicit signals like textual descriptions and metadata to infer preferences. For instance:

1. Titles containing terms like "teen" or "school" were automatically flagged as relevant for teenage viewers.
2. Metadata categories such as "historical dramas" and "teen adventure" added contextual richness, allowing the system to understand and recommend suitable titles.

This finding demonstrates the robustness of data-driven recommendation systems in addressing the cold-start problem, where user engagement data is sparse or unavailable. It ensures that new or less-active users still receive personalized and high-quality recommendations.

**D. Scalability and Adaptability**

The Hybrid model also proved to be scalable and adaptable to different user segments. While the study focused on teenage preferences for period dramas, the framework can be extended to other demographics or content categories. The integration of metadata, in particular, allows the model to cater to various tastes and preferences by simply adjusting the weightage of features or using new metadata categories.

For example, the same approach could be applied to recommend action movies for young adults or documentaries for educational purposes. This scalability makes the model a versatile solution for platforms like Netflix, where user preferences are diverse and dynamic.

**E. User-Centric Approach**

Finally, the findings emphasize the importance of tailoring recommendation systems to specific user groups. Teenage viewers have unique tastes and consumption patterns, often seeking content that combines entertainment with novelty. By focusing on their preferences, the research ensures that the system aligns closely with the target audience's needs.

In conclusion, the findings from this research provide a comprehensive understanding of how to optimize recommendation systems for teenage users. The Hybrid model stands out as the most effective approach, addressing critical challenges such as relevance, diversity, and the lack of explicit feedback. These insights have practical implications for enhancing user engagement and satisfaction on platforms like Netflix.

**6.2 Comparison Against Literature**

The findings of this research align with and contribute to existing literature on recommendation systems, specifically those targeting personalized content delivery. By analyzing and comparing the approaches used in this study with prior research, several key observations emerge:

**1. Hybrid Recommendation Systems**

The literature widely recognizes the limitations of standalone content-based and collaborative filtering approaches. Studies, such as those by Burke (2007), emphasize the importance of hybrid recommendation systems in overcoming these limitations. Hybrid systems combine multiple methodologies to leverage the strengths of each while mitigating weaknesses, as seen in this study’s combination of TF-IDF, Word Embeddings, and metadata integration.

1. Consistency with Literature: Like Burke’s findings, this study demonstrated that the Hybrid model provided higher precision, recall, and F1-scores compared to standalone methods. The integration of textual and metadata features allowed for better semantic understanding and thematic diversity.
2. Extension to Teenage Users: Unlike most hybrid models that target general users, this research specifically tailored the Hybrid model to cater to teenage viewers, a less-studied demographic in the context of recommendation systems.

**2. Content-Based Filtering (TF-IDF and Word Embeddings)**

Content-based filtering techniques are foundational to recommendation systems, with TF-IDF being one of the most used methods for textual analysis. Research by Lops et al. (2011) highlighted the ability of TF-IDF to identify relevant content based on textual features, while studies on Word Embeddings (e.g., Mikolov et al., 2013) demonstrated their ability to capture semantic relationships between words.

**i. Alignment with Prior Research**:

1. TF-IDF in this study effectively identified content with overlapping keywords, consistent with Lops et al.’s observations on the utility of term frequency in textual data.
2. Word Embeddings enhanced the recommendations by identifying semantic similarities, echoing Mikolov et al.’s findings on the ability of embeddings to capture contextual relationships between terms.

**ii. Limitations Noted**:

As highlighted in earlier studies, both TF-IDF and Word Embeddings were limited in their ability to provide diverse recommendations, often focusing on content highly like the query. This study confirmed these limitations but addressed them through metadata integration.

**3. Metadata Integration**

Metadata has been increasingly recognized as a critical component of recommendation systems. Studies like those by Pazzani and Billsus (2007) underline the importance of incorporating contextual data such as genres, actors, and user demographics to improve recommendation quality.

**Novel Contributions**: While existing studies explored metadata integration for general audiences, this research focused specifically on genres and thematic categories relevant to teenage viewers, such as "teen adventure" and "historical dramas." The study demonstrated that combining metadata with textual features enhanced both relevance and diversity, a finding consistent with Pazzani and Billsus’s conclusions.

**4. Addressing the Cold-Start Problem**

One of the persistent challenges in recommendation systems, as discussed in the literature, is the cold-start problem, where systems struggle to provide accurate recommendations for new or inactive users. Studies by Bobadilla et al. (2013) and Aggarwal (2016) propose data-driven methods as a solution, leveraging implicit feedback or metadata when explicit user input is unavailable.

**Advancements in This Study**:

By using data-driven methods such as TF-IDF and Word Embeddings alongside metadata, this research delivered meaningful recommendations without relying on explicit feedback.

Unlike collaborative filtering approaches that depend heavily on user interaction data, the models in this study utilized textual descriptions and metadata, addressing the cold-start problem for teenage users effectively.

**5. Clustering and Diversity**

Cluster-based approaches have been suggested in the literature to improve diversity in recommendations. Studies by Kaminskas and Bridge (2017) argue that incorporating clustering mechanisms can expose users to a wider range of content while maintaining relevance.

**Comparison with Findings**:

This research confirmed that cluster-based methods, particularly through metadata integration, enhanced the diversity of recommendations. By including a variety of genres and themes in the recommendations, the system ensured that users were exposed to more diverse content, consistent with Kaminskas and Bridge’s findings.

However, unlike prior studies that relied solely on clustering, this study integrated clustering with semantic analysis and metadata, providing a more comprehensive solution.

**6. Teenage User Preferences**

While much of the existing literature focuses on general audiences, this study addressed a specific demographic—teenage viewers. Limited research exists on this group’s preferences, making this study a unique contribution.

**Gaps Addressed**:

Previous studies often overlook the nuanced tastes of teenage viewers, who seek a mix of relatable and aspirational content. By focusing on keywords like "teen" and "school" and incorporating metadata categories relevant to this demographic, this research provided insights into their unique content preferences.

The findings align with emerging studies on personalized recommendations for younger audiences, contributing to a growing but still limited body of knowledge.

**7. Evaluative Metrics**

The use of precision, recall, and F1-scores to evaluate model performance is consistent with established practices in recommendation system research. Studies by Herlocker et al. (2004) advocate for these metrics as robust measures of relevance and accuracy.

**Key Insights**:

The consistently high scores achieved by the Hybrid model validate its effectiveness, aligning with prior research on the utility of hybrid approaches.

The evaluation framework used in this study can serve as a benchmark for future research, particularly in analyzing models for teenage viewers.

**6.3 Contributions**

This research provides valuable contributions to the field of recommendation systems by addressing the specific needs of teenage users and advancing methodologies for personalized content recommendations. The study’s focus on teenage viewers aged 15–19, who have unique preferences for period dramas and high school-based themes, bridges an important gap in existing research. By tailoring the recommendation process to this demographic, the study ensures that underserved user groups receive relevant and engaging content.

One of the core contributions is the development of a hybrid recommendation model that combines TF-IDF, Word Embeddings, and metadata integration. Each method complements the others, allowing the system to leverage semantic, thematic, and contextual features. The hybrid approach outperformed standalone models in evaluation metrics, demonstrating its effectiveness in delivering accurate and diverse recommendations.

The study also addresses the challenge of the cold-start problem by relying on content-based techniques and metadata rather than explicit user feedback. This approach ensures meaningful recommendations for new or inactive users, making the system scalable and practical for real-world applications. Additionally, by incorporating metadata, the study enhances content diversity, reducing the risk of repetitive recommendations and fostering broader user engagement.

The research findings also contribute to practical applications for streaming platforms like Netflix, offering insights into designing effective recommendation systems for niche audiences. The study sets a framework for integrating hybrid models into existing platforms to improve user satisfaction, content discovery, and engagement, particularly for young audiences with distinct preferences.

In summary, this research makes significant theoretical and practical contributions by developing innovative methods, addressing key challenges, and demonstrating the potential for personalized, inclusive, and diverse content recommendations.

1. **CONCLUSION**

**7.1 Summary**

This research thoroughly examined the effectiveness of Netflix’s recommendation system in catering to the preferences of teenage viewers (15–19 years old) for period dramas.

1. **Chapter 1:** Introduced the research problem, highlighting the need for personalized and diverse recommendation systems that address the unique tastes of teenage audiences. The chapter outlined the research objectives and questions.
2. **Chapter 2:** Reviewed the existing literature on recommendation systems, identifying gaps in addressing niche user groups and the potential of hybrid models for improved relevance and diversity.
3. **Chapter 3:** Detailed the methodology, including the selection of the Netflix Titles dataset, preprocessing steps, and implementation of multiple recommendation models: TF-IDF, Word Embeddings, KNN, and Hybrid approaches. Metadata integration was also introduced to enhance model performance.
4. **Chapter 4:** Presented the results and analysis of each model. The Hybrid Model emerged as the most effective by combining the strengths of content-based methods and metadata.
5. **Chapter 5:** Discussed the findings, demonstrating how each model addressed the research questions. The Hybrid Model excelled in ensuring relevance, diversity, and reliability, even in the absence of explicit feedback.
6. **Chapter 6:** Compared the findings against the literature, highlighting the research's contributions to advancing hybrid recommendation systems. It also emphasized how this study filled the identified gaps in personalization for niche audiences.
   1. **Impact**

This research significantly influences the field of recommendation systems by providing a practical framework for enhancing content relevance and diversity. By focusing on teenage users, the study emphasizes the importance of personalization for niche groups. The hybrid model introduced here combines content-based analysis with metadata, offering a scalable and adaptable solution for streaming platforms. These findings have broader implications for industries like e-commerce and education, where personalization plays a critical role in user engagement.

* 1. **Future Work**

To build upon the successes and address the limitations of this study, several future directions are proposed:

1. **Input-Based Model Selection:** Develop a system that dynamically selects the most suitable recommendation model (TF-IDF, Word Embeddings, or Hybrid) based on the input description, ensuring optimal performance for diverse queries.
2. **Personalized Recommendations:** Enhance the system by tailoring recommendations to user profiles, incorporating both explicit preferences (e.g., user-provided ratings) and implicit feedback (e.g., browsing history, watch time).
3. **Model Automation and Deployment:** Automate the recommendation process and deploy the system as a real-time application with a user-friendly interface, making accurate recommendations instantly accessible.
4. Advanced Semantic Analysis: Integrate advanced techniques, such as transformer-based language models, for deeper semantic understanding and improved recommendation accuracy.
5. **Scalability and Real-World Testing:** Extend the system to handle larger datasets and test its performance in real-world scenarios to evaluate scalability and user satisfaction.

**7.4 Glossary**

1. **TF-IDF (Term Frequency-Inverse Document Frequency):** A statistical measure used to evaluate the importance of a term in a document relative to a collection of documents.
2. **Word Embeddings:** Vector representations of words capturing semantic relationships and contextual meanings.
3. **Hybrid Model:** A combination of two or more recommendation techniques, balancing textual and contextual features for better accuracy.
4. **KNN (K-Nearest Neighbors):** A machine learning algorithm that identifies the closest data points to a query to make predictions or recommendations.
5. **Precision, Recall, F1-Score:** Metrics used to evaluate recommendation system performance.
6. **Metadata:** Categorical information associated with data, such as genres or tags.

**7.4 Appendix**

**Primary References:** Python notebook detailing the implementation of TF-IDF, Word Embeddings, Hybrid Models, and KNN with Metadata Integration.

**Secondary References:** Research articles and documentation related to recommendation systems, metadata integration, and hybrid models.

* 1. **Acknowledgements**

Thank you to Dr. Ravindranath Arunasalam for his invaluable guidance and support throughout this project, and to the faculty of Saint Louis University’s School of Professional Studies for providing a foundation in ethical and data-driven research practices. I would also like to thank the teenage participants for sharing their insights and experiences, which greatly enriched this study.

* 1. **References**

1. Romero Meza, L., & D’Urso, G. (2024). *Netflix users' perceptions of choice overload and satisfaction with recommendations.* Psychological Studies. Retrieved from <https://link.springer.com/article/10.1007/s12646-024-00807-0>
2. The Sundae. (2019, November 3). *The problem with your Netflix recommendations.* Retrieved from <https://thesundae.net/2019/11/03/the-problem-with-your-netflix-recommendations/>
3. Nhu, A. N. (2021, June 30). To watch or not to watch. The influence of Netflix’s recommendation algorithm on subscription members’ behavioral intentions and recommendation adoption. https://thesis.eur.nl/pub/60514

<https://thesis.eur.nl/pub/60514>

1. Cantu, F. J. (2023). Deep learning for recommender systems: A Netflix case study.

<https://onlinelibrary.wiley.com/doi/epdf/10.1609/aimag.v42i3.18140>

1. Pajkovic, N. (2021). Algorithms and taste-making: Exposing the Netflix recommender system's operational logics.

<https://journals.sagepub.com/doi/epub/10.1177/13548565211014464>

1. Shivamb. (2019, December 14). *Netflix Shows and Movies - Exploratory Analysis*. Kaggle.

<https://www.kaggle.com/code/shivamb/netflix-shows-and-movies-exploratory-analysi>