Recommendation System Analysis & Modeling - Overview and Summary

# 📌 Project Overview

This project focuses on developing a Recommendation System capable of providing personalized suggestions across multiple domains such as e-commerce, media, and subscription services. By leveraging user interaction data (views, add-to-cart actions, transactions) and item metadata, the system delivers personalized, scalable, and efficient recommendations. The methodology follows the CRISP-DM framework, ensuring a structured and reproducible workflow.

# 📊 Project Summary

1. Business Understanding:  
 - Goal: Enhance personalization to improve engagement and sales.  
 - Challenges: Scalability, cold-start problems, and balancing accuracy with diversity.  
  
2. Data Understanding:  
 - Datasets include user events (views, add-to-cart, purchases), item properties, and category hierarchy.  
 - Views dominate interactions, but only a small fraction convert to transactions.  
  
3. Data Preparation:  
 - Events cleaned and mapped to weighted scores.  
 - Item properties pivoted into metadata for content-based filtering.  
 - Train-test split performed at the user level to ensure realistic evaluation.  
  
4. Modeling:  
 - Collaborative Filtering (item-item similarity using cosine similarity).  
 - Content-Based Filtering (TF-IDF on item properties).  
 - Hybrid approaches explored to combine both.  
  
5. Evaluation:  
 - Precision@K and Recall@K used to assess recommendation quality.  
 - Results show item-item collaborative filtering captures co-purchase behavior effectively.  
  
6. Key Insights:  
 - Views dominate the funnel; purchases are a small subset.  
 - Popular items follow a skewed distribution (Pareto principle).  
 - Peak activity observed in specific time ranges (evening/lunch hours).  
 - Content-based similarity helps mitigate cold-start issues.  
  
7. Recommendations:  
 - Promote items during peak interaction hours.  
 - Use hybrid models to balance collaborative and content signals.  
 - Improve engagement by diversifying recommendations.  
 - Continuously retrain models with fresh data.  
 - Implement monitoring dashboards for real-time recommendation performance.

# 👩‍💻 Author

Jessica William