

INTEREST-BASED CONTENT RECOMMENDATION SYSTEM

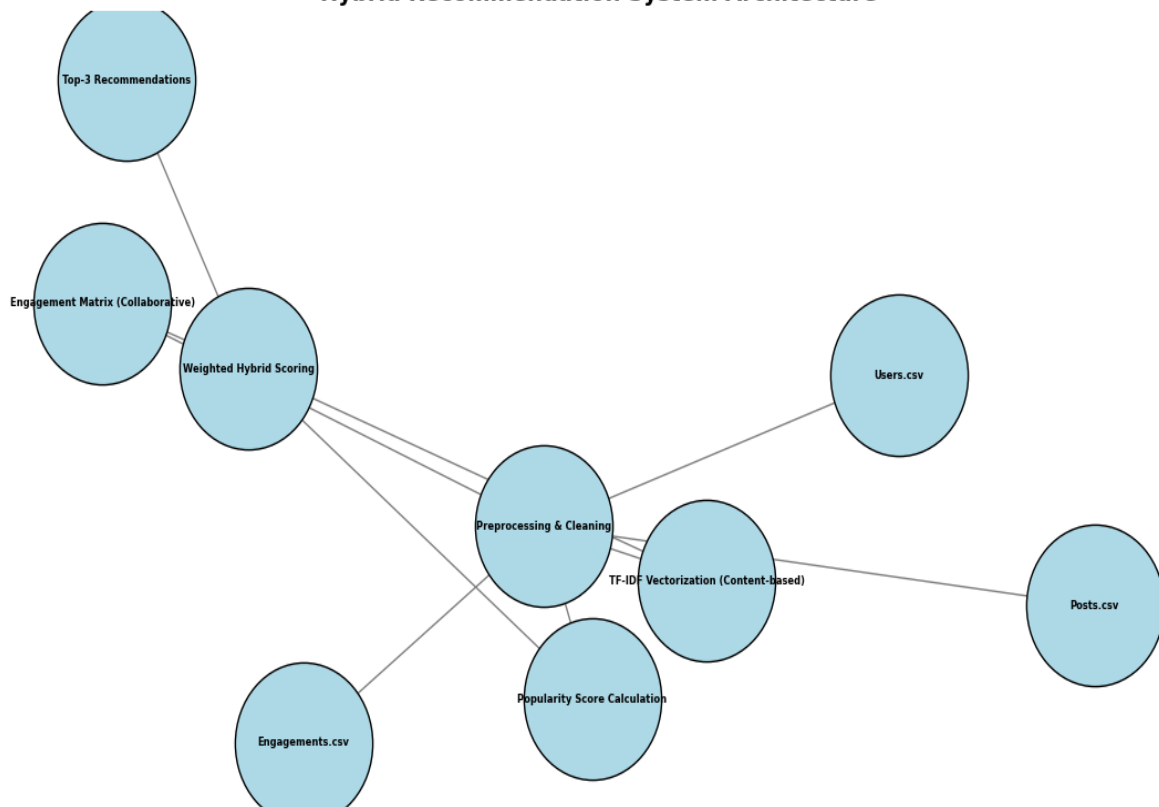
Introduction

The project aims to design and develop a personalized content recommendation system that suggests the top three posts to each user based on their unique preferences and engagement behaviour. In today's digital platforms, information overload makes it difficult for users to identify relevant content, and a recommender system bridges this gap by providing meaningful and personalized suggestions. Our system combines multiple approaches rather than depending on a single method to improve accuracy, coverage, and reliability.

Objectives

The main objective of this project is to build a **hybrid recommendation engine that balances personalization with popularity so that each user receives posts which are both interesting to them and contextually relevant**. The system is expected to integrate content-based similarity, collaborative filtering, and popularity-based ranking in order to generate recommendations that adapt dynamically to different users.

Hybrid Recommendation System Architecture



Methodology and Approaches

We adopted a hybrid methodology consisting of three major approaches. The first was **content-based filtering**, where the declared interests of users and post metadata such as tags and textual descriptions were vectorized using TF-IDF, and cosine similarity was applied to measure the closeness between a user's profile and available posts. This helped in aligning recommendations with user-declared preferences. The second approach was **collaborative filtering**, where user-post engagement patterns were converted into a matrix, and item-item similarity was calculated. This allowed the system to recommend posts similar to those previously engaged with, even if they did not directly match the user's stated interests. The third approach was **popularity-based ranking**, where posts were assigning scores based on normalized engagement counts, ensuring that widely appreciated content also received visibility. All three approaches were then combined using weighted aggregation, where each method contributed proportionally to the final recommendation score.

```
Hybrid Top-3 Recommendations for U1:
  post_id      text
77      P78  sports art video
21      P22  sports art audio
0        P1  sports food video
Scores: [0.38818993 0.38601534 0.23754012]
```

Implementation

The datasets were pre-processed to remove inconsistencies and normalize textual attributes. TF-IDF vectorization was used for handling tags and content features, while the user-item matrix was derived from engagement logs. Cosine similarity functions and matrix computations were implemented using Python libraries such as scikit-learn and pandas. Popularity scores were calculated by normalizing likes, shares, and comments, and these were blended with similarity scores to generate a final ranked list of posts for each user. The system was designed to output the top three recommendations per user in a consistent manner.

```
user_id      text
0      U1    sports art gaming
1      U2    travel food fashion
2      U3    sports travel fashion
3      U4      fashion music tech
4      U5    fashion food fitness
post_id      text
0      P1    sports food video
1      P2    music travel video
2      P3    sports travel text
3      P4    music gaming image
4      P5    food fashion image
```

Results and Discussion

The hybrid recommendation model successfully generated personalized suggestions by balancing between user interest, engagement similarity, and popularity. The inclusion of collaborative filtering helped discover hidden patterns that pure content-based filtering would miss, while the popularity measure ensured that trending content was not ignored. Compared to single-approach systems, the hybrid method produced recommendations that were more relevant and diverse, with improved coverage across users.

	user_id	recommended_posts	scores
0	U1	[P78, P22, P1]	[0.38818992927601115, 0.38601533602027516, 0.2...
1	U2	[P5, P80, P46]	[0.42718843593319855, 0.35999032290531136, 0.2...
2	U3	[P42, P3, P39]	[0.5194141549434416, 0.35162111295680704, 0.23...
3	U4	[P37, P16, P53]	[0.2841452453718329, 0.23519409610758035, 0.23...
4	U5	[P7, P26, P69]	[0.4611984689453777, 0.41811938804383586, 0.37...

CONCLUSION

The project demonstrates the effectiveness of a hybrid recommender system in addressing personalization challenges in online platforms. By integrating content-based filtering, collaborative filtering, and popularity-based ranking, the system ensures that recommendations are not only relevant to individual users but also capture broader engagement trends. The work can be extended in the future by incorporating deep learning–based embeddings and real-time feedback mechanisms to further enhance accuracy.