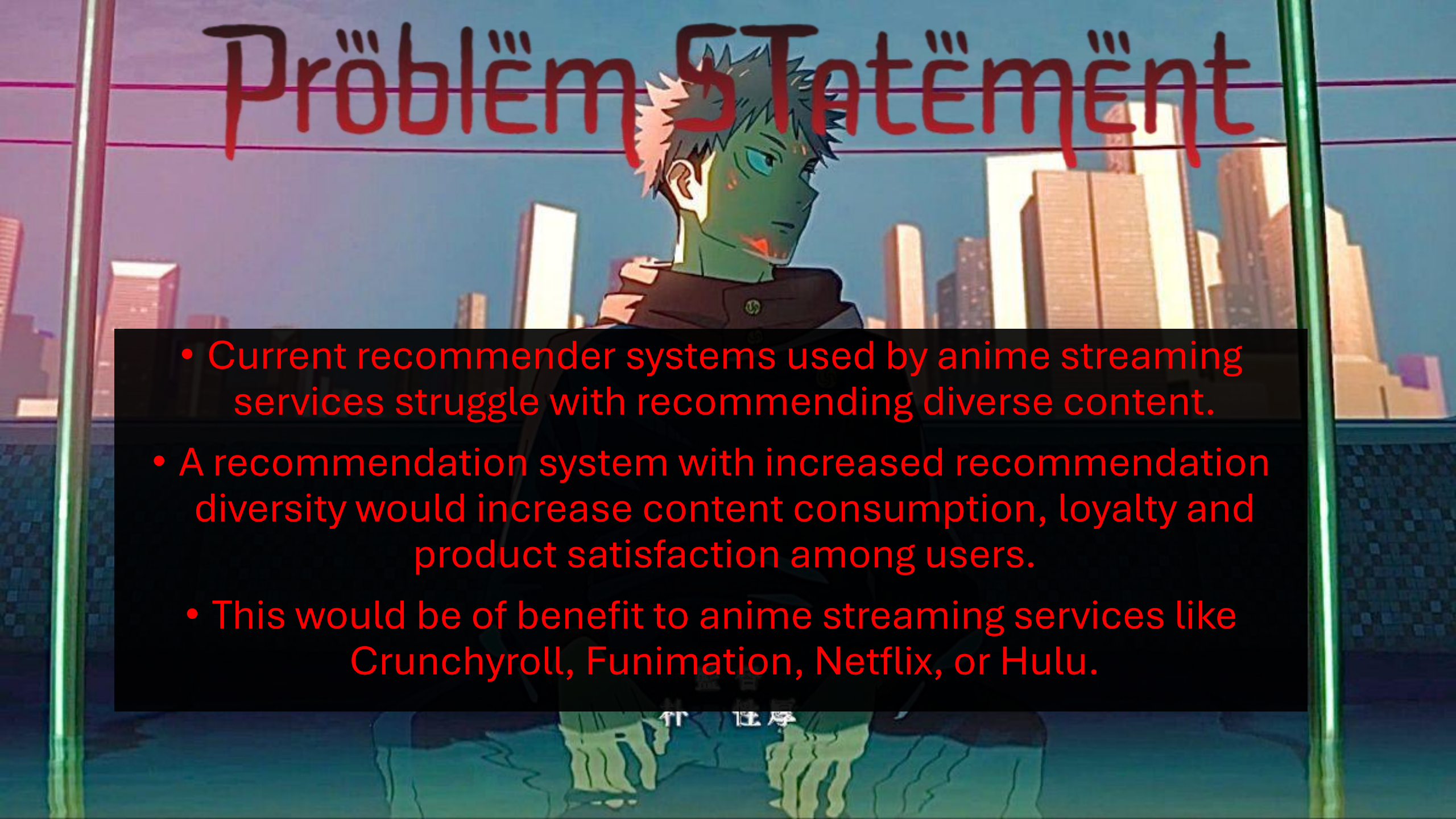
Seemeen Wookey
(201703165)

Problem Statement



- Current recommender systems used by anime streaming services struggle with recommending diverse content.
- A recommendation system with increased recommendation diversity would increase content consumption, loyalty and product satisfaction among users.
- This would be of benefit to anime streaming services like Crunchyroll, Funimation, Netflix, or Hulu.


Objectives
Recommend 10 anime titles based on what other people enjoy
Recommend 10 anime titles based on the features of other similar anime

Methodology

1. Business Understanding

- Develop recommendation system with increased recommendation diversity and improved exposure for lesser-known titles

2. Data Understanding

- User preferences
- Content popularity
- User-anime interactions

CRISP-DM

3. Data Preparation

- Data Cleaning
- Data Preprocessing
- Exploratory Data Analysis

4. Modelling

- Content-based filtering
- Collaborative filtering

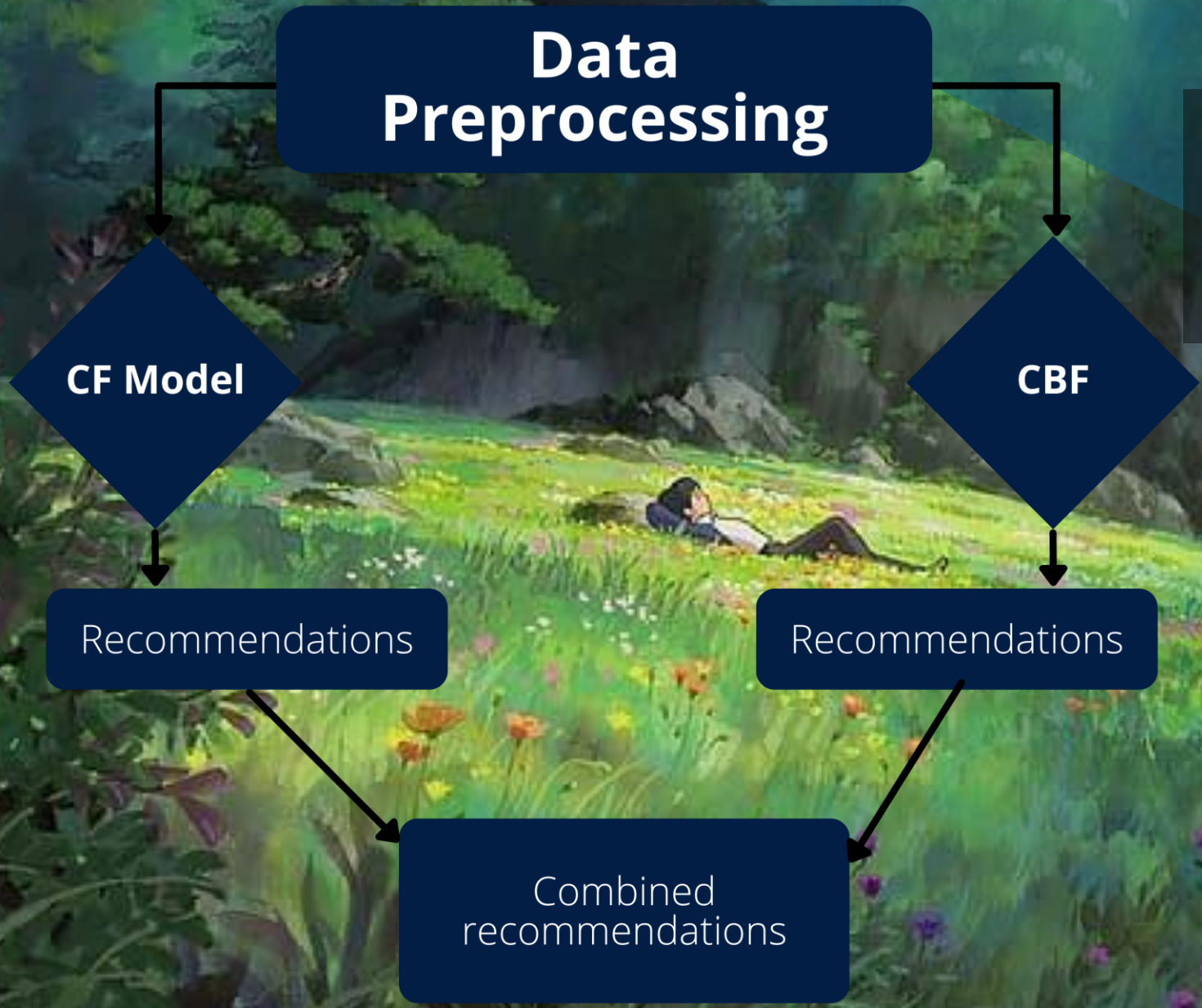
5. Evaluation

- Mean Squared Error and Mean Absolute Error

6. Deployment

- Recommendations

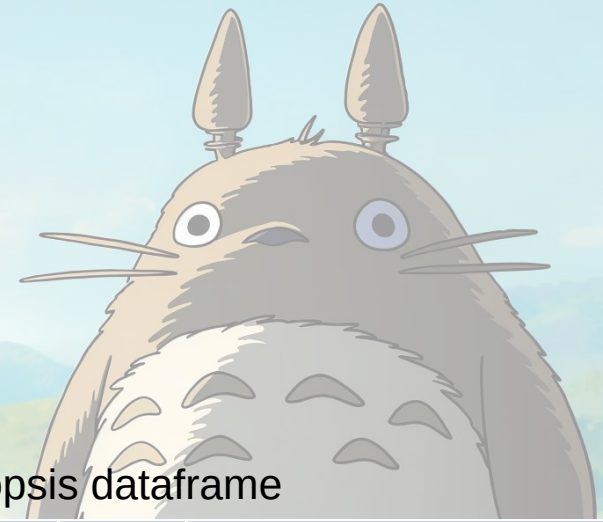
Process Flow



Final Model Data

Rating dataframe

Attribute	Data type	description	count
User_id	String	ID generated by MyAnimeList for its users at random	24573682
Anime_id	String	ID from MyAnimeList website that uniquely identi_es each anime	24573682
Rating	Integer	Rating assigned to an anime by a particular user	24573682



Anime dataframe

Attribute	Data type	description	count
Name	String	Full name of the anime	12309
Genre	String	Comma separated values identifying the genre/genres that each anime belongs to	12309
Type	String	Series or movie etc	12309
Rating	String	Parental guidance rating	12309
Score	Integer	average score of the anime given from all users in MyAnimelist database	12309
Popularity	Float	position based in the the number of users who have added the anime to their list.	12309

Synopsis dataframe

Attribute	Data type	description	count
MAL_ID	String	ID from MyAnimeList website that uniquely identifies each anime	16214
Name	String	Full name of the anime	16214
Genres	String	Comma separated values identifying the genre/genres that each anime belongs to	16214
Synopsis	String	Text in the form of string providing the synopsis of the anime	16206

Feature Selection

Method	Features
Collaborative Filtering	**Raw Features:** User-Anime Interactions User ID Anime ID User Ratings **Derived Features:** Latent Factors for Users Latent Factors for Anime Item-Based Similarity Matrix Anime Weight Similarity Predicted Ratings
Content-based Filtering	**Raw Features:** Genres Type Rating Score Popularity Rank **Derived Features:** TF-IDF for Genres TF-IDF for Type TF-IDF for Rating Combined Score and Popularity Rank Metric Weighted Feature Multiplication Cosine Similarity

Algorithm and Model Details

Collaborative Filtering

```
Model: "model"
__________________________________________________________________________________________________
 Layer (type)                Output Shape              Param #   Connected to
==================================================================================================
 user (InputLayer)           [(None, 1)]               0         []

 anime (InputLayer)          [(None, 1)]               0         []

 user_embedding (Embedding)  (None, 1, 128)            4517376   ['user[0][0]']

 anime_embedding (Embedding)  (None, 1, 128)           2156032   ['anime[0][0]']

 dropout (Dropout)           (None, 1, 128)            0         ['user_embedding[0][0]']

 dropout_1 (Dropout)         (None, 1, 128)            0         ['anime_embedding[0][0]']

 dot_product (Dot)           (None, 1, 1)              0         ['dropout[0][0]',
                                                                  'dropout_1[0][0]']

 flatten (Flatten)           (None, 1)                 0         ['dot_product[0][0]']

 dense (Dense)               (None, 1)                 2         ['flatten[0][0]']

 batch_normalization (BatchNorm  (None, 1)             4         ['dense[0][0]']
 alization)

 activation (Activation)     (None, 1)                 0         ['batch_normalization[0][0]']

==================================================================================================
Total params: 6,673,414
Trainable params: 6,673,412
Non-trainable params: 2
```

This architecture is designed for collaborative filtering in recommendation systems, where user-item interactions are predicted based on learned embeddings.

Content-Based strategy

involves creating a recommendation system based on the content features of anime titles. We focus on several columns, such as Genres, Type, Rating, Score, and Popularity Rank, to build our system.

cosSinDF

	0	1	2	3	4	5	6	7	8	9	...	12299	12300
0	1.000000	0.572665	0.764340	0.411153	0.391319	0.419351	0.392709	0.329025	0.407034	0.341701	...	0.000000	0.372404
1	0.572665	1.000000	0.329893	0.264342	0.000000	0.073948	0.068274	0.000000	0.116972	0.162034	...	0.000000	0.000000
2	0.764340	0.329893	1.000000	0.496330	0.553409	0.636315	0.475864	0.504760	0.491760	0.347837	...	0.269591	0.453335
3	0.411153	0.264342	0.496330	1.000000	0.435935	0.423591	0.418492	0.357143	0.462252	0.590426	...	0.071429	0.357143
4	0.391319	0.000000	0.553409	0.435935	1.000000	0.459210	0.301863	0.421548	0.301863	0.285714	...	0.452047	0.301863
...	...	...	...	...	...	...	...	...	...	...	...	...	...
12304	0.167841	0.000000	0.257666	0.071429	0.016149	0.289518	0.210015	0.237316	0.071429	0.000000	...	0.071429	0.403465
12305	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.137312	0.000000	0.000000	0.000000	...	0.000000	0.410030
12306	0.056983	0.000000	0.079377	0.016149	0.071429	0.090191	0.424352	0.504766	0.016149	0.000000	...	0.301863	0.128876
12307	0.071429	0.206211	0.082028	0.202962	0.224627	0.096057	0.000000	0.073065	0.000000	0.405435	...	0.431143	0.000000
12308	0.073111	0.000000	0.152553	0.071429	0.016149	0.572094	0.131797	0.452255	0.333031	0.000000	...	0.357143	0.216063

12309 rows × 12309 columns

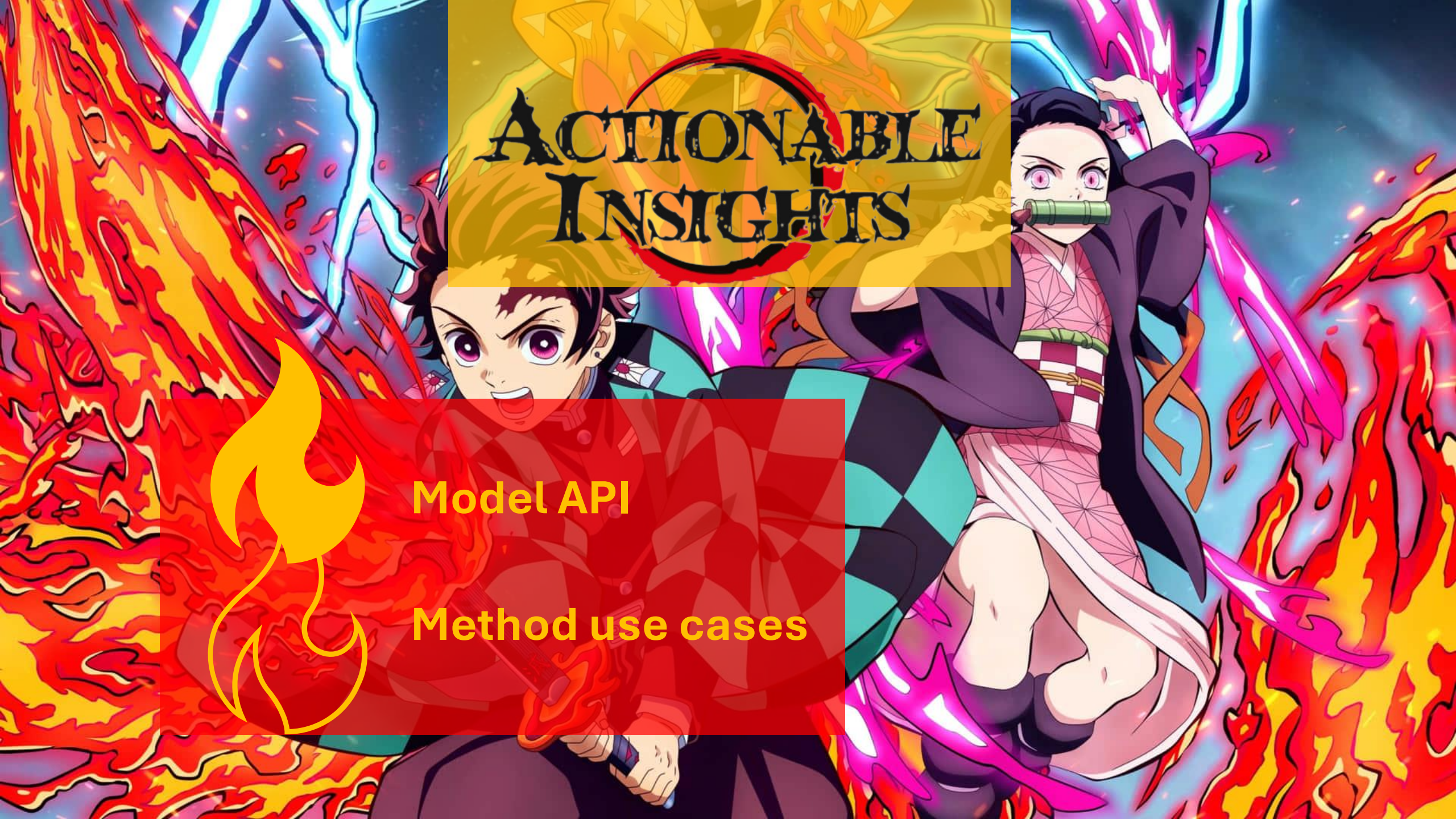
Model Performance

Metric	Training	Validation
Loss	0.5919	0.5877
Mean Absolute Error (MAE)	0.1315	0.1204
Mean Squared Error (MSE)	0.0261	0.0235
Learning Rate (LR)	0.000052261	-

Results

	name	similarity	genre
10	Dragon Ball Z	0.885399	Action, Adventure, Comedy, Fantasy, Martial Arts, Shounen, Super Power
9	Rurouni Kenshin	0.787129	Action, Adventure, Comedy, Historical, Romance, Samurai, Shounen
8	One Piece Film Z	0.768603	Action, Adventure, Comedy, Drama, Fantasy, Shounen
7	Hunter x Hunter	0.767178	Action, Adventure, Super Power, Fantasy, Shounen
6	Yu Yu Hakusho:Ghost Files	0.766122	Action, Comedy, Demons, Supernatural, Martial Arts, Shounen
5	Reborn!	0.753528	Action, Comedy, Shounen, Super Power
4	One Piece Film Strong World	0.751829	Action, Adventure, Comedy, Drama, Fantasy, Shounen
3	Hunter x Hunter: Greed Island Final	0.739457	Action, Adventure, Super Power, Fantasy, Shounen
2	Hajime No Ippo:The Fighting!	0.737633	Comedy, Sports, Drama, Shounen
1	Hunter x Hunter: Greed Island	0.737105	Action, Adventure, Super Power, Fantasy, Shounen

	name	Cosine similarity	genre
10	Dragon Ball Z	0.983883	Action Adventure Comedy Fantasy Martial-Art
9	Dragon Ball Super	0.983883	Action Adventure Comedy Super-Power Martial-Ar...
8	Dragon Ball Kai (2014)	0.983883	Action Adventure Comedy Fantasy Martial-Arts S...
7	Dragon Ball Kai	0.983883	Action Adventure Comedy Super-Power Martial-A
6	Naruto: Shippuuden	0.961626	Action Adventure Comedy Super-Power Martial-Ar.
5	Naruto	0.961626	Action Adventure Comedy Super-Power Martial-Ar
4	Flame of Recca	0.948297	Action Adventure Martial-Arts Shounen Super-Power
3	Boruto:Naruto Next Generations	0.948297	Action Adventure Super-Power Martial-Arts Shounen
2	Shadow Skill: Eigi	0.930216	Adventure Fantasy Magic Martial-Arts Super-Pow
1	Wolverine	0.893521	Action Martial-Arts Super-Power


ACTIONABLE INSIGHTS
Model API
Method use cases

References

Du, Y., Ranwez, S., Sutton-Charani, N., & Ranwez, V. (2021). Is diversity optimization always suitable? Toward a better understanding of diversity within recommendation approaches. *Information processing & management, 58*(6).

Jena, A., Jaiswal, A., Lal, D., Rao, S., Ayubi, A., & Sachdeva, N. (2022). Recommendation System For Anime Using Machine Learning Algorithms. *International Conference on Innovative Computing & Communication*.

Marana, F. T., Fernandes, R. d., Guzmán, J. L., Carvalho, A. C., & Cúri, M. (2023). A Cognitive Diagnosis Approach for Recommending Items Based on Polytomous Responses and Latent Attributes. *Expert Systems With Applications*.

Mutteppagol, V. M. (2021). A deep learning recommender system for anime.


THANK
YOU