Book Recommendation System

Team Members

Daria Zinchenko

Neha Sharma

Seeke Olu-Daniel

Valentyna Kravets

Outline

Project Goals

Methodology

Recommendation Algorithm

Website Demo & Data visualization

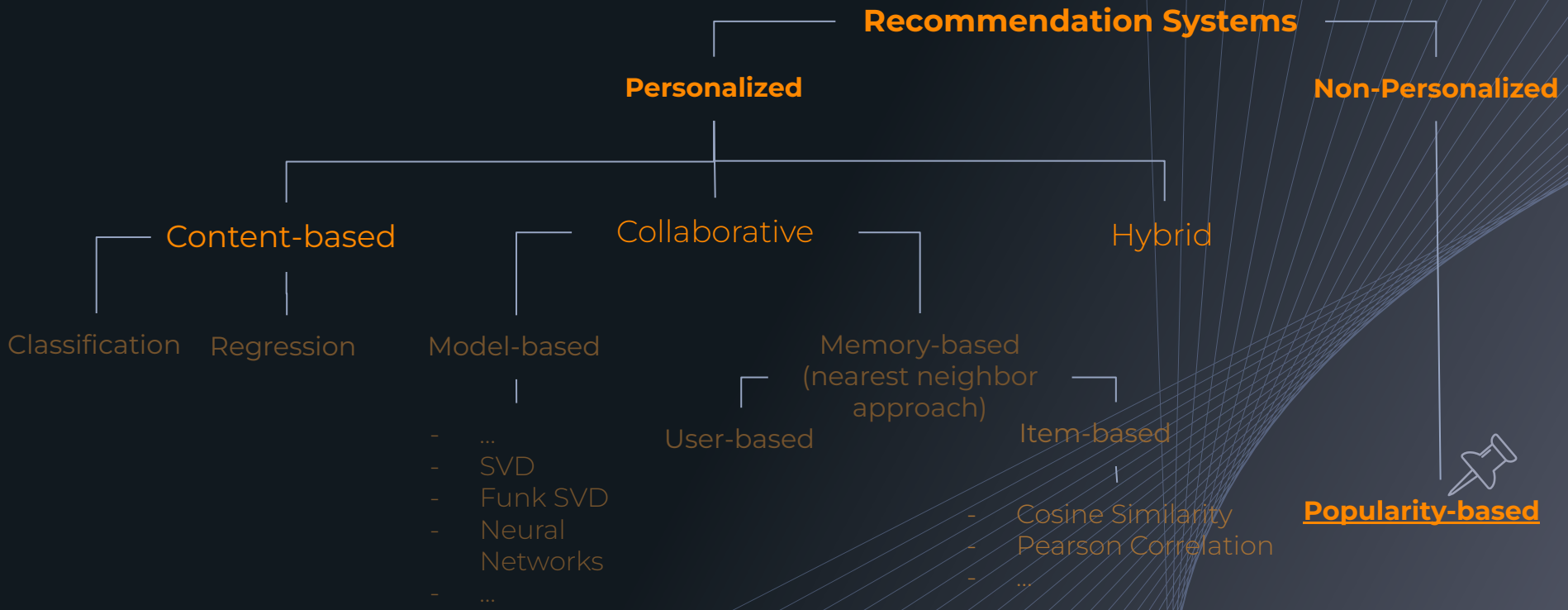
Insights, Conclusion & Next Steps

Project Goals

- Increased user engagement by providing personalized recommendations aligned with their interests and preferences.
- Boost book sales.
- Maximize inventory utilization by promoting niche books to users likely to be interested in such.
- Enhance user experience.
- Optimize recommendation algorithms.

Recommendation Algorithm

Recommendation Algorithm



Recommendation Algorithm

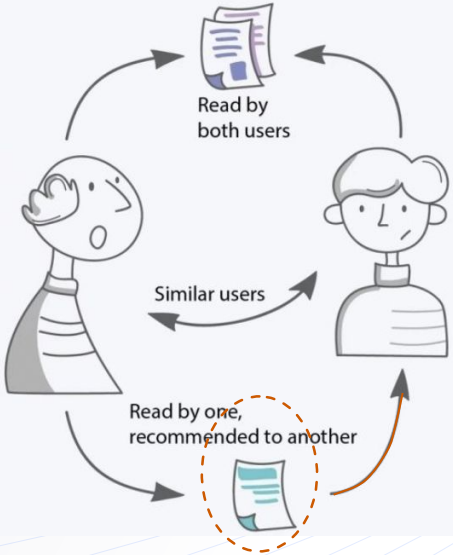
Behavior-driven

Serendipity

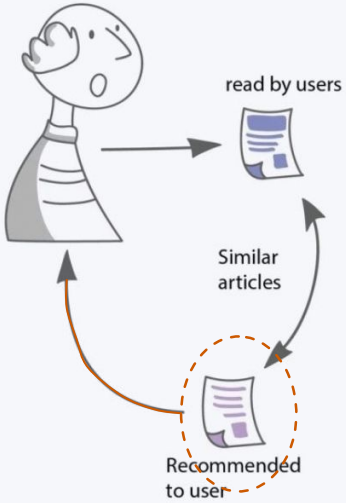
Cold Start

Popularity Bias

COLLABORATIVE FILTERING



CONTENT-BASED FILTERING



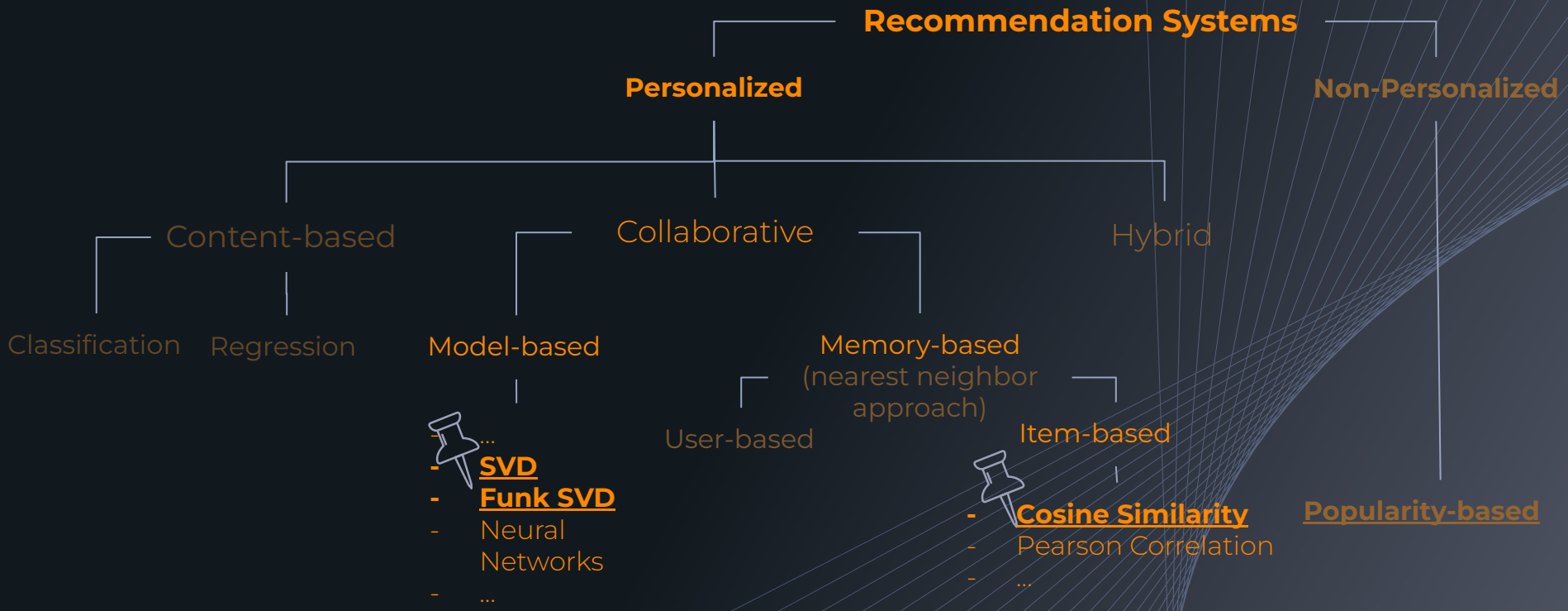
New Items
among
Recommendations

Transparency

Limited Analysis

Over-specialization

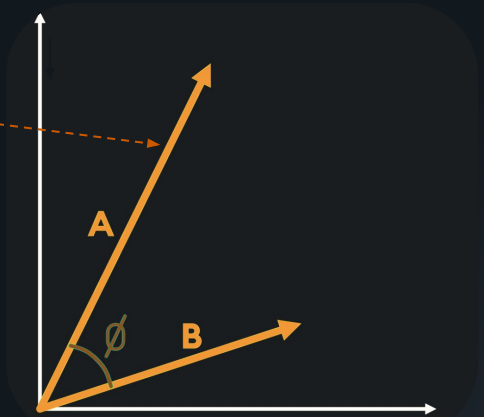
Recommendation Algorithm



Cosine Similarity

	User_1	...	User_n
Book_1	a_1	...	a_n
...	...	...	...
Book_m	b_1	...	b_n

a_i, b_i ∈ [1,10]



$$\cos(\theta) = \frac{a \bullet b}{\|a\|\|b\|} = \frac{(a_1 * b_1) + (a_2 * b_2) + \ldots + (a_n * b_n)}{\sqrt{(a_1^2 + a_2^2 + \ldots + a_n^2) * (b_1^2 + b_2^2 + \ldots + b_n^2)}}$$

Why?

Since overall ratings and popularity can create *different distances*, but the *direction of similar vectors remains close.*

Singular Value Decomposition

	Book_1	...	Book_n
User_1	x_{11}	...	x_{1n}
...	...	...	...
User_m	x_{m1}	...	x_{mn}

Data normalization

Matrix Factorization (Decomposition)

*Extracted top **r** latent factors*

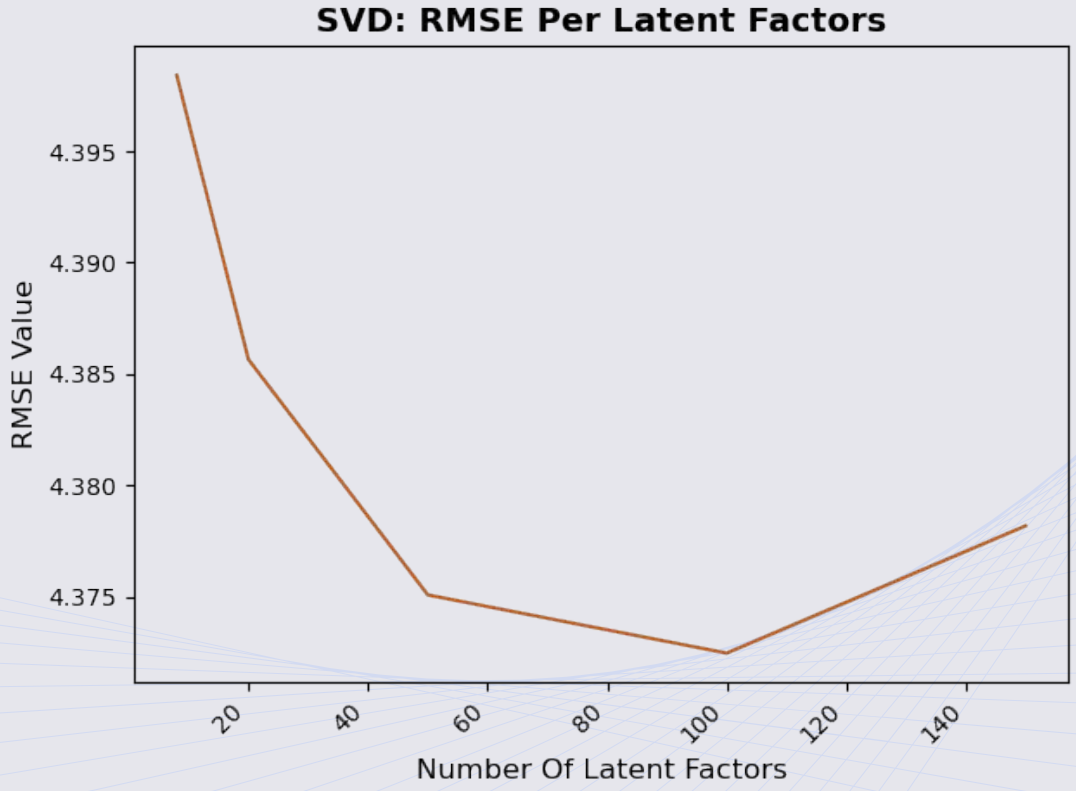
Predictions (Dot product)

De-normalization

Recommendations (Top rated titles, which were not read)

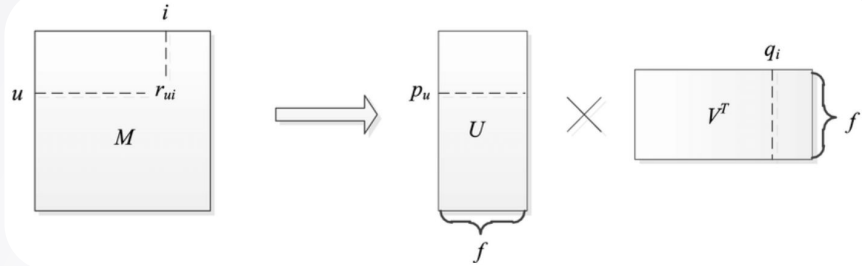
$$X = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & \\ \vdots & \vdots & \ddots & \\ x_{m1} & & & x_{mn} \end{pmatrix}_{m \times n} = \begin{pmatrix} u_{11} & \cdots & u_{1r} \\ \vdots & \ddots & \\ u_{m1} & & u_{mr} \end{pmatrix}_{m \times r} \begin{pmatrix} s_{11} & 0 & \cdots \\ 0 & \ddots & \\ \vdots & & s_{rr} \end{pmatrix}_{r \times r} \begin{pmatrix} v_{11} & \cdots & v_{1n} \\ \vdots & \ddots & \\ v_{r1} & & v_{rn} \end{pmatrix}_{r \times n}$$

Model Evaluation

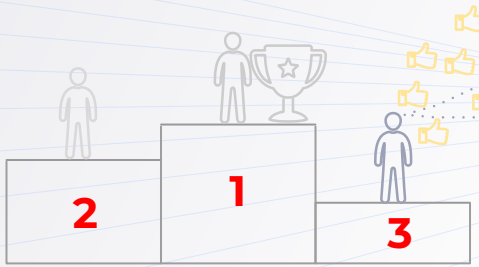


$$RMSE=\sqrt{(\sum(Predicted\_i-Actual\_i)^2/N)}$$

SVD Funk



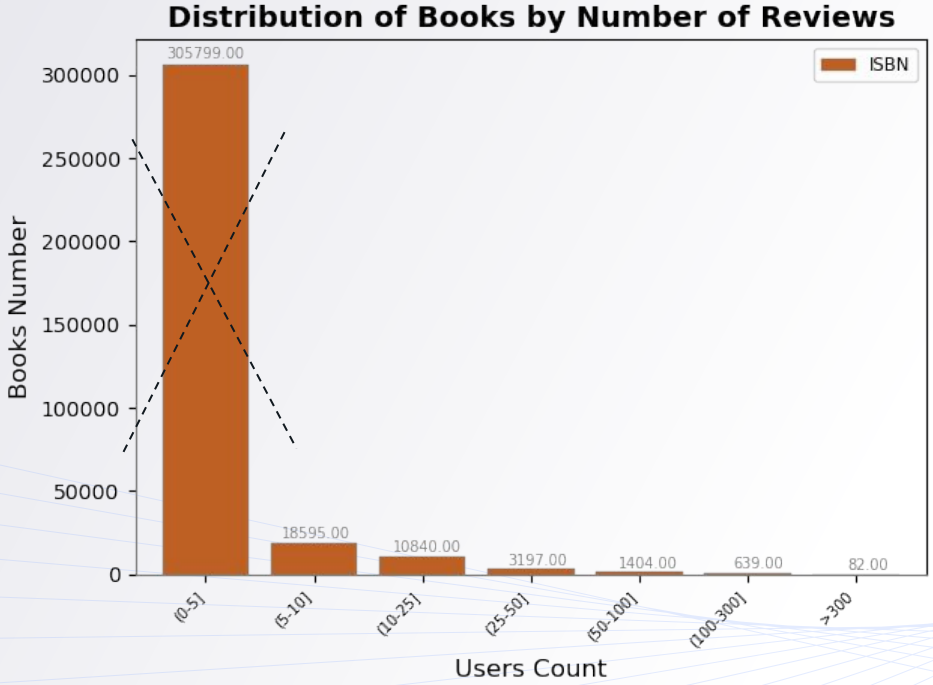
- Well-suited for Sparse Matrix
- Focuses on Accuracy Improvement

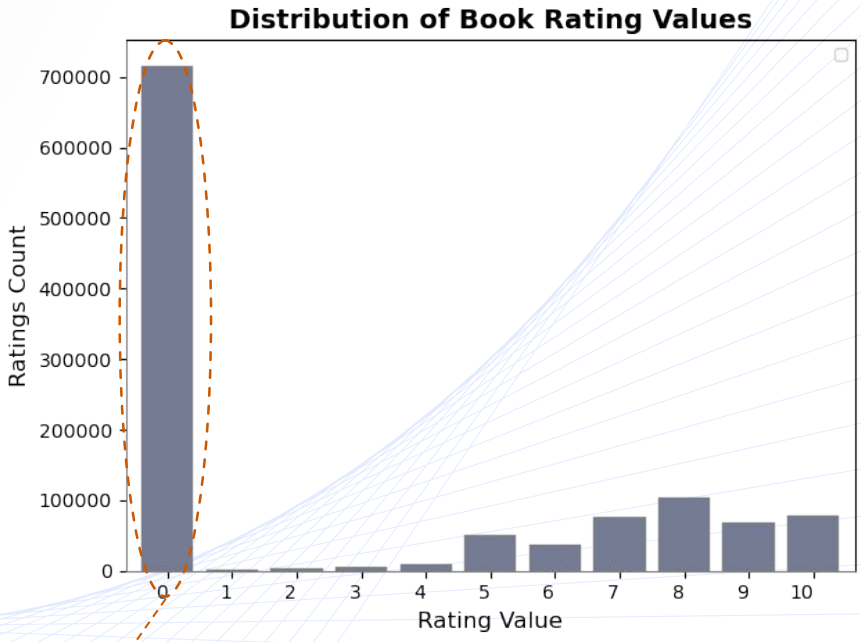


surprise

A Python scikit for recommender systems.

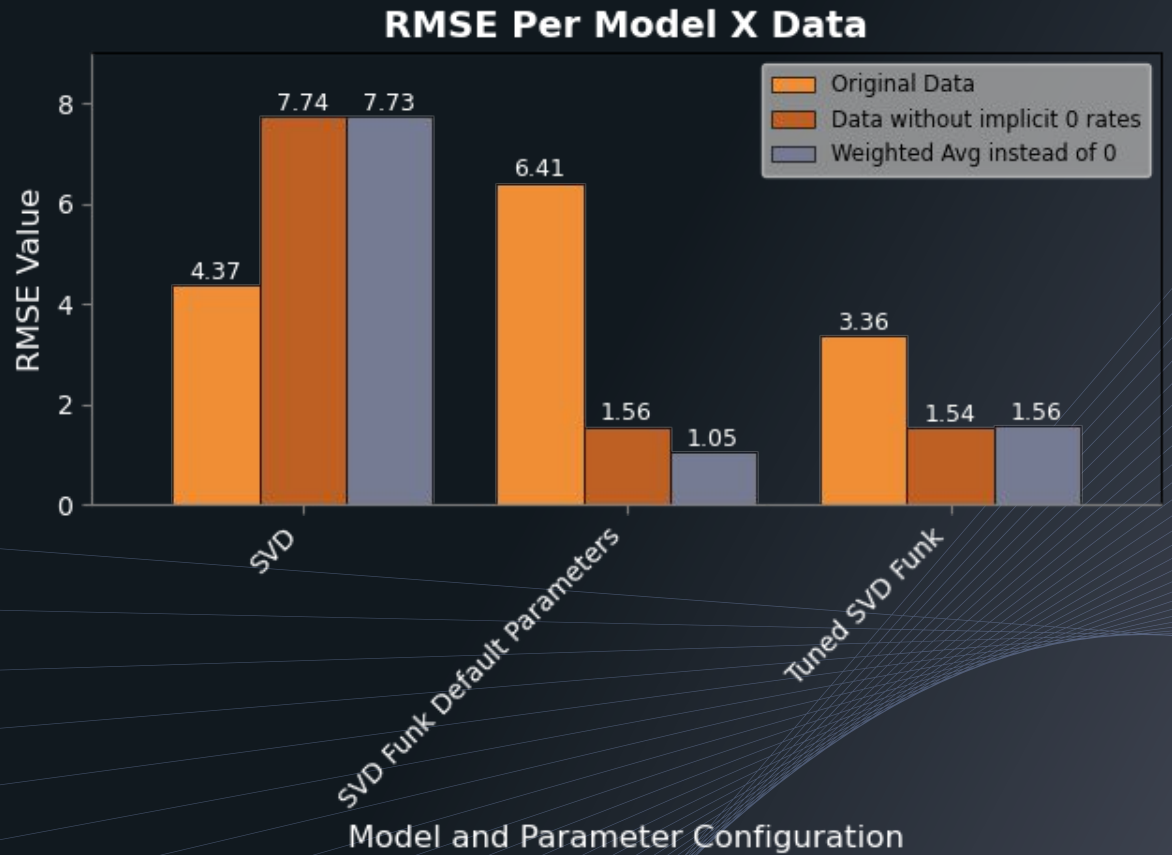
Algo Input Data





$$\frac{\text{Book\_Average\_Rating} * \text{Ratings\_Count} + \text{Min\_Ratings\_Count} * \text{Neutral\_Score}}{\text{Ratings\_Count} + \text{Min\_Ratings\_Count}}$$

Model Evaluation



Demo

Future Steps and Improvements

- ❑ Incorporate newly rated books
- ❑ Sentiment Analysis
- ❑ Tags and genre-based search functionalities
- ❑ Expand database to include more languages
- ❑ Feedback loop
- ❑ Predictive analytics
- ❑ Prescriptive analysis

Thanks!

Questions are now welcome.