BAX493 - FINAL PROJECT

**TEAM 2** Candice Her   Hitesh Kunchakuri   Mia Lai   Shriya Yegalapati

01

PRODUCT

INTELLIGENT BOOK RECOMMENDATION SYSTEM

PRODUCT

Book Recommendation System

- This end-to-end machine learning application is designed to provide seamless integration and predictive analytics capabilities through **an online prediction API**.

- Suggest books to users based on their individual preferences and reading habits
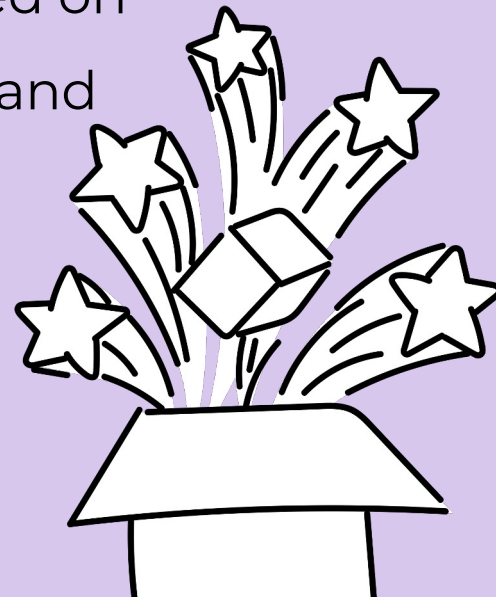
PROBLEM STATEMENT

- **Challenge:**

Readers often struggle to find books that match their preferences.

- **Impact:**

This leads to frustration and decreased reading engagement.

- **Opportunity:**

An intelligent system can enhance user experience by providing personalized book recommendations.

SOLUTION OVERVIEW

Product
An AI-powered book recommendation system.

Functionality
Suggests books based on user preferences, reading history, and similar user profiles.

02

LET'S TALK ABOUT DATA
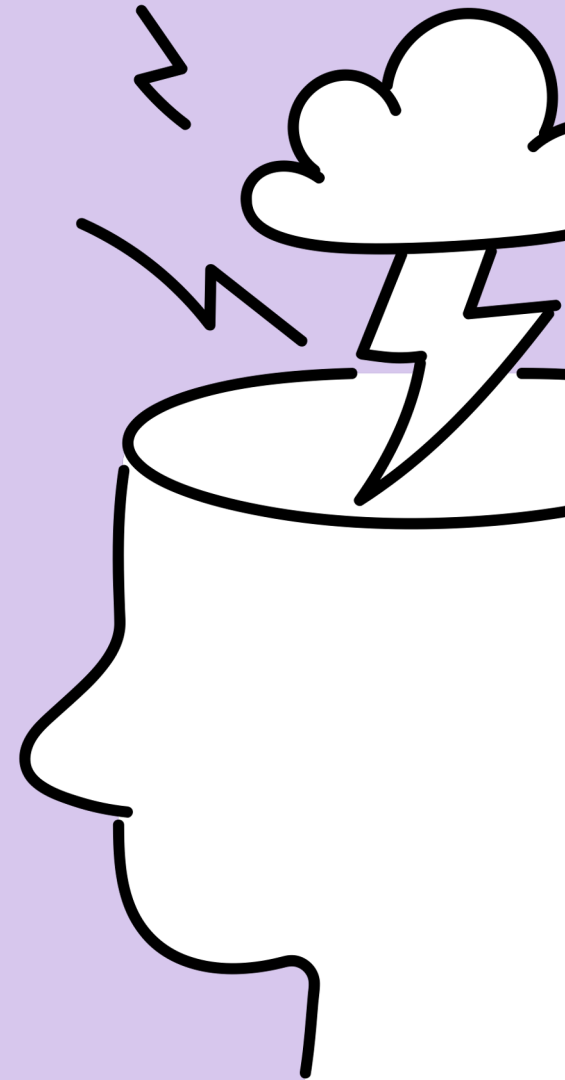
DATA SELECTION

- **Data Source:**

Goodbooks-10k dataset from Kaggle.

- **Data Includes:**

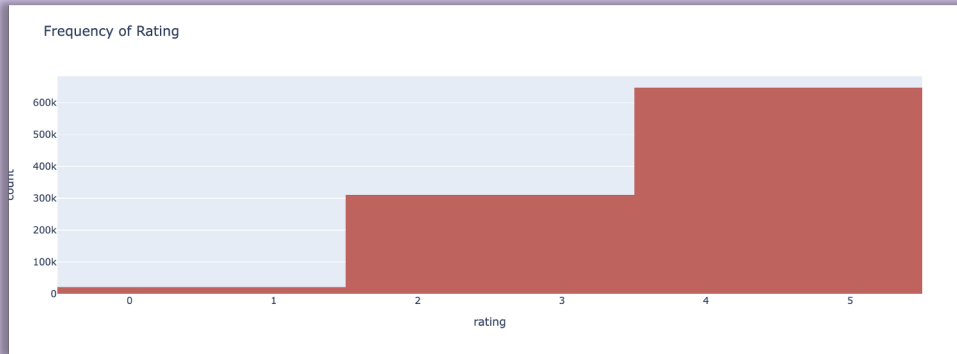Book details, ratings, tags, and user interactions.

- **Data Size:**

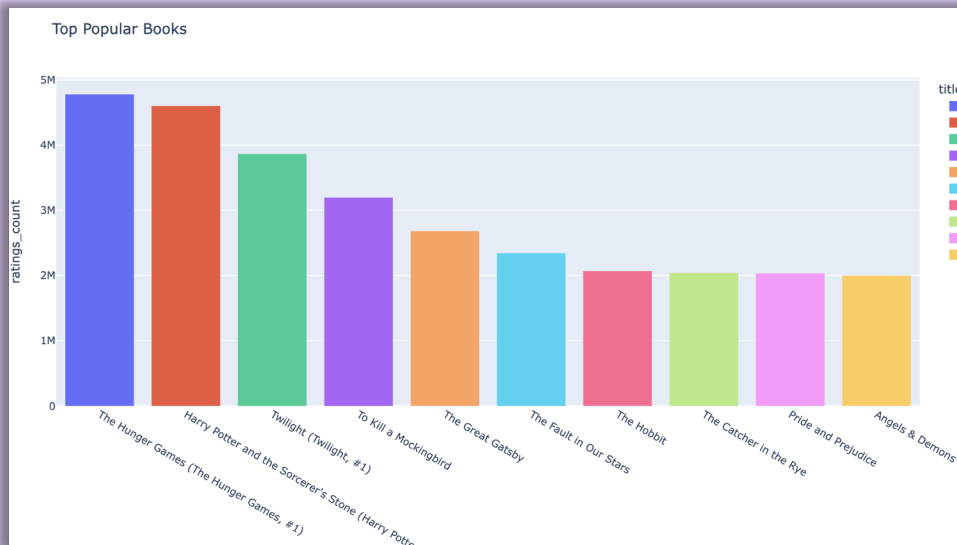10,000 books, 50,000+ Users, 100,000+ user ratings.

DISTRIBUTION

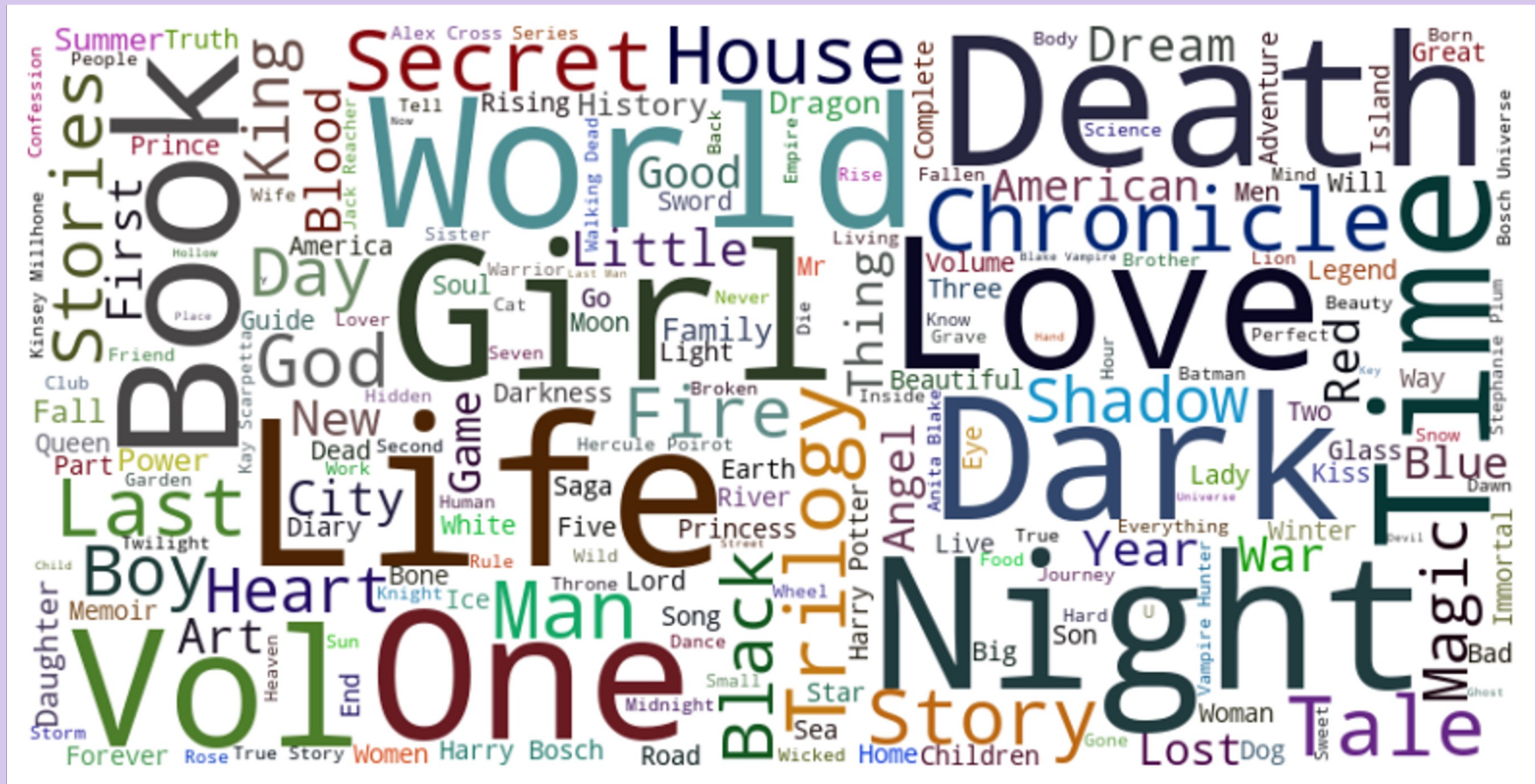**High ratings dominate.**

Steady growth till 2012.

"Hunger Games" and "Harry Potter" lead.

Bill Watterson prominent.

WORD CLOUD

Popular book titles often feature words like "Life," "Death," "Love," and "World."

DATA PROCESSING

- **Cleaning Steps:**

Handling missing values, normalizing text

fields, and encoding categorical data.

- **Tools Used:**

Big Query on Google Cloud

- **Outcome:**

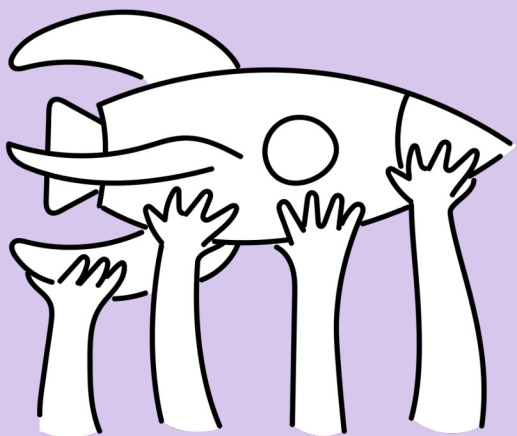Cleaned dataset ready for model training.

03

MODEL

Model Training

- **ML technology:**

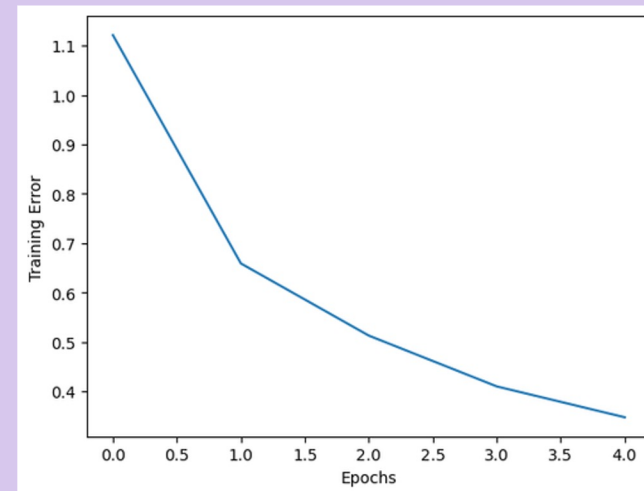 o TensorFlow on Python

  o AutoML on Google Cloud

- **Training using Vertex AI Step**:

  o Split data into training and test sets.

  o Choose Rating as Outcome, while User_Id, and Book_id as features.

  o Train the collaborative filtering model on user-item interactions.
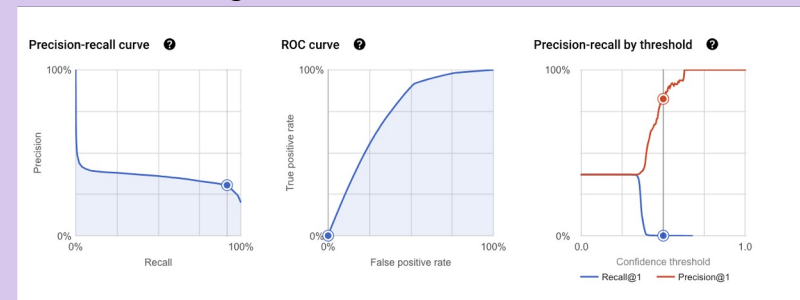
Evaluation

TensorFlow Model on Python

Loss: 0.3175   **MSE: 0.2359**

AutoML on Google Cloud

ROC-AUC: 0.742   **Precision: 82.5%**
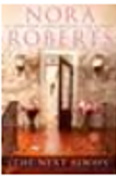
Model Deployment

- **API Integration**

  o **Endpoint ID:** 7641260566392078336

  o **Project ID:** 343320691785

- **API Request Format**

```
{
 "instances": [
  { "feature_column_a": "value", "feature_column_b": "value" },
  { "feature_column_a": "value", "feature_column_b": "value" },
  …
 ]
}
```

- **Execute the request**

```
curl -X POST \
  -H "Content-Type: application/json" \
  -d @INPUT-JSON \
"https://ml.googleapis.com/v1/projects/343320691785/models/7641260566392078336:predict"
```

Image	Title	Authors	Year	Ratings Count
	Nausea	Jean-Paul Sartre, Lloyd Alexander, Hayden Carruth	1938.0	60910
	The Next Always (Inn BoonsBoro, #1)	Nora Roberts, MacLeod Andrews	2011.0	63081
	Red Storm Rising	Tom Clancy	1986.0	58980
	The Sign of Four	Arthur Conan Doyle, Peter Ackroyd, Ed Glinert	1890.0	72378
	The Mitten	Jan Brett	1929.0	67185

04

BUSINESS OPPORTUNITY

- **Revenue Streams:**
Subscription-based model, partnership with bookstores, in-app purchases.

- **Value Proposition**
Enhanced reading experience, and increased customer loyalty.

COMPTETIVE ATAVANTAGES

- **Unique Features:** Advanced machine learning algorithms, personalized recommendations, and real-time updates.

- **Differentiators:** Superior user experience, extensive dataset, and hybrid recommendation model.

- **Path to Profitability:** Scaling User Base, Cost Management, Diversifying Revenue Streams

- **Investment Opportunity** Growing Market, First-Mover Advantage and Return on Investment (ROI).

THANK YOU

TEAM 2