

# BOOK RECOMMENDATION SYSTEM BUILT ON AMAZON REVIEWS

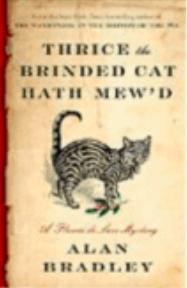
Kenneth Hua

# Background

- Recommendation Systems are useful tools for Booksellers to sell books to users who have a purchase history
- Typically, recommendation systems are built upon User Ratings. Can we build one using User Reviews that is more effective?

**Recommended for you**

[See more recommendations ›](#)



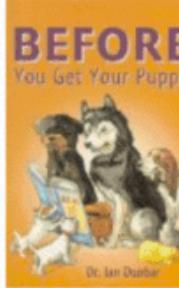
**Thrice the Brinded Cat Hath Mew'd: A...**  
› Alan Bradley  
List price: \$26.00  
Kindle price: \$12.99  
[Why recommended?](#)



**Egg: Nature's Perfect Package**  
› Steve Jenkins, Robin Page  
 (9)  
List price: \$16.99  
Kindle price: \$12.99  
[Why recommended?](#)



**The Untethered Soul: The Journey...**  
› Michael A. Singer  
 (2,792)  
List price: \$16.95  
Kindle price: \$8.64  
[Why recommended?](#)



**Before You Get Your Puppy**  
› Ian Dunbar  
 (23)  
List price: \$7.95  
Kindle price: \$5.95  
[Why recommended?](#)



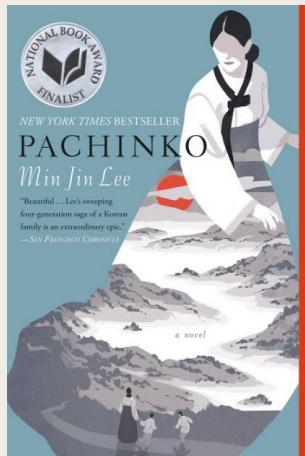
**How to Be Your Dog's Best Friend:**  
A...  
Of New Skete Monks  
 (368)  
List price: \$27.00  
Kindle price: \$8.99  
[Why recommended?](#)

# Design

User: Kenneth Hua

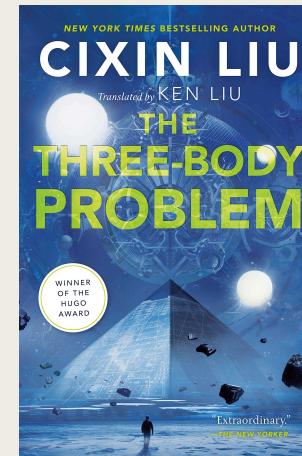
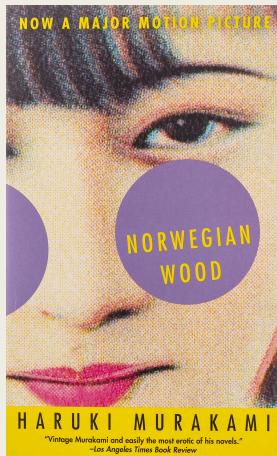
Book: Pachinko, by Min Jin Lee

Review: Very engrossing, historical fiction epic. Can be quite long at times, but overall you are taken along the journey of a family through multiple generations and different countries. Very relatable for someone with an immigrant background and I really enjoyed it. Some characters were more interesting than others, which made some parts feel slow, but overall it felt like a dense epic.



Sentiment (0-1): 0.744

Review: Very *engrossing*, historical fiction epic. *Can be quite long at times*, but overall you are taken along the journey of a family through multiple generations and different countries. *Very relatable for someone with an immigrant background and I really enjoyed it. Some characters were more interesting than others, which made some parts feel slow, but overall it felt like a dense epic.*



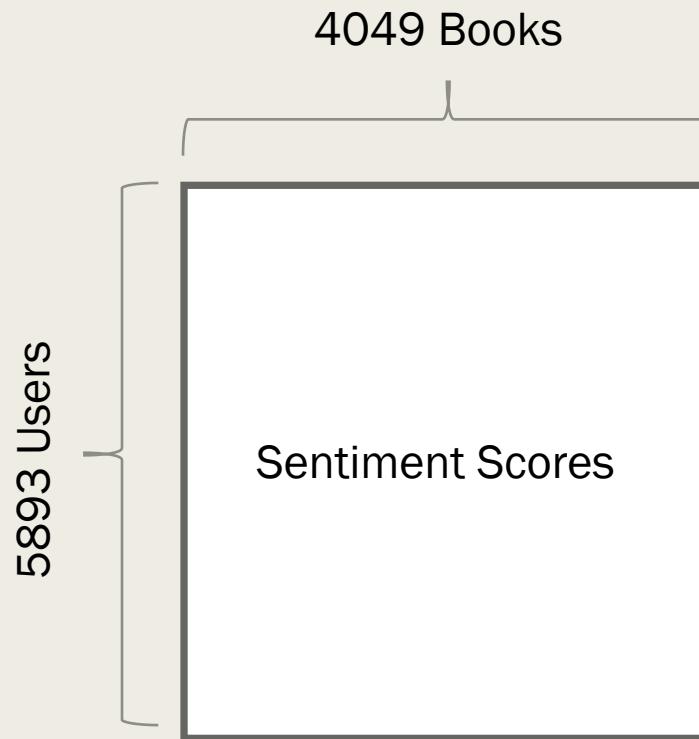
User: Kenneth Hua

Likes: Pachinko, Norwegian Wood, Three Body Problem

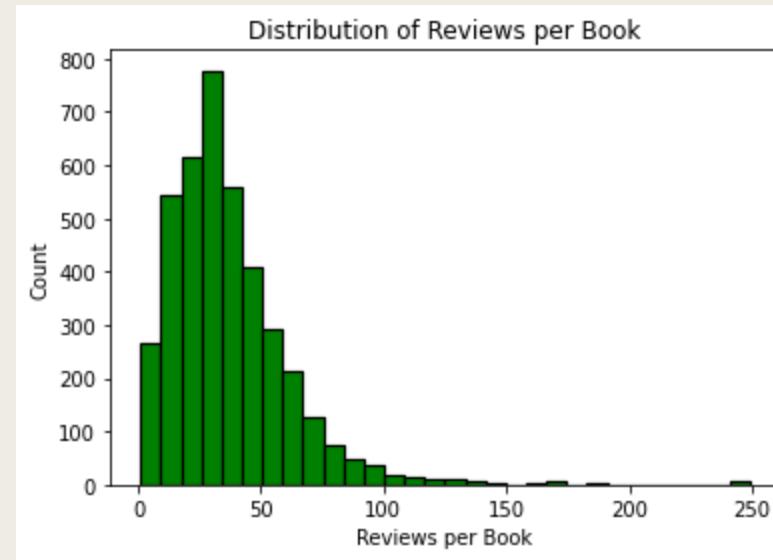
Dislikes: MiddleMarch, The Wind-Up Bird Chronicle

What are the books that this user would enjoy that they have not read?

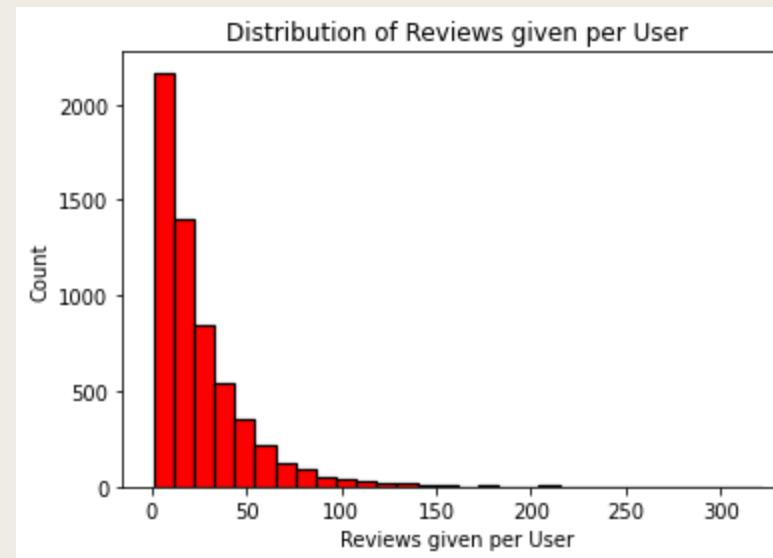
# Data



Data can be organized in a  
5893 by 4049 pivot table



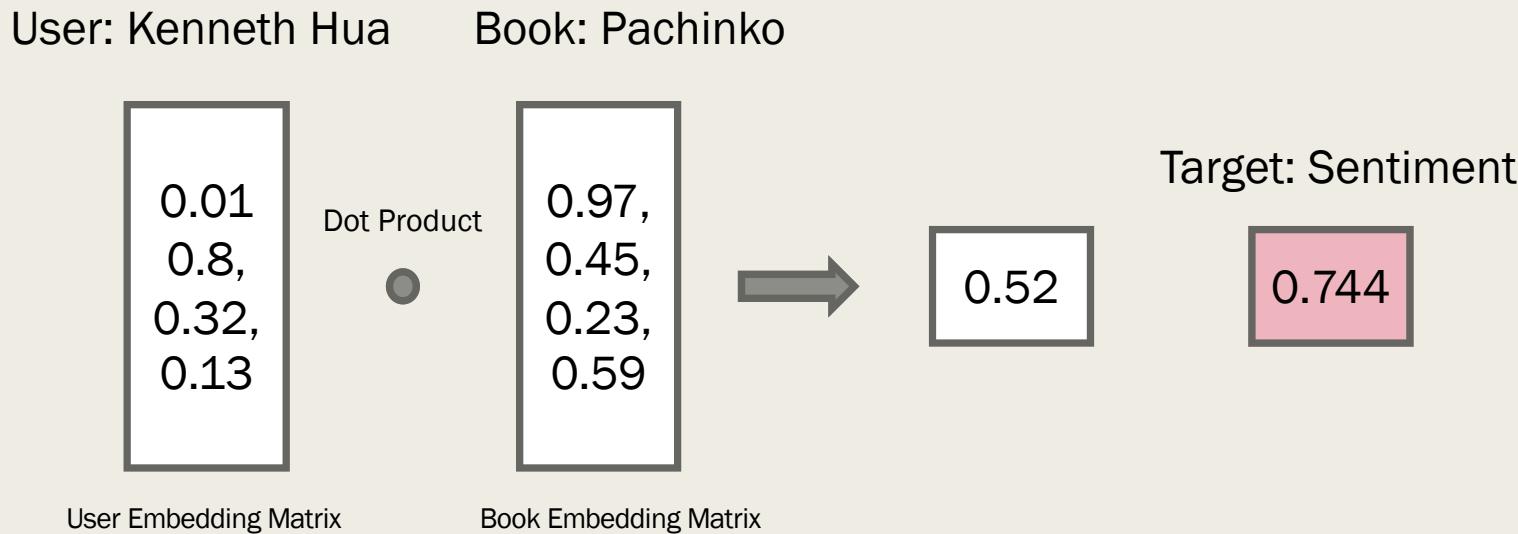
Median: 32 Reviews per Book



Median: 17 Reviews given per User

# Model Architecture

Matrix Factorization – Use SVD to break User-Book Matrix into embeddings, then take dot product of embeddings



Matrix Factorization is too simple!

Need a more advanced model that can map out the complexities  
of User-Book Relationships

Average RMSE: 0.131

# Model Architecture

Neural  
Network  
Model

User: Kenneth Hua

Input Layer

User Embedding Matrix

0.85,  
0.8,  
0.32,  
0.13

2048

Book: Pachinko

Book Embedding Matrix

0.97,  
0.45,  
0.23,  
0.59

Hidden Dense Layers – 100 epochs

1024

512

256

64

16

Target: Sentiment

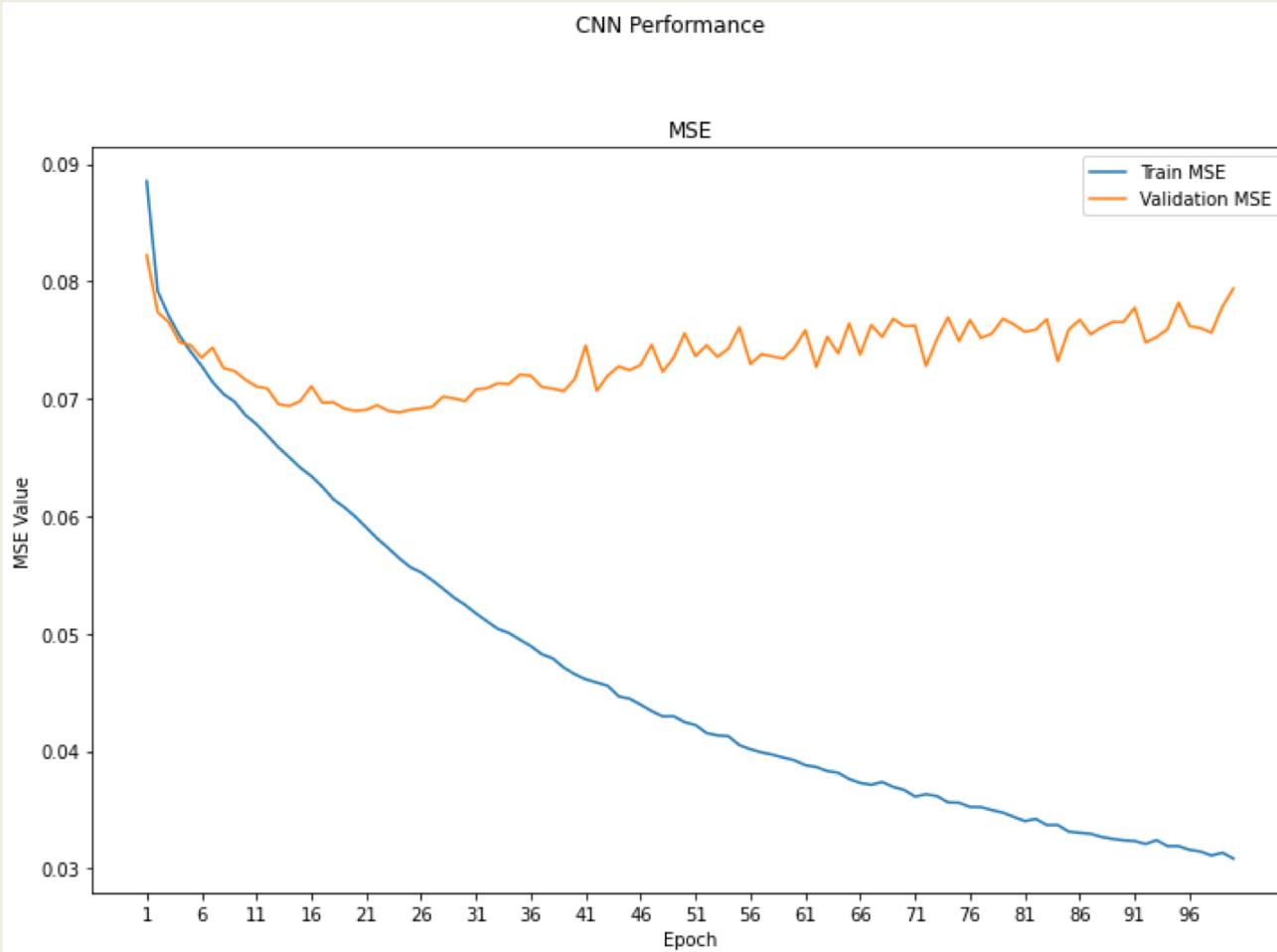
0.744

Output Layer

0.69

Average RMSE:  
0.07

# Model Performance

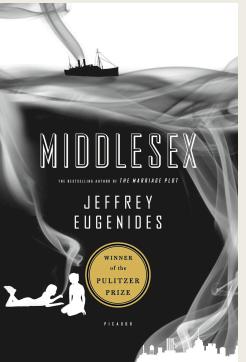


- Validation MSE is optimal around epoch 25
- Training loss continues to decrease, some overfitting occurs.
- Model can be improved with more data, hyperparameter tuning

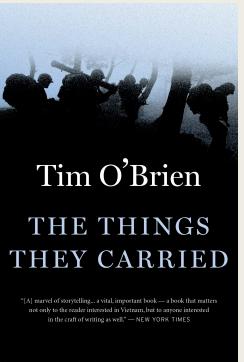
# Let's try the recommender in action:

Name: Craig Wood, Read 53 Books in the Dataset

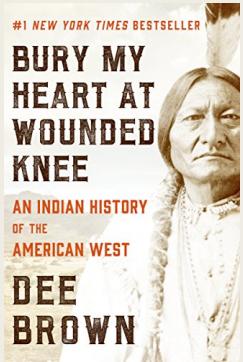
## ■ Top 5 Books: History, War, Novel, Politics



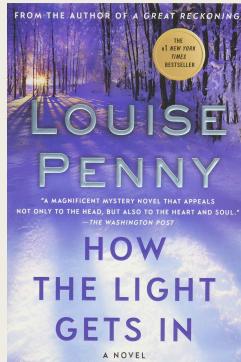
Sentiment: 1



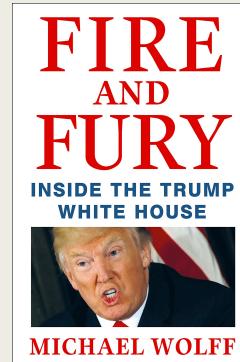
Sentiment: 0.94



Sentiment: 0.94

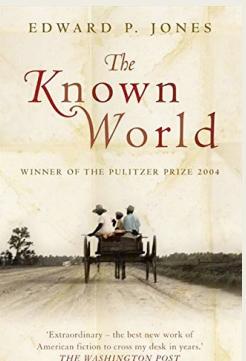


Sentiment: 0.88

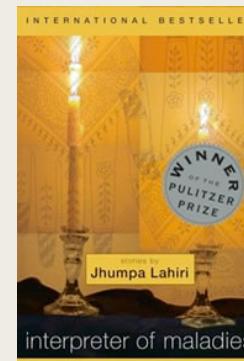


Sentiment: 0.88

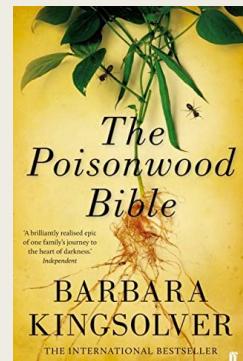
## ■ Recommended 5 Books: History, culture, mystery, missionary



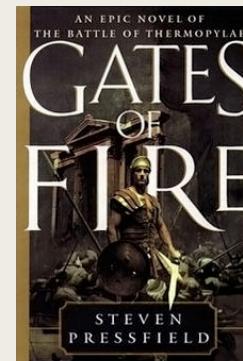
Prediction: 0.97



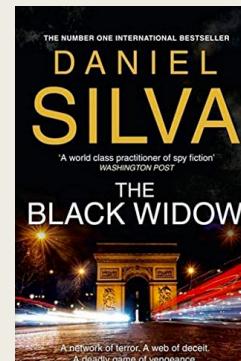
Prediction: 0.96



Prediction: 0.94



Prediction: 0.93

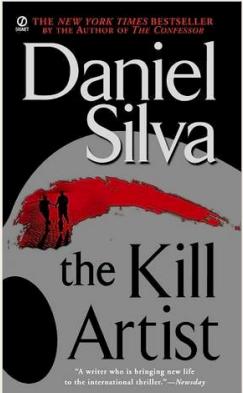


Prediction: 0.92

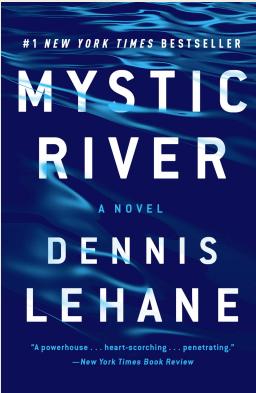
# Let's try the recommender in action:

Name: nobizinfla, Read 41 Books in the Dataset

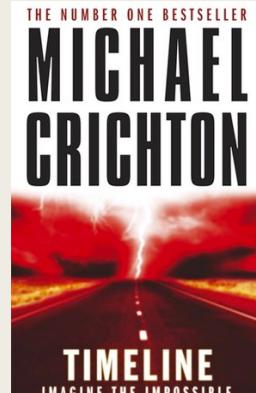
- Top 5 Books: Mystery, Thriller, Detective



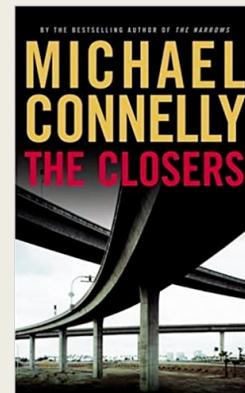
Sentiment: 1



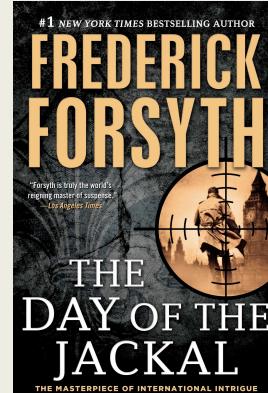
Sentiment: 1



Sentiment: 1

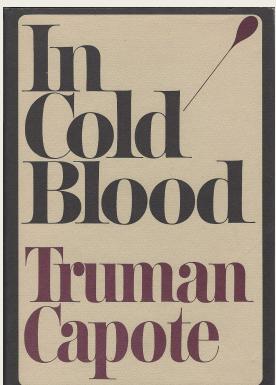


Sentiment: 1

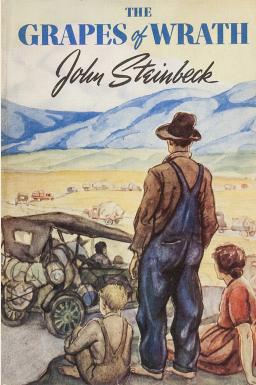


Sentiment: 0.94

- Recommended 5 Books: Mystery, Detective, Small Town America, Michael Connelly



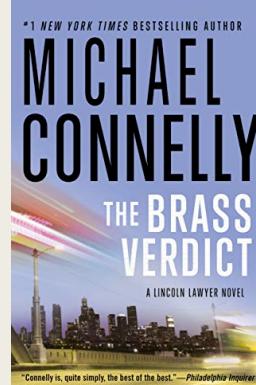
Prediction: 1.04



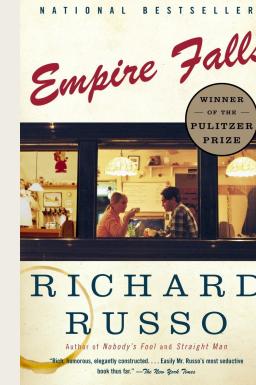
Prediction: 1.00



Prediction: 1.00



Prediction: 0.98



Prediction: 0.97

# Next Steps

- Implement model in production, evaluate it's performance on increasing customer engagement
- If successful, find ways to continuously train the model w/ new reviews
- If the model performs poorly, we can refine the Neural Network Architecture or adjust the process for data cleaning/preparation