

BOOK RECOMMENDER SYSTEM GROUP 1

Prepared by:

BAI XINYUE

LIANG YINGFANG

LIN TZU-CHUN

WANG JIALI





AGENDA

- 1 BACKGROUND
- 2 DATASET
- 3 MODEL
- 4 EXPERIMENT RESULT
- 5 WORKING SYSTEM
- 6 INSIGHTS



BACKGROUND

The screenshot shows the Goodreads homepage with a navigation bar at the top. The main content area is titled "Recommendations > Read Shelf". It displays a grid of book covers with their titles and authors. Each book has a "Want to Read" button below it. A sidebar on the right provides options like "Recommendation overview", "read", "Recommendation genres", "More Actions", "Recommendations", "Give Recommendations", "Ask for Recommendations", and "Books Marked as".

goodreads

Home My Books Browse ▾ Community ▾ Search books

Explore Sci-Fi & Fantasy Week >

Recommendations > Read Shelf

Here are some books we think you'll like based on the books you've added to [this shelf](#). Other readers with similar interests have enjoyed them. [How to improve your recommendations...](#)

updated: 0 minutes ago

View: covers | list

Quantum Roots - Kyle Keyes
George R.R. Martin's A Game of Thrones - George R.R. Martin
Stephen King's The Stand - Stephen King
Stephen King's IT - Stephen King
One Flew Over the Cuckoo's Nest - Ken Kesey

Want to Read ★★★★☆ Not interested

Thomas Harris' Red Dragon - Thomas Harris
Chuck Palahniuk's Fight Club - Chuck Palahniuk
Stephen King's The Shining - Stephen King
The Lost Starship - Vaughn Heppner
Robert A. Heinlein's The Moon Is a Harsh Mistress - Robert A. Heinlein

Want to Read ★★★★☆ Not interested

H. Beam Piper's Little Fuzzy - H. Beam Piper
William Peter Blatty's The Exorcist - William Peter Blatty
Dan Simmons' Hyperion - Dan Simmons
Van West's Ready Player One - Van West

GOODREADS

- A platform providing user recommendations for various books based on ratings by the users.
- When creating a new account, goodreads asks you to choose genres you like and to rate 20 books.

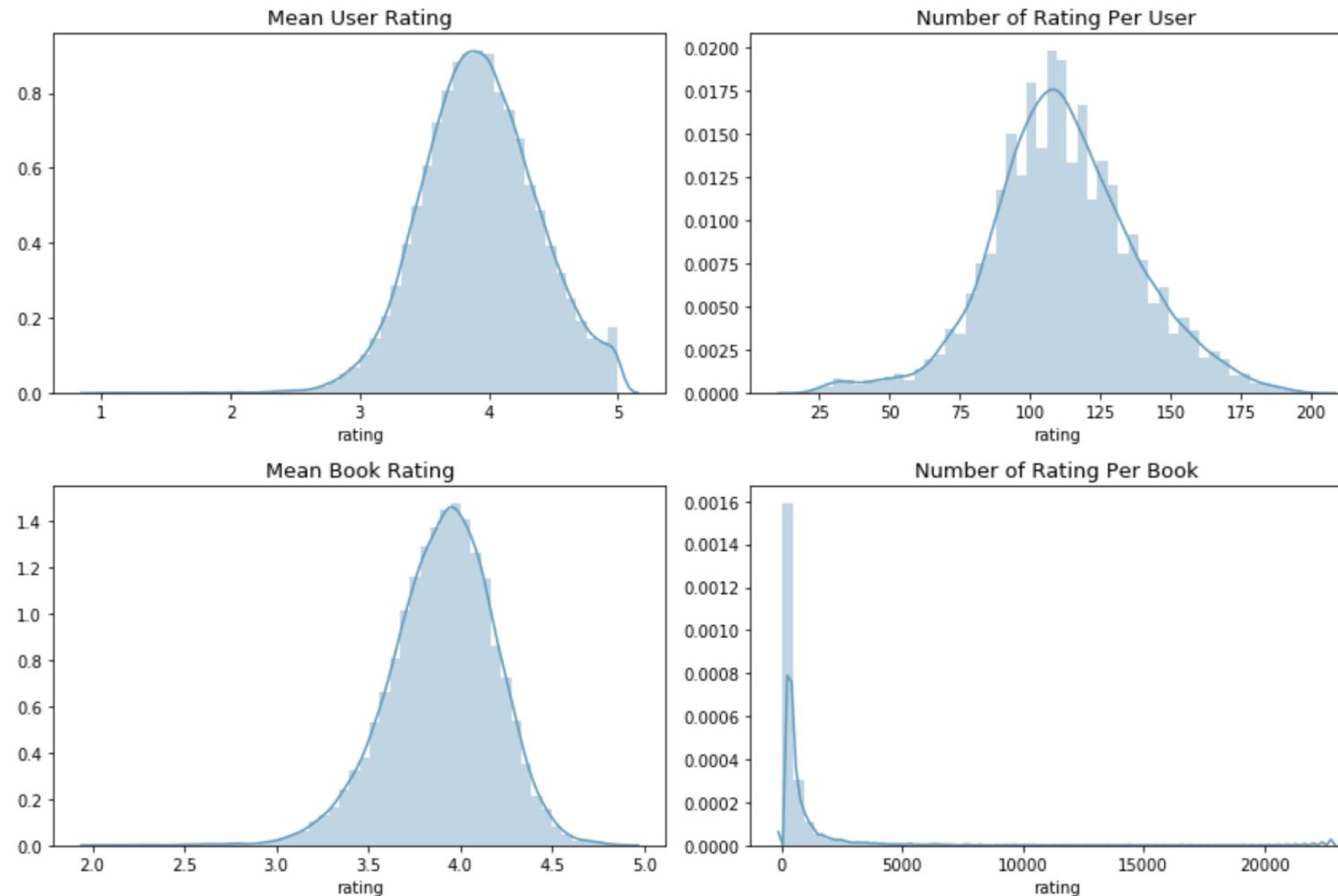
Problem

- Cold start: Users need to rate more than 20 books then they can get recommendations.
- Trying different algorithms for different situations.



DATASET

RATING DATASET



- Including 3 columns: rating, book_id, user_id
- Number of users: 53424
- Number of items: 10000
- Number of available ratings: 5,976,479
- Number of all possible ratings: 534,240,000
- SPARSITY: 98.88%

BOOK TAG DATASET

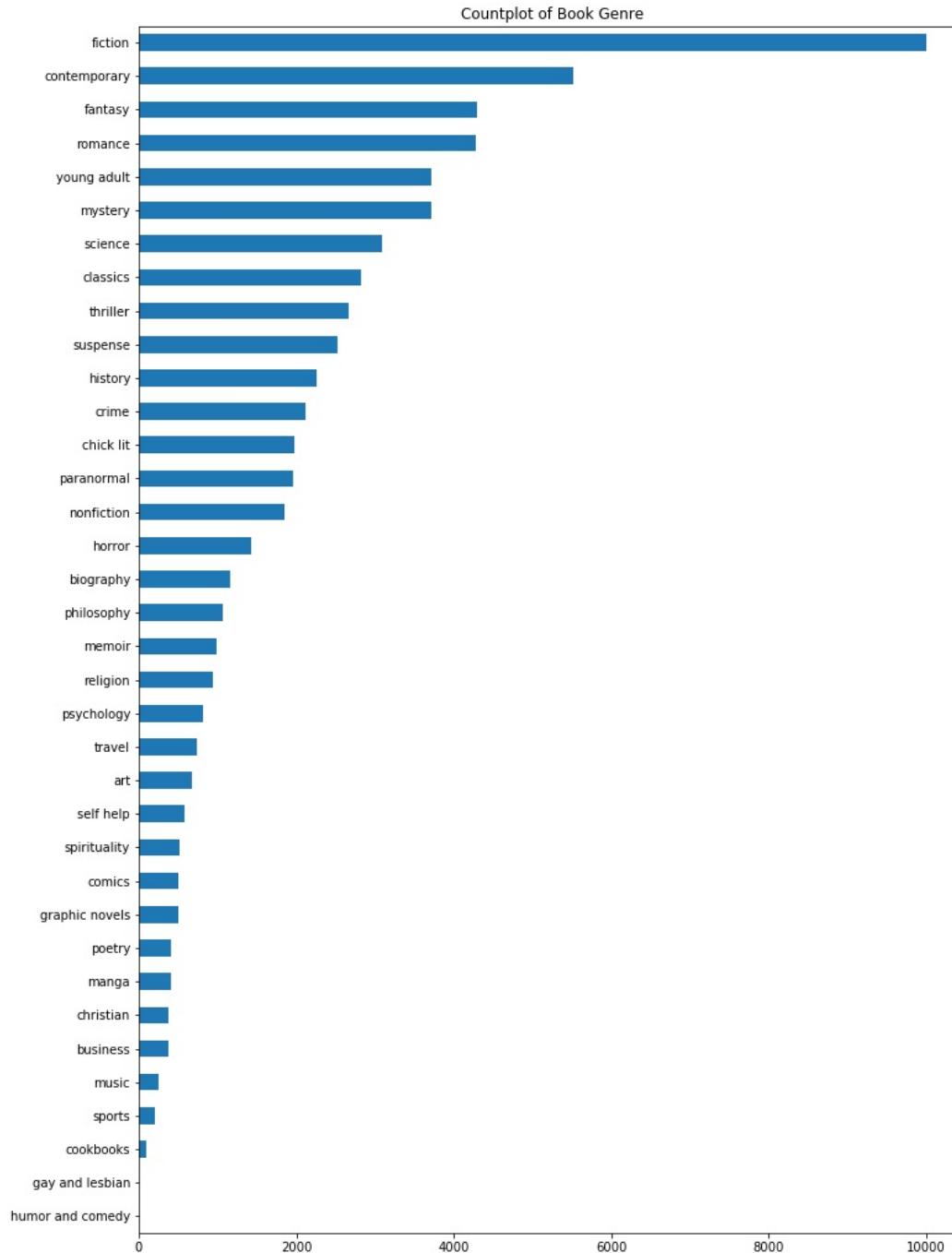
Users are allowed to tag books, so there are too many tags. Hence, we keep only 37 tags.

Before

- Number of Book Tag: 34252

After

- Unique genre number: 37
- Average tag number: 6.4795
- Unique book id number with tags: 10000





MODEL

ALGORITHM



Models

- Baseline (**Most Popular**)
- Simple Models (**BPR**)
- Deep-learning Models (**BiVAE**)
- Multimodality Models (**VBPR, CTR**)

MOST POPULAR

- Recommend books by their popularity

BPR & BiVAE

- Hyper-parameter Tuning

BPR

K=	300
Lambda_reg =	0.01
Learning_rate =	0.01

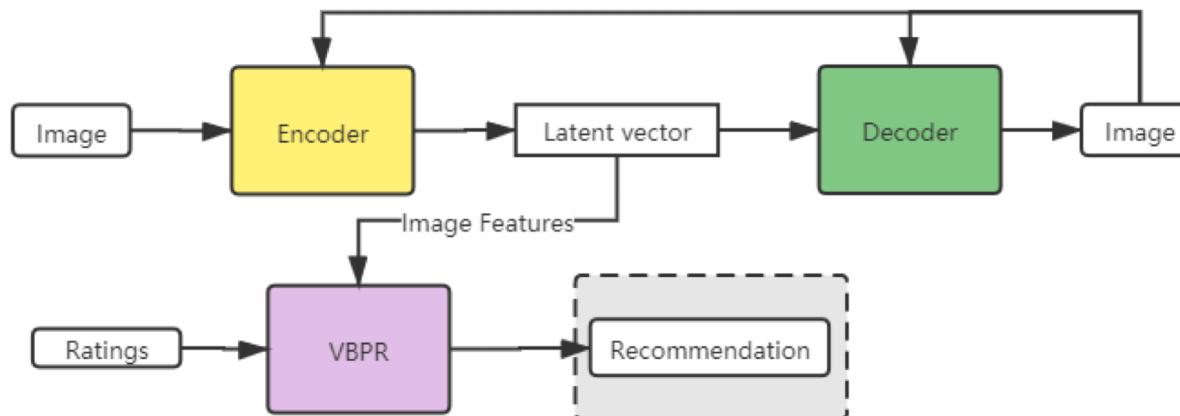
BiVAE

num_factors=	50
AE_layers=	[100]
Activation_func =	tanh
Likelihood=	pois
Learning_rate=	5.00E-04

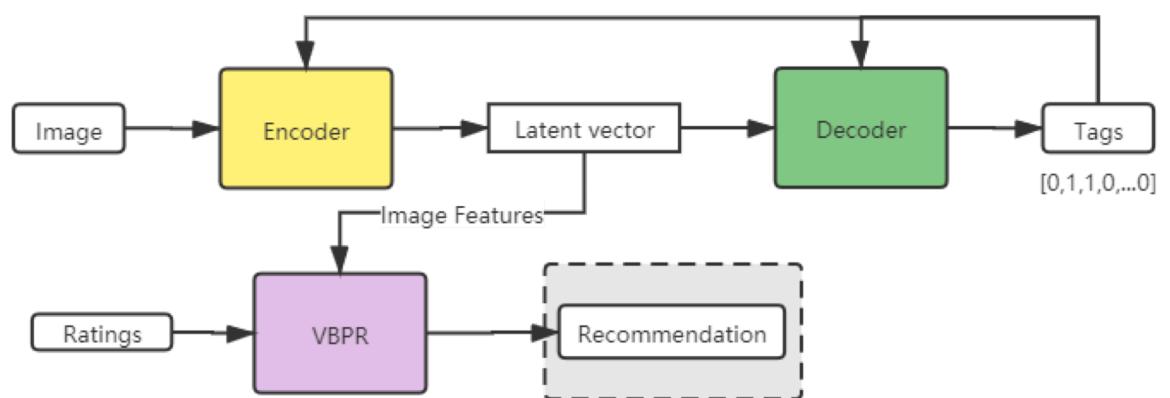
VBPR

- **Proposed Structure**

- ResNet + Dense -> Latent Vector -> Image



- ResNet + Dense -> Latent Vector -> Tag vector (book genre)



- The results for the two models are similar and not outperforming the other model.
- Overall, the image features might be hard to contribute to the recommendation.

CTR

- Generate the introduction of the book

The title of the book is <book_title>.
The original title of the book is <original_book_title>.
The authors of the book is <authors>.
<book_genre>.

- Hyperparameter Tuning

K =	[20,30,50,100,300]
a =	[1,2,5,10]
b =	[0.01,0.05,0.1]

* Numbers in **yellow** are those gives the best results

CTR

- Topic Words Analysis

Although the larger k gives better results, but when k=30, the top words in each topic show more associations.

25	love, chronicles, young, daniel, romance, steve, kate, c, end, rose
26	thriller, adult, mystery, classics, good, day, tom, princess, club, l
27	young, chick, dark, white, dont, raymond, d, know, hope, moon
28	philosophy, classics, vampire, harry, war, blood, children, potter, charles, like
29	suspense, history, robert, world, business, la, jim, janet, classics, empire
30	suspense, history, mark, lisa, mr, koontzfantasy, long, men, soul, andrew

* Topic words when k=30



EXPERIMENT RESULT

MODEL RESULT

Model	relevant item matters		ranking position matters		Harmonic Mean
	Recall@20	NDCG@20	NCRR@20		
Most Pop	0.0544	0.1083	0.1177	0.0831	
BPR	0.1118	0.0730	0.0876	0.0881	
VBPR	0.1353	0.0528	0.0266	0.0469	
CTR	0.2501	0.1025	0.0553	0.0942	
BiVAE	0.2262	0.1016	0.0598	0.0968	

- As we can see from the result table, BiVAE has the best overall performance with the highest harmonic mean. But surprisingly, our baseline model, item-based most pop has the highest NDCG and NCRR, this shows that our baseline model has the most accurate ranking prediction/ranking position prediction. (list of items recommended by Most Top is shown in the next slide).
- CTR model has the highest recall, implying that this model captures more relevant items -> book attributes does help in our recommendation
- VBPR has a very low NCRR and NDCG, implying poor ranking position prediction -> image as an additional modality, doesn't improve the ranking recommendation performance

BASELINE - TOP20 BOOK

book_id	authors	original_publication_year	title	language_code
0	Suzanne Collins	2008.0	The Hunger Games (The Hunger Games, #1)	eng
1	J.K. Rowling, Mary GrandPré	1997.0	Harry Potter and the Sorcerer's Stone (Harry P...)	eng
2	Harper Lee	1960.0	To Kill a Mockingbird	eng
3	Stephenie Meyer	2005.0	Twilight (Twilight, #1)	en-US
4	F. Scott Fitzgerald	1925.0	The Great Gatsby	eng
5	Suzanne Collins	2009.0	Catching Fire (The Hunger Games, #2)	eng
6	Suzanne Collins	2010.0	Mockingjay (The Hunger Games, #3)	eng
7	J.K. Rowling, Mary GrandPré, Rufus Beck	1999.0	Harry Potter and the Prisoner of Azkaban (Harr...	eng
8	J.K. Rowling, Mary GrandPré	1998.0	Harry Potter and the Chamber of Secrets (Harry...	eng
9	J.R.R. Tolkien	1937.0	The Hobbit	en-US
10	J.K. Rowling, Mary GrandPré	2000.0	Harry Potter and the Goblet of Fire (Harry Pot...	eng
11	J.K. Rowling, Mary GrandPré	2007.0	Harry Potter and the Deathly Hallows (Harry Po...	eng
12	J.K. Rowling, Mary GrandPré	2003.0	Harry Potter and the Order of the Phoenix (Har...	eng
13	J.K. Rowling, Mary GrandPré	2005.0	Harry Potter and the Half-Blood Prince (Harry ...	eng
14	George Orwell, Erich Fromm, Celâl Üster	1949.0	1984	eng
15	J.D. Salinger	1951.0	The Catcher in the Rye	eng
16	Stieg Larsson, Reg Keeland	2005.0	The Girl with the Dragon Tattoo (Millennium, #1)	eng
17	George Orwell	1945.0	Animal Farm	eng
18	William Golding	1954.0	Lord of the Flies	eng
19	Dan Brown	2000.0	Angels & Demons (Robert Langdon, #1)	en-CA

These are all very classic/well known book -> would have a higher rank for most of the users -> high NDCG, NCRR



WORKING SYSTEM

BOOK RECOMMENDATION WEBSITE

We developed a simple website to display our recommendation result by Django framework. Users can login into our system by their user id and see their own reading profile. In this reading profile page, we shows different statistics of user and those books we recommend to this user.

Search [Logout](#)

Your Interesting Stats

Your favourite books

Title	Authors	Average Rating
The Vampire's Assistant (Cirque Du Freak, #2)	Darren Shan	4.12
Tunnels of Blood (Cirque Du Freak, #3)	Darren Shan	4.14
Clara and Mr. Tiffany	Susan Vreeland	3.71
Hunters of the Dusk (Cirque Du Freak, #7)	Darren Shan	4.17
Allies of the Night (Cirque du Freak, #8)	Darren Shan	4.19

Your Favorite Genres

- horror
- paranormal
- thriller
- mystery
- young adult
- fantasy
- contemporary
- fiction

Your Average Rating
3.78

User 8563

You have read 18 books

Your Wishlist

Lord of the Flies
William Golding

Gone Girl
Gillian Flynn

The Hitchhiker's Guide to the Galaxy (Hitchhiker's Guide to the Galaxy, #1)

Books

Picks for you (7)

BOOK RECOMMENDATION WEBSITE

Our recommendation is based on the result of BiVAE model for users who already read more than 10 books. For those users who read less than 10 books, we use the Most Popular model to recommend some trending books for them.

Title	Authors	Average Rating
The Famished Road	Ben Okri	3.73
Miracles	C.S. Lewis	4.08
The Complete Tales and Poems	Edgar Allan Poe	4.48

Your Favorite Genres

- religion
- classics
- fantasy
- contemporary
- fiction
- christian
- philosophy
- nonfiction

Your Average Rating
3.67

You have read 3 books

Your Wishlist

- The Fellowship of the Ring (The Lord of the Rings, #1) J.R.R. Tolkien
- War and Peace Leo Tolstoy, Henry Gifford, Aylmer Maude, Louise Maude
- The Pilgrim's Progress John Bunyan
- Dead Souls Nikolai Gogol, Robert A. Maguire
- The Nicomachean Ethics Aristotle, J.A.K. Thomson, Jonathan Barnes, Hugh Tredennick

Picks for you (20)

The Hunger Games (The Hunger Games, #1)
Suzanne Collins

Harry Potter and the Sorcerer's Stone (Harry Potter, #1)
J.K. Rowling

To Kill a Mockingbird
Harper Lee

Twilight (Twilight, #1)
Stephenie Meyer

The Great Gatsby
F. Scott Fitzgerald



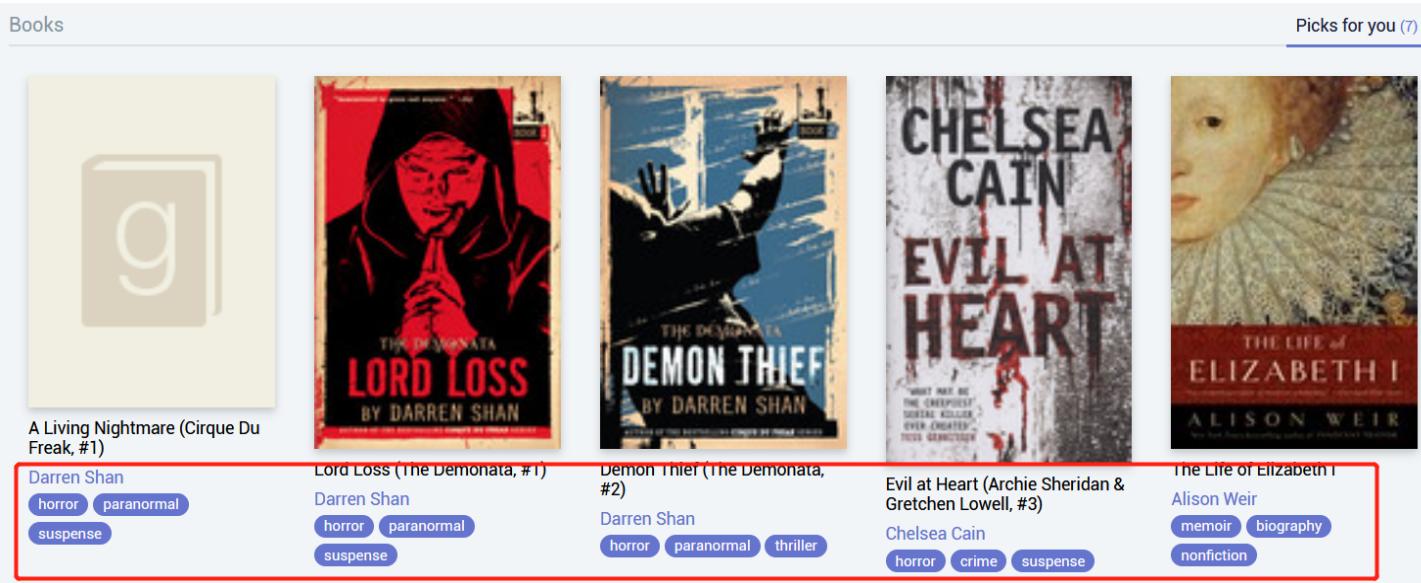
INSIGHT

INSIGHT

Your Interesting Stats

Your favorite books			
Title	Authors	Average Rating	Your Favorite Genres
The Vampire's Assistant (Cirque Du Freak, #2)	Darren Shan	4.12	horror paranormal thriller
Tunnels of Blood (Cirque Du Freak, #3)	Darren Shan	4.14	mystery young adult
Clara and Mr. Tiffany	Susan Vreeland	3.71	fantasy contemporary
Hunters of the Dusk (Cirque Du Freak, #7)	Darren Shan	4.17	fiction
Allies of the Night (Cirque du Freak, #8)	Darren Shan	4.19	

We can see that our model can provide recommendation consist with users' reading history.



And we also noticed that for this kind of book recommendation system based on users' historical data, sufficient amount of data is very important.

In order to solve this problem, we combined the Most Popular model with our deep learning model to provide different kind of users.

1

O

Thank You

INC

Y

H

E A

R

T

Do you wear a watch? Are you ever more comprehensively surrounded by madnesses. Do you get out there at a crack of dawn? We had a few crazy days, finished a crazy year. It's a new one that seems to be everything except normal. Extraordinary. It's wonderful, putting 'the crack of dawn' into context. Time is a subjective experience. What was your 'eureka' moment last year? Where does one year end and another begin? Were you wearing a watch?