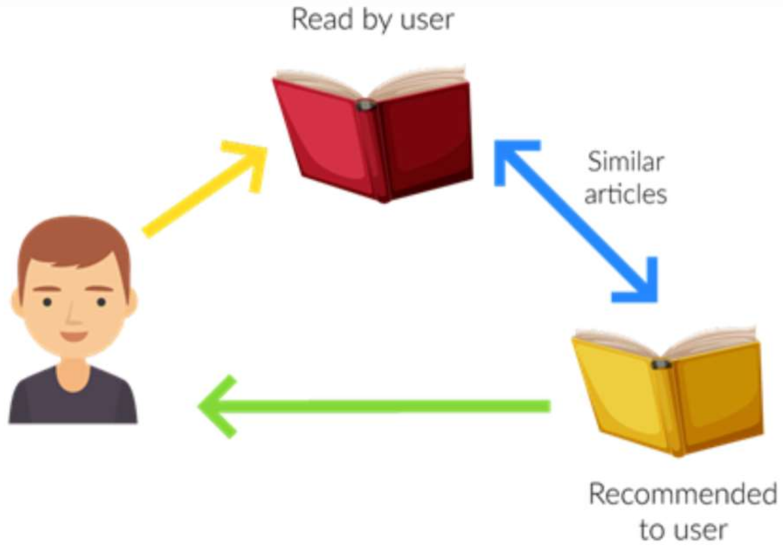
Recommendation System



Group 3
Ankita Bagaria, Hao Chun Niu (Eric), Jing Xie, Riya Khurana, Travis St. Louis

Outline

01 Business Problem

02 Collaborative Filter Recommender System

03 Technology Used

04 System Instruction

05 Future Steps

Business Problem

Value for Customer

Improve Customer Experience

Discover new things

Reduced transaction costs

Complementary items

Value for Business

Improved Customer Loyalty

Personalized Service

Increased sales & conversion rate

More customer knowledge

Item Based Recommender System

Users	Harry Potter	A Tale of Two Cities	The Silent Patient	Alice in Wonderland
	5	1	5	4
	4	4	3	3
	3	5	2	3
	1	2	4	?

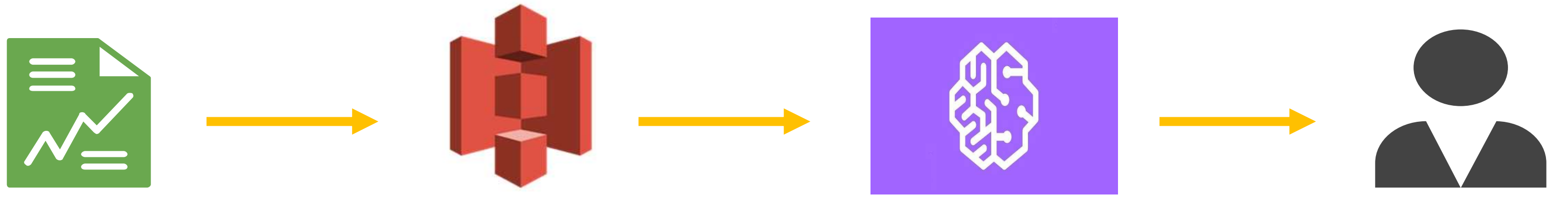
Item to Item Similarity

$$Similarity(\vec{A},\vec{B}) = \frac{\vec{A}\cdot\vec{B}}{||\vec{A}||*||\vec{B}||}$$

Prediction Computation

$$rating(U,I_i) = \frac{\sum_j rating(U,I_j)*s_{ij}}{\sum_j s_{ij}}$$

Technology Used



Data Set

Historical data obtained on books rating from Kaggle

Amazon S3

Stored data into S3 for further processing

Amazon Sagemaker

Built memory based recommendation system within Sagemaker

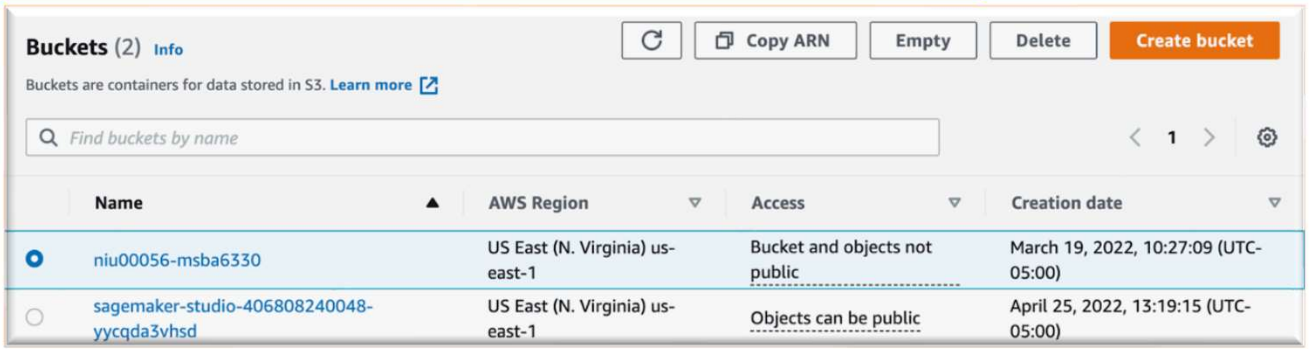
End users

Promote cross-selling by showing top recommended books to customers

The implementation of recommendation system is quite straightforward and requires only a storage platform and Machine Learning tool

S3 Data Storage Process Flow

Step1. Create S3 Storage Bucket

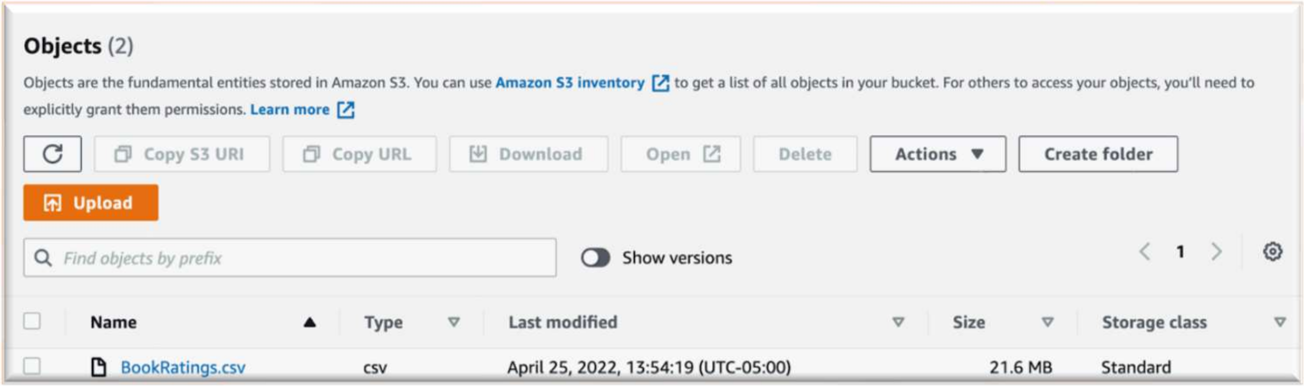


Buckets (2) Info

Buckets are containers for data stored in S3. Learn more

	Name	AWS Region	Access	Creation date
●	niu00056-msba6330	US East (N. Virginia) us-east-1	Bucket and objects not public	March 19, 2022, 10:27:09 (UTC-05:00)
○	sagemaker-studio-406808240048-yycqda3vhsd	US East (N. Virginia) us-east-1	Objects can be public	April 25, 2022, 13:19:15 (UTC-05:00)

Step2. Upload Data into Bucket



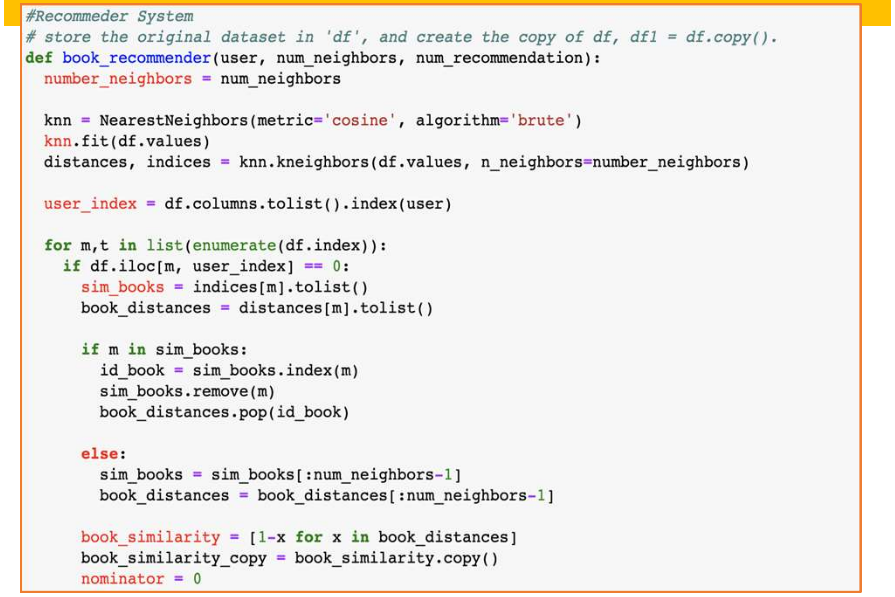
Objects (2)

Objects are the fundamental entities stored in Amazon S3. You can use Amazon S3 inventory to get a list of all objects in your bucket. For others to access your objects, you'll need to explicitly grant them permissions. Learn more

□	Name	Type	Last modified	Size	Storage class
□	BookRatings.csv	csv	April 25, 2022, 13:54:19 (UTC-05:00)	21.6 MB	Standard

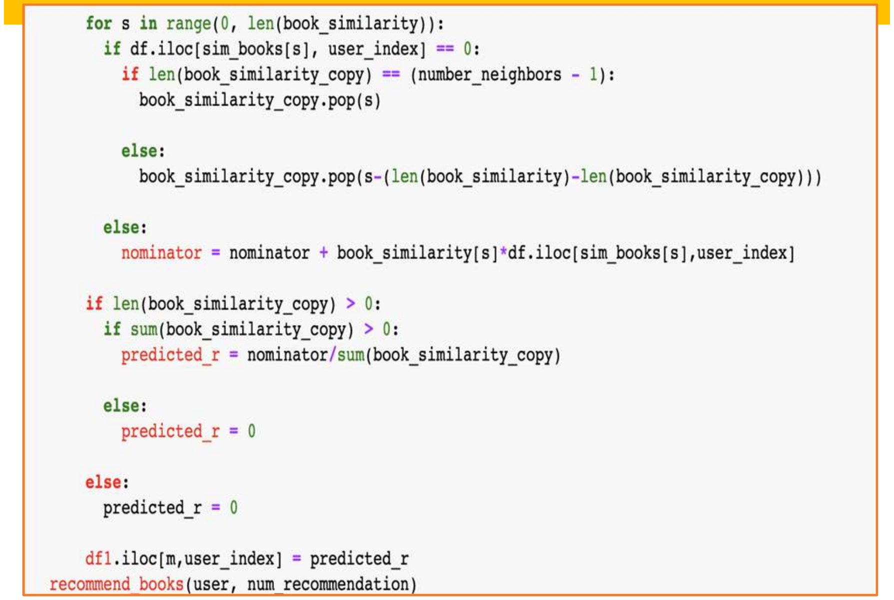
Model Development Using Sagemaker

Step1. Calculate Book Similarity



```
#Recommeder System
# store the original dataset in 'df', and create the copy of df, df1 = df.copy().
def book_recommender(user, num_neighbors, num_recommendation):
  number_neighbors = num_neighbors

  knn = NearestNeighbors(metric='cosine', algorithm='brute')
  knn.fit(df.values)
  distances, indices = knn.kneighbors(df.values, n_neighbors=number_neighbors)

  user_index = df.columns.tolist().index(user)

  for m,t in list(enumerate(df.index)):
    if df.iloc[m, user_index] == 0:
      sim_books = indices[m].tolist()
      book_distances = distances[m].tolist()

      if m in sim_books:
        id_book = sim_books.index(m)
        sim_books.remove(m)
        book_distances.pop(id_book)

      else:
        sim_books = sim_books[:num_neighbors-1]
        book_distances = book_distances[:num_neighbors-1]

      book_similarity = [1-x for x in book_distances]
      book_similarity_copy = book_similarity.copy()
      nominator = 0
```

Step2. Make Predictions



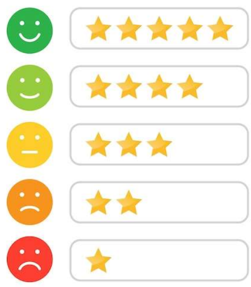
```
    for s in range(0, len(book_similarity)):
      if df.iloc[sim_books[s], user_index] == 0:
        if len(book_similarity_copy) == (number_neighbors - 1):
          book_similarity_copy.pop(s)

        else:
          book_similarity_copy.pop(s-(len(book_similarity)-len(book_similarity_copy)))

      else:
        nominator = nominator + book_similarity[s]*df.iloc[sim_books[s],user_index]

    if len(book_similarity_copy) > 0:
      if sum(book_similarity_copy) > 0:
        predicted_r = nominator/sum(book_similarity_copy)

      else:
        predicted_r = 0

    else:
      predicted_r = 0

    df1.iloc[m,user_index] = predicted_r
recommend_books(user, num_recommendation)
```

System Instruction - Result

Input

User ID

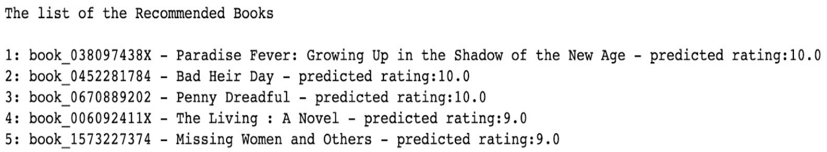
Items



Ratings

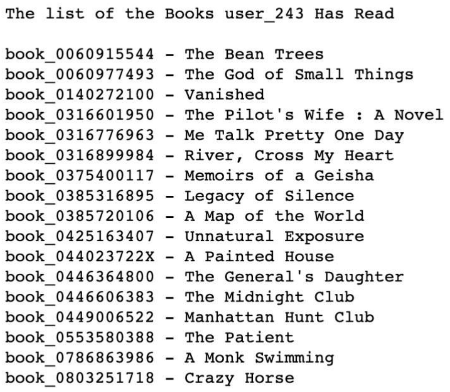
Output

Personalized Recommendations



```
The list of the Recommended Books

1: book_038097438X - Paradise Fever: Growing Up in the Shadow of the New Age - predicted rating:10.0
2: book_0452281784 - Bad Heir Day - predicted rating:10.0
3: book_0670889202 - Penny Dreadful - predicted rating:10.0
4: book_006092411X - The Living : A Novel - predicted rating:9.0
5: book_1573227374 - Missing Women and Others - predicted rating:9.0
```

Books Already Read



```
The list of the Books user_243 Has Read

book_0060915544 - The Bean Trees
book_0060977493 - The God of Small Things
book_0140272100 - Vanished
book_0316601950 - The Pilot's Wife : A Novel
book_0316776963 - Me Talk Pretty One Day
book_0316899984 - River, Cross My Heart
book_0375400117 - Memoirs of a Geisha
book_0385316895 - Legacy of Silence
book_0385720106 - A Map of the World
book_0425163407 - Unnatural Exposure
book_044023722X - A Painted House
book_0446364800 - The General's Daughter
book_0446606383 - The Midnight Club
book_0449006522 - Manhattan Hunt Club
book_0553580388 - The Patient
book_0786863986 - A Monk Swimming
book_0803251718 - Crazy Horse
```

Limitations

1. Limited computational resources

2. Number of recommendations

3. Measurement of recommendation system

4. Finding niche items

Conclusion

- Business value
- How recommendation system works
- Our book recommendation system
  - Technology used
  - Prediction process
  - Recommendation output
- Limitations of recommender system
- Benefits
  - Improved customer experience
  - Personalized service
  - Increased sales