E-COMMERCE RECOMMENDER SYSTEM

Pranavi Kolipaka, Sai Vuruma

UNIVERSITY OF
South Carolina

College of Engineering and Computing

AGENDA

- Introduction
- Related Work
- Dataset
- Model Building
- Evaluation
- Q&A

Engineering and Computing
UNIVERSITY OF SOUTH CAROLINA

INTRODUCTION

- Recommender Systems (or RecSys as they are widely known) can be found in many modern day applications ranging from e-commerce websites like Amazon to music platforms like Spotify.

- They leverage machine learning algorithms to give users a list of items that are relevant to the item that they are currently looking for.

- Criteria: past purchases, user demographics, behavioral data, etc.

RELATED WORK

- **ShopperBERT [1]** extracts a universal user representation pre performs the pretext tasks predicting masked information of the [MASK] token with the final hidden vector of that token and averages the final hidden vectors of the behavior tokens which will be used as a user representation.

- **Islek [2]** uses a DeepIDRS approach with a two-level hierarchy. This hierarchical structure provides a personalized, agile and explainable recommendation system which uses bidirectional encoder representations to provide more accurate recommendation results.

RELATED WORK

- **Chu [3]** adopts Word2Vec to extract information from the comments users have made on the items bought. PCA is used to reduce the dimensionality and a clustering algorithm is used to group similar data points. Finally, ICRRS is applied to make item recommendations to the users

- **Loukili [4]** developed a Machine Learning based recommender algorithm to suggest personal recommendations to customers using association rules via the Frequent Pattern-Growth (FP-growth) algorithm. The limitation of this work is that some of the evaluation characteristics of a recommender system, such as diversity and explainability, are difficult to define.

DATASET

- Behavioral data collected from around 285 million users of a large eCommerce store, publicly available on Kaggle.
- Time period: Oct-Nov 2019
- Schema:
  - event_time, event_type, category_code, brand and user_session are object data
  - product_id, category_id and user_id are integers
  - price is floating-point data.
- 2019-Oct dataset consists of 42448764 rows and 2019-Nov dataset consists of 67501979 rows.

DATASET: EDA

```
df1.describe()
```

	product_id	category_id	price	user_id
count	42,448,764.0	42,448,764.0	42,448,764.0	42,448,764.0
mean	10,549,932.375842676	2.0574042379407987e+18	290.3236606849143	533,537,147.50816935
std	11,881,906.970608113	1.843926466140415e+16	358.26915533940263	18,523,738.17465412
min	1,000,978.0	2.0530135522261076e+18	0.0	33,869,381.0
25%	1,005,157.0	2.0530135554641106e+18	65.98	515,904,318.0
50%	5,000,470.0	2.0530135556318828e+18	162.93	529,696,452.0
75%	16,000,305.0	2.0530135634248998e+18	358.57	551,578,838.25
max	60,500,010.0	2.1754195950939676e+18	2,574.07	566,280,860.0

```
df2.describe()
```

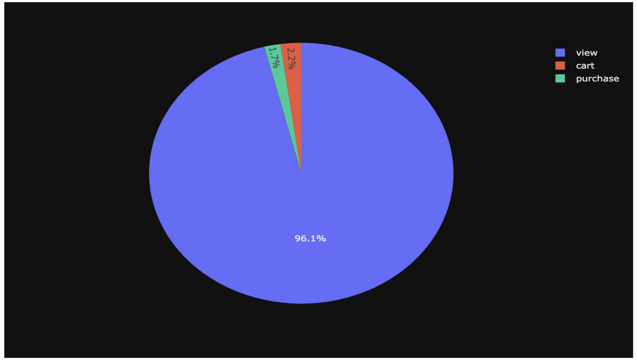
	product_id	category_id	price	user_id
count	67,501,979.0	67,501,979.0	67,501,979.0	67,501,979.0
mean	12,514,064.889882294	2.0578976443220687e+18	292.45931656462915	538,639,745.6296743
std	17,257,413.629846174	2.012549032884276e+16	355.67449958606664	22,885,161.05152215
min	1,000,365.0	2.0530135522261076e+18	0.0	10,300,217.0
25%	1,305,977.0	2.0530135553215043e+18	69.24	516,476,241.0
50%	5,100,568.0	2.0530135556318828e+18	165.77	535,057,264.0
75%	17,300,752.0	2.0530135636513923e+18	360.34	561,079,379.0
max	100,028,554.0	2.187707861038007e+18	2,574.07	579,969,851.0

The standard deviation value of product_id and price is higher than their respective mean value. This indicates that there is a high variation between values, and abnormal distribution for data.

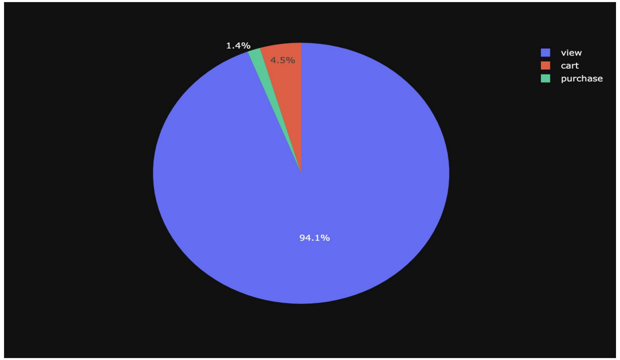
Engineering and Computing
UNIVERSITY OF SOUTH CAROLINA

DATASET: EDA

	event_type	count_of_users	prcnt
0	view	40779399	96.1
1	cart	926516	2.2
2	purchase	742849	1.7

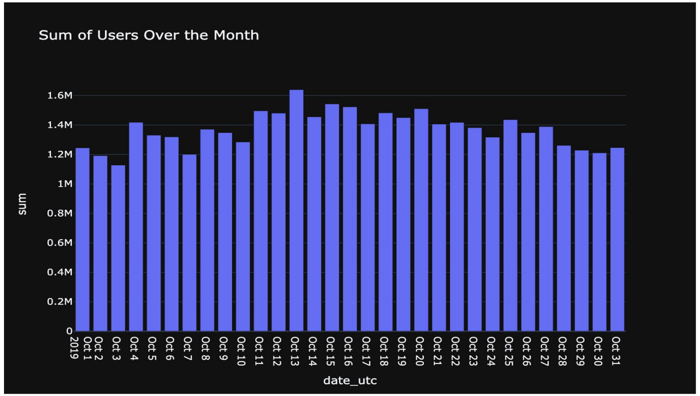


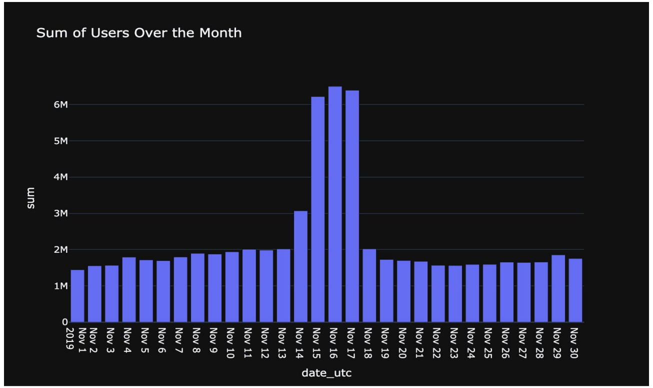
	event_type	count_of_users	prcnt
0	view	63556110	94.2
1	cart	3028930	4.5
2	purchase	916939	1.4



People who just viewed the items are highest. Second being those who just carted their items, then the people who bought the items are least.

DATASET: EDA





In Oct, over all days there are many users who engaged with the shop. Whereas in Nov, the number of users engaged with the shop during mid-month is more compared to other days.

Engineering and Computing
UNIVERSITY OF SOUTH CAROLINA

MODEL: SEQUENTIAL RECOMMENDATION

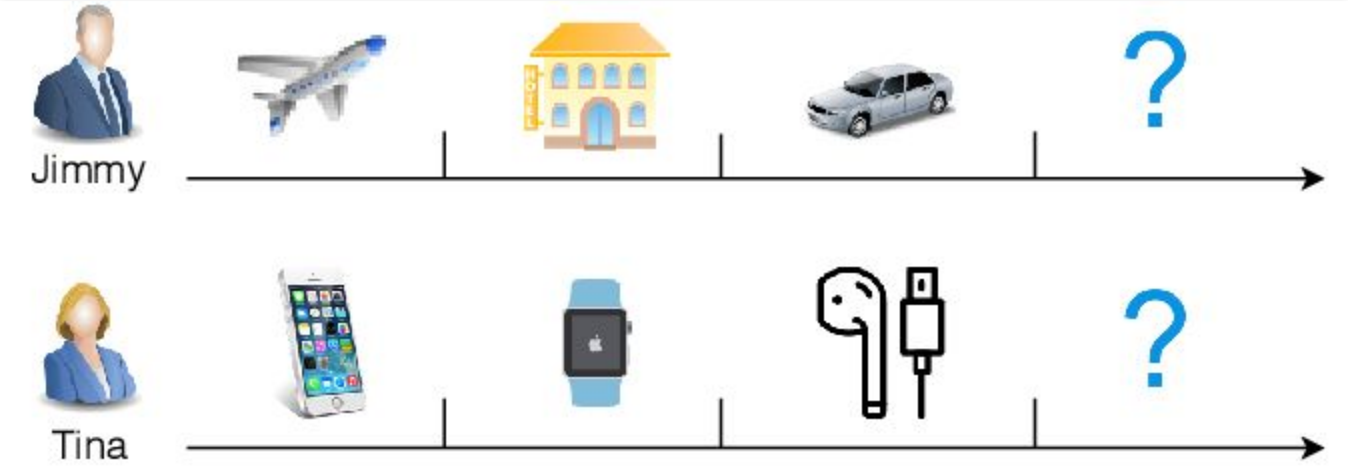


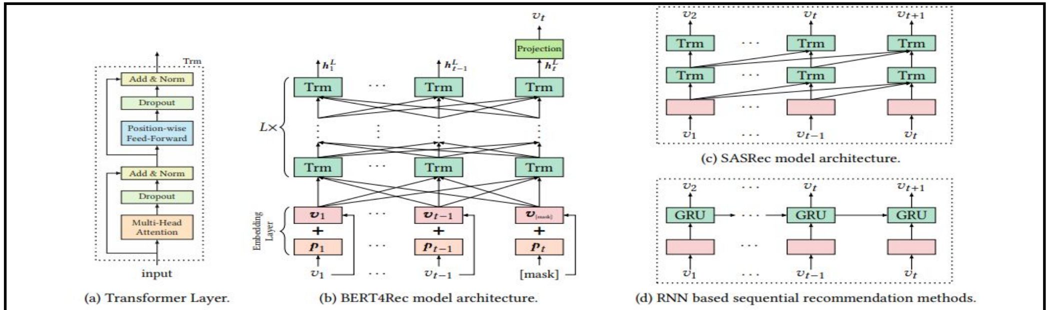
Figure 1: Two examples of SRSs: (1) After Jimmy has booked a flight, a hotel and rented a car, what will be his next action? (2)

Source: Sequential Recommender Systems: Challenges, Progress and Prospects

Engineering and Computing
UNIVERSITY OF SOUTH CAROLINA

MODEL: BERT4REC

- Building on the strong points for the original BERT model, BERT4Rec adopts transformers to sequential recommendation tasks.
- It uses bidirectional self-attention to model users' behavior sequences to improve quality of recommendations.



(a) Transformer Layer. (b) BERT4Rec model architecture. (c) SASRec model architecture. (d) RNN based sequential recommendation methods.

Figure 1: Differences in sequential recommendation model architectures. BERT4Rec learns a bidirectional model via Cloze task, while SASRec and RNN based methods are all left-to-right unidirectional model which predict next item sequentially.

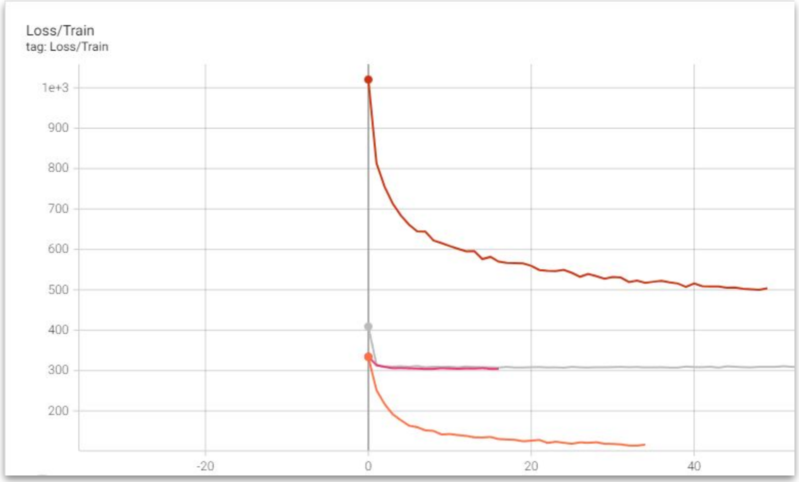
Source: [5]

Engineering and Computing
UNIVERSITY OF SOUTH CAROLINA

MODEL: TRAINING

Specs:

- embedding_size: 64
- n_layers: 2
- n_heads: 2
- dropout_prob: 0.3
- loss_type: 'CE'
- mask_ratio: 0.2
- hidden_act: 'gelu'
- epochs: 50
- train_batch_size: 4096
- learner: adam
- learning_rate: 0.05



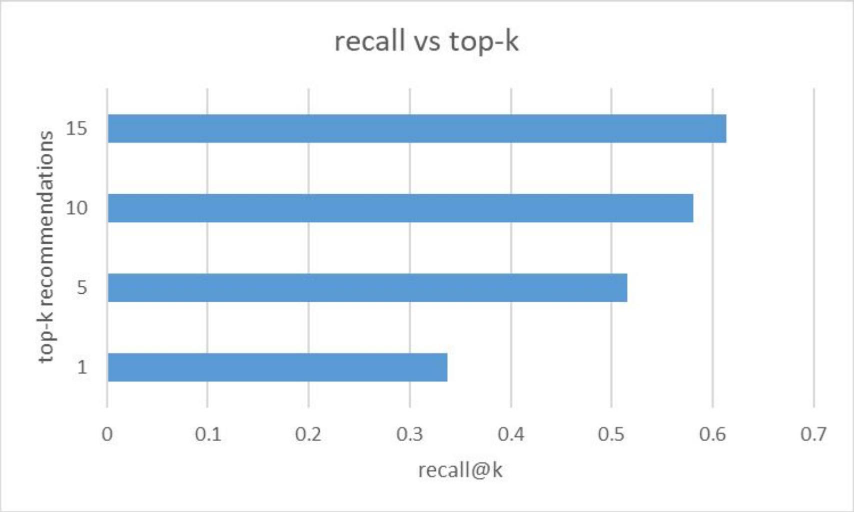
EVALUATION: METRICS

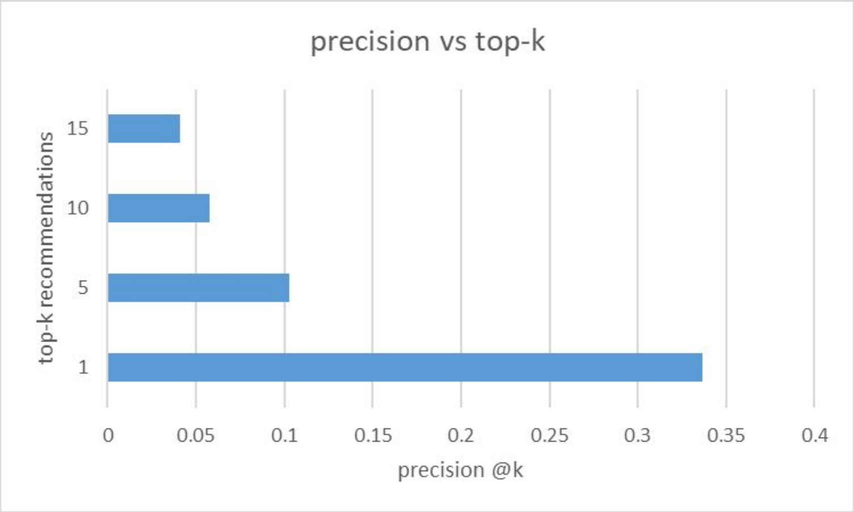
1. **Precision @K, Recall @K:** Similar to how precision and recall are usually calculated, but adjusted for the top-K number of predictions made by the model.

2. **Item Coverage @K:** It measures the proportion of items that a recommender system can recommend, and the measure increases as the size of the recommendation list increases.

3. **Normalized Discounted Cumulative Gain @K:** Also known as NDCG, this metric can be used to calculate a cumulative score of an ordered set of items. Items of higher relevance appearing at lower ranks are penalized thus resulting in a more grounded score.

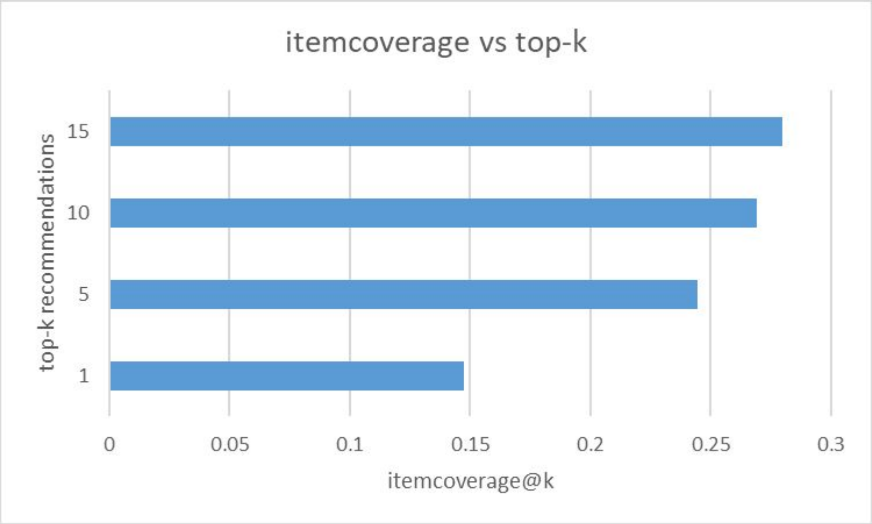
EVALUATION: RESULTS

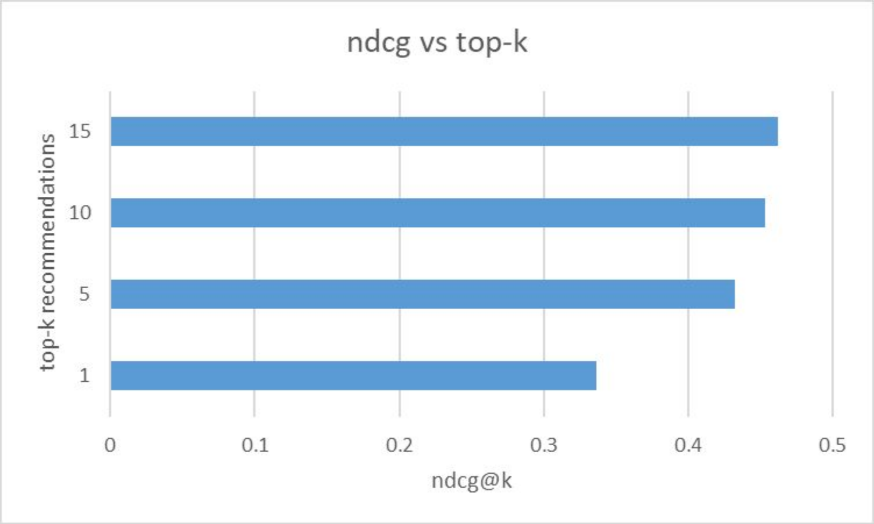




Engineering and Computing
UNIVERSITY OF SOUTH CAROLINA

EVALUATION: RESULTS





CONCLUSION

- We built a BERT4Rec model for sequentially recommending products on a ecommerce dataset with the following performance metrics: Recall@10 of about 0.6.

- In the future, this work can be expanded to include data from the other product-specific columns such as price, brand, etc. for contextual recommendation.

Engineering and Computing
UNIVERSITY OF SOUTH CAROLINA

REFERENCES

1. Shin, Kyuyong, et al. "One4all user representation for recommender systems in e-commerce." *arXiv preprint arXiv:2106.00573* (2021).
2. Islek, Irem, and Sule Gunduz Oguducu. "A hierarchical recommendation system for E-commerce using online user reviews." *Electronic Commerce Research and Applications* 52 (2022): 101131.
3. Chu, Pang-Ming, and Shie-Jue Lee. "A novel recommender system for E-commerce." *2017 10th international congress on image and signal processing, biomedical engineering and informatics (CISP-BMEI)*. IEEE, 2017.
4. Loukili, Manal, Fayçal Messaoudi, and Mohammed El Ghazi. "Machine learning based recommender system for e-commerce." *IAES International Journal of Artificial Intelligence* 12.4 (2023): 1803-1811.
5. Sun, Fei, et al. "BERT4Rec: Sequential recommendation with bidirectional encoder representations from transformer." Proceedings of the 28th ACM international conference on information and knowledge management. 2019.

THANK YOU!

Pranavi Kolipaka
kolipakv@email.sc.edu

Sai Vuruma
svuruma@email.sc.edu

Engineering and Computing
UNIVERSITY OF SOUTH CAROLINA