Latex file: <https://www.overleaf.com/project/6088b0ad33233523cd3c1a9f>

Data set citation: <https://cseweb.ucsd.edu/~jmcauley/datasets.html#multi_aspect>

**Problem Statement**

*what is your problem formulation?*

It is difficult for people to choose a new beer they would enjoy online.

*why is this problem important and interesting? what are the challenges in this problem?*

According to [American Heart Association](https://www.heart.org/en/news/2020/07/01/covid-19-pandemic-brings-new-concerns-about-excessive-drinking), when the pandemic started, alcohol consumption overall went up by 60%, with the increase of 500% in online alcohol sales. In these stressful and lonely months of quarantine, a lot of people choose to up their alcohol consumption to cope with this crazy world. However, drinking became much lonelier : breweries, tasting spaces ( taproom or retail stores), bars… they are no longer widely safe and available options for many people to try out new options with their friends. When people get sick of their usual drinks, it is now much harder to discover the next new favorite. With the increase in online sale of beers and alcohol, and the difficulties for consumers to try new options - we want to build the future of beer discovery: a beer recommendation system that is your personalized online bartender, “pouring” you your next favorite cold beer - bye bye, Corona!

**A comprehensive literature review**

*what are the typical methods to solve your problem?*

* Popularity-based recommendation system: recommend items that are in trend now.
* Content-based recommendation system: recommend item based on similarity between items
* Collaborative filtering: recommend item based on other similar users’ preferences
* Matrix decomposition: recommend item based on latent features that represent similarity between users and items

*what are the state-of-the-art methods to solve your problem? for those methods, what are their pros and cons?*

Link to the Amazon paper (This is way too hard for our class): <https://assets.amazon.science/96/71/d1f25754497681133c7aa2b7eb05/temporal-contextual-recommendation-in-real-time.pdf>

The state-of-the-art model is the HRNN-meta recommendation system proposed in the paper “[Temporal-Contextual Recommendation in Real-Time](https://assets.amazon.science/96/71/d1f25754497681133c7aa2b7eb05/temporal-contextual-recommendation-in-real-time.pdf)”. The model preserves the temporal order of the data and adapts to new user trends. The session-based model also predicts the best timing to make recommendations based on the user’s interaction during the session. In addition, the model uses item state evolution as well as contextual meta-data other than time to overcome the cold-start problem of recommendation systems. Unlike other RNN models that find it hard to train with a large number of items, the HRNN-meta model uses importance sampling to speed up the training process. However, this model is not suitable for our dataset as our dataset does not contain user session data.

*If you plan to use some complex techniques/methods in your own method, please also discuss these techniques/methods.*

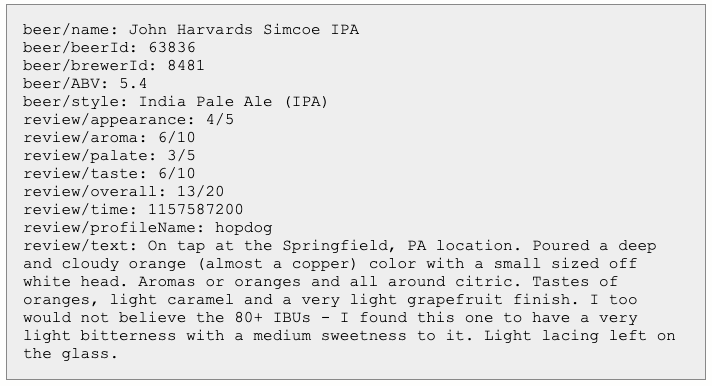
We can probably build a RNN based recommendation system to predict future recommended item based on users’ preferences in the present.

**Work plan**

*what are the data you plan to use?*

Data set citation: <https://cseweb.ucsd.edu/~jmcauley/datasets.html#multi_aspect>

These datasets (1998-2011) include reviews with multiple rated dimensions. The most comprehensive of these are beer review datasets from Ratebeer and Beeradvocate, which include sensory aspects such as taste, look, feel, and smell.

An example of beer data: 

*which evaluation metric do you plan to use?*

Some applicable metrics copy-pasted from the lecture note: Precision@k, recall@k; AUC, Area under Precision-Recall Curve, MRR

*what are the baselines you plan to compare against?*

K-Nearest-Neighbor recommender or random recommender (randomly recommend items to users).

**Potential schedule**

Week 5: Proposal + initial scope of project

Week 6: Research recommender models

Week 7: Preprocess dataset (EDA, parsing, cleaning, train-test split, and etc), begin implementing models

Week 8: Model implementation

Week 9: Fine tune models and improve model performances, work on report and record presentation

Week 10: June 3 (Report Due) up to 8 pages project report (exclude reference)

**Division of work (plan)**

Research models - Everyone

Model preprocessing - Person A, B

Model 1 implementation - Person A, C

Model 2 implementation - Person B, D

Fine Tune Models - Person C, D

Report, Presentation - Everyone

**Reference**

The state-of-the-art recommendation system: <https://towardsdatascience.com/decoding-state-of-the-art-recommender-system-38ee800f6fe><https://assets.amazon.science/96/71/d1f25754497681133c7aa2b7eb05/temporal-contextual-recommendation-in-real-time.pdf>

Similar projects:

<https://mohitatgithub.github.io/2018-03-20-Building-a-Beer-Recommendation-System/>

<https://medium.com/@medfordxie/what-to-drink-next-a-simple-beer-recommendation-system-using-collaborative-filtering-b65dd32b600d>

<https://towardsdatascience.com/creating-a-brewery-recommender-with-doc2vec-15ca20e28e7a>