

Coding Practices

All codes must be pushed to a remote git repository following standard coding practices.

- Code formattings: [Black](#), [Flake8](#)
- [Pre-commit](#)
- [PEP8](#)
- [Writing good commit messages](#)

One sprint means a week.

Sprint 1

This sprint focuses on the basics of any ML project: handling the data. It covers the necessary steps required in any recommendation project. It also touches on NLP to cover text data manipulation.

| Week | Goal | Resource | Additional Resource |
|--------|---|--|--|
| Week 1 | Data Wrangling | Book Chapters 1, 2, 3 | |
| | Text Preprocessing <ul style="list-style-type: none">- Tokenization- N-grams- Stemming and Lemmatization- Stop words, punctuation removal- Encoding:<ul style="list-style-type: none">- Bag of words- TF-IDF | Text Cleaning Text Preprocessing Text Encoding | Recommendation Notes (Download to view) Book Pg. 43, 44, 45 |
| | Embeddings <ul style="list-style-type: none">- Self-trained- W2V (Skip-gram & CBOW)- sBERT | E-commerce product embeddings Illustrated Word2Vec sBERT | Cbow and skip gram |
| | Association Rule Mining (ARM) based recommendation for Next item recommendation | ARM Kaggle ARM Rec (Only ARM part) | Resource Link |
| | Git and GitHub | | |

Deliverables

1. A notebook that shows all text preprocessing steps for a dataset of your choice.
2. A recommendation system using ARM on the Instacart dataset.

Sprint 2

This sprint focuses on different types of recommendations. It covers everything required, from building a basic recommendation system to evaluating them and finally deploying through APIs.

| Week | Goal | Resource | Additional Resource |
|--------|--|---|---|
| Week 2 | Content-based recommender | Book Chapter 4 | Content Content & collaborative Recommender Systems 101 |
| | Similarity measures, dimensionality reduction, | Book Chapter 5 | |
| | Collaborative Filtering | Book Chapter 6 | |
| | Hybrid Recommendation | Book Chapter 7 | |
| | Cold Start, Long Tail | Cold Start Problem Long Tail Problem | |
| | Evaluation Metrics | Metrics Book Pg. 92 Evaluation Metrics(metrics@k) | |
| | Deployment using FastAPI | Blog | |
| | Containerization using Docker | Videos Blog | |

Deliverables

API endpoints that are containerized for:

3. Movie recommendation using content-based recommendation
4. Movie recommendation using collaborative filtering

5. Movie recommendation using a hybrid method

Sprint 3

This sprint focuses on building a recommendation system using deepCTR, a library for click-through-rate(CTR) prediction. It covers various aspects, including data preparation, model training, and MLflow integration.

| Week | Goal | Resource | Additional Resource |
|--------|--|--|---|
| Week 3 | Need for factorization machines | Blog | |
| | Deep Factorization Machines | Neural Field Aware Factorization Machines | |
| | MLFlow | Mlflow | Blog |
| | DeepCTR <ul style="list-style-type: none">- SparseFeat vs DenseFeat- Embeddings- Converting data to model format- Training model- Evaluating model- Inference | Examples | |
| | Development of a landing page model | | |
| Week 4 | Sequential Recommender | Sequential Recommender Systems: Challenges, Progress and Prospects | Meal Recommendation (Optional) |
| | Next Item Recommendation | Next-item-recommendation in short sessions | Next-item Rec in short sessions |
| | RecBole <ul style="list-style-type: none">- Config Settings- Atomic Files- Training Model- Evaluating Model- Inference | RecBole | |

Deliverables

6. Landing page food recommendation using DeepCTR
7. Add-to-cart food Recommendation using RecBole.

Sprint 4

This sprint focuses on building a recommendation system using the Learning-to-rank(LTR) approach.

| Week | Goal | Resource | Additional Resource |
|--------|-------------------------------------|---|--|
| Week 6 | Google Recommendation System Course | Course | |
| | Pointwise, Pairwise, and Listwise | Blog | |
| | Introduction to Kaggle Challenge | Challenge Problem and Winning Solution Walkthrough | |
| | Getting started | Intro to Learning to Rank Model (Code) | Winning Solution Runner up Solution When to use Deep models vs ML techniques |
| Week 7 | | | |
| | | | |
| | | | |
| | | | |

Deliverables

8. Fashion product recommendation using LTR on H&M dataset.

Other Resources

Courses

- [Google Recommender Systems Course](#)

Resources

Recommendation

- [HomePage Recommendation using Exploration and Exploitation](#)

Papers

For Fashion Recommendation

- [A Review of Modern Fashion Recommender Systems](#)

Paper Notes

- [Vinija's Notes on Recommender Systems](#)

Blogs

- [RecSys](#)