

# 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	<a href="#">Book</a> Chapters 1, 2, 3	
	Text Preprocessing <ul style="list-style-type: none"><li>- Tokenization</li><li>- N-grams</li><li>- Stemming and Lemmatization</li><li>- Stop words, punctuation removal</li><li>- Encoding:<ul style="list-style-type: none"><li>- Bag of words</li><li>- TF-IDF</li></ul></li></ul>	<a href="#">Text Cleaning</a> <a href="#">Text Preprocessing</a> <a href="#">Text Encoding</a>	<a href="#">Recommendation Notes</a> (Download to view) <a href="#">Book</a> Pg. 43, 44, 45
	Embeddings <ul style="list-style-type: none"><li>- Self-trained</li><li>- W2V (Skip-gram &amp; CBOW)</li><li>- sBERT</li></ul>	<a href="#">E-commerce product embeddings</a> <a href="#">Illustrated Word2Vec</a> <a href="#">sBERT</a>	<a href="#">Cbow and skip gram</a>
	Association Rule Mining (ARM) based recommendation for Next item recommendation	<a href="#">ARM Kaggle</a> <a href="#">ARM Rec</a> (Only ARM part)	<a href="#">Resource Link</a>
	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	<a href="#">Book</a> Chapter 4	<a href="#">Content</a> <a href="#">Content &amp; collaborative</a> <a href="#">Recommender Systems 101</a>
	Similarity measures, dimensionality reduction,	<a href="#">Book</a> Chapter 5	
	Collaborative Filtering	<a href="#">Book</a> Chapter 6	
	Hybrid Recommendation	<a href="#">Book</a> Chapter 7	
	Cold Start, Long Tail	<a href="#">Cold Start Problem</a> <a href="#">Long Tail Problem</a>	
	Evaluation Metrics	<a href="#">Metrics</a>  <a href="#">Book</a> Pg. 92  <a href="#">Evaluation Metrics(metrics@k)</a>	
	Deployment using FastAPI	<a href="#">Blog</a>	
	Containerization using Docker	<a href="#">Videos</a> <a href="#">Blog</a>	

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

## 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	<a href="#">Course</a>	
	Pointwise, Pairwise, and Listwise	<a href="#">Blog</a>	
	Introduction to Kaggle Challenge	<a href="#">Challenge</a> <a href="#">Problem and Winning Solution Walkthrough</a>	
	Getting started	<a href="#">Intro to Learning to Rank Model (Code)</a>	<a href="#">Winning Solution</a> <a href="#">Runner up Solution</a> <a href="#">When to use Deep models vs ML techniques</a>
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](#)