



Recommendation Systems and Personalisation: Collaborative Filtering, Embeddings and Ranking



Recommendation Systems and Personalisation: Collaborative Filtering, Embeddings and Ranking

Introduction:

Recommendation systems and personalisation engines decide what a customer sees next in a shop, a streaming catalogue or a banking app, yet many organisations still ship popularity lists or opaque vendor widgets that nobody can evaluate, tune or explain. Sparse interaction logs, new users and new items make matters worse. This Core Concept course trains data scientists and ML engineers to build, measure and serve recommender engines in Python, from collaborative filtering to two-tower retrieval and ranking. Participants produce a Recommender Prototype Notebook and Offline Evaluation Report on a realistic interaction dataset.

Course Objectives:

- Frame a personalisation problem in retail, media or financial services as a retrieval and ranking task with a measurable business target
- Prepare explicit ratings and implicit interaction logs into sparse user-item matrices with leakage-free temporal splits
- Build neighbourhood, matrix factorisation, content-based and hybrid recommenders in Python and compare their behaviour
- Evaluate ranked lists offline with precision@k, recall@k, MAP and NDCG alongside coverage, novelty and diversity measures
- Design an online experiment and a low-latency serving path with embedding indexes, feature stores and caching
- Assess cold start, popularity bias, privacy and fairness risks and specify controls before release

Target Audience:

- Data science functions that model customer behaviour and own personalisation experiments
 - Machine learning engineering functions that train, deploy and monitor ranking models in production
 - Digital product analytics functions that define recommendation targets and read experiment results
 - Marketing analytics functions responsible for next-best-offer and cross-sell targeting logic
 - Data platform functions that maintain event pipelines, feature stores and vector indexes feeding recommenders
-

Course Outline:

Day 1: Recommender Use Cases and Interaction Data

- Recommendation Use Cases: Retail Basket, Media Streaming and Banking Next-Best-Offer
- Explicit Ratings Versus Implicit Signals: Clicks, Dwell Time and Purchases
- Sparse User-Item Matrix Construction with pandas and SciPy
- Popularity and Recency Baselines as Benchmarks
- Temporal Train-Test Splitting and Leakage Checks for Interaction Logs

Day 2: Collaborative Filtering, Content-Based and Hybrid Architectures

- User-Based and Item-Based k-Nearest Neighbour Filtering with Cosine Similarity
- Matrix Factorisation with Alternating Least Squares for Implicit Feedback
- Content-Based Filtering with TF-IDF and Item Attribute Vectors
- Hybrid Recommender Designs: Weighted, Switching and Feature-Combination
- Candidate Generation, Scoring and Re-Ranking Pipeline Architecture

Day 3: Embeddings, Deep Recommenders and Ranking Evaluation

- Item and User Embeddings Learned from Interaction Sequences
- Two-Tower Retrieval Models and Deep Ranking Networks at Overview
- Learning-to-Rank Objectives: Pointwise, Pairwise and Listwise Losses
- Offline Ranking Metrics: Precision@k, Recall@k, MAP and NDCG
- Beyond-Accuracy Measures: Catalogue Coverage, Novelty and Serendipity

Day 4: Cold Start, Bias, Online Testing and Real-Time Serving

- Cold Start Remedies: Attribute Priors, Onboarding Questions and Multi-Armed Bandits
- Popularity Bias, Filter Bubbles and Diversity Re-Ranking with Maximal Marginal Relevance
- Online A/B Test Design: Guardrail Metrics, Sample Size and Novelty Effects
- Real-Time Serving: Feature Store, Approximate Nearest Neighbour Index and Response Caching
- Privacy-Preserving Personalisation and Fairness Audits for Provider and User Groups

Day 5: Recommender Build Lab and Evaluation Report

- E-Commerce Lab: Implicit-Feedback ALS Model with Hybrid Content Fallback
 - Media Catalogue Lab: Embedding Retrieval with an Approximate Nearest Neighbour Index
 - Model Comparison Leaderboard on NDCG, Coverage and Diversity
 - Experiment Plan and Serving Design Sheet for Production Rollout
 - Recommender Prototype Notebook and Offline Evaluation Report Peer Review
-

Skills You Will Gain:

- Recommender Problem Framing
- Interaction Data Engineering
- Collaborative Filtering Modelling
- Embedding-Based Retrieval
- Ranking Evaluation
- Online Experiment Design
- Low-Latency Model Serving
- Responsible Personalisation

Why Attend This Course:

- Return with a Recommender Prototype Notebook and Offline Evaluation Report built and benchmarked during the labs
- Know which recommender family suits sparse logs, rich item metadata or new catalogues before committing engineering effort
- Spot misleading offline results caused by leakage, popularity bias or the wrong metric before they reach an A/B test
- Compare recommender designs with practitioners from retail, media, telecom and banking teams

Conclusion:

A recommender earns trust when its data, model and evaluation can be inspected and repeated. The course moves from use cases and interaction data, through neighbourhood methods, matrix factorisation, content-based and hybrid designs, to embeddings, two-tower retrieval and ranking metrics, then to cold start, bias, online testing, serving and privacy. The final day runs hands-on build labs and produces a Recommender Prototype Notebook and Offline Evaluation Report that teams can extend into production.