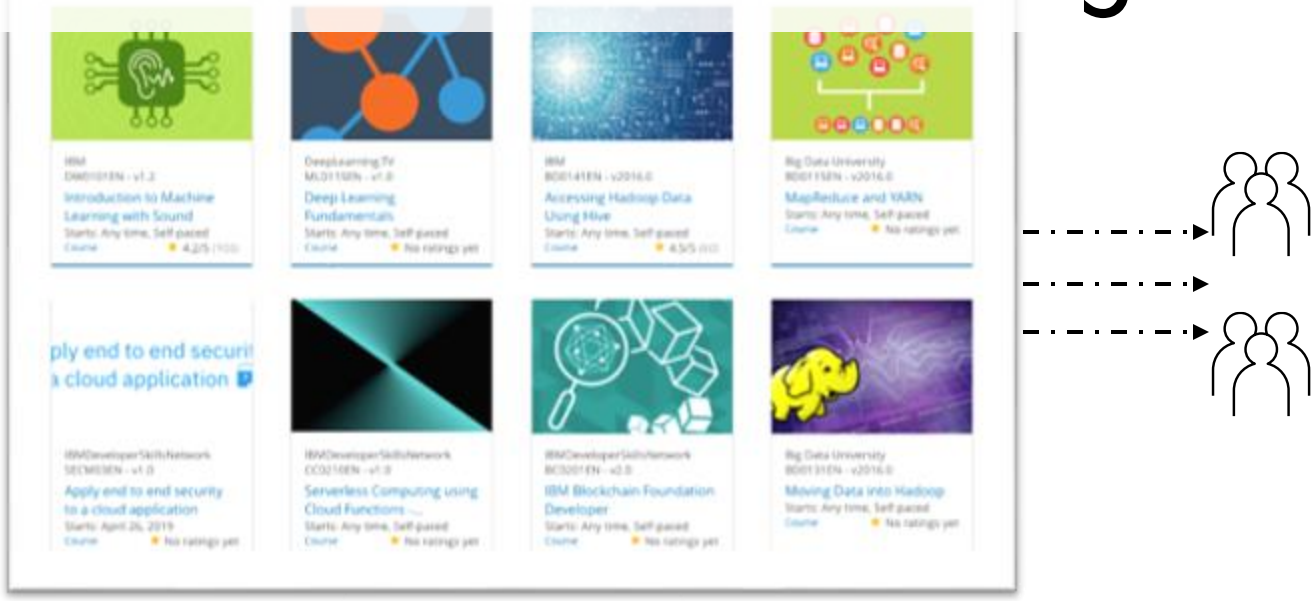
Build a Personalized Online Course Recommender System with Machine Learning

Mohamed MZAOUALI

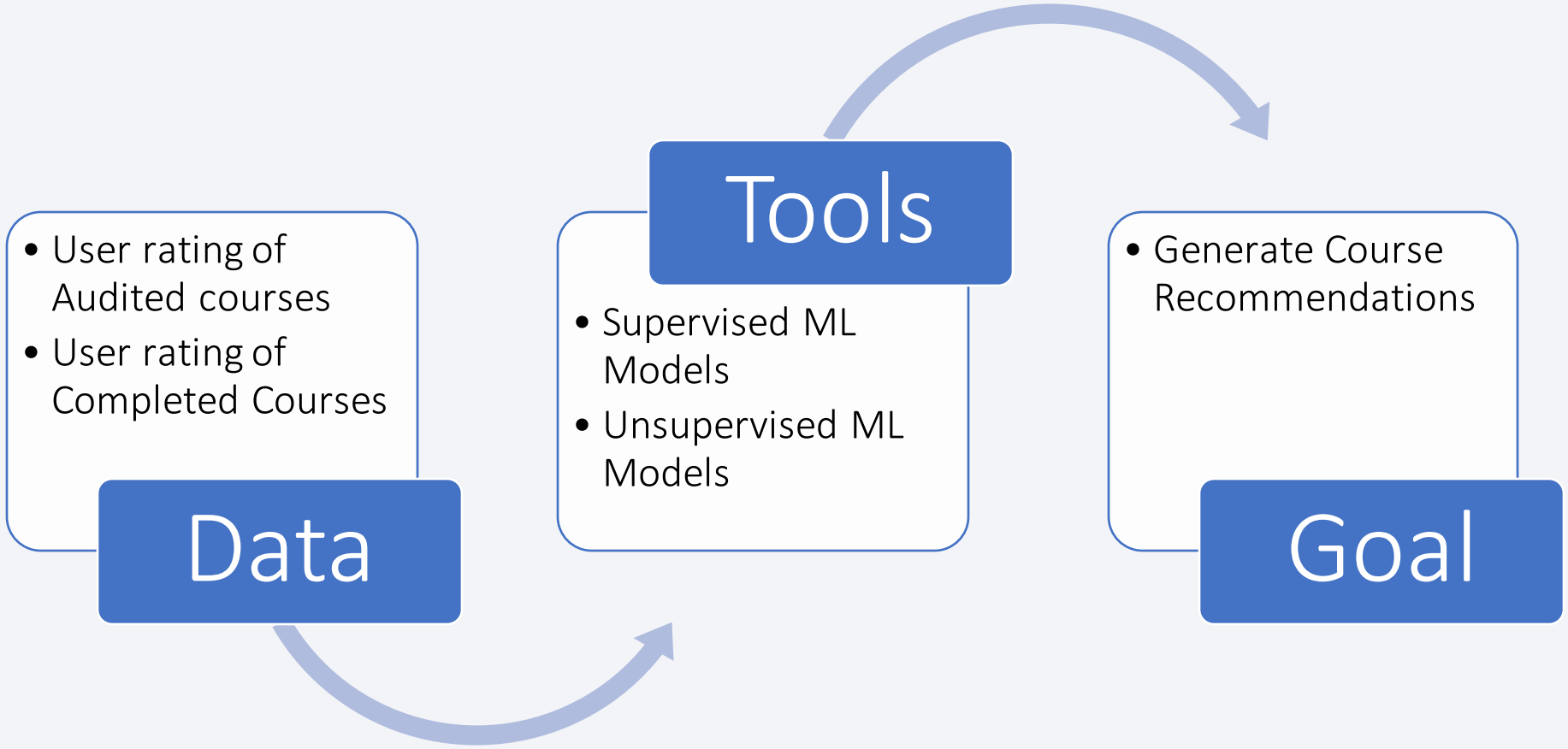
August the 21th, 2022



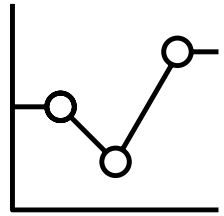
Outline

- Introduction and Background
- Exploratory Data Analysis
- Content-based Recommender System using Unsupervised Learning
- Collaborative-filtering based Recommender System using Supervised learning
- Conclusion
- Appendix

Introduction

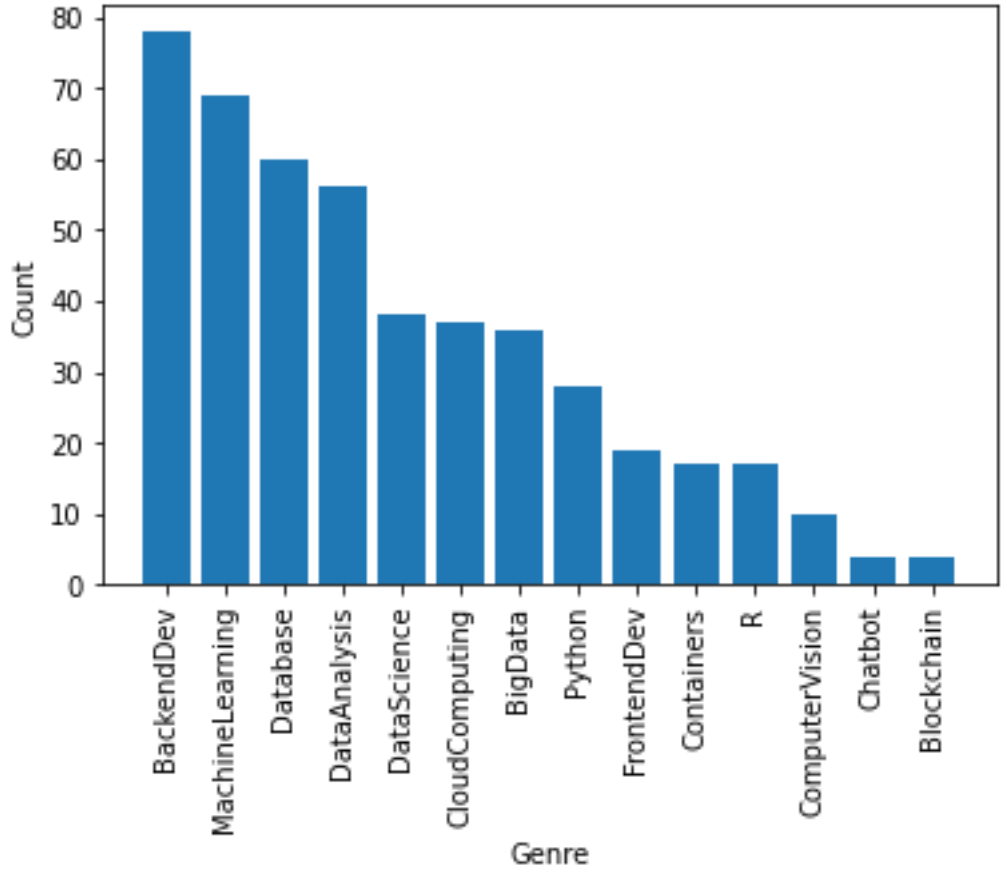


Exploratory Data Analysis



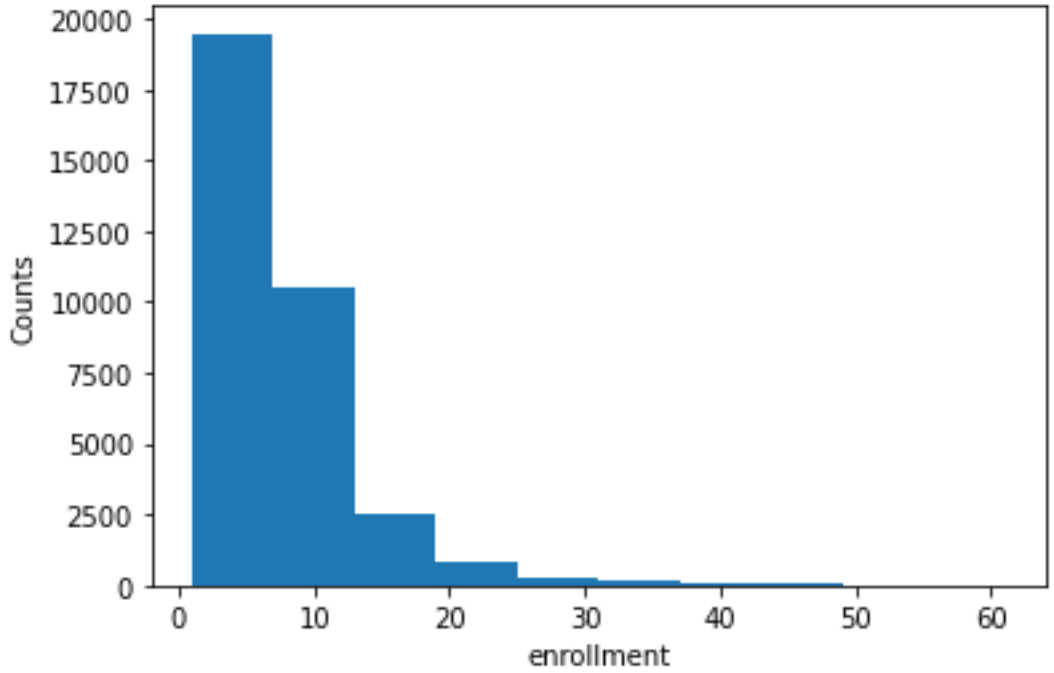
Course counts per genre

Using seaborn barplot method to plot course genre counts using a barchart.



Course enrollment distribution

Using the Matplotlib hist methods to get a histogram showing the enrollment distributions



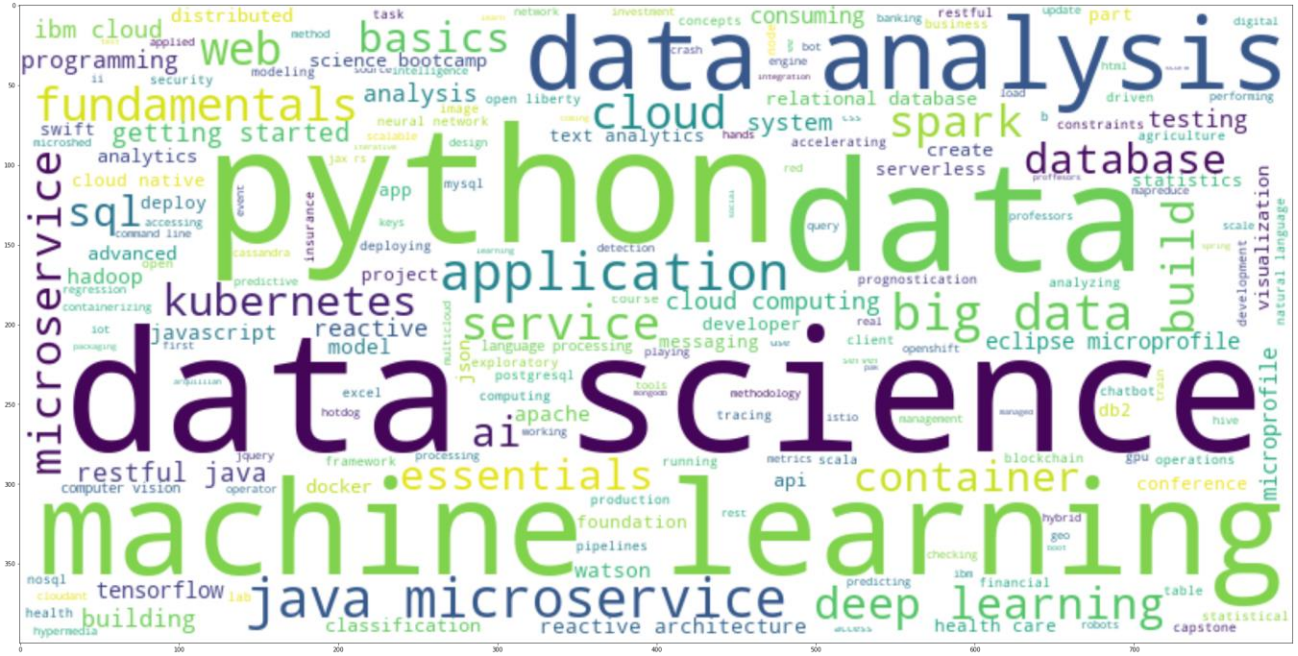
20 most popular courses

The 20 items with the most rating counts

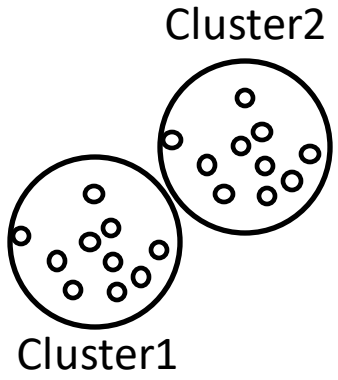
TITLE	ENROLLS
python for data science	14936
introduction to data science	14477
big data 101	13291
hadoop 101	10599
data analysis with python	8303
data science methodology	7719
machine learning with python	7644
spark fundamentals i	7551
data science hands on with open source tools	7199
blockchain essentials	6719
data visualization with python	6709
deep learning 101	6323
build your own chatbot	5512
r for data science	5237
statistics 101	5015
introduction to cloud	4983
docker essentials a developer introduction	4480
sql and relational databases 101	3697
mapreduce and yarn	3670
data privacy fundamentals	3624

Word cloud of course titles

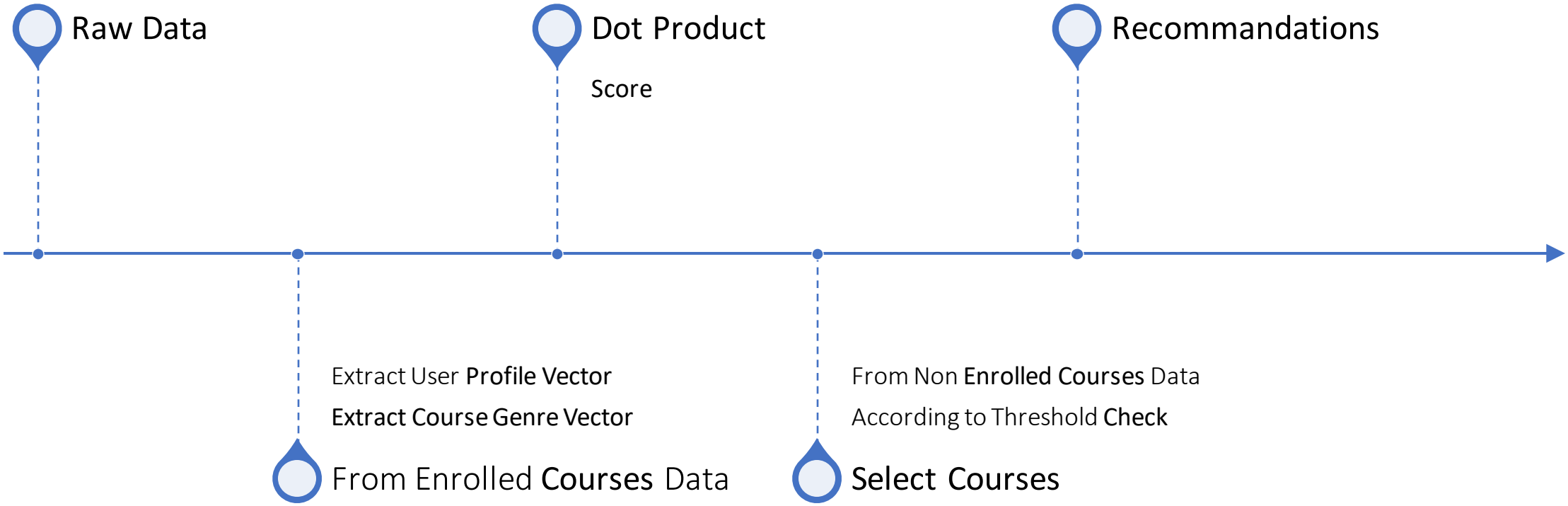
Popular IT keywords from the course titles



Content-based Recommender System using Unsupervised Learning



Flowchart of content-based recommender system using user profile and course genres



Evaluation results of user profile-based recommender system

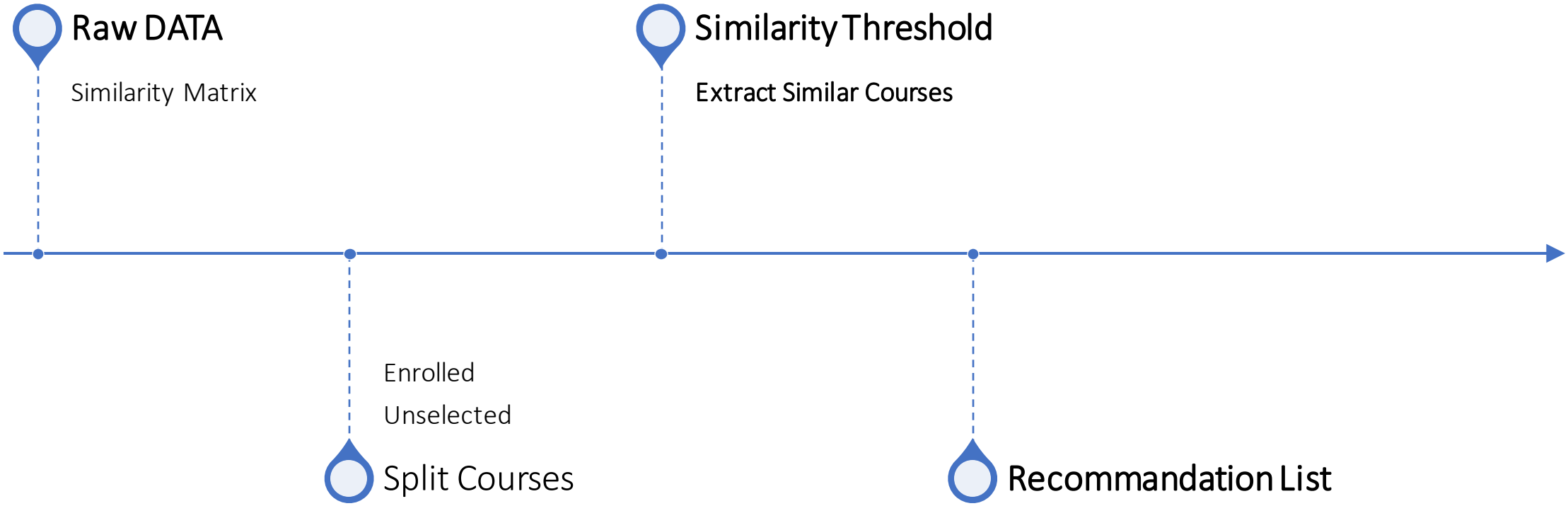
score_threshold = 10.0

On average, 61 new/unseen courses have been recommended per user (in the test user dataset)

- The top-10 commonly recommended courses across all users are:

1. TA0106EN
2. GPXX0IBEN
3. excourse22
4. excourse21
5. ML0122EN
6. GPXX0TY1EN
7. excourse04
8. excourse06
9. excourse31
10. excourse73

Flowchart of content-based recommender system using course similarity



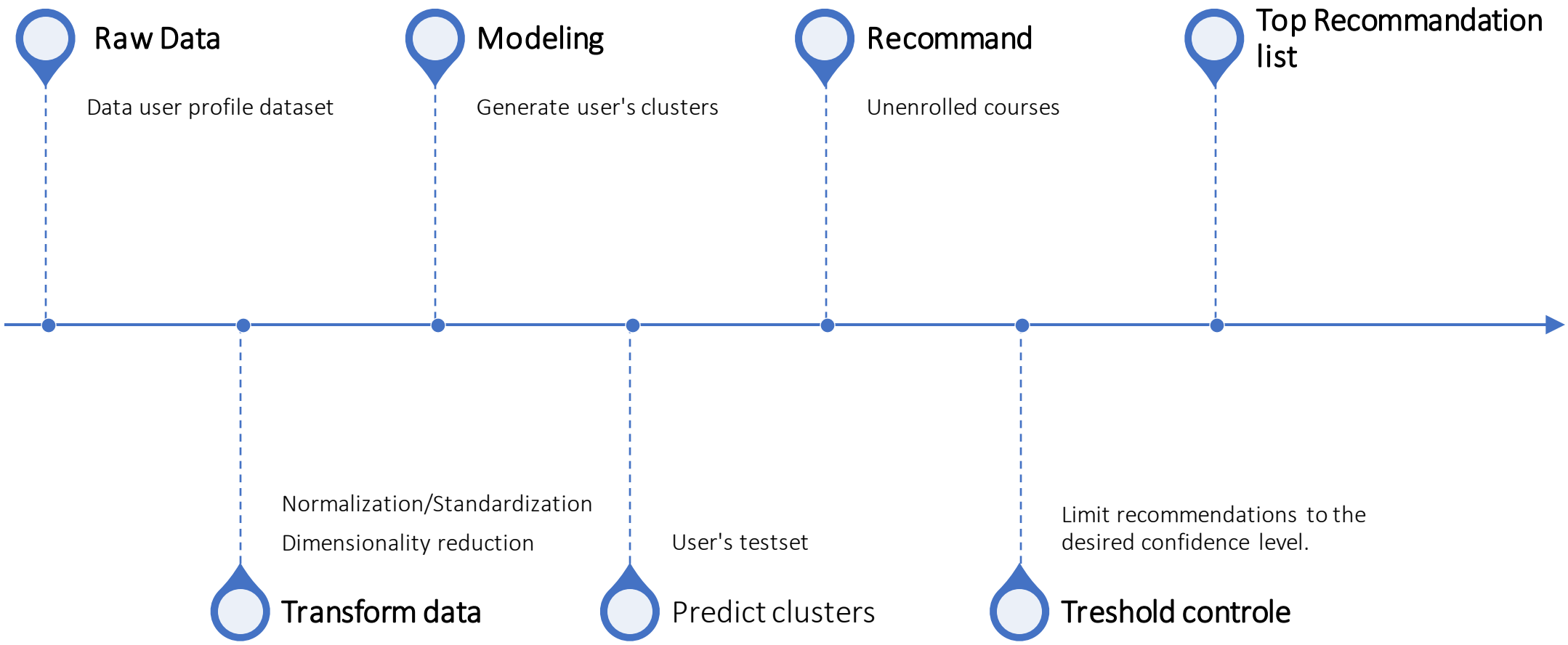
Evaluation results of course similarity based recommender system

score_threshold = 0.6

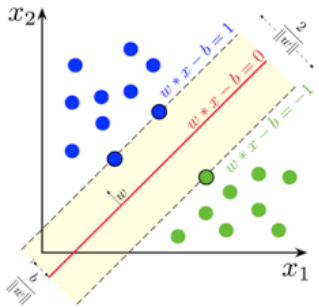
On average, 12 new/unseen courses have been recommended per user (in the test user dataset)

1)'BD0101EN'
2)'DS0101EN'
3)'DS0110EN'
4)'excourse04'
5)'excourse23'
6)'excourse32'
7)'excourse33'
8)'excourse36'
9)'excourse63'
10)'excourse67'
11)'excourse68'
12)'excourse72'

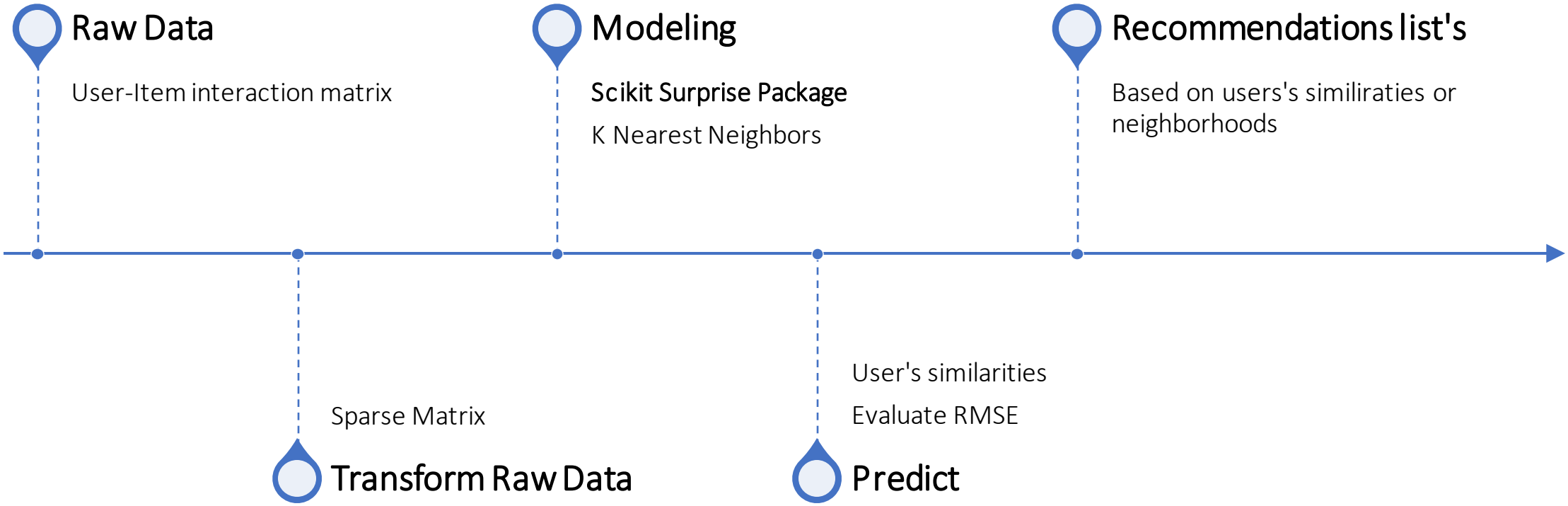
Flowchart of clustering-based recommender system



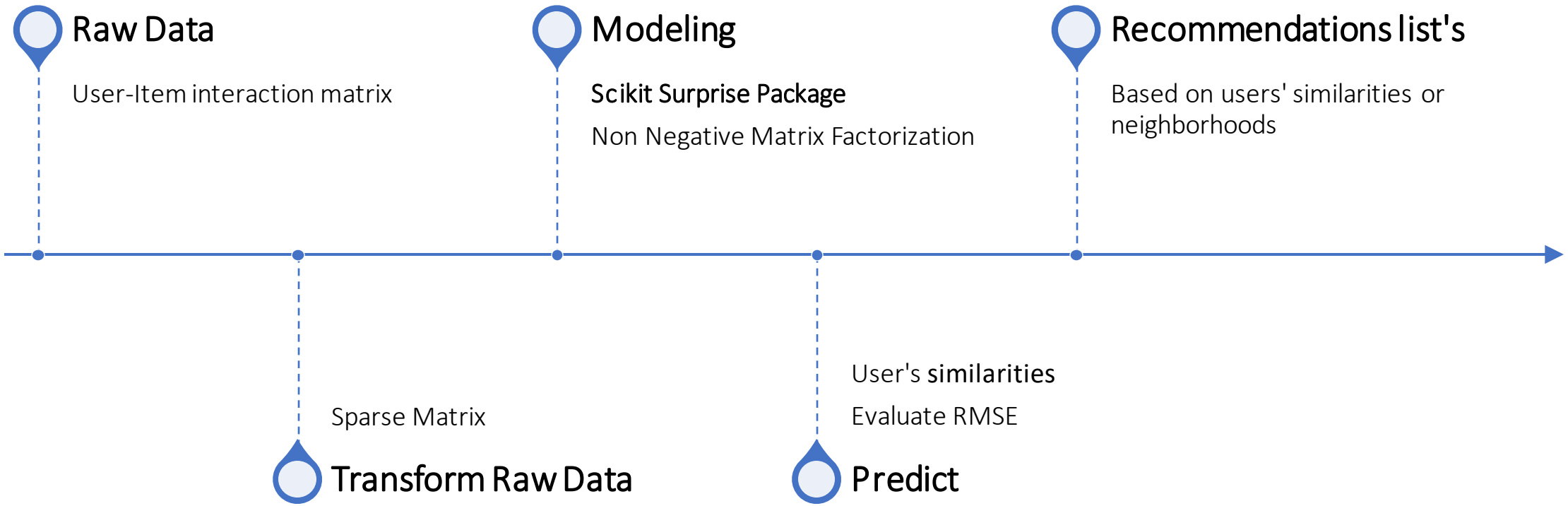
Collaborative-filtering Recommender System using Supervised Learning



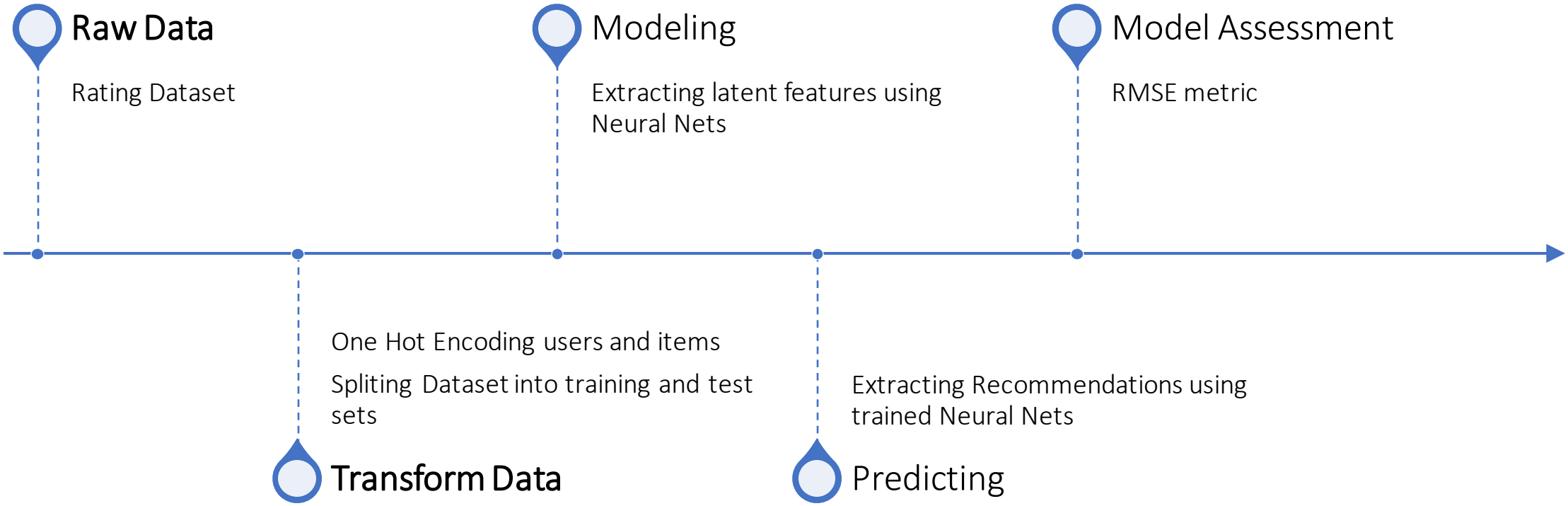
Flowchart of KNN based recommender system



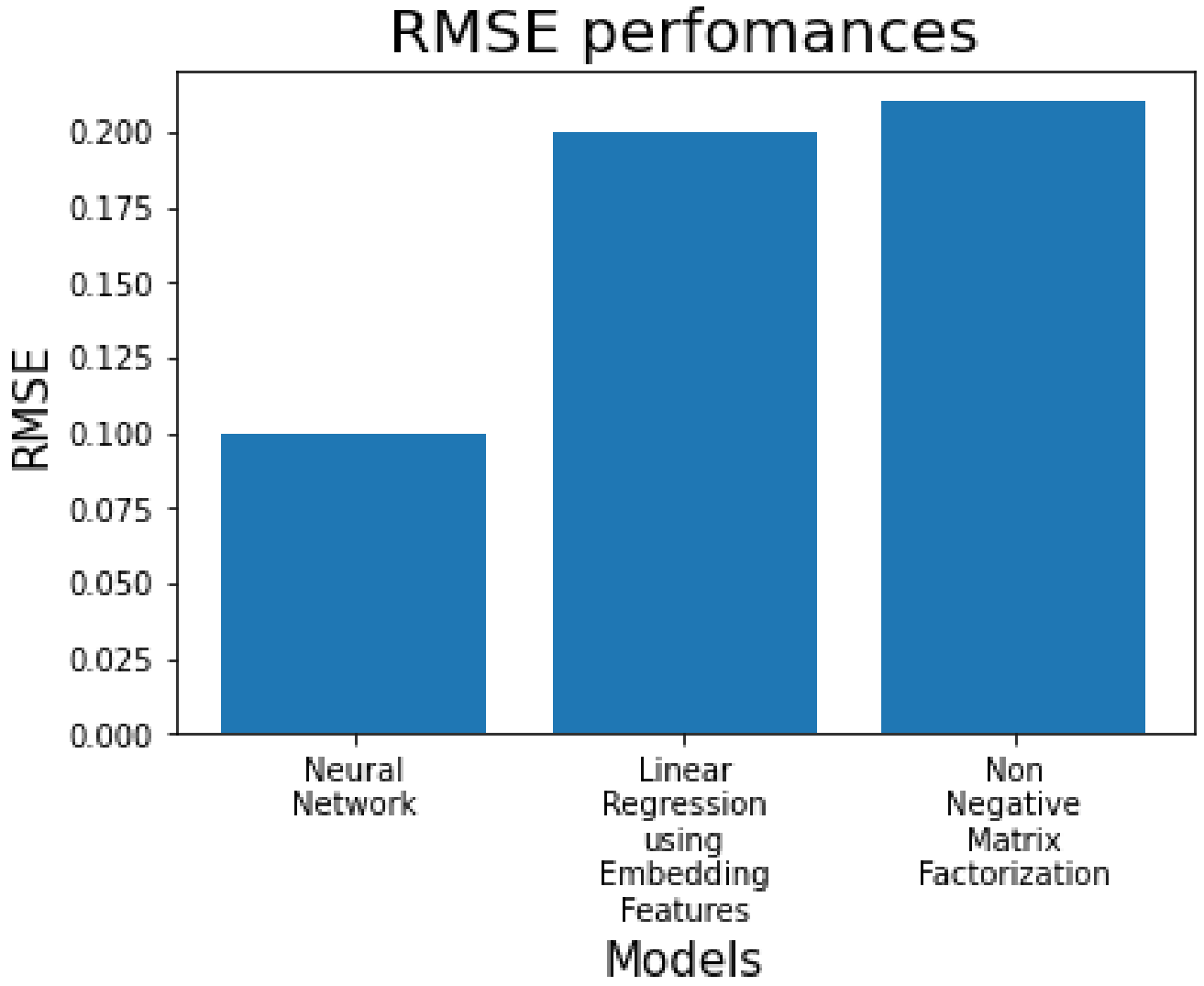
Flowchart of NMF based recommender system



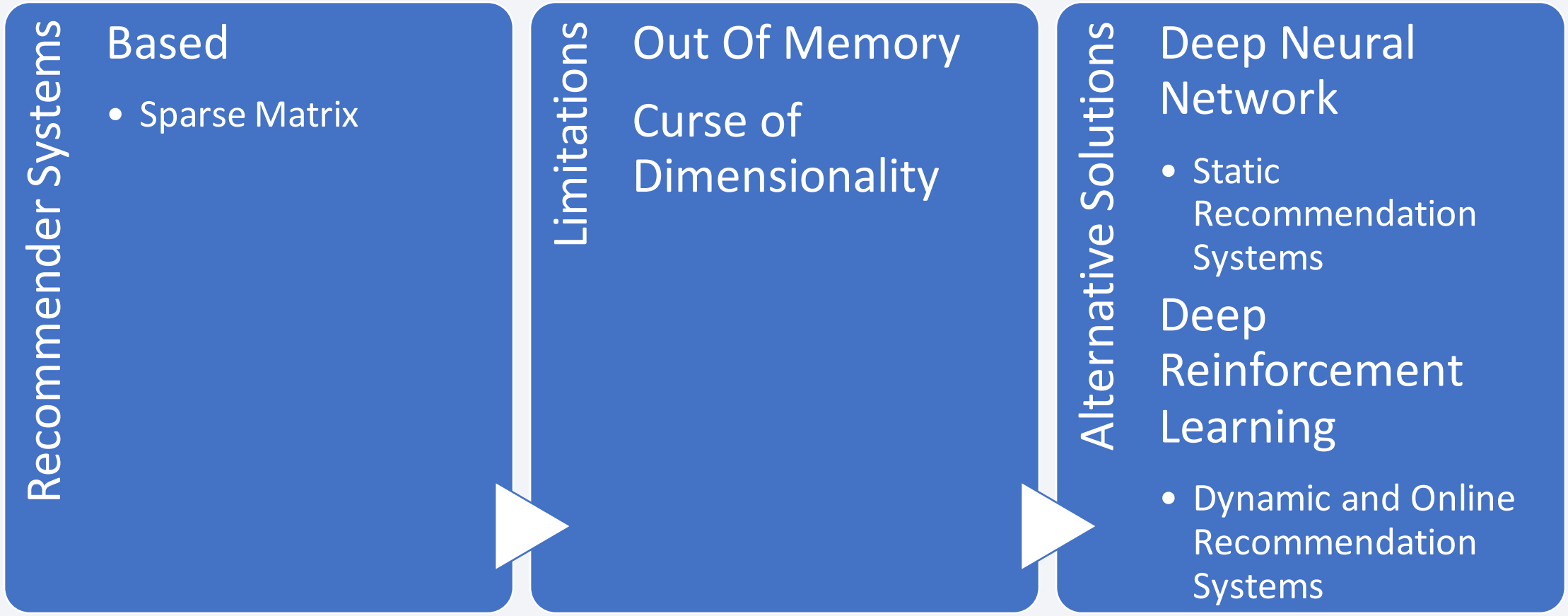
Flowchart of Neural Network Embedding based recommender system



Compare the performance of collaborative-filtering models



Conclusions



Appendix

- The entire project capstone can be fount in the followin Github repo:
- https://github.com/mzaoualim/Coursera_IBM_Machine_Learning_Professional_Certificate/tree/main/Machine%20Learning%20Capstone