Recommender Systems || Lecture 13

Summer 2022

Dr. Karthik Mohan

UW, Seattle

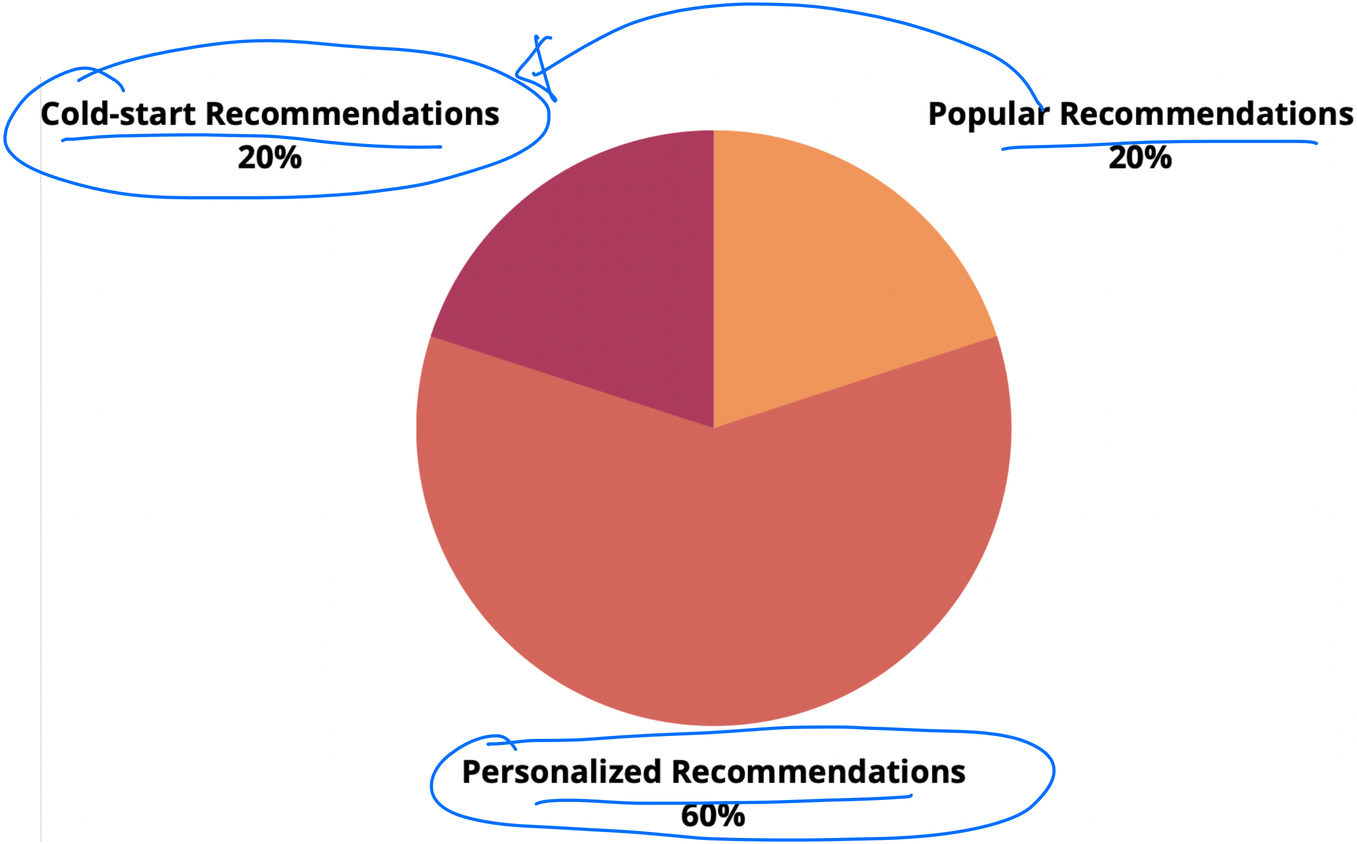
August 17, 2022

Today

1. Recap and Review of Recommender Systems

Recommender Systems are Ubiquitous

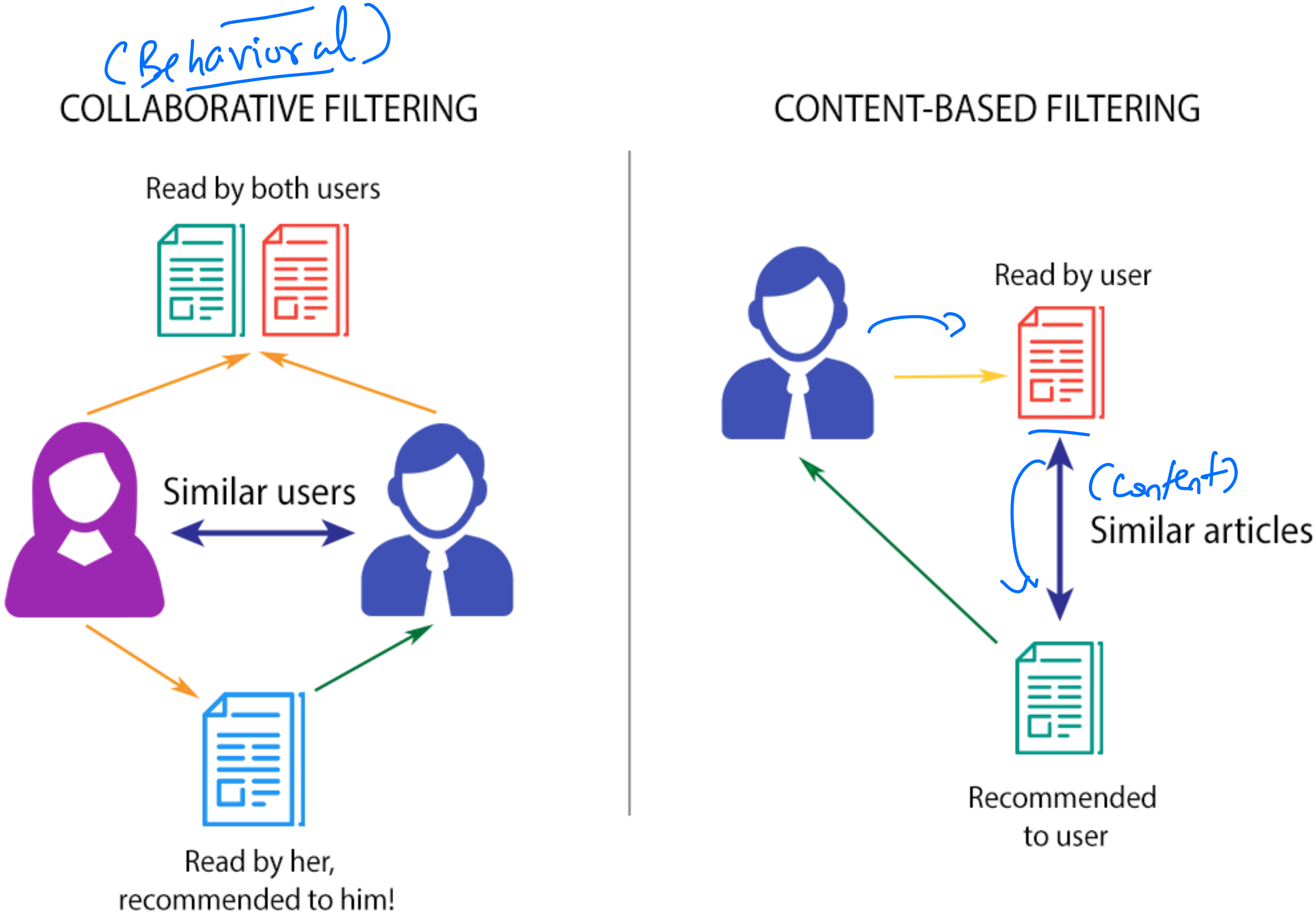
Recommender Types



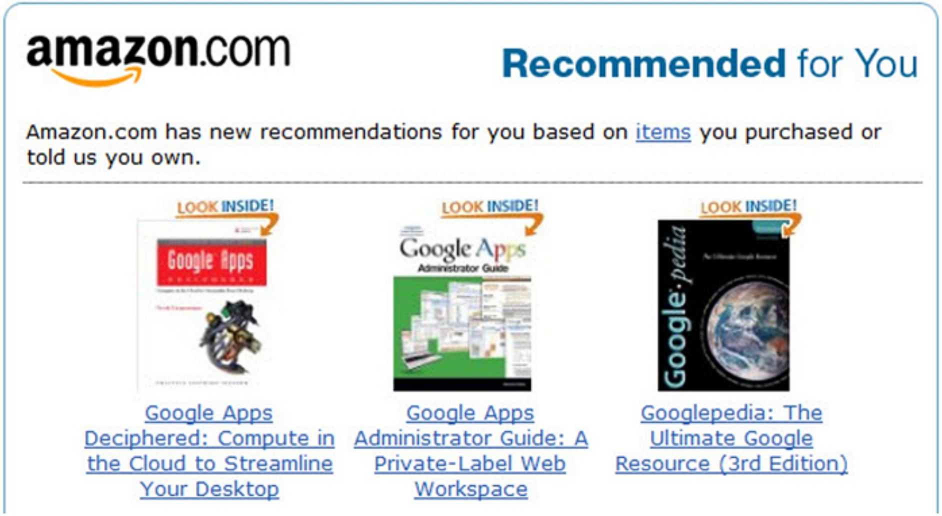
Recommender Types

1. Batch Recommender (offline) → Compute Recs before a Customer shows up
2. Online/Sessions Recommender → Recs Can change during session

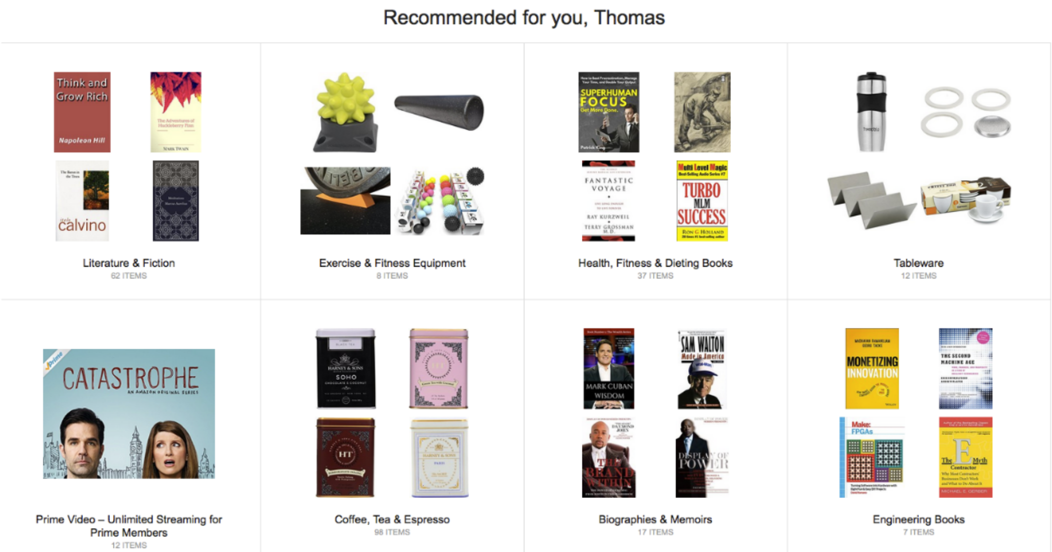
Collaborative filtering vs Content Based Filtering



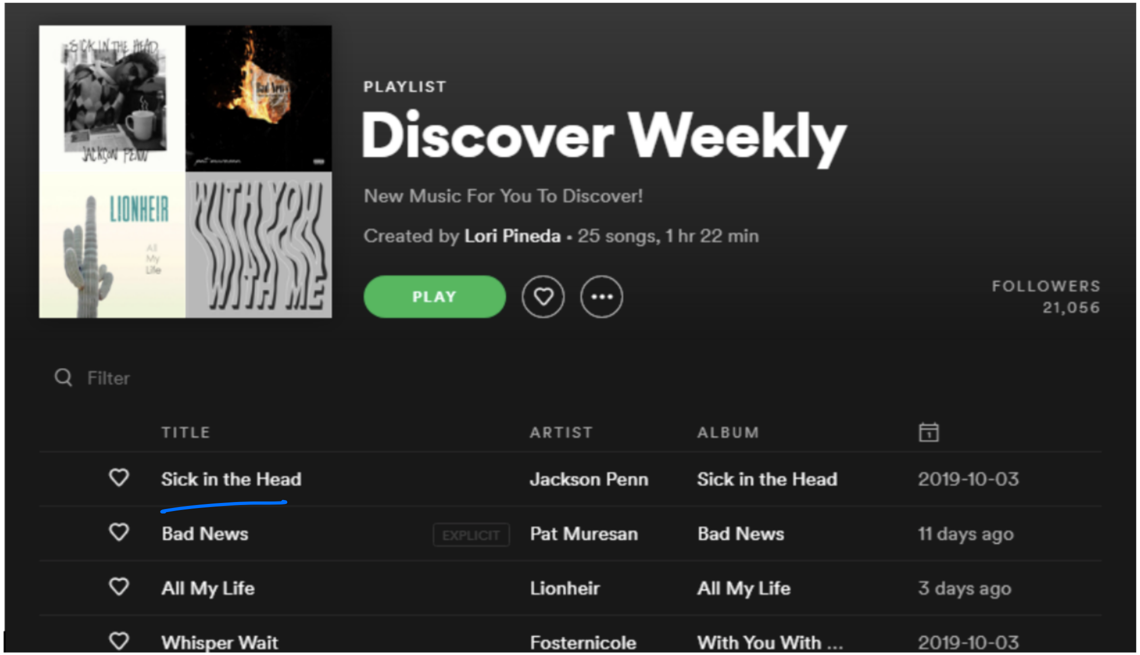
Item based recommendations



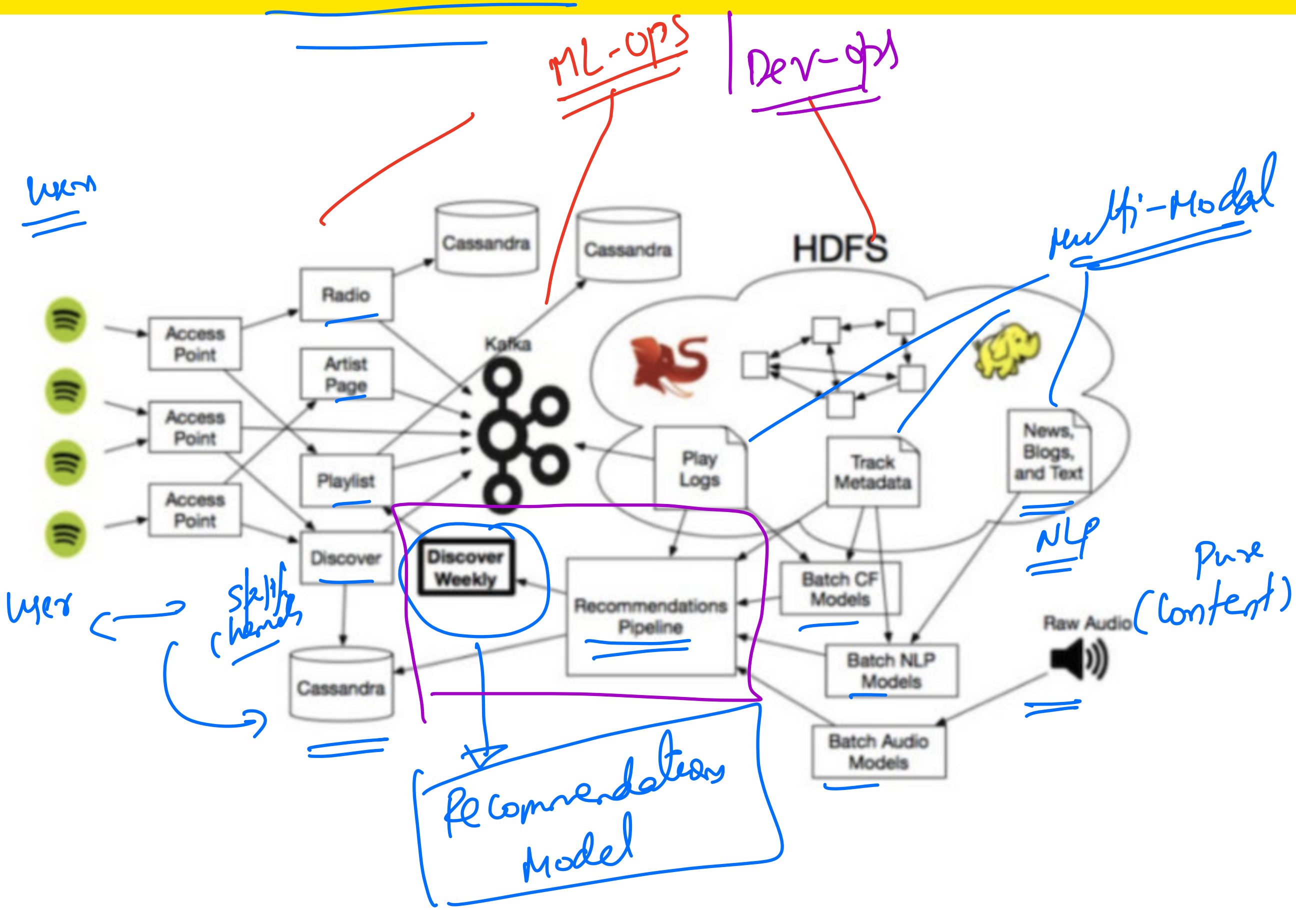
Amazon Recommendations



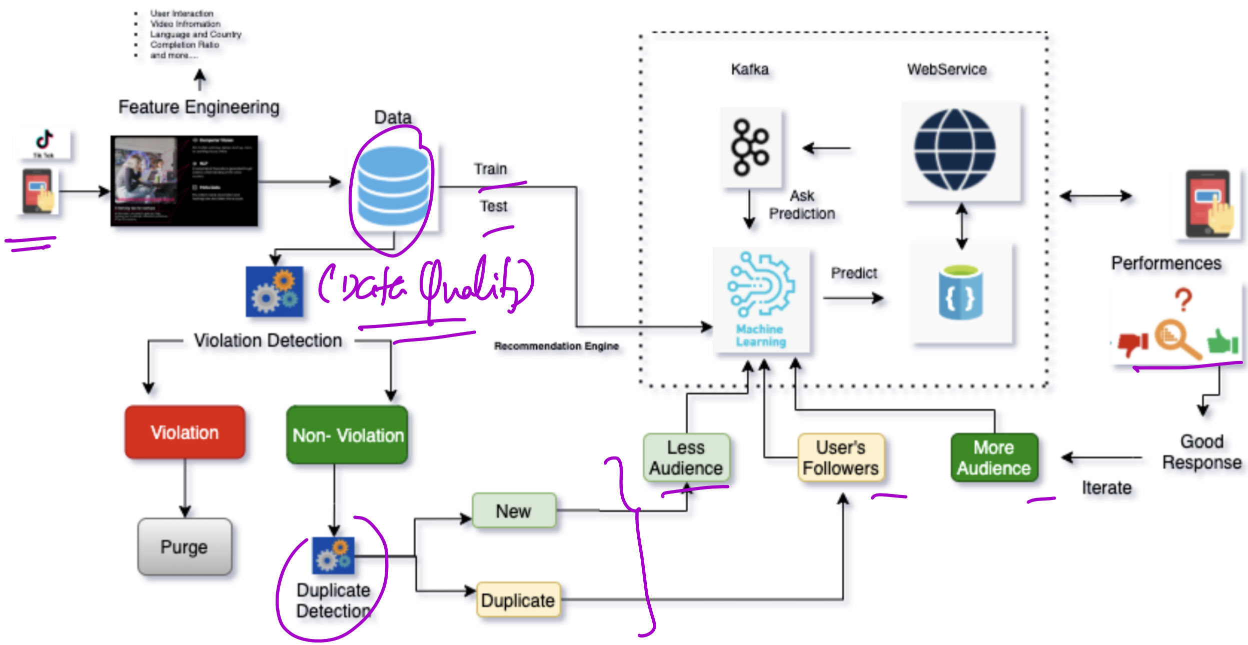
Spotify Recommendations



Spotify Pipeline System Design



TikTok Recommendations



Recommender Modeling Pipeline Considerations

1. Data Quality ✓ (Log Data → Filter → Denoise → Extract features, etc → High Data quality
2. Freshness of Recommendations
3. Relevance
4. Diversity
5. Personalized vs Popular
6. Offline vs Online Evaluation
7. Recommendation Metrics vs Business Metrics

Recommender Modeling Pipeline Considerations — Data Quality

video-user interaction data (behavioral),

1. Often Noisy
2. Multiple Data streams - E.g. video meta data, video content, video text description, etc for video recommendations
3. Click Data vs Engagement Data

useful not useful Bots

Recommender Modeling Pipeline Considerations — Freshness

1. Fashion
2. Music (Trending)
3. News

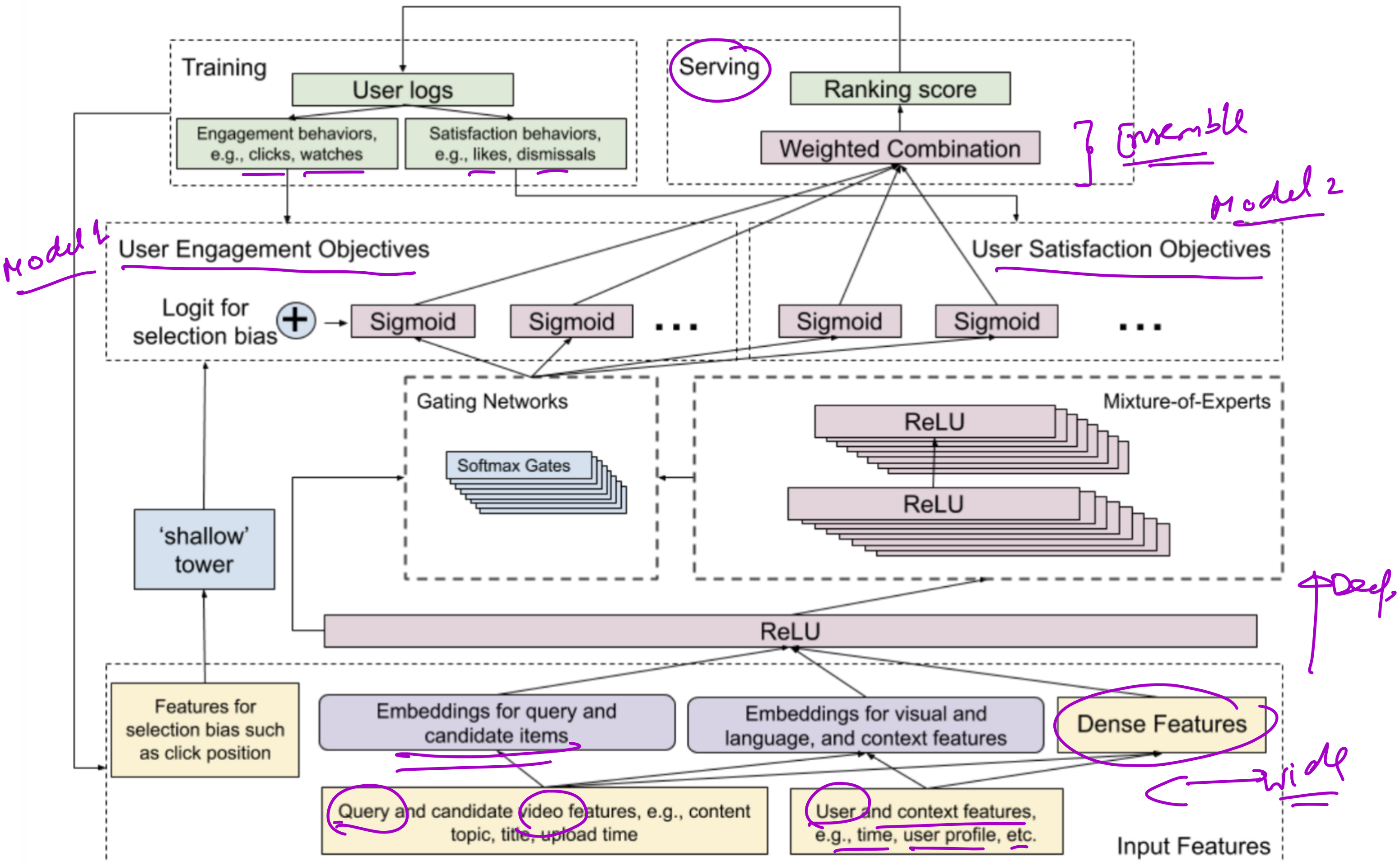
Recommender Modeling Pipeline Considerations — Relevance

1. Bread and butter of recommendations modeling
2. Relevance to the point of "The app understands me"
3. Relevance models for both cold start and regular customers
4. Deep Relevance models - Using deep learnning for Relevance modeling

Recommender Modeling Pipeline Considerations — Deep Relevance

1. AutoRec
2. DeepRec
3. Multi-Layer perceptron
4. Session based recommendations: Uses LSTM/Sequential models
5. YouTube Recommendations is a good example of multi-modal, multi-objective, ensemble based deep and wide recommendations model

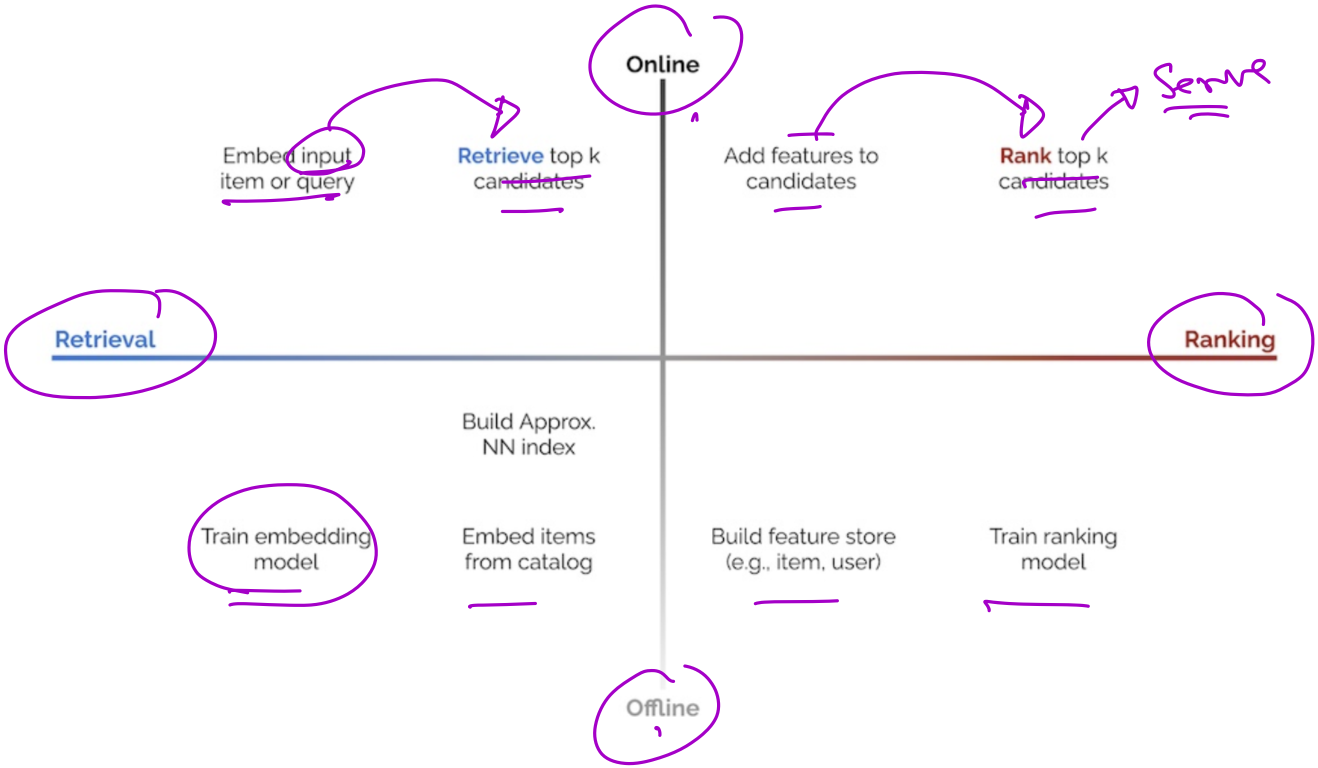
YouTube Recommender Design



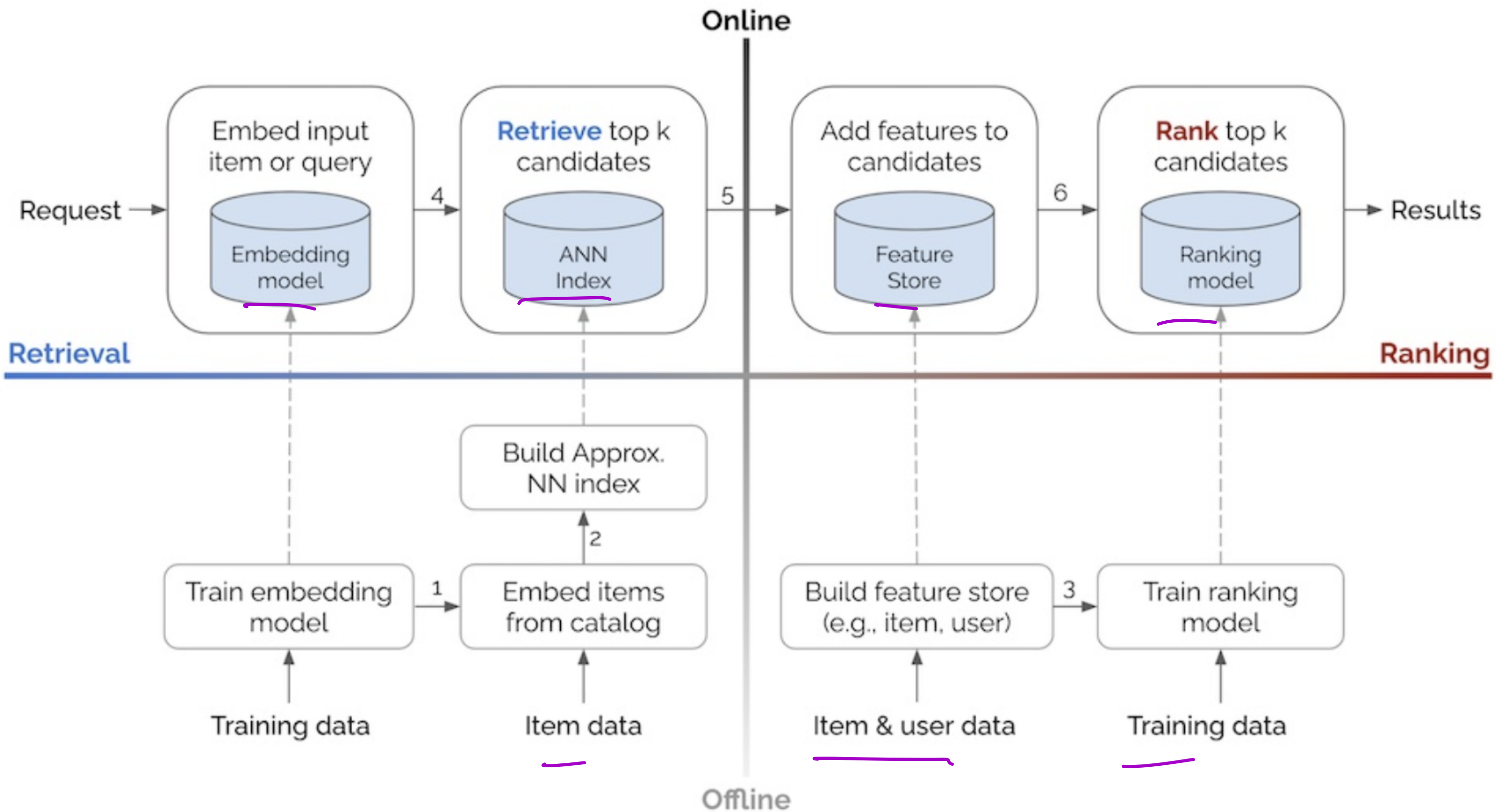
Recommender Modeling Pipeline Considerations

1. Data Quality
2. Freshness of Recommendations
3. Relevance
4. Diversity
5. Personalized vs Popular
6. Offline vs Online Evaluation
7. Recommendation Metrics vs Business Metrics

Recommender System Design High-level



Recommender System Design High-level



Recommender System Level Considerations

1. Scalability for Model Training and Model Serving

1MM | 100MM+

Recommender System Level Considerations

1. Scalability for Model Training and Model Serving
2. Latency for Model Serving

Recommender System Level Considerations

1. Scalability for Model Training and Model Serving
2. Latency for Model Serving
3. Data integration after data quality checks

(Continuous)

Recommender System Level Considerations

1. Scalability for Model Training and Model Serving
2. Latency for Model Serving
3. Data integration after data quality checks
4. More on this in ML Ops Guest Lecture tomorrow!

Metrics for Recommender Systems

Recommendation vs Ranking Metrics

1. Offline Recommendation metrics help us understand how well offline recommendations are on test set

Recommendation vs Ranking Metrics

1. Offline Recommendation metrics help us understand how well offline recommendations are on test set
2. Ranking metrics in addition help us understand how well the offline recommendations are ordered/ranked

Recommendation vs Ranking Metrics

1. Offline Recommendation metrics help us understand how well offline recommendations are on test set
2. Ranking metrics in addition help us understand how well the offline recommendations are ordered/ranked
3. In practice - Combination of Recommendation and Ranking metrics used

Summary of Offline Metrics for Recommender Systems

Rec metrics

Metric	Description
Precision	$\frac{TP}{TP+FP}$
Recall	$\frac{TP}{TP+FN}$
Precision @ k	$\frac{TP@k}{k}$
Recall @ k	$\frac{TP@k}{N}$
F-score	HM of Precison, Recall
Average Precision @ k	Average of 'active' precisions in k
DCG @ k	Discounted Cumulative Gain
NDCG @ k	Normalized DCG

Ranking metrics

1) Recommendation Bias:-

2) Click Bias:-

Online Evaluation

1. Internal Testing
2. A/B Testing - Control vs Treatment
3. A/B Testing Example - Existing Page for Groceries (Control) vs Existing Page + Your Recommendations widget (Treatment)
4. Bandit Testing (Recent development) - Testing multiple options simultaneously instead of just 2

Online Metrics

1. Usually business metrics
2. Number of purchases
3. Increase in customers?
4. Revenue/Sales?

Offline-Online Gap

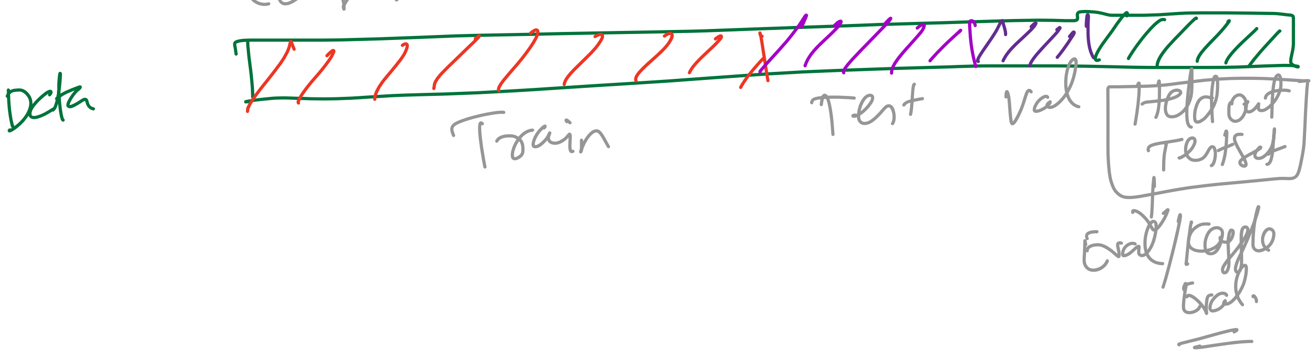
1. There is always some gap between offline and online metrics

Offline-Online Gap

1. There is always some gap between offline and online metrics
2. How to correct for this?

1. There is always some gap between offline and online metrics
2. How to correct for this?
3. Held-out testing } Coverage

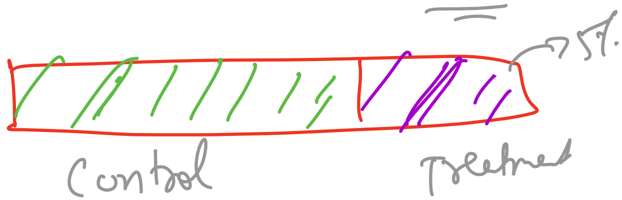
(Large)



Offline-Online Gap

1. There is always some gap between offline and online metrics
2. How to correct for this?
3. Held-out testing
4. Internal testing

1. There is always some gap between offline and online metrics
2. How to correct for this?
3. Held-out testing
4. Internal testing
5. 5% A/B test - Can do fixes if something is off



Offline-Online Gap

1. There is always some gap between offline and online metrics
2. How to correct for this?
3. Held-out testing
4. Internal testing
5. 5% A/B test - Can do fixes if something is off
6. Gradually increase the percentage as confidence builds

Building your Recommender System!

1. **Data sources?**

Labels

Targets

How rich init??

multi-modal??

Building your Recommender System!

1. **Data sources?**
2. **Data quality checks**

Building your Recommender System!

1. **Data sources?**
2. **Data quality checks**
3. **Offline vs online recommendations**] Rec Design

Building your Recommender System!

1. **Data sources?**
2. **Data quality checks**
3. **Offline vs online recommendations**
4. **Collaborative, content or hybrid recs?**] modeling

Building your Recommender System!

1. **Data sources?**
2. **Data quality checks**
3. **Offline vs online recommendations**
4. **Collaborative, content or hybrid recs?**
5. **Baseline and gaps in baseline?**

Building your Recommender System!

1. **Data sources?**
2. **Data quality checks**
3. **Offline vs online recommendations**
4. **Collaborative, content or hybrid recs?**
5. **Baseline and gaps in baseline?**
6. **Evaluating your system right**

Iterate

Building your Recommender System!

1. **Data sources?**
2. **Data quality checks**
3. **Offline vs online recommendations**
4. **Collaborative, content or hybrid recs?**
5. **Baseline and gaps in baseline?**
6. **Evaluating your system right**
7. **Use cases where your model needs improvement?**

Building your Recommender System!

1. **Data sources?**
2. **Data quality checks**
3. **Offline vs online recommendations**] System Design
4. **Collaborative, content or hybrid recs?**
5. **Baseline and gaps in baseline?**
6. **Evaluating your system right**
7. **Use cases where your model needs improvement?**
8. **Scalability and latency considerations**

Questions/Thoughts??