PERFORMANCE FACTORS

1. User Interface (UI) x User Experience (UX)

2. Data: accuracy, coverage, clean

3. Domain Knowledge: requirement understanding ➡ purchase/visit

4. Algorithm/Model Iteration: A/B test on hyper-parameter with online feedback

SSENSE

SUMMARY OF ALGORITHM

‣ Popularity: not just hot

‣ Collaborative Filtering

  - Memory Based / Neighbourhood Based: User-based x Item-based (kNN)

  - Model Based:

    ◉ Association Rule: Apriori, **FP-Growth***

    ◉ **SVD/SVD++***, Factorization Machine (**ALS***, SGD, MCMC)[libFM]: Matrix Decomposition

    ◉ Clustering*: User-based x Item-based

◉ Classification*/Regression Model*: purchase/visit as Y

◉ LSA / pLSA / LDA: vectorize items

◉ Restricted Boltzmann Machine (RBM): 2-layer neural network

◉ Graph-based algorithm

‣ Content-based / Knowledge-based

  - Item Attribut Vector x User Profile Vector

  - Distance Measurement `Cosine`

‣ Context-aware: when, where, mood

*: implemented in Apache Spark