Fundamentals of Deep Recommender Systems

Wenqi Fan
The Hong Kong Polytechnic University
https://wenqifan03.github.io, wenqifan@polyu.edu.hk

Tutorial website: https://deeprs-tutorial.github.io
A General Architecture of Deep Recommender System

**Embedding layer**

**Hidden layer** (e.g., MLP, CNN, RNN, etc.)

**Prediction layer**

Field 1, Field m, Field M

User, Item, Context, Interaction
NeuMF unifies the strengths of MF and MLP in modeling user-item interactions.

- **MF** uses an inner product as the interaction function
- **MLP** is more sufficient to capture the complex structure of user interaction data
The wide linear models can memorize seen feature interactions using cross-product feature transformations.

The deep models can generalize to previously unseen feature interactions through low-dimensional embeddings.
Neural Factorization Machines (NFM)s “deepens” FM by placing hidden layers above second-order feature interaction modeling.
Neural FM

Neural Factorization Machines (NFMs) “deepens” FM by placing hidden layers above second-order feature interaction modeling.

“Deep layers” learn higher-order feature interactions only, being much easier to train.

Bilinear Interaction Pooling:

\[ f_{BI}(V_x) = \sum_{i=1}^{n} \sum_{j=i+1}^{n} x_i v_i \odot x_j v_j \]
DeepFM ensembles FM and DNN and to low- and high-order feature interactions simultaneously from the input raw features.

**Prediction Model:**

\[
\hat{y} = \text{sigmoid}(y_{FM} + y_{DNN})
\]
DeepFM

DeepFM ensembles FM and DNN and to low- and high-order feature interactions simultaneously from the input raw features.

**Prediction Model:**

\[
\hat{y} = \text{sigmoid}(y_{FM} + y_{DNN})
\]
Collaborative Filtering with users’ social relations
(Social Recommendation)

US could see millions of coronavirus cases and 100,000 or more deaths
Collaborative Filtering with users’ social relations (Social Recommendation)

Users might be affected by direct/distant neighbors.
- Information diffusion
- Users with high reputations

Dr. Anthony Fauci

US could see millions of coronavirus cases and 100,000 or more deaths

Deep Social Collaborative Filtering, RecSys, 2019
Collaborative Filtering with users’ social relations
(Social Recommendation)

Users might be affected by direct/distant neighbors.
- Information diffusion
- Users with high reputations

Social Sequences via Random Walk techniques

Bi-LSTM with attention mechanisms

Deep Social Collaborative Filtering, RecSys, 2019
User behave and interact **differently** in the item/social domains.
Collaborative Filtering with users’ social relations
(Social Recommendation)

- User behave and interact differently in the item/social domains.

Learning separated user representations in two domains.
Collaborative Filtering with users’ social relations
(Social Recommendation)

- User behave and interact differently in the item/social domains.
- Learning separated user representations in two domains.

Bidirectional Knowledge Transfer with Cycle Reconstruction

\[
P_i^I \rightarrow h^I \rightarrow S (P_i^I) \rightarrow h^S \rightarrow I (h^I \rightarrow S (P_i^I)) \simeq P_i^I
\]

\[
\mathcal{L}_{cyc}(h^S \rightarrow I, h^I \rightarrow S) = \sum_{i=1}^{N} (\|h^S \rightarrow I (h^I \rightarrow S (P_i^I)) - P_i^I\|_2 + \|h^I \rightarrow S (h^S \rightarrow I (P_i^S)) - P_i^S\|_2)
\]

Deep Adversarial Social Recommendation, IJCAI, 2019
Optimization for Ranking Tasks

- **Negative Sampling’s Main Issue:**
  - It often generates low-quality negative samples that do not help you learn good representation.
Optimization for Ranking Tasks

- **Negative Sampling’s Main Issue:**
  - It often generates *low-quality negative samples* that do not help you learn good representation [Cai and Wang, 2018; Wang et al., 2018b].

---

Deep Adversarial Social Recommendation, IJCAI, 2019
Optimization for Ranking Tasks

- **Negative Sampling’s Main Issue:**
  - It often generates low-quality negative samples that do not help you learn good representation [Cai and Wang, 2018; Wang et al., 2018b].

- Dynamically generate “difficult" negative samples

Deep Adversarial Social Recommendation, IJCAI, 2019
Deep Adversarial Social Recommendation, IJCAI, 2019
Deep Adversarial Social Recommendation, IJCAI, 2019
Item Domain Discriminator Model

- **Discriminator**
  
  **Goal:** distinguish real user-item pairs (i.e., real samples) and the generated “fake” samples (relevant)

  \[
  D^I(u_i, v_j; \phi_D^I) = \sigma(f_{\phi_D^I}^I(x_i^I, y_j^I)) = \frac{1}{1 + \exp(-f_{\phi_D^I}^I(x_i^I, y_j^I))} \quad \text{(Sigmoid)}
  \]

  **Score function:**
  
  \[
  f_{\phi_D^I}^I(x_i^I, y_j^I) = (x_i^I)^T y_j^I + a_j,
  \]

Deep Adversarial Social Recommendation, IJCAI, 2019
Item Domain Generator Model

Generator Model

Goal:
1. Approximate the underlying real conditional distribution $p^I_{\text{real}}(v | u_i)$
2. Generate (select/sample) the most relevant items for any given user $u_i$.

$G^I(v_j | u_i; \theta^I_G) = \frac{\exp(g^I_{\theta^I_G}(p^S_{SI}, q^I_j))}{\sum_{v_j \in \mathcal{V}} \exp(g^I_{\theta^I_G}(p^S_{SI}, q^I_j))}$

$g^I_{\theta^I_G}(p^S_{SI}, q^I_j) = (p^S_{SI})^T q^I_j + b_j$

Optimization with Policy Gradient

- Real Samples
- Generated Samples
- Loss
- Yes/No
- Reward

$p=0.20$
$p=0.03$
$p=0.09$
$p=0.58$
Sequential (Session-based) Recommendation

user’s sequential behavior

Next Item

0.8
0.6
0.1

Session-based Recommendations with Recurrent Neural Networks, ICLR, 2016.
BERT4Rec: Sequential Recommendation with Bidirectional Encoder Representations from Transformer, CIKM, 2019.
Sequential (Session-based) Recommendation

Session-based Recommendations with Recurrent Neural Networks, ICLR, 2016.
BERT4Rec: Sequential Recommendation with Bidirectional Encoder Representations from Transformer, CIKM, 2019.
Sequential (Session-based) Recommendation

Session-based Recommendations with Recurrent Neural Networks, ICLR, 2016.
BERT4Rec: Sequential Recommendation with Bidirectional Encoder Representations from Transformer, CIKM, 2019.
Shortcomings of Existing Deep Recommender Systems

Recommendation Policies
- Offline optimization
- Short-term reward
Shortcomings of Existing Deep Recommender Systems

- Recommendation Policies
  - Offline optimization
  - Short-term reward

- Graph-structured Data
  - Information Isolated Island Issue: ignore implicit/explicit relationships among instances
Shortcomings of Existing Deep Recommender Systems

Recommendation Policies
- Offline optimization
- Short-term reward

Graph-structured Data
- Information Isolated Island
  - Issue: ignore implicit/explicit relationships among instances

Manually Designed Architectures
- Expert knowledge
- Time and engineering efforts