Modeling Prediction in Recommender Systems Using Restricted Boltzmann Machine

Hanene Ben Yedder, Umme Zakia, Aly Ahmed, and Ljiljana Trajković
Communication Networks Laboratory, Simon Fraser University, Vancouver, British Columbia, Canada

INTRODUCTION

- Recommender systems (RSs) significantly enhance the users' experience when accessing online services.
- RSs are used to generate lists of suggestions using approaches such as collaborative filtering, content-based filtering, or hybrid methods.
- Collaborative filtering (CF) predicts a user's selection of a new advertisement based on past viewing history of users.
- CF prediction accuracy significantly decreases when ratings are very sparse thus limiting the extraction of useful features.

PROPOSED MODEL

- Employs the Restricted Boltzmann Machine (RBM) for collaborative filtering.
- The Neighborhood-Conditional RBM (N-CRBM) model is based on joint distributions of similarity and popularity scores.
- The model is trained and evaluated based on the number of hidden units, learning rates, and activation functions.

RECOMMENDER SYSTEMS

- The Clusters-Based RBM model is extended to incorporate advertisements based on the neighborhood content.

N-CRBM MODEL

- The proposed Neighborhood Conditional RBM (N-CRBM) model with similarity S and popularity P layers. Visible layer V and hidden layer H form a clusters-based RBM for a group of users viewing or selecting the same advertisements.
- Similarity score of an advertisement / for the user i:
  \[ S_i(u, i) = \frac{\sum_{a \in N(S)} \text{Rating}_x(i)}{N(u)} \]
- Popularity score of an advertisement / for the user i:
  \[ P_i(d, i) = \frac{\sum_{v \in D_u} \text{Rating}_y(i, l)}{D(u) \cdot R_{\text{max}}} \]
- Joint distribution of scores (V, H), conditional on the similarity score S and popularity score P:
  \[ P(H_i | V, S, P) = \sigma \left( a_j + \sum_{i \in V} V_i W_{ij} + \sum_{i \in S} S_i N_{ij} + \sum_{i \in P} P_i D_{ij} \right) \]

CRBM MODEL

- We first design a Clusters-Based RBM model by clustering users who viewed the same advertisements.

N-CRBM MODEL TRAINING AND TESTING

- Parameters required to calculate the gradient descent in the log-likelihood for model training:
  \[ \Delta W_{ij} = \varepsilon \left( < V_i H_j >_{\text{data}} - < V_i H_j >_{\text{rep}} \right) \]
  \[ \Delta S_{ij} = \varepsilon \left( < S_i H_j >_{\text{data}} - < S_i H_j >_{\text{rep}} \right) \]
  \[ \Delta D_{ij} = \varepsilon \left( < P_i H_j >_{\text{data}} - < P_i H_j >_{\text{rep}} \right) \]
- Predicted ratings of an advertisement q are used for model testing:
  \[ P(q_i = 1 | V, S, P) = \sigma \left( a_q + \sum_{j \in V} V_j W_{ij} + \sum_{j \in S} S_j N_{ij} + \sum_{j \in P} P_j D_{ij} \right) \]

KAGGLE DATASET

- Navigation histories of advertisements viewed or selected by users. Each viewed or selected advertisement is accompanied by semantic attributes of the visited documents.

REFERENCES