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

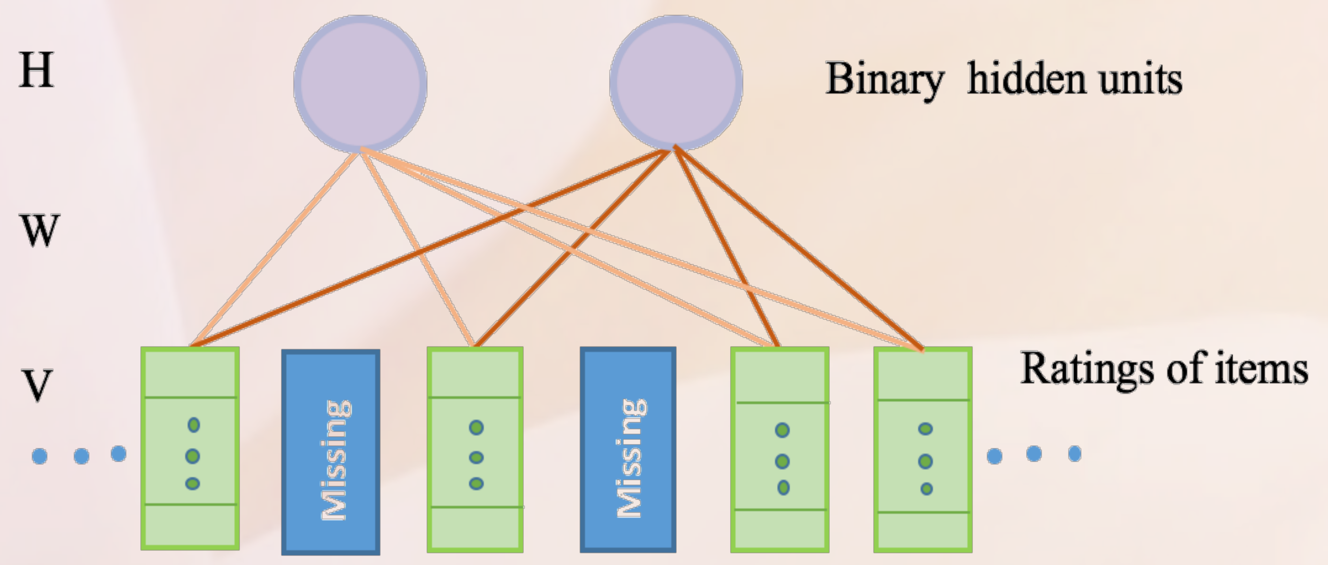
- Recommender systems (RSs) significantly enhance the users' experience when accessing online services.
- RSs are used to generate lists of suggestions using approaches such as collaborative filtering, content-based filtering, or hybrid methods.
- Collaborative filtering (CF) predicts a user's selection 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.

Restricted Boltzmann Machine

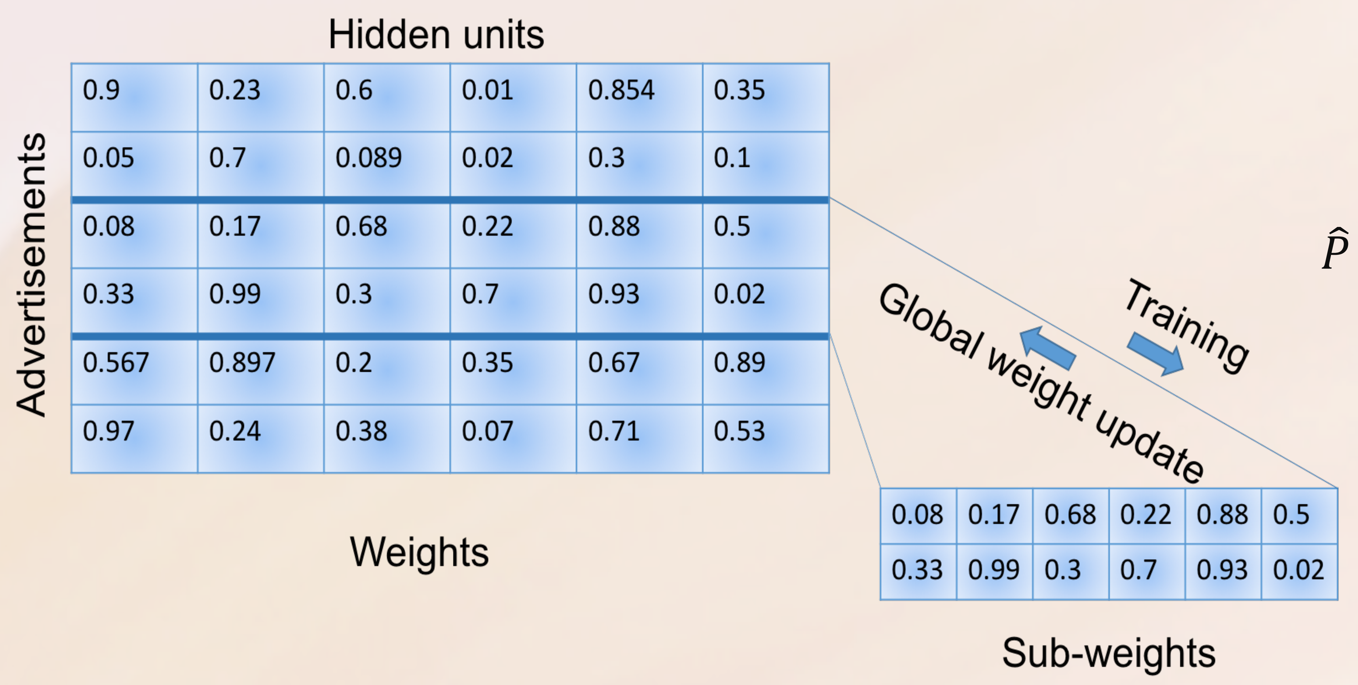
- RBM is a generative stochastic artificial neural network that learns a probability distribution over a set of inputs.
- Employs gradient descent approximation algorithms such as contrastive divergence.
- Belongs to energy-based models and consists of Bernoulli-valued (binary) hidden and visible units.
- RBM models are useful when designing deep learning models and offer more accurate results using automatically learned features while hiding the details.
- Conditional Restricted Boltzmann Machine (CRBM) is a probabilistic model that considers in visible layer both rated and unrated items along with additional information.



- The RBM model: Each visible unit corresponds to an item that is rated. Items that are not rated are considered missing.

CRBM Model

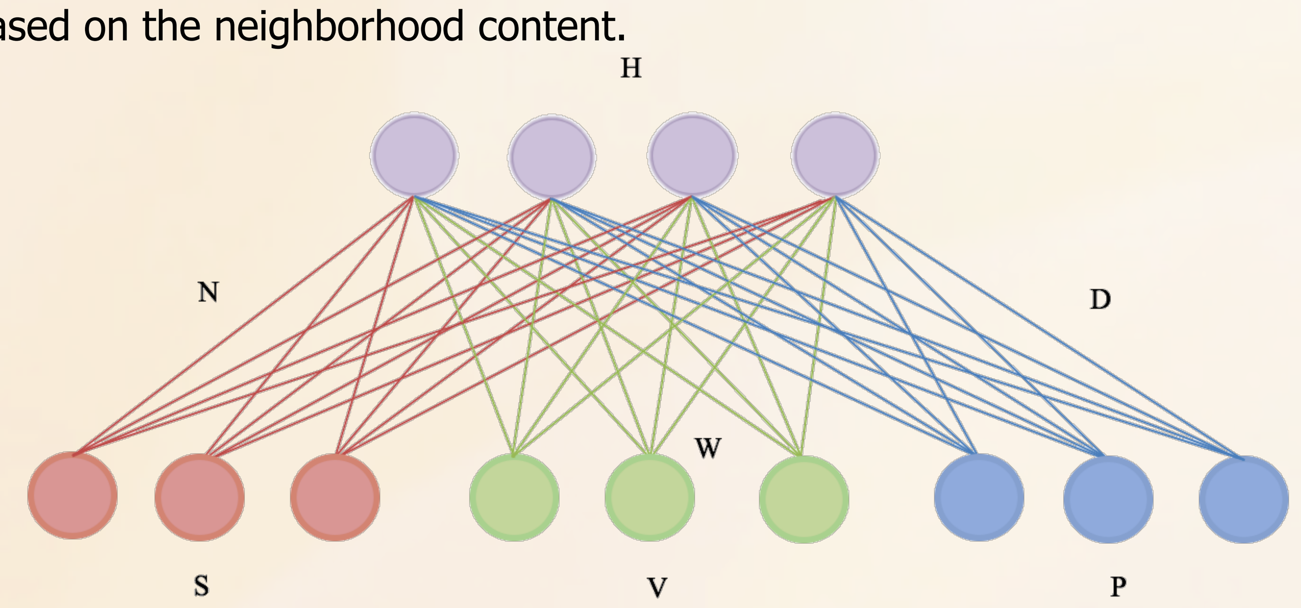
- We first design a Clusters-Based RBM model by clustering users who viewed the same advertisements.



- Sub-weights learned by an RBM that is designed for each cluster of users viewing the same advertisements. The weight matrix is shared among all RBMs.

N-CRBM Model

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



- 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 *i* for the user *u*:

$$S_i\,(u,i) = \frac{\sum_{x\in N_u} Rating_{x,i}}{N(u)}$$

- Popularity score of an advertisement *i* for the user *u*:

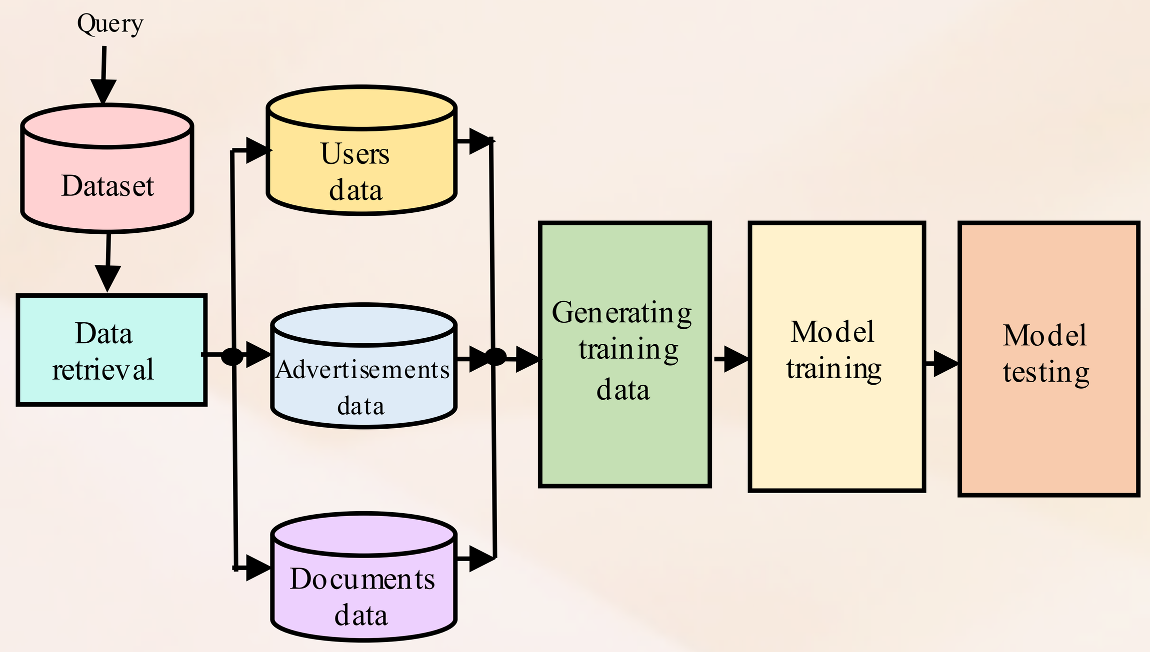
$$P_i\,(d,i) = \frac{\sum_{y\in D_u} Rating_{y,i}}{D(u)\,R_{max}}$$

- Joint distribution of scores (V, H), conditional on the similarity score S*i* and popularity score P*i*:

$$P(H_j = 1|V,S,P\,) = \sigma\left(a_j + \sum_{i\in V} V_i\,W_{ij} + \sum_{i\in S} S_i\,N_{ij} + \sum_{i\in P} P_i\,D_{ij}\right)$$

N-CRBM Model Training and Testing

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



- Parameters required to calculate the gradient descent in the log-likelihood for model training:

$$\Delta W_{ij} = \epsilon(< V_i\,H_j >_{data} - < V_i\;H_j >_T)$$
$$\Delta N_{ij} = \epsilon(< S_i\,H_j >_{data} - < S_i\,H_j >_T)$$
$$\Delta D_{ij} = \epsilon(< P_i\,H_j >_{data} - < P_i\,H_j >_T)$$

- Predicted ratings of an advertisement q are used for model testing:

$$\hat{P} = P(H_j = 1|V,S,P) = \sigma(a_j + \textstyle\sum_{i\in V} V_i\,W_{ij} + \sum_{i\in S} S_i\,N_{ij} + \sum_{i\in P} P_i\,D_{ij})$$
$$P(V_q = 1|\hat{P}) = \sigma(a_q + \textstyle\sum_{j=1}^{F} \hat{P}_j W_{qj})$$

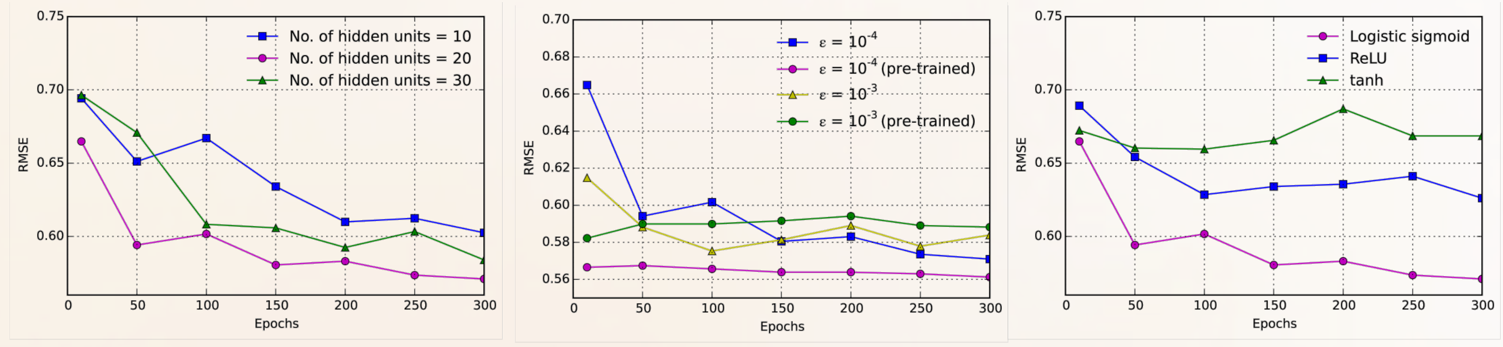
Kaggle Dataset

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

Number	Sample subset	Kaggle dataset
Unique advertisements	$3\text{x}10^3$	$330\text{x}10^3$
Unique users	$9\text{x}10^3$	$2\text{x}10^6$
Records	$27\text{x}10^3$	$22\text{x}10^6$

Selection of Model Parameters

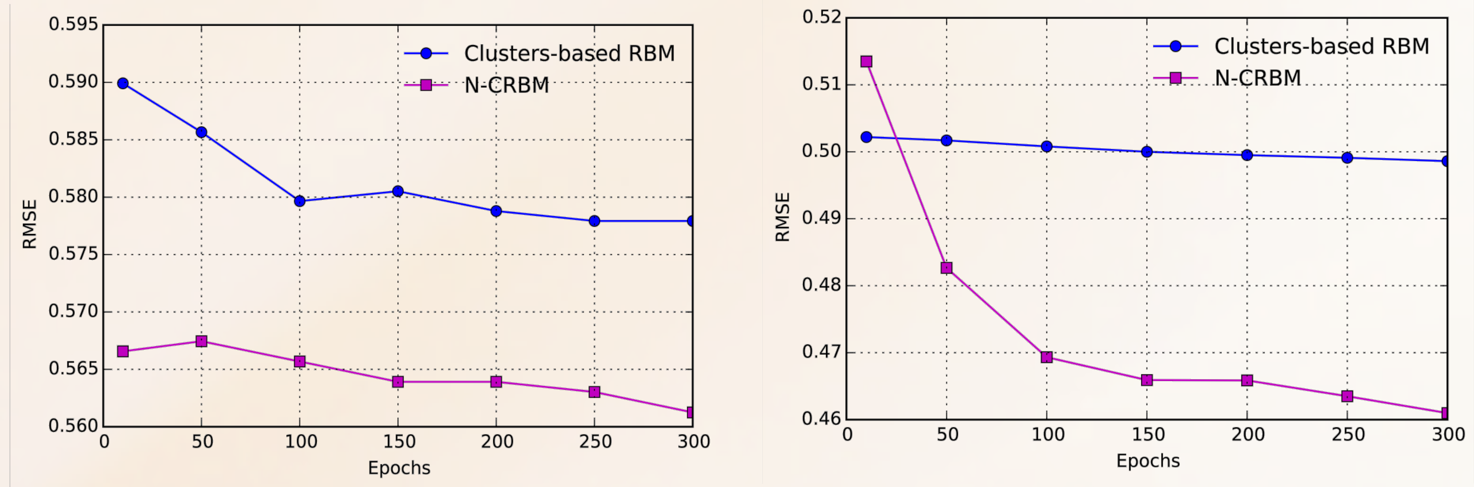
- The sample subset was used to determine the optimal number of hidden units, learning rates, random vs. learned weights, and activation functions for the Clusters-Based RBM model.
- The best learning results: 20 hidden units, learning rate 10^{-4} using pre-trained models (learned weights), and logistic sigmoid for activation functions.



Hidden Units | Learning Rates | Activation Functions

Performance of RBMs

- RMSE:



Sample Subset | Kaggle Dataset

- Accuracy:

Dataset	Clusters-Based RBM (%)	N-CRBM (%)
Sample subset	66.6	68.9
Kaggle dataset	76.0	78.5

- Sensitivity:

Dataset	Clusters-Based RBM (%)	N-CRBM (%)
Sample subset	65.1	69.5
Kaggle dataset	18.7	29.4

Conclusion

- Despite the sparsity of data in the Kaggle dataset, the N-CRBM model outperforms the clusters-based RBM model in terms of RMSE, accuracy and sensitivity due to the similarity and popularity scores.
- The additional neighborhood features help to overcome the CF cold-start problem and enhance the ability of N-CRBM model in recommending advertisements to a user.

References

- C. M. Bishop, *Pattern Recognition and Machine Learning.* New York, NY, USA: Springer, 2006, Ch. 2, p. 133.
- J. Bobadilla, F. Ortega, A. Hernando, and A. Gutierrez, "Recommender systems survey," *J. Knowledge Based Systems*, vol. 46, pp. 109–132, July 2013.
- G. E. Hinton, "A practical guide to training restricted Boltzmann machines," in *Lecture Notes in Computer Science, Neural Networks: Tricks of the Trade*, G. Montavon, G. B. Orr, K. R. Müller, Eds. Springer, 2012, vol. 7700, pp. 599–619.
- R. Salakhutdinov, A. Mnih, and G. Hinton, "Restricted Boltzmann machines for collaborative filtering," in *Proc. Int. Conf. Mach. Learning*, Corvallis, OR, USA, June 2007, pp. 791–798.
- (Apr. 04, 2017) Restricted Boltzmann machines. [Online]. Available: https://github.com/echen/restricted-boltzmann-machines.

SFU SIMON FRASER UNIVERSITY
THINKING OF THE WORLD

SMC 2017
The IEEE International Conference on Systems, Man, and Cybernetics
October 2017, Banff, Alberta, Canada

Communication Networks Laboratory
http://www.ensc.sfu.ca/~ljilja/cnl