A non-IID Framework for Collaborative Filtering with Restricted Boltzmann Machines

Kostadin Georgiev and Preslav Nakov (2013)

Presented by Boyoung Kim

January 10, 2017

Contents

1. User/Item based RBM (U-RBM/ I-RBM)
2. Hybrid non-IID RBM (UI-RBM)
3. Neighborhood method boosted by I-RBM (I-RBM + INB)
4. Experimental Results

RBM for CF(Salakhutdinov et al.)

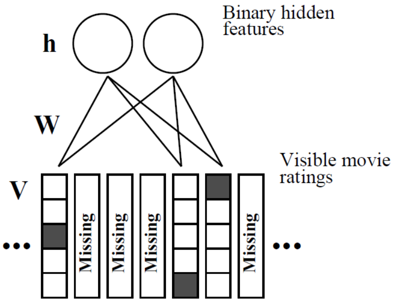

Figure 1: RBM with multinomial visible units

- User-based : use different RBM for each user.
- Ratings as multinomial variables.

User based RBM model

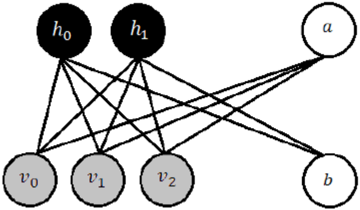

Figure 2: The user based RBM model with real-valued visible units.

- For N users, M items, F hidden units,
- $\{v_1, v_2, \cdots, v_M\}$: M real-valued visible units.
- $\{h_1, h_2, \cdots, h_F\}$: F binary hidden units.
- It has modeled the correlation between item ratings only.

User based RBM model

- Conditional Bernoulli distribution for modeling "hidden" user feature:

$$P(h_j = 1|\mathbf{V}) = \sigma(b_j + \sum_{i=1}^{M} W_{ij} v_i)$$

.

- Conditional normal distribution for modeling "visible" rating :

$$P(v_i|\mathbf{h}) = \mathcal{N}(a_i + \sum_{j=1}^{F} W_{ij} h_j, \sigma_i^2)$$

where σ_i^2 is the variance of the *ith* visible unit, bias a_i for the visible unit v_i, and bias b_j for the hidden unit h_j.

- For prediction, $\hat{v}_i = a_i + \sum_{j=1}^{F} W_{ij} h_j$.

Item based RBM model

- $\{v_1, v_2, \cdots, v_N\}$ represent all ratings for an item.
- It has modeled the correlation between user ratings only.

Non-IID Hybrid RBM model

- Combination of the two RBM models into a single unifed model.
- Takes into account both item-item and user-user correlations.

- Let v_{ij} be the visible unit that corresponds to the rating of user i on item j.
- The unit v_{ij} is connected to two independent hidden layers - one user-based and another item-based.

Non-IID Hybrid RBM model

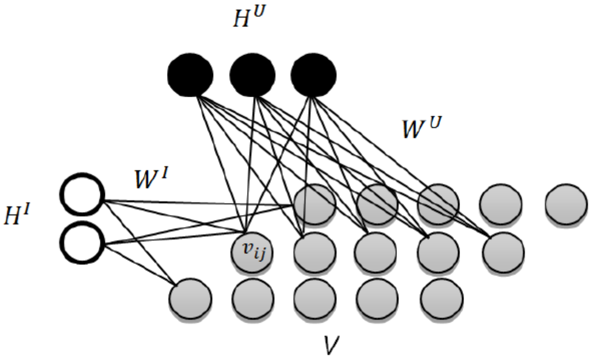

Figure 3: User-item based RBM model.

Non-IID Hybrid RBM model

- Let the upper index U/I be related to the user-based/item-based hidden layer.
- For example, for the upper index U,
    - F^U : size of the user-based hidden layer
    - W_{jk}^U : weight between item j and user-based hidden unit k
    - a_i^U : bias for the user i
    - b_k^U : bias for the user-based hidden unit k

Learning/ Prediction

- "Contrastive Divergence" (CD) :

$$\Delta W_{jk}^{U} = \epsilon \frac{1}{N} \sum_{i=1}^{N} (< v_{ij} h_k^U >_{data} - < v_{ij} h_k^U >_T)$$

$$\Delta W_{ik}^{I} = \epsilon \frac{1}{M} \sum_{i=1}^{M} (< v_{ij} h_k^I >_{data} - < v_{ij} h_k^I >_T)$$

- We average the predictions of two models :

$$\hat{r}_{ij} = \frac{1}{2} \left[a_i^U + \sum_{k=1}^{F^U} W_{jk}^U h_k^U + a_j^I + \sum_{k=1}^{F^I} W_{ik}^I h_k^I \right].$$

Neighborhood method boosted by I-RBM

- Combine the predictions of an RBM model with those of a standard neighborhood-based model.

- Let r'_{ui} be rating of user u for item i by the I-RBM model.
- let $\overline{r_i}$ be average rating for item i.
- Pearson correlation

$$Corr_{ij} = \frac{\sum_{u=1}^{N}(r'_{ui} - \overline{r_i})(r'_{uj} - \overline{r_j})}{\sqrt{\sum_{u=1}^{N}(r'_{ui} - \overline{r_i})^2}\sqrt{\sum_{u=1}^{N}(r'_{uj} - \overline{r_j})^2}}$$

- Prediction

$$\hat{r_{ui}} = \overline{r_i} + \frac{\sum_{j=1}^{M}(r_{uj} - \overline{r_j})Corr_{ij}}{\sum_{j=1}^{M}|Corr_{ij}|}$$

Experimental Results with MovieLens data (100k)

CF Model	MAE
SVD PCA (Vozalis et al., 2010)	0.793
H-NLPCA (Vozalis et al., 2010)	0.784
U-RBM	**0.779**
I-RBM	**0.775**
SVD (Sarwar et al., 2002)	0.733
Item-based CF (Sarwar et al., 2001)	0.726
Iter PCA + K-means (Kim & Yum, 2005)	0.712
Iter PCA + RRC (Kim & Yum, 2005)	0.700
I-RBM+INB	**0.699**
UI-RBM	**0.690**
Latent CF (Langseth & Nielsen, 2012)	0.685

Table 1: Comparison (on 100k) of the prediction quality of various CF models and our RBM-based models (in bold).

- Error comp. : U-RBM > I-RBM > SVD > I-RBM+INB > UI-RBM

Experimental Results with MovieLens data (1M)

CF Model	MAE
Real U-RBM	**0.762**
Real I-RBM	**0.761**
L8 (Taranto et al., 2012)	0.720
Multinomial U-RBM	0.711
Multinomial I-RBM	0.710
Multinomial UI-RBM	**0.685**
GAUSS-UI-BM (Truyen et al., 2009)	0.675
Real I-RBM+INB	**0.669**
ORD-UI-BM (Truyen et al., 2009)	0.657
Real UI-RBM	**0.645**
ORD-UI-BM-CORR (Truyen et al., 2009)	0.640

Table 2: Comparison (on 1M) of the prediction quality of various CF models and our RBM-based models (in bold).

- real U-RBM/I-RBM > multinomial U-RBM/I-RBM > real I-RBM+INB
- multinomial UI-RBM > real UI-RBM