A Generative Model for Review-Based Recommendations

INTUIT Ai

Oren Sar Shalom, Guy Uziel, Amir Kantor

IBM

Problem

Setting:
- input: textual reviews and explicit ratings
- Output: personalized predicted rating

Intuition 1:
- Reviews don't necessarily reveal user preferences or item traits
- They describe the *matching* between them
- Spans a single review which is easier to process

Intuition 2:
- Reviews might be ambiguous or simply do not cover some important aspects
- There could be a varied level of uncertainty
- Infer the *distribution* of the matching

IMDb

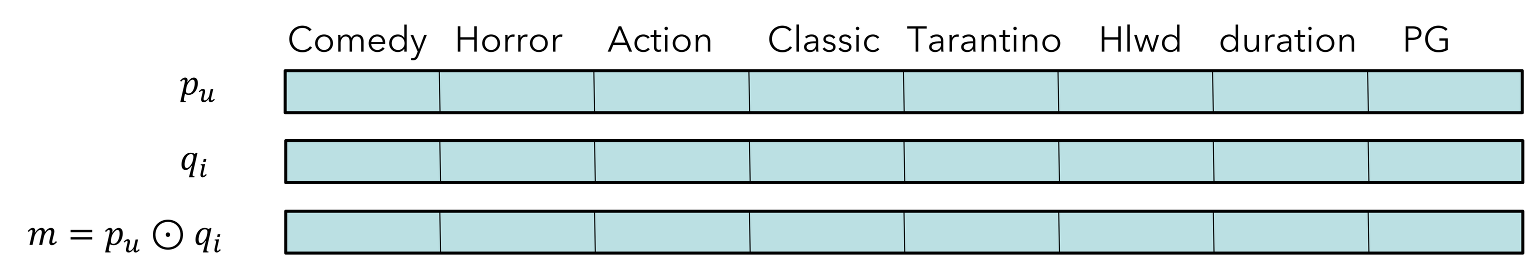
10/10

"The best movie I ever watched in any form or any in language".

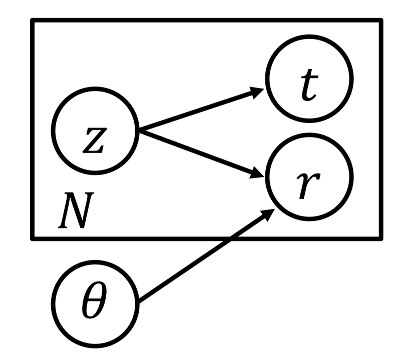
No, I am not exaggerating! It is the best movie ever, whether in human form or in animated. The story is nice and free flowing and without any unnecessary bumpers. One can not recognized the voices behind the characters, because the characters are so prominent. Music! Let me say I am not very fond of English pop songs, but I could hum almost most of the songs of this movie, so I would say, music is that nice. Animation is superb, landscapes are eye catching, dialogues are sharp and comedy is at its best. What more might I say?

Solution – Phase I

Matching Vectors



Generative Model



Vector z influences the user while generating the observed r and t

Matching vectors distribution

Objective

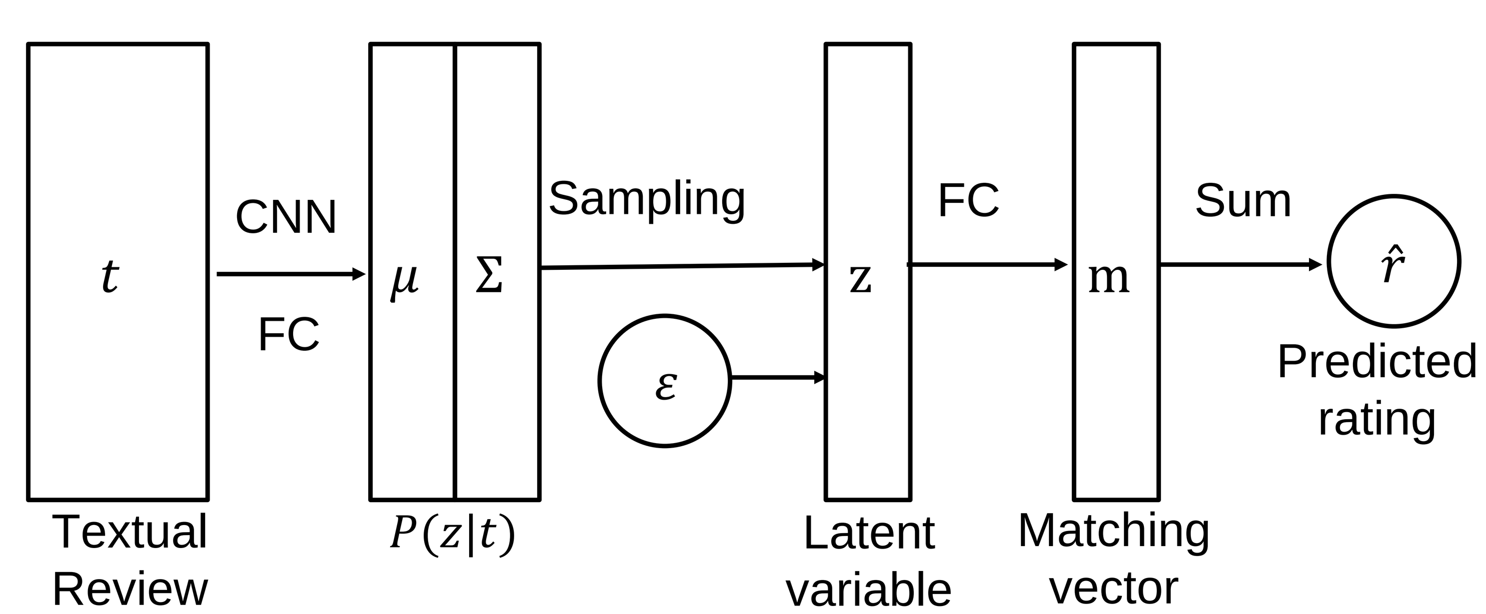
$$\sum_{(r,t)\in D} \ln P(r|t;\theta) = \sum_{(r,t)\in D} \ln \int_z P(r|z,t)P(z|t;\theta) =$$

$$\sum_{(r,t)\in D} \ln \int_z P(r|z;\theta)P(z|t) \geq N \cdot \mathbb{E}_{(r,t)\sim D}\mathbb{E}_{z\sim P(\cdot|t;\theta)}[\ln P(r|z)]$$

Lower bound for the objective

Decoder: $P(r|z)$
- z is an intermediate step in the process of inferring the matching vector. $m = g(z;\theta)$
- Overall matching: $\hat{r} = \sum_f m^{[f]}$
- Randomness effect: $P(r|z;\theta) = N(\hat{r}, \sigma_r^2)$

Encoder: $P(z|t)$
- We assume $P(z|t) \sim N(\mu_t, \Sigma_t)$
  - Σ_t is diagonal
- μ_t, Σ_t are implemented via a CNN-based architecture
  - Separate only in the last layer



Solution – Phase II

Augmented CF

- Common to latent factor models – inner product indicates affinity
- $\hat{r}_{ui} = p_u \cdot q_i = \sum_f (\boldsymbol{p_u} \odot \boldsymbol{q_i})^{[f]}$
- Objective: $\min_\theta \mathrm{E}[l(r, \hat{r}; \theta)]$
  - Single scalar label

- Augmented objective:
  - Matching vectors: apply Phase I and take the mode of the distribution
  - $\min_\theta \mathrm{E}\left[l(r, \hat{r}; \theta) + \alpha \cdot \|p_u \odot q_i - m\|_2^2\right]$
  - A per-dimension label
  - $\hat{r} = \sum_f p_u^{[f]} \cdot q_i^{[f]} = 4, r = 3$
  - All factors $\left\{p_u^{[f]} \cdot q_i^{[f]}\right\}$ will be decreased

Results

	SVD	DeepCoNN	DeepCoNN-rev$_{AB}$	TransNet	TransNet-Ext	MDR	Improved
Yelp	1.8661	1.8984	1.7045	1.6387	1.5913	**1.4257**	10.4%
A-Electronics	1.8898	1.9704	2.0774	1.8380	1.7781	**1.5329**	13.8%
A-Clothes	1.5212	1.5487	1.7044	1.4487	1.4780	**1.2837**	13.1%
A-Movies	1.4324	1.3611	1.5276	1.3599	1.2691	**1.1782**	7.2%

MSE comparison with baselines

	MSE
Yelp	0.52
A-Electronics	0.76
A-Clothes	0.43
A-Movies	0.29

MSE obtained by Phase I

Summary

- A robust and fast approach for rating prediction
- Achieved by capturing the dynamics of user generated reviews