Principle to Program: Neural Review mining for recommendation

Muthusamy Chelliah
Flipkart

Sudeshna Sarkar
IIT Kharagpur

Vishal Kakkar
Flipkart

Agenda

1. Introduction/topic models - cold start/explainability (10 min.s) (Chelliah)
2. Basic neural methods for recommendation (20 min.s) (Chelliah)
3. Attention - review ranking/aspect-based recommendation (10 min.s) (Vishal)
4. Programs - attention (20 min.s) (Vishal)
4. Generative approaches - review/tips (5 min.s) (Vishal)

5. Programs - generation (10 min.s) (Vishal)
6. Attention/generative approaches (15 min.s) (Sudeshna)

1.Introduction/background

Collaborative filtering

People who share similar preferences in past

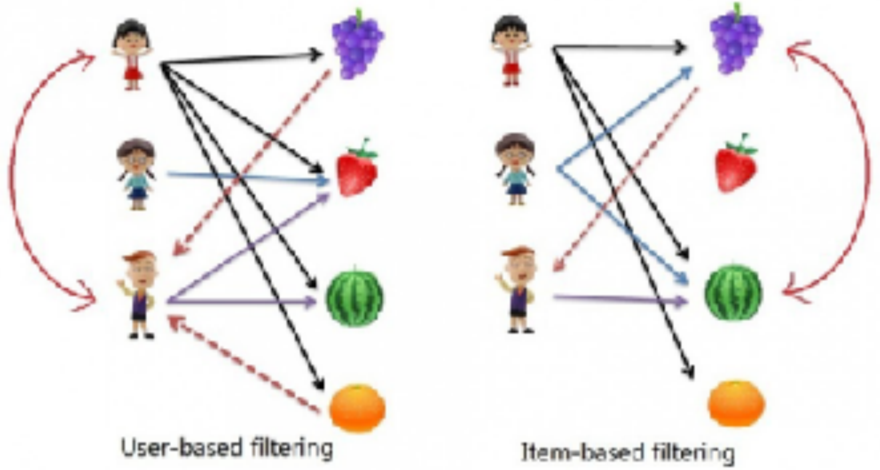
- Exercise same choices in future

Predict user preference for a product

- Based on other products/users
- Rating of unobserved user/ product pair

Side information

- User/product characteristics
- Natural language reviews



Matrix factorization

Observed ratings as elements

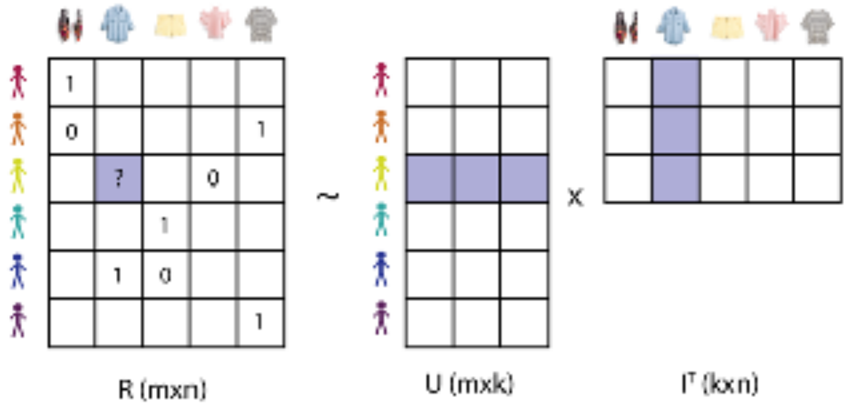
- Given by set of users to set of items

Row/column correspond to user/item

Fill in missing values

Find common factors

- Underlying reasons for ratings
- E.g., movie recommendation
- (genre, actor, director)



Flipkart

Matrix factorization (contd.)

Give importance of each factor/user

Factorized into product of 2 matrices

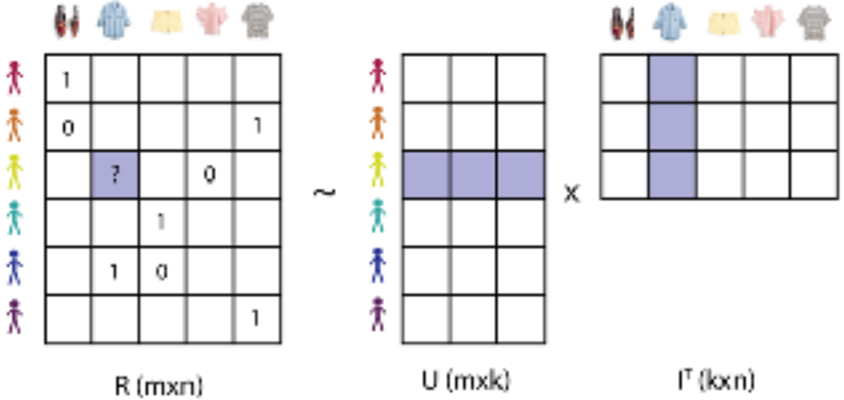
- Of lower rank representing item/ user

Rating reconstructed as dot product

- Of item/user vectors

Naive factorization overfits

- Training set of observed ratings



Flipkart

Data sparsity

No. of items rated by users

- insignificant to total no.

Netflix Challenge

- 100M out of 8B possible ratings
- 500K users; 18K movie

Amazon review datasets

- 99.99% rating info. missing

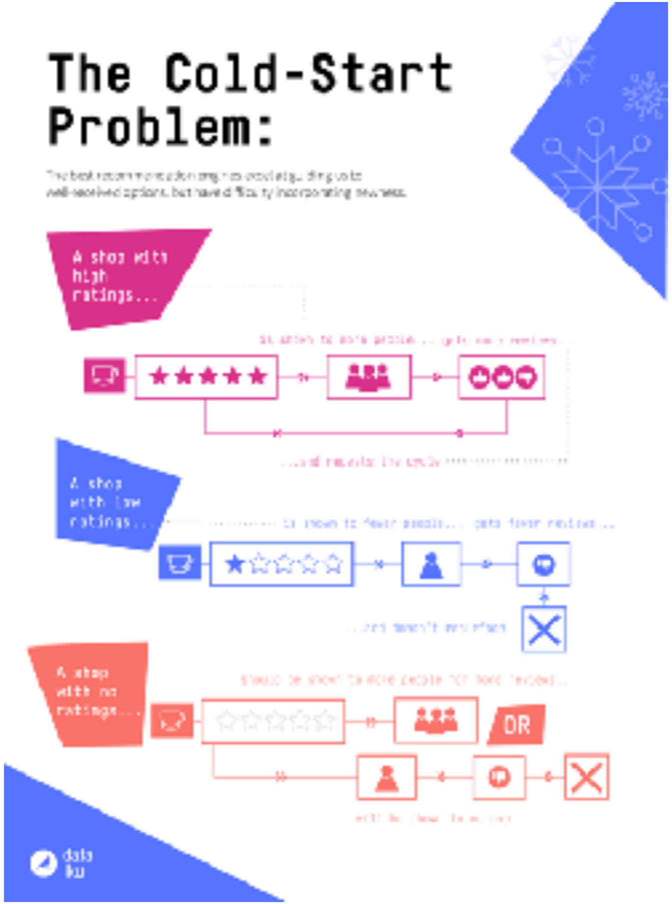
Cold start

New users/items with fewer ratings not exposed much

- Popular items recommended more

Reviews to improve prediction accuracy

- Extract/leverage element characteristics/ user choices



Flipkart

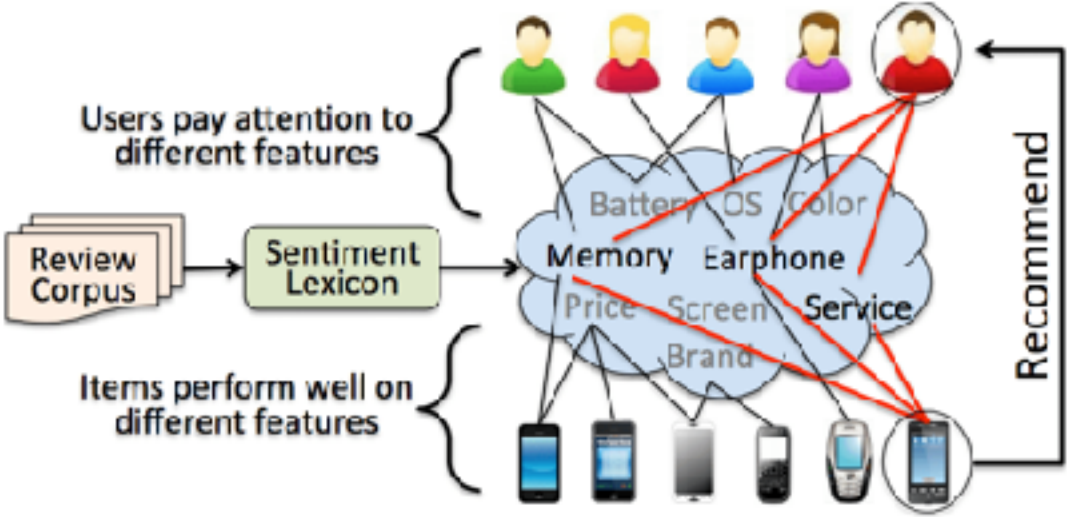
Explainable recommendation

Reviews justify user rating

- Describe features that affect opinions
- With details why an item is liked

Rich semantic, textual information

- Vs. interaction data/ logs



Multi-task learning

- Rating prediction
- Natural language modelling

Capture author preferences

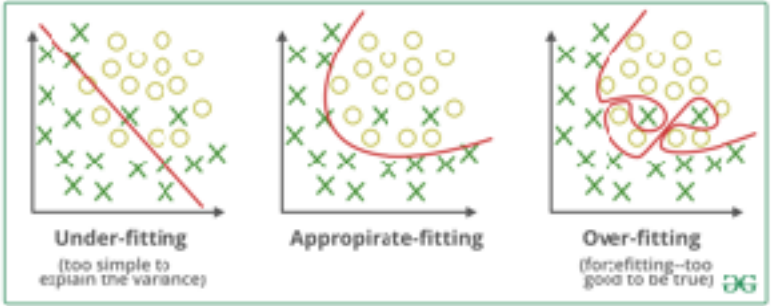
- Improve generalisation performance

Small subset of whole vocabulary

- Differ vastly from each other

$$r_{u,i} \approx \hat{r}_{u,i} = \mu + \beta_u + \beta_i + \gamma_u^\top \gamma_i,$$

Regularization with reviews



Higher rating to **Free Willie** movie

- If review is *this is a great movies children and adults like would love*

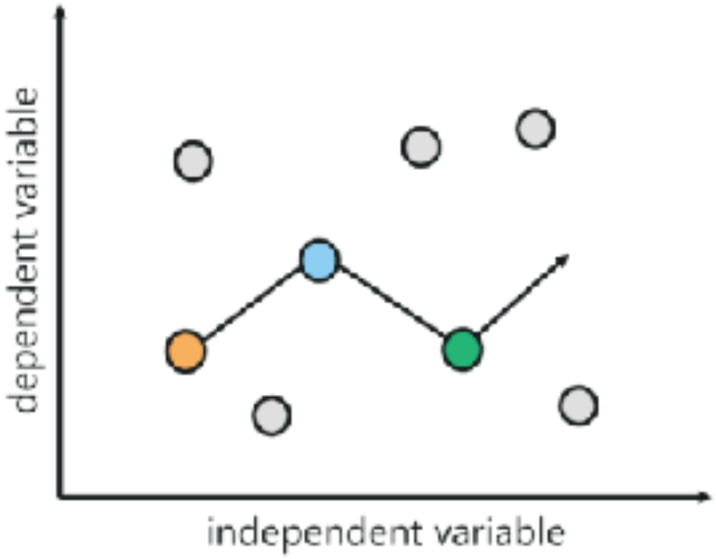
Conditional distribution

- Puts most probability on few product-specific words
- Leaving others with zero chance

Parameterize word conditional probability

- Given a product representation

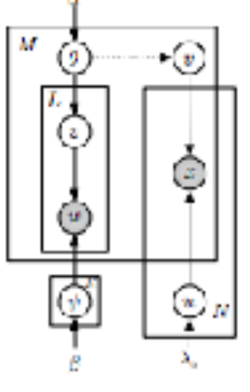
$$p\left(d_{u,i} = \left(w_{u,i}^{(1)}, \cdots, w_{u,i}^{(n_{u,i})}\right) \mid \boldsymbol{\gamma}_i, \boldsymbol{\theta}_D\right)$$



[Almahairi 15] Learning distributed representations from reviews for collaborative filtering, RecSys

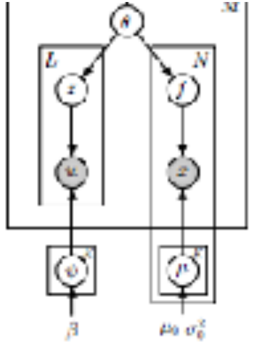
Topic models - early work (Cold start)

[Mcauley 13] Transform function to learn latent factors/topics together towards interpretable textual labels which justify ratings with reviews



(c) HFT

[Ling 14] Ratings modelled using mixture of Gaussian to scale parameters with flexibility



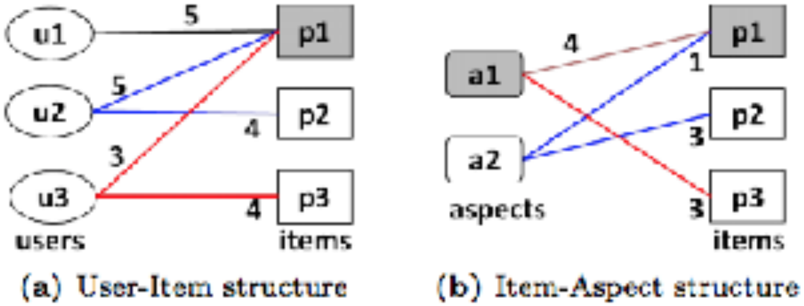
(d) RMR

[Mcauley 13] Hidden factors and topics: understanding rating dimensions with review topics, RecSys

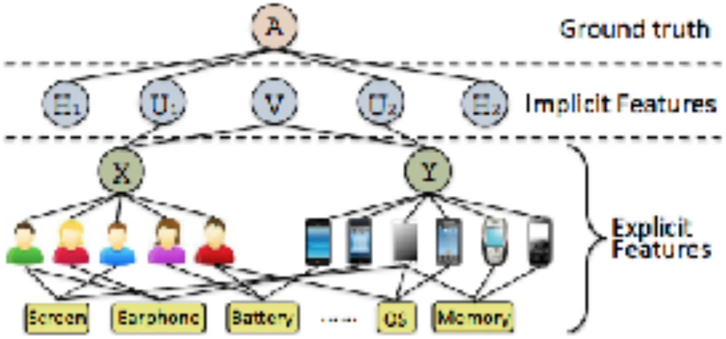
[Ling 14] Ratings meet reviews, a combined approach to recommend, RecSys

Explainability - early work

[Zhang 14] Factor graph model with explicit features as latent variables



External sentiment analysis tools uncover aspects/sentiments



[He 15] user-item-aspect ternary relation modelled as tripartite graph casts recommendation task to vertex ranking problem

[Zhang 15] Explicit factor models for explainable recommendation based on phrase-level sentiment analysis, SIGIR

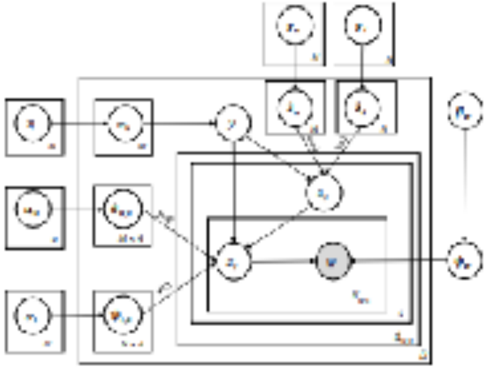
[He 15] Trirank: review-aware explainable recommendation by modelling aspect, CIKM

Topic models - recent advances (explainability)

[Bauman 17] Sentiment utility logistic model (SULM) recommends valuable aspects (e.g., seafood) based on user's potential experience with item (e.g., restaurant)

Meat	Fish	Dessert	Money	Service	Decor
beef	cod	tiramisu	price	bartender	design
meat	salmon	cheesecake	dollars	waiter	ceiling
bbq	catfish	chocolate	cost	service	decor
ribs	tuna	dessert	budget	hostess	lounge
veal	shark	ice cream	charge	manager	window
pork	fish	macaroons	check	staff	space

[Cheng 18] Aspect-aware latent factor model (ALFM) integrates aspect importance into fine-grained rating prediction with a weighted matrix



[Bauman 17] Aspect based recommendations: Recommending items with the most valuable aspects based on user reviews. KDD

[Cheng 18] Aspect-aware latent factor model: Rating prediction with ratings and reviews. WWW

Topic model for rating prediction

Probabilistic models extract from reviews

- syntactic/hidden semantic information

Transfer features from textual to rating space

BoW for representation

- Considers only stat.s not sequences

[Almahairi 15] Learning distributed representations from reviews for collaborative filtering, RecSys

Flipkart

Topic model for rating prediction (contd.)

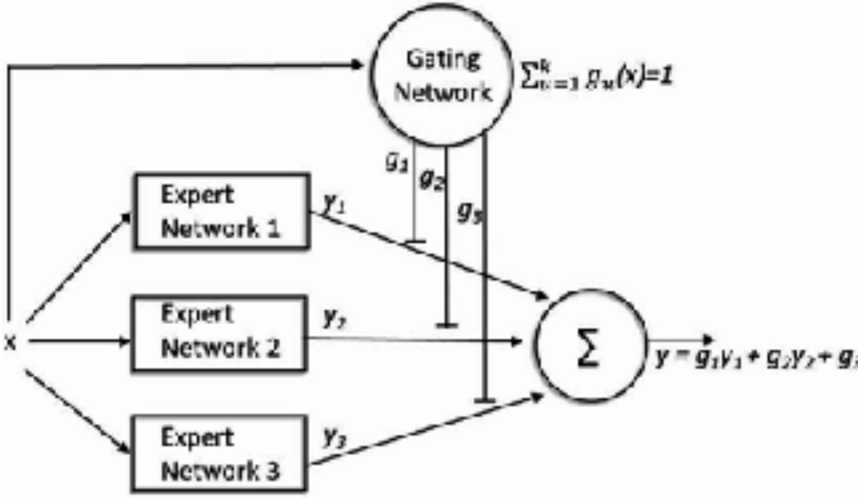
Only document-level co-occurrence

- Models words in review text

Mixture of experts on a per-word basis

- Topics are mixture components

No mechanism to suppress probability mass assigned to word



$$p(d_{u,i} \mid \gamma_i) = \prod_{t=1}^{n_{u,i}} \sum_{k=1}^{\dim(\gamma_i)} p(w_{u,i}^{(t)} \mid z_k = 1) p(z_k = 1 \mid \gamma_i) \tag{11}$$

$$= \prod_{t=1}^{n_{u,i}} \sum_{k=1}^{\dim(\gamma_i)} \pi_k p(w_{u,i}^{(t)} \mid z_k = 1)$$

$$p(w_{u,i}^{(t)} = j \mid \gamma_i) = \sum_{k=1}^{\dim(\gamma_i)} w_{j,k}^{*} \frac{\exp\{\kappa\gamma_{i,k}\}}{\|\exp\{\kappa\gamma_i\}\|_1}.$$

$$w_{j,k}^{*} = \frac{\exp\{q_{j,k}\}}{\sum_l \exp\{q_{l,k}\}},$$

[Almahairi 15] Learning distributed representations from reviews for collaborative filtering, RecSys

Neural network based document models

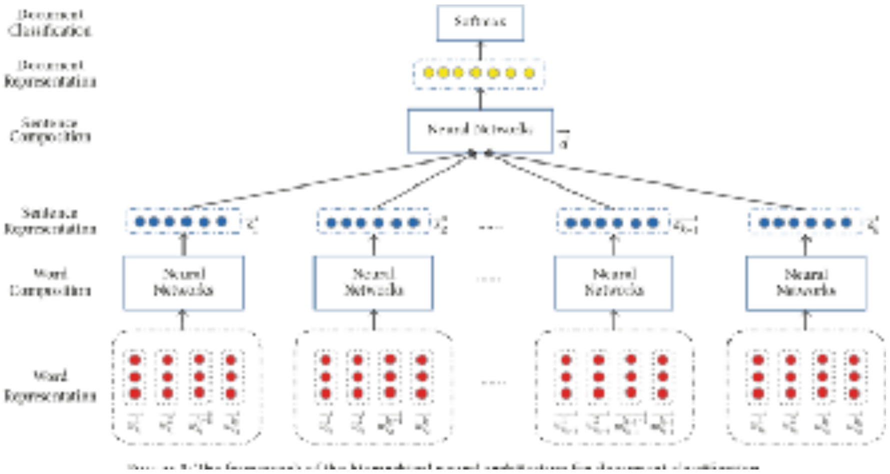
Simple matrix-vector product representation

- With no restriction on weight matrix

Elements can assume -ve values

Exercise strong influence

- Suppressing expression of given set of words



[Almahairi 15] Learning distributed representations from reviews for collaborative filtering, RecSys

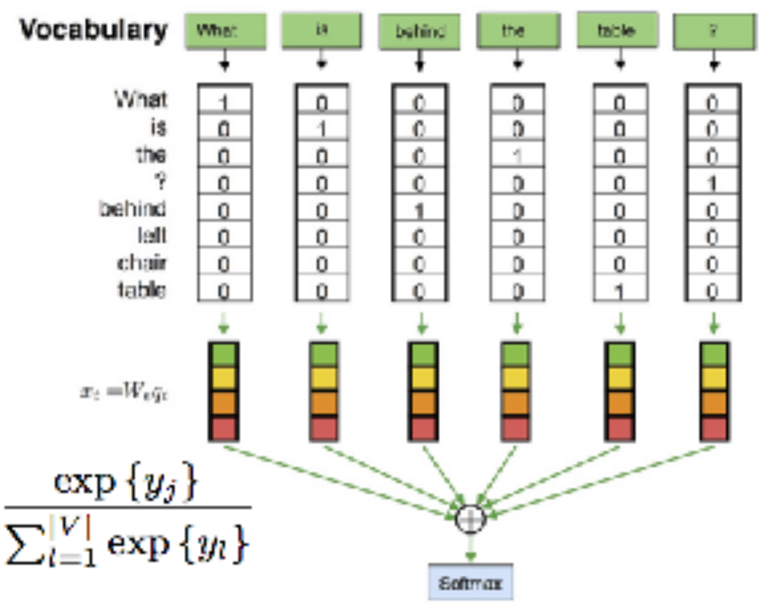
Product of experts (PoE)

- Models better peaky distribution in high-dimensional space

With a single affine transformation

- weight matrix plus offset vector

BoWLF: distributed bag of words



$$p(d_{u,i} \mid \boldsymbol{\gamma}_i) = \prod_{l=1}^{n_{u,i}} p(w_{u,i}^{(l)} \mid \boldsymbol{\gamma}_i).$$

$$p(w_{u,i}^{(l)} = j \mid \boldsymbol{\gamma}_i) = \frac{\exp\{y_j\}}{\sum_{l=1}^{|V|} \exp\{y_l\}}$$

$$\mathbf{y} = \mathbf{W}\boldsymbol{\gamma}_i + \mathbf{b}$$

[Almahairi 15] Learning distributed representations from reviews for collaborative filtering, RecSys

Chain of PoE

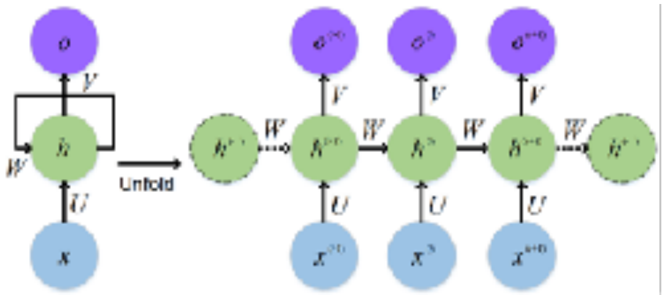
LMLF: RNN

Takes a set of words preserving its order

- Not clear if product-specific or language feature

Does not encode inside representation

- Underlying product related structure
- For rating prediction model to easily decode



$$p\left(w_{u,i}^{(t)} = j \mid w_{u,i}^{(<t)}, \gamma_i\right) = \frac{\exp\left\{y_j^{(t)}\right\}}{\sum_{l=1}^{|V|} \exp\left\{y_l^{(t)}\right\}}$$

$$p(d_{u,i} = (w_{u,i}^{(1)}, \cdots, w_{u,i}^{(n_{u,i})}) \mid \gamma_i)$$

$$= p\left(w_{u,i}^{(1)} \mid \gamma_i\right) \prod_{t=2}^{n_{u,i}} p\left(w_{u,i}^{(t)} \mid w_{u,i}^{(1)}, \cdots, w_{u,i}^{(t-1)}, \gamma_i\right)$$

$$\mathbf{y}^{(t)} = \mathbf{W}\mathbf{h}^{(t)} + \mathbf{b}$$

$$\mathbf{h}^{(t)} = \phi\left(\mathbf{h}^{(t-1)}, w_{u,i}^{(t-1)}, \gamma_i\right).$$

[Almahairi 15] Learning distributed representations from reviews for collaborative filtering, RecSys

Flipkart

Evaluation: Nearest neighbours based on product representations

Product	HFT	BoWLF	LMLF
Extra Spearmint Sugarfree Gum	Hong Kong Fu Xiang Yuan Moon Cakes French Chew - Vanilla Peck's Anchovette	Dubble Bubble Gum Trident Sugarless White Gum Gold Mine Nugget Bubble Gum	Gumballs Special Assorted Bazooka Bubble Gum Gourmet Spicy Beef Jerky
Dark Chocolate Truffle	Tastykake Kreamies Kakes Ceam .. Miko - Awase Miso Soyabean Paste Haribo Berries Gummi Candy	Ritter Sport Corn Flakes Chocolate Chocolate Dobosh Torte Sugar Free, Milk Chocolate Pecan Turtles	Fantis Grape Leaves Grape Flavoring Tutti Fruitti Flavoring
MTR Simply Tomato Soup	Wellington Cracked Pepper Crackers Maggi Instant Noodles Haribo Gummi Candy	MTR Mulligatawny Soup hai Kitchen Coconut Ginger Soup Miko - Awase Miso Soyabean Paste	Muir Glen Organic Soup Soy Ginger Saba Noodles Alessi Soup

Cosine similarity

- Estimated for Gourmet Foods dataset

Much broader for HFT

- Gummy bears/crackers for soup

LMLF results worse qualitatively

[Almahairi 15] Learning distributed representations from reviews for collaborative filtering, RecSys

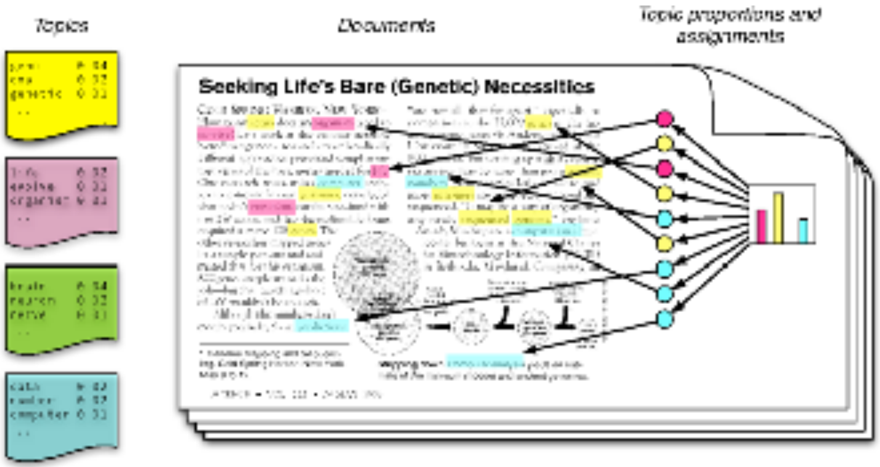
Flipkart

LDA limitations

Ignorance of word local context

Direct equivalence between user/item and review topic factors

- Not all words/topics relevant to rating
- Brings irrelevant information to some dimensions of embeddings
  - Thus affecting prediction accuracy



[Zhang 16] Collaborative multi-Level embedding learning from reviews for rating prediction, IJCAI

Flipkart

Surrounding text for each word

Word local context

Topic model considers word co-occurrence at document level

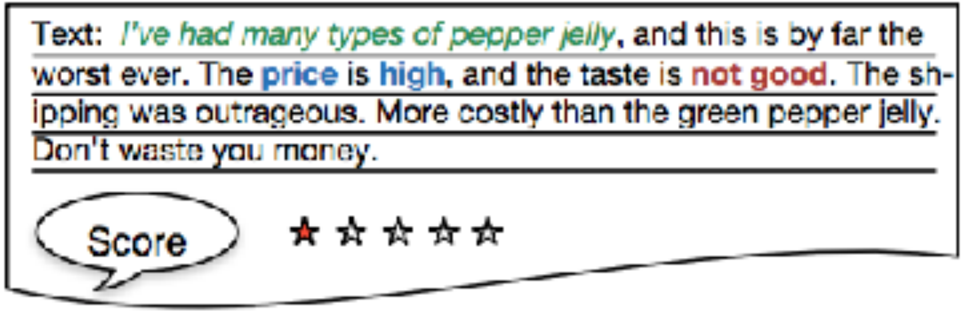
- Regards all word in each document equally

Local relation cannot be captured

- E.g., No, not for good

Sentiment polarity depends on aspect described

- E.g., High price vs. quality



Word embedding constructed considering context words/review

[Zhang 16] Collaborative multi-Level embedding learning from reviews for rating prediction, IJCAI

Between user/item/review embedding too restrictive

Direct Equivalence

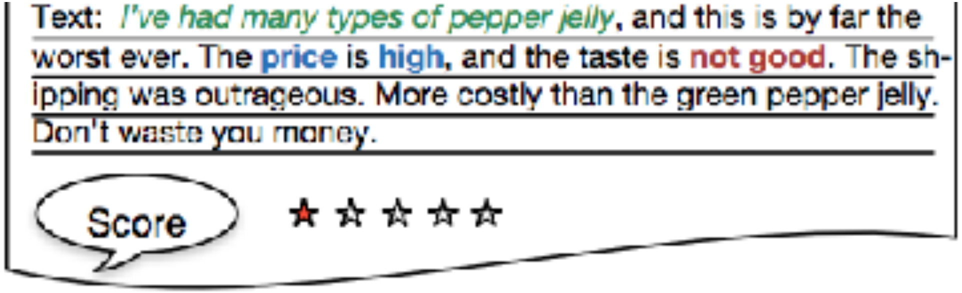
Some words/segments irrelevant to sentiment polarity

- Just play a transitional role (italicized)

Not each topic shown in review contributes to rating

User/item characteristics cannot be fully revealed

- Review length/sparsity



Projection layer connects review representation to user/item

- Thus relaxing restriction

[Zhang 16] Collaborative multi-Level embedding learning from reviews for rating prediction, IJCAI

Flipkart

Captures syntactic/semantic information

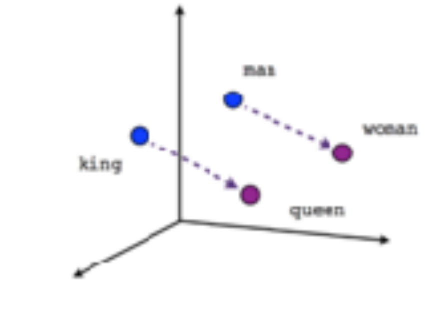
CBOW leverages succeeding/ preceding words

- To predict current word

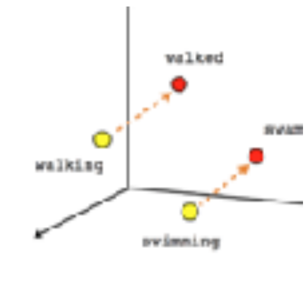
Skip-gram vice-versa

Still need to incorporate intrinsic characteristics

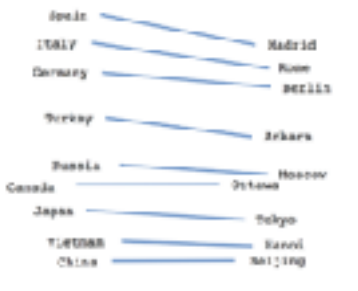
- Among rating behaviors of users



Male-Female



Verb tense



Country-Capital

[Zhang 16] Collaborative multi-Level embedding learning from reviews for rating prediction, IJCAI

Construct review embedding with word embedding

- CBOW to model each word of review text
- Bridges word- with user/item- embedding

Learn user embedding from review embedding/rating

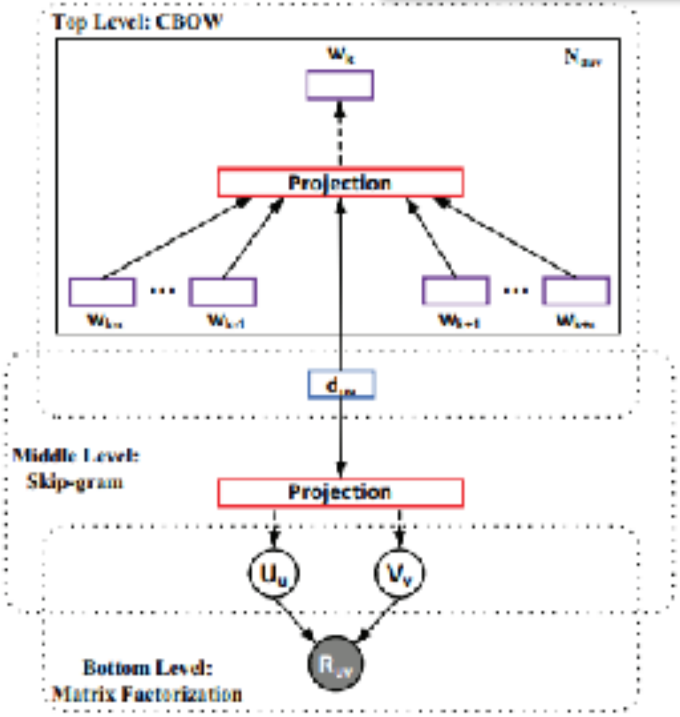
- With a projection similar to skip gram

Inner product vs. direct, normalized exponential transformation

- Relax equivalence restriction

CMLE

Collaborative multi-level embedding



[Zhang 16] Collaborative multi-Level embedding learning from reviews for rating prediction, IJCAI

CNN: text processing

Input: sequence of words

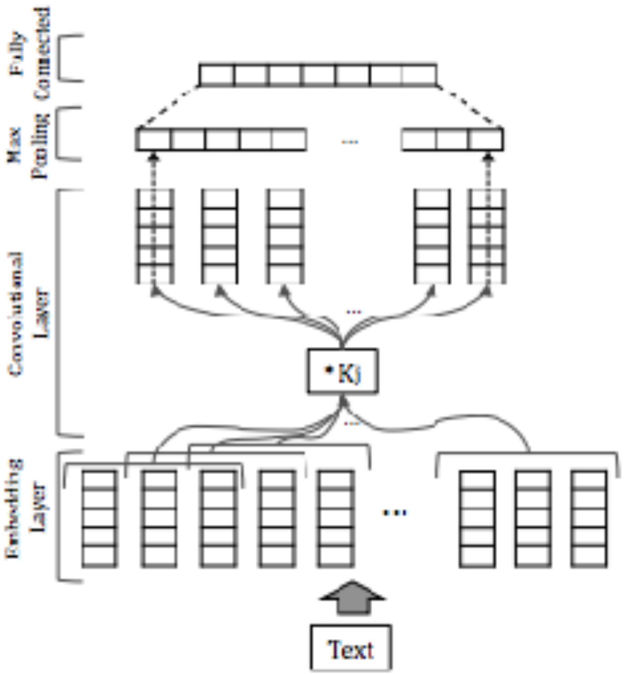
Output: distributed representations

Embedding function maps each word

- In M-sized vocabulary into d-dimensional vector
- Pre-trained (e.g., word2vec on Google News, GloVe on Wikipedia

Max-pooling enables feature detection

- Irrespective of where in text it appears



[Zheng 17] Joint deep modeling of users and items using reviews for recommendation, WSDM

2 parallel neural networks coupled in the last layers for learning

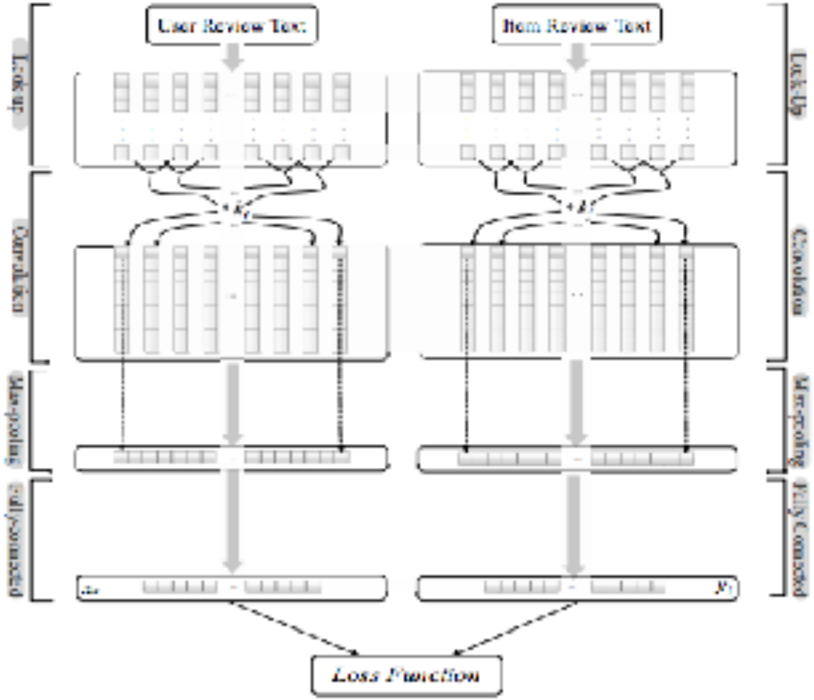
Deep CoNN

Cooperative Neural Networks

- Behaviors with reviews written by user
- Properties with review written for item

Shared layer on top enables latent factors to interact with each other

- For user/items similar to factorization machines



[Zheng 17] Joint deep modeling of users and items using reviews for recommendation, WSDM

Drop-out suppresses some neurons

- Regularize network

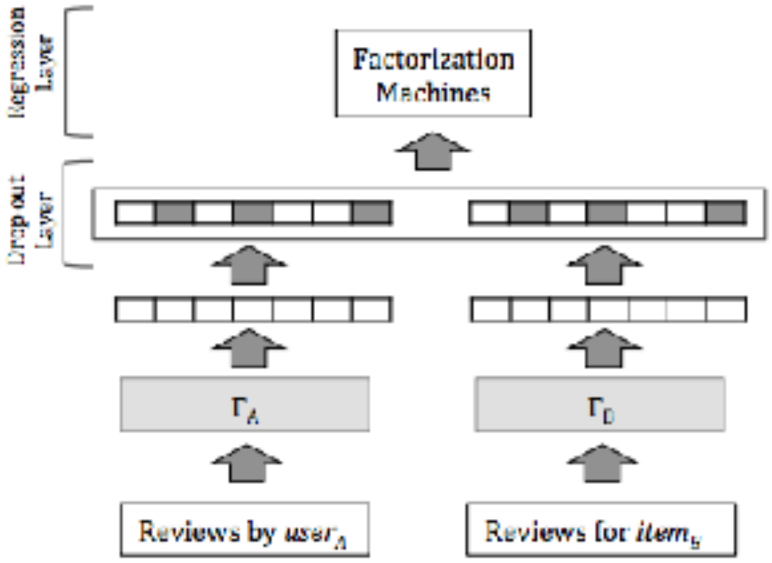
Regression computes interaction

- between input vector elements

Performs better only when target review is available at test time

- Written by target user on target item
- Item recommended to user before experience not reasonable to expect

Deep CoNN: predict rating



[Zheng 17] Joint deep modeling of users and items using reviews for recommendation, WSDM

TransNet

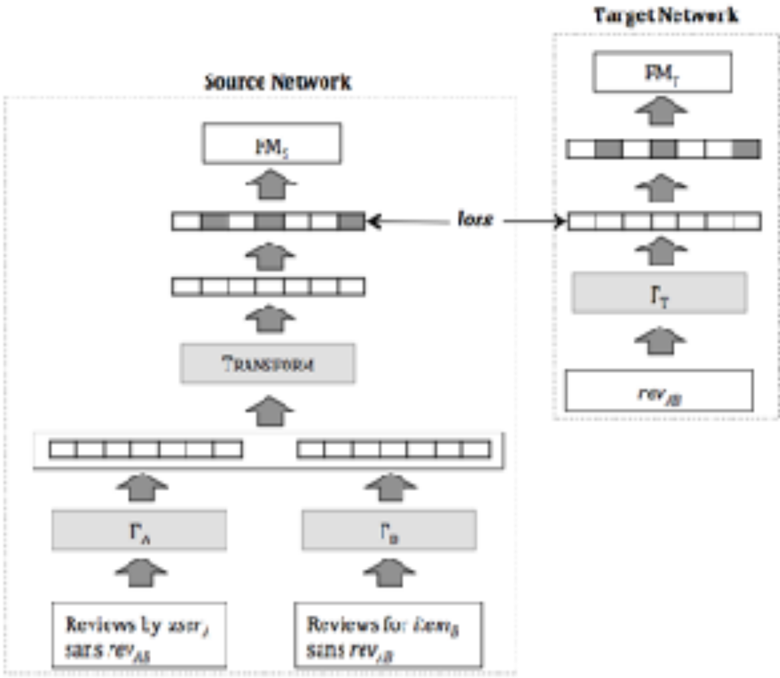
DeepCoNN exploits fact that test reviews were leaked into training set

Augment this neural network with additional multi-task learning

- Reducing recommendation problem to document-level sentiment analysis

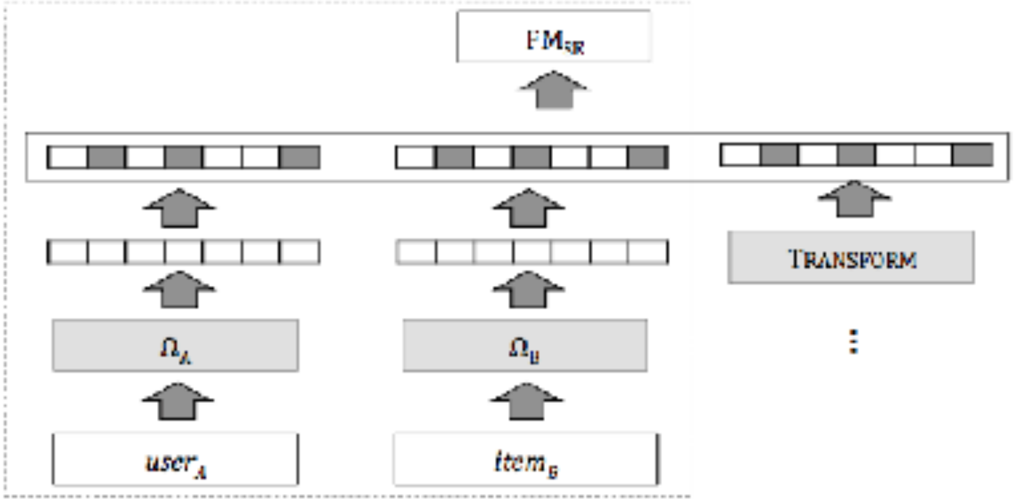
Transform penultimate hidden layer into CNN-encoded representation

- Of target review (useful signal)



[Catherine 17] Transnets: Learning to transform for recommendation, RecSys

Extended TransNet



TransNet identity agnostic

- User/item fully represented using just review text

Recommender systems identity-aware

- Learn latent representation similar to matrix factorization

Source network with new embedding matrices for user/item

- Map string representing id. into n-dimensional representation

[Catherine 17] Transnets: Learning to transform for recommendation, RecSys

Flipkart

Evaluation: Original vs. Predicted Reviews

Original Review	Predicted Review
my laptop flatlined and i did n't know why . just one day it did n't turn on . i cam here based on the yelp reviews and happy i did . although my laptop could n't be revived due to the fried motherboard , they did give me a full explanation about what they found and my best options . i was grateful they did n't charge me for looking into the problem , other places would have . i will definitely be coming back if needed . .	my hard drive crashed and i had to buy a new computer . the store where i bought my computer could n't get data off my old hard drive . neither could a tech friend of mine . works could ! they did n't charge for the diagnosis and only charged $ 100 for the transfer . very happy .
excellent quality korean restaurant . it 's a tiny place but never too busy , and quite possibly the best korean dumplings i 've had to date .	for those who live near by islington station you must visit this new korean restaurant that just opened up . the food too good to explain . i will just say i havent had a chance to take picture since the food was too grat .
this place is so cool . the outdoor area is n't as big as the fillmore location , but they make up for it with live music . i really like the atmosphere and the food is pretty spot on . the sweet potato fry dip is really something special . the vig was highly recommended to me , and i 'm passing that recommendation on to all who read this .	like going on monday 's . happy hour for drinks and apps then at 6pm their burger special . sundays are cool too , when they have live music on their patio .
i have attempted at coming here before but i have never been able to make it in because it 's always so packed with people wanting to eat . i finally came here at a good time around 6ish ... and not packed but by the time i left , it was packed ! the miso ramen was delicious . you can choose from add on 's on your soup but they charge you , i dont think they should , they should just treat them as condiments . at other ramen places that i have been too i get the egg , bamboo shoot , fire ball add on 's free . so i am not sure what their deal is .	hands down top three ramen spots on the west coast . right up there with , and the line can be just as long .

Give user predicted reaction more detailed than rating (quality vs. price)

Predicted reviews talk about particulars which original reviews highlight

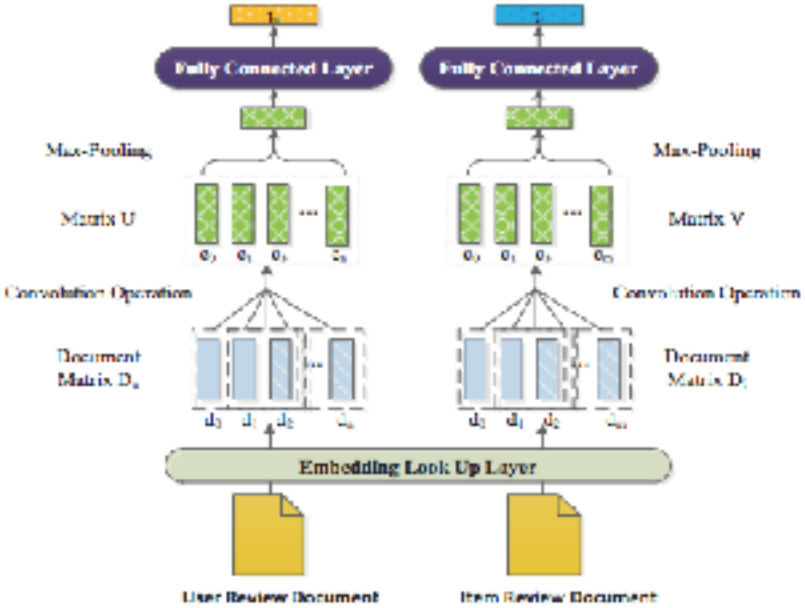
[Catherine 17] Transnets: Learning to transform for recommendation, RecSys

Flipkart

Performance of review-based rating prediction

Deep CoNN

- Robust with few rating records
- Suffers due to lack of reviews



Extracts non-linear latent features

- Hampered when review is incomplete

[Wu 18] PARL: Let Strangers Speak Out What You Like, CIKM

Reviews are sparse

- No comments for purchased products
- Short reviews cannot reflect user interests fully

Leverage users with similar rating to same item

- complementary/fresh informative features uncovered

Reviews by like-minded users

Jayden is a superhero fan
Shawn may rate AntMan higher

[Wu 18] PARL: Let Strangers Speak Out What You Like, CIKM

PARL

Pair-wise dependent features from auxiliary reviews of like-minded users

No domination by popular products

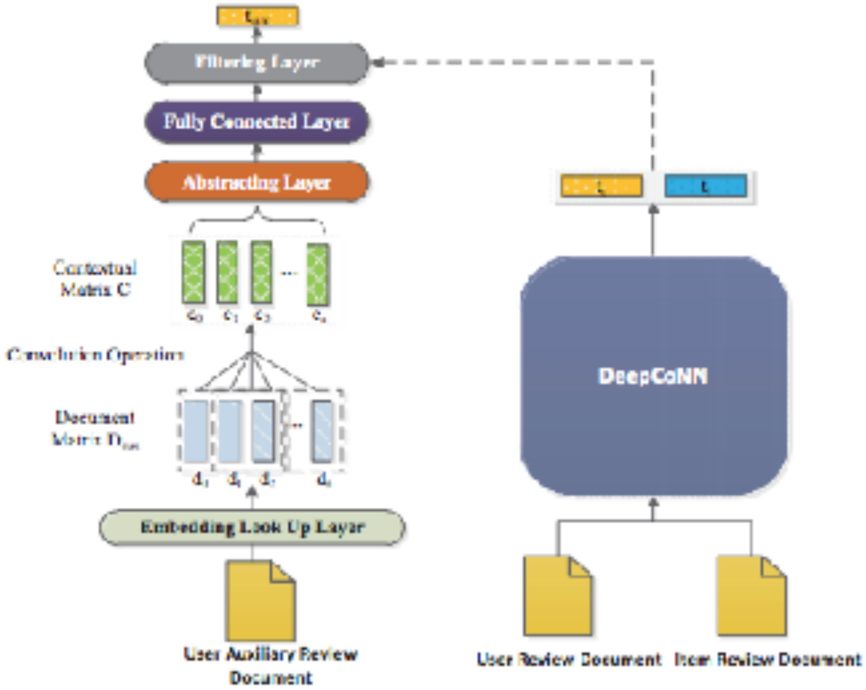
Prefer other users with higher ratings

Not all information in auxiliary review documents is useful

- High rating given for product with aesthetics and practicability
- By both appearance-valuing and pragmatist users

Extra semantic features from heterogeneous reviews

- With no explicit relation to target user
- E.g., describe same Infinity War movie
- Shawn: smart, funny, brave
- Jayden: blood, sweat, tears



[Wu 18] PARL: Let Strangers Speak Out What You Like, CIKM

Flipkart

Evaluation: Target vs. auxiliary reviews

With different ranking ordered by prediction accuracy

Target Reviews	Rank	Auxiliary Reviews
Pours dark brown with a very thin white head that quickly recedes. Aroma is of sweet malt. Taste is also of sweet malt with a lingering sweet finish. A nice, solid beer that I will definately drink again.	1	Amber with small head. Incredibly easy to drink. A little bit of hops. Would make an excellent brew for a night of billiards with friends.
Pours dark brown with a very thin white head that quickly recedes. Aroma is of sweet malt. Taste is also of sweet malt with a lingering sweet finish. A nice, solid beer that I will definately drink again.	2	poured a cloudy amber with an off white head. Had an aroma of citrus, hops and grain. Had a nice bitter finish.
⋮	⋮	⋮
Pours dark brown with a very thin white head that quickly recedes. Aroma is of sweet malt. Taste is also of sweet malt with a lingering sweet finish. A nice, solid beer that I will definately drink again.	6	On tap at Founders. Nice aroma of chocolate, coffee, vanilla, and oak. Great dark appearance with a creamy, everlasting head. Great beer, something I come to expect from Founders!!
⋮	⋮	⋮
Pours dark brown with a very thin white head that quickly recedes. Aroma is of sweet malt. Taste is also of sweet malt with a lingering sweet finish. A nice, solid beer that I will definately drink again.	11	Way overrated on this site. Fortehouses description is very accurate. Maybe this was old cause it does indeed taste like vinegar. The foamy head takes forever to simmer down. If someone knows how to find a good bottle let me know.
⋮	⋮	⋮
Pours dark brown with a very thin white head that quickly recedes. Aroma is of sweet malt. Taste is also of sweet malt with a lingering sweet finish. A nice, solid beer that I will definately drink again.	15	The foam crown does not look so good, in color it is gold-yellow. In the taste hoppy s and a little alcoholic but fits well! Ingenious Saftigkeit, such a Sifligen Bock I did not have! The aftertaste resembles the taste in the mouth.

Informative/incompatible features
Highlighted in orange/grey color

Extraction of useful, semantic features to make rating prediction accurate

- synonymous words, positive sentiment correlation

Noisy information penalizes rank of long reviews

Flipkart

[Wu 18] PARL: Let Strangers Speak Out What You Like, CIKM

Interpret part of image looked at

- While generating a description word

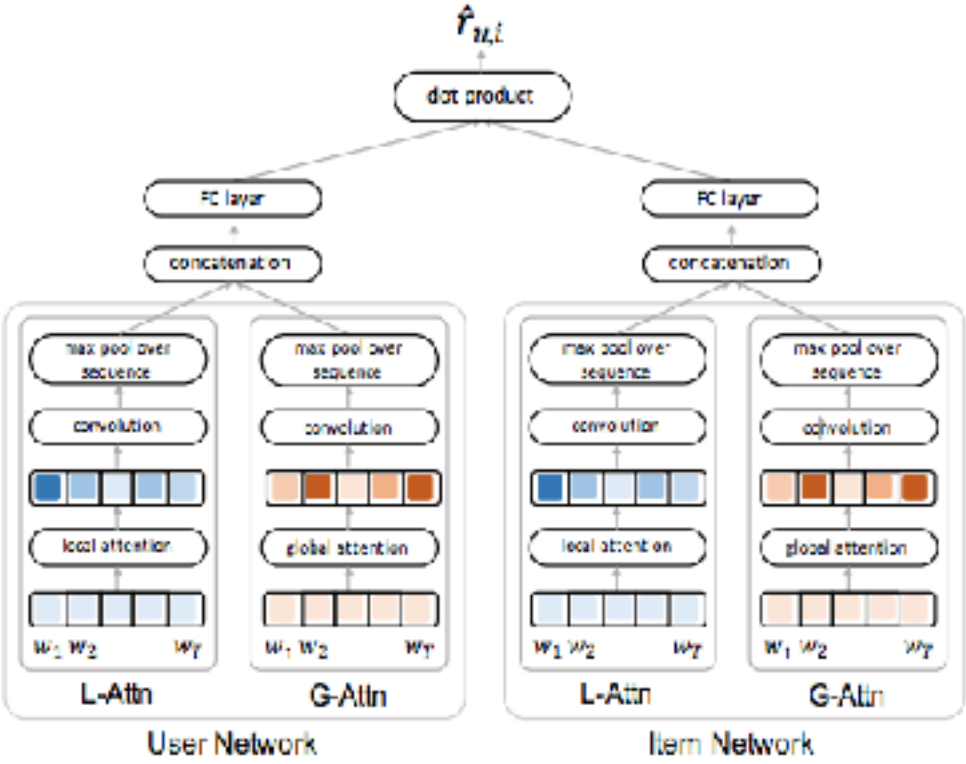
Select informative words

- that contribute to rating
- Local attention

Ignore noisy/irrelevant words from long reviews

- Global attention

D-Attn



[Seo 17] Interpretable convolutional neural networks with dual local and global attention for review rating prediction. RecSys

User preference/item properties

Evaluation: Local score

Word with smaller score is less important

- On individual attributes

Highest-score: dark green

- Consecutive words including trivial pronouns (e.g., good people, great atmosphere)

Higher-score: light green

- Neutral (e.g., Arizona, that)

Yelp (user), L-Attn-only model: local attention

They carry some rare things that you can't find anywhere else. The staff is pretty damn cool too best in Arizona. I prefer ma-and-pa. They treat you the best and they value your business extreme. They are good people great atmosphere and music. I definitely believe that Lux has the best coffee I've ever had at this point. Screw all my previous reviews. This place has coffee down, they make damn good toast too.

Yelp (user), D-Attn model: local attention

They carry some rare things that you can't find anywhere else. The staff is pretty damn cool too best in Arizona. I prefer ma-and-pa. They treat you the best and they value your business extreme. They are good people great atmosphere and music. I definitely believe that Lux has the best coffee I've ever had at this point. Screw all my previous reviews. This place has coffee down, they make damn good toast too.

[Seo 17] Interpretable convolutional neural networks with dual local and global attention for review rating prediction. RecSys

Flipkart

Evaluation: Global score

Focus on semantic meaning

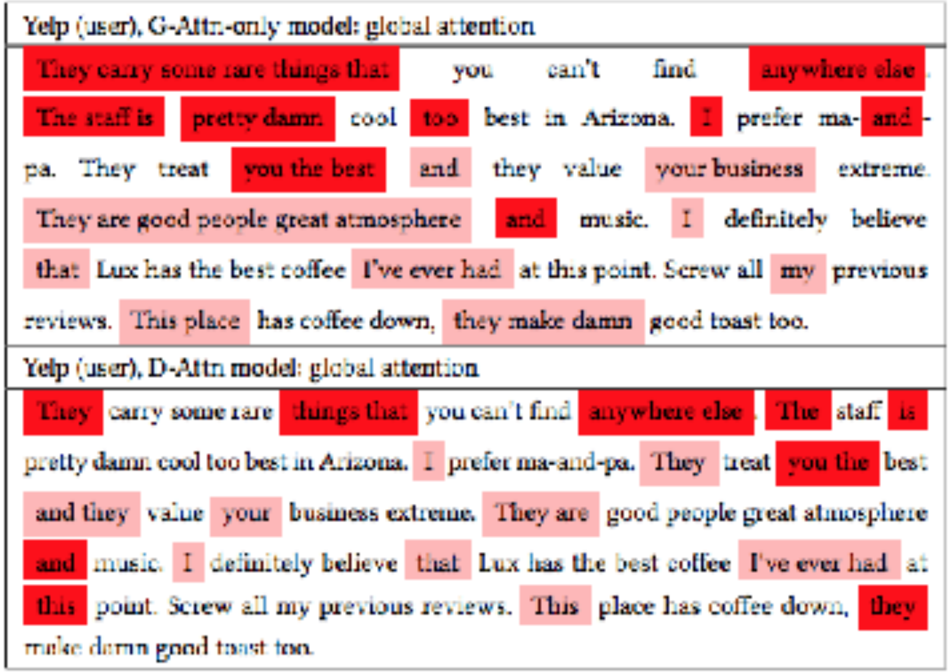
- For whole review text

Effects of uninformative words are diminished

Lowest-score: dark red

- preferences word rare, best grouped with neighboring pronouns

Lower-score: light red



Global removes neutral words whereas local emphasizes preference words

Flipkart

[Seo 17] Interpretable convolutional neural networks with dual local and global attention for review rating prediction. RecSys

Evaluation: Attention words

Highlighted by Local model in different networks

- User: soon, great, unfair, waiter
- Item: disappointing, suffer

More words highlighted in user network

- To show preferences

Yelp (user), D-Attn: local attention
A disappointing meal and a very disappointing service . I won't be coming back anytime soon . If I was the manager I would demote the waiter and promote the busboy to a waiter as he was great tonight and he was the only reason I gave a 20% tip as it is unfair for him to suffer because of the waiters.
Yelp (item), D-Attn: local attention
A disappointing meal and a very disappointing service. I won't be coming back anytime soon. If I was the manager I would demote the waiter and promote the busboy to a waiter as he was great tonight and he was the only reason I gave a 20% tip as it is unfair for him to suffer because of the waiters.

[Seo 17] Interpretable convolutional neural networks with dual local and global attention for review rating prediction. RecSys

Flipkart

Evaluation: Recommended items without review history

Query	Visited items	Rank	HFT Recommendations [Categories]	ConvMF+ Recommendations [Categories]	D-Attn Recommendations [Categories]
user1	**The Duce** [Bars, Nightlife, Lounges, Restaurants]	1	**Yoshi's Asian Grill** [Asian Fusion, Restaurant]	**FnB** [Gastropubs, American]	**O'connor's Pub** [Pubs, Bars, Nightlife]
		2	**Nathans Famous Hotdogs** [Hot Dogs, Restaurants]	**Petite Maison** [French, Restaurants]	**Rosie McCaffrey's** [Pubs, Bars, Nightlife, Restaurants]
		3	**Wendy's** [Fast Food, Restaurants]	**Lox Stock & Bagel** [Bagels, Breakfast & Brunch, Restaurants]	**The Vig** [Pubs, Bars, Nightlife, Restaurants]
		4	**The Coffee Bean & Tea Leaf** [Food, Coffee & Tea]	**True Food Kitchen** [American, Restaurants]	**Arcadia Tavern** [Pubs, Bars, Nightlife, Sports Bars]
user2	**Matador Restaurant** [Mexican, Greek, Restaurants]	1	**The Saguaro** [American, Mexican]	**Rocket Burger & Subs** [Burger, Hot Dogs, Sandwiches]	**Sofia's Mexican** [Mexican, Restaurants]
		2	**The Grapevine** [American, Karaoke]	**Roka Akor** [Steakhouses, Sushi Bars, Japanese]	**Tacos Jalisco** [Mexican, Restaurants]
		3	**Citizen Public House** [Gastropubs, American]	**The Fry Bread House** [American, Restaurants]	**Carolina's Mexican** [Mexican, Restaurants]
		4	**AZ 88** [Bars, American, Lounges]	**Five Guys** [Burgers, Restaurants]	**El Taco Tote** [Mexican, Restaurants]

Nearest-neighbor in embedding space

- Have similar taste as query user

Same category as already visited (e.g., Mexican, Matador)

Common features (price ($), ambience (casual)

[Seo 17] Interpretable convolutional neural networks with dual local and global attention for review rating prediction. RecSys

Mixture of Gaussian Model (GMM)

Imitate rating behavior of users to items

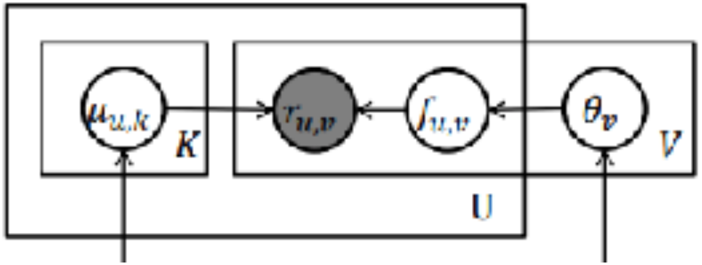
Users consider weights over different item factors

- when giving evaluations

Unknown ratings by weighted summation

Each rating generated from a GMM

- Which models user preference over item factors



Each user/item represented as a mixture of latent variables

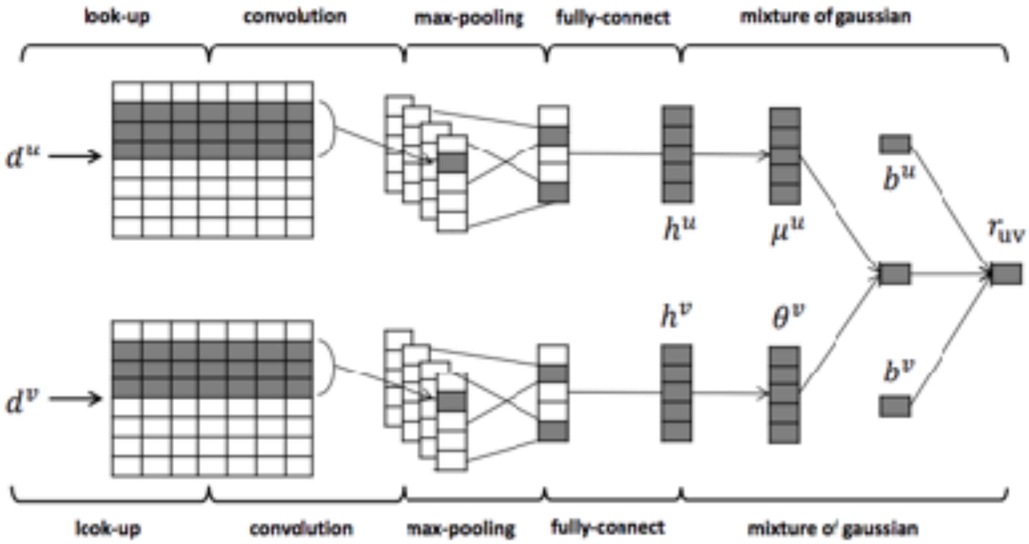
- Expected value of user preference
- Component of an item it belongs

Conditional probability of rating given user/item

- Computed as combination of component-item & rating-gaussian distribution

[Deng 18] Neural gaussian mixture model for review-based rating prediction RecSys

Neural GMM



Gaussian component has zero variance

- Mean described by corresponding component in user's latent vector
- Weight by item's vector

Shared GMM layer on top to simulate model parameters

- Mean and mixture proportion

[Deng 18] Neural gaussian mixture model for review-based rating prediction RecSys

CONCLUSION

4.1 RevGen
CF-GGN [Ni17]
MT [Lu 18]
MRG [Truong 19]

4.2 TipsGen
NRT [Li17]
PATG [Li 19]

3.1 UsefulRev
NARRE [Chen18]
MPCN [Tay 18]

3.2 AspectReco
A3CNF [Cheng18]
ANR [Chin 18]

1.3 TM-Explainable
SULM [Bauman 17]
ALFM [Cheng 18]

2.1 personalisation
DeepCoNN [Zheng 17]
TransNet [Catherine 17]
PARL [Wu 18]
NGMM [Deng 18]

2.2 Cold-start
BoWLF/LMLF
[Almahairi 15]
CMLE [Zhang 16]

1.2 Explainable
EFM [Zhang 14]
TriRank [He 15]

1.1 TM-Coldstart
HFT [McAuley 13]
RnR [Ling 14]

Attention based approaches for recommendation using Reviews

Sudeshna Sarkar

Attention based Recommendation using Reviews

- Reviews are useful source of information as they contain product aspects such as quality, material, colour.
- The task is to select useful reviews and extract information from these reviews to get a better recommendation.
- While most of the models integrate reviews to enhance the performance of the recommendation, considering the contribution of each review helps in better modelling.
- There have been two aspect to using reviews for recommendation:

1. **Measuring Usefulness**: Giving attention to the usefulness of each reviews for rating prediction
2. **Aspect based approach:** Measuring user's preference on different aspects of the item.

Flipkart

Aspect based Recommendation using Reviews

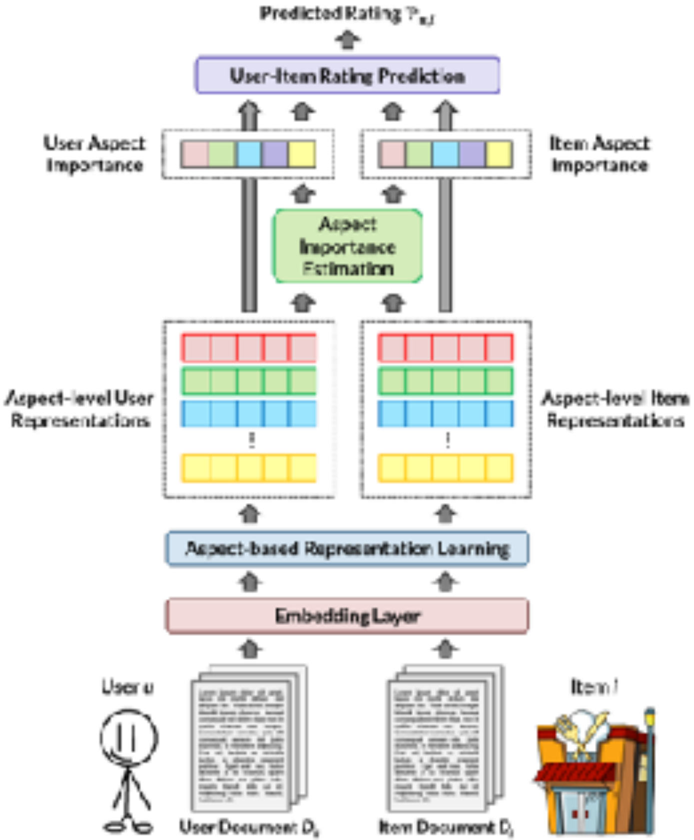
- Each part of the user review focuses on different facet of the user's overall experience, such as *location of the restaurant, taste of dish, attitude of the service staff.*
- However, to model rich semantics of the review content, it's important to go beyond surface level representations. For example;

  (1) The laptop has *long* battery life.

  (2) The laptop requires a *long* startup time.
- Both these sentences bear the word long, yet it is positive sentiments towards the target aspect (battery life) in the first sentence, while it has negative sentiments for the exact same item for a different facet (startup time).
- Similarly one may focus on the food of a particular restaurant and others may focus on the ambience.

ANR: Aspect based Neural Network

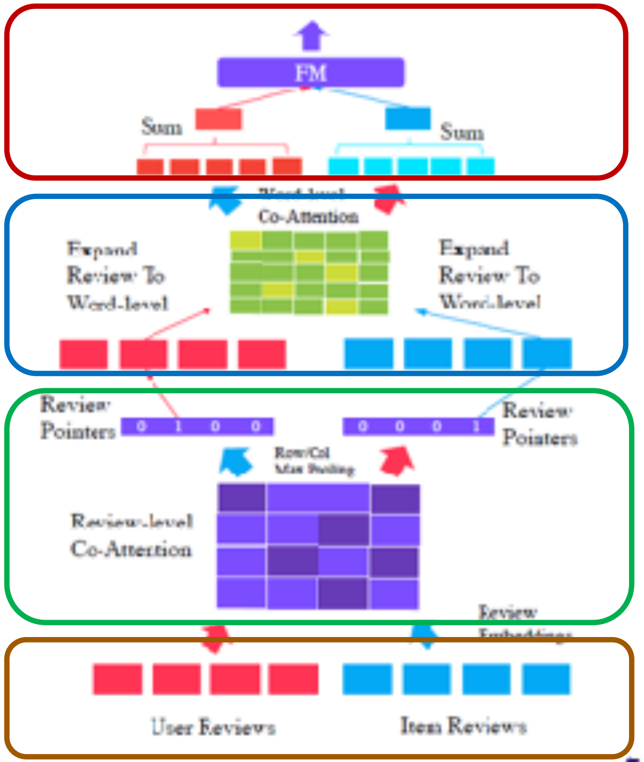
- While all parts are equally important, some choice of the words may reflect different meanings based on the context.
- ANR performs aspect based representation learning for both users and items via attention based components.
- The sentiment bearing words (expensive, delicious, high) of each facets (price, taste, ambience) could be completely different for two different aspect in the same domain. Also, these two sets of words are often in close proximity.



[Chin 18] ANR - Aspect-based Neural Recommender. CIKM

MPCN: Overview

- Input encoding: Use gating mechanism to control the relevance of the reviews
- Review Level Co-attention: To select most informative review from the review bank of each user and item respectively.
- Word Level Co-attention: To model finer granularity at word level we select important words from the useful reviews just as the review level.
- Multi-Point Learning and Prediction: To eliminate noisy reviews, MPCN uses multiple-pointer compositing mechanism. The prediction layer uses Factorization Machine to predict the ratings.



[Tay 18] Multi-pointer co-attention networks for recommendation. KDD

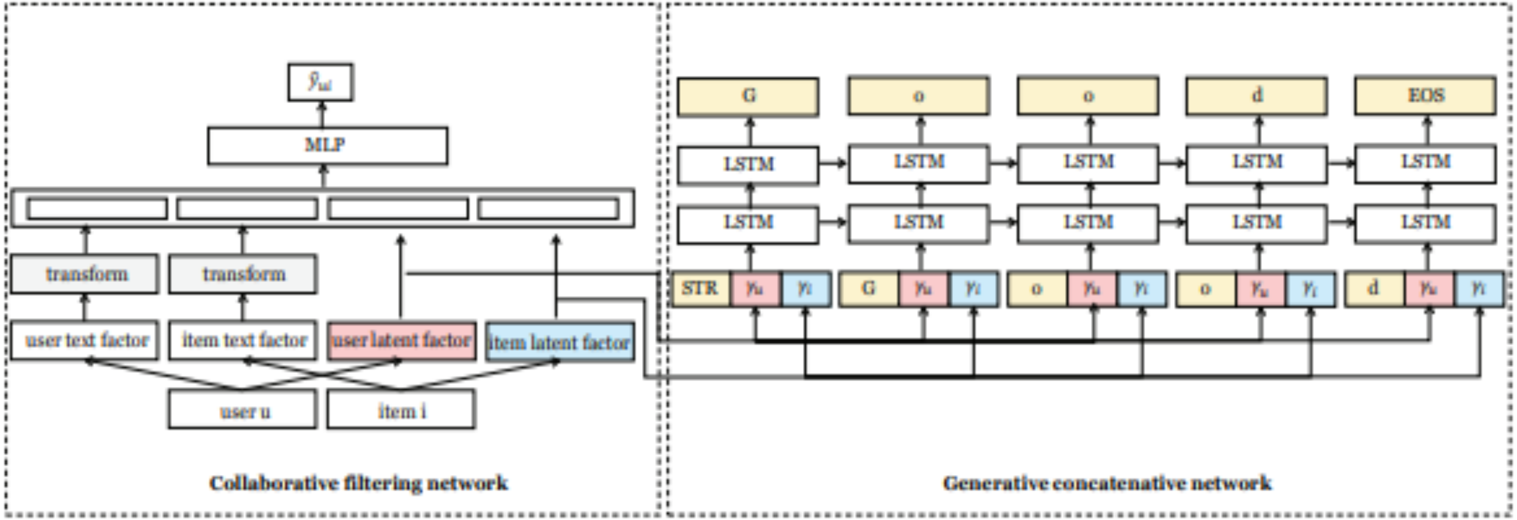
Generative Approaches: Generating reviews/explanations and tips.

- Review text provides valuable information about user and item attributes and interaction between them.
- User written reviews can be viewed as explanations of ratings given by the users.
- Therefore, in an attempt to provide explainable recommendations, there has been a growing interest in the task of review generation.
- The following works will be discussed on this topic.

1. Estimating Reactions and Recommending Products with Generative Models of Reviews

Given an item that a user hasn’t interacted with, the objective is to

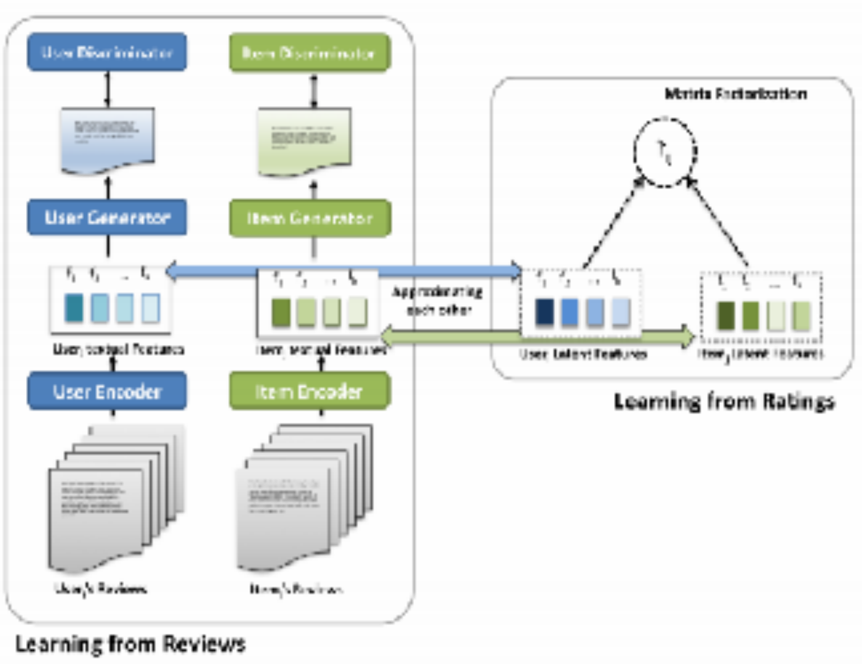
- **Generate a plausible review**, in order to estimate the user’s nuanced reaction



Ni, Jianmo, et al. "Estimating reactions and recommending products with generative models of reviews." *Proceedings of the Eighth International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*. 2017.

2. Why I like it: Multi-task Learning for Recommendation and Explanation

- Design a multi-task learning framework that jointly learns to perform rating prediction and recommendation explanation from user-generated reviews
- Enforces consistency between the suggested recommendation and the provided explanation.
- Employ a matrix factorization model for rating prediction, and a sequence-to-sequence learning model for explanation generation, by generating personalized reviews for a given recommendation, user pair.



Lu, Yichao, Ruihai Dong, and Barry Smyth. "Why I like it: multi-task learning for recommendation and explanation." *Proceedings of the 12th ACM Conference on Recommender Systems*. ACM, 2018.

Examples of Reviews and tips

(1) Really satisfied with this shoes. Fits well actually my foot size was 24cm and this size was closest to it so i doubted whether it will be loose but fits well. The comfort is awesome. Whole foot is comforted with the shoes, no congested feeling at all. The only problem was Amazon it says 999 whereas on shoes its MRP was 949 :(But overall shoes, delivery, everything all good.

(2) I m not happy with material, doesn't look like buy from vero moda

(3) I got the bag yesterday. It looks average to me, and the synthetic leather looks old (though few of my friends commented it looks good, and more expensive than what it cost me). Build quality looks good. Has enough space as stated (edited, I missed one chain earlier). The straps thankfully are wider than my previous backpack, so it is far easy on my shoulders esp. since I walk a couple of kilometres daily with my laptop.

Example reviews from Amazon website

Tips	Rating
(1) Great fit and finish for shower.	5
(2) I selected this radio for myself several years ago and i have found that all claims for it are true.	5
(3) If your looking for a radio for your shower then look no further.	5
(4) Easy to set up stations.	5
(5) Excellent design and quality construction.	5
(6) First one lasted years just bought another one.	5

(a) Tips for the item "Sony Weather Band Shower Radio".

Tips	Rating
(1) Works perfectly in my msi wind.	5
(2) Perfect size for a home office.	5
(3) Excellent player for price.	5
(4) Wonderful docking speaker with full sound.	4
(5) I like it when it not dropping the signal.	4
(6) Works fine in a pinch.	3
(7) Piece of crap do bother.	1
(8) Revised star piece of crap.	1

(b) Tips for different items written by a particular user.

Example tips from [4]

[4] Li, Piji, Zihao Wang, Lidong Bing, and Wai Lam. "Persona-Aware Tips Generation?." In *The World Wide Web Conference*, pp. 1006-1016. ACM, 2019.

Flipkart

Tips Generation Approaches:

- Tips are typically single-topic nuggets of information, and shorter than reviews with a length of about 10 words on average.
- Tips can give other people quick insights, saving the time of reading long reviews.

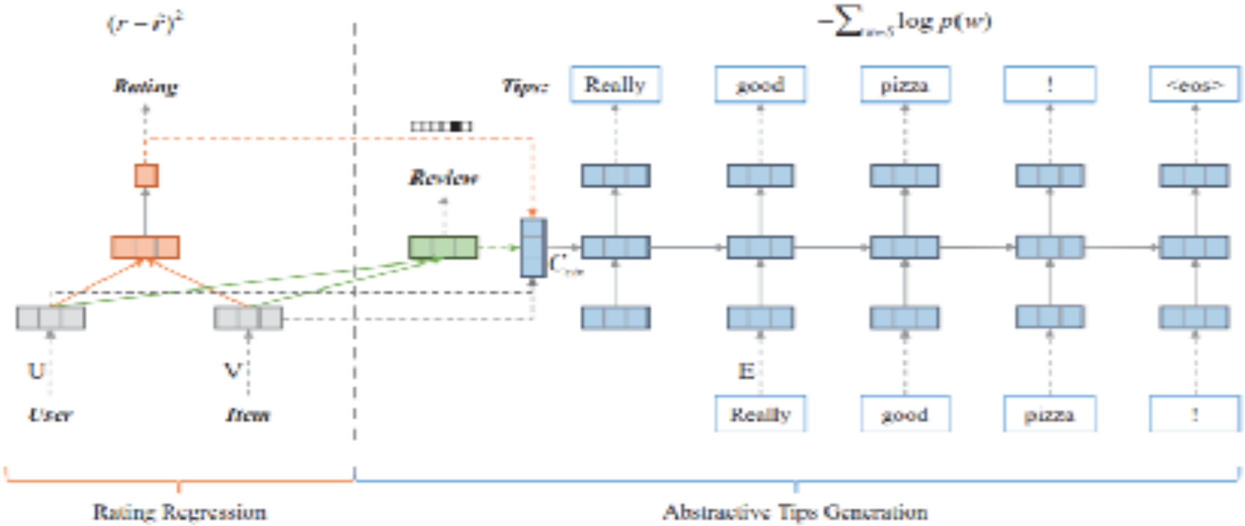
[Li 19] generated persona-aware tips, where persona information such as writing style and vocabulary preference, is considered.

Flipkart

[Li 19] Persona-Aware Tips Generation. WWW.

1. Neural Rating Regression with Abstractive Tips Generation for Recommendation

- Design a deep learning based framework to predict precise ratings and generate abstractive tips.



Li, Piji, et al. "Neural rating regression with abstractive tips generation for recommendation." *Proceedings of the 40th International ACM SIGIR conference on Research and Development in Information Retrieval*. ACM, 2017.

NRR-Tips Generation: Overview

- Major modules: (1) Neural Rating Regression
  (2) Neural Abstractive Tips Generation
- Given user, item vectors , a multi-layer perceptron network based regression model is employed to predict the rating.
- For abstractive tips generation, GRU based sequence decoding model is used to "translate" the combination of a user latent factor u and an item latent factor v into a sequence of words, representing tips.
- Two kinds of context information generated based on u and v are also fed into the sequence decoder model : (1) Hidden variable from the rating regression component, which is used as sentiment context information.
  (2) Hidden output of a generative model for review texts.
- At the time of testing, beam search is used for decoding/generating the best tips given a trained model.

Flipkart