Product recommendations enhanced with reviews

Muthusamy Chelliah

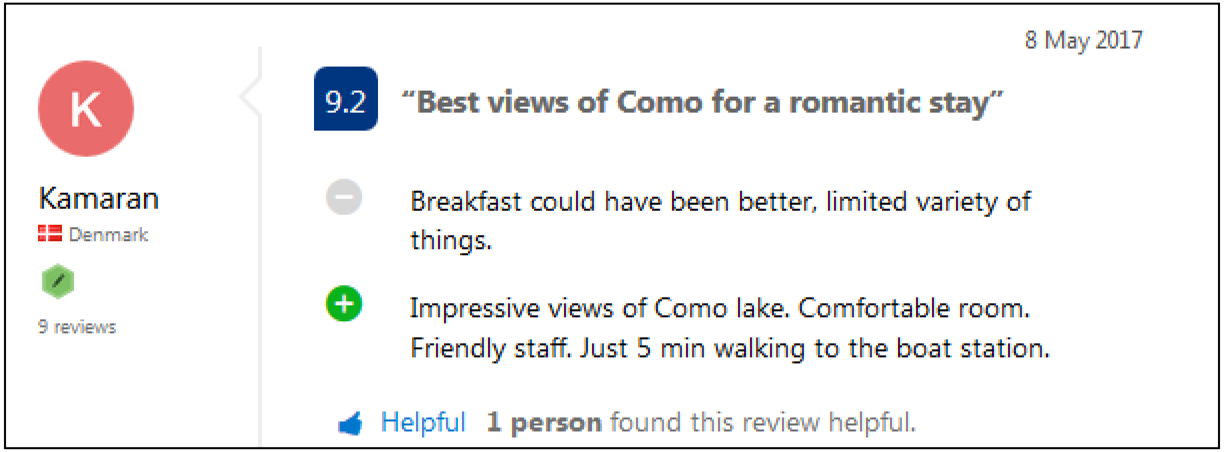
Director, Academic Engagement, Flipkart

Sudeshna Sarkar

Professor, CSE, IIT Kharagpur

Email: sudeshna@cse.iitkgp.ernet.in

Reviews



Which aspect is more important?

Food or Location?

What value of the aspect is desired?

Location near the lake or

Location near the conference venue?

Sell on Flipkart | Advertise | Gift Card | Download App | 24x7 Customer Care | Track Order | Signup | Log In

Flipkart

Search for Products, Brands and More

CART 0

15,900 + ₹798 = ₹16,698

k to top

CART

ADD TO CART

BUY NOW

4.1 ★

42,844 Ratings

10,516 Reviews

5 ★		22,331
4 ★		10,714
3 ★		3,927
2 ★		1,620
1 ★		4,252

Rate and Review product

Most Helpful | Positive | Negative | Most Recent | By Certified Buyers

4 ★ **Decent Performance**

Pros:
- Decent camera - good in day light
- Battery lasts a business day with full 4G on
- Fingerprint scanner works well

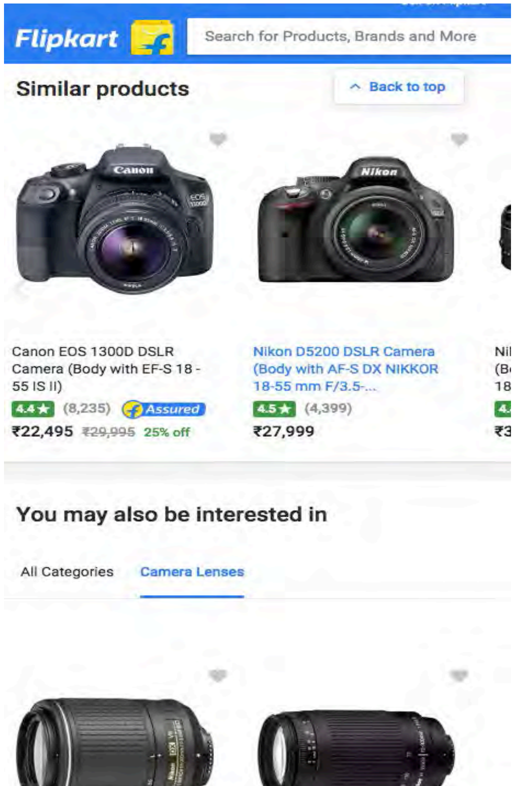
Cons:
- Occasionally hangs
- Bit expensive for these feature set. Worth buy at 15k.

Ram N Certified Buyer 19 Dec, 2016 1729 200

2 ★ **Expected a better product**

Specs n performance r ok ...but d phone is too fragile ...a single drop from a vry height caused its display to shatter ... d display cost me a lil under 8k to replace at

Product Recommendation



- Suggest new products for current selection
  - Substitute
  - Complementary
- Help offer
  - Up-sell
  - Cross-sell
  - Bundle

Rich content in user reviews

About Looks : Very very impressive like a premium phone but back side is not so attractive, it's look a like Redmi Series Phone's as well as back camera edge style is boring.

I'm not impress with the picture quality coz its not what Samsung provide even in his low range phones, if you compare with moto series phones i will give 10/7 to samsung and moto 10/9, also the shake reduction rate on this phone is very bad, After taking the picture (naturally our hand shake little bit) the picture is blurry or shaked.

Sound : Not for who listening the music on speakers, While ringing it's good but not excellent

Well I'm shocked in packaging too as soon as i see the samsung box it's look they given me a 2 year old box

User Modeling

Recommendation Generation

Flipkart

Review Enhanced Recommendation

1. Introduction
2. Review mining
    a) Background: Features and Sentiment, Latent features
    b) Statistical approaches. LDA
    c) Recent topic models
    d) Deep Learning (ABSA)
    e) Flipkart ABSA/Recommendation
3. Review based Recommendation
    a) Handling Data Sparsity
    b) Topic Model, Matrix Factorization and Mixture Model based Recommendation
    c) Deep learning based Recommendation
4. Explainable Product Recommendations
5. Summary, Future Trend

Recommender System

Reviews

INTRODUCTION TO RECOMMENDER SYSTEMS

The Recommendation Task

- Rating Prediction: Estimate a utility function to predict how a user will like an item
- Recommendation: Recommend a set of items to maximize user's utility

Recommender Systems

Data

- Content
  - Product Information
    - Product Taxonomy
    - Product Attributes
    - Product Description
  - Customer Demographics
- User Activity
  - Purchase
  - Rating, Preference
  - Click / Browse
  - User generated reviews
- Social data

Methods

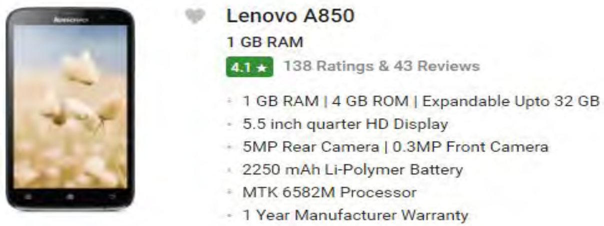
- Content based
  - Recommend items similar to those a user has liked in the past
  - Issues: Does not use quality judgments of users
- Collaborative filtering
  - Finds users with similar tastes as target user to make recommendation
  - Use ratings
  - Issues: Sparsity, Cold Start

Content based recommender systems

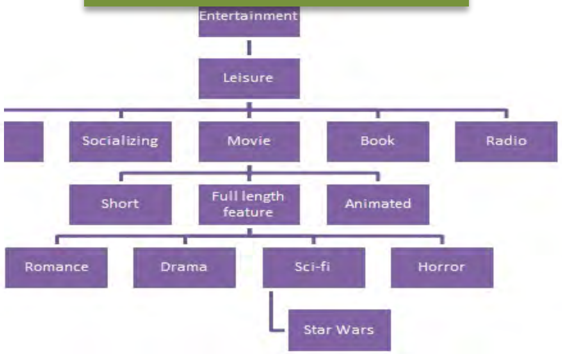
- Each item defined by a profile vector obtained from the item's textual description, metadata, etc.
- Weighting methods such as TF-IDF

- User profile vector : by aggregating the profile vectors of the items that the user liked or purchased in the past.
- Recommend items with high similarity between user vector and item vector.

Product description



Product Ontology



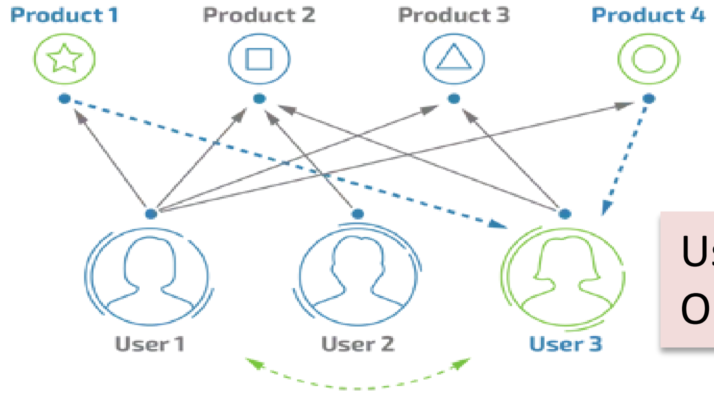
User Review

Flipkart

Collaborative filtering

Predict user preferences on new items by collecting rating information from many users (collaborative).

- Many items to choose from
- Few data per user
- No data for new user



Difficulties

- Data sparsity
  - too few ratings
  - 90%+ items with less than 10 reviews
- Cold-start
  - New user/item appear continuously

User Reviews provide additional rich data.
One Review worth much more than one rating.

Flipkart

Collaborative filtering

<u>Memory Based CF</u>:

1. User-based CF
2. Item-based CF

<u>Model-based CF</u>:

- Train a parametric model with the rating matrix
- Applied to predict ratings of unknown items

o Ex: latent factor model

Latent Factor Model: Matrix Factorization

- Map both users and items to a joint latent factor space
- item i : associated vector q_i
- user u : vector p_u
- ${q_i}^T p_u$ captures the interaction between u and i
- This approximates r_{ui} : user u's rating of item i

- Factor rating matrix using SVD: obtain , ,
- Reduce the matrix to dimension
- Compute two resultant matrices: and
- **Predicting task**

Latent Factor Models

- Goal: to find Q and P such that:

$$\min_{P,Q} \sum_{(i,u)\in R} \left(r_{ui} - q_i \cdot p_u\right)^2$$

users (columns) × items (rows):

1		3			5			5		4	
		5	4			4			2	1	3
2	4		1	2		3		4	3	5	
	2	4		5			4			2	
		4	3	4	2					2	5
1		3		3			2			4	

≈

Q: items (rows) × factors (columns):

.1	-.4	.2
-.5	.6	.5
-.2	.3	.5
1.1	2.1	.3
-.7	2.1	-2
-1	.7	.3

P^T: factors (rows) × users (columns):

1.1	-.2	.3	.5	-2	-.5	.8	-.4	.3	1.4	2.4	-.9
-.8	.7	.5	1.4	.3	-1	1.4	2.9	-.7	1.2	-.1	1.3
2.1	-.4	.6	1.7	2.4	.9	-.3	.4	.8	.7	-.6	.1

Dealing with Missing Entries

1	3	4			
	3	5			5
		4	5		5
		3			
		3			
2			?		?
				?	
	2	1			?
	3			?	
1					

- To address overfitting, introduce regularization:

$$\min_{P,Q} \underbrace{\sum_{\text{training}} (r_{ui} - q_i p_u)^2}_{\text{error}} + \underbrace{\left[\lambda_1 \sum_x \|p_u\|^2 + \lambda_2 \sum_x \|q_i\|^2\right]}_{\text{length}}$$

$\lambda_1, \lambda_2 \ldots$ user set regularization parameters

Why use reviews?

- Reviews provide rich information.
- Reviews help to explain users' ratings.
- Reviews are useful at modeling new users and products: one review tells us much more than one rating.
- Reviews help us to discover the aspects of users' opinions
  - Modeling these factors improve recommendation

- Value for customer
  - Help explore space of options
  - Discover new things

- Value for company
  - Increase trust and loyalty
  - Increase conversion
  - Opportunities for promotion, persuasion
  - Obtain more knowledge about customers

The Value of Reviews

- Recommender systems played an important part in helping users find products.
- Reviews were used by users but only relevant ones.
- Any algorithmic recommender needs to communicate results to the user in a meaningful manner.

Case Amazon: Ratings and Reviews as Part of Recommendations
Juha Leino and Kari-Jouko Räihä, RecSys 2007

- The volume of reviews affects product conversion (upto 270% increase)
- High priced items have a higher value for reviews.

The Value of Online Customer Reviews
Georgios Askalidis, Edward C. Malthouse
RecSys 2016

The importance of Explanations

- Explanations improve user engagement and satisfaction with a recommender system.
- Controllability and transparency.
- Why are these items recommended?

Reviews can help

- Trust
  – user's confidence
- Persuasive
  – Convince users to try/buy
- Satisfaction
  – user's enjoyment
- Help compose judgment
  – attributes to a rating

Product recommendations enhanced with reviews Part I

Muthusamy Chelliah

Director, Academic Engagement, Flipkart

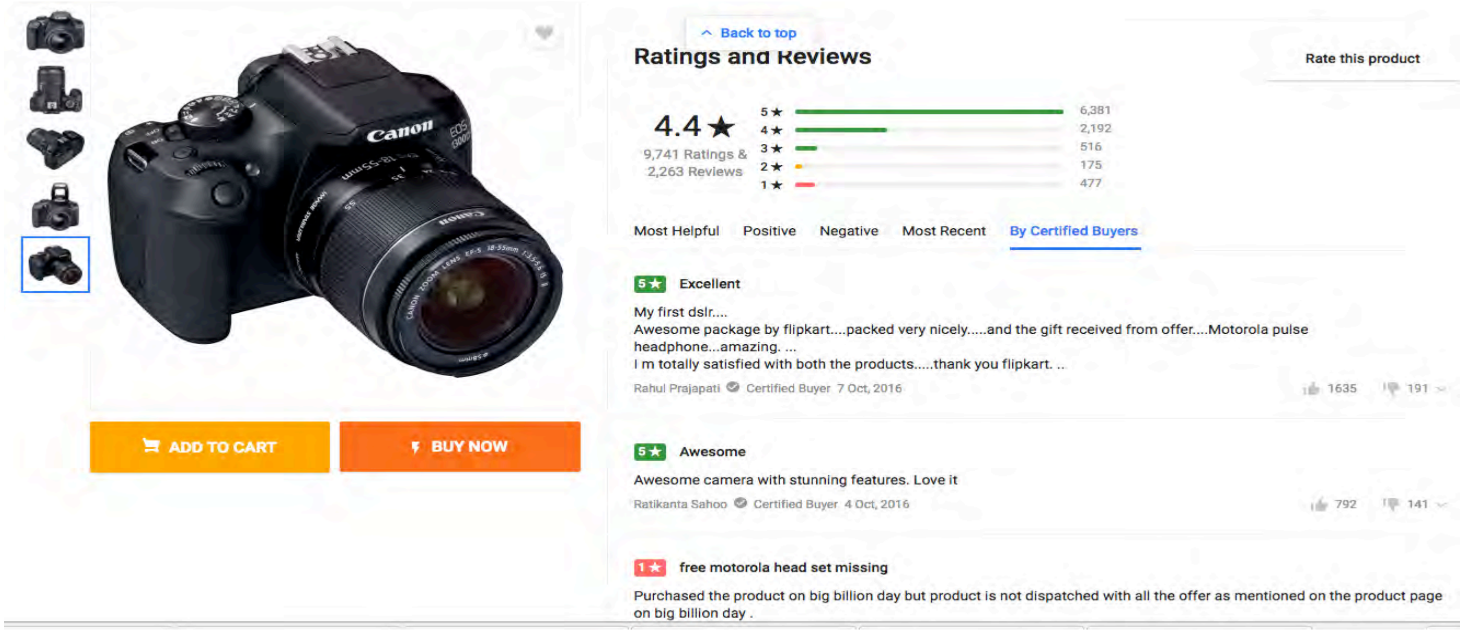
Sudeshna Sarkar

Professor, CSE, IIT Kharagpur

Email: sudeshna@cse.iitkgp.ernet.in


I GOT A SMILEY FACE STICKER ON MY QUIZ.
WHOOP-DE-DOO FOR YOU.
WHAT DID YOU GET?
NONE OF YOUR DARN BUSINESS!
YOU GOT A FROWNY FACE, DIDN'T YOU?
NO, AS A MATTER OF FACT, I DIDN'T!
I DIDN'T EVEN KNOW THEY MADE BARFING FACE STICKERS.

Product Reviews



Product specification

ADD TO CART

BUY NOW

Highlights

- 18 megapixel APS-C CMOS sensor & DIGIC 4+, Wi Fi and NFC supported, 9 point AF with 1 centre cross type AF point
- Effective Pixels: 18 MP
- Sensor Type: CMOS
- WiFi Available

Services

- 2 Years Canon India Warranty ?
- 10 Days Replacement Policy ?
- Cash on Delivery available ?

Seller RetailNet

View 5 sellers starting from ₹23,499

Specifications

In The Box

EF-S18-55mm f/3.5-5.6 IS II EOS 1300D Body + Eyecup Ef + Body Cap Battery Charger LC-E10E Battery Pack LP-E10 Interface Cable Wide Strap EW-400D

General

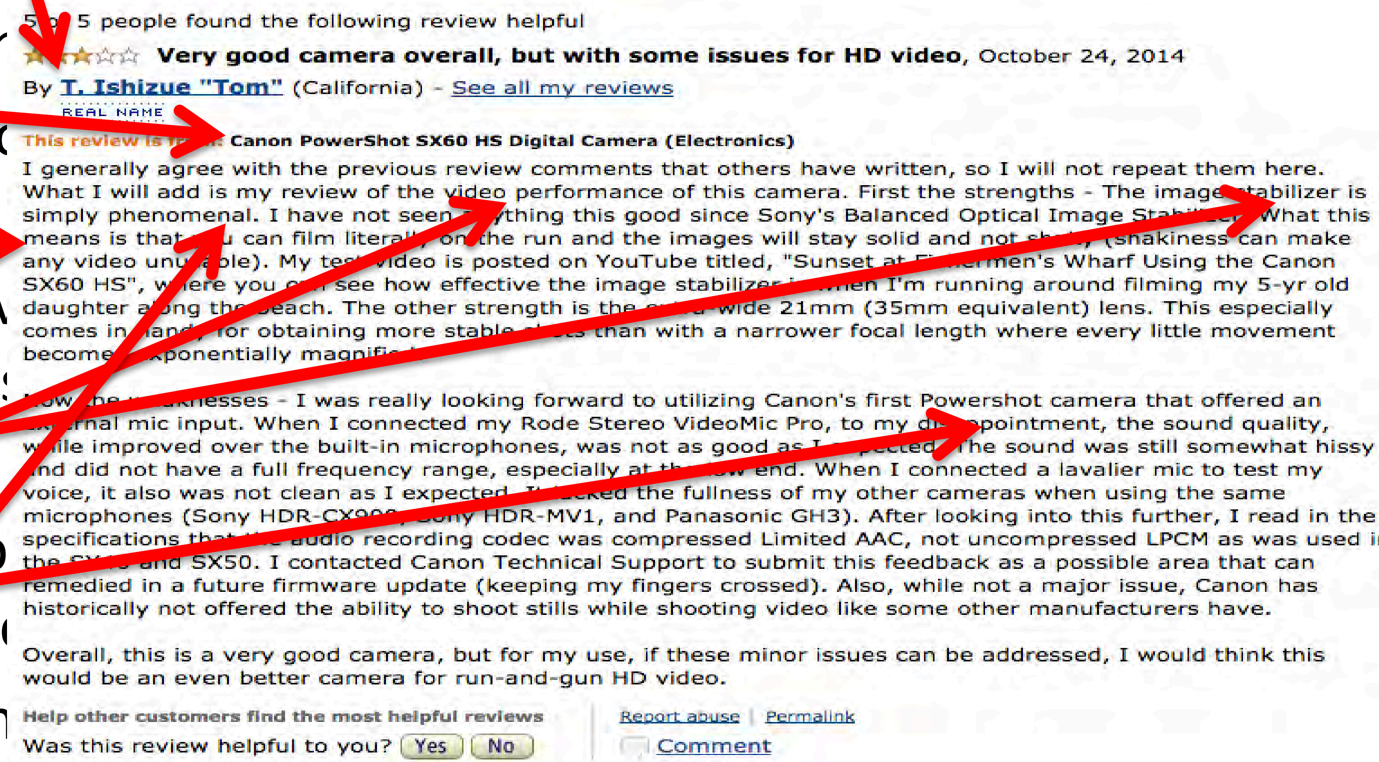
Brand	Canon
Model Number	EOS 1300D
Model Name	Canon EOS 1300D
SLR Variant	(Body with EF-S 18 - 55 IS II)
Brand Color	Black
Type	DSLR
Color	Black
Effective Pixels	18 MP
Shooting Modes	Basic Zone Modes: Scene Intelligent Auto, Flash Off, Creative Auto, Portrait, Landscape, Close-up, Sports, Food, Night Portrait, Creative Zone Modes: Program AE, Shutter Priority AE, Aperture Priority AE, Manual exposure
Tripod Socket	Yes
Wifi	Yes
Face Detection	Yes

Definitions

- **Contributor:** The person or organization who is expressing their opinion in text.
- **Object:** An entity which can be a product, service, person, event, organization, or topic.
- **Review:** A contributor-generated text that contains their opinions about some **aspects** of the object.
- **Aspect:** The component, attribute or feature of the object that contributor has commented on.
- **Opinion:** An opinion on an aspect is a positive, neutral or negative view, attitude, emotion or appraisal on that aspect from a contributor.

Definitions

- **Contributor:** Contributor is the person or organization who is expressing h
- **Object:** An
  person, eve
- **Review:** Rev
  the opinion
  object.
- **Aspect:** Asp
  object that
- **Opinion:** An
  negative view, attitude, emotion or appraisal on that aspect from a contributor.



Flipkart

Terminologies

- Features / Attributes / Aspect / Properties of a product or service
- Opinion
  - Overall opinion
  - Feature opinion
- Review helpfulness:
  - the number of “helpful" votes given by readers
  - Can be used to determine quality score

Aspect-based sentiment analysis (ABSA)

[Moghaddam '10]

Input

Canon GL2 Mini DVD Camcorder
... excellent zoom ... blurry lcd ... great picture quality ... accurate zooming ... poor battery ... inaccurate screen ... good quality ... affordable price ... poor display ... inadequate battery life... fantastic zoom ... great price ...

Output

Aspect	Head Terms	Rating
zoom	zoom, zooming	5
price	price	4
picture quality	picture quality, quality	4
battery life	battery life, battery	2
screen	screen, lcd, display	1
...		...

Prediction

- [Liu '16]
- screen is clear and great
- **Aspect extraction**: screen
- **Opinion identification**: clear, great
- Polarity classification: clear is +ve
- Opinion separation/generality: clear (aspect-specific), great (general)
  - Understand consumer taste of products

Flipkart

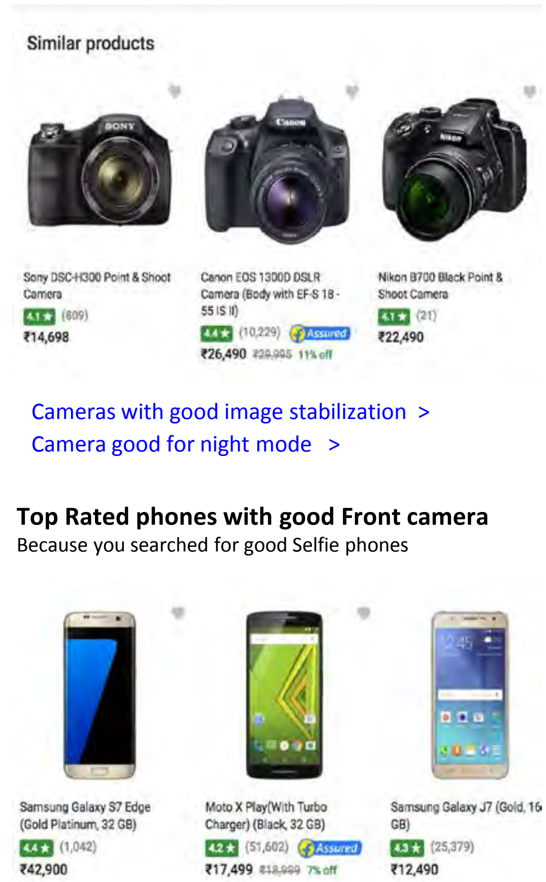
FLIPKART CUSTOMER ASKS

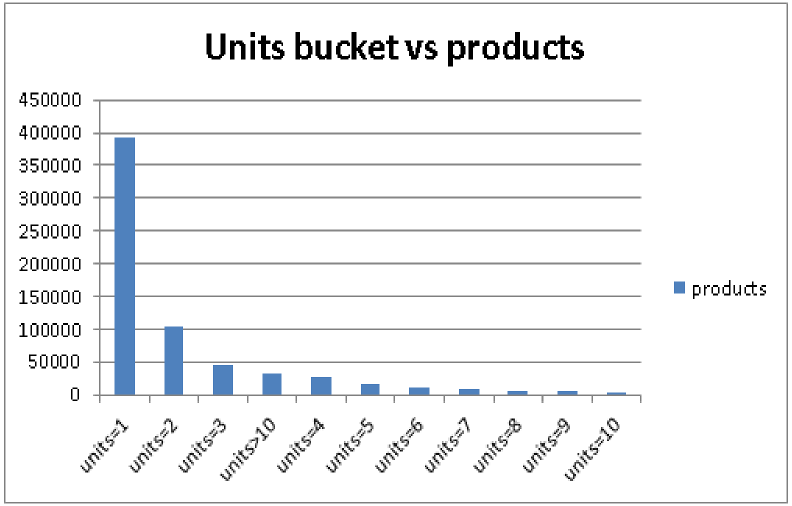
Recommendation: User Needs

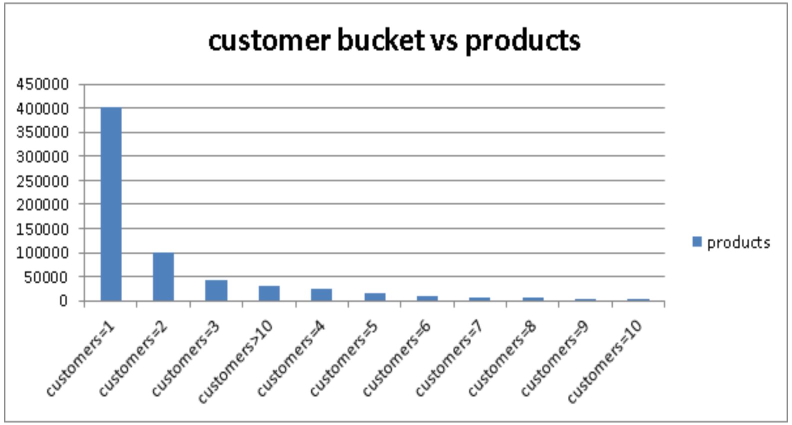
- **Help me narrow down on my choice**
  - I’ve shortlisted few cameras and need help deciding one
  - I’m almost leaning towards buying this speaker system but want to get further confidence and assurance on sound quality.

Flipkart Ratings-based Recommendations

- Product discovery with Rating as a signal
  - rank products with high quality ratings at the top
- Highlight Ratings to narrow down purchase
  - show ratings attributes that map to user intent







Long tail (both in users/products):
induces sparsity/cold-start

Recommendations Leveraging Reviews

- Good coverage
- High quality and detailed reviews
- Parsing of reviews into meaningful clusters for Discovery
  - Picture Quality (4.5), Sound (3.8), Ease of Installation (2.5)

ABSA - User needs

- Individual preferences
  - easy way to shortlist products matching a criteria
  - compare two products on certain aspects
- Too many reviews to read
  - interested in only key elements of a product
  - be aware of product shortcomings before making the final buying decision

ABSA - Partner Needs

- Refine available selection per user preference
- Brands can validate future product design
- Correlation between aspect and sales is a good indicator of market demand

ABSA - agenda

- **Statistics-based text mining (5 minutes)**
- Topic models
  – Early (5 minutes)
  – Recent (10 minutes)
- Deep learning (10 minutes)

Research landscape – ABSA

Aspect extraction, Rating prediction
[Moghadddam '10]
Normalization [Guo '09][Bing '16]

Statistical text mining

ILDA [Moghaddam '11]
MG-LDA, JST, Maxent-LDA,
ASUM, CFACTS:overview
[Liu '11, Moghaddam '12]

Early topic models

HASM [Kim '13],
SpecLDA [Park '15]
LAST [Wang '16]

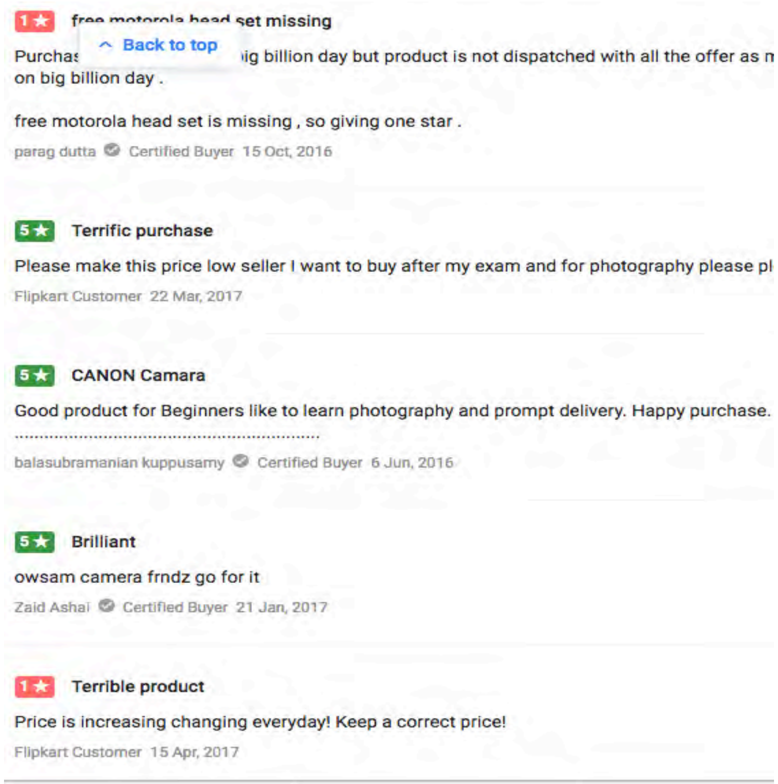
Recent topic models

Rating prediction
- RecursiveNN [Socher'13]

Aspect extraction
- CNN [Lu '14]
- RecurrentNN [Liu '15]
- LSTM/attention [Tang '16]

Deep learning

Rating prediction



Flipkart

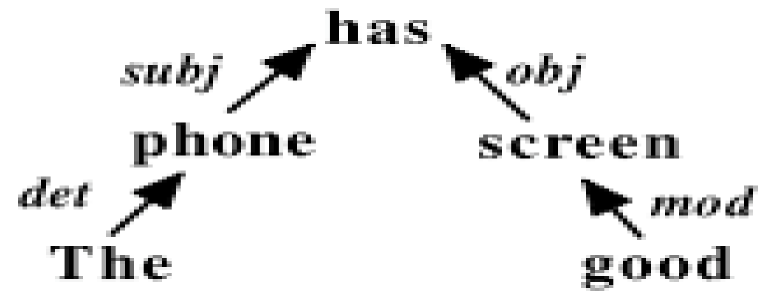
[Moghaddam ’10]

- Estimate score [1-5]
  - With kNN algorithm
  - Wordnet for similarity
- Aggregate rating
  - aspect-level estimate

Aspect extraction

- Frequency-based
  - Extract frequently occurring nouns based on seed words [Moghaddam ‘10]
- Pattern mining
  - Rules for identifying templates of POS-tagged tokens

    (e.g., _JJ_ASP)
- Clustering
  - Group aspect nouns [Guo ‘09]



Category	Product-feature Terms
Battery	*battery life; Lithium battery; AA Alkaline; AA batteries; battery charger; battery capability; battery pack; battery adapter; AA Lithium batteries; rechargeable battery;*
Memory	*memory; flash card; flash memory; memory card; memory capacity; SD card; sd memory; camera flash; digital memory; CF memory card;*

Attribute-aspect normalization

	Concept 1	Concept 2	Concept 3
	wireless, network, wi-fi, internet, connection	screen, resolution, display, brightness, monitor	battery, life, 6-cell, charge, capacity
Product 1	"draft-n wi-fi network," "the integrated bluetooth connectivity"	"10.1 inches ultrawide display," a superbright 10.1-inch lcd	"extended battery life," "the traditional lithium battery"
Product 2	"fast ethernet connection," "the bluetooth 2.0+edr technology"	"normal tft lcd," "10.1 inches widescreen"	"65 hours battery life," "a 6 cells battery"
Product 3	"ultimate wireless accessibility," "wired ethernet lan"	"widescreen trubrite display," "a 1024 × 600-pixel resolution"	"long battery life," "6-cell lithium-ion battery"

The second row shows the top-five weighted terms in the popular features discovered from the customer reviews. Each cell of the table contains the popular attribute text fragments extracted from product description pages.

[Bing '16]

- Extract product attributes from description page
- Infer aspect popularity in reviews/map to feature
- Hidden CRF bridges vocabulary gap

References

- Bing, L., Wong, T.L. and Lam, W. Unsupervised extraction of popular product attributes from e-commerce web sites by considering customer reviews. ACM TOIT '16.
- Guo, H., Zhu, H., Guo, Z., Zhang, X., & Su, Z. (2009, November). Product feature categorization with multilevel latent semantic association. CIKM 2009.
- Moghaddam, S., & Ester, M. Opinion digger: an unsupervised opinion miner from unstructured product reviews. CIKM 2010

ABSA - agenda

- Statistics-based text mining
- Topic models
  - **Early (5 minutes)**
  - Recent (10 minutes)
- Deep learning (10 minutes)

Topic Modeling: Latent Dirichlet Allocation (LDA)

- Given a document set D, assume for each , there is an associated K-dimensional topic distribution , which encodes the fraction of words in d that discuss each of the K topics. That is, words in document d discuss topic k with probability .
- Each topic k has an associated word distribution , which encodes the probability that a particular word is used for that topic k.
- and are random vectors each has a Dirichlet prior distribution.

Topic 1 to K

Document d $\left[\theta_{d,1}, \theta_{d,2}. \dots, \theta_{d,k}, \dots, \theta_{d,K}\right]$

Topic k

$$\begin{bmatrix} \phi_{k,1} \\ \phi_{k,2} \\ \vdots \\ \phi_{k,j} \\ \vdots \\ \phi_{k,N} \end{bmatrix}$$

Word j

Aspect-based sentiment analysis (ABSA)

[Moghaddam '10]

Input

Canon GL2 Mini DVD Camcorder
… excellent zoom … blurry lcd … great picture quality … accurate zooming … poor battery … inaccurate screen … good quality … affordable price … poor display … inadequate battery life… fantastic zoom … great price …

Output

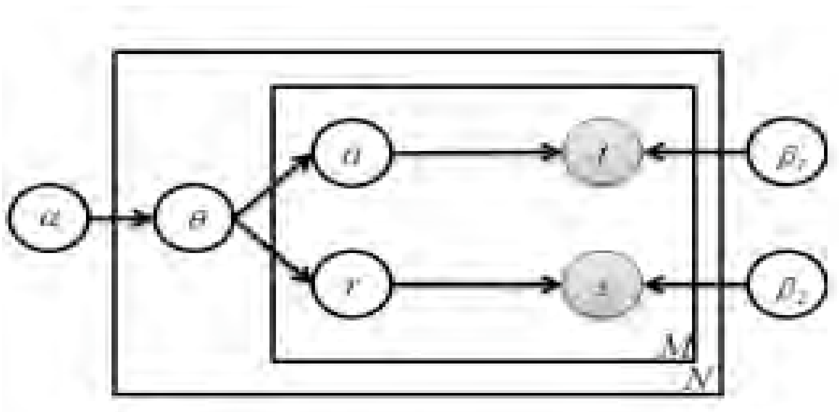
Aspect	Head Terms	Rating
zoom	zoom, zooming	5
price	price	4
picture quality	picture quality, quality	4
battery life	battery life, battery	2
screen	screen, lcd, display	1
…		…

Prediction

- [Liu '16]
- screen is clear and great
- **Aspect extraction**: screen
- **Opinion identification**: clear, great
- Polarity classification: clear is +ve
- Opinion separation/generality: clear (aspect-specific), great (general)
  - Understand consumer taste of products

Flipkart

LDA – review mining



- Each item: finite mixture over an underlying set of latent variables
- Generative, probabilistic model for collection of discrete items
  - Review corpora generated first by choosing a value for aspect/rating pair $<a_m, r_m>$ (θ)
  - Repeatedly sample M opinion phrases $<t_m, s_m>$ conditioned on θ

Extract aspects/ratings together

Category	#Reviews	#Opinion Phrases	#Phrases per Product
Camcorder	197	1,694	211.77
Cel. Phone	630	7,642	955.23
Dig. Camera	707	10,435	1304.41
DVD Player	324	3,707	463.32
Mp3 Player	625	6,131	766.41
Overall	2,483	29,609	3701.14

- Topic modeling helps only with document clustering
- Independent aspect identification/ rating prediction leads to errors [Moghaddam '11]
  - Same sentiment word shows different opinions for various aspects
- Dataset
  - 29609 phrases, 2483 reviews, 5 categories from Epinions

Aspect-based sentiment analysis (ABSA)

[Moghaddam ’10]

Input

Canon GL2 Mini DVD Camcorder

… excellent zoom … blurry lcd … great picture quality … accurate zooming … poor battery … inaccurate screen … good quality … affordable price … poor display … inadequate battery life … fantastic zoom … great price …

Output

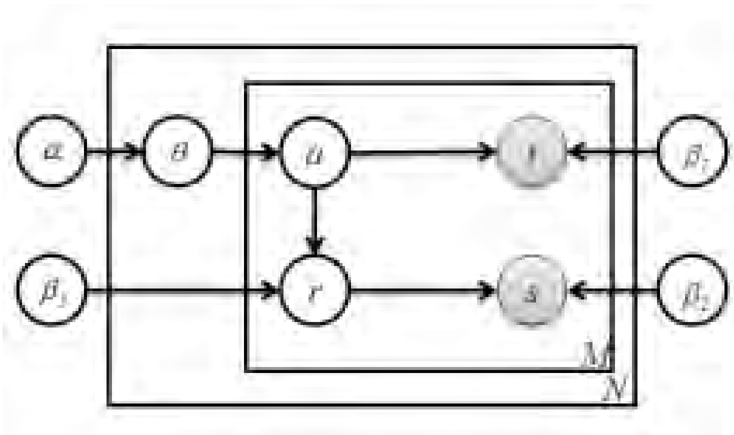
Aspect	Head Terms	Rating
zoom	zoom, zooming	5
price	price	4
picture quality	picture quality, quality	4
battery life	battery life, battery	2
screen	screen, lcd, display	1
…		…

Prediction

- [Liu ‘16]
- screen is clear and great
- **Aspect extraction**: screen
- **Opinion identification**: clear, great
- Polarity classification: clear is +ve
- **Opinion separation/generality**: clear (aspect-specific), great (general)
  - Understand consumer taste of products

Flipkart

Interdependent LDA model



- [Moghaddam '11]
- Generate an aspect a_m from LDA
- Generate rating r_m conditioned on sample aspect
- Draw head term t_m and sentiment s_m conditioned on a_m and r_m respectively

References

Liu, B. Sentiment analysis. AAAI 2011.

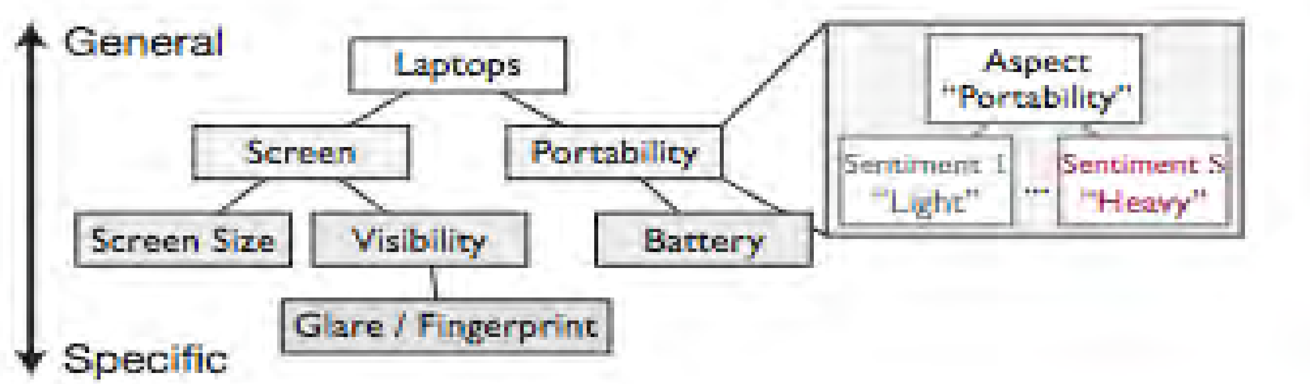
Moghaddam S, Ester M. ILDA: interdependent LDA model for learning latent aspects and their ratings from online product reviews. SIGIR ’11

Moghaddam, S., & Ester, M. Aspect-based opinion mining from online reviews. SIGIR 2012.

Aspect-based sentiment analysis (ABSA)

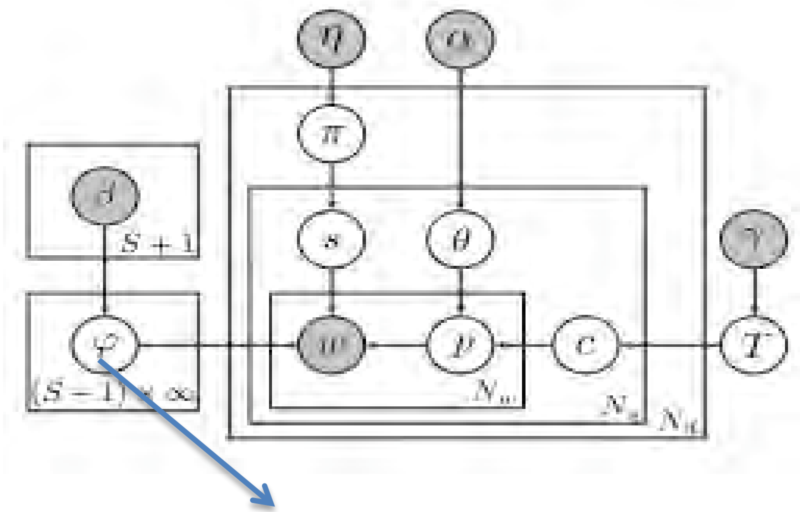
- Statistics-based text mining
- Topic models
  - Early
  - **Recent (10 minutes)**
- Deep learning (10 minutes)

Structure in Aspect-Sentiment



- Different consumers interested in different aspect granularities
- Polarity of general to aspect-specific sentiment words through sparse co-occurrence [Kim '13]
- Dataset
  - 10414 (laptop)/20862 (digital SLR) reviews, 825/260 targets: Amazon

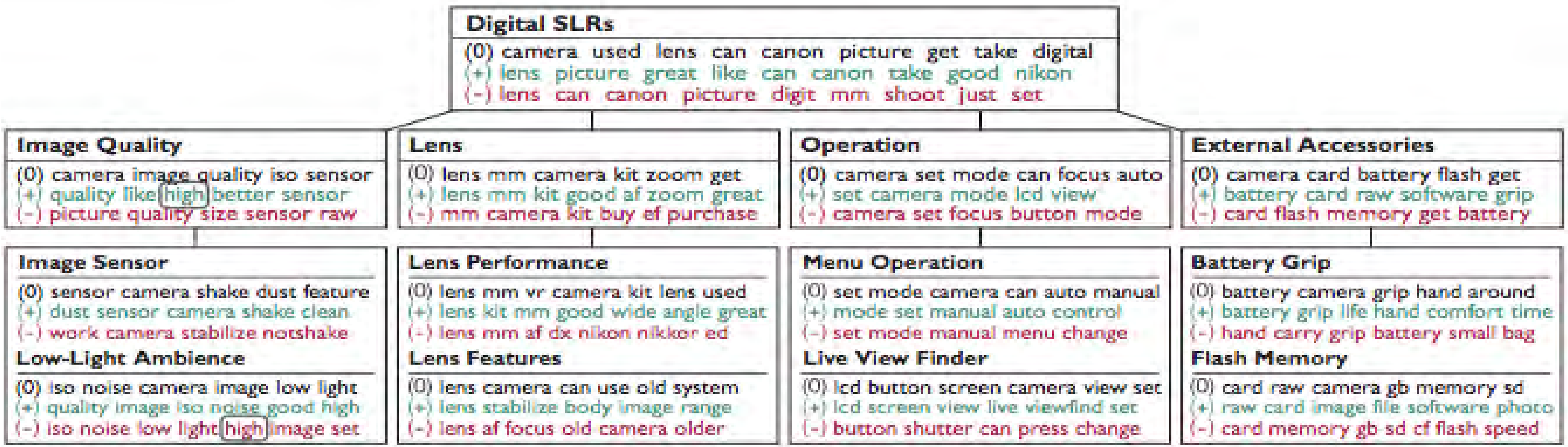
Hierarchical aspect-sentiment model



Light..Heavy
Portability
|
Battery

- [Kim '13]
- Jointly infer aspect-sentiment tree
  - with a bayesian non-parametric model as prior
- Aspect-sentiment node ϕ_k itself is a tree
  - aspect topic at root and sentiment-polar (S) in leaves
- Topics share semantic theme
  - generated from dirichlet (beta) distribution

Sentiment-aspect hierarchy: Digital SLRs [Kim '13]



- [Kim '13]
- Polarity of sentiment words depend on aspect at various granularities

Aspect-based sentiment analysis (ABSA)

[Moghaddam '10]

Input

Canon GL2 Mini DVD Camcorder
... excellent zoom ... blurry lcd ... great picture quality ... accurate zooming ... poor battery ... inaccurate screen ... good quality ... affordable price ... poor display ... inadequate battery life... fantastic zoom ... great price ...

Output

Aspect	Head Terms	Rating
zoom	zoom, zooming	5
price	price	4
picture quality	picture quality, quality	4
battery life	battery life, battery	2
screen	screen, lcd, display	1
...		...

Prediction

- [Liu '16]
- screen is clear and great
- **Aspect extraction**: screen
- **Opinion identification**: clear, great
- **Polarity classification**: clear is +ve
- **Opinion separation/generality**: clear (aspect-specific), great (general)
  - Understand consumer taste of products

Flipkart

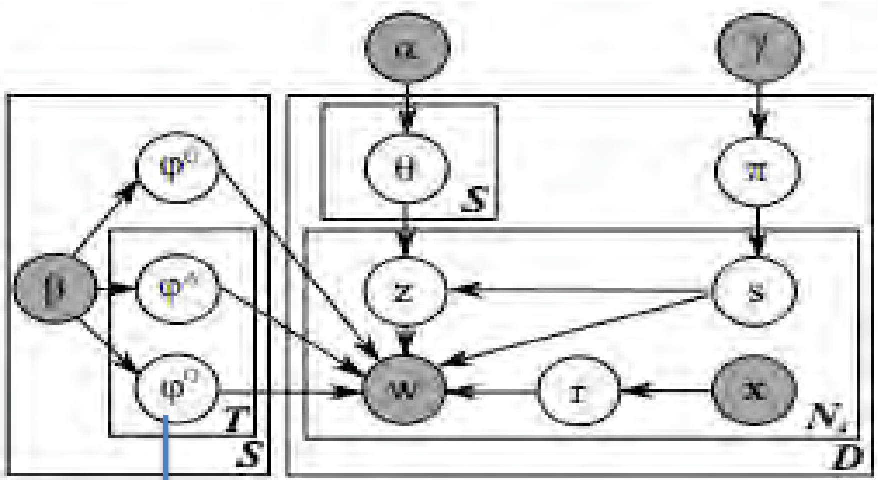
Aspect-sentiment topic

- [Wang '16]
- Generated topics inconsistent with human judgment
  - Wrong identification of opinion words (nice screen)
- Models based on co-occurrence suffer
  - Opinions (smooth screen) NOT identified if mixed in same topic

Lifelong learning across categories

- [Wang '16]
- Extract/accumulate knowledge from past
- Use discovered knowledge for future learning
  - Like human experience
- Dataset:
  - 50 product domains with 1000 reviews each - Amazon

JAST Model



screen

[Wang '16]

S	the number of sentiment polarities
D	the number of documents
T	the number of aspect topics
V	the number of words or terms in vocabulary
N_d	the number of words in document d
s, d, z	sentiment polarity, document, topic
w, x, r	word, lexicon indicator, word type
π	multinomial distribution over sentiments
θ	multinomial distribution over topics or aspects
φ^G	multinomial distribution over general opinion words
φ^A	multinomial distribution over aspect words
φ^O	multinomial distribution over aspect-specific opinion words
α, β, γ	Dirichlet prior for θ, φ, π

clear

great

Topic quality evaluation: opinion

Battery (Negative)			Shipping&Order (Positive)		
LAST	JAST	ASUM	LAST	JAST	ASUM
die	old	*problem*	new	free	*great*
dead	die	hot	free	*happy*	*good*
short	fail	*bad*	fast	fast	quickly
drain	*suck*	die	quick	*pleased*	*well*
fail	useless	*original*	refund	refund	*love*
old	hassle	old	promptly	*recommend*	*perfect*
hassle	*bad*	*new*	original	new	*nice*
wrong	*concern*	*long*	correct	works	*perfectly*
useless	bother	break	works	quick	new
complain	*nervous*	*hate*	accurate	promptly	fast

- [Wang ‘16]
- Incorrect (aspect-specific/polarity) (new, original)
- Incorrect (aspect-specific) (bad, suck)

Topic quality evaluation: aspect

Battery			Shipping&Order		
LAST	JAST	ASUM	LAST	JAST	ASUM
battery	battery	charge	order	arrive	*screen*
charge	charge	battery	receive	receive	receive
hour	life	recharge	arrive	order	arrive
life	hour	*iphone*	shipping	purchase	order
power	*device*	*sd*	ship	*expect*	*privacy*
charger	cable	*card*	today	send	cost
recharge	*phone*	receive	delivery	ship	money
night	*ipad*	replacement	usual	shipping	*monitor*
outlet	power	*purcharse*	*expect*	back	purchase
aaa	plug	*star*	*manner*	*seller*	*seller*

- [Wang '16]
- Joint modeling improves aspect quality while mining coherent opinions

Attribute-aspect normalization

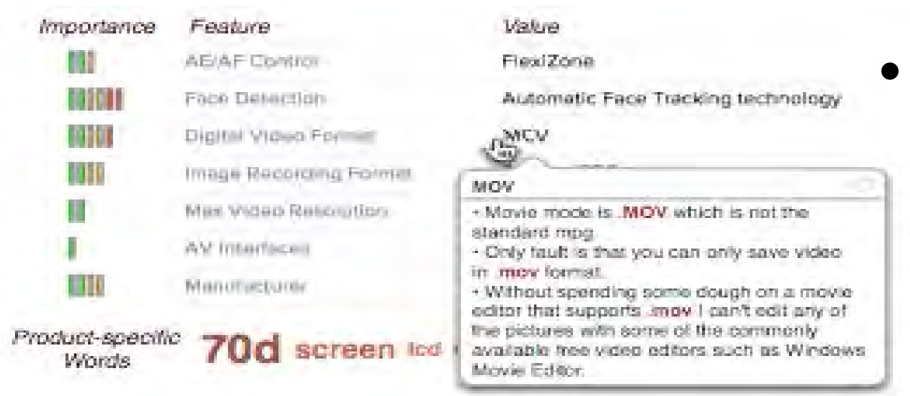
	Concept 1	Concept 2	Concept 3
	wireless, network, wi-fi, internet, connection	screen, resolution, display, brightness, monitor	battery, life, 6-cell, charge, capacity
Product 1	"draft-n wi-fi network," "the integrated bluetooth connectivity"	"10.1 inches ultrawide display," a superbright 10.1-inch lcd	"extended battery life," "the traditional lithium battery"
Product 2	"fast ethernet connection," "the bluetooth 2.0+edr technology"	"normal tft lcd," "10.1 inches widescreen"	"65 hours battery life," "a 6 cells battery"
Product 3	"ultimate wireless accessibility," "wired ethernet lan"	"widescreen trubrite display," "a 1024 × 600-pixel resolution"	"long battery life," "6-cell lithium-ion battery"

The second row shows the top-five weighted terms in the popular features discovered from the customer reviews. Each cell of the table contains the popular attribute text fragments extracted from product description pages.

~~[Bing '16]~~

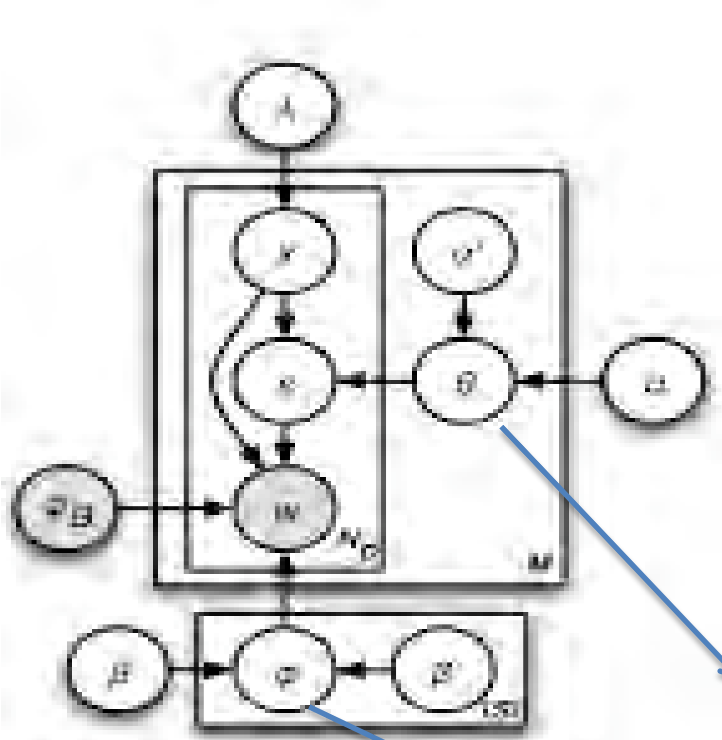
- ~~Extract product attributes from description page~~
- ~~Infer aspect popularity in reviews/map to feature~~
- ~~Hidden CRF bridges vocabulary gap~~

Augmented specification: camera



- [Park '15]
- Advanced features hard for novice customers
- Specs can be enhanced with reviews
  - Opinions on feature value/importance
  - Product-specific words

DuanLDA



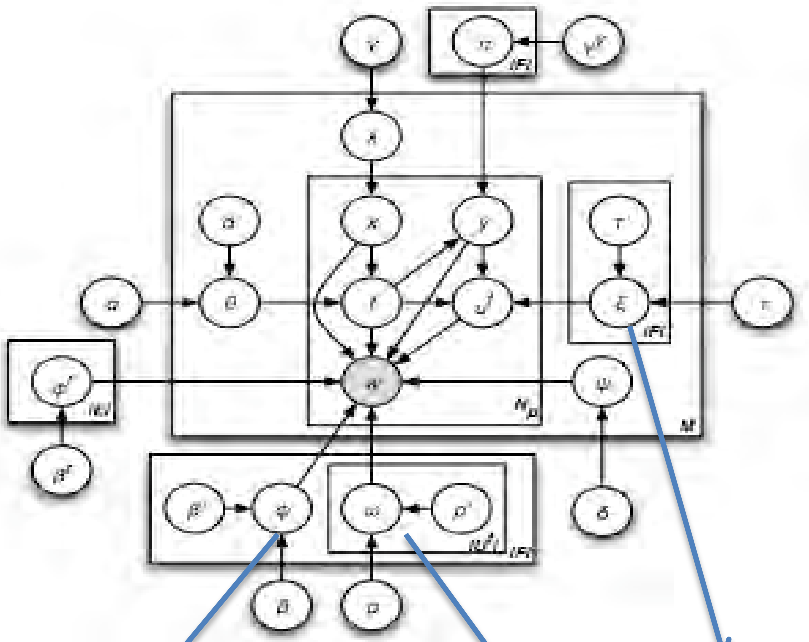
Feature - value

Review aspect

[Park '15]

- Each document is concatenated review
- s is feature-value pair/specification topic
- Switch x decides background/spec. topic

SpecLDA



[Park '15]

- Feature-value pairs with same feature share info
- Organized as hierarchy for better topic estimation
- Capture user preference to indicate a pair
  - for feature-/value-related words
- Employ value-word prior in addition to feature-word

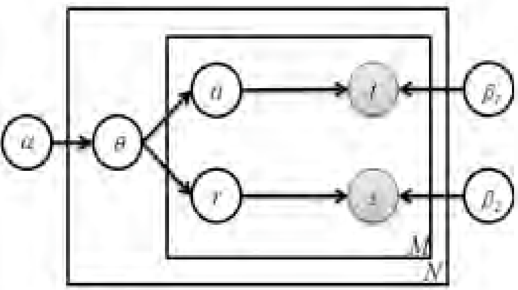
Evaluation: top words for a spec. topic

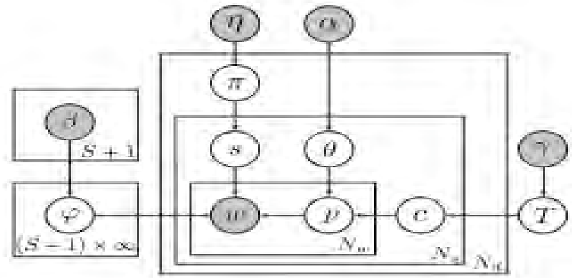
DuanLDA	SpecLDA–	SpecLDA–_V	SpecLDA
display	digital	lcd	lcd
type	lcd	display	2.5
phtography	2.5	2.5	display
font	display	screen	screen
florescent	huge	type	large
informative	technology	large	inch
channel	expensive	inch	monitor
triple	type	monitor	articulate
resultant	large	phtography	1.8
lcd	outstanding	font	230,000
printout	icon	crack	panel
a10	silver	symbol	bright
colorful	follow	1.8	icon
information	rubber	icon	rotatable
info	phtography	bright	sunlight
horizontal	salesman	brightness	brightness
picture/video	font	articulate	viewfinder
infinite	info	salesman	tilt
2aa	blessing	range	viewer
3.0	symbol	informative	flippable

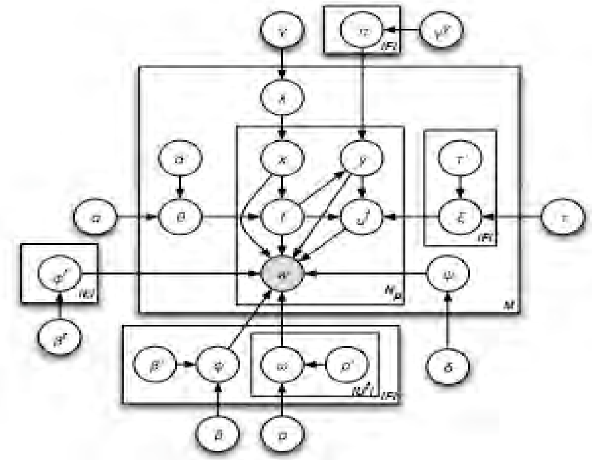
[Park '15]

- Spec. ("Display – type","2.5in LCD display")
- DuanLDA – does not use value word prior
- SpecLDA – feature and value topics form a cluster

ABSA Summary (Topic models)





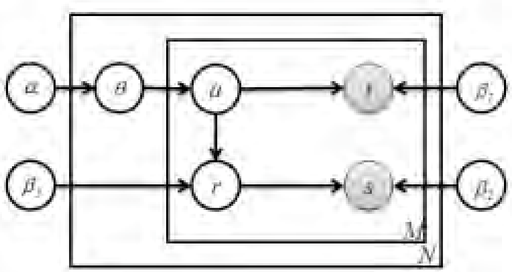


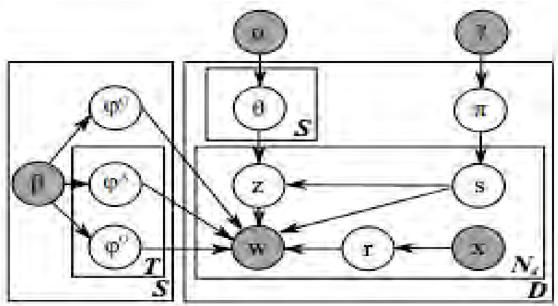
LDA: sample opinion phrases, not all terms

HASM: aspect-sentiment tree at each with topic root and S sentiment leaves

ILDA: Correspondence betn. aspect/rating

SpecLDA: feature-, value- word, aspect





JAST: aspect: (term, opinion), polarity, generic opinion

References

Kim, S., Zhang, J., Chen, Z., Oh, A. H., & Liu, S. A Hierarchical Aspect-Sentiment Model for Online Reviews. AAAI 2013.

Park, D.H., Zhai, C. and Guo, L. Speclda: Modeling product reviews and specifications to generate augmented specifications. SDM 2015.

Wang, S., Chen, Z., & Liu, B. Mining aspect-specific opinion using a holistic lifelong topic model. WWW 2016.

Aspect-based sentiment analysis (ABSA)

- Statistics-based text mining
- Topic models
  - Early
  - Recent
- **Deep learning (10 minutes)**

Aspect-based sentiment analysis (ABSA)

[Moghaddam ’10]

Input

Canon GL2 Mini DVD Camcorder
… excellent zoom … blurry lcd … great picture quality … accurate zooming … poor battery … inaccurate screen … good quality … affordable price … poor display … inadequate battery life … fantastic zoom … great price …

Output

Aspect	Head Terms	Rating
zoom	zoom, zooming	5
price	price	4
picture quality	picture quality, quality	4
battery life	battery life, battery	2
screen	screen, lcd, display	1
…		…

Prediction

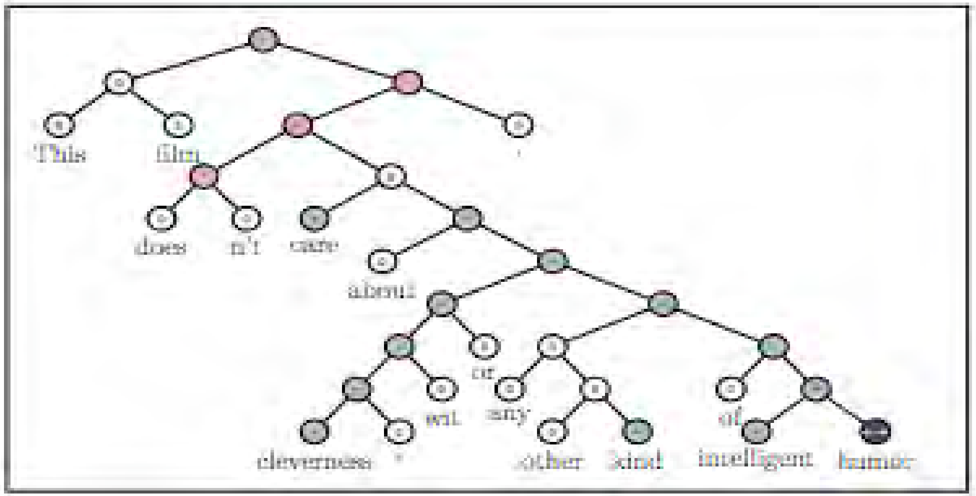
- [Liu ‘16]
- screen is clear and great
- **Aspect extraction**: screen
- **Opinion identification**: clear, great
- Polarity classification: clear is +ve
- Opinion separation/generality: clear (aspect-specific), great (general)
  - Understand consumer taste of products

Flipkart

ABSA/Deep learning: Overview

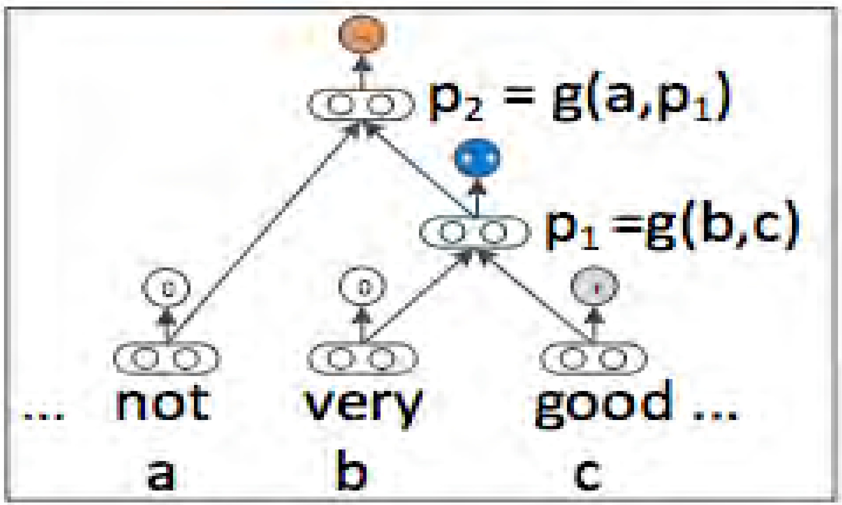
- Rating prediction
  – Recursive neural networks
- Aspect extraction
  – CNN
  – Recurrent neural networks
  – LSTM/attention

Semantic compositionality



- Rating prediction [Socher '13]
- Word spaces cannot express meaning of longer phrases in a principled way
- Sentiment detection requires powerful models/supervised resources

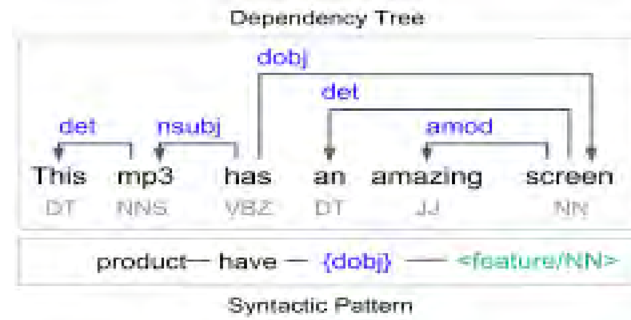
Recursive deep models



- Rating prediction [Socher '13]
- Parent vectors computed bottom-up
- compositionality function varies for different models
- Node vectors used as features for classifier

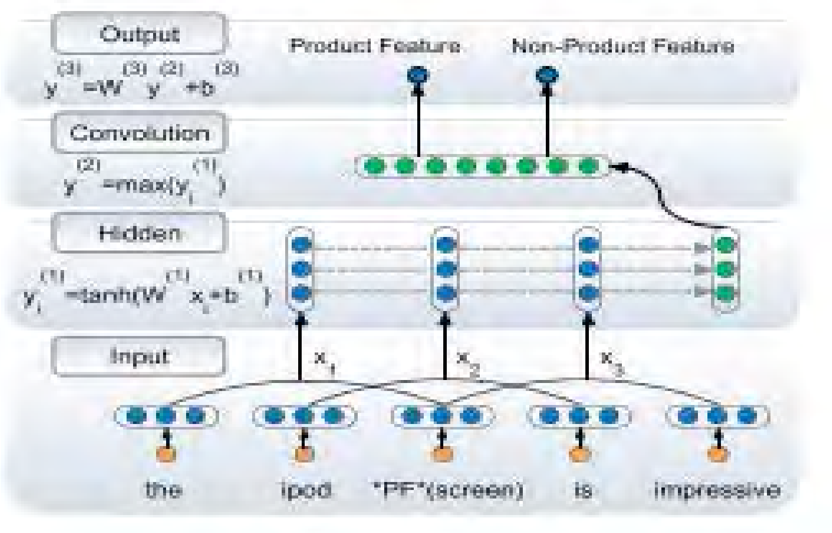
Semantic product feature mining

- Aspect extraction [Xu ‘14]
- Reviews features mined using syntactic patterns
- Syntax helps leverage only context
  - Suffering from sparsity of discrete information
- Term-product association needs semantic clues



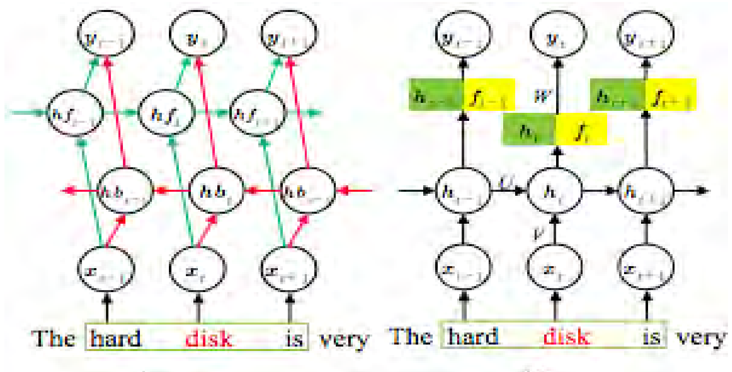
a) *This player has an fm tuner.*
b) *This mp3 supports wma file.*
c) *This review has helped people a lot.*
d) *This mp3 has some flaws.*

Convolutional neural network



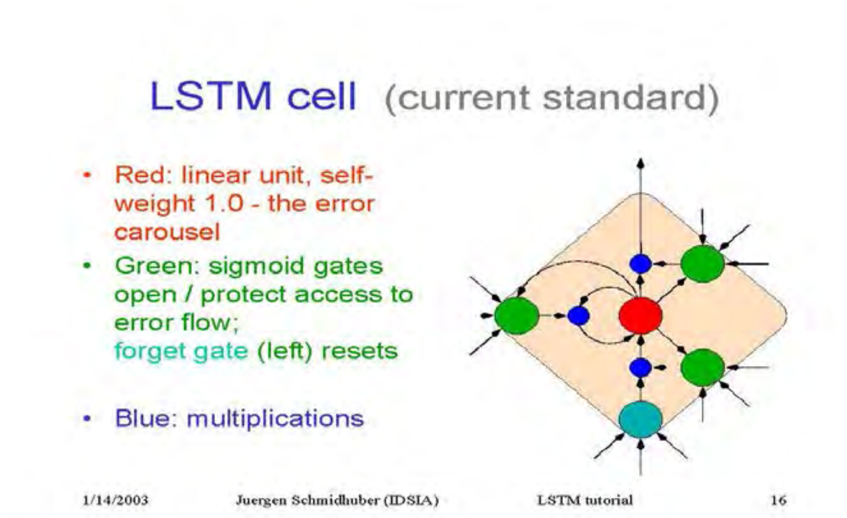
- Aspect extraction [Xu ’14]
- Similarity graph
  - Encodes lexical clue
- CNN
  - Captures context
- Label propagation
  - Captures both semantic clues

Recurrent neural networks



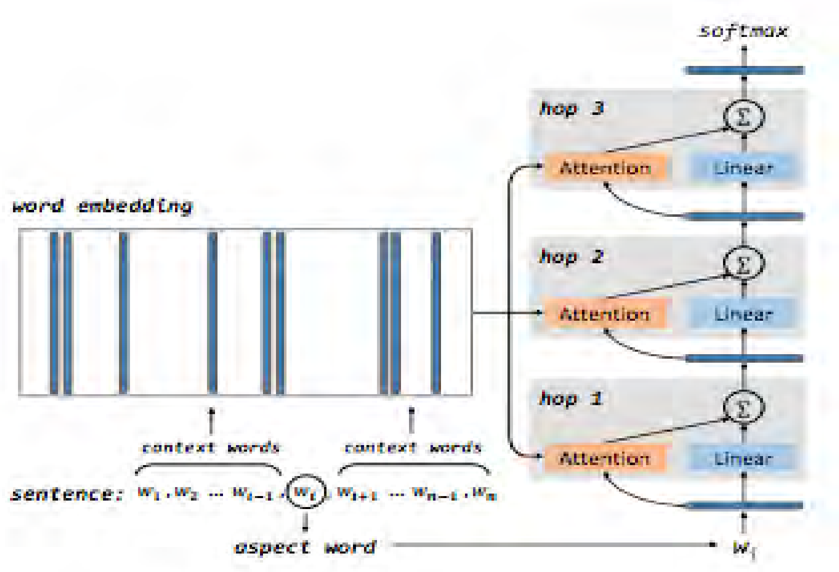
- Aspect extraction [Liu '15]
- window-based approach for short-term dependencies
- bidirectional links for long-range dependencies
- linguistic features to guide training

Deep memory network: LSTM



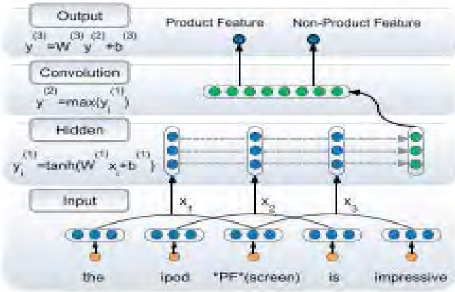
- Aspect extraction [Tang '16]
- *Fast processor but small memory*
- Need to capture semantic relations of aspect/context words
- LSTM incapable of exhibiting context clues
- Selectively focus instead on context parts

Deep memory network: attention



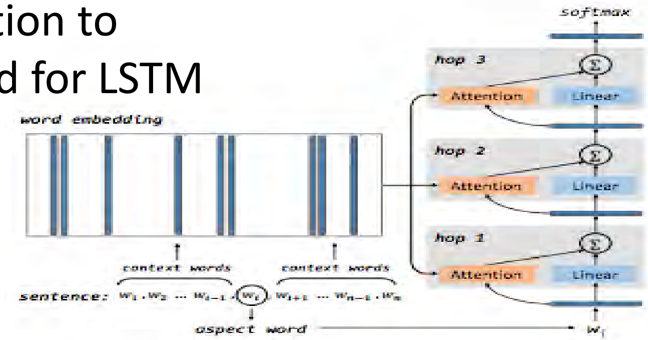
- Aspect extraction [Tang '16]
- Multiple hops are stacked
  - so abstractive evidence is selected from external memory
- Attention mechanism
  - weight to lower position while computing upper representation
- linear/attention layers with shared parameters in each hop
- Context word importance based on distance to aspect

ABSA summary (deep learning)

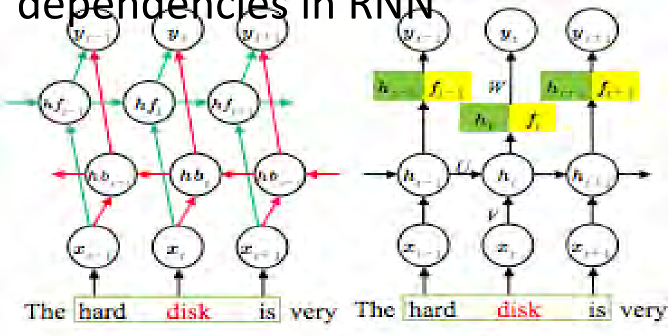


Similarity graph for lexical and CNN for context semantics

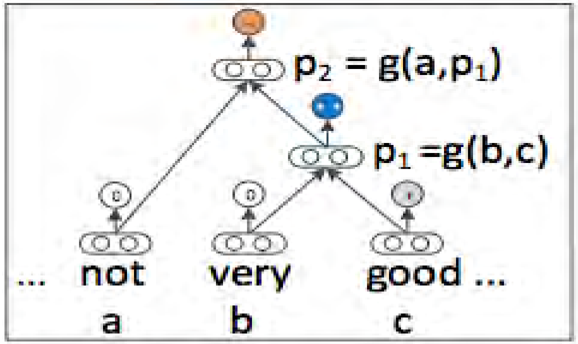
Content/location attention to emphasize context word for LSTM



Window for short-term and bidirectionality for long-range dependencies in RNN



Parent vectors computed bottom-up in RecNN for semantic compositionality



References

Liu, P., Joty, S. R., & Meng, H. M. Fine-grained Opinion Mining with Recurrent Neural Networks and Word Embeddings. EMNLP ’15.

Socher, R., Perelygin, A., Wu, J.Y., Chuang, J., Manning, C.D., Ng, A.Y. and Potts, C., Recursive deep models for semantic compositionality over a sentiment treebank, EMNLP ’13.

Tang, D., Qin, B. and Liu, T. Aspect level sentiment classification with deep memory network. EMNLP ’16

Xu, L., Liu, K., Lai, S. and Zhao, J.. Product Feature Mining: Semantic Clues versus Syntactic Constituents. ACL 2014.

FLIPKART ABSA

Trusted advisor

Objective

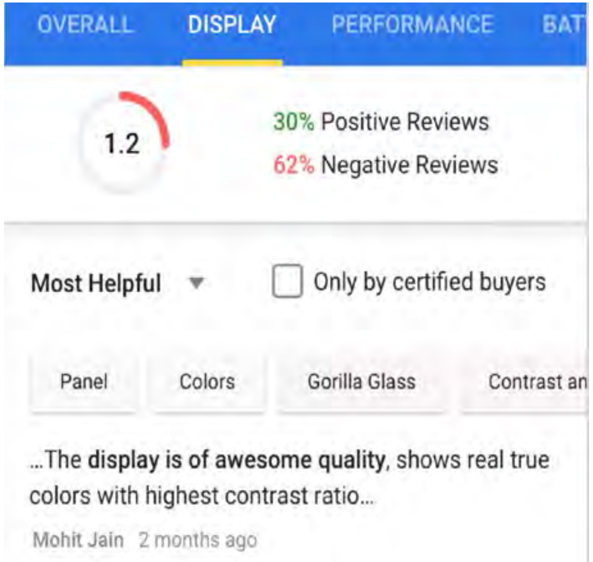
- Transform Flipkart into a research destination
- Shorten the purchase cycle for a user

Aspect Reviews

- Identify what users care about while buying a specific product
- Organize/classify reviews into these aspects & sub-aspects

Aspect Ratings

- Organize sentiments (polarity & strength)
- Summarize sentiments in the form of a user rating at an aspect level



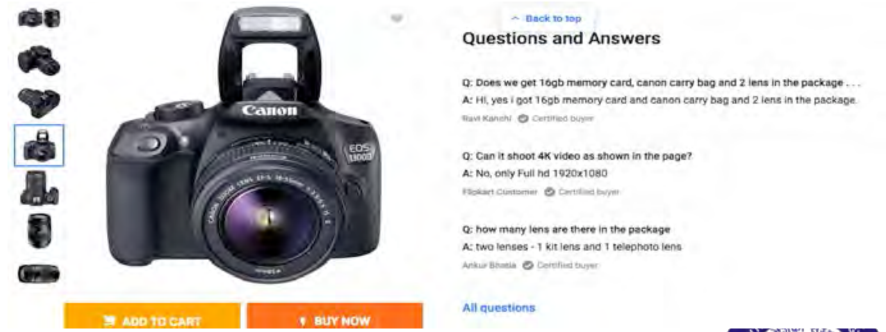
Flipkart

Future directions

- Product comparison
  - Model entities assessed directly in a single review sentence [Sikchi '16]
  - Leverage product spec. to learn attribute importance

Aspects	Camcorder1	Camcorder2
zoom	4	2
sound	1	3
screen	2	4
price	3	3
size	4	2
Overall Rating	3	3

- Question/answer
  - Summarize reviews with real-world questions
  - Rerank initial candidates for diversity [Liu '16]



References

- Liu, M., Fang, Y., Park, D.H., Hu, X. and Yu, Z. Retrieving Non-Redundant Questions to Summarize a Product Review. SIGIR 2016.
- Sikchi, A., Goyal, P. and Datta, S., Peq: An explainable, specification-based, aspect-oriented product comparator for e-commerce. CIKM 2016.

FLIPKART RECOMMENDATION

Traditional View

User-Item
- Purchase history from logs

Recommender
- Collaborative Filtering
- Content based
- Hybrid

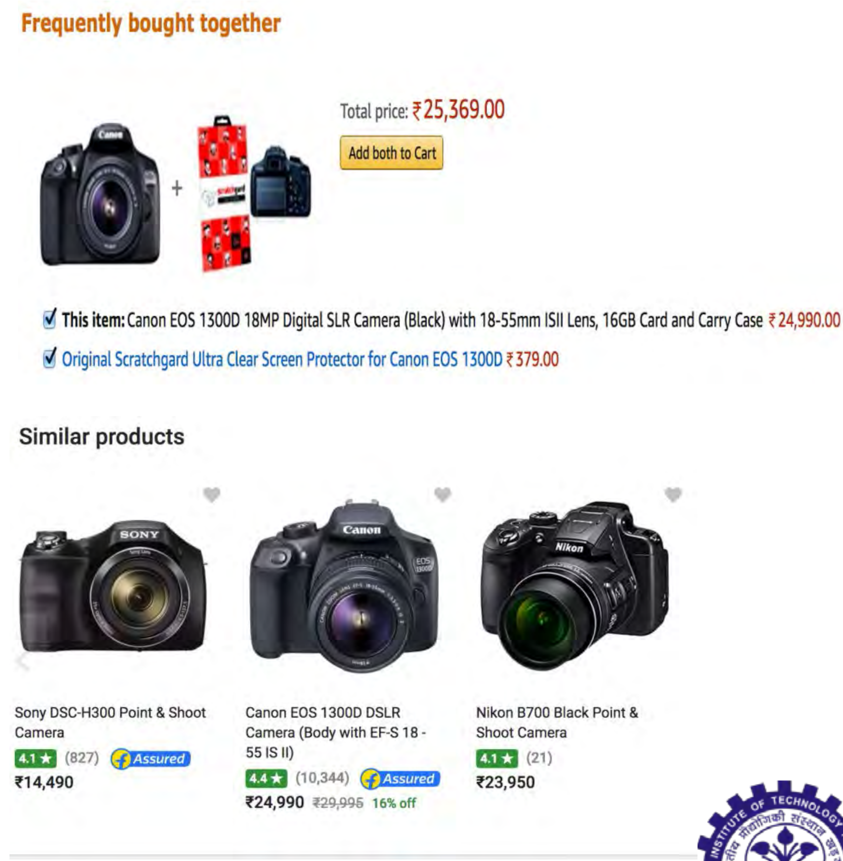
App/Website
- Present to user based on score from recommenders

Handling Cold-Start

- L1 ranking – leverage content (e-commerce has structured data), business provided attributes (CTR, conversion rate)
- L2 ranking - many online learning algorithms – AdaGrad, FTRL, AdPredictor for real-time learning (per-coordinate learning) , can handle new users/products very soon
- Other sources: Reviews, ratings
  - Some issues
    - Reviews/ratings are always biased to extremes
    - Ratings: Users very satisfied or unsatisfied
    - Reviews: more information to type. Biased towards angst/unsatisfactory experience

Future directions

- Bundle [Pathak ‘17]
  - Personalized/not considered in isolation
  - Assess size/compatibility
- Upsell [Lu ‘14]
  - Maximize revenue thru utility, saturation and competing products

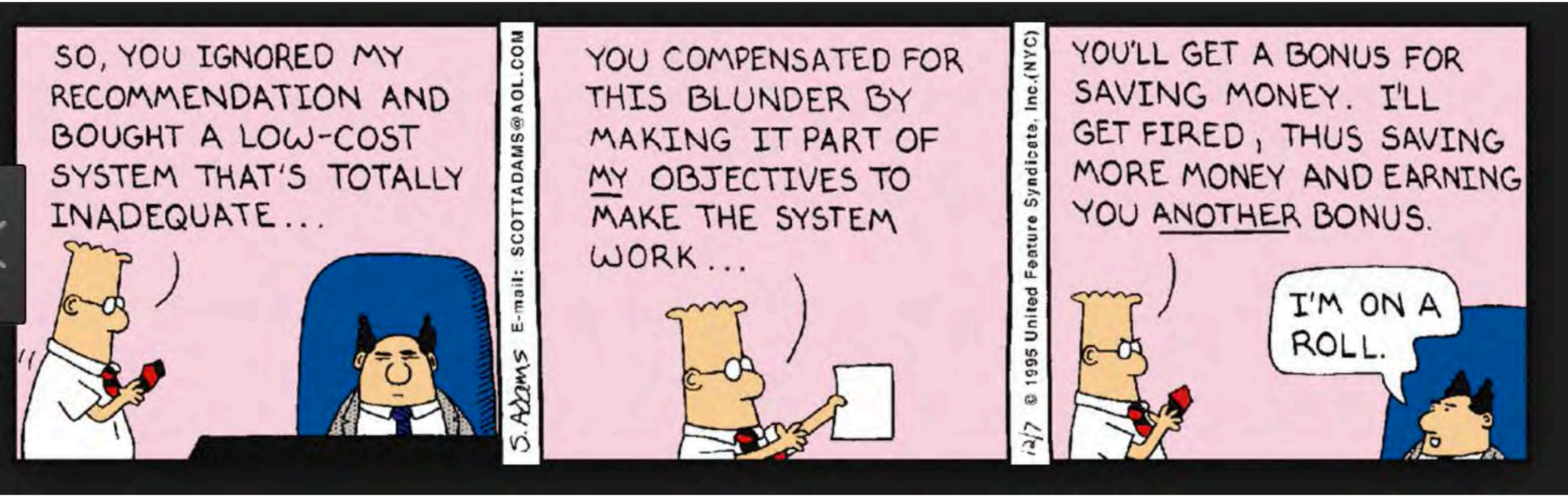


Flipkart

References

- Pathak, A., Gupta, K., & McAuley, J. Generating and Personalizing Bundle Recommendations on Steam. SIGIR '17.

- Lu, W., Chen, S., Li, K., & Lakshmanan, L. V. Show me the money: Dynamic recommendations for revenue maximization, VLDB '14.


SO, YOU IGNORED MY RECOMMENDATION AND BOUGHT A LOW-COST SYSTEM THAT'S TOTALLY INADEQUATE...
S. Adams E-mail: SCOTTADAMS@AOL.COM
YOU COMPENSATED FOR THIS BLUNDER BY MAKING IT PART OF MY OBJECTIVES TO MAKE THE SYSTEM WORK...
12/7 © 1995 United Feature Syndicate, Inc.(NYC)
YOU'LL GET A BONUS FOR SAVING MONEY. I'LL GET FIRED, THUS SAVING MORE MONEY AND EARNING YOU ANOTHER BONUS.
I'M ON A ROLL.

Product recommendations enhanced with reviews
Part 2

Muthusamy Chelliah

Director, Academic Engagement, Flipkart

Sudeshna Sarkar

Professor, IIT Kharagpur

Email: sudeshn@cse.iitkgp.ernet.in

Review Enhanced Recommendation

1. Introduction
2. Review mining
   a) Background: Features and Sentiment, Latent features
   b) Statistical approaches. LDA
   c) Recent topic models
   d) Deep Learning (ABSA)
   e) Flipkart ABSA/Recommendation
3. Review based Recommendation
   a) Handling Data Sparsity
   b) Topic Model, Matrix Factorization and Mixture Model based Recommendation
   c) Deep learning based Recommendation
4. Explainable Product Recommendations
5. Summary, Future Trend

Handling Data Sparsity

Explainable Recommendation

REVIEW AWARE RECOMMENDATION

Issues with Recommender Systems

Collaborative Filtering

- Cold Start
  - New items, New Users
- Sparsity
  - it is hard to find users who rated the same items.
- Popularity Bias
  - Cannot recommend items to users with unique tastes.

Content Based

- Hard to identify quality judgement.
- Ignores user’s ratings.

Review Based

- Text information from reviews address sparseness of rating matrix.
  - Bag of words model
- More sophisticated models required to capture contextual information in documents for deeper understanding: CNN, etc.
- Enable Explanations

Review Aware Recommender Systems

Traditional Models

- Weight on quality of ratings, Recsys'12
- Use reviews as content, Recsys'13
- Aspect Weighting, WI'15
- TriRank, CIKM'15

Topic Models and Latent Factors

- HFT, Recsys '13
- RMR, Recsys'14
- CMR, CIKM'14
- TopicMF, AAAI'14
- FLAME, WSDM'15

Joint Models of Aspects and Ratings

- JMARS,KDD'14
- EFM, SIGIR'14
- TPCF, IJCAI'13
- SULM, KDD'17

Deep Learning Models

- ConvMF, Recsys'16
- DeepCONN, WSDM'17
- TransNets, Recsys'17
- D-Attn, Recsys'17

Using weights based on reviews

- Attach weights or quality scores to the ratings.
- The quality scores are determined from the corresponding review.
- Incorporate quality scores as weights for the ratings in the basic PMF framework

Review Quality Aware Collaborative Filtering, by Sindhu Raghavan, Suriya Gunasekar, Joydeep Ghosh, Recsys'12

- Propose a unified framework for performing the three tasks:
  - aspect identification
  - aspect-based rating inference
  - weight estimation.

Review Mining for Estimating Users' Ratings and Weights for Product Aspects, by Feng Wang and Li Chen, WI'15

Sentimental Product Recommendation

1. Convert unstructured reviews into rich product descriptions
   a) Extract Review Features using shallow NLP
   b) Evaluate Feature Sentiment by opinion pattern mining
2. Content-based recommendation approach based on feature similarity to a query product
   a) Similarity-Based Recommendation
   b) Sentiment-Enhanced Recommendation: seek products with better sentiment than the query product.

Sentimental Product Recommendation, by Ruihai Dong, Michael P. O'Mahony, Markus Schaal, Kevin McCarthy, Barry Smyth, Recsys 2013

Review Aware Recommender Systems

Topic Models and Latent Factors

- HFT, McAuley & Leskovec, Recsys 13, LDA+MF
- RMR, Ling etal,Recsys'14, LDA+PMF
- CMR, CIKM'14, Xu etal, LDA+PMF+User factors
- TopicMF, Baoetal, AAAI'14 NMF+MF

Joint Models of Aspects and Ratings

- JMARS, Diao etal,KDD'14, Graphical Models
- EFM, Zhang etal,SIGIR'14, Collective NMF
- TPCF, Musat etal,IJCAI'13 User topical profiles
- SULM, KDD'17

Flipkart

HFT: Linking Reviews

1. Latent Dirichlet Allocation (LDA) finds low dimensional structure in review text (topic representation)
2. SVD learns latent factors for each item.

- Statistical models combine latent dimensions in rating data with topics in review text.
- Use a transform that aligns latent rating and review terms.
- Link them using an objective that combines the accuracy of rating prediction (in terms of MSE) with the likelihood of the review corpus (using a topic model).

Hidden Factors and Hidden Topics: Understanding Rating Dimensions with Review Text: McAuley and Leskovec, Recsys 2013

HFT: Optimization

- A standard recommender system : $\hat{r}_{u,i} = p_u . q_i$
- Minimize MSE:

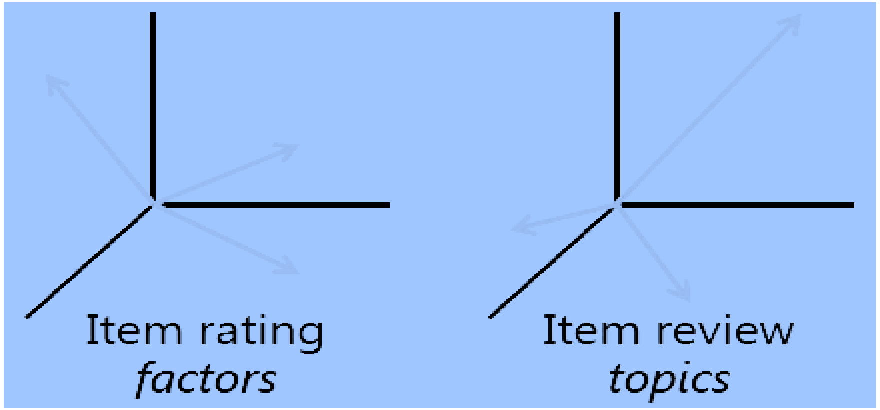
$$\min_{P,Q} \frac{1}{\mathcal{T}} \sum_{|\mathcal{T}|} (r_{ui} - q_i p_u)^2 + \left[\lambda_1 \sum_x \|p_u\|^2 + \lambda_2 \sum_x \|q_i\|^2 \right]$$

- Instead use review text as regularizer:

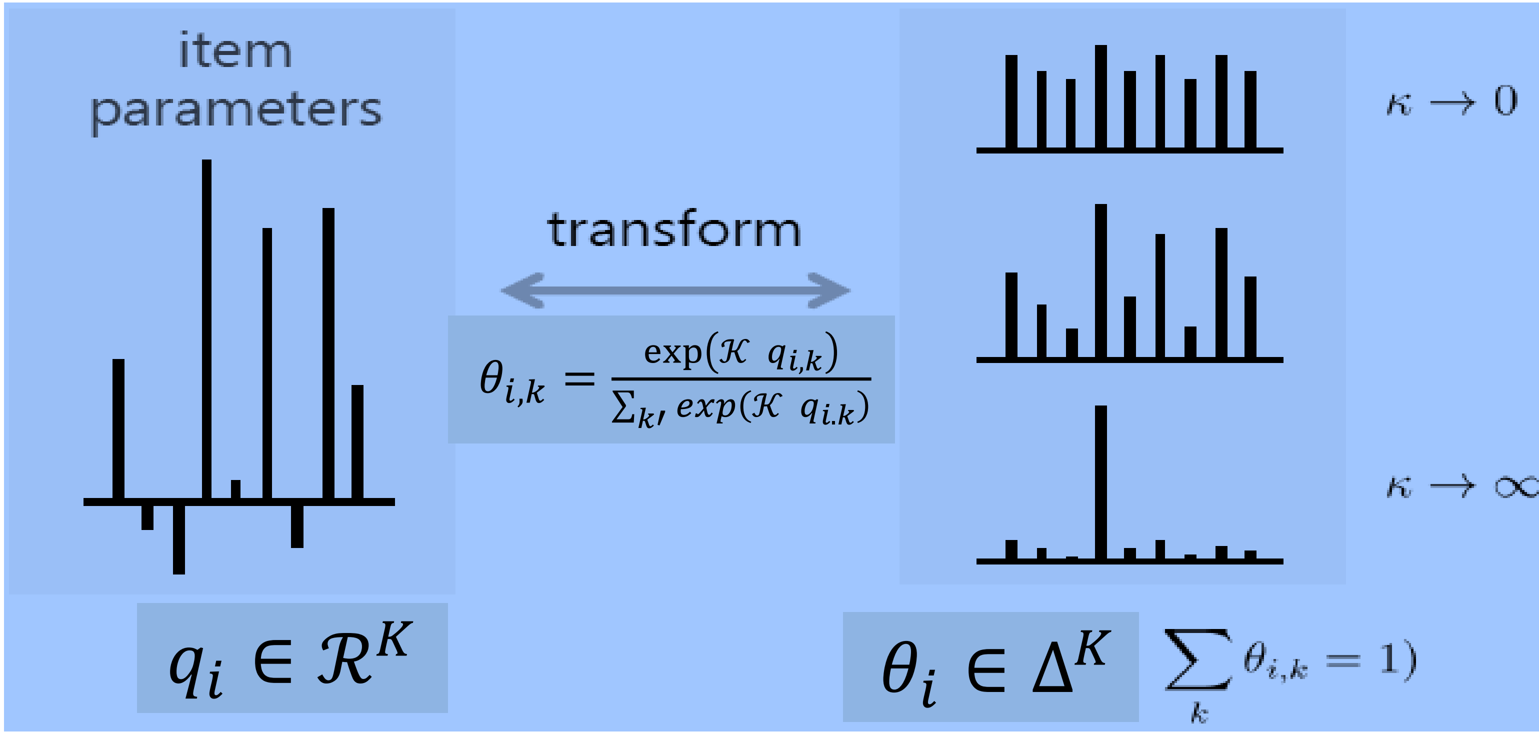
$$\min_{P,Q} \frac{1}{\mathcal{T}} \sum_{|\mathcal{T}|} \underbrace{(r_{ui} - q_i p_u)^2}_{\text{Rating error}} - \underbrace{\lambda\, l(\mathcal{T} | \Theta, \phi, z)}_{\text{Corpus likelihood}}$$

HFT: Combining ratings and reviews

- Matrix factorization and LDA project users and items into low-dimensional space
- Align the two



HFT: Combining ratings and reviews



By linking rating and opinion models, find topics in reviews that inform about opinions

HFT: Model Fitting

Repeat Step 1 and Step 2 until convergence:

Step 1: Fit a rating model regularized by the topic

(solved via gradient ascent using L-BFGS)

Step 2: Identify topics that "explain" the ratings.

Sample with probability

(solved via Gibbs sampling)

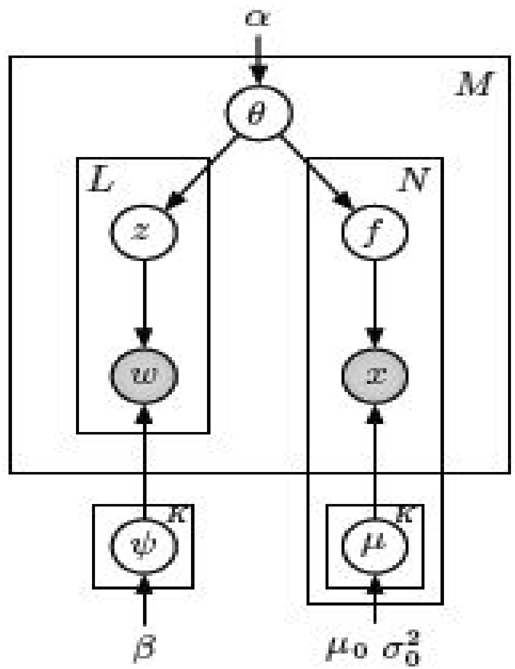
Ratings Meet Reviews, a Combined Approach to Recommend

- Use LDA to model reviews
- Use mixture of Gaussians to model ratings
- Combine ratings and reviews by sharing the same topic distribution

- N users $\{u_i\}$
- M items $\{v_j\}$
- (u_i, v_j) defines the observed ratings $X = \{x_{ij}\}$,
- Associated review r_{ij}

- Ratings meet reviews, a combined approach to recommend, Ling et al., Recsys 2014.

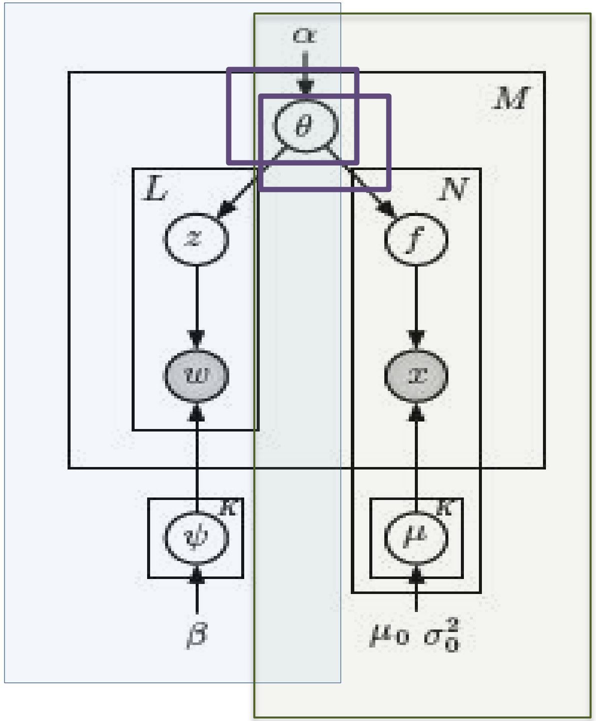
Graphical Model of RMR



$$P(\mathbf{w}, \mathbf{x} | \Theta; \alpha, \beta, \mu_0, \sigma_0^2, \sigma^2) \propto \prod_{j=1}^{M} P(\theta_j|\alpha) \prod_{i \in \mathcal{U}_j} \left(\prod_{l=1}^{L_{i,j}} \sum_{z=1}^{K} P(z|\theta_j) P(w_l|\psi_z) \right) \left(\sum_{f=1}^{K} P(f|\theta_j) P(x_{i,j}|\mu_{i,f}, \sigma^2) \right)$$

$$P(\mathbf{w},\mathbf{x}|\Theta;\alpha,\beta,\mu_0,\sigma_0^2,\sigma^2) \propto \prod_{j=1}^{M} P(\theta_j|\alpha) \prod_{i\in\mathcal{U}_j} \left(\prod_{l=1}^{L_{i,j}} \sum_{z=1}^{K} P(z|\theta_j)P(w_l|\psi_z) \right) \left(\sum_{f=1}^{K} P(f|\theta_j)P(x_{i,j}|\mu_{i,f},\sigma^2) \right)$$

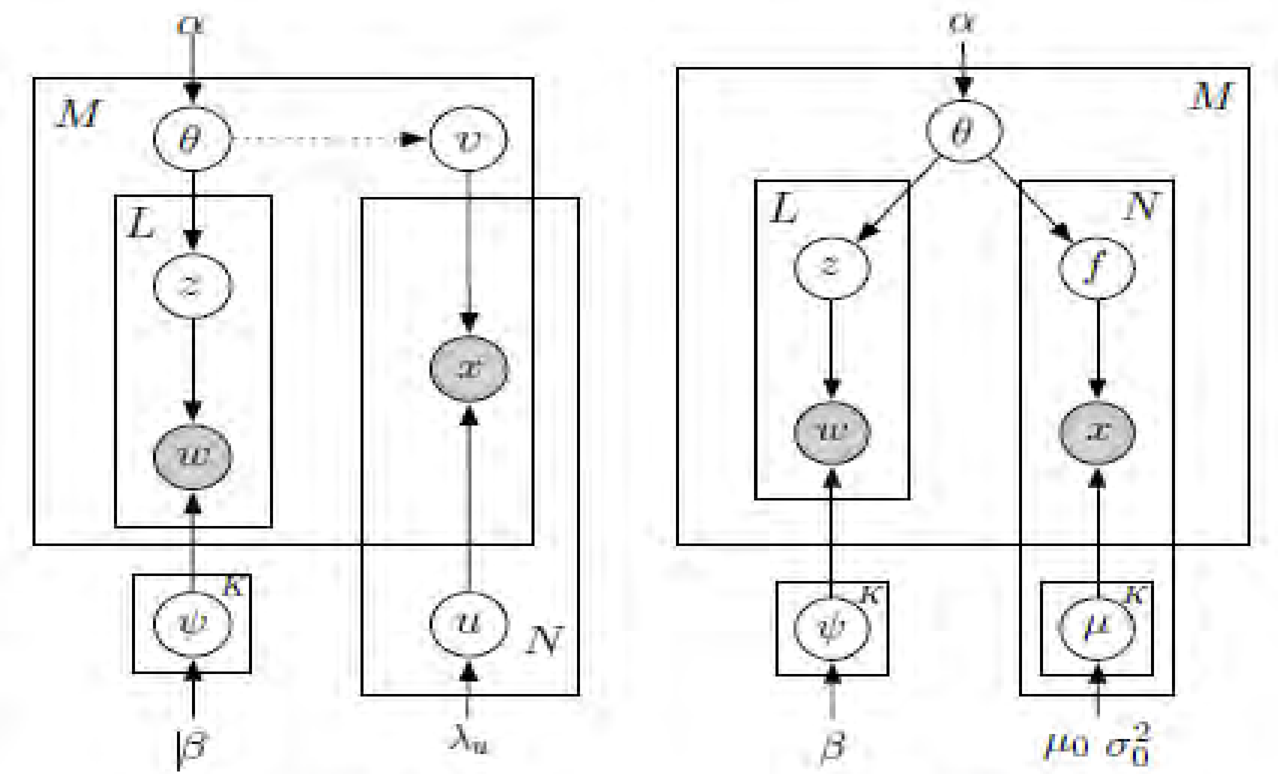
Use LDA to model the reviews



Use the same topic distribution to connect the rating part and the review part

Use mixture of Gaussians to model the ratings

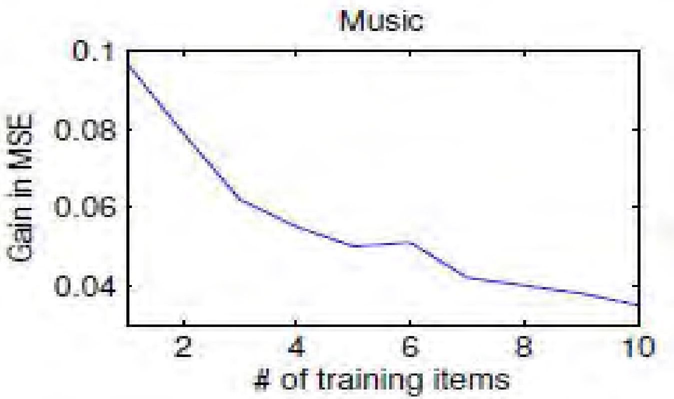
Comparison of HFT and RMR

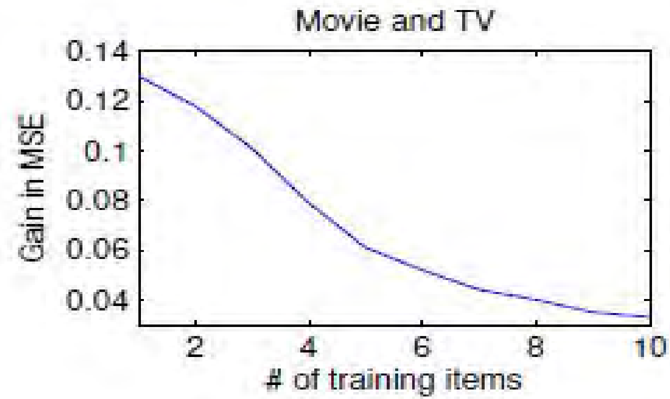


HFT RMR

Results for Cold Start Setting

- The improvement of HFT and RMR over traditional MF is more significant for datasets that are sparse.



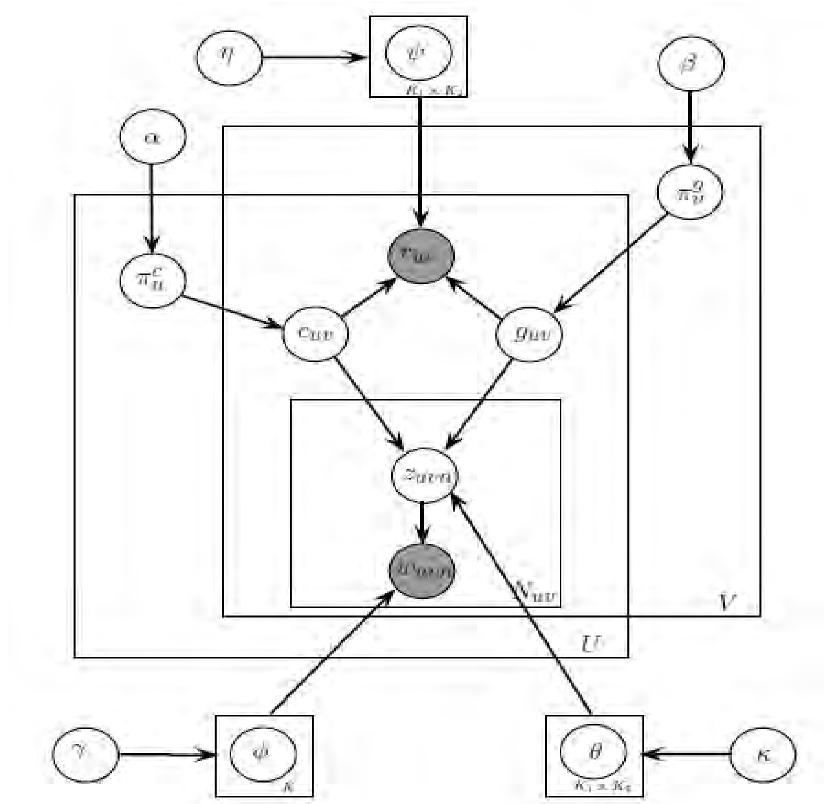


CMR

- Predict user's ratings by considering review texts as well as hidden user communities and item groups relationship.
- Uses Co-clustering to capture the relationship between users and items.
- Models as a mixed membership over community and group respectively.
- Each user or item can belong to multiple communities or groups with varying degrees.

Collaborative Filtering Incorporating Review Text and Co-clusters of Hidden User Communities and Item Groups, Yinqing Xu Wai Lam Tianyi Lin, CIKM 2014

CMR Graphical Model



U: user collection
V: item collection
r_{uv}: u's rating of item v
d_{uv}: u's review of item v
w_{uvn}: nth word of d_{uv}
z_{uvn}: topic of w_{uvn}
π_u^c: community membership of u
π_v^g: group membership of v
c_{uv}: user community of d_{uv} and r_{uv}
g_{uv}: item group of d_{uv} and r_{uv}
θ_{cg}: topic distribution of co-cluster (c,g)
ψ_{cg}: rating distribution of co-cluster (c,g)
ϕ_k: topic distribution of topic k

Collapsed Gibbs sampling algorithm to perform approximate inference.

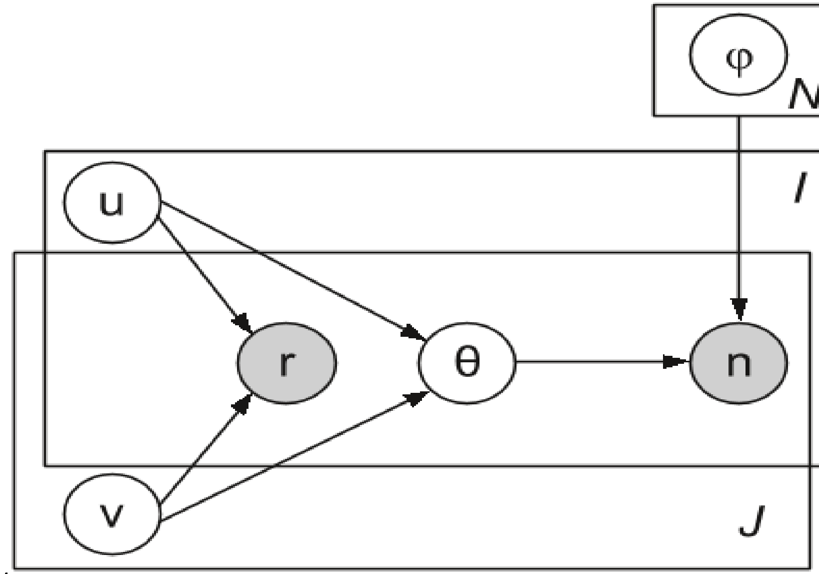
TopicMF: Simultaneously Exploiting Ratings and Reviews for Recommendation

Yang Bao† Hui Fang Jie Zhang, AAAI,2014

TOPIC-MF

TopicMF

- Matrix Factorization (MF) factorizes user-item rating matrix into latent user and item factors.
- Simultaneously topic modeling technique with Nonnegative Matrix Factorization (NMF) models the latent topics in review texts.

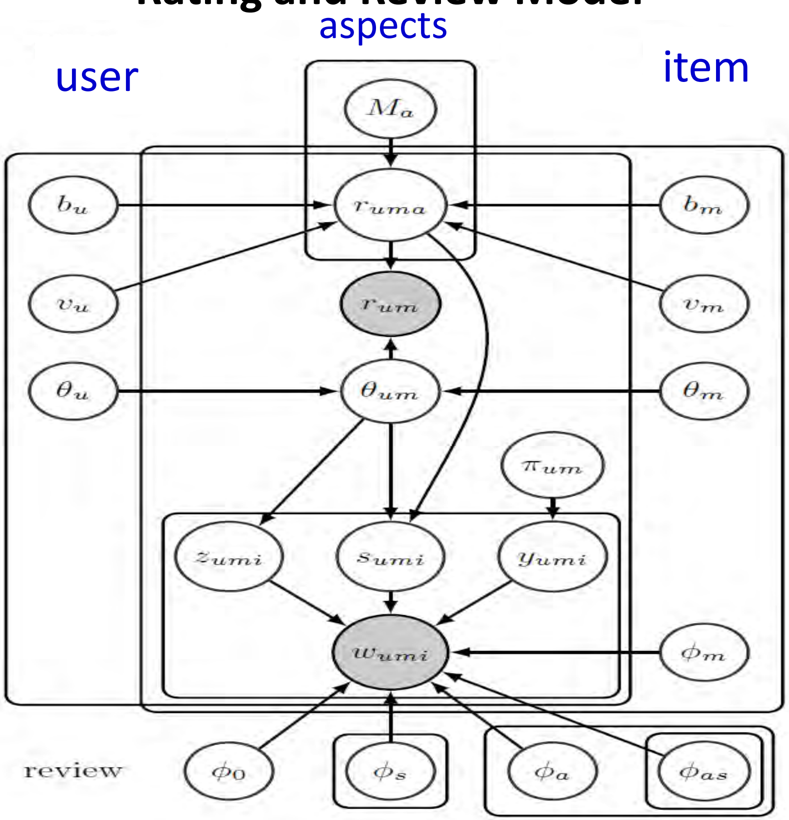


Align these two tasks by using a transform from item and user latent vectors to topic distribution parameters so as to combine latent factors in rating data with topics in user-review text.

TopicMF: Simultaneously Exploiting Ratings and Reviews for Recommendation, by Yang Bao, Hui Fang, Jie Zhang, AAAI,2014

JMARS Model

Rating and Review Model



Aspect Modeling

(Interest of the user/ Property of the item)

$$\theta_u \sim \mathcal{N}\left(0, \sigma^2{}_{\text{user}-\text{aspect}} I\right)$$

$$\theta_m \sim \mathcal{N}\left(0, \sigma^2{}_{\text{item}-\text{aspect}} I\right)$$

$$\theta_{um} \sim exp(\theta_u + \theta_m)$$

Rating Modeling

Aspect-specific rating

$$v_u{}^T M_a v_m + b_u + b_m$$

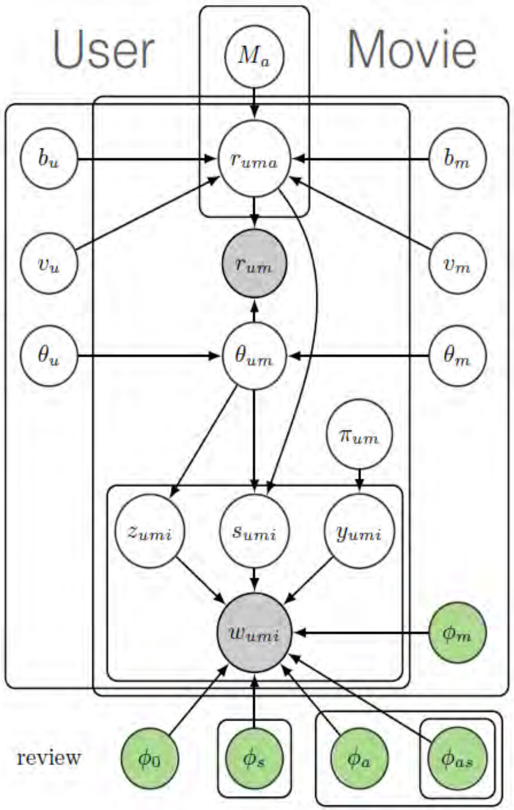
Overall Rating

$$\hat{\boldsymbol{r}}_{\boldsymbol{um}} = \sum_{\boldsymbol{a}} \boldsymbol{\theta}_{\boldsymbol{uma}} \boldsymbol{r}_{\boldsymbol{uma}}$$

M_a emphasizes the aspect specific properties.

Jointly Modeling Aspects, Ratings and Sentiments for Movie Recommendation (JMARS), by Qiming Diao, Minghui Qiu, Chao-Yuan Wu, KDD 2014

JMARS Language Model



We assume that the review language model is given by a convex combination of five components.

1. **Background ϕ_0**: words uniformly distributed in every review.
2. **Sentiment ϕ_s**: not aspect specific content such as great, good, bad.
3. **Item-specific ϕ_m**: any term that appears only in the item.
4. **Aspect-specific ϕ_a**: words associated with specific aspects. Music, sound, singing
5. **Aspect-specific sentiment words ϕ_{as}**: to express positive or negative sentiments.

Each of the language models is a multinomial distribution.

Inference and Learning
(hybrid sampling/optimization)

- The goal is to learn the hidden factor vectors, aspects, and sentiments of the textual content to accurately model user ratings and maximize the probability of generating the textual content.
- Use Gibbs-EM, which alternates between collapsed Gibbs sampling and gradient descent, to estimate parameters in the model.
- The Objective function consist of two parts :
  1. The prediction error on user ratings.
  2. The probability of observing the text conditioned on priors.

$$\mathcal{L} = \sum_{r_{um} \in \mathcal{R}} [\epsilon^{-2}(r_{um} - \hat{r}_{um})^2 - \log p(w_{um} | \Upsilon, \Omega)]$$

Qualitative Evaluation

To evaluate if the model is capable of interpreting the reviews correctly, the learned aspect ratings are examined.

Aspect rating, Sentiment and Movie-Specific Words

Aspect	Director	History	War	Life	Character
Rating	9.36	8.55	8.51	9.20	9.50
Prob	0.12	0.10	0.09	0.09	0.08

... what an excellent piece of cinema ... the actors are great and directing incredible ... in 300, Gerard Butler dominates the screen ... battle scenes are incredible ...

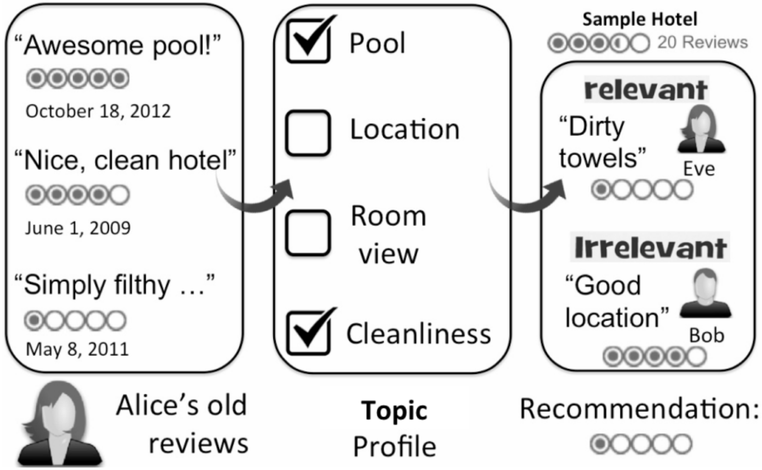
Table 4: The learnt aspect-specific ratings and latent sentiment identified by our model for a review.

Positive words

Movie-specific

TPCF

- Topic creation and manual verification and grouping.
- Faceted opinion extraction
- Based on the topics a user writes about, create an individual interest topic profile.
- Personalize the product rankings for each user, based on the reviews that are most relevant to her profile.



Recommendation Using Textual Opinions
Claudiu-Cristian Musat, Yizhong Liang, Boi Faltings, IJCAI 2013

Sentiment Utility Logistic Model

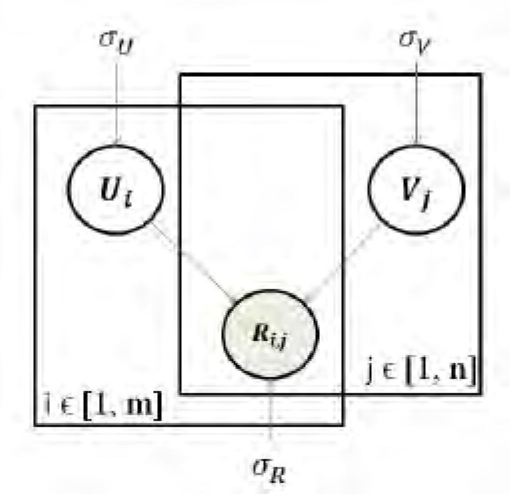
- SULM builds user and item profiles
  - Predicts overall rating for a review
  - Estimating sentiment utilities of each aspect.
- SULM identifies the most valuable aspects of future user experiences.
- Uses Opinion Parser for aspect extraction and aspect sentiment classification.
- SULM estimates the sentiment utility value for each of the aspects k using the matrix factorization approach.
- Tested on actual reviews across three real-life applications.

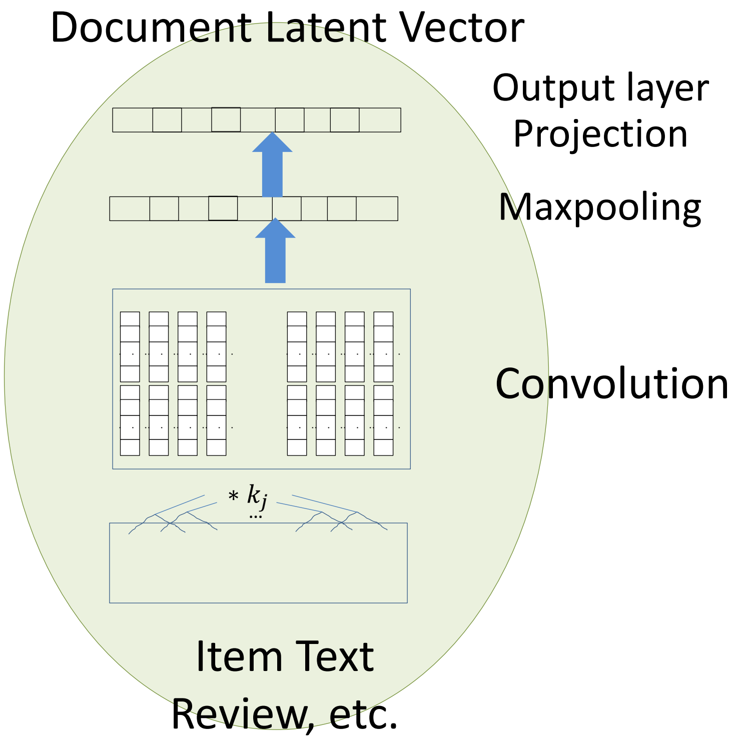
Aspect Based Recommendations: Recommending Items with the Most Valuable Aspects Based on User Reviews, by Konstantin Bauman, Bing Liu, Alexander Tuzhilin, KDD'17

DEEP LEARNING BASED REVIEW AWARE RECOMMENDATION

CNN to find Document Latent Vector PMF

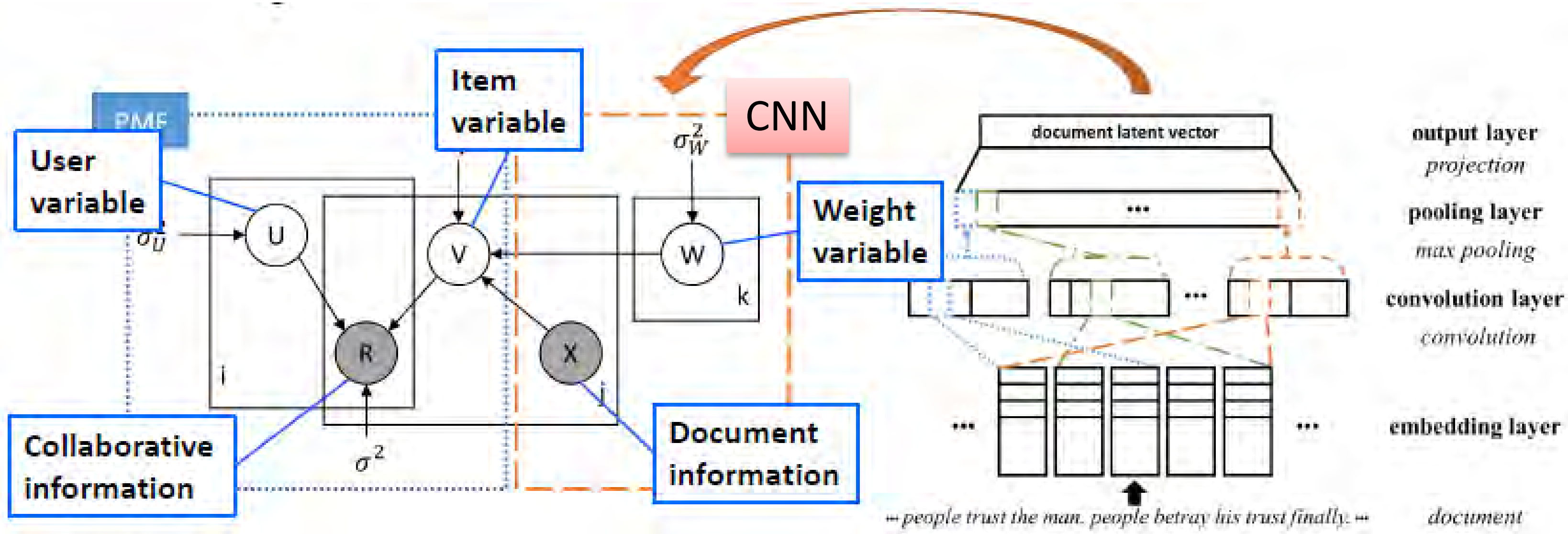
- CNN Architecture to generate document latent vector.
- Ratings approximated by probabilistic methods.





ConvMF

- Integrate CNN into PMF for the recommendation task.
- Item variable plays a role of the connection between PMF and CNN in order to exploit ratings and description documents.



Optimization

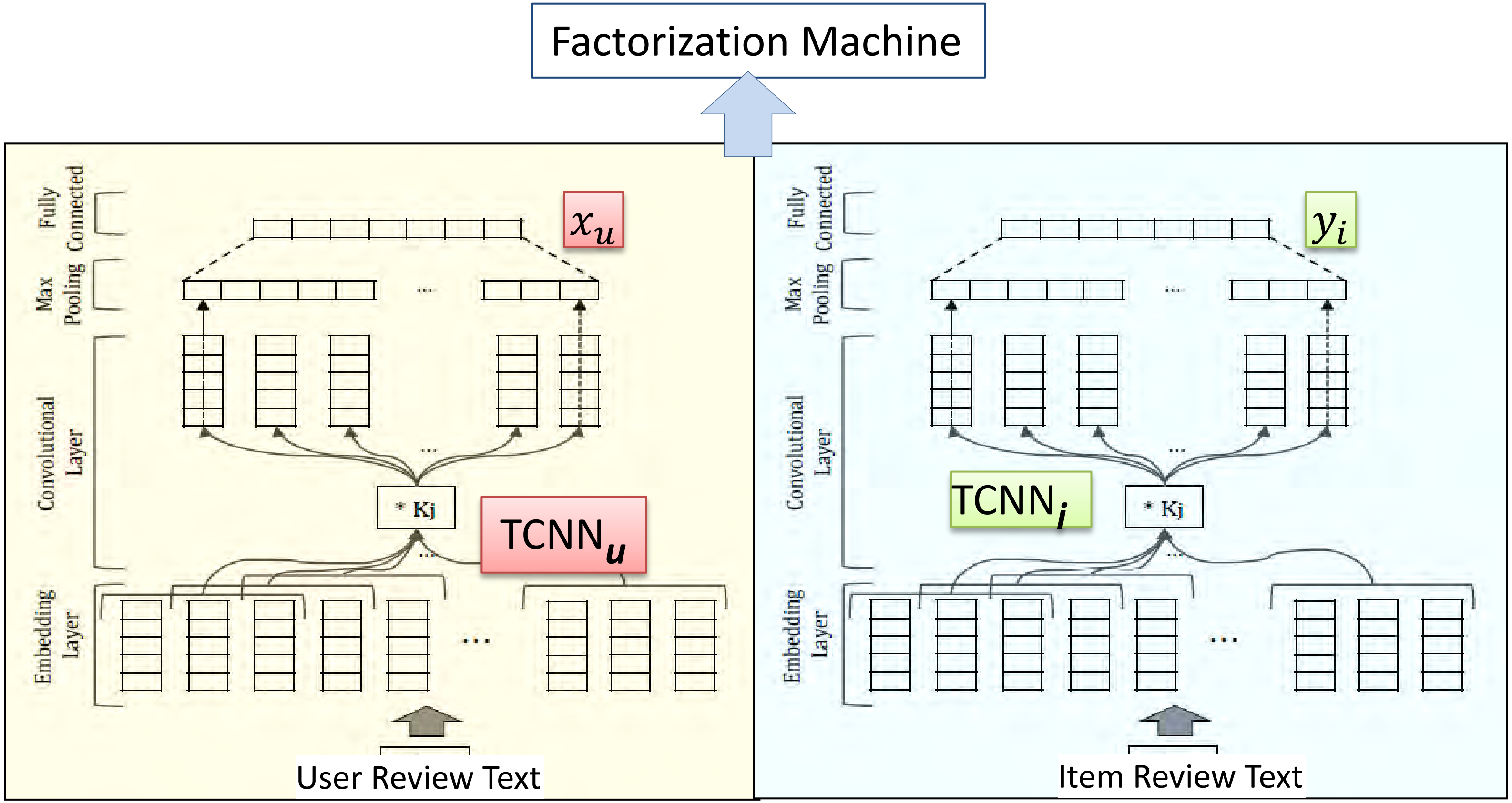
- Use Maximum a posteriori to solve U, V, W
- When U, V are temporarily fixed, to optimzie W, use backprop with given target variable v_j
- ConvMF+ used pretrained word embedding.

Deep Cooperative Neural Network

- (DeepCoNN) models users and items jointly.
- Two parallel convolutional neural networks to model user behaviors and item properties from review texts.
- Utilizes a word embedding technique to map the review texts into a lower-dimensional semantic space as well as keep the words sequence information.
- The outputs of the two networks are concatenated.
- The factorization machine captures their interactions for rating prediction.

Joint deep modelling of Users and Items using Review for Recommendation, Lei Zheng, Vahid Noroozi, Philip S. Yu, WSDM 2017

DeepCONN Architecture



User Network

Item Network

User and Item Network

User Network

Lookup Layer:

- Reviews are represented as a matrix of word embeddings.
- All reviews written by user u are merged into a single document

$$V_{1:n}^{u} = \phi\left(d_1^u\right) \oplus \phi\left(d_2^u\right) \oplus \phi\left(d_3^u\right) \oplus \cdots \oplus \phi\left(d_n^u\right)$$

Item Network

Lookup Layer:

- All reviews written on item i are merged into a single document

CNN Layers:

Each neuron j in the convolutional layer uses filter $\mathrm{k_j}$ on a window of words size t.

Max-pooling layer

$$o_j = \max\{z_1, z_2, \cdots, z_{n-t+1}\}$$

Output from multiple filters:

$$O = \{o_1, o_2, \cdots, o_{n1}\}$$

Fully connected layer: weight matrix W.

The output of the fully connected layer is considered as features of U.

Shared Layer

- A shared layer to map output of user and item network into the same feature space.

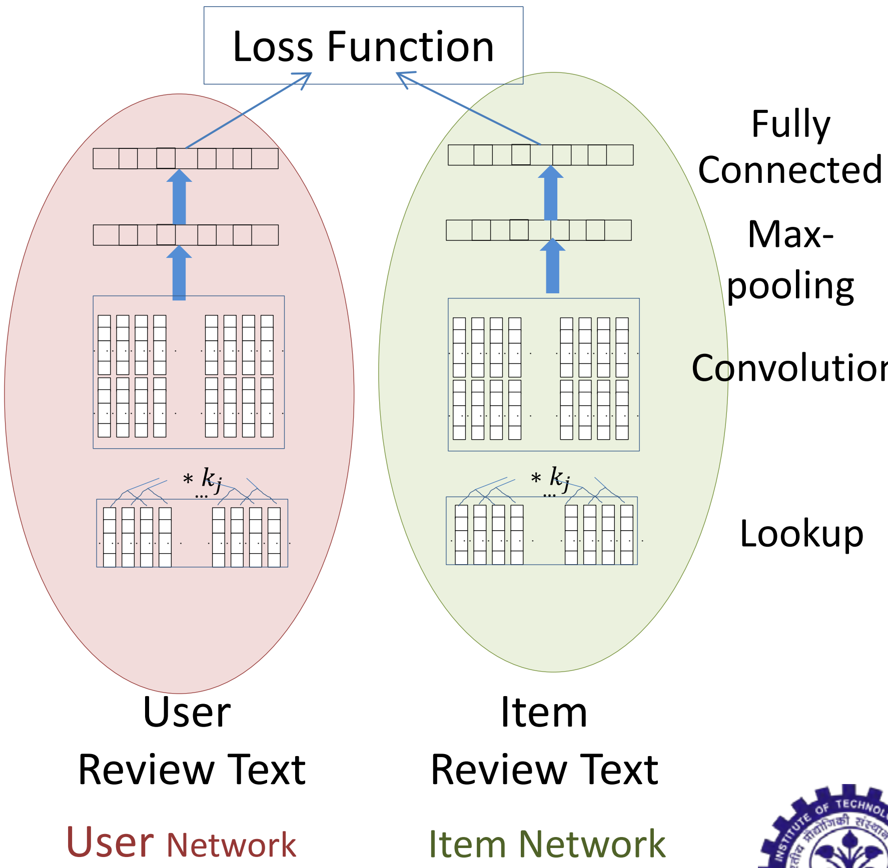
$$\hat{z} = (x_u, y_i)$$

- Factorization Machine(FM) computes the second order interactions between the elements of the input vector.

$$\hat{r}_{ui} = \widehat{w}_0 + \sum_{i=1}^{\hat{z}} \widehat{w}_i \hat{z}_i + \sum_{i=1}^{\hat{z}} \sum_{j=i+1}^{\hat{z}} \langle \hat{v}_i, \hat{v}_j \rangle \hat{z}_i \hat{z}_j$$

- L1 loss:

$$J = \sum_{(u,i,r_{ui}) \in D} |r_{ui} - \hat{r}_{ui}|$$



Experimental Results

Dataset	MF	PMF	LDA	CTR	HFT-10	HFT-50	CDL	DeepCoNN	Improvement of DeepCoNN (%)
Yelp	1.792	1.783	1.788	1.612	1.583	1.587	1.574	**1.441**	8.5%
Amazon	1.471	1.460	1.459	1.418	1.378	1.383	1.372	**1.268**	7.6%
Beer	0.612	0.527	0.306	0.305	0.303	0.302	0.299	**0.273**	8.7%
Average on all datasets	1.292	1.256	1.184	1.112	1.088	1.09	1.081	**0.994**	8.3%

TransNets: Learning to Transform for Recommendation

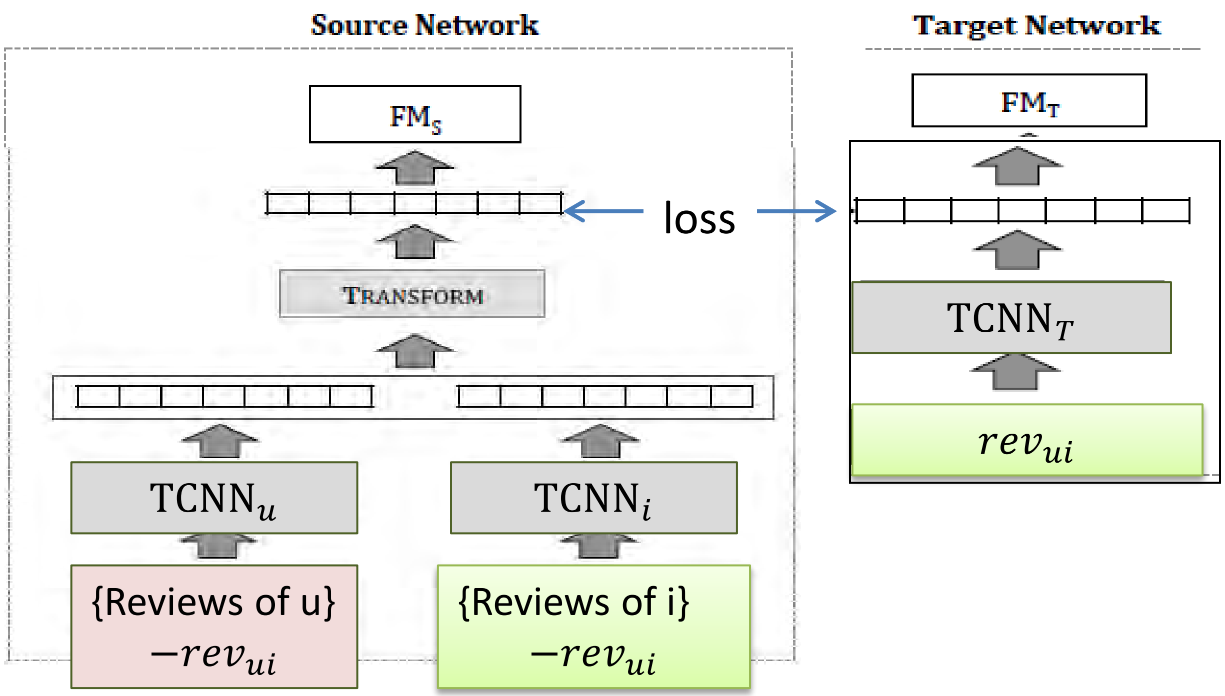
Rose Catherine William Cohen, RecSys 2017

TRANSNETS

TransNets

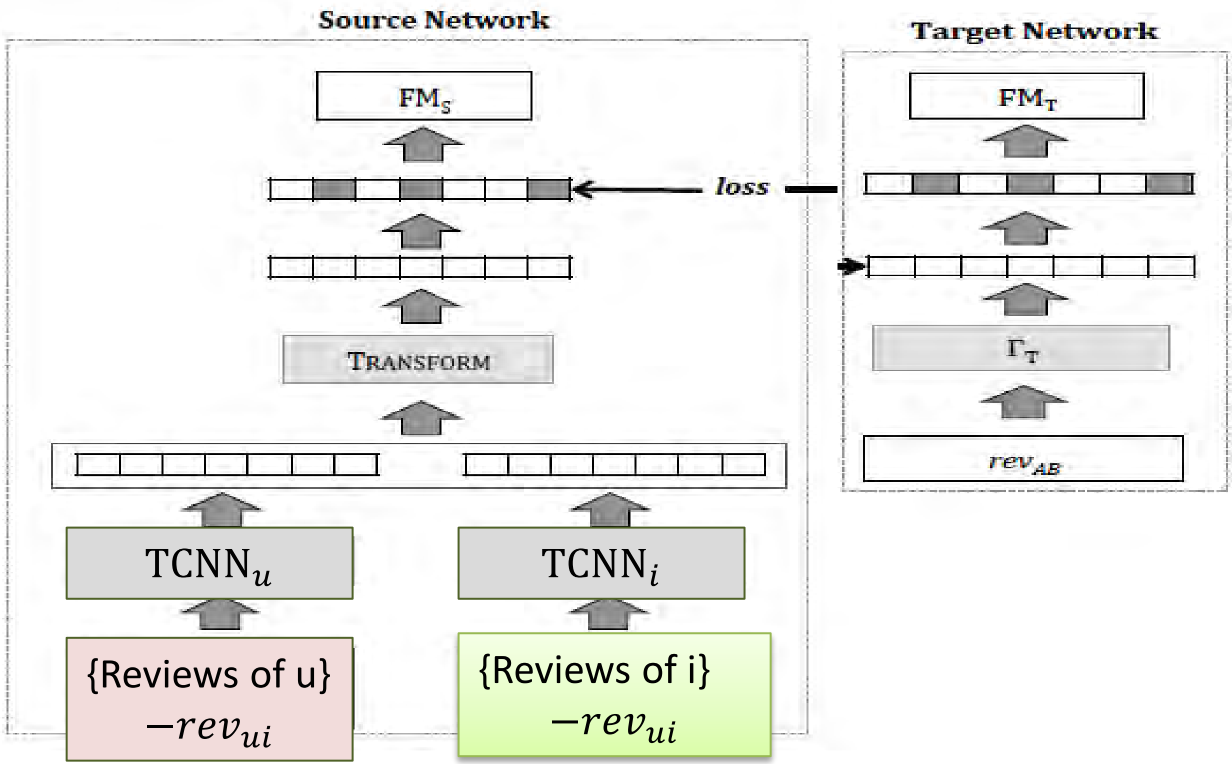
- DeepCONN performance is lower when the corresponding review is not in the training set.
- TransNets extends the DeepCoNN model by introducing an additional latent layer representing an approximation of the review corresponding to the target user-target item pair.
- Regularize this layer at training time to be similar to another latent representation of the target user's review of the target item.

TransNet



Source Network
Target Network
FM_S
FM_T
loss
TRANSFORM
$TCNN_T$
rev_{ui}
$TCNN_u$
$TCNN_i$
{Reviews of u}
$-rev_{ui}$
{Reviews of i}
$-rev_{ui}$

TransNets



Source Network
Target Network
FM_S
FM_T
loss
TRANSFORM
Γ_T
rev_{AB}
$TCNN_u$
$TCNN_i$
{Reviews of u}
$-rev_{ui}$
{Reviews of i}
$-rev_{ui}$

Training TransNets

- During training, force the Source Network's representation z_L to be similar to the encoding of rev_{ui} produced by the Target Network.

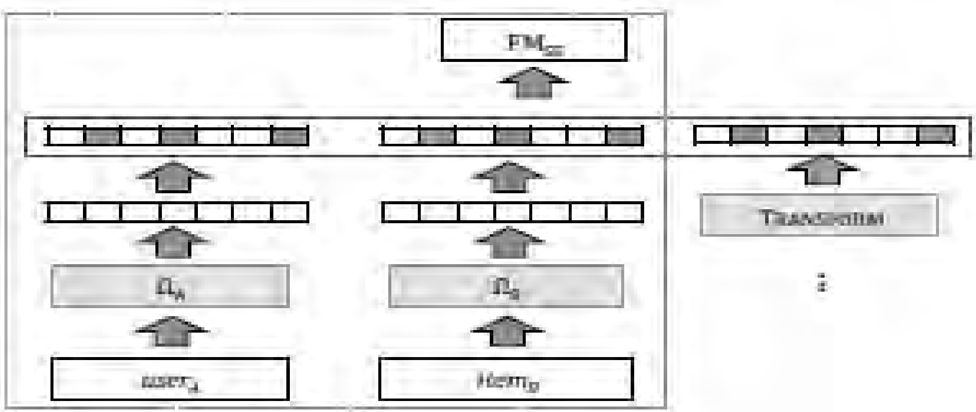
Repeat for each review:

- Step 1: Train Target Network on the actual review. Minimize L1 loss: $|r_{ui} - \hat{r}_T|$
- Step 2: Learn to Transform: minimize a L2 loss computed between z_L and the representation x_T of the actual review.
- Step 3: Train a predictor on the transformed input: Paramaters of FM_S are updated to minimize a L1 loss : $|r_{ui} - \hat{r}_S|$

At test time, TransNet uses only the Source Network.

Extended TransNet

- TransNet does not use user, item identity.
- Ext-TN learn a latent representation of the users and items, similar to Matrix Factorization methods.
- Input to FM_S: latent user and item representation concatenated with the output of Transform layer.

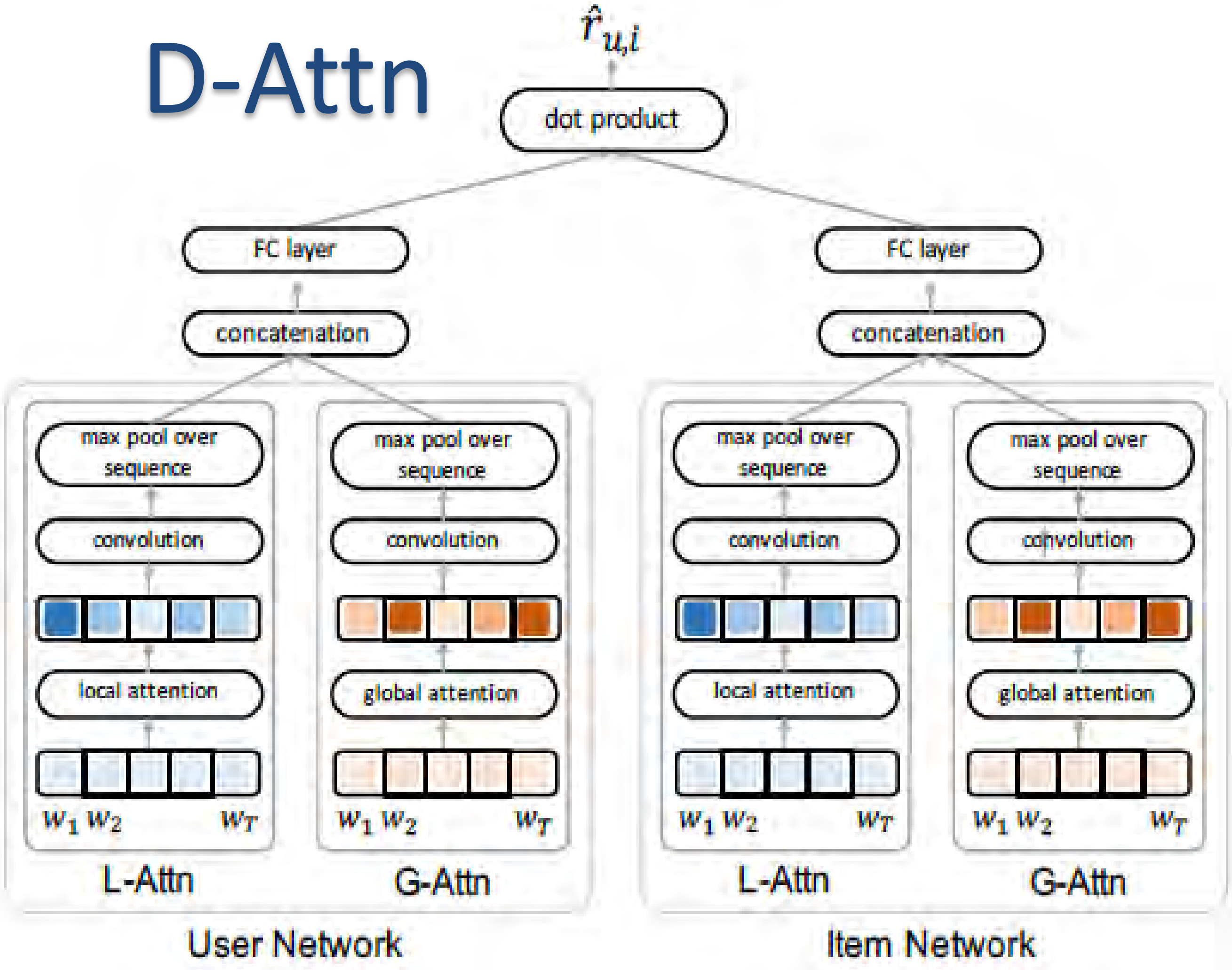


Experimental Results

Performance comparison using MSE

Dataset	DeepCoNN + Test Reviews	MF	DeepCoNN	DeepCoNN-rev$_{AB}$	TransNet	TransNet-Ext
Yelp17	1.2106	1.8661	1.8984	1.7045	1.6387	1.5913
AZ-Elec	0.9791	1.8898	1.9704	2.0774	1.8380	1.7781
AZ-CSJ	0.7747	1.5212	1.5487	1.7044	1.4487	1.4780
AZ-Mov	0.9392	1.4324	1.3611	1.5276	1.3599	1.2691

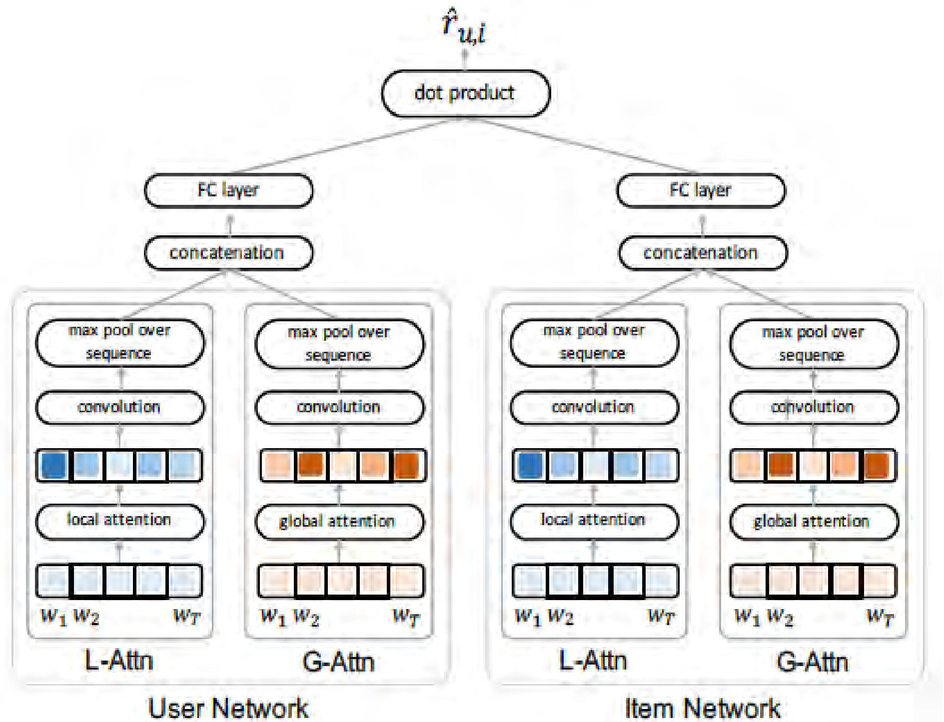
D-Attn



Interpretable Convolutional Neural Networks with Dual Local and Global Aention for Review Rating Prediction Sungyong Seo, Jing Huang, Hao Yang, Yan Liu

D-Attn

- The local attention model (L-Attn) selects informative keywords from a local window.
- The global attention model (G-Attn) ignores irrelevant words from long review sequences.
- The outputs of L-Attn and G-Attn are concatenated and followed by a FC layer.
- Dot product of the vectors from both networks estimate ratings.



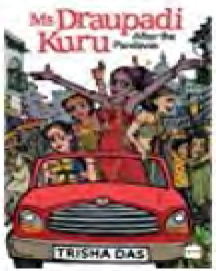
EXPLAINABILITY

Explainable Recommendation

- Why are these items recommended?

Your recently viewed items and featured recommendations

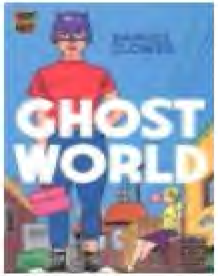
Inspired by your purchases



Ms Draupadi Kuru: After the Pandavas
› Trisha Das
★★★★☆ 18
Paperback

Legends Of Halahala
Appupen
★★★★☆ 4
Paperback

Maid in India: Stories of Inequality and...
› Tripti Lahiri
★★★★☆ 4



Ghost World
› Daniel Clowes
★★★★☆ 4
Paperback

Flipkart

Reviews Explain Ratings

Use Different Charger

By RAKESHBS on 23 August 2017

Colour: Gold

Nice Phone ..

Pros:

1) Comfortable to hold in one hand
2) Good Camera Clarity
3) Finger print sensor works fast
4) Performance is good

Cons:

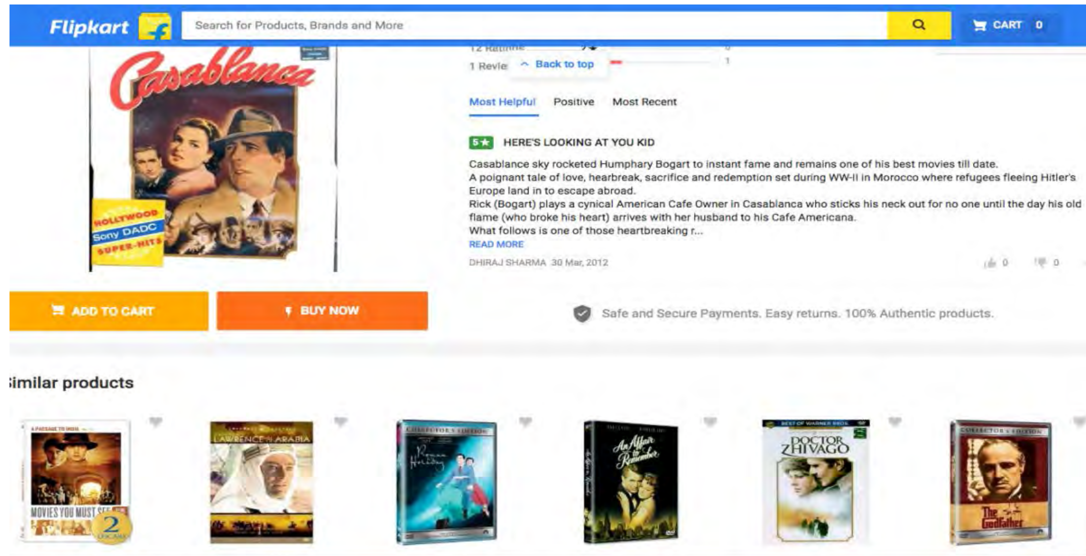
1) it will get heated up when used with Charger shipped with Xiomi ..So I suggest to use other charger ...Like Nilkin 2A Charger or any other good 2A charger ...
2) Getting a good Tempered Glass will be costly and most of the time it is difficult to find one for 2.5D glass
3) Hybrid sim slot ...

- Reviews justify a user's rating:
  - by discussing the specific properties of items (aspects)
  - by revealing which aspects the user is most interested in.

Flipkart

Hidden factor/topics

- Interpretable textual labels for rating justification.
- Topics obtained explain variation in rating/review data



Hidden factors and hidden topics: understanding rating dimensions with review text, Mcauley et al, RecSys 2013.

EFM

Explicit Factor Models for Explainable Recommendation based on Phrase-level Sentiment Analysis

Yongfeng Zhang, Guokun Lai, Min Zhang, Yi Zhang, Yiqun Liu,Shaoping Ma, SIGIR 2014

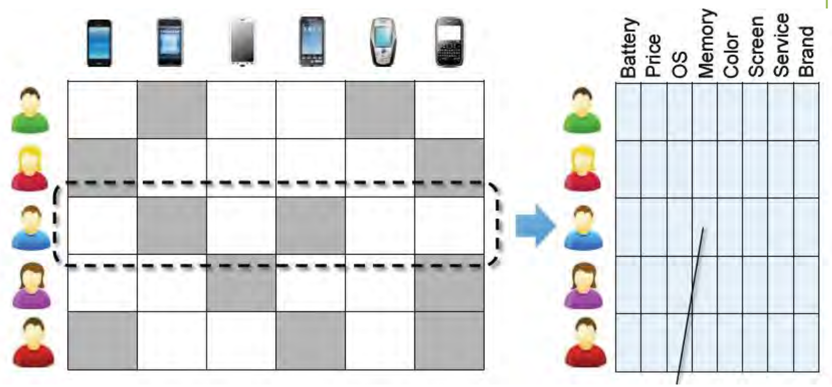
EFM

User-Feature Attention Matrix X

- Suppose feature F_j is mentioned by user u_i for t_{ij} times.

$$X_{ij} = \begin{cases} 0, if\ u_i \text{ did not mention } F_j \\ 1 + (N-1)\left(\dfrac{2}{1+e^{-t_{ij}}} - 1\right) & \text{otherwise} \end{cases}$$

- Rescale t_{ij} into the same range [1,N] as the rating matrix A by reformulating the sigmoid function.



Item-Feature Quality Matrix Y

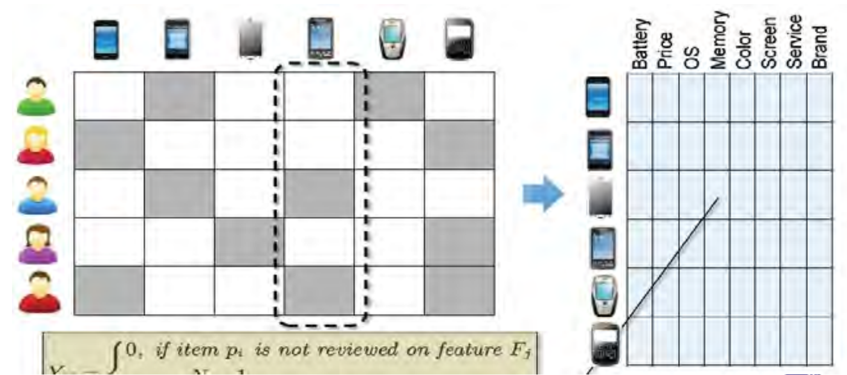
For each item p_i,

extract the corresponding (F, S') pairs form all its reviews.

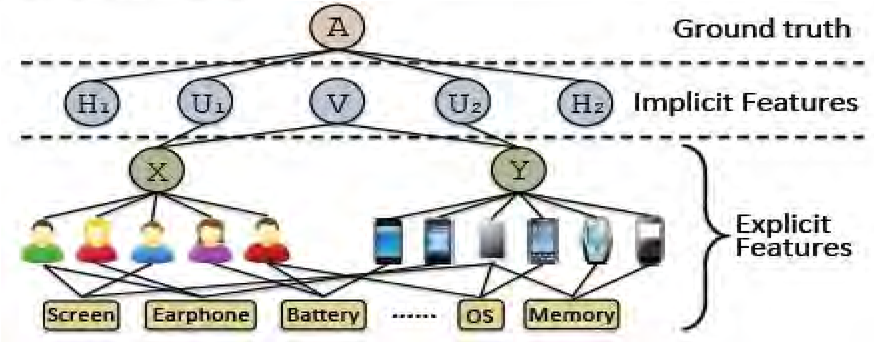
If feature F_j is mentioned for k times on item p_i, and the average of sentiment of feature F_j in those mentions are s_{ij}.

$$Y_{ij} = \begin{cases} 0, \text{if item } p_i \text{ is not reviewed on } F_j \\ 1 + (N-1)\left(\dfrac{2}{1+e^{-k.s_{ij}}} - 1\right) & \text{otherwise} \end{cases}$$

Rescaled to [1,N]



Factorization Model: Integrating Explicit and Implicit Features



- Build a factorization model over user-feature attention matrix X and item-feature quality matrix Y.

$$\underset{U_1, U_2, V}{\text{minimize}} \left\{\lambda_x \|U_1.V - X\|_F^2 + \lambda_y \|U_2.V - Y\|_F^2\right\}$$

- Introduce r' latent factors: H_1 and H_2.
- We use $P = [U_1 H_1]$ and $Q = [U_2 H_2]$ to model the overall rating matrix A.
- U1: explicit factors
- H1: Hidden factors

$$\underset{P, Q}{\text{minimize}} \left\{ \|P.Q - A\|_F^2 \right\}$$

$$\boxed{\underset{U_1, U_2, V, H_1, H_2}{\text{minimize}} \left\{\|P.Q - A\|_F^2 + \lambda_x \|U_1.V - X\|_F^2 + \lambda_y \|U_2.V - Y\|_F^2 + \cdots\right\}}$$

Explanation Generation

- Template based explanation and an intuitional word cloud based explanation.
- Feature-level explanation for a recommended item.
- Provide disrecommendations by telling the user why the current browsing item is disrecommended.

You might be interested in <feature> on which this product performs well

You might be interested in <feature> on which this product performs poorly.

Flipkart

Trirank Approach to Aspect based Recommendation

Offline Training

- Extract aspects from user reviews
- Build tripartite graph of User, Product, Aspect
- Graph Propagation: Label propagation from each vertex
- Machine learning for graph propagation

TriRank: Review-aware Explainable Recommendation by Modeling Aspects: Xiangnan He, Tao Chen, Min-Yen Kan, Xiao Chen

Inputs:

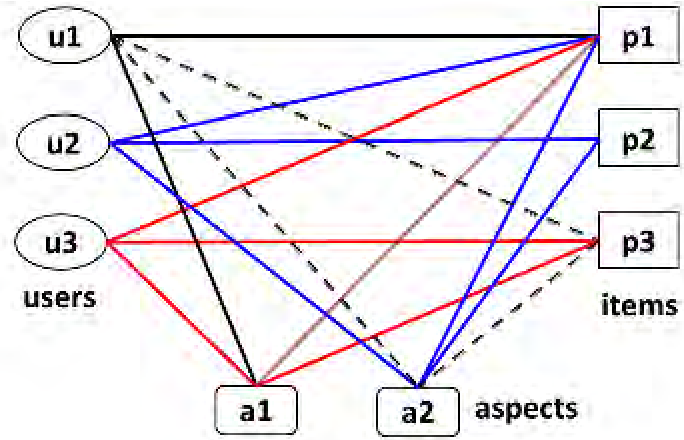
<u1, p1, a1>

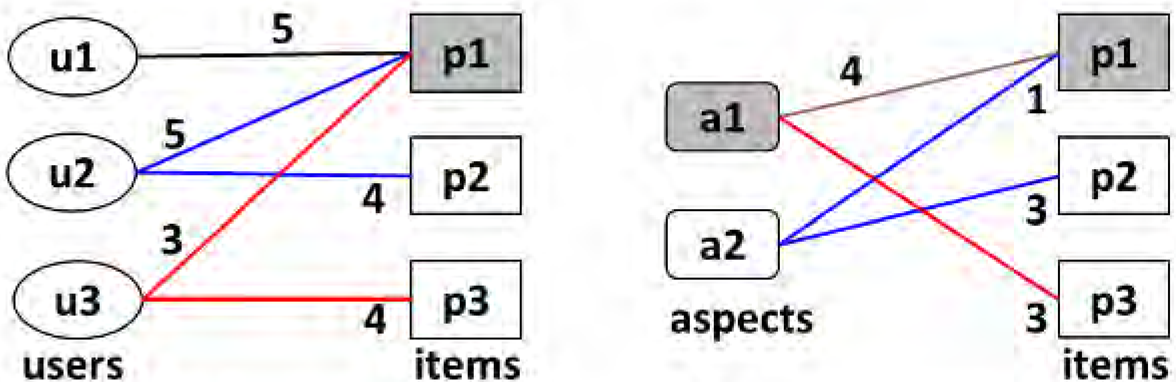
<u2, p1, a2>

<u2, p2, a2>

<u3, p1, a1>

<u3, p3, a1>



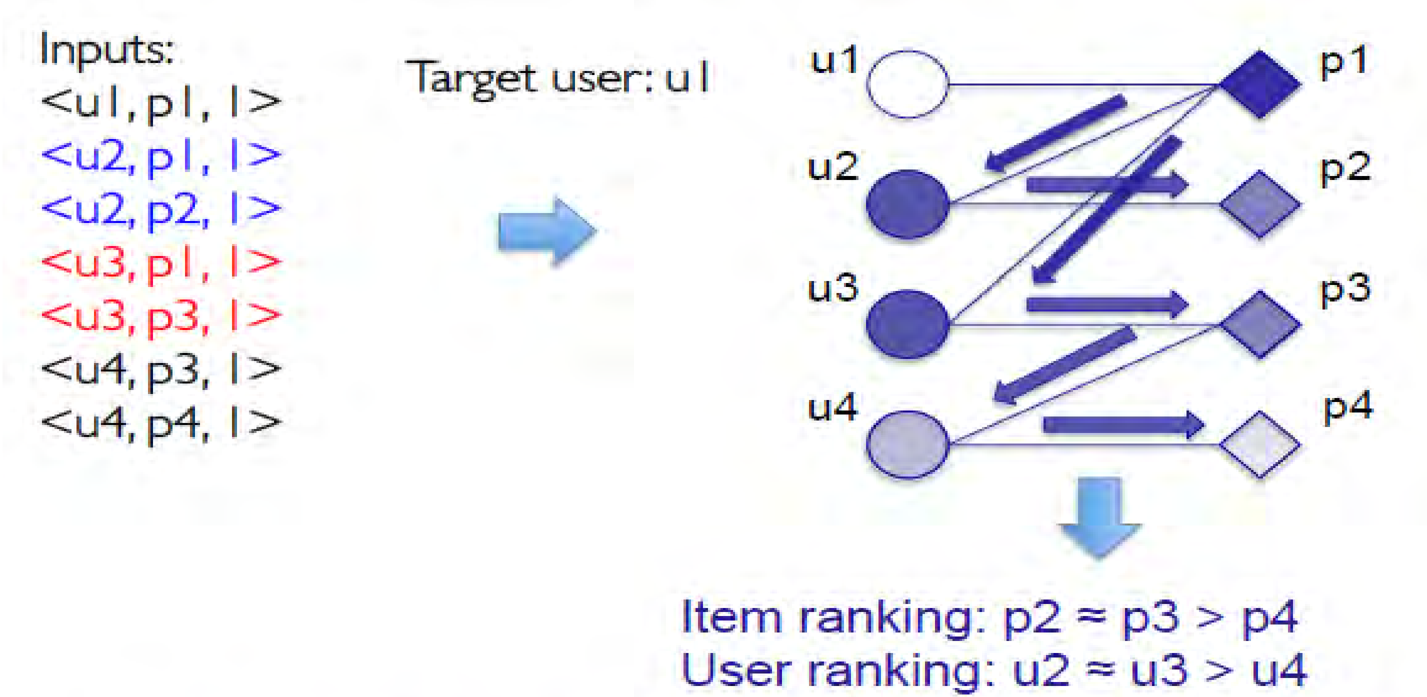


(a) User-Item structure (b) Item-Aspect structure

Trirank – Review Aware Recommendation

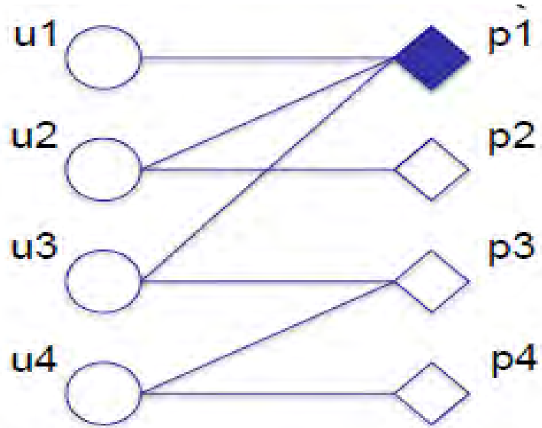
- Uses User-Item-Aspect ternary relation
  - Represented by heterogeneous tripartite graph
- Model reviews in the level of aspects
- Graph-based method for vertex ranking on tripartite graph accounting for
  - the structural smoothness (encoding collaborative filtering and aspect filtering effects)
    Smoothness implies local consistency: that nearby vertices should not vary too much in their scores.
  - fitting constraints (encoding personalized preferences, prior beliefs).
- Offline training + online learning.
  - Provide instant personalization without retraining.

Basic Idea: Graph Propagation



Label Propagation from the target user’s historical item nodes captures Collaborative Filtering

Machine Learning for Graph Propagation



- Input:
  - Graph Structure (Matrix Y)
  - Initial labels to propagate (vectors p^0)
- Output:
  - Scores of each vertex (vectors u, p)

Smoothness Regularizer: Nearby vertices should not vary too much

$$\sum_{i \in U} \sum_{j \in P} y_{ij} \left(\frac{u_i}{\sqrt{d_i}} - \frac{p_j}{\sqrt{d_j}} \right)^2$$

Fitting Constraints (initial labels):
Ranking scores should adhere to the initial labels.

$$\sum_{j \in P} \left(p_j - p_j{}^0\right)^2$$

Optimization (Gradient Descent):

$$p = S_Y u + p^0$$

$$u = S_Y{}^T p, \qquad \text{where } S_Y = \left[\frac{y_{ui}}{\sqrt{d_u d_i}} \right]$$

[He et al, SIGIR 2014]

Connection to CF models

- Recap: ranking loss function (for a target user):

$$\underbrace{\boxed{\sum_{j \in P} (p_j - p_j^0)^2}}_{\text{Prediction loss}} + \underbrace{\lambda \sum_{i \in U} \sum_{j \in P} y_{ij} \left(\frac{u_i}{\sqrt{d_i}} - \frac{p_j}{\sqrt{d_j}} \right)^2}_{\text{Regularizations}}$$

- Traditional machine learning-based CF models:
  1. Prediction model:

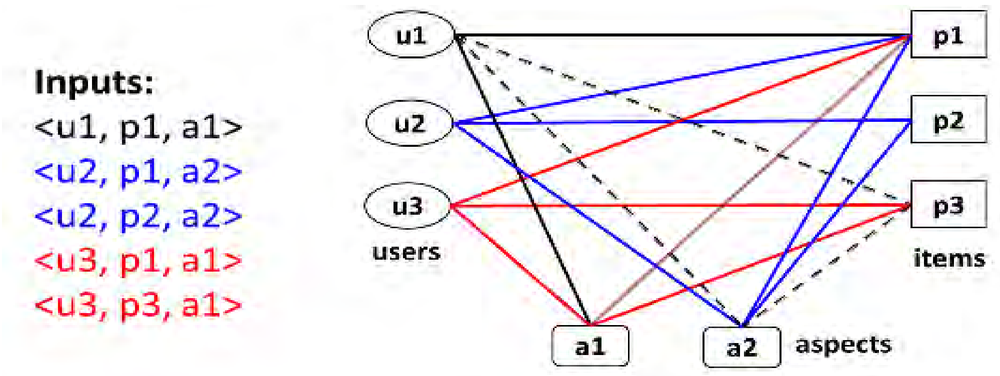
     E.g., matrix factorization: $\hat{y}_{ui} = < v_u, v_i >$
  2. Loss function:

$$\sum_{u \in U} \boxed{\sum_{i \in I} (y_{ui} - \hat{y}_{ui})^2}$$

Prediction loss on all items (include imputations).
(important for top-K recommendation)

Trirank Approach

- Graph propagation in the tripartite graph:



$$\mathbf{u} = \mathbf{u}^0 + \lambda_1 UP \cdot \mathbf{p} + \lambda_2 UA \cdot \mathbf{a}$$
$$\mathbf{p} = \mathbf{p}^0 + \lambda_3 PU \cdot \mathbf{u} + \lambda_4 PA \cdot \mathbf{a}$$
$$\mathbf{a} = \mathbf{a}^0 + \lambda_5 AP \cdot \mathbf{p} + \lambda_6 AU \cdot \mathbf{u}$$

Initial labels should encode:

- Target user's preference on aspects/ items/ users:
  - a^0: reviewed aspects.
  - p^0: ratings on items.
  - u^0: similarity with other users (friendship).

Online Learning

<u>Offline Training</u> (for all users):

1. Extract aspects from user reviews
2. Build the tripartite graph model (edge weights)
3. Label propagation from each vertex and save the scores.
   - *i.e.* store a |V|×|V| matrix $f(v_i, v_j)$.

Online Recommendation (for target user u_i):

1. Build user profile (*i.e.*, L_u vertices to propagate from).
2. Average the scores of the L_u vertices:

$$y_j = \frac{1}{|L_u|} \sum_{v_u \in L_u} f(v_u, v_j)$$

Flipkart

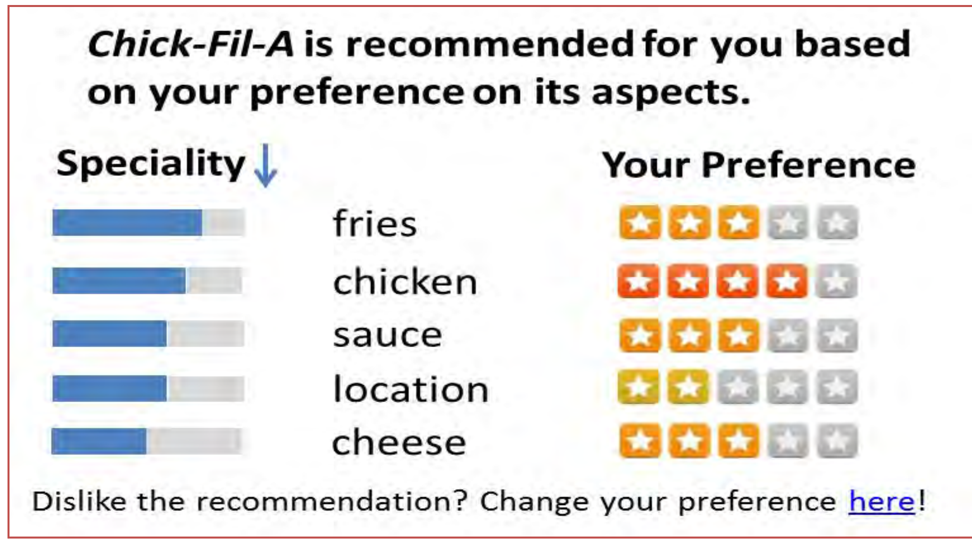
Explainability

- Collaborative filtering + Aspect filtering

(Similar users also choose the item)

(Reviewed aspects match with the item)

  - Item Ranking
  - Aspect Ranking
  - User Ranking

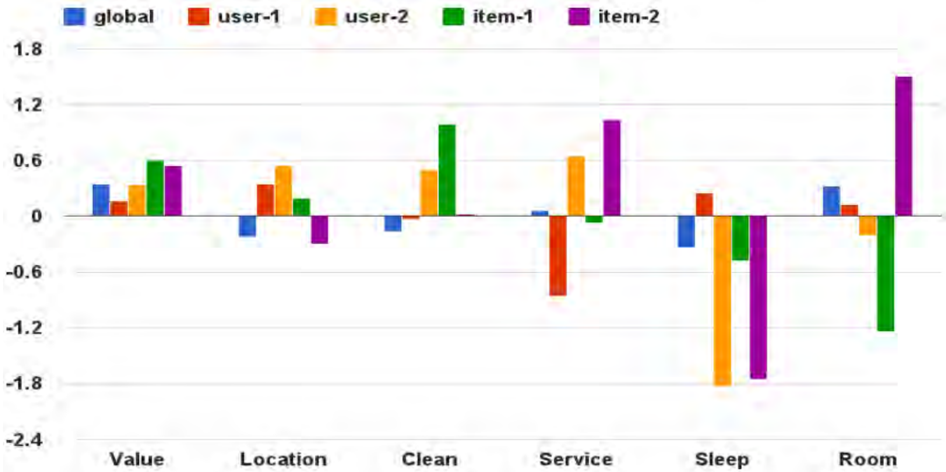


FLAME

- FLAME learns users' personalized preferences on different aspects from their past reviews, and predicts users' aspect ratings on new items by collective intelligence.
- Propose a new model combining aspect-based opinion mining and collaborative filtering to collectively learn users' preferences on different aspects.
- Preference diversity: For example, the food of the same restaurant might be delicious for some users but terrible for others.
- To address the problem of Personalized Latent Aspect Rating Analysis, we propose a unified probabilistic model called Factorized Latent Aspect ModEl (FLAME), which combines the advantages of both collaborative filtering and aspect-based opinion mining

FLAME: Probabilistic Model Combining Aspect Based Opinion Mining and Collaborative Filtering, Yao Wu and Martin Ester, WSDM'15

Aspect Weights, Aspect Values



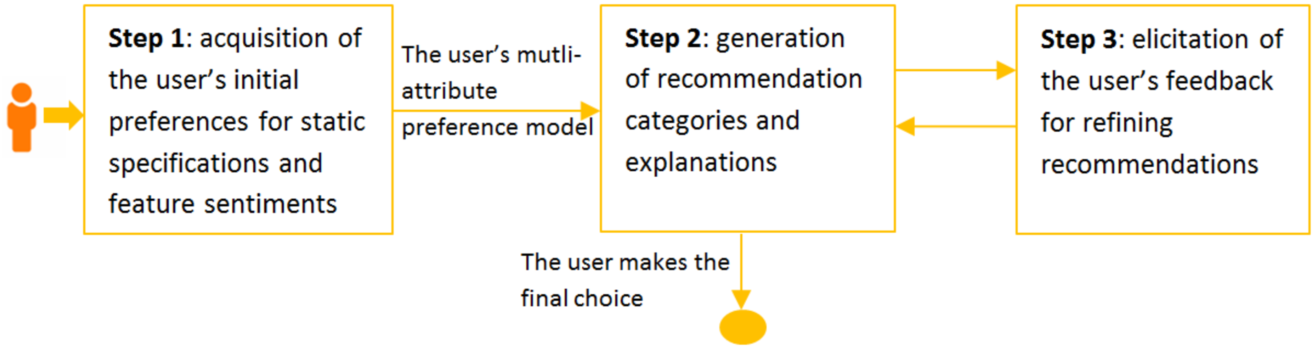
- user-1 likes to comment on the Location and Sleep
- user-2 cares more about the Service, Clean and Location.

Produce more persuasive recommendation explanations by the predicted aspect ratings and some selected reviews written by similar users.

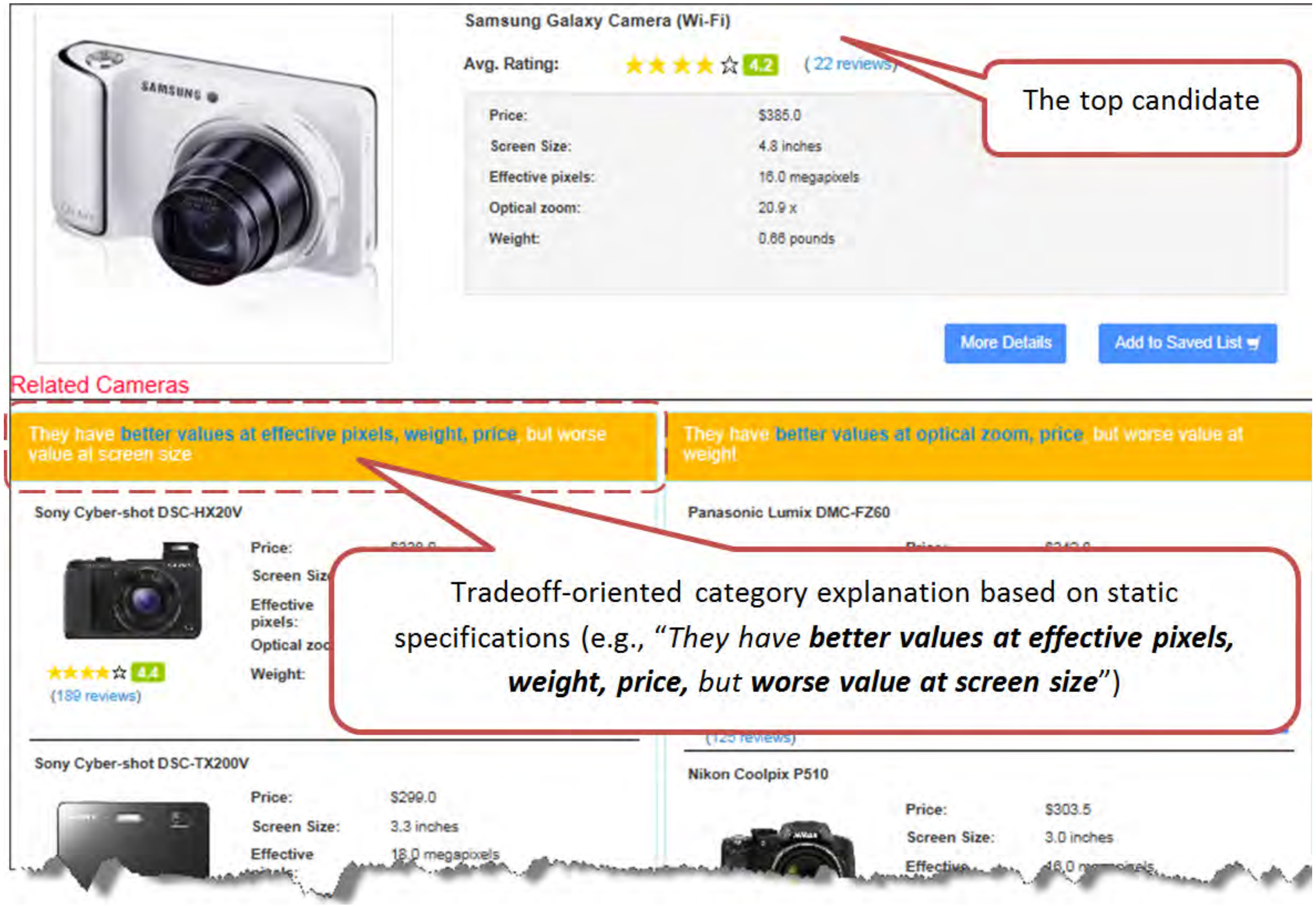
(a) Location (b) Location 2-star (c) Location 5-star

Senti-ORG: Explanation Interface

- An explanation interface that emphasizes explaining the tradeoff properties within a set of recommendations in terms of both their static specifications and feature sentiments extracted from product reviews.
- Assist users in more effectively exploring and understanding product space.



Explaining Recommendations Based on Feature Sentiments in Product Reviews
Li Chen, F Wang, IUI 2017


Samsung Galaxy Camera (Wi-Fi)
Avg. Rating: 4.2 (22 reviews)
Price: $385.0
Screen Size: 4.8 inches
Effective pixels: 16.0 megapixels
Optical zoom: 20.9 x
Weight: 0.66 pounds
More Details
Add to Saved List
Related Cameras
They have better values at effective pixels, weight, price, but worse value at screen size
They have better values at optical zoom, price, but worse value at weight
Sony Cyber-shot DSC-HX20V
Price:
Effective pixels:
Weight:
4.4
(189 reviews)
Panasonic Lumix DMC-FZ60
Sony Cyber-shot DSC-TX200V
Price: $299.0
Screen Size: 3.3 inches
Effective
Nikon Coolpix P510
Price: $303.5
Screen Size: 3.0 inches
The top candidate
Tradeoff-oriented category explanation based on static specifications (e.g., "*They have* ***better values at effective pixels, weight, price,*** *but* ***worse value at screen size***")

Future Trends

- Explanation Interfaces
- User control on recommendations.
- Use with other knowledge sources: context, ontology, …

Use of reviews beyond recommendations.

- Addressing Complex and Subjective Product-Related Queries, (McAuley WWW'16)

Thank You