Proposal for RecSys '18 tutorial

1. Tutorial Title: Product review recommendation: ranking, selection and personalization
2. Tutorial Length: 90 minutes
3. Motivation for proposing this tutorial

Reviews ranked by their helpfulness facilitate purchase decisions by potential buyers. For a balanced coverage though, such selection of top-k reviews should span the entire spectrum of aspects and sentiments of relevant products. Personalization makes review recommendation more effective as user preference varies not only across authors but also voters of reviews and raters of products. Review text helps capture user idiosyncrasies; a variety of techniques exist to represent this diverse data across users/products thus leading to further advances in factorization and AI methods for showing useful reviews to consumers - an important need in online shopping. Proposed tutorial thus covers trending topics of interest to RecSys community: ranking, elicitng/interpreting user preferences and deep learning with a unique perspective synergizing NLP(review text) with data mining (utility prediction).

4. Name, email address, and affiliation of tutorial instructor(s).

a. Dr.Muthusamy Chelliah,
Director,Academic engagement, Flipkart, Bangalore, India
(muthusamy.c@flipkart.com) (+91 9845197259)

b. Dr. Sudeshna Sarkar,
Professor and Head, Department of Computer Science & Engineering, and Head, Centre of Artificial Intelligence, IIT Kharagpur, India
(sudeshna@cse.iitkgp.ernet.in) (+91 9434016956)

5. Detailed bulleted outline of the tutorial

1) Introduction (10 minutes)

a) Role of Reviews: Reviews form a low-cost and efficient feedback channel
   i) enable buyers to record their sentiments about the product and its aspects
   ii) thus helping other consumers make purchase decisions and
   iii) Provide feedback to merchants.

b) Addressing Information Overload: Too many reviews for a product cannot be read and hence show only top K reviews

c) Review recommendation: Select a small representative set of high-quality reviews which covers various aspects of the product.

d) Reviews may be ranked based on their quality scores. But the K reviews with the highest scores may contain redundancy in content or opinion [1].

e) Personalized review recommendations are more helpful as different individuals may be interested in different aspects of the product.

f) Identifying product aspects/sentiments: Topic models and word embedding can capture product features, user preferences and reviewer characteristics.

g) Outline: In this tutorial we discuss methods of representing reviews and user interests in terms of the fine grained aspects of product and associated sentiments. We present methods to select reviews for recommendation to a user. We also discuss evaluation methodologies.

2) Review scoring: Helpfulness prediction (10 minutes)

a) Utility score of reviews is typically computed through a function learnt from social (e.g., number of past reviews, PageRank) and structural (e.g., length, readability) features and content of reviews [2].
   i) Social: [3] explores two roles users play in review helpfulness score, e.g., authors with high reputations are likely to write helpful reviews, and product raters find reviews from their connected authors more helpful.
   ii) Structural features of reviews: number of tokens, verbs vs. nouns
   iii) Ranking of reviews by content using convolutional neural network.

b) Personalized scoring: The quality of a review is not independent from its readers - latent factors of whom affect the evaluation.

3) Review Selection: 10 minutes

A user has to be recommended a subset of reviews with following characteristics:

a) Comprehensive: Coverage over al important product aspects
b) Representative: Representative distribution of product aspects and associated opinions as in the entire set of reviews
c) Personalized: Emphasis on aspects of interest to the user.

[4] proposes strategies for pruning review search space and objective functions/greedy algorithms for diverse, quality and representative review sets. [5] leverages concise/ focused snippets for review selection - matching as few review sentences to as many micro-reviews.

4) Context Modeling: 25+20 minutes

a) **Aspect Based Sentiment Analysis (Topic Models):** Review Selection depends on identifying the aspects and their sentiments from individual reviews and aggregating them to form product and user representation. There are different methods for the aspect and semantic extraction
   i) Rule Based
   ii) Topic Model based: We will discuss some of the topic model based methods which are general and portable across domains.

Matrix and tensor factor models help capture latent features of reviews, authors, raters and products [2]. A topic model may be used to jointly discover the underlying aspects and sentiments guided by review helpfulness voting information [6]. Demographic information of review authors is incorporated into topic modeling process in order to discover associations between market segments, topical aspects and sentiments [7]. Latent factors like review expertise, writing style and judgement about fine-grained product aspects are explored for utility prediction through a hybrid HMM-LDA model [8]. [9] helps identify most valuable aspects of user’s potential experience with an item for item recommendation with aspects over which they have control.

b) **Distributed Representation**: A shared layer helps couple latent factors learned for user behavior and item properties from reviews by two parallel networks [10]. Gated recurrent neural networks help translate user and item latent representations into concise abstractive tips with good linguistic quality simulating user experience and feelings [11]. By translating various sources (e.g., review, rating) into a unified representation space, heterogeneous information can be integrated for informed recommendation [12]. Highly-useful reviews are obtained through a novel, attention mechanism and provide explanations for users to make better and faster decisions [13]. [14] combines embedding method with an easy-to-interpret attention network, for explainable recommendations, with a tree-based model learning decision rules from side information (e.g., user demographics).

5) Review Recommendation Systems in an online shopping platform (10 minutes)

a) User-independent [15]
b) Personalized

6) Conclusion (5 minutes)

1. Maroun LB, Moro MM, Almeida JM, Silva AP. Assessing review recommendation techniques under a ranking perspective, ACM Hypertext 2016.
2. Moghaddam, S, Mohsen J, and Martin E. ETF: extended tensor factorization model for personalizing prediction of review helpfulness, WSDM 2012.
3. Tang J, Gao H, Hu X, Liu H. Context-aware review helpfulness rating prediction, RecSys 2013.
4. Xu N, Liu H, Chen J, He J, Du X. Selecting a representative set of diverse quality reviews automatically, SIAM SDM 2014.
5. Nguyen TS,Lauw HW, Tsaparas P.Using micro-reviews to select an efficient set of reviews, CIKM 2013.
6. Hai Z, Cong G, Chang K, Liu W, Cheng P. Coarse-to-fine review selection via supervised joint aspect and sentiment model, SIGIR 2014.
7. Yang Z, Kotov A, Mohan A, Lu S. Parametric and non-parametric user-aware sentiment topic models, SIGIR 2015.
8. Mukherjee S, Popat K, Weikum G. Exploring Latent Semantic Factors to Find Useful Product Reviews, SIAM SDM 2017.
9. Bauman K, Liu B, Tuzhilin A. Aspect based recommendations: Recommending items with the most valuable aspects based on user reviews, KDD 2017.
10. Zheng L, Noroozi V, Yu PS. Joint deep modeling of users and items using reviews for recommendation, WSDM 2017.
11. Li P, Wang Z, Ren Z, Bing L, Lam W. Neural rating regression with abstractive tips generation for recommendation, SIGIR 2017.
12. Zhang Y, Ai Q, Chen X, Croft WB. Joint representation learning for top-n recommendation with heterogeneous information sources, CIKM 2017.
13. Chen C, Zhang M, Liu Y, Ma S. Neural Attentional Rating Regression with Review-level Explanations, WWW 2018.
14. Wang X, He X, Feng F, Nie L, Chua TS. Tem: Tree-enhanced embedding model for explainable recommendation, WWW 2018.
15. Paul D, Sarkar S, Chelliah M, Kalyan C, Nadkarni P. Recommendation of High Quality Representative Reviews in e-commerce. RecSys 2017.

6. Targeted audience: Intermediate.
   Researchers / industry practitioners with Computer Science / Statistics background; exposure to Machine Learning techniques desirable.

7. Importance of topic for RecSys community (relevant tutorials from others)

Ranking is at the core of recommender systems which provide a list of items per user preference [1]. Content constantly added to news portals triggers recommendation of such streamed data to new users visiting the portal [2]. Cross-domain recommendation leverages knowledge of user preferences from one domain to another [3]. Embedding methods (e.g., Word2Vec) and networks (feed-forward, recurrent, convolutional) for music and session-based recommendation are covered in [4]. Our own proposal now complements our earlier tutorial on review-based item recommendation [5]. We focus instead here on score prediction which ranks review quality and recommendation factoring in item aspects and user profile leveraging review text and user interaction with it (e.g., vote).

1. Karatzoglou A, Baltrunas L, Shi Y. Learning to rank for recommender systems. RecSys 2013.
2. Cantador I, Cremonesi P. Cross-domain recommender systems. RecSys 2014.
3. Hopfgartner F, Kille B, Heintz T, Turrin R. Real-time recommendation of streamed data. RecSys 2015.
4. Karatzoglou A, Hidasi B. Deep Learning for Recommender Systems. RecSys 2017.
5. Chelliah M, Sarkar S. : Product Recommendations Enhanced with Reviews. RecSys 2017.

8. Teaching experiences and history of prior tutorials by the presenter(s).

Muthusamy Chelliah heads external research collaboration for Flipkart – who is a pioneer in the nascent online shopping market in a vibrant, emerging economy (India). He holds a PhD degree in Computer Science from Georgia Tech., Atlanta with a focus in distributed systems. He then spent 15 years with HP as a scientist and architect in US and India working on various areas like middleware, OS security, fault-tolerant systems and cloud computing. He then moved to Yahoo where he held a similar role as now engaging academia on solving problems relevant to industry leveraging research in ML, IR, NLP and data mining. He's passionate about catalyzing industry-relevant data science in global universities. He has published articles in conferences like IEEE SRDS as well as journals like TKDE and delivered tutorials in ICWS.

Sudeshna Sarkar has a long experience in experience in the Indian Institute of Technologies where she has taught for the last 22 years. She has taught many courses on Artificial Intelligence, Machine Learning, Natural Language Processing, Information Retrieval and Deep Learning among others. She has a popular video course on Artificial Intelligence which had been used by a large number of students, and a MOOC on Machine Learning which has run three times and catered to several thousand students each time. She has given several invited talks and short courses throughout her career. She has presented a Tutorial in Recsys' 2017 - along with Chelliah as below:

(https://www.slideshare.net/maranlar/product-recommendations-enhanced-with-reviews)

9. List of relevant publications by the presenter(s).

a) Muthusamy Chelliah, Sudeshna Sarkar: Product Recommendations Enhanced with Reviews. RecSys 2017 Tutorial: 398-399

b) Debanjan Paul, Sudeshna Sarkar, Muthusamy Chelliah, Chetan Kalyan, Prajit Prashant Sinai Nadkarni: Recommendation of High Quality Representative Reviews in e-commerce. RecSys 2017: 311-315

c) Maunendra S. Desarkar, Sudeshna Sarkar, and Pabitra Mitra: Aggregating Preference Graphs for Collaborative Rating Prediction, in 4th ACM Conference on Recommender Systems, RECSYS 2010, Barcelona, Sep 26-30, 2010.

d) Maunendra Sankar Desarkar, Sudeshna Sarkar, Pabitra Mitra: Preference Relations Based Unsupervised Rank Aggregation for Metasearch Expert Systems with Applications Volume 49, 1 May 2016, Pages 86-98

e) Maunendra Sankar Desarkar, Roopam Saxena, Sudeshna Sarkar: Preference Relation Based Matrix Factorization for Recommender Systems in 20th International Conference on User Modeling, Adaptation and Personalization (UMAP) 2012: 63-75.

f) Maunendra Sankar Desarkar, Sudeshna Sarkar: Rating prediction using preference relations based matrix factorization. In UMAP Workshops 2012.

g) Maunendra Sankar Desarkar, Sudeshna Sarkar: User based Collaborative Filtering with Temporal Information for Purchase Data. KDIR 2012: 55-64.

h) Bhattacharya, Paheli, Pawan Goyal, and Sudeshna Sarkar. "Query Translation for Cross-Language Information Retrieval using Multilingual Word Clusters." WSSANLP 2016 (2016): 152.

i) Bhattacharya, Paheli, Pawan Goyal, and Sudeshna Sarkar. "Using Word Embeddings for Query Translation for Hindi to English Cross Language Information Retrieval." Computación y Sistemas 20.3 (2016): 435-447.

j) Rajendra Prasath, Sudeshna Sarkar and Philip O'Reilly: "Improving Cross Language Information Retrieval Using Corpus Based Query Suggestion Approach.", CICLING 2015, Springer LNCS, Vol 9042, pp 448-457

10. A 2-minute video where the presenters introduce themselves and pitch their

tutorial: https://youtu.be/AZys0duDYJI

11. Statement :

We, Muthusamy Cheliiah and Sudeshna Sarkar, certify that the materials (slides, readings, and/or code) used/ mentioned in the tutorial will be publicly available after the tutorial.