## LDA with multiple data streams for fashion applications

Doris Jung-Lin Lee (jlee782@illinois.edu)

Temporal changes are important in fashion, as trends and fads come and go quickly and new items are introduced every few months or even every few days for fast-fashion brands. Prior work in computational approaches to fashions have focussed on using static data repositories to characterizing styles and build visual or textual recognition and understanding of the relationship between clothing items and styles[3-9]. While there has been existing work that studies trends in fashion [3-4], these studies are still based partitioning static datasets into separate time periods, and then modelling these subsets independently.

This is a research track project where we will be looking at ways to integrate multiple data streams for updating the Latent Dirchlet Allocation (LDA) model. More specifically, we will be looking at how recent advances in streaming LDA could be applied to fashion applications to combine multiple heterogenous data streams.

Streaming fashion data is a challenging problem because:

- data may have various different update rates and noisy level (e.g. fashion-related hashtags versus highly-curated fashion magazine articles). From an applications point-of-view, it is important to incorporate these priors into model in a sensible manner.
- need to constantly update the model to ensure that the downstream applications reflects the most up-to-date trends.
- The "bag-of-words" assumption (exchangeability hypothesis) employed by most topic models no longer hold for data streams, since the ordering of data based on their time of arrival an important aspect.

I am currently reviewing existing literature on how various inference methods for LDA could be modified to take in data streams including work on updating population posterior [1] and applying stochastic variational inference[2]. The data sources that we have in mind includes new product information, seasonal runway data, magazine reviews and daily social media content from famous designers.

Different types of trends caters to different use cases. For example, fashion consumers are often interested in knowing the current trend (used for recommendation) , while fashion designers and buyers want to predict long-term future trends in upcoming seasons to commercialize the effects. After understanding how single streams are handled in existing work, we will then look at how to work with multiple data streams for modelling different kinds of trends into a single learned model. Past work has shown that learning these models on single streams requires non-trivial adaptations [1],[2] and therefore contribute to the novel research aspect of this project.

## Tentative Timeline:

- Now ~11/7 : Literature review and understanding feasible inference methods for streaming LDA (This will also help prepare content for presentation for the project-based workshop in late Nov-early Dec)
- 11/7~11/18 : Aggregating various data sources and streams to feed into the model.
- 11/18~12/2 : Implementing inference algorithm.
- 12/2~12/9 : Experiments for testing efficacy of model compared trends discovered by static analysis.

## References:

[1] Mcinerney, J., & Blei, D. M. (2016). The Population Posterior and Bayesian Inference on Streams, 1–10.
[2] Boyd, N., Wibisono, A., Wilson, A. C., Jordan, M. I., & Broderick, T. (2013). Streaming Variational Bayes, 1–9.
[3] He, R., & Mcauley, J. (2016). Ups and Downs : Modeling the Visual Evolution of Fashion Trends with One-Class Collaborative Filtering.
[4] McAuley, J., & Leskovec, J. (2013). From amateurs to connoisseurs: modeling the evolution of user expertise through online reviews. In Proceedings of the 22nd International Conference on World Wide Web, 897–908.
[5] Vittayakorn, S., Yamaguchi, K., Berg, A. C., & Berg, T. L. (2015). Runway to realway: Visual analysis of fashion. Proceedings - 2015 IEEE Winter Conference on Applications of Computer Vision, WACV 2015, 951–958.
[6] Chen, H., Xu, Z. J., Liu, Z. Q., & Zhu, S. C. (2006). Composite templates for cloth modeling and sketching. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1, 943–950.
[7] Yamaguchi, K., Brook, S., Hill, C., Berg, T. L., & Hill, C. (2013). Paper Doll Parsing: Retrieving Similar Styles to Parse Clothing Items.
[8] Arora, S. (2016). Decoding Fashion Contexts Using Word Embeddings.
[9] Kristen Vaccaro, Sunaya Shivakumar, Ziqiao Ding, Karrie Karahalios, and R. K. (2016). The Elements of Fashion Style.