# Aspect Based Recommendations: Recommending Items with the Most Valuable Aspects Based on User Reviews

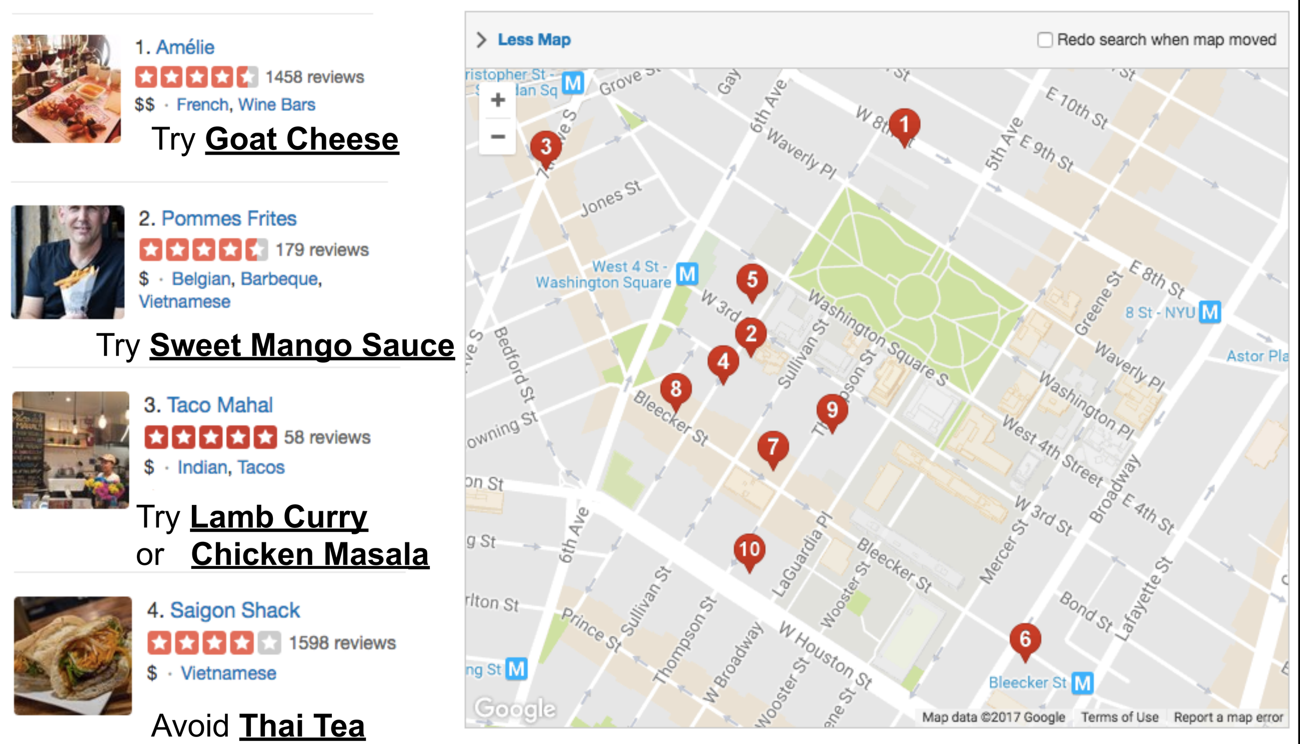
K.Bauman TEMPLE UNIVERSITY    B.Liu UIC    A.Tuzhilin NYU

**Research Idea:**

Recommending not only a product but also the most important (positive or negative) *aspects* that can *enhance customer experience* with the product.

**Examples:**

*Positive:* visit "Aquagrill" and order "FISH" there.
*Negative:* visit "Cafe X" but do not order "DESSERT".






| | |
|---|---|
| 1. minestrone soup | 1. antipasto |
| 2. mushroom risotto | 2. pasta carbonara |
| 3. panna cotta | 3. tiramisu |
| Predicted rating: **3.4** | Predicted rating: **4.5** |

***We recommend to order Pasta!***

## Method of Recommending the Most Valuable Aspects

1. Extracting aspects from the reviews
2. Training ***Sentiment Utility Logistic Model (SULM)***
   - aspect sentiments
   - overall satisfaction
3. Calculating aspect impact on rating
4. Recommending products and aspects

## 1. Extracting Aspects and Sentiments

"(1) Had lunch in Taqueria today. (2) Ordered the taco with rice and beans and it was great. (3) The service was quick. (4) The atmosphere was dark and soothing."

- FOOD – positive
- SERVICE – positive
- ATMOSPHERE – positive

## 2(a) SULM: Aspect Sentiments

Expressed sentiment (OP output): $o_{ui}^k \in \{0,1\}$

Sentiment utility $s_{ui}^k \in \mathbb{R}$ - level of satisfaction of user $u$ with aspect $k$ of item $i$, latent variable.

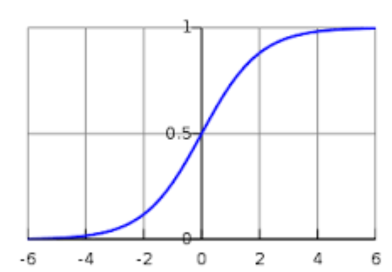
Use Matrix Factorization to estimate:

$$\hat{s}_{ui}^k(\theta_s) = \mu^k + b_u^k + b_i^k + p_u^k \cdot (q_i^k)^T$$

$$\theta_s = \{\mu^k, b_u^k, b_i^k, p_u^k q_i^k\}$$

Logistic function:

$$g(t) = \frac{1}{1+e^{-t}}$$




Estimation of explicit sentiment:

$$\hat{o}_{u,i}^k(\theta_s) = g\left(\hat{s}_{u,i}^k(\theta_s)\right)$$

Maximize log-likelihood:

$$l_s(S|\theta_s) = \sum_{u,i,k}\left(o_{ui}^k \log\left(\hat{o}_{u,i}^k(\theta_s)\right) + (1-o_{ui}^k)\log\left(1-\hat{o}_{u,i}^k(\theta_s)\right)\right)$$

## 2(b) SULM: Overall Satisfaction

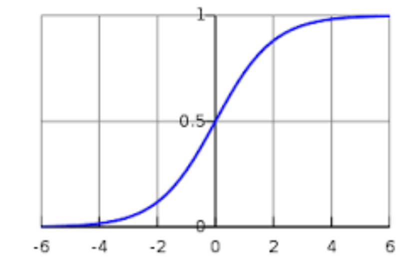
Overall rating (binary): $r_{ui} \in \{0,1\}$

Overall level of satisfaction (latent): $d_{ui} \in \mathbb{R}$

$$\hat{d}_{ui}(\theta) = \sum_{k \in \mathbb{A}} \hat{s}_{u,i}^k(\theta_s) \cdot (z^k + \underbrace{w_u^k}_{\text{user}} + \underbrace{v_i^k}_{\text{item}})$$

Binary rating estimation:

$$\hat{r}_{u,i}(\theta) = g\left(\hat{d}_{u,i}(\theta)\right)$$




Maximize log-likelihood:

$$l_r(R|\theta) = \sum_{u,i}(r_{ui} \cdot \log(\hat{r}_{u,i}(\theta)) + (1-r_{ui}) \cdot \log(1-\hat{r}_{u,i}(\theta)))$$

**Training SULM**

*Simultaneously* fits ratings *and* sentiments:

$$Q(\theta) = -\alpha \cdot \underbrace{l_r(R|\theta)}_{\text{rating part}} - (1-\alpha) \cdot \underbrace{l_s(S|\theta_s)}_{\text{sentiment part}} + \underbrace{\frac{\lambda_r}{2} \cdot \|\theta_r\|^2 + \frac{\lambda_s}{2} \cdot \|\theta_s\|^2}_{\text{regularization}}$$

Use Stochastic Gradient Descent to fit the model.

## 3. Aspect Impact on Rating

For a new potential review compute the impact of each aspect $k$ on the predicted rating as the corresponding summand from the rating prediction part of the model.

$$impact_{ui}^k = \hat{s}_{ui}^k \cdot (z^k + \underbrace{w_u^k}_{\text{user}} + \underbrace{v_i^k}_{\text{item}})$$

## 4. Recommending Products and Aspects

- identify group of aspects over which *the customer has control* (e.g. Dish vs. Decor)
- identify the most valuable aspects
- recommend a product *and* its corresponding suggestions to experience (positive) or do not experience (negative) a particular aspect

## Experiments

yelp

| | Restaurants | Hotels | Beauty&Spa |
|---|---|---|---|
| Initial | 1,344,405 | 96,384 | 104,199 |
| Filtered | 602,112 | 5,669 | 5,065 |

**Performance Measures**

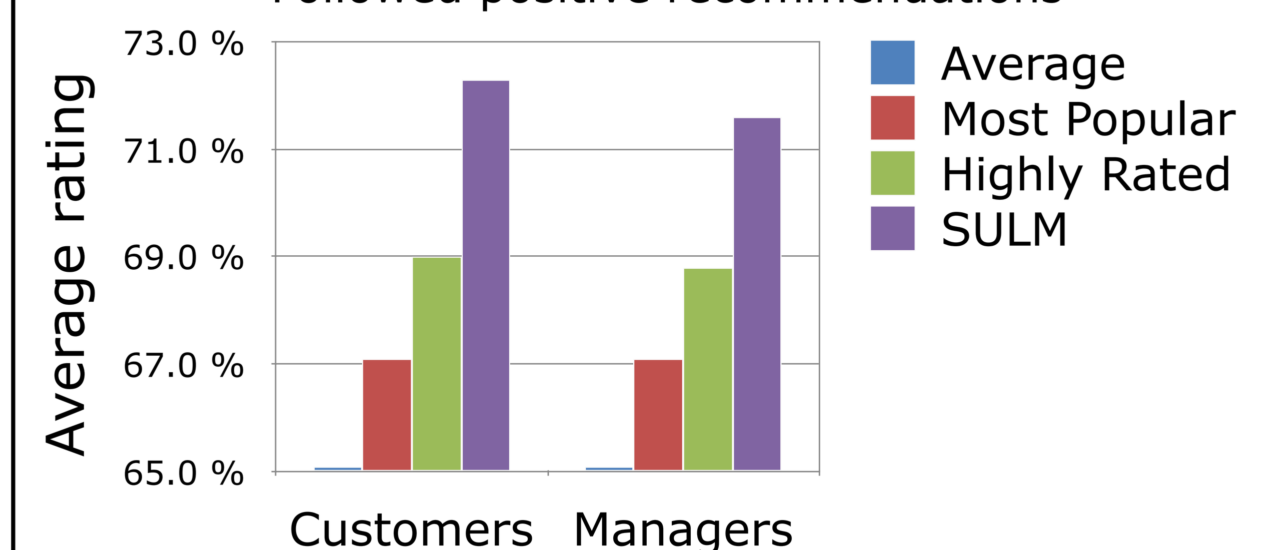
1. **Recommendations of Aspects**: how much the average rating is changed for those customers who followed the recommendations of aspects.
2. **Rating prediction**: *AUC, Precision@Top3*
3. **Aspect ranking**: *Precision@Top3, Precision@Top5*

**Baselines**

- Recommending the most popular aspect of a product
- Recommending highly rated aspect of a product
- Hidden Factors as Topics (**HFT**) (McAuley and Leskovec 2013)
- Learning to Rank User Preferences Based on Phrase-Level Sentiment Analysis Across Multiple Categories (**LRPPM**) (Chen et al. 2016)

## Results (Restaurants)

Followed positive recommendations




Did *not* follow negative recommendations


**Conclusions:**

- Our method **(SULM)** of recommending products with the most valuable aspects help
  - to get better customer experience vs. baselines
  - to avoid negative customer experience better than baselines
  - has rating prediction and aspect ranking performances comparable to the baselines.