Case study
Data Mining and Predictive Analytics

Temporal Modeling of Reviewer Expertise
Why do Americans like Pale Ale?

Lupulin threshold shift:
People become accustomed to hops over time, and can recognize more subtle flavors.

American Pale Ales:
- Hopsecutioner
- Hoptimus Prime
- Smooth Hoperator
- Red Hoptober
- Hoppy ending
- Hoptopus
- Hopsickle
- Tricerahops
Users and products evolve over time

“Classics” are rated better (Koren, 2010); new products cause users to change focus (Koller & Malouf, 2007).

Users influence each other (Ma et al., 2011); communities shift over time (Xiong et al., 2010).

How can we effectively characterize **acquired tastes** or **expertise**?

Age of the product

Age (development) of the user

Age (zeitgeist) of the community
Data

- **ratebeer**: 3M reviews, 100K beers, 40K users
- **Beeradvocate**: 1.5M reviews, 60K beers, 30K users
- **Cellartracker**: 2M reviews, 500K wines, 45K users
Replace the ‘standard’ model

\[ r_{ec}(u, i) = \alpha + \beta_u + \beta_i + \gamma_u \cdot \gamma_i \]

With one whose parameters change as a function of time (t)

\[ r_{ec_t}(u, i) = \alpha(t) + \beta_u(t) + \beta_i(t) + \gamma_u(t) \cdot \gamma_i(t) \]

How we define \( t \) determines what type of evolution we model
Models of user and community evolution

user review timelines  stages of community evolution

user review timelines  stages of user evolution

time →

time →
Observation:
People evolve and develop at different rates. We must **learn** the rate of development for each user.
Models of user and community evolution

**rows:** models of increasingly “experienced” users

**columns:** review timeline for one user

Each user’s evolution can be thought of as a **monotonic** path through a graph.
Optimization problem & fitting

**Model:**

\[ rec_{e,u,i}(u,i) = \alpha(e) + \beta_u(e) + \beta_i(e) + \gamma_u(e) \cdot \gamma_i(e) \]

experience at time of review  
offset, bias (user/item), and latent factors

**Optimization problem:**

\[ \arg \min_{\Theta, \varepsilon} \frac{1}{|T|} \sum_{r_{u,i} \in T} (rec_{e,u,i}(u,i) - r_{u,i})^2 + \Omega(\Theta) \]

model & experience parameters  
rating error  
regularizer  
smoothness & l2 regularity
Optimization problem & fitting

Repeat steps (1) and (2) until convergence:

**Step 1:**
fit expertise progression
(solved using dynamic programming)

**Step 2:**
fit rating models for each expertise level

\[
\arg\min_\Theta \frac{1}{|T|} \sum_{r_{ui} \in T} (\text{rec}_{e}(u, i) - r_{u,i})^2 + \Omega(\Theta)
\]

solved via gradient ascent using L-BFGS
(see e.g. Koren & Bell, 2011)
Outcomes – applications

**Rating prediction:**
- Beer: 6% improvement over state-of-the-art
- Wine: 13% improvement
- Movies (Amazon): 23% improvement

**User retention:**
What happens to users who **fail** to acquire taste for a product?

Users who acquire tastes slowly are more likely to quit the community.
Outcomes – understanding

1. Entire categories of products tend to be preferred by experts or beginners.

2. Experts rate the top products more generously, and the bottom products more harshly.

---

RateBeer

- Lagers
- Mild Ales
- Strong Ales

difference between expert and novice ratings (stars)

Firestone XV

Bud Light

average product rating (stars)
Outcomes – understanding

What are experts?
Predictability and agreement are two necessary conditions (Einhorn, 1974)

Experts are more predictable than beginners. They are also more inclined to agree with each other (right).
1. We extended our model to allow for multiple “classes” of progression:

“Class 1” users

Pliny The Elder
Velvet Merlin (Merkin)
Lukcy 13asartd Ale
Parabola

Double Jack
Stone 15th Anniversary Escondidian Imperial Black IPA

Sofie
Cellar Door
Daisy Cutter Pale Ale
Mongo

“Class 2” users
Extensions: Other types of data

2. We modeled to data from other domains, including web navigation traces

Browsing strategies for two classes of articles (from “Wikispeedia”)

WWW 2014 (w/ Yang, Leskovec, LePendu & Shah)
Extensions: Other types of data

2. We modeled to data from other domains, including medical records

Stages of Chronic Kidney Disease

© ultracare-dialisys.com
Questions?