Business Data Analytics

MTAT.03.319

Lecture 12

The slides are available under creative common license. The original owner of these slides is the University of Tartu.
Cross Selling & Upselling

How to sell more?
(Cross/Up) SELLING

Here is a simple but powerful rule - always give people more than they expect to get.

Nelson Boswell
Tips

• You already know about following tips
  • “The cost of acquiring a new customer is often around 4 times more expensive than it is to sell to an existing customer.”

• Something new:
  • The most successful business practices to achieve this are by up-selling and cross-selling.
Fast Food Seller

Would you like Estonian potatoes with that? 😊

Cross Selling
Cross Selling: Amazon Shopping
Cross Selling: Definition

• Cross-sell involves the sale of multiple products offered by a single product/service provider to a new or existing customer.
UpSelling

Listen free or subscribe to Spotify Premium.

Spotify Free
$0.00/month
- Shuffle play
- Ad free
- Unlimited skips
- Listen offline
- Play any track
- High quality audio

GET FREE

Spotify Premium
$9.99/month
Start your 30 day free trial
- Shuffle play
- Ad free
- Unlimited skips
- Listen offline
- Play any track
- High quality audio

GET PREMIUM
Tips For Cross/Up selling?

**Cross Selling**
- Peers Also Bought
- Incentives
- On Product Copy
- Discounted Second Buy
- Build A Relationship And Then Ask

**Up Selling**
- Sell the benefits of the up-sell
- Keep The Up-Sell Below 25% Of The Original Order
- Highlight Your Up-sell
Return of Cross/Up Selling Strategy?

- Amazon reportedly attributes as much as 35 percent of its sales to cross-selling through its options on every product page
  - “customers who bought this item also bought” and
  - “frequently bought together”.
How to Increase revenues from existing or new customers?

Up-selling

Cross-selling

Cross Selling or Up-Selling?
Cross Selling or Up-Selling?
## Cross Selling | Up Selling

### Who?
- Identify the customer or a cluster for a better approach.
- Present relevant offers based on his buying history and/or social-demographics characteristics.

### What?
- Identify the products or services which best fit the buying situation.
- Constantly analyze buying behavior in order to identify new trends (predictive models).

### When, Where?
- Identify the best moment during the buying flow to offer another product or service.
- Respect the users main objective.

### How?
- Identify the best position on the screen
- Identify the best model (text based, txt+img, advertising, radio buttons, checkboxes, etc)

Source: https://www.slideshare.net/dmedeiros/crosssellecommerce07
Customer Life Cycle:
How to solve this problem?

• What products to recommend to whom?
  • Solution: Technical based approach

• At what stage of the browsing process to show other options?
  • Solution: Psychological based approach
  • Out of scope of this course.
What products to recommend to whom?: Recommender Systems

Goal of a Recommender System: Identify products most relevant to the user (Eg. Top n offers).
Friend Recommendations
Product Recommendations

Python Machine Learning  Paperback
2015
by Sebastian Raschka  (Author)

Customers Who Bought This Item Also Bought

- Data Science from Scratch: First Principles with Python
  - Joel Grus
- Python for Data Analysis:
  - Wes McKinney
- Fundamentals of Machine Learning for Predictive Data Analytics:
  - John D. Kelleher
Job Recommendations

10 jobs that match your preferences

Front End Engineer & Growth Maker
Gliffy
San Francisco Bay Area
3 days ago

Principal Data Scientist
Move, Inc
San Francisco Bay Area
3 hours ago

NCG - Research Scientist - Data Analytics
Visa
A Naive understanding of Recommender Systems
Types of Recommendation Systems

- Popularity Based System
- Classification based
- Collaborative Filtering
  - Nearest Neighbor (remember KNN classification technique ?)
  - Matrix Factorization (we will not cover)
Solution 1: Popularity based Recommender System

Recommend items viewed/purchased by most people
Recommendations: Ranked list of items by their purchase count
Better would be something like this
Solution 2: Classification Model

Use features of both products as well as users in order to predict whether a user will like a product or not.

User Features (Eg. Age, Gender)
Product Features (Eg. cost, quality)
Purchase History

Classifier
Like or Not?

Limitation: Difficult to collect high quality information about products and users.
Solution 3: Nearest neighbor Collaborative Filtering

User-based

Find users who have a similar taste of products as the current user.

Similarity is based upon similarity in users’ purchasing behaviour.

“User x is similar to user y because both purchased items A, B and C.”

Item-based

Recommend items that are similar to the items the user bought.

Similarity is based upon co-occurrence of purchases.

“Items A and B were purchased by both users x and y, so they are similar.”

Fig. Source: http://www.salemmarafi.com/code/collaborative-filtering-with-python/
User – User Collaborative Filtering
Consider users x and y with rating vectors \( r_x \) and \( r_y \)

We need similarity metric \( \text{Sim}(x,y) \)

Capture the intuition that \( \text{Sim}(A,B) > \text{Sim}(A,C) \)

### Similar Users

<table>
<thead>
<tr>
<th>Users</th>
<th>Movies</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HP1</td>
</tr>
<tr>
<td>A</td>
<td>4</td>
</tr>
<tr>
<td>B</td>
<td>5</td>
</tr>
<tr>
<td>C</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>3</td>
</tr>
</tbody>
</table>
Similar Users: Jaccard Similarity

- Jaccard similarity \( (A,B) = \frac{r_A \cap r_B}{r_A \cup r_B} \)
- Jaccard distance = \( 1 - \frac{r_A \cap r_B}{r_A \cup r_B} \)
- \( \text{Sim} (A,B) = 1/5 \); \( \text{Sim} (A,C) = 2/4 \)
  - \( \text{Sim}(A,B) < \text{Sim}(A,C) \) : Ignores the rating values

<table>
<thead>
<tr>
<th>Users</th>
<th>HP1</th>
<th>HP2</th>
<th>HP3</th>
<th>TW</th>
<th>SW1</th>
<th>SW2</th>
<th>SW3</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>4</td>
<td></td>
<td></td>
<td>5</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>4</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Movies</th>
<th>HP1</th>
<th>HP2</th>
<th>HP3</th>
<th>TW</th>
<th>SW1</th>
<th>SW2</th>
<th>SW3</th>
</tr>
</thead>
<tbody>
<tr>
<td>HP1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HP2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HP3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TW</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SW1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SW2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SW3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Similar Users: Cosine Similarity

Users

<table>
<thead>
<tr>
<th></th>
<th>HP1</th>
<th>HP2</th>
<th>HP3</th>
<th>TW</th>
<th>SW1</th>
<th>SW2</th>
<th>SW3</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>B</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>C</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>4</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>D</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
</tbody>
</table>

Movies

\[
\text{Sim}(A,B) = 0.38 \ ; \ \text{Sim}(A,C) = 0.32
\]

\[
\text{Sim}(A,B) > \text{Sim}(A,C) : \text{but not much}
\]

\[
\cos(\theta) = \frac{A \cdot B}{\|A\|_2 \|B\|_2} = \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2} \sqrt{\sum_{i=1}^{n} B_i^2}}
\]

\[
\text{NOTE: Fill empty values by 0}
\]

Problem: Treat missing values as negative
## Similar Users: Centered Cosine

### Normalized ratings by subtracting the row mean

<table>
<thead>
<tr>
<th>Users</th>
<th>HP1</th>
<th>HP2</th>
<th>HP3</th>
<th>TW</th>
<th>SW1</th>
<th>SW2</th>
<th>SW3</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>4</td>
<td>5</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>C</td>
<td></td>
<td></td>
<td>2</td>
<td>4</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>3</td>
<td>5</td>
<td></td>
<td>1</td>
<td>4</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Avg. Rat</th>
</tr>
</thead>
<tbody>
<tr>
<td>10/3</td>
</tr>
<tr>
<td>14/3</td>
</tr>
<tr>
<td>11/3</td>
</tr>
<tr>
<td>6/2 =3</td>
</tr>
</tbody>
</table>

In each row: original value \(-\text{Avg. Rat}\)

Each row addition = 0

Ratings are centered around 0.

+ : users liked it
- : users did not like it
### Similar Users: Centered Cosine (2)

#### Movies

<table>
<thead>
<tr>
<th></th>
<th>HP1</th>
<th>HP2</th>
<th>HP3</th>
<th>TW</th>
<th>SW1</th>
<th>SW2</th>
<th>SW3</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>2/3</td>
<td>0</td>
<td>0</td>
<td>5/3</td>
<td>-7/3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>B</td>
<td>1/3</td>
<td>1/3</td>
<td>-2/3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>C</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-5/3</td>
<td>1/3</td>
<td>4/3</td>
<td>0</td>
</tr>
<tr>
<td>D</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

#### Users

- Sim (A,B) = 0.09 ; Sim (A,C) = -0.56
  - Sim(A,B) >> Sim(A,C) : but not much
- Captures intuition better
  - Missing ratings treated as “average”
  - Handles “tougher raters” and “easy raters”

Also known as pearson correlation.
Rating Predictions

• Goal: Prediction for user $X$ and item $i$
• What we need:
  • Let $r_X$ be the rating for the user $X$.
  • Let $N$ be the set of $k$ users most similar to $X$, who have rated item $i$.
• Option 1: $r_{xi} = \frac{1}{k} \sum_{y \in N} r_{yi}$ \hspace{1cm} (Average)
  
  For a neighbor $y$ in ($\in$) the set $N$

  $s$ is the similarity of the user $x$ and its neighbor $y$

• Option 2: $r_{xi} = \frac{\sum_{y \in N} s_{xy} r_{yi}}{\sum_{y \in N} s_{xy}}$ \hspace{1cm} (Weighted Average)
Item – Item Collaborative Rating (1)

• For item $i$, find other similar items.
• Estimate rating for item $i$ based on ratings for similar items
• Can use some similarity metrics and prediction functions as in user-user model.

$$r_{x_i} = \frac{\sum_{j \in N(i:x)} s_{ij} r_{xj}}{\sum_{j \in N(i:x)} s_{ij}}$$

- $s_{ij}$: similarity of items $i$ and $j$
- $r_{xj}$: ratings of item $i$ by the user $x$
- $N(i:x)$: set of items similar to $i$, rated by user $x$. 

Rating function
Item – Item Collaborative Filtering (2)

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>3</td>
<td>5</td>
<td>?</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Ratings are between 1 to 5
Empty boxes: unknown rating

? : Estimate the rating of movie 1 by the user 5
Item – Item Collaborative Filtering (3)

Sim = Pearson Correlation.

1) Subtract mean rating $m_i$ from each movie $i$.
   1) $m_1 = (1+3+5+5+4)/2 = 3.6$
   2) Row 1 = (-2.6, 0, -0.6, 0, 0, 1.4, 0, 0, 1.4, 0, 0.4,0)
2) Compute Cosine similarities between row1 and other rows

Remember $N = 2$
Select 2 movies similar to 1 and rated by user 5.

$$r_{xi} = \frac{\sum_{j \in N(i)} s_{ij} r_{xj}}{\sum_{j \in N(i)} s_{ij}}$$

Weighted Average = \frac{(0.41*2 + 0.59*3)/ (0.41*0.59)} = 2.6
Performance Metric for Recommendation Systems

All Recommendations (made on training dataset)

<table>
<thead>
<tr>
<th>Relevant Items that are also recommended</th>
<th>Irrelevant Items that are recommended</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relevant items that are not recomms</td>
<td></td>
</tr>
</tbody>
</table>

Precision: Measure of Exactness

\[
\text{Number of relevant products being recommended} \over \text{number Items being recommended.}
\]

Recall: Measure of Completeness

\[
\text{#relevant products being recommended} \over \text{total number relevant items.}
\]
User to User Vs Item to Item

• User to User
  • Problem: Sparse: Users have limited interests (in buying)

• Item-Item outperforms User-User
  • Users are more complex than Items
  • Items have limited genre than Users.
  • Item similarity makes more sense than Users similarity
References

• [https://www.youtube.com/watch?v=39vJRxlPSxw](https://www.youtube.com/watch?v=39vJRxlPSxw)
• [https://www.youtube.com/watch?v=3SInFQbLQA](https://www.youtube.com/watch?v=3SInFQbLQA)
• [Collaborative filtering:](https://www.youtube.com/watch?v=h9gpufJFF-0)
• Google 😊