Business Data Analytics

MTAT.03.319

Lecture 6

The slides are available under creative common license. The original owner of these slides is the University of Tartu.
Till now and today!

Lecture 1: Intro BDA

Lecture 2: Data Exploration prerequisite task

Lecture 3: Customer segmentation

Lecture 4: Customer Lifecycle Management (CLM) Regression problems

Lecture 5: Customer Lifecycle Management (CLM) Classification problems

Lecture 6: Cross-sell/Up-selling
Cross Selling & Upselling

How to sell more?
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 potatoes with that? 😊

Cross Selling
Cross Selling: Amazon Shopping
Cross Selling: Definition

• To sell related or complementary products to a new or existing customer.

  https://www.investopedia.com/terms/c/cross-sell.asp

• Cross-selling is a sales technique used to get a customer to spend more by purchasing a product that’s related to what’s being bought already.

  Source: https://www.shopify.com/encyclopedia/cross-selling
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

**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 FREE

GET PREMIUM
Up-Selling
Cross and Up Selling: Definition

• Up-selling: is a sales technique where a seller induces the customer to purchase more expensive items, upgrades or other add-ons in an attempt to make a more profitable sale.

  Source: https://en.wikipedia.org/wiki/Upselling

• Cross-selling: To sell related or complementary products to a new or existing customer.
Tips For Cross/Up selling?

Cross Selling

• Peers Also Bought
• Incentives
• 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
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”.
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/crosssalecommerce07
Classification in Customer Lifecycle Management

- Lead propensity
- Cross-sell/up-sell propensity
- Response modeling / buying propensity
- Fault/Complaint prediction
- Churn propensity
- Win-back propensity
How to solve this problem?

- **Question:** What products to recommend to whom?
- **Solution:** Recommendation Systems

- **What products?**
  - Popularity
  - Market Basket Analysis

- **What products to whom?**
  - Collaborative filtering
What products to recommend to whom? : Recommender Systems

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

Platforms

Recommendations

Friends

Books

Jobs
Solution 1: Popularity based Recommender System

Recommend items viewed/purchased by most people
Recommendations: Ranked list of items by their purchase count
Pareto Law

- 20% (highly valued customers) of customers bring 80% of profit
- 20% of products bring 80% of the profit
  - But what about rest of the 80% products? Popularity based techniques are not helpful
Popular products!

There are many popular ecommerce products to sell them online like:

- Smartwatches.
- Video Doorbells.
- Facial Masks.
- Highlighters.
- Phone Cases.
- Avocado Oil.
- Bluetooth Speakers.
- Enamel Pins.
Popularity is safe but what about association among the products you recommend?
Solution 2: Market Basket Analysis

Market

Basket

Analysis

MBA
MBA put sense while recommending products
Market Basket Analysis (MBA)

- It is a technique or algorithm to identify the association rules from your data

- Input
  - List of purchases by customers over different visits

- Output
  - What items purchased together?

Association rules -> Generates rules
Example: (X -> Y)

Market Basket -> Assigns business outcome to those rules
Example: X,Y could be sold together
MBA: Terminologies

• **Items**: Objects that we are identifying associations between.

• **Examples**:  
  • In a supermarket, each item is a product.  
  • For a publisher, each item might be an article, a blog post, a video etc.

• **A group of items is an item set.**  
  • \{i_1, i_2, i_3 \ldots , i_k\}
MBA: Terminologies Cont..

- **Items**: Objects that we are identifying associations between.
- **Transactions**: Transactions are instances of groups of items co-occurring together.
- **Examples**:
  - For an online retailer, a transaction is, generally, a group of items bought together
  - For a publisher, a transaction might be the group of articles read in a single visit to the website.
  - **NOTE**: It is up to the analyst to define over what period to measure a transaction.
- **For each transaction, we have an item set**.
  - \( t_n = \{i_1, i_2, i_3, ..., i_k\} \)
• **Items**: Objects that we are identifying associations between.

• **Transactions**: Transactions are instances of groups of items co-occurring together.

• **Rules**: are statements of the form
  
  - \{i_1, i_2, i_3 \ldots\} \Rightarrow \{i_k\}
  
  - if you have the items in item set (on the left hand side (LHS) of the rule i.e. \{i_1, i_2, \ldots\}, then it is likely that a visitor will be interested in the item on the right hand side (RHS i.e. \{i_k\}).

  - In the example above, rule would be:
    
    - \{flour, sugar\} \Rightarrow \{eggs\}
• **Items:** Objects that we are identifying associations between.

• **Transactions:** Transactions are instances of groups of items co-occurring together.

• **Rules:** are statements of the form
  
  • \( \{i_1, i_2, i_3, \ldots \} \Rightarrow \{i_k\} \)
  
  • if you have the items in item set (on the left hand side (LHS) of the rule i.e. \( \{i_1, i_2, i_3, \ldots \} \), then it is likely that a visitor will be interested in the item on the right hand side (RHS i.e. \( \{i_k\} \)).

• In the example above, rule would be:
  
  • \{flour, sugar\} \Rightarrow \{eggs\}
MBA: Terminologies Cont..

- **Items**: Objects that we are identifying associations between.
- **Transactions**: Transactions are instances of groups of items co-occurring together.
- **Rules**: Find associated items for sale
- **Output of MBA?**

<table>
<thead>
<tr>
<th>Association rules -&gt; Generates rules</th>
</tr>
</thead>
<tbody>
<tr>
<td>Example: (X -&gt; Y)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Market Basket -&gt; Assigns business outcome to those rules</th>
</tr>
</thead>
<tbody>
<tr>
<td>Example: X,Y could be sold together</td>
</tr>
</tbody>
</table>
MBA: Terminologies Cont..

• It works basically on following concepts
  • Support: \( P(A \cup B)/T \)
    • What is co-occurrence of two items name A and B
    • (Co-occurrence of A and B)/ Total Transactions.
  • Confidence: \( P(A \cup B)/P(A) \):
    • The proportion of transactions which contain A and also contain B.
    • How confident we are that B is present in presence of A.
    • Ratio of Support of (A and B), and Support of A.
  • Expected Confidence:
    • How confident we are that B is present in absence of A (or do not care about A).
    • # transactions where B is present /Total transactions
MBA: Terminologies Cont..

• **Lift:**
  • Ratio of Confidence and Expected Confidence
  • Ratio of \((B \text{ in presence of } A)\) and \((B \text{ in absence of } A)\)
  • Explains the change in probability of \(B\) over “presence of \(A\)” and “absence of \(A\)”
  • Lift \(\leq 1\)
    • A has no impact on \(B\)
  • Lift \(> 1\)
    • Relationship between \(A\) and \(B\) is significant
    • Larger the lift ratio, the more significant the association.

Source: https://www.youtube.com/watch?v=WxDV9WEYqPw
## Possible shopping Baskets (T)

<table>
<thead>
<tr>
<th>Transaction</th>
<th>Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Beer, Diaper, Chips, Aspirin</td>
</tr>
<tr>
<td>2</td>
<td>Diaper, Beer, Chips, Lotion, Juice, Milk</td>
</tr>
<tr>
<td>3</td>
<td>Soda, Chips, Milk</td>
</tr>
<tr>
<td>4</td>
<td>Soup, Beer, Diaper, Milk, Icecream</td>
</tr>
<tr>
<td>5</td>
<td>Soda, Coffee, Milk, Bread</td>
</tr>
<tr>
<td>6</td>
<td>Beer, Chips</td>
</tr>
</tbody>
</table>
# Retail Case Study

### Possible shopping Baskets (T)

<table>
<thead>
<tr>
<th>Transaction</th>
<th>Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transaction 1</td>
<td>Beer, Diaper, Chips, Aspirin</td>
</tr>
<tr>
<td>Transaction 2</td>
<td>Diaper, Beer, Chips, Lotion, Juice, Milk</td>
</tr>
<tr>
<td>Transaction 3</td>
<td>Soda, Chips, Milk</td>
</tr>
<tr>
<td>Transaction 4</td>
<td>Soup, Beer, Diaper, Milk, Icecream</td>
</tr>
<tr>
<td>Transaction 5</td>
<td>Soda, Coffee, Milk, Bread</td>
</tr>
<tr>
<td>Transaction 6</td>
<td>Beer, Chips</td>
</tr>
</tbody>
</table>

### Frequent Items (based on Ms = 30)

- (Beer, Diaper) : with support 50 %
- (Beer, Chips) : with support 50 %

Support = (Co-occurrence of A and B)/ Total Transactions.
Retail Case Study

Possible shopping Baskets (T)

<table>
<thead>
<tr>
<th>Transaction</th>
<th>Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transaction 1</td>
<td>Beer, Diaper, Chips, Aspirin</td>
</tr>
<tr>
<td>Transaction 2</td>
<td>Diaper, Beer, Chips, Lotion, Juice, Milk</td>
</tr>
<tr>
<td>Transaction 3</td>
<td>Soda, Chips, Milk</td>
</tr>
<tr>
<td>Transaction 4</td>
<td>Soup, Beer, Diaper, Milk, Icecream</td>
</tr>
<tr>
<td>Transaction 5</td>
<td>Soda, Coffee, Milk, Bread</td>
</tr>
<tr>
<td>Transaction 6</td>
<td>Beer, Chips</td>
</tr>
</tbody>
</table>

Frequent Items (based on Ms = 30)

<table>
<thead>
<tr>
<th>Items</th>
<th>Support</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Beer, Diaper)</td>
<td>50 %</td>
</tr>
<tr>
<td>(Beer, Chips)</td>
<td>50 %</td>
</tr>
</tbody>
</table>

A is Beer and B is either Diaper or Chips
Confidence: \( \frac{P(A \cup B)}{P(A)} \)
Expected Confidence: \( \frac{\text{# transactions where B is present}}{\text{Total transactions}} \)
Lift = \( \frac{\text{Confidence}}{\text{Expected Confidence}} \)

Rule 1 Beer \( \rightarrow \) Diaper
Confidence = 3/4 , Expected Confidence = 3/6
Lift = \( \frac{3/4}{3/6} = 1.5 \)

Rule 2 Beer \( \rightarrow \) Chips
Confidence = 3/4, Expected Confidence = 4/6
Lift = \( \frac{3/4}{4/6} = 1.1 \)

Support = \( \frac{\text{Co-occurrence of A and B}}{\text{Total Transactions}} \).
Let us summarize about the problem*

• Generate set of rules that link two or more products together.
• Each of these rules should have a lift greater than one.
• Also, we are interested in the support and confidence of those rules:
  • Higher confidence rules are ones where there is a higher probability of items on
    the RHS being part of the transaction given the presence of items on the LHS.
• Recommendations based on these rules to drive a higher response rate.
  • We’re also better off actioning rules with higher support first, as these will be
    applicable to a wider range of instances.

*Problem: How to find rules which can help us in finding patterns ?
MBA is through association rules

- We have to generate rules
- 3 Types
  - Actionable Rules: On which you can take action.
  - Trivial Rules: Interesting and need to do more research on.
  - Inexplicable Rules: Complex or incomprehensible (does not make sense)

Association rules -> Generates rules
Example: (X -> Y)

Market Basket -> Assigns business outcome to those rules
Example: X,Y could be sold together
Netflix Challenge!

DVD Rental
Utilizing the Inventory

Grand prize
of US$1,000,000
September 21, 2009
How to make Personalized recommendation

Users who liked “Love Simon” also liked following movies
Solution 3: 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 user x, so they are similar.”

Fig. Source: http://www.salemmarafi.com/code/collaborative-filtering-with-python/
User – User Collaborative Filtering
Similar Users

• 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 $\text{Sim}(A,B) > \text{Sim}(A,C)$

<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

\[
\text{Jaccard similarity}(A,B) = \frac{r_A \cap r_B}{r_A \cup r_B}
\]

\[
\text{Jaccard distance} = 1 - \frac{r_A \cap r_B}{r_A \cup r_B}
\]

\[
\text{Sim} (A,B) = 1/5 ; \text{Sim} (A,C) = 2/4
\]

- Sim(A,B) < Sim(A,C) : Ignores the rating values

<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: Cosine Similarity

• **Cosine similarity** ($A, B) = \cos(r_A, r_B)$
  • -1 : dissimilar, 0: orthogonal; +1: similar
  • Sim ($A, B) = 0.38$ ; Sim ($A, C) = 0.32$
    • Sim($A, B) >$ Sim($A, C$) : but not much

<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>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>

*NOTE: Fill empty values by 0*

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

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

### Normalized ratings by subtracting the row mean

<table>
<thead>
<tr>
<th>Users</th>
<th>Movies</th>
<th>Avg. Rat</th>
</tr>
</thead>
<tbody>
<tr>
<td>HP1</td>
<td>HP2</td>
<td>HP3</td>
</tr>
<tr>
<td>A</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>

In each row, original value – Avg. Rat

Each row addition = 0

Ratings are centered around 0.
+ : users liked it
- : users did not liked it
Similar Users: Centered Cosine (2)

<table>
<thead>
<tr>
<th>Users</th>
<th>Movies</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HP1</td>
</tr>
<tr>
<td>A</td>
<td>2/3</td>
</tr>
<tr>
<td>B</td>
<td>1/3</td>
</tr>
<tr>
<td>C</td>
<td>-5/3</td>
</tr>
<tr>
<td>D</td>
<td>0</td>
</tr>
</tbody>
</table>

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: $r_{xi} = \frac{1}{k} \sum_{y \in N} r_{yi}$ (Average)  
  For a neighbor $y$ in ($\in$) the set $N$

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

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

• 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.

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

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

Ratings are between 1 to 5
Empty boxes: unknown rating

?): Estimate the rating of movie 1 by the user C
Sim = Pearson Coeff.

1) Subtract mean rating $m_i$ from each movie $i$.
   1) $m_1 = (1+3+5+5+4)/5 = 3.6$
   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 rows

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

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

Weighted Average = $(0.41*2 + 0.59*3)/( 0.41*0.59)$ = 2.6
User to User Vs Item to Item

• Item-Item outperforms User-User

• Users are more complex than Items
  • Sparse: Users have limited interests (in buying)
  • Not all users can have likes/interests about all the items

• Items are simple: example: limited genres.

• Item similarity makes more sense than Users similarity
Evaluation

Test Data
If you analyse it as regression problem

- MAE
- Mean Square Error
- Root Mean Square Error

- Alternative: Precision at top-k
  - Percentage of predictions in the user’s top-k withheld ratings
Comparison of MBA & CF

Market Basket Analysis
• Association Rule Mining
• Lacks the personalized approach
• Clustering problem
• Unsupervised approach
• No labeled data is provided
• Scalable
• No serendipity
• Mostly look for popular items

Collaborative Filtering
• User-user or Item-Item Filtering
• Can be used for personalized recomm.
• More of a regression problem
• Supervised approach
• Labels (ratings etc) are provided.
• Computationally expensive
• Serendipity possible
• Looks at products in the long tail
Demo time!

https://courses.cs.ut.ee/2018/bda/fall/Main/Practice