RECOMMENDER SYSTEM IN RETAIL
Traditionally **Non-personalized** channels and mass offers have had the strongest presence.

- Success is difficult to measure
- Future of retail focuses on customer needs
SHIFT TO CUSTOMER CENTRICITY

Target Customers
Based on propensity to purchase product

Personalization
Personalize each unique product based on relevance to Target customers

High Response
Significantly Increase Response

Incremental Revenue
Focused on Customer Centricity
Who Else?
EVERYDAY APPLICATIONS OF PERSONALIZATION

RECOMMENDER SYSTEMS
RECOMMENDER SYSTEMS

- System to recommend items to users based on examples of their preferences/purchases/ratings
CURRENT STATE

STRATEGY AT SDM?
RECOMMENDER SYSTEM AT SDM

DIRECT MAIL

DIGITAL
CUSTOMER-CENTRIC APPROACH TO PERSONALIZATION

- Ultimate goal: Build personalization with 100% variability across all levers (currency, product, etc…) and types of offers
- Leverage Customer segments & prediction to drive offer type pairing
### METHODOLOGY EXAMPLE - CATEGORY AFFINITY

<table>
<thead>
<tr>
<th>Category 1</th>
<th>Category 2</th>
<th>Affinity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bath Products</td>
<td>Mens Fragrance</td>
<td>High</td>
</tr>
<tr>
<td>Cosmetic Accessories</td>
<td>Vitamins</td>
<td>High</td>
</tr>
<tr>
<td>Cosmetic Accessories</td>
<td>Ladies Fragrance</td>
<td>High</td>
</tr>
<tr>
<td>Derm Face &amp; Body</td>
<td>Skin Care</td>
<td>Medium</td>
</tr>
<tr>
<td>Derm Face &amp; Body</td>
<td>Prestige Faci Trtmt</td>
<td>High</td>
</tr>
<tr>
<td>Face &amp; Body Colours</td>
<td>Skin Care</td>
<td>High</td>
</tr>
</tbody>
</table>

Diagram:
- **Cx Access**: High Affinity with **Ladies Frag** and **Vitamins**.
- **Low Affinity** between **Cx Access** and **Light Bulbs**.

**Example**: Cx Access is highly related to Ladies Frag and Vitamins, but less so with Light Bulbs.
Increase Incremental Sales and Profitability by focusing on customers purchasing more products per basket.
CURRENT STATE

HOW WE DO IT AT SDM?
TWO APPROACHES TO RECOMMENDER SYSTEMS:

**Content Based**
- Focuses on **properties** of **items**
- Similarity of items is determined by measuring the similarity in their properties
- Example: Profiling of Internet Movie Database (IMDB) - assigns a genre to every movie

**Collaborative-Filtering**
- Focuses on the **relationship** between **users** and **items**
- Similarity of items is determined by the similarity of the ratings of those items by the users who have rated both items
- Example: Recommending Products / Movies

**General Application**

**Specialized Application**
COLLABORATIVE FILTERING

- Collaborative Filtering
- Non-probabilistic Algorithms
- Probabilistic Algorithms
- Bayesian-network models
- User-based nearest neighbor
- Item-based nearest neighbor
- Dimension Reduction
COLLABORATIVE FILTERING

Transactions / Ratings

Active User

Correlation Match

Extract Recommendations

C
Cosine Similarity measures similarity between two vectors of an inner product space that measures the cosine of the angle between them.

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

Jaccard Similarity measures similarity between finite sample sets, and is defined as the size of the intersection divided by the size of the union of the sample set.

\[
J(A, B) = \frac{|A \cap B|}{|A \cup B|}, \quad 0 \leq J(A, B) \leq 1.
\]

The Pearson correlation measures the Pearson correlation similarity of two users \( x, y \).

\[
\text{simil}(x, y) = \frac{\sum_{i \in I_{xy}} (r_{x,i} - \bar{r}_x)(r_{y,i} - \bar{r}_y)}{\sqrt{\sum_{i \in I_{xy}} (r_{x,i} - \bar{r}_x)^2} \sum_{i \in I_{xy}} (r_{y,i} - \bar{r}_y)^2}
\]
### Example:

<table>
<thead>
<tr>
<th>Users</th>
<th>Product 1</th>
<th>Product 2</th>
<th>Product 3</th>
<th>Product 4</th>
<th>Product 5</th>
<th>Product 6</th>
<th>Average - User</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tom</td>
<td></td>
<td>3</td>
<td>5</td>
<td>4</td>
<td>1</td>
<td></td>
<td>3.25</td>
</tr>
<tr>
<td>Matt</td>
<td></td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>1</td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>Ben</td>
<td>4</td>
<td>3</td>
<td></td>
<td>3</td>
<td>1</td>
<td></td>
<td>2.75</td>
</tr>
<tr>
<td>Jade</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td></td>
<td>1</td>
<td></td>
<td>3.25</td>
</tr>
<tr>
<td>Torri</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td></td>
<td>3</td>
<td></td>
<td>4.25</td>
</tr>
</tbody>
</table>

**Average Prod**

<table>
<thead>
<tr>
<th>Product 1</th>
<th>Product 2</th>
<th>Product 3</th>
<th>Product 4</th>
<th>Product 5</th>
<th>Product 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.3</td>
<td>3.4</td>
<td>4.5</td>
<td>3.5</td>
<td>1.4</td>
<td></td>
</tr>
</tbody>
</table>

**Similarity between Product 1 & 2**

\[
\text{Similarity} = \frac{(4 - 2.75)(3 - 2.75) + (4 - 3.2)(4 - 3.2) + (5 - 4.25)(4 - 4.25)}{\sqrt{(4 - 2.75)^2 + (4 - 3.2)^2 + (5 - 4.25)^2}}
\]

\[
= \frac{0.7650}{\sqrt{2.765 \times 0.765}} = \frac{0.7650}{1.6628 \times 0.8746} = \frac{0.7650}{1.4543} = 0.5260
\]
We are looking for products or product groups that fall in the area that is ‘Above Average Odds’. This means that the product or product group is most likely associated with ‘Incremental basket’.
RECOMMENDER APPROACHES

Attribute-based recommendations
(You like action movies, starring Clint Eastwood, you might like “Good, Bad and the Ugly” Netflix)

Item Hierarchy
(You bought Printer you will also need ink - BestBuy)

Collaborative Filtering – Item-Item similarity
(You like Godfather so you will like Scarface - Netflix)

Collaborative Filtering – User-User Similarity
(People like you who bought beer also bought diapers - Target)

Social+Interest Graph Based
(Your friends like Lady Gaga so you will like Lady Gaga, PYMK – Facebook, LinkedIn)

Model Based
Training SVM, LDA, SVD for implicit features
MEASUREMENT
TEST & LEARN

MEASURE Effectiveness of Recommendations

- Evaluation of Predicted Ratings (Mean Average Error, RMSE)
- Evaluation of top-N reccos
- MAE
- Accuracy
- Precision and Recall (F1 Score)
- ROC Curves
- Test vs Control

Next Steps

- Incorporate New Methodologies into current Recommender Systems
- Enhance contribution of LifeTime Value Models
- Bundling of Product
- Feed Results to SDM portal:
CHALLENGES
Problems with Collaborative Filtering method:

- **Cold Start**: There needs to be enough users already in the system to find a match.

- **Sparsity**: If there are many items to be recommended, even if there are many users, the user/ratings matrix is sparse, and it is hard to find users that have rated the same items.

- **First Rater**: Cannot recommend an item that has not been previously rated.
  - New items
  - Esoteric items

- **Popularity Bias**: Cannot recommend items to someone with unique tastes.
  - Tends to recommend popular items.
CHALLENGES

Problems with Content-based method:

- Requires content that can be encoded as meaningful features
- Users’ tastes must be represented as a learnable function of these content features
- Unable to exploit quality judgments of other users
  - Unless these are somehow included in the content features
MAIN CHALLENGES

Sparsity of Data

Size of Data
(Data Computation)

SERVERS!

SERVERS!

SERVERS!