ISSS603 – CUSTOMER FOCUSED ANALYTICS AND IT

Santander

PRODUCT RECOMMENDATION FOR DIGITAL MARKETING

GROUP 4

CHIN | JINGLONG | KAUSHIK | PRIYADARSAN | VAISHNAVI | VISHALI

INTRODUCTION & BUSINESS MOTIVE

Banco Santander

- Spanish multinational commercial bank in Santander, Spain
- Offers an array of financial services and products
  Retail banking, mortgages, corporate banking, cash management, credit card, capital markets, trust and wealth management, and insurance.

Current system

- Small number of Santander's customers receive many recommendations while many others rarely see any, resulting in an uneven customer experience

Business Analytics Objective

- To analyse and profile Santander's bank customers, products and services
- Analyse product ownership associations to aid targeted mass marketing and website design strategies
- Improve personal product recommendation system for its customers to aid relationship managers reach out to customers effectively

DATA DESCRIPTION AND EDA

BANCO SANTANDER

13M customers

Product ownership and demographics data

24 products

Accounts and services (Current account, Home, Credit cards, securities, funds, etc.)

18 months

- Monthly updates of activation and deactivation
- One but last month chosen for analysis and last month for scoring

Demographics

- From 54 provinces
- 50-50% ratio of Male and Females
- 24.6% missing values for Gross Income
- Highly skewed nominal variables (residency index, employee index, etc.)

Relationship

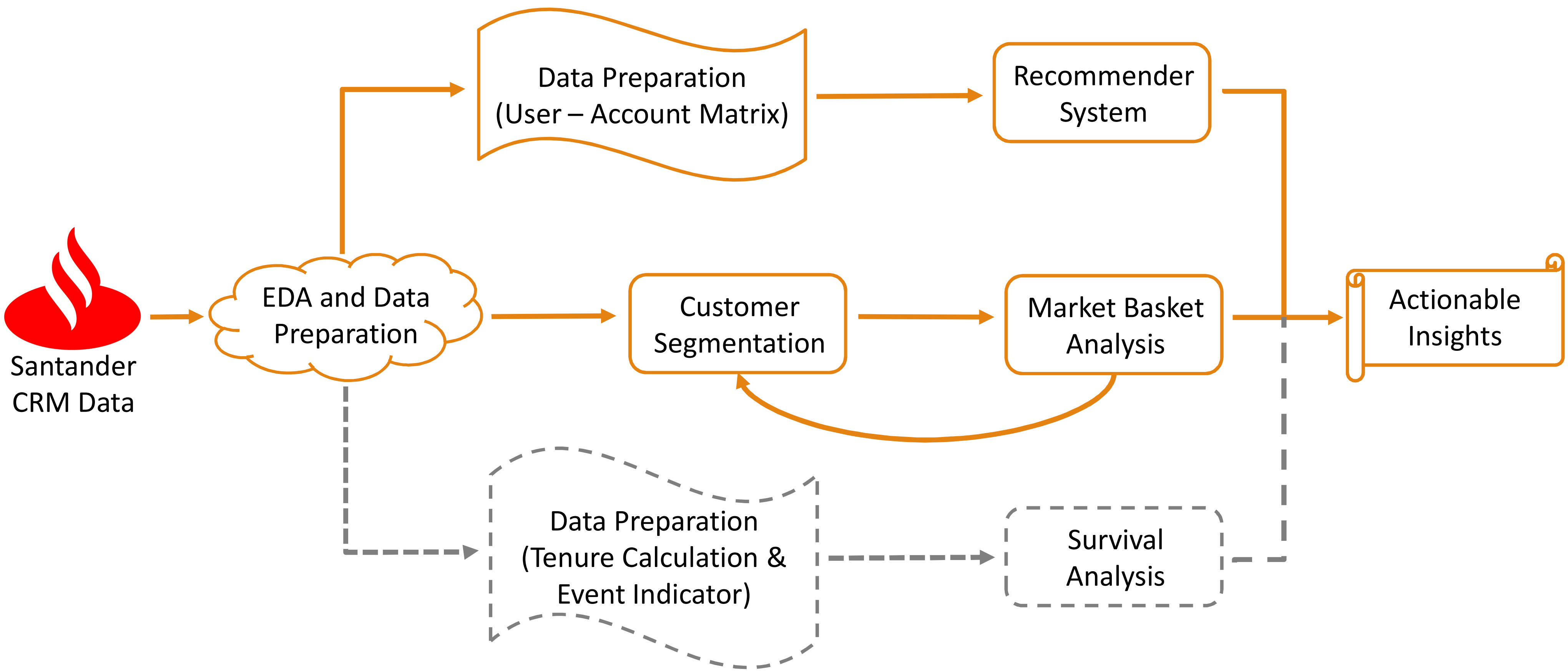
- Seniority – relation with bank in months
- New Customer Index- indicates if customer joined in the last 6 months
- 99.8% missing value for Last date as primary variable

DATA PREPARATION

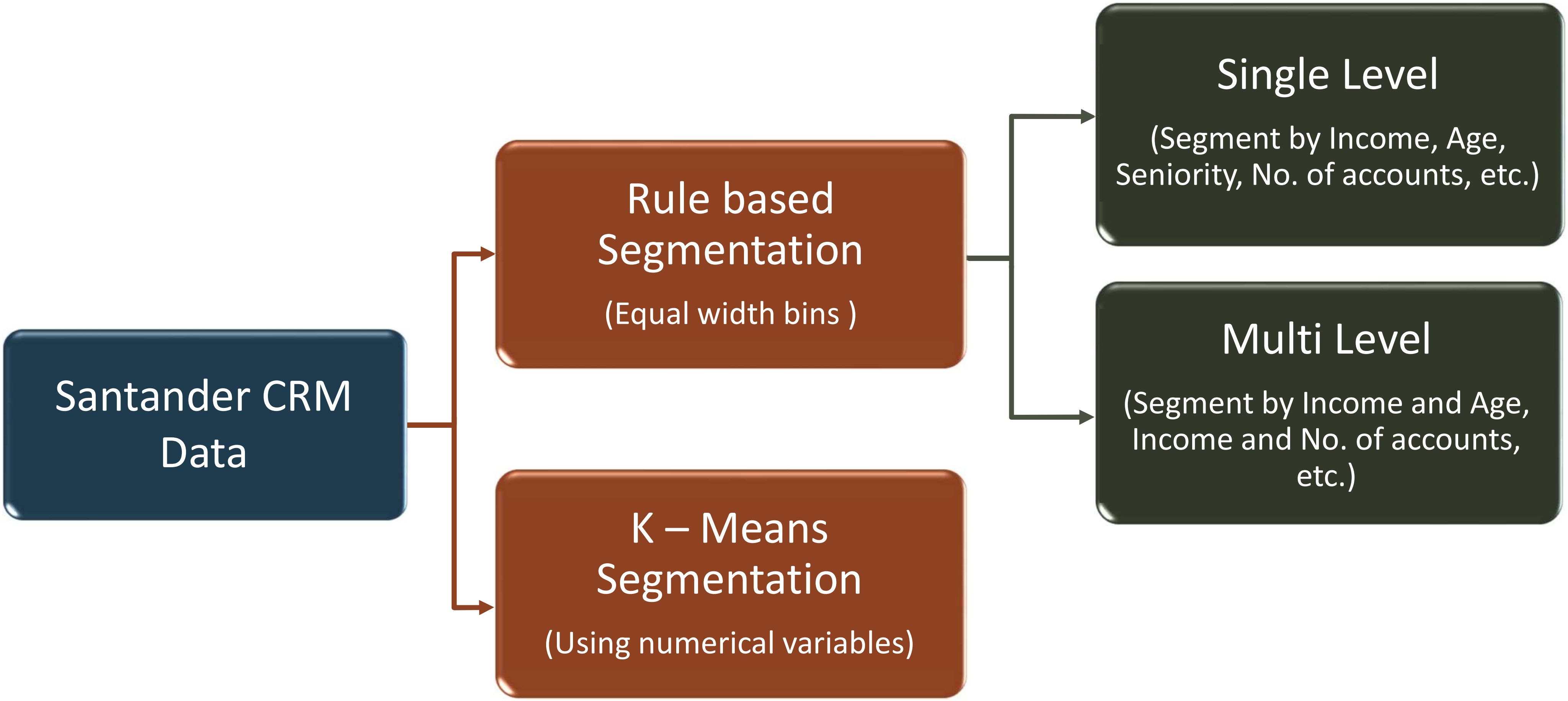
1. Subset the data by selecting one month data for analysis
2. Remove rows with invalid values for variables and impute missing values
3. Exclude skewed variables from analysis and recode the variables
4. Feature Engineering

   (*# of Accounts, Province Segment*)

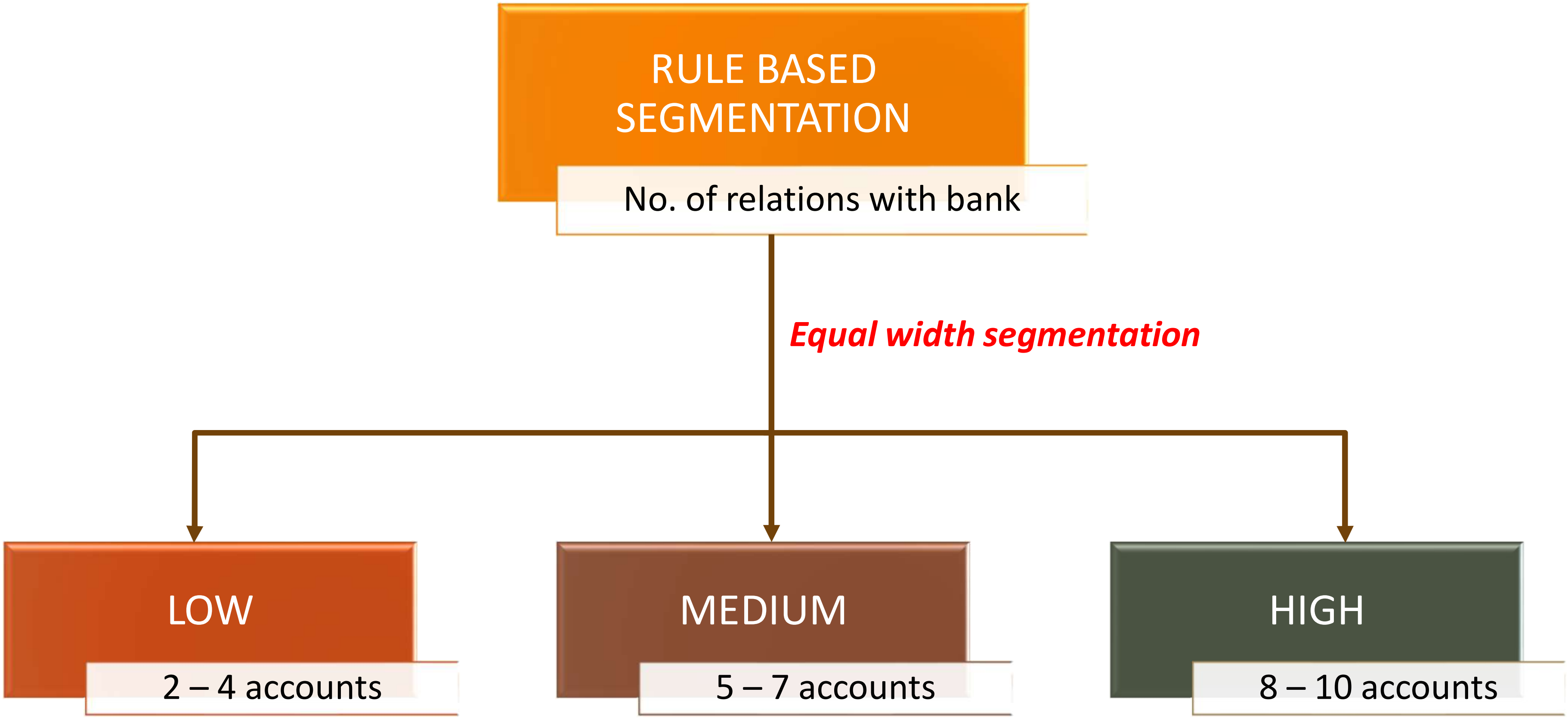
ANALYSIS PROCESS FLOW



CUSTOMER SEGMENTATION



CUSTOMER SEGMENTATION



MARKET BASKET ANALYSIS

RULE BASED SEGMENTATION

# of relations with bank

LOW

2 – 4 accounts

ASSOCIATION RULES

Potential accounts

- ❑ Pensions
- ❑ Credits Card

MEDIUM

5 – 7 accounts

ASSOCIATION RULES

Potential accounts

- ❑ Funds
- ❑ Long Term Deposits

HIGH

8 – 10 accounts

ASSOCIATION RULES

Potential accounts

- ❑ Home Account

ACTIONABLE STRATEGY 1

Website flyers have to be updated starting of each month

Santander

Support Talk to us Search

Current accounts Credit cards Insurance Loans Mortgages Savings and investments

Presenting a way to save

With our eISA
For 1|2|3 World or Santander Select customers.

1.50%
AER/Tax-free (variable) for 12 months

Click here to apply

Save from £500. You can transfer ISAs you have with other providers to Santander at any time. This is an instant access cash ISA that is managed in Online or Mobile Banking. Subject to availability.

ACTIONABLE STRATEGY 1

Website flyers have to be updated starting of each month

Ms. Alice
(Marketing Manager Santander)

Strategy for May - I want to push potential products for each customer segment

Mr. KJ
(Head of Analytics Santander)

Based on association rules, I can provide you with potential products to market

*Practical Implementation – Market **Pensions** and **Credit Cards** using website flyers for customers in Low segment*

ACTIONABLE STRATEGY 2

When Product Manager wants to market only a specific product

Q2 Target
Increase number of **Direct Debits** by 5%

Mr. Chin
(Direct Debits Manager Santander)

Can provide you list of potential customers to target

Mr. KJ
(Head of Analytics Santander)

Ms. Alice
(Marketing Manager Santander)

We can increase visibility of Direct Debits by sending email brochures for potential customers

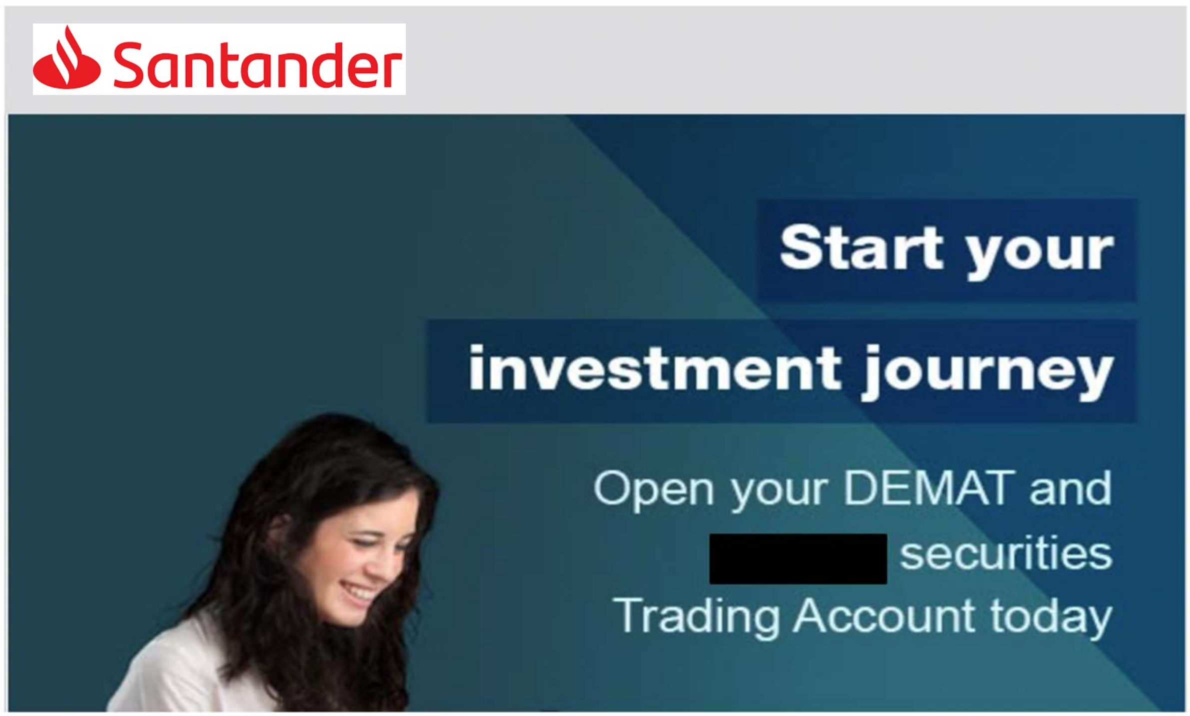
ACTIONABLE STRATEGY 2

When Product Manager wants to market only a specific product

Special investment benefits for you Inbox ×

[redacted]bank.net> Unsubscribe Mar 28, 2019, 10:31 PM

[redacted]



ACTIONABLE STRATEGY 2

Shortlisted rules for 5 products

Marketing Account (Target Account)	Antecedents (Recommend if customer already uses these accounts and doesn't have target account)	Customer Segment
Direct Debits	Payroll Account	Low
Long Term Deposit	Funds	Medium
	Current Account	Medium
	Eaccount	Low
Payroll Account	Pensions	Low
	Direct Debits	Low
Home Accounts	Particular Account	High
Funds	Long Term Deposits	Medium
	Securities	Medium
	Long Term Deposits	High

Santander Product Manager Dashboard

Select your target account to find potential customers to send mass marketing campaigns

Select account | Direct Debit Account

This excel sheet consists of customer ID's to whom you can market

Download Excel sheet of potential customers

RECOMMENDER SYSTEM

- Santander bank's core business involves a client relationship manager guiding a client's investment decisions
- Recommender systems can aid relationship managers with making personalized and automated selection of next best products for private banking clients

CHALLENGES

- CRM data only comprises of indicators for 24 products offered by the bank
- Data has no explicit feedback of preference, only indicators of product purchase or service use
- Implicit feedback is also not directly quantifiable to confidence as it's unary.

RECOMMENDER SYSTEM APPROACH

Collaborative Filtering Recommender Systems

- Implicit Feedback CF with Alternating Least Squares
- Baseline: User – User similarity based
- Enhanced: User – User similarity based enhanced by demographic correlations

Data split

- TRAIN – One but last month data
- TEST – Latest month data (De-activations ignored)

Evaluation metrics

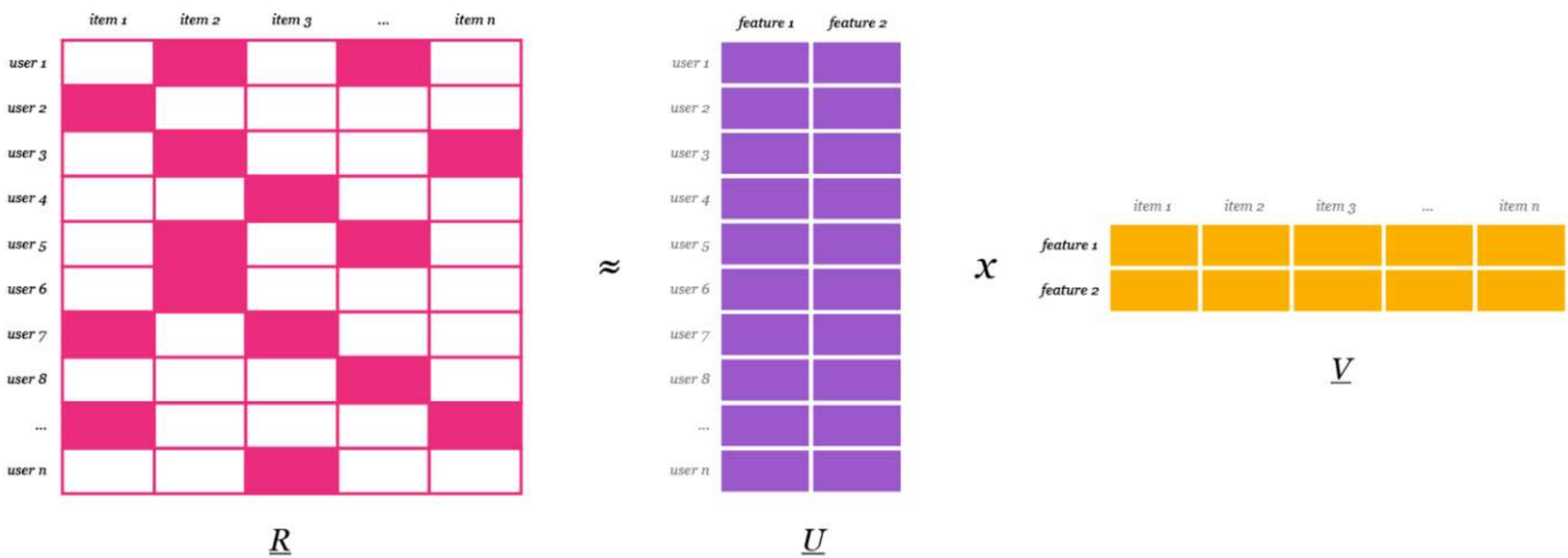
- Precision @ K
- Recall @ K

COLLABORATIVE FILTERING WITH IMPLICIT FEEDBACK

- Generic collaborative filtering using latent factorization represents user ratings as:

$$\hat{r}_{ui} = q_i^T \cdot p^u$$

Where $\hat{\mathbf{r}}_{\mathbf{ui}}$ is the predicted rating of user u for item i, and $\mathbf{q}_i^T$ and $\mathbf{p}^u$ are the latent factors



COLLABORATIVE FILTERING WITH IMPLICIT FEEDBACK

- Generic collaborative filtering using latent factorization represents user ratings as:

$$\hat{r}_{ui} = q_i^T \cdot p^u$$

Where $\hat{\mathbf{r}}_{\mathbf{ui}}$ is the predicted rating of user u for item i, and $\mathbf{q}_i^T$ and $\mathbf{p}^u$ are the latent factors

- The optimization involves the minimization of the objective function:

$$min \sum (r_{u,i} - q_i^T \cdot p^u)^2 + \lambda(||q_i||^2 + ||p_u||^2)$$

- For implicit feedback, "ratings" are modelled as a function of Preference & Confidence:

$$min \sum c_{ui}(p_{ui} - q_i^T \cdot p^u)^2 + \lambda(||q_i||^2 + ||p_u||^2)$$

where, Confidence $c_{ui} = 1+\alpha * r_{ui}$ and Preference $p_{ui} = \begin{cases} 1, r_{ui} > 0 \\ 0, r_{ui} = 0 \end{cases}$

- Alternating Least Squares is used for the optimization as the $q_i^T \cdot p^u$ term makes the loss function non-convex

USER BASED CF ENHANCED BY DEMOGRAPHIC CORRELATION

USER – PRODUCT MATRIX

	CURRENT	*PARTICULAR*	*PENSIONS*	*PAYROLL*	*HOME*
USER 1					
USER 2					
USER 3					
USER 4					
USER 5					

	USER 1	*USER 2*	*USER 3*	*USER 4*	*USER 5*
USER 1		0.55	0.88	0.67	0.29
USER 2	0.55		0.65	0.51	0.23
USER 3	0.88	0.65		0.23	0.65
USER 4	0.67	0.51	0.23		0.29
USER 5	0.29	0.23	0.65	0.29	

USER – DEMOGRAPHICS MATRIX

	REGION *MADRID*	REGION *GIRONA*	SEX *M*	SEX *F*	AGE *< 25*	AGE *> 25*
USER 1						
USER 2						
USER 3						
USER 4						
USER 5						

	USER 1	*USER 2*	*USER 3*	*USER 4*	*USER 5*
USER 1		0.62	0.83	0.04	0.38
USER 2	0.62		0.77	0.86	0.21
USER 3	0.83	0.77		0.10	0.62
USER 4	0.04	0.86	0.10		0.54
USER 5	0.38	0.21	0.62	0.54	

- Calculate pairwise USER-USER similarity using COSINE similarity measure using USER-ITEM vectors
- Calculate pairwise USER-USER similarity using COSINE similarity measure using USER-DEMOGRAPHY vectors

USER BASED CF ENHANCED BY DEMOGRAPHIC CORRELATION

	USER 1	USER 2	USER 3	USER 4	USER 5
USER 1		0.55	0.88	0.67	0.29
USER 2	0.55		0.65	0.51	0.23
USER 3	0.88	0.65		0.23	0.65
USER 4	0.67	0.51	0.23		0.29
USER 5	0.29	0.23	0.65	0.29	

	USER 1	USER 2	USER 3	USER 4	USER 5
USER 1		0.62	0.83	0.04	0.38
USER 2	0.62		0.77	0.86	0.21
USER 3	0.83	0.77		0.10	0.62
USER 4	0.04	0.86	0.10		0.54
USER 5	0.38	0.21	0.62	0.54	

$$UU\ Sim_{enhanced} = UU\ Sim_{user\_item} + (UU\ Sim_{user\_item} * UU\ Sim_{user\_demography})$$

	USER 1	USER 2	USER 3	USER 4	USER 5
USER 1		0.65	1.49	0.47	0.19
USER 2	0.65		0.93	0.69	0.10
USER 3	1.49	0.93		0.08	0.82
USER 4	0.47	0.69	0.08		0.24
USER 5	0.19	0.10	0.82	0.24	

- Compute the enhanced User – User similarity using the formula stated above
- Identify K (=1000 in this case) nearest neighbours for each user based on the enhanced similarity scores
- Predict items to recommend as a weighted average of the preferences from each user's neighbourhood
- Remove the items already owned by user and return the recommendations ranked by score

MODEL COMPARISON

PRECISION @ K **RECALL @ K**

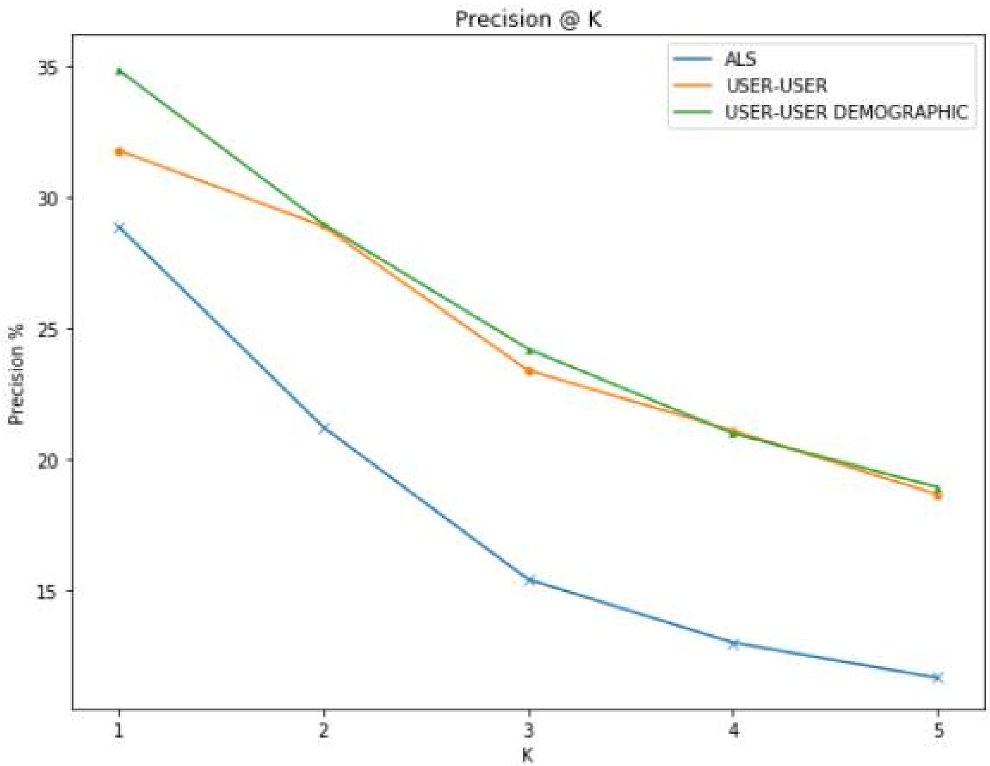
	K=1	K=2	K=3	K=4	K=5
IMPLICIT FB LATENT FACTORIZATION BY ALS	0.29	0.21	0.15	0.13	0.12
USER BASED CF FOR BINARY IMPLICIT FB	0.32	0.29	0.23	0.21	0.19
USER BASED CF WITH DEMOGRAPHIC CORRELATION	0.35	**0.29**	0.24	0.21	0.19

RECALL @ K = (# of recommended items @k that were purchased) / (total # of purchased items)

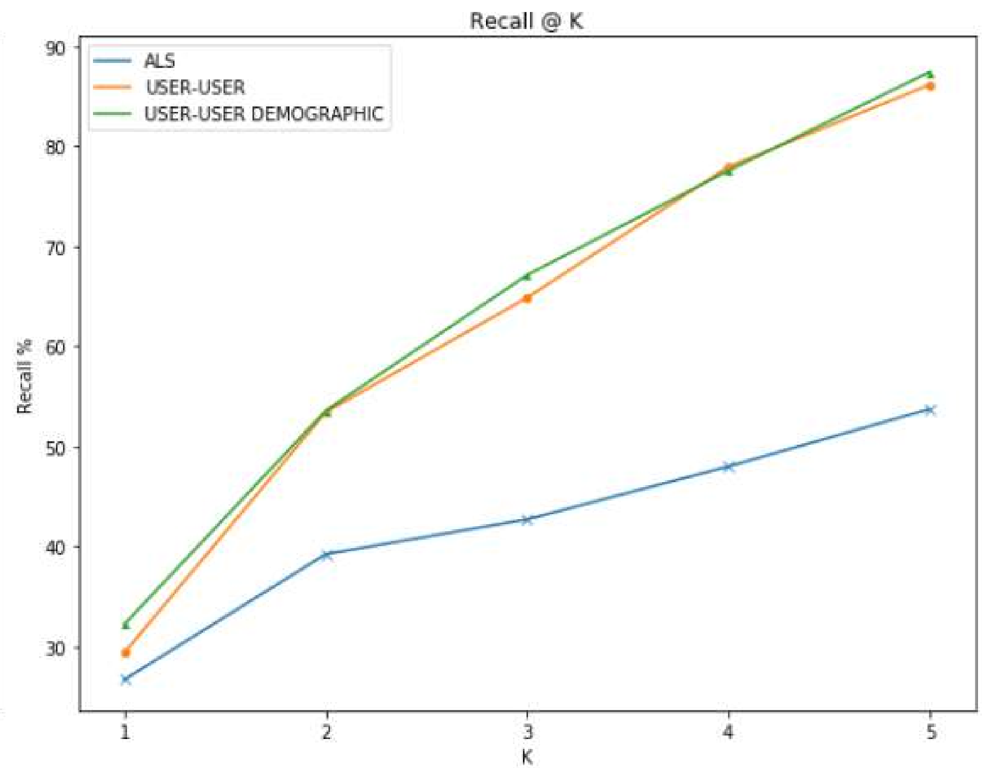
	K=1	K=2	K=3	K=4	K=5
IMPLICIT FB LATENT FACTORIZATION BY ALS	0.27	0.39	0.43	0.48	0.54
USER BASED CF FOR BINARY IMPLICIT FB	0.29	0.53	0.65	0.78	0.86
USER BASED CF WITH DEMOGRAPHIC CORRELATION	0.32	**0.54**	0.67	0.78	0.87

MODEL COMPARISON

PRECISION @ K



RECALL @ K



RECOMMENDER SYSTEM FOR ASSISTED DECISION MAKING

Relationship Manager A wants to reach out to his client X with a product offer

He pulls out his client's personal details and bank relations from the CRM database

Inputs the details into the CF recommender system model packaged as an API and run the model to get the ranked recommendations

Select the NEXT BEST OFFER account based on his business knowledge

Manager reaches out to the client with an investment offer

Customer information	ID - 302370
Gender	M
Age	47
New customer?	N
Seniority in months	195
Foreigner Index	N
Province name	MADRID
Income	126,850

+

Customer held accounts
Current Account
Particular Plus Account
Securities Account

=

Rank	Recommended Accounts
1	**Particular Account**
2	**E-Account**
3	Direct Debit Account
4	Taxes Account
5	Credit card

THANK YOU