

Intelligent Miner for Data and Telecoms



Jelena Mamcenko
Vilnius Gediminas Technical
University



What are the telecoms business issues to address?

- ✍ What are the characteristics of my customers?
- ✍ Can I predict the customers who are likely to leave voluntarily and should we try to keep them?
- ✍ Can I identify the characteristics of customers who need special action as suspension or deactivation?
- ✍ ...



The Business Issue

 **How can we discover the characteristics of customers?**



1. Translate the business issue into a question

- ✍ Define segments based on the customer behavior



2. Identify the data

- ✍ Behaviour data (call, services)
- ✍ Demographic data
- ✍ Additional data to profile each segment

Data model for segmentation

	Variable name	Description
	CALLING BEHAVIOUR	
1	Regular_dur	Minutes of call in regular time frame
2	Discount_dur	Minutes of call in discount time frame
3	Night_dur	Minutes of call in night time frame
...		
	SERVICE BEHAVIOUR	
7	Svc_data	Number of internet services
8	Svc_info	Number of information services
9	Svc_free	Number of free services
...		
	DEMOGRAPHIC DATA	
16	Age	Customer age
17	Gender	Customer gender
18	Job	Customer job
	ADDITIONAL DATA	
19	Agent	Agent office where the phone was initially bought
20	Price_plan	Price plan
21	Tenure	Account age (how many month customer with your company)
...		

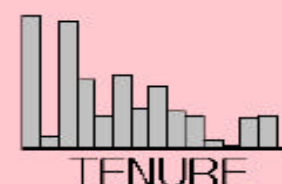
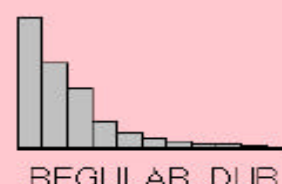
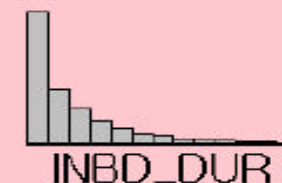
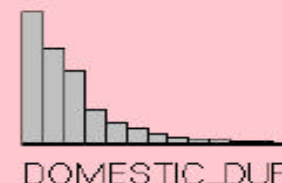
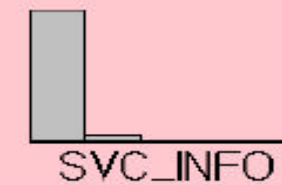
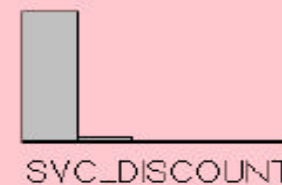
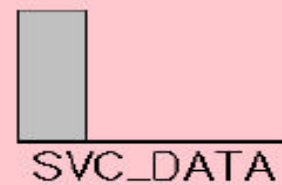
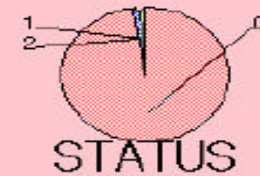
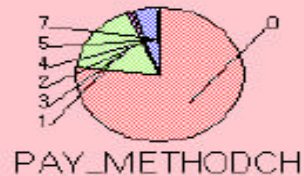
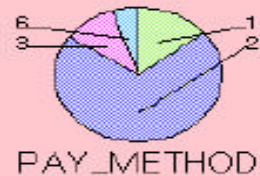
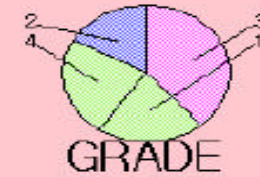
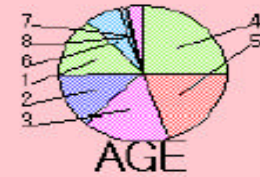
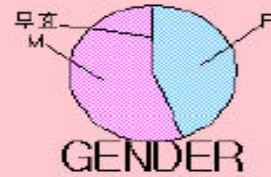


3. Sourcing and preprocessing the data

- ✍ To create our data model we have to take the raw data that we collect and convert it into the format required by the data models.

4. Evaluating the data

segmentation data explore

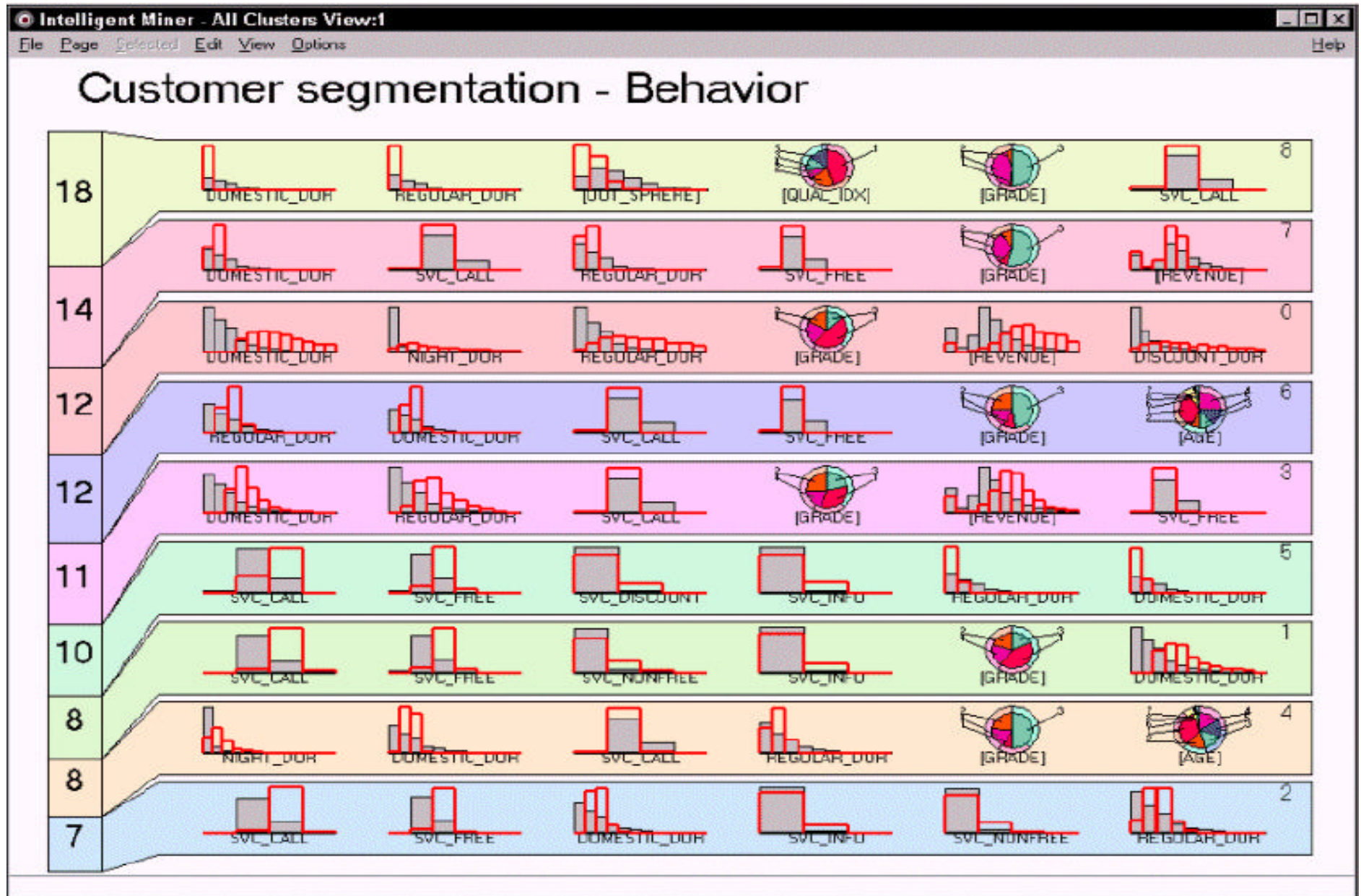




5. Choosing the mining techniques

- ✍ The segmentation technique is chosen to identify customer behaviour.

6. Interpreting the results





Interpreting the results from business perspective

- ✍ The customers who have high call usage and low service utilization (Segment 8)
- ✍ The customers who have low call usage but high service utilization (Segment 5)
- ✍ The customers who have high call usage at certain time frame (Segment 4)



Segment 4 has customers who have characteristics, such as:

- ✍ Mostly call at night (NIGHT_DUR)
- ✍ Not interested in most services (SVC_CALL, SVC_FREE) except friends and family type of discount service (SVC_DISCOUNT)
- ✍ ...
- ✍ Teenage students (AGE, JOB)

Summary

Segment	Segment name	Relative size	Call behaviour	Service behaviour	Revenue
8	Basic	18,1%	VeryLow	Low	10,6
7	Economic	13,6%	Low	Low	16
0	Premium - Young	12,0%	VeryHigh	Medium	41,8
6	Conservative	11,9%	Medium	Low	21,3
3	Inbound	11,1%	High	Low	31,1
5	Service - oriented	9,7%	Low	High	15
1	True mobile	8,5%	VeryHigh	High	37,7
4	Night friends	7,8%	Medium	Low	20,2
2	Regular teen	7,5%	Medium	High	21,8

Summary

Rank	Segment size	Revenue share
1	Segment 8 (18,1%)	Segment 0 (21,8%)
2	Segment 7 (13,6%)	Segment 3 (15,0%)
3	Segment 0 (12,0%)	Segment 1 (13,9%)
4	Segment 6 (11,9%)	Segment 6 (11,1%)
5	Segment 3 (11,1%)	Segment 7 (9,5%)
6	Segment 5 (9,7%)	Segment 8(8,4%)
7	Segment 1 (8,5%)	Segment 2 (7,1%)
8	Segment 4 (7,8%)	Segment 4 (6,9%)
9	Segment 2 (7,5%)	Segment 5 (6,3%)



???