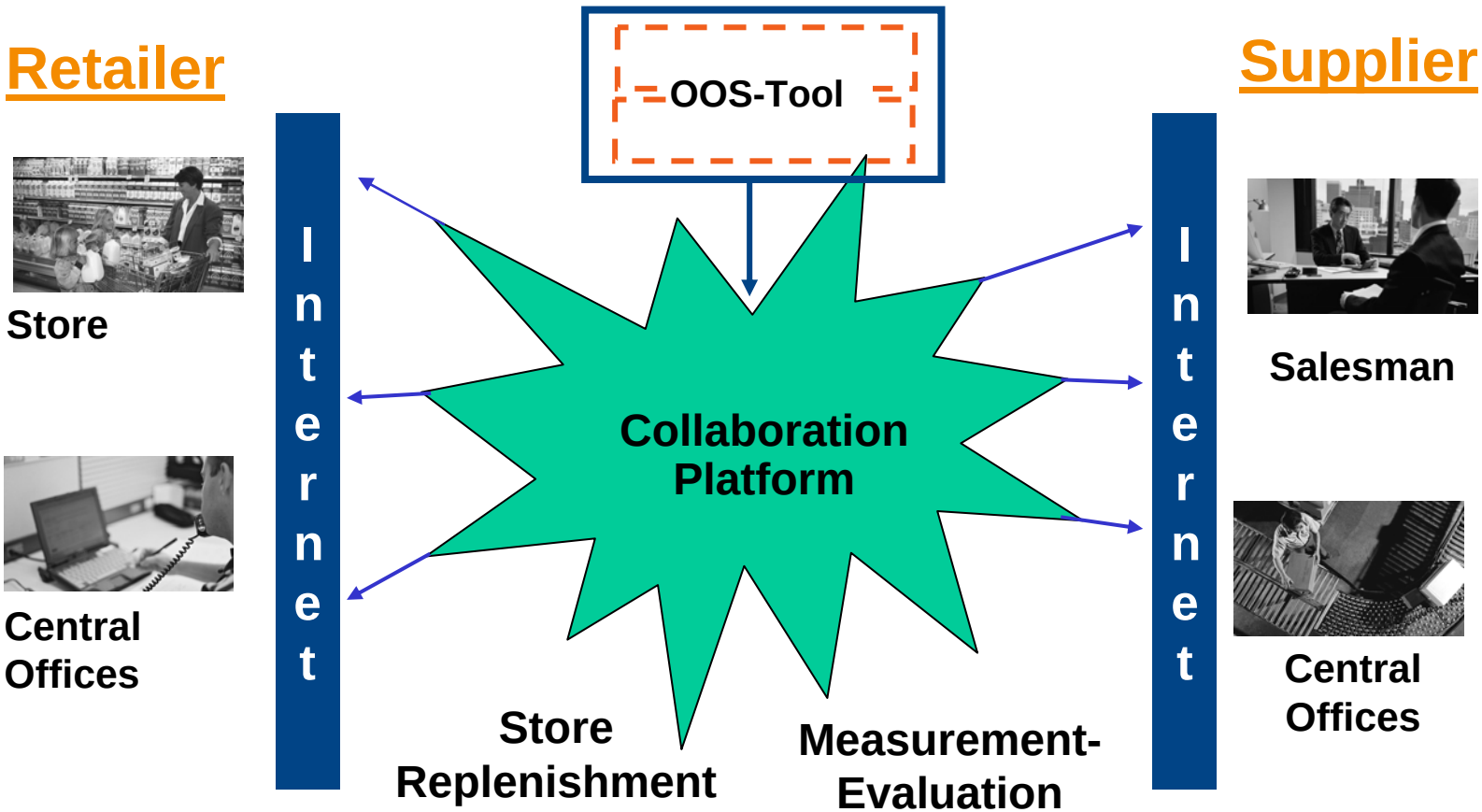
A collaborative approach for measuring and solving the Out-Of-Stock problem

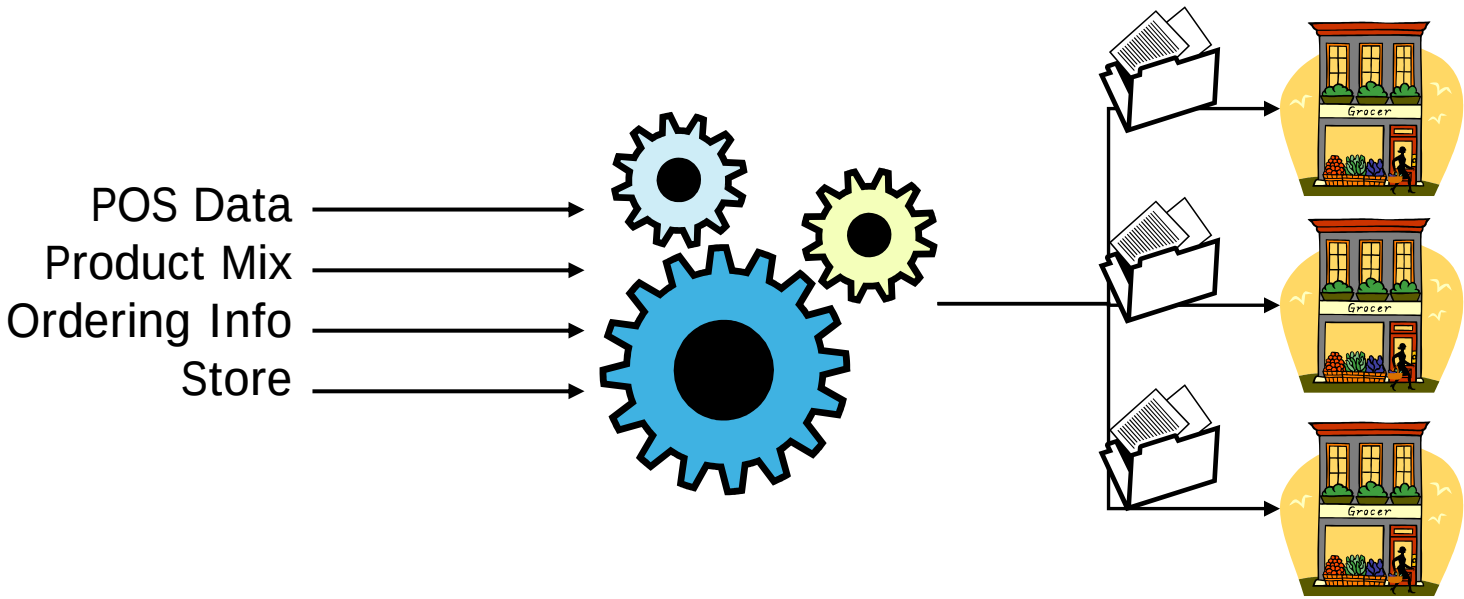
Dimitris Papakiriakopoulos
dpap@aueb.gr
ELTRUN
Department of Management Science and Technology
Athens University of Economics and Business

A Collaborative approach



OOS Tool-Research Question

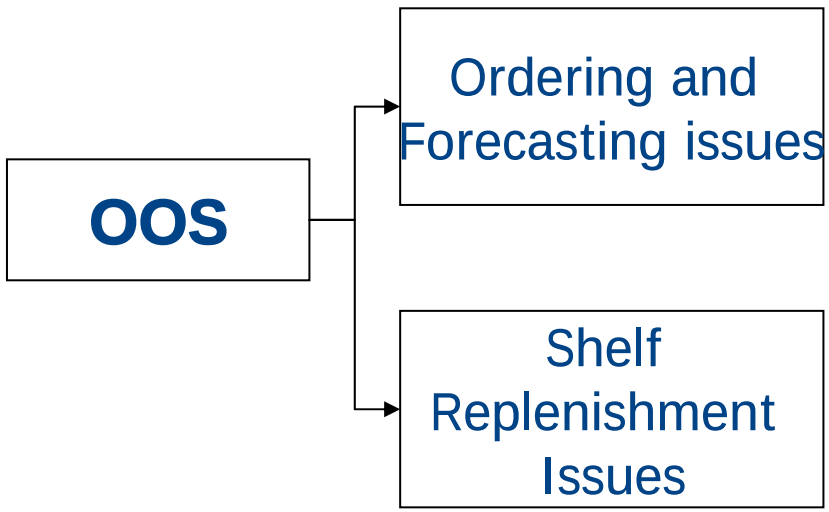
- Can we construct an algorithm to detect OOS situations, without having to visit the store?



Output:

Product Name	Code	Description
Product 1	111111	Stock Out
Product 2	222222	Too many days without sale
Product 3	333333	Too many days without sale
Product 4	444444	Unexpected low sales

Causes of the Out-Of-Shelf Problem



•The product is out of stock at the store
•The Central Warehouse could not deliver the product because it was out of stock
•The ordered quantity couldn't meet consumer demand

•Product is at the store but not on the shelf
•During Saturday, the personnel of the store is not enough to replenish all the shelves

- Marketing Perspective
  - Focuses on the consumer reaction
  - Proposes that OOS negatively affects consumer loyalty
  - Most of the studies agree that the normal OOS rate is between 5%-10%

- Operational Research Perspective
  - Focuses on inventory levels (Stock-Out-Problem)
  - Proposes various models for determining ordering/replenishment parameters

Thinking about OOS-cases

- Based on the duration of the OOS
  - Full OOS: For the whole day the shelf was empty
  - Partial OOS: For some hours the shelf was empty
- Based on the type
  - A product is not on the shelf and is not at the store
  - A product is not on the shelf but there are some units in the backroom
  - A product has never been in the store although it is in the assortment
- Based on the identification
  - Easy to capture (e.g. fast moving items)
  - Difficult to capture (e.g. products at the top of the shelf)
- Based on the result
  - Daily anomaly: the OOS product sold less than expected
  - Long term anomalies: brand shift or left the store
- Based on the store managers intentions
  - Deliberately: store manager is waiting for 3 days in order to place an order with more products
  - Overlooking: the store manager forgot to place the order

Approaches and Barriers

- One day before
  - Forecast the OOS situations
    - Demand forecasting
- At the end of the day
  - Estimate Stock Level (Inbound – Outbound)
    - Unreliable stock data
    - Doesn't capture shelf out-of-stocks
  - Classify new cases utilizing prior knowledge
    - Statistical methods
      - Classical Statistics (LDA)
      - Modern Statistics – Decision Trees (Naive Bayes)
    - Machine Learning (C4.5, ADTree, RandomForest)
      - Decision Trees
      - Rule Based Algorithms
    - Neural Networks (MLP, RBF)

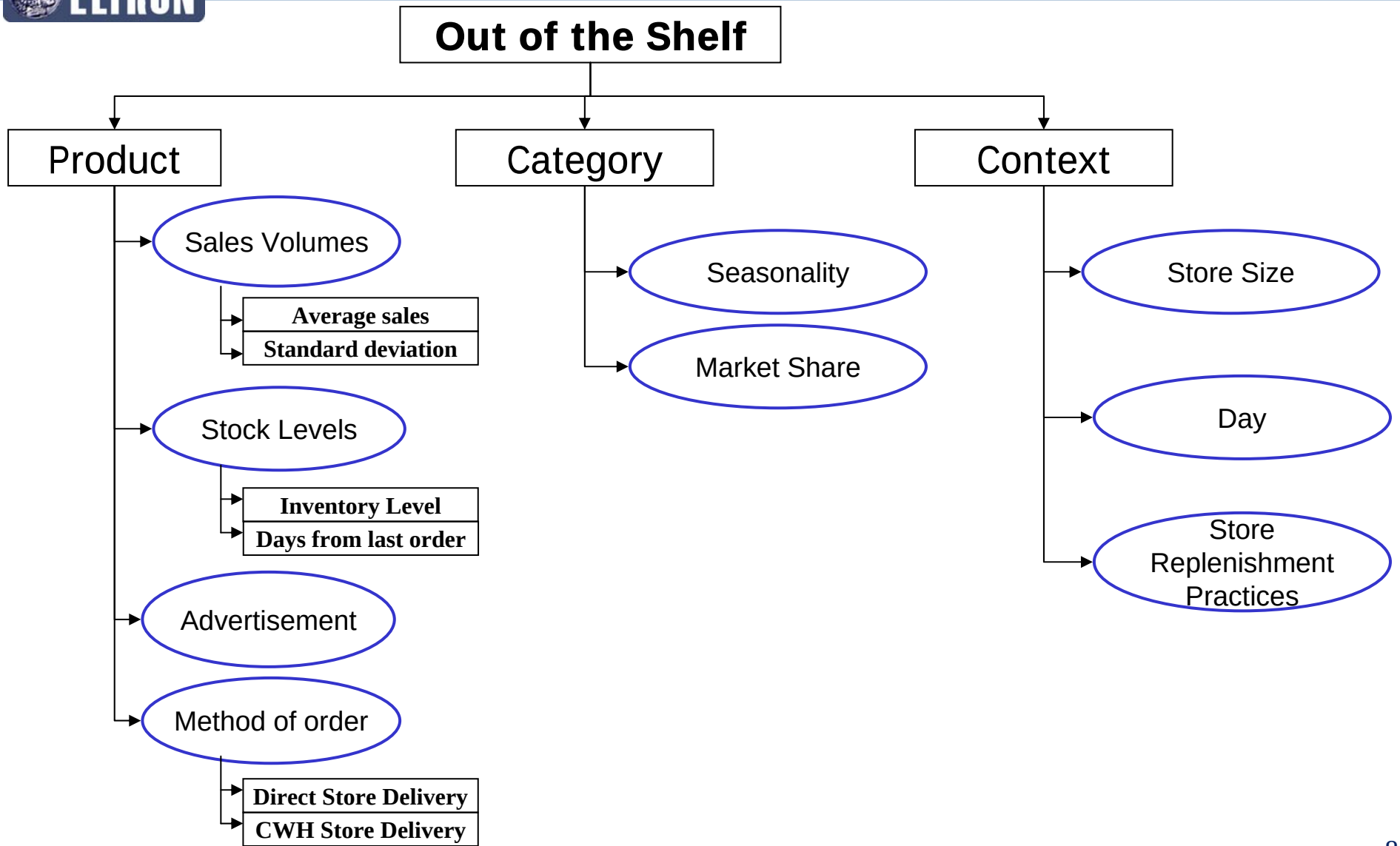
OOS as a classification problem

- Establish **rules** whereby we can classify a new observation into one of the existence classes (Supervised Learning).
- We need a training data set

Variable 1	Variable 2	Variable 3	Variable 4	Variable 5	Result
0.3328	YES	1	0	TRUE	EXISTS
1.333	YES	1	3	TRUE	EXISTS
1.0223	NO	1	1	TRUE	EXISTS
0.0732	NO	2	0	FALSE	OOS

- Examples of rules
  - **IF** (Variable2='NO' **AND** Variable4<0.5) **THEN** RESULT = OOS
  - **IF** (Variable5='TRUE') **THEN** RESULT = EXISTS
- The algorithm for detecting the OOS situations includes a set of rules

Variables related with the OOS



From variables to features

ELTRUN

Product Related

- Sales Volumes
  - Daily Mean & standard deviation
  - Daily Mean of selling and standard deviation
  - No selling periods {zero average, zeros std, last zero sequence}
  - Fast Moving Index [0..1]
  - Today's sales
- Stock Levels
  - Inventory Exists = {Stock Exists, ?}
  - Last order = {date,?}
  - Pending Order = {Pending Order, ?}
- Advertisement = {Promotional, Regular}
- Method of Order = {Direct Store Delivery, CWH Delivery, Unknown Delivery}

Category Related

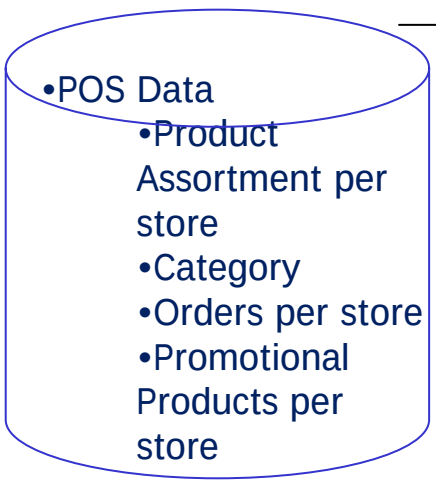
- Seasonality = {Yes, No}
- Market Share = products sells / category sales

Context Related

- Store size = {small, medium, large}
- Day= {Mon, Tues, Wedn, Thur, Frid, Sat}

Experiment Description

- Duration : 2/2/2004 – 7/2/2004
- Store Selection
  - Representative stores provided by retailer (9 stores)
- Product Selection
  - 109 products were selected
  - Stratified Sampling (9 strata based on sales average and sales standard deviation)
  - Add one stratum with promotional products
- Physical store visits to capture OOS



Sales Average	Sales Std. deviation	.....	Store Size	Day	Result
0.382	1.13	…….	Large	Monday	EXISTS
1.837	4.329		Large	Tuesday	EXISTS
12.372	9.382		Large	Monday	EXISTS
0	?		Medium	Monday	OOS
			Small	Saturday	EXISTS

Selecting a Classification Algorithm

Algorithm	Correctly Classified	Mean Absolute Error	Precision	Confusion Matrix
ADTree	90.401 %	0.2398	0.936 0.691	4792 237 a=ex 328 529 b=oos
DecisionStump	89.4326%	0.1677	0.936 0.642	4733 296 a=ex 326 531 b=oos
C4.5	93.4591 %	0.0962	0.948 0.835	4913 116 a=ex 269 588 b=oos
LMT	92.2528%	0.104	0.944 0.771	4860 169 a=ex 287 570 b=oos
NBTree	92.9664 %	0.1117	0.941 0.841	4962 103 a=ex 311 546 b=oos
RandomForest	93.1702 %	0.1073	0.941 0.856	4937 92 a=ex 310 547 b=oos
REPTree	92.8304 %	0.1102	0.941 0.853	4910 119 a=ex 303 554 b=oos

Selecting the appropriate rules (1/2)

Rule	Class	Length	Support	Confidence	Capture
0	exists	0	100.0%	84.9%	100.0%
1	oos	3	3.7%	91.3%	22.4%
2	exists	1	22.3%	84.9%	22.3%
3	oos	3	1.4%	89.4%	8.5%
4	oos	1	13.7%	64.5%	58.9%
5	exists	4	0.6%	66.7%	0.5%
6	oos	4	0.1%	100.0%	0.7%
7	exists	3	1.9%	96.4%	2.1%
8	oos	2	3.8%	55.8%	14.0%
9	exists	1	3.2%	96.3%	3.6%
10	exists	1	21.6%	97.7%	24.9%
11	exists	2	9.4%	99.1%	10.9%
12	exists	2	16.3%	96.5%	18.6%
13	exists	1	29.0%	96.0%	32.8%
14	exists	1	25.3%	72.3%	21.6%
15	exists	3	0.5%	93.8%	0.6%
16	exists	2	2.7%	81.5%	2.6%
17	exists	3	3.6%	99.1%	4.2%
18	oos	5	0.1%	100.0%	0.9%
19	oos	2	5.9%	82.6%	32.5%
20	exists	2	17.7%	89.9%	18.7%

Rule1 IF store_type = 1 AND zero_avg >= 31.25 AND zero_seq < 4 THEN RES = oos

Rule2 IF store_type = 2 THEN RES = exists

Rule3 IF store_type = 3 AND zero_avg >= 31.25 AND zero_seq < 7 THEN RES = oos

Rule4 IF zero_avg >= 31.25 THEN RES = oos

Rule5 IF store_type = 1 AND zero_avg >= 31.25 AND zero_seq < 7 AND zero_seq >= 4 THEN RES = exists

Rule6 IF store_type = 2 AND zero_avg >= 31.25 AND zero_seq < 23 AND zero_seq >= 14 THEN RES = oos

Rule7 IF store_type = 3 AND zero_seq < 12 AND zero_seq >= 7 THEN RES = exists

Rule8 IF zero_avg >= 63 AND zero_seq < 3 THEN RES = oos

Rule9 IF zero_avg < 1.230769 THEN RES = exists

Rule10 IF zero_avg < 1.583333 THEN RES = exists

Rule11 IF store_type = 3 AND zero_avg < 1.775 THEN RES = exists

Rule12 IF store_type = 1 AND zero_avg < 31.25 THEN RES = exists

Rule13 IF avg >= 1.384615 THEN RES = exists

Rule14 IF zero_seq >= 4 THEN RES = exists

Rule15 IF zero_avg < 31.25 AND zero_avg >= 21 AND zero_seq >= 42 THEN RES = exists

Rule16 IF zero_avg < 31.25 AND zero_avg >= 15.85714 THEN RES = exists

Rule17 IF store_type = 4 AND zero_avg < 5.823529 AND zero_avg >= 3.470588 THEN RES = exists

Rule18 IF store_type = 4 AND zero_avg < 31.25 AND zero_avg < 6.052631 AND zero_avg >= 5.823529 AND zero_seq >= 4 THEN RES = oos

Rule19 IF store_type = 1 AND zero_avg >= 31.25 THEN RES = oos

Rule20 IF store_type = 2 AND zero_avg < 51.5 THEN RES = exists

Selecting the appropriate rules (2/2)

- We have examined 12 classification algorithms.
- Each algorithm produces 140 rules on average
  - Only 15% of the rules refer to the OOS situations
- Most of the rules are common
- Various measures have been proposed to asses the prediction accuracy of the rules
  - Accuracy (Impurity)
  - Confidence
  - Simplicity
  - ....

Validation Method

- Validation methods
  - Through physical store visits
    - Develop the rule based algorithm
    - Run it in a trial phase and get the list of the OOS products
    - The same date visit the store and consolidate the results
  - Utilizing existing Machine Learning validation techniques
    - Re-sampling method
      - Repeated Holdout Method
      - Cross-validation
      - Bootstrap
    - Three way data partition
      - From the one single dataset construct
        - » A training set
        - » A validation set
        - » A test set

We have two more data sets from the same retailer
•September 2002
•August 2004

Results and Limitations

- Current results
  - A rule based algorithm has been develop to monitor the OOS situations, capturing approximately 60%-70% of the real OOS cases.

- Future work
  - Utilize the other two datasets
  - Further evaluation for
    - New products because of the limited history
    - Promotional products because of the demand fluctuations
  - Resample after six month to evaluate the “new form” of OOS