



# Change Mining of Customer Profiles based on Transactional Data

Edward Apeh and Bogdan Gabrys

Smart Technology Research Centre, School of Design, Engineering and Computing  
Bournemouth University , UK

[eapeh@bournemouth.ac.uk](mailto:eapeh@bournemouth.ac.uk) & [bgabrys@bournemouth.ac.uk](mailto:bgabrys@bournemouth.ac.uk)

# Outline

- ❑ The Problem: Static Model v. Time Variant Data
- ❑ Introduction to Change Mining
- ❑ Change Mining of Customer profiles
  - ⇒ Description of customer profiles
  - ⇒ Description of the model
  - ⇒ Description of the proposed approach
- ❑ Experiments & Results
  - ⇒ Effects of window size on prediction accuracy
  - ⇒ Classifier stability over time
- ❑ Conclusion and Future Work

## The Problem: Static Model v. Time Variant Data

- ❑ Data mining models tend to be static.
- ❑ Complex data structures evolve over time.
- ❑ Research is currently focused on model adaptation mechanisms.
- ❑ But approaches for analysing how and when the models and their decisions change are also desired.
- ❑ Change mining is concerned with these issues.

# Change Mining

- ❑ Change mining encompasses techniques for mining evolving data for the purpose of:
  - ⇒ Capturing the process of change;
  - ⇒ Analysing how models have changed;
  - ⇒ Predicting changes that will emerge.

# Customer Profiles



## □ Customer Profiles:

- are a collection of data describing individual users or groups;
- are either factual, rule-based or collaborated/verified;
- can be sourced electronically or off-line;
- provide businesses with vital information about their customers such as:
  - who their valuable customers are; and
  - how they behave.

# Customer Profiles

- Formally, given:

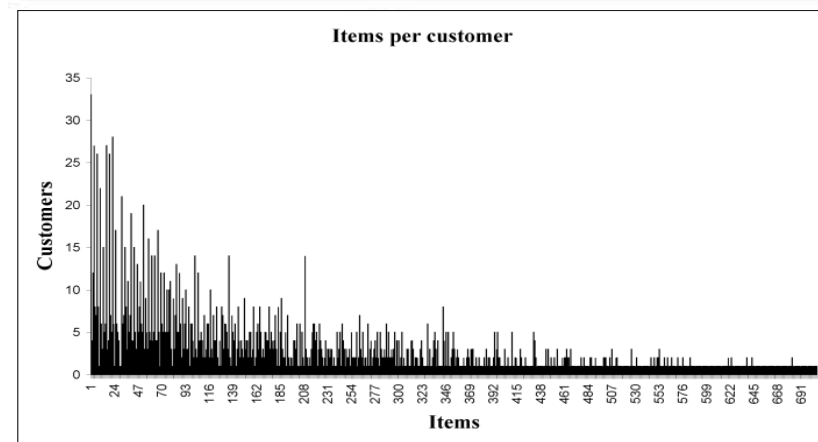
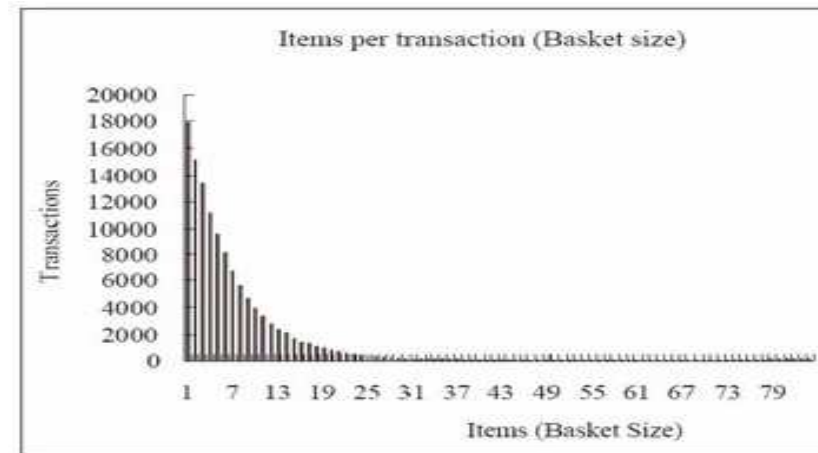
$$\mathbf{T} = \begin{bmatrix} t_{1,1} & t_{1,2} & \cdots & t_{1,d} \\ t_{2,1} & t_{2,2} & \cdots & t_{2,d} \\ \vdots & \vdots & \cdots & \vdots \\ t_{N,1} & t_{N,2} & \cdots & t_{N,d} \end{bmatrix}$$

- We define a set of  $n$  customer profiles as:

$\mathbf{P} = \{\mathbf{P}_1, \mathbf{P}_2, \dots, \mathbf{P}_n\}$  where  $\mathbf{P}_j = (t_1, t_2, \dots, t_M)^T$  and  $t_i \in \mathbf{T}_j$  is an aggregated transaction from the set  $\mathbf{T}_j$  of all the transactions of the  $j$ -th customer.

# The Issue with Transactional Data

- ❑ Transactional data is inherently:
  - Large
  - Sparse
  - Skewed
- ❑ Many traditional pre-processing and data mining techniques do not work.

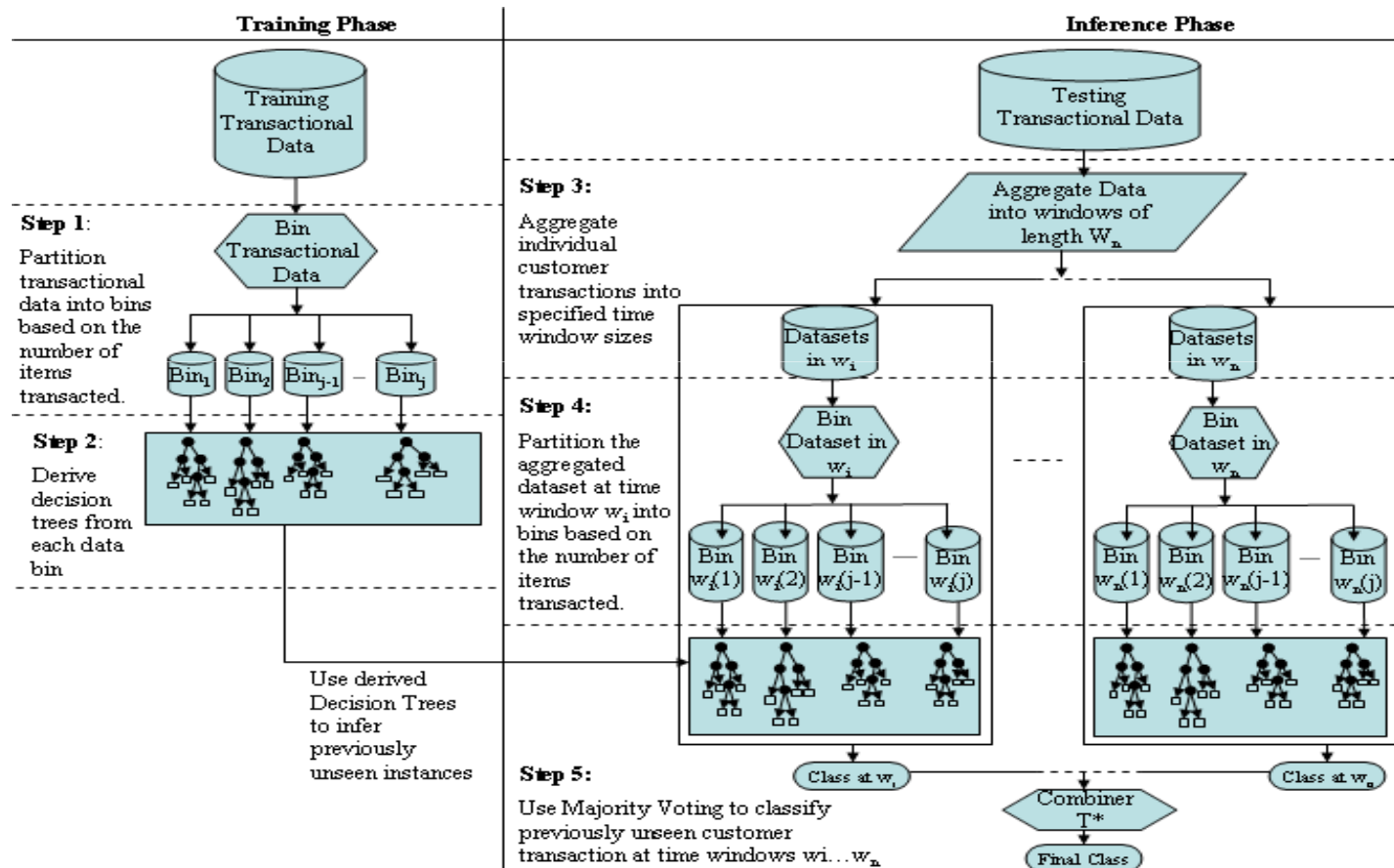


# Our Approach

- ❑ Partition data using data binning.
- ❑ Induce an ensemble of decision trees from the binned data.
- ❑ Monitor the change of customer profiles over time to gauge the effect of time on the classifier's:
  - ⇒ Accuracy &
  - ⇒ Stability.



# Proposed classification framework



## Experiment: Goals

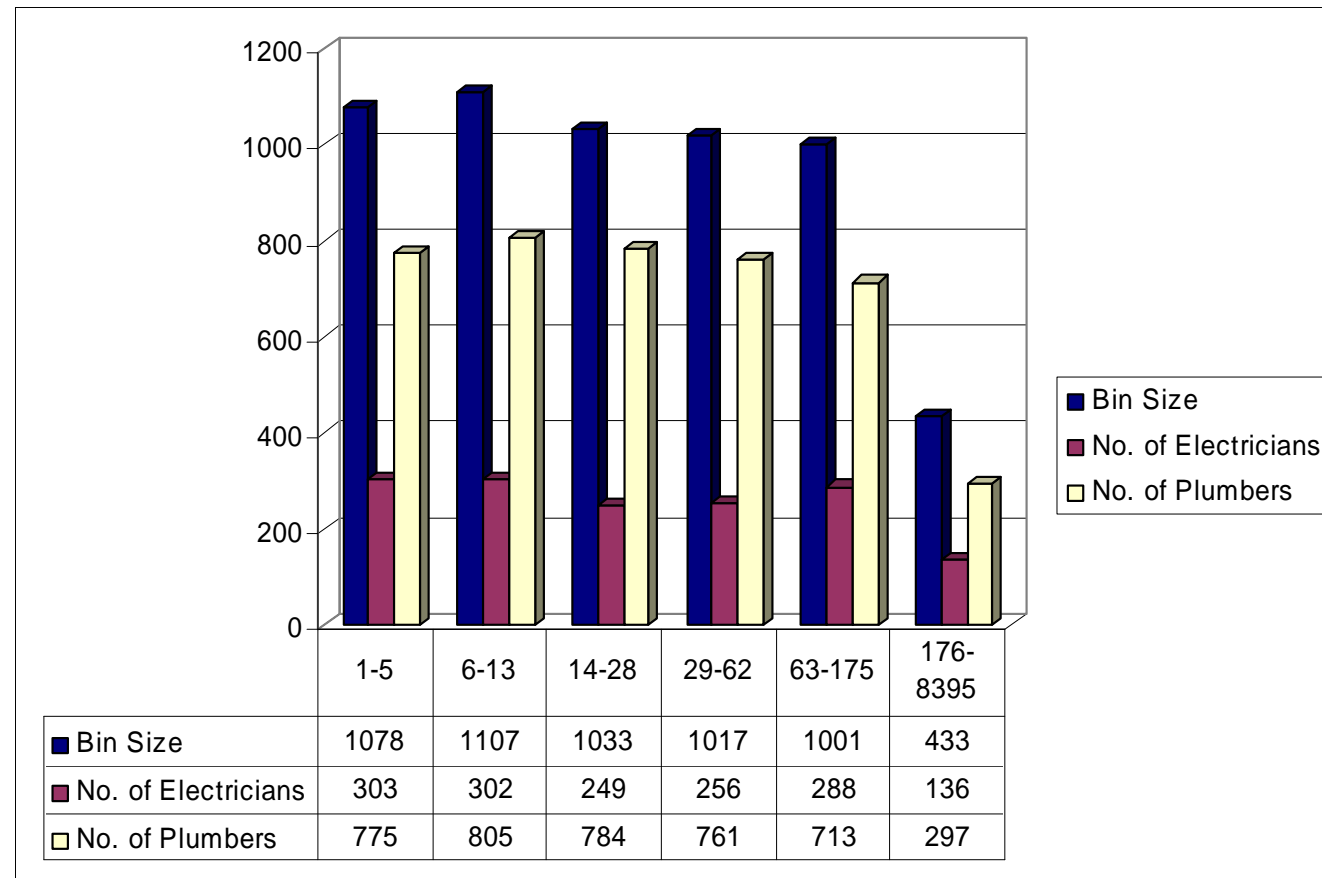
- ❑ Quantitatively capture and analyze the effect of evolving customer profiles on a classifier model over time by investigating:
  - ⇒ the effect of varying time windows on the prediction quality of a classifier over time;
  - ⇒ the prediction stability of a classifier in varying transaction time windows.

## Experiment: Data Description

| Profile Name | No. of Customers | No. of Transactions | No. of Items Transacted |
|--------------|------------------|---------------------|-------------------------|
| Electricians | 1537             | 32063               | 111730                  |
| PlumbHeaters | 4135             | 68715               | 230542                  |

- ❑ Data provided by Screwfix – a UK retail company.
- ❑ Data consists of transactions covering 30 months for 2 trade types: Electricians and PlumberHeaters which were verified with third party.

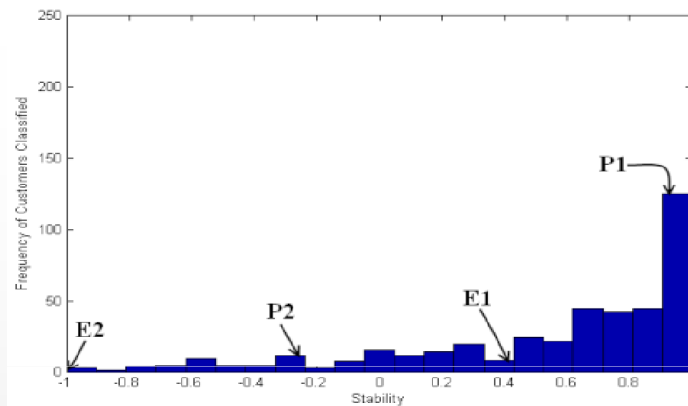
# Experiment: Data Bins



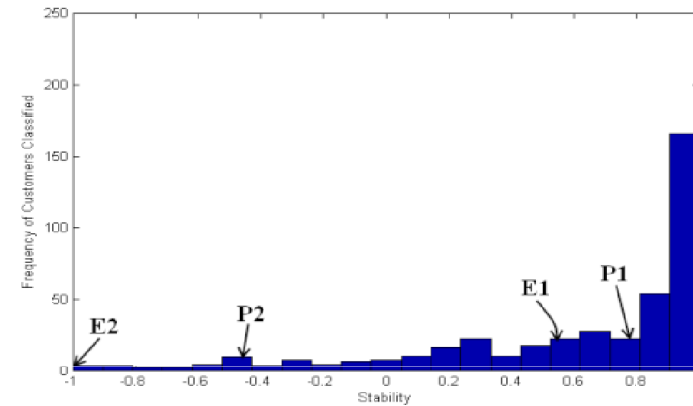
# Result & Analysis: Classifier Stability Overtime

| Sliding Window Dataset Partitions | Electrician (E)          |                          |                          |                           |                              | PlumbHeater (P)          |                          |                          |                           |                              |
|-----------------------------------|--------------------------|--------------------------|--------------------------|---------------------------|------------------------------|--------------------------|--------------------------|--------------------------|---------------------------|------------------------------|
|                                   | Class in 3 Months Window | Class in 6 Months Window | Class in 9 Months Window | Class in 12 months Window | Majority Vote Across Windows | Class in 3 Months Window | Class in 6 Months Window | Class in 9 Months Window | Class in 12 months Window | Majority Vote Across Windows |
| 1                                 | E                        |                          |                          |                           | E                            | P                        |                          |                          |                           | P                            |
| 2                                 | E                        |                          |                          |                           | E                            | P                        |                          |                          |                           | P                            |
| 3                                 | P                        |                          |                          |                           | P                            | P                        |                          |                          |                           | P                            |
| 4                                 | P                        | E                        |                          |                           | E                            | E                        | P                        |                          |                           | E                            |
| 5                                 | E                        | E                        |                          |                           | E                            | P                        | P                        |                          |                           | P                            |
| 6                                 | P                        | E                        |                          |                           | E                            | P                        | P                        |                          |                           | P                            |
| 7                                 | E                        | E                        | E                        |                           | E                            | P                        | E                        | P                        |                           | P                            |
| 8                                 | E                        | E                        | -                        |                           | E                            | P                        | P                        | P                        |                           | P                            |
| 9                                 | E                        | E                        | E                        |                           | E                            | P                        | P                        | P                        |                           | P                            |
| 10                                | P                        | E                        | E                        | -                         | E                            | P                        | P                        |                          | P                         | P                            |
| 11                                | E                        | E                        | E                        | -                         | E                            | P                        | P                        | P                        | P                         | P                            |
| 12                                | P                        | P                        | E                        | -                         | P                            | P                        | P                        | P                        | P                         | P                            |
| 13                                | E                        | P                        | E                        | -                         | E                            | P                        | P                        | P                        | P                         | P                            |
| 14                                | P                        | P                        | E                        | -                         | P                            | P                        | E                        | P                        | P                         | P                            |
| 15                                | E                        | P                        | E                        | -                         | E                            | P                        | E                        | P                        | P                         | P                            |
| 16                                | E                        | E                        | E                        | -                         | E                            | P                        | P                        | P                        | P                         | P                            |
| 17                                | E                        | E                        | P                        | -                         | E                            | P                        | P                        | P                        | P                         | P                            |
| 18                                | P                        | E                        | P                        | -                         | P                            | P                        | P                        | P                        | P                         | P                            |
| 19                                | E                        | E                        | P                        | -                         | E                            | P                        | P                        | P                        | P                         | P                            |
| 20                                | E                        | E                        | E                        | P                         | E                            | P                        | P                        | P                        | P                         | P                            |
| 21                                | E                        | E                        | E                        | E                         | E                            | P                        | P                        | P                        | P                         | P                            |
| 22                                | E                        | E                        | E                        | E                         | E                            | P                        | P                        | P                        | P                         | P                            |
| 23                                | E                        | E                        | E                        | E                         | E                            | P                        | P                        | P                        | P                         | P                            |
| 24                                | E                        | P                        | E                        | E                         | E                            | P                        | P                        | P                        | P                         | P                            |
| 25                                | E                        | E                        | E                        | E                         | E                            | P                        | P                        | P                        | P                         | P                            |
| 26                                | P                        | P                        | E                        | E                         | E                            | P                        | P                        | P                        | P                         | P                            |
| 27                                | P                        | E                        | E                        | E                         | E                            | P                        | P                        | P                        | P                         | P                            |
| 28                                | E                        | E                        | E                        | E                         | E                            | P                        | P                        | P                        | P                         | P                            |
| Majority Vote in Windows          | E                        | E                        | E                        | E                         | E                            | P                        | P                        | P                        | P                         | P                            |
| Accuracy Stability Measure        | 0.38                     | 0.52                     | 0.71                     | 0.78                      | 0.71                         | 0.93                     | 0.76                     | 1                        | 1                         | 0.93                         |

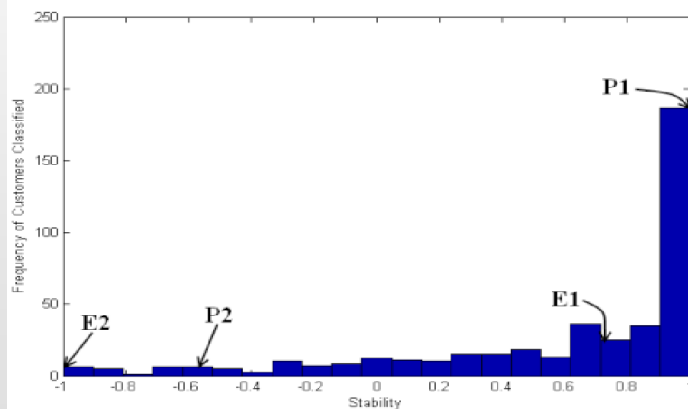
# Result & Analysis: Effect of Window Size on Classification Confidence



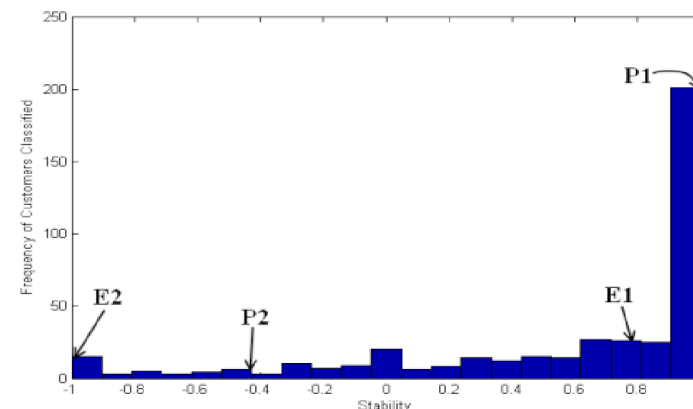
Three (3) Months Sliding Window



Six (6) Months Sliding Window

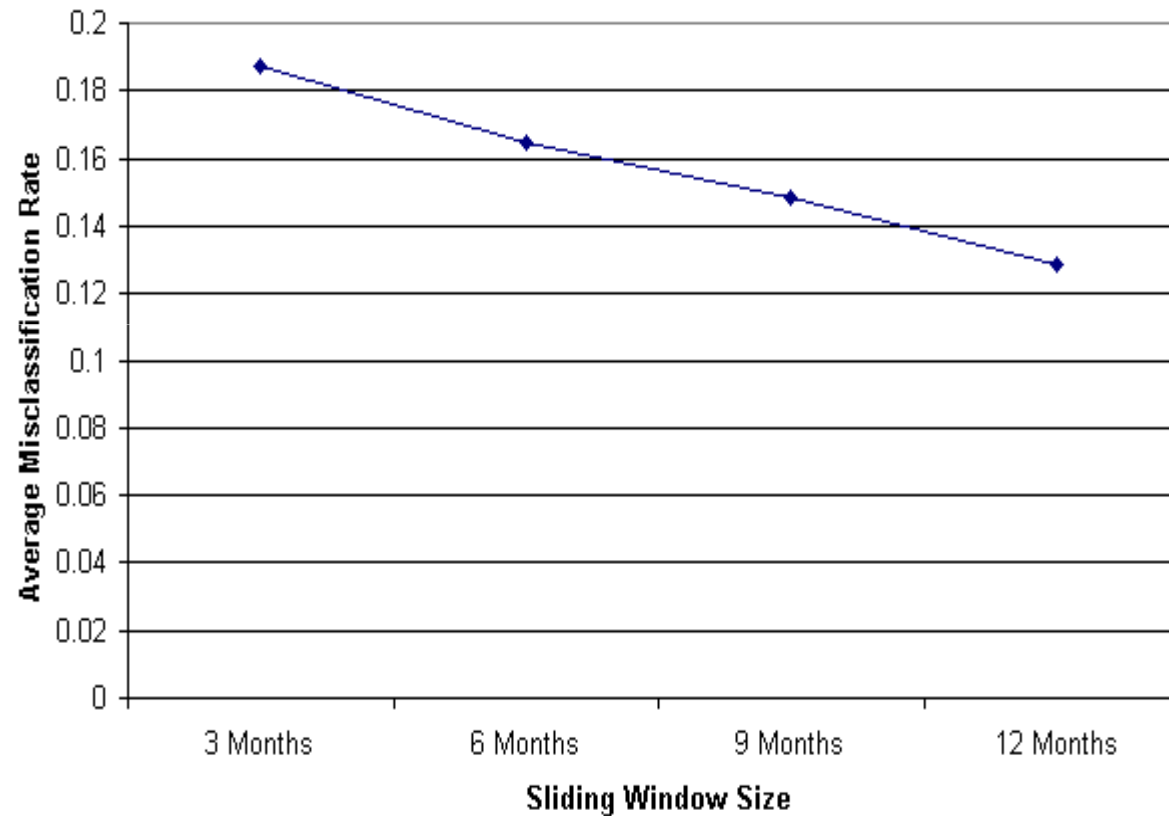


Nine (9) Months Sliding Window



Twelve (12) Months Sliding Window

# Results & Analysis: Effect of Window Size on Prediction Accuracy



## Conclusion and Future Work

- ❑ Our approach to change mining can be used to monitor and analyze the prediction accuracy and stability of a classifier over time for the purpose of:
  - ⇒ identifying the change in buying behaviour;
  - ⇒ verifying customer label;
  - ⇒ monitoring classifier performance over time.
- ❑ Future work focusing on a major challenge in case of seemingly mis-classified samples:
  - ⇒ Re-label an instance or adapt a classifier? How to tell a difference between the two scenarios?



# Q & A

