Tracking Recurring Concepts with Meta-Learners

J. Gama[1] P. Kosina[2]

[1]LIAAD-INESC Porto, FEP-University of Porto

[2]Faculty of Informatics, Masaryk University, Brno

Fourteenth Portuguese Conference on Artificial Intelligence

Contents

Introduction

- Meta-learning
  - Information about relation between tasks/domains and learning strategies
  - Finding proper model
- Data streams
  - Real world problems
  - Continuous data
- Concept drift
  - Change over time
- Recurrent concepts
  - Seasonal change

Drift Detection

- Distribution of data is stationary
  - Error-rate decreases with increasing number of examples
- Error-rate increases - warning/drift is reported
  - warning

$$p_i + s_i \geq p_{min} + 2 * s_{min}$$

  - drift

$$p_i + s_i \geq p_{min} + 3 * s_{min}$$

  - where p_i is the error-rate and s_i is standard deviation

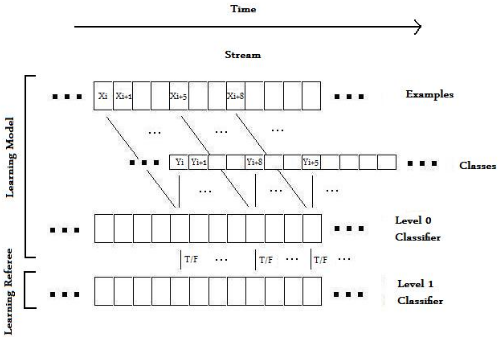
Motivation

- Presence of delay
  - Between arrival of example and obtaining label
  - Unlabeled items are usualy unused
- Could we use just attributes to predict change
  - Referees

Referee

- What is a referee
  - A meta-learning model (level 1 classifier)
  - Makes decisions about performance of primary (level 0) classifier
- How it learns
  - Examples with new class labels
    - *false* when level 0 prediction is incorrect
    - *true* when level 0 prediction is correct

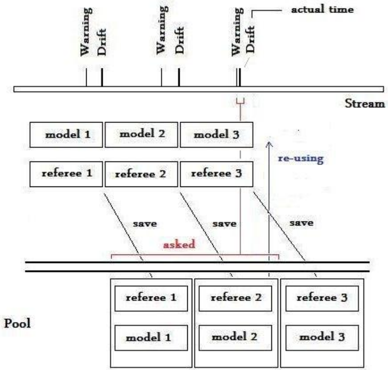
Overview of Learning the Referee



Method Strategy

- One referee for one concept model
- Before concept drift - ask referees
  - After warning level is reached
  - Proactive approach
    - Select historical (in advance) - does not need class label
    - or continue and learn new one
- After concept drift store old model with referee

Overview of Strategy

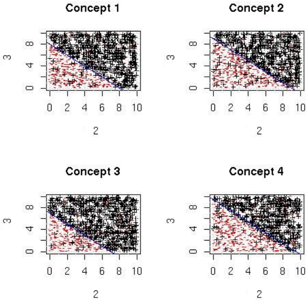


Problems

- Distribution of referee's examples = error-rate of level 0 classifier
  - Skewness of data
- Classes were not very discriminative
  - Mean of attributes
- Better to start new classifier that use wrong one

Evaluation - data

- SEA Concepts
  - Frequently used benchmark dataset with concept drift
  - 3 attributes $\rightarrow$ 2 relevant (sum $>$ threshold)
  - 4 different concepts (thresholds) repeated twice
  - 120,000 examples



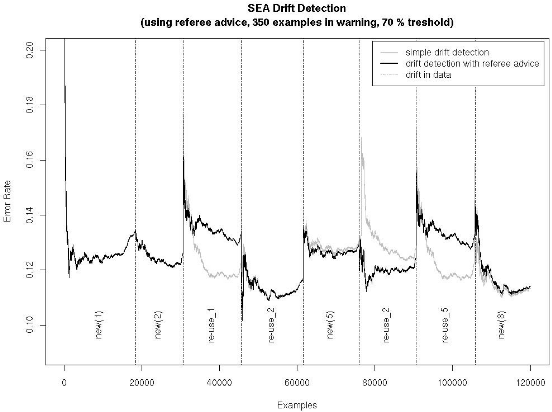
Evaluation - data

- Hyperplane
  - Represents continuously moving hyperplane in d-dimensional space
  - Recurrence?
- LED data
- Proteins
- STAGGER
- Intrusion

Evaluation Hypothesis

- After drift detection a new model always takes place
- Referees are asked and older model could be re-used
- Models itself are asked and older model could be re-used

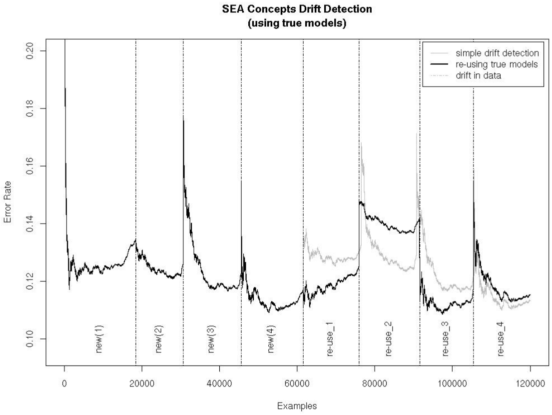
Evaluation - referees



Evaluation - notes

- Re-used models were from similar concepts → difference in error was not very significant
- Detection was faster
  - 4 times re-used and 3 times drift was sooner (183.5 examples on average)
  - Considering all the warning phases, number of examples in them was decreased by 80 on average (9.25 %)

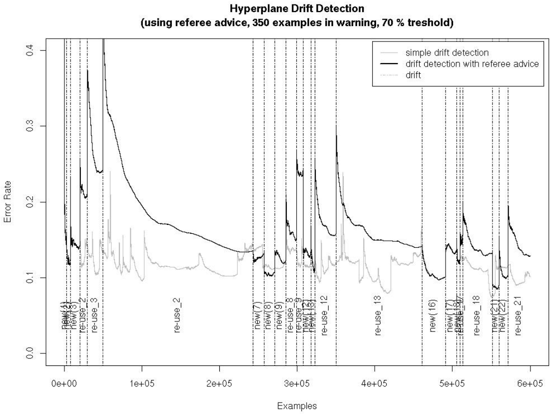
Evaluation - true models



Evaluation - notes

- Manually re-used models
  - Slower increase of error-rate
  - Early warning $\longrightarrow$ learning from examples of previous concept
  - Drift times were not better than with referee (except the last one)

Evaluation - Hyperplane



Conclusions

- It is not easy task to estimate performance without class labels - unusable for certain types of data
- We worked with only one classifier, ensemble could improve performance
- Pros
  - Can detect change faster
  - Can improve accuracy
- Cons
  - Wrong decision can lead to considerable decrease in accuracy

Thank you for your attention!