# Hierarchical Concept Drift Detection and Adaptation

**Zubin Abraham**
Robert Bosch LLC
zubin.abraham@us.bosch.com

**Shujian Yu**
University of Florida, FL
yusjlcy9011@ufl.edu

## Abstract

A change in the relationship of the response and predictor variables results in the deterioration of the predictive performance of a classifier. This paper presents Hierarchical Linear Four Rates (HLFR), that detects these concept drifts and adapts the base clasifier (soft-margin SVM) to the new concept. The results show that HLFR significantly outperforms benchmark approaches in terms of recall, accuracy and delay in detection and adaptation to concept drifts.

## 1 Introduction

For a binary classification problem, concept drift is said to occur when the joint distribution $P(\mathbf{X}_t, y_t)$ changes over time, where $\mathbf{X}_t \in \mathbb{R}^d$ are the $d$ predictor variables at time step $t$ and $y_t \in \{0, 1\}$ the corresponding binary response variable. Drift Detection Method (DDM), Drift Detection Method for Online Class Imbalance (DDM-OCI), Early Drift Detection Method (EDDM), and PerfSim are popular concept drift detection algorithms. However, they fare poorly in imbalanaced classification tasks; when concept drift occurs without affecting the recall of the minority class; or in streaming environments where decisions are made instantly. To address these limitations, we present Heirarchical Linear Four Rates (HLFR) for detecting the drift of $P(\mathbf{X}_t, y_t)$. HLFR is a two-layer hierarchical hypothesis testing framework that detects and adapts to all possible variants of concept drift, even in the presence of imbalanced class labels, as shown in Section 3.

## 2 Concept Drift Detection with Hierarchical Hypothesis Testing

HLFR is a two-layer hierarchical hypothesis testing framework for detecting and adapting to concept drift. HLFR analysis the confusion matrix streams (true positive rate (TPR), true negative rate (TNR), false positive rate (FPR) and false negative rate (FNR)) for detection. Each new received observation is analyzed by the Linear Four Rates [1] as part of the first-layer test in the HLFR. If the first-layer test identifies a potential concept drift point ($T_{potential}$) among the incoming observations, the second-layer test is trigered to assess differences in the statistical learning behavior of the streaming classification process before and after $T_{potential}$. Only if the second-layer test states that the samples generated before and after $T_{potential}$ comes from two different concepts, the detection raised by the first-layer test is confirmed. Else the first-layer test is reconfigured to improve performance. A soft-margin SVM is used as the baseline classifier due to its universality and stability. HLFR identifies the change in $P(\hat{f}(\mathbf{X}_t), y_t)$, where $\hat{f}$ is the classifier used for prediction. This is motivated by the fact that any drift of $P(\hat{f}(\mathbf{X}_t), y_t)$ would imply a drift in $P(\mathbf{X}_t, y_t)$, with probability 1.

Given the efficacy of the $P_\star$ (where, $\star \in \{TPR, TNR, PPV, NPV\}$) to detect concept drift, the first-layer test uses estimators of the rates in $P_\star$ as test statistics to conduct statistical hypothesis testing at each time step. At each time step $t$, statistical tests with the following null and alternative hypotheses is conducted. $[H_0 : \forall\star, P(\text{estimator of } P_\star^{(t-1)}) = P(\text{estimator of } P_\star^{(t)})]$, $[H_A : \exists\star, P(\text{estimator of } P_\star^{(t-1)}) \neq P(\text{estimator of } P_\star^{(t)})]$. The concept is stable under $H_0$ and is

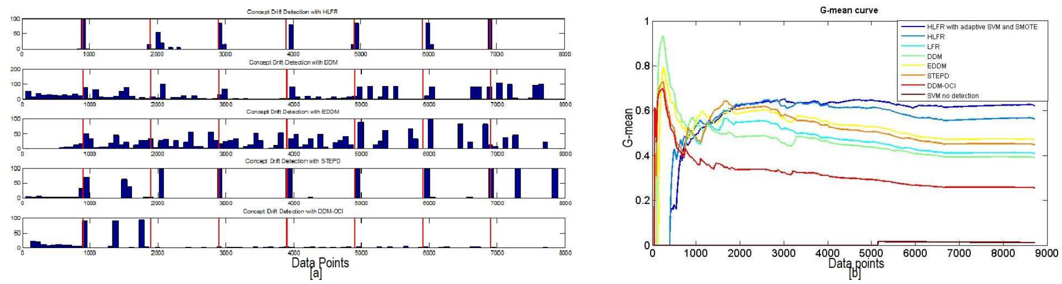


Figure 1: HLFR demonstrating superior performance compared to baseline approaches in terms of $Precision$ and $Recall$ (sub-figure [a]) and $G - mean$ (sub-figure [b]). The 'red' vertical lines in [a] corresponds to the observed point where concept drift occurs. The blue columns correspond to the probability of the model detecting a concept drift. As seen HLFR has the highest probability of detecting any observed concept drift with the least false positives and delay.

considered to have drifted if $H_0$ is rejected. The second-level test validates the detections raised by the first-layer through permutation tests. Permutation tests are theoretically well founded and does not require apriori information about the monitored process or nature of the drift. The second layer test reduces the false positives of the first-layer test by providing a feedback mechanism to adjuct the model parameters used for detetion to adapt to the incoming stream of data. Once the concept drift is confirmed, Adaptive SVM is used to update the soft-margin SVM[2]

## 3 Experiments

We compare the performance of HLFR with state-of-the art baseline approaches such as DDM, EDDM, DDM-OCI and STEPD [1,2]. We use both popular concept drift synthetic datasets (USENET1, Checkerboard, SEA, HYPERPLANE) as well as real data (SPAM), controlling drift points and allowing precise performance analysis [1]. The warning and detection thresholds of DDM (EDDM) was set to $\alpha$=3 ($\alpha$=0.95) and $\beta$=2 ($\beta$=0.90), the warning and detection significant levels of STEPD to $w$=0.05, $d$=0.01, and the parameters of DDM-OCI to $\alpha$=3 and $\beta$=2 .

A True Positive (TP) detection is defined as a detection within a fixed delay range after the precise concept change time. The detection quality is then measured both by the Recall and Precision of the detector. In the following synthetic classification tasks the base algorithm was soft margin SVM with linear kernel, each stream was independently repeated 100 times, and P=1000 reshuffling splits were used in HLFR. In all the datasets, as shown in Figure 1, HLFR outperfomed the various baseline approaches in terms of Recall, Precision, F-measure, Accuracy, F-value, G-mean, Kappa and Kappa-plus statistics.

## 4 Conclusion

The proposed HLFR framework detects the occurance of a concept drift and adapts the base classifier to the new concept and significantly outperforms existing benchmark approaches in terms of early detection of concept drifts, high detection rate and low false alarm rate across the types of concept drifts. The second layer test reduces the false positives and provides a feedback mechanism to adjuct the model parameters used for detetion to adapt to the incoming stream of data.

## 5 References

[1] Wang, H., & Abraham, Z. (2015, July). Concept drift detection for streaming data. In *2015 International Joint Conference on Neural Networks (IJCNN)* (pp. 1-9). IEEE.

[2] Yang, Jun, Rong Yan, and Alexander G. Hauptmann. Cross-domain video concept detec-tion using adaptive svms. In *Proceedings of the 15th ACM international conference on Multimedia*, pp. 188-197. ACM, 2007.