The Venn-ABERS Testing for Change-Point Detection

Ilia Nouretdinov I.R.NOURETDINOV@RHUL.AC.UK
Alex Gammerman A.GAMMERMAN@RHUL.AC.UK
Centre for Reliable Machine Learning
Royal Holloway, University of London, Surrey, UK

Editor: Harris Papadopoulos, Khuong An Nguyen, Henrik Boström and Lars Carlsson

1. Introduction

A recurrent problem in many domains is the accurate and rapid detection of a change in the distribution of observed variables. This is important since the algorithms have been trained for a certain data distribution and if the distribution has changed, the results will not be accurate and/or valid any longer. Instances of this problem, which are generally referred to as change-point detection, are found in fault detection in vehicle control systems, detection of the onset of an epidemic and many other applications. It has been a subject of intensive research, with many publications in the statistical literature. Among well-known methods, there are CumulativeSum (CUSUM) and Shiryaev-Roberts procedures for the detection of changes.

However, many of the methods would require complete or partial knowledge of the distribution of observed variables before and after the distribution has changed. Recent work in Conformal Testing and the introduction of Conformal Test Martingales (CTM) allows us to avoid this limitation and obtained valid results without information about used distributions Vovk et al. (2022). This is done in online mode with the assumption that data are exchangeable and the corresponding martingale accumulates evidence against this assumption. This paper considers an approach to the problem using the Venn-ABERS testing. It allows us to find deviations in the distributions when IID was violated.The Venn-ABERS approach satisfies the property of validity, makes calibrated probabilistic predictions and would allow reducing the number of false alarms. In application to the change-point prediction we also use e-values instead of p-values which makes some computational savings Using e-values allows us to avoid making an additional step from the non-conformity scores to the p-values. This makes it more convenient to combine with Venn-ABERS scores. As it has been pointed out in the above reference, a large e-value can be interpreted as evidence against the IID (or exchangeability) assumption: the evidence is strong when e-value exceeds 10 and decisive when it exceeds 100.

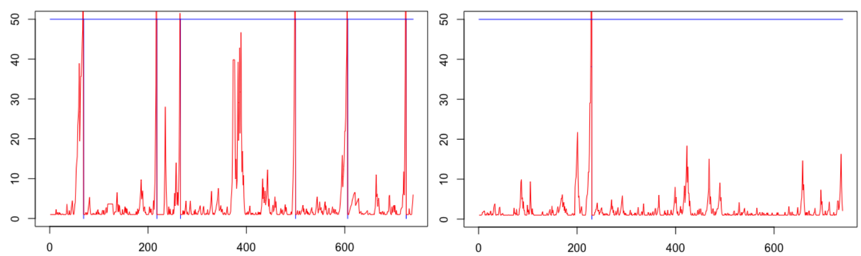
2. Results

Our idea is to run *Modified MUSUC* procedure from Vovk (2020) with the scores extracted from Venn-ABERS machine. When this is applied to a data set, the results look as a sequence of change points detected whenever the level of accumulated evidence against I.I.D. exceeds a threshold C.

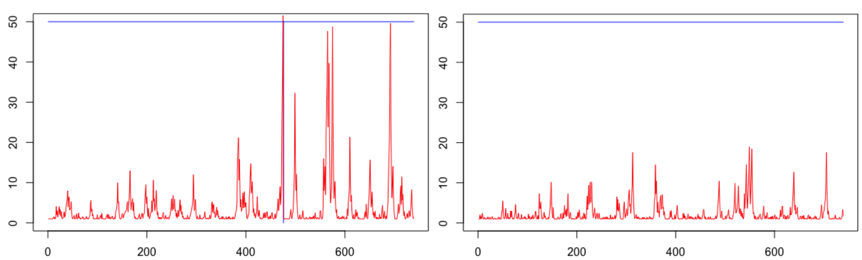
The procedure is asymmetric (one-sided), therefore two separate complementing versions are run in parallel. They refer to two kinds of IID violation, detecting whether probabilistic predictions of Venn-ABERS are over or below the background truth.

In the following examples, we use 'Absenteeism at work' from UCI repository. Each record in this database is related to a case when a person was absent from work for the whole day or for a part of it. The label (for prediction) is the number of hours missed (in a binary form using a 3.5 hours threshold). The number of other attributes is 20.

According to the following results, 7 change points are detected for $C = 50$ in total, distributed as 6+1 over different directions of violation, shown separately on the left and right sides.



As a baseline (control) experiment, we can apply the same procedure to a reshuffled (randomly re-ordered) version of the data set, so that there are no real deviations from IID, and all the detected change points are really 'false alarms'. As it can be seen from the next plots, there are just one of them in such a case.



References

Vladimir Vovk. Conformal e-prediction for change detection. Available at http://alrw.net/articles/29.pdf, 2020.

Vladimir Vovk, Alex Gammerman, and Glenn Shafer. *Algorithmic learning in a random world*. Springer Science and Business Media, 2022.