

Classifier selection for imbalanced data streams with *Minority Driven Ensemble*

Paweł Zyblewski^[0000–0002–4224–6709], Michał Woźniak^[0000–0003–0146–4205], and
Paweł Ksieniewicz^[0000–0001–9578–8395]

Department of Systems and Computer Networks,
Faculty of Electronics, Wrocław University of Science and Technology,
Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland
{pawel.zyblewski,michal.wozniak,pawel.ksieniewicz}@pwr.edu.pl

Abstract. The abstract should briefly summarize the contents of the paper in 15–250 words.

Keywords: First keyword · Second keyword · Another keyword.

1 Introduction

Klasyfikacja taka ciężka. Niezbalansowana jeszcze cięższa. A przy strumieniach to już w ogóle.

To może selekcja klasyfikatorów [1].

2 Method

3 Set-up of experiments

3.1 Datasets

3.2 Experiments

Experiment 1 — hyperparameters optimization Ustalamy typowy strumień zawierający problem o dużej skali niezbalansowania (1:9) i szumie na poziomie udziału klasy mniejszościowej (.1). Testujemy uśrednioną miarę i zależność statystyczną przebiegów dla różnych wielkości komitetów i różnej wartości parametru odcięcia klasyfikatorów o przestarzałej granicy decyzyjnej. Testujemy osobno strumień z dryfami nagłymi i inkrementalnymi.

Oczekiwania: - powiększanie komitetu początkowo stabilizuje jakość, ale z czasem degraduje zdolność do reakcji na dryf koncepcji, - zwiększanie progu odcięcia początkowo kompensuje spowolnienie reakcji na dryf, ale z czasem wpływa negatywnie na stabilność jakości. - obie wartości powinny być relatywnie niskie. - czas trwania badań to około pół godziny.

Opis generatora

Przegląd zbiorów syntetycznych

Wypadałoby chyba jednak dodać te rzeczywiste, ale zrobmy to w następnym kroku

Experiment 2 — comparative analysis of classifier selection methods

4 Experimental evaluation

Experiment 1 — hyperparameters optimization Wyniki na Figure 1. Na osi Y wielkość komitetu, na X parametr alpha.

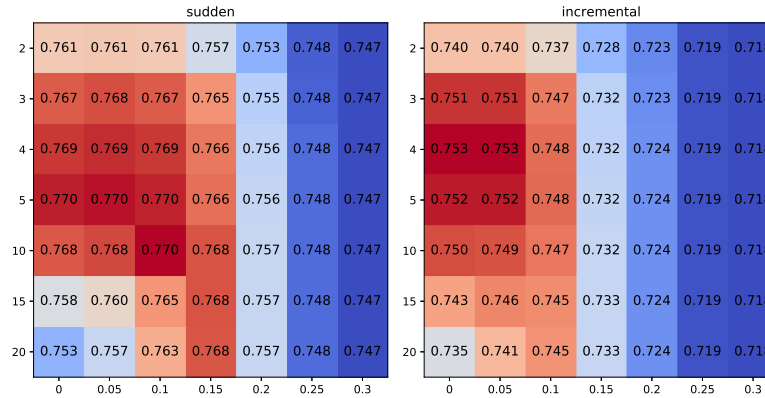


Fig. 1. Podpis

Do dalszych badań wybrano komitet trzech klasyfikatorów i parametr alpha .05.

Experiment 2 — comparative analysis of classifier selection methods

5 Conclusions

Jest super, akurat dla takich wspaniale nietypowych danych, które wcale nie są takie nietypowe.

Acknowledgement

This work was supported by the statutory funds of the Department of Systems and Computer Networks, Faculty of Electronics, Wrocław University of Science and Technology and by the Polish National Science Centre under the grant No. ???.

Numer grantu

References

1. Cruz, R.M.O., Hafemann, L.G., Sabourin, R., Cavalcanti, G.D.C.: DESlib: A Dynamic ensemble selection library in Python. arXiv preprint arXiv:1802.04967 (2018)