

# Lokalne klasyfikatory jako narzędzie analizy i klasyfikacji sygnałów

Streszczenie rozprawy doktorskiej

Wit Jakuczun

Rozprawa została poświęcona problemowi analizy i klasyfikacji *sygnałów cyfrowych* rozumianych jako ciąg liczb powstałych w wyniku pomiarów jakiegoś fizycznego zjawiska, np. sygnał EEG czy sygnał mowy. Z sygnałami cyfrowymi i problemem ich analizy badacze różnych dziedzin spotykają się coraz częściej. Dlatego też, pomimo istnienia wielu metod, które sprawdziły się w praktyce, uważam, że warto zajmować się tworzeniem nowych podejść dających nowe możliwości analizy danych.

Metoda opisana w niniejszej rozprawie należy do grupy metod, w której analizowane sygnały rozpisywane są w bazie przestrzeni liniowej i jest rozwinięciem idei *lokalnych baz dyskryminacyjnych* [11]. Opisywana metoda jest oparta na nowym podejściu konstruowania biortogonalnych baz falkowych drugiej generacji [12] i [3], będących uogólnieniem falek klasycznych [4] i [14]. Przy konstrukcji bazy korzystam ze znanego i jednego z najlepszych algorytmów konstrukcji klasyfikatorów zwanego *maszyną wektorów wspierających* (ang. *Support Vector Machines - SVM*) [13].

Konstruowana baza ma dwie cechy, które moim zdaniem wyróżniają ją na tle innych tego typu metod:

**Lokalność** Elementy bazy są lokalne w czasie. Dzięki temu mogę opisywać analizowane sygnały korzystając jedynie z częściowej informacji.

**Dyskryminacja** Elementy bazy są tak konstruowane, aby współczynniki rozwinięcia dla sygnałów z dwu różnych klas decyzyjnych były jak najbardziej różne.

W ramach współpracy z zespołem prof. Andrzeja Wróbla<sup>1</sup> proponowana przeze mnie metoda została z sukcesem zastosowana do analizy tzw. *lokalnych potencjałów wywołanych* [8]. Otrzymane wyniki okazały się zgodne z

---

<sup>1</sup>Instytut Biologii Doświadczalnej im. Marcelego Nenckiego PAN

poprzednimi wynikami uzyskanymi przez wspomniany zespół [15], [10]. Potwierdza to, że proponowana metoda może służyć jako narzędzie *eksploracji danych*, mogące pomóc w zrozumieniu dlaczego analizowane sygnały są różne.

Z punktu widzenia problemu konstrukcji klasyfikatorów metodę można zaliczyć do grupy algorytmów opartych na *lokalnych klasyfikatorach* [11], [7], [9]. Podejścia te dokonują klasyfikacji sygnałów w oparciu o lokalny opis analizowanych danych. Typowy schemat wygląda tak, że jest generowany pewien zestaw *lokalnych deskryptorów (klasyfikatorów)*, a ostateczna klasyfikacja jest dokonywana na podstawie łączenia odpowiedzi owych *lokalnych deskryptorów* poprzez głosowanie albo stosowanie dodatkowego klasyfikatora. W pracy opisuję jak wykorzystać proponowaną metodę do konstrukcji tego typu klasyfikatorów z wykorzystaniem *zespołu klasyfikatorów* [6], [1], [2], [5].

Metoda została dokładnie przetestowana na szerokim spektrum danych, tak sztucznych jak i rzeczywistych. Dane zawierały zarówno sygnały jednowymiarowe, dwuwymiarowe (np. baza USPS cyfr pisanych ręcznie) sygnały wielokanałowe ( np. dane Auslan [9] oraz EEG<sup>2</sup>). Uzyskane wyniki świadczą o tym, że metoda warta jest rozważenia dla badacza chcącego analizować lub klasyfikować dane w postaci sygnałów.

Oprócz testów praktycznych przeprowadziłem też prostą analizę metody z punktu widzenia *statystycznej teorii uczenia się* poprzez oszacowanie wymiaru Vapnika-Czerwonenkisa [13].

Wszelkie testy przeprowadziłem w środowisku Matlab, korzystając z własnoręcznie napisanych skryptów.

## Literatura

- [1] L. Breiman. Arcing classifiers. 1998.
- [2] Leo Breiman. Random forests. *Machine Learning*, 45(1):5–32, 2001.
- [3] R. Claypoole, R. Baraniuk, and R. Nowak. Adaptive wavelet transforms via lifting. 1998.
- [4] Ingrid Daubechies. *Ten Lectures on Wavelets*. SIAM, 1992.
- [5] Thomas G. Dietterich. An experimental comparison of three methods for constructing ensembles of decision trees: Bagging, boosting, and randomization. *Machine Learning*, 1999.

---

<sup>2</sup>[http://ida.first.fhg.de/projects/bci/bbci\\_official/](http://ida.first.fhg.de/projects/bci/bbci_official/)

- [6] Yoav Freund and Robert E. Schapire. Experiments with a new boosting algorithm. In *International Conference on Machine Learning*, pages 148–156, 1996.
- [7] Carlos Alonso Gonzalez and Juan J. Rodriguez Diez. Time series classification by boosting interval based literals. *Inteligencia Artificial, Revista Iberoamericana de Inteligencia Artificial*, 11:2–11, 2000.
- [8] W. Jakuczun, A. Wróbel, D. Wójcik, and E. Kublik. Classifying evoked potentials with local classifiers. *Acta Neurobiologiae Experimentalis*, 2005.
- [9] Mohammed Waleed Kadous. *Temporal Classification: Extending the Classification Paradigm to Multivariate Time Series*. PhD thesis, School of Computer Science & Engineering, University of New South Wales, 2002.
- [10] E. Kublik, P. Musiał, and A. Wróbel. Identification of principal components in cortical evoked potentials by brief surface cooling. *Clinical Neurophysiology*, 2001.
- [11] Naoki Saito. *Local Feature Extraction and Its Application Using a Library of Bases*. PhD thesis, Yale University, 1994.
- [12] Wim Sweldens. The lifting scheme: A construction of second generation wavelets. *SIAM Journal on Mathematical Analysis*, 29(2):511–546, 1998.
- [13] Vladimir Vapnik. *Statistical Learning Theory*. John Wiley & Sons, 1998.
- [14] Przemysław Wojtaszczyk. *Teoria falek*. Wydawnictwo Naukowe PWN SA, 2000.
- [15] M. Wypych, E. Kublik, P. Wojdyło, and A. Wróbel. Sorting functional classes of evoked potentials by wavelets. *Neuroinformatic*, 2003.