

Politechnika Poznańska
Wydział Informatyki i Zarządzania
Instytut Informatyki

Praca dyplomowa magisterska

**OPTYMALIZACJA KLASYFIKATORA SVM ZA POMOCĄ PROGRAMOWANIA
GENETYCZNEGO**

Tomasz Ziętkiewicz

Promotor
dr hab. Krzysztof Krawiec

Poznań, 2012

Tutaj przychodzi karta pracy dyplomowej;
oryginał wstawiamy do wersji dla archiwum PP, w pozostałych kopiach wstawiamy ksero.

Spis treści

1	Wprowadzenie	1
1.1	Cel i zakres pracy	1
1.2	Struktura pracy	1
2	Podstawy teoretyczne	2
2.1	SVM	2
2.2	Programowanie genetyczne	2
2.3	Ewolucja kerneli	2
2.4	Obrazowanie mózgu przy pomocy fMRI	3
2.5	ADHD	3
3	Opis implementacji algorytmu Kernel GP	4
3.1	Opis rozszerzeń algorytmu	4
3.1.1	Lokalna optymalizacja stałych	4
3.2	Złożoność obliczeniowa	4
4	Wyniki działania algorytmu na popularnych zbiorach danych	6
4.1	Dokładność	6
5	Case study - klasyfikacja danych ADHD 200	8
5.1	Opis zbioru danych	8
5.1.1	Surowe dane	8
5.1.2	Preprocessing	8
5.2	Konstrukcja i selekcja cech	8
5.3	Wyniki klasyfikacji	8
5.3.1	Kernel GP	8
5.3.2	Porównanie z innymi algorytmami	8
6	Podsumowanie i kierunki dalszego rozwoju	9
	Zasoby internetowe	10

Rozdział 1

Wprowadzenie

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Nulla purus purus, fermentum in condimentum nec, sodales nec enim. Fusce auctor auctor porta. Proin tempus lacinia tortor, eget aliquam ante condimentum id. Morbi viverra congue posuere. Nunc non odio eros, sollicitudin pulvinar metus. Sed eget ligula ligula, a congue orci. Proin laoreet aliquet vulputate. Vivamus ut enim sed diam pretium fringilla. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos himenaeos. Etiam convallis mi nec dolor pretium sed aliquam orci sagittis. Maecenas aliquam dictum neque vel mollis. Morbi vel vehicula mauris.

1.1 Cel i zakres pracy

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Nulla purus purus, fermentum in condimentum nec, sodales nec enim. Fusce auctor auctor porta. Proin tempus lacinia tortor, eget aliquam ante condimentum id. Morbi viverra congue posuere. Nunc non odio eros, sollicitudin pulvinar metus. Sed eget ligula ligula, a congue orci. Proin laoreet aliquet vulputate. Vivamus ut enim sed diam pretium fringilla. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos himenaeos. Etiam convallis mi nec dolor pretium sed aliquam orci sagittis. Maecenas aliquam dictum neque vel mollis. Morbi vel vehicula mauris.

1.2 Struktura pracy

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Nulla purus purus, fermentum in condimentum nec, sodales nec enim. Fusce auctor auctor porta. Proin tempus lacinia tortor, eget aliquam ante condimentum id. Morbi viverra congue posuere. Nunc non odio eros, sollicitudin pulvinar metus. Sed eget ligula ligula, a congue orci. Proin laoreet aliquet vulputate. Vivamus ut enim sed diam pretium fringilla. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos himenaeos. Etiam convallis mi nec dolor pretium sed aliquam orci sagittis. Maecenas aliquam dictum neque vel mollis. Morbi vel vehicula mauris.

Rozdział 2

Podstawy teoretyczne

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Nulla purus purus, fermentum in condimentum nec, sodales nec enim. Fusce auctor auctor porta. Proin tempus lacinia tortor, eget aliquam ante condimentum id. Morbi viverra congue posuere. Nunc non odio eros, sollicitudin pulvinar metus. Sed eget ligula ligula, a congue orci. Proin laoreet aliquet vulputate. Vivamus ut enim sed diam pretium fringilla. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos himenaeos. Etiam convallis mi nec dolor pretium sed aliquam orci sagittis. Maecenas aliquam dictum neque vel mollis. Morbi vel vehicula mauris.

2.1 SVM

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Nulla purus purus, fermentum in condimentum nec, sodales nec enim. Fusce auctor auctor porta. Proin tempus lacinia tortor, eget aliquam ante condimentum id. Morbi viverra congue posuere. Nunc non odio eros, sollicitudin pulvinar metus. Sed eget ligula ligula, a congue orci. Proin laoreet aliquet vulputate. Vivamus ut enim sed diam pretium fringilla. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos himenaeos. Etiam convallis mi nec dolor pretium sed aliquam orci sagittis. Maecenas aliquam dictum neque vel mollis. Morbi vel vehicula mauris.

SVM (ang. *Support Vector Machine*) [] jest ...

2.2 Programowanie genetyczne

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Nulla purus purus, fermentum in condimentum nec, sodales nec enim. Fusce auctor auctor porta. Proin tempus lacinia tortor, eget aliquam ante condimentum id. Morbi viverra congue posuere. Nunc non odio eros, sollicitudin pulvinar metus. Sed eget ligula ligula, a congue orci. Proin laoreet aliquet vulputate. Vivamus ut enim sed diam pretium fringilla. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos himenaeos. Etiam convallis mi nec dolor pretium sed aliquam orci sagittis. Maecenas aliquam dictum neque vel mollis. Morbi vel vehicula mauris.

2.3 Ewolucja kerneli

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Nulla purus purus, fermentum in condimentum nec, sodales nec enim. Fusce auctor auctor porta. Proin tempus lacinia tortor, eget aliquam ante condimentum id. Morbi viverra congue posuere. Nunc non odio eros, sollicitudin pulvinar metus. Sed eget ligula ligula, a congue orci. Proin laoreet aliquet vulputate. Vivamus ut enim sed diam

pretium fringilla. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos himenaeos. Etiam convallis mi nec dolor pretium sed aliquam orci sagittis. Maecenas aliquam dictum neque vel mollis. Morbi vel vehicula mauris.

2.4 Obrazowanie mózgu przy pomocy fMRI

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Nulla purus purus, fermentum in condimentum nec, sodales nec enim. Fusce auctor auctor porta. Proin tempus lacinia tortor, eget aliquam ante condimentum id. Morbi viverra congue posuere. Nunc non odio eros, sollicitudin pulvinar metus. Sed eget ligula ligula, a congue orci. Proin laoreet aliquet vulputate. Vivamus ut enim sed diam pretium fringilla. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos himenaeos. Etiam convallis mi nec dolor pretium sed aliquam orci sagittis. Maecenas aliquam dictum neque vel mollis. Morbi vel vehicula mauris.

2.5 ADHD

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Nulla purus purus, fermentum in condimentum nec, sodales nec enim. Fusce auctor auctor porta. Proin tempus lacinia tortor, eget aliquam ante condimentum id. Morbi viverra congue posuere. Nunc non odio eros, sollicitudin pulvinar metus. Sed eget ligula ligula, a congue orci. Proin laoreet aliquet vulputate. Vivamus ut enim sed diam pretium fringilla. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos himenaeos. Etiam convallis mi nec dolor pretium sed aliquam orci sagittis. Maecenas aliquam dictum neque vel mollis. Morbi vel vehicula mauris.

Rozdział 3

Opis implementacji algorytmu Kernel GP

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Nulla purus purus, fermentum in condimentum nec, sodales nec enim. Fusce auctor auctor porta. Proin tempus lacinia tortor, eget aliquam ante condimentum id. Morbi viverra congue posuere. Nunc non odio eros, sollicitudin pulvinar metus. Sed eget ligula ligula, a congue orci. Proin laoreet aliquet vulputate. Vivamus ut enim sed diam pretium fringilla. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos himenaeos. Etiam convallis mi nec dolor pretium sed aliquam orci sagittis. Maecenas aliquam dictum neque vel mollis. Morbi vel vehicula mauris.

3.1 Opis rozszerzeń algorytmu

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Nulla purus purus, fermentum in condimentum nec, sodales nec enim. Fusce auctor auctor porta. Proin tempus lacinia tortor, eget aliquam ante condimentum id. Morbi viverra congue posuere. Nunc non odio eros, sollicitudin pulvinar metus. Sed eget ligula ligula, a congue orci. Proin laoreet aliquet vulputate. Vivamus ut enim sed diam pretium fringilla. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos himenaeos. Etiam convallis mi nec dolor pretium sed aliquam orci sagittis. Maecenas aliquam dictum neque vel mollis. Morbi vel vehicula mauris.

3.1.1 Lokalna optymalizacja stałych

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Nulla purus purus, fermentum in condimentum nec, sodales nec enim. Fusce auctor auctor porta. Proin tempus lacinia tortor, eget aliquam ante condimentum id. Morbi viverra congue posuere. Nunc non odio eros, sollicitudin pulvinar metus. Sed eget ligula ligula, a congue orci. Proin laoreet aliquet vulputate. Vivamus ut enim sed diam pretium fringilla. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos himenaeos. Etiam convallis mi nec dolor pretium sed aliquam orci sagittis. Maecenas aliquam dictum neque vel mollis. Morbi vel vehicula mauris.

3.2 Złożoność obliczeniowa

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Nulla purus purus, fermentum in condimentum nec, sodales nec enim. Fusce auctor auctor porta. Proin tempus lacinia tortor, eget aliquam ante condimentum id. Morbi viverra congue posuere. Nunc non odio eros, sollicitudin pulvinar metus. Sed eget ligula ligula, a congue orci. Proin laoreet aliquet vulputate. Vivamus ut enim sed diam pretium fringilla. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos hi-

menaeos. Etiam convallis mi nec dolor pretium sed aliquam orci sagittis. Maecenas aliquam dictum neque vel mollis. Morbi vel vehicula mauris.

Rozdział 4

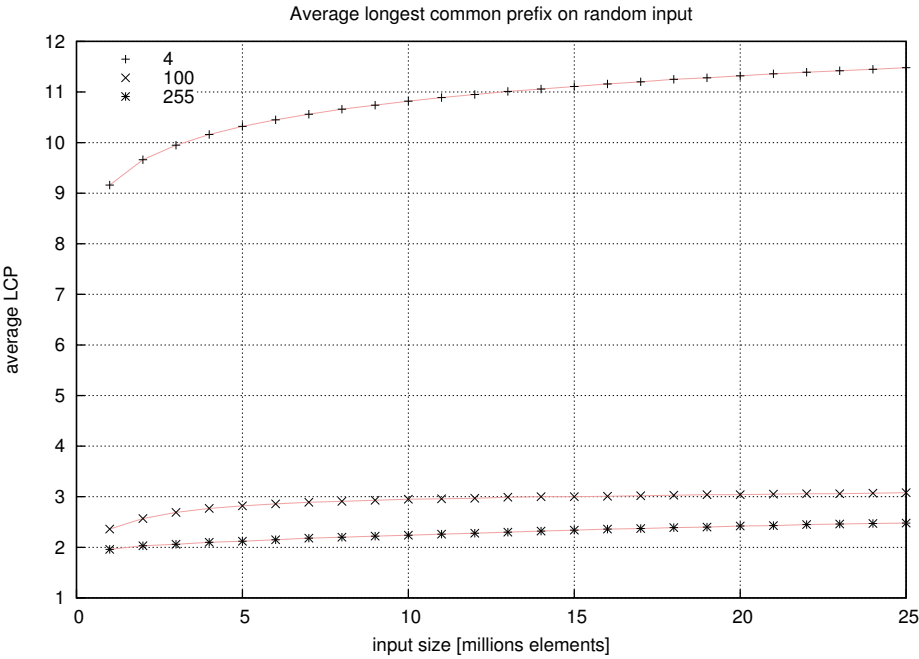
Wyniki działania algorytmu na popularnych zbiorach danych

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Nulla purus purus, fermentum in condimentum nec, sodales nec enim. Fusce auctor auctor porta. Proin tempus lacinia tortor, eget aliquam ante condimentum id. Morbi viverra congue posuere. Nunc non odio eros, sollicitudin pulvinar metus. Sed eget ligula ligula, a congue orci. Proin laoreet aliquet vulputate. Vivamus ut enim sed diam pretium fringilla. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos himenaeos. Etiam convallis mi nec dolor pretium sed aliquam orci sagittis. Maecenas aliquam dictum neque vel mollis. Morbi vel vehicula mauris.

4.1 Dokładność

<i>j</i>	<i>SA[j] = i</i>	<i>x[i..n - 1]</i>
0	5	\$
1	4	o\$
2	1	orro\$
3	3	ro\$
4	2	rro\$
5	0	zorro\$

RYSUNEK 4.1: Tablica sufiksów dla sufiksów ciągu „zorro\$”.



RYSUNEK 4.2: Opis wykresu.

Rozdział 5

Case study - klasyfikacja danych ADHD 200

5.1 Opis zbioru danych

5.1.1 Surowe dane

5.1.2 Preprocessing

5.2 Konstrukcja i selekcja cech

5.3 Wyniki klasyfikacji

5.3.1 Kernel GP

5.3.2 Porównanie z innymi algorytmami

SVM

Inne klasyfikatory

Rozdział 6

Podsumowanie

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Nulla purus purus, fermentum in condimentum nec, sodales nec enim. Fusce auctor auctor porta. Proin tempus lacinia tortor, eget aliquam ante condimentum id. Morbi viverra congue posuere. Nunc non odio eros, sollicitudin pulvinar metus. Sed eget ligula ligula, a congue orci. Proin laoreet aliquet vulputate. Vivamus ut enim sed diam pretium fringilla. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos himenaeos. Etiam convallis mi nec dolor pretium sed aliquam orci sagittis. Maecenas aliquam dictum neque vel mollis. Morbi vel vehicula mauris.

Zasoby internetowe

- [A] Project Gutenberg.
<http://www.gutenberg.net>
- [B] The AspectJ Project.
<http://www.eclipse.org/aspectj/>
- [C] The Gauntlet (Universal Robustness Corpus).
<http://www.michael-maniscalco.com/testset/gauntlet/>
- [D] Manzini's Large Corpus.
<http://www.mfn.unipmn.it/~manzini/lightweight/corpus/>
- [E] Yuta Mori, Suffix Array Construction Benchmark
<http://homepage3.nifty.com/wpage/benchmark/index.html>
- [F] Michael Maniscalco, The MSufSort Algorithm.
<http://www.michael-maniscalco.com/msufsort.htm>
- [G] Peter Sanders, Skew algorithm.
<http://www.mpi-inf.mpg.de/~sanders/programs/suffix/>
- [H] M. Douglas McIlroy, ssort.c
<http://cm.bell-labs.com/cm/who/doug/source.html>
- [I] Dmitry A. Malyshev, Archon
<http://kvgate.com/index.php?root/comp/arch/archon/>
- [J] Giovanni Manzini, A Lightweight Suffix Array and BWT Construction Algorithm
<http://web.unipmn.it/~manzini/lightweight/ds.tgz>
- [K] Yuta Mori, libdivsufsort project homepage.
<http://code.google.com/p/libdivsufsort/>
- [L] N.Jesper Larsson, qsufsort.c
<http://www.larsson.dogma.net/qsufsort.c>
- [M] Klaus-Bernd Schürmann i Jens Stoye, bpr downloadpage.
<http://bibiserv.techfak.uni-bielefeld.de/download/tools/bpr.html>



© 2012 Tomasz Ziętkiewicz

Instytut Informatyki, Wydział Informatyki i Zarządzania
Politechnika Poznańska

Skład przy użyciu systemu \LaTeX .

Bib \TeX :

```
@mastersthesis{ mnowak-masterthesis,  
  author = "Tomasz Ziętkiewicz",  
  title = "{Optymalizacja klasyfikatora SVM za pomocą programowania genetycznego}",  
  school = "Poznan University of Technology",  
  address = "Pozna{\n}, Poland",  
  year = "2012",  
}
```