



Wykorzystanie metody SVM i PCA do przetwarzania danych pomiarowych w spektroskopii ramanowskiej

Adam Władziński

Krew



Dlaczego krew?



Czym jest krew?

Krew jest rodzajem tkanki łącznej o charakterystycznych właściwościach.

Jest mieszaniną kilku płynów, często opisywaną w literaturze jako emulsja.

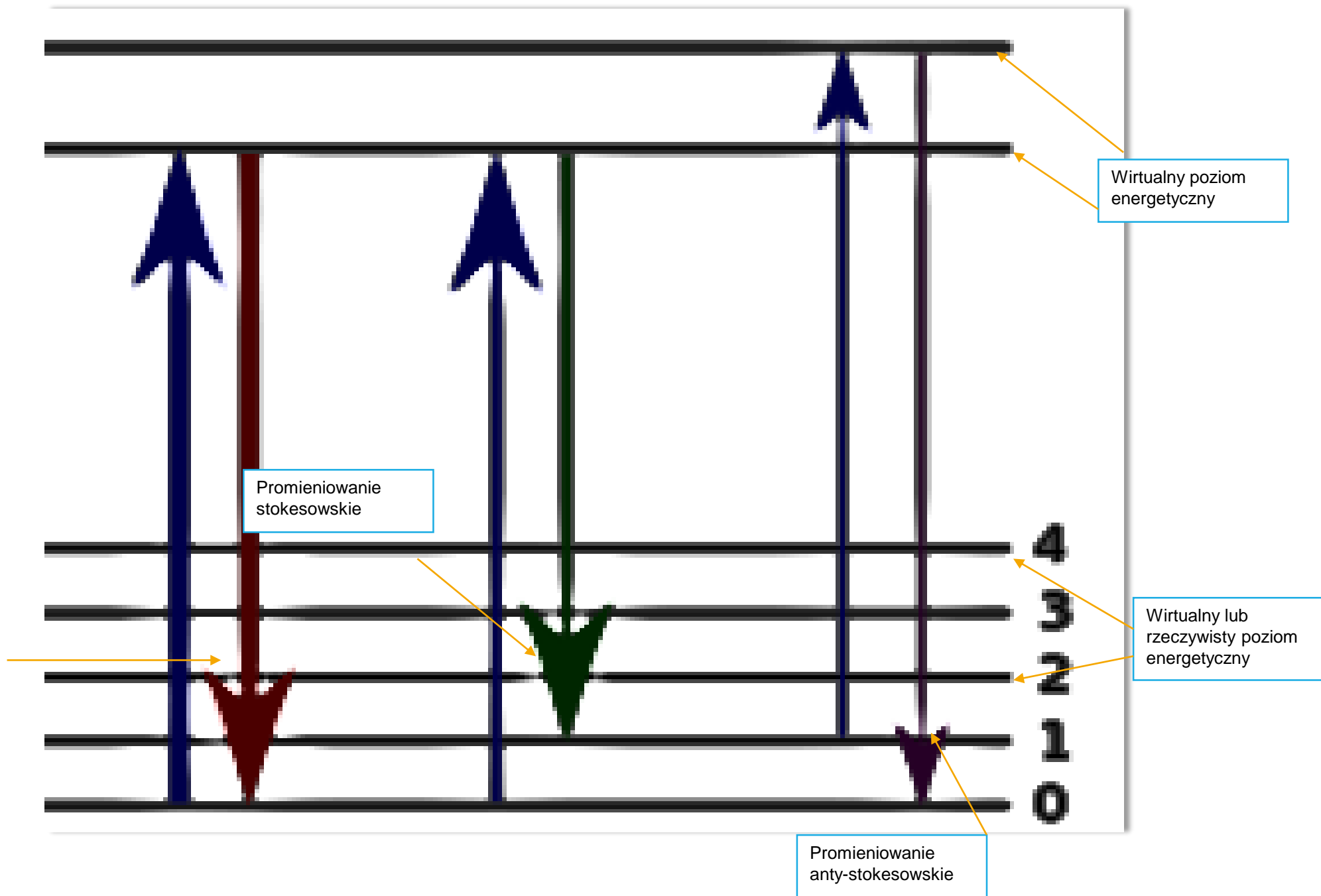
Jest zamknięta w organizmach żywych i praktycznie odizolowana od środowiska zewnętrznego.

Promieniowanie Ramana



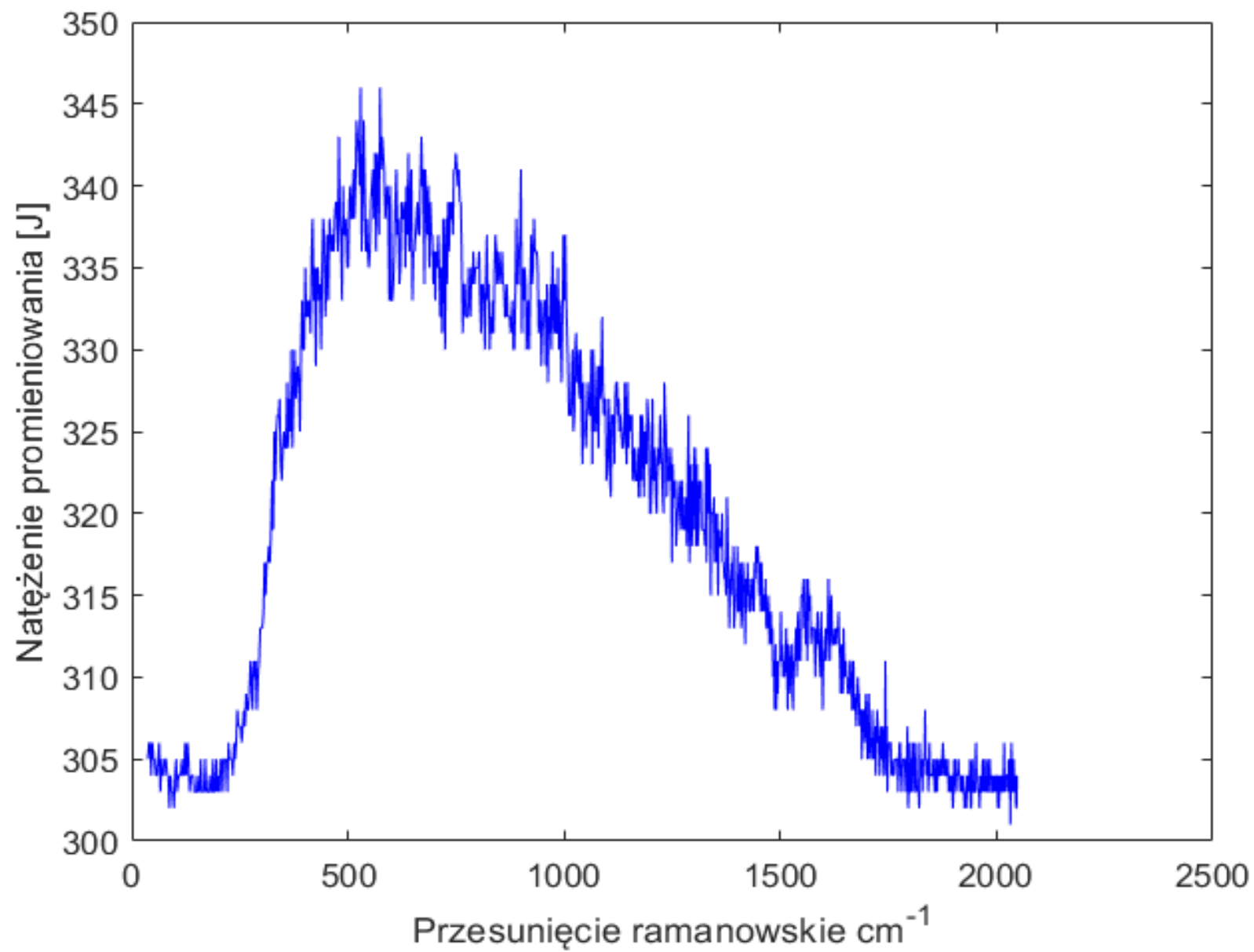
Promieniowanie ramanowskie

- Promieniowanie jest rozproszonym nieelastycznie promieniowaniem monochromatycznym.
- Rozpraszanie następuje, w wyniku oddziaływania pomiędzy dipolami, indukowanymi w drgającym układzie atomowym, a padającą wiązką.
- W następstwie tego zjawiska, materia zostaje wzbudzona do wyższego poziomu energetycznego, który może być wirtualny lub rzeczywisty.
- W wyniku powrotu do stanu równowagi, czyli do stanu sprzed wzbudzenia, następuje emisja fali o takiej samej długości λ , jak promieniowania wzbudzającego. Samo to zjawisko rozpraszania nosi nazwę rozpraszania Rayleigha.



Dlaczego to promieniowanie Ramana?

- Różnica częstości drgań odpowiada, drganiom własnym układów atomowych , oraz drganiom wiązań pomiędzy nimi.
- Jest to cecha unikatowa dla każdego wiązania i atomu.
- Drgania atomowe mają swoje źródło w stanie fizycznym danej substancji.

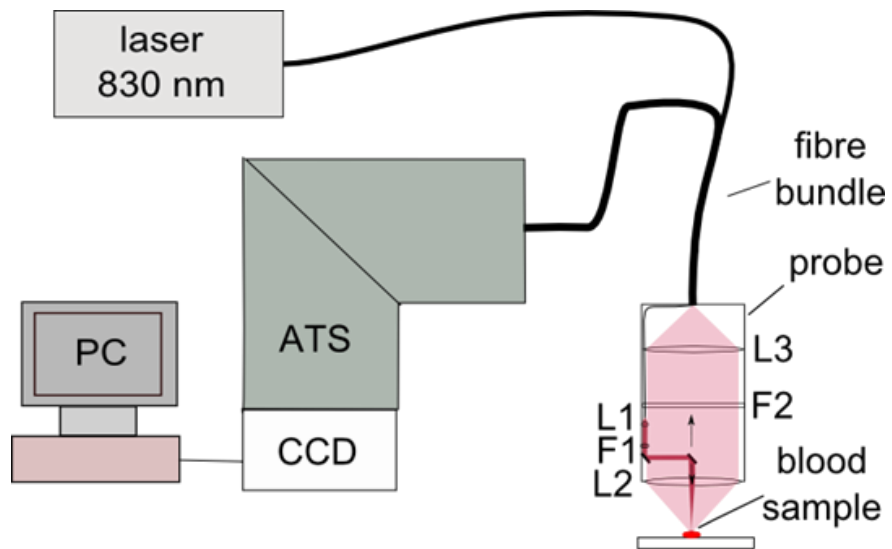


Widmo ramanowskie Alkoholu

Pomiar



Układ pomiarowy



L1,L2, L3 - lenses;
F1- laser line filter
F2- longpass filter;
ATS- axial transmissive
spectrograph

- Laser
- Specjalna zamknięta przestrzeń z detektorem
- Stanowisko komputerowe.





ANDOR

TECHNOLOGY



Model No.:

DV401-BV

AM

Serial No.:

CCD-4676

Date:

JUNE 2004

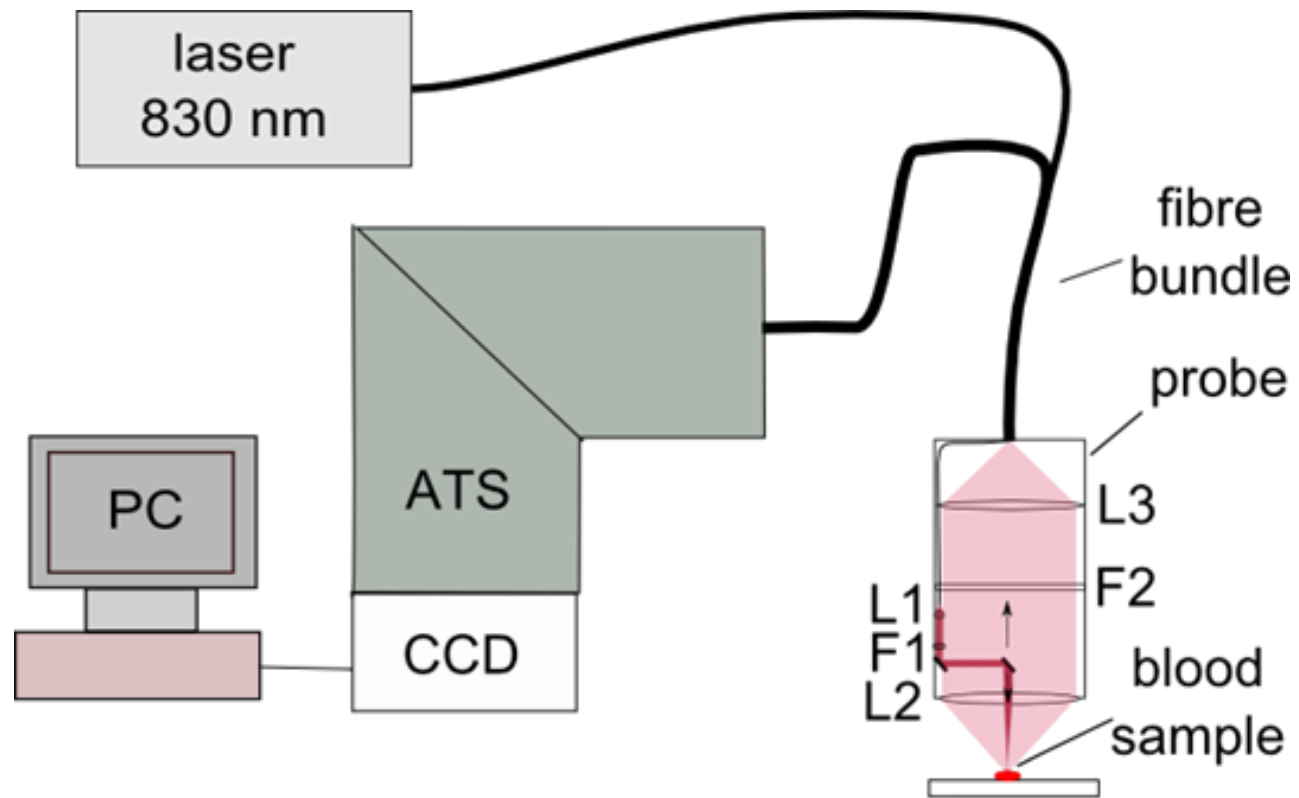
www.andor-tech.com

Made in EU



WARNING

ENSURE PC IS
SWITCHED OFF BEFORE
CONNECTING OR
DISCONNECTING CABLE



L1,L2, L3 - lenses;
F1- laser line filter
F2- longpass filter;
ATS- axial transmissive
spectrograph

Rodzaje próbek

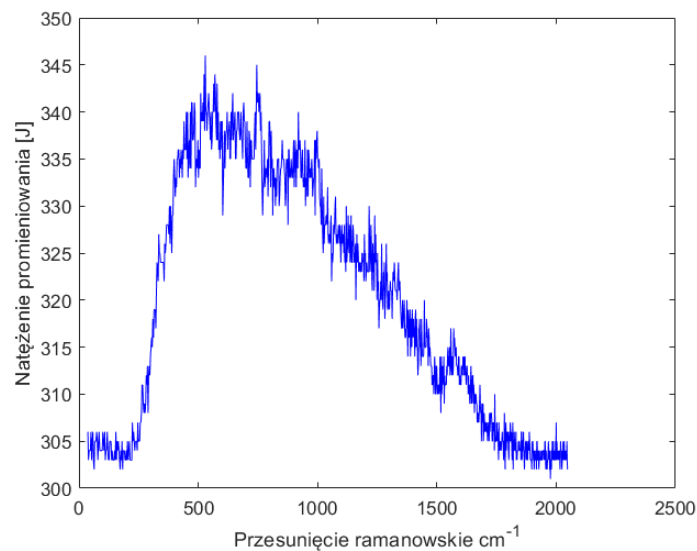




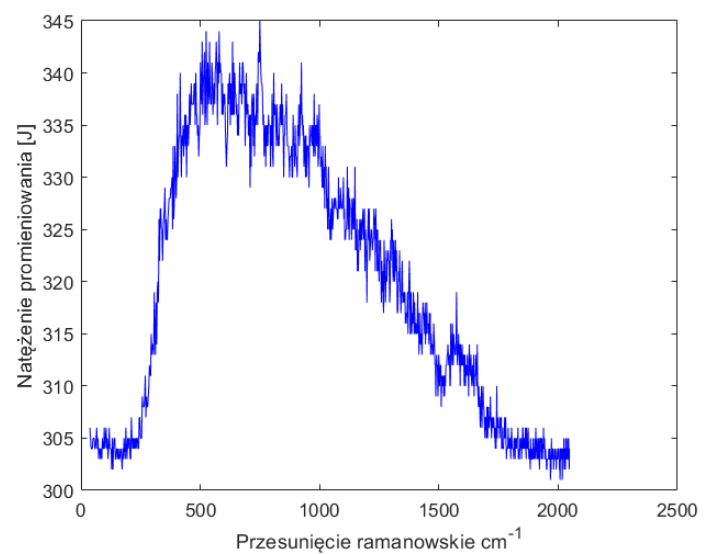
Dane



Jakie próbki badaliśmy?



Widmo ramanowskie psiej krwi.



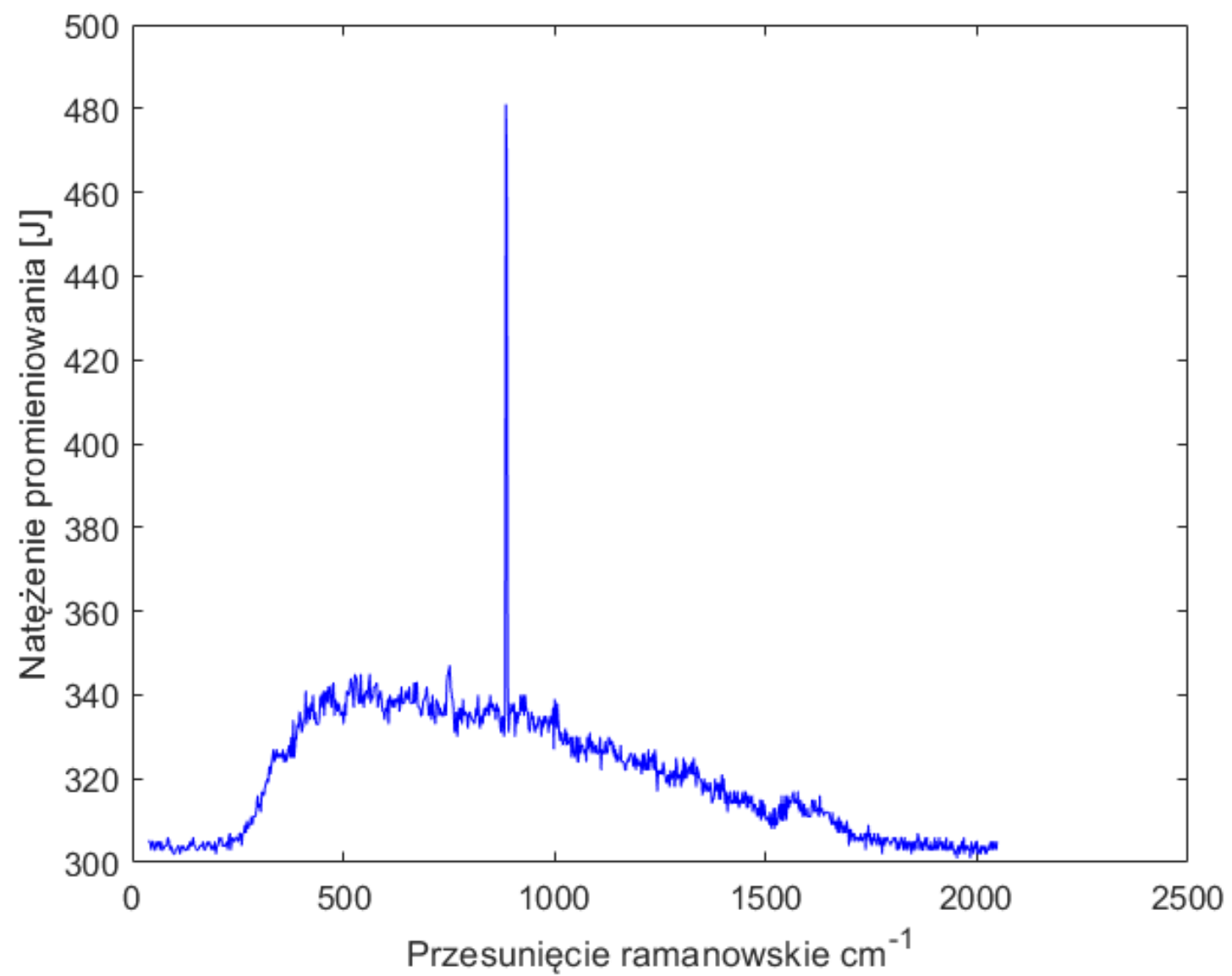
Widmo ramanowskie ludzkiej krwi.

dane1024_cz1

1025x13210 double

	2765	2766	2767	2768	2769	2770	2771	2772	2773
1000	304	305	304	304	305	303	305	304	304
1001	303	306	305	303	303	303	306	306	305
1002	305	304	305	306	304	303	305	304	305
1003	305	304	305	305	304	304	306	304	304
1004	305	305	303	303	303	304	305	305	303
1005	303	304	304	305	304	304	305	305	305
1006	303	304	303	304	305	305	305	303	303
1007	304	304	305	304	306	304	305	306	306
1008	304	305	306	305	304	305	304	304	304
1009	304	304	304	303	303	305	305	304	304
1010	306	304	304	305	304	303	303	305	303
1011	304	305	303	306	305	306	303	304	304
1012	305	304	304	304	305	303	304	305	304
1013	306	304	305	305	307	303	304	305	305
1014	304	304	305	305	305	305	306	305	304
1015	305	304	304	305	304	304	305	303	306
1016	304	304	304	306	303	305	304	305	304
1017	304	304	304	305	305	303	306	304	302
1018	305	305	304	303	303	305	306	305	303
1019	304	305	303	304	304	305	305	304	304
1020	305	304	303	304	306	304	305	306	304
1021	304	305	306	305	306	304	304	305	304
1022	305	304	304	304	304	303	304	304	305
1023	305	304	305	305	303	305	305	304	304
1024	305	304	304	306	305	307	306	306	305
1025	19	19	19	19	19	19	19	19	19

Fragment macierzy z wartościami do przetwarzania.

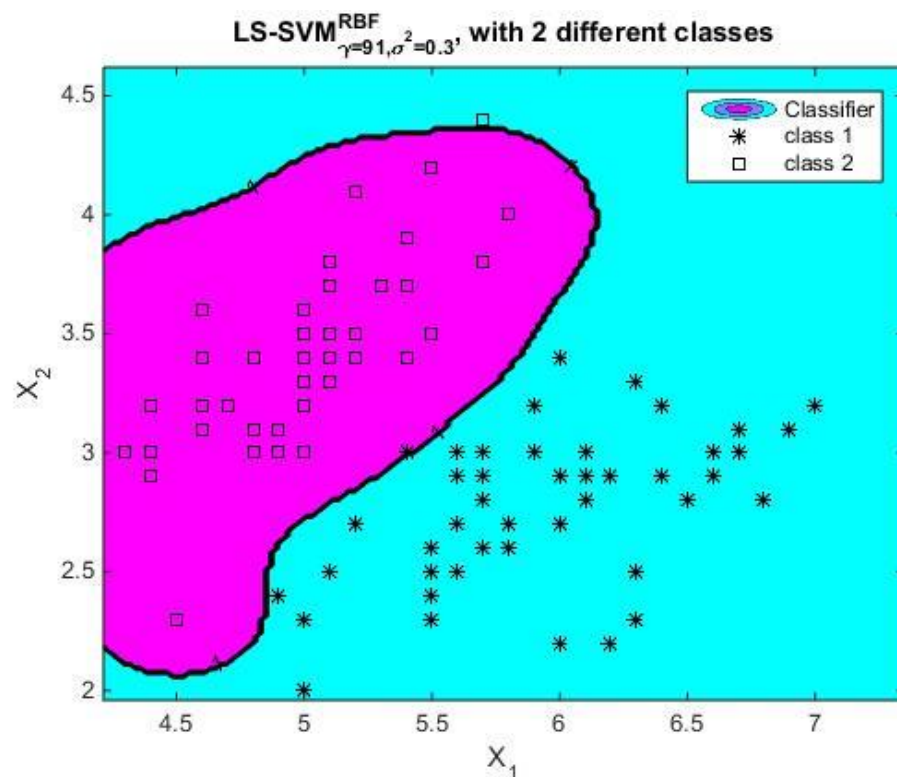


Matlab



TOOLBOX LSSVM (Least Squares Support Vector Machines)

- Toolbox jest udostępniony za darmo na licencji MIT.



Metoda PCA

Algorytm w kilku krokach

1. Zmienne są normalizowane metodą Zscore(standaryzowane)
2. Obliczane są składowe główne. Generowana jest macierz kowariancji zmiennych wejściowych.
3. Poszukiwane są wartości własne macierzy kowariancji, oraz sortowanie w kierunku malejącym w celu wybrania najbardziej znaczących wektorów.
4. Wybór k najbardziej znaczących wartości, czyli reprezentujących największą zmienność danych wejściowych



Metoda PCA

- Została wykorzystana do zmniejszenia liczby wymiarów.
- Pozwoliła oczyścić dane z szumów.
- Umożliwiła unifikację danych z kilku systemów pomiarowych i różnych kamer.

Metoda SVM

Metoda SVM

- Zasada działania algorytmu:
- Znalezienie optymalnych parametrów prostej, która podzieli zbiór na dwie części.
- Jeżeli prosta nie istnieje, szukamy prostej w hiperprzestrzeni, o liczbie wymiarów równej liczbie wymiarów macierzy wejściowych pomniejszonej o jeden.
- Znalezione proste przenoszone są do 2 wymiarów, tworząc łamaną, która jest granicą oddzielającą dane.

Metoda SVM

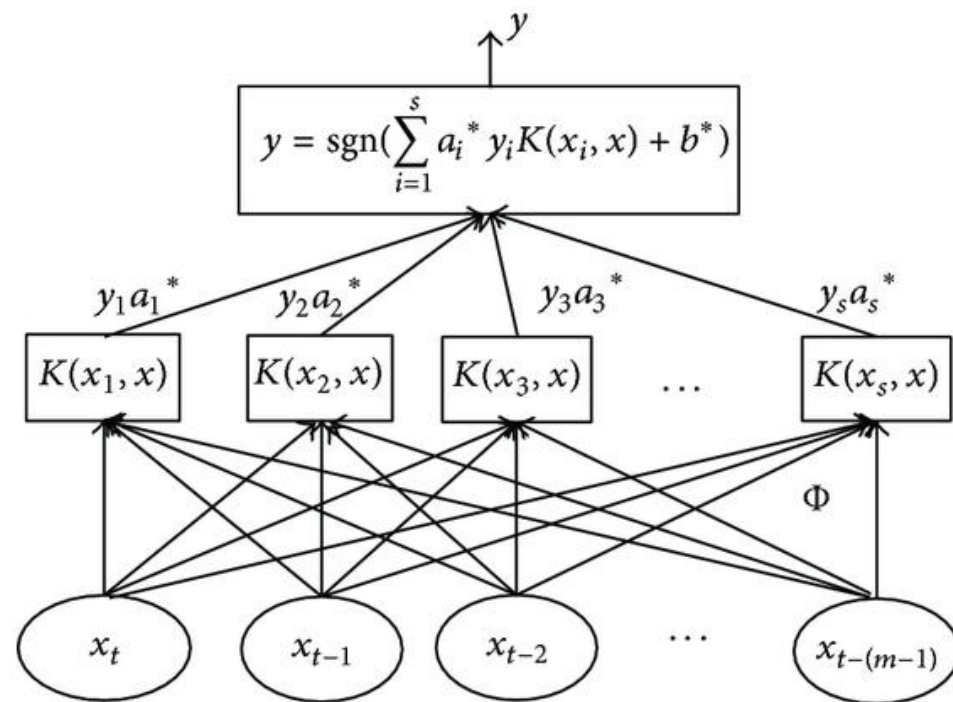
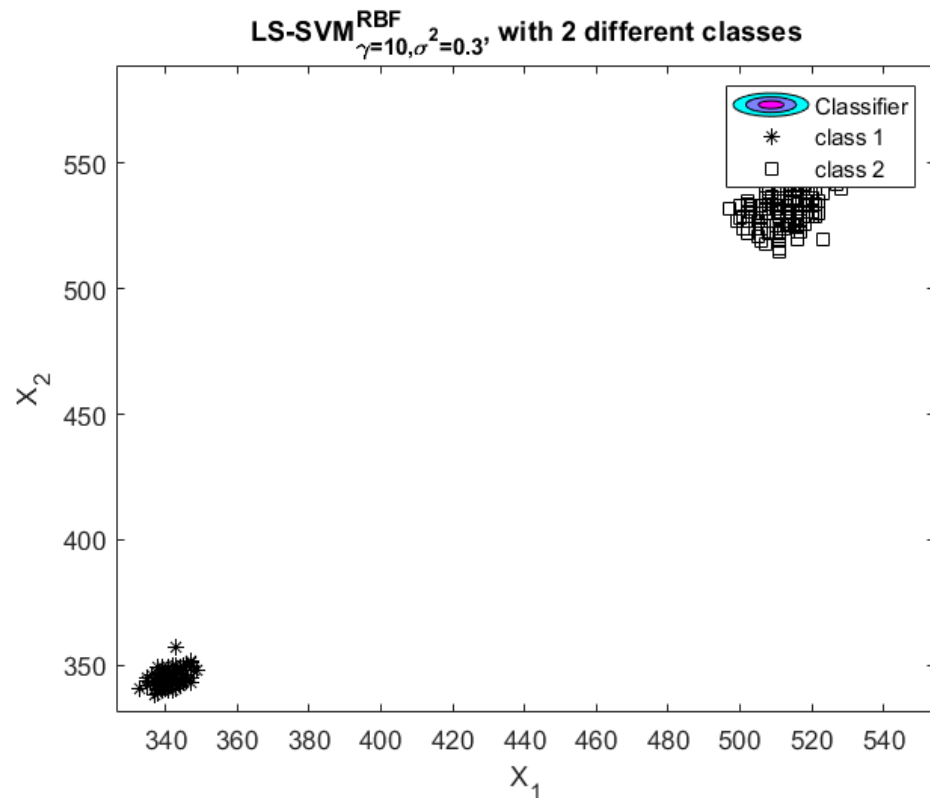


Diagram algorytmu SVM dla wszystkich próbek.

Wyniki klasyfikacji



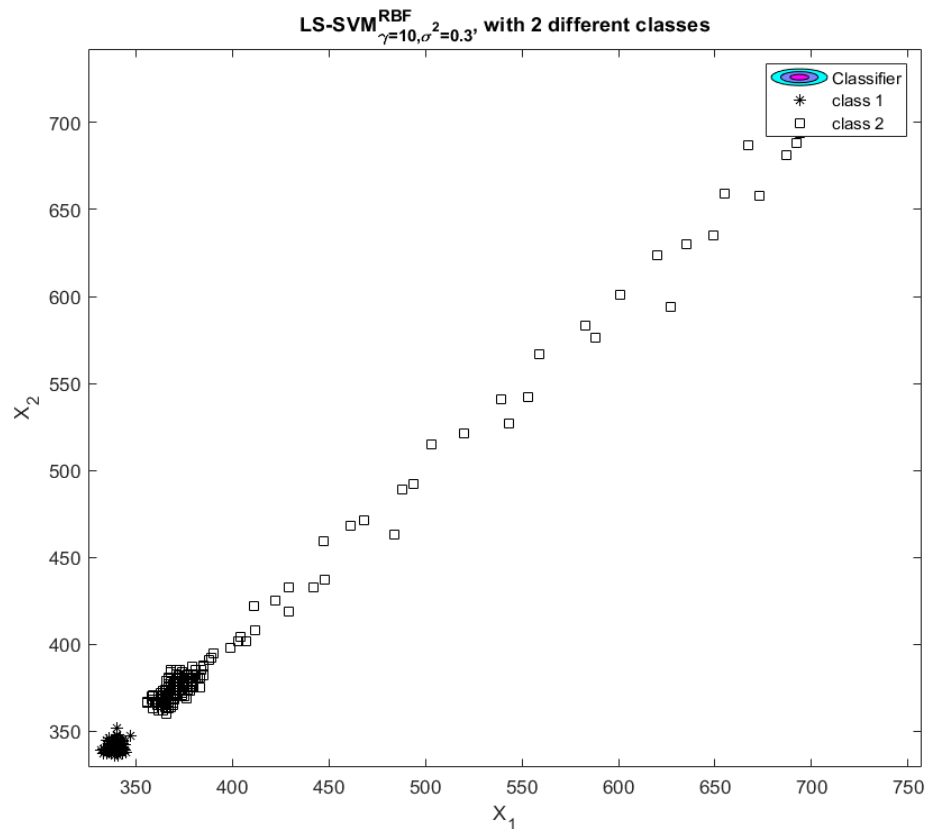
Wynik klasyfikacji



Sklasyfikowane dane alkoholu i
pustej płytki pomiarowej

Otrzymany wynik to 100%

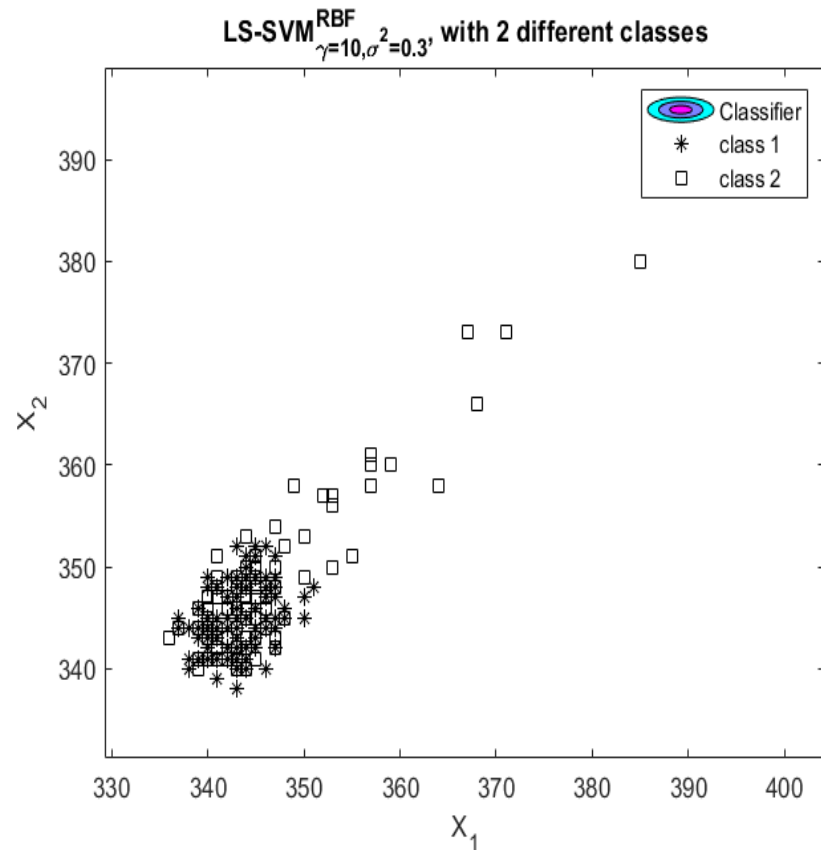
Wynik klasyfikacji



Sklasyfikowane dane krwi psiej
(kwadreaty) i ludzkiej (gwiazdki)

Otrzymany wynik to 100%

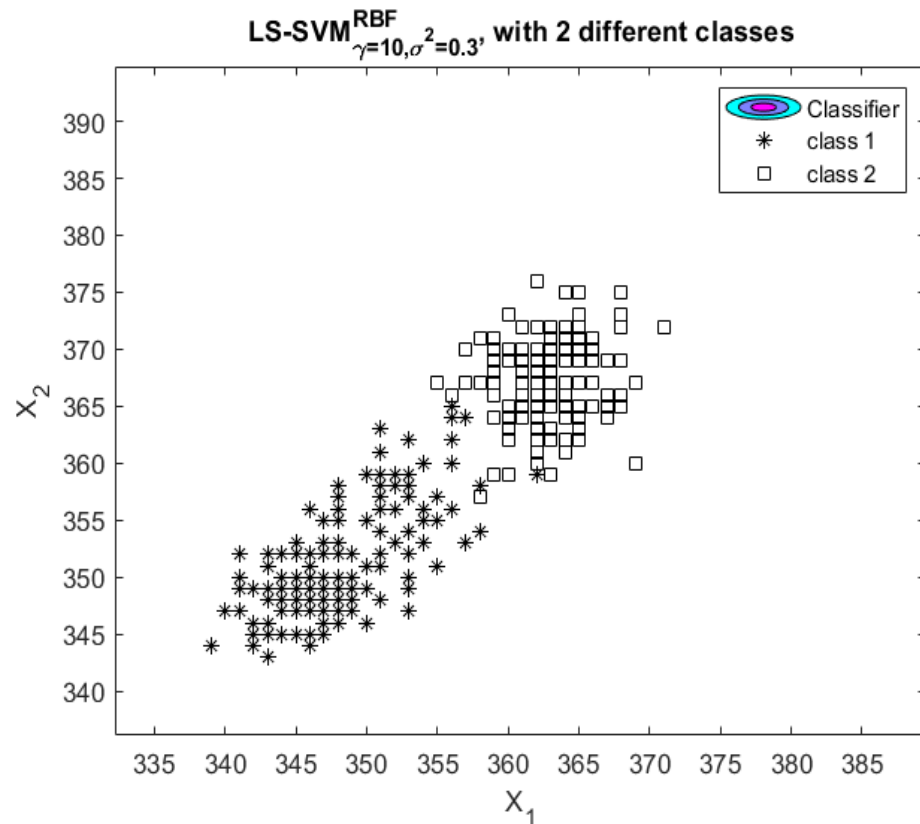
Wynik klasyfikacji



Sklasyfikowane dane krwi 2 różnych psów.

Otrzymany wynik to 61.3%

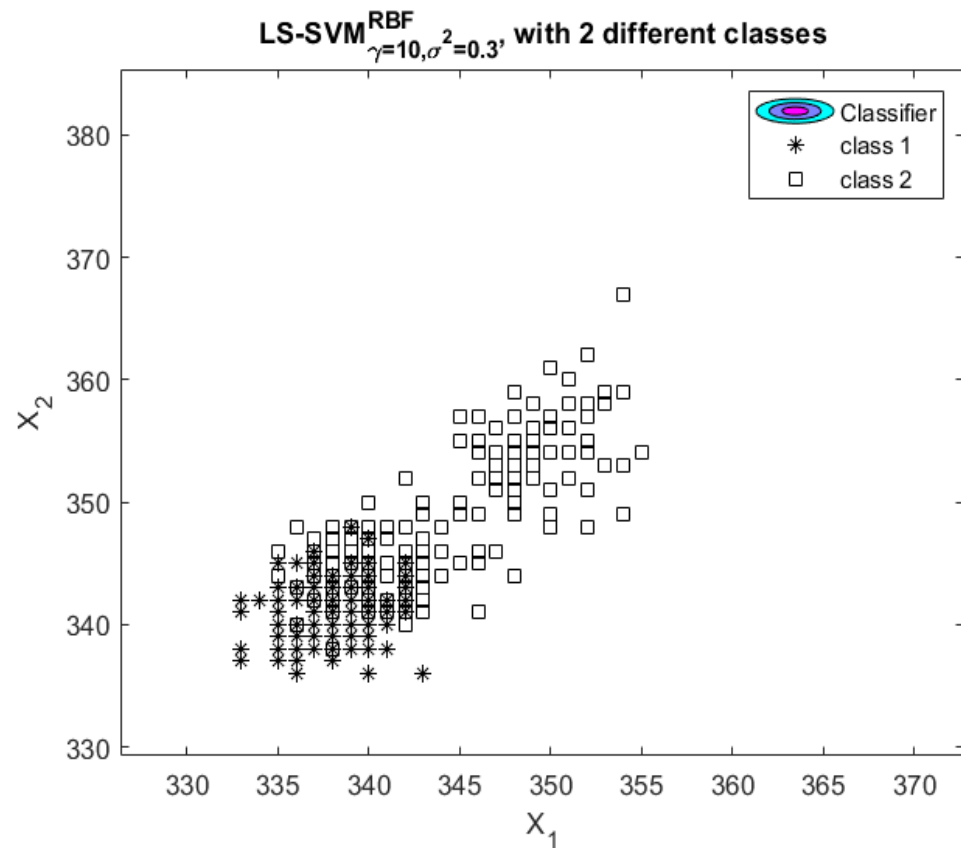
Wynik klasyfikacji



Sklasyfikowane dane krwi z
lekiem i bez leku.

Otrzymany wynik to 91.3%

Wynik klasyfikacji



Pogrupowanie krwi różnych dawców na starym systemie pomiarowym

Otrzymany wynik to 93.3%

Co dalej?



Czy da się rozwijać badania?

- Środowisko badawcze.
- Badanie bezinwazyjne.
- Badanie inwazyjne.
- Prosty układ badawczy
- Więcej danych!

Literatura i pytania

Literatura

- J. Sadlej, Spektroskopia Molekularna, Warszawa: Wydawnictwo Naukowo-Techniczne, 2002.
- W. Demtröder, Spektroskopia Laserowa, Warszawa: Wydawnictwo Naukowe PWN, 1993.
- M. Drozdowski, Spektroskopia ciała stałego, Poznań: Wydawnictwo Politechniki Poznańskiej, 1996.
- Gardiner and Graves, Practical Raman Spectroscopy, Springer-verlag, 1991.
- Z. Kęcki, Podstawy spektroskopii molekularnej, Warszawa: Wydawnictwo PWN, 1998.
- Kuptsov i A. H. Zhizhin, Handbook of Fourier Transform Raman and Infrared Spectra of Polymers, Elsevier, 1998.
- H. Barańska, A. Łabudzińska i J. Terepiński, Laserowa Spektrometria Ramanowska, Warszawa: Państwowe Wydawnictwo Naukowe, 1981.
- M. KERANEN, M. GNYBA i P. RAERINNE, „Synthesis and Characterization of Optical Sol-Gel Adhesive for Military,” *Journal of Sol-Gel Science and Technology*, tom 0.1023, nr JSST.0000048019.00322., 2004.
- M. Gnyba, M. Wróbel, R. Urniaż, T. Myllylä i M. Jędrzejwska-Szczerska, „Detection of propofol concentrations in blood by Raman spectroscopy,” Munich, 2015.
- K. Cieśllicki, Hydrodynamiczne uwarunkowania krążenia mózgowego, Warszawa: Akademicka Oficyna Wydawnicza, 2001.
- M. Szeliga, Data Science i uczenie maszynowe, Warszawa: Wydawnictwo Naukowe PWN SA, 2017.
- K. D. Brabanter, „lssvmlab,” 30 9 2019. [Online]. Available: <https://www.esat.kuleuven.be/sista/lssvmlab/>. [Data uzyskania dostępu: 30 9 2019].
- R. Fisher, „The use of multiple measurements in taxonomic problems,” *Annual Eugenics*, pp. 179-188, 1936.
- R. Fisher, „<https://archive.ics.uci.edu/>,” 30 wrzesień 2019. [Online]. Available: <https://archive.ics.uci.edu/ml/datasets/iris>. [Data uzyskania dostępu: 30 wrzesień 2019].
- D. McIlwraith, H. Marmanis i D. Babenko, Inteligentna sieć. Algorytmy przyszłości. Wydanie II, Gliwice: Helion, 2017.
- R. L. McCreery, Raman Spectroscopy for Chemical Analysis, John Wiley & Sons, 2005.