Klasyfikatory nieparametryczne

- Brak wiedzy o postaci rozkładów warunkowych cech w klasach
- Do konstrukcji klasyfikatora używamy tylko i wyłącznie wiedzy zawartej w zbiorze uczącym
- Przykłady klasyfikatorów
  - Klasyfikator minimalno-odległościowy
  - Klasyfikatory oparte na nieparametrycznej estymacji funkcji gęstości
  - Klasyfikator SVM
  - Klasyfikatory neuronowe

Klasyfikator minimalno-odległościowy

- Każda z klas reprezentowana jest przez jeden „typowy” dla niej obiekt.
- W przypadku algorytmów wykorzystujących zbiór uczący do konstrukcji funkcji klasyfikujących w roli tego reprezentanta występuje na ogół „środek” obiektów z danej klasy, to znaczy:

$$\bar{x}_j = \frac{1}{N_j}\sum_{x\in V_j} x \text{ dla } j=1,...,K,$$

gdzie V_j oznacza podzbiór zbioru uczącego, złożony z elementów należących do klasy *j*.

Klasyfikator minimalno-odległościowy

- Reguła decyzyjna klasyfikatora minimalno-odległościowego zalicza obiekt do klasy reprezentowanej przez najbliższy w sensie przyjętej metryki obiekt reprezentant $\bar{x}_j$
- funkcje klasyfikujące są postaci:

$$g_j(x) = -\left\|x - x_j\right\|^2 \text{ dla } j = 1,..., K$$

- W przypadku metryki euklidesowej:

$$g_j(x) = 2\bar{x}_j^T x - \bar{x}_j^T \bar{x}_j \,.$$

Klasyfikator minimalno-odległościowy…

Przykład:

Dla każdej z klas, rolę obiektu „reprezentanta”, pełni obiekt z wektorem cech równym wektorowi średnich, wyliczony za pomocą zbioru uczącego według wzoru:

$$\bar{x}_j = \frac{1}{|V_j|}\sum_{k\in V_j} x_k \,.$$

Wartości te ynoszą dpowiednio:

$$\bar{x}_1 = \begin{bmatrix} -2{,}734 \\ -2{,}963 \end{bmatrix},\ \bar{x}_2 = \begin{bmatrix} 3{,}254 \\ 3{,}142 \end{bmatrix}$$

Klasyfikator minimalno-odległościowy…

- Funkcja klasyfikująca *j*-tej klasy dla przyjętej normy euklidesowej

$$g_1\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = 2\cdot\begin{bmatrix} -2,734 & -2,963 \end{bmatrix}\cdot\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \begin{bmatrix} -2,734 & -2,963 \end{bmatrix}\cdot\begin{bmatrix} -2,734 \\ -2,963 \end{bmatrix} =$$
$$= -5,468x_1 - 5,926x_2 - 16,254$$

$$g_2\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = 2\cdot\begin{bmatrix} 3,254 & 3,142 \end{bmatrix}\cdot\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \begin{bmatrix} 3,254 & 3,142 \end{bmatrix}\cdot\begin{bmatrix} 3,254 \\ 3,142 \end{bmatrix} = .$$
$$= 6,508x_1 + 6,284x_2 - 20,461$$

Klasyfikator minimalno-odległościowy…

- Powierzchnia rozdzielająca w przypadku algorytmu minimalno-odległościowego wyraża się wzorem

$$2\left(\bar{x}_1^T - \bar{x}_2^T\right)x + \bar{x}_1^T\bar{x}_1 - \bar{x}_2^T\bar{x}_2 = 0$$

czyli:

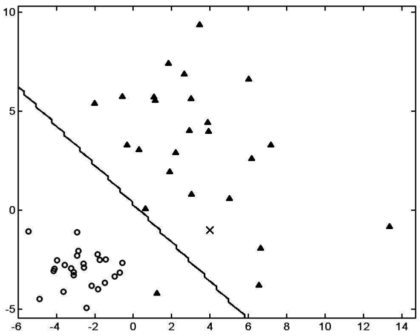
$$2\left(\begin{bmatrix} -2,734 \\ -2,963 \end{bmatrix}^T - \begin{bmatrix} 3,254 \\ 3,142 \end{bmatrix}^T\right)\cdot\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 3,254 \\ 3,142 \end{bmatrix}^T\cdot\begin{bmatrix} 3,254 \\ 3,142 \end{bmatrix} - \begin{bmatrix} -2,734 \\ -2,963 \end{bmatrix}^T\cdot\begin{bmatrix} -2,734 \\ -2,963 \end{bmatrix} = 0$$

$$-11,976x_1 - 12,21x_2 + 20,461 - 16,254 = 0$$

ostatecznie:

$$-11,976x_1 - 12,21x_2 + 4,207 = 0\,.$$

Klasyfikator minimalno-odległościowy…



Powierzchnia rozdzielająca $-11,976x_1 - 12,21x_2 + 4,207 = 0$

Klasyfikator minimalno-odległościowy…

- Działanie reguły decyzyjnej klasyfikatora minimalno-odległościowego dla obiektu opisanego wektorem cech $\begin{bmatrix}4 & -1\end{bmatrix}^T$:

$$g_1\left(\begin{bmatrix}4 & -1\end{bmatrix}^T\right) = 6,508 \cdot 4 + 6,284 \cdot (-1) - 20,461 = -0,713$$

$$g_2\left(\begin{bmatrix}4 & -1\end{bmatrix}^T\right) = -5,468 \cdot 4 - 5,926 \cdot (-1) - 16,254 = -32,2$$

$$\psi(x) = 2$$

Klasyfikatory liniowe

- Otrzymana w przykładzie powierzchnia decyzyjna jest prostą dla przypadku dwuwymiarowej przestrzeni cech
- W ogólnym przypadku jest to hiperpłaszczyzna
- Klasyfikatory, których powierzchnia decyzyjna jest hiperpłaszczyzną nazywamy liniowymi

Miary odległości

Poprzez odległość rozumiemy funkcję, która parze obiektów przyporządkowuje liczbę nieujemną, spełniającą warunki metryki

Miary odległości…

Wśród najczęściej spotykanych miar odległości wyróżniamy:

1. Odległość euklidesową:

   $$d^2(x, y) = (x - y)^T (x - y)$$

2. Odległość Minkowskiego (norma L_p)

   $$d_p(x_i, x_j) = \sqrt[p]{\sum_{k=1}^{m} \left|x_{ik} - x_{jk}\right|^p}$$

3. Odległość Mahalanobisa

   $$d_M(x_i, x_j) = (x_i - x_j)^T \Sigma^{-1} (x_i - x_j),$$

   gdzie Σ jest macierzą kowariancji

Miary odległości…

4. Odległość Sebestyena

   $$d_S(x_i, x_j) = (x_i - x_j)^T W (x_i - x_j),$$

   gdzie W jest diagonalną macierzą wag atrybutów.

- Zauważmy, że odległość euklidesowa jest szczególnym przypadkiem odległości Minkowskiego dla $p = 2$.
- W przypadku stosowania odległości euklidesowej zaleca się normalizację danych, z uwagi na wpływ skali atrybutu na jej wartość
- Odległość Sebestyena, pozwala na zróżnicowanie wpływu poszczególnych cech na odległości między obiektami, poprzez wprowadzenie diagonalnej macierzy wag współrzędnych.

Klasyfikator kNN

- Przyjmujemy pewną naturalną liczbę k i dla klasyfikowanego obiektu x poszukujemy k w sensie przyjętej odległości najbliższych mu obiektów zbioru uczącego V.
- Reguła decyzyjna zwana algorytmem k najbliższych sąsiadów zalicza obiekt do klasy najliczniej reprezentowanej wśród tych k znalezionych najbliższych obiektów zbioru uczącego. Funkcje klasyfikujące mają zatem postać:

$$g_j(x) = k_j,$$

gdzie k_j-ilość spośród k najbliższych sąsiadów należących do klasy j.

Klasyfikator kNN…

- Uczenie algorytmu kNN ogranicza się do zapamiętania zbioru uczącego
- Klasyfikator kNN wymaga ustalenia parametru k
- Brak ogólnej metody, jedną z częściej stosowanych jest dobór zgodnie ze wzorem:

  $k = c\sqrt{N}$, gdzie c jest stałą dodatnią
- Przypadkiem szczególnym algorytmu kNN jest algorytm najbliższego sąsiada – 1NN, jego funkcje klasyfikujące są postaci:

$$g_j(x) = -\min_{x_k : d_k = j} \|x - x_k\|$$

Klasyfikator kNN…

Przykład

- Działanie algorytmu zilustrujemy na przykładzie klasyfikacji obiektu, o wektorze cech równym $\begin{bmatrix} 4 & -1 \end{bmatrix}^T$ dla kilku wartości k.
- Wartością funkcji klasyfikującej i-tej klasy na wektorze x jest ilość elementów zbioru uczącego należących do klasy i, wśród k najbliższych sąsiadów x.
- Tabela zawiera posortowane rosnąco wartości odległości wspomnianego obiektu od obiektów zbioru uczącego. Wartość i-tej funkcji klasyfikującej dla zadanego k, to zatem ilość wystąpień klasy i w pierwszych k wierszach tak utworzonej tabeli

Klasyfikator kNN…

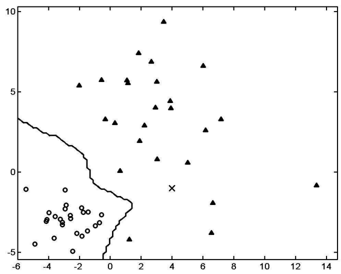
	x_1	x_2	Klasa	$x_1 - 4$	$x_2 + 1$	$\|\cdot\|$
1	5,028	0,581	2	1,028	1,581	1,886
2	3,049	0,808	2	-0,951	1,808	2,043
3	6,636	-1,928	2	2,636	-0,928	2,795
4	0,636	0,08	2	-3,364	1,08	3,533
5	1,895	1,942	2	-2,105	2,942	3,618
6	6,56	-3,787	2	2,56	-2,787	3,784
7	6,18	2,612	2	2,18	3,612	4,219
8	1,241	-4,208	2	-2,759	-3,208	4,231
9	2,21	2,911	2	-1,79	3,911	4,301
10	-0,565	-2,653	1	-4,565	-1,653	4,855
⋮	⋮	⋮	⋮	⋮	⋮	⋮

Klasyfikator kNN…

1. ***k*=1**

$g_1\left(\begin{bmatrix} 4 \\ -1 \end{bmatrix}\right) = 0$, $g_2\left(\begin{bmatrix} 4 \\ -1 \end{bmatrix}\right) = 1$, zatem wektor $\begin{bmatrix} 4 \\ -1 \end{bmatrix}$ zaliczony zostaje do klasy 2.

Powierzchnia rozdzielająca algorytmu najbliższego sąsiada:

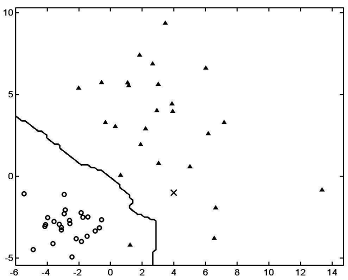


Klasyfikator kNN…

***k*=3**

$g_1\left(\begin{bmatrix} 4 \\ -1 \end{bmatrix}\right) = 0$, $g_2\left(\begin{bmatrix} 4 \\ -1 \end{bmatrix}\right) = 3$, zatem wektor $\begin{bmatrix} 4 \\ -1 \end{bmatrix}$ zaliczony zostaje do klasy 2.

Powierzchnia rozdzielająca algorytmu 3NN:

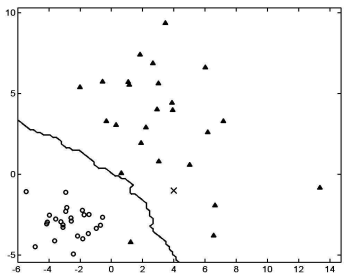


Klasyfikator kNN…

***k*=7**

$g_1\left(\begin{bmatrix} 4 \\ -1 \end{bmatrix}\right) = 0$, $g_2\left(\begin{bmatrix} 4 \\ -1 \end{bmatrix}\right) = 7$ zatem klasyfikowany wektor ten należy do klasy 2.

Powierzchnia rozdzielająca algorytmu 7NN:



Ocena klasyfikatora

- Jak wybrać klasyfikator najlepszy, dobrać parametr/parametry?
- Naturalnym sposobem wybrania najlepszego klasyfikatora jest wybór tego dla którego prawdopodobieństwo błędnego zaklasyfikowania nowej obserwacji jest minimalne
- Prawdopodobieństwo powyższe nie jest znane – szacujemy eksperymentalnie
- Sprawność klasyfikatora szacujemy na podstawie niezależnej od zbioru uczącego próby, zwanej **próbą walidacyjną.**
- Sprawność klasyfikatora – prawdopodobieństwo poprawnej klasyfikacji

Ocena klasyfikatora

- Jeśli do oceny klasyfikatora wykorzystalibyśmy zbiór uczący otrzymana sprawność będzie znacznie zawyżona
- Uczenie klasyfikatora polega na „dopasowaniu” go do danych zawartych w zbiorze uczącym
- Ostatecznej oceny sprawności klasyfikatora powinno się dokonywać w oparciu o jeszcze jedną próbę – **próbę testową**
- W przypadku gdy ocenie podlega tylko jeden klasyfikator wydzielenie próby walidacyjnej nie jest konieczne
- Konstrukcja i ocena sprawności klasyfikatora wymaga dysponowania odpowiednio licznymi zbiorami, zwłaszcza uczącym.

Ocena klasyfikatora…

- W przypadku dysponowania niewielką liczbą obserwacji podział na trzy części pozbawiony jest sensu
- W przypadku gdy podziału można dokonać jak powinien on wyglądać? Najczęściej:
    - 50% ucząca, po 25% testowa i walidacyjna
    - 60% ucząca, po 20% testowa i walidacyjna
- W przypadku braku dostatecznej liczebności w celu oceny klasyfikatora wykorzystuje się je możliwie wiele razy dla osiągnięcia możliwie niewielkiego obciążenia szacowanej sprawności

Walidacja krzyżowa

- Walidacja krzyżowa – kroswalidacja

- Idea *k* krotnej kroswalidacji:
    1. Podział próby uczącej na *k* równych części
    2. Dokonujemy uczenia klasyfikatora, przy czym jedna z wydzielonych części nie bierze udziału w uczeniu
    3. Na wydzielonej próbie dokonujemy oceny klasyfikatora
    4. Powtarzamy 2-3 *k* krotnie odkładając za każdym razem inną z wydzielonych części
    5. Uśredniamy uzyskane sprawności (błędy) klasyfikatora

Walidacja krzyżowa…

- W każdym kroku ocena sprawności odbywa się na obserwacjach nie biorących udziału w procesie uczenia
- Procedura walidacji krzyżowej jest kosztowana obliczeniowo
- Po wybraniu klasyfikatora bądź oszacowaniu parametrów metodą walidacji krzyżowej ostateczny klasyfikator budujemy na podstawie pełnego zbioru uczącego
- Na ogół *k* przyjmuje się 5 lub 10
- Często stosowanym wariantem walidacji krzyżowej jest *N* krotna walidacja krzyżowa (ang. *leave one out cross-validation*)

Metoda bootstrap

- Wielokrotne repróbkowanie oryginalnego zbioru uczącego metodą losowania ze zwracaniem
- Za każdym razem poprzez losowanie ze zwracaniem wybiera się tą samą liczbę obserwacji co w zbiorze uczącym
- Dokonujemy wielokrotnego repróbkowania, np. 1000 razy
- Średnio ok. 1/3 obserwacji nie zostaje wylosowana, dokładnie $\left(1-\frac{1}{n}\right)^n$ zatem w przybliżeniu $e^{-1} = 0,368$
- Na podtawie wylosowanych prób bootstrapowych konstruuje się klasyfikator, ocenia jego sprawność i uśrednia uzyskiwane wyniki

Ocena klasyfikatora – uwagi

- Załóżmy, że dany jest problem klasyfikacji o dwóch klasach
- W praktyce w pewnych przypadkach, np. medycznych koszty błędnej klasyfikacji zależą od klasy do której obserwacja zostanie błędnie zaklasyfikowana
- W zastosowaniach medycznych z dwojga złego lepiej zdrowego zaklasyfikować jako chorego niż na odwrót.
- Wynik działania klasyfikatora w zast. medycznym na próbie testowej można ocenić za pomocą czułości i specyficzności.

Ocena klasyfikatora – uwagi…

	Próbka zaklasyfikowana jako zdrowa	Próbka zaklasyfikowana jako chora
Próbka zdrowa	TN	FP
Próbka chora	FN	TP

TN -true negative
FP -false positive
FN -false negative
TP -false positive

Ocena klasyfikatora – uwagi…

- Prawdopodobieństwo błędnej klasyfikacji:

$$\frac{FP+FN}{TN+FP+FN+TP}$$

- Czułość – prowdopodobieństwo przewidzenia choroby, pod warunkiem, że pacjent jest na nią chory

$$\frac{TP}{TP+FN}$$

- Specyficzność – prawdopodobieństwo przewidzenia, że pacjent jest zdrowy, pod warunkiem że faktycznie jest zdrowy

$$\frac{TN}{TN+FP}=1-\frac{FP}{TN+FP}$$

Ocena klasyfikatora – uwagi…

- Pojęcie czułości / specyficzności nie dotyczy tylko zastosowań medycznych
  - Zadaniem klasyfikatora może być ocena czy dany moduł pracuje poprawnie czy został uszkodzony
  - Klasyfikacja klientów ze względu na zdolność kredytową

Klasyfikator z estymatorem jądrowym gęstości

- Klasyfikator oparty o jądrową estymację funkcji gęstości rozkładu cech w klasach
- Estymator jądrowy funkcji gęstości:

$$\hat{f}(x)=\frac{1}{Nh^p}\sum_{j=1}^{N}K(u_j),$$

gdzie $u_j=\frac{(x-x_j)^T(x-x_j)}{h^p}$, K – funkcja jądra spełniająca

warunki: $\int_{-\infty}^{\infty}K(x)=1,\ \int_{-\infty}^{\infty}|K(x)|<\infty,\ \sup_{-\infty<x<\infty}|K(x)|<\infty,$

$\lim_{\|x\|\to\infty}x\cdot K(x)=0$

Klasyfikator z estymatorem jądrowym gęstości

Przy zastosowaniu jądra gaussowskiego gęstość warunkowego rozkładu cech w klasie *j* wyraża się wzorem:

$$f_j(x)=\frac{1}{N_j}\cdot\frac{1}{h^p\sqrt{(2\pi)}}\sum_{i\in V_j}\cdot\exp\left(-\frac{(x-x_i)^T(x-x_i)}{2h^2}\right),$$

gdzie: V_j -podzbiór zbioru uczącego, złożony z elementów należących do klasy *j*.

N_j -liczba elementów zbioru uczącego należących do klasy *j*.

h -parametr, liczba dodatnia.

Klasyfikator z estymatorem jądrowym gęstości

Funkcje klasyfikujące tego klasyfikatora maja postać:

$$g_j(x) = \frac{1}{h^p N\sqrt{(2\pi)}} \sum_{i \in V_j} \cdot \exp\left(-\frac{(x - x_i)^T (x - x_i)}{2h^2} \right),$$

gdzie:

N -liczba elementów zbioru uczącego.

- Dla obliczenia wartości funkcji klasyfikujących dla klasyfikowanego wektora $x = \left[x^{(1)}, ..., x^{(p)} \right]^T$, konieczne jest pamiętanie całego zbioru uczącego.

Klasyfikator z estymatorem jądrowym gęstości…

- Niech $f_j^{(l)}(x)$ oznacza gęstość warunkowego rozkładu cech w klasie j ze względu na l-tą składową wektora x, wtedy:

$$f_j^{(l)}(x) = \frac{1}{N_j} \cdot \frac{1}{h\sqrt{(2\pi)}} \sum_{i \in V_j} \exp\left(-\frac{\left(x^{(l)} - x_i^{(l)} \right)^2}{2h^2} \right) =$$

$$= \frac{1}{N_j} \cdot \frac{1}{h\sqrt{(2\pi)}} \cdot \sum_{i \in V_j} \exp\left(\frac{-\left(x^{(l)} \right)^2}{2h^2} \right) \exp\left(\frac{-\left(x_i^{(l)} \right)^2}{2h^2} \right) \exp\left(\frac{x^{(l)} x_i^{(l)}}{h^2} \right)$$

Klasyfikator z estymatorem jądrowym gęstości…

Rozwijając w szereg Taylora w punkcie 0 ostatni składnik sumy:

$$f_j^{(l)}(x) = \frac{1}{N_j} \cdot \frac{1}{h\sqrt{(2\pi)}} \cdot \sum_{i \in V_j} \exp\left(\frac{-\left(x^{(l)} \right)^2}{2h^2} \right) \exp\left(\frac{-\left(x_i^{(l)} \right)^2}{2h^2} \right) \sum_{n=0}^{\infty} \frac{\left(x^{(l)} \right)^n \left(x_k^{(l)} \right)^n}{n! h^{2n}} =$$

$$= \frac{1}{N_j} \cdot \frac{1}{h\sqrt{(2\pi)}} \cdot \exp\left(\frac{-\left(x^{(l)} \right)^2}{2h^2} \right) \sum_{n=0}^{\infty} \left(x^{(l)} \right)^n \sum_{i \in V_j} \frac{\left(x_k^{(l)} \right)^n}{n! h^{2n}} \cdot \exp\left(\frac{-\left(x_i^{(l)} \right)^2}{2h^2} \right) =$$

$$= \frac{1}{N_j} \cdot \frac{1}{h\sqrt{(2\pi)}} \cdot \exp\left(\frac{-\left(x^{(l)} \right)^2}{2h^2} \right) \sum_{n=0}^{\infty} \left(x^{(l)} \right)^n c_{jn}^{(l)},$$

gdzie $c_{jn}^{(l)} = \sum_{i \in V_j} \frac{\left(x_k^{(l)} \right)^n}{n! h^{2n}} \cdot \exp\left(\frac{-\left(x_i^{(l)} \right)^2}{2h^2} \right).$

Klasyfikator z estymatorem jądrowym gęstości…

Zakładając niezależność warunkowych rozkładów cech w klasie j ze względu na składowe wektora x otrzymujemy:

$$f_j(x) = \prod_{l=1}^{p} f_j^{(l)}(x) = \frac{1}{\left(N_j \right)^p} \cdot \frac{1}{h^p \sqrt{(2\pi)^p}} \cdot \prod_{l=1}^{p} \exp\left(\frac{-\left(x^{(l)} \right)^2}{2h^2} \right) \sum_{n=0}^{\infty} \left(x^{(l)} \right)^n c_{jn}^{(l)}, ,$$

Czyli mnożąc przez prawdopodobieństwo a priori wystąpienia klasy j i opuszczając stałą $\frac{1}{N} \cdot \frac{1}{h^p \sqrt{(2\pi)^p}}$, która nie zależy od klasy, otrzymujemy funkcje klasyfikujące postaci

Klasyfikator z estymatorem jądrowym gęstości…

- Funkcje klasyfikujące

$$g_j(x) = \frac{1}{\left(N_j \right)^{p-1}} \prod_{l=1}^{p} \exp\left(\frac{-\left(x^{(l)} \right)^2}{2h^2} \right) \sum_{n=0}^{\infty} \left(x^{(l)} \right)^n c_{jn}^{(l)},$$

gdzie $c_{js}^{(l)} = \frac{1}{s! h^{2s}} \sum_{j \in V_j} \left(x_j^{(l)} \right)^s \exp\left[-\frac{\left(x_j^{(l)} \right)^2}{2h^2} \right].$

W takim przedstawieniu w celu wyliczenia wartości funkcji dyskryminacyjnych dla wektora x wystarczy pamiętanie tylko współczynników c_{js} wyliczonych z tego zbioru w procesie uczenia.

Klasyfikator z estymatorem jądrowym gęstości…

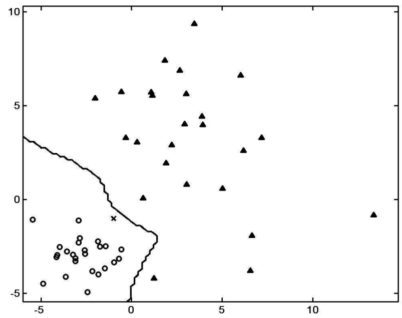
Przykład

- Algorytm Parzena z jądrem Gaussowskim, na przykładzie klasyfikacji obiektu opisanego wektorem cech $x = [-1 \quad -1]^T$ dla $h = 0,5$.
- Sposób wyliczenia funkcji klasyfikujących - w dwóch pierwszych kolumnach współrzędne wektorów powstałych przez odjęcie od wektorów zbioru uczącego klasyfikowanego wektora, w kolejnych kwadraty tych współrzędnych, wyrażenie $-\frac{(x_1 + 1)^2 + (x_2 + 1)^2}{2h^2}$, dla którego wartość funkcji exp znajduje się w ostatniej kolumnie.

Klasyfikator z estymatorem jądrowym gęstości…

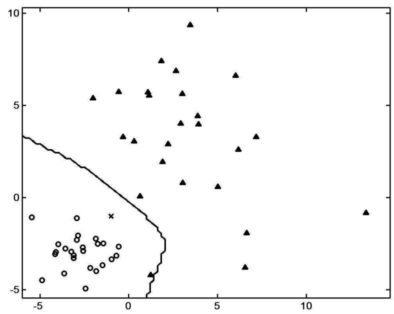
Klasa1					
x_1+1	x_2+1	$(x_1+1)^2$	$(x_2+1)^2$	$-\frac{z^T z}{2h^2}$	$\exp(...)$
-3,887	-3,489	15,109	12,173	-54,564	0,0000
-1,928	-0,104	3,717	0,011	-7,456	0,0006
0,293	-2,145	0,086	4,601	-9,374	0,0001
-1,566	-1,883	2,452	3,546	-11,996	0,0000
-1,442	-3,940	2,079	15,524	-35,206	0,0000
-0,861	-1,228	0,741	1,508	-4,498	0,0111
-2,582	-1,765	6,667	3,115	-19,564	0,0000
-2,976	-1,535	8,857	2,356	-22,426	0,0000
…	…	…	…	…	…
0,435	-1,653	0,189	2,732	-5,842	0,0029
0,026	-2,332	0,001	5,438	-10,878	0,0000
				Suma	**0,0272**

Klasyfikator z estymatorem jądrowym gęstości…



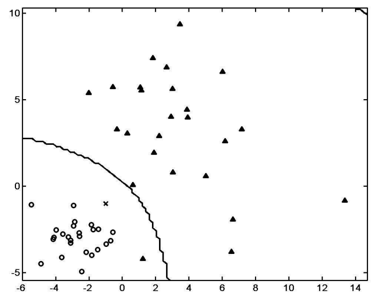
Powierzchnia decyzyjna algorytmu Parzena
z Gaussowską funkcją jądra, z parametrem h=0,5.

Klasyfikator z estymatorem jądrowym gęstości…



Powierzchnia decyzyjna algorytmu Parzena
z Gaussowską funkcją jądra, z parametrem h=2.

Klasyfikator z estymatorem jądrowym gęstości…



Powierzchnie decyzyjne klasyfikatora Parzena, h=2,
z rozwinięciem w szereg Taylora. Stopień rozwinięcia 3.