

Metoda wektorów wspierających

Celem zadania jest zapoznanie się z metodą wektorów wspierających (ang. Support Vector Machines, SVM). Pracować będziemy na zbiorach:

- jain: <http://cs.joensuu.fi/sipu/datasets/jain.txt>
- flame: <http://cs.joensuu.fi/sipu/datasets/flame.txt>

Są to proste zbiory punktów w przestrzeni dwuwymiarowej. Każdy punkt należy do jednej z do dwóch klas, mamy zatem do czynienia z klasyfikacją binarną.

Celem zadania jest zaobserwowanie:

- Jak wygląda granica decyzyjna – które obszary klasyfikator przypisuje do klasy pierwszej, a które do klasy drugiej (i z jakim poziomem pewności).
Przydatny będzie kod ze sztandarowego przykładu z biblioteki scikit-learn: https://scikit-learn.org/stable/auto_examples/classification/plot_classifier_comparison.html. Warto zwiększyć rozdzielczość skali kolorów, tak by wykrywanych było więcej poziomów.
- Które obserwacje ze zbioru wejściowego zostały wybrane jako wektory wspierające.
- Jaka jest aktualna szerokość marginesu między klasami i jak zależy oddanego parametru – przynajmniej dla liniowego SVM.

Należy rozważyć następujące rodzaje klasyfikatora SVM:

1. Liniowy SVM

$$K(x,y) = x^T y$$

...z dużym naciskiem na próbki po właściwej stronie (wysokie kary za nieprawidłowości)... ...oraz z dużym naciskiem na regularyzację i szeroki margines (niskie kary). Sprawdź szersze spektrum wartości parametrów, w tym wartości skrajne.

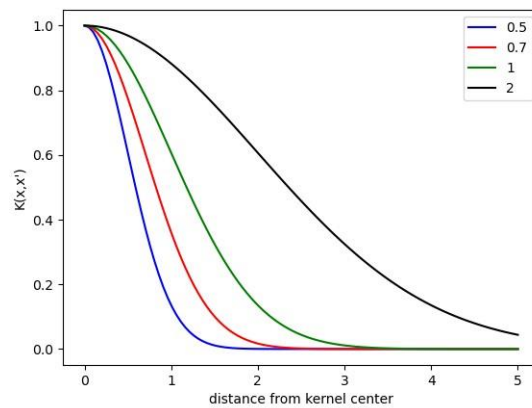
2. Kernel SVM

- (a) Kernel typu niepełny naturalny iloczyn skalarny, tj. iloczyn pierwszej składowej obu wektorów, pomijający drugą składową:

$$K(x, y) = x_1 y_1$$

- (b) Kernel typu RBF (ang. Radial Basis Function):

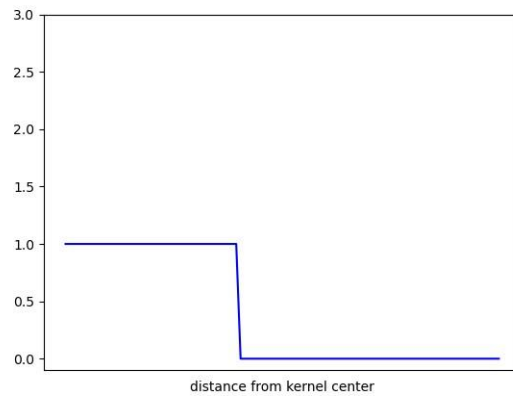
$$K(x, x') = \exp\left(-\frac{\|x - x'\|^2}{2\sigma^2}\right)$$



Na jej pracę wpływa czynnik normalizujący sigma, σ . Sprawdź skutki użycia omawianego kernela dla szerokiego spektrum wartości tego parametru.

- (c) Kernel typu skokowego:

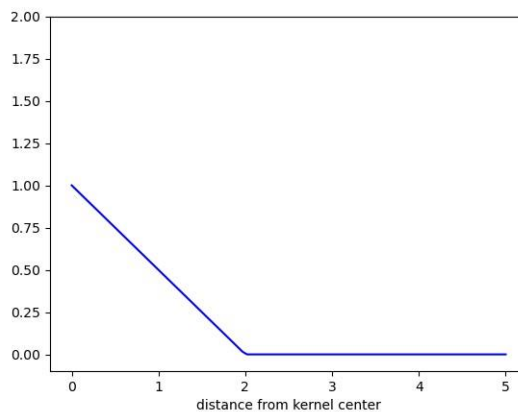
$$K(x, y) = \begin{cases} 1 & \text{dla } \|x - y\| \leq R \\ 0 & \text{dla } \|x - y\| > R \end{cases}$$



Funkcję zwraca 1 jeżeli wektory są odległe o mniej niż pewna ustalona minimalna odległość R (leżą koło siebie) i 0 w przeciwnym wypadku (nie leżą w pobliżu).

(d) Kernel typu funkcja zawiasowa (ang. hinge function):

$$K(x, y) = \begin{cases} 1 - \|x - y\|/R & \text{dla } \|x - y\| \leq R \\ 0 & \text{dla } \|x - y\| > R \end{cases}$$



Materiały:

<https://github.com/rasbt/machine-learning-book/blob/main/ch03/ch03.ipynb>