

1. Jeśli problem jest liniowo separowalny to ile granic decyzyjnych może oddzielić pozytywne przykłady od negatywnych? Którą granicę decyzji wybierze SVM?
2. Zdefiniuj problem optymalizacyjny SVM dla następującego zbioru danych.

$x_1$	$x_2$	$y$
1	1	+1
3	2	+1
1	4	+1
2	4	-1
5	1	-1
6	3	-1
5	5	-1
5	4	-1

3. Prawo Heapsa w języku naturalnym określa zależność pomiędzy liczbą unikalnych słów  $|V|$  występujących w tekście a jego długością  $N$ . Zależność ta jest określana wzorem

$$|V| = \alpha N^\beta$$

gdzie  $\alpha$  i  $\beta$  są parametrami zależącymi od języka (angielski, polski) ale też rodzaju tekstu (czat, literatura piękna) i muszą być szacowane empirycznie. Mając do dyspozycji jeden długi tekst (np. „Pan Tadeusz”) zaproponuj metodę estymacji parametrów  $\alpha$  i  $\beta$  technikami uczenia maszynowego.

4. W jaki sposób moglibyśmy uzyskać nieliniową granicę decyzji używając liniowego SVM? Jakie są ograniczenia podanego sposobu?
5. Dokonaj transformacji problemu (hard) SVM do postaci dualnej.
6. W formulacji dualnej problemu SVM optymalizowane są zmienne, które zwykle oznaczamy jako  $\alpha_i$  i nie są to wagi. W jaki sposób zatem znajdowana jest hiperpłaszczyzna separująca?
7. Czym jest tzw. trick jądrowy? Czym jest funkcja jądrowa?
8. Pokaż, że wielomianowa funkcja jądrowa  $K(x, z) = (x^T z + 1)^2$  policzona na dwuwymiarowych przykładach uczących tj.  $x = [x_1, x_2]$ , daje taki sam wynik jak funkcja liniowa  $K(x, z) = x^T z$  policzona na rozszerzonej przestrzeni cech transformatą  $\Theta(x) = [1, \sqrt{2}x_1, \sqrt{2}x_2, x_1^2, x_2^2, \sqrt{2}x_1x_2]$ .
9. Podsumuj wady i zalety stosowania metody SVM.