

Tutorial 4: SVM

1. Krótka prezentacja klasyfikatora SVM:
 - a. Klasyfikator liniowy
 - b. Klasyfikator SVM
 - c. Trik kernelowy
 - d. Opisać badanie jakości klasyfikacji (cross-validation) przy pomocy krzywych ROC i Precision-Recall, pól pod tymi wykresami (także F1). Porównać dwie metody. Wytłumaczyć jaka jest między nimi różnica.
2. Na zbiorze WBC (wybierając odpowiednią liczbę przykładów) zademonstrować działanie klasyfikatora SVM dla różnych kerneli. Dokonać walidacji krzyżowej. Porównać krzywe ROC i PR a także miary F-1 oraz powierzchni pod nimi.
3. Podglądnąć wyniki używając metody PCA.
4. ZADANIE dla studentów: Na wybranych przez prowadzących tutorial zbiorach danych dokonać klasyfikacji metodą SVM dla różnych kerneli, oraz klasyfikacji metodą k-NN dla $k=1,3$ i 5. Podglądnąć otrzymane wyniki przy pomocy PCA.

Skomentować otrzymane wyniki.

Zrobić to samo, ale redukując wcześniej wymiarowość danych metodą PCA. Np. w przypadku zbioru MNIST z 784 do 30. Czy wyniki uległy poprawie? Jeżeli tak (nie) to dlaczego? Skomentować otrzymane wyniki.