

Podstawy uczenia maszynowego: plan kursu

Krzysztof Ślot

13 maja 2019

1. Podstawowe koncepcje klasyfikacji danych

(a) Klasyfikacja danych

- Metody bazujące na ocenie odległości od klasy (niparametryczne: NN, k-NN, parametryczne: NM)
 - Koncepcje: [model klasy i trening modelu](#), [odległość próbki od klasy](#)
- Metody bazujące na ocenie prawdopodobieństwa przynależności próbki do klasy (parametryczne: klasyfikacja Bayesowska)
 - Koncepcje: [p-stwo warunkowe](#), [p-stwa a priori i a posteriori](#), [Likelihood](#), [MLE](#)

(b) Regresja

- Liniowa
 - Koncepcje: [kryterium uczenia - błąd średniokwadratowy \(MSE\)](#), [błędy uchybu \(bias\) i wariancji](#)

(c) Klasyfikacja danych, c.d.

- Podział przestrzeni cech: powierzchnie decyzyjne i regiony decyzyjne
 - Regresja logistyczna - [funkcja logistyczna](#), [rozkład dwumianowy](#), [kryterium log-prob](#)

2. Klasyfikacja w przestrzeniach wielowymiarowych

- (a) Klasyfikacja minimalnoodległościowa i Bayesowska danych wielowymiarowych - [macierz kowariancji](#)
- (b) Klasyfikacja przez podział przestrzeni cech: drzewa decyzyjne i maszyny wektorów wspierających (SVM)

3. Uczenie nienadzorowane

- (a) Grupowanie (klasteryzacja)

- Bazujące na reprezentacji skupień (k-means, EM, k-medoids)
 - Bazujące na podobieństwie próbek (Spektralne - SC, DBSCAN)
- (b) Budowa przestrzeni cech
- Selekcja cech (kryterium Fishera)
 - Ekstrakcja cech (PCA, LDA, MDS, ICA)
4. Nieliniowe problemy klasyfikacji
- (a) Agregacje klasyfikatorów
- Łasy decyzyjne (*bagging*)
 - Wzmacnianie (*Ada-boost*)
- (b) Rozbudowa klasycznych metod klasyfikacji
- Probabilistycznych - GMM (*Gaussian Mixture Models*)
 - Minimalnoodległościowych (drzewa k-wymiarowe + k-NN, NM dla wielu modów)
 - SVM z wykorzystaniem funkcji jądra (*kernel methods*)

Zalecana literatura

- Materiały kursu
- Podręczniki dostępne w sieci:
 1. Understanding Machine Learning: From Theory to Algorithms, Shai Shalev-Shwartz, Shai Ben-David, Cambridge University, 2014 (dostępne on-line za darmo)
 2. The Hundred-Page Machine Learning Book, Andriy Burkov (dostępne on-line za darmo na stronie Autora)
 3. “Deep Learning”, Ian Goodfellow, Yoshua Bengio, Aaron Courville, and Francis Bach, (dostępne rozdziały)
 4. “The Elements of Statistical Learning”, Trevor Hastie, Robert Tibshirani, and Jerome Friedman, 2013 (dostępne)