

Tuning hiperparametrów

20 grudnia 2016

1 Dobór wartości parametrów uczenia

Zajęcia mają na celu zaprezentowanie wybranych parametrów uczenia oraz przykładowych metod ich doboru. We wszystkich zadaniach należy użyć danych z konkursu Let's Roq.

2 Stała regularyzacji

★

Treść

Naucz regresję logistyczną http://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LogisticRegression.html. Przeprowadź dobór wartości parametru C , regulującego siłę regularyzacji.

3 Liczba epok/iteracji uczenia

★

Treść

Użyj algorytmu optymalizacji Stochastic Gradient Descent ucząc model liniowy ze stratą logistyczną. http://scikit-learn.org/stable/modules/generated/sklearn.linear_model.SGDClassifier.html#sklearn.linear_model.SGDClassifier. Dla ustalonego `learning_rate` i `eta0` przeprowadź 30 epok uczenia. Sprawdź wartość błędu logistycznego na zbiorze treningowym i walidacyjnym po każdej epoce. Następnie użyj `learning_rate = 'constant'`, `eta0 = 0.0001` (te wartości nie są optymalne, dobrane aby można było łatwo analizować wyniki) i zobacz jak zachowuje się błąd w kolejnych epokach.

4 Przeuczenie

★

Treść

Naucz drzewo decyzyjne scikit-learn.org/stable/modules/generated/sklearn.tree.DecisionTreeClassifier.html#sklearn.tree.DecisionTreeClassifier.

Sprawdź wpływ parametrów `min_samples_split` oraz `min_samples_leaf` na `accuracy` na zbiorze treningowym i walidacyjnym.