

Zadania dzień 4 – implementacja algorytmu k-nn

Piotr Pasza Storożenko

1. Wczytaj ramkę danych `iris`.
2. Zamień kolumnę z gatunkiem na zmienną typu `categorical`.
3. Oddziel zmienne opisujące od zmiennej opisywanej i zamień powstałą macierz oraz wektor na `ndarray` z `numpy`'a.
 - Nazywamy zmienne opisujące jako `X`.
 - Zmienną opisywaną jako `y`.
4. Podziel zbiór na treningowy i testowy w stosunku 3:1. Nazywamy je:
 - `X_train`
 - `y_train`
 - `X_test`
 - `y_test`
5. Zaimplementuj klasyfikator 3-nn.
 - Dla każdej obserwacji ze zbioru testowego, znajdź odległości tej obserwacji od obserwacji ze zbioru treningowego.
 - Wybierz klasę która występuje najczęściej wśród `k` najbliższych obserwacji.
6. Sprawdź jakość `accuracy` napisanego klasyfikatora (wykorzystaj funkcję z poprzednich zajęć).
7. Przedstaw wyniki na wykresach.
8. Porównaj wyniki 1-nn, 3-nn, 5-nn, 7-nn.