Zadania - ewaluacja wyników

1. Dla zbioru danych z zadania 2 z listy 3 przygotuj model (klasyfikator) używając metody k najbliższych sąsiadów (K-Neighbors Regressor) z parametrem liczby sąsiadów równym 2.

A. Wyznacz prognozowane klasy (y\_prediction) za pomocą wytrenowanego modelu używając części walidacyjnej zbioru.

B. Wyświetl 15 pierwszych prognozowanych klas i porównaj z pierwszymi 15 wartościami ze zbioru walidacyjnego (y\_val)

C. Używając walidacyjnej części zbioru wyznacz dokładność (accuracy) dla przygotowanego modelu.

D. Zinterpretuj uzyskaną dokładność (sprawdź wcześniej co oznaczają wartości target dla wczytanego zbioru).

E. Porównaj dokładność wyznaczoną dla zbioru walidacyjnego i dla zbioru treningowego - jak rozumiesz różnicę pomiędzy uzyskanymi wynikami?

2. Dla przygotowanego wcześniej modelu (zad.1) używając zbioru walidacyjnego zbadaj dodatkowe metryki:

A. wyznacz i zinterpretuj: precyzję klasyfikacji (ang. precision),

B. wyznacz i zinterpretuj: czułość klasyfikacji (ang. recall = sensitivity = hit rate),

C. wyznacz i zinterpretuj odporne metryki F1 score ir: (F1-score - średnia harmoniczna pomiędzy precyzją i czułością; r = współczynnik Mathewsa),

D. wyznacz i wyświetl macierz pomyłek (ang. confusion matrix), zinterpretuj uzyskane wartości, pamiętając co oznaczają poszczególne 3 badane klasy,

E. wyznacz i wyświetl raport klasyfikacji (ang. classification report), zinterpretuj uzyskane wartości,

3. Dla zbioru danych z zadania 7 z listy 3 przygotuj model (regresor) używając metody k najbliższych sąsiadów (K-Neighbors Regressor) z parametrem liczby sąsiadów równym 2.

A.Wyznacz prognozy (y\_prediction) za pomocą wytrenowanego modelu używając części walidacyjnej zbioru.

B. Wyświetl 10 pierwszych prognoz i porównaj z pierwszymi 10 wartościami ze zbioru walidacyjnego (y\_val)

C. Używając walidacyjnej części zbioru wyznacz średni błąd bezwzględny - MAE (mean absolute error) dla przygotowanego modelu.

D. Zinterpretuj wyznaczony błąd (sprawdź wcześniej co oznaczają wartości target dla wczytanego zbioru).

E. Porównaj średni błąd bezwzględny wyznaczony dla zbioru walidacyjnego i dla zbioru treningowego - jak rozumiesz różnicę pomiędzy uzyskanymi wynikami?

4.Dla przygotowanego wcześniej modelu (zad.3) używając zbioru walidacyjnego zbadaj dodatkowe metryki:

A. wyznacz i zinterpretuj: błąd średniokwadratowy - MSE (ang. mean square error),

B. wyznacz i zinterpretuj: podstawowy błąd średniokwadratowy - RMSE (ang. root mean square error) - jako kwadrat MSE,

C. wyznacz i zinterpretuj: średni bezwzględny błąd procentowy - MAPE (ang. mean absolute percentage error)

D. wyznacz i zinterpretuj: współczynnik determinacji - R2 (ang. coefficient of determination)

5. Dla zbioru danych z zadania 2 z listy 3 przygotuj model (klasyfikator) używając metody k najbliższych sąsiadów (K-Neighbors Regressor) z parametrem liczby sąsiadów równym 2.

A. Wyznacz prognozowane klasy (y\_prediction) za pomocą wytrenowanego modelu używając części walidacyjnej zbioru.

B. Wyświetl 15 pierwszych prognozowanych klas i porównaj z pierwszymi 15 wartościami ze zbioru walidacyjnego (y\_val)

C. Używając walidacyjnej części zbioru wyznacz dokładność (accuracy) dla przygotowanego modelu.

D. Zinterpretuj uzyskaną dokładność (sprawdź wcześniej co oznaczają wartości target dla wczytanego zbioru).

E. Porównaj dokładność wyznaczoną dla zbioru walidacyjnego i dla zbioru treningowego - jak rozumiesz różnicę pomiędzy uzyskanymi wynikami?

6. Wykonaj polecenia (zad.1 i zad.2 z tej listy) dla zbioru danych z zadania 3 z listy 3

7. Wykonaj polecenia (zad.1 i zad.2 z tej listy) dla zbioru danych z zadania 4 z listy 3

8. Wykonaj polecenia (zad.1 i zad.2 z tej listy) dla zbioru danych z zadania 5 z listy 3

9. Wykonaj polecenia (zad.1 i zad.2 z tej listy) dla zbioru danych z zadania 6 z listy 3

• W zadaniach warto skorzystać z modułu: sklearn.metrics - https://scikit-learn.org/stable/modules/classes.html#module-sklearn.metrics