

Technologie informatyczne II

Projekt-

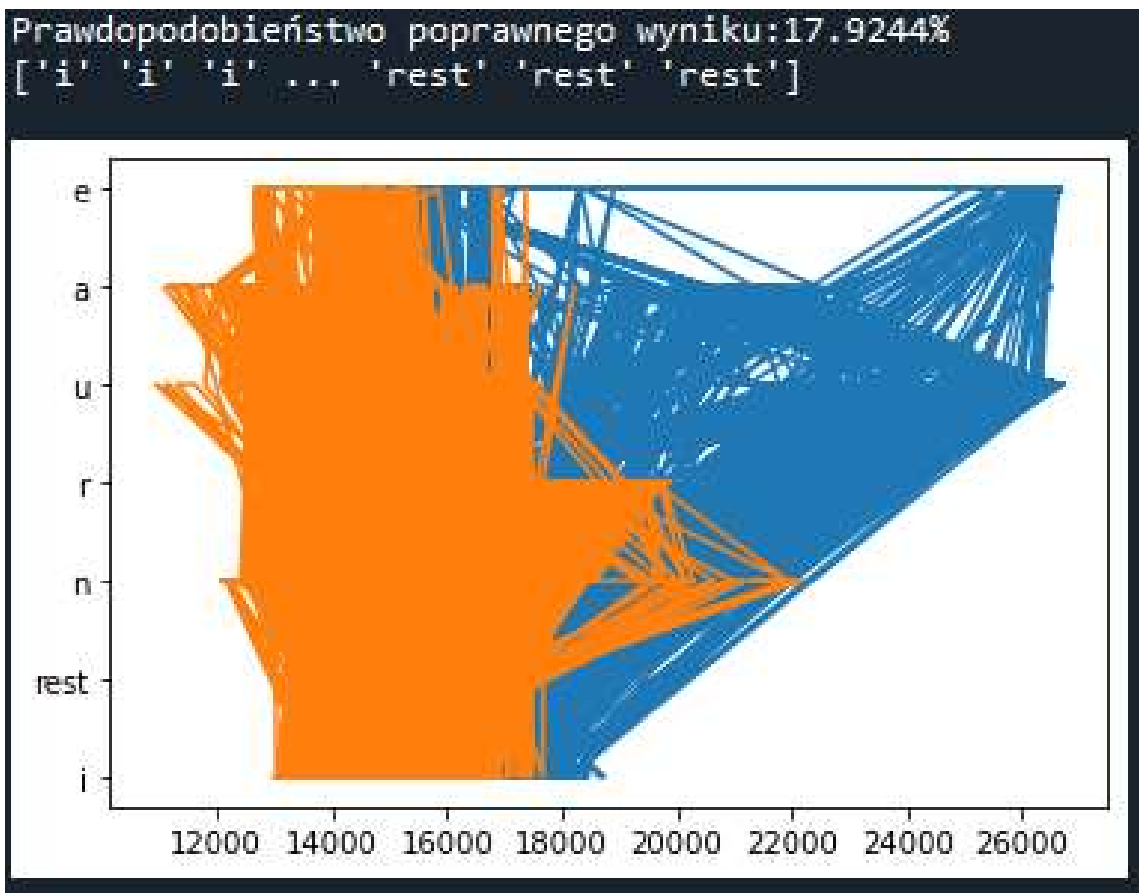
Grupowanie danych m.in. sposobem neuronowym

Sprawozdanie

Hubert Ciszek 240859, Julian Błyskal 240854

„Neural network models”

Zrzut ekranu wyniku



Uwagi:

Program przewiduje literę na podstawie przedstawionych mu danych aczkolwiek nie udało się nam stworzyć wykresu pomiędzy danymi. Pomimo prób łączenia danych za pomocą komendy np. meshgrid a następnie stworzenia wykresu za pomocą komendy plt.contourf za każdym razem wyskakiwał błąd uniemożliwiający stworzenie wykresu. Możliwe że jest on spowodowany błędnym łączeniem danych co skutkuje problemem w stworzeniu wykresu.

Program również podaje prawdopodobieństwo na to czy wynik jest poprawny

Uwagi do metody „Gaussian Process”:

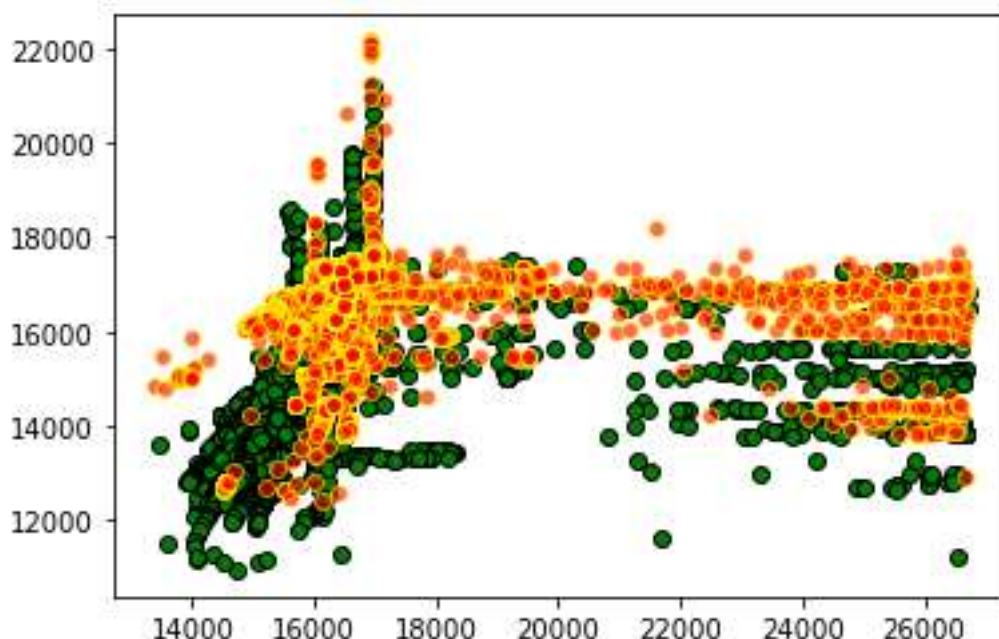
Ta metoda nie chciała z niewiadomych nam przyczyn przypisać wartości do jakiegokolwiek litery

Uwagi do metody „SVM” ,

W/w metoda nie funkcjonowała ponieważ komenda **from sklearn.inspection import DecisionBoundaryDisplay** nie działała.

„Stochastic Gradient Descent”

Wykres punktowy dla wybranych danych:



Tak samo jak z poprzednią metodą (oraz następną) nie udało się nam się stworzyć wykresu przedstawionego na stronie scikit-learn.org co jest w naszej opinii spowodowane błędnym połączeniem wartości funkcją `np.meshgrid`

Sam program natomiast z sukcesem przewiduje literę na podstawie danych

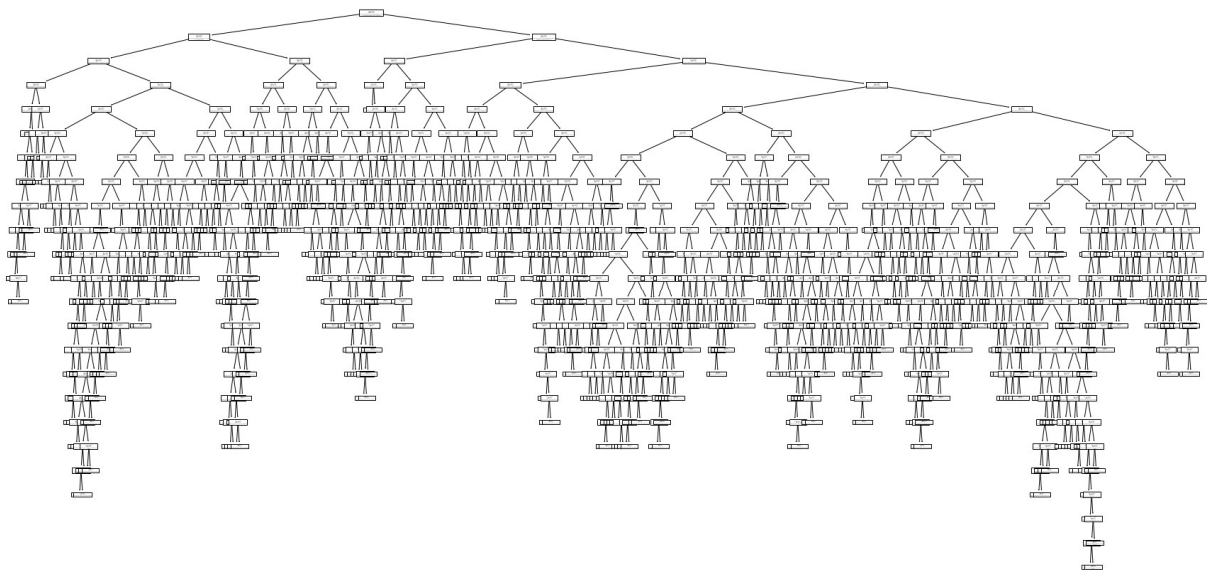
„Nearest Neighbors regression”

„module” object is not callable – ten błąd również pojawiał się przy próbach zastosowania metody gaussa do naszych danych przez co nie byłem w stanie zmusić programu do odczytania poprawnie danych z pliku

DRZEWO DECYZYJNE

Wykres metodą „decision tree”:

Jest to wykres dla wszystkich danych



Wnioski:

- Funkcja `[:,0]` zmienia 2d na 1d co umożliwia stworzenie wykresu funkcji za pomocą komendy `plt.contourf`
- Neuronowa metoda wydaje się być najlepszą metodą do grupowania danych ponieważ było z nią najmniej jakichkolwiek problemów (pomijając wykresy).
- Przy dużej ilości danych korzystanie z komendy np. `meshgrid()` może powodować zawieszenie się zarówno programu python oraz całego komputera.

- Ilość danych jest wprost proporcjonalna do prawdopodobieństwa poprawnego wyniku obliczonego przez metodę neuronową
- Nie dało się zbalansować rozmiaru czcionki tak aby tekst był czytelny i jednocześnie na tyle mały aby komórki na siebie wzajemnie nie nachodziły

Podział pracy:

Hubert Cisek – Metoda neuronowa oraz metoda SGD

Julian Błyskal – Drzewo decyzyjne oraz metoda bliskich sąsiadów

Bibliografia:

<https://scikit-learn.org/stable/>

W załączniku do raportu przesyłamy programy:

tree.py – Drzewo decyzyjne

SGD3.py – metoda „Stochastic Gradient Descent”

NEURONOWA4.py – metoda neuronowa