**PODSTAWY SZTUCZNEJ INTELIGENCJI – PROJEKT 2**

Joanna Kiesiak, Katarzyna Rzeczyca

1. TEMAT

Wybrano temat PZ.U11 „Testy akustyczne”. Zadanie polegało na porównaniu dwóch algorytmów: perceptronu wielowarstwowego oraz regresji liniowej.

1. DANE

Dane wykorzystywane w projekcie pochodzą z wykonanych przez NASA testów akustycznych płatów nośnych przeprowadzonych w dźwiękoszczelnym tunelu aerodynamicznym. Są one dostępne na stronie Center for Machine Learning and Intelligent Systems pod adresem:

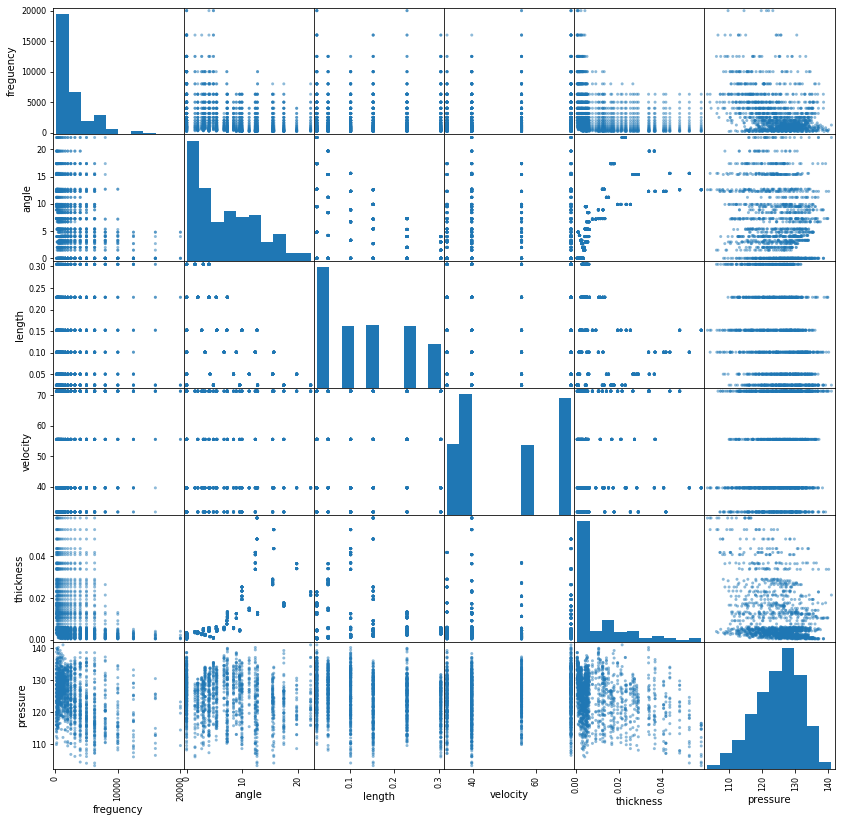
[https://archive.ics.uci.edu/ml/datasets/Airfoil+Self-Noise#](https://archive.ics.uci.edu/ml/datasets/Airfoil+Self-Noise)

Dane wejściowe stanowiły kolejno:

1. Częstotliwość [Hz]
2. Kąt natarcia [o]
3. Długość cięciwy [m]
4. Prędkość strumienia powietrza [m/s]
5. Grubość przesunięcia warstwy ssącej [m]

Wartościami wyjściowymi był poziom przeskalowanego ciśnienia dźwięku [dB]. Liczba pomiarów wyniosła 1503. Poniżej przedstawiono współczynniki korelacji dla poszczególnych danych oraz macierz wykresów zależności poszczególnych parametrów względem siebie.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **frequency** | **angle** | **length** | **velocity** | **thickness** | **pressure** |
| **frequency** | 1.000000 | -0.272765 | -0.003661 | 0.133664 | -0.230107 | -0.390711 |
| **angle** | -0.272765 | 1.000000 | -0.504868 | 0.058760 | 0.753394 | -0.156108 |
| **length** | -0.003661 | -0.504868 | 1.000000 | 0.003787 | -0.220842 | -0.236162 |
| **velocity** | 0.133664 | 0.058760 | 0.003787 | 1.000000 | -0.003974 | 0.125103 |
| **thickness** | -0.230107 | 0.753394 | -0.220842 | -0.003974 | 1.000000 | -0.312670 |
| **pressure** | -0.390711 | -0.156108 | -0.236162 | 0.125103 | -0.312670 | 1.000000 |



Na podstawie tabeli zamieszczonej powyżej można zauważyć, że dużą korelacją dodatnią charakteryzują się wyłącznie kąt natarcia oraz grubość przesunięcia warstwy ssącej. Z kolei niemała korelacja ujemna jest obserwowana pomiędzy długością cięciwy i kątem natarcia. Pozostałe parametry wydają się być niezależne liniowo. Na podstawie wykresów można ocenić, że duża część z nich przyjmuje wartości dyskretne, a więc w czasie eksperymentu zmieniały się one skokowo.

1. REALIZACJA PROJEKTU

Projekt napisano w pythonie, co pozwoliło na wykorzystanie bibliotek pandas, sklearn, numpy oraz scipy dostarczających implementacji wszystkich potrzebnych funkcji i algorytmów.

Zbiór danych podzielono losowo na część treningową i testową w proporcji 4:1. Następnie wypróbowano skuteczność algorytmów wymienionych w treści zadania dla domyślnych parametrów oraz porównano ją z wynikami otrzymanymi dla lasu losowego i regresji liniowej połączonej z regularyzacją (zarówno L1 i L2). To samo powtórzono dla znormalizowanych danych wejściowych. Na końcu wykorzystano metodę kroswalidacji i przeszukiwania losowego dla dokładniejszej eksploracji możliwości regresji liniowej z regularyzacją oraz perceptronu wielowarstwowego w zależności od zmieniających się parametrów algorytmów.

Jako miarę dopasowania przyjęto współczynnik R2.

1. WYNIKI

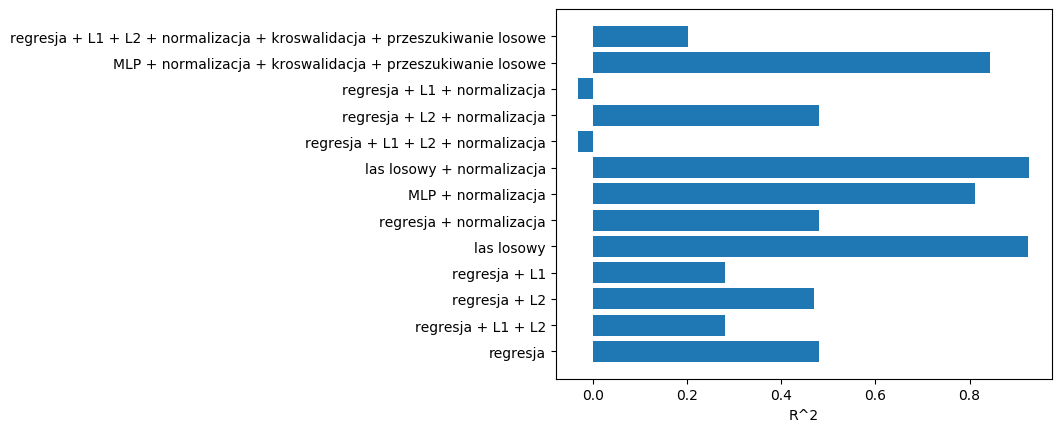
Obserwacje wskazują, że dla niemodyfikowanych danych i domyślnych parametrów regresja liniowa zapewnia wyraźnie lepsze dopasowanie (R2≈0,5) niż perceptron wielowarstwowy (R2<0). Regularyzacja L1 w przypadku regresji liniowej ma negatywny wpływ na wyniki (R2≈0,2). Najlepszym dopasowaniem w takiej sytuacji charakteryzuje się algorytm lasu losowego (R2≈0,9). Trzeba jednak zaznaczyć, że dokładne wartości końcowe zależą m.in. od przypadkowego podziału zbioru próbek na zbiory treningowy i testowy.

Po znormalizowaniu danych sprawdzono wpływ tej operacji na otrzymywane wyniki. W przypadku regresji liniowej nie miało to wpływu na wartość R2. Pogorszeniu jednak uległo dopasowanie dla regresji uwzględniającej regularyzację (R2≈0,0). Zauważalnej poprawie uległy wyniki otrzymywane na wyjściu perceptronu wielowarstwowego (R2≈0,75). Charakteryzują się one jednak dużymi wahaniami wartości przy ponownym uruchomieniu algorytmu w zależności od zainicjalizowanych współczynników początkowych sieci i jest to naturalnym zachowaniem tego modelu.

W dalszej części projektu zastosowano metodę kroswalidacji oraz przeszukiwania losowego w celu dokładnego dopasowania hiperparametrów. Zdecydowano się na przetestowanie regresji liniowej z regularyzacją (zamiast wersji klasycznej), ponieważ daje to możliwość zmiany większej liczby elementów wejściowych.

Pod wpływem zastosowanych metod udało się otrzymać modele MLP, spośród których najlepszy osiągnął średni wynik R2 ok. 0,9. Finalna uzyskana wartość na danych testowych była jednak porównywalna do otrzymanej poprzednio (R2≈0,8). Z kolei udało się poprawić dopasowanie modelu regresji liniowej z regularyzacją i ponownie osiągnąć wartość R2 porównywalną do początkowej (R2≈0,2).

Poniżej umieszczono przykładowy wykres porównujący otrzymane współczynniki R^2 po jednym wykonaniu każdego z wyżej wymienionych wariantów. Pominięto jedynie wynik otrzymany dla MLP bez normalizacji aby zachować przejrzystość ilustracji (duża wartość bezwzględna).



1. WNIOSKI

Perceptron wielowarstwowy jest skutecznym narzędziem estymującym skomplikowane zjawiska, jednak jego użyteczność jest w bardzo dużym stopniu uzależniona od odpowiedniego doboru parametrów wejściowych. Warto więc poświęcić czas na ich optymalizację. Szczególnie duże znaczenie miała w tym przypadku normalizacja danych wejściowych. Nie miała ona jednak wpływu na skuteczność dopasowania modelu regresji liniowej. Wykazano też, że nie zawsze regularyzacja pozytywnie wpływa na otrzymywane wyniki. Należy również zaznaczyć, że w celu otrzymania jak najlepiej dopasowanego modelu warto wypróbować dodatkowo inne algorytmy, jak np. las losowy, który w tej sytuacji okazał się najtrafniej przybliżać przedstawiony problem.

W przypadku tego konkretnego zastosowania spośród dwóch algorytmów wymienionych w temacie projektu lepszy okazał się perceptron wielowarstwowy.