

Raport zespołowy nr 2 - Reklama

Dominik Sobieraj – 308208

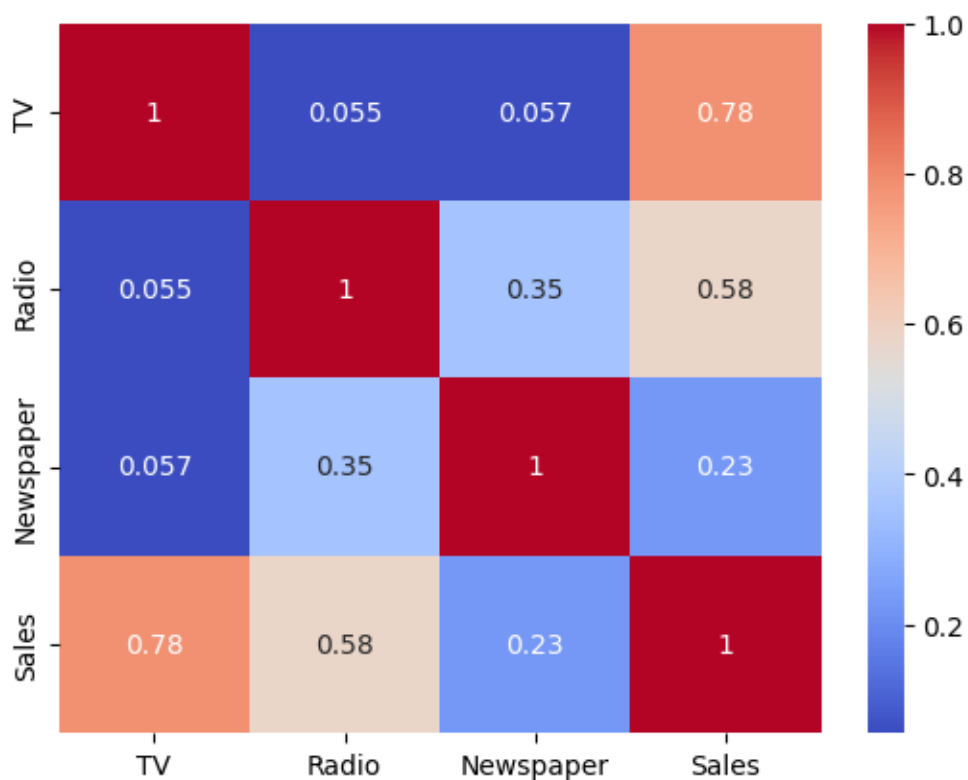
Jan Łukaszewicz – 308187

Unikalny parametr do generatorów – 308197

Wszystkie informacje dot. przeróbki danych, jak i kroki podejmowane, aby osiągnąć wyniki przedstawione poniżej wraz z ich rozumowaniem można znaleźć w osobnym pliku w załącznikach pod nazwą „**models.ipynb**”.

Macierz korelacji danych:

- `Sales` z `Tv` to aż 0.78, zatem inwestycja w reklamie TV musi przynosić korzyści, skoro wraz ze wzrostem inwestycji w Tv rośnie sprzedaż.
- `Sales` z `Radio` 0.58, inwestycja w radio również się opłaca.
- `Sales` z `Newspaper` 0.23 dzięki reklamie w gazecie sprzedaż również rośnie, lecz najmniej przyczynia się do wzrostu sprzedaży.



Model 1 – sieć neuronowa

Opis sieci:

1. Sieć składa się z 3 warstw, warstwy wejściowej, jednej ukrytej i wyjściowej.
2. Liczba neuronów w warstwie wejściowej wynosi 3, ponieważ na wejściu otrzymuje trzy wartości: `TV`, `Radio` i `Newspaper`.
3. Domyślną liczbę neuronów w warstwie ukrytej.
4. Jest to funkcja aktywacji, która aktywuje neuron.
5. Sieć ma jedną jednostkę w warstwie wyjściowej, jakim jest przewidywany wynik.
6. Jest to funkcja tożsamościowa, która nie zmienia swojego wejścia, $f(x) = x$. Oznacza to, że wyjście z neuronu jest równoważne jego wejściu.

Zbiór uczący

Średni błąd bezwzględny (MAE) 2.676

Błąd średniokwadratowy (MSE) 12.328

Pierwiastek błędu średniokwadratowego (RMSE) 3.511

Średni bezwzględny błąd procentowy (MAPE) 23.351 %

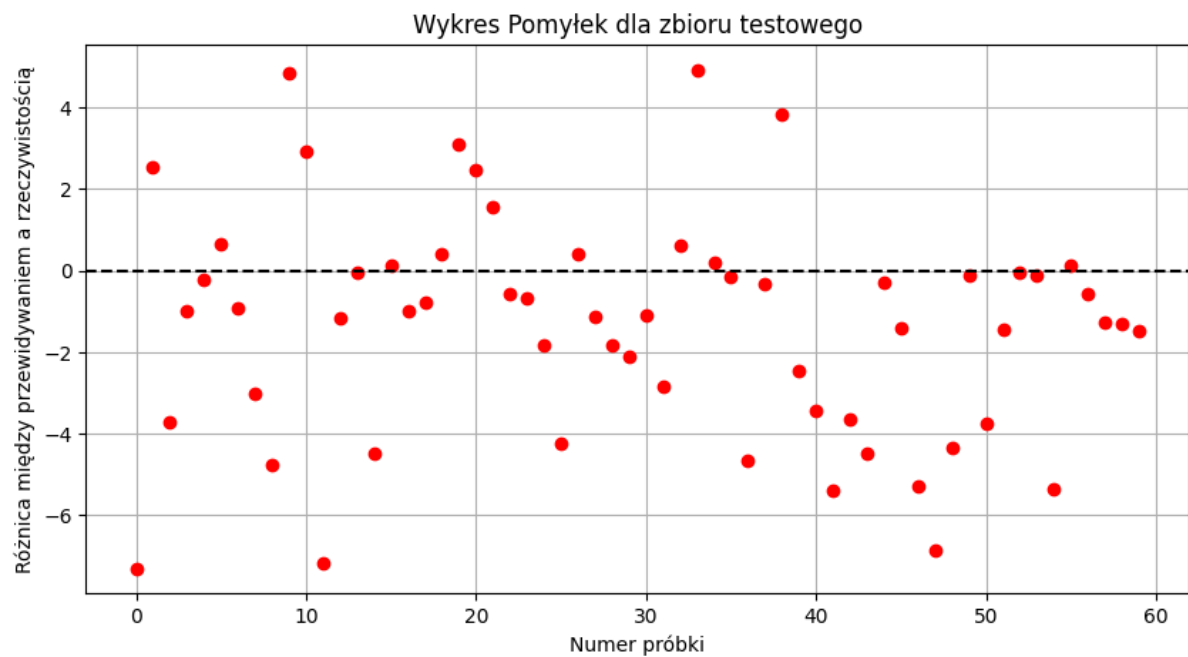
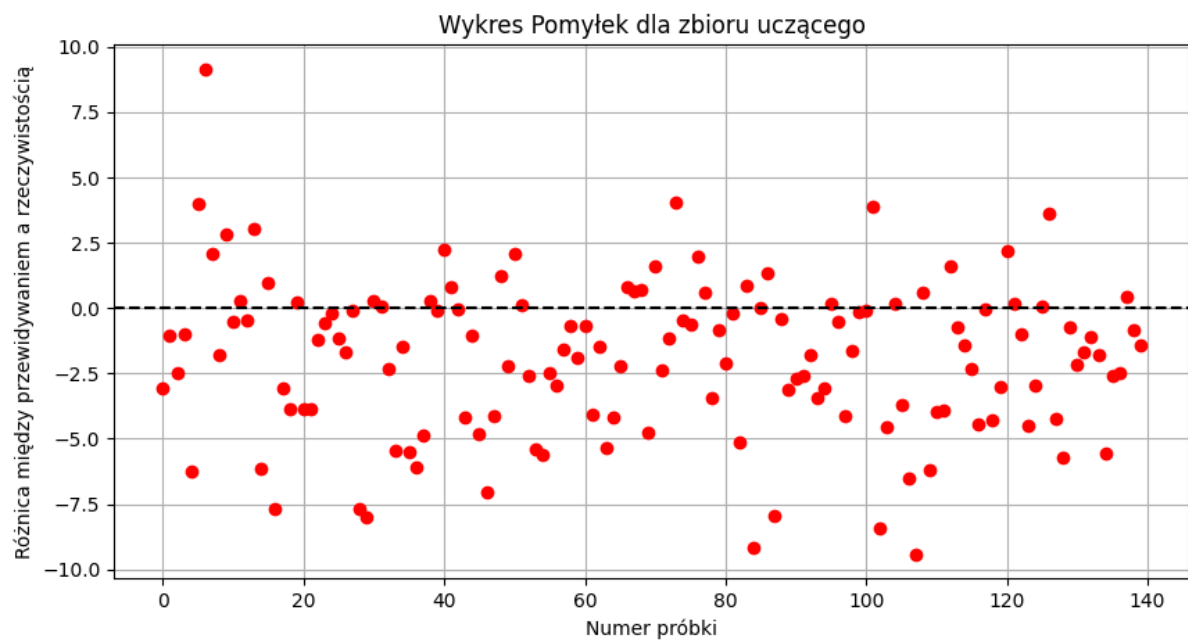
Zbiór testowy

Średni błąd bezwzględny (MAE) 2.312

Błąd średniokwadratowy (MSE) 9.36

Pierwiastek błędu średniokwadratowego (RMSE) 3.059

Średni bezwzględny błąd procentowy (MAPE) 17.628 %



Model 2 – Lasy losowe

Model został zbudowany za funkcji RandomForestRegressor w sklearn.ensemble

Głębokość lasu losowego została ograniczona.

Udało się odnaleźć idealne parametry dla niego na podstawie wczytanych danych:

```
{'max_depth': 7, 'min_samples_split': 10}
```

Zbiór uczący

Średni błąd bezwzględny 0.501

Błąd średniokwadratowy 0.475

Pierwiastek błędu średniokwadratowego 0.689

Średni bezwzględny błąd procentowy 5.416 %

Zbiór testowy

Średni błąd bezwzględny 0.828

Błąd średniokwadratowy 1.134

Pierwiastek błędu średniokwadratowego 1.065

Średni bezwzględny błąd procentowy 6.948 %

Jak widać dane zostały dobrze dobrane bo ciężko zauważyć przeuczenie modelu.

Ważność predyktorów:

