

# Sprawozdanie z listy 3

Eksploracja danych

Marta Stankiewicz (282244)

Paweł Nowak (282223)

2025-05-06

## Spis treści

1	Klasyfikacja na bazie modelu regresji liniowej	1
---	--	---

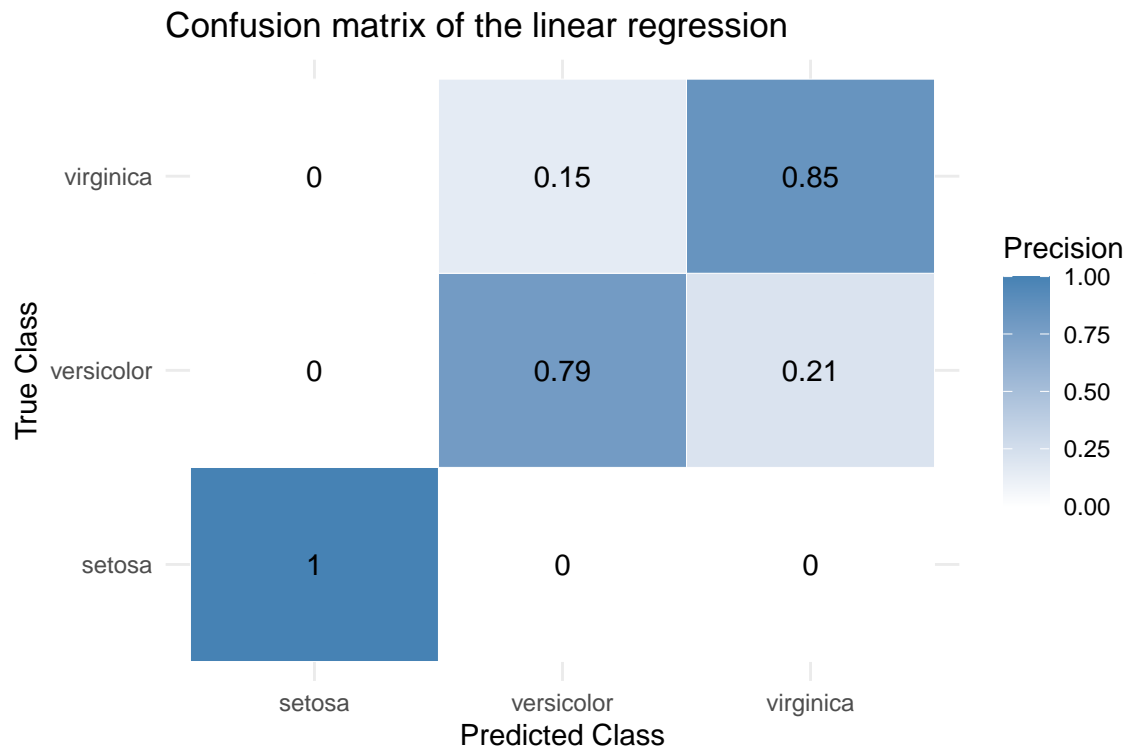
## Spis rysunków

1	Confusion matrix of Linear Regression for training samples . . . . .	2
2	Confusion matrix of Linear Regression for testing samples . . . . .	3

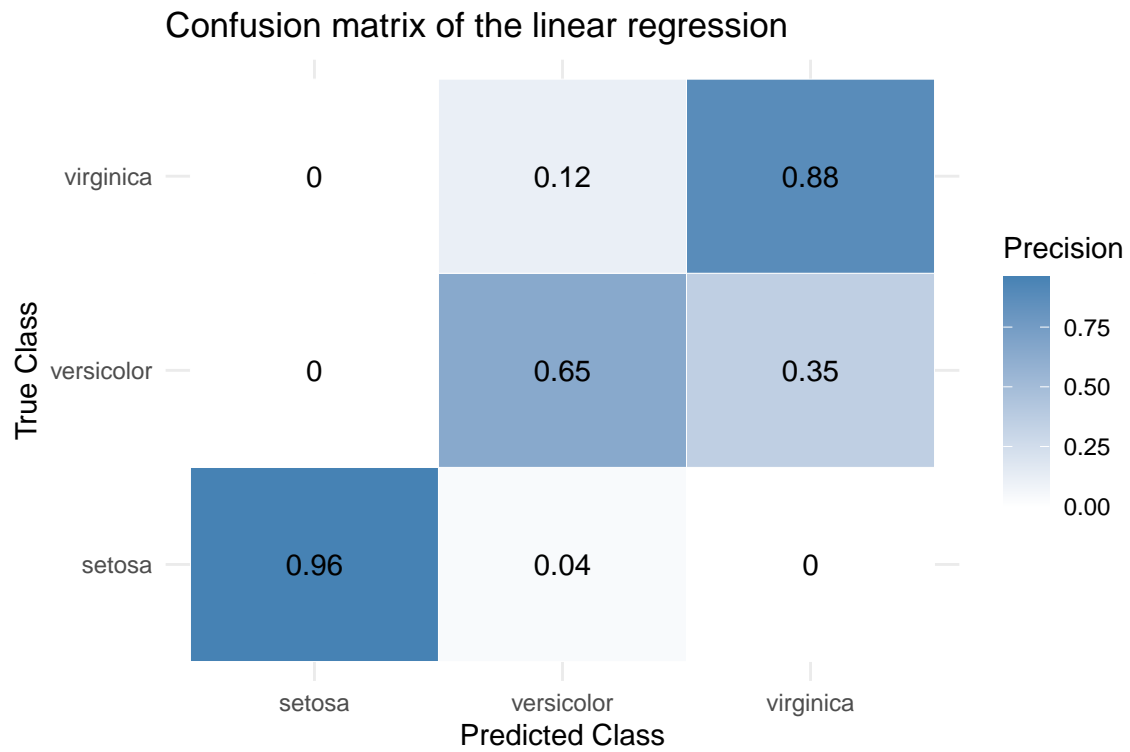
## Spis tabel

### 1 Klasyfikacja na bazie modelu regresji liniowej

Aby ocenić skuteczność klasyfikatora opartego na modelu regresji liniowej, wykorzystamy zbiór danych iris, w którym zmienną objaśnianą jest Species, zawierająca 3 unikalnych klas. W celu uniknięcia problemu wycieku danych (ang. data leakage), przeprowadzimy podział oryginalnego zbioru na zbiór treningowy oraz zbiór testowy, przy czym zbiory te będą zawierały odpowiednio 60 obserwacji z danych iris. Po wytrenowaniu modelu na danych treningowych, przeprowadzimy ewaluację jego skuteczności na podstawie zbioru testowego. Wyniki klasyfikacji zaprezentujemy za pomocą znormalizowanej macierzy pomyłek, w której wartości w każdej kolumnie zostaną podzielone przez sumę elementów tej kolumny, co pozwoli na lepszą interpretację skuteczności klasyfikacji dla poszczególnych klas.



Rysunek 1: Confusion matrix of Linear Regression for training samples



Rysunek 2: Confusion matrix of Linear Regression for testing samples