

Sprawozdanie Laboratorium 4 Regresja liniowa z regularyzacją L1 i L2

Piotr Gierżatowicz-Sierpień 331376

Cel zadania

W zadaniu należało zaimplementować model regresji liniowej z możliwością użycia regularyzacji L1 i L2. Dodatkowo, trzeba było wykorzystać wybrany algorytm optymalizacji gradientowej do trenowania modelu. Celem zadania jest przewidywanie wagi ryb, na podstawie gatunku i wymiarów.

Eksperymenty

Test różnych wartości parametrów regularyzacji i ocena na zbiorze walidacyjnym:

Przeprowadziłem eksperyment mający na celu wyłonienie najlepszej wartości λ dla regularyzacji L1 i L2, oraz najlepszego parametru $learning_rate$ dla modelu. Wartości dla których testowany był model to odpowiednio $[0.01, 0.02, 0.03, 0.04, 0.05, 0.1, 0.2, 0.3, 0.5]$ dla parametru λ oraz $[0.001, 0.01, 0.1]$ dla $learning_rate$

Wartościami dla których modele osiągnęli najlepsze wyniki okazały się:

Dla modelu z regularyzacją L1: $\lambda = 0.01$, $learning_rate = 0.01$

Dla modelu z regularyzacją L2: $\lambda = 0.01$, $learning_rate = 0.1$

Wartości MSE:

Dla modelu L1 około 4659

Dla modelu L2 około 4763

Wartości R2:

Dla modelu L1 około 0.9592

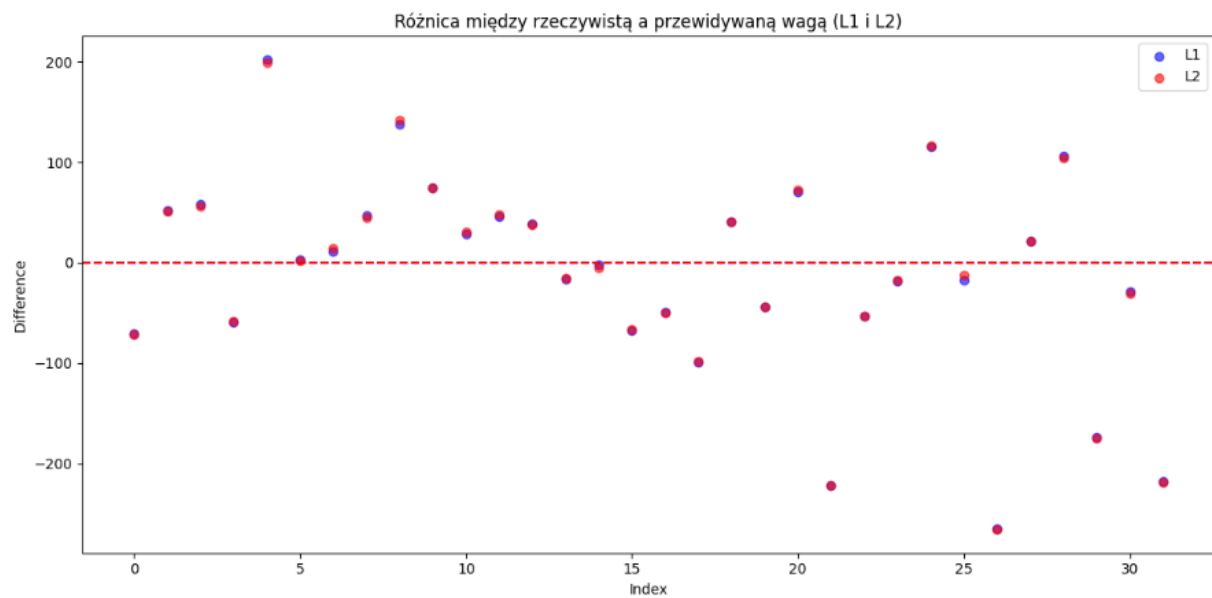
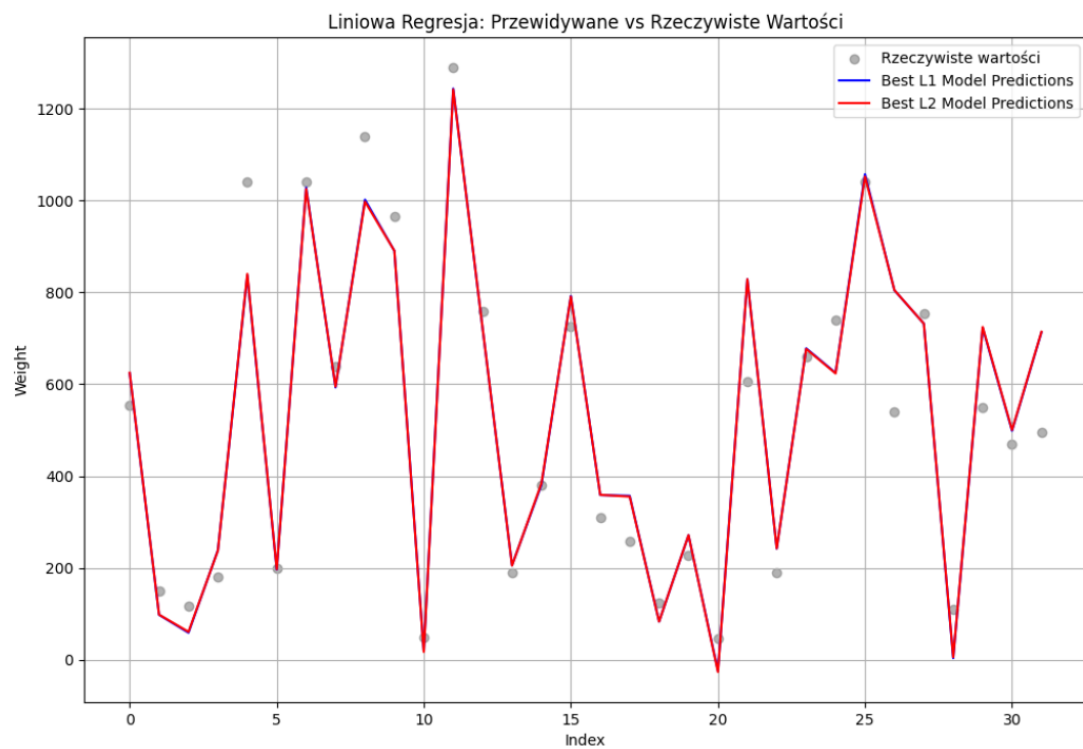
Dla modelu L2 około 0.9583

```
Best L1 Model on Validation Set: 4659.694659923203 0.9592494565543151
Best L2 Model on Validation Set: 4763.93921638157 0.9583378040455212
```

Wagi modeli:

```
Best L1 Model Weights: [ -8.59302161  7.35517866 -8.68746425 22.38061526 32.02581118
120.31654482 123.21696281 117.85044223 -12.93705211 19.38760007
 8.09601153 -60.79740196 -4.72062633 65.13355685 1.4304643 ]
Best L2 Model Weights: [ -8.94084427  7.04271464 -8.86639322 24.7686628 36.52069376
117.3398833 119.61895149 114.28624131 -13.380718 18.17352339
 7.49968706 -55.81425148 -5.45386989 63.69595674 1.14634278 ]
```

Test modeli na zbiorze testującym



Dla modeli na zbiorze testującym, zmianie uległy wskaźniki MSE i R2, wyniki które zostały uzyskane to:

Dla modelu L1:

MSE = 10577

R2 = 0.9145

Dla modelu L2:

MSE = 10602

R2 = 0.9143

```
L1 Test MSE: 10577.983009353797, R2: 0.9144785627784167
L2 Test MSE: 10602.604223249855, R2: 0.9142795038844225
```

Wnioski i analiza wyników

Model wykorzystujący regularyzację L1:

```
Best L1 Model Weights: [ -8.59302161  7.35517866 -8.68746425  22.38061526  32.02581118
120.31654482 123.21696281 117.85044223 -12.93705211  19.38760007
 8.09601153 -60.79740196 -4.72062633  65.13355685  1.4304643 ]
```

Warto zwrócić uwagę, na wartości wag, dla poszczególnych cech. Im wyższa wartość wagi, tym większe znaczenie dla klasyfikacji zyskuje dany parametr obiektu. Widzimy także skrajne wartości wag jak np. 123 czy -60, co może świadczyć o wpływie regularyzacji L1 na ich wartości. Z drugiej strony waga np. 1.43 świadczy o znikomym wpływie tej cechy na regularyzację. Możemy również zauważyć przedział wag odpowiadający za pole „Species”, jego wagi wahają się od -8.68 do 120, co sugeruje, że regularyzacja wybierała najbardziej znaczące cechy wpływające na predykcje.

Parametr MSE = 10577, świadczy o niewielkim błędzie i dobrej precyzji w przewidywaniu wagi ryby. Potwierdza to współczynnik R2 = 0.9143 na zbiorze testującym.

Model wykorzystujący regularyzację L2:

```
Best L2 Model Weights: [ -8.94084427  7.04271464 -8.86639322  24.7686628  36.52069376
117.3398833 119.61895149 114.28624131 -13.380718  18.17352339
 7.49968706 -55.81425148 -5.45386989  63.69595674  1.14634278]
```

Niektóre z wag różnią się w porównaniu z modelem L1, wskazuje to na wybór innych cech na te bardziej znaczące w modelu L2. Parametr MSE = 10602, delikatnie wyższy niż w przypadku L1, również świadczy o stosunkowo dobrym dopasowaniu modelu do przewidywania wagi ryb. R2 również jest lekko niższy w przypadku modelu L2.

Podsumowanie

Model L1 osiągnął minimalnie lepszą jakość przewidywania wagi ryb, o czym świadczy wyższy współczynnik R2 i niższy MSE. Może być spowodowane to niższym parametrem `learning_rate` w przypadku modelu L1.