

Softmax Regression

Procedure 方法

Softmax Regression adalah model machine learning yang digunakan untuk klasifikasi multiclass. Model ini menggunakan fungsi logistik untuk memodelkan interaksi dari fitur ke target dan menghasilkan nilai probabilitas. Loss function yang digunakan adalah cross-entropy loss.

1. Inisialisasi

Model di-inisialisasi dengan parameter learning rate, penalty, max_iter, mu, dan verbose. Weights dan bias di-inisialisasi dengan 0.

2. One hot encode variable target

Matriks onehot ini digunakan untuk menghitung cross-entropy loss dan gradiennya.

3. Prediksi y

Prediksi dilakukan dengan softmax dari $X @ \text{weights} + \text{bias}$, dimana $@$ adalah matrix multiplication.

4. Menghitung gradient

Gradient dihitung dari error, yaitu perbedaan y_{true} dengan prediksi, menggunakan hasil turunan parsial cross-entropy loss.

5. Gradient Descent

Weights dan biases model dikurang dengan $\text{learning_rate} * \text{gradient}$ masing-masing. Langkah 3-5 diulangi hingga max_iter

VS Sklearn

Hasil di notebook DoE menghasilkan nilai F1 yang sedikit lebih rendah. Hal tersebut mungkin dikarenakan sklearn menggunakan solver yang lebih pintar yang dapat mengubah learning rate secara dinamis.

Potential Improvements

- Solver yang lebih pintar dengan learning rate adjustment
- Hyperparameter tuning dapat mencari kombinasi yang optimal
- Early stopping dapat menjadi fitur untuk menghentikan training jika validation loss tidak menurun