

Softmax regression merupakan pengembangan dari logistic regression. Pada dasarnya, softmax regression bekerja dengan melakukan optimisasi dengan metode gradient descent untuk meminimasi cross-entropy loss function. Besaran optimasi bergantung pada learning rate. Algoritma akan melakukan iterasi sebanyak `num_iter` dan pada tiap iterasinya mengurangi besaran parameter θ dari tiap feature.

Jika dibandingkan dengan implementasi scikit-learn, implementasi logistic regression dari scratch bekerja serupa. Keduanya cukup baik memprediksi kelas, dengan perbedaan F1 score 0.05. Hal ini mungkin disebabkan dengan learning-rate yang sudah tetap diberikan oleh pemrogram saat memanggil fungsi softmax regression yang dibuat sendiri.

Agar lebih baik, dapat dilakukan hyperparameter tuning. Namun, karena softmax regression dijadikan baseline model, hal tersebut tidak dilakukan.