HW1 & HW2

Project

MLE, Loss function, Gradient descent

Gradient of Loss functions

Calculus Review

Machine Learning pipeline

What we've done so far

1. Collect some data
2. Use Perceptron
3. Evaluate the estimate

More principled

1. Collect some data
2. Choose a model
3. Choose a loss function
4. Choose an optimization procedure
5. Evaluate the estimate

Perceptron

Perceptron = Linear model + Hinge loss + Stochastic gradient descent

(model) + (loss function) + (optimization method)

Loss functions

- Hinge Loss (Perceptron)
- Logistic Loss (Logistic Regression)
- Squared error Loss (Linear Regression)
- and more ...

Loss functions

Ignore “1 - “ for Perceptron.

Hinge Loss: $\ell_i(w) = \max\{0, 1 - y_i x_i^T w\}$

Logistic Loss: $\ell_i(w) = \log(1 + \exp(-y_i\, x_i^T w))$

Squared error Loss: $\ell_i(w) = (y_i - x_i^T w)^2$