Logistic Regression with Regularization

LF

$$J(\theta_0,\theta_1) = -y \log\left(h_\theta(x)\right) - (1-y_i)\log\left(1-h_\theta(x)\right)$$

$$h_\theta(x) = \frac{1}{1+e^{-(\theta_0+\theta_1 x_1)}}$$

Ridge

$$J(\theta_0,\theta_1) = -y \log\left(h_\theta(x)\right) - (1-y_i)\log\left(1-h_\theta(x)\right) + L_2 \text{ regularization}$$

L_2 regularization ⇓ reduce overfitting

Lasso

$$J(\theta_0,\theta_1) = -y \log\left(h_\theta(x)\right) - (1-y_i)\log\left(1-h_\theta(x)\right) + L_1 \text{ regularization}$$

L_1 regularization ⇓ feature selection

Elastic net

$$J(\theta_0,\theta_1) = -y \log\left(h_\theta(x)\right) - (1-y_i)\log\left(1-h_\theta(x)\right) + L_1 + L_2$$

L_1 Reg. L_2 regular.

L_2 regularization = Reduce overfitting

$$J(\theta_0,\theta_1) = -y_i \log\left(h_\theta(x)\right) - (1-y)\log\left(1-h_\theta(x)\right) + \lambda \sum_{i=1}^{n} (\text{slope})^2$$

L_1 regularization ⇒ Feature selection

$$J(\theta_0,\theta_1) = -y \log\left(h_\theta(x)\right) - (1-y)\log\left(1-h_\theta(x)\right) + \lambda \sum_{i=1}^{n} |\text{slope}|$$

Elastic Net ($L_1 + L_2$)

$$J(\theta_0,\theta_1) = -y_i \log\left(h_\theta(x)\right) - (1-y_i)\log\left(1-h_\theta(x)\right) + \lambda_1 \underbrace{\sum_{i=1}^{n} (\text{slope})^2}_{L_2} + \lambda_2 \underbrace{\sum_{i=1}^{n} |\text{slope}|}_{L_1}$$

$$C = \frac{1}{k}$$