

Multiclass Perceptron + Logistic Regression $y = \{1, 2, 3\}$

Multiclass Perceptron

for t in epochs

for i in data diff weights

$$y_{\text{pred}} = \underset{y}{\text{argmax}} \bar{w}_y^T f(\bar{x})$$

if $y_{\text{pred}} \neq y^{(i)}$

$$\bar{w}_{y_{\text{pred}}} \leftarrow \bar{w}_{y_{\text{pred}}} - \alpha f(\bar{x})$$

$$\bar{w}_{y^{(i)}} \leftarrow \bar{w}_{y^{(i)}} + \alpha f(\bar{x})$$

DF: $\bar{w} \leftarrow \bar{w} + \alpha f(\bar{x}, y^{(i)})$

Ex: $[1, 1, 0]$ $y=1$ health
too many drug trials, too few patients

$[1, 0, 1]$ $y=2$ sports

baseball players taking drugs

Assume "default" pred is $y=3$

$$\left[\underbrace{0 \ 0 \ 0}_{\bar{w}_{y_1}} \ \{ \ \underbrace{0 \ 0 \ 0}_{\bar{w}_{y_2}} \ \{ \ \underbrace{0 \ 0 \ 0}_{\bar{w}_{y_3}} \} \right]$$

Update on $[1, 1, 0], 1$

$$y_{\text{pred}} = 3 \quad y^{(i)} = 1 \quad \alpha = 1$$

$$[1, 1, 0] \ \{ \ 0 \ 0 \ 0 \ \{ \ -1 \ -1 \ 0 \}$$

$$y_{\text{pred}} = 1 \quad y^{(i)} = 2$$

$$[0, 1, -1] \ \{ \ 1, 0, 1 \ \{ \ -1 \ -1 \ 0 \}$$

Multiclass LR

DF

$$P(y = \hat{y} | \bar{x}) = \frac{e^{\bar{w}^T f(\bar{x}, \hat{y})}}{\sum_{y' \in \mathcal{Y}} e^{\bar{w}^T f(\bar{x}, y')}}$$

in binary LR:

$$y = +1: e^{\bar{w}^T f(\bar{x})}$$

$$y = -1: 1 = e^{\bar{w}^T \bar{0}}$$

↑
must sum to
1 over \mathcal{Y}

$$\frac{\partial}{\partial \bar{w}} \text{loss}(\bar{x}^{(i)}, y^{(i)}, \bar{w}) = -f(\bar{x}, y^{(i)}) + \sum_{y' \in \mathcal{Y}} P(y' | \bar{x}) f(\bar{x}, y')$$

$$P(y^{(i)} | \bar{x}) \approx 1: -f(\bar{x}, y^{(i)}) + f(\bar{x}, y^{(i)}) \approx 0$$

$$P(y_{\text{bad}} | \bar{x}) \approx 1: -f(\bar{x}, y^{(i)}) + f(\bar{x}, y_{\text{bad}})$$

$$\text{SGD: } -\alpha: \frac{\partial}{\partial \bar{w}}$$