Introduction to Statistical Learning

Omid Safarzadeh

January 19, 2022

(ロ) (리) (리) (본) (본) (본) (인()

Table of Contents

- Activation Functions
 - Exponential Linear Unit (ELU)
 - Exponential activation function
 - Gaussian error linear unit (GELU)
 - Hard sigmoid
 - Rectified Linear Unit (ReLU)
 - Scaled Exponential Linear Unit (SELU)
 - Sigmoid
 - Softplus
 - Softsign
 - Swish
 - Hyperbolic Tangent
- 2 Loss Functions
 - Mean Absolute Error (MAE)
 - Mean Absolute Percentage Error (MAPE)
 - Mean Squared Error (MSE)
 - Mean Squared Error (MSE)
 - Indicator function



Exponential Linear Unit (ELU)

$$ELU(x) = \begin{cases} x & x \ge 0 \\ \alpha(e^x - 1) & x < 0 \end{cases}$$

$$ELU'(x) = \begin{cases} 1 & x < 0 \\ \alpha e^x & x < 0 \end{cases}$$

Omid Safarzadeh Inscribertor to Servisited Learning January 19, 2022 3

Exponential activation function

$$exp(x) = exp'(x) = x$$

Omid Safarzadeh Introdustron to Stensorer Ustrania January 19, 2022

Gaussian error linear unit (GELU)

$$GELU(x) = xP(X \le x) = x\Phi(x)$$

$$\approx 0.5x(1 + \tanh\left[\sqrt{2/\pi}(x + 0.044715x^3)\right]$$

$$GELU'(x) \approx 0.5(\tanh\left[\sqrt{2/\pi}(0.044715x^3 + x)\right] + 1)$$

$$+ \frac{1}{\sqrt{2\pi}}[x(0.134145x^2 + 1) sech^2[\sqrt{2/\pi}(0.044715x^3)]$$

Omid Safarzadeh Involución to Statistical Iscarding January 19, 2022 5

Hard sigmoid

$$f(x) = \min(1, \max(0, \frac{(2x+5)}{10}))$$

$$f'(x) = \begin{cases} 0 & \text{if } |x| > 2.5 \\ 0.2 & \text{otherwise} \end{cases}$$

Rectified Linear Unit (ReLU)

$$ReLU(x) = \max(0, x)$$

$$ReLU'(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{otherwise} \end{cases}$$

Omid Safarzadeh Invodustion to Statistical Security January 19, 2022

Scaled Exponential Linear Unit (SELU)

$$SELU(x) = \begin{cases} x & x > 0 \\ \alpha(e^{x} - 1) & x \le 0 \end{cases}$$
$$SELU'(x) = \begin{cases} 1 & x > 0 \\ \alpha e^{x} & x < 0 \end{cases}$$

Sigmoid

$$Sigmoid(x) = rac{1}{1 + e^{-x}}$$
 $Sigmoid'(x) = rac{e^{x}}{(1 + e^{x})^{2}}$

$$Sigmoid'(x) = \frac{e^{x}}{(1+e^{x})^{2}}$$



Omid Safarzadeh January 19, 2022

Softplus

$$Softplus(x) = \log(e^x + 1)$$

 $Softplus'(x) = \frac{e^x}{1 + e^x}$



Omid Safarzadeh Universität va Sieuringung January 19, 2022

Softsign

$$Softsign(x) = \frac{x}{|x|+1}$$

$$Softsign'(x) = \frac{1}{(|x|+1)^2}$$



Swish

Swish(x) = x * Sigmoid(x) =
$$\frac{x}{1 + e^{-x}}$$

Swish'(x) = $\frac{1}{1 + e^{-x}} + \frac{xe^{-x}}{(1 + e^{-x})^2}$



Hyperbolic Tangent

$$\tanh(x) = \frac{\sinh(x)}{\cosh x} = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}}$$
$$\tanh'(x) = (\frac{2e^{x}}{1 + e^{2x}})^{2}$$

Mean Absolute Error (MAE)

$$MAE = \frac{1}{n} \sum_{i=0}^{n} |y_{-}\hat{y}_{i}|$$

Where \hat{y} is the predicted value



Mean Absolute Percentage Error (MAPE)

$$MAPE = \frac{100\%}{n} \sum_{i=0}^{n} |\frac{y_i - \hat{y}_i}{y_i}|$$

Where \hat{y} is the predicted value

(ロ) (個) (差) (差) (差) のQ()

Mean Squared Error (MSE)

$$MSE = \frac{1}{n} \sum_{i=0}^{n} (y - \hat{y}_i)^2$$

Where \hat{y} is the predicted value



Mean Squared Logarithmic Error (MSLE)

$$MSE = \frac{1}{n} \sum_{i=0}^{n} (\log(y_i + 1) - \log(\hat{y}_i + 1))^2$$

Where \hat{y} is the predicted value



Omid Safarzadeh https://doi.org/10.1001/j.j.com/pg. January 19, 2022

Indicator function

$$I(X \in A) := \begin{cases} 1 & \text{if } X \in A \\ 0 & \text{otherwise} \end{cases}$$

• $E[I(X \in A)] = P(X \in A)$

