$x \in R^d$
$y \in [C]$
$f : R^d \to [C]$
$C = 10$
$C = 26$
Step 1
$\{\pm 1\}$
$\{1, 2\}$

$$f(x) = \{ 1\ if w^T x 0 2 otherwise$$

$$f(x) = \{ 1\ if w_1^T x w_2^T x 2 if w_1^T x < w_2^T x =_{k=1,2} w_k^T x$$

w_1, w_2
$w = w_1, w_2$
$w_k^T x$
score
k
Step 2
w_3
$f(x) =_{1k3} w_k^T x.$

Step 3
$\mathcal{F} = \{f(x) =_{k \in [C]} w_k^T x : w_k \in R^d\} = \{f(x) =_{k \in [C]} (Wx)_k : W \in R^{C \times d}\}.$

$w = w_1, w_2$

$$P(y = 1 \mid x; w) = \sigma(w^T x) = \frac{1}{1 + e^{-w^T x}} = \frac{e^{w_1^T x}}{e^{w_1^T x} + e^{w_2^T x}} \propto e^{w_1^T x}.$$

$$P(y = 2 \mid x; w) \propto e^{w_2^T x}.$$

w_k
C

$$P(Y = k \mid x; W) = \frac{\exp(w_k^T x)}{\sum_{j \in [C]} \exp(w_j^T x)} \propto \exp(w_k^T) x.$$

$w_k^T x \mapsto P(y = k \mid x; w)$
soft-max func-tion
$y_1, \ldots, y_n$
$x_1, \ldots, x_n$

$$P(W) = \prod_{i=1}^{n} P(y_i \mid x_i; W) = \prod_{i=1}^{n} \frac{\exp w_{y_i}^T x_i}{\sum_{k \in [C]} \exp(w_k^T x_i)}.$$

neg-a-tive
mul-ti-class lo-gis-tic loss

$$F(W) = \sum_{u=1}^{n} \log\left(\frac{\sum_{k \in [C]} \exp(w_k^T x_i)}{\exp(w_{y_i}^T x_i)}\right) = \sum_{k=1}^{n} \log\left(1 + \sum_{k \neq y_i} \exp((w_k - w_{y_I})^T x_i\right) .*$$

$0-1$
mis-clas-si-fi-ca-