Video 10

Sunday, October 4, 2020   3:33 PM

Linear classifier:

$$\boxed{z = \vec{w}^T \vec{x}} + \cancel{b}$$

weight   input   bias

b can be eliminated if we set the first dimension of $\vec{x}$ to be a dummy variable.

$$y = \begin{cases} 1 & \text{if } z \geq 0 \\ \cancel{0} & \text{if } z < 0 \end{cases}$$

-1 ← prediction model makes for negative examples. (notation specific to the perceptron learning rule)

How to check whether prediction is correct?

$\boxed{y^{(i)} = t^{(i)}}$: when $x^{(i)}$ lies exactly on the decision boundary, it will be classified as positive

Instead, use $\boxed{z^{(i)} t^{(i)} > 0}$

If $t^{(i)} = -1$
$$\begin{cases} \text{if } z^{(i)} > 0, & z^{(i)} t^{(i)} < 0, & y^{(i)} \neq t^{(i)} \\ \text{if } z^{(i)} < 0, & z^{(i)} t^{(i)} > 0, & y^{(i)} = t^{(i)} \\ \text{if } z^{(i)} = 0, & z^{(i)} t^{(i)} = 0, & y^{(i)} \neq t^{(i)} \end{cases}$$

If $t^{(i)} = 1$
$$\begin{cases} \text{if } z^{(i)} > 0, & z^{(i)} t^{(i)} > 0, & y^{(i)} = t^{(i)} \\ \text{if } z^{(i)} < 0, & z^{(i)} t^{(i)} < 0, & y^{(i)} \neq t^{(i)} \\ \text{if } z^{(i)} = 0, & z^{(i)} t^{(i)} = 0, & y^{(i)} \neq t^{(i)} \end{cases}$$

How do we update the linear model's weights?

If $t = 1$ $(t = -1)$ and $z = \vec{w}^T \vec{x} > 0$ $(z < 0)$, we don't make any updates.

If $t = 1$ $(t = -1)$ and $z < 0$ $(z > 0)$, then $y = -1$ $(y = 1)$ is wrong (we need z to be larger)

Update: $\vec{w}' \longleftarrow \vec{w} + \vec{x}$ $\quad (\vec{w}' \longleftarrow \vec{w} - \vec{x})$

Justification:
$$\vec{w}'^T \vec{x} = (\vec{w} + \vec{x})^T x$$
$$= \vec{w}^T x + x^T x$$
$$= \vec{w}^T x + \underbrace{\|x\|^2}_{\geq 0}$$

$$\Rightarrow \underbrace{\vec{w}'^T x}_{z'} \geq \underbrace{\vec{w}^T x}_{z}$$

z becomes larger.

$\underbrace{\vec{w}'^T x \leq \vec{w}^T x}$ z becomes smaller

Perceptron algorithm   targets $t^{(i)} \in \{-1, 1\}$ rather than $\{0, 1\}$ as is the case in the rest of the course

For each training example $(\vec{x}^{(i)}, t^{(i)})$:   input   target

$$z^{(i)} \longleftarrow \vec{w}^T \vec{x}^{(i)}$$

If $z^{(i)} t^{(i)} \leq 0$:
$$\vec{w} \longleftarrow \vec{w} + t^{(i)} \vec{x}^{(i)}$$
} Perceptron learning rule

$t^{(i)}$ tells us the sign which we have to multiply $\vec{x}^{(i)}$ by to get the corresponding update.

If $t^{(i)} = 1$ we need a plus sign

If $t^{(i)} = -1$ we need a minus sign

Stop when the weights $\vec{w}$ were not updated during the last epoch

epoch: one pass through the entire training set

Perceptron is guaranteed to find a feasible solution after a finite number of steps if the problem is feasible