One-Hot-Categorical

Weather (W) → Cloudy, Rainy, Sunny

→ one-hot encoding

$W \in \{C, R, S\}$
$\quad\quad\;\; 0 \;\; 1 \;\; 2$

$W \in \{C, R, S\}$
$[1,0,0]^T \quad [0,1,0]^T \quad [0,0,1]^T$

encoding as a one-hot encoding

$\rightarrow \cancel{p(W)} \rightarrow p(\underline{W})$

↑ Random variable as a vector

We have to save:

$$p(\underline{W} = [1,0,0]^T) = \theta_0 \quad \text{Cloudy}$$
$$p(\underline{W} = [0,1,0]^T) = \theta_1 \quad \text{Rainy}$$
$$p(\underline{W} = [0,0,1]^T) = \theta_2 \quad \text{Sunny}$$

$$\underline{\theta} = \begin{bmatrix} \theta_0 \\ \theta_1 \\ \theta_2 \end{bmatrix}$$

Pmf:

$$p(\underline{W}) = \underline{\theta}^T \underline{W} \quad (= \underline{\theta} \cdot \underline{W})$$

Example

$$\underline{\theta} = \begin{bmatrix} 0.2 \\ 0.3 \\ 0.5 \end{bmatrix}$$

$$p(\underline{W} = \underline{R}) = p(\underline{W} = [0,1,0]^T)$$
$$= \begin{bmatrix} 0.2 \\ 0.3 \\ 0.5 \end{bmatrix}^T \cdot \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$
$$= 0.2 \cdot 0 + 0.3 \cdot 1 + 0.5 \cdot 0$$
$$= \underline{\underline{0.3}}$$