INSEA

33

Machine learning classique

Intro to neural nets

En pratique, ce modèle ne fonctionne pas pour ces données complexes. Pourquoi à votre avis ?

Idée:

$$
\begin{aligned}
z_1 &\stackrel{\text{def}}{=} h(\mathbf{x}^\top \boldsymbol{\beta}^{\mathbf{1}} + \alpha_1) \\
&\vdots \\
z_p &\stackrel{\text{def}}{=} h(\mathbf{x}^\top \boldsymbol{\beta}^{\boldsymbol{p}} + \alpha_p)
\end{aligned}
$$

Linéarités

$$z_1 \stackrel{\text{def}}{=} \mathbf{x}^\top \boldsymbol{\beta}^{\mathbf{1}} + \alpha_1$$

$$\vdots$$

$$z_p \stackrel{\text{def}}{=} \mathbf{x}^\top \boldsymbol{\beta}^{\boldsymbol{p}} + \alpha_p$$

non-linéarité

$$\max(z_1, \ldots, z_p) \longrightarrow \text{sigmoid}$$

1. On n'utilise qu'une seule non-linéarité

2. Elle est fixée par la fonction max: on ne l’apprend pas

1. Appliquer plusieurs non-linéarités h plus tôt

2. Combiner les z_j linéairement avec w_j à optimiser

$$\sum_{j=1}^{p} \omega_j z_j + \omega_0$$

sigmoid

Il faudrait donc: utiliser plusieurs non-linéarités simples + les combiner pour apprendre des fonctions non-linéaires complexes

Linéarités

$$z_1 \stackrel{\text{def}}{=} \mathbf{x}^\top \boldsymbol{\beta}^{\mathbf{1}} + \alpha_1$$

$$\vdots$$

$$z_p \stackrel{\text{def}}{=} \mathbf{x}^\top \boldsymbol{\beta}^{\boldsymbol{p}} + \alpha_p$$

non-linéarité

$$\max(z_1, \ldots, z_p) \longrightarrow \text{sigmoid}$$

En pratique, ce modèle ne fonctionne pas pour ces données complexes. Pourquoi à votre avis ?

1. On n'utilise qu'une seule non-linéarité
2. Elle est fixée par la fonction max: on ne l'apprend pas

Il faudrait donc: utiliser plusieurs non-linéarités simples + les combiner pour apprendre des fonctions non-linéaires complexes

Idée:

1. Appliquer plusieurs non-linéarités h plus tôt
2. Combiner les z_j linéairement avec w_j à optimiser

$$z_1 \stackrel{\text{def}}{=} h(\mathbf{x}^\top \boldsymbol{\beta}^{\mathbf{1}} + \alpha_1)$$

$$\vdots$$

$$z_p \stackrel{\text{def}}{=} h(\mathbf{x}^\top \boldsymbol{\beta}^{\boldsymbol{p}} + \alpha_p)$$

$$\sum_{j=1}^{p} \omega_j z_j + \omega_0$$

$$\downarrow$$

sigmoid

INSEA

$$z_1 \stackrel{\text{def}}{=} h(\mathbf{x}^\top \boldsymbol{\beta}^{\mathbf{1}} + \alpha_1)$$

$$\vdots$$

$$z_p \stackrel{\text{def}}{=} h(\mathbf{x}^\top \boldsymbol{\beta}^{\boldsymbol{p}} + \alpha_p)$$

$$\longrightarrow \text{sigmoid}(\sum_{j=1}^{p} \omega_j z_j + \omega_0)$$

INSEA