EE-311—Apprentissage et intelligence artificielle

4. Complément Labo Machines à vecteurs de support

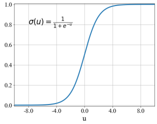
Michael Liebling

https://moodle.epfl.ch/course/view.php?id=16090

14 mars 2025 (compilé le 14 mars 2025)

Régression Logistique : comment afficher l'équation de la frontière

Fonction logistique



(B) Fonction logistique.

Azencott

$$\sigma(u) = \frac{1}{1+e^{-u}} = \frac{e^u}{1+e^u}$$

Régression logistique quand $\vec{x} = (x_1, x_2)$, fonction de décision :

$$g(x_1, x_2) = \sigma(\beta_0 + \beta_1 x_1 + \beta_2 x_2)$$

On fixe le seuil en spécifiant la probabilité p d'être dans la classe 1 :

$$p = \sigma(\beta_0 + \beta_1 x_1 + \beta_2 x_2) \quad \Big| \quad \mathrm{logit}(\cdot)$$

$$\mathrm{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2, \quad \mathrm{logit}(p) = \log \frac{p}{1-p}$$

$$\underbrace{x_2}_{y} = \underbrace{-\frac{\beta_1}{\beta_2} x_1}_{ax} + \underbrace{\frac{\mathrm{logit}(p) - \beta_0}{\beta_2}}_{b}$$

En utilisant les notations `sklearn`

Équation de la frontière de décision sous forme "$y = mx + h$"

$$\underbrace{x_2}_{y} = \underbrace{-\frac{\beta_1}{\beta_2}x_1}_{mx} + \underbrace{\frac{\text{logit}(p) - \beta_0}{\beta_2}}_{h}$$

```
from sklearn.linear_model import LogisticRegression
classif = LogisticRegression().fit(X_train, y_train)
```

Les coefficients β_0 β_1 et β_2 peuvent être récupérés dans les variables suivantes de l'objet `classif` :

- β_0 : `classif.intercept_`
- β_1 : `classif.coef_[0,0]`
- β_2 : `classif.coef_[0,1]`

SVM : comment afficher l'équation de l'hyperplan H ?

L'hyperplan H est donné par l'équation :

$$\langle \vec{x}, \vec{w} \rangle + b = 0$$

$$w_1 x_1 + w_2 x_2 + b = 0$$

$$w_2 x_2 = -b - w_1 x_1$$

$$\underbrace{x_2}_{y} = \underbrace{-\frac{w_1}{w_2} x_1}_{mx} \underbrace{-\frac{b}{w_2}}_{+h}$$

```
from sklearn.svm import SVC
classif_svm = SVC(kernel='linear', C=10000).fit(X_train, y_train)
```

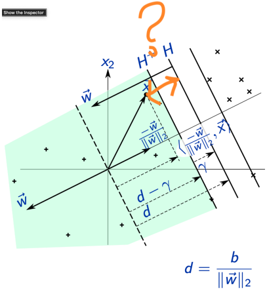
- b : `classif_svm.intercept_`
- w_1 : `classif_svm.coef_[0,0]`
- w_2 : `classif_svm.coef_[0,1]`

Calculer la distance d'un point à une droite

On veut la distance d'un point à une droite spécifiée par :

$$b + w_1 x_1 + w_2 x_2 = 0$$

Ce problème est équivalent à chercher la distance d_H d'un point arbitraire $\vec{x}$ à l'hyperplan H dans l'équation de la SVM :



$$d_H = \left| \underbrace{d}_{\frac{b}{\|\vec{w}\|}} - \left\langle \frac{-\vec{w}}{\|\vec{w}\|_2}, \vec{x} \right\rangle \right|$$

($=$ la distance d de l'hyperplan à l'origine moins la longeur de la projection du vecteur $\vec{x}$ sur la droite perpendiculaire à l'hyperplan)

Il s'ensuit :

$$d_H = \left| \frac{b}{\|\vec{w}\|} + \frac{w_1}{\|\vec{w}\|} x_1 + \frac{w_2}{\|\vec{w}\|} x_2 \right| = \left| \frac{\langle \vec{w}, \vec{x} \rangle + b}{\|\vec{w}\|_2} \right|$$