1.4 Apprentissage génératif

Un modèle génératif essaie d'abord d'apprendre comment les données sont générées en estimant $P(x|y)$, nous permettant ensuite d'estimer $P(y|x)$ par le biais du théorème de Bayes.

1.4.1 Gaussian Discriminant Analysis

❒ **Cadre** – Le Gaussian Discriminant Analysis suppose que y et $x|y=0$ et $x|y=1$ sont tels que :

$$\boxed{y \sim \textrm{Bernoulli}(\phi)}$$

$$\boxed{x|y=0 \sim \mathcal{N}(\mu_0,\Sigma)} \quad \textrm{et} \quad \boxed{x|y=1 \sim \mathcal{N}(\mu_1,\Sigma)}$$

❒ **Estimation** – Le tableau suivant récapitule les estimations que l'on a trouvées lors de la maximisation de la vraisemblance :

$\widehat{\phi}$	$\widehat{\mu_j} \quad (j=0,1)$	$\widehat{\Sigma}$
$\displaystyle\frac{1}{m}\sum_{i=1}^m 1_{\{y^{(i)}=1\}}$	$\displaystyle\frac{\sum_{i=1}^m 1_{\{y^{(i)}=j\}}x^{(i)}}{\sum_{i=1}^m 1_{\{y^{(i)}=j\}}}$	$\displaystyle\frac{1}{m}\sum_{i=1}^m (x^{(i)}-\mu_{y^{(i)}})(x^{(i)}-\mu_{y^{(i)}})^T$

1.4.2 Naive Bayes

❒ **Hypothèse** – Le modèle de Naive Bayes suppose que les caractéristiques de chaque point sont toutes indépendantes :

$$\boxed{P(x|y)=P(x_1,x_2,...|y)=P(x_1|y)P(x_2|y)...=\prod_{i=1}^n P(x_i|y)}$$

❒ **Solutions** – Maximiser la log vraisemblance donne les solutions suivantes, où $k \in \{0,1\}, l \in [\![1,L]\!]$

$$\boxed{P(y=k)=\frac{1}{m}\times\#\{j|y^{(j)}=k\}} \quad \textrm{et} \quad \boxed{P(x_i=l|y=k)=\frac{\#\{j|y^{(j)}=k \textrm{ et } x_i^{(j)}=l\}}{\#\{j|y^{(j)}=k\}}}$$

Remarque : Naive Bayes est couramment utilisé pour la classification de texte et pour la détection de spams.

1.5 Méthode à base d'arbres et d'ensembles

Ces méthodes peuvent être utilisées pour des problèmes de régression et de classification.

❒ **CART** – Les arbres de classification et de régression (en anglais *CART - Classification And Regression Trees*), aussi connus sous le nom d'arbres de décision, peuvent être représentés sous la forme d'arbres binaires. Ils ont l'avantage d'être très interprétables.