

Exercice Bayes Naïf

6 novembre 2018

L'objectif de cet exercice est :

- utiliser un classifieur naïf de Bayes sur un exemple réel simple,
- étudier le modèle lorsque l'assomption d'indépendance est fausse.

Partie I

1. Charger l'ensemble de données `iris` disponible dans le module `datasets` de `sklearn`.
2. Visualiser les données (afficher la taille de l'ensemble, afficher un exemple, afficher la longueur des pétales en fonction de leur largeur, etc.).
3. Séparer les données en deux ensemble (80% pour l'apprentissage et 20% pour le test). Attention à respecter la distribution des classes dans les deux sous-ensembles.
4. Entraîner un classifieur naïf de Bayes (fonction `GaussianNB` de `sklearn`) sur les données d'apprentissage et évaluer la précision du modèle sur les données de test (fonction `score` de l'objet `GaussianNB`).

Partie II

1. Générer un ensemble de données de taille 1000 avec la fonction `make_blobs` du module `datasets` de `sklearn`.
2. Séparer les données en deux ensemble (50% pour l'apprentissage et 50% pour le test). Attention à respecter la distribution des classes dans les deux sous-ensembles.
3. Entraîner un classifieur naïf de Bayes sur les données d'apprentissage et évaluer la précision sur les données de test.
4. Tracer les données de test :
 - avec comme couleur la classe réelle.
 - avec comme couleur la classe prédite.
5. Générer un ensemble de données de taille 1000 avec la fonction `make_moons` du module `datasets` de `sklearn`. Faire la même étude.
6. Commenter.