Module 1
Introduction à l'apprentissage machine

Sommaire

Dernière mise à jour le 3 janvier 2019

1.1 Introduction

L'**apprentissage machine** (aussi appelé apprentissage artificiel ou automatique, en anglais *machine learning*) est le « processus par lequel un ordinateur acquiert de nouvelles connaissances et améliore son mode de fonctionnement en tenant compte des résultats obtenus lors de traitements antérieurs »(Office québécois de la langue française, 2010). Ceci permet de regrouper « toutes les méthodes qui permettent de construire un modèle de la réalité à partir de données, soit en améliorant un modèle existant moins général, soit en créant un nouveau modèle représentatif de nouvelles données »[4]. Ces modèles servent souvent à prendre des décisions.

Il existe deux approches principales en apprentissage machine. La première est issue de l'intelligence artificielle *syntaxique* ou *symbolique*. Elle est fondée sur la modélisation du raisonnement logique et sur la représentation et la manipulation de la connaissance par des symboles formels. La deuxième est issue de l'intelligence artificielle statistique ; elle est qualifiée de statistique aussi parfois numérique parce que, souvent, la représentation et la manipulation de la connaissance sont sous une forme numérique [4].

Le cours INF 1421 s'intéresse à l'apprentissage machine statistique.

1.2 Techniques d'apprentissage machine

Les algorithmes d'apprentissage sont catégorisés selon les techniques d'apprentissage qu'ils emploient. Nous en citons l'apprentissage supervisé, l'apprentissage non supervisé, l'apprentissage par renforcement et l'apprentissage semi-supervisé.

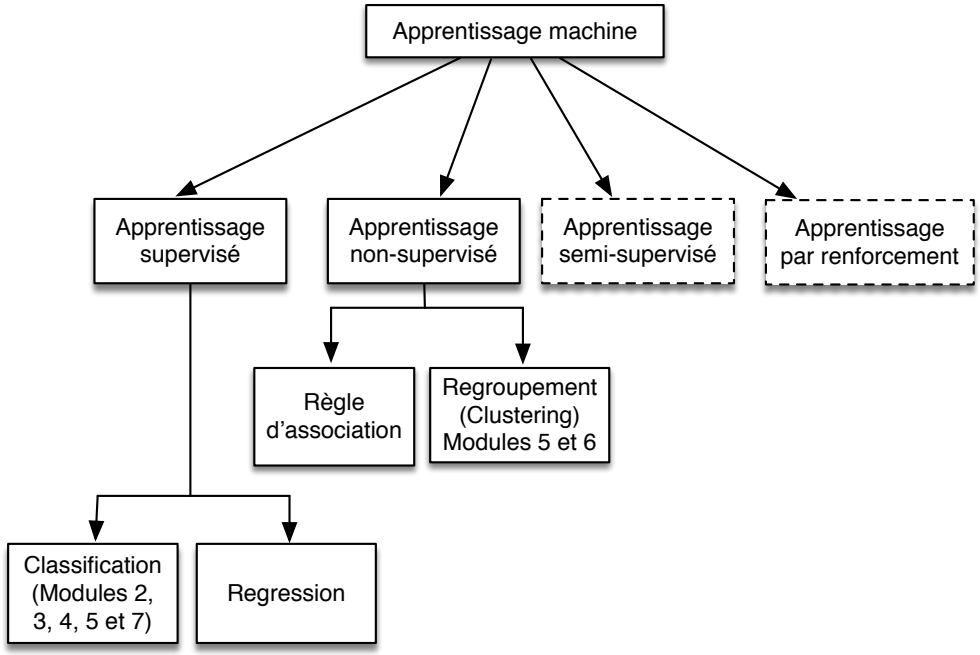

Figure 1.1: Techniques d'apprentissage machine.

Apprentissage supervisé

Dans le cas d'un apprentissage supervisé (en anglais *supervised learning*), le système observe des couples de types entrée-sortie et apprend une fonction (un modèle) qui permet d'aboutir à la sortie à partir de l'entrée. Cette phase est appelée *phase d'apprentissage* ou *d'entraînement*. C'est en ce sens que l'apprentissage est appelé supervisé, métaphore qui signifie qu'un professeur apprend au système la sortie à fournir pour chaque entrée. Les données d'entrée et les données de sortie correspondantes, aussi appelées classes, sont connues (aussi dites labellisées ou étiquetées). Telles que décrits précédemment, elles sont regroupées dans un ensemble de données appelées *données d'apprentissage* ou

d'entrainement qui se présentent sous la forme de couples $(x_i, y_i)_{1 \leq i \leq N}$ avec N le nombre d'échantillons.

Apprentissage non supervisé

Contrairement à l'apprentissage supervisé, dans le cas non supervisé les données de sortie ne sont pas connues. Le système apprend alors de lui-même à organiser les données ou à déterminer des structures dans les données. La tâche d'apprentissage la plus courante est le regroupement (*clustering* en anglais) qui consiste à regrouper les données d'entrées selon leurs caractéristiques communes. Ce type d'apprentissage est utilisé dans le but de visualiser ou explorer des données.

Apprentissage semi supervisé

L'apprentissage semi-supervisé se base sur un mélange de données étiquetées et non-étiquetées. Ceci permet, d'une part, d'améliorer la qualité de l'apprentissage, et d'autre part, de réduire le temps de préparation des données pour leur étiquetage.

Apprentissage par renforcement

L'apprentissage par renforcement se base sur des données d'entrée similaires à celles utilisées en apprentissage supervisé. Cependant dans ce cas, l'apprentissage est guidé par l'environnement sous la forme de récompenses (positive ou négative) calculées en fonction de l'erreur commise lors de l'apprentissage. En robotique, l'apprentissage par renforcement a permis de mettre au point des robots plus autonomes et adaptatifs que ceux existants.

Le cours INF1421 présente les deux techniques les plus utilisées en apprentissage à savoir l'apprentissage supervisé et l'apprentissage non supervisé.

1.3 Champs d'application de l'apprentissage machine

Les champs d'application de l'apprentissage machine peuvent être regroupés selon deux grands axes : la *reconnaissance de formes* et la *fouille de données*.

1.3.1 Reconnaissance de formes

La reconnaissance ou classification de formes (en anglais, *pattern recognition*) est « l'ensemble des techniques permettant à l'ordinateur de détecter la présence de formes visuelles ou auditives spécifiées, en comparant leurs caractéristiques avec celles de motifs de référence »(Office québécois de la langue française, 2010). Les domaines d'application de la reconnaissance des formes sont très variés. Ils regroupent la reconnaissance de textes imprimés (hors ligne) ou manuscrits (en ligne), la reconnaissance de signature, la reconnaissance vocale et la programmation de robots. Certaines de ces applications seront détaillées dans la section 1.3.3.

1.3.2 Forage de données

Le forage de données (en anglais, *data mining*) est « l'ensemble des techniques de recherche et d'analyse de données qui permet de dénicher des tendances ou des corrélations cachées parmi des masses de données, ou encore de détecter des informations stratégiques ou de découvrir de nouvelles connaissances, en s'appuyant sur des méthodes de traitement statistique »(Office québécois de la langue française, 2010). De ce fait, nous retrouvons toutes les applications qui utilisent l'informatique décisionnelle pour déterminer, par exemple, quels sont les critères qui regroupent les consommateurs de certaines régions ou bien quels sont les indicateurs qui ont entrainé le déclenchement de certaines maladies.

1.3.3 Exemples d'applications

L'apprentissage machine est un domaine de recherche multidisciplinaire qui permet de créer des systèmes non biologiques qui imitent les capacités des systèmes biologiques. Dans ce qui suit nous citerons quelques exemples :

Applications médicales

Le domaine médical est devenu parmi les principaux utilisateurs de l'apprentissage machine. La segmentation et l'annotation automatique de structures dans les images biomédicales utilisent des techniques d'apprentissage machine pour plusieurs applications clés dont l'aide au diagnostic, le suivi de structure anatomique et le suivi de pathologies. La figure 1.2 illustre la segmentation d'image de colonnes vertébrales pour l'aide au diagnostic et le suivi de la scoliose idiopathique [3].

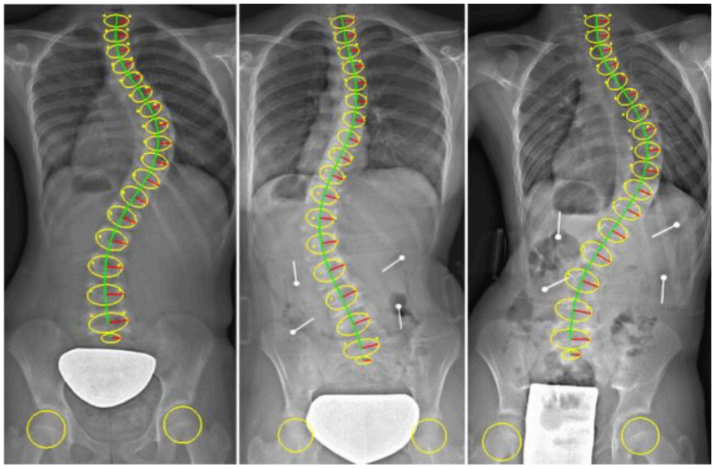

Figure 1.2: Segmentation de colonnes vertébrales basée sur des réseaux de neurones [3].

Reconnaissance de l'écriture manuscrite

Les systèmes de reconnaissance de l'écriture manuscrite ont pour but de traduire un texte écrit en un texte codé numériquement. La reconnaissance de l'écriture manuscrite regroupe deux types de systèmes : les systèmes hors-ligne et les systèmes en ligne.

Dans les systèmes de reconnaissance d'écriture hors-ligne ou statiques, l'écriture se présente sur un support classique tel que le papier. Après numérisation, on obtient une image. Dans ce cas, l'objectif est de traiter des documents de manière automatique. À l'opposé de ce mode d'acquisition, la reconnaissance de l'écriture en ligne s'effectue souvent au moment même où le scripteur écrit. L'acquisition est réalisée au moyen d'un stylet et d'une tablette électronique qui peut être assimilée à un papier électronique. L'information acquise correspond au suivi de la trajectoire de la pointe du stylet sur la tablette, laquelle est mémorisée sous forme de signaux dépendants du temps, c'est-à-dire une séquence de coordonnées de points ordonnées dans le temps $(x(t), y(t))$.

Il existe sur le marché plusieurs applications de reconnaissance d'écriture manuscrite hors-ligne (Exemple : GOCR, OCRopus et Tesseract) et en-ligne (Google Handwiting, Figure 1.3). Ces applications permettent la reconnaissance de chèque, la reconnaissance de codes postaux et la reconnaissance d'adresse.

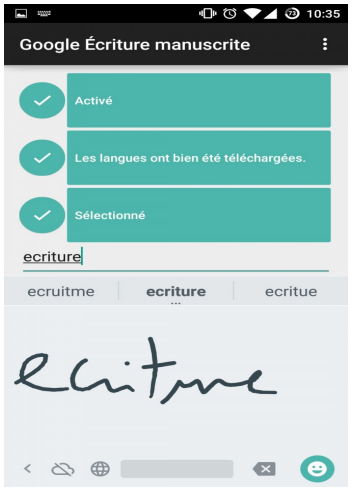

Figure 1.3: Google Handwiting.

Reconnaissance de visage

La reconnaissance de visage est un domaine de la vision par ordinateur qui consiste à reconnaître automatiquement une personne à partir d'une image de son visage. La reconnaissance de visage trouve son application dans de nombreuses applications en vidéosurveillance, biométrie, robotique, indexation d'images et de vidéos.
Il existe actuellement plusieurs applications commercialisées ou en cours de développement. À titre d'exemple, nous citons le système de reconnaissance faciale, DeepFace, développé par le réseau social Facekook dont la précision est à peine inférieure à ce que peut faire un humain [2]. Cette technologie, qui demeure encore à l'étape de projet, a été développée

Figure 1.4: Système de reconnaissance faciale pour la vidéosurveillance dans les aéroports. [1].

par trois chercheurs internes de Facebook et un professeur de l'université de Tel-Aviv. Le système de classification est capable de déterminer si deux photos contiennent le même visage avec un taux de réussite de 97,25 %.

Apprentissage de l'apprenant (*Learning analytics*)

Ce section d'application est relativement récent. Son objectif est de mettre en place des méthodes d'apprentissage personnalisées selon le besoin de l'apprenant. Grâce à l'apprentissage automatique, il est désormais possible de connaître l'apprenant : les algorithmes collectent et analysent les données d'apprentissage pertinentes afin d'adapter le processus de formation [5].

Foresterie

Les techniques d'apprentissage machine touche également de domaine de la foresterie. En effet ces techniques sont utilisées afin d'identifier et de quantifier les arbres dans les forêts,

de manières automatiques à partir d'une application installée sur un terminal mobile (Fig. 1.5). L'objectif étant d'aider les propriétaires fonciers, les groupes de conservation et les entreprises forestières à gérer leur inventaire et à préserver des habitats naturels précieux [6].

Figure 1.5: Application *Craze Branches* dans le domaine de la foresterie [6].

Applications militaires

Le secteur du militaire est aussi l'un des principaux utilisateurs de l'apprentissage machine. Nous citons, à titre d'exemple, BigDog est un robot quadrupède à l'allure d'un chien créé en 2005 par la société américaine Boston Dynamics en collaboration avec l'université Harvard (Figure 1.7). Il s'agit d'un robot destiné à accompagner les soldats en leur transportant du matériel dans des terrains trop irréguliers pour les véhicules. Le déplacement de BigDog est contrôlé par un système embarqué intelligent qui utilise un réseau de capteurs. Le lien BigDog permet de visualiser une vidéo de ce robot sur Youtube. Les drones utilisent également des techniques d'apprentissage machine pour pouvoir effectuer de nombreuses tâches complexes de manière autonome telle que le décollage, l'atterrissage, la navigation ou même la destruction de cible. (Figure 1.6).

Figure 1.6: Drones.

Figure 1.7: Le robot BigDog.

1.4 Principe de l'apprentissage machine

La définition présentée à la section 1.1 permet de ressortir les éléments importants suivants sur le principe de l'apprentissage machine : il s'agit d'un ensemble de méthodes qui permettent de construire un modèle de la réalité à partir de données, soit en améliorant un modèle existant moins général, soit en créant un nouveau modèle représentatif de nouvelles données. Les modèles servent souvent à prendre des décisions.

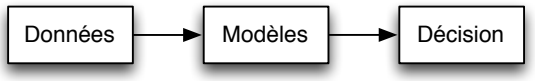

Figure 1.8: Principe de l'apprentissage machine.

Cette définition permet de schématiser le principe de l'apprentissage machine selon la figure 1.8 dans laquelle nous avons en intrant un ensemble de données qui permettent de construire un modèle. Ce modèle est utilisé pour la prise de décision.
Nous allons décrire dans la suite chacun de ces éléments puisqu'ils sont à la base du principe de l'apprentissage machine.

1.4.1 Données

Le mot **donnée** se définit de différentes manières dans la littérature, selon les domaines et les champs d'application. Mentionnons quelques exemples : Une donnée est...

- un **enregistrement** caractérisé par un ensemble de champs (terminologie des bases de données).
- un **individu** défini par un ensemble de caractéristiques ou de variables (terminologie issue de la statistique et du forage de données).
- une **forme** définie par un ensemble de caractéristiques ou de variables (terminologie issue de la reconnaissance de formes).
- une **instance** caractérisée par un ensemble d'attributs (terminologie orientée objet en informatique).

Étant donné que l'apprentissage machine trouve son application dans les axes de reconnaissance de forme et de forage de données, nous utiliserons souvent les terminologies issues de ces deux axes. Une donnée est donc un individu ou une forme caractérisée par

un ensemble de variables ou de caractéristiques.

La détermination du type de chaque variable est une étape nécessaire avant leur analyse. Cette étape permet de décider des méthodes d'analyse appropriées.

Variables qualitatives

Une variable est dite **qualitative** si ses valeurs ne sont pas mesurables. Les valeurs d'une variable qualitative sont appelées **modalités**. Le sexe, la profession, l'état matrimonial sont quelques exemples de variables qualitatives. Les modalités de la variable sexe sont `Féminin` et `Masculin`.

Une variable qualitative est dite **ordinale** si ses modalités suivent une relation d'ordre. Par exemple, une pathologie peut être « légère », « modérée » ou « sévère ». Ces valeurs peuvent être ordonnées : légère $<$ modérée $<$ sévère, donc on parle d'une variable ordinale.

Une variable qualitative est dite **nominale** si ses modalités ne sont pas ordonnées naturellement. Par exemple, dans une population de personnes actives, la profession est une variable nominale.

Variables quantitatives

Une variable quantitative est dite **discrète** si elle ne peut prendre que des valeurs qui peuvent être énumérées. Par exemple, le nombre de professeurs universitaires dans un département ou le nombre d'étudiants dans une classe.
La variable quantitative est dite **continue** si ses valeurs potentielles ne peuvent pas être énumérables (ou sont difficilement énumérables).

Variables binaires

Les variables binaires sont des variables quantitatives discrètes qui possèdent des propriétés particulières.
Nous distinguons deux types de données binaires :

- Données **symétriques** : une variable binaire est dite symétrique si ses deux modalités ont la même importance, c'est-à-dire si celles-ci peuvent être indifféremment codées par 0 ou 1. Par exemple, la variable sexe peut être codée par 0 ou 1 pour masculin (de même que pour féminin).
- Données **asymétriques** : une variable binaire est dite asymétrique si les deux modalités n'ont pas la même importance. Par exemple, le résultat d'un examen médical ne peut pas être codé par 0 (négatif) ou 1 (positif) en ce qui concerne son importance.

▶ **Exemple 1.1** Prenons l'ensemble de données décrites dans le tableau suivant :

Tableau 1.1: Un ensemble de données E décrivant la possibilité de jouer au tennis selon les conditions météorologiques.

Individus	Ciel	Temp. (en °C)	Humidité	Vent	Jouer
J1	Soleil	38	Élevée	Faible	Non
J2	Soleil	39	Élevée	Fort	Non
J3	Couvert	37.7	Élevée	Faible	Oui
J4	Pluie	20	Élevée	Faible	Oui
J5	Pluie	15	Normale	Faible	Oui
J6	Pluie	18	Normale	Fort	Non
J7	Couvert	18	Normale	Faible	Oui
J8	Soleil	21	Élevée	Faible	Non
J9	Soleil	15	Normale	Faible	Oui
J10	Pluie	21	Normale	Fort	Oui
J11	Soleil	23.5	Normale	Fort	Oui
J12	Couvert	23	Élevée	Fort	Oui
J13	Couvert	40	Normale	Faible	Oui
J14	Pluie	20	Élevée	Fort	Non

L'analyse de ce tableau permet de voir que nous disposons de 14 individus (14 jours : J1 à J14). Chaque individu est caractérisé par 4 variables qui correspondent aux conditions météorologiques (État du ciel, Température, Humidité dans l'air, Force du vent). La classe `Jouer` peut prendre les valeurs `Oui` ou `Non`.

La description des variables peut être résumée dans le tableau suivant :

Tableau 1.2: Description des variables.

Variables	Type	Unités/Modalitées
Ciel	Qualitative	Soleil, Couvert, Pluie
Temp.	Qunatitative continue	°C
Humidité	Qualitative	Élevée, Normale
Vent	Qualitative	Fort, Faible

La figure 1.9 représente la variation de température au cours des 14 jours et la figure 1.10 le diagramme circulaire des variables Ciel, Humidité et Vent.

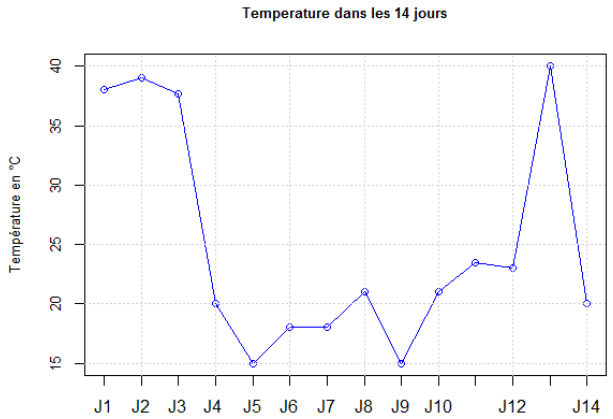

Figure 1.9: Température des 14 jours.

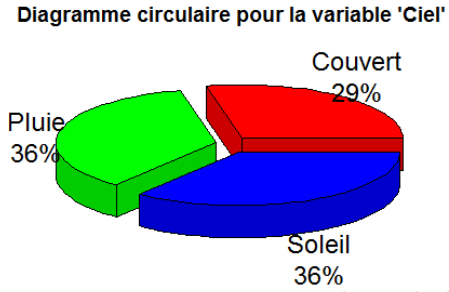

Figure 1.10: Diagramme circulaire de la variable `Ciel`.

1.4.2 Modèles

La détermination du modèle est une étape importante de l'apprentissage machine. Cette étape permet de déterminer une fonction qui permet de renvoyer une décision à partir des données d'entrée (données d'apprentissage). Il existe plusieurs modèles possibles dans la littérature parmi lesquels nous citons le modèle de Bayes (qui sera traité au Module 2) et le modèle neuronal (qui sera traité au Module 5).

1.4.3 Décisions

La décision à prendre dépend essentiellement de la problématique de l'apprentissage machine à résoudre. Par exemple, s'il s'agit d'un problème de reconnaissance de forme, la

décision consiste souvent à reconnaître la classe de la forme (plus de détails sont fournis dans le Module 2). Par contre s'il s'agit d'un problème de regroupement, la décision consistera à identifier les différents regroupements (plus de détails sont fournis dans le Module 6).

1.5 Étapes de l'apprentissage machine

1.5.1 Collecte des données

La collecte de données est une étape très importante en apprentissage machine. Cette étape est souvent coûteuse en temps et en argent. Les données collectées servent aussi bien pour le développement du modèle que pour sa validation (comme décrit dans la suite).

1.5.2 Extraction et sélection des caractéristiques

En apprentissage machine, nous travaillons rarement avec des données brutes. Une étape d'extraction des caractéristiques est souvent nécessaire. Elle consiste à projeter l'ensemble des données originales de dimension N dans un autre espace de dimension d via une transformation $\mathbf{A}$. Cet espace est appelé espace des caractéristiques.

Une caractéristique peut être par exemple : la couleur dans une image ou sa texture, la surface d'un objet, son périmètre, la longueur des axes majeur et mineur, les angles des axes majeur et mineur etc.

Dans certains cas, nous réalisons, aussi, une sélection des caractéristiques qui consiste à trouver les d' caractéristiques parmi les d possibles qui discriminent le mieux les formes à étudier.

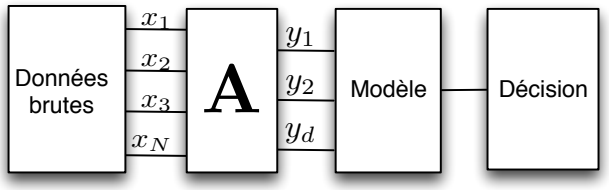

Figure 1.11: Extraction des caractéristiques.

La caractéristique est, souvent, notée par x s'il s'agit d'un scalaire et par $\mathbf{x} = (x_1, x_2, ..., x_d)^t$ s'il s'agit d'un vecteur de d variables ou caractéristiques.

1.5.3 Choix du modèle

Le choix du modèle dépend essentiellement des données à analyser et de la problématique en question. Les paramètres du modèle sont déterminés durant la phase d'apprentissage en utilisant l'algorithme qui lui est spécifique.

1.5.4 Entraînement, test et validation du modèle

L'ensemble des données considérées pour une analyse par apprentissage machine supervisé peut être réparti en trois sous-ensembles :

- Un ensemble d'entraînement qui est utilisé pour l'apprentissage (entraînement) des paramètres du modèle.
- Un ensemble de validation. Il s'agit d'un ensemble de données qui permet d'évaluer le modèle pendant la phase d'entraînement. Cette étape, appelée validation du modèle peut être omise en passant à la phase de test directement.
- Un ensemble de test. Une fois le modèle construit à partir de l'ensemble d'entraînement, le modèle est évalué en utilisant un ensemble de test : un ensemble d'échantillons n'ayant pas servi pour l'apprentissage.

Il existe différents types d'algorithmes pour réaliser le partage des données en données d'apprentissage et données de test. Parmi ces méthodes on trouve la validation par K-fold.

1.5.5 Validation par K-fold

Le principe de cette méthode de validation consiste à diviser l'échantillon original en K échantillons de même taille. Puis on prend un échantillon pour procéder à la validation. On répète le processus jusqu'à l'atteinte des K échantillons. En d'autres termes,

on divise l'ensemble des données original en K échantillons, puis on sélectionne un des K échantillons comme ensemble de tests et les $(K-1)$ autres échantillons constitueront l'ensemble d'apprentissages pour le développement du a conception du système de classification. À la fin, on prend la moyenne des résultats de validations pour avoir un seul résultat.
La répartition la plus communément utilisée entre ces ensembles est une proportion de 2/3 pour l'apprentissage et 1/3 pour le test, c'est-à-dire avec un $K = 3$. Dans le cas de grande taille de base de données, la validation 10-folds est plus appropriée.

▶ **Exemple 1.2** Soit un ensemble de données de 1200 échantillons répartis équitablement sur 4 classes. Nous désirons faire une partition 2/3 pour l'entraînement et 1/3 pour le test. Dans ce cas, nous devons considérer 800 échantillons pour l'entraînement et 400 échantillons pour le test (soit 100 échantillons par classe pour le test).

1.6 Évaluation d'un système de reconnaissance de forme

Lorsque l'apprentissage machine a pour objectif de développer un système de reconnaissance (classification) de formes, la validation se fait en calculant le taux de bonne classification et à travers la matrice de confusion.

1.6.1 Taux de classification

Soit S un ensemble d'échantillons d'apprentissage, et T un ensemble d'échantillons de test. L'estimation du taux de bonne classification est mesurée sur l'ensemble de test selon :

$$\tau = \frac{\text{nbr bien classifié(T)}}{\text{nbr(T)}}$$

C'est-à-dire le rapport entre le nombre d'échantillons de l'ensemble de test qui sont bien

classifiés par rapport au nombre total d'échantillons de la base de test.

Le taux de bonne classification est généralement donné en pourcentage. Il lui correspond une valeur complémentaire à 100 correspondant au taux d'erreur.

▶ **Exemple 1.3** Reprenons l'ensemble de données décrites précédemment. Admettons que la totalité des données a servi pour l'entraînement (J1 à J14) et que nous disposons d'un ensemble de données de test constitué de 5 individus (5 jours : J15 à J19). Soit Pred-Jouer la classe prédite par un classificateur quelconque.

Individus	Ciel	Temp. (en °C)	Humidité	Vent	Jouer	Pred-Jouer
J15	Soleil	38	Élevée	Faible	Non	Oui
J16	Soleil	39	Élevée	Fort	Non	Non
J17	Couvert	37.7	Élevée	Faible	Oui	Oui
J18	Pluie	20	Élevée	Faible	Oui	Non
J19	Pluie	15	Normale	Faible	Oui	Oui

L'analyse de ce tableau permet de voir que la classe réelle Jouer est égale à la classe prédite Pred-Jouer pour trois individus. Nous pouvons donc conclure que le taux de bonne classification est de 3/5 ou bien 60%. Nous pouvons aussi dire que le taux d'erreur est de (100 - 60) = 40%.

1.6.2 Matrice de confusion

La matrice de confusion est aussi connue sous les termes matrice d'erreur, tableau de contingence ou matrice d'erreur de classification.

C'est une matrice (ou un tableau) affichant les statistiques de la précision de classification et plus particulièrement les taux de classification par classes. Généralement, L'information des lignes (données horizontales) correspond aux classes réelles des formes. Quant aux colonnes (données verticales), elles contiennent l'information prédite résultant de la

classification.

Les valeurs de la diagonale de la matrice représentent le nombre de formes correctement classifié. La somme des valeurs par ligne correspond au nombre d'échantillons de test par classe. Le taux de classification par classes est donné par la valeur à la diagonale divisée par la somme des valeurs par ligne.

▶ **Exemple 1.4** Nous désirons dans cet exemple construire la matrice de confusion du tableau de données précédent sachant que `Jouer` est la classe réelle et `Pred-Jouer` est la classe prédite.

Individus	Ciel	Temp. (en °C)	Humidité	Vent	Jouer	Pred-Jouer
J15	Soleil	38	Élevée	Faible	Non	Oui
J16	Soleil	39	Élevée	Fort	Non	Non
J17	Couvert	37.7	Élevée	Faible	Oui	Oui
J18	Pluie	20	Élevée	Faible	Oui	Non
J19	Pluie	15	Normale	Faible	Oui	Oui

L'analyse de ce tableau montre que la classe réelle est égale à la classe prédite sauf pour J15 et J18. Nous pouvons donc représenter la matrice de confusion suivante :

		`Pred-Jouer`	
		`Oui`	`Non`
`Jouer`	`Oui`	2	1
	`Non`	1	1

Ensuite, nous pouvons calculer le taux de bonne classification de `Jouer=Oui` et `Jouer=Non` :

$$\tau(\texttt{Oui}) = \frac{2}{3} * 100 = 66.6\%$$

$$\tau(\texttt{Non}) = \frac{1}{2} * 100 = 50\%$$

Ainsi, nous retrouvons le taux de bonne classification global (situé sur la diagonale) :

$$\tau = \frac{(2+1)}{5} * 100 = 60\%$$

1.7 Annexes

Outre les fonctions standards de R décrites dans le Guide de notions générales , nous présentons, ici, quelques fonctions utiles pour la compréhension et l'application des notions étudiées précédemment en utilisant le logiciel R.

- `readxl` : est le paquet de transfert d'un fichier excel vers R.
  https://cran.r-project.org/web/packages/readxl/index.html.
- `boxplot` : est la fonction qui permet d'afficher le graphique boite à moustache.
- `plotix` (*Various Plotting Functions*) : Il s'agit d'un ensemble de fonctions permettant la génération de différents types de graphiques.
- `confusionMatrix` : est la fonction qui permet de calculer la matrice de confusion à partir d'un tableau de données qui contient la classe réelle et la classe prédite.
  https://www.rforge.net/doc/packages/SDMTools/confusion.matrix.html.

Références

[1] Caméra de vidéosurveillance à reconnaissance faciale. http://www.inter-assistance.com/videosurveillance-reconnaissance-faciale.html. Consulté : 20 septembre 2016.

[2] "deepface", le nouveau système de reconnaissance faciale de facebook qui fait froid dans le dos. http://www.huffingtonpost.fr/2014/03/20/deepface-reconnaissance-faciale-facebook_n_5000872.html. Consulté : 29 septembre 2016.

[3] B. Aubert, C. Vazquez, T. Cresson, S. Parent, and J. De Guise. Automatic spine and pelvis detection in frontal x-rays using deep neural networks for patch displacement learning. In *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*, pages 1426–1429, April 2016.

[4] A. Cornuejois and L. Miclet. *Eyrolles*. Apprentissage artificiel - Concepts et algorithmes, 2003.

[5] E. Friburg. "le machine learning pour personnaliser l'apprentissage. https://domoscio.com/le-machine-learning-pour-personnaliser-lapprentissage/. Consulté : 29 septembre 2016.

[6] A. E. Thessen. "adoption of machine learning techniques in ecology and earth science. https://peerj.com/preprints/1720.pdf. Consulté : 29 septembre 2016.

