

الجمهورية الجزائرية الديمقراطية الشعبية

République Algérienne Démocratique et Populaire
Ministère de l'enseignement Supérieur et de la Recherche scientifique

Université Mohamed El Bachir El Ibrahimi Bordj Bou Arréridj Faculté des
Sciences et de la Technologie Département d'informatique



MEMOIRE DE FIN D'ETUDES

Réalisé en vue de l'obtention du diplôme de MASTER 2

Specialité: T.I.C

Thème

**Prévision du désabonnement de clients
dans le secteur de télécommunication**

Présenté par :

- Ammar Boudjelal Amina.

- Benchikh Chaima.

Encadré par :

– Saad Saoud Manel.

Année universitaire : 2020/202

Table des figures

Figure 1.3 : gestion reaction client.....	3
Figure 2.4 : Modèle de neurons biologie	23
FIGURE 3.6 : présentation réseau neurone	26
Figure 4.7 :shéma de perceptron multicouches	33
Figure 5.7 : Algorithme de Rétro propagation	34
Figure 6.4 : Représentation de dataset	44
Figureé 7.4 : changement de type de données.....	45
Figure 8.6 : création d'arbre de décision	46
Figure 9 .6 : matrice de confusion.....	47
Figure 10 .6 : création de réseaux neurones.....	48
Figure 11 .6 : matrice de confusion.....	49
Figure 11 .6 : Présentation roc curv.....	50

Table des tableaux

Tableau 1.6 : type des fonction d'activation	30
Tableau 2.4 : description de base de donné.....	43
Tableau 2.7 : comparaison des résultats des mesures d'évaluation pour les algorithmes d'apprentissage supervisé.....	

Introduction général

Aujourd'hui les entreprises sont face à une perte importante de revenus en raison de la concurrence croissante et la perte clientèle, c'est pour quoi , la prédiction du taux de désabonnement des clients (churn) est devenu l'une des Principales préoccupations de l'industrie de télécommunication problèmes de l'industrie de télécommunication .

L'objectif de ce projet est créé deux modèles de prédiction du churn des clients dans le secteur de telecommunication, afin d'identifier les clients les plus proches d'abandonner leur operateur telephonique actuel.

Les méthodes que nous avons adoptées est celles des réseaux de neurone artificielles et arbre de decision .

Structure de mémoire

Ce mémoire s'article autour de trois chapitres principaux :

- Dans le premier chapitre « secteur de télécommunication », une première partie est consacrée aux secteur de télécommunications avec quelques définitions. technique de prévision, Une deuxième partie présente la gestion de client et La relation de client et entreprise .
- Dans le deuxième chapitre qui s'intitule « apprentissage automatique », nous présenterons l'apprentissage automatique avec ses différentes méthodes. Ensuite, nous focaliserons sur les deux méthodes de prévision qu'on a choisie.
- Le dernier chapitre « développement et implémentation », comprend deux parties: La première partie est consacrée à la definition de la base de donnée utilisée et sa préparation .dans la deuxième partie nous exposerons les resultats obtenus.

Chapitre 1



**Secteur de
télécommunication**

1. Introduction

L'histoire des télécommunications a commencé à se développer vers la première moitié du XIXe siècle, avec le télégraphe électrique (qui permettait d'envoyer des messages avec des lettres et des chiffres). Plus tard est apparu le téléphone grâce auquel il est devenu possible de communiquer en faisant usage de la voix.

Avec les ondes de radio, la communication sans fil est venue pour révolutionner les habitudes de l'humanité.

Il est clair que les innovations technologiques dans le domaine de la télécommunication ne se sont jamais arrêtées. Le modem a rendu possible la transmission de données entre des ordinateurs et d'autres dispositifs, constituant ainsi le point de départ pour le développement d'Internet et d'autres réseaux informatiques.

À l'heure actuelle, les télécommunications font partie d'un secteur industriel qui fait brasser des millions d'euros chaque année dans le monde entier.

2. Télécommunication

Les télécommunications sont aujourd'hui définies comme la transmission à distance

D'information avec des moyens électroniques. Ce terme est plus utilisé que le terme.

Synonyme officiel communication électronique. Elles se distinguent ainsi de la poste.

qui transmet des informations ou des objets sous forme physique.

2.1 Télécommunications et sciences

Ce secteur connaît une forte croissance depuis les années 1980, grâce aux progrès technologiques réalisés dans les domaines scientifiques dont ce secteur est le carrefour :

- mathématiques : le traitement du signal, la cryptographie, la théorie de l'information et le numérique.
- physique : l'électromagnétisme, les semi-conducteurs, l'électronique et l'opto-électronique .
- informatique : le génie logiciel et la diffusion de la micro-informatique.
- chimie : Réaction d'oxydo-réduction (réduction de poids et autonomie prolongée des batteries des appareils portatifs) .

2.2 Télécommunications et société

Les télécommunications sont un élément crucial de la société moderne. En 2006, l'industrie des télécommunications représentait un revenu de 1200 milliards de dollars, soit 3% du revenu mondial

A l'échelle microéconomique, les entreprises utilisent les télécommunications pour construire leur activité, comme les ventes en ligne, ou améliorer leur efficacité, comme les magasins traditionnels. Dans le monde entier, des services à domiciles peuvent être obtenus sur simple appel téléphonique, des livraisons de pizzas au dépannage. Dans les communautés les plus pauvres, le téléphone mobile sert aussi bien au Bangladesh qu'en Côte d'Ivoire pour négocier les ventes agricoles au meilleur prix du marché

En raison des avantages économiques d'une infrastructure correcte de télécommunications , laquelle une grande partie du monde n'a pas accès, l'écart de développement peut se creuser par manque de télécommunications.

3. Gestion de relation client

3.1 Définition

La gestion de la relation client est une stratégie de compréhension et un processus.

D'acquisition, de rétention, et de partenariat d'une/avec une clientèle sélective dont le but est de créer une valeur supérieure à la relation de l'entreprise avec ses clients la gestion de la relation client est un système complet qui fournit :

- Un moyen ainsi qu'une méthode qui renforce l'expérience de chaque client –sur un niveau individuel- pour en faire un client à vie.
- Des moyens technologiques et fonctionnels qui permettent d'identifier, capturer et retenir les clients.
- Une vue cohérente de la contribution du client dans l'entreprise.



Figure1.3 : gestion realtion client.

La gestion clientèle repose donc sur 3 mots d'ordre :

- Qualité de la relation client.
- Fidélisation des clients.
- Augmentation du chiffre d'affaires.

3.2 Les objectifs de la gestion de la relation client

Une bonne relation client implique de concentrer prioritairement sa stratégie sur le client, de passer d'une vision centrée sur le produit à une vision centrée sur le client.

Par ailleurs, les décisions des clients ou prospects sont fondées sur des éléments factuels, mais aussi sur leurs émotions. La dimension humaine ne doit pas être négligée afin de lier une relation durable avec le client.

3.3 Amélioration la stratégie de gestion de la relation client

Pour soigner votre relation client, prêtez attention aux points suivants :

- un service client disponible via plusieurs canaux.
- un service après-vente ou un support efficace .
- une communication claire.
- des messages et offres personnalisés.

l'implication des clients dans la vie de votre entreprise, via la prise en compte de leurs avis ou demandes.

4.1 La littérature sur le désabonnement

4.1 Définition

Le désabonnement est l'action qui permet à un abonné à un service de mettre fin à son abonnement. Selon les cas, la procédure de désabonnement peut être plus ou moins longue et complexe.

Le phénomène est généralement mesuré par le taux d'attrition ou désabonnement. L'observation du taux d'attrition est une donnée très importante quand on connaît l'impact de la fidélisation sur la rentabilité.

4.2 Taux de désabonnement

Le taux de désabonnement des clients est l'une des mesures les plus importantes à évaluer pour une entreprise en croissance. Bien que ce ne soit pas la mesure la plus heureuse, c'est un chiffre qui peut donner à l'entreprise la dure vérité sur sa fidélisation de la clientèle .

4.3 Les causes Du désabonnement

S'il est difficile d'empêcher les clients de quitter la plateforme, on peut en revanche identifier les causes des désabonnements précédents et tenter d'y remédier. Plusieurs outils permettent de collecter et de retrouver ces données :

- Le sondage en phase de désabonnement, qui doit proposer plusieurs choix de réponse au client
- Des outils de prédiction et d'analyse des données liées à l'attrition
- Une stratégie de suivi afin de prévenir le départ des clients, notamment à travers des campagnes personnalisées et des enquêtes .

4.4 L'importance de taux désabonnement

Le taux de désabonnement est l'une des mesures commerciales les plus critiques pour les entreprises utilisant un modèle commercial basé sur un abonnement de modélisation financière. Par exemple, un taux de désabonnement élevé ou un taux de désabonnement en constante augmentation dans le temps peut nuire à la rentabilité d'une entreprise et limiter son potentiel de croissance. Ainsi, la capacité de prédire le taux de désabonnement est essentielle au succès de l'entreprise. De nombreuses entreprises s'appuient sur des analyses prédictives qui permettent de créer des modèles qui prévoient les taux de désabonnement.

Afin de réduire le taux de désabonnement, les entreprises utilisent différentes méthodes et stratégies. En général, les stratégies sont axées sur l'amélioration de la fidélisation et de la satisfaction de la clientèle. L'objectif de la liaison client est de développer une communication proactive avec les clients, d'obtenir un retour d'information constant des clients sur les performances de l'entreprise et d'améliorer les opérations.

Formul du calcul :

$$\text{Taux d' attrition} = \text{Nombre Clients Perdus} / \text{Nombre Clients Total}$$

4.5 Types des clients désabonner

Il existe deux type sont :

- **Les clients churn involontaires :**

sont les abonnées que la société de télécommunication d'écide de supprimer pour des raisons telles que la fraude et le non-paiement.

- **Le client churn volontaire**

peut être décrit comme la fin du service par l'abonné. Le taux de désabonnement volontaire peut également être considéré comme un désabonnement incontrôlé et délibéré.

5. Prédiction de désabonnement

5.1 Définition

La prédiction est une étude générale d'une situation donnée, dont on peut, par déduction, calcul, mesure scientifique, connaître par avance l'évolution; Ce que l'on prévoit, ce qui est prévisible, ce que l'on juge devoir être .

D'une façon générale, c'est la science de la description de l'avenir. Dans un sens plus restrictif, en épistémologie contemporaine, la prédiction se distingue de la prédiction qui est issue d'une loi ou théorie scientifique hautement confirmée ou corroborée tandis que la prédiction découle d'hypothèses ou de conjectures moins assurées.

Selon Fayol, le père de la direction moderne: **«La prédiction est l'essence même de la gestion. Le succès d'une entreprise dépend en grande partie de l'efficacité des prévisions et de la préparation aux événements futurs.**

5.2 Pourquoi la prédiction ?

les prévisions aujourd'hui on peut se trouvent dans la plupart des sphères de l'activité humaines et sont essentielles pour le bon fonctionnement dans l'entreprise car ils ont aidé à la prise de décision.

5.3 comment faire prédiction ?

nous avons besoin de collecter un ensemble de données de passées pour analyser et utiliser pour effectuer une prédiction.

5.4 Type de prévision

- Prévision d'aérodrome.
- Prévision des crues.
- Prévision de la demande.
- Prévision économique.
- Prévision des épidémies.
- Prévision numérique du temps.
- Prévision météorologique.
- Prévision des orages violents.
- Prévision de qualité de l'air.
- Prévision volcanologique

5.5 Les étapes d'une prévision

- Préciser le but de la prévision.
- Déterminer l'horizon de temps.
- Choisir une méthode de prévision.
- Recueillir des données pertinentes des sources fiables.
- Établir un modèle de prévision.
- Évaluer le modèle.
- Implanter les résultats du modèle.
- Entretenir le modèle de prévision.

5.6. Technique de prévision

La capacité de prédire qu'un client particulier présente un risque élevé d'être un churn, alors qu'il est encore temps de faire quelque chose à ce sujet, représente une énorme source de revenus potentiels supplémentaires pour chaque

entreprise. Ainsi, un énorme investissement a été réalisé dans ce domaine, et de multiples approches ont été étudiées et testées. L'exactitude des techniques utilisées est essentielle à la réussite de tout effort de rétention. Dans les sections suivantes, certaines de ces techniques appliquées au problème de baratte seront présentées .

5.6.1 Prédire le taux du churn des clients dans l'industrie des télécommunications, en utilisant SAS

Le but de cette étude de Lu [2002] est d'appliquer des techniques d'analyse de survie pour prédire le taux de désabonnement des clients en utilisant les données d'une entreprise de télécommunications. En fait, cette étude fournit un autre outil statistique SAS qui analyse de survie pour prévoir le taux de désabonnement des clients. En classant les probabilités de survie prédites par les clients dans l'ordre croissant, les deux d'éciles supérieurs capturent 55-60% des clients churn et les cinq premiers d'éciles capturent près de 90% des clients churn. Les résultats de cette étude aideront les entreprises de télécommunications à comprendre le risque de désabonnement des clients au cours de la durée de leur mandat afin d'optimiser leurs ressources de rétention et/ou de traitement de la clientèle dans leurs efforts de réduction du roulement. Dans l'ensemble, cette étude est utile pour personnaliser les communications marketing et les programmes de traitement des clients pour optimiser leur marketing efforts d'intervention.

5.6.2 Une recherche Prédiction du churn dans l'industrie des télécommunications en utilisant R

La recherche proposée par Kaur and Mahajan [2015] a utilisé les techniques du data mining et le logiciel R pour prédire les résultats du churn des clients sur l'ensemble de données de référence de Churn. Elle a évalué, le nombre des clients churn en utilisant les techniques de classification. L'outil R a représenté le grand volume de données sous la forme de graphiques qui représentent les résultats de manière claire et d'une manière unique de visualisation de modèle. Le facteur de désabonnement est utilisé dans de nombreuses fonctions pour représenter les différentes zones ou scénarios lorsque le taux de churn est élevé. L'étude prédit qu'il y a un énorme écart dans le graphique des clients churn lorsque les appels de service à la clientèle sont mesurés. Les graphiques sont faits en prenant les facteurs de désabonnement comme les paramètres décisifs. ils représentent les différentes manières d'observer le nombre des clients churn de l'ensemble de données.

5.6.3 Prédire le taux du churn des clients dans l'industrie des télécommunications en utilisant réseaux neuronaux récurrents

Cette thèse de Ljunghed [2017] propose une nouvelle approche de la prédiction du churn. Le modèle proposé utilise un réseau de neurones récurrent pour identifier les clients churn en fonction de la régression des séries temporelles de la valeur à vie du client. En effet, cette thèse a étudié la possibilité de combiner RNN avec une prédiction de séries chronologiques CLV, une nouvelle approche de prédiction de churn. Les résultats expérimentaux montrent que le modèle proposé a obtenu de meilleurs résultats que l'aléatoire lors de la prédiction des clients churn. Il a été constaté que même si la précision de la régression était faible, le RNN était capable d'identifier la tendance parmi tous les membres, ce

qui est essentiel pour prédire le taux du churn. Ces résultats indiquent que le RNN formé pour effectuer la prédiction des séries chronologiques CLV est une approche prometteuse pour la prédiction du churn. Il a également proposé que l'algorithme K-means puisse être utilisé pour effectuer une analyse plus complète des clients churn prédits. En effet, cet algorithme a donné des classes prometteurs et significatifs, ce qui conclut qu'il y avait des modèles à distinguer dans l'ensemble de données. En raison d'une période de temps limitée, une seule solution de regroupement a été étudiée. Les résultats de l'évaluation montrent que K-means est applicable pour étudier la prédiction du churn sans avoir besoin d'un algorithme d'extraction de règle, ce qui résulte en un temps d'exécution plus court et une implémentation plus simple.

5.6.4 Prédiction du churn dans les services d'abonnement

une application de Machines vectorielles de support (SVM) L'une des approches étudiées par Coussement and Van den Poel [2008] consiste à appliquer la SVM à un contexte de baratte d'abonnement à un journal. L'objectif était de construire un modèle précis de baratte en utilisant et en mettant au point cette technique. Les performances du modèle en matière de prédiction du taux du churn ont été comparées à la régression logistique et aux prévisions aléatoires. Les auteurs ont choisi comme métrique principale pour évaluer leurs modèles Area under curve (AUC). La meilleure valeur concernant leurs modèles SVM était 85.14%, et le modèle de forêt aléatoire utilisé comme référence a conquis une valeur finale AUC de 87.21 .

6. Conclusion

Dans le premier chapitre, nous avons présenté les informations générales sur le secteur de télécommunication et ainsi le désabonnement des clients, aussi les différents techniques et objectif de la prévision de désabonnement.

Dans le prochain chapitre, nous allons présenter les algorithmes d'apprentissage utilisés pour la prédiction du churn des clients. Malheureusement l'entreprise a refusé de fournir sa propre base de données car elles sont privées et protégées donc on utilise une base de données publique téléchargée à partir d'un site sur internet.

Chapitre 2



APPRENTISSAGE
Automatique

1. Introduction

Dans sa nature, l'Homme a toujours tendance à vouloir apprendre de ses expériences passées, de ses échecs et de ses réussites, afin de pouvoir contrôler et prédire au mieux possible sa vie future. Ainsi, chacune de nos expériences est représentée par ses caractéristiques et son contexte, qui sont emmagasinées dans notre mémoire au fur et à mesure. Ces informations sont alors utilisées pour nous aider à prendre une décision lorsque l'on rencontre une nouvelle situation. Dans un contexte scientifique, cet apprentissage humain est plutôt connu sous le nom de l'apprentissage automatique.

Dans ce chapitre nous présentons les différentes définitions du l'apprentissage automatique et nous citons quelques algorithmes les plus populaires.

2. Apprentissage automatique

2.1 Définition

- «L'apprentissage automatique est la science qui consiste à faire en sorte que les ordinateurs agissent sans être explicitement programmés.» - Stanford [1]
- L'apprentissage automatique également appelé apprentissage machine ou apprentissage artificiel et en anglais machine learning, est une forme d'intelligence artificielle qui permet à un système d'apprendre à partir des données et non à l'aide d'une programmation explicite. Cependant, l'apprentissage automatique n'est pas un processus simple. Au fur et à mesure que les algorithmes ingèrent les données de formation, il devient possible de créer des modèles plus précis basés sur ces données. Un modèle de machine learning est le résultat généré lorsque vous entraînez votre algorithme d'apprentissage automatique avec des données. Après la formation, lorsque vous fournissez des données en entrée à un modèle, vous recevez un résultat en sortie.
- est un champ d'étude de l'intelligence artificielle qui se fonde sur des approches mathématiques et statistiques pour donner aux ordinateurs la capacité d'apprendre .

à partir de données, c'est-à-dire d'améliorer leurs performances à résoudre des tâches sans être explicitement programmés pour chacune. Plus largement, il concerne la conception, l'analyse, l'optimisation, le développement et l'implémentation de telles méthodes.

2.2 Type d'apprentissage automatique

Les algorithmes sont les moteurs qui propulsent l'apprentissage automatique. En général, deux grands types d'algorithmes d'apprentissage automatique sont utilisés aujourd'hui : l'apprentissage supervisé et l'apprentissage non supervisé. La différence entre les deux est définie par la manière dont chacun apprend sur les données pour faire des prédictions.

- **Apprentissage non supervisé**

L'apprentissage non supervisé utilise une approche plus indépendante, dans laquelle un ordinateur apprend à identifier des processus et des modèles complexes sans l'encadrement étroit et constant d'un humain. L'apprentissage non supervisé comporte un entraînement fondé sur des données qui n'ont pas d'étiquettes ou une sortie précise et définie.

L'objectif de l'apprentissage non supervisé est de modéliser la structure ou la distribution sous-jacente dans les données afin d'en apprendre davantage sur les données.

- **Apprentissage supervisé**

est le paradigme d'apprentissage le plus populaire en Machine Learning et en Deep Learning. Comme son nom l'indique, cela consiste à superviser l'apprentissage de la machine en lui montrant des exemples (des données) de la tâche qu'elle doit réaliser. Les applications sont nombreuses : Reconnaissance vocale, vision par ordinateur, régressions, classifications... La grande majorité des problèmes de Machine Learning et de Deep Learning

Utilisent l'apprentissage supervisé. Il est donc essentiel de bien comprendre le fonctionnement de cette mécanique.

L'apprentissage supervisé fonctionne en 4 étapes :

1. Importer un Dataset qui contient nos exemples
2. Développer un Modèle aux paramètres aléatoires
3. Développer une Fonction Coût qui mesure les erreurs entre le modèle et le Dataset
4. Développer un Algorithme d'apprentissage pour trouver les paramètres du modèle qui minimisent la Fonction Coût.

L'apprentissage supervisé nécessite une base d'apprentissage où chaque instance est préalablement étiquetée avec sa classe respective. Supposons que nous disposons d'un échantillon de données formulés de N variables, nous aurons alors deux types d'informations :

- a. Un vecteur de valeurs $X = (x_1, x_2, \dots, x_N)$ prises par chaque variable.
- b. une valeur de sortie Y appelée valeur supervisée (qui peut être une classe pour le cas d'une classification).

Nous pouvons représenter le problème décrit précédemment par un ensemble d'entrée sortie

(X, Y) , avec $i = 1 \dots n$, n le nombre d'observations disponibles. On appelle alors fonction d'apprentissage la fonction notée : $X \rightarrow Y$ qui associe un résultat supervisé à toute nouvelle entrée qui pourrait lui être présentée.

2.3 Objectifs des méthodes d'apprentissage

Extraire et exploiter automatiquement l'information présente dans un jeu de données.

- **La classification**

La classification consiste à inférer, à partir d'un échantillon d'objets classés, une procédure de classification. Donc les classes sont connues et l'on dispose d'exemple de chaque classe , par exemple, classer si un email est un spam ou non ou classer une image de scanographie comme étant maligne ou bénigne .

- **La régression**

Similaire à la classification, sauf que la classification se rapporte à des événements discrets. L'estimation porte sur des variables continues. Par exemple en prévoyant les prix d'un logement en fonction de l'âge de la maison, du nombre de pièces, de l'emplacement du quartier.

3 Les algorithmes d'apprentissage automatique

- **Algorithme d'apprentissage supervisé**

Parmi les méthodes de classification supervisées les plus populaires on peut citer

3.1 Plus proche voisins

L'algorithme k-voisins les plus proches utilise l'intégralité du data set en tant qu'entraînement, au lieu de diviser se dernier en un training et testing set. Quand un résultat est requis pour une nouvelle instance de données, l'algorithme K-voisins parcourt l'intégralité du data set pour rechercher les k-instances les plus proches de la nouvelle instance ou le nombre k d'instances les plus similaires au nouvel enregistrement, puis renvoie la moyenne de les résultats ou le classe à laquelle appartient cette instance si c'est un problème de classification. L'utilisateur spécifie lui meme la valeur de k.

3.2. Classifieur Naïf Bayésien

Naïve Bayes Classifier est l'une des méthodes d'apprentissage supervisé les plus populaires parmi celle utilisant les similarités, qui s'appuie sur le populaire théorème de probabilité de Bayes. En particulier pour la prédiction des maladies et la classification des documents. Il s'agit d'une classification simple de mots basée sur le théorème de probabilité de Bayes pour l'analyse subjective du contenu.

L'avantage du classifieur bayésien naïf est qu'il requiert relativement peu de données d'entraînement pour estimer les paramètres nécessaires à la classification, à savoir moyennes et variances des différentes variables. En effet, l'hypothèse d'indépendance des variables permet de se contenter de la variance de chacune d'entre elles pour chaque classe, sans avoir à calculer de matrice de covariance.

3.3. Arbre de décision

Un arbre de décision est une structure qui est souvent utilisée pour représenter des connaissances. Il permet de remplacer ou d'assister un expert humain dans la détermination des propriétés d'un objet, c'est l'opération de classification. Un arbre de décision est une représentation d'une procédure de décision pour déterminer la classe d'un objet donnée.

Un avantage majeur des arbres de décision est qu'ils peuvent être calculés automatiquement à partir de bases de données par des algorithmes d'apprentissage supervisé. Ces algorithmes

sélectionnent automatiquement les variables discriminantes à partir de données non structurées et potentiellement volumineuses. Ils peuvent ainsi permettre

d'extraire des règles logiques de cause à effet (des déterminismes) qui n'apparaissent pas initialement dans les données brutes.

3.3.1 Les composants d'un arbre de décision

La racine :

La racine c'est la décision à laquelle on fait face. C'est la question principale à laquelle on cherche à répondre

Les branches porteuses :

Les branches porteuses représentent nos principaux choix. Ce sont des grandes options.

Les sous-branches:

Les sous-branches sont des extensions de branches. Ce sont des sous-choix si vous voulez .

3.3.2 les étapes pour créer un arbre de décision

❖ Définissez la problématique

La première étape de l'arbre de décision est de définir la problématique. Pour rappel c'est la racine de décision, c'est de là que tout part. Pour définir la problématique, demandez-vous quelle décision vous devez prendre et formulez-la par une question.

❖ Définissez les branches porteuses

Une fois que vous avez formulé votre décision par une question, vous devez lister les différents choix auxquels vous êtes confrontés. Ces choix constitueront les branches porteuses.

❖ Définissez des sous-branches

Si le choix est simple et que vous n'avez besoin que des branches porteuses pour faire le choix.

❖ Choisissez les critères discriminants

Pour l'étape 4 on peut choisir un ou plusieurs critères discriminants.

Un critère discriminant est un élément que vous utilisez pour départager vos choix. Par exemple si vous hésitez entre plusieurs plats au restaurant, un critère discriminant serait le prix. C'est ce que vous utilisez pour faire votre choix.

❖ Faites les calculs

Une fois que vous avez choisi les critères discriminants, il vous suffit de procéder au calcul final. Généralement une simple multiplication suffit. Donc ici on fera le calcul suivant :

(% de probabilité de réussite) x (profit que je pense pouvoir dégager la 1ère année)

❖ Prenez votre décision

Pour prendre votre décision, il vous suffit de repérer le plus grand résultat pour savoir quoi choisir.

3.3.3 Les Avantages d'arbre de décision

La popularité des arbres de décision se justifie par les raisons suivantes :

- Ils sont faciles à comprendre.
- Ils peuvent être utiles avec ou sans données concrètes, et les données quelles qu'elles soient nécessitent une préparation minimale.
- De nouvelles options peuvent être ajoutées aux arbres existants.
- Ils permettent de sélectionner l'option la plus appropriée parmi plusieurs.
- Il est facile de les associer à d'autres outils de prise de décision.

3.3.4 Algorithme d'apprentissage d'arbre de décision

Echantillon A avec m enregistrements étiquetés .

Initialisations : Arbre : vide ; nœud courant : racine ; échantillon courant : A ;

Répéter

Décider si le nœud courant est terminal ;

Si le nœud courant est terminal **alors**

Etiqueter le nœud courant par une feuille

Sinon

Sélectionner un test et créer le sous arbre

FinSi ;

Nœud courant : un nœud non encore étudié .

Echantillon courant : échantillon atteignant le nœud courant .

Passer au nœud suivant non exploré s'il en existe .

Jusqu'à obtenir un arbre de décision h .

Sortie : arbre de décision.

3.4 Réseaux de neurone :

En intelligence artificielle, les réseaux de neurones sont des assemblages de plusieurs couches de neurones. Il existe des centaines de types de réseaux de neurones différents. Le principe est toujours le même, c'est l'architecture qui est différente à chaque fois.

De manière générale, les réseaux de neurones sont un ensemble de neurones, organisés en plusieurs couches et reliés entre eux. La première couche est appelée input layer, la

dernière output layer et les couches qui se situent entre les deux sont appelées hidden layers. Il peut y en avoir un très grand nombre, selon les domaines d'applications.

Pourquoi réseaux de neurones ?

Les réseaux de neurones ont la capacité de traiter des problèmes divers et variés. Le résultat peut être une prédiction, une classification ou une analyse de

données. Ils permettent de traiter des problèmes non structurés, c'est-à-dire des problèmes sur lesquels on ne dispose d'aucune information au préalable.

Dans notre étude nous avons choisi cette méthode car Les réseaux de neurones ont la capacité de traiter des problèmes divers et variés. Le résultat peut être une prédiction, une classification ou une analyse de données. Ils permettent de traiter des problèmes non structurés, c'est-à-dire des problèmes sur lesquels on ne dispose d'aucune information au préalable.

3.4.1. Neurone biologique

Structure

On pense que le système nerveux compte plus de 1000 milliards de neurones interconnectés. Bien que les neurones ne soient pas tous identiques, leur forme et certaines caractéristiques permettent de les répartir en quelques grandes classes. En effet, il est aussi important de savoir, que les neurones n'ont pas tous un comportement similaire en fonction de leur position

dans le cerveau. Avant de rentrer plus en avant dans les détails, examinons un neurone.

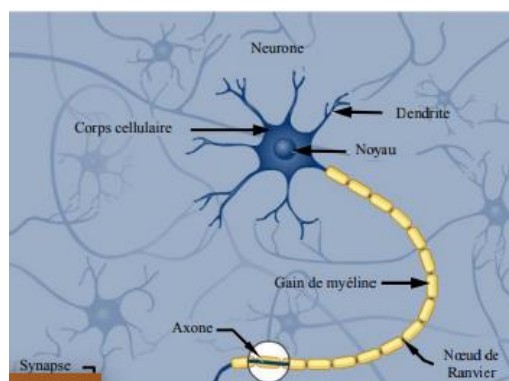


Figure 2.4 Modèle de neurones biologique

On peut le décomposer en trois régions principales:

- Le corps cellulaire : Il contient le noyau du neurone ainsi que la machine biochimique nécessaire à la synthèse des enzymes. Ce corps cellulaire de forme sphérique ou pyramidale contient aussi les autres molécules essentielles à la vie de la cellule. Sa taille est de quelques microns de diamètre.
- Les dendrites : Ce sont de fines extensions tubulaires qui se ramifient autour du neurone et forment une sorte de vaste arborescence. Les signaux envoyés au neurone sont captés par les dendrites. Leur taille est de quelques dizaines de microns de longueur.
- L'axone : C'est le long de l'axone que les signaux partent du neurone. Contrairement aux dendrites qui se ramifient autour du neurone, l'axone est plus long et se ramifie à son extrémité ou il se connecte aux dendrites des autres neurones. Sa taille peut varier entre quelques millimètres à plusieurs mètres.
- Synapse : Une synapse est une jonction entre deux neurones, et généralement entre l'axone d'un neurone et un dendrite d'un autre neurone (mais il existe aussi des synapses axo-axonales par exemple)

3.4.2 Neurone formel

1. Présentation

Le neurone formel est donc une modélisation mathématique qui reprend les principes du fonctionnement du neurone biologique, en particulier la sommation des entrées. Sachant qu'au niveau biologique, les synapses n'ont pas toutes la même «valeur» (les connexions entre les neurones étant plus ou moins fortes), les auteurs ont donc créé un algorithme qui

pondère la somme de ses entrées par des poids synaptiques (coefficients de pondération). De plus, les 1 et les -1 en entrée sont là pour figurer une synapse excitatrice ou inhibitrice.

2. Interprétation mathématique

D'un point de vue mathématique, le neurone formel peut être représenté de la manière suivante: Pour un nombre compris entre $j (=1)$ et un nombre quelconque n , le neurone formel va calculer la somme de ses entrées (x_1, \dots, x_n) , pondérées par les poids synaptiques (w_1, \dots, w_n) , et la comparer à son seuil $teta$. Si le résultat est supérieur au seuil, alors la valeur renvoyée est 1, sinon la valeur renvoyée est 0. D'où la formule (avec $f =$ fonction seuil).

3. Fonctions d'activation

Dans sa première version, le neurone formel était donc implémenté avec une fonction à seuil, mais de nombreuses versions existent. Ainsi le neurone de McCulloch et Pitts a été généralisé de différentes manières, en choisissant d'autres fonctions d'activations, comme les fonctions linéaires par morceaux, des sigmoïdes ou des gaussiennes.

4. réseaux de neurones artificiels

4.1 Définition

- Un réseau de neurones artificiels, est un système informatique matériel et / ou logiciel dont le fonctionnement est calqué sur celui des neurones du cerveau humain.

- Il s'agit là d'une variété de technologie Deep Learning qui fait elle même partie de sous-catégorie d'intelligence artificielle du Machine Learning. Un réseau de neurones artificiels est un modèle de calcul dont la conception est très schématiquement inspirée du fonctionnement des neurones biologiques. Les réseaux de neurones sont généralement optimisés par des méthodes d'apprentissage de type probabiliste, en particulier bayésiens.

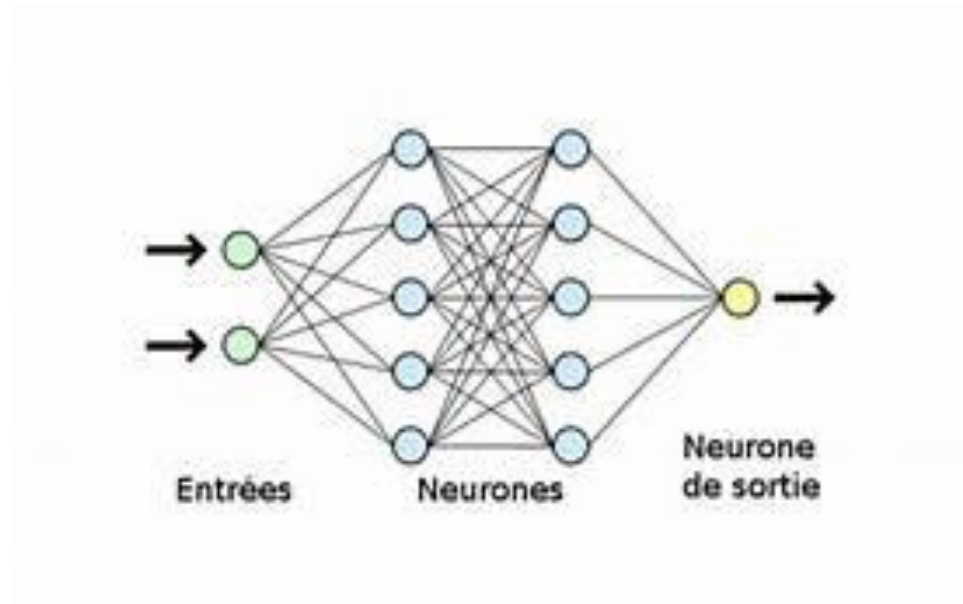


Figure 3.6 présentation réseau neurone

Ils existe quelques grands principes :

- **Le parallélisme** : les neurones sont des entités réalisant une fonction très simple, mais ils sont très fortement interconnectés entre eux, ce qui rend le traitement du signal massivement parallèle.
- **Les poids synaptiques** : les connexions entre les neurones ont des poids variables, qui déterminent la force de l'interaction entre chaque paire de neurones.

L'apprentissage : ces coefficients synaptiques sont modifiables lors de l'apprentissage, dans le but de faire réaliser au réseau la fonction désirée.

4.2 Architecture des réseaux neurones

L'architecture d'un réseau de neurones est l'organisation des neurones entre eux au sein d'un même réseau. Autrement dit, il s'agit de la façon dont ils sont ordonnés et connectés. La majorité des réseaux de neurones utilise le même type de neurones. Quelques architectures plus rares se basent sur des neurones dédiés. L'architecture d'un réseau de neurones dépend de la tâche à apprendre. Un réseau de neurones est en général composé de plusieurs couches de neurones, des entrées jusqu'aux sorties. On distingue deux grands types d'architectures de réseaux de neurones : les réseaux de neurones non bouclés et les réseaux de neurones bouclés.

▪ Les réseaux de neurones non bouclés

Un réseau de neurones non bouclé réalise une (ou plusieurs) fonctions algébriques de ses entrées, par composition des fonctions réalisées par chacun de ses neurones. Un réseau de neurones non bouclé est représenté graphiquement par un ensemble de neurones "connectés" entre eux, l'information circulant des entrées vers les sorties sans "retour en arrière"; si l'on représente le réseau comme un graphe dont les nœuds sont les neurones et les arêtes les "connexions" entre ceux-ci, le graphe d'un réseau non bouclé est acyclique. Le terme de "connexions" est une métaphore : dans la très grande majorité des applications, les réseaux de neurones sont des formules algébriques dont les valeurs numériques sont calculées par des programmes d'ordinateurs, non des objets physiques (circuits électroniques spécialisés) ; néanmoins, le terme de connexion, issu des origines biologiques des réseaux de neurones, est

passé dans l'usage, car il est commode quoique trompeur. Il a même donné naissance au terme de connexionnisme.

▪ Réseaux de neurones monocouches

La structure d'un réseau monocouche est telle que des neurones organisés en entrée soient entièrement connectés à d'autres neurones organisés en sortie par une couche modifiable de poids .

▪ Réseaux de neurones multicouches

Les neurones sont arrangés par couche. Il n'y a pas de connexion entre neurones d'une même couche, et les connexions ne se font qu'avec les neurones de couches avales. Habituellement, chaque neurone d'une couche est connecté à tous les neurones de la couche suivante et celle-ci seulement. Ceci nous permet d'introduire la notion de sens de parcours de l'information (de l'activation) au sein d'un réseau et donc définir les concepts de neurone d'entrée, neurone de sortie. Par extension, on appelle couche d'entrée l'ensemble des neurones d'entrée, couche de sortie l'ensemble des neurones de sortie. Les couches intermédiaires n'ayant aucun contact avec l'extérieur sont appelées couches cachées.

4.3 fonctionnement du réseau neurone artificiel

La conception des réseaux de neurones artificiels s'appuie sur la structure des neurones biologiques du cerveau humain.

Les réseaux de neurones artificiels peuvent être décrits comme des systèmes composés d'au moins deux couches de neurones - une couche d'entrée et une couche de sortie - et comprenant généralement des couches intermédiaires (« hidden layers »). Plus le problème à résoudre est complexe, plus le réseau de

neurones artificiels doit comporter de couches. Chaque couche contient un grand nombre de neurones artificiels spécialisés.

➤ Les fonctions d'activation

Les fonctions d'activations les plus utilisées sont les fonctions «seuil», «linéaire» et «sigmoïde». Comme son nom l'indique, la fonction seuil applique un seuil sur son entrée. Plus précisément, une entrée négative ne passe pas le seuil, la fonction retourne alors la valeur 0 (on peut interpréter ce 0 comme signifiant faux), alors qu'une

entrée positive ou nulle dépasse le seuil, et la fonction retourne à 1 (vrai). Utilisée dans le contexte d'un neurone, . Nous verrons plus loin que cette fonction permet de prendre des décisions binaires.

➤ La fonction linéaire

est très simple, elle affecte directement son entrée à sa sortie : $y = s$ Appliquée dans le contexte d'un neurone, . Dans ce cas, la sortie du neurone correspond à son niveau d'activation dont le passage à zéro se produit lorsque $w \cdot x + b = T$. La fonction de transfert sigmoïde est donnée par : $s = \frac{1}{1 + \exp(-y)}$ (7) Elle ressemble soit à la fonction seuil, soit à la fonction linéaire, selon que l'on est loin ou près de b , respectivement. La fonction seuil est très non linéaire car il y a une discontinuité lorsque $w \cdot x + b = T$. De son côté, la fonction linéaire est tout à fait linéaire. Elle ne comporte aucun changement de pente. La sigmoïde est un compromis intéressant entre les deux précédentes. Notons finalement, que la fonction «tangente hyperbolique (tanh)» est une version symétrique de la sigmoïde.









Nom de la fonction	Relation d'entrée/sortie	Icône
seuil	$a = 0$ si $n < 0$ $a = 1$ si $n \geq 0$	
seuil symétrique	$a = -1$ si $n < 0$ $a = 1$ si $n \geq 0$	
linéaire	$a = n$	
linéaire saturée	$a = 0$ si $n < 0$ $a = n$ si $0 \leq n \leq 1$ $a = 1$ si $n > 1$	
linéaire saturée symétrique	$a = -1$ si $n < -1$ $a = n$ si $-1 \leq n \leq 1$ $a = 1$ si $n > 1$	
linéaire positive	$a = 0$ si $n < 0$ $a = n$ si $n \geq 0$	
sigmoïde	$a = \frac{1}{1+\exp^{-n}}$	
tangente hyperbolique	$a = \frac{e^n - e^{-n}}{e^n + e^{-n}}$	

Tableau 1.6 type des fonction d'activation

5.Apprentissage

L'apprentissage est un processus dynamique et itératif permettant de modifier les paramètres d'un réseau en réaction avec les stimuli qu'il reçoit de son environnement. Le type d'apprentissage est déterminé par la manière dont les changements de paramètre surviennent

Cette définition implique qu'un réseau se doit d'être stimulé par un environnement, qu'il subisse des changements en réaction avec cette stimulation, et que ceux-ci provoquent dans le futur une réponse nouvelle vis-à-vis de l'environnement. Ainsi, le réseau peut s'améliorer avec le temps .

5.1 Stratégie d'apprentissage :

L'apprentissage au sein des différentes architectures dépend de l'architecture du réseau et de l'environnement du problème. Les deux règles d'apprentissage pour mettre à jour les poids d'un neurone (règle de Hebb et de Widrow) ne concernent qu'un neurone seul. Ces règles peuvent servir pour mettre à jour les poids d'un neurone, de certains réseaux de neurones, mais ne peuvent être généralisées et s'appliquer à n'importe quelle architecture .

La notion d'apprentissage recouvre deux réalités :

- La mémorisation : le fait d'assimiler sous une forme dense des exemples éventuellement nombreux ;
- La généralisation : le fait d'être capable, grâce aux exemples appris, de traiter des exemples distincts, encore non rencontrés, mais similaires.

✓ Les règles d'apprentissage

Il existe plusieurs règles de modification :

- Loi de Hebb : $\Delta w_{ij} = R a_i a_j$
- Règle de Widrow-Hoff (delta rule) : $\Delta w_{ij} = R (d_i - a_i) a_j$
- Règle de Grossberg : $\Delta w_{ij} = R (a_j - w_{ij}) a_i$.

✓ Loi de Hebb :

Si deux unités connectées sont actives simultanément, le poids de leur connexion est augmenté ou diminué. R est une constante positive qui représente la force d'apprentissage (learning rate)

- $a_i = -1 \quad a_i = 1 \quad a_j = -1 \quad \Delta W_{ij} = R \quad \Delta W_{ij} = -R$
- $a_j = 1 \quad \Delta W_{ij} = -R \quad \Delta W_{ij} = R$
- $\Delta W_{ij} = R a_i a_j$

✓ **Loi de Widrow-Hoff (delta rule) :**

activation produite par le réseau de réponse désirée par l'expert humain Par exemple si la sortie est inférieure à la réponse désirée, il $W_{ij} = R(d_i - a_i)$ à j Par exemple si la sortie est inférieure à la réponse désirée, il va falloir augmenter le poids de la connexion à condition

bien sûr que l'unité j soit excitatrice (égale à 1). On est dans l'hypothèse d'unités booléennes $\{0,1\}$.

$$a_i = 0 \quad a_i = 1$$

$$d_i = 0 \quad \Delta W_{ij} = 0 \quad \Delta W_{ij} = -R$$

$$d_i = 1 \quad \Delta W_{ij} = R \quad \Delta W_{ij} = 0$$

✓ **Loi de Grossberg :**

On augmente les poids qui entrent sur l'unité gagnante a_i s'ils sont trop faibles, pour les rapprocher du vecteur d'entrée a_j . C'est la règle d'apprentissage utilisée dans les cartes auto-organisatrices de Kohonen

$$\Delta W_{ij} = R a_i (a_j - W_{ij})$$

Type de réseau de neurone utilisé dans notre domaine d'étude est Le perceptron multi-couches où perceptron multicouches à été le premier réseau de neurones à trouver des applications industrielles.

6. Le perceptron multicouches

Le perceptron multicouche où encore multilayers perceptron en anglais est le premier réseau de neurones à avoir trouvé de nombreuses applications pratiques telles que la reconnaissance de fleurs, la détection de fraudes, etc.. Il peut être utilisé pour toutes tâches de classification

supervisées. De nos jours, il est l'un des modèles les plus populaires, et est implémenté par de nombreuses bibliothèques telles que TensorFlow, Weka, Scikit-Learn, etc.

PMC ou (Multi-Layer Perceptron) réseau organisé en couches :

une couche : un groupe de neurones uniformes sans connexion les uns avec les autres .

réalise une transformation vectorielle :

une couche reçoit un vecteur d'entrée et le transforme en vecteur de sortie

une couche au moins n'est pas linéaire les dimensions d'entrée et de sortie peuvent être différentes au moins 2 couches (une couche dite "cachée" et une couche de sortie)

structure sans cycle (feed-forward) : une couche ne peut utiliser que les sorties des couches précédentes.

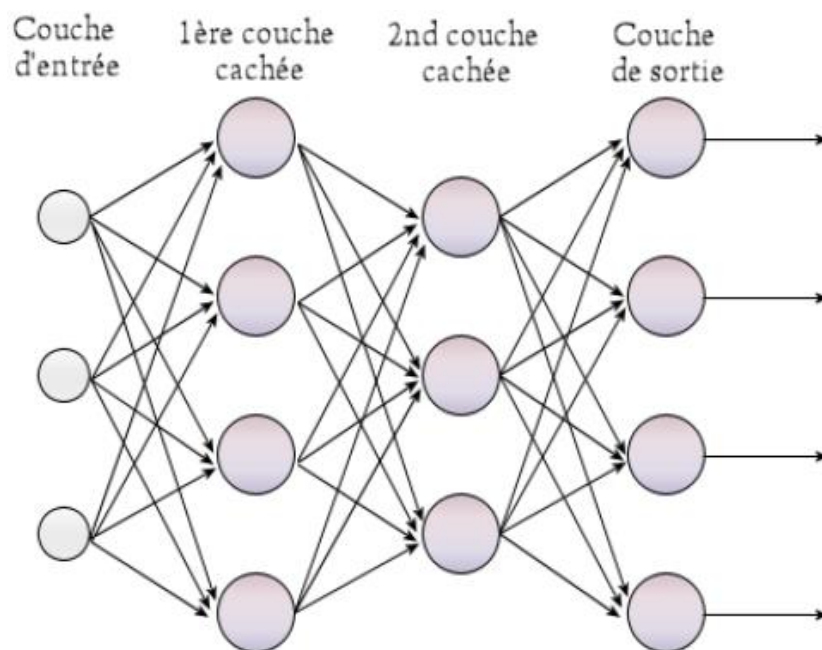


Figure 4.7 schéma de perceptron multicouches

7-Algorithmme de la rétro propagation

La rétropropagation permet de corriger et d'améliorer les erreurs de prédictions qui représentent la différence entre les résultats obtenus et les sorties désirées, avec un simple , retour arrière qui permet de modifier les poids et ceci à chaque fois que la sortie ne correspond au résultat voulu. L'objectif principal de la rétropropagation est de minimiser le taux d'erreurs.

La figure suivante représente un exemple d'un réseau à deux entrées, trois neurones dans une couche cachée, et deux neurones dans la couche de sortie :

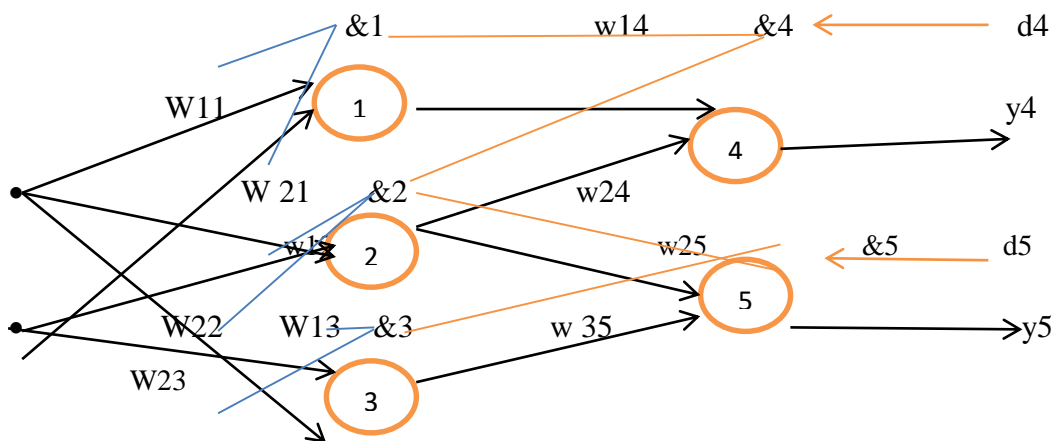


Figure 5.7 Algorithme de Rétro propagation

- Propagation
- Modification des poids.
- Rétropropagation

7. Présentation de l'algorithme de rétropropagation

on utilise la règle d'apprentissage du delta généralisé pour chaque neurone i ,
Pour réaliser l'apprentissage d'un réseau multicouche.

On a Pour chaque neurone de sortie une règle d'apprentissage delta suivantes

$$w_{k+1} = w_k + 2\mu e f'(z)x$$

Pour chaque neurone de sortie, il n'y a pas de problème d'implémentation de la règle, car l'erreur ($e = d - y$) est directement disponible. Par contre pour les neurones de la couche cachée (si l'on considère un réseau a une couche cachée), on est dans l'obligation de

rétro propagé l'erreur de sortie pour obtenir une erreur effective à la sortie de chaque neurone de la couche cachée.

8-Procédure de développement d'un réseau de neurones

Le cycle classique de développement d'un réseau de neurones peut être décomposé en sept étapes :

- **Collecte des données**

L'objectif de cette étape est de recueillir des données suffisantes pour constituer une base représentative.

- **Analyse des données**

Il est souvent préférable d'effectuer une analyse des données de manière à déterminer et différencier celles ci; cette détermination des caractéristiques a des conséquences sur la taille du réseau, ses performances et sur le temps de développement (temps d'apprentissage).

- **Séparation des bases de données**

Il est nécessaire de disposer de deux bases de données : une base pour effectuer l'apprentissage et une autre pour tester le réseau obtenu et déterminer ses performances.

- **Choix du réseau de neurones**

Il existe un grand nombre de types de réseaux de neurones, avec pour chacun des avantages et des inconvénients qui rendent le choix très important.

- **Mise en forme des données**

De manière générale, les bases de données doivent subir un prétraitement afin d'être adaptées aux entrées et sorties du réseau en question.

- **Apprentissage**

Tous les modèles de réseaux de neurones requièrent un apprentissage. Plusieurs types d'apprentissage peuvent être adaptés à un même type de réseau de neurones. Les critères de choix sont souvent la rapidité de convergence ou les performances de généralisation.

- **Validation**

Il est nécessaire de tester le neurone sur une base de données différente de celles utilisées pour l'apprentissage. La validation permet à la fois d'apprécier les performances du système neuronal et de détecter le type de données qui pose problème.

9. Méthodes de normalisation des données

Permet plusieurs techniques de normalisation des données on choisit les plus courantes

X comme référence de variable originale et X* comme variable normalisée.

9-1 La normalisation min-max

Fonctionnement :

Regarde de combien une valeur donnée est plus grande que la valeur minimum X(min) et en échelonnant cette différence par l'amplitude. Ce qui donne

$$X^* = \frac{X - \min(X)}{\text{amplitude}(X)} = \frac{X - \min(X)}{\max(X) - \min(X)}$$

9-2 La normalisation par le test Z

Fonctionnement :

prenant la différence entre la valeur de la variable étudiée et la moyenne des valeurs de cette variable et en échelonnant cette différence par la variance des

valeurs du champ, ce qui donne

$$X^* = \frac{X - \text{moyenne}(X)}{\text{écart type}(X)}$$

9.3 La normalisation max

Fonctionnement :

divise toutes les valeurs d'une variable sur la valeur maximale afin d'obtenir des valeurs normalisées comprises dans l'intervalle 0 et 1

$$X^* = \frac{X}{\max(X)}$$

10. les limites de réseau du neurone :

Le réseau de neurones représente l'intelligence artificielle des programmes informatiques mais de nombreux éléments limitent encore ses capacités :

- Plus le problème est complexe, plus il nécessite une base de données importante pour la « base d'apprentissage » : coûteux en temps et en argent.
- La détermination des descripteurs nécessite une analyse réalisée par un expert, et la même base de données brutes, selon les descripteurs qui en sont extraits, peut engendrer des réseaux de neurones aux performances très différentes .

11. Conclusion

Nous avons présenté une vue générale sur les méthodes de classification supervisée et qu'elles peuvent être basées sur des hypothèses probabilistes (classifieur naïf bayésien..) ou sur des notions de proximité (plus proche voisins) ou bien encore sur des recherches dans des espaces d'hypothèses (arbre de décision, réseau de neurone et SVM).

Nous avons vu les arbres de décision et les différentes composantes étapes comment réaliser, Après nous avons vu les types des réseaux de neurones, l'architecture des neurones ainsi que les fonctions d'activation

Nous avons choisi le type de perceptron-multicouches dans lequel on utilise l'algorithme de rétro-propagation ,

A la fin on représente les limites des réseaux de neurones.

Chapitre 3

**Développement
et
Implémentation**



1 .Introduction

Ce chapitre est consacré à la partie réalisation et tests de notre projet. En effet nous commençons par présenter les logiciels et bibliothèques utilisés ainsi que la base de données sur laquelle nous avons construit et testé notre système, pour ensuite exposer les résultats obtenus. Au final, nous discuterons ces résultats .

2.Outile de développement

C'est très important Le choix du bon environnement et langage de programmation pour le développement des projets.

2.1 Langage de programmation

- **python** : Python est un langage dynamique et interactif orienté sur les objets. Il vous permet d'utiliser modules, exceptions, syntaxe dynamique, classes et des types de données de haut niveau. Il est souvent comparé à d'autres langages comme Perl, Ruby ou Java.

Python se définit comme un langage facile à utiliser. Il utilise une syntaxe élégante qui permet au code source d'être facilement lu. Il est multiplateforme, donc il fonctionne sous Windows, MacOS, Linux, etc.

2.2 Environnement de développement

- **Google colab Colaboratory**

permet d'écrire et d'exécuter du code Python dans votre navigateur. Il offre les avantages suivants :

- Aucune configuration requise
- Accès gratuit aux GPU
- Partage facile

Les notebooks Colab exécutent ce code sur les serveurs cloud de Google. Nous avons donc à votre disposition toute la puissance du matériel Google, y compris les GPU et TPU, quelle que soit la puissance de notre ordinateur. nous n'avons besoin que d'un navigateur.

Colab est très largement utilisé par la communauté du machine learning, par exemple dans les applications suivantes :

- Premiers pas avec TensorFlow
- Développement et entraînement de réseaux de neurones
- Expérimentation avec les TPU
- Dissémination de la recherche en IA
- Création de tutoriels

3.Librairies

1. **Numpy** est une librairie fondamentale pour effectuer des calculs numériques avec Python. Elle facilite grandement la gestion des tableaux de données avec un grand nombre de fonctions permettant de générer des objets de type ndarray. Ces objets ont l'avantage d'être facilement manipulable et contiennent des éléments de même type
2. **Pandas** est également facile à prendre en main et permet de lire simplement des données sous différents formats : fichiers CSV et texte, Microsoft Excel, bases de données SQL.
3. **Kersa** est idéal si vous souhaitez implémenter des modèles complexes de Deep learning. Il est plus souple et plus facile à prendre en main qu'un autre framework très connu en Deep Learning : Tensorflow. Grâce à Keras, vous apprendrez par

4. exemple à implémenter simplement un réseau de neurones convolutif (CNN) reconnu dans les problèmes de classification liés à l'image.
5. **Seaborn** la visualisation de données, spécialisée dans l'analyse statistique. Basée sur la bibliothèque Matplotlib, elle est totalement adaptée aux data frames de Pandas. Ainsi, au-delà d'une interface visuellement améliorée, Seaborn permet de produire rapidement et intuitivement des graphiques statistiques de qualité.

5. Matplotlib est un puissant outil pour tracer des graphiques et visualiser des données. Elle est destinée à tracer des graphiques de toute sorte (diagramme circulaire, histogramme, nuage de points...). Matplotlib contient une sous-bibliothèque pyplot qui crée une interface proche du logiciel commercial Matlab qui contient des fonctions très similaires à celui-ci.

4. Mise en œuvre

4.1 Compréhension du problème

L'objectif de la prédiction du churn est d'identifier les clients qui sont susceptibles de se désabonner et de quitter l'entreprise afin de d'améliorer la prise de décision et de mettre en place des actions de rétention. Pour ce faire, nous considérons les clients qui ont tourné et analysons leurs données.

4.2 Compréhension du données

L'ensemble des données disponibles à cette recherche contient des informations concernant l'historique des opérations d'appel, des sms et de l'utilisation de données des clients .

Pour notre étude nous avons choisi une base de données sur le site «kaggle .com » Notre ensemble de données contient 3334 lignes et 21 variables.

Le tableau ci-dessous explique le rôle de chaque fonctionnalité.

Le nom de fonctionnalité	Description	Type
State	identifiant	string
Account length	Duré de compte	numeric
Area code	Indicatif régionl	numeric
Phone	Numéro téléphone	numeric
Int'l plan	Plan international	boolean
Vmail plan	Vocal mail plan	Boolean
Vmail message	Vocalmail message	Numeric
Day mins	Minute total jours	Numeric
Day calls	Applle total jours	Numeric
Day charge	Charge total jours	Numeric
Eve Mins	Minute total matin	Numeric
Eve calls	Applle total matin	Numeric
Eve charge	Charge total matin	Numeric
Night Mins	Minutetotal soir	Numeric
Night calls	Applle total soir	Numeric
Night charge	Charge total soir	Numeric
Int Mins	Internation minute	Numeric
Intcalls	Internationall applle	Numeric
Int charge	Internationall charge	Numeric
Crust service calls	Appl service client	Numeric
Churn	Churn	boolean

Tableaué2.4 description de base de données

4.3 Préparation du jeu de données

Le traitement et le nettoyage des données est l'objectif de cette section. le traitement du nombre de doubles et les valeurs manquantes .

```

State Account Length Area Code ... Intl Charge CustServ Calls Churn?
0 KS 128 415 ... 2.70 1 False.
1 OH 107 415 ... 3.70 1 False.
2 NJ 137 415 ... 3.29 0 False.
3 OH 84 408 ... 1.78 2 False.
4 OK 75 415 ... 2.73 3 False.

[5 rows x 21 columns]
Type de données dans les colonnes:
State object
Account Length int64
Area Code int64
Phone object
Int'l Plan object
VMail Plan object
VMail Message int64
Day Mins float64
Day Calls int64
Day Charge float64
Eve Mins float64
Eve Calls int64
Eve Charge float64
Night Mins float64
Night Calls int64
Night Charge float64
Intl Mins float64
Intl Calls int64
Intl Charge float64
CustServ Calls int64
Churn? object
dtype: object

```

Figure 6.4 Représentation de dataset

4.4 Transformation de données

Certains algorithmes dans l'exploration de données requièrent que l'ensemble de données ait des caractéristiques spécifiques. La première étape de la transformation des données consistait à convertir les types de données en types de données adéquats. est représentée la nouvelle structure de données après cette conversion. La plupart des variables ont été converties en type numérique pour rendre l'ensemble de données facilement adaptable `à plusieurs algorithmes. L'exception à cette conversion était la variable Churn. Ces variables ont été

définies comme un facteur. Les variables factorielles sont des variables catégorielles qui peuvent être numériques ou une chaîne.

```
#changer le type
df['Int\l Plan'].replace(to_replace='yes', value=1, inplace=True)
df['Int\l Plan'].replace(to_replace='no', value=0, inplace=True)

df['VMail Plan'].replace(to_replace='yes', value=1, inplace=True)
df['VMail Plan'].replace(to_replace='no', value=0, inplace=True)

df['Churn'].replace(to_replace='True.', value=1, inplace=True) # True
df['Churn'].replace(to_replace='False.', value=0, inplace=True)
```

Figureé 7.4 changement de type de données

4.5 Diviser base de données

On a divisé les données d'origine en deux sous-ensembles : ensemble de train et de test définies avec un rapport 70/30. L'ensemble de train est formé par 70% des données originales, et il vise à former l'algorithme. L'ensemble de test contient 30% de l'ensemble de données initial et il est utilisé pour minimiser le surajustement et pour ajuster les paramètres d'algorithmes.

```
#diviser le dataset en training et testing
X_train, X_test, y_train, y_test = train_test_split(features, labels, test_size=0.2, random_state=42) # train_size=2500
```

5 Les critères d'évaluation

La première approche consistait à utiliser la précision "Accuracy" comme une métrique pour évaluer les modèles, ainsi les métriques d'évaluation choisies sont AUC ou courbe ROC .

Les algorithmes appliqués ont été choisis en raison de leur diversité de représentation et de style d'apprentissage, et de leur application commune sur ce

type de problèmes. Nous avons également pris en considération les études concernant la popularité et l'efficacité [Wu et al., 2008]. Deux modèles différents d'apprentissage automatique ont.

été formés et comparés entre eux. Ces algorithmes sont :

1. Arbre de décision
2. Réseau de neurones

6 .les Résultats

6-1.Implémentation de modèle d'arbre décision

Dans cette figure on a une petite représentation de code pour crée une arbre dde décidion .

```
print("Creation du modele Arbre de Decision...")
dt_clf = tree.DecisionTreeClassifier(max_features='sqrt', max_depth=15)
dt_clf.fit(X_train, y_train)
results = dt_clf.predict(X_test)
print("Score Arbre de Decision: ", dt_clf.score(X_test, y_test))
dt_predic = dt_clf.predict(X_test)

Creation du modele Arbre de Decision...
Score Arbre de Decision:  0.964957264957265

# les metriques
metrics = getConfMatrixMetrics(y_test, dt_predic)
print('Métriques du modèle Arbre de Decision')
printMetrics(metrics)

Métriques du modèle Arbre de Decision
ACC : 96.50, TNR : 94.97, FNR : 2.02, FPR : 5.03, precision : 95.25, recall : 97.98, F1 : 96.60,
```

Figure 8.6 création d'arbre de décision

Lorsque nous avons testé avec les nouvelles données, le modèle a continué à montrer

une bonne performance. La matrice de confusion du modèle était la suivante :

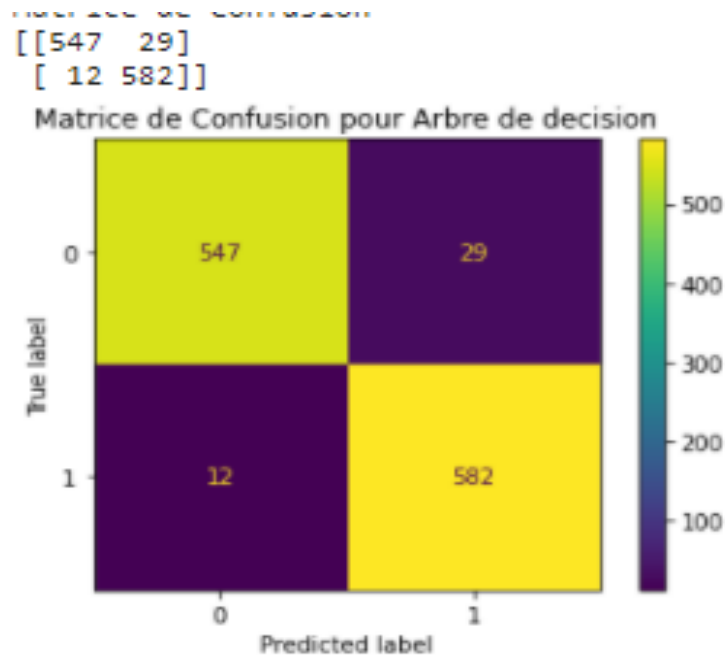


Figure 9 .6 matrice de confusion

On a dans les lignes classe réel et classe prédite donc on a trouve quatre catégorie du résultat :

TP : prédiction et la valeur réel sont positive (547)

TN : prédiction et valeur réelle sont négative (582)

FP : Prédiction est positive alors que la valeur réelle est négative (12)

FN : prédiction est négative alors que la valeur réelle est positive(29)

Alors en trouve que erreur et entre (12 ,29)

6-2. Implémentation de modèle réseau neurone

Dans ce partie on voit une partie de code pour créer modèle de réseaux neurones

```

# perceptron multi couche
print("Création du modèle avec Réseau de Neurone...")
dimension = X_train.shape[1]
|

Création du modèle avec Réseau de Neurone...

[20] # model = tf.keras.Sequential()
model = keras.Sequential()
model.add(Dense(64, activation='relu', input_dim=dimension)) # input_shape=(,) #512

model.add(Dropout(0.1)) # drop 10% of neuron to prevent overfitting
model.add(Dense(32, activation='relu'))

model.add(Dropout(0.1))
model.add(Dense(16, activation='relu'))
model.add(Dropout(0.1))

model.add(Dense(8, activation='relu'))
model.add(Dropout(0.1))

model.summary()

model.compile(loss='binary_crossentropy', optimizer='adam', metrics=['accuracy']) # optimizer : rmsprop, adam
history = model.fit(X_train, y_train, batch_size=100, epochs=200, verbose=2, validation_data=(X_test, y_test))

score = model.evaluate(X_test, y_test, verbose=0)
dnn_predict = model.predict_classes(X_test, verbose=1)

# les metriques
metrics = getConfMatrixMetrics(y_test, dnn_predict)
print('Métriques du modèle Réseau de neurone :')
printMetrics(metrics)

```

Figure 10.6 création de réseaux neurones

Lorsque nous avons testé avec les nouvelles données, le modèle a continué à montrer une bonne performance. La matrice de confusion du modèle était la suivante :

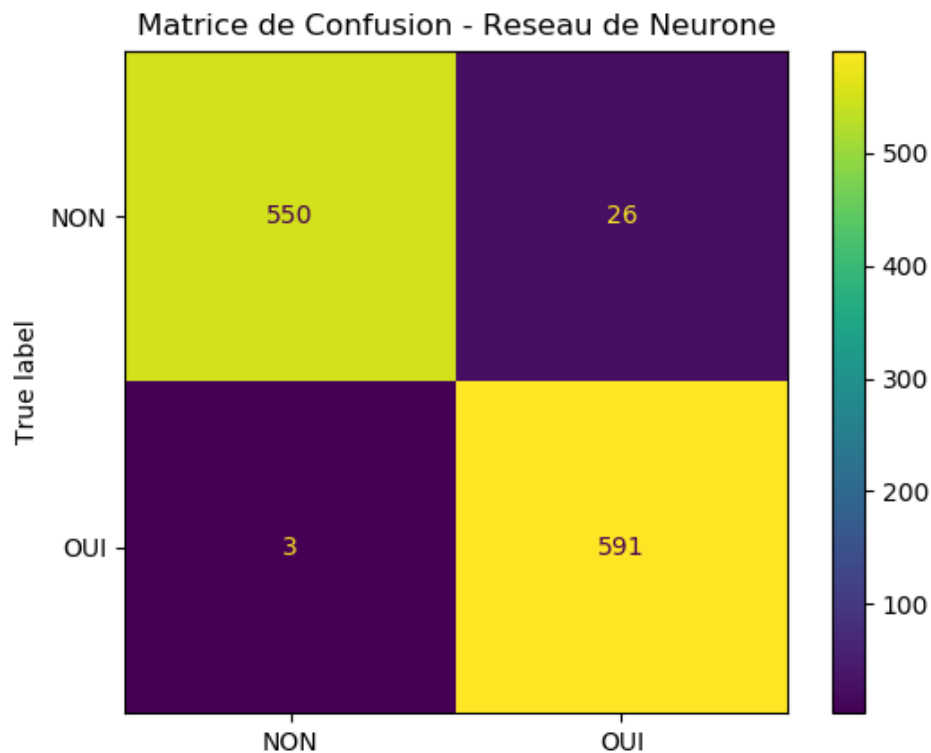


Figure 11 .6 matrice de confusion

On a dans les lignes classe réel et classe prédite donc on a trouve quatre catégorie du résultat :

TP : prédiction et la valeur réel sont positive (550)

TN : prédiction et valeur réelle sont négative (591)

FP : Prédiction est positive alors que la valeur réelle est négative (3)

FN : prédiction est négative alors que la valeur réelle est positive(26)

Alors en trouve que erreur et entre (3 ,26)

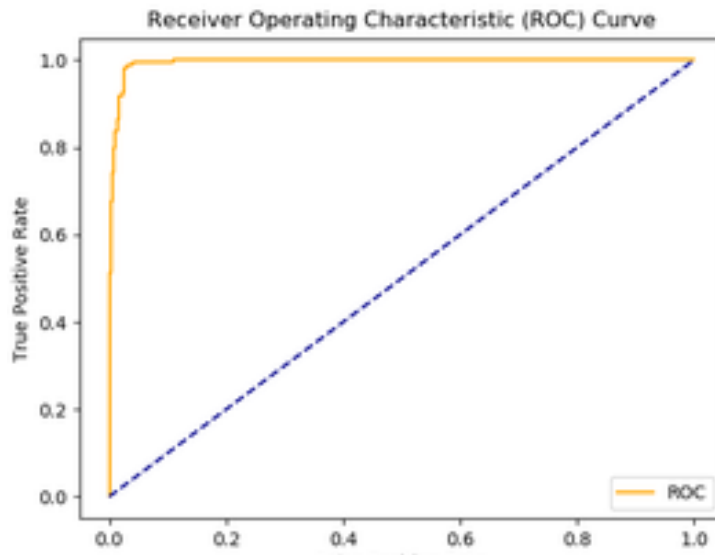


Figure 11 .6 Présentation roc curve

La courbe Présenté dans la figure nous montre les résultats de validation On voie clairement que les valeurs prédites de désabonnement sont très proches a des valeurs réelles ce qui rassure que notre système est fiable et performant

7 .Comparaison de résultat

	Arbre décision	Réseaux neurone
Accuracy	96.75	97.52
Décision	94.41	95 ,79

Tableau 2.7 : comparaison des résultats des mesures d'évaluation pour les algorithme d'apprentissage supervisé.

Comme indiqué dans le tableau, nous observons que les réseaux neurones ont surperformé l'autres algorithme testés avec une valeur de précision de 97.52%. L'arbre de décision avec une valeur de précision de 96.75%.

Les deux Modèles sont acceptables pour cette étude. Ils ont un bon pouvoir de prédiction plus que 90%

8 - conclusion :

nous avons présenté notre base de données sur laquelle nous avons testé notre système de prédiction du churn.

Dans ce chapitre nous avons présenté les différentes étapes qu'on a suivies pour arriver au bon développement et fonctionnement de notre système en utilisant les Réseaux neurone(pmc) et arbre décision qui sont des outils puissants pour modéliser la prédiction de désabonnement des client diCe chapitre présente aussi un ensemble des codes concernant notre système avec interprétations et explications des différents résultats obtenus. On peut dire que nous avons obtenu de bons résultats en termes d'erreur et de performance.

Conclusion général

Notre étude qui s'est porté sur « l'utilisation d'une méthode d'apprentissage automatique pour la détection et la prédiction de l'attrition des clients dans le secteur de télécommunications »

Le présent manuscrit montre l'importance de ce type d'études sur le marché des télécommunications dans le but d'aider les entreprises à accroître leurs bénéfices.

Ce projet a été réalisé en utilisant langage de programmation python, sous l'environnement de développement Google colaboratory

Afin obtenir de bons résultats en termes d'erreur nous avons implémenté deux algorithmes d'apprentissage, le perceptron multi couche et l'arbre de décision .

Nous avons observé que les résultats sur la base de test sont satisfaisants bien que les valeurs prédites approchent les valeurs réelles.

Résumé

Au cours de ces dernières années ,la prévision de l'attrition (churn) client dans les industries de télécommunications est devenue un sujet de recherche primordial. Les algorithmes d'apprentissage automatique ont été appliqués pour prédire le désabonnement des clients et ont montré des performances prometteuses par rapport aux méthodes traditionnelles. L'objectif de travail est de réaliser un système de prédiction de l'attrition de client dans le secteur de télécom en utilisant les arbre de décision , et réseaux de neurones artificiels.

Mot clé :

Désabonnement des clients, prévision, télécommunication, arbre de décision, réseaux de neurones artificiels.

abstract

in recent year , forecasting customer churn in the telecommunications industries has become a major reseaech topic. Machine learning algorithms have been applied to predict customer churn and have shown promising performance over traditional methods. The objective of this work is to realize system for predicting customer attrition in the telecom sector using decision trees, and artificial neural networks.

Key words :

Customer churn, forecasting, telecommunications, decision trees, artifical neural networks.

الملخص

في السنوات الأخيرة أصبح التنبؤ بتغير العملاء في صناعات الاتصالات موضوعا بحثيا رئيسا. تم تطبيق خوارزميات التعلم الآلي للتنبؤ بضغط العملاء و أظهرت اداء و اعدا مقارنة بالطرق التقليدية الهدف من هذا العمل هو تحقيق نظام للتنبؤ بتناقص العملاء في قطاع الاتصالات باستخدام أشجار القرار و الشبكات العصبية الاصطناعية .

الكلمات المفتاحية:

انسحاب العميل ، التنبؤ، الاتصالات السلكية و اللاسلكية ، أشجار القرار ، الشبكات العصبية الاصطناعية.

Remerciement

Nous remercies Allah de nous avoir donné le courage, la santé et la motivation
pour finir ce

projet de fin d'études.

Nous adressons nos premiers remerciements à notre professeur, Mme Saad
saoud Manel, pour

L'aide et les conseils que vous nous avez apportés tout au long du projet. Ce
travail ne sera pas

Riche sans ses idées fortes, ses discussions constructives et ses solutions
pertinentes

et l'échange de connaissances scientifiques. elle a toujours été très accueillante

Et le soutien est inestimable.

Nous tiens à remercier les membres du jury d'avoir bien accepté de consacrer
leur temps

pour évaluer ce travail modeste.

nous souhaitons vivement remercier nos famille pour leur soutien et leur
compréhension tout

au long de nos années d'études.

Enfin, NOUS n'oublions pas de remercier nos amis d'études et tous ceux qui
ont contribué de

près ou de loin dans ce projet

Bibliographie

- http://elearning.univjijel.dz/elearning/pluginfile.php/4333/mod_resource/content/1/SupportCours_Mokhtar_Taffar_ApprAuto.pdf.
- <https://tel.archives-ouvertes.fr/tel-00335313v2/document>. F.Souam Ait El Hadj, cour fouille de données, Master I, CPI, UMMTO 2018.
- [<http://chercheurs.lille.inria.fr/pgermain/publis/memoire.pdf>]
- http://eric.univlyon2.fr/~jahpine/cours/m2_dm-ml/cm.pdf.
- http://www.dmi.usherb.ca/~larocheh/cours/ift725_A2014/diapositives/ift615apprentissageautomatique.pdf.
- <http://Kaggle.com>
- <http://GOOGleschoolar.com>
- <http://wékepedia>
- C.-F. Tsai and Y.-H. Lu. Customer churn prediction by hybrid neural networks. Expert Systems with Applications, 36(10) :12547–12553, 2009a
- C.-F. Tsai and Y.-H. Lu. Customer churn prediction by hybrid neural networks. Expert Systems with Applications, 36(10) :12547–12553, 2009b.
- W. Verbeke, K. Dejaeger, D. Martens, J. Hur, and B. Baesens. New insights into churn prediction in the telecommunication sector : A profit driven data mining approach. European Journal of Operational Research, 218(1) :211–229, 2012
- G.M. ApurvaSree, S. Ashika, S. Karthi, V. Sathesh, M. Shankar, J. Pamina, Mar-Apr 2019 : "Churn Prediction in Telecom using".
- T. Vafeiadis, K. I. Diamantaras, G. Sarigiannidis, K. Ch. Chatzisavvas, February 2011 : A "Comparison of machine learning techniques for customer churn prediction".