université de Carthage

Institut National de sciences appliquées et de technologies



Rapport de Projet

L'intelligence artificielle utilisée dans l'amélioration des performances des réseaux haut débits

Réalisé par :

Amira Bedoui Arij Hraghi

Enseignant:

мr, Kamel karoui

Année Universitaire

2022/2023

Table des matières

Introduction	générale	4
Chapitre 1:	cadre général du projet	5
I.	Introduction	5
I.	Présentation du projet	5
II.	Concepts fondamentaux	5
	i. Réseau haut débit	5
	ii. Intelligence Artificielle AI	6
III.	Conclusion	9
Chapitre 2:	Etude du rôle le l'AI dans l'amélioration des performances de réseaux	
II.	Introduction	10
III.	Relation entre l'IA et les réseaux HD	
IV.	Avantage de l'IA dans les réseaux HD	10
V.	Exemple de l'application de l'IA pour renforcer les réseaux HD	11
	i. La maintenance prédictive	11
	ii. L'optimisation des équipements et du réseau en temps réel	11
	iii. L'optimisation de la qualité de service	11
	iv. Sécurité	12
IV.	Conclusion	13
Chapitre 3:	Réalisation d'un modèle de détection du spam	14
I.	Introduction	14
II.	Implémentation	14
	i. Méthodologie	14
	ii. Dataset	14
	iii. Algorithme utilisé	16
	iv. Outils utilisés	17
	v. Explication du modéle	17
III.	Conclusion	19
Conclusion 9	rénérale	20

Liste des figures

Figure 1 : qualité	é de service	9
Figure 2 : Métho	dologie de travail	15
Figure 3 : Youtul	be Dataset	16
Figure 4 : formul	le de naive bayes	16
Figure 5 : Jupyer	notebook	17
Figure 6 : Pandas	s Library	17
Figure 7 : Numpy	y Library	17
Figure 8 : Maplo	tlib library	17
Figure 9 : Bibliot	théques	17
Figure 10 : Datase	et importation	18
Figure 11 : Train_	_test_split	18
Figure 12 : Tokeni	ization	18
Figure 13 : Model		19
Figure 14 : Predict	tion	19
Figure 15 : Perfori	mances	19

Introduction générale

Les utilisateurs cherchent un accès à Internet performant permettant d'envoyer et de recevoir un grand nombre de données (documents, photos, vidéos, etc.) en **temps réel** et en des durées très courtes.

Cette augmentation exponentielle de l'usage de l'internet, des architectures de **type client et serveurs** et du **contenu multimédia** demande une énorme bande passante exige l'apparition des réseaux haut débit pour satisfaire ce besoin gourmand.

Ainsi, la révolution technique du à l'apparition de **l'intelligence artificielle** peut servir dans le domaine des réseaux haut débit en **améliorant ses performances**.

En énumérant les différentes techniques utilisées, nous allons mettre en relief et souligner le besoin de l'emploi des techniques de l'IA afin de sécuriser et booster le rendement des réseaux haut débit.

Chapitre 1:

Cadre général du projet

I. Introduction:

Dans ce chapitre, on présentera le cadre générale du projet ainsi que les définitions des termes génériques réseaux haut débit et l'intelligence artificielle et leurs concepts fondamentaux en mettant en évidence leur importance et leurs rôles primordiaux.

II. Présentation du projet :

Ce projet est un travail de recherche dont lequel on a étudié l'évolution des réseaux haut débit et son utilisation inévitable par la suite on a mis en relief l'importance de l'intelligence artificielle et son rôle pour augmenter les performances des réseaux haut débit en prenant multiples exemples comme la sécurisation des réseaux haut débit .

III. Concepts fondamentaux:

i. Réseaux Haut Débit

Un réseau à très haut débit (THD) est un réseau d'accès à Internet qui permet d'envoyer et de recevoir un énorme volume de données, des multimédias dans des courtes durées.

Débit : la quantité d'informations que le réseau permet de transférer en un temps donné. Il est exprimé en bit /s.

En augmentant le débit, la vitesse de téléchargement des données(DOWNLINK) est plus rapide. Ainsi, le débit et la vitesse de téléchargement des données sont proportionnels.

Les technologies:

Les principales technologies utilisées par les clients aux réseaux haut-débit sont :

• FTTH, fiber to the home, c'est la fibre optique chez l'abonné.

- le câble, introduit le premier historiquement par les fournisseurs de TV par câble, présent surtout dans les zones urbaines
- le satellite, qui a un débit soutenu intéressant en voie DOWNLINK (voie descendante), mais a des désavantages énormes principalement la latence qui affecte l'utilisation des applications en communication temps réel citant à titre d'exemple les jeux vidéo joué en ligne (Sensible à la variation de délai) et la téléphonie IP.
- la DSL, technologie s'appuyant sur le réseau téléphonique traditionnel.
- WiFi
- les courants porteurs en ligne, qui utilisent le réseau électrique de distribution.[1]

Théoriquement, ces techniques sont capables d'apporter des services multiples multimédia à haut débit mais pratiquement, la qualité de service n'est pas stable. Cette instabilité est due à plusieurs causes :

Les principales métriques des réseaux haut débit:

- Taux de perte : Nombre de paquets perdu lors de paquets.
 - o Pertes à cause d'un médium de mauvaises qualités
 - o Congestion.
- Latence
- l'affaiblissement du signal : offrir un débit faible dans le cas où le client est loin du nœud du raccordement.
- La modification de l'ordre d'arrivée des paquets.

ii. Intelligence Artificielle

Intelligence artificielle

 $m{L}$ 'intelligence artificielle (IA) est un processus d'imitation de l'intelligence humaine qui repose sur la création et l'application d'algorithmes exécutés dans un environnement informatique dynamique. Son but est de permettre à des ordinateurs de penser et d'agir comme des êtres humains.[2]

De nos jours, êtres humains et machines génèrent des données plus vite qu'il n'est humainement possible de les absorber et de les interpréter pour prendre des décisions complexes. L'intelligence artificielle est la base de tout apprentissage par un ordinateur et représente l'avenir des processus décisionnels complexes.

Fonctionnement de l'intelligence artificielle

L'IA fait recours a plusieurs à plusieurs techniques notamment les machine Learning qui désignent la science par laquelle une machine peut apprendre sans être explicitement programmée , les Deep Learning qui repose sur un réseau de neurones artificiels qui imitent notre cerveau et un ensemble d'algorithmes qui réalise des taches complexes .

L'IA peut fonctionner qu'avec des ordinateurs avec des capacités de calcul performantes grâce au cloud et des algorithmes révolutionnaires.

Machine Learning:

Le machine learning est la branche de l'IA qui connaît le plus de développement, et c'est également celle qui est la plus confondue avec l'IA elle-même. C'est un champ assez ancien de l'IA, dont les premières avancées techniques datent des années 1950. Depuis une quinzaine d'années, le domaine connaît un fort renouveau grâce à l'augmentation de la puissance de calcul et de la quantité de données disponibles: c'est l'avènement du Big Data. D'un point de vue technique, le machine learning est une technologie basée sur les statistiques permettant à des systèmes d'apprendre de manière autonome à partir de données, sans avoir été programmés au préalable pour cet apprentissage. Au lieu de formaliser des règles manuellement, c'est le modèle qui les déduira à partir des données.

Fonctionnement du machine learning

Le machine learning ou encore apprentissage automatique en français est un ensemble d'algorithme essentiellement :

• *Apprentissage supervisé* : la machine apprend à partir d'un ensemble de données étiquetées

L'apprentissage supervisé répond essentiellement à deux types de problèmes :

Régression

La régression sert à trouver la relation d'une variable par rapport à une ou plusieurs autres. Dans l'apprentissage automatique, le but de la régression est d'estimer une valeur (numérique) de sortie à partir des valeurs d'un ensemble de caractéristiques en entrée.

Il existe plusieurs algorithmes pour la régression:

- Régression linéaire
- Régression polynomiale
- Régression logistique
- Régression quantile

Classification

La classification est le processus de recherche ou de découverte d'un modèle ou d'une fonction qui aide à séparer les données en plusieurs classes catégorielles, c'est-à-dire des valeurs discrètes. Dans la classification, les données sont classées sous différentes étiquettes en fonction de certains paramètres donnés en entrée, puis les étiquettes sont prédites pour les données

- . Il existe plusieurs algorithmes pour la classification:
- Naïve bayes
- Support Vector Machine SVM

- Arbre de décision
- Foret d'arbres
- *Apprentissage non supervisé* : la machine apprend à partir d'un ensemble de données non étiquetées, il détermine lui-même la structure de l'entrée

Il existe plusieurs types d'algorithmes non supervisé

- Clustering : l'objectif consiste à trouver des regroupements dans les données.
- Association : l'objectif consiste à identifier les règles qui permettront de définir de grands groupes de données
- Apprentissage par renforcement : Il s'agit de prendre des mesures appropriées pour maximiser la récompense dans une situation particulière.

Il existe plusieurs types d'algorithmes par renforcement

- Monte carlo : le programme reçoit ses récompenses à la fin de l'état « terminal ».
- Machine learning par différence temporelle : les récompenses sont évaluées et accordées à chaque étape.

Deep learning

L'apprentissage en profondeur est un sous-ensemble de l'apprentissage automatique, qui est essentiellement un réseau de neurones à trois couches ou plus.[3]

Fonctionnement du deep learning

Les réseaux d'apprentissage en profondeur apprennent en découvrant des structures complexes dans les données qu'ils expérimentent .

Chaque algorithme de la hiérarchie applique une transformation non linéaire à son entrée et utilise ce qu'il apprend pour créer un modèle statistique en sortie. Les itérations se poursuivent jusqu'à ce que la sortie ait atteint un niveau de précision acceptable. Le nombre de couches de traitement par lesquelles les données doivent passer est ce qui a inspiré le label deep.

Utilisation de l'IA par les entreprises

Selon « Harvard Business Review », les entreprises utilisent principalement l'IA pour :

- Détecter et dissuader les intrusions de sécurité (44 %)
- Résoudre les problèmes technologiques des utilisateurs (41 %)
- Réduire les tâches de gestion de la production (34 %)
- Évaluer la conformité interne lorsqu'elle fait appel à des fournisseurs approuvés (34 %) [4]

IV. Conclusion

Dans ce chapitre, nous avons présenté l'idée générale du projet, nous avons aussi élaborés concepts fondamentaux .Le prochain chapitre est consacré à l'analyse de la relation de l'AI avec les réseaux Haut débit

Chapitre 2:

Rôle de l'IA dans l'amélioration des performances des réseaux haut débit

I. Introduction:

Dans ce chapitre, nous allons présenter l'intervention de l'AI dans les réseaux haut débit en développent les avantages ainsi que les inconvénients.

II. Relation entre L'AI et le réseau HD

Sous le but d'optimiser la gestion du réseau et la QoS vue la complexité des réseaux et la difficulté dans leurs surveillances, l'intelligence artificielle est devenue une nécessité dans l'amélioration des performances de leurs services.

Grâce à la quantité très importante de données que fournissent les réseaux de télécommunications sur leur fonctionnement, et grâce aux modèles prédictifs de l'IA, sont apparus de nombreux cas d'utilisation de l'IA dans la gestion des réseaux.

III. Avantage de l'AI pour les réseaux HD

L'IA est parfaitement adaptée pour résoudre certains de nos problèmes quotidiens dont le réseau et l'un de ces problèmes.

L'IA peut aider à acheminer le trafic et à équilibrer les charges dans le réseau. Elle améliore ses performances, ainsi que la simplification des déploiements et la qualité du service.

L'intégration de l'IA vise principalement à réduire les dépenses d'investissement, optimiser les performances des réseaux et créer de nouvelles sources de revenus.

Interrogées par Ericsson, 55 % des décideurs d'entreprises mondiales ont déclaré que l'IA est déjà adoptée pour améliorer l'expérience client. En effet, l'IA renforce la qualité du réseau et propose des services personnalisés.[5]

IV. Exemple de l'application de l'IA pour renforcer les réseaux HD

i. La maintenance Prédictive

Certaines défaillances techniques (par exemple indisponibilité de ressources, dysfonctionnement de VPN, etc.) peuvent déjà être identifiées, voire anticipées. Cette anticipation s'appuie sur des jeux de données décrivant les dysfonctionnements antérieurs, et en connaissant les impacts utilisateurs.

ii. L'optimisation des équipements et du réseau en temps réel

Comme lorsqu'elle cherche à répondre à des incidents, l'intelligence artificielle peut considérer l'ensemble des modifications qu'elle peut apporter à la configuration du réseau et simuler le déploiement de ces modifications afin d'en évaluer les effets. L'intelligence artificielle peut notamment servir à optimiser la consommation énergétique des réseaux en mettant en veille les équipements (antennes, serveurs, fibres...) ayant peu de chance d'être utilisés ou au contraire en leur allouant dynamiquement plus de ressources. L'apprentissage automatique peut aussi être associé au Software Defined Network (SDN) afin de configurer le réseau de manière globale et cohérente (placement des fonctions dans le réseau, estimation des ressources à allouer, etc.).

iii. L'optimisation de la qualité de service

La qualité de service :

La qualité de service (QoS) fait référence à toute technologie qui gère le trafic de données pour réduire la perte de paquets , la latence et la gigue sur un réseau. QoS contrôle et gère les ressources du réseau en définissant des priorités pour des types de données spécifiques sur le réseau.



- Figure1 : qualité de service -

un incident désigne est un tout événement qui ne fait pas partie du fonctionnement standard d'un service et qui cause une interruption ou une diminution de la qualité de ces services (QoS) et dans ce contexte, la gestion des incidents est caractérisée par un objectif qui vise à « Restaurer aussi vite que possible le fonctionnement normal des services, minimiser l'impact négatif sur les activités Métier et s'assurer ainsi que les meilleurs niveaux de qualité de service et de disponibilité sont maintenus »[6]

• Intervention de l'IA dans la gestion et la réduction des accidents :

La puissance de l'intervention d'une l'IA « forte » analysant les champs non structurés de la base des incidents / tickets et s'interconnectant avec les logs serveurs ou réseau, permet d'effectuer une corrélation des symptômes remontés et donc de définir les points accidente gènes des applications étudiées ou du SI analysé permettant de réaliser une entimiention de la redirection des tielets vers les bons

de réaliser une optimisation de la redirection des tickets vers les bons interlocuteurs (c'est-à-dire ceux en capacités de résoudre l'incident) sans passe-plat pour diminuer le nombre d'interlocuteurs et donc le temps de traitement des incidents ou à plus long terme à de la maintenance préventive par anticipation.

De plus, grâce au développement exponentiel du deep learning (méthode d'apprentissage approfondi et automatique fondée sur l'apprentissage de modèles de données) l'efficacité de l'analyse faite par la machine sera en constante progression de sorte qu'au bout d'un certain temps, plus aucune intervention humaine ne sera nécessaire dans la résolution des incidents.

Cette méthode permet d'automatiser jusqu'à 80% des tickets / incidents d'une entreprise ce qui permettrait de repositionner les hommes sur des tâches à plus fortes valeur ajoutée.

iv. Sécurisation du réseau

L'IA et l'apprentissage automatique (ML) [4]sont devenus des technologies essentielles pour la sécurité de l'information, car ils sont capables d'analyser rapidement des millions d'événements et d'identifier de nombreux types de menaces différents, des logiciels malveillants exploitant les vulnérabilités du jour zéro à l'identification des comportements à risque pouvant conduire à un hameçonnage. Ces technologies apprennent au fil du temps, s'inspirant du passé pour identifier dès maintenant de nouveaux types d'attaques.

En utilisant des algorithmes sophistiqués, les systèmes d'IA sont formés pour détecter les logiciels malveillants, exécuter la reconnaissance de modèles et détecter même les comportements les plus infimes des attaques de logiciels malveillants ou de rançongiciels avant qu'ils n'entrent dans le système.

V. Conclusion

Dans ce chapitre, nous avons spécifié comment l'AI améliore le fonctionnement du réseau haut en citant plusieurs exemples . Le prochain chapitre on va détailler l'exemple de sécurité .

Chapitre 3:

Réalisation d'un modèle de détection du spam

I. Introduction:

 $m{D}$ ans cette partie, on va expliquez le modèle qu'on a implémenté qui montre le rôle joué par l'AI pour assurer la sécurité en détaillant la démarche et l'algorithme qu'on a utilisé.

II. Implémentation:

Youtube est l'un des sites les plus utilisés et les plus populaires, les utilisateurs de youtube sont alloués de partager, de télécharger, d'évaluer, d'abonner ou encore de commenter sur un n'importe quel contenu .Cependant cela s'est transformé en une opportunité pour les utilisateurs malveillants de partager du contenu promotionnel connu sous le nom de spam. Youtube a tenté de lutter contre le spam avec des méthodes basiques mais les spammeurs trouvent toujours des techniques pour faire se contourner .

Cela indique la nécessité de disposer d'une technique capable de classer un commentaire comme spam ou non en se basant sur le contenu tout en restant résistant aux transformations.

Le travail est réalisé avec le youtube dataset qu'on va la détailler par la suite et l'algorithme de naïve bayes.

i. Méthodologie

1-Dataset: importation du dataset

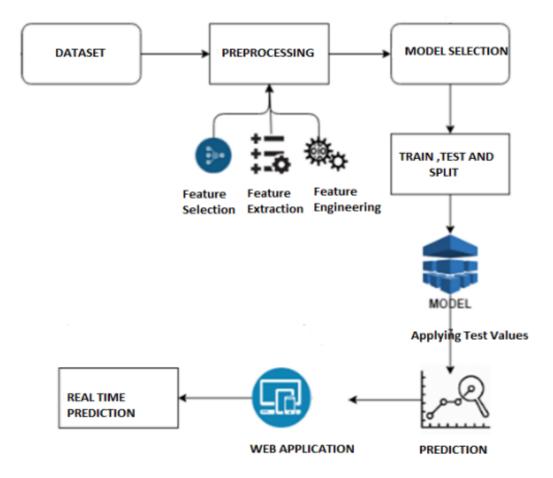
2- Preprocessing : c'est l'ensemble des transformations qu'on doit les effectuer pour que notre algorithme soit fonctionnel

3-Train_test_split : division du dataset en des échantillons pour le train et les autres pour le test pour que notre modèle soit testé sur des échantillons dont il ne connait pas

4-Model: application du modèle

5-Prédiction

6-Evaluation : c'est la ou l'évaluation de la performance de notre modèle aura lieu



- Figure2 : Méthodologie de travail-

ii. Dataset

On a utilisé la dataset du youtube qui est une dataset étiquetée qui est composé de :

- Un attribut « CONTENT » qui représente le contenu du commentaire.
- Un attribut « «COMMENT_ID »
- Un attribut « AUTHOR »
- Un attribut « DATE »
- La classe qui présente deux valeurs 0 ou 1, 0 pour désigner un spam et 1 pour désigner un ham.
- 1955 échantillons.

Data		
	CONTENT	CLASS
0	Huh, anyway check out this you[tube] channel:	1
1	Hey guys check out my new channel and our firs	1
2	just for test I have to say murdev.com	1
3	watch?v=vtaRGgvGtWQ Check this out .	1
4	Hey, check out my new website!! This site is a	1
1950	I love this song because we sing it at Camp al	0
1951	I love this song for two reasons: 1.it is abou	0
1952	wow	0
1953	Shakira u are so wiredo	0
1954	Shakira is the best dancer	0

1955 rows x 2 columns

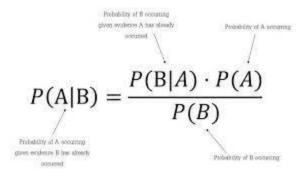
- Figure3 : Youtube Dataset-

iii. Algorithme du Naïve bayes

L'algorithme de naïve bayes est un algorithme de classification qui est basé sur une condition passée. Thomas bayes est celui qui a pensé a cet algorithme. Il est considérer comme naïve car il traite les ordres des mots de la même manière.

L'idée derrière cet algorithme est de calculer la probabilité a posteriori qui présente la probabilité qu'un échantillon x appartient à un classe y en se basent sur la probabilité à priori c'est-à-dire la probabilité avant La preuve et la probabilité de likelihood qui est la probabilité d'un échantillon dans une classe connue.

Naïve bayes est considéré parmi les algorithmes facile à implémenter, et qui travailler sur des applications en temps réel .



- Figure4 : formule de naïve bayes-

iv. Outils utilisés



- Figure5 : Jupyter Notebook-



- Figure6: Pandas library-





-Figure7: NumPy Library-

- Figure8 : MatplotLib lirary-

v. Explication du modèle

Pour l'implémentation du modèle on a suit la démarche suivante

• Importation des bibliothèques nécessaires

Entrée [98]:

import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import glob
from sklearn.feature_extraction.text import CountVectorizer
from sklearn.model_selection import train_test_split
from sklearn.aive_bayes import MultinomialNB
from sklearn.metrics import classification_report,confusion_matrix

- Figure9 : Bibilothéques-

 Importation du dataset : elle est composé de plusieurs portions donc on accède au chemin d'accès et on garde uniquement la colonne « CONTENT » dont on a besoin .

- Figure 10 : Dataset importation-

 Puisque la dataset est composé de plusieurs, On a subdiviser la data en test et train avec la méthode train_test_split du sickit learn

- Figure11 : Train_test_split-

• On a convertit l'ensemble des échantillons en une matrice avec l'outil CountVectorizer() pour évaluer l'apparition de chaque mots.

- Figure 12: Tokenization-

• On a réalisé la phase d'entrainement avec l'algorithme Naïve Bayes

```
Entrée [92]: M clf=MultinomialNB()

Entrée [94]: M clf=clf.fit(x_train_vect,y_train)## P(Y|xi)=P(Xi|y)*p(y)
```

- Figure 13: Model-

• Phase de prédiction

```
Entrée [95]: My_predicted=clf.predict(x_test_vect)
```

- Figure14 :Prédiction-
- Phase d'évaluation en utilisant la matrice de confusion qui nous permet d'interpréter la correspondance entre les classes prédites et celles réellement classées
- Avec Classification_report on peut voire que notre modèle présente une accuracy de 91%.

```
Entrée [101]: ▶ confusion_matrix(y_test,y_predicted)
   precision recall f1-score support
               1
                   0.89
                         0.95
                               0.92
                               0.91
                                     391
           accuracy
         macro avg
weighted avg
                    0.92
                          0.91
                        0.91
                    0.91
                               0.91
```

- Figure 15: Performances-

III. Conclusion:

Au bout de ce chapitre, on a implémenté un modèle de machine learning qui montre bien l'effet de l'intelligence artificielle dans la sécurité.

Conclusion générale

Ce projet était une bonne occasion pour entrer profondément dans l'univers de la machine learning en liaison avec les réseaux haut débit.

En effet effectuer une recherche dans un domaine d'actualité, est une tache ardue qui nécessite beaucoup d'effort pour parvenir à un résultat de qualité.

Sous ce rapport, la première phase était d'effectuer une recherche sur les mots clés du projet pour comprendre le contexte générale du projet

La deuxième phase était de faire la liaison entre entre l'intelligence artificielle et les réseaux haut débit dont on a découvert l'effet primordiale que L'AI peut porter afin d'améliorer les performances du réseau haut débit, on a focalisé le plus sur l'aspect de la sécurité.

C'est pourquoi dans la troisième phase on a implémenté un modèle qui permet de mettre en valeur en cet aspect.

Finalement, ce projet nous a permis d'approfondir nos connaissances dans les bonnes pratiques de la gestion de projet vu que nous avons eu l'opportunité d'organiser son déroulement dès le début. En Outre, il nous a permis de mettre en œuvre les acquis théoriques que nous avons appris tout le long de notre cursus universitaire et de les enrichir ainsi que l'approfondissement de connaissance dans le domaine de l'AI et du Réseaux.

Bibliographie

- [1]:https://www.technoscience.net/definition/3987.html, derniére visite le 20/11/2022
- [2]: https://www.netapp.com/fr/artificial-intelligence/what-is-artificial-intelligence/, derniére visite le 20/11/2022
- [3]: https://www.techtarget.com/searchenterpriseai/definition/deep-learning-deep-neural-network, derniére visite le 20/11/2022
- [4]: https://www.oracle.com/fr/artificial-intelligence/what-iai/#:~:text=l'intelligence%20artificielle,Intelligence%20artificielle%20(IA)%20%2D%20Exp lication,des%20informations%20qu'ils%20recueillent., derniére visite le 21/11/2022
- [5]: https://www.geeksforgeeks.org/ml-classification-vs-regression/, derniére visite le 21/11/2022
- [6]: https://www.silicon.fr/avis-expert/intelligence-artificielle-quel-role-dans-la-qualite-deservices-du-si , derniére visite le 21/11/2022

