الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne Démocratique et Populaire
وزارة التعليم العالي والبحث العلمي
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
جامعة غرداية
Université de Ghardaïa
كلية العلوم والتكنولوجيا
Faculté des Sciences et de la Technologie
قسم الرياضيات والاعلام الآلي
Département de Mathématiques et d'Informatique

N° d'enregistrement
/…./…./…./…./….

Mémoire

Pour l'obtention du diplôme de master

Domaine: Mathématiques et Informatique

Filière: Informatique

Spécialité: Systèmes Intelligents pour l'Extraction des Connaissances (SIEC)

Thème

Algorithme Bio-Inspiré Pour La Sécurité Dans Twitter

Présenté par
Boughellaba wissal & Djebrit hanane

Devant le jury composé de:

M. Slimane BELLAOUAR	MCB	Université de Ghardaïa	Président
M. Messaoud BETKA	MAA	Université de Ghardaïa	Examinateur
M. Abdelkader BOUHANI	MAA	Université de Ghardaïa	Encadreur
Mme. Nacéra BRAHIM	MAA	Université de Ghardaïa	Examinateur

Année Universitaire 2020/2021

ملخص

أصبحت مواقع التواصل الاجتماعي شائعة جدًا في السنوات الأخيرة. يستخدمها الاشخاص للعثور على أصدقاء جدد ، وتحديث الأصدقاء الحاليين بآخر أفكارهم وأنشطتهم. من بين هذه المواقع ، يعد تويتر هو الأسرع نموًا. تجذب شعبيتها أيضًا العديد من مرسلي البريد العشوائي الذين يتسللون إلى حسابات المستخدمين الشرعيين بكميات كبيرة من رسائل البريد العشوائي.

في هذا العمل ، نناقش بعض الميزات المستندة إلى المستخدم والقائمة على المحتوى والتي تختلف بين مرسلي البريد عشوائي والمستخدمين الشرعيين ، ثانيًا نناقش الطرق الفعالة لاكتشاف البريد العشوائي. يهدف هذا العمل إلى تطبيق الإلهام الحيوي على هذه المشكلة ، على وجه الخصوص خوارزمية التحسين (EHO) ، ومناقشة فعاليتها. بالرغم من عدم كفاءة الأجهزة ومحدودية بيئة العمل والوقت إلا أننا تحصلنا على نتائج مرضية.

الكلمات المفتاحية : تحسين رعي الأفيال (EHO) ، مستوحى من الحيوية ، اكتشاف البريد العشوائي ، ميتا هيوريستيك ، هيوريستسك ، تحسين اندماجي

ABSTRACT

Social networking sites have become very popular in recent years. people use them to find new friends, update existing friends with their latest thoughts and activities. Among these sites, Twitter is the fastest growing, its popularity also attracts many spammers who infiltrate the accounts of legitimate users with large amounts of spam messages.

In this work, we discuss some user-based and content-based features that are different between spammers and legitimate users. Secondly, we discuss effective methods for spam detection. This work aims to apply bio-inspired to this problem, in particular, optimization algorithm (EHO), discussing their effectiveness. Despite the inefficiency of the devices and the limited working environment and time, we obtain satisfactory results.

Key words : Elephant herding optimization (EHO), Bio-inspired, Detected spam, Meta-Heuristic, Heuristic, Combinatorial optimization

RÉSUMÉ

Les sites de réseaux sociaux sont devenus très populaires ces dernières années. Les gens les utilisent pour trouver de nouveaux amis, mettre à jour leurs amis existants avec leurs dernières pensées et activités. Parmi ces sites, Twitter est celui qui connaît la croissance la plus rapide, sa popularité attire également de nombreux spammeurs qui infiltrent les comptes d'utilisateurs légitimes avec une grande quantité de messages de spam.

Dans ce travail, nous discutons certaines caractéristiques basées sur l'utilisateur et sur le contenu qui différent selon la nature de l'utilisateur légitime ou spammeur. Ensuite, nous discutons des méthodes efficaces pour détection des spam. Ce travail vise à appliquer la bio-inspiration à ce problème, en particulier l'algorithme d'optimisation (EHO), en discutant leur efficacité. Malgré l'inefficacité des dispositifs et la limite de l'environnement et le temps de travail nous avons obtenu des résultats satisfaisants.

Mots clés : Optimisation d'élevage d'éléphants (EHO), Bio-inspiré, Spam détecté, Méta-Heuristique, Heuristique, Optimisation combinatoire

TABLE DES MATIÈRES

TABLE DES FIGURES

Remerciement

Nous remercions tout d'abord Allah SWT, tout puissant, de nous avoir donné la volonté, l'aide, la patience et le courage pour accomplir ce projet, ainsi que la force et l'audace pour dépasser toutes les difficultés.

Nous tenons à remercier l'encadreur de cette thèse, M.BOUHANI Abdelkader pour sa suggestion de ce sujet et ses précieux conseils.

Nous remercions profondément M.ZIADi Djaloul pour son soutien et aide.

Nous remercions les membres du jury d'avoir accepté l'évaluation de notre travail. Nous tenons également à exprimer nos sincères remerciements à notre professeurs M.BELLAOUAR Slimane et M.OULAD NAOUI Slimane , qui a toujours été une source d'inspiration pour nous, et que nous avons beaucoup d'appréciation et de gratitude pour son enseignement, ses conseils et encouragements.merci vraiment pour vos efforts incroyables.

Enfin, nous adressons nos grands remerciements à tous ceux qui nous ont aidés et encouragés lors de la réalisation de ce travail.

Dédicacions

Avec l'expression de ma reconnaissance, je dédie ce modeste travail à ceux qui, quels que soient les termes embrassés, je n'arriverais jamais à leur exprimer mon amour sincère.

A l'âme de mon père

A ma chère mère,La femme qui a souffert sans me laisser souffrir, qui n'a jamais dit non àmes exigences et qui n'a épargné aucun effort pour me rendre heureuse

A mes frère "Moaad" & "Ahmed", qui n'ont pas cessée de me conseiller, encourager et soutenir tout au long de mes études. Que Dieu les protège et leurs offre la chance et le bonheur

A Ma sœur "Soulef" et ma petite chérie "Fatima",la joie et le bonheur de la maison

A mes oncles et mes tantes. Que Dieu leur donne une longue et joyeuse vie

A mes chère amis que j'ai connu jusqu'à maintenant, Merci pour leurs amours et leurs encouragements

Sans oublier ma binôme et ma chère amie "hanane" pour sa soutien moral, sa patience et sa compréhension tout au long de ce projet

Wissal Boughellaba

Dédicacions

Je dédie ce travail :
A mon cher père
mon idole, la source de ma joie et espoirs A ma chère mère, qui m'a donné tout le soutien et l'amour
A ma frères "immad" , Ma sœur "meriem"
A toute la famille "djebrit "
A toute la famille "Boukanoun "
A mon partenaire et compagnon "wissal"
Tous mes chers amis "Safa","Ihsen", mes collègues
Et enfin, je remercie Dieu tout puissant pour ma vie et ce voyage dont je suis fier.

DJEBRIT HANANE

INTRODUCTION

Depuis la démocratisation de l'Internet, les sites de réseautage social ont commencé avec sixdegrees.com en 1997, puis makeoutclub.com en 2000, Sixdegrees.com et d'autres sites de ce type n'ont pas pu survivre et ont disparu très rapidement, mais de nouveaux sites comme MySpace, LinkedIn, Bebo, Orkut, Twitter, Avec des bases de données d'utilisateurs plus importantes dans les réseaux sociaux en ligne (Online Social Network ONS), ils deviennent des cibles plus intéressantes pour les spammeurs /utilisateurs malveillants.

Le spam peut prendre différentes formes sur les sites Web sociaux et n'est pas facile à détecter. Toute personne qui connaît Internet a été confrontée à une forme ou une autre de spam, qu'il s'agisse de spam par e-mail, de spam sur des forums, des groupes de discussion, etc. Le spam est défini comme l'utilisation d'un système de messagerie électronique pour envoyer des messages non sollicités en masse.

Ainsi, les spammeurs sont attirés par l'utilisation de Twitter comme outil pour envoyer des messages non sollicités à des utilisateurs légitimes, publier des liens malveillants et détourner des sujets d'actualité. Le spam devient un problème croissant sur Twitter ainsi que sur d'autres sites de réseaux sociaux en ligne.

Notre mémoire se compose de trois chapitres, Dont le premier chapitre contient les notions de base de l'intelligence artificielle, l'apprentissage automatique et le problème d'optimisation combinatoire puis nous exposons la bio-inspiré dans laquelle nous présentons son historique et ses algorithmes, enfin nous indiquons l'algorithme que nous allons

utilisé.

Le deuxième chapitre, nous présentons l'état de l'art de la détection spam dans TWITTER, Nous expliquons plusieurs concepts tels que le problème de sécurité, les types de spam, les caractéristiques de la détection spam et comment le TWITTER traite les spam Puis nous parlons de différentes méthodes pour la détection à savoir des méthodes d'apprentissage automatique et profonde.

Le dernier chapitre présenté le développent d'algorithme EHO et ses étapes pour résoudre le problème de détection spam dans TWITTER et enfin nous terminons par une conclusion qui résume tout ce qui précède.

CHAPITRE 1

NOTIONS PRÉLIMINAIRES

1.1 Introduction

Les nouvelles approches de l'intelligence artificielle découlent de l'idée que l'intelligence émerge autant des cellules, des corps et des sociétés, plus que du développement et de l'apprentissage. Traditionnellement, les nouvelles approches s'inspirent d'un éventail plus large de structures biologiques capables d'auto-organisation autonome. Des exemples de ces nouvelles approches comprennent le calcul évolutif et l'électronique évolutive, les réseaux de neurones artificiels, les systèmes immunitaires artificiels, la biorobotique et l'intelligence en essaim[13].

Dans ce chapitre nous verrons des définitions et des concepts sur l'intelligence artificielle et l'apprentissage automatique en montrant la relation entre eux, nous mettons aussi en évidence l'efficacité d'algorithme d'essaim à partir de ses origines, puis nous expliquons les problèmes combinatoires NP- difficiles et les méthodes de ses résolutions, par la suite Nous passons également de présenter la Bio-inspiration et son historique, finalement définie l'algorithme d'optimisation EHO et ses principes de base.

1.2 Intelligence Artificielle

On peut connaître la météo, mettre une alarme, être prévenu de nos événements, avec les itinéraires pour s'y rendre et des résultats de plus en plus précis au fil du temps.Tel est le nouveau visage de l'intelligence artificielle dans notre vie courante [14], L'intelligence artificielle peut être définie comme un ensemble de techniques visant à permettre aux machines d'imiter une forme d'intelligence réelle.

L'intelligence artificielle est utilisée dans un nombre grandissant de domaines d'application.La propriété principale de l'intelligence artificielle est sa capacité a rationaliser et a prendre des mesures qui ont les meilleures chances d'atteindre un objectif spécifique.[15]

1.3 Apprentissage Automatique

L'apprentissage automatique (machine learning) est un sous-domaine de l'intelligence artificielle, Elle permet aux machines d'apprendre sans être programmées spécialement pour le faire. Les algorithmes d'apprentissage automatique adoptent un modèle mathématique basé sur un ensemble de données, appelé " données d'apprentissage ", afin d'en extraire les règles qui les rendent capables d'apprendre et de prédire dans le futur [16].Les algorithmes d'apprentissage automatique sont classés en plusieurs types, nous expliquons ici quelques types communs mentionnés dans Figure 1.1.

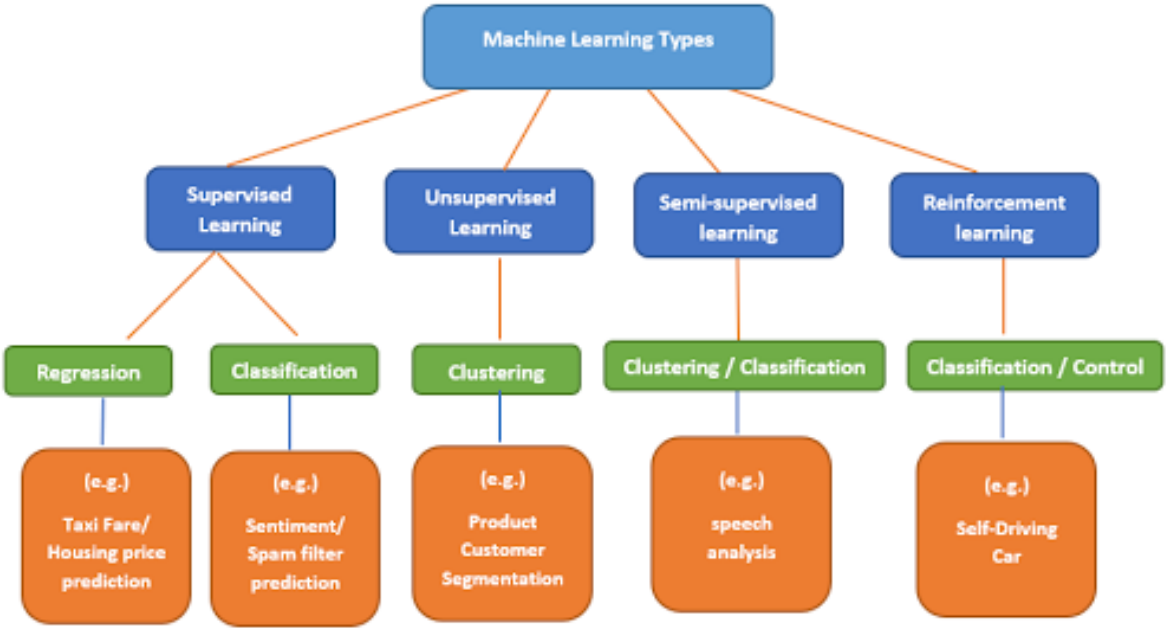

FIGURE 1.1 – Type d'apprentissage[1]

1.3.1 Apprentissage Supervisé

Ce type d'apprentissage utilise comme entrée des données qui sont étiquetés avec les sorties souhaitées. L'algorithme apprend en comparant la sortie réelle de l'entrée avec la sortie prédite, pour trouver des erreurs et modifier le modèle en conséquence jusqu'à arriver au meilleur taux de reconnaissance[16].

1.3.2 Apprentissage Non Supervisé

Ce type d'apprentissage utilise comme entrée des données non-étiquetées, de sorte que l'algorithme d'apprentissage trouve tout seul des points communs entre ses données. L'objectif de l'apprentissage non supervisé peut être aussi simple que découvrir des modèles cachés dans un ensemble de données, mais il peut aussi avoir un objectif d'apprentissage des caractéristiques, qui permet à la machine intelligente de découvrir automatiquement les représentations nécessaires pour classer les données brutes [16].

1.3.3 Apprentissage Semi-Supervisé

Ce type d'apprentissage utilise à la fois des données étiquetées et non-étiquetées pour s'entraîner. En règle générale, une petite quantité de données étiquetées est utilisée avec une grande quantité de données non-étiquetées. Pour cause, les données non étiquetées sont moins chères et plus faciles à obtenir. Ce type d'apprentissage peut être utilisé avec des méthodes comme la classification, la régression et la prédiction[16].

1.3.4 Apprentissage Par Renforcement

Ce type d'apprentissage multiplie les tentatives pour tenter de découvrir quelles actions apportent les plus grandes récompenses. Ce type d'apprentissage regroupe trois principaux composants : l'agent (qui apprend ou prend les décisions), l'environnement (tout ce avec quoi l'agent interagit), et les actions (ce que peut faire l'agent)[16].

1.4 Problème d'optimisation Combinatoire

Les Problèmes d'optimisation combinatoire sont des problèmes d'optimisation dont les ensembles réalisables sont finis mais combinatoires. Aussi le nombre de solutions réalisables des problèmes combinatoires augmente de façon exponentielle avec la taille du problème, ce qui exclut les méthodes de résolution basées sur l'énumération de toutes les solutions réalisables[17].

En mathématiques, l'optimisation combinatoire désigne l'ensemble des méthodes qui permettent de déterminer l'optimum d'une fonction avec ou sans contraintes.[18]

Un problème d'optimisation combinatoire (PIC) peut être défini comme suit :

- Un ensemble de solutions X
- Un ensemble de contraintes C
- Un sous-ensemble S de X correspondant aux solutions réalisables qui satisfont les contraintes C
- Une fonction de coût f (fonction objective) qui attribue à chaque solution $s \in S$ une valeur f(s).

Le but est de trouver une solution optimale $s^* \in S$ qui optimise (minimise ou maximise) la fonction de coût f

$$\mathbf{f(s^*) \leqslant f(s_i), \forall s_i \in S}$$

La résolution d'un problème d'optimisation combinatoire nécessite l'étude des éléments suivants :

- Définir l'ensemble des solutions "X"
- Exprimer l'ensemble des contraintes du problème "C" pour définir l'ensemble des solutions réalisables "S"
- Exprimer la fonction objective "f" à optimiser
- Choisir la méthode de résolution à appliquer

Les trois premiers points sont liés à la modélisation du problème, le dernier élément à sa résolution.

Chaque problème d'optimisation peut être associé à un problème de décision dont le but est de déterminer s'il existe une solution pour laquelle la fonction objectif est inférieure (ou supérieure) à une valeur donnée.

La complexité d'un problème d'optimisation est relative à la complexité du problème de décision associé. En effet, si le problème de décision est NP-complet, alors le problème d'optimisation est NP-difficile Cela a encouragé les chercheurs à développer des méthodes de résolution en recherche opérationnelle et en intelligence artificielle : Les méthodes exactes, Les méthodes approchée.

1.4.1 Méthodes Exactes

Ils sont généralement basés sur une recherche complète de l'espace de combinaison pour trouver une solution optimale.

Les Algorithmes exacts les plus performants de la littérature appartiennent aux quatre paradigmes : Les méthodes de séparation et d'évaluation, méthodes rétrospectives, La programmation dynamique et la programmation linéaire[18].

1.4.2 Méthodes Approchées

Ils permettent de retrouver une bonne solution (pas nécessairement optimale) en un temps raisonnable. Son objectif est de trouver une solution acceptable en un temps raisonnable, mais sans garantir l'optimalité de cette solution. Le principal avantage de ces méthodes est qu'elles peuvent être appliquées à n'importe quelle classe de problèmes, faciles ou très difficiles. De plus, elles ont démontré leur robustesse et leur efficacité face à plusieurs problèmes d'optimisation combinatoire. Elles regroupent deux classes : Les heuristiques et les méta-heuristiques[18].

Heuristique

Ce sont de simples règles empiriques basées sur l'expérience, ne fournissant pas nécessairement une solution optimale[18].

Méta-Heuristique

Le mot Méta-Heuristique est dérivé de la composition de deux mots grecs :
- Heuristique qui vient du verbe heuriskein () et signifie "trouver".
- Méta qui est un suffixe signifiant "au-delà", "à un niveau supérieur".

La Méta-heuristique peut être définie comme une méthode algorithmique capable de guider le processus de recherche dans un espace de solutions (souvent très grand) vers des régions riches en solutions optimales dans le but de trouver des solutions, peut-être pas toujours optimales, mais en tout cas très proches de l'optimum, dans un temps raisonnable[18].

Ils sont classés en deux classe principales en fonction du nombre de solutions :[2]

- Méta-heuristiques à solution unique.
- Méta-heuristiques à population de solutions.

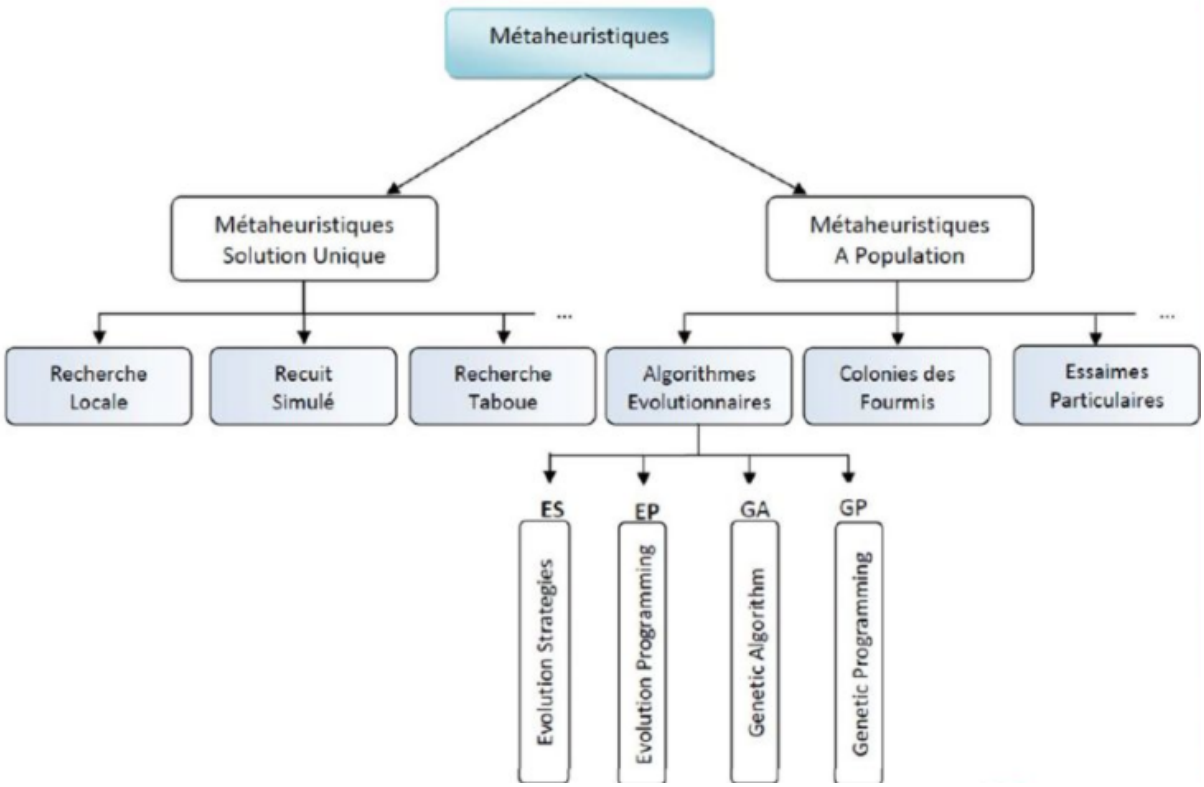

FIGURE 1.2 – Classe principaux de Méta-heuristique[2]

Application en intelligence artificielle

Dans cette partie, nous évoquons leur application à deux problèmes centraux en intelligence artificielle sont[2] :

1. La satisfiabilité de formules booléennes (sat) : est un problème de décision défini par des formules logiques. Il s'agit, étant donné une formule de logique propositionnelle, de décider si cette formule possède une solution, c'est-à-dire s'il existe une assignation des variables rendant la formule vraie.

2. La satisfaction de contraintes (csp) : est un problème modélisé sous la forme d'un ensemble de contraintes posées sur des variables, chacune de ces variables prenant ses valeurs dans un domaine. Résoudre un CSP consiste à affecter des valeurs aux variables, de telle sorte que toutes les contraintes soient satisfaites.

1.5 Bio-Inspirés

La bio-inspiration est un transfert de paradigme qui conduit les concepteurs à s'inspirer de la nature pour développer de nouveaux systèmes. La bio-inspiration est souvent basée sur le bio-mimétisme, Lequel elle peut s'inspirer du monde végétal, animal, fongique, bactérien et viral, et a déjà contribué à des applications dans des domaines aussi variés que l'aéronautique, la pharmacie, les sciences marines, la médecine, la chimie verte, les matériaux composites, la robotique, l'intelligence artificielle et les nanotechnologies[19].

1.5.1 Historique

Au fur et à mesure des connaissances scientifiques sur le vivant progressent, Toutes ces innovations technologique sont conçues en s'inspirant de phénomènes naturels : les fleurs de bardane pour le velcro, le système de thermorégulation des termitières pour l'eastgate building et les pouvoir hydrophobes de la fleur de lotus pour la peinture. Comme l'écrivait il y a 500 ans, **Léonard de Vinci** « scrute la nature, c'est là qu'est ton futur ». C'est à partir de ce précepte qu'est née la bio-inspiration et le Toscan est l'un des premiers à s'être inspiré des êtres vivants pour concevoir des machines comme l'ornithoptère, une aile volante issue de l'observation du vol de l'oiseau.

Aujourd'hui, la bio-inspiration est un véritable phénomène qui inspire à la fois les chercheurs qui inventent de nouveaux matériaux, En décodant et en imitant certaines propriétés du vivant, les chercheurs et les ingénieurs cherchent des réponses et des solutions à des défis technologiques. L'exemple le plus parlant est le nez du train à grande vitesse japonais, le Shinkansen. Il a été copié sur le bec du martin-pêcheur afin de réduire la perte de vitesse et la pollution sonore dues à la compression de l'air lors du passage du train dans les tunnels.

La bio-inspiration s'incarne dans trois grands domaines : dans celui des formes et des structures (architecture, design), dans celui des procédés et des matériaux et dans celui des organisations et des systèmes. L'objectif est de proposer des innovations performantes. Le bio-mimétisme est défini par **Janine Benyus** du Biomimicry Institue comme une philosophie dont l'ambition est de prendre la nature pour modèle afin de relever les défis du développement durable[20].

1.5.2 Informatique Bio-inspirés

Les méthodes bio-inspirées ont récemment gagné en importance en informatique en raison de la nécessité de disposer de moyens flexibles et adaptables pour résoudre les problèmes d'ingénierie. Les algorithmes bio-inspirés sont basés sur la structure et le fonctionnement des systèmes naturels complexes et permettent de résoudre les problèmes de manière adaptative et distribuée[21].

1.5.3 Motivation de L'utilisation Bio-inspirés

La nature de ces phénomènes extraordinaires nous fournit des solutions grâce à des caractéristiques telles que :

- L'émergence : des éléments simples quand interagissent entre eux sont réaliser des tâches extra-ordinaires.
- L'auto-organisation : l'organisation interne du système se structure automatiquement sans être dirigée par une source externe.
- La modularité : le système est composé d'éléments simples qui coopèrent entre eux pour atteindre l'objectif global.
- Décentralisation : elle assure la robustesse du système, capable de continuer à fonctionner en cas de défaillance d'un de ses composants.
- Réactivité : les éléments du système coopèrent et communiquent entre eux par des interactions locales.
- L'auto-adaptation : la capacité d'un système à modifier ses paramètres afin de continuer à fonctionner de manière satisfaisante malgré les variations de son environnement.

1.5.4 Processus de Création d'Algorithme Inspiré du la Nature

L'homme s'inspire de la nature pour développer une observation sur un phénomène naturel.[3] Il commence par sa modélisation en utilisant des simulations mathématiques. Une fois le modèle est raffiné, il sera utilisé pour extraire une méta-heuristique,nous expliquons ici le passage d'un phénomène naturel à un algorithme inspiré de la nature dans Figure1.3.

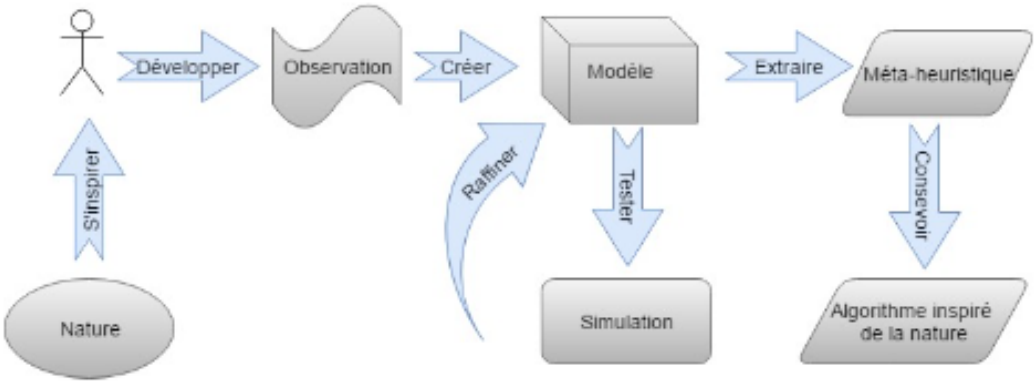

FIGURE 1.3 – Passage d'un phénomène naturel à un algorithme inspiré de la nature[3]

1.5.5 Classification d'Algorithmes Bio-inspirés

Les méthodes bio-inspirés peuvent être réparties en deux grandes classes selon la source d'inspiration de la méthode bio-inspiré, est représenté sur la Figure 1.4.

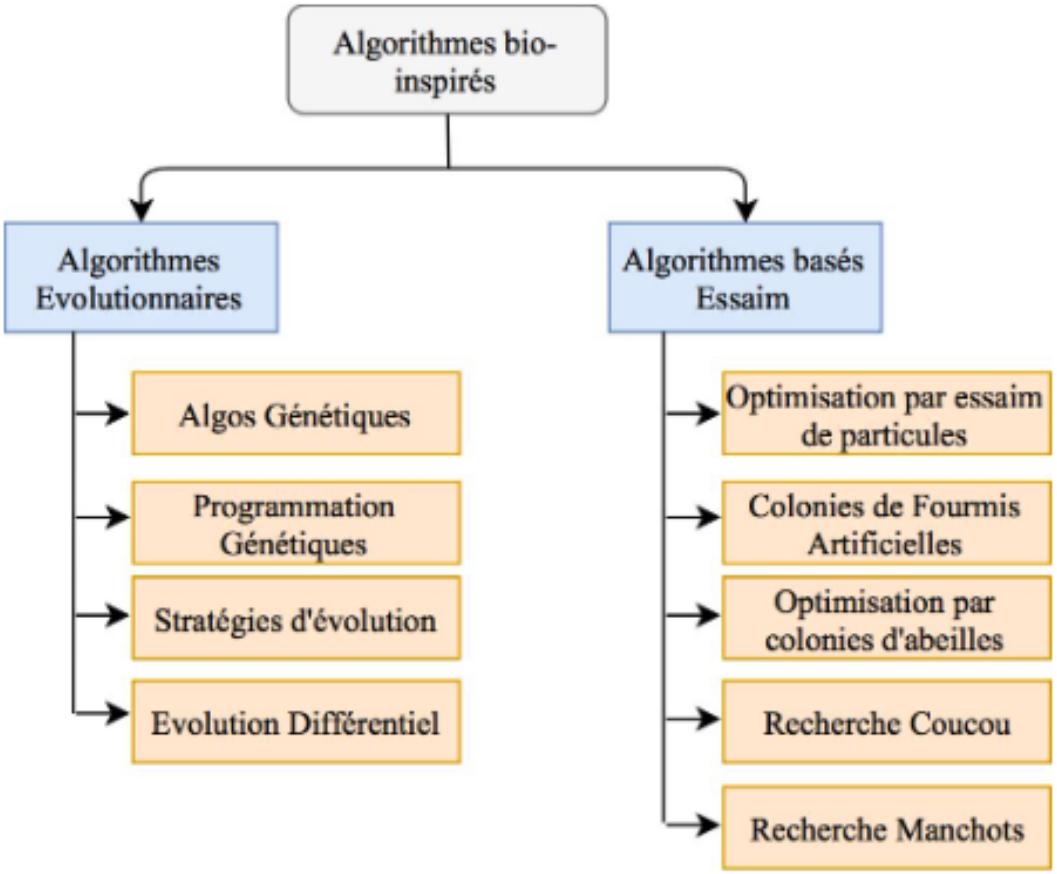

FIGURE 1.4 – Classification Algorithmes bio-inspirés[4]

Algorithmes évolutionnaires

Les algorithmes évolutionnaires sont des techniques de recherche inspirées par l'évolution biologique des espèces. Ils s'inspirent de l'évolution des êtres vivants (la théorie Darwi-nienne de la sélection naturelle des espèces) pour résoudre des problèmes d'optimisation Un algorithme évolutionnaire simule un processus d'évolution sur une population d'individus, dans le but de faire évoluer vers les optimums globaux du problème d'optimisation considéré un algorithme évolutionnaire typique réunit trois composants [4] :

1. Une population constituée de plusieurs individus représentant des solutions potentielles du problème posé .

2. Une fonction d'adaptation (fitness) qui évalue la performance d'un individu par rapport au milieu.

3. Un mécanisme d'évolution de la population composé de plusieurs opérateurs de modification et de sélection permettant (grâce à ces opérateurs prédéfinis), d'éliminer certains individus et d'en créer de nouveaux.

Algorithmes basés essaim

L'intelligence en essaim (SI) est l'une des techniques d'intelligence computationnelle utilisées pour résoudre des problèmes complexes proposé par **Dor-igo** comme «l'intelligence collective émergente de groupes d'agents simples».

Les algorithmes basés essaim sont des techniques d'optimisation inspirés du comportement collectif chez les espèces sociaux Plusieurs techniques d'optimisation basées sur les principes du SI ont été inspirées de systèmes de comportement collectif réels dans la nature, notamment Ant Colony Optimizations (ACO) par **M. Dorgo** en 1992, Particles Swarm Optimizations (PSO) par **Eberhart et Kenedy** en 1995, Artificial Bee Colony (ABC) de **Karabaga** en 2005 Glowworm Swarm Optimization, elephant herding optimization (EHO) par **Wang fin** en 2015 et d'autres algorithmes d'optimisations.

1.6 Optimisation L'élevage d'éléphants(EHO)

Optimisation de L'élevage d'éléphants (EHO), est une méthode de recherche méta-heuristique basée sur un essaim intelligent proposée par **Wang** fin 2015 pour résoudre des problèmes d'optimisation. L'algorithme dérive de la modélisation du comportement grégaire de vrais éléphants dans la nature. Le comportement grégaire peut être résumé comme suit [22] :

- Les essaims d'éléphants sont constitués d'un certain nombre de sous-groupes, appelés clans, qui sont composés d'un certain nombre d'éléphants femelles et mâles.

- Chaque clan se déplace sous la supervision (leadership) d'une matriarche (éléphant femelle adulte).

- Les veaux mâles qui atteignent l'âge adulte quittent le clan auquel ils appartiennent.

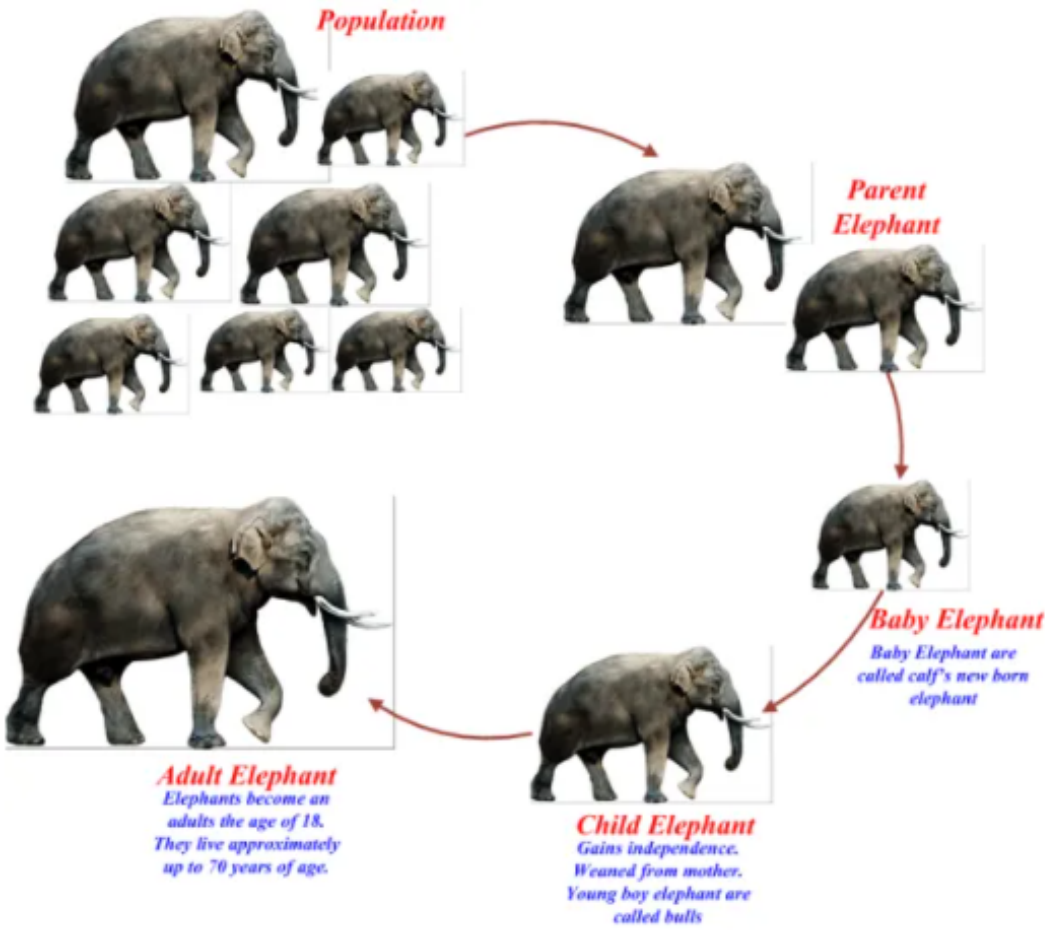

FIGURE 1.5 – Cycle de vie d'éléphants[5]

1.6.1 Source d'Inspiration

Dans la nature, les éléphants sont des animaux sociaux, et ils ont des structures sociales complexes de femelles et d'éléphanteaux. Un groupe d'éléphants est composé de plusieurs clans sous la direction d'une matriarche montré dans la figure 1.6 souvent la femelle la plus âgée. Un clan est composé d'une femelle avec ses éléphanteaux ou certaines femelles apparentées. Les femelles préfèrent vivre en groupes familiaux, alors que les éléphants mâles ont tendance à vivent dans l'isolement, et ils quitteront leur groupe familial lorsqu'ils seront grands. Bien que les éléphants mâles vivent loin de leur groupe familial, ils peuvent rester en contact avec les éléphants de leur clan par le biais de vibrations à basse fréquence[6].

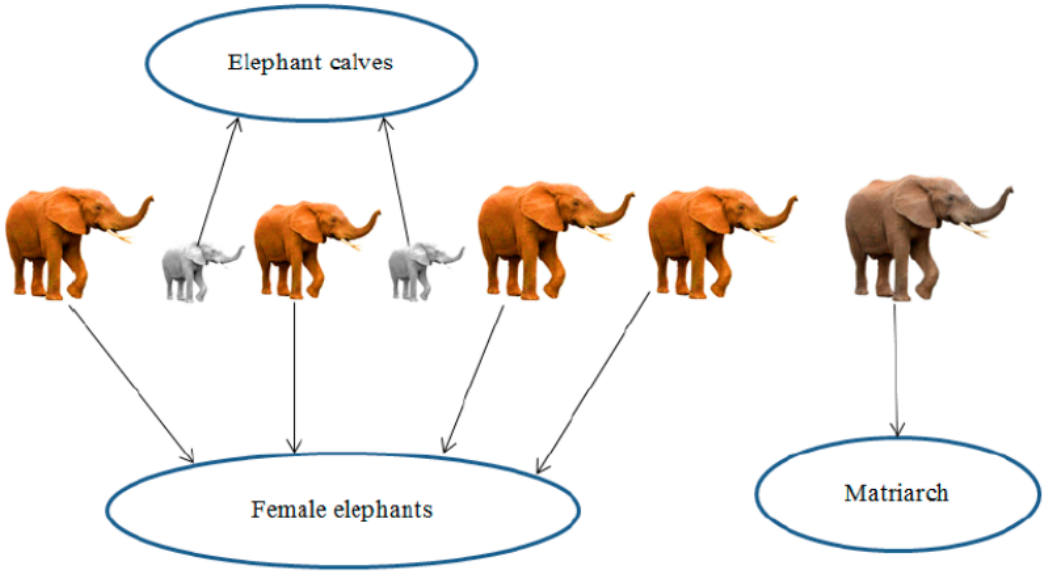

FIGURE 1.6 – Population d'éléphants[6]

1.6.2 Principes de Base

EHO modélise les comportements d'élevage des éléphants en deux opérations[22] :

1. Mise à jour du clan :

   Tous les éléphants vivent ensemble sous la direction d'une matriarche dans chaque clan. Par conséquent, pour chaque éléphant du clan ci, sa prochaine position est influencée par la matriarche ci. Pour l'éléphant j du clan ci, il peut être mis à jour comme suit :

$$\mathbf{x_{new,ci,j}} = \mathbf{x_{ci,j}} + \alpha \times (\mathbf{x_{best,ci}} - \mathbf{x_{ci,j}}) \times \mathbf{r} \tag{1.1}$$

où $x_{new,ci,j}$ et $x_{ci,j}$ sont la position nouvellement mise à jour et l'ancienne position pour l'éléphant j dans le clan ci, respectivement. $\alpha \in [0, 1]$ est un facteur d'échelle qui détermine l'influence de la matriarche ci sur x_{cij}, $x_{best,ci}$ représente la matriarche ci, qui est l'individu le plus apte à devenir éléphant dans le clan ci.$r \in [0, 1]$. On utilise ici une distribution uniforme. L'éléphant le plus apte dans chaque clan ne peut pas être mis à jour par l'équation 1.1 , c'est-à-dire que $x_{ci,j} = x_{best}$,ci. Pour l'éléphant le plus apte, il peut être mis à jour comme suit :

$$\mathbf{x_{new,ci,j}} = \beta \times \mathbf{x_{center,ci}} \tag{1.2}$$

où $\beta \in [0, 1]$ est un facteur qui détermine l'influence de le $x_{center,ci}$ sur $x_{new,ci,j}$. Nous pouvons voir que le nouvel individu $x_{new,ci,j}$ dans l'équation 1.2 est généré par les informations obtenues par tous les éléphants du clan ci. $x_{center,ci}$ est le center du clan ci, et pour la d-ième dimension, il peut être calculé comme suit :

$$\mathbf{x_{center,ci,d}} = \frac{\mathbf{1}}{\mathbf{n_{ci}}} \times \sum_{\mathbf{j=1}}^{\mathbf{n_{ci}}} \mathbf{x_{ci,j,d}} \tag{1.3}$$

où $1 \leqslant d \leqslant D$ indique la d-ième dimension, D est sa dimension totale. n_{ci} est le nombre d'éléphants dans le clan ci. $x_{ci,j,d}$ est le d-ième de l'individu éléphant $x_{ci,j}$. Le centre du clan ci, $x_{center,ci}$ peut être calculé par l'intermédiaire des calculs D par l'équation 1.3 sur la base de la description ci-dessus[22].

2. Opérateur de séparation :

   Dans les groupes d'éléphants, les éléphants mâles quittent leur groupe familial et vivent seuls lorsqu'ils atteignent la puberté. Ce processus de séparation peut être modélisé par un opérateur de séparation lors de la résolution de problèmes d'optimisation. Afin d'améliorer encore la capacité de recherche de la méthode EHO, supposons que les éléphants avec la plus mauvaise forme physique mettront en œuvre l'opérateur de séparation à chaque génération, comme le montre l'équation 1.4.

$$\mathbf{x_{worst,ci}} = \mathbf{x_{min}} + (\mathbf{x_{max}} - \mathbf{x_{min}} + \mathbf{1}) \times \mathbf{rand} \tag{1.4}$$

où x_{max} et x_{min} sont respectivement les limites supérieure et inférieure de la position de l'individu éléphant. $x_{worst,ci}$ est le pire individu éléphant du clan ci. $rand \in [0, 1]$ est une sorte de distribution stochastique et une distribution uniforme dans l'intervalle [0, 1] est utilisée dans notre travail actuel[22].

1.6.3 Pseudo Code d'Algorithme EHO

Algorithm : Elephant herding optimization

Begin
Initialization. Set the initialize iterations G = 1 ; initialize the population P randomly ; set maximum generation MaxGen
while stopping criterion is not met **do**
  Sort the population according to fitness of individuals.
  for all clans ci **do**
    for elephant j in the clan ci **do**
      Generate x_{new}, ci,j and update $x_{ci,j}$by Equation 1.1
      if $x_{ci,j} = x_{best,ci}$ **then**
        Generate x_{new}, ci,j and update $x_{ci,j}$by Equation 1.2
      end if
    end for
  end for
  for all clans ci **do**
    Replace the worst individual ci by Equation 1.4
  end for
  Evaluate each elephant individual according to its position.
  T = T + 1.
end while

1.6.4 Applications de l'algorithme

- Problème du voyageur de commerce
- Applications industrielles
- Contrôle de la fréquence de la charge
- Détection des spams
- Machine à vecteur de support (SVM)
- Sélection de fonctionnalités
- Planification de la trajectoire d'un véhicule
- Réseaux neuronaux artificiels

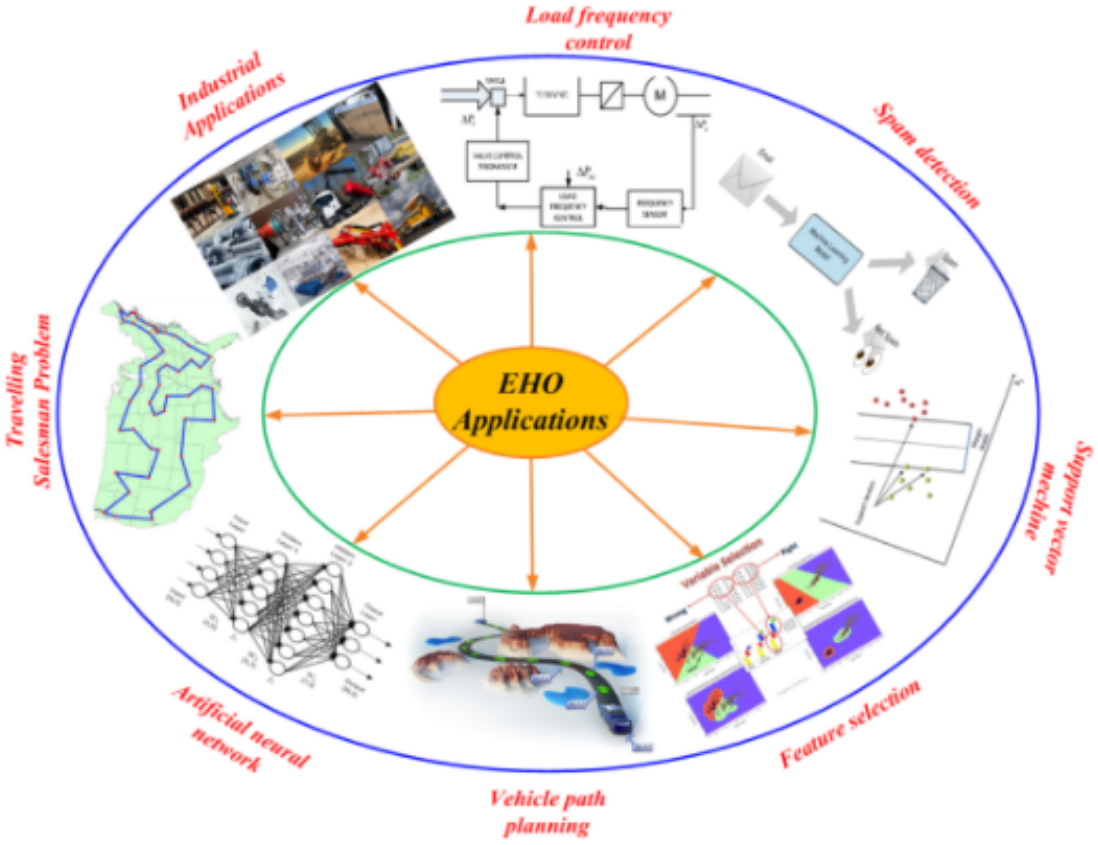

FIGURE 1.7 – Applications de l'algorithme[5]

1.7 Conclusion

Dans ce chapitre, nous avons vu le principe de fonctionnement des différents types d'apprentissage automatique et le problème d'optimisation combinatoire puis nous sommes concentrés en particulier sur le méta-heuristique et intelligence d'essaim ainsi que la clarification du choix de bio-inspiré pour résoudre notre problème.

Dans le chapitre suivant on présente les différentes fonctionnalité et approches utilisées pour la détection des spam dans TWITTER.

CHAPITRE 2

ÉTAT DE L'ART

2.1 Introduction

Le spam est une véritable menace pour l'utilité du web. Les spammeurs masquent leur contenu en tant que contenu utile ou pertinent, et par conséquent est délivré à l'utilisateur. Les utilisateurs légitimes consomment ce spam en considérant qu'ils répondent à leurs besoins d'information.

Shirky [23] a remarqué qu'un canal de communication ne vaut pas la peine son sel jusqu'à ce que les spammeurs descendent. Les spammes ne sont pas faciles à arrêter. Depuis plusieurs années, les services de messagerie comme Gmail, Microsoft et d'autres, ont réussi détecter les courriels non sollicités, mais ceux-ci continuent de circuler sur le web.

Twitter est l'un des réseaux sociaux les plus populaires réseau social en ligne le plus populaire qui a été très affecté par le spamme sur Twitter est plus menaçant parce qu'il est plus ciblé sur les sujets tendances de Twitter et donc un peu plus facile à pénétrer, en particulier en raison de l'existence d'un réseau de spamme. Plus facile à pénétrer, notamment en raison de l'opérateur hashtag, un autre fait qui fait de Twitter une cible plutôt facile et fructueuse pour les spammeurs est la variété de son public.

Dans ce chapitre, nous essayons d'introduisons brièvement le problème de sécurité dans le Twitter ça sera comme un type de spam puis étudié les caractéristiques de la détection spam. Dans la dernière section, nous rapportons quelques méthodes utilisées par Twitter pour traiter les spam et leur différentes approches évoquées pour résoudre le problème en question.

2.2 Réseau Social Twitter

TWITTER est un site de réseau social tout comme Facebook et MySpace, sauf qu'il ne fournit qu'un service de microblogging. où les utilisateurs peuvent envoyer de courts messages (appelés tweets) qui apparaissent sur les pages de leurs amis.

Un utilisateur de TWITTER n'est identifié que par un nom d'utilisateur et éventuellement par un nom réel. Un utilisateur de TWITTER peut commencer à suivant un autre utilisateur X, Par conséquent, cet utilisateur reçoit les tweets de l'utilisateur X sur sa propre page. L'utilisateur X qui est abonné peut le suivant à son tour s'il le souhaite[24].

Les tweets peuvent être regroupés à l'aide de hashtags qui sont des mots populaires, commençant par le caractère "#". Les hashtags permettent aux utilisateurs de rechercher efficacement des tweets sur la base de sujets d'intérêt. Lorsqu'un utilisateur aime le tweet d'une autre personne, il peut "retweeter" ce message. En conséquence, ce message est montré à tous ses abonnés (followers). Un utilisateur peut décider de protéger son profil[25].

Ce faisant, tout utilisateur qui souhaite abonné cet utilisateur privé doit obtenir son autorisation. TWITTER est le site de réseau social qui connaît la croissance la plus rapide, avec un taux de croissance de 660 % en 2009[7].

2.2.1 Spam Social

Le spam social est un contenu textuel absurde ou non-sens qui apparaît sur les réseaux sociaux en ligne et tout site Web qui gère le contenu généré par les utilisateurs tels que les chats et les commentaires[26]. Le spam social peut prendre plusieurs formes, Y compris les blasphèmes, les insultes, les discours de haine, les critiques frauduleuses, les faux amis, la messagerie de groupe, le hameçonnage et les liens malveillants, et le matériel pornographique. Vous pouvez traiter le spam social comme une information non pertinente.

Cependant, cette interprétation est complètement fausse. Nous prouvons que cette explication est raisonnable Selon la définition d'un système de recherche d'informations (RI)[27], où la disponibilité des documents dans le système RI dépend de Requête d'entrée de recherche. Par conséquent, les fichiers qui ne sont pas liés à des fichiers de demande entrante n'est pas nécessairement un spam.

Le spam social peut être défini comme n'appartenant pas à pertinent, aucune explication dans aucun contexte jusqu'à ce que l'entrée soit déterminée spam.

2.2.2 Problème de Sécurité

L'ouverture et la commodité de la plateforme TWITTER attirent également comptes criminels (spammeurs), afin d'attaquer la plateforme dans le but de gagner de l'argent de manière illégitime.

Comme il existe une restriction sur la longueur des tweets, il est courant que les spammeurs diffusent des messages non sollicités, qui peuvent rédiger les utilisateurs vers des sites web externes malveillants.

Par rapport au spam traditionnel qui se propage par courriels, le spam TWITTER est plus dangereux et plus sophistiqué pour attirer les internautes et les tromper. Selon un rapport récent, le taux de clics des spam TWITTER atteint 0,13%, alors qu'il n'atteint que 0,0003% 0,0006% dans le spam par e-mail.

Afin de résoudre le problème du spam TWITTER, de nombreux outils de détection ont vu le jour ces dernières années et de nombreux systèmes de détection ont été proposés[28].

2.2.3 Types des Spam

Les spammeurs sont des utilisateurs malveillants qui contaminent les informations présentées par les utilisateurs légitimes et qui à leur tour, constituent un risque pour la sécurité et la confidentialité de TWITTER. Les spammeurs appartiennent à l'une des catégories suivantes[29] :

1. Hameçonneurs (Phishers) : sont les utilisateurs qui se comportent comme un utilisateur normal pour acquérir les données personnelles d'autres utilisateurs authentiques.

2. Faux utilisateurs : ce sont les utilisateurs qui se font passer pour des utilisateurs authentiques afin d'envoyer des spams à leurs amis ou à d'autres utilisateurs du réseau.

3. Promoteurs : ce sont ceux qui envoient des liens malveillants de publicité ou d'autres liens promotionnels à d'autres personnes afin d'obtenir leurs informations personnelles.

2.2.4 Caractéristiques Pour Détection Des Spam

Les caractéristiques pour la détection des spam sur TWITTER sont classées comme suit : caractéristiques basées sur l'utilisateur, caractéristiques basées sur le contenu, et relation entre l'émetteur et le récepteur du tweet. Ces caractéristiques constituent l'ossature des caractéristiques utilisées par les travaux connexes dans la littérature. Chaque catégorie de caractéristiques est présentée ci-dessous[7].

A) Caractéristique Basées Sur L'utilisateur (Compte)

caractéristique basée sur l'utilisateur sur TWITTER, vous pouvez créer votre propre réseau social en suivant vos amis et en autorisant les autres à vous suive . Les comptes de spam tentent de suivre un grand nombre d'utilisateurs afin de gagner leur attention. La politique de TWITTER en matière de spam et d'abus stipule que "si vous avez un petit nombre d'abonnée (follower) par rapport au nombre de personnes que vous abonnement (following) [7], il peut être considéré comme un compte de spam.

Trois caractéristiques basées sur l'utilisateur, à savoir le nombre d'abonnée (follower), le nombre d'abonnement (followers) et la réputation d'un utilisateur. Cette dernière est calculés comme suit pour la détection du spam :

$$\mathbf{R\,(j)} = \frac{\mathbf{n_i\,(j)}}{\mathbf{n_i\,(j) + n_0(j)}} \tag{2.1}$$

où $n_i(j)$ représente le nombre d'abonnés (followers) de l'utilisateur j et $n_o(j)$représente le nombre d'abonnement (following) de l'utilisateur j [7].

B) caractéristiques Basées Sur Le Contenu (Tweet)

Pour les caractéristiques basées sur le contenu, nous utilisons certaines caractéristiques évidentes, par exemple la longueur moyenne d'un tweet. D'autres caractéristiques basées sur le contenu sont décrites ci-dessous.

- **Nombre D'URL**

Étant donné que Twitter n'autorise qu'un message avec une longueur maximale longueur de 140 caractères, de nombreuses URL incluses dans les tweets sont URL raccourcies. Les spammeurs incluent souvent des URL raccourcies dans leurs tweets pour inciter les utilisateurs légitimes à y accéder.

Twitter filtre les URL liées à des sites malveillants connus. Cependant, les URL raccourcies peuvent masquer les URL sources et masquer les sites malveillants qui se cachent derrière elles. Bien que Twitter ne vérifie pas ces URL raccourcies pour les logiciels malveillants, les mises à jour de tout utilisateur qui consistent principalement en des liens sont des spam selon la politique de Twitter[28].

- **Réponses/Mentions**

Un utilisateur est identifié par un nom d'utilisateur unique @username dans les tweets sur Twitter. Chaque utilisateur peut envoyer un message de réponse à un autre utilisateur en utilisant le format @username+message où @username est le destinataire du message. Chaque utilisateur peut répondre à n'importe qui sur Twitter, qu'il s'agisse de ses abonnés(followers)/abonnements
(following) ou non. Il peut également mentionner un autre @username n'importe où dans son tweet, plutôt que juste au début. Twitter collecte automatiquement tous les tweets contenant un nom d'utilisateur au format @username dans son onglet de réponses. Les caractéristiques de réponse et de mention sont conçues pour aider les utilisateurs à abonné la conversation et à se découvrir sur Twitter.

Cependant, les spammeurs abusent souvent de cette caractéristique en incluant de nombreux @noms d'utilisateur en tant que réponses ou mentions non sollicitées dans leurs tweets. Si un utilisateur inclut trop de réponses/mentions dans ses tweets, Twitter considérera ce compte comme suspect[29].

- **Retweets**

Twitter permet aux utilisateurs de retweeter les tweets générés par d'autres utilisateurs. Tous les retweets commencent par le symbole @RT. Le nombre de retweets dans les 20 à 100 tweets les plus récents d'un utilisateur est également utilisé comme l'une des caractéristiques basées sur le contenu de système de détection spam[29].

- **Hashtags**

Les sujets tendance sont les termes les plus fréquemment mentionnés sur Twitter à l'adresse ce moment, cette semaine ou ce mois-ci. Les utilisateurs peuvent utiliser le hashtag, qui est le #symbole suivi d'un terme décrivant ou nommant les sujets, à un tweet. S'il existe de nombreux tweets contenant le même terme, le terme deviendra un sujet tendance. Les spammeurs publient souvent de nombreux tweets sans rapport qui contiennent les sujets tendances pour inciter les utilisateurs légitimes à lire leurs tweets. Twitter considère un compte comme spam «si un utilisateur publie plusieurs mises à jour non liées à un sujet en utilisant le symbole #»[7].

Le nombre de tweets contenant le symbole # dans les 100 tweets les plus récents d'un utilisateur est utilisé comme l'une des caractéristiques basées sur le contenu.

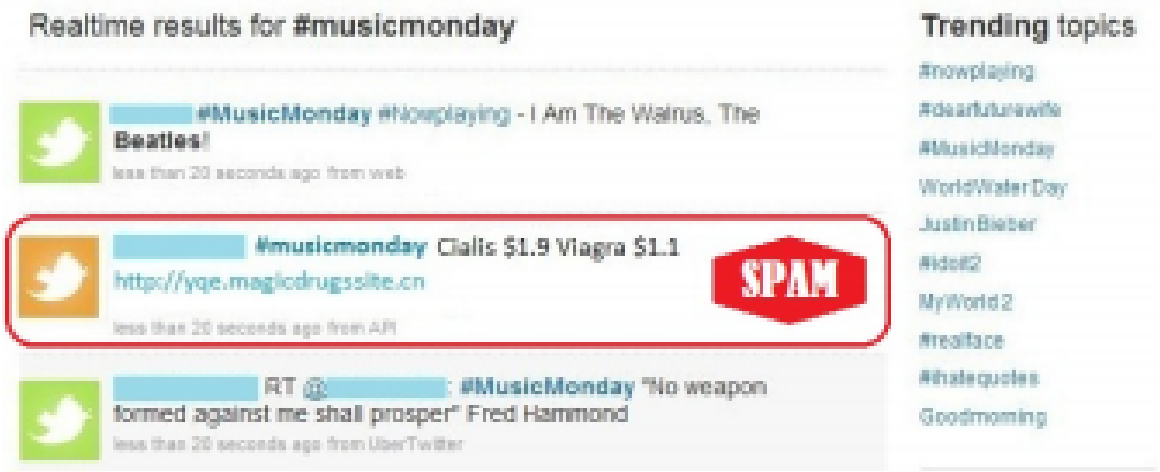

Figure 2.1 – Exemple illustratif d'une Recherche Sur Twitter Pour hashtag musicmonday[7]

2.3 Comment Twitter Traite le Spam

TWITTER utilise à la fois des services manuels et automatisés pour lutter contre les spammeurs afin de fournir un environnement sans spam. La méthode manuelle consiste à permettre aux utilisateurs de signaler les spammeurs via les pages de profil des spammeurs.

TWITTER fournit une interface utilisateur comme celle présentée à la Figure 2.2 pour signaler le compte en sélectionnant la raison. Une autre méthode couramment citée dans la littérature consiste à signaler les spammeurs au compte officiel "@spam", mais selon un rapport récent de TWITTER, cette méthode de signalement du spam est dépassée. De plus, **Wang** signale que cette méthode est utilisée de manière abusive par les canulars et les spams.

Ces approches manuelles demandent beaucoup de travail et ne suffiraient pas à détecter tous les spammeurs compte tenu des milliards d'utilisateurs. TWITTER utilise divers facteurs tels que[8] :

1. La publication de messages en double sur plusieurs comptes ou de multiples messages en double sur un seul compte.
2. suivre/ne plus suivre un grand nombre de comptes en peu de temps,
3. Le fait d'avoir un grand nombre de plaintes pour spam déposées contre le compte.
4. aimer, abonnés (Followers) et retweeter de manière agressive.
5. Poster des liens malveillants.
6. Poster des tweets qui consistent principalement en des liens au lieu de poster également des mises à jour personnelles.
7. Poster des tweets sans rapport avec un sujet d'actualité pour déterminer quelle conduite est considérée comme du spamming.

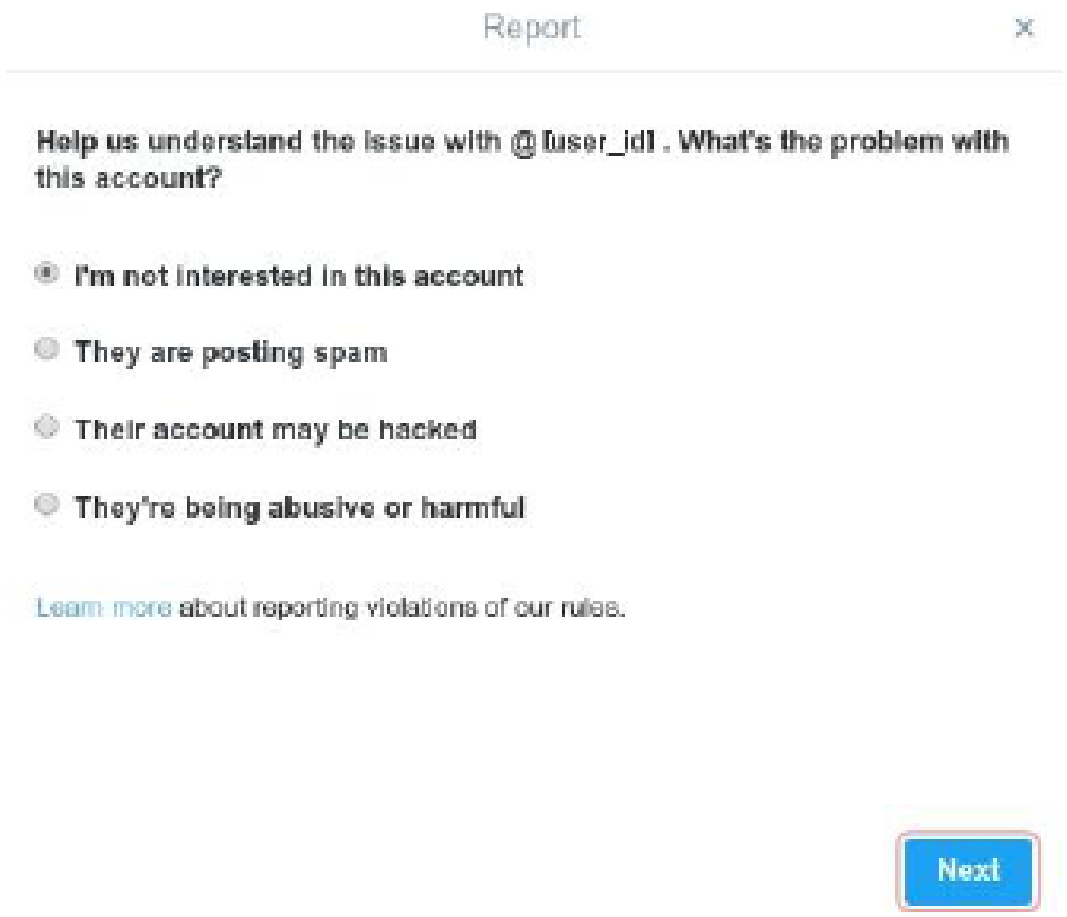

FIGURE 2.2 – L'interface utilisateur de TWITTER qui est utilisée pour signaler un compte en sélectionnant la raison[8]

2.4 Détection Des Spam Par l'Apprentissage Automatique (ML)

La plupart des méthodes de détection des spam ont utilisé des algorithmes d'apprentissage automatique supervisé à deux niveaux de détection, répartis entre la détection au niveau du tweet, et celle au niveau du compte. Nous mentionnons les plus utilisées dans la section suivante.

2.4.1 Naïve Bayes

L'algorithme Naïve Bayes est un algorithme d'apprentissage supervisé, basé sur le théorème de Bayes et utilisé pour résoudre des problèmes de classification. Il est principalement utilisé dans la classification de texte qui comprend un ensemble de données d'entraînement de grande dimension. La formule du théorème de Bayes est donnée par[30] :

$$P(c/x) = \frac{P(x/c)P(c)}{P(x)} \tag{2.2}$$

Naïve Bayes Classifier est l'un des algorithmes de classification les plus simples et les plus efficaces qui aide à créer des modèles d'apprentissage automatique rapides qui peuvent faire des prédictions rapides. C'est un classificateur probabiliste, ce qui signifie qu'il prédit sur la base de la probabilité d'un objet.

Le modèle proposé par Deepali M.Gohi et Ashwini Athawale [9] pour détecte des spam utilise un algorithme d'apprentissage automatique (Naïve Bayes). pour identifié le spam dans Twitter, ils utilisent des caractéristique basée sur le contenu et sur l'utilisateur pour analyser le comportement de spam de l'utilisateur.

Ils ont utilisé l'API Twitter (Application Protocole Interface) pour obtenir tous les détails de l'utilisateur de Twitter, la détection est effectuée par ensemble de données, les étapes de prétraitement des informations comprennent le filtrage et la suppression des mots vides.

Puis ils ont comparé les résultats obtenus par la méthode de Naïve Bayes avec les résultats de SVM comme le montre la Figure 2.3.

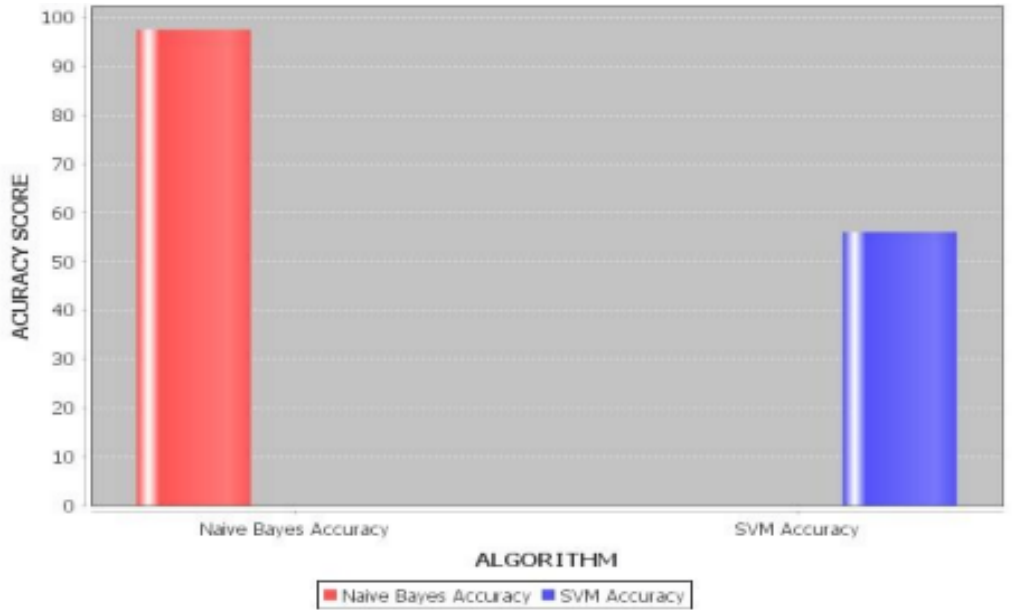

FIGURE 2.3 – Le résultat obtenu[9]

2.4.2 Machines à Vecteurs de Support

Machines à vecteurs de support(SVM) sont des algorithmes d'apprentissage automatique proposés par **V.N. Vapnik** et **A.Ya. ChervonenkisLes** , décrivent une approche de classification supervisée basée sur une interprétation géométrique, en s'appuyant sur la notion de marge maximale.Qui est une méthode de pointe en matière de classification et qui a obtenu les meilleurs résultats parmi un ensemble de classificateurs testés[31].

L'objectif d'un SVM est de trouver l'hyperplan qui sépare de manière optimale et avec une marge maximale les données d'apprentissage en deux parties d'un espace à N dimensions[32].

le SVM effectue une classification en mettant en correspondance des vecteurs d'entrée dans un espace à N dimensions et en vérifiant de quel côté de l'hyperplan défini se trouve le point.

Fabricio Benevenuto et Gabriel Magn [33] ont présenté une méthode basée sur la classification d'apprentissage pour la détection des spam en utilisant SVM non linéaire, avec le noyau de la fonction de base radiale (RBF). A base d'ensemble de données presque complet de TWITTER ainsi qu'une collection de spammeurs construite manuellement et les non-spammeurs.

Le modèle dépendait d'un mélange de caractéristique basée sur les tweets et de caractéristique basées sur l'utilisateur, pour couvrir plus de contexte .et il a évalué son travail

en utilisant les métriques standard de recherche d'informations de rappel,précision, Micro-F1 et Macro-F1. Les résultats ont montré que la méthode est efficace pour détecte spam et le système atteint une précision de 87.6%.

Yash Thigale et Prof [10] Ont présenté une comparaison entre les algorithme d'apprentissage (SVM et Naïve Bayes) par la détection des spam TWITTER. Dans cette expérience, l'ensemble de données TWITTER en temps réel a été utilisé pour obtenir des tweets en temps réel.

Il ont pris six paramètres d'évaluation, à savoir l'exactitude, la précision, le rappel, la mesure F pour l'analyse en profondeur des tweets spams et non spams, comme le tableau suivant 2.4.

	Naïve Bayes	SVM
Precision	68.45	78.70
Recall	79.44	65.64
F-Measure	72.11	74.31
Accuracy	80.29	87.26
Execution Time (ms)	435	245

FIGURE 2.4 – Paramètres d'évaluation[10]

Le modèle a trouvé que la capacité des classificateurs à détecter le spam augmente plus que la méthode SVM.

Yash Thigale et tous ont trouvé que le Naïve Bayes donné un bon rappel par apport au SVM mais s' ils regardons le temps d'exécution le SVM prend moins de temps par rapport à Naïve Bayes.

Comme ils ont trouvé que l'utilisateur devrait essayer d'apporter plus des caractéristiques ou des paramètres pour construire le modèle de meilleure façon afin d'exceller dans la recherche de tweets spam sur TWITTER.

2.4.3 Arbre de décision

Les arbres de décision sont un type d'apprentissage automatique supervisé. L'arbre peut être expliqué par deux entités, à savoir les nœuds de décision et les feuilles. Les feuilles sont les décisions ou les résultats finaux. Et les nœuds de décision sont l'endroit

où les données sont divisées. Il est simple de convertir des arbres de décision en règles de classification[34].

L'apprentissage de l'arbre de décision utilise un arbre de décision en tant que modèle prédicatif qui mappe les observations sur un élément à des conclusions sur la valeur de l'objet de l'élément. C'est l'une des approches de modélisation prédicative utilisées dans les statistiques, l'exploration de données et l'apprentissage automatique [35].

Po-Ching Lin et Po-Min Huang [36] ont présenté une méthode basée sur la classification d'apprentissage pour la detection des spam en utilisant l'arbre de décision. Ils sélectionnent le taux d'URL et le taux d'interaction comme des caractéristiques principales, et vérifier leur efficacité avec 26 758 comptes avec 508403 tweets sur la période de septembre 2011 à Mars 2012 collectés par les API Twitter. Les résultats ont montré que la méthode est efficace pour détecter le spam, la précision de la détection est estimée entre 0,829 et 0,885.

2.5 Détection Des Spam Par l'Apprentissage Profond (DL)

L'apprentissage profond a été utilisé pour résoudre de nombreux problèmes, notamment les problèmes de sécurité dans TWITTER en particulier la détection des spam. Le réseau de neurones artificiels et réseaux neuronaux convolutifs ont été utilisés dans de nombreux travaux pour améliorer leur efficacité, nous décrivons quelques-uns ci-dessous.

2.5.1 Réseau Neurones Artificiels

Un réseau de neurones artificiels ou Neural Network est un système informatique s'inspirant du fonctionnement du cerveau humain pour apprendre.

Un neurone formel, ou neurone, est une fonction algébrique non linéaire et bornée, dont la valeur dépend de paramètres appelés coefficients ou poids[11]. Les variables de cette fonction sont habituellement appelées « entrées » du neurone, et la valeur de la fonction est appelée « sortie ». Un neurone est donc avant tout un opérateur mathématique, dont on peut calculer la valeur numérique par quelques lignes de programme informatique. Les neurones les plus fréquemment utilisés sont ceux dont la fonction est calculée en deux étapes :

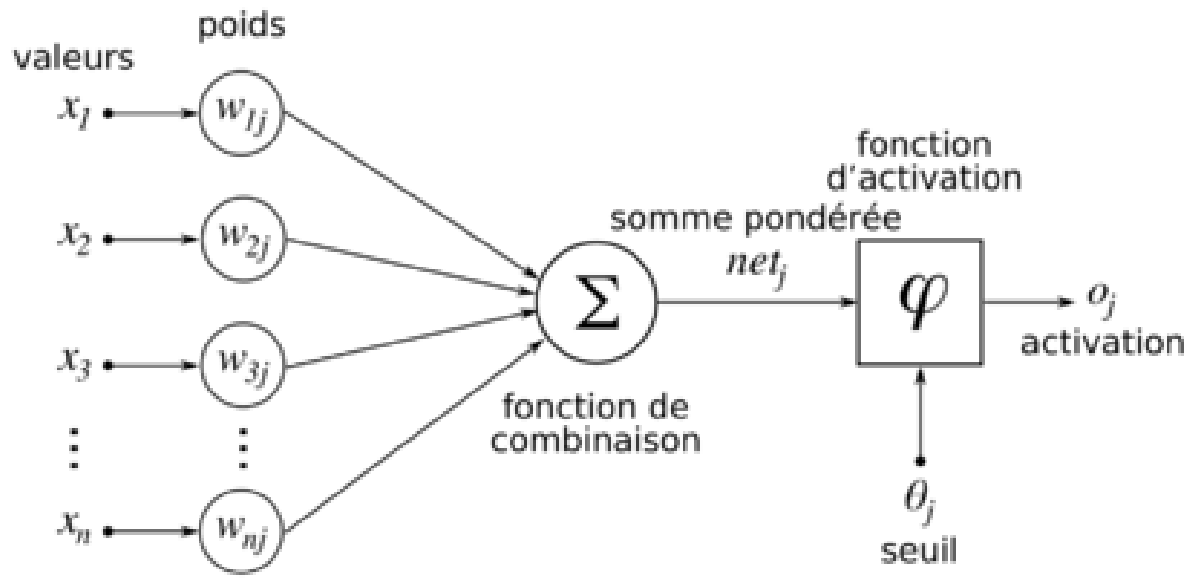

FIGURE 2.5 – Réseau de neurones artificiels[11]

–calcul de la somme v des entrées pondérées par les paramètres du réseau ;

–calcul d'une fonction non linéaire (y) de cette somme (dite « fonction d'activation ») par exemple une fonction sigmoïde. y = (1/(1 + e(—v))) ou encore y = th v. Les xi sont les variables (ou entrées) du neurone, les ci sont des paramètres ajustables (coefficients

ou poids).

Le modèle proposé par Mehmet ŞİMŞEK et Oğuzhan YILMAZ [12] pour détecte spam utilise un réseau de neurones artificiels. Ils ont utilisé 10 caractéristiques basées sur les comptes (âge du compte, nombre d'abonnés, nombre d'abonnés, nombre de favoris des utilisateurs, nombre de listes, nombre de tweets, nombre de retweets, nombre de hashtags, nombre de mentions d'utilisateurs et URL).

L'ensemble de données utilisé dans cette étude comprend 100000, 95000 comptes ne sont pas des spammeurs et 5000 comptes sont des spammeurs.

Ils ont utilisé quatre fonctions d'activation différentes et comparé les efficacités de différentes combinaisons de ces fonctions d'activation. ses fonctions d'activation sont la fonction Softmax, la fonction sigmoïde, le redresseur (ReLu) et la tangente hyperbolique (Tanh). L'ANN construit a 1 couche d'entrée avec 10 neurones,1 couche cachée avec 10 neurones et 1 couche de sortie avec 1 neurone.

Ils ont réalisé deux groupes d'expériences, Dans le premier groupe, ils utilisé la même fonction d'activation à toutes les couches (couche d'entrée, couche cachée et couche de sortie).

La fonction sigmoïde a donné les meilleurs résultats. Résultats de la tangente hyperbolique pas aussi bon que les résultats de la fonction sigmoïde, la variance est trop élevée et la moyenne des précisions est trop faible.

Les résultats de la fonction de redressement sont meilleurs que les résultats de la fonction tangente hyperbolique mais toujours pas bons que les résultats de la fonction sigmoïde. La fonction Softmax a les pires résultats car la fonction Softmax a également été utilisée pour la couche de sortie. En tant que couche de sortie, la fonction Softmax ne convient pas.

Dans le deuxième groupe d'expériences, ils ont utilisé différentes fonctions d'activation à différentes couches.

Activation Functions (input layer, hidden layer, output layer	Mean	Variance
Softmax-Softmax-Sigmoid	0,9772	0,0015
Tanh-Tanh-Sigmoid	0,9768	0,0039
Rectifier- Rectifier- Sigmoid	0,9721	0,0023
Rectifier- Rectifier-Tanh	0,8593	0,1846
Tanh-Tanh-Rectifier	0,9594	0,0113
Softmax-Softmax-Tanh	0,9524	0,0091
Softmax-Softmax-Rectifier	0,9601	0,0124

FIGURE 2.6 – Résultat de deuxièmes expériences[12]

Softmax-Softmax-Sigmoid a donné les meilleurs résultats. Les performances de Tanh-Tanh-Sigmoid et Rectifier-Redresser-Sigmoid sont plus proches des performances de Softmax-Softmax-Sigmoïde.

2.5.2 Réseaux Neurones Convolutifs

Les réseaux de neurones convolutifs (CNN) sont l'un des réseaux de neurones profonds les plus connus et le plus utilisé, notamment dans le domaine de la vision par ordinateur, où sa structure simule le fonctionnement du cortex visuel dans le cerveau du chat [37].

Sreekanth Madisetty et Maunendra Sankar Desarkar [11] ont présente trois méthode pour classer les tweets comme spam ou non spam basé sur le réseau de neurones à convolution avec entrée de mots de longueur variable en utilisant modèle word2vec pour obtenir embeddings de mots qui sont utilisé dans ce travail et aussi le modèle classique basé sur les caractéristique suivant :

- Basé sur l'utilisateur en utilisant l'ensemble de caractéristiques pour entraîner le modèle (Longueur de la description, emplacement donné ou non, Nombre d'amis, Nombre d'abonnés...ect).
- Basé sur le contenu (Nombre de mots, Longueur du Tweet, Nombre de liens URL...ect).
- N-grammes(Unigrammes, Bigrammes avec fréquence de terme (tf)).

Enfin ils ont combine cinq convolutions modèles de réseaux de neurones (CNN) et un modèle basé sur les caractéristiques montrer en Figure 2.7 via un réseau de neurones formé par chaque CNN avec des mots de différentes dimensions, ensuit sélectionnez la meilleure méthode parmi chacun des CNN sur la base de F-mesure, enfin ils appliquent le Méta-classificateur pour combiner les sorties. Ils ont utilisé deux ensembles de données

HSpam (data set équilibré), et **1KS10KN** (data set déséquilibré).

La méthode proposé a atteint une précision 92% pour data set 1KS10KN et 94% pour data set Hspam.

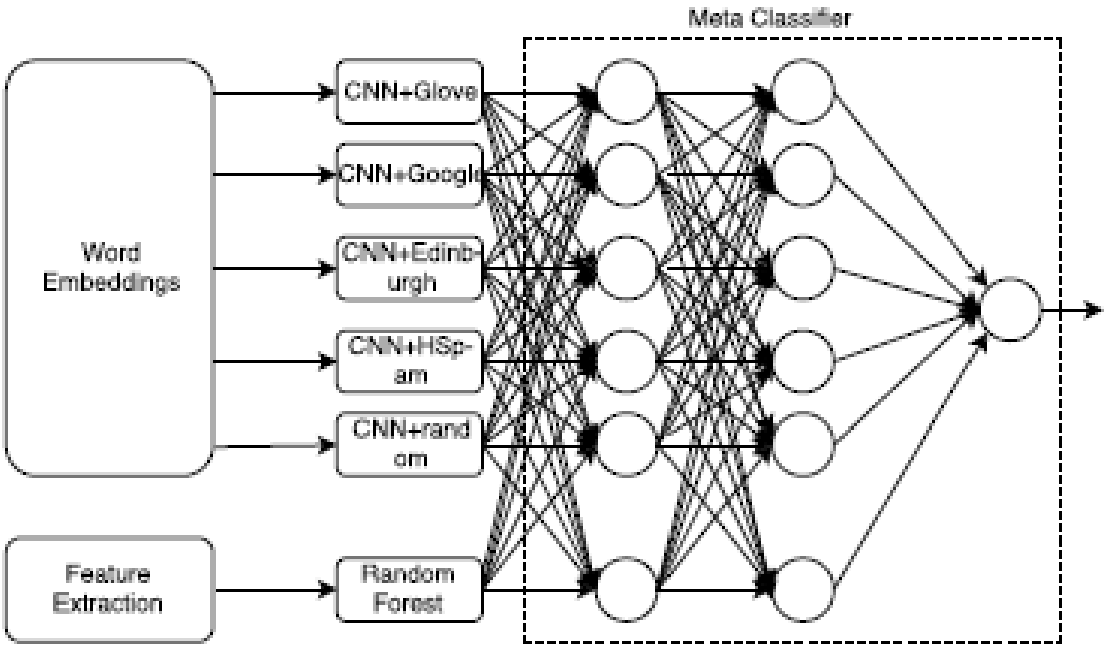

FIGURE 2.7 – Architecture D'ensemble Basée Sur Les Réseaux Neuronaux[11]

2.6 Conclusion

Dans ce chapitre, nous avons donné un aperçu de types de spam puis nous avons présenté les caractéristiques les plus célèbres utilisées dans la détection spam, comme nous avons mentionné un certain nombre de recherches et nous avons discuté leurs résultats atteints.

Dans le dernier chapitre nous utilisons et développons un algorithme d'optimisation EHO pour résoudre ce problème, et ce dernier n'a jamais été utilisé dans des situations pareilles.

CHAPITRE 3

EXPÉRIMENTATION

3.1 Introduction

Dans ce chapitre nous allons mettre en œuvre une implémentation pour la détection de spam dans le Twitter baseé sur la bio-inspiration à l'aide de l'algorithme d'optimisation L'élevage d'éléphants (EHO), puis nous expliquons comment appliquer l'algorithme EHO dans notre travail et nous présentons l'environnement, l'ensemble de données sur lesquels nous avons travaillés, par la suite nous passons aux étapes de mise en œuvre, à la fin nous discutons les résultats obtenus dans cette expérimentation.

3.2 Application d'algorithme EHO

L'algorithme est déduit du comportement des groupes d'éléphants dans la nature qui est utilisé pour résoudre divers des problèmes d'optimisation. EHO modélise les comportements d'élevage des éléphants en deux opérations. Premièrement, le mécanisme de mise à jour du clan. Deuxièmement, le mécanisme de séparation. Dans notre cas, nous utilisons EHO pour faire la détection spam dans TWITTER.

3.2.1 Fonction De Fitness Appliquée

Nous calculons le fitness pour chaque utilisateur à partir de nombre abonnés (followers) et abonnement (following), comme il n'est pas nécessaire qu'il y ait une forte augmentation entre le nombre d' abonnés (followers) et abonnement (following), la fonction de fitness est définie comme suit :

$$\mathbf{R\,(j)} = \frac{\mathbf{n_i\,(j)}}{\mathbf{n_i\,(j) + n_0(j)}} \tag{3.1}$$

Où $n_i(j)$ représente le nombre d'abonnés (followers) de l'utilisateur j et $n_o(j)$ représente le nombre d'abonnement (following).

3.2.2 Population

La population d'éléphants est composée par des clans, et chaque clan a un nombre d'éléphants fixe. Dans notre cas, on va représente la population par l'identificateur d'utilisateur (ID) et leur fitness.

3.2.3 Mise à Jour De Population

Dans chaque clan, la prochaine position de chaque éléphant est influencée par la matriarche. Pour l'éléphant j du clan ci, il peut être mise à jour par1.1, L'éléphant le plus apte (matriarch) peut être mise à jour par1.2. Dans notre cas on fait la mise à jour de population on utiliser le fitness d'utilisateur .

3.2.4 Séparation d'Utilisateur

Les éléphants mâles quittent leur groupe familial et vivent seuls après avoir atteint l'âge adulte. L'éléphant mâle adulte et la plus mauvaise se sépare à chaque génération, mettre en œuvre l'opérateur de séparation à chaque génération comme indiqué dans l'équation 1.4. Pour notre cas nous avons séparé l'utilisateur de mauvaise fitness.

3.3 Environnement

L'environnement matériel utilisé dans cette expérimentation est un ordinateur **hp** qui possède un processeur Intel Core i3 à 2.2 GHZ avec 4 GB RAM, le système d'exploitation est Windows 8.1 professionnel avec 64 bit.

L'environnement d'implémentation utilisé dans cette expérimentation est C# éditeur dans Visual Studio version 2019.

3.3.1 C Sharp (C#)

C (C sharp) est un langage de programmation orienté objet, commercialisé par Microsoft depuis 2003 et destiné au développement sur la plateforme Microsoft.NET. Il est dérivé du C++ et est très proche de Java, dont il reprend la syntaxe et les concepts généraux, en y ajoutant des notions telles que la surcharge d'opérateurs, les indexeurs et les délégués.

FIGURE 3.1 – Logo de C#

3.3.2 Visual Studio

Visual Studio est un ensemble complet d'outils de développement qui permet de créer des applications Web ASP.NET, des services Web XML, des applications de bureau et des applications mobiles. Visual Basic, Visual C++, Visual C et Visual J utilisent tous le même environnement de développement intégré (IDE), ce qui leur permet de partager des outils et facilite la création de solutions utilisant plusieurs langages.

3.4 Ensemble de Donnée

Dans ce travail, nous expérimentons notre méthode avec la collaboration du réseau social Twitter. Twitter. Ci-après sont décrits les ensembles de données exploités.

Chaque ligne représente un tweet de notre ensemble. La dernière colonne est la classe du tweet (spammer ou non-spammer) et les autres colonnes sont les valeurs des caractéristiques.

L'ensemble de donnée est disponible sur le lien suivant :http://nsclab.org/nsclab/resources/ .

3.4.1 Caractéristiques

Les caractéristiques sont listées comme ci-dessous :

- **account age (âge du compte) :**L'âge (jours) d'un compte depuis sa création jusqu'au moment de l'envoi du dernier tweet.
- **no_follower (nombre des abonnés) :**Le nombre des abonnés(followers) de cet utilisateur de Twitter.
- **no_following (nombre des abonnement) :**Le nombre des abonnement(following) de cet utilisateur de Twitter.
- **no_userfavourites (favoris des utilisateurs) :**Le nombre de favoris que cet utilisateur Twitter a reçu.
- **no_lists :**Le nombre de listes que cet utilisateur de Twitter a ajoutées.
- **no_tweets :**Le nombre de tweets envoyés par cet utilisateur de Twitter.
- **no_retweets :**Le nombre de retweets de ce tweet.

- **no_hashtag (nombre de hashtags) :**Le nombre de hashtags disponibles dans le texte du tweet.
- **no_usermention (nombre de mentions) :**Le nombre de comptes (utilisateurs) mentionnés dans le tweet
- **no_urls :**Le nombre d'URLs inclus dans ce tweet.
- **no_char (nombre de caractères) :**Le nombre de caractères numériques dans le texte tweet
- **no_digits (nombre de chiffres) :**Le nombre de chiffres dans ce tweet.

3.5 Implémentation

Le nettoyage de notre ensemble de donné est négligé, nous avons commencé par définir notre population où nous spécifions le nombre de classe (nclan) et le nombre d'utilisateur par classe (nel), chaque utilisateur possède un identificateur(ID) et une valeur de fitness qui est calculé à partir de nombre d'abonnées (followers) et abonnement (following). Puis nous fixons le nombre maximale pour la génération (GenMax), et pour les paramètres suivant :

- α : le facteur d'échelle

- r : l'uniforme distributeur

- β : le facteur qui détermine l'influence d'utilisateur le plus apte (matriarch) sur les autres utilisateurs.

Sont tous des valeurs aléatoires entre 0..1.

3.5.1 Prétraitement

Pendant la phase d'apprentissage on choisir d'utiliser seulement 2% de l'ensemble des données que nous avons divisé en deux classes (spam et non spam), chaque classe possède 100 utilisateur et on fixe le nombre de génération maximale à cinq (GenMax=5).

```
int GenMax = 5;
int[] id = getcol(path, 1);
int[] flower = getcol(path, 2);
int[] flowing = getcol(path, 3);
int nclan =  2;
```

FIGURE 3.2 – Les paramètre initial

Nous affichons la population initial (ID ⇒ Valeur de fitness) comme la Figure 3.3.

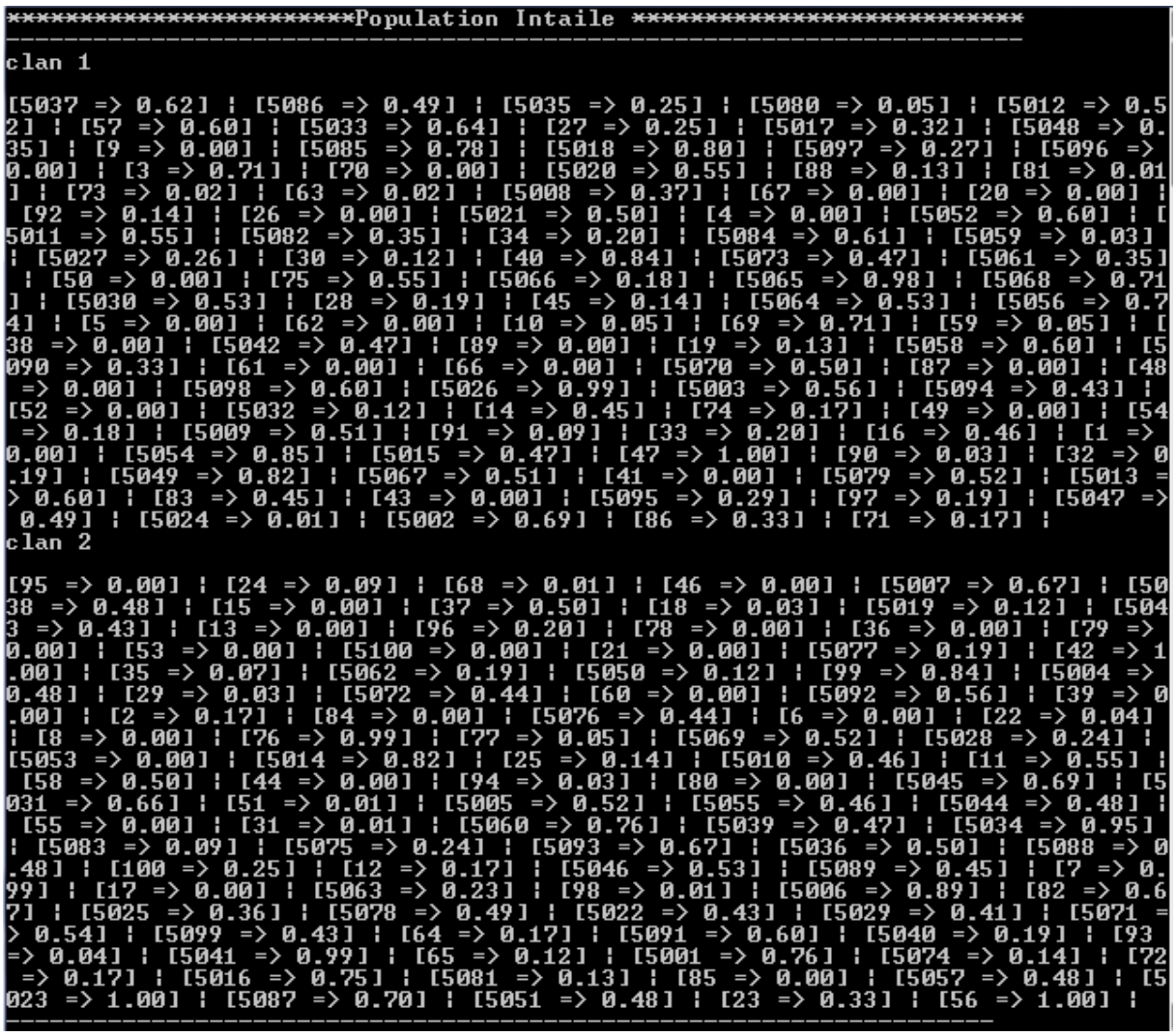

FIGURE 3.3 – Population initial

3.5.2 Résultat et Performance

Pour expérimenter l'algorithme (EHO) et voir son adéquation, nous déterminons la matriarche (meilleur valeur de fitness) et appliquons les étapes de mise à jour pour chaque classe (clan). Une fois le triage est fait en réalise la séparation par le remplacement des valeurs de mauvaise fitness avec des valeur obtenu du l'équation 1.4 après son triage.

Ensuit le programme détecte le spam et non spam par la valeur de fitness, inférieur à 0.5 ($\prec 0.5$) pour classe spam et supérieure ou égal à 0.5 ($\geqslant 0.5$) pour classe non spam, pour la première expérimentation le résultat de programme est satisfiable avec une précision atteinte 47.50%, Voir la Figure 3.4

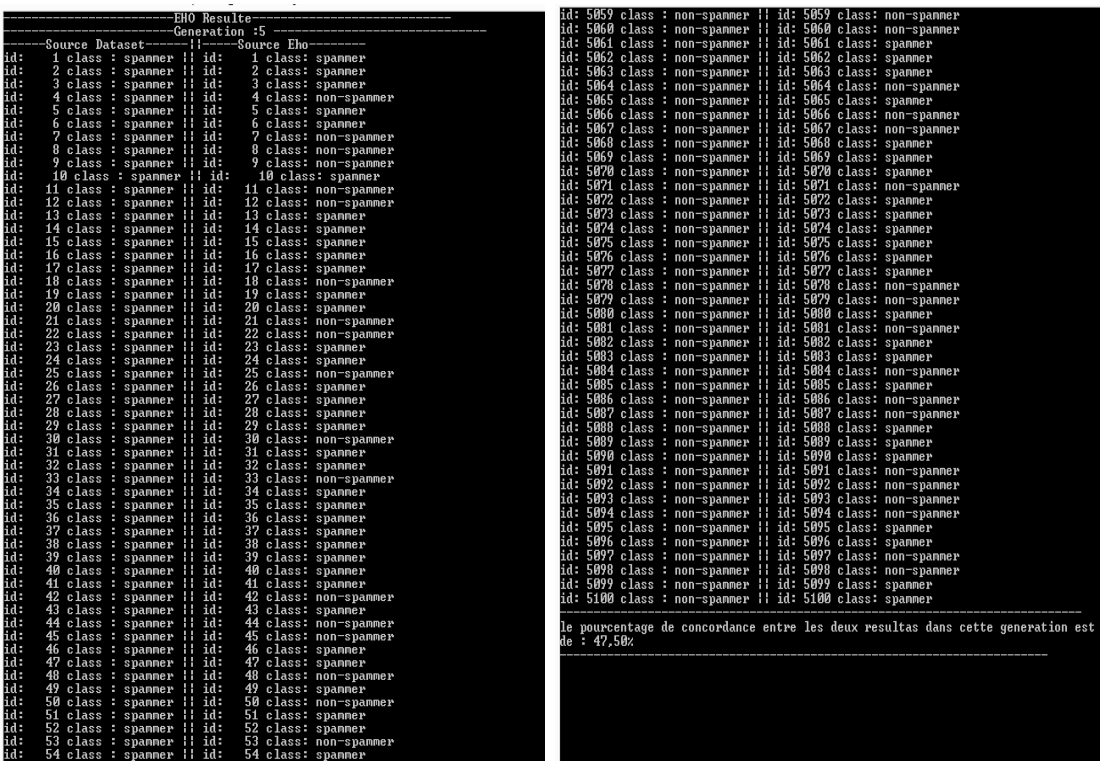

FIGURE 3.4 – Résultat de détection spam pour première expérimentation

Après plusieurs expérimentation avec des valeur de paramètre de génération maximal suivants (GenMax=15 et GenMax=27), on a abouti à des meilleurs résultats comme les Figures 3.5, 3.6.

Pour la valeur de génération maximal GenMax=15 une précision atteinte 49%.

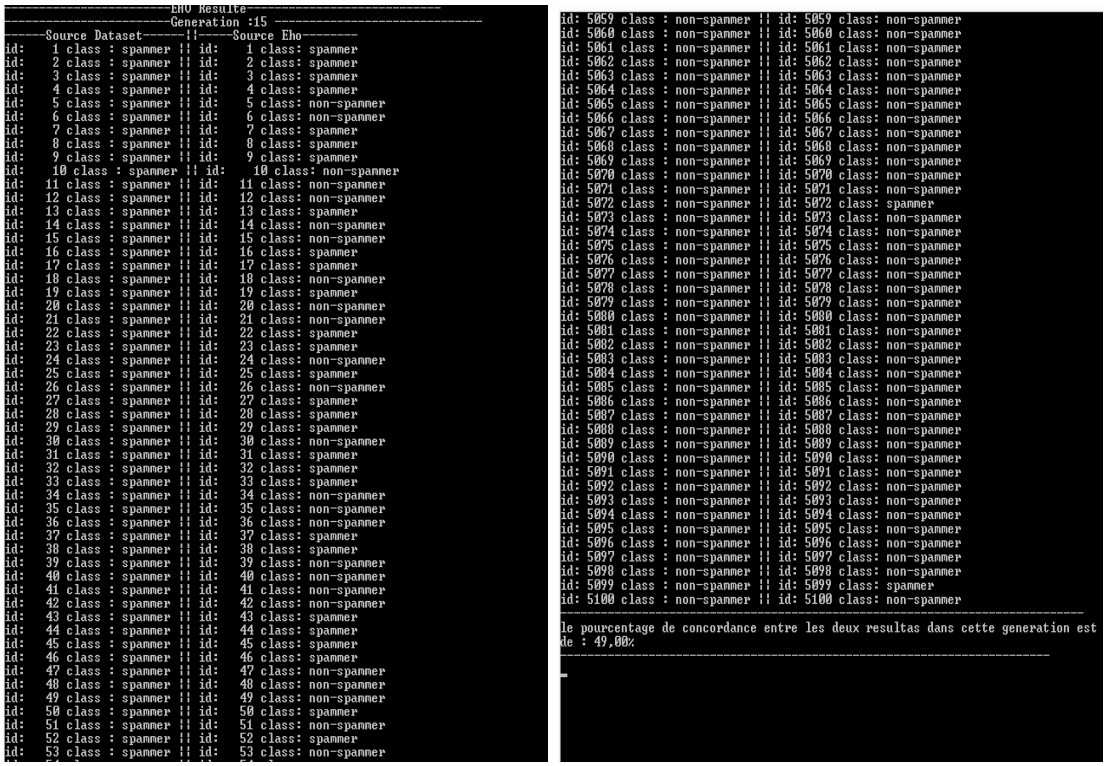

FIGURE 3.5 – Résultat de détection spam pour GenMax=15

Pour la valeur de génération maximal GenMax=27 une précision atteinte 64%.

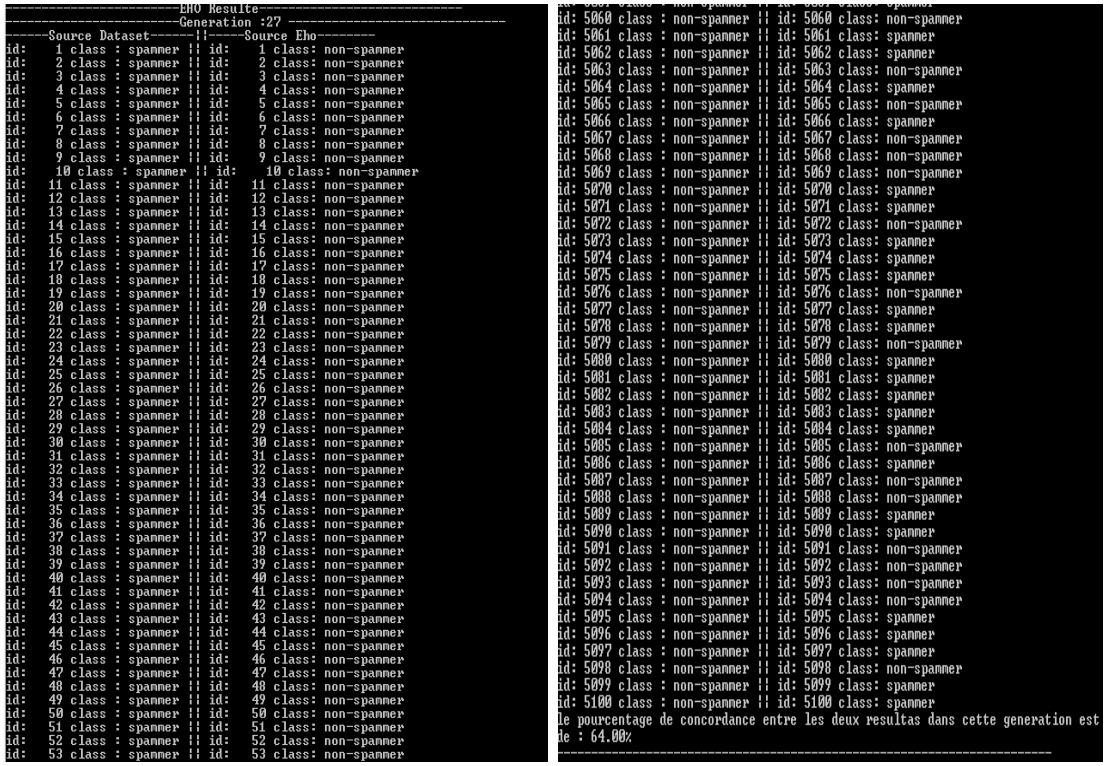

FIGURE 3.6 – Résultat de détection spam pour GenMax=27

3.5.3 Discussion

Après de nombreuses expérimentations et d'après ce que nous avons appris sur le problème de détection, le résultat obtenu par notre travail est satisfiable d'après sa première utilisation dans se genre de problème. Sa pourrait être mieux, si les environnements matériel n'est pas limité et le temps limité accordé à notre recherche, ainsi que les jeux d'essai de données sur lesquels ont a fait nos expérimentations.

La précision de l'implémentation est affectée par de nombreux paramètres :

- Le valeur de générations maximal (GenMax)

- Le nombre d'utilisateur dans classe (nel)

- Les paramètres de facteur d'échelle α, uniforme distributeur r, facteur qui détermine l'influence de l'utilisateur le plus apte (matriarch) sur utilisateur β.

Nous ne pouvons pas garantir des meilleurs précision, donc Nous continuons d'essayer et changement dans des paramètres d'évaluation.

Le résultat peut être amélioré en ajoutant plus des caractéristique comme le nombre d'URL, nombre de retweet, hashtag.

Un autre voie à suivre afin d'améliorer les performances du travail est d'incorporer une méthode de classification avec l'algorithme (EHO).

3.6 Conclusion

Dans ce chapitre, nous avons utilisé l'algorithme EHO pour la détection des spams, et nous avons sélectionné deux caractéristiques basés sur l'utilisateur par le nombre d'abonnés (followers) et d'abonnements (following).

Le résultat obtenu par notre travail est satisfiable d'après sa première utilisation dans ce genre de problème, et peut être amélioré par l'augmentation de la valeur de génération maximal (GenMax) et avec un plus grand nombre de data set.

CONCLUSION

Le but de ce travail est de détecter le spam dans Twitter on utilisant les algorithmes de Bio inspiration, nous avons commencé par un aperçu de l'intelligence artificielle et le méta-heuristique, puis nous avons introduit le domaine de Bio-inspiré en concentrant sur les algorithmes d'optimisations et spécifiquement l'algorithme EHO.

Après nous avons passé à l'état de l'art, des techniques de l'apprentissage automatique et l'apprentissage profond pour la détection de spam dans Twitter. Par la suite, nous avons présenté notre modèle de détection où nous avons utilisé deux caractéristiques basés sur l'utilisateur, le nombre d'abonnés (followers) et d'abonnements (following) par l'application de l'algorithme EHO.

Nous avons utilisé un ensemble de données du 10000 comptes, et nous avons sélectionné 200 comptes pour faire l'expérimentation de l'efficacité du notre implémentation, la précision de la détection est estimée à être 47.5 % obtenue à partir de la première expérimentation, et 49 %,64 % pour les autres.

Afin d'obtenir des meilleurs résultats, nous suggérons d'utiliser plusieurs caractéristiques pour améliorer la précision, nous pouvons fusionner l'algorithme de EHO avec un autre algorithme de classification.

BIBLIOGRAPHIE

[1] Anthony McGregor, Mark Hall, Perry Lorier, and James Brunskill. Flow clustering using machine learning techniques. In *International workshop on passive and active network measurement*, pages 205–214. Springer, 2004.

[2] BAKHOUIA Roqiya Abed Djemaa. *Heuristique Et Métaheuristique*. PhD thesis, Université d'Adrar Faculté des Sciences et de la Technologie, 2016.

[3] L Saïd. Méthodes bio-inpirées hybrides pour la résolution de problèmes complexes. *Université Constantine*, 2, 2013.

[4] Abbas El Dor. *Perfectionnement des algorithmes d'optimisation par essaim particulaire : applications en segmentation d'images et en électronique*. PhD thesis, Université Paris-Est, 2012.

[5] Gai-Ge Wang, Suash Deb, and Leandro dos S Coelho. Elephant herding optimization. In *2015 3rd International Symposium on Computational and Business Intelligence (ISCBI)*, pages 1–5. IEEE, 2015.

[6] Juan Li, Hong Lei, Amir H Alavi, and Gai-Ge Wang. Elephant herding optimization : variants, hybrids, and applications. *Mathematics*, 8(9) :1415, 2020.

[7] Michael Mccord and M Chuah. Spam detection on twitter using traditional classifiers. In *international conference on Autonomic and trusted computing*, pages 175–186. Springer, 2011.

[8] Abdullah Talha Kabakus and Resul Kara. A survey of spam detection methods on twitter. *International Journal of Advanced Computer Science and Applications*, 8(3) :29–38, 2017.

[9] Deepali M. Gohi Ashwini Athawale. Spam detection on collection of twitter data using naive bayes algorithm. *International Journal of Innovative Research in Science, Engineering and Technology*, 6 :7105–7110, 2018.

[10] Yash Thigale and Prof. Ms. Deipali Gore. Twitter spam classification and detection using svm. *AEGAEUM JOURNAL*, (6) :78–84, 2020.

[11] Sreekanth Madisetty and Maunendra Sankar Desarkar. A neural network-based ensemble approach for spam detection in twitter. *IEEE Transactions on Computational Social Systems*, 5(4) :973–984, 2018.

[12] Mehmet ŞİMŞEK, Oğuzhan YILMAZ, Asena Hazal KAHRİMAN, and Levent SABAH. Detecting fake twitter accounts with using artificial neural networks. *Artificial Intelligence Studies*, 1(1) :26–29, 2018.

[13] Dario Floreano and Claudio Mattiussi. Bio-inspired artificial intelligence. *Ch*, 5 :335–396, 2008.

[14] Cédric Villani, Yann Bonnet, Charly Berthet, François Levin, Marc Schoenauer, Anne Charlotte Cornut, and Bertrand Rondepierre. *Donner un sens à l'intelligence artificielle : pour une stratégie nationale et européenne*. Conseil national du numérique, 2018.

[15] Semestre Printemps and Jacques Savoy. Intelligence artificielle (3in1007).

[16] Baarir Nihel Fatima. Fake news detection using machine learning. Master's thesis, biskra, 2019.

[17] Lyes Belhoul. *Résolution de problèmes d'optimisation combinatoire mono et multi-objectifs par énumération ordonnée*. PhD thesis, Université Paris Dauphine-Paris IX, 2014.

[18] Me Aroussi. *HEURISTIQUES ET MÉTA-HEURISTIQUES*. PhD thesis, blida1, 2015.

[19] Hadjer Gahgah. *Problème d'emploi de temps : proposition un algorithme bio-inspiré*. PhD thesis, FACULTE DES MATHEMATIQUES ET DE L'INFORMATIQUE DEPARTEMENT D'INFORMATIQUE, 2018.

[20] Perig Pitrou, Anne Dalsuet, and Bérengère Hurand. Modélisation, construction et imitation des processus vitaux. approche pluridisciplinaire du biomimétisme, 2015.

[21] S Mohamed. Traitement d'images par des approches bio-inspirées application à la segmentation d'images, 2014.

[22] Gai-Ge Wang, Suash Deb, Xiao-Zhi Gao, and Leandro Dos Santos Coelho. A new metaheuristic optimisation algorithm motivated by elephant herding behaviour. *International Journal of Bio-Inspired Computation*, 8(6) :394–409, 2016.

[23] Clay Shirky. Blog explosion and insider's club : Brothers in cluelessness. 2004.

[24] Prabhjot Kaur, Anubha Singhal, and Jasleen Kaur. Spam detection on twitter : A survey. In *2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom)*, pages 2570–2573. IEEE, 2016.

[25] Rutuja Katpatal and Aparna Junnarkar. An efficient approach of spam detection in twitter. In *2018 International Conference on Inventive Research in Computing Applications (ICIRCA)*, pages 1240–1243. IEEE, 2018.

[26] Nitin Agarwal and Yusuf Yiliyasi. Information quality challenges in social media. In *ICIQ*, 2010.

[27] Christopher D Manning, Prabhakar Raghavan, and Hinrich Schütze. Text classification and naive bayes. *Introduction to information retrieval*, 1(6), 2008.

[28] Tingmin Wu, Sheng Wen, Yang Xiang, and Wanlei Zhou. Twitter spam detection : Survey of new approaches and comparative study. *Computers & Security*, 76 :265–284, 2018.

[29] Monika Verma and Sanjeev Sofat. Techniques to detect spammers in twitter-a survey. *International Journal of Computer Applications*, 85(10), 2014.

[30] Kevin P Murphy et al. Naive bayes classifiers. *University of British Columbia*, 18(60), 2006.

[31] Ingo Steinwart and Andreas Christmann. *Support vector machines*. Springer Science & Business Media, 2008.

[32] Marti A. Hearst, Susan T Dumais, Edgar Osuna, John Platt, and Bernhard Scholkopf. Support vector machines. *IEEE Intelligent Systems and their applications*, 13(4) :18–28, 1998.

[33] Fabricio Benevenuto, Gabriel Magno, Tiago Rodrigues, and Virgilio Almeida. Detecting spammers on twitter. In *Collaboration, electronic messaging, anti-abuse and spam conference (CEAS)*, volume 6, page 12, 2010.

[34] Ricco Rakotomalala. Arbres de décision. *Revue Modulad*, 33 :163–187, 2005.

[35] MC Pull, CB Pull, and P Pichot. Des critères empiriques français pour les psychoses : Iii. algorithmes et arbre de décision. *L'Encéphale : Revue de psychiatrie clinique biologique et thérapeutique*, 1987.

[36] Po-Ching Lin and Po-Min Huang. A study of effective features for detecting long-surviving twitter spam accounts. In *2013 15th International Conference on Advanced Communications Technology (ICACT)*, pages 841–846. IEEE, 2013.

[37] Samira Pouyanfar, Saad Sadiq, Yilin Yan, Haiman Tian, Yudong Tao, Maria Presa Reyes, Mei-Ling Shyu, Shu-Ching Chen, and SS Iyengar. A survey on deep learning : Algorithms, techniques, and applications. *ACM Computing Surveys (CSUR)*, 51(5) :1–36, 2018.