

MS/003-48/01

الجمهورية الجزائرية الديمقراطية الشعبية
وزارة التعليم العالي والبحث العلمي

Université Abou Bekr Belkaid



جامعة أبي بكر بلقايد

تلمسان الجزائر

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
UNIVERSITE ABOU BEKR BELKAID
Faculté des sciences
Département d'Informatique

*Mémoire pour l'obtention du grade de Master
En Informatique*

Option : système d'informations et de connaissances (SIC)



Réalisation d'un système immuno-artificiel pour le diagnostic du diabète

Soutenu le :

03 Juillet 2012

Par :

Mr. BELKHITER Mohammed

Devant la Commission d'Examen

Jury

Mr. BENZAOUZ.M

Président

Mr. CHIKH.A

Encadreur

Mr ELHABIB DAHO.M

Co-encadreur

Mme BEREMDANE.D

Examinatrice

Mr. BENZIANE.Y

Examineur

2011 – 2012

TABLE DES MATIERES

INTRODUCTION GENERALE.....	4
----------------------------	---

CHAPITRE I : PRESENTATION GENERALE DU DIABETE

1.1 Introduction.....	7
1.2 Définition	7
1.2.1 Diabète de type 1.....	8
1.2.2 Diabète de type 2.....	8
1.3 Cause du diabète.....	9
1.4 Symptômes et Complications.....	10
1.5 Critères diagnostiques.....	11
1.6 Traitement.....	11
1.6.1 Traitement non pharmacologique du DM2.....	12
1.6.2 Traitement avec des antidiabétiques oraux.....	12
1.6.2.1 Sulfonylurée ou sulfamides hypoglycémiants.....	12
1.6.2.2 Les méglitinides.....	12
1.6.2.3 Biguanides, Glucophage, Diabamyl, Metformine ou Stagid.....	13
1.6.2.4 Les thiazolidinediones ou glitazones.....	13
1.6.2.4 Les inhibiteurs des alpha-glucosidases.....	13
1.6.3 Insulinothérapie.....	13
1.7 Complications.....	14
1.8 Prévention.....	15
1.8.1 Prévention primaire.....	15
1.8.2 Prévention secondaire.....	15
1.8.3 Prévention tertiaire.....	15
1.9 Facteurs de risque de diabète de type 2.....	16
1.10 Conclusion.....	17

CHAPITRE II : LES SYSTEMES IMMUNITAIRES

A. Systèmes immunitaires naturels	
1.1 Introduction.....	18
1.2 Systèmes immunitaires naturels.....	18
1.3 Historique.....	18
1.4 Eléments du système immunitaire.....	19
1.4.1 Les organes du système immunitaire.....	19
1.4.2 Les cellules du système immunitaire.....	20
1.5 Architecture du système immunitaire.....	22
1.5.1 L'immunité innée (naturelle non spécifique ou naïve).....	22
1.5.2 L'immunité acquise (spécifique).....	23
1.6 Propriétés du système immunitaire.....	24
1.7 Fonctionnement du système immunitaire.....	25
1.8 Détection du soi et non-soi.....	27

1.8.1 Sélection positive.....	27
1.8.2 Sélection négative.....	27
1.9 Théorie de la sélection clonal.....	28
1.10 Déroulement d'une réponse immunitaire.....	29

B. Système Immunitaire Artificiel

2.1 Introduction.....	31
2.2 Systèmes immunitaires naturels.....	31
2.3 Les algorithmes immunitaires.....	32
2.3.1 La sélection clonale.....	32
2.3.2 La sélection négative.....	34
2.3.3 Les réseaux immunitaires (ou idiotypiques).....	36
2.4 AIRS: Artificial Immune Recognition System.....	37
2.4.1 Introduction.....	37
2.4.2 Définition.....	39
2.5 Implémentation d'un SIA.....	41
2.6 Domaines d'application des SIAs.....	42
2.6.1 Robotique.....	42
2.6.2 Optimisation.....	42
2.6.3 Sécurité des ordinateurs.....	43
2.6.4 La détection et l'élimination des virus informatiques.....	43
2.6.5 Reconnaissance de formes.....	43
2.6.6 Diagnostic médical.....	44
2.6.7 Autres applications.....	45
2.7 Conclusion.....	45

CHAPITRE III : RESULTATS ET DISCUSSION

1.1 Introduction.....	47
1.2 Description de la base de données utilisée.....	47
1.2.1 L'intérêt de la base de données.....	47
1.2.2 Base de données utilisées.....	49
1.3 Algorithme modified-AIRS.....	50
1.3.1 Principe.....	50
1.3.2 Etapes d'apprentissage.....	52
1.3.2.1 Initialisation.....	52
1.3.2.2 Identification des cellules B et génération des ARBs.....	52
1.3.2.3 Compétition des ressources et développement des cellules mémoires candidates... 54	
1.3.2.4 Introduction des cellules mémoires.....	55
1.3.3 Classification.....	55
1.4 Résultats.....	55
1.4.1 Choix des paramètres.....	55
1.4.1.1 Pas de mutation.....	56
1.4.1.2 Seuil de stimulation.....	56
1.4.1.3 Scalaire du seuil d'affinité (ATS).....	56
1.4.1.4 Taux de clonage.....	56
1.4.1.5 Hyper taux de clonage.....	56
1.4.1.6 Nombre de ressources.....	56

1.4.2 Expérimentations.....	57
1.5 Comparaison.....	58
1.6 Conclusion.....	58

LISTE DES TABLEAUX

Tableau 1 : Taux de diabétiques en 2000 et les prévisions pour le 2030.....	7
Tableau 2 : Correspondance entre AIRS et le système immunitaire.....	39
Tableau 3 : Description des attributs de la base.....	50
Tableau 4 : Temps d'apprentissage en millisecondes.....	54
Tableau 5 : Les résultats obtenus.....	57
Tableau 6 : Etude comparative.....	58

LISTE DES FIGURES

Figure 1 : Fonctionnement de l'insuline.....	9
Figure2 Organes du système immunitaire.....	20
Figure 3 Structure d'un antigène avec ses épitopes.....	21
Figure 4 Un lymphocyte, principale composante du système immunitaire.....	21
Figure5 Architecture du système immunitaire.....	22
Figure 6 La structure multicouche du système immunitaire.....	23
Figure 7 Déroulement de la réponse immunitaire.....	26
Figure 8 Réponses immunitaire primaire et secondaire.....	26
Figure 9 Sélection négative dans le thymus.....	28
Figure 10 Processus de sélection clonale.....	29
Figure11: Génération de l'ensemble de détecteurs.....	35
Figure12 : Surveillance d'éléments du non-soi.....	35
Figure 13 : Zones de présence des indiens Pima.....	49
Figure 14 : Représentation de l'algorithme m-AIRS.....	51

Introduction générale

L'apprentissage est la capacité de s'améliorer avec l'expérience, de se rappeler les décisions antérieures et les résultats afin de faire de meilleurs choix à l'avenir dans des situations similaires. L'apprentissage automatique est une discipline de l'intelligence artificielle. D'après Tom Mitchell, l'apprentissage automatique cherche à trouver le moyen de construire des programmes informatiques qui s'améliorent automatiquement avec l'expérience. Quand à H. Simon, il a présenté l'une des métaphores utilisées dans le domaine de l'apprentissage automatique et qui considère la résolution de problèmes comme un type d'apprentissage qui, une fois le problème résolu, est capable de reconnaître la problématique et réagir en utilisant la stratégie apprise.

L'une des motivations les plus importantes de la construction des systèmes d'apprentissage automatique réside dans le fait que dans plusieurs domaines l'expérience est insuffisante et la codification de la connaissance qui la décrit est limitée, fragmentée et donc incomplète. En plus doter un agent avec toute la connaissance nécessaire est une tâche très complexe, coûteuse, longue et dans laquelle l'élimination des erreurs possibles est difficile et nécessite une attention spécialisée [1]. Il n'est pas surprenant que les informaticiens se soient souvent inspirés des systèmes biologiques afin de construire des systèmes qui travaillent avec l'intelligence ou l'efficacité que possèdent les systèmes naturels. Ces systèmes permettent de résoudre des problèmes complexes tels que la reconnaissance d'objets, des langages naturels, de l'écriture, de formes syntaxiques, l'aide au diagnostic médical, la bioinformatique, l'analyse financière, la classification des séquences d'ADN, etc. Il existe plusieurs modèles qui s'inspirent de la biologie tels que les réseaux de neurones, les algorithmes génétiques, les algorithmes de colonies de fourmis, les algorithmes d'optimisation par essais particuliers, les systèmes immunitaires artificiels, etc.

Le système immunitaire est un ensemble complexe de cellules et organes divers interdépendant qui protège le corps. Pour les vertébrés, le système immunitaire est composé de deux couches: le système immunitaire inné qui est assez statique et représente la première ligne de défense et le système immunitaire adaptatif qui est un système en constante évolution dynamique qui réagit et interagit avec l'environnement dans une variété de façons. Le système immunitaire offre des possibilités alléchantes d'un point de vue informatique. Dans les dernières années plusieurs algorithmes s'inspirant des systèmes immunitaires sont apparus, le nom collectif de l'ensemble de ces méthodes est Systèmes Immunitaires Artificiels. Les SIAs

Introduction générale

sont utilisés pour résoudre plusieurs problèmes tels que la robotique, la sécurité des ordinateurs, l'optimisation, l'analyse de données, l'apprentissage artificiel et l'aide au diagnostic médical. C'est dans ces deux dernières catégories que se range notre travail.

Le diagnostic médical est un processus de classification. L'utilisation de l'informatique pour la réalisation de cette classification devient de plus en plus fréquente. Même si la décision de l'expert est le facteur le plus important lors du diagnostic, les systèmes de classification fournissent une aide substantielle, car elles réduisent les erreurs dues à la fatigue et le temps nécessaire pour le diagnostic. Actuellement, la plupart des hôpitaux modernes sont bien équipés avec des dispositifs de collecte de données. Ces données seront partagées en inter- et intra-systèmes d'information hospitaliers. Ce qui était avant une base de données isolées ou un système d'information de laboratoire est maintenant intégrée dans un système d'information médicale à plus grande échelle (ministères, hôpitaux, ou à base communautaire). L'augmentation du volume de données entraîne des difficultés à extraire des informations utiles pour l'aide à la décision. Les méthodes traditionnelles d'analyse de données sont devenues insuffisantes, et les méthodes dites intelligentes sont indispensables. [2]

Dans ce mémoire nous nous intéresserons au diagnostic du diabète qui est un dysfonctionnement du système de régulation de la glycémie.

La méthode appliquée dans ce travail pour la reconnaissance automatique du diabète est une hybridation de deux méthodes. La première méthode est l'AIRS (Artificial Immune Recognition System) qui est un algorithme des systèmes immunitaire artificiel à apprentissage supervisé. Cette méthode effectue un apprentissage sur la base de données Pima Indian Diabetes (PID) et donnera comme résultat un ensemble réduit de vecteur de données qui nous permettra d'effectuer la classification par la suite. La deuxième méthode est l'algorithme du k-plus proche voisin elle effectue la classification grâce à la nouvelle base de données réduit généré à la fin de l'apprentissage.

Le manuscrit est organisé en trois principaux chapitres encadré par une introduction générale et une conclusion générale.

Le 1er chapitre présente un aperçu général sur la maladie du diabète en citant les causes, le diagnostic, les traitements et la prévention. Le 2eme chapitre résume l'essentiel des principes des systèmes immunitaires artificiels (SIAs). Il commence par un aperçu du système immunitaire naturel (SIN) avec ses différentes théories. En deuxième lieu, le chapitre présente les quatre principaux algorithmes des SIAs.

Introduction générale

Le 3eme chapitre détaille l'approche proposée dans ce travail en commençant par présenter les algorithmes utilisés, m-AIRS et le K-ppv. Ensuite, nous présentons la base de données PID utilisée pour évaluer et comparer les approches. Finalement les résultats sont présentés, comparés et interprétés. Des pistes de recherche futures ouvertes pour ce travail sont présentées dans la conclusion générale.

Présentation générale du diabète

1.1 Introduction

Le diabète est l'une des maladies les plus dangereuses, connu aussi avec le nom de "tueur silencieux". Cette maladie est un problème majeur de santé dans les pays industrialisés et en développement, et son incidence est en augmentation avec plus de 220 millions de personnes diabétiques dans le monde. Il est la quatrième cause de décès. L'augmentation du nombre de diabétique est tellement rapide que l'organisation mondiale de la santé (OMS) l'a identifié comme étant une épidémie. D'après l'OMS, le nombre d'adulte diabétiques atteindra les 300 millions de malades en 2025 (Tableau 1).

Allemagne	2.626.842	3.770.815
Autriche	238.930	366.120
Belgique	317.342	461.439
Danemark	156.505	2.324.28
Finlande	158.580	239.282
Espagne	2.717.401	3.751.632
France	1.753.243	2.645.444
Grèce	853.246	1.077.022
Irlande	85.787	156.835
Italie	4.252.036	5.373.724
Luxemburg	12.057	21.193
Pays Bas	425.676	719.753
Portugal	662.283	882.428
Royaume Unis	1.804.943	2.665.884
Suisse	291.908	404.414

Tableau 1 : Taux de diabétiques en 2000 et les prévisions pour le 2030 [3]

1.2 Définition

Le diabète est une maladie dans laquelle la production d'insuline est insuffisante pour répondre aux besoins de l'organisme et ou la réponse des cellules à l'insuline est anormale. Le carburant principal de l'organisme est un type de sucre appelé glucose, qui provient des aliments (une fois ceux-ci dégradés). Le glucose passe dans le sang, puis les cellules l'utilisent pour obtenir de l'énergie. Pour pouvoir se servir du glucose, l'organisme a besoin d'une hormone appelée insuline, qui est synthétisée par le pancréas.

La classification du diabète se base principalement sur son étiologie et caractéristique physiopathologique. Le diabète est classé en quatre types :

- Diabète type 1 (DM1);
- Diabète type 2 (DM2);

Présentation générale du diabète

d'insuline pour survivre, près du 40 % des malades finissent par en avoir besoin pour contrôler la glycémie. L'hyperinsulinisme de plus de 80 % de diabétique de type 2 est la conséquence de leur obésité (la graisse abdominal est la plus dangereuse). [4]

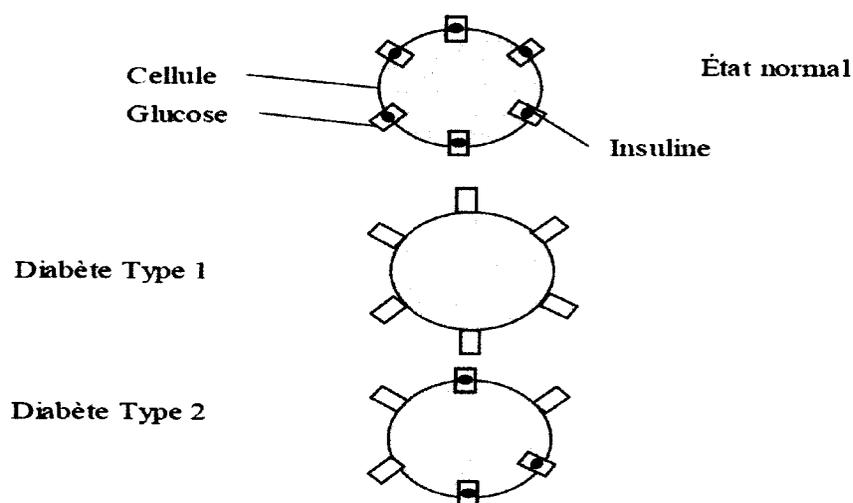


Figure 1 : Fonctionnement de l'insuline

1.3 Cause du diabète

La prévalence de cette maladie a été multipliée par cinq en moins de cinquante ans. Cette augmentation progressive est due à divers facteurs : [3]

- le vieillissement global de la population,
- l'augmentation de l'espérance de vie du diabétique,
- l'augmentation de la fécondité des femmes diabétique,
- l'augmentation de l'obésité,
- l'incrémentation de la consommation des sucres raffinés.

Ainsi que d'autre facteurs qui peuvent servir comme déclencheur tels que :

- le sédentarisme,
- les régimes riches en graisse et protéine,
- la consommation réduite de fibre,
- une alimentation déficiente en hydrate de carbone complexe et vitamine E,
- le stress chronique.

Présentation générale du diabète

- Autre types spécifique de diabète;
- Diabète gestationnel (DMG).

1.2.1 Diabète de type 1

Quand la maladie est diagnostiquée, la sécrétion d'insuline est déficiente mais pas inexistante. La sécrétion d'insuline est insuffisante aussi bien à jeun qu'en réponse aux différents stimulus, ceci est une conséquence de l'autodestruction progressive et sélective des cellules Béta des îlots de Langerhans, ce qui affecte l'utilisation des hydrates de carbone, protéine et graisse.

Puisque un pancréas sain sécrète une quantité d'insuline beaucoup plus élevée de ce dont le corps a besoin, des mois ou des années peuvent passer avant que la maladie soit diagnostiquée. La vitesse à laquelle les cellules se détruisent dépendra de l'âge du malade, étant plus rapide pour les bébés et les enfants et plus lente pour les adultes. Une fois le traitement avec insuline établi, l'organisme passera par une période, allant jusqu'à un an, durant laquelle les sécrétions d'insuline se réinitialisent et les besoins exogènes diminuent, sans pour autant abandonner le traitement. Après une dizaine d'années, les cellules Béta seront entièrement détruites, donc toute l'insuline nécessaire devra être administrée par injection. [3]

1.2.2 Diabète de type 2

Dans ce type de diabète l'altération métabolique n'est pas aussi intense que pour DM1 et l'évolution de la maladie est progressive. Ce diabète est caractérisé par la résistance ou la faible sensibilité du corps à l'insuline, c'est-à-dire que le taux d'insuline endogène peut se trouver dans les paramètres normaux, mais les tissus sont incapables de l'assimiler et par conséquent le taux de glucose dans le sang augmente. [4]

L'insuline agit à niveau cellulaire à travers de certains récepteurs de membrane (Figure 1). La liaison insuline-récepteur active un deuxième messager qui induit la synthèse des protéines et l'activation et inhibition des enzymes intracellulaire. Les malades souffrant de diabète de type 2 ont des altérations dans les mécanismes post-récepteurs, ce qui oblige l'organisme à augmenter la sécrétion d'insuline pour compenser et ceci peut conduire à l'épuisement des cellules Béta. Pour des personnes avec une certaines prédisposition les cellules ne seront pas capable de maintenir un taux de glucose normale, ce qui conduit à l'apparition du diabète. Même si les diabétiques de type 2 ne nécessitent pas les injections

1.4 Symptômes et Complications

Les personnes atteintes du diabète de type 2 peuvent ne présenter aucun symptôme pendant des années ou des dizaines d'années, mais, au fur et à mesure que la maladie évolue, les symptômes apparaissent. Les signes et les symptômes possibles sont les suivants :

- Soif et faim accrues,
- besoin fréquent d'uriner,
- fatigue,
- peau sèche accompagnée de démangeaisons,
- cicatrisation lente des coupures et des lésions,
- vision trouble,
- infections urinaires et vaginales fréquentes,
- diminution des sensations ou engourdissement dans les mains et les pieds,
- impuissance chez l'homme (dysfonction érectile).

Malheureusement, dans beaucoup de cas, le diabète de type 2 passe inaperçu pendant plusieurs années et il est diagnostiqué seulement lorsque la personne consulte le médecin en raison de complications liées au diabète.

Un taux élevé de sucre dans le sang (hyperglycémie) peut mener à un trouble appelé intoxication au glucose, qui aggrave l'atteinte du pancréas et réduit la capacité de l'organisme à produire de l'insuline. Sans insuline, le taux de sucre continue d'augmenter et atteint un niveau qui provoque des lésions dans divers organes tels que les yeux, les nerfs et les reins. Une atteinte des yeux, plus particulièrement de la rétine, s'appelle rétinopathie diabétique et représente la première cause de cécité. Une atteinte des reins, un trouble appelé néphropathie diabétique, peut mener à l'insuffisance rénale et à la dialyse. Une atteinte des nerfs qui innervent les bras, les jambes et les voies digestives est qualifiée de neuropathie. Enfin, certains diabétiques qui ont une neuropathie périphérique (atteinte des nerfs dans les jambes) et une mauvaise circulation sanguine dans les jambes auront besoin d'une amputation un jour ou l'autre.

Les personnes diabétiques sont exposées à un risque accru de durcissement des grosses artères (athérosclérose), ce qui peut provoquer une crise cardiaque, un accident vasculaire cérébral ou une mauvaise circulation sanguine dans les jambes. Les lésions des petits vaisseaux sanguins peuvent nuire aux yeux, aux reins et aux nerfs.

Si la glycémie devient très élevée, plus particulièrement en présence d'autres facteurs de stress telle qu'une infection, les personnes atteintes du diabète de type 2 peuvent

Présentation générale du diabète

présenter des épisodes de confusion et d'étourdissements, voire des convulsions. On parle alors d'un coma hyperosmolaire, sans acidocétose et avec hyperglycémie, qui exige des soins médicaux immédiats.

Heureusement, il est possible de prévenir les complications du diabète, de retarder leur apparition ou de freiner leur évolution en maîtrisant la glycémie et en la maintenant le plus près possible des valeurs normales (glycémie à jeun normale : de 4 µl/L à 6 µl/L). [5]

1.5 Critères diagnostiques

Seule une prise de sang faite en laboratoire déterminera avec certitude l'état de santé. Elle mesurera le taux de glucose (sucre) dans le sang.

Les critères diagnostiques du diabète proposés par Les Lignes Directrices de Pratique Clinique 2008 pour la prévention et le traitement du diabète au Canada sont :

- Glycémie, à jeun : 1,26 g/L (7,0 mmol/L) et plus

Ou

- Glycémie, 2 heures après l'ingestion de 75 g de glucose : 2 g/L (11,1 mmol/L) et plus

Ou

- Glycémie, à tout moment de la journée : 2 g/L (11,1 mmol/L) et plus, avec les symptômes classiques

Chez une personne non-diabétique, on considère normal:

- Glycémie, à jeun: moins de 5,6 mmol/L (entre 0,7 et 1,10 g/l)
- Glycémie, 2 heures après l'ingestion de 75 g de glucose : moins de 7,8 mmol/L

Parfois, une deuxième analyse est nécessaire afin de confirmer les résultats obtenus.

À partir de 40 ans, il est recommandé que toute personne qui ne présente pas de facteurs de risques fasse mesurer sa glycémie à jeun à tous les trois ans. [5]

1.6 Traitement

Le contrôle strict du taux de sucre dans le sang est le principe fondamental du traitement du diabète. Le traitement de référence pour le diabète est constitué d'un ensemble de mesures diététiques et d'hygiène de vie, des antidiabétiques oraux et parfois des injections d'insuline.

1.6.1 Traitement non pharmacologique du DM2

Le traitement non pharmacologique et en particulier la réduction du poids pour le patient obèse, est le seul traitement intégral capable de contrôler simultanément la plupart des problèmes métaboliques du patient atteint de diabète de type 2, ainsi que l'hyperglycémie, la résistance à l'insuline, l'hypertriglycéridémie et l'hypertension artérielle. Il est possible d'obtenir des changements significatifs avec une réduction de 5 à 10 % du poids, et par conséquent ceci doit être l'un des premiers objectifs du traitement du diabétique obèse. Le traitement non pharmacologique comprend trois aspects de base :

- Un plan d'alimentation: il doit être personnalisé et adapté aux conditions de vie du patient. Le patient doit recevoir des instructions diététiques en fonction de son âge, sexe, état métabolique, activité physique, maladie intercurrente et situation économique. Il est recommandé de consommer des aliments riches en fibres solubles qui améliorent le contrôle glycémique et réduisent l'hyperinsulinémie et les taux de lipides.
- Faire du sport et adopter une activité physique régulière : marche, natation...
- Hygiène de vie: arrêter de fumer et diminuer l'alcool.

1.6.2 Traitement avec des antidiabétiques oraux

Si l'exercice et le régime ne sont pas suffisants pour contrôler le niveau de glucose, le médecin peut juger nécessaire l'utilisation des antidiabétiques. Il existe plusieurs classes de médicaments antidiabétiques oraux, parmi lesquelles se trouvent : [6]

1.6.2.1 Sulfonylurée ou sulfamides hypoglycémiantes

Leur principale fonction est d'augmenter la sécrétion d'insuline par le pancréas donc ils ne sont pas efficaces pour le diabète de type 1. Les sulfonylurées de dernière génération comme le glimépiride et le gliclazide ont aussi des effets favorables sur le système cardiovasculaire, la fonction endothéliale et les cellules Béta. Son principal effet secondaire est l'hypoglycémie qui est moins fréquente avec les médicaments de dernière génération et en particulier avec celle d'action prolongée qui sont administrés une fois par jour. Parfois ils peuvent produire des réactions cutanées, hématologiques ou hépatiques. Ce médicament est contre-indiqué pour les femmes enceintes ou allaitantes, ainsi que pour les patients souffrant d'insuffisance rénale.

1.6.2.2 Les méglitinides

De même que les sulfonylurées ils stimulent la sécrétion d'insuline. Leur principal effet secondaire est l'hypoglycémie mais elle se produit moins souvent que pour les sulfonylurées

Présentation générale du diabète

car les méglitinides ont une durée d'action plus courte. Ce médicament est contre-indiqué pour les patients souffrant d'insuffisance rénale ou hépatique.

1.6.2.3 Biguanides, Glucophage, Diabamyl, Metformine ou Stagid

Ce médicament réduit le niveau de glucose dans le sang en diminuant la production de glucose par le foie. Il peut aussi incrémenter la sensibilité des cellules musculaires à l'insuline et retarder l'absorption du glucose par l'intestin. Ce traitement est efficace pour les diabètes de type 1 et 2 car il ne dépend pas de l'insuline pour fonctionner. Il peut produire une intolérance gastro-intestinale. Il est contre-indiqué pour les patients souffrant d'insuffisance rénale ou hépatique et pour les alcooliques.

1.6.2.4 Les thiazolidinediones ou glitazones

Ce fut le premier agent oral pour le DM2 qui réussit à inverser réellement la résistance à l'insuline. Ceci en produisant des changements dans les cellules musculaires et adipeuses où réside la résistance à l'insuline. Ces changements surviennent après plusieurs semaines. Son principal effet secondaire est l'oedème (en particulier en combinaison avec l'insuline) et le gain de poids. La troglitazone fut retirée du marché à cause de sa toxicité hépatique et même si les autres glitazones, la rosiglitazone et la pioglitazone, n'ont pas eu ce problème, ce traitement est contre-indiqué pour les patients souffrant d'insuffisance cardiaque.

1.6.2.5 Les inhibiteurs des alpha-glucosidases

Ces médicaments diminuent l'absorption des sucres de l'intestin vers le sang (sucres absorbés au cours des repas) : ils diminuent donc la glycémie après les repas. Leurs principaux effets secondaires sont des problèmes digestifs, des ballonnements et des gaz intestinaux parfois abondants.

6.3 Insulinothérapie

Au début l'insuline ne pouvait s'obtenir qu'à partir des pancréas d'animaux. Ce qui provoquait des réactions immunologiques. En 1978, des chercheurs ont réussi à modifier une bactérie appelé *E coli* pour qu'elle produise de l'insuline humaine.

Dans le corps humain, l'insuline répond constamment aux niveaux de glucose dans le sang. Actuellement, il n'existe pas de dispositif capable de mesurer le glucose dans le sang et d'approvisionner l'insuline comme le fait naturellement le pancréas. Afin d'éviter plusieurs injections par jour, plusieurs types d'insuline furent développés qui dépendent essentiellement de leur rapidité de passage dans la circulation sanguine: [7]

Présentation générale du diabète

- **Les insulines ultra-rapides** : agissent entre 2 et 4h et sont injectées juste avant les repas. Elles permettent de faire entrer les sucres apportés par le repas dans les cellules. Ces insulines sont des insulines humaines légèrement modifiées afin de raccourcir la durée d'action et ainsi se rapprocher de l'action normale de l'insuline pendant un repas chez une personne non diabétique.
 - **Les insulines rapides** : agissent entre 4 et 6h et sont injectées une quinzaine de minutes avant un repas. Elles permettent également de faire entrer les sucres apportés par le repas dans les cellules. Ces insulines sont des insulines humaines non modifiées.
 - **Les insulines intermédiaires** : agissent entre 10 et 16h. Elles peuvent avoir des actions différentes : elles agissent soit de façon prolongée et équilibrée pendant toute leur durée d'action, soit de façon plus importante pendant les 6 premières heures que pendant les heures suivantes.
 - **Les insulines lentes** : agissent entre 20 et 24h. Ces insulines permettent de couvrir les besoins du métabolisme durant toute la journée. Ce sont des insulines qui ont été modifiées afin d'allonger leur durée d'action et/ou une libération prolongée pour s'approcher au maximum d'une sécrétion basale la plus constante possible.
- Il existe également des **mélanges** tout prêt d'insuline à la fois rapide (ou ultra-rapide) et intermédiaire (ou lente) pour ne faire qu'une injection au lieu de deux.

1.7 Complications

Le diabète sucré peut présenter des symptômes caractéristiques comme la soif, la polyurie, une vision floue, et la perte de poids. Dans ses formes les plus graves, une acidocétose ou un état hyperosmolaire non cétosique peut se développer et conduire à la stupeur, le coma et, en l'absence de traitement efficace, la mort.

Les effets à long terme du diabète sucré comprennent le développement progressif des complications spécifiques de la rétinopathie et potentiellement de cécité, la néphropathie qui peut entraîner une insuffisance rénale, et / ou de neuropathie avec des risques d'ulcères du pied, amputation, articulations de Charcot, et les caractéristiques de disfonctionnement d'autonomie. Les personnes diabétiques courent un risque accru de maladie cardiovasculaires, cérébrovasculaires et vasculaires périphériques. [8]

Présentation générale du diabète

1.8 Prévention

La prévention du diabète et ces complications implique adopter un ensemble d'actions afin d'éviter son apparition ou progression. Cette prévention peut être réalisée en trois niveaux : [6]

1.8.1 Prévention primaire

Elle a comme objectifs d'éviter la maladie. Dans la pratique c'est toute activité qui à lieu avant l'apparition de la maladie dans le but d'éviter son apparition. Deux stratégies d'intervention primaire sont proposées:

- Dans la population en général pour éviter et contrôler l'établissement du syndrome métabolique comme facteur de risque aussi bien du diabète que des maladies cardiovasculaires, plusieurs facteurs de risque sont potentiellement modifiable tels que l'obésité, le sédentarisme, la dyslipidémie, l'hypertension artérielle, le tabagisme et la nutrition.
- Dans la population qui a un haut risque d'apparition du diabète, les actions suivantes sont proposées afin d'éviter la maladie: une éducation sanitaire principalement à travers de brochures, magazines, bulletins, etc. Prévention et traitement de l'obésité en promouvant la consommation des régimes à faible teneur en matières grasses, de sucres raffinés et riche en fibres. Faire attention à l'indication des médicaments diabétogènes tels que les corticoïdes et enfin la stimulation de l'activité physique.

1.8.2 Prévention secondaire

Elle permet principalement d'éviter les complications, en mettant l'accent sur la détection précoce du diabète comme stratégie de prévention. A ce niveau les objectifs sont :

- Chercher la remise, si c'est possible.
- Prévenir l'apparition de complications aiguës et chroniques.
- Ralentir la progression de la maladie.
- Les actions sont basées sur le contrôle métabolique optimal du diabète.

1.8.3 Prévention tertiaire

Elle vise à éviter l'invalidité fonctionnelle et sociale et la réhabilitation des patients handicapés. Ses objectifs sont :

- Arrêter ou ralentir la progression des complications chroniques de la maladie.
- Prévenir l'invalidité du patient causé par des complications en phase terminale comme l'insuffisance rénale, la cécité, l'amputation, et ainsi de suite.

Présentation générale du diabète

- Prévenir la mortalité précoce.
- Ces actions nécessitent la participation de professionnels spécialisés dans les différentes complications du diabète.

1.9 Facteurs de risque de diabète de type 2

Dans les cas de diabète de type 2, soit le pancréas ne synthétise pas suffisamment d'insuline ou bien l'organisme n'utilise pas l'insuline correctement. Bien que la cause exacte du diabète de type 2 demeure mal comprise, il a été établi que ce trouble est plus fréquent chez les personnes présentant les caractéristiques suivantes :

- Age supérieur à 40 ans : Même si de plus en plus de jeunes sont atteints de diabète de type 2, l'âge reste un des principaux facteurs de risque. Ceci est dû au fait qu'avec l'âge, l'effet de l'insuline diminue, surtout chez les personnes en surpoids. Pour y remédier, le pancréas augmente sa production d'insuline. Lorsque l'organe a atteint ses limites de production, le diabète se manifeste.
- Obésité.
- Antécédents familiaux de diabète : les facteurs génétiques jouent également un rôle important dans le diabète de type 2. Par exemple, si un des parents est diabétique, les enfants ont un risque de 30 à 40 % de développer un diabète de type 2.
- Appartenance à une population à risque élevé (exemple : personnes d'origine autochtone, hispanique, sud-asiatique, asiatique ou africaine).
- Maladie cardiaque.
- Taux de cholestérol anormal.
- Hypertension artérielle : Souvent l'hypertension artérielle va de pair avec la surcharge pondérale et reste longtemps méconnue parce qu'elle ne présente pas de symptômes. Les personnes diabétiques souffrent trois fois plus souvent d'hypertension que les non-diabétiques. Si elle n'est pas traitée, l'hypertension augmente le risque d'infarctus du myocarde et d'accident vasculaire cérébral.
- Diabète de grossesse.
- Accouchement d'un bébé de plus de 4 kg.

Un bon nombre des facteurs énumérés (par exemple : taux de cholestérol anormal, hypertension artérielle, excès de poids) ont tendance à se manifester à la même période, ce qui crée un groupe de facteurs de risque auquel on donne le nom de « syndrome métabolique ». L'apparition de ces facteurs de risque a probablement une cause commune, soit une accumulation excessive de graisse abdominale. On croit qu'une

Présentation générale du diabète

diminution de cette réserve « entraine » de gras peut réduire la résistance à l'insuline et donc avoir un effet bénéfique sur un bon nombre des facteurs de risque énumérés.

1.10 Conclusion

Nous avons vu dans ce chapitre les types de diabète, les différents traitements et tests ainsi que les complications dû à cette maladie. Même s'il existe des méthodes de prévention qui permettent de réduire le risque d'avoir le diabète, parfois il est impossible de l'éviter comme pour le diabète de type 1. Dans ces cas là, la seule solution est de pouvoir le diagnostiquer très tôt et faire tout son possible pour combattre les complications.

Notre travail, dans le cadre de ce mémoire, consiste à implémenter un algorithme basé sur le Système immunitaire naturel pour la reconnaissance automatique du diabète.

1.1 Introduction

Il existe de nombreux problèmes complexes qui ne peuvent pas être résolus par un algorithme dans un temps polynomial. A partir de 1980, des méthodes nommées métaheuristiques ont commencé à apparaître pour résoudre au mieux les problèmes complexes. La création des métaheuristiques s'est inspirée de divers domaines tels que la biologie ou l'éthologie (étude du comportement des animaux). Parmi ces métaheuristiques nous pouvons citer par exemple: les algorithmes génétiques qui s'inspirent des théories de l'évolution, les algorithmes de colonie de fourmis qui s'inspirent du phénomène de suivi de piste chez les fourmis ou les systèmes immunitaires artificiels qui se basent sur le fonctionnement du système immunitaire naturel. Les systèmes immunitaires artificiels essaient de reproduire des caractéristiques intéressantes du système immunitaire naturel tel que la capacité d'adaptation, la mémorisation, la reconnaissance de formes et l'apprentissage.

1.2 SYSTEMES IMMUNITAIRES NATURELS

Le système immunitaire est responsable de la protection du corps contre les "agressions" d'organismes extérieurs qui présentent un danger, ces entités sont appelées antigènes. Le système immunitaire se compose de différents types de cellules qui se déplacent dans l'organisme et dont chaque catégorie a un rôle spécifique.

Pour être effectif, le système immunitaire doit être capable de distinguer entre le soi (toute molécule appartenant au corps) et le non-soi (les antigènes) donc il doit répondre uniquement aux entités étrangères. La discrimination du soi et non-soi est une caractéristique essentielle du système immunitaire car une réponse inappropriée contre les molécules du corps peut être fatale (les maladies auto-immunes).

1.3 Historique

L'immunologie est une science relativement nouvelle. Son origine est attribuée à Edward Jenner qui a découvert en 1796 (il y a eu plus de 2 siècles), que la vaccine ou cow-pox, induit une protection contre la variole humaine, une maladie souvent mortelle (Janeway Jr. & Travers, 1997). Ce chercheur l'a baptisée « vaccination hisprocess », expression qui décrit

Les systèmes immunitaires

l'inoculation des individus en bonne santé par des échantillons affaiblis ou atténués d'agents qui provoquent des maladies, en vue d'obtenir une protection contre ces maladies.

Lorsqu'Edward Jenner introduisit la vaccination, on ne savait encore rien sur l'agent ethnologique de l'immunologie. Au XIXe siècle, Robert Koch confirma que les maladies infectieuses sont causées par des microorganismes pathogènes, dont chacun est responsable d'une pathologie donnée.

Au début des années 1900, Jules Bordet et Karl Landsteiner portent la discussion sur la notion de « spécificité immunologique ». Il montre que le système immunitaire est capable de produire des anticorps spécifiques contre les produits chimiques synthétisés artificiellement, et qui donc n'ont jamais existé à l'état naturel.

Les années soixante sont en général considérées comme le début de l'époque moderne de l'immunologie. Rodney Porter et Gerald Edelman réussirent à élucider la structure des anticorps entre 1959 et 1961, et furent lauréats du prix Nobel de médecine en 1972.

Vers 1960, la communauté scientifique découvrait, grâce aux travaux de Jacques Miller, d'autres caractéristiques fondamentales des cellules immunitaires.

En 1989, Charles Janeway propose un modèle selon lequel, ce serait l'immunité innée qui serait la véritable gardienne des clefs du déclenchement d'une réponse immunitaire.[9]

1.4 .Éléments du système immunitaire

Le système immunitaire est un système complexe composé de plusieurs éléments. Il est possible de classer ses éléments en deux classes : les organes, les cellules.

4.1.1 Les organes du système immunitaire :

Les organes du système immunitaire sont nombreux et peuvent être divisé en deux classes:

➤ *Les organes lymphoïdes primaires* : ils constituent le site de développement et de maturation des cellules : [10]

• *Moelle osseuse* : Lieu de développement des cellules progénitrice lymphoïde qui après leur croissance se diviseront pour former les cellules précurseurs des lymphocytes B et T. Les cellules précurseurs des lymphocytes T quittent la moelle osseuse pour le thymus mais celles des lymphocytes B restent dans la moelle osseuse où elles continuent leur maturation.

Les systèmes immunitaires

- *Thymus* : Dans le bas du cou, constitue le site de maturation des lymphocytes T, Mais la plupart d'entre eux meurent sur place, seul 5 % quittent le thymus.

- *Vaisseaux lymphatiques* : Transportent la lymphe, les vaisseaux lymphatiques sont situés dans tout le corps.

➤ *Les organes lymphoïdes secondaires* : ils constituent le lieu d'interaction entre l'antigène et le lymphocyte. [10]

- Les *amygdales*.

- Les *ganglions lymphatiques*.

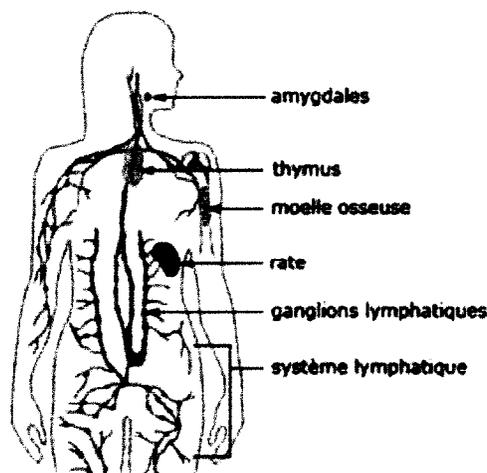


Figure2 Organes du système immunitaire

1.4.2 Les cellules du système immunitaire :

Les globules blancs aussi appelées leucocytes qui se trouvent dans le sang, la lymphe et les organes lymphoïdes sont les cellules responsables de protéger le corps des organismes étrangers. Il existe différents types de leucocytes dont les lymphocytes B et T.

➤ **Les lymphocytes B** : la fonction principale des lymphocytes B est de produire des protéines appelées anticorps qui se fixent sur les protéines étrangères des antigènes. La partie de l'anticorps responsable de la reconnaissance de l'antigène est appelée paratope. Le paratope se lie à une partie spécifique de l'antigène appelée épitope. La liaison entre un paratope et un épitope est d'autant plus forte que leurs formes sont complémentaires. La force de cette liaison est appelée affinité. Les cellules B qui vont mieux reconnaître l'antigène vont

Les systèmes immunitaires

proliférer en se clonant. Les lymphocytes B ont également la capacité de se comporter en cellule présentant le corps étranger (cellule mémoire).

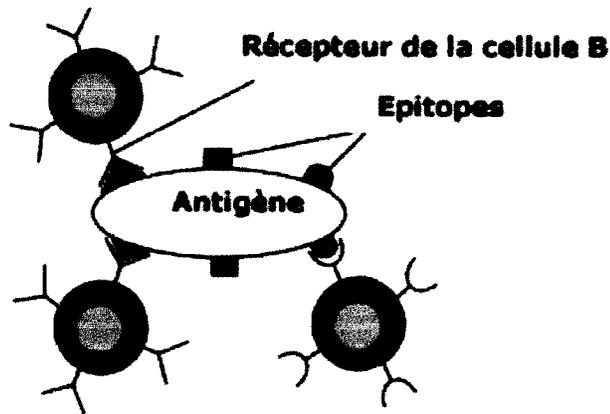


Figure 3 Structure d'un antigène avec ses épitopes.

➤ Les **lymphocytes T** : Il existe plusieurs types de lymphocytes T : [11]

- *Les lymphocytes T cytotoxiques* qui, lorsqu'elles sont activées, détruisent directement les cellules infectées par des virus et les cellules tumorales.

- *Les lymphocytes T auxiliaires*, qui se chargent de la coordination d'autres aspects de la réponse immunitaire c'est-à-dire elles déclenchent l'expansion clonale et stimulent ou suppriment la formation d'anticorps.

- *Les lymphocytes T suppresseurs* sont des régulateurs de l'immunité et luttent contre les réactions auto-immunes.



Figure 4 Un lymphocyte, principale composante du système immunitaire.

1.5 Architecture du système immunitaire

La défense de l'organisme contre le milieu extérieur comporte une immunité dite innée ou naturelle, c'est-à-dire existante en absence de tout contact avec un antigène, et une immunité dite adaptative ou acquise, c'est-à-dire apparaissant après contact de l'organisme avec des molécules étrangères qui sont des antigènes.

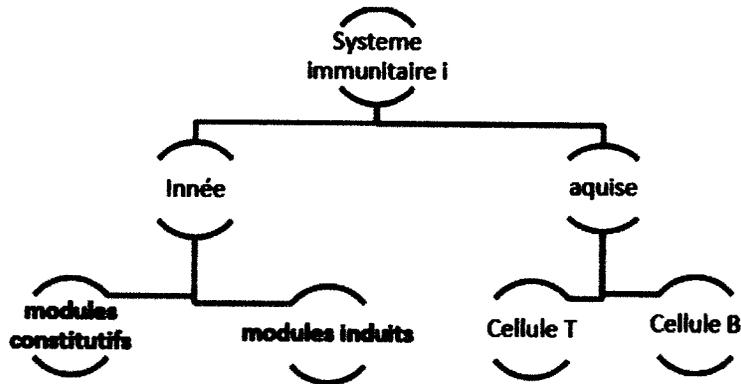


Figure5 Architecture du système Immunitaire

1.5.1L'immunité innée (naturelle non spécifique ou naïve)

Elle est présente dès la naissance, elle est constituée de la barrière anatomique (la peau et le système respiratoire), les conditions physiologiques telles que le pH, la température ainsi que les cellules tueuses naturelles NK (natural Killer) et les cellules phagocytaires (Figure 6). L'immunité innée permet de : [12] [13]

- Agir sans tenir compte du type de l'entité qu'elle combat, elle possède une perception intrinsèque de ce qui est microbien, ou hostile.
- Effectuer une distinction globale du soi et du non-soi.
- Agir d'une manière immédiate est non adaptative.
- Initialiser et de régulariser la réponse immunitaire adaptative. Ce système ne peut pas assurer la protection complète du corps.

Les systèmes immunitaires

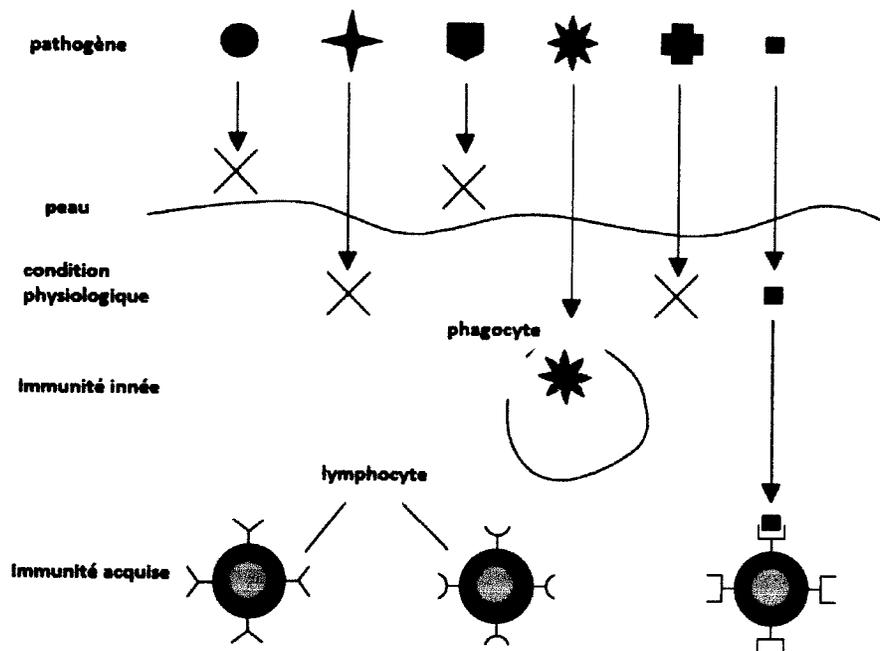


Figure 6 La structure multicouche du système immunitaire.

1.5.2L'immunité acquise (spécifique)

Elle se base sur la reconnaissance de l'entité à combattre et la mémorisation de cet événement. Lors de la réponse immunitaire spécifique, interviennent les lymphocytes, un type de globules blancs, qui s'adaptent rapidement afin de combattre les antigènes qui se trouve dans le corps. Les lymphocytes ont une plus grande capacité de mémoire que les phagocytes responsables de la réponse immunitaire innée.

Quand un lymphocyte entre dans la circulation sanguine pour la première fois, il porte un récepteur d'antigène spécifique, cette spécificité a été déterminée par un mécanisme spécial de réorganisation de gènes capables de générer des millions de récepteurs différents qui intervient lors du développement des lymphocytes dans la moelle osseuse et le thymus. C'est ainsi que chaque lymphocyte porte un récepteur différent des autres : [12] [13]

➤ Le système immunitaire spécifique protège le corps des organismes étrangers qui n'ont pas été détectés par le système immunitaire inné.

➤ Le système immunitaire spécifique est adressé à des intrus spécifiques.

➤ Le système immunitaire spécifique possède une mémoire immunologique qui permet aux cellules de se souvenir des intrus déjà rencontrés. Ainsi lors de la première intrusion d'un antigène la réponse immunologique est lente et difficile, cependant elle permet de mémoriser l'antigène grâce à ses marqueurs (épitopes). Donc si le même antigène pénètre une seconde

fois le corps, la réponse sera plus rapide et bien spécifique avec une production d'anticorps particuliers pour cet antigène.

1.6 Propriétés du système immunitaire :

Avec des propriétés très importantes, il devient une référence précieuse et une source d'inspiration pour les nouvelles branches de l'informatique. Beaucoup de travaux de recherche ont vu le jour en s'inspirant du fonctionnement de ce dernier.

Voici quelques unes des propriétés les plus importantes du système immunitaire :

Multicouche : Le système immunitaire possède une architecture multicouche qui consiste en deux sous-systèmes inter-liés qui sont : le système immunitaire inné et le système immunitaire adaptatif. Ces deux systèmes combinent leurs tâches et responsabilités pour assurer la protection et la sécurité globale.

Unicité : Chaque élément dans le système immunitaire assume des responsabilités particulières.

Autonomie : Le système immunitaire humain ne possède aucun contrôle central ou un gestionnaire particulier. Il possède une autonomie globale dans la détection et l'élimination des intrus.

Distribution : Les cellules immunitaires et les molécules sont distribuées dans le corps humain pour assurer sa protection. Il n'existe pas un point de contrôle centralisé.

Parallélisme : Le système immunitaire est capable de produire plusieurs réponses immunitaires en même temps à des endroits dispersés.

Tolérance au soi : Le système immunitaire humain peut différencier entre les cellules de soi et les cellules de non soi.

Apprentissage : Le système immunitaire augmente la capacité d'identification des anticorps à un antigène sélectif (les réponses primaire et secondaire). Il apprend continuellement les structures de pathogènes.

Adaptabilité : Le système immunitaire humain permet la production des cellules de plus en plus spécialisées pour l'identification des antigènes. Cela est garanti par la théorie de la sélection clonale suivie par le mécanisme de l'hyper mutation somatique.

Dynamique : Le système immunitaire change continuellement par la création de nouvelles cellules et molécules, l'élimination des cellules vieilles ou endommagées. Un bon exemple de la dynamique du système immunitaire est la théorie du réseau idiotypique.

Les systèmes immunitaires

Mémorisation : Après une réponse immunitaire à un antigène donné, un ensemble de cellules constituent l'ensemble des cellules mémoires qui seront dotées pour une durée de vie longue afin de fournir des réponses immunitaires plus rapides et plus puissantes aux rencontres suivantes d'un même antigène.

Coopération: Les cellules immunitaires coopèrent leurs capacités pour assurer une meilleure détection et également une protection puissante par exemple les cellules T d'aide, les molécules MHC, etc.

Détection: Le système immunitaire est capable d'identifier et de détecter les intrus dans le corps sans aucune connaissance antérieure de la structure de ces intrus.

1.7 Fonctionnement du système immunitaire

Le corps est protégé par un ensemble de cellules et molécules qui coopèrent et dont la cible est l'antigène, une molécule étrangère provenant d'une bactérie, un virus ou tout autre envahisseur.

On peut voir dans la figure 7 un schéma simplifié de la réponse immunitaire : [11]

I. Des cellules spécialisées appelées cellules présentatrices d'antigène (APC), tel que les macrophages, parcourent le corps, en ingérant et digérant des antigènes qu'ils trouvent et les fragmentant en peptides antigéniques.

II. Des fragments de ces peptides sont joints au complexe majeur d'histocompatibilité (MHC : major histocompatibility complex) et sont affichées sur la surface de la cellule.

III. Les cellules T, possèdent des récepteurs qui permettent à chacun d'entre eux de reconnaître une combinaison différente de peptide-MHC.

IV. Les cellules T activés par la reconnaissance, se divisent et sécrètent des lymphokines ou des signaux chimiques, qui mobilisent d'autres composantes du système immunitaire.

V. Les lymphocytes B, qui ont également des récepteurs d'une seule spécificité sur leur surface, répondent à ces signaux. Contrairement aux récepteurs des cellules T, toutefois, celles des cellules B peuvent reconnaître les pièces d'antigènes libres, sans les molécules MHC.

VI. Lorsqu'elles sont activées, les cellules B se divisent et se différencient en cellules plasma qui sécrètent des anticorps, qui sont la forme soluble de leurs récepteurs.

VII. En se liant à des antigènes qu'ils trouvent, les anticorps peuvent les neutraliser ou précipiter leur destruction.

Les systèmes immunitaires

VIII. Certains lymphocytes T et B deviennent des cellules mémoires qui persistent dans la circulation, elles permettent une réponse immunitaire plus rapide du même antigène dans de prochaine exposition. Parce que les gènes d'anticorps dans les cellules B sont souvent victimes de mutation, la réponse immunitaire s'améliore après immunisations répétées, ce phénomène est appelé la maturation d'affinité. (Figure Réponses immunitaire primaire et secondaire).

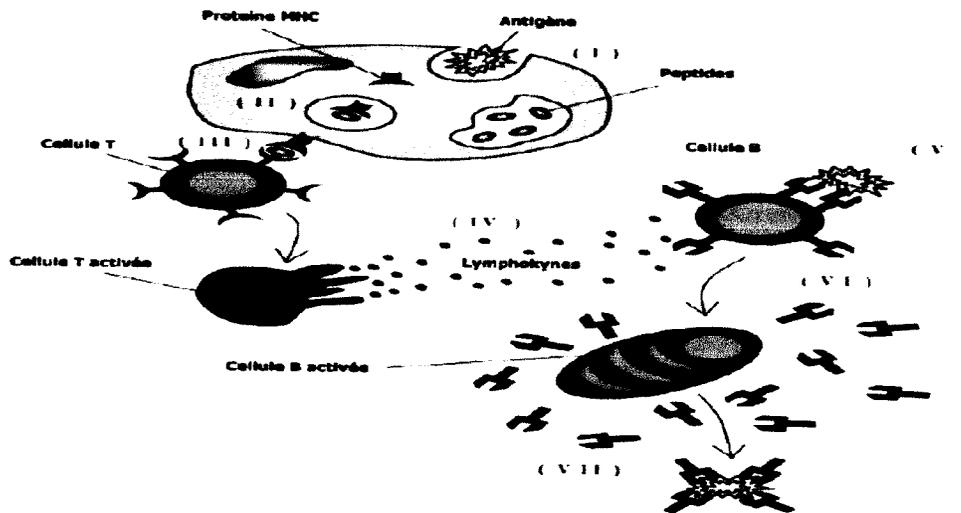


Figure 7 Déroulement de la réponse immunitaire.

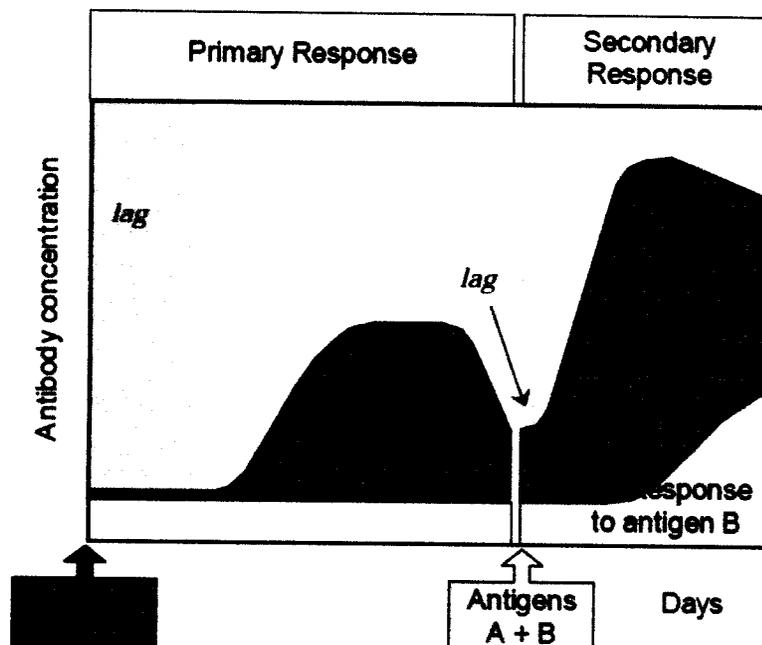


Figure 8 Réponses Immunitaire primaire et secondaire.

1.8 Détection du soi et non-soi

Pour que le système immunitaire fonctionne correctement, il faut qu'il soit capable de faire la différence entre les cellules du soi et les cellules nuisibles (non-soi) qui sont a priori impossibles à distinguer, cette capacité est appelée la tolérance de soi. Si le système immunitaire n'est pas capable d'effectuer cette distinction, alors une réponse immunitaire sera déclenchée contre les antigènes du soi, provoquant des maladies auto-immunes. Pour palier à ce genre de problème, le système immunitaire soumet les lymphocytes à deux types de sélection :

1.8.1 Sélection positive

Lors de la création des lymphocytes un mécanisme spécial de réorganisation de gènes génère des millions de récepteurs différents (TCR) capable de reconnaître des centaines de molécules du CMH présents dans la population humaine. Cependant ce n'est qu'un nombre restreint de lymphocytes T qui possèdent des récepteurs pouvant interagir avec les molécules du CMH d'un individu et sont capables de répondre aux antigènes présentés par ces mêmes molécules du CMH. Donc la Sélection positive permet à cette population de lymphocytes T d'arriver à maturation laissant la plupart des autres cellules mourir par apoptose.

1.8.2 Sélection négative

Après la production des cellules T naïves dans la moelle osseuse, elles migrent vers le thymus où elles subiront un processus appelé sélection négative. Ce processus permettra d'éliminer les cellules T capable de reconnaître un antigène du soi et ceci grâce à la barrière qui évite que tout antigène du non-soi accède à l'environnement du thymus. [14]

Malgré tout, la sélection négative n'est pas parfaite et certaines cellules T arrivent à passer. Ces cellules pourraient causer des lésions, mais pour qu'un lymphocyte T soit activé, il a besoin de recevoir d'autres signaux que ceux de la liaison avec les complexes peptide-CMH2, ces cellules ne réussissent pas à s'activer. [11]

La tolérance du soi pour les cellules T n'est pas suffisante pour assurer la protection de l'organisme et c'est pour ça que la sélection négative est aussi appliquée sur les cellules B dans la moelle osseuse, quand les cellules B immatures identifient les cellules du soi, elles seront éliminées. [15]

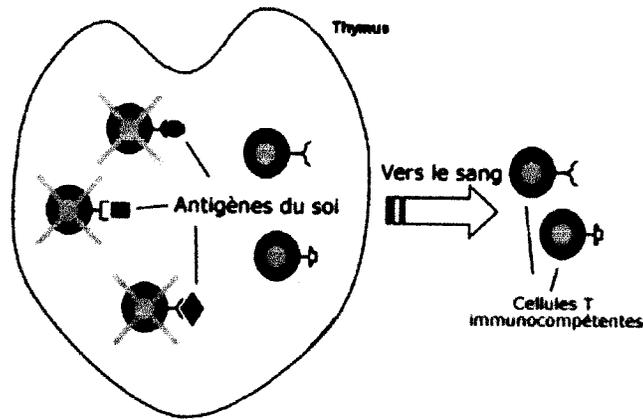


Figure 9 Sélection négative dans le thymus

1.9 Théorie de la sélection clonale

Quand le corps est exposé à un antigène, les lymphocytes B produisent les anticorps. Chaque lymphocyte secrète un seul type d'anticorps, qui est relativement spécifique à l'antigène. Quand l'anticorps s'associe à l'antigène à l'aide des récepteurs (épitopes-paratopes) et reçoit un signal des cellules telles que les T auxiliaires, les cellules B sont stimulées à proliférer (division cellulaire) et murir vers des cellules terminales (non divisibles) sécrétrices d'anticorps appelés **cellules plasma**. Le processus de division cellulaire (mitose) génère des clones c'est-à-dire une ou un ensemble de cellules qui sont la progéniture d'une seule cellule. D'autre part, les cellules T jouent un rôle central dans la régulation de la réponse des lymphocytes B et sont par excellence dans les réponses immunitaires à médiation cellulaire. [15]

En plus de la prolifération, les lymphocytes subissent un second processus qui permettra de sélectionner parmi les nouvelles cellules celles présentant une grande affinité afin d'en faire des cellules B mémoires. Les cellules mémoires circulent dans le sang, la lymphe et les tissus, et lorsqu'elles sont exposées à un stimulus antigénique pour la seconde fois elles commencent à se différencier en lymphocytes capables de produire des anticorps de haute affinité, présélectionnées pour l'antigène spécifique qui a stimulé la réponse primaire. [15] Les principales caractéristiques de la théorie de la sélection clonale sont les suivantes :

➤ La prolifération et différenciation des cellules après avoir été stimulé par des antigènes ;

Les systèmes immunitaires

- La création de nouveaux changements génétiques aléatoires, représentant des profils d'anticorps différents, par une forme de mutation somatique accélérée (ce processus est appelé la maturation d'affinité) ;
- L'élimination des nouveaux lymphocytes qui portent des récepteurs de faible affinité antigénique.

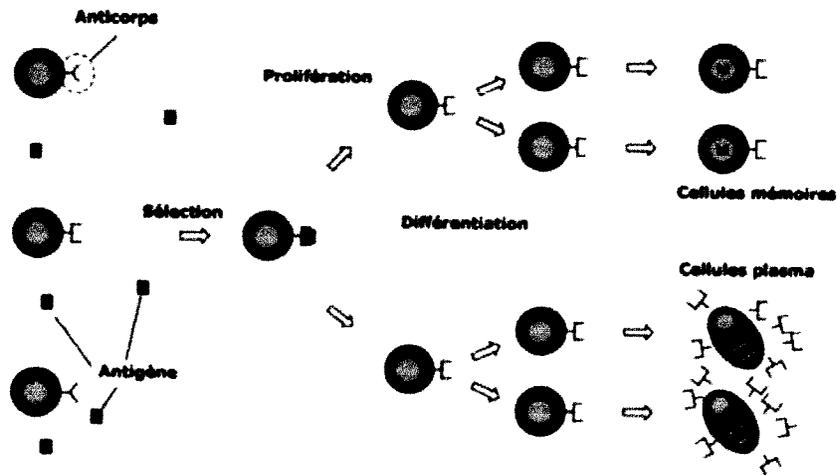


Figure 10 Processus de sélection clonale.

1.10 Déroulement d'une réponse immunitaire :

Une importante collection de cellules et de molécules protège notre corps. Ces dernières collaborent contre n'importe quelle molécule étrangère comme les bactéries ou d'autres envahisseurs :

- ✓ Quand un intrus envahit le corps, les cellules de présentation antigénique comme les macrophages procèdent à l'ingestion et la digestion de l'antigène rencontré pour le présenter comme des fragments de peptides antigéniques, ces peptides seront liés avec les molécules MHC pour permettre leurs liaisons avec les cellules T qui ont la capacité de reconnaître la combinaison de peptide / MHC.
- ✓ Les cellules T activées par cette identification produisent et sécrètent des lymphokines ou des signaux chimiques pour mobiliser d'autres composants du système immunitaire.
- ✓ Les cellules B qui ont aussi des molécules de récepteur complémentaires répondent à ces signaux. À la différence des récepteurs de cellules T, ceux de cellules B peuvent reconnaître les parties d'antigènes libres sans les molécules MHC. Après cette

Les systèmes immunitaires

activation, les cellules B prolifèrent et se différencient et sécrètent des protéines d'anticorps.

- ✓ La liaison entre les anticorps et les antigènes disponibles mènent à la destruction et la suppression des antigènes.
- ✓ Certaines cellules B et T deviennent des cellules mémoires qui ont une durée de vie illimitée, en permettant l'élimination rapide de l'antigène s'il se présente une autre fois dans l'avenir.

Système Immunitaire Artificiel

2.1 Introduction

L'intérêt pour les SIAs est apparu au milieu des années 80 avec la publication de l'article de Farmer, Packard et Perelson "The immune system, adaptation and machine learning" [24] et celui de Bersini et de Varela sur les réseaux immunitaire en 1990. Cependant, c'est seulement au milieu des années 90 que les SIAs sont devenus un domaine de recherche. En 1994, Dasgupta a entrepris des études étendues sur des algorithmes de sélection négatives. Hunt et Cooke ont commencé des travaux sur les modèles de réseau immunitaire en 1995, Timmis et Neal ont continué ce travail et ont apporté quelques améliorations. Le travail de De Castro, Von Zuben, Nicosia et de Cutello sur la sélection clonal est devenu notable en 2002. Le premier livre sur les systèmes immunitaires artificiels a été édité par Dasgupta en 1999. Maintenant, de nouvelles idées telles que la théorie de danger et les algorithmes inspirés par le système immunitaire inné, sont explorés également. Bien qu'il subsiste un certain doute qu'ils offrent quelque chose au delà des algorithmes existants des SIAs.

2.2 Définition

Dasgupta a défini le système immunitaire artificiel comme suit : « Le système immunitaire artificiel est la composition de méthodologies intelligentes inspirées par le système immunitaire naturel afin de résoudre des problèmes du monde réel » [4].

De leur côté, Timmis et De Castro ont donné la définition suivante : « Les systèmes immunitaires artificiels sont des systèmes adaptatifs inspirés par des théories immunologiques et des observations de fonctions immunitaires, des principes et des modèles, qui seront appliqués à la résolution des problèmes ». [6]

Les systèmes immunitaires artificiels sont une nouvelle branche de l'intelligence artificielle. Destinés à résoudre des problèmes divers, ils s'inspirent des propriétés et concepts remarquables du système immunitaire biologique.

Les systèmes immunitaires

Les systèmes immunitaires artificiels sont une implémentation mathématique ou informatique du fonctionnement du système immunitaire naturel. Cette implémentation reprend les plus grandes lignes de son fonctionnement. Cependant il reste presque impossible de modéliser le comportement complet des systèmes biologiques. [16]

2.3 Les algorithmes immunitaires:

Il existe différents algorithmes des théories immunitaires selon le contexte et le problème à résoudre.

2.3.1 La sélection clonale

La théorie de la sélection clonale a été utilisée comme source d'inspiration pour le développement des SIAs qui effectuent des tâches d'optimisation et de reconnaissance de formes. En particulier, l'inspiration a été prise du processus de maturation d'affinité des cellules B et de son mécanisme d'hypermutation. Ces SIA font souvent appel à l'idée de cellules mémoires afin de conserver les bonnes solutions du problème à résoudre. Dans leur livre, Castro et Timmis mettent en évidence deux aspects importants de la maturation d'affinité dans les cellules B qui peuvent être exploitées à partir du point de vue informatique:

- Le premier est que la prolifération des cellules B est proportionnelle à l'affinité de l'antigène auquel elles se sont liées, donc les cellules avec le plus grand taux d'affinité, sont celles qui produisent le plus de clones.
- En second lieu, les mutations subies par les anticorps d'une cellule-B sont inversement proportionnelles à l'affinité de l'antigène auquel ils se sont liés, donc plus l'affinité est grande, plus le taux de mutation est petit.

Utilisant ces deux caractéristiques, de Castro et Von Zuben ont développé l'un des algorithmes de SIA inspiré de la sélection clonale les plus populaires et largement utilisés appelé CLONALG, qui a été utilisé pour effectuer les tâches de filtrage et d'optimisation multimodale. [17]

Entrée : S = ensemble d'antigènes.

Sortie : M = ensemble de détecteurs de mémoire capable de classer de nouveaux modèles.

Les systèmes immunitaires

Début:

Créer un ensemble aléatoire d'anticorps, A . Pour chaque élément de S faire:

1. Déterminer l'affinité avec chaque anticorps de l'ensemble A .
2. Créer un sous-ensemble K des anticorps avec le plus grand taux d'affinité.
3. Générer des clones des anticorps du sous-ensemble K . Le nombre de clones pour un anticorps est proportionnel à son affinité
4. Appliquer le processus de mutation sur les clones pour augmenter leur degré de correspondance avec l'antigène.
5. Exposer les clones de nouveau à l'antigène et recalculer leurs affinités.
6. Les meilleurs clones seront placés dans l'ensemble M .
7. Remplacer les n anticorps de plus faible affinité dans A par de nouveaux anticorps générés aléatoirement.

Fin

Bien que performant le CLONALG présente certains inconvénients, par exemple il ne permet pas de capitaliser les informations générées par chaque population de clone, car dès qu'une cellule mémoire est sélectionnée le reste des cellules mutées seront éliminées, alors que cette population pourrait contenir un certain nombre de cellules de haute affinité. En préservant une grande partie de la population de cellules mature, l'algorithme pourrait construire à partir d'une base solide de liens de haute affinité et devrait théoriquement aboutir à une solution optimale en moins de générations. Toutefois, ceci introduit le risque de convergence vers un minimum local. Ceci pourrait être évité en générant de façon aléatoire de nouveaux anticorps et les ajouter à la population. En 2003, White et Garrett dans leur article "Improved Pattern Recognition with Artificial Clonal Selection" utilisent l'algorithme CLONALG pour le problème de reconnaissance de formes et après avoir examiné la performance de la technique sur la classification des caractères non vue, ils ont proposé une version enrichie du CLONALG appelé CLONCLAS (CLONal selection algorithm for CLASsification). Le principe est d'allouer une classe à chaque anticorps de la population, ce qui lui permet d'effectuer une classification en attribuant sa classe à un antigène. Quand un nouvel antigène est exposé à la population, il est alloué à la classe de l'anticorps avec la plus grande affinité pour cet antigène.

Entrée :

S = ensemble d'antigènes.

Sortie : M = ensemble de détecteurs de mémoire capable de classer de nouveaux modèles.

Début

1. Générer aléatoirement une population initiale d'anticorps Ab . Il est composé de deux sous-ensembles Ab_m (population de mémoire) et Ab_r (population réservoir).

Les systèmes immunitaires

2. Sélectionnez un antigène A_{gi} de la population S .
 3. Pour G générations faire
 - Effectuer les étapes 1-6 de CLONALG
 - Remplacez les anticorps dans Abr avec le même nombre d'anticorps de la population triés des clones qui ont subi une mutation.
 - Supprimer les anticorps avec une faible affinité dans la population Abr et les remplacer avec de nouveaux membres générés de façon aléatoire.
- Fin
4. Revenir à l'étape 2 jusqu'à ce que tous les antigènes aient été présentés.
- Fin

2.3.2 La sélection négative

Les algorithmes de la sélection négative sont inspirés par le principal mécanisme dans le thymus qui produit un ensemble de cellules T matures capables de se lier seulement aux antigènes du non-soi. Le premier algorithme de la sélection négative a été proposé par Forrest en 1994 pour détecter la manipulation de données causée par un virus dans un système informatique. Le point de départ de cet algorithme est de produire un ensemble de chaînes de soi, S , qui définissent l'état normal du système. La tâche est alors de générer un ensemble de détecteurs, D , qui ne se lient (reconnaissent) que le complément de S . Ces détecteurs peuvent ensuite être appliqués à de nouvelles données afin de les classer comme étant soi ou non-soi, ou dans le cas de l'oeuvre originale de Forrest, détecter si les données ont été manipulées. L'algorithme de Forrest produit l'ensemble de détecteurs via le processus suivant:

Entrée : S = ensemble d'éléments du soi.

Sortie : D = ensemble de détecteurs généré.

Début

Répéter jusqu'à ce que les critères d'arrêt aient été atteints:

1. Générer aléatoirement des détecteurs potentiels et les placer dans un ensemble P .
2. Déterminer l'affinité de chaque membre de P avec chaque membre de l'ensemble S .
3. Si un élément de S est reconnu par un détecteur de P selon un seuil de reconnaissance T :

- ✓ alors le détecteur est rejeté,
- ✓ sinon il est ajouté à l'ensemble de détecteurs disponibles D .

fin

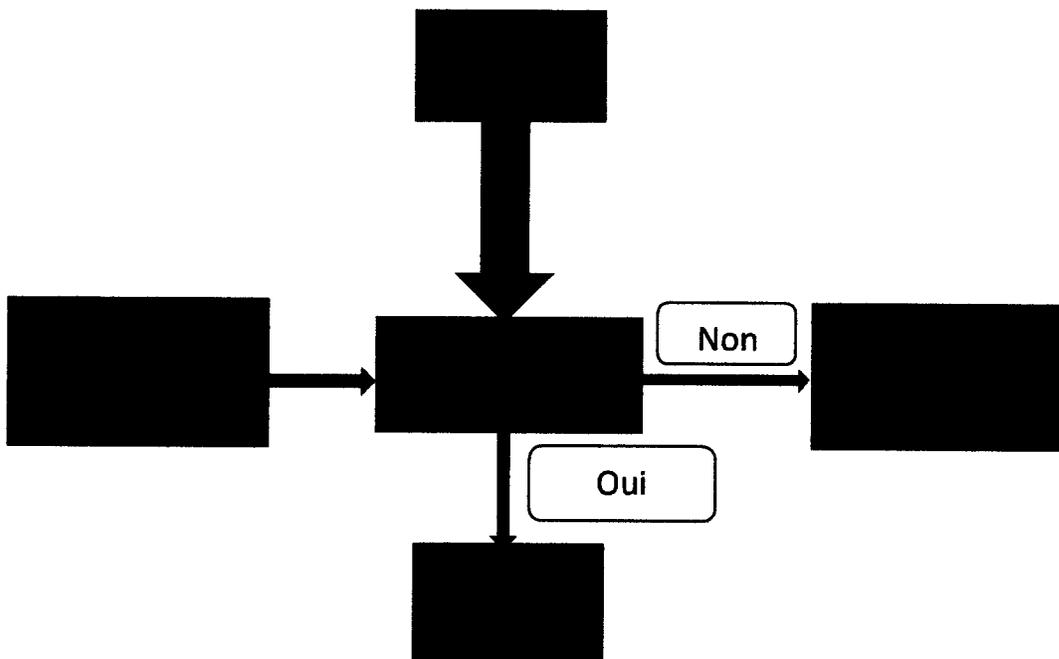


Figure11: Génération de l'ensemble de détecteurs.

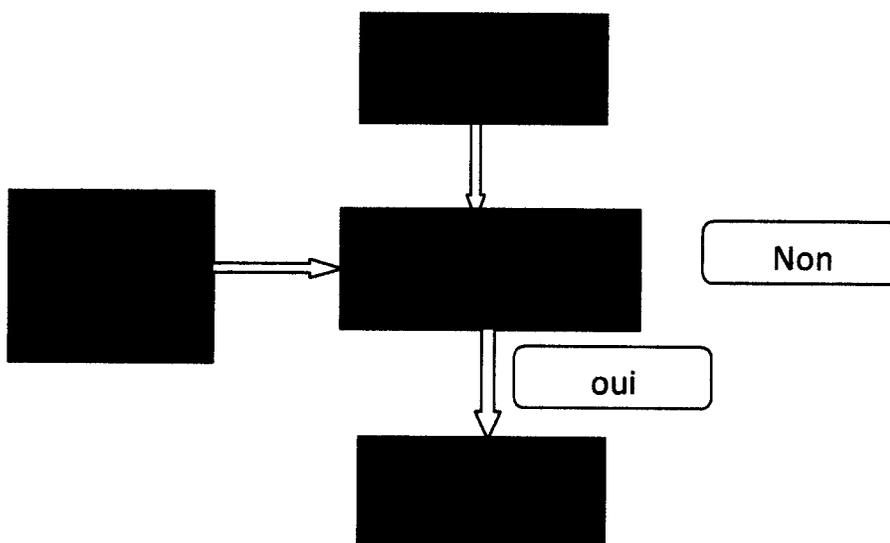


Figure12 : Surveillance d'éléments du non-soi.

2.3.3 Les réseaux immunitaires (ou idiotypiques)

La théorie du réseau immunitaire a suggéré un système immunitaire avec un comportement dynamique même en absence d'un antigène de non soi. Il existe plusieurs modèles du réseau immunitaire dans cette section nous nous pencherons sur le modèle nommé aiNet « artificial immune NETWORK » qui a été proposé par De Castro & Von Zuben. [18] [27] Le modèle aiNet sera composé d'un ensemble d'anticorps, reliés entre eux par des liens avec leurs forces de connexion associées. Les anticorps aiNet sont censés représenter les images du réseau interne des agents pathogènes contenues dans l'environnement auquel ils sont exposés. Les connexions entre les anticorps détermineront leurs interrelations, fournissant un degré de similitude entre eux: plus les anticorps sont proches, plus ils sont similaires.

L'algorithme de base est le suivant: [17]

Entrée : S = ensemble d'antigène à reconnaître, nt seuil d'affinité réseau, ct seuil piscine clonale, h le nombre de clones avec la plus grande affinité, a le nombre de nouveaux anticorps à introduire.

Sortie : N = ensemble de détecteurs mémoire capable de classer les modèles non vus.

Début

1. Créer un premier ensemble aléatoire d'anticorps, N .

2. Répéter jusqu'à ce qu'un des critères d'arrêt soit atteint

Pour chaque individu de S faire

- ✓ Déterminer l'affinité avec chaque anticorps en N .
- ✓ Générer des clones des anticorps avec la plus haute affinité.
- ✓ Muter les attributs de ces clones, Sélectionner quelques clones de plus haute affinité pour constituer l'ensemble mémoire, C ;
- ✓ Éliminez tous les éléments de C dont l'affinité avec l'antigène est inférieur à un seuil prédéfini ct ;
- ✓ Déterminer l'affinité entre tous les anticorps en C et éliminer les anticorps dont l'affinité avec les autres est inférieur au seuil ct ;

Incorporer les clones restants de C dans N ;

✓ Fin

3. Déterminer l'affinité entre chaque paire d'anticorps dans N et éliminer tous les anticorps dont l'affinité est inférieur au seuil nt ;

4. Introduire un nombre a d'anticorps générés aléatoirement et les placer dans N ;

Fin

Fin

2.4 AIRS: Artificial Immune Recognition System

2.4.1 Introduction

AIRS (système de reconnaissance immunitaire artificiel) est inspiré d'algorithme d'apprentissage supervisé.

Notre motivation pour cette étude vient du fait qu'il y a un manque important de recherches qui ont exploré l'utilisation de la métaphore du système immunitaire pour l'apprentissage supervisé. En effet, le seul travail identifié est celui de Carter [7].

Toutefois, il a été noté que dans la communauté AIS, il y a eu un certain nombre d'enquêtes sur l'exploitation des mécanismes immunitaires à l'apprentissage non supervisé.

Dans son étude, Castro [14] a examiné le rôle du processus de sélection clonale au sein du système immunitaire. Il a développé un apprentissage non supervisé connu sous le nom CLONALG (cité précédemment).

Nous décrivons brièvement les aspects les plus pertinents de l'immunologie qui nous ont servi de source d'inspiration pour ce travail.

Tout au long de la vie d'une personne, le corps est exposé à une grande variété de pathogènes (potentiellement dangereux). Le système immunitaire contient des lymphocytes connus sous le nom de cellules B et T ; chacune de ces cellules a un type unique de récepteur moléculaire (localisation dans un espace de forme). Les récepteurs dans cet espace de forme, permettent la liaison des Antigènes avec une plus grande affinité entre le récepteur et l'antigène qui indique une forte liaison.

Dans AIRS, l'idée de liaison antigène / anticorps est connue et employée comme la présentation antigénique. Lors du traitement des algorithmes d'apprentissage, elle est utilisée pour mettre en oeuvre l'idée de l'adéquation entre les données d'apprentissage (antigènes) et les solutions possibles (cellules B). Dans leur travail, Timmis, J. and M.Neal [11] ont employé l'idée de « artificial recognition ball » (ARB), qui a été inspirée par le travail de Farmer [12] décrivant l'interaction antigénique au sein d'un réseau immunitaire.

Autrement dit, un ARB peut être considéré comme représentant un certain nombre de cellules B identiques ; c'est un mécanisme utilisé pour réduire les redondances. Une fois l'affinité entre une cellule B et un antigène déterminé, la cellule B concernée se transforme en une cellule plasma et va être clonée. Pendant le processus d'expansion clonale, les cellules B subissent une prolifération rapide (clonage) en proportion de la façon dont elles correspondent à l'antigène. Cette réponse est l'antigène spécifique.

Les systèmes immunitaires

Ces clones passent par la maturation d'affinité, où certains engagent une hyper mutation somatique (la mutation est ici inversement proportionnelle à l'affinité antigénique) et finiront par passer par un processus de sélection par lequel une cellule donnée peut devenir une cellule mémoire. Ces cellules mémoire sont conservées pour permettre une réponse plus rapide au même antigène, ou à un antigène similaire. Dans l'algorithme AIRS, l'idée de l'expansion clonale et la maturation d'affinité, sont employées pour encourager la production de cellules de mémoire potentiel.

Ces cellules mémoires sont ensuite utilisées pour la classification. S'inspirant des travaux de Timmis [11], AIRS utilise l'idée de niveau de stimulation pour un ARB, qui, encore une fois, a été calculé à partir des équations d'un réseau immunitaire. Bien que AIRS a été inspiré par ce travail sur les réseaux immunitaires, le développement du classificateur conduit à l'abandon des principes du réseau.

Toutefois, AIRS ne prend pas en compte le principe de la mutation d'affinité proportionnelle. Lorsque les ARB ont été créés, ils ont été soumis à un processus de mutation aléatoire avec une certaine probabilité et ont ensuite été intégrés à l'ensemble des cellules de mémoire lorsque leur affinité satisfait à certains critères. Dans le système AIRS, les ARB survivants reposent sur l'idée d'un système de ressources limitées.

Un nombre prédéfini de ressources existent, pour lesquels les ARB participent en fonction de leur niveau de stimulation: plus la valeur de stimulation d'un ARB est haute plus le nombre de ressources allouées est grand. Les ARB qui ne peuvent pas résister à la concurrence pour les ressources, seront retirés du système. Le terme métadynamique du système immunitaire se réfère à l'évolution constante des cellules B à travers la prolifération cellulaire et la mort. Cela a été présenté dans AIRS à la production continue et à l'élimination des ARB de la population. Ce tableau résume la correspondance entre le système immunitaire et AIRS :

Système immunitaire	AIRS
Anticorps	Vecteur de caractéristique
Expansion clonale	Reproduction des ARB qui correspondent le mieux avec des antigènes
Antigènes	Donnée d'apprentissage
Maturation d'affinité	Mutation aléatoire de l'ARB et le retrait

Les systèmes immunitaires

	des ARB les moins stimulé
Mémoire immunitaire	L'ensemble mémoire des ARB mutés
Metadynamics	L'enlèvement continu et la création des ARB et des cellules de mémoire

Tableau 2 Correspondance entre AIRS et le système immunitaire

2.4.2 Définition

L'AIRS est un algorithme inspiré du système immunitaire mais à la différence des autres algorithmes des SIAs est l'apprentissage supervisé.

Les AIRSs utilisent une terminologie légèrement différente des autres algorithmes SIAs, les termes et concepts qu'on trouve dans l'algorithme AIRS sont: [18]

➤ **Artificial Recognition Ball (ARB):** ou cellules B. Chaque ARB contient des informations sur la cellule-B (anticorps, un comptage du nombre de ressources détenues par la cellule, et la valeur de stimulation courante de la cellule).

➤ **Seuil d'Affinité (AT):** La valeur moyenne d'affinité entre tous les antigènes dans l'ensemble d'apprentissage ou au sein d'un sous-ensemble sélectionné de ces antigènes d'apprentissage.

➤ **Seuil d'Affinité scalaire (ATS):** Une valeur comprise entre 0 et 1 qui, multiplié par le seuil d'affinité AT, fournit une valeur seuil pour le remplacement de cellules de mémoire dans la routine d'apprentissage AIRS.

➤ **Cellule mémoire candidate:** Les cellules mémoires de l'ensemble ARB, de la même classe que l'antigène de l'apprentissage, qui était le plus stimulée après l'exposition à l'antigène donné.

➤ **Taux de clonage:** Une valeur entière utilisée dans le produit " taux de Clonage * valeur stimulation " qui détermine le nombre de clones à produire par une ARB après avoir répondu à un antigène donné. Ce produit est également utilisé dans l'affectation de ressources à un ARB.

➤ **Taux d'hypermutation:** Une valeur entière utilisée pour déterminer le nombre de clones mutés qu'une cellule mémoire donné est autorisée à injecter dans la population cellulaire. Dans l'implémentation actuelle, la cellule mémoire sélectionnée injecte au moins " *taux d'hypermutation * taux de clonage * valeur de stimulation* " de clones mutés dans la population cellulaire au moment de l'introduction d'antigène.

➤ **Ressources:** Un paramètre qui limite le nombre d'ARB permis dans le système. On attribut à chaque ARB un certain nombre de ressources en fonction de la valeur de stimulation

Les systèmes immunitaires

et du taux de clonage. Le nombre total des ressources système à l'échelle est réglée à une certaine limite. Si le nombre de ressources consommés est supérieur au nombre de ressources autorisé à exister dans le système, alors les ressources sont retirées des cellules les moins stimulées jusqu'à ce que le nombre de ressources dans le système retourne au nombre autorisé. Si toutes les ressources d'une ARB donnée sont retirées, alors l'ARB est retiré de la population de cellules.

➤ **Fonction de stimulation:** Une fonction qui permet de mesurer la réponse d'un ARB à un antigène ou à un autre ARB.

➤ **Valeur de stimulation:** La valeur retournée par la fonction de stimulation.

➤ **Seuil de stimulation (ST) :** Un paramètre entre 0 et 1 utilisée comme un critère d'arrêt pour l'apprentissage sur un antigène spécifique.

L'algorithme AIRS comporte quatre étapes:

- **Étape 1:** initialiser et normaliser les données.
- **Étape 2:** identification des cellules mémoire et génération des ARBs.
- **Étape 3:** compétition des ressources et développement de la cellule mémoire candidate.
- **Étape 4:** introduction d'une cellule mémoire candidate dans l'ensemble des cellules mémoires.

L'algorithme AIRS est le suivant: Notons que:

- *agi* et *agj*: deux antigènes.
- *affinité* (*a,b*): la distance euclidienne normalisée entre *a* et *b*.
- *mc*: un élément de l'ensemble MC.
- *agc*: la classe de l'antigène *ag*.
- *AB*: représente un ensemble d'ARBs.
- *MU*: un ensemble de clones d'ARBs muté.
- $\text{stimulation}(x,y) = 1 - \text{affinité}(x,y)$.

Entrées: S: ensemble des antigènes.

Sortie : MC: ensemble des cellules mémoire

1. Initialisation

- Normaliser les données d'apprentissage (antigène), les vecteurs caractéristiques appartiendront à l'intervalle $[0, 1]$.

- Calculer le seuil d'affinité (AT) selon l'équation (1).

$$AT = \frac{\sum_{i=1}^n \sum_{j=i+1}^n \text{affinité}(ag_i, ag_j)}{\frac{n(n-1)}{2}} \quad (1)$$

2. Pour chaque antigène faire:
 - a. *Identification des cellules mémoire et génération des ARB*: Pour chaque élément de MC
 - Déterminer son affinité avec l'antigène.
 - Sélectionner *mc* la cellule avec la plus haute affinité.
 - Générer des clones de *mc* et les ajouter à l'ensemble des ARB (P).
 - muter chaque clone de *mc* et ajouter ces ARB muté à P.
 - b. *Compétition des ressources*:
 - Calculer les ressources de chaque ARB et éliminer ceux qui n'ont pas de ressource.
 - Calculer la moyenne de stimulation.
 - c. *Expansion clonal et Maturation d'affinité*: cloner et muter les ARB restant dans P.
 - d. *Cycle*: si la moyenne de stimulation est inférieure au seuil de stimulation revenir à b sinon arrêter l'apprentissage sur cet antigène.
 - e. *Introduction des cellules mémoire dans MC*:
 - Sélectionnez l'ARB avec la plus forte affinité et de la même classe que l'antigène.
 - Si la stimulation de la cellule candidate est supérieure à la stimulation de toutes les cellules mémoire dans MC alors ajouter cette cellule à MC.
 - Si l'affinité de la cellule candidate avec la meilleure cellule de MC (*mc*) est inférieure au seuil d'affinité alors supprimer *mc*.

À la fin de l'algorithme présenté, nous obtenons un ensemble de cellules mémoires qui ont subi un apprentissage leur permettant d'effectuer une classification et cela grâce à la méthode **k-plus proche voisin**.

2.5 Implémentation d'un SIA

Afin d'implémenter un système immunitaire artificiel, il faut suivre les étapes suivantes :

- Identifier les éléments qui feront partie du SIA.
- Choisir la meilleure représentation (encoding) pour ces éléments.

- Déterminer la mesure d'affinité appropriée.
- Choisir l'algorithme immunitaire qui convient le mieux au type du problème à résoudre.

2.6 Domaines d'application des SIAs

Le système immunitaire artificiel s'inspire du système immunitaire naturel et chaque processus de ce système sert de base pour un modèle différent. Cette diversité de modèle permet que les SIAs soient utilisés pour résoudre plusieurs problèmes différents tel que : [18]

2.6.1 Robotique

L'une des tâches les plus difficiles de la robotique est le problème de navigation autonome, où un robot (un ensemble de robots) doit pouvoir accomplir certaines tâches sans aucune indication extérieure. Les articles "A Robot with a Decentralized Consensus-Making Mechanism Based on the Immune System" (1997) et "Immunoid: A Robot with a Decentralized Behavior Arbitration Mechanisms Based on the Immune System" (1996) de Ishiguro et al. portaient sur l'élaboration d'un mécanisme dynamique consensus de décision décentralisé, fondé sur la théorie des réseaux immunitaires où l'intelligence devait émerger des interactions réciproques entre les agents (appelé modules de compétence) ou entre un robot et son environnement. Leur but était la construction d'un mécanisme qui peut prendre les décisions appropriées entre plusieurs modules et la façon de préparer ces modules. Cette méthode a été évaluée sur un problème de collecte des ordures en tenant compte de l'autonomie, c'est à dire un robot devait recueillir un ensemble d'ordures et le mettre dans une poubelle, sans manquer d'énergie (niveau de batterie).

2.6.2 Optimisation

Les problèmes d'optimisation sont présents dans plusieurs domaines d'applications. Le but dans ce type de problème est de trouver l'ensemble des meilleures conditions admissibles pour atteindre un certain objectif.

En 2000, De Castro & Von Zuben ont présenté un algorithme de sélection clonale, qui prend en compte la maturation de l'affinité de la réponse immunitaire, afin de résoudre des

problèmes complexes, comme l'apprentissage et l'optimisation multimodal. Leur algorithme constitue une mise en oeuvre des processus biologiques et ne tient pas compte de toute sophistication mathématique pour améliorer ses performances dans des tâches particulières.

2.6.3 Sécurité des ordinateurs

La protection des ordinateurs contre les virus, les utilisateurs non autorisés, etc., constitue un champ riche de la recherche pour les systèmes de détection d'anomalie. En 1994, Forrest et al. dans leur article "*Self Non-self Discrimination in a Computer*" comparent le problème de la protection des systèmes informatiques à celui de l'apprentissage de la distinction entre soi et le non-soi des systèmes immunitaires. Ils ont décrit une stratégie de détection basée sur la *sélection négative* intrinsèque à notre système immunitaire.

2.6.4 La détection et l'élimination des virus informatiques

Dans le système développé par Kephart en 1994 dans son article "*A Biologically Inspired Immune System for Computers*", un ensemble d'anticorps de virus informatiques ou des vers qui n'ont pas été rencontrés ont été générés de façon à favoriser une réponse plus rapide et plus forte contre de futures infections. Il était également préoccupé par la réduction du risque d'une réaction auto-immune, dans laquelle le système immunitaire informatique aurait à tort identifié les logiciels légitimes comme étant indésirables.

En 1999, Okamoto et Ishida dans leur article ont proposé un système multi agent basé sur les SIAs, plus précisément l'algorithme de la sélection négative. La détection des virus est réalisée en effectuant une correspondance entre les informations propres d'un fichier (tel que les premiers bits de l'entête du fichier, sa taille, le chemin d'accès) et le fichier de l'hôte. La neutralisation des virus est faite par la réécriture des informations initiales sur le fichier infecté.

2.6.5 Reconnaissance de formes

La reconnaissance de formes est le domaine de recherche qui étudie le fonctionnement et la conception de systèmes capables de reconnaître des tendances dans les données. Il renferme des sous-disciplines comme l'analyse discriminante, l'extraction de caractéristiques, estimation de l'erreur, inférence grammaticale et syntaxique (appelé reconnaissance des

formes syntaxiques). Quelques domaines d'application : l'analyse d'image, la reconnaissance de caractères, l'analyse de la parole, le diagnostic, l'identification des personnes et l'inspection industrielle.

Forrest et al dans "Using Genetic Algorithms to Explore Pattern Recognition in the Immune System" en 1993 utilisent un modèle binaire du système immunitaire, afin d'étudier la reconnaissance de formes et l'apprentissage dans le système immunitaire. Ils utilisent aussi un algorithme génétique pour étudier le maintien des capacités de diversité et généralisation du modèle de chaîne de bits du système immunitaire, lors de la généralisation des moyens de détection de schémas communs qui sont partagés entre de nombreux antigènes.

2.6.6. Diagnostic médical

L'utilisation des systèmes de classification pour le diagnostic médical augmente graduellement. Il n'y a aucun doute que le facteur le plus important lors du diagnostic est la décision de l'expert mais les systèmes intelligents de diagnostic apportent une aide non négligeable puisqu'ils réduisent les erreurs dues à la fatigue. Ces systèmes peuvent aussi réduire le temps nécessaire au diagnostic. Parmi ces systèmes, se trouvent les systèmes immunitaires artificiels.

En examinant les différents articles qui concernent la classification de données médicales, il est possible de distinguer qu'un grand nombre de problèmes utilise comme algorithme de base l'algorithme AIRS. Parmi ces méthodes nous trouvons :

Dans [19] Hung-Chun Lin & al. utilisent AIRS pour la prédiction du diabète du type 2 chez les femmes souffrant de diabète gestationnel, ils ont obtenu une précision de 62.8 %.

Dans [20], Seral Ozsena & al. proposent un algorithme baptisé SAMA (Supervised Affinity Maturation Algorithm) pour le diagnostic l'athérosclérose en utilisant des échographie Doppler, ils ont obtenu une précision de classification de 98.93 %, tandis que dans [21] les auteurs ont traité le même problème avec un AIRS et ont obtenu 99,29 %. Dans [22], l'introduction d'une étape de prétraitement (PCA-k-NN) avant l'exécution du AIRS a augmenté la précision jusqu'à 100 %.

Dans plusieurs publications une étape de prétraitement a été introduite avant l'exécution du AIRS, ce qui a contribué à augmenter la précision de classification. Parmi ces publications, nous trouvons [23] où les auteurs utilisent une étape de sélection de caractéristiques afin de réduire le nombre d'attributs utilisés, ils ont obtenu une précision de 98.51 % pour le problème du cancer du sein, alors que dans [24] ils ont utilisé la méthode d'analyse en composantes principales pour le cas du cancer de poumon et ont obtenu une précision de 100 %.

Dans les cas précédents, les prétraitements effectués consiste à réduire le nombre de caractéristiques de la base de données par contre dans ([25], [26]), Salih Güneş & al. attribuent des poids aux caractéristiques en fonction de la pertinence de l'information qu'ils apportent au problème. La méthode utilisée en [25] avec la base de données "thyroid disease" a donné comme résultat une précision de 95.90 %, tandis que la méthode utilisée en [26] sur la base de données "Pima Indians Diabetes" a donné une précision de 75.87 %.

2.6.7 Autres applications

Il existe d'autres domaines d'application pour les SIAs tels que:

- **Segmentation d'images:** McCoy & Devarajan. "*Artificial Immune Systems and Aerial Image Segmentation*", en 1997.
- **Apprentissage:** Potter & De Jong. "*The Coevolution of Antibodies for Concept Learning*", en 1998.
- **Ordonnancement:** Hart et al. "*Producing Robust Schedules Via An Artificial Immune System*", en 1998.
- **Data mining:** Hunt & Fellows "*Introducing an Immune Response into a CBR system for Data Mining*", en 1996.
- **Système de classification :** Forrest & Hofmeyr. "*John Holland's Invisible Hand: an Artificial Immune System* ", en 1999.

2.7 Conclusion

Ce chapitre a mis l'accent sur un système d'apprentissage supervisé basé sur des principes immunologiques. Le système de reconnaissance immunitaire artificiel (AIRS) présenté

Les systèmes immunitaires

comme un succès initial d'algorithme de classification. Cependant, comme avec n'importe quel système initial, il y a des révisions et des améliorations qui peuvent être apportées au AIRS qui réduisent la complexité du système

La seconde partie de ce projet sera consacrée à la conception de notre application basée sur l'algorithme AIRS inspiré des systèmes immunitaires artificiels. Nous finirons par une série de tests qui nous permettront d'évaluer notre algorithme et surtout de mettre en relief l'apport des systèmes immunitaires pour la classification médicale.

1. Introduction

Ce chapitre présente la méthode M-AIRS, utilisée pour effectuer la classification du diabète, qui est une modification de l'algorithme AIRS (présenté dans le chapitre 2). Le deuxième algorithme est le K-plus proche voisin. AIRS a été introduit en 2001 comme l'une des premières approches du système immunitaire dédié à la classification. À la différence des autres algorithmes qui s'inspirent du système immunitaire, l'AIRS est un algorithme à apprentissage supervisé, le seul autre algorithme qui utilise les systèmes immunitaires pour l'apprentissage supervisé a été présenté par Carter en 2000 dans son article "The immune systems as a model for pattern recognition and classification".

Notre travail est basé sur l'algorithme AIRS, inspiré du principe de la sélection clonale, développé par A. Watkins [WAT, 01]. Nous essayerons d'améliorer l'efficacité et l'efficacité de cet algorithme en intervenant principalement au niveau de l'algorithme de mutation et la méthode de classification.

2. Description de la base de données utilisée

Dans ce mémoire nous avons utilisé la base Pima Indian Diabetes (PID), disponible sur le site officiel de l'UCI (machine learning repository), offerte par Vincent Sigillito. Cette base est une collection de rapports de diagnostic médicaux de 768 femmes âgé de plus de 21 ans, ces patientes sont des indiennes Pima, une population vivant près de Phoenix, Arizona, USA.

2.1 L'intérêt de la base de données

Des études ont montré que les Indiens Pima d'Arizona, dont le régime alimentaire et le mode de vie est semblable à la plupart des Américains, ont un taux de diabète de type 2 beaucoup plus élevé que la moyenne nationale, ce qui les rend le groupe le plus prédisposé au diabète dans le monde, elle atteint 70 % pour les personnes âgées entre 55 et 64 ans. Les modifications du mode de vie au cours du siècle dernier sont à l'origine de l'apparition du diabète chez les Indiens Pimas mais elles ne peuvent être tenues pour responsables de son caractère épidémique. Cette population présente une prédisposition génétique particulière au diabète de type 2. Avec un degré d'obésité et un niveau glycémique similaire, les Indiens

Résultats et discussion

Pimas, comparés aux autres populations, présentent une résistance à l'action de l'insuline. Cette caractéristique est certainement en partie génétiquement déterminée. [27]

Les recherches sur les indiens Pima ont commencé en 1963 lorsque le NIDDK (National Institute of Arthritis, Diabetes and Digestive and Kidney Diseases) a effectué une étude sur les rhumatismes et ils découvrirent une fréquence invraisemblable de diabète. Deux ans plus tard la communauté Pima, le NIH (National Institutes of Health) et le IHS (Indian Health Service) lancent une étude épidémiologique pour découvrir les raisons de ce taux élevé et ils espèrent expliquer pourquoi les Indiens Américains, les hispaniques et d'autres peuples non-blancs ont un taux de diabète qui atteint jusqu'à dix fois plus que les Blancs. La concentration à Gila River de gens qui ont un patrimoine génétique homogène et qui, en moins de cinquante ans, ont vu leur mode de vie passer brutalement du néolithique à la société de consommation, a fait de cette tribu un modèle idéal pour étudier l'interaction de la génétique et de l'environnement dans le diabète. [27]

Le généticien James Neel de l'université du Michigan explique que la cause de cette prédisposition est le «génotype économe». Toutes les régions frappées de plein fouet par le diabète de type 2 ont été victimes de la «coca-colonisation» il y a cinquante ans. En quelques années, des populations entières sont passées d'un mode de vie traditionnel cueilleur/chasseur à une société d'abondance où l'activité physique est réduite au minimum et le régime alimentaire indexé sur le fast-food. Donc le gène qui favorise le stockage de graisse et qui leur a permis de survivre en période de famine s'est retourné contre eux en causant l'obésité qui est l'un des facteurs de risque du diabète de type 2. [[27], [28]]

L'interaction des gènes et de l'environnement a été confirmée par les chercheurs de Phoenix grâce au indiens Pima du Mexique. Les Indiens Pima en Arizona ont été génétiquement lié à un village d'Indiens Pima vivant un mode de vie plus traditionnel dans une région montagneuse du Mexique. Une étude en 1995 sur les Pimas mexicains a révélé une incidence rare de diabète. Plusieurs recherches ont été menées afin de déterminer la prédisposition génétique des indiens Pima à l'obésité et le diabète ainsi que des anomalies en cause. [[27], [28]]

Résultats et discussion

Des années d'efforts collectifs par les scientifiques et les volontaires ont jeté les bases pour éventuellement guérir ou de prévenir le diabète et ses complications. Ces travaux ont permis de fixer les critères de diagnostic utilisés pour identifier et traiter le diabète et d'anticiper la façon dont elle est susceptible de se développer. [27]

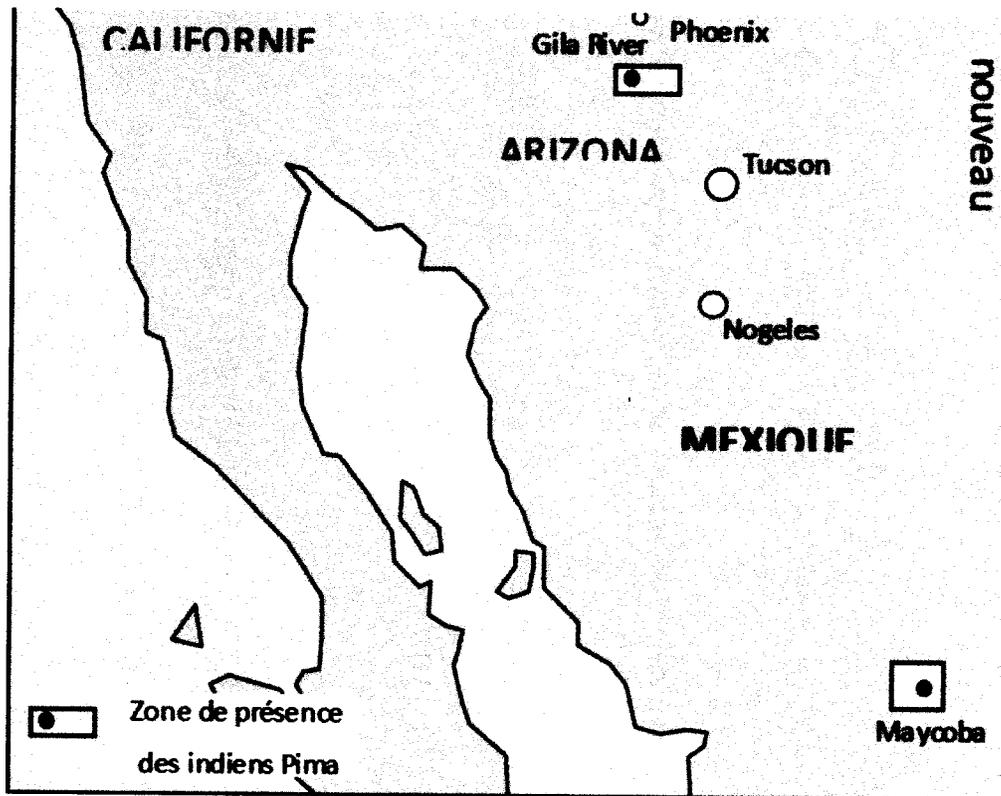


Figure 13 : Zones de présence des indiens Pima

2.2 Base de données utilisées

La base Pima Indian Diabetes est constitué de 768 cas dont 268 sont diabétique et 500 non diabétique. Chaque cas est formé de 9 attributs, dont 8 représentent des facteurs de risque et le 9eme représente la classe du patient, le tableau 3 présente une description de ces attributs:

<i>N° Attribut</i>	<i>Description attribut</i>	<i>Moyenne</i>	<i>Déviatiion standard</i>
1	Nombre de grossesses (Ngross)	3.8	3.4
2	Concentration du glucose plasmatique (mg/dl)	120.9	32.0
3	Pression artérielle diastolique (mm Hg)	69.1	19.4

Résultats et discussion

4	(PAD) Épaisseur de la peau au niveau du triceps (mm) (Epai)	20.5	16.0
5	Taux d'insuline au bout de 2 heures (mU/ml) (INS)	79.8	115.2
6	Indice de masse corporelle (poids en kg/ m ²) (IMC)	32.0	7.9
7	Fonction pédigrée du diabète (Ped)	0.50	0.3
8	Age (années)	33.20	11.8

Tableau3 : Description des attributs de la base

Malheureusement, il existe quelques cas avec des données manquantes qui ont été remplacé par des zéros, ce qui donne des valeurs biologiquement impossible tel qu'une pression artérielle égal à 0. Après élimination de ces cas, nous obtenons une base de 392 patientes dont 262 non-diabétiques et 130 diabétiques.

3. Algorithme modified-AIRS

3.1 Principe

Le présent algorithme est inspiré des travaux de A. Watkins nommé AIRS pour la reconnaissance des formes. Tout d'abord, l'ensemble d'apprentissage représente les antigènes dont il aura besoin pour le déroulement du programme, et le résultat d'apprentissage est l'ensemble d'anticorps (cellules B mémoire) qui seront utilisés pour la phase de test (les nouveaux antigènes introduits dans le système).

Résultats et discussion

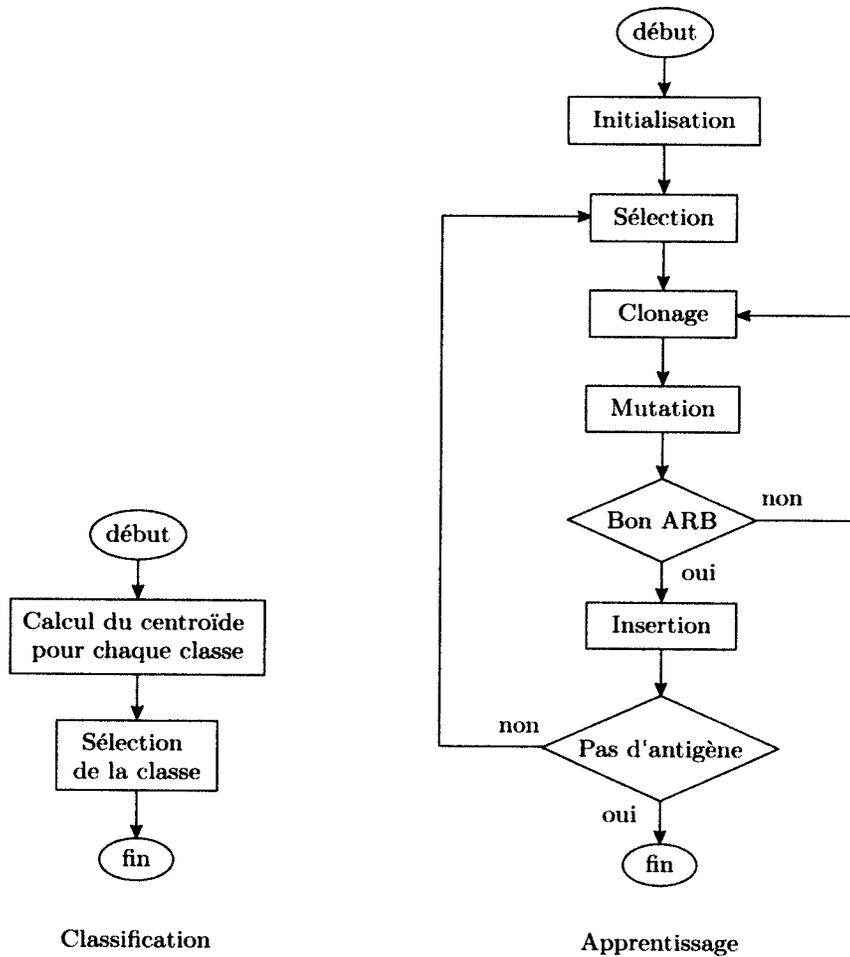


Figure 14 : Représentation de l'algorithme m-AIRS

Pour ce qui va suivre nous utilisons la notation suivante :

Paramètres	Description
MC	représente l'ensemble des cellules mémoires, mc est un individu de cet ensemble
ag	est l'antigène en cours de traitement
AB	est l'ensemble des ARBs, ab représente un ARB où $ab \in AB$
nc	est le nombre de classes

L'apprentissage du système immunitaire artificiel se déroule en quatre étapes, chaque sortie est une introduction à l'étape suivante. Les étapes sont détaillées dans les points qui suivent:

3.2 Etapes d'apprentissage

3.2.1 Initialisation

La phase d'initialisation peut être divisée en deux parties principales, le prétraitement de la base d'apprentissage et le calcul du seuil d'affinité. Durant cette phase les valeurs de la base sont toutes normalisées dans l'intervalle $[0,1]$. Après la normalisation, le seuil d'affinité est calculé par la moyenne de toutes les valeurs d'affinité entre les paires de vecteurs appartenant à la base d'apprentissage. Le seuil d'affinité est calculé par l'équation suivante :

$$\text{seuil d'affinité} = \frac{\sum_{i=1}^{n-1} \sum_{j=i+1}^n \text{affinite}(ag_i, ag_j)}{\frac{n(n-1)}{2}}$$

Où n est le nombre d'antigènes dans la base d'apprentissage, ag_i et ag_j sont respectivement le i ème et le j ème antigène de la base d'apprentissage et $\text{affinite}(x,y)$ retourne la distance Euclidienne entre les vecteurs des deux antigènes.

Il faut noter ici qu'il est possible d'utiliser d'autres mesures de stimulation telle que la distance de Manhattan ou d'autres.

La dernière étape de l'initialisation est la génération de l'ensemble initiale des cellules mémoires. L'initialisation de cet ensemble se fait par l'introduction de 0 ou plus d'antigènes choisis aléatoirement de la base d'apprentissage.

3.2.2 Identification des cellules B et génération des ARBs

Une fois que l'initialisation est achevée, cette étape aura lieu pour chaque nouvel antigène introduit. Une cellule mémoire est sélectionnée de l'ensemble entier de cellules B est nommée mc_{match} , cette dernière est tirée de telle sorte qu'elle ressemble le plus à l'antigène en cours de traitement selon la formule :

$$mc_{match} = \arg \max_{mc \in MC} \text{stimulation}(ag, mc)$$

Où mc est une cellule mémoire appartenant à l'ensemble MC , ag est l'antigène en cours de traitement et $\text{stimulation}(x, y)$ retourne $1 - \text{affinite}(x,y)$.

Résultats et discussion

Une fois que la cellule *mcmatch* est sélectionnée, elle sera utilisée pour générer l'ensemble ARB. L'ensemble ARB contient les clones de *mcmatch* correspondante à l'antigène en cours de traitement.

Le nombre de clones de la cellule *mcmatch* généré par mutation est limité selon la formule:

$$\text{nombre_de_clones} = \text{hyper_taux_clonage} * \text{taux_clonage} * \text{stimulation}(mc_{\text{match}}, ag)$$

```
Mutation(X,b){
  Pour chaque xi de X faire
    change = random()
    changer_vers = random()
    SI(change < taux_mutation)
      xi = changer_vers
      b = vrai;
    FinSI
  FinPour
  SI(b == vrai)
    change = random()
    changer_vers = lrandom() mod nc
    SI(change < taux_mutation)
      x.c = changer_vers
    FinSI
  FinSI
  Retourner X
}
```

L'algorithme original de la mutation [29] consiste à générer à chaque fois des nouvelles valeurs aléatoires pour la cellule mutée. Cette méthode ne permet pas de savoir ni d'estimer le temps nécessaire pour avoir une cellule qui a une bonne stimulation avec l'antigène car elle utilise des modifications totalement aléatoire pour avoir la cellule mutée. Dans notre travail nous allons s'inspirer de la mutation du système immunitaire naturel pour essayer de résoudre ce problème.

L'algorithme de mutation utilisé dans ce travail est le suivant :

Résultats et discussion

```

Mutation(X){
  Pour chaque xi de X faire
    opération = random()
    Si(opération dans [0, 0.33])
      xi = xi+pas_mutation
    FinSi
    Si(opération dans ]0.33, 0.66])
      xi = xi
    FinSi
    Si(opération dans ]0.66, 1])
      xi = xi-pas_mutation
    FinSi
  FinPour
  Retourner X
}
  
```

Cet algorithme consiste à générer des cellules voisines à la cellule initiale en modifiant partiellement les valeurs de cette dernière ce qui nous permet de se rapprocher de la meilleure stimulation avec l'antigène.

Seuil de stimulation	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Algorithme original	3300	3300	3300	6570	224070	3.5x10 ⁶	10.8x10 ⁶
Nouvel algorithme	313	344	328	328	344	469	1031

Tableau 4 : Temps d'apprentissage en millisecondes

Les résultats obtenus démontrent que la durée d'apprentissage appliquée sur la base PIMA a diminué à un temps moyen égal à 2s pour un seuil de stimulation égale à 0.9 et c'est une valeur qui peut être négligée par rapport au temps nécessaire pour l'algorithme original qui dépasse 180 mn pour la même valeur du seuil.

3.2.3 Compétition des ressources et développement des cellules mémoires candidates

Cette partie se base sur les résultats de la précédente, elle complète les informations des ARBs générés en calculant les ressources pour chaque anticorps avec l'antigène en cours de traitement. Ces ressources sont mises à jour durant l'apprentissage, et chaque ARB n'ayant

Résultats et discussion

pas de ressources sera supprimé de l'ensemble des ARBs. Le nombre de ressources pour chaque ARB est calculé par la formule :

$$ressources = stimulation(ag, ab) * taux_clonage$$

Les étapes 2 et 3 seront exécutées pour un antigène donné jusqu'à ce que la condition $s \geq Seuil_stimulation$ soit vérifiée, avec

$$s = \frac{\sum_{i=1}^{|AB|} ab_i.stim}{|AB|}, ab_i \in AB$$

Une fois que la condition précédente est vérifiée, on peut dire que nous avons pu avoir un ensemble d'ARBs optimal.

3.2.4 Introduction des cellules mémoires

Cette étape consiste à choisir parmi les ARBs la cellule candidate qui convient le plus à l'antigène ; en d'autres termes celle qui le reconnaisse mieux que *mcmatch*. La cellule candidate *mccandidate* sera additionnée à l'ensemble des cellules mémoires seulement si elle retourne une stimulation plus élevée que la cellule de *mcmatch* avec l'antigène en cours de traitement. Si l'affinité entre *mcmatch* et *mccandidate* est inférieur au produit entre le seuil d'affinité et le scalaire du seuil d'affinité (ATS) définis par l'utilisateur, dans ce cas *mcmatch* sera supprimée de l'ensemble des cellules mémoires.

3.3 Classification

A la fin de la phase d'apprentissage, l'ensemble des cellules mémoires sont utilisés pour la classification.

Dans [29] ils ont proposé d'utiliser l'approche des k plus proches voisins. Dans cette approche les k cellules mémoires qui ont les meilleures stimulations avec le nouvel antigène à classer sont choisis, ensuite un vote aura lieu pour choisir la classe de l'antigène.

4. Résultats

4.1 Choix des paramètres

Dans cette section, nous discuterons les choix des paramètres d'apprentissage, et la validité de ces choix.

Résultats et discussion

La manière la plus simple et la plus naïve est de tester toutes les combinaisons des valeurs possibles des paramètres. Cette méthode possède l'inconvénient qu'elle n'admet pas de sémantique lors d'une affectation des différents paramètres. Cependant, il est possible d'utiliser des heuristiques pour guider l'affectation.

4.1.1 Pas de mutation

Le pas de modification de chaque attribut est défini par *pas_mutation*. Puisque les valeurs de chaque cellule appartiennent à $[0,1]$ la valeur de *pas_mutation* doit appartenir au même intervalle. La valeur prise définit le degré de voisinage entre la cellule originale et la cellule mutée, plus la valeur est grande, plus les deux cellules sont plus éloignées. Elle a aussi un impact sur le temps d'exécution.

4.1.2 Seuil de stimulation

Le seuil de stimulation est le paramètre qui raffine l'ensemble des cellules mémoires et il définit aussi la qualité de l'ensemble ARB. Une grande valeur du seuil de stimulation veut dire que les cellules ont une bonne stimulation avec l'antigène.

4.1.3 Scalaire du seuil d'affinité (ATS)

ATS est le scalaire du seuil de l'affinité, il est utilisé pour comparer la distance entre la nouvelle cellule mémoire et *mcmatch*, si cette distance est petite (i.e. les deux cellules reconnaissent le même antigène) la *mcmatch* est supprimé. Plus le scalaire est grand plus les cellules mémoires obtenus auront la possibilité de reconnaître divers antigènes et c'est ce qu'on veut dans notre cas.

4.1.4 Taux de clonage

Le taux de clonage va déterminer à la fois le nombre de clones pour chaque cellule et le nombre de ressources prises par chaque cellule dans ARB. Nous avons fixé le taux de clonage à 10 pour que le nombre de ressources de chaque cellule soit un entier entre 1 et 10.

4.1.5 Hyper taux de clonage

L'hyper taux de clonage est un scalaire du taux de clonage. Il est utilisé pour déterminer le nombre de clones pour chaque cellule. La valeur choisie est égale à 2.

4.1.6 Nombre de ressources

Le nombre de ressources définit le volume de l'ensemble ARB par rapport à son qualité. Plus le nombre de ressources est grand plus le nombre de cellules de ARB est grand.

Résultats et discussion

Si on veut avoir par exemple, un ensemble de cellules avec une stimulation entre 0,8 et 0,9, le nombre de cellules ARB, pour un taux de clonage égale à 10, sera compris entre $\frac{Nbr_ressources}{10 \times 0,9}$ et $\frac{Nbr_ressources}{10 \times 0,8}$.

4.2 Expérimentations :

Dans cette première expérimentation nous effectuons la classification sur notre base de diabète qui contient 392 cas (sans les valeurs manquantes). C'est-à-dire que nous effectuons l'apprentissage de cette base avec l'algorithme M-AIRS.

Les performances de classification obtenues ont été évaluées en utilisant le classificateur Knn.

Les résultats sont résumés dans le tableau ci-dessous :

ATS	Seuil de stimulation	Nombre de ressources	Taux %	SN %	SP %	Temps Milliseconde
0.1	0.9	200	80.45	75.38	82.81	1074
0.1	0.8	100	75.91	68.71	78.80	803
0.1	0.9	300	81.36	62.36	89.08	1178
0.2	1	300	80.63	63.01	87.88	1634
0.2	0.75	150	77.27	58/06	84.5	731

Tableau 5 : Les résultats obtenus

Suite aux résultats présentés au tableau ci-dessus, nous constatons que

L'augmentation de la valeur du seuil de stimulation améliore le taux de classification plus précisément avec la valeur de $k=0.9$. Le meilleur choix pour la base de diabète, c'est de prendre un seuil de stimulation entre 0.90 et 0.94.

Résultats et discussion

Les mêmes résultats montrent que les taux de classification avec l'ATS=0.1 et l'ATS= 0.2, sont presque similaires et ceci malgré la réduction de la taille de cellules mémoires enregistrée entre les deux valeurs. Si on augmente la valeur d'ATS, le nombre de cellules mémoires va diminuer et nous constatons une amélioration du taux de classification (jusqu'à certain seuil).

Les meilleurs résultats sont obtenus avec ATS = 0.1, seuil = 0.9 et nombre de ressource égale à 300.

5 Comparaison :

Les travaux de ces dernières années menés sur la classification du diabète sont résumés dans le tableau suivant :

Auteurs	Méthode	Taux de classification
Aibinu & al. (2010)	CVNN-based CAR	81.00 %
Aibinu & al. (2010)	RVNN- based AR	80.65 %
Ganji & al. (2010)	fuzzy Ant Colony Optimization	79.48 %
Polat & al. (2007)	AIRS	79.22 %
Notre travail	M-AIRS	81.36 %

Tableau 6 : Etude comparative

Nous remarquons que le taux de classification obtenu par notre méthode est parmi les meilleurs résultats obtenus jusqu'à maintenant pour la classification du diabète avec un taux de 81.36 %.

5. conclusion

Dans ce chapitre, nous avons présenté une des approches des systèmes immunitaires. Le point de départ était les travaux de A.Watkins nommés AIRS qui est un algorithme de la sélection clonale. Dans notre travail, nous avons essayé de faire quelques améliorations sur cet algorithme, soit en termes de rapidité d'exécution, ou bien, en termes de taux de classification.

Les points essentiels touchés dans ce chapitre sont les suivants :

1. Nous avons modifié l'algorithme original de mutation, ce qui nous a permis d'améliorer le temps d'exécution de l'algorithme et le taux de classification.
2. Nous avons aussi changé la méthode classification ce qui nous a donné jusqu'à 50% d'améliorations de taux de classification.

Résultats et discussion

Notre algorithme M-AIRS a atteint son objectif puisque nous avons eu une amélioration, soit en efficacité ou bien en efficacité et en temps d'exécution par rapport à l'algorithme original AIRS, nous avons eu aussi le meilleur taux de classification après l'approche hybride neuro-rule.

Résultats et discussion

Notre algorithme M-AIRS a atteint son objectif puisque nous avons eu une amélioration, soit en efficacité ou bien en efficacité et en temps d'exécution par rapport à l'algorithme original AIRS, nous avons eu aussi le meilleur taux de classification après l'approche hybride neuro-rule.

Conclusion générale

Le diabète est considéré actuellement comme la maladie du siècle vu le nombre de diabétiques qui ne cesse d'augmenter. Selon une enquête menée par l'institut national de santé publique, le diabète se situe dans la quatrième place des maladies chroniques non transmissibles en Algérie. La prévalence du diabète de type 2 dans l'est et l'ouest du pays varie entre 6.4 % et 8.2 % chez des patients allant de 30 à 64 ans . Le diabète est l'une des maladies les plus répandues au monde avec plus de 220 millions de personnes diabétiques. L'augmentation du nombre de diabétiques est tellement rapide que l'organisation mondiale de la santé (OMS) l'a identifié comme étant une épidémie.

Beaucoup de travaux ont été menés afin d'effectuer la classification ou le diagnostic du diabète. Dans ce mémoire , nous avons présenté une méthode pour le diagnostic du diabète basée sur les systèmes immunitaires artificiels et les systèmes flous. L'algorithme des SIAs effectue un apprentissage supervisé et génère un ensemble de données appelé cellules mémoires qui représentent des prototypes des données de la base d'apprentissage.

Nous avons testé notre méthode sur la base Pima Indian . Le taux de classification obtenu avec notre méthode est parmi les meilleurs résultats obtenus jusqu'à maintenant pour la classification du diabète, ce taux est aussi le meilleur résultat obtenu par rapport aux autres algorithmes des systèmes immunitaires artificiels. Les principales perspectives de recherche qui apparaissent à l'issue de ce travail sont l'utilisation d'autres méthodes pour le calcul de la fonction d'affinité à la place de la distance Euclidienne, l'utilisation d'autres méthodes de classification tel que les méthodes à base de connaissance pour augmenter l'interprétabilité du diagnostic. Aussi nous souhaitons trouver une solution efficace pour traiter les données manquantes de la base.

Bibliographie

- [1] A. Moreno et al., *Aprendizaje Automático*.: UPC, 1994.
- [2] N. Lavrac, E. Keravnou, and B. Zupan, *Intelligent Data Analysis in Medicine*., 1997.
- [3] D. Altit, "Diabetes (1^a Parte)," *Correo Farmacéutico*, 2003.
- [4] Medical News Today. [Online]. <http://www.medicalnewstoday.com/info/diabetes/>
- [5] K. Polat and S. Güneş, "An improved approach to medical data sets classification: artificial immune recognition system with fuzzy resource allocation mechanism.," *Expert Systems*, pp. 252-270, 2007.
- [6] Association Latino-Américaine de Diabète ALAD, *Guías ALAD de diagnóstico, control y tratamiento de la Diabetes Mellitus Tipo 2.*, 2007.
- [7] Le Portail des Diabétique Francophones. [Online]. http://www.diabetenaute.net/article.php3?id_article=26
- [8] World Health Organization WHO, "Definition, Diagnosis and Classification of Diabetes Mellitus and its Complications. Report of a WHO Consultation. Part 1: Diagnosis and Classification.," Geneva, 1999.
- [9] mémoire de fin d'étude "classification des données médicales avec le système AIRS2"
- [10] M. Gharbi, *Optimisation grâce aux Systèmes Immunitaires Artificiels*. CERV : Centre Européen de Réalité Virtuelle. EA.3883. EBV. Ecosystémique et Biologie Virtuelles, 2006.
- [11] L.N. De Castro and F.J. Von Zuben, "Artificial Immune Systems: Part I, Basic Theory and Application," Technical Report 1999.
- [12] I. LABED, "Proposition d'un système immunitaire artificiel pour la détection d'intrusions," Université Mentouri De Constantine, Thèse de magister 2006.
- [13] M. Moutschen, "Immunologie générale," Université de Liège, Cour 2009.
- [14] L.N. De Castro and Jonathan Timmis, "In Artificial Neural Networks in Pattern Recognition Artificial Immune Systems: A Novel Paradigm to Pattern Recognition," *University of Paisley, UK*, pp. 67-84, 2002.
- [15] L. N. De Castro and F.J. Von Zuben, "Learning and Optimization Using the Clonal Selection Principle," *IEEE Transactions on Evolutionary Computation; Special Issue on Artificial Immune Systems*, 2001.
- [16] L. Rubi Dr. Alan, *Diabetes Para Dummies, 2a Edición*. Indianápolis, Indiana: Wiley Publishing, Inc, 2007.
- [17] AISWeb; the Online Home of Artificial Immune Systems. [Online]. <http://www.artificial-immune-systems.org/algorithms.shtml>

- [18] A. Watkins, L. Boggess, and J. Timmis, "Artificial Immune Recognition System (AIRS): An Immune-Inspired Supervised Learning Algorithm," *Genetic Programming and Evolvable Machines*, pp. 291–317, 2004.
- [19] Hung-Chun Lin, Chao-Ton Su, and Pa-Chun Wang, "An Application of Artificial Immune Recognition System for Prediction of Diabetes Following Gestational Diabetes," *Journal of Medical Systems*, 2009.
- [20] S. Ozsena, S Karab, F. Latifoglu, and S. Güneş, "A new supervised classification algorithm in artificial immune systems with its application to carotid artery Doppler signals," *Computer methods and programs in biomedicine*, pp. 246–255, 2007.
- [21] F. Latifoglu, K. Polat, S. Kara, and S. Güneş, "Medical application of Artificial Immune Recognition System (AIRS): diagnosis of atherosclerosis from Carotid Artery Doppler Signals," *Journal of Biomedical Informatics*, pp. 1092–1099, 2007.
- [22] F. Latifoglu, K. Polat, S. Kara, and S. Güneş, "Medical diagnosis of atherosclerosis from Carotid Artery Doppler Signals using principal component analysis (PCA); k-NN based weighting pre-processing and Artificial Immune Recognition System (AIRS)," *Journal of Biomedical Informatics*, pp. 15–23, 2008.
- [23] K. Polat, S. Sahan, H. Kodaz, and S. Güneş, "A New Classification Method for Breast Cancer Diagnosis: Feature Selection Artificial Immune Recognition System (FS-AIRS)," in *ICNC 2005, LNCS 3611*. Berlin Heidelberg: Springer-Verlag, 2005, pp. 830 – 838.
- [24] K. Polat and S. Güneş, "Computer aided medical diagnosis system based on principal component analysis and artificial immune recognition system classifier algorithm.," *Expert Systems with Applications*, pp. 773–779, 2008.
- [25] H. Kodaz, S. Ozsen, A. Arslan, and S. Güneş, "Medical application of information gain based artificial immune recognition system (AIRS): Diagnosis of thyroid disease," *Expert Systems with Applications*, pp. 3086–3092, 2009.
- [26] S. Şahan, K. Polat, S. Güneş, and H. Kodaz, "The Medical Applications of Attribute Weighted Artificial Immune System (AWAIS): Diagnosis of Heart and Diabetes Diseases.," in *ICARIS 2005, LNCS 3627*. Berlin Heidelberg: Springer-Verlag, 2005, pp. 456 – 468.
- [27] The Pima Indian pathfinders for Health. The Pima Indian pathfinders for Health. [Online]. <http://diabetes.niddk.nih.gov/dm/pubs/pima/index.htm>
- [28] N. LEVISALLES, "En Arizona; les Indiens pimas recordmen du diabète. En Arizona; les Indiens pimas recordmen du diabète. Enquête sur la population la plus touchée au monde," *Liberation.fr*, 2000. [Online]. <http://www.liberation.fr/sciences/0101335856-en-arizona-les-indiens-pimas-recordmen-du-diabete-enquete-sur-la-population-la-plus-touchee-au-monde>