



# Classification Automatique de l'hypothyroïdie

Khalida Douibi, Nesma Settouti

► **To cite this version:**

Khalida Douibi, Nesma Settouti. Classification Automatique de l'hypothyroïdie. 2013. <hal-00851096>

**HAL Id: hal-00851096**

**<https://hal.archives-ouvertes.fr/hal-00851096>**

Submitted on 12 Aug 2013

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Classification Automatique de l'hypothyroïdie

Khalida Douibi\* & Nesma Settouti\*

\*Département Génie Électrique et Électronique,  
Laboratoire Génie Biomédical  
Université Abou Bekr Belkaid – Tlemcen,  
B.P.230- Tlemcen 13000, Algérie

## Table des matières

<b>I</b>	<b>Présentation de la maladie de l'hypothyroïdie</b>	<b>3</b>
1	Définition	3
2	Comment la glande thyroïde est-elle contrôlée ?	4
3	Détection d'une hypothyroïdie	4
4	C'est quoi l'hypothyroïdie ?	4
4.1	Les types de l'hypothyroïdie . . . . .	5
5	Les symptômes de la maladie	5
6	Les causes principales de l'hypothyroïdie	5
7	Personnes à risques	6
8	Quelles sont les complications de l'hypothyroïdie ?	6
9	Démarche diagnostique	6
10	Le traitement de l'hypothyroïdie	6
11	Quelques conseils lorsqu'on souffre d'hypothyroïdie	6
<b>II</b>	<b>La Classification automatique de l'hypothyroïdie par les Arbres de décisions (CART)</b>	<b>8</b>

<b>1</b>	<b>Problématique</b>	<b>8</b>
<b>2</b>	<b>Pourquoi les arbres de décision</b>	<b>8</b>
<b>3</b>	<b>Études sur la banque de données</b>	<b>8</b>
3.1	Préparation de la banque de données . . . . .	8
<b>4</b>	<b>Classification</b>	<b>10</b>
4.1	Plan . . . . .	10
<b>5</b>	<b>Application</b>	<b>10</b>
<b>6</b>	<b>Résultats et Interprétation</b>	<b>11</b>
6.1	Les grandeurs Basiques pour mesurer les performances d'un classifieur . . . . .	11
6.2	Interprétation . . . . .	13
<b>7</b>	<b>Travail effectué sur la même banque de données</b>	<b>14</b>
<b>8</b>	<b>Conclusion</b>	<b>14</b>
<b>9</b>	<b>Perspectives</b>	<b>15</b>

## Table des figures

1	Schéma représentatif de la thyroïde[dp11] . . . . .	3
2	Schéma représentatif de la procédure de fonctionnement du cerveau afin de contrôler la thyroïde . . . . .	4
3	Graphique représentatif de la procédure d'apprentissage utilisée . . . . .	11
4	Graphique de l'arbre de décision résultant. . . . .	13

## Liste des tableaux

1	Matrice de confusion.[EPS+05] . . . . .	11
2	Quelques formules résultantes de la matrice de confusion.[LSG+04] . . . . .	12
3	Matrice de confusion du classifieur CART. . . . .	12
4	Tableau représentatif des paramètres correspondants aux nœuds. . . . .	13

## Introduction

L'objectif dans ce rapport de recherche est de déterminer les critères de la maladie de la thyroïde en particulier l'hypothyroïdie, et de plus l'application d'une technique de classification sur une banque de données de l'hypothyroïdie qui sera décrite par la suite de ce rapport.

## Première partie

# Présentation de la maladie de l'hypothyroïdie

## 1 Définition

La thyroïde est la plus volumineuse glande endocrine (sécrétant des hormones) elle sécrète principalement 2 hormones [cdIT11] :

- T3 (triiodothyronine)
- T4 (thyroxine ou tétra-iodothyronine)

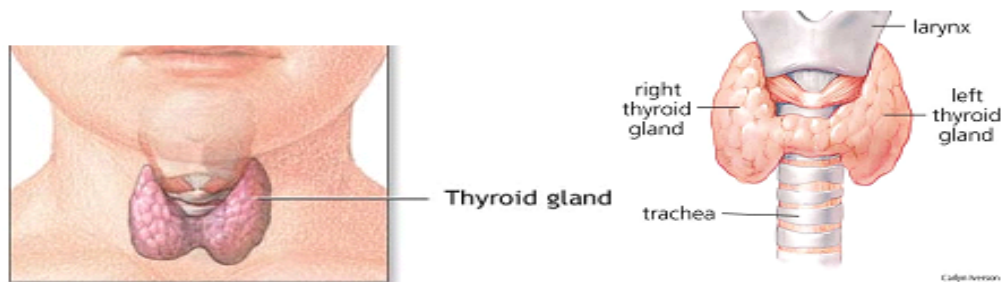


FIGURE 1 – Schéma représentatif de la thyroïde[dP11]

Rôle de ces deux hormones [dP11] :

- Agissent sur le corps pour augmenter le métabolisme basal (circulation sanguine, fonctionnement de cerveau, respiration, digestion, maintien de température du corps ...)
- Agissent sur la synthèse des protéines.
- T3 et T4 stimulent les métabolismes lipidiques, glucidiques, protidiques ainsi que la croissance.
- Contrôler l'énergie musculaire, l'humeur, la concentration...
- Déterminer la vitesse de fonctionnement de nos cellules.

## 2 Comment la glande thyroïde est-elle contrôlée ?

Le schéma ci dessous explique le fonctionnement du cerveau afin de contrôler la thyroïde

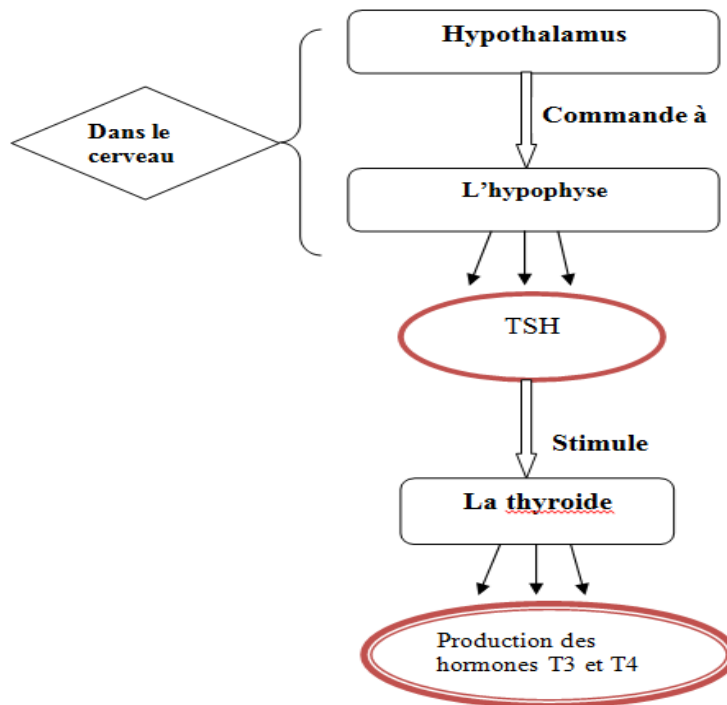


FIGURE 2 – Schéma représentatif de la procédure de fonctionnement du cerveau afin de contrôler la thyroïde

- Si la quantité d'hormones T3 et T4 sécrété est grande donc on a hyperthyroïdie
- Si la quantité d'hormones T3 et T4 sécrété est petite donc on a hypothyroïdie

## 3 Détection d'une hypothyroïdie

On peut détecter une hypothyroïdie en mesurant le taux de TSH dans le sang, et on le trouve élevé car l'hypophyse réagit au manque d'hormones thyroïdiennes (T3 et T4) en sécrétant d'avantage TSH en tentant de stimuler la thyroïde pour qu'elle produise plus d'hormones.

## 4 C'est quoi l'hypothyroïdie ?

L'hypothyroïdie est la conséquence d'une faible production d'hormones par la glande thyroïde, cette maladie touche les femmes après l'âge de 50 ans.

## 4.1 Les types de l'hypothyroïdie

**Primaire :** dysfonctionnement au niveau de la glande thyroïde.

**Secondaire :** dysfonctionnement de l'hypophyse qui sécrète une quantité insuffisante de TSH (hormone de stimulation de la thyroïde).

**L'hypothyroïdie compensée :** est caractérisée par un niveau élevé de TSH avec les niveaux normaux ou seulement légèrement réduits d'hormones thyroïdiennes

**Hypothyroïdie tertiaire :** L'hypothalamus ne fabrique pas assez de TRH. La TRH est une hormone de l'hypothalamus pour stimuler l'anté-hypophyse à produire de la TSH qui elle doit stimuler la thyroïde. La TSH est alors elle aussi basse, comme les hormones.

## 5 Les symptômes de la maladie

- Fatigue et manque d'énergie.
- Gain de poids inexplicable malgré un faible appétit.
- Rythme cardiaque ralenti pouvant causer un essoufflement.
- Irritabilité (hyper-sensibilité) parfois état dépressif.
- Visage et yeux enflés.
- Peau pâle et sèche.
- Perte de cheveux, ongles cassants.
- Cycles menstruels irréguliers.
- Goitre.
- Taux élevé de cholestérol sanguin.

## 6 Les causes principales de l'hypothyroïdie

1. Peut être le résultat du traitement à l'iode radioactif (prise d'I131) pour soigner une hyperthyroïdie, après une chirurgie pour enlever la glande thyroïde. [dEdDemM11]
2. Une thyroïdite de Hashimoto (maladie auto-immune) qui provoque la destruction de la glande thyroïde par le système immunitaire.
3. Peut être résultat d'un défaut congénital (absence de glande thyroïde à la naissance).
4. Ablation chirurgicale de glande thyroïde (traitement du cancer de la thyroïde) entraîne nécessairement l'hypothyroïdie.
5. Le lithium utilisé pour stabiliser l'humeur et traiter les personnes maniaco-dépressives, peut provoquer une hypothyroïdie.
6. Carence en iode (qui est nécessaire à la production du T3 et T4). [dP11]
7. Thyroïdite post-partum (c'est-à-dire après accouchement).

## 7 Personnes à risques

- Les femmes âgées de plus de 50 ans, étant plus facilement atteintes que les hommes à cause des différents événements de leur vie, qui peuvent être à l'origine de bouleversements hormonaux comme : cycle menstruel, grossesse, accouchement, allaitement, pré-ménopause et ménopause. [dEdDemM11]
- Les personnes qui ont des antécédents personnels ou familiaux de maladie de la thyroïde ou de maladie auto-immune (diabète de type 1, maladie coeliaque, l'arthrite rhumatoïde, etc.).
- Les femmes qui ont enfanté au cours de l'année.

## 8 Quelles sont les complications de l'hypothyroïdie ?

Il semblerait qu'une hypothyroïdie fruste non soignée puisse augmenter le risque d'accidents cardiovasculaires (infarctus du myocarde, accident vasculaire cérébral, etc.). Chez les enfants, une hypothyroïdie non soignée entraîne des complications graves : petite taille (nanisme) et retard mental important (le « crétinisme »).

## 9 Démarche diagnostique

La Démarche diagnostique devant une suspicion d'hypothyroïdie primaire [dP11] :

Examens de 1ère intention : TSH

Examens de 2ème intention : T4 libre : détermine la profondeur de l'hypothyroïdie

**Bilan étiologique :**

- Anticorps anti-TPO, anticorps anti-Tg si anti-TPO négatifs
- Echographie thyroïdienne : goitre ? thyroïdite ?

## 10 Le traitement de l'hypothyroïdie

Si l'on ne peut pas guérir d'une hypothyroïdie, il est tout à fait possible de la compenser en prenant chaque jour des hormones thyroïdiennes de remplacement. Une fois le traitement instauré, les symptômes vont diminuer progressivement en quelques semaines et les taux sanguins d'hormones thyroïdiennes vont revenir à la normale. Une fois la posologie efficace identifiée, le traitement est poursuivi à vie. Parce que les hormones thyroïdiennes augmentent le travail du cœur, le médecin va d'abord s'assurer de l'absence de risque de maladie cardiaque (en particulier, de risque d'angine de poitrine ou d'infarctus du myocarde) afin d'éviter que les hormones thyroïdiennes ne déclenchent ce type d'accident [cdIT11].

## 11 Quelques conseils lorsqu'on souffre d'hypothyroïdie

- Suivez scrupuleusement votre traitement, même si vous vous sentez mieux.

- Faites-vous expliquer par votre médecin les signes d'hyperthyroïdie qui pourraient apparaître en cas de surdosage du traitement.
- Un bilan sanguin est pratiqué chaque année afin de vérifier si les doses d'hormones de substitution que vous prenez sont bien adaptées.
- Signalez à votre médecin des symptômes évocateurs d'une hypothyroïdie qui persisteraient alors que vos examens de sang sont normaux.
- Prenez votre médicament chaque jour à la même heure, idéalement trente à soixante minutes avant votre petit-déjeuner.
- Évitez de prendre votre médicament avec des aliments ou des boissons à base de soja ou de café qui diminueraient l'absorption des hormones thyroïdiennes.
- Ayez une alimentation saine et variée, apportant une quantité suffisante d'iode, de sélénium et de zinc.
- Attention aux aliments et aux compléments alimentaires riches en iode (en particulier ceux contenant des algues). Ne les consommez pas sans avis préalable de votre médecin.



## Deuxième partie

# La Classification automatique de l'hypothyroïdie par les Arbres de décisions (CART)

## 1 Problématique

L'hypothyroïdie est une maladie très fréquente chez les femmes après l'âge de 50 ans et elle peut aussi affecter les hommes. Après plusieurs recherches sur le net et un questionnaire fait avec un médecin généraliste DR. Ghechetal salima du Centre Hospitalier de Mostaghanem-Algérie, nous avons recueillis les problèmes que peut rencontrer un médecin lors de son diagnostic général, comme nous l'avons expliqué dans la première partie du mini projet (étude physiopathologique sur la maladie) les symptômes de l'hypothyroïdie sont divers et poussent le médecin à penser à d'autres maladies sans prêter une grande attention à la thyroïde en vue que les hormones de la glande thyroïde sont impliquées dans la plupart des fonctions vitales de l'organisme (respiration, circulation sanguine . . .), la vitesse de fonctionnement des cellules et autres rôles très important pour le fonctionnement normale de notre organisme et un problème touchant la thyroïde peut détruire le corps humain et même conduit à la mort de la personne ;et donc le but que nous visons d'après cette étude est d'appliquer une méthode de classification qui est les arbres de décisions pour classier la maladie selon plusieurs paramètres qu'utilise les médecins pour détecter la présence de l'hypothyroïdie.

## 2 Pourquoi les arbres de décision

Le choix des arbres de décision comme méthode de classification reviens a leurs représentation graphique qui peut être représenter par un ensemble de règles qui sont aisément interprétables [RL02], de plus il est essentiel de pouvoir produire des procédures de classification compréhensibles par l'expert alors à partir des règles de l'arbre on peut expliquer au médecin comment le classifieur a fait pour donner une décision soit le patient est atteint on non par une hypothyroïdie et par la suite lui donner des arguments bien précis et faciles à comprendre puisque mêmes les paramètres utilisés dans cette étude de la banque de données sont utilisés par le médecin pour diagnostiquer l'hypothyroïdie (mesure des paramètres présentant dans le sang).

## 3 Études sur la banque de données

### 3.1 Préparation de la banque de données

- Les paramètres de la banque de données [Qui10] :
- 'age' integer
  - 'sex' F, M codage(F=2,M=1 -> femme=2,homme=1)

- 'on thyroxine' f, t f=0,t=1 (f=false=0,t=true=1)  
La thyroxine est une hormone contenant de l'iode, sécrétée par la glande thyroïde, dont le rôle est d'assurer la régulation de la vitesse du métabolisme et des processus de croissance et de différenciation des tissus. Elle aide également à réguler le développement physique et mental ainsi que la maturité sexuelle
- 'query on thyroxine' f, t
- 'on antithyroid medication' f, t
- 'sick' f, t
- 'pregnant' f, t
- 'thyroid surgery' f, t
- 'I131 treatment' f, t
- 'query hypothyroid' f, t
- 'query hyperthyroid' f, t
- 'lithium' f, t
- 'goitre' f, t
- 'tumor' f, t
- 'hypopituitary' f, t
- 'psych' f, t
- 'TSH measured' t, f
- 'TSH' real
- 'T3 measured' t, f
- 'T3' real
- 'TT4 measured' t, f
- 'TT4' real
- 'T4U measured' t, f
- 'T4U' real colonne x
- 'FTI measured' t, f
- 'FTI' real
- 'TBG measured' f supprimé
- 'TBG' real supprimé
- 'referral source' SVHC, other, SVI, STMW, SVHD // supprimé
- 'Class' negative, compensated-hypothyroid, primary-hypothyroid, secondary-hypothyroid

En observant le jeu de données, nous avons constaté que certaines variables étaient inutiles.

- Suppression de l'attribut *referral source* : il n'apporte aucune information intéressante concernant la maladie de la thyroïde, puisqu'il s'agit de l'endroit où les données ont été récoltées.
- Suppression de l'attribut *secondary hypothyroid* : il n'y a que deux individus qui appartiennent à cette classe.
- Nous avons supprimé la variable *TBG measured* car toutes les valeurs de cet attribut dans la base de données sont nulles donc cet attribut n'a pas d'impact sur la classification car n'apporte pas une différence entre les classes.
- Supprimer la variable *TBG* car toutes les données sont manquantes et donc on conclue que ce paramètre n'est pas pris en considération lors de l'étude et recueil de données de base de données.
- Nous avons remplacé les *données manquantes*, par la moyenne des deux données suivantes et

précédentes.

## 4 Classification

### 4.1 Plan

1. Recherche sur la fonction `classregtree` (CART) pour faire l'application en Matlab .
2. Diviser la Banque de données en deux parties :
3. Une pour l'apprentissage (2/3 de la banque de données).
4. Une pour le test (1/3).
5. Faire la classification.
6. Calculer le taux de classification, spécificité, sensibilité, la matrice de confusion.

## 5 Application

**Recherche sur CART** L'apprentissage de règles par arbres de décision consiste à commencer par construire un arbre de décision à partir des données de l'ensemble d'apprentissage, puis à traduire cet arbre en un ensemble de règles. Pendant la phase de test ;les Arbres de décisions sont des méthodes d'induction utilisant une structure d'arbre pour classer les données. chaque nœud de l'arbre correspond à un test sur un attribut.une question qui se pose est celle du choix de « l'attribut racine ». [LAU] Pour y répondre, on établit un test statistique, qui détermine à quel point tel attribut peut discriminer l'ensemble d'apprentissage

### L'algorithme CART

Indice de Gini

$$I = 1 - \sum_i^N f_i^2$$

$N$  = nombre de classes à prédire

$f_i$  = fréquence de la classe  $i$  dans le nœud Plus l'indice de Gini est bas, plus le nœud est pur.

Conditions d'arrêts existantes [AEBP] :

- Profondeur de l'arbre atteint une limite fixée (=nombre de variables utilisées)
- Nombre de feuilles atteint un maximum fixé
- L'effectif de chaque noeud est inférieur à un Seuil fixé
- La qualité de l'arbre est suffisante
- La qualité de l'arbre n'augmente plus de façon Sensible

Processus d'élagage de CART :

- Création de l'arbre maximum

- Toutes les feuilles des extrémités sont pures
- Élagages successifs de l'arbre
- Retient l'arbre élagué pour lequel le taux d'erreur estimé mesurer sur un échantillon test est le plus bas possible

Pour Comprendre mieux la procédure d'apprentissage utilisée sur la banque de données, nous avons constitué le schéma suivant :

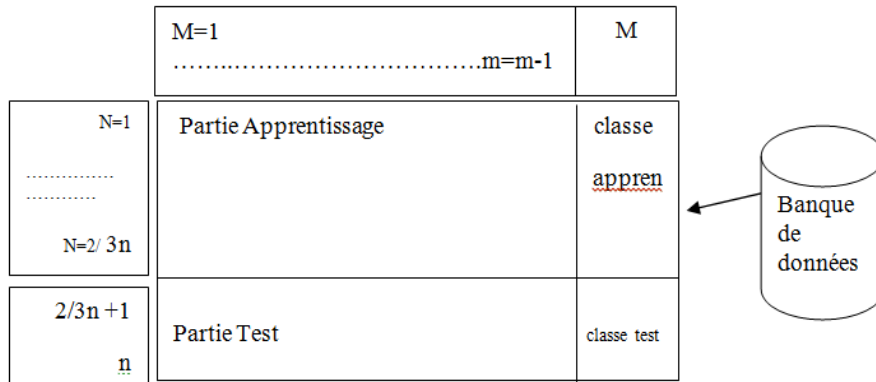


FIGURE 3 – Graphique représentatif de la procédure d'apprentissage utilisée

Avec

N : le nombre de lignes de la banque de données.

M : le nombre de colonnes de la banque de données.

## 6 Résultats et Interprétation

### 6.1 Les grandeurs Basiques pour mesurer les performances d'un classifieur

Calcul des vrai positif (VP), des Vrai négatifs, les faux positifs, faux négatifs, La sensibilité, la spécificité...

		PREDICTED LABEL	
		Positive	Negative
KNOWN LABEL	Positive	True Positive (TP)	False Negative (FN)
	Negative	False Positive (FP)	True Negative (TN)

TABLE 1 – Matrice de confusion.[EPS+05]

MESURE	FORMULE	INTUITIVE MEANING
<b>Recall/Sensitivity</b>	$TP/(TP+FN)$	The percentage of positive labeled instances that were predicted as positive.
<b>Specifity</b>	$TN/(TN+FP)$	The percentage of negative labeled instances that were predicted as negative.
<b>Accuracy</b>	$(TP+TN)/(TP+TN+FP+FN)$	The percentage of predictions that are correct

TABLE 2 – Quelques formules résultantes de la matrice de confusion.[LSG+04]

Les résultats de classification :

**Le taux de classification :** un pourcentage très encourageant=99,68% Ce qui permet de dire que notre classifieur est performant mais il nous reste à analyser encore les autres grandeurs de la Table 2.

**Spécificité :**  $Sp=99,83\%$  C'est-à-dire le pourcentage de cas négatifs qui sont prédits par le classifieur comme étant négatifs.

**Sensibilité :**  $Se=99,65\%$  : C'est-à-dire le pourcentage de cas positifs qui sont prédits par le classifieur comme étant positifs.

	PRÉSENCE D'ÉVÉNEMENT	ABSENCE D'ÉVÉNEMENT
<b>Classification Positive</b>	83	2
<b>Classification Négative</b>	2	1168

TABLE 3 – Matrice de confusion du classifieur CART.

**La matrice de confusion :** De la matrice de confusion, on remarque qu'il y a que 4 cas où le classifieur s'est trompé de prédire leurs classes  $FN=2$  (cela veut dire qu'il y a 2 cas qu'ils ont une hypothyroïdie (primaire ou composated) et le classifieur les a considérés comme hypothyroïdie négative c'est-à-dire absence de l'hypothyroïdie et ce cas est le plus dangereux.

On a  $FP=2$  (c'est-à-dire il existe deux autres cas qui ne sont pas atteint de l'hypothyroïdie et il les a considérés comme Malades en fait c'est une erreur de classification mais elle est moins alarmante que la précédente.

Les  $VP=83$  (cela signifie qu'il ya 83 cas qui sont atteint d'une hypothyroïdie et ils sont bien classés par notre classifieur) c'est un chiffre très encourageant.

Et enfin les  $VN=1168$  (cela signifie qu'il ya 1168 cas qui ne sont pas atteint d'une hypothyroïdie et le classifieur les a bien classés).

## 6.2 Interprétation

L'arbre de décisions généré après la classification étant comme suit Figure 4 :

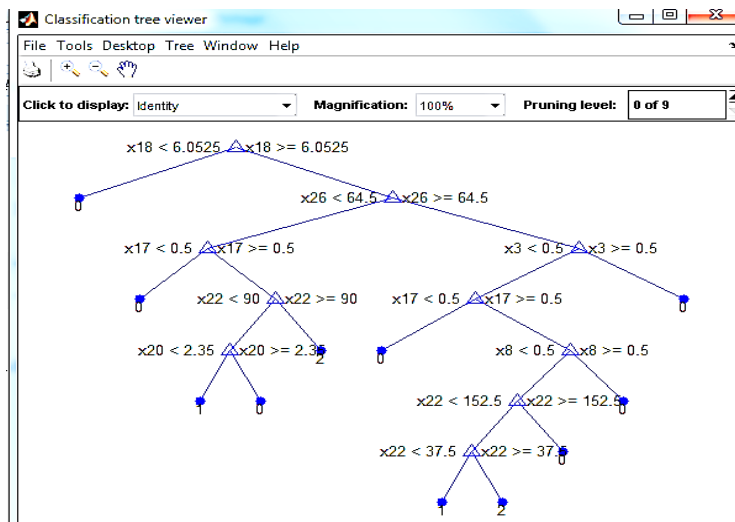


FIGURE 4 – Graphique de l'arbre de décision résultant.

Les règles de l'arbre selon la Figure 4 :

LE NŒUD	PARAMÈTRE CORRESPONDANT	INFORMATION CORRESPONDANTE
X18	TSH.	Hormone thyroïde stimulante.
X26	FTI.	Index libre de thyroxine.
X17	TSH measured.	Vrai ou faux.
X3	On thyroxine.	Hormone contient de l'iode.
X8	Thyroid surgery.	Ablation chirurgical l'un des causes de la tyroïde.
X20	T3.	Triiodothyronine.
X22	TT4.	Thyroxine total.

TABLE 4 – Tableau représentatif des paramètres correspondants aux nœuds.

Les règles de l'arbre de décisions sont :

- Si le taux de TSH < 6.0525 alors la personne n'est pas atteinte d'une hypothyroïdie.
- Si le taux de TSH >= 6.0525 et le taux de FTI (index libre de thyroxine) < 64.5 et TSH < 0.5 alors la personne n'est pas atteinte d'une hypothyroïdie.
- Si le taux de TSH >= 6.0525 et le taux de FTI (index libre de thyroxine) >= 64.5 et On thyroxine (hormone) >= 0.5 et TT4 >= 90 alors la personne est atteinte par une hypothyroïdie compensé.

- Si le taux de TSH  $\geq 6.0525$  et le taux de FTI (index libre de thyroxine)  $\geq 64.5$  et On thyroxine(hormone)  $\geq 0.5$  et TT4  $< 90$  et T3  $< 2.35$  alors on est dans le cas d'une hypothyroïdie primaire.
- Si le taux de TSH  $\geq 6.0525$  et le taux de FTI (index libre de thyroxine)  $\geq 64.5$  et On thyroxine(hormone)  $\geq 0.5$  et TT4  $< 90$  et T3  $\geq 2.35$  alors la personne n'est pas atteinte d'une hypothyroïdie.
- Si le taux de TSH  $\geq 6.0525$  et le taux de FTI (index libre de thyroxine)  $\geq 64.5$  et On thyroxine (hormone)  $< 0.5$  et TSH measured  $< 0.5$  alors la personne n'est pas atteinte d'une hypothyroïdie.
- Si le taux de TSH  $\geq 6.0525$  et le taux de FTI (index libre de thyroxine)  $\geq 64.5$  et On thyroxine(hormone)  $< 0.5$  et TSH measured  $\geq 0.5$  et Thyroid surgery  $\geq 0.5$  alors la personne n'est pas atteinte d'une hypothyroïdie.
- Si le taux de TSH  $\geq 6.0525$  et le taux de FTI (index libre de thyroxine)  $\geq 64.5$  et On thyroxine(hormone)  $< 0.5$  et TSH measured  $\geq 0.5$  et Thyroid surgery  $< 0.5$  et TT4  $\geq 152.5$  alors on a hypothyroïdie négative.
- Si le taux de TSH  $\geq 6.0525$  et le taux de FTI (index libre de thyroxine)  $\geq 64.5$  et On thyroxine(hormone)  $< 0.5$  et TSH measured  $\geq 0.5$  et Thyroid surgery  $< 0.5$  et TT4  $< 152.5$  et SI TT4  $< 37.5$  alors on a une hypothyroïdie primaire sinon une hypothyroïdie secondaire.

## 7 Travail effectué sur la même banque de données

Un réseau bayésien a été réalisé par l'équipe de Aline Conseil et al. [CFFM06] pour le diagnostic de l'hypothyroïdie. Leur objectif était de réaliser un système automatique pour diagnostiquer l'hypothyroïdie. En comparant entre les deux travaux, nous avons travaillé tout d'abord avec les mêmes paramètres et sur la même banque de données (mêmes individus) eux ils ont essayé de faire plusieurs structures de réseau bayésiens pour déterminer le plus fiables avec un taux de classification élevé et bien sûr en regardant d'autres critères, en ce qui concerne le taux de classifications était de 99% en général mais leurs vision de classification est un petit peu différente que la notre, leurs système fait la classification sur 3 classes et classe selon les 3 types, or le mien ne précise pas le type de l'hypothyroïdie dans le cas ou un malade est hypothyroïdien.

## 8 Conclusion

Dans ce rapport et dans le cadre de la classification automatique de l'hypothyroïdie, les arbres de décisions (CART) nous a permis d'obtenir de bons résultats. D'une part un très bon taux de classification et le nombre de cas où le classifieur s'est trompé est minime et d'autre part les résultats donnés sont interprétables et les règles de l'arbre construit correspond exactement au mêmes

paramètres qu'utilise le médecin pour déclarer si la personne est hypothyroïdienne ou pas et à la fin on peut dire qu'on a pu atteindre l'objectif ciblé dès le début de ce travail.

## 9 Perspectives

Dans ce travail nous avons envisagé de regrouper les 2 catégories (hypothyroïdie primaire et compensé) et ainsi de classer les personnes malades ou saines. Nous aurions pu choisir de classer nos individus selon leur type de maladie et ça sera un model plus détaillé et plus satisfaisant qu'on pourra l'appliquer dans un système automatique d'aide au diagnostique de l'hypothyroïdie.

## Références

- [AEBP] Alexandre Aussem, Haytham Elghazel, Khalid Benabdeslem, and Marc Plantevit. Cours msad : Modèles statistiques pour l'aide à la décision. Technical report, Université Claude Bernard Lyon 1.
- [cdIT11] La Fondation canadienne de la Thyroïde. Série de dépliants sur les affections thyroïdiennes. In *Conscience-Soutien-Recherche* [http://www.thyroid.ca/fr/clinical\\_diagnosis.php](http://www.thyroid.ca/fr/clinical_diagnosis.php), Juin 2011.
- [CFFM06] Aline CONSEIL, Violaine FERANDEZ, Nicolas FERRAFIAT, and Olivier MONTANGUON. Réseaux bayésiens pour le diagnostic de l'hypothyroïdie. Technical report, Master 2 GIS MIS, D2006.
- [dEdDemM11] Collège des Enseignants d'Endocrinologie Diabète et maladies M. Enseignement du 2ème cycle polycopie nationale. Technical report, Université Virtuelle Francophone, 2010-2011.
- [dP11] Salle de Presse. Hypothyroïdie congénitale : grâce au dépistage systématique, les personnes atteintes et traitées vivent normalement. Technical report, Assistance Publique-Hôpitaux de Paris. <http://www.aphp.fr/>, 08/07/2011.
- [EPS<sup>+</sup>05] Roman Eisner, Brett Poulin, Duane Szafron, Paul Lu, and Russell Greiner. Improving protein function prediction using the hierarchical structure of the gene ontology. In *CIBCB*, pages 354–363. IEEE, 2005.
- [LAU] PHILIPPE BESSE & BÉATRICE LAURENT. Apprentissage statistique : modélisation, prévision et data mining. Technical report, Département Génie Mathématique et Modélisation Institut National des Sciences Appliquées de Toulouse 31077 Toulouse cedex 4.
- [LSG<sup>+</sup>04] Z. Lu, D. Szafron, R. Greiner, P. Lu, D.S. Wishart, B. Poulin, J. Anvik, C. Macdonnell, and R. Eisner. Predicting subcellular localization of proteins using machine-learned classifiers. *Bioinformatics*, 20(4) :547–556, March 2004.
- [Qui10] Ross Quinlan. Thyroid disease data set. In *UCI Machine Learning Repository*, 2010.
- [RL02] Ricco Rakotomalala and Stéphane Lallich. Construction d'arbres de décision par optimisation. *Revue d'Intelligence Artificielle*, 16(6) :685–703, 2002.