



**HAL**  
open science

# outil interactif d'aide à la décision basé sur les réseaux bayésiens: application au diagnostic du cancer de la prostate

Mohamed Ali Mahjoub

## ► To cite this version:

Mohamed Ali Mahjoub. outil interactif d'aide à la décision basé sur les réseaux bayésiens: application au diagnostic du cancer de la prostate. 5èmes Journées Francophones sur les Réseaux Bayésiens (JFRB2010), May 2010, Nantes, France. hal-00468266

**HAL Id: hal-00468266**

**<https://hal.science/hal-00468266>**

Submitted on 23 Apr 2010

**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.

---

# **Outil interactif d'aide à la décision basé sur les réseaux bayésiens : application au diagnostic du cancer de la prostate**

**Mohamed Ali Mahjoub<sup>\*,\*\*</sup> - Imed Boudriga<sup>\*\*</sup>**

*\* Institut Préparatoire aux Etudes d'Ingénieurs de Monastir  
Rue Ibn Eljazzar Monastir 5019  
Tunisie  
Medali.mahjoub@ipeim.rnu.tn*

*\*\* Unité de recherche en Signaux, image et documents, Ecole nationale d'ingénieurs de Sousse Eniso Tunisie  
Boudriga\_imed@hotmail.com*

---

*RÉSUMÉ. L'objectif de l'article est de présenter une application d'inférence exacte dans les réseaux bayésiens basée sur l'algorithme JLO (arbre de jonction) dans le but de développer un outil interactif d'aide au diagnostic médical. Il s'agit d'une application développée entièrement en Java implémentant l'algorithme d'inférence exacte JLO avec une interface graphique conviviale et souple d'emploi. Elle permet essentiellement une gestion facile des nœuds du réseau à savoir la création, la suppression, l'édition des probabilités marginales et la manipulation des liens. En outre, le calcul et l'affichage des probabilités d'inférence peut être assuré et visualisé au niveau d'un nœud particulier, au niveau de tout le réseau ou encore au niveau du réseau transformé en arbre. D'autre part un format de sauvegarde compatible matlab de réseau a été mis au point pour assurer une passerelle entre notre application et la bibliothèque BNT de Murphy. Pour la validation de notre application, nous avons conçu avec la coopération de médecins urologues un réseau bayésien modélisant le cancer de la prostate. L'intérêt de l'application réside dans d'une part dans l'apprentissage des paramètres du réseau qui est réalisé par des approches statistiques basée sur l'étude de la base des dossiers des patients disponible, et d'autre part sur l'inférence instantanée dans le réseau permettant de propager les données d'un patient dans le réseau déjà établi permettant d'obtenir des états probabilistes sur le diagnostic du cancer de la prostate. L'inférence dans le réseau nous a permis grâce à une confrontation avec des dossiers déjà traités et classés de confirmer la bonne qualité de l'application développée..*

*MOTS-CLÉS : Réseau bayésien, inférence, Algorithme JLO, apprentissage statistique, cancer de la prostate.*

---

## 1. Introduction

Les Réseaux Bayésiens sont des modèles graphiques qui représentent les relations probabilisées entre un ensemble des variables, ayant pas mal d'avantages par rapport à d'autres techniques, ces réseaux sont devenus pendant ces dernières années des outils de travail indispensables et complémentaires pour représenter et manipuler des connaissances dans un système expert.

Les outils, bibliothèques et logiciels, manipulant les réseaux bayésiens sont en cours d'évolution, de ce fait, ils passent à des interfaces vives et conviviales pour mieux représenter un réseau bayésien. De plus, les algorithmes d'inférence et d'apprentissage utilisés par ces applications sont très variés.

L'objectif de ce travail est de concevoir, puis implémenter, une application d'apprentissage des réseaux bayésiens. Une application qui sera axée dans un premier temps sur deux algorithmes d'inférence exacte (Arbre de jonction & élimination des variables). Comme validation du travail une application générique construite à l'aide d'un urologue, la maladie du cancer de la prostate sera développée. L'essentiel de la solution logicielle est un module en cours d'exploration est le développement de méthodes d'apprentissage automatique de structures de réseaux.

Ce papier est consacré d'une part à la présentation des fonctionnalités de notre application, et d'autre part à l'étude pratique de la modélisation de la maladie de cancer de la prostate.

## 2. Présentation de la solution logicielle

Dans les sections précédentes, nous avons étudié les différentes caractéristiques techniques de l'application et les algorithmes d'inférence utilisés.

La technologie Java est aujourd'hui présente partout : dans les ordinateurs de bureau, sur internet, dans les téléphones portables, les cartes à puce, etc. Un seul langage, une seule syntaxe, qui se décline selon les besoins en J2EE (applications d'entreprise), J2SE (applications de taille plus modeste), J2ME (applications embarquées) et Java Card (cartes à puces). La multitude de composants disponibles permet de répondre à tous les besoins.

Choisir la technologie Java pour le développement d'une application présente de multiples avantages : développement très structuré (orienté objet), portabilité du code sur toutes les plateformes (Windows, Unix, etc.), rapidité d'exécution, robustesse et sécurité, maintenance, etc.

Dans cette section, nous allons nous concentrer sur le principe de l'utilisation de l'application et présenter un exemple générique utilisé dans l'urologie et approuvé par un expert dans le domaine. Cet exemple est un diagnostic médical pour la maladie du cancer de la prostate qui apparaît fréquemment chez les individus de sexes masculins.

### 3. Organigramme de l'application

L'organigramme de l'application peut être défini comme schématisé dans la figure 2.

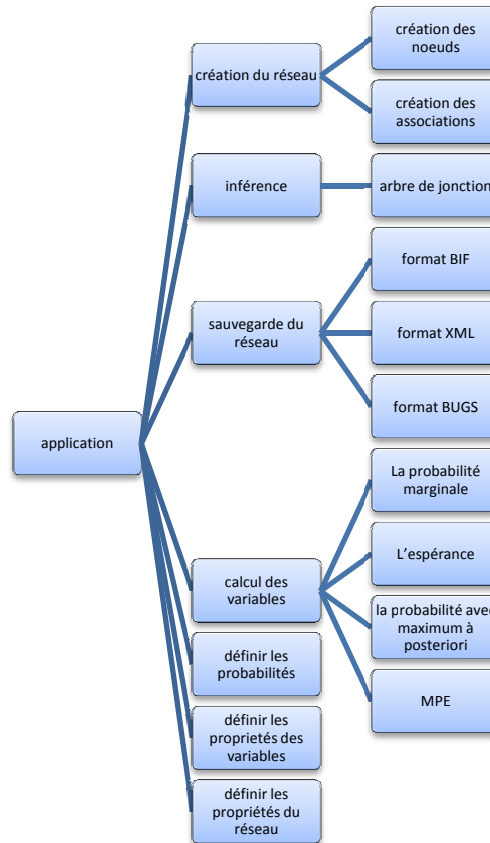


Figure 2 : Organigramme de l'application

L'application nous fournit des options de sauvegarde et d'ouverture de réseau construit en gardant bien sur les évidences ajoutées, aussi elle nous permet de choisir le mode, l'algorithme d'inférence et le format de sauvegarde du réseau. La figure suivante (figure 3) présente l'interface de l'application.

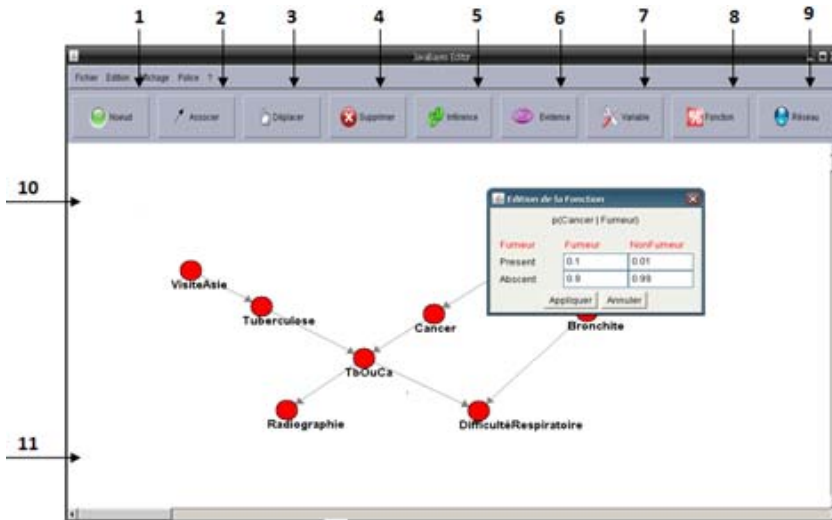


Figure 3 : Interface logicielle

Où :

- 1 : Nouveau nœud.
- 2 : Nouvelle association.
- 3 : Déplacement des nœuds et des associations.
- 4 : Suppression des nœuds et des associations.
- 5 : calculer l'inférence.
- 6 : ajouter une évidence.
- 7 : configurer les propriétés d'un nœud.
- 8 : ajouter les probabilités d'un nœud.
- 9 : configurer les propriétés du réseau construit.
- 10 : zone d'édition de réseau.
- 11: zone d'affichage de résultat et de help.

#### 4. Manipulation de l'application

Lorsqu'on démarre l'application, nous pouvons ajouter un nœud par un simple clic sur le bouton nœud et par la suite, un deuxième clic dans la zone d'édition de réseau, comme nous pouvons ajouter plusieurs nœuds. Ensuite, pour définir les relations entre les nœuds, nous devons cliquer sur le bouton associer et puis nous pourrons ajouter toutes les relations dans le réseau. Il s'agit d'une interface très souple d'emploi, en effet, il est possible d'organiser la disposition des nœuds tout en gardant les relations. C'est une gestion complètement intuitive.

En outre, l'application développée permet la sauvegarde du réseau construit sous deux formats (BIF et XML) qui nous permettent d'avoir une interaction avec d'autres bibliothèques et logiciels manipulant les réseaux bayésiens. D'autre part, des modules d'exportation du réseau sous le format BNT (bibliothèque de Murphy

[17]) de Matlab ainsi qu'une importation d'un réseau par un simple éditeur de texte ont été également conçus et implémentés dans la solution.

## **5. Etude de cas : Aide au diagnostic du cancer de la prostate**

### **5.1. Cancer de la prostate**

L'étude du cancer de la prostate est un domaine de recherche très actif en urologie, et ce pour plusieurs raisons. C'est une maladie qui se prête bien à un dépistage généralisé. Il s'agit d'une affection fréquente, entraînant une morbidité et une mortalité significatives, le plus souvent après une longue période asymptomatique. Des tests de dépistage simples (prostate specific antigen PSA) sont disponibles, permettant une détection durant la période asymptomatique. En outre, le cancer de la prostate représente environ 25 % des cancers masculins. Etant donné que l'incidence croît exponentiellement avec l'âge, la fréquence et la mortalité risquent d'augmenter en raison du vieillissement de la population. D'autre part, la prévention primaire n'est pas possible, car les facteurs de risque principaux sont l'âge et l'hérédité. Il s'agit d'un cancer hormonodépendant : une diminution significative du taux de PSA et un retard dans l'apparition du cancer peuvent être observés sous finastéride (au prix d'effets secondaires et d'un risque accru de cancers peu différenciés). Le rôle de phénomènes inflammatoires est encore mal connu. Des études sur l'efficacité préventive de suppléments en vitamine E et Sélénium sont en cours. Les symptômes initiaux du cancer de la prostate sont le plus souvent secondaires à l'extension extraprostatique et aux métastases, notamment osseuses ce qui complique encore la prévention et les traitements préventifs.

Pour pouvoir valider notre travail dans un domaine persistant et pour prouver le bon fonctionnement de l'application, nous avons procédé à l'étude et à la modélisation d'expertise du cancer de la prostate. Elle se caractérise par certains facteurs spécifiques qui ne pourront être donnés que par un expert du domaine vu le manque de base de données pour établir le réseau automatiquement. Dans le cadre de notre travail, nous avons contacté un professeur urologue à l'hôpital Fattouma Bourguiba de Monastir (Tunisie). Il nous a fourni en premier lieu l'arbre de diagnostic explicatif chez un individu en bon état général puis il nous a défini les différentes probabilités de ce réseau comme schématisé dans les figures 4 et 5. Notre objectif est d'aider les médecins à établir un diagnostic différentiel entre certaines maladies liées à la prostate.

### **5.2. Modèle bayésien**

Comme dans d'autres domaines, on distingue trois méthodes de construction de réseaux bayésiens [1] :

- **Automatique** : par application d'un algorithme d'apprentissage à une base de données. Les algorithmes d'apprentissage peuvent identifier à la fois la structure (le graphe) du modèle et les paramètres (les probabilités conditionnelles).
- **Manuelle** : on interroge les médecins et ajoutent les noeuds, les liens et les probabilités conditionnelles au réseau bayésien sur la base de la connaissance recueillie.
- **Hybride** : dans cette approche, la structure du réseau est décrite avec l'aide des médecins et les probabilités sont obtenues à partir d'une base de données. C'est cette approche que nous avons adopté dans notre application. Ainsi la structure du réseau est déterminée une fois pour toute par le médecin, et donc le réseau est saisi manuellement à l'aide de la souris. Par contre toutes les probabilités sont déterminées d'une manière automatique à partir de la base des cas dont on dispose.

Nous nous sommes donc concentrés avec l'assistance du médecin sur la variable représentant le cancer de la prostate pour identifier les principaux signes, symptômes et facteurs de risque associés. Ce processus a conduit à un modèle comportant 31 liens et 25 nœuds comme le montre la figure 4 : la variable principale (cancer de la prostate), les principaux facteurs de risque et les symptômes, signes, tests et les autres maladies pouvant être provoquées par des complications.

### **5.3. Problèmes liés à l'estimation des paramètres**

Ces problèmes ont d'abord, les bases de données médicales ne contiennent généralement que quelques observations accompagnées du diagnostic final, tandis que la construction d'un réseau bayésien nécessite l'identification d'un grand nombre de variables intermédiaires, afin de satisfaire les hypothèses d'indépendances conditionnelles. Certes, il existe des algorithmes capables de trouver les variables dites cachées, mais il subsiste un problème majeur quand les variables identifiées ne correspondent à aucun concept médical ce qui peut remettre en cause la validité du modèle. D'autre part, beaucoup d'algorithmes d'apprentissage nécessitent que la base de données ne comporte aucune donnée absente. Cependant, dans la pratique, toutes les bases de données médicales sont incomplètes, et la proportion de données manquantes est souvent importante. Pour remédier à tous ces problèmes, nous nous sommes contentés dans un premier temps de collecter la maximum de dossiers (informations), ainsi les données manquantes ont été estimées par le médecin.

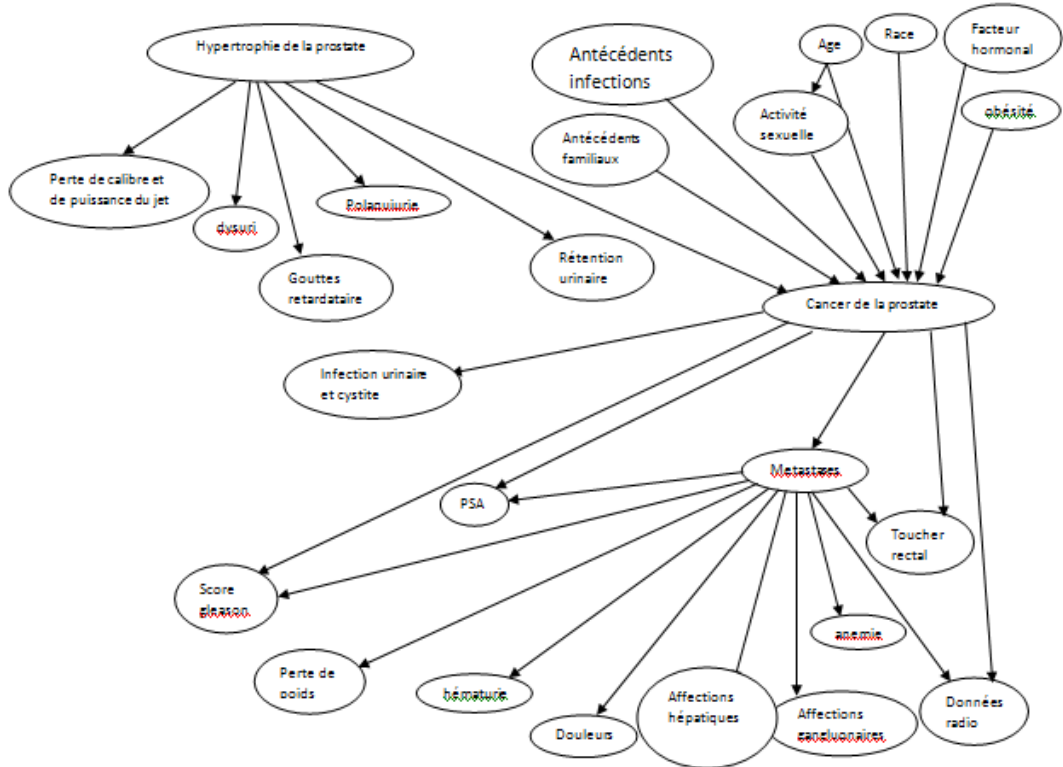


Figure 4 : Réseau modélisant le cancer de la prostate

- AS : activité sexuelle (Hpr : Hyper, Hpo : Hypo, N : normal).
- AI : antécédents infections (Oui, Non).
- HP : hypertrophie de la prostate (Oui, Non).
- Race (C : caucasienne, NC : Non Caucasienne).
- FH : facteur hormonal (+ : Positif, - : Négatif).
- Obésité (Oui, Non).
- AF : antécédents familiaux (P : présent, A : Absent).
- Age ( $\geq 50$ ,  $< 50$ ).

#### 5.4 Apprentissage des paramètres

Nous cherchons ici à estimer les distributions de probabilités à partir de données disponibles. L'estimation de distributions de probabilités, paramétriques ou non, est un sujet très vaste et complexe. Nous décrivons ici les méthodes les plus utilisées dans le cadre des réseaux bayésiens, selon que les données à notre disposition sont complètes ou non.

Dans le cas où toutes les variables sont observées, la méthode la plus simple et la plus utilisée est l'estimation statistique qui consiste à estimer la probabilité d'un



événement par la fréquence d'apparition de l'événement dans la base de données. Cette approche, appelée maximum de vraisemblance (MV), nous donne alors :

$$P(X_i = x_k | \text{parent}(X_i) = c_j) = \theta_{i,j,k} = \frac{N_{i,j,k}}{\sum_k N_{i,j,k}}$$

Sachant que :

- $N_{i,j,k}$  est le nombre d'événements dans la base de données pour lesquels la variable  $X_i$  est dans l'état  $x_k$  et ses parents sont dans la configuration  $c_j$

En outre, dans la pratique, et comme on l'a signalé dans la section précédente les bases de données sont très souvent incomplètes. La méthode d'estimation de paramètres avec des données incomplètes, la plus couramment utilisée, est fondée sur l'algorithme itératif EM (Expectation- Maximisation) proposé par Dempster [DLR77].

Soit :

- $X_v = \{X_v^{(l)}\}_{l=1..N}$  l'ensemble des données observées.
- $\theta^{(t)} = \{\theta_{i,j,k}^{(t)}\}$  les paramètres du réseau bayésien à l'itération  $t$ .

L'algorithme EM s'applique à la recherche des paramètres en répétant jusqu'à convergence les deux étapes Espérance et Maximisation décrites ci-dessous :

- **Espérance** : estimation des  $N_{i,j,k}$  manquants en calculant leur moyenne conditionnellement aux données et aux paramètres courants du réseau :

$$N_{i,j,k}^* = \sum_{l=1}^N p(X_i = x_k | \text{parent}(X_i) = c_j, X_v^{(l)}, \theta^{(t)})$$

- **Maximisation** : en remplaçant les  $N_{i,j,k}$  manquants par leur valeur moyenne calculée précédemment, il devient possible de calculer de nouveaux paramètres  $\theta^{(t+1)}$  par maximum de vraisemblance

$$\theta_{i,j,k}^{(t+1)} = \frac{N_{i,j,k}^*}{\sum_k N_{i,j,k}^*}$$

Calculons à titre d'exemple les probabilités du nœud « Toucher rectal » de la figure 4:

Soient A= toucher rectal , B= cancer de la prostate et C= métastase. Pour le calcul de cette probabilité, nous avons étudié 400 dossiers de patients et nous avons estimé à l'aide du médecin les données manquantes. De ce fait, le calcul de ces probabilités se fait de la manière suivante :

$$\frac{p(A = \text{suspect} | B = \text{présent ET } C = \text{présent})}{\text{le nombre de personne où } B = \text{présent ET } C = \text{présent}} = \frac{99}{100} = 0,99$$

$$p(A = \text{Non suspect} | B = \text{présent ET } C = \text{présent}) = \frac{\text{le nombre de personne où } A = \text{Non suspect}, B = \text{présent ET } C = \text{présent}}{\text{le nombre de personne où } B = \text{présent ET } C = \text{présent}} = \frac{1}{100} = 0.01$$

$$p(A = \text{suspect} | B = \text{présent ET } C = \text{absent}) = \frac{\text{le nombre de personne où } A = \text{suspect}, B = \text{présent ET } C = \text{absent}}{\text{le nombre de personne où } B = \text{présent ET } C = \text{absent}} = \frac{98}{100} = 0.98$$

$$p(A = \text{Non suspect} | B = \text{absent ET } C = \text{absent}) = \frac{\text{le nombre de personne où } A = \text{Non suspect}, B = \text{absent ET } C = \text{absent}}{\text{le nombre de personne où } B = \text{absent ET } C = \text{absent}} = \frac{2}{100} = 0.02$$

$$p(A = \text{suspect} | B = \text{absent ET } C = \text{présent}) = \frac{\text{le nombre de personne où } A = \text{suspect}, B = \text{absent ET } C = \text{présent}}{\text{le nombre de personne où } B = \text{absent ET } C = \text{présent}} = \frac{1}{100} = 0.01$$

$$p(A = \text{Non suspect} | B = \text{absent ET } C = \text{présent}) = \frac{\text{le nombre de personne où } A = \text{Non suspect}, B = \text{absent ET } C = \text{présent}}{\text{le nombre de personne où } B = \text{absent ET } C = \text{présent}} = \frac{99}{100} = 0.99$$

$$p(A = \text{suspect} | B = \text{absent ET } C = \text{absent}) = \frac{\text{le nombre de personne où } A = \text{suspect}, B = \text{absent ET } C = \text{absent}}{\text{le nombre de personne où } B = \text{absent ET } C = \text{absent}} = \frac{6}{100} = 0.06$$

$$p(A = \text{Non suspect} | B = \text{absent ET } C = \text{absent}) = \frac{\text{le nombre de personne où } A = \text{Non suspect}, B = \text{absent ET } C = \text{absent}}{\text{le nombre de personne où } B = \text{absent ET } C = \text{absent}} = \frac{94}{100} = 0.94$$

Nous pouvons schématiser ces résultats sous forme d'un tableau.

Toucher rectal :

|                |                       |         |         |         |        |
|----------------|-----------------------|---------|---------|---------|--------|
|                | Cancer de la prostate | Présent | Présent | Absent  | Absent |
|                | Métastase             | Présent | Absent  | Présent | Absent |
| Toucher rectal | Suspect               | 0.99    | 0.98    | 0.01    | 0.06   |
|                | Non suspect           | 0.01    | 0.02    | 0.99    | 0.94   |

Par la suite, notre expert a défini les différentes probabilités de ce réseau traitées sous forme de tableaux comme illustré dans les 3 exemples suivants :

**Antécédents infections :**

|                        |     |
|------------------------|-----|
| Antécédents infections | P   |
| Oui                    | 0.5 |
| Non                    | 0.5 |

**Age :**

|     |     |
|-----|-----|
| Age | P   |
| <50 | 0.7 |
| ≥50 | 0.3 |

**Race :**

|                 |     |
|-----------------|-----|
| Race            | P   |
| Caucasienne     | 0.7 |
| Non caucasienne | 0.3 |

**Facteur hormonal :**

|                  |     |
|------------------|-----|
| Facteur hormonal |     |
| Positif          | 0.4 |
| Négatif          | 0.6 |

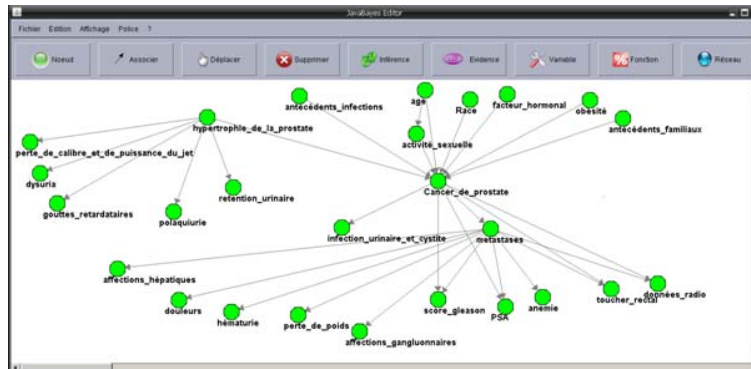


Figure 5 : interface logicielle modélisant le cancer de la prostate

Pour mesurer quantitativement l’efficacité de notre application, nous avons procédé à la modélisation de 320 dossiers de patients déjà classés. Comme le montre le tableau suivant 82% des dossiers ont été bien classés et ce par comparaison avec les dossiers archivés. En outre ce taux est comparable avec ceux obtenus avec d’autres applications commerciales manipulant les réseaux bayésiens.

| Solutions                            | <i>Notre application</i> | Elvira | BayesiaLab | Netica |
|--------------------------------------|--------------------------|--------|------------|--------|
| <b>Taux de dossiers bien classés</b> | 82%                      | 83%    | 80%        | 85%    |

Tableau 1 : Taux de dossiers bien classés

## 6. Conclusion

Dans ce papier nous avons présenté notre application logicielle basée sur les réseaux bayésiens et les algorithmes d’inférence associés dans un but de modélisation de connaissances et d’aide à la décision. Il s’agit d’une application qui peut être utilisée d’une manière intuitive, en effet elle offre une interface graphique permettant la modélisation de connaissances à l’aide d’un réseau bayésien structuré en nœuds (connaissance quantitative) et liens entre les nœuds (connaissances qualitative).

Si la structure du réseau est déterminée manuellement à l’aide d’un expert, les probabilités conditionnelles sont calculées d’une manière automatique par un apprentissage dit statistique. Une fois le réseau est établi, un module d’inférence peut être exécuté après avoir injecté les observations sur le modèle permettant de propager ses évidences dans le réseau, il s’agit d’un processus qui va converger sur

des résultats concrets. Des valeurs de probabilités qui vont guider le médecin dans sa définition de diagnostic.

Pour une meilleure validation, nous avons conçu notre travail sur une étude pratique connu chez les urologues qui concerne l'aide au diagnostic du cancer de la prostate. Dans cet exemple, la structure du réseau a été définie manuellement tandis que les différentes probabilités conditionnelles de tous les nœuds ont été déterminées automatiquement. C'est ainsi qu'on a pu confronter les résultats obtenus par notre application avec ceux déjà établis par les médecins concernant 320 patients. Le calcul des paramètres de rappel et de précision ont permis de valider d'une part la structuration du réseau bayésien conçu et d'autre part la qualité des résultats (diagnostic médical). En effet, dans 80% des dossiers traités on a eu un diagnostic correct.

Les prochaines mises en œuvre de ces idées permettront à notre application de gagner le pari de la qualité dans les applications des réseaux bayésiens tout en relevant le défi enchanté pour ceci. En effet, nous sommes en plein étude de méthodes d'apprentissage automatique de structures de réseaux à partir d'observations (données patients).

### **Remerciements**

*Nous remercions le Docteur Bahri Anis, urologue à l'hôpital universitaire Fattouma Bourguiba de Monastir pour son aide précieuse dans la collecte des informations et le développement du modèle de diagnostic du cancer de la prostate.*

### **Bibliographie**

- [1] Patrick Naïm, Pierre-Henri Wuillemin, Philippe Leray, Oliver Pourret et Anna Becker, "Réseaux Bayésiens", 3<sup>ème</sup> Edition Eyrolles.
- [2] Camille Séka Kotchi, Véronique Delcroix et Sylvain Piechowiak, "Etude de la performance des algorithmes d'inférence dans les réseaux bayésiens", Laboratoire d'Automatique, de Mécanique et d'Informatique industrielles et Humaines, Université de Valenciennes et du Hainaut Cambrésis.
- [3] Philippe Leray, " Réseaux Bayésiens : Définition - Inférence", INSA Rouen – Département ASI – Laboratoire PSI.
- [4] Philippe Leray, " Réseaux Bayésiens : Apprentissage de la Structure", INSA Rouen – Département ASI – Laboratoire PSI.
- [5] Martin Pelikan and Kumara Sastry, "Fitness Inheritance in the Bayesian Optimization Algorithm", Février 2004.
- [6] Philippe Leray and Olivier Franois, "Bayesian Network Structural Learning and Incomplete Data", INSA Rouen - PSI Laboratory.
- [7] John M. Winn, " Variational Message Passing and its Applications", Thèse de doctorat Janvier 2004.
- [8] Tomi Silander and Petri Myllymäki, "A Simple Approach for Finding the Globally Optimal Bayesian Network Structure", Complex Systems Computation Research Group (CoSCo).

- [9] Michael Feeley, Norman Hutchinson et Suprio Ray, “De l’identification de structure de réseaux bayésiens à la reconnaissance de formes à partir d’informations complètes ou incomplètes”, Thèse de doctorat Novembre 2006.
- [10] Linda Smail, “Algorithmique pour les Réseaux Bayésiens et leurs Extensions”, these de doctorat Avril 2004.
- [11] Philippe Leray et Sourour Ammar, “ Étude Comparative des outils manipulant les réseaux bayésiens Délivrable n°7”, Université de Nantes 2006.
- [12] Philippe Leray, “Réseaux bayésiens: apprentissage et modélisation de systèmes complexes”, Maître de Conférences Novembre 2006.
- [13] Philippe Leray, “Apprentissage”, INSA Rouen – Département ASI – Laboratoire PSI.
- [14] Edward Jay Kreps, “Learning Structure in Directed Graphical Models”, these de doctorat Juin 2004.
- [15] Markus Harva, Tapani Raiko, Antti Honkela, Harri Valpola, and Juha Karhunen, “ Bayes Blocks: An Implementation of the Variational Bayesian Building Blocks Framework”, Neural Networks Research Centre.
- [16] Michel Gagnon,, “ Réseaux bayésiens”, cours 7 avril 2005.
- [17] Kevin Murphy, “Software for Graphical models: a review”, Décembre 2008.
- [18] Benoît Lavoie, “ Apprentissage bayésien”, Doctorat avril 2006.
- [19] D.M. Chickering (1996): Learning bayesian network is NP-complete, In Learning from data: artificial intelligence and statistics V, pages 121-130. Springer-Verlag, New York
- [20] G.F.Cooper. Computational complexity of probabilistic inference using bayesian belief networks. Artificial Intelligence, vol.42, p.393-405, 1990.
- [21] A.Dempster, N.Laird, et D.Rubin. Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistical Society, pages 1-38, 1977.
- [22] O.Francois. De l’identification de structure de réseaux bayésiens à la reconnaissance de formes à partir d’informations complètes ou incomplètes. Thèse de doctorat, Institut National des Sciences Appliquées de Rouen, 2006.