

Resumen:

Introducción: Las redes Bayesianas son un algoritmo de Inteligencia Artificial utilizado en todo campo en donde la incerteza es moneda corriente. El modelo combina teoría de grafos, topología y probabilidades permitiendo armar representaciones gráficas que representan determinaciones causales y probabilidades relativas.

Objetivo: Se intentarán enunciar los elementos básicos requeridos para el armado de una red destinada a la resolución de problemas diagnósticos a utilizar en el campo de la psiquiatría y la psicopatología. El modelo debe emular en alguna medida el modo de realización del diagnóstico por un experto.

Conclusiones: El campo médico-psiquiátrico es sin duda uno de los más prometedores para la utilización de estas redes. Aún desde el armado básico de una red-prototipo puede apreciarse la utilidad que esta herramienta puede aportar a la disciplina. Junto con el uso diagnóstico existen algoritmos que permiten “aprender” de los datos y de esa manera extraer relaciones que en la confección previa de la red pueden no haberse reconocido aún estando presente en los datos con que se la ha cargado.

Introduction: Bayesian Networks are an Artificial Intelligence algorithm used in fields in where the uncertainty is current currency. The model combines theory of graphs, topology and probabilities allowing of arming graphical representations they represent causal determinations and relative probabilities.

Objective: We will tried to enunciate the required basic elements for make a Bayes network useful for resolution of diagnosis problems to be used in psychiatry and psychopathology. The model must emulate the way an expert can accomplish some diagnosis.

Conclusions: The medical-psychiatric is without a doubt one of most promising field for the use of these Bayesian networks. Still from the basic one of a network-prototype the utility can be appraised to think how much this tool can contribute to the discipline. Along with the diagnosis use also exist algorithms that allow "to learn" from the data and extract relations that in the previous preparation of the network can not have recognized, still being present in the data loaded to it.

Palabras claves:

Diagnóstico psiquiátrico, Diagnóstico psicopatológico, Bayesian Net, Red Bayesiana, Psychiatric diagnose, Psychopathological diagnose.

Bayesian Networks en diagnóstico psiquiátrico o psicopatológico.

Lic. Mariano Acciardi
mariano[at]acciardi.com.ar

Introducción:

Sin lugar a dudas, la salud y el diagnóstico son frecuentes campos de aplicación de estas redes. En un dominio en donde la mayor parte del conocimiento que una red expresa es manejado implícitamente por los más experimentados especialistas. Estos dominios han sido históricamente los principales objetivos de los sistemas expertos. Originariamente la conexión entre estos campos de aplicación y las Redes Bayesianas ha sido establecida por investigadores como David Heckerman y Eric Horvitz. Cabe agregar que estos investigadores son dos de los principales impulsores del desarrollo en investigación que han tenido dichas redes en los últimos años. Asimismo Pearl acuña el concepto de “Bayesian Belief Network” (Pearl J. 1986) abriendo las puertas a su aplicación en los más diversos campos en donde la incerteza sea moneda corriente, debiendo lidiar con ella para toda resolución de problemas ulterior. Específicamente en diagnóstico médico las investigaciones de Beinlich (Beinlich et al, 1989) son un aporte imprescindible, así como también los avances que los algoritmos de aprendizaje y refinamiento de los datos de la red que ha propuesto Spiegelhalter and Lauritzen (Spiegelhalter and Lauritzen, 1993) para el procesamiento paralelo de una red demasiado grande como para ser procesada secuencialmente, lo que suele ser el caso del diagnóstico de ciertas enfermedades que pueden planteadas en términos de un sistema complejo (García R. 1984) en razón de la complejidad de las determinaciones que subyacen a la presencia de una u otra afección. Estos son solo algunos de los muchos investigadores que en este momento se encuentran estudiando los diversos modos de utilización de estas redes para el diagnóstico médico.

Como antecedentes importantes a mencionar (aunque no los únicos) en este campo se encuentran funcionando de manera exitosa: es el sistema PATHFINDER, disponible hoy comercialmente como sistema experto de decisión para patólogos quirúrgicos; el sistema MUNIN, constituido por una red múltiplemente conectada para el diagnóstico de desórdenes neuromusculares; o el sistema PROMEDAS ampliamente utilizado para diagnóstico en endocrinología.

Redes Bayesianas

Una Red Bayesiana (BN) es una estructura gráfica que permite representar el modo en que un humano razona en ciertos campos que se caracterizan por estar plagados de incerteza. El hecho del que el ser humano sea capaz de razonar e inclusive hacerlo correctamente aún sin disponer de todos los elementos que serían necesarios para realizar una inferencia o tomar una decisión correcta, es lo que motiva las investigaciones de los últimos años en inteligencia artificial, en el diseño y desarrollo de modelos expertos. Tal como se menciona más arriba, en ciertos campos muy particulares, hace años que estas redes están brindando asombrosos resultados consecuencia de la capacidad cada vez más potenciada de las máquinas para emular el razonamiento humano.

Una red Bayesiana se compone básicamente de “Nodos”, que representan variables generalmente discretas, cuya fuerza de relación entre ellas se cuantifica mediante una distribución de probabilidades condicionales que determinan el valor final de aquellos nodos que no se han cargado como evidencia. La relación entre dichas variables se establece mediante “arcos” que representan una determinación causal entre nodos.

La determinación de los nodos como representación de información relevante es realizada por un experto. Los valores que pueden adoptar dichos nodos deben ser discretos, aunque el número de valores que puede adoptar prácticamente no tiene límites teóricos relevantes.

Es importante destacar aquí una mención epistemológica: La causalidad en términos de “Explicación Causal”, depende necesariamente de una teoría. Es decir que una “Explicación Causal” en ciencia implica poder deducir como necesidad lógica un hecho B a partir de un hecho A. Es decir no se trata de una simple secuencia de fenómenos sino que es preciso que exista una explicación causal en los términos de un teoría de por qué el hecho B es derivado del hecho A por lo que puede considerarse que A es causa de B. Sin embargo, como se menciona sobre el final del texto, las redes Bayesianas poseen algoritmos capaces de desentrañar “determinaciones causales” por fuera de toda teoría, o por lo menos no totalmente explicables en el marco de la teoría que permitió el armado de la red. Esta es una característica tanto peligrosa como emocionante, ya que por la simple consideración de una cantidad de datos suficiente, es posible extraer de ellos tendencias o determinaciones no pensadas previamente.

Los modelos que se aplican realmente en la práctica pueden ser: a) muy complejos y tomar la forma de varias redes interconectadas; o bien, según la situación a modelar, b) muy simples tomando la forma de una red bien sencilla, caso bastante común en el uso de estas redes como modos de clasificar enfermedades en la que un “Nodo” puede representar en sus “estados” diversas patologías.

Aspectos cualitativos (estructura) y cuantitativos (probabilidades)

Una BN combina conjuntamente topología y probabilidades, junto a un modo de relacionar estas probabilidades entre sí, mediante un algoritmo que se conoce como la fórmula de Bayes. Es por eso que es crucial en una red Bayesiana la determinación de cuales serán los “nodos” o “variables” a tener en cuenta, como así también la forma de relacionarse entre ellos (arcos), es decir las determinaciones causales que los unen. Esto es la “topología” de la red.

En la definición de esta topología deben tenerse en cuenta diversos criterios:

- 1) Dos nodos deben estar conectados entre sí si uno afecta o causa el otro.
- 2) No es posible establecer lazos cíclicos entre nodos (haría imposible el cálculo de las probabilidades relativas, ingresando el cálculo en un bucle infinito)

A partir de estos simples criterios que acuerdan completamente con el sentido común que uno daría a un elemento conectado por una flecha con otro, es posible complejizar esta red de manera de poder representar el modo de razonamiento seguido por el experto para sacar una conclusión partiendo de cierta información básica, y siguiendo las consecuencias del agregado progresivo de nuevas evidencias. Esto es claramente lo mismo que realiza cualquier médico cuando utiliza su formación previa (Universidad) para establecer lazos causales entre síntomas y diversas enfermedades. A esa formación básica, se suman evidencias que en cada caso en particular van ajustando cada vez más las probabilidades de que haber contraído tal o cual afección. Podríamos decir que una BN no es más que una representación casi natural de la forma de razonamiento llevada a cabo por un humano en determinadas situaciones o campos.

La estructura de una red de diagnóstico puede separarse teóricamente en tres partes: 1) Los nodos de distintos cuadros o enfermedades interconectados entre sí -si corresponde-; 2) Los nodos que representan el conocimiento base previo existente del campo “antecedente” teórico o epidemiológico; y ,en el otro extremo, 3) los nodos que representan los resultados de tal cuadro o enfermedad. (incluyendo tanto resultados de test como los síntomas producidos por la enfermedad).

Dentro del conocimiento de base se incluyen todos aquellos factores que pueden ser considerados como factores de predisposición y la información previa de distribución epidemiológica local para los cuadros o enfermedades.

En estos nodos que forman parte del conocimiento de base (background knowledge) suele estar representado todo el conocimiento de una disciplina en un momento dado que forma lo que en BN se conoce como “probabilidades previas” (Prior probabilities).

Los nodos de “resultados” constituyen las probabilidades a ser cargadas/calculadas en referencia a un caso particular bajo la forma de lo que se denomina “evidencia”.

Razonamiento probabilístico:

Una red bayesiana puede emular diferentes tipos de razonamiento que implican diversos modos de recorrer la red. El razonamiento que se utiliza para problemas diagnósticos, recorre la red partiendo de los resultados (que recordemos incluyen tanto resultados de los tests como así también los síntomas fundamentales.) y dirigiéndose al cuadro o “causa” de esos resultados o “evidencias”. Esto implica un recorrido en sentido inverso al de los arcos que representan la dependencia causal.

Es decir, las enfermedades son la “causa” de tal síntoma o tal resultado patológico de un test o estudio. Si aparece el resultado, es de esperar que exista previamente su causa. Esta es la forma en que una red bayesiana razona para realizar un diagnóstico. Cuanto mayor sea la evidencia, si las probabilidades relativas están correctamente cargadas para relacionar estos nodos, mayor será la aproximación del resultado diagnóstico.

El agregado de evidencia permite realizar otro tipo de razonamiento también utilizado para resolver problemas diagnósticos que es el “predictivo” en el que progresivamente se agrega nueva información que genera nuevas “creencias” (belief) respecto de los efectos que cada evidencia cargada produce.

Entrenamiento de una red Bayesiana:

De acuerdo a lo antedicho, las BN pueden utilizarse para calcular nuevas creencias o “Posterior Belief” en base a información que progresivamente se agrega a los nodos de la red (evidencias).

Una evidencia es un valor determinado que toma un Nodo en cierto momento. La evidencia puede tener un valor certero, por ejemplo: “Tiene alucinaciones=True”. O bien un valor no certero, en donde comienza a jugar lo que en BN se denomina: “porcentaje de creencia”: “Hay un 30% de probabilidades que haya tomado alcohol recientemente”. Es precisamente en este punto en donde pueden utilizarse los conocimientos previos epidemiológicos en todos aquellos nodos en donde no se disponga de información suficiente como para considerarlos “evidencia”. Esto produciría una especie de generalización de los elementos no ciertos de acuerdo a una preponderancia epidemiológica independiente de la existencia de una afección. El manejo de estos “agujeros de incerteza” en base a un cuadro de probabilidades ajustado localmente es uno de los elementos que hace de la BN un instrumento tan poderoso. Ya que aún en donde no hay información, el conocimiento de la disciplina puede permitir otorgar un cierto valor relativo a ese factor que ni siquiera es conocido para el caso particular (no hay evidencia de él)

En una BN para diagnóstico, es imprescindible contar con este conocimiento previo a la recopilación de toda evidencia. Este tipo de conocimiento previo se denomina “Prior Belief” o “Creencia previa” y suele generarse a partir de estudios epidemiológicos de campo o bien operacionalizando factores teóricos relevantes extraídos de la bibliografía existente. Por ejemplo: “De acuerdo a estudios previos, la probabilidad de que un paciente que viene a consultar sea un paranoico es del 2%”; “En esta comunidad específica la probabilidad de que un sujeto consuma marihuana es del 30 %”

A medida de que a esta información de base se van agregando evidencias del caso concreto, las probabilidades relativas correspondientes a los diferentes cuadros se va modificando. Este tipo de lazos causales se establece a partir del conocimiento previo del especialista para operacionalizar las evidencias a rastrear, su relación causal con el cuadro, y el porcentaje de condicionamiento de cada nodo sobre el siguiente se realiza en dirección a los arcos establecidos.

Un modo de utilización de estas redes que erizaría a más de un teórico es su capacidad de, mediante ciertos algoritmos particulares, “aprender” de datos aparentemente sin ningún formato y sin realizar ninguna asunción previa (Prior Belief). Esta es una poderosa característica digna de una película de ciencia ficción, en donde sin otorgar probabilidades diferenciales previas a cada uno de los nodos, y alimentando la red con una suficiente cantidad de datos de caso, con el único formato de la red causal previa, se han definido algoritmos que son capaces de aislar y detectar tendencias allí en donde en una primera aproximación a los mismos solo había caos e indeterminación. Pero este tema será objeto de un trabajo posterior.

Conclusiones:

A partir de los datos correspondientes a la “creencia previa”, ya es posible, aún con anterioridad a toda evidencia, realizar el cálculo de las “Probabilidades condicionales” de todos los nodos de la red. Este “conocimiento”, se basa únicamente en las probabilidades calculadas de acuerdo al conocimiento previo y a las relaciones causales establecidas por el especialista en el armado de la red.

Luego, un cambio en las probabilidades de alguno de los nodos correspondientes a “evidencias”, el recálculo de la red relativizará, a la manera de un juego de ajedrez, todos los valores de los restantes nodos.

Para finalizar ilustrando lo antedicho, se tomará el ejemplo antedicho complejizándolo un poco para entender el modo en que las probabilidades relativas se actualizan ante el agregado de nueva evidencia: “Tomo alcohol recientemente” le cambiamos el valor “epidemiológico” local, es decir de la comunidad a la que pertenece este individuo, supongamos que lo subimos de 30% a 100% ya que tenemos la “evidencia” de que ello ha ocurrido, desde luego se relativizaran los valores correspondientes a

Paranoia (disminuyendo su porcentaje) y aumentará el valor de “Síndrome alucinatorio alcohólico). Si a ello le agregamos el valor 100% de “Tiene alucinaciones visuales”, sin duda en el próximo recálculo aumentará el valor del nodo “Síndrome alcohólico”; si a esto agregamos el valor de una “Temperatura elevada” y los resultados de un test de vitamina B, entonces aumentarán mucho más las probabilidades calculadas para el cuadro “Delirium tremens” y disminuirá el valor correspondiente al nodo “Paranoia”. Como así también desde luego valor calculado del nodo “Esquizofrenia”. Si por ejemplo en lugar de aumentar la “Temperatura” y el “Síndrome deficitario de Vitamina B”, las cambiamos a 0, entonces aumentarán las probabilidades calculadas para el cuadro “Esquizofrenia” disminuyendo las correspondientes al síndrome “Delirium Tremens”.

Bibliografía:

Berger Marcel, Les mathématiques aujourd'hui, Pour la Science, Paris, 1984.

Cowell Robert G y otros, Probabilistic Networks and Expert Systems, Springer, New York, 1999.

Glymour C., "The Mind's Arrows: Bayes Nets and Graphical Causal Models in Psychology", MIT Press, 2001.

Jensen FV. "Bayesian Networks and Decision Graphs". Springer. 2001

Heckerman David; A tutorial on learning with Bayesian Networks; 1996

Korb K, Nicholson A , "Bayesian Artificial Intelligence", CRC Press 2004.

Naïm Patrick y otros, Réseaux bayésiens, Eyrollles, Paris, 2007.