

# On the Road to a Method for Forecasting Performance in Informal Structure Domains Tasks

Macario Ruiz-Grijalva

Maestrante del programa de Maestría en Cómputo aplicado DEyC, IIT, UACJ

150625@alumnos.uacj.mx

Jorge Rodas-Osollo

Laboratorio Nacional en Tecnología de la Información (sede UACJ)

jorge.rodas@uacj.mx

Karla Olmos-Sánchez

Laboratorio Nacional en Tecnología de la Información (sede UACJ)

kolmos@uacj.mx

## ABSTRACT

How to make an assessment of the skills necessary for a certain person to successfully perform a task or acquire knowledge is a long-standing subject of current interest for which some theories, approaches and models have been proposed that aim to achieve it in different areas. The difficulty of carrying out a meaningful evaluation focuses on the large amount of uncertainty and tacit knowledge that are present during the valuation process. Therefore, as a first step, it is necessary to structure, as far as possible, what is known of the skills to later carry out a validation of whether knowledge would be acquired or not. The objective of this article is to communicate a series of ideas that in the end allowed to predict the performance in tasks or specific knowledge pertaining to an Informal Structure Domain (DEI) through the assessment of skills. Of the ideas reviewed such as: Bayesian Networks (RB), Rough Set, Probabilistic Models, Cognitive Approach among others, it is believed that RB are a promising option that would allow a representation of the scope and prognostic knowledge, from skills, of the achievement or the acquisition of some knowledge.

## RESUMEN

**Título de artículo: En Camino a un Método para Pronosticar el Desempeño en Tareas de Dominios de Estructura Informal**

El cómo realizar una valoración de las habilidades necesarias para que cierta persona logre desempeñar satisfactoriamente alguna tarea o bien adquirir un conocimiento, es un tema añejo y de interés actual para el cual, se han planteado algunas teorías, enfoques y modelos que pretenden lograrlo en diferentes ámbitos. La dificultad de realizar una evaluación significativa se centra en la gran cantidad de incertidumbre y conocimiento tácito que están presentes durante el proceso de valoración. Por tanto, como un primer paso, es necesario dar estructura, en la medida de lo posible, a lo que se conoce de las habilidades para posteriormente llevar a cabo una validación de si el conocimiento se adquiriría o no.

El objetivo del presente artículo es comunicar una serie de ideas que a la postre permitieran pronosticar el desempeño en tareas o conocimiento específico pertenecientes a un Dominio de Estructura Informal (DEI) por medio de la valoración de habilidades.

De las ideas revisadas como son: Redes Bayesianas (RB), Rough Set, Modelos Probabilísticos, el Enfoque Cognitivo entre otros, se cree que las RB son una opción prometedora que permitiría una representación del conocimiento del ámbito y el pronóstico, a partir de habilidades, del logro o de la adquisición de algún conocimiento.

## CCS Concepts

• Education—Learning Management Systems

## Palabras Clave

**Valoración de habilidades; Redes Bayesianas; Dominios de Estructura Informal.**

## 1. INTRODUCCIÓN

De acuerdo con el Diccionario de la lengua española, evaluar significa estimar, apreciar, calcular el valor de algo; por consiguiente, evaluar el saber consiste en la estimación de los conocimientos, aptitudes y rendimiento de alguna persona respecto a un objetivo de evaluación.

De acuerdo a [9] la evaluación tradicional se apoya en los exámenes, pruebas y otros instrumentos basados mayormente en la medición cuantitativa. La evaluación cualitativa, por otra parte, se orienta a las siguientes funciones: 1) diagnosticar, si se sitúa en el desarrollo de la estrategia de enseñanza y aprendizaje; 2) adaptar, cuando facilita información para que se puedan realizar adecuaciones y acomodar los diferentes métodos a los distintos ritmos de aprendizaje; y 3) controlar, cuando certifica la cantidad y calidad de las competencias adquiridas. Por lo tanto, en un contexto más amplio se puede definir el proceso de evaluación, como la estimación de conocimientos que realiza un elemento evaluador (modelo, método, persona, etc.) a un elemento evaluado desde alguno o ambos enfoques cuantitativo o cualitativo y respecto a un objetivo de evaluación.

La evaluación en el ámbito educativo es una tarea fundamental, ya que permite medir el grado de avance que se tiene respecto algún dominio, cabe mencionar que la evaluación intenta reconocer el grado de conocimiento que se tiene sobre algún área para emitir un resultado que facilite información para entender qué contenidos se han asimilado, cómo se procesan y cuál es el grado de eficiencia en su uso sobre los contextos de evaluación propuestos [9].

En un dominio donde la estructura es formal las tareas a evaluar pueden dividirse fácilmente en sub tareas más simples, permitiendo que la validación de conocimientos se convierta en una actividad trazable, es decir, una solución única que se pueda representar mediante una serie de pasos bien definidos. Sin embargo, este proceso es diferente en los Dominios de Estructura Informal (DEI): Dominios con un alto grado de informalidad, donde el conocimiento es informalmente establecido, parcialmente completo, implícitamente asumido, tácito y no estructurado [2]. Además, el conocimiento está disperso entre los especialistas del dominio. Por ejemplo, la evaluación de programación de computadoras, ética, leyes, etc.

El objetivo del presente artículo es dar a conocer el interés que se tiene por desarrollar un método que permita elegir de una extensa cantidad de opciones las más adecuadas para crear una estructura de conocimientos que facilite la verificación de habilidades en algún sujeto al desempeñar una actividad. Para el que se propone un enfoque orientado a problemas donde se puedan representar y

vincular las habilidades necesarias para llevar a cabo una tarea de forma satisfactoria.

Se desea generar un pronóstico de habilidades en base a una estructura que represente el conocimiento y verifique las habilidades por medio de evidencia parcial desde algunas o pocas interacciones, las cuales pueden ser preguntas, aplicación de conocimiento a problemas, etc. Algunas de las áreas de interés de la aplicación del método son: la evaluación en el sector educativo, en el ámbito industrial para la evaluación de candidatos durante el proceso de reclutamiento y selección, en el análisis de riesgos para las Pymes entre otras.

Existen diversos obstáculos cuando se desea evaluar conocimiento en DEI, debido a la gran cantidad de informalidad e incertidumbre. En este caso se resaltan dos en particular: 1) la gran cantidad de conocimiento tácito que se debe educir y 2) la necesidad de sintetizar el conocimiento de los especialistas.

Los modelos probabilísticos para la evaluación de habilidades juegan un papel clave en los avances de los entornos de aprendizaje. El modelado y la validación de habilidades de sujetos nos permiten evaluar de manera eficaz el conocimiento en dominios de estructura mayormente formal. Básicamente lo que se busca es inferir a partir de acciones y respuestas que habilidades están presentes para así evaluar el conocimiento de algún tema [1].

El artículo está organizado de la siguiente manera: la primera sección incluye una breve contextualización respecto a la caracterización que se decide adoptar de los DEI. En la segunda sección se describe el paradigma de RB y su aporte en la Ingeniería del Conocimiento. Se mencionan además algunas de las características de las RB y la racionalidad de su empleo en el presente proyecto. Por último, se explica la interpretación del conocimiento en la RB. La cuarta sección expone la percepción general del método propuesto. Finalmente, la quinta sección incluye la discusión generada a partir de esta investigación.

## 2. DOMINIOS DE ESTRUCTURA INFORMAL (DEI)

De acuerdo a RAE formal significa definido, ordenado y metódico; esto no necesariamente implica que sea lógico o probado. Informal es, por lo tanto, una categoría residual. Si un dominio es formal, los especialistas de dominio pueden construir una estructura formal para resolver sus problemas. Sin embargo, si un dominio es informal, esto no significa que los especialistas de dominio no puedan construir una estructura; definitivamente lo hacen al proponer soluciones.

Existen múltiples y variadas opiniones respecto al concepto de los dominios y su caracterización de acuerdo al grado de estructura que poseen. En el presente trabajo se consideró la caracterización de los DEI propuesta por Olmos & Rodas [2], donde se menciona que:

- Los conceptos y sus relaciones tienden a ser definidos basado en un consenso, en lugar de formalmente.
- Los problemas en estos dominios tienen las siguientes características: 1) Las soluciones son diversas, consensuales e inverificables, y 2) No existen algoritmos para llegar a soluciones.
- Para obtener la solución de un problema los especialistas generalmente construyen estructuras parciales con el conocimiento explícito. Sin embargo, siempre se requieren grandes cantidades de conocimiento tácito para obtener una solución aceptable.

Los DEI tienen una estructura o contenido que hace difícil evaluar el conocimiento. No obstante, se debe evitar confundir con los Dominios Complejos. Los dominios complejos contienen una gran

cantidad de elementos del conocimiento y relaciones entre ellos, pero pueden ser bien definidos. Por ejemplo, en el campo de la Geografía gestionar los países y sus capitales (miles de pares). Sin embargo, muchos DEI sí son considerados dominios complejos.

## 3. ENFOQUE PROBABILÍSTICO

### 3.1 Redes Bayesianas

Las Redes Bayesianas, (también conocidas como redes causales probabilísticas, redes causales, sistemas expertos bayesianos, redes de creencia, sistemas expertos estadísticas o diagramas de influencia) son herramientas estadísticas que representan un conjunto de incertidumbres asociadas sobre la base de las relaciones de independencia condicional que se establecen entre ellas. Siguiendo a Charniak [5], una red bayesiana es un conjunto de variables, una estructura gráfica que conecta estas variables y un conjunto de distribuciones de probabilidad condicional (ver figura 1), que codifica incertidumbre asociada a cada variable por medio de probabilidades y, empleando el teorema de Bayes, esta incertidumbre es susceptible de ser modificada con base en observaciones (o evidencias) sobre el modelo.

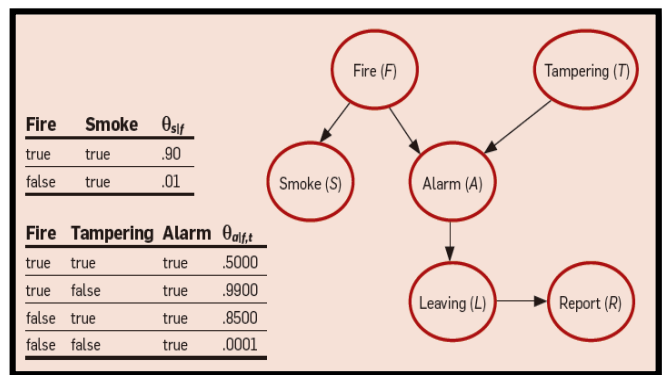


Figura 1. Red bayesiana con algunas de sus tablas de probabilidad condicional. Tomada de "What are Bayesian Networks and why are their Applications Growing Across all Fields?" [6].

Aunque se podría establecer una distinción más específica de los elementos que componen una red bayesiana, se suele decir que una red bayesiana tiene dos dimensiones: una cualitativa y otra cuantitativa [8].

### 3.2 Dimensión Cualitativa

Una RB es un grafo; esto es, una representación gráfica de un problema. Aunque la definición de grafo, y la terminología que la acompaña varía en función de los autores se puede definir grafo como un par  $G = (V, E)$ , donde  $V$  es un conjunto finito de vértices, nodos o variables y  $E$  es un subconjunto del producto cartesiano  $V \times V$  de pares ordenados de nodos llamados enlaces o aristas. Por otro lado, una red bayesiana es un tipo concreto de grafo que se denomina grafo dirigido a cíclico (GDA)[8][5][6].

### 3.3 Dimensión Cuantitativa

Existen tres elementos esenciales que caracterizan la dimensión cuantitativa de una RB: el concepto de probabilidad como un grado de creencia subjetiva relativa a la ocurrencia de un evento, un conjunto de funciones de probabilidad condicionada que definen a cada variable en el modelo y el teorema de Bayes como herramienta básica para actualizar probabilidades con base en experiencia [8][5][6].

### 3.4 Propiedades de las Redes Bayesianas

Las RB se utilizan en problemas en los que interviene la incertidumbre, es decir, donde no se tiene un completo conocimiento del estado del sistema, pero que sin embargo, se pueden realizar observaciones (obtener evidencias) y actualizar las probabilidades del resto del sistema [8]. Hay dos tipos de evidencia:

1) Evidencia firme o específica (instanciación), se da cuando se asigna un valor concreto a una variable, es decir, se tiene certeza del estado de dicha variable. Por ejemplo, imagínese que la variable A representa el resultado de un partido de baloncesto, con dos posibles estados (ganar, perder); si se conoce que el equipo ha ganado el partido, se podría asignar la probabilidad 1 (certeza absoluta) al estado ganar y 0 al estado perder (ver figura 2).

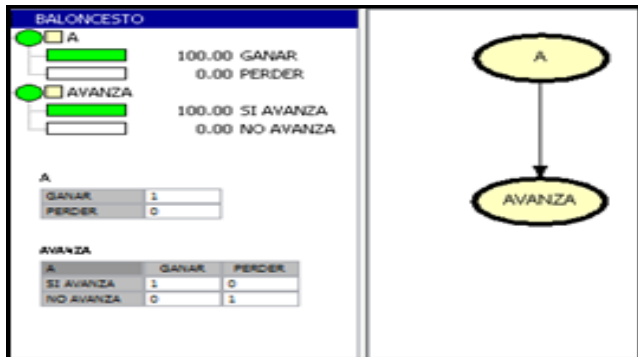


Figura 2. Red bayesiana que ilustra el ejemplo de la evidencia firme en un partido de baloncesto.

2) Evidencia parcial o virtual de un nodo, permite actualizar las probabilidades a priori de los estados que puede tomar la variable. Por ejemplo, a la variable resultado del partido de baloncesto, si el equipo pierde por 20 puntos a pocos minutos del final, se podría asignar una probabilidad muy alta a perder y muy baja a ganar [8].

Una RB en la que se proporciona un sistema de inferencia, una vez encontradas nuevas evidencias sobre el estado de ciertos nodos, se modifican sus tablas de probabilidad; y a su vez, las nuevas probabilidades son propagadas al resto de los nodos. La propagación de probabilidades se conoce como inferencia probabilística, es decir, la probabilidad de algunas variables puede ser calculada dadas evidencias en otras variables [1].

Existen varias propiedades de las RB que las convierten en una opción atractiva para la creación de estructuras que representan el conocimiento. De acuerdo a Rodríguez & Dolado [8] algunas de estas características son:

**Representación gráfica.** Las RB ofrecen una representación gráfica de las relaciones explícitas de dependencia del dominio y nos permiten modelar sistemas complejos permitiéndonos entender las relaciones causales visualizándolas por medio del grafo.

**Modelado cualitativo y cuantitativo.** Las RB cuentan con un enfoque cualitativo, el grafo, y uno cuantitativo, las tablas de probabilidades, que permiten utilizar criterios objetivos y subjetivos.

**Inferencia bidireccional.** Las RB pueden hacer inferencia en ambos sentidos, es decir, las variables de entrada pueden ser usadas para predecir las variables de salida y viceversa. Fijando las variables de salida con los valores deseados, es posible predecir qué valores de las variables de entrada permiten dicha salida.

**Análisis de sensibilidad.** Dado un conjunto de evidencias, las redes Bayesianas permiten fácilmente calcular la sensibilidad de ciertas variables, simplemente modificando las evidencias.

**Incetidumbre.** Las redes bayesianas pueden modelar grados de certidumbre, en vez de valores exactos. Por tanto, permiten modelar la incertidumbre de manera efectiva y explícitamente, por lo que pueden realizar buenas predicciones con información incompleta.

### 4. EN CAMINO DEL MÉTODO PARA LA REPRESENTACIÓN, VALORACIÓN E INTERPRETACIÓN DE HABILIDADES (REVIN-H)

La elaboración de una red bayesiana no es una tarea sencilla. Esto se debe a que es necesario educir, representar e ingresar el conocimiento y la experiencia de los especialistas como valores porcentuales, donde cada valor representa las creencias del especialista respecto a la variable analizada. De esta forma se ve afectado el grado de efectividad del pronóstico que emita la red bayesiana y por ende su eficiencia.

Existen un método propuesto por Xiao-xuan & Hui [7] que permite educir el conocimiento y la experiencia del especialista del dominio para la creación de una red bayesiana. En este método se sugiere la construcción de redes bayesianas en dos fases: (1) determinar la estructura de la red, y (2) la obtención de las tablas de probabilidad condicional de cada nodo.

La primera fase se centra en determinar la relación causal entre las variables. Esta fase se apoya en la creación de una tabla que incluye a cada variable y sus relaciones. El objetivo es facilitar la colaboración entre especialistas y llegar a un consenso.

Para la segunda fase es necesario especificar las tablas de probabilidad condicional de los nodos. Esta es la tarea más crítica en la construcción de redes bayesianas, ya que es muy difícil para los expertos proporcionar directamente probabilidades numéricas.

Como apoyo a esta fase se propone el uso de una "escala de probabilidad" (ver figura 3) para facilitar la interpretación del conocimiento que el especialista posee y traducirlo a una distribución de probabilidad. El resultado del uso de este método promete mejores resultados para la creación de redes bayesianas en base a conocimiento de especialistas [7].

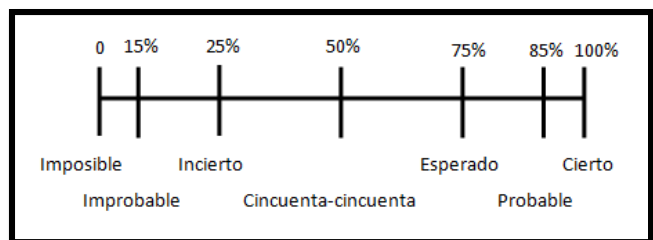


Figura 3. Escala de Probabilidad. Tomada de "Using Expert's Knowledge to Build Bayesian Networks" [7].

El método propuesto por Xiao-xuan & Hui [7] ofrece una opción interesante para la educación de conocimientos de especialistas. Sin embargo no incluye ningún aspecto respecto a la estructura necesaria para una mejor representación del conocimiento en la RB, así como la elección de las técnicas necesarias para la evaluación del conocimiento. Se limita a representar los datos que se pueden ingresar a la RB como valores probabilísticos y a apoyar el proceso de consenso entre los especialistas del dominio. No obstante puede ser de utilidad para la fase de interpretación que se desea

implementar en el método propuesto para la verificación de habilidades.

El método a desarrollar contará con tres etapas principales: educación de conocimientos, representación de conocimientos y valoración de habilidades. Basándose en un enfoque incremental donde se partirá de la creación de una estructura parcial de conocimientos de los especialistas del dominio.

Posteriormente se pretende aplicar un ciclo de educación de conocimiento tácito a explícito generando una primera representación de conocimientos de tipo grafo causal que represente el área de interés para el objetivo de evaluación y las posibles tareas/ problemas que podrán evaluarse. Luego se desea sintetizar las tareas/ problemas a una representación de habilidades necesarias para desempeñar satisfactoriamente el objetivo de evaluación. Obteniendo un mapa de conocimiento que deberá entrar al ciclo de educación de conocimientos para ser validado por los especialistas del dominio.

Una vez obtenido el mapa de conocimientos se pretende crear una representación formal por medio de un GDA que conforme la parte cualitativa de la RB. Luego apoyarse en los especialistas del dominio para generar las tablas de probabilidad de los nodos que forman el GDA y representan la parte cuantitativa de la RB.

Finalmente la aplicación de una escala de probabilidad como la propuesta por Xiao-xuan & Hui [7] será de utilidad para la creación de una red bayesiana. Además se realizará una adecuación a la escala de forma que permita la interpretación del pronóstico emitido por la red bayesiana. Con el interés de ofrecer la traducción de probabilidades a un esquema de información en palabras o números que nos facilite el entendimiento de las inferencias obtenidas por la red bayesiana.

El resultado esperado de la aplicación del método deberá ser una estructura capaz de representar el conocimiento, así como la posibilidad de elegir alguna de las técnicas más adecuadas para verificar e interpretar las habilidades.

De esta forma se pretende generar un método útil para quienes se encuentren con la necesidad de realizar la evaluación de habilidades y requieran de una guía para el desarrollo de una estructura capaz de representar conocimientos y elegir las técnicas que permitan verificar e interpretar habilidades.

## 5. DISCUSIÓN

Algunos enfoques y modelos generados a partir de la IA como el enfoque cognitivo, el modelo probabilístico, las técnicas Rough Set y las Redes Bayesianas entre otras, pueden ser más apropiados para ciertas necesidades de evaluación facilitando la construcción de estructuras que representen el conocimiento deseado. Sin embargo, será necesario elegir un modelado del conocimiento del dominio y técnicas de razonamiento para evaluar dicho conocimiento, en consecuencia se supone pertinente la propuesta de un método para la verificación de habilidades en DEI.

Algunas técnicas del área de Soft Computing pueden utilizarse para la inferencia, clasificación y generación de conocimientos de manera efectiva. No obstante, las características y estructura que posea el conocimiento que se desea evaluar serán independientes

de las técnicas empleadas. Es importante destacar que el grado de asertividad de las inferencias dependerá en gran medida de la correcta representación, transferencia, verificación e interpretación del conocimiento que se desea evaluar en el dominio en cuestión.

La eficiencia de la evaluación en el enfoque cognitivo depende en gran parte de la relación entre la cantidad de pruebas con la amplitud de la evaluación. En problemas donde se tiene una amplia gama de habilidades el reto es generar una evaluación utilizable. Esta evaluación requerirá de un modelo que pueda establecer vínculos entre habilidades (requisitos previos del conocimiento del dominio) capaz de inferir el conocimiento necesario del sujeto en términos del dominio abordado [1].

Hasta el momento el modelo Bayesiano parece ser la opción más acertada para inferir el conocimiento deseado en la verificación de habilidades. Además la RB es capaz de representar de manera cualitativa y cuantitativa el conocimiento necesario para temas de DEI donde se busca emitir un pronóstico del posible desempeño del sujeto respecto a un objetivo de evaluación.

## 6. REFERENCIAS

- 1 Desmarais M. C., Baker R. S. J. d. , 2011. *"A Review of Recent Advances in Learner and Skill Modeling in Intelligent Learning Environments"*. Review. Polytechnique Montréal, Worcester Polytechnic Institute.
- 2 Olmos K. M. , 2015. *"KMoS-RE: Knowledge Management on a Strategy for Requirements Engineering"*. Tesis de Doctorado, Doctor en Ciencias en Ingeniería. Universidad Autónoma de Ciudad Juárez.
- 3 Lynch C. F. & Ashley K. D., Alevent V. & Pinkwart N. *"Defining Ill-Defined Domains; a Literature Survey"*. Survey. University of Pittsburgh, Carnegie Mellon University.
- 4 Olmos K. M., Osollo J. E., Fernández L. F. , 2010. *"Pertinencia de la Formalización de Dominios Semi-Formalmente Definidos en el Análisis Inteligente de Datos"*. Culcyt Universidad Autónoma de Ciudad Juárez.
- 5 Charniak E. , 1991. *"Bayesian Networks Without Tears"*. A.I. Magazine Volume 12 Number 4.
- 6 Darwiche A. , 2010. *"What are Bayesian Networks and why are their Applications Growing Across all Fields?"*. Communications of the ACM Volume 53 Number 12. DOI 10.1145/1859204.1859277.
- 7 Xiao-xuan H., Hui W., Shuo W. , 2007. *"Using Expert's Knowledge to Build Bayesian Networks"*. International
- 8 Rodríguez D., Dolado J., *"Redes Bayesianas en la Ingeniería del Software"*. Universidad de Alcalá, Universidad del País Vasco.
- 9 Chahuán-Jiménez K., 2009. *"Evaluación cualitativa y gestión del conocimiento"*. Educación y educadores, volumen 12, No. 13. Universidad de la Valparaíso.