Predicción del rendimiento académico de los estudiantes de física a través de las redes bayesianas en la unidad de cantidad de movimiento lineal

Miguel E. López Balanzátegui, Jorge Flores Herrera,
Bolívar Flores Nicolalde, Francisca Flores Nicolalde
Escuela Superior Politécnica del Litoral, ESPOL, Facultad de Ciencias Naturales y Matemáticas, Campus Gustavo Galindo Km 30.5 Vía Perimetral, P.O. Box 09-01-5863, Guayaquil, Ecuador.

E-mail: mblopez@espol.edu.ec

Resumen
En este trabajo se presenta una investigación de probabilidades estadísticas con Redes Bayesianas, basándonos en los resultados de las evaluaciones realizadas a estudiantes de física de una universidad ecuatoriana, con lo que se pudo inferir resultados futuros del desempeño de los estudiantes y la relación de los conocimientos teóricos con la resolución de problemas en la materia de física. Se realizó el cálculo de las probabilidades *a priori*, a posteriori e inferencias con fórmulas y con el software para Redes Bayesianas llamado "ELVIRA", el cual se encuentra disponible en la web. Adicionalmente se presentaron los gráficos de las Redes Bayesianas y sus probabilidades generadas por el programa.

Palabras claves: Redes Bayesianas, Probabilidades *a priori* y *a posteriori*, Inferencias.

Abstract
This paper presents an investigation of probability statistics with Bayesian Networks, based on the results of assessments made on physics students of a Ecuadorian university, with what could be infer results future performance of the students and the relationship of the theoretical knowledge with the resolution of problems in the field of physics. It performs the calculation of probabilities *a priori*, *a posteriori* and inferences with formulas and with the software for Bayesian Networks called "ELVIRA" which is available on the web. Additionally graphics Bayesian Networks and their probabilities generated by the program are presented.

Keywords: Bayesian networks, Prior and posterior probabilities, Inferences.

PACS: 01.40.gb, 02.70.Rr, 07.05.Mh **ISSN 1870-9095**

I. INTRODUCCIÓN

La educación en nuestro país presenta algunos problemas a nivel de instrucción secundaria y superior, en lo que respecta a los conocimientos iniciales que se van concatenando a los conocimientos posteriores.

La Física es una ciencia de carácter jerárquico, en donde los estudiantes necesitan ir formando los conocimientos progresivamente, y los maestros necesitan que sus alumnos dominen un tema para poder impartir uno nuevo, teniendo en cuenta los conceptos previos; pero no siempre entienden la teoría, y hay falta de estrategias en la resolución de problemas.

Estos conocimientos concatenados son los que nos permitieron inferir resultados del rendimiento de los estudiantes, utilizando la probabilidad estadística de Bayes.

Este trabajo está dirigido a los maestros de nuestro país o de cualquier lugar del mundo, en donde las condiciones y las realidades en temas de la educación, impliquen adquirir información sobre la jerarquía de los conocimientos y el uso de las probabilidades de Bayes. Estos resultados pueden ser utilizados para retroalimentación de los maestros y para mejorar el desempeño en forma individual y grupal de los estudiantes.

En los casos de evaluación convencional los maestros realizan inferencias sobre el rendimiento de todo un curso.

Este problema podría solucionarse a través de la aplicación de las Redes Bayesianas, que permiten analizar a los estudiantes en forma individual e inferir acerca de grupos de individuos. [1]

Russell Almond y Valeria Shute (2008), en su trabajo Bayesian Networks: A teacher´s view, describen:

> *"Los profesores estiman competencias basados en Redes Bayesianas a una clase llena de estudiantes, que se enfrentan a un problema diferente, de un tutor mirando a un estudiante a la vez. Afortunadamente, estimaciones de competencias individuales pueden ser agregadas en el aula y otras estimaciones de grupo a través de sumas y promedios".*

II. MARCO TEÓRICO

A. Inteligencia Artificial

La Inteligencia Artificial ha sido muy importante como fuente inagotable de técnicas, métodos, modelos y algoritmos, tanto para el análisis de datos, como para el modelado y simulación de sistemas. Técnicas tales como redes neuronales artificiales, algoritmos evolutivos, autómatas celulares, Redes Bayesianas y modelos ocultos de Markov, resultan ser enfoques ideales para dominios que se caracterizan por una explosión de datos y muy poca teoría, como es el caso de las Redes Bayesianas [2].

La estructura de una Red Bayesiana puede resultar ser un problema de optimización combinatoria, ya que consiste en encontrar la mejor red de todas las posibles, en un espacio en el que intervienen *N* atributos para identificar los objetos del dominio de aplicación.

B. Redes Bayesianas

El origen del concepto de la obtención de probabilidades "*a posteriori*" con información limitada se le atribuye al reverendo Thomas Bayes (1702-1761). Él estudió el problema de la determinación de la probabilidad de las causas a través de los efectos observados.

Actualmente, con base en su obra, "Essay towards solving a problem in the Doctrine of Chances", en el que trataba el problema de las causas a través de los efectos observados, y donde se enuncia el Teorema que lleva su nombre. El trabajo fue entregado a la Royal Society por Richard Price, y es la base de la inferencia bayesiana.

La fórmula básica para la probabilidad condicional en circunstancias de dependencia se conoce como Teorema de Bayes.

$$P\,(B/A) = P\,(A/B).P\,(B) \,/\, P\,(A)$$

El Teorema de Bayes ofrece un método estadístico poderoso para evaluar nueva información, y revisar nuestras anteriores estimaciones de la probabilidad de que las cosas se encuentran en un estado o en otro [3].

Antes de presentar formalmente la teoría matemática de las redes bayesianas, explicaremos mediante un ejemplo sencillo el significado intuitivo de los conceptos que después introduciremos.

En una red bayesiana, cada nodo corresponde a una variable, que a su vez representa una entidad del mundo real.

Por tanto, de aquí en adelante hablaremos indistintamente de nodos y variables, y los denotaremos con letras mayúsculas, como X. Utilizaremos la misma letra en minúscula, x, para referirnos a un valor cualquiera de la variable X. Los arcos que unen los nodos indican relaciones de influencia causal.

Una Red Bayesiana está compuesta de variables independientes, y variables dependientes, las independientes representan a los nodos que son "padres" y las dependientes representan a los nodos que son "hijos".

FIGURA 1. Red bayesiana

Es decir, el nodo con la variable *X*, es el padre, y el nodo con la variable *Y*, es el hijo.

Ejemplo 1. La red bayesiana más simple [4].

La red bayesiana no trivial más simple que podemos imaginar consta de dos variables, que llamaremos *X* e *Y1*, y un arco desde la primera hasta la segunda.

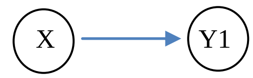

FIGURA 2. Red Bayesiana simple

Para concretar el ejemplo, supongamos que *X* representa el sida e *Y1* representa el test de Elisa, que es la prueba más habitual para detectar la presencia de dicha enfermedad.

Cuando *X* sea una variable binaria, denotaremos por *+x* la presencia de aquello a lo que representa, y por *-x* a su ausencia.

Así, por ejemplo en este caso, *+x* significara "el paciente tiene sida", y *-x* "el paciente no tiene sida"; *+y1* significará un resultado positivo del test de Elisa, y *-y1* un resultado negativo.

La información cuantitativa de una red bayesiana viene dada por:

• La probabilidad *a priori* de los nodos que no tienen padres.

• La probabilidad condicionada de los nodos con padres.

Por tanto, en nuestro ejemplo, los datos que debemos conocer son P(x) y P (y1/x).

Así, la red bayesiana completa sería:

FIGURA 3. Red bayesiana completa (con probabilidades).

$$P\left(+x\right) = 0.005.$$

$$P\left(+y1/+x\right) = 0.985.$$

$$P\left(+y1/\neg\, x\right) = 0.0009.$$

Veamos qué significado tienen en este caso estos valores:

- P (+x) = 0.005 indica que, *a priori*, un 0.5% de la población padece de sida. En medicina, esto se conoce como prevalencia de la enfermedad.
- P(+y1/+x) = 0.985 indica que cuando hay sida, el test de Elisa da positivo en el 98.5% de los casos. Esto se conoce como sensibilidad del test.

- P(+y1/ -x) = 0.0009 indica que, cuando no hay sida, el test de Elisa da positivo en el 0.09% de los casos, y negativo en el 99.91%. A esta segunda probabilidad se la llama especificidad del test.

Alternativamente, se habla también de las tasas de falsos positivos (probabilidad de que el test dé positivo aunque la persona no está enferma) y tasas de falsos negativos (probabilidad de test negativo cuando la persona está enferma).

Conociendo estos datos, podemos calcular:

a) La probabilidad *a priori* de *Y1*:

$$P\left(+y1\right) = P\left(+y1/+x\right) P\left(+x\right) + P\left(+y1/\neg x\right) P\left(\neg x\right) = 0.00582.$$

$$P\left(\neg y1\right) = P\left(\neg y1/+x\right) P\left(+x\right) + P\left(\neg y1/\neg x\right) P\left(\neg x\right) = 0.99423.$$

b) Las probabilidades *a posteriori* dada una evidencia observada, e, es: *P*(x) = P(x/e).*

Supongamos que el test de Elisa ha dado positivo. ¿Qué probabilidad hay ahora de que la persona padezca la enfermedad? Si la prueba tuviese fiabilidad absoluta, esta probabilidad sería del 100%. Pero, como existe la posibilidad de que haya habido un falso positivo, buscamos *P*(+x) = P (+x/+y1).* Para calcularla, podemos aplicar el teorema de Bayes:

$$P^*\left(+x\right) = P\left(+x/+y1\right) =$$

$$\frac{P\left(+x\right) P\left(+y1/+x\right)}{P\left(+y1\right)} =$$

$$\frac{\left(0.005\right)\left(0.985\right)}{0.00582} = 0.8462.$$

Es decir, de acuerdo con el resultado de la prueba, hay un 84,62% de probabilidad de que el paciente tenga sida.

De la misma forma, podríamos calcular P (¬ x):

$$P^*\left(\neg x\right) = P\left(\neg x/+y1\right) =$$

$$\frac{P\left(\neg x\right) P\left(+y1/\neg X\right)}{P\left(+Y1\right)} = 0.1539.$$

Que, por supuesto, es la probabilidad complementaria.

La expresión general del Teorema de Bayes que hemos utilizado es:

$$P^*\left(x\right) = P\left(x/y\right) = \frac{P\left(x\right) P\left(y/x\right)}{P\left(y\right)}.$$

C. Definición formal de red bayesiana

Antes de definir formalmente las redes bayesianas, vamos a definir algunos conceptos de teoría de grafos y teoría de la probabilidad:

D. Definiciones previas

1. *Arco*. Es un par ordenado (*X*, *Y*). Esta definición de arco corresponde a lo que en otros lugares se denomina arco dirigido. En la representación gráfica, un arco (*X*, *Y*) viene dado por una flecha desde *X* hasta *Y*.

FIGURA 4. Arco dirigido (flecha).

2. *Grafo dirigido.* Es un par *G = (N, A)* donde *N* es un conjunto de nodos y un conjunto de arcos definidos sobre los nodos.

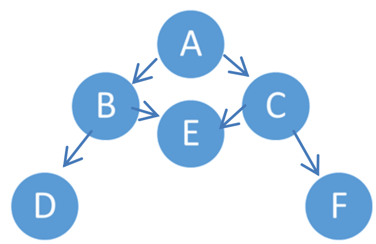

FIGURA 5. Grafo dirigido.

3. *Camino dirigido.* Es una secuencia ordenada de nodos.

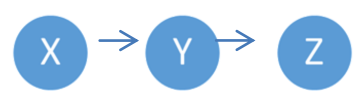

FIGURA 6. Camino dirigido.

4. *Ciclo:* es un camino no dirigido que empieza y termina en el mismo nodo *X*.

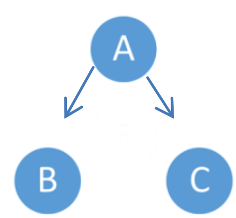

FIGURA 7. Ciclo.

5. **Grafo acíclico:** es un grafo que no contiene ciclos.

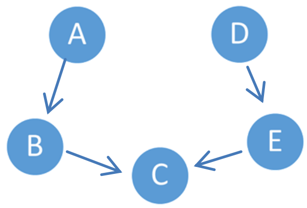

FIGURA 8. Gráfico acíclico.

6. **Padre:** X es un padre de Y si y sólo si existe un arco $X \rightarrow Y$. Se dice también que Y es hijo de X.

Dos variables X e Y son independientes, si se tiene que $P(X/Y) = P(X)$. De esta definición, se tiene una caracterización de la independencia, que se puede utilizar como definición alternativa: X e Y son independientes sí y sólo sí: $P(X, Y) = P(X)P(Y)$.

E. Aplicación de las redes bayesianas en la educación

Las redes bayesianas representan explícitamente nuestro conocimiento sobre los elementos en el sistema y las relaciones que existen entre ellos. Estas relaciones operan propagando conocimiento a través de la red una vez que se tiene evidencia sobre alguno de los objetos o eventos del sistema.

De esta manera, se pueden "aprender" las probabilidades de todos los elementos de la red a partir del conocimiento de algunos de ellos y de las relaciones condicionales entre ellos.

Por tal razón se pueden aplicar en la educación porque infiere resultados futuros, útiles para que el maestro tenga la ubicación del estudiante dependiendo del rendimiento de cada uno de ellos.

F. Cantidad de movimiento lineal

La cantidad de movimiento, momento lineal, ímpetu o momentum, es una magnitud física fundamental de tipo vectorial que describe el movimiento de un cuerpo.

En mecánica clásica la cantidad de movimiento se define como el producto de la masa del cuerpo y su velocidad en un instante determinado.

$$\vec{P} = m\vec{V}.$$

Donde $\vec{P}$ es el símbolo con que se representa la cantidad de movimiento. $\vec{P}$ es un vector que apunta en la misma dirección que $\vec{V}$.

Una experiencia común indica que, todo cuerpo con masa posee inercia, propiedad que representa la oposición que ofrece dicho cuerpo a que le cambien su estado de movimiento.

A mayor masa más oposición a detenerlo, así como a mayor rapidez es más difícil cambiar su cantidad de movimiento.

G. Variación en la cantidad de movimiento

Cuando ocurre un cambio en la masa y/o en la velocidad, existirá un cambio en la cantidad de movimiento del cuerpo considerado.

Si la masa permanece constante, pero la velocidad del cuerpo cambia, se tendrá que:

$\vec{P}_1 = m\vec{V}_1$, en el primer instante de tiempo.

$\vec{P}_2 = m\vec{V}_2$, en el segundo instante de tiempo.

La variación de la cantidad de movimiento será:

$$\vec{P}_2 - \vec{P}_1 = m\vec{V}_2 - m\vec{V}_1 \Rightarrow \vec{P}_2 - \vec{P}_1 = m(\vec{V}_2 - \vec{V}_1).$$

Luego,

$$\Delta\vec{P} = m\Delta\vec{V}.$$

H. Conservación de la cantidad de movimiento

Para deducir el enunciado de este principio, se parte de la Tercera Ley de Newton (ley de acción y reacción).

Considere dos esferas de masa m_1 y m_2, las cuales se hayan dotadas inicialmente, de velocidades, y al chocar las velocidades serán diferentes.

Como las esferas están en contacto mutuo durante un intervalo de tiempo muy pequeño, el impulso inicial debe ser igual y opuesto al impulso final, escribiéndose:

$$\vec{F}_1\Delta t = -\vec{F}_2\Delta t. \quad (1)$$

Por otra parte

$$\vec{F}_1\Delta t = m_1(\vec{V}'_1 - \vec{V}_1); -\vec{F}_2\Delta t = -m_2(\vec{V}'_2 - \vec{V}_2). \quad (2)$$

Sustituyendo (2) en (1) se tiene que:

$$m_1(\vec{V}'_1 - \vec{V}_1) = -m_2(\vec{V}'_2 - \vec{V}_2).$$

Aplicando la propiedad distributiva se tiene que:

$$m_1\vec{V}'_1 - m_1\vec{V}_1 = -m_2\vec{V}'_2 + m_2\vec{V}_2;$$

trasponiendo términos se obtiene:

$$m_1\vec{V}'_1 + m_2\vec{V}'_2 = m_1\vec{V}_1 + m_2\vec{V}_2,$$

$$\vec{P}'_1 + \vec{P}'_2 = \vec{P}_1 + \vec{P}_2.$$

El primer miembro de la ecuación representa la suma vectorial de las cantidades de movimiento después del choque, y el segundo miembro, representa la suma de las cantidades de movimiento antes del choque.

I. Rendimiento académico

El rendimiento escolar es parte esencial en todo acto didáctico, según Cortez Bohiga es:

"El nivel de conocimiento de un alumno medido en una prueba de evaluación".

Y considera también que:

"En el rendimiento escolar intervienen además del nivel intelectual, variables de personalidad (extroversión, introversión, ansiedad...) y motivacionales, cuya relación con el rendimiento escolar no siempre es lineal, sino que esta modulada por factores como nivel de escolaridad, sexo, aptitud" [4].

La combinación del rendimiento con los resultados de las evaluaciones, permite o proporciona índices de fiabilidad; siendo ésta a su vez, una aproximación al verdadero desempeño académico. Es decir, que estos indicadores son fundamentales, ya que proporcionan un control en las diferentes entidades de educación. Además, permiten tomar decisiones en la planificación educativa y en las políticas a aplicarse.

García-Valcárcel, (2007) describe:

"La tendencia para medir el rendimiento académico es hacerlo desde un punto de vista práctico, que vincule el éxito o el fracaso con resultados inmediatos, es decir, con las calificaciones de los alumnos en un determinado tiempo" [5].

J. Software "ELVIRA"

El software "ELVIRA" nace como un proyecto, cuyo principal objetivo era la construcción de un entorno que sirviera por un lado, para la investigación de nuevos métodos y algoritmos de razonamiento probabilístico, y por otro, para la implementación de sistemas expertos bayesianos.

El programa resultante se llamó Elvira, tomando el antiguo nombre de la ciudad de Granada, a cuya Universidad están vinculados la mayor parte de los investigadores del proyecto.

Este programa cuenta con un formato propio para la codificación de los modelos, un lector-intérprete para los modelos codificados, una interfaz gráfica para la construcción de redes, con opciones específicas para modelos canónicos (puertas OR, AND, MAX, etc.), algoritmos exactos y aproximados de razonamiento tanto para variables discretas como continuas, métodos de explicación del razonamiento, algoritmos de toma de decisiones, aprendizaje de modelos a partir de bases de datos, fusión de redes, etc.

Elvira está escrito y compilado en Java, lo cual permite que pueda funcionar en diferentes plataformas y sistemas operativos (Linux, MS-DOS/Windows, Solaris, etc.).

III. MÉTODO

Esta investigación se fundamenta y diseña en un modelo de intervención en Enseñanza de las Ciencias e Investigación Educativa, con un enfoque constructivista, para su aplicación en Educación Superior. El área temática a trabajar en este estudio se centra en problemas referentes a Enseñanza de la Física, Estadística e Investigación.

El trabajo de investigación se desarrolla a través de una metodología cualitativa y cuantitativa, en especial para predecir el rendimiento académico de los estudiantes.

La metodología es de tipo correlacional, es decir en la relación entre los resultados de aprendizaje basados en pruebas, y su efecto en el rendimiento académico de los estudiantes. Se efectúan pruebas para predecir el rendimiento académico de los alumnos, y basado en estos resultados aplicar la técnica de las redes bayesianas idónea para determinar el desempeño académico futuro de ellos.

Chain (2003) aporta una metodología, utilizando el análisis de la información obtenida en el proceso de ingreso y trayectoria académica de los estudiantes [6], por medio de la utilización de redes bayesianas, estimando cuantitativamente las relaciones existentes entre las variables estudiadas.

A. Sujetos

Participaron en este estudio 27 estudiantes de una institución de educación superior ecuatoriana, que cursan el primer año de ingenierías en informática, en la materia de Física; en donde se los evaluó mediante pruebas conceptuales y resolución de problemas en la unidad Cantidad De Movimiento Lineal, para poder inferir los resultados futuros de los estudiantes.

B. Tareas instruccionales y materiales

La tarea instruccional utilizada en este estudio fue, la unidad Cantidad de Movimiento Lineal. Para esto, se dispuso de doce horas de clases, en el dictado de la materia por parte del profesor, y dos horas asignadas a dos evaluaciones. Las dos evaluaciones se tomaron a medida que el profesor iba desarrollando el contenido de la materia. Estas evaluaciones estuvieron basadas en cada uno de los objetivos y, de acuerdo al desempeño, se infirió el éxito o fracaso de los estudiantes.

IV. RESULTADOS

Se realizó una tabulación de los datos de todos los alumnos en una hoja de Excel con los resultados de la primera y segunda evaluación.

Éstas se encuentran separadas por: la nota de las preguntas conceptuales, la nota de los problemas resueltos, la nota total, actividades adicionales del profesor, la nota final de la materia y el resultado de aprobación o reprobación de la materia.

Cabe señalar que, todos estos datos fueron procesados al final de la investigación, para poder inferir en forma futura el resultado de las evaluaciones, relacionando estos con la probabilidad de que el estudiante apruebe o no la materia.

Se utilizaron distintos colores en la hoja de Excel, para representar y evaluar los distintos casos y probabilidades condicionales.

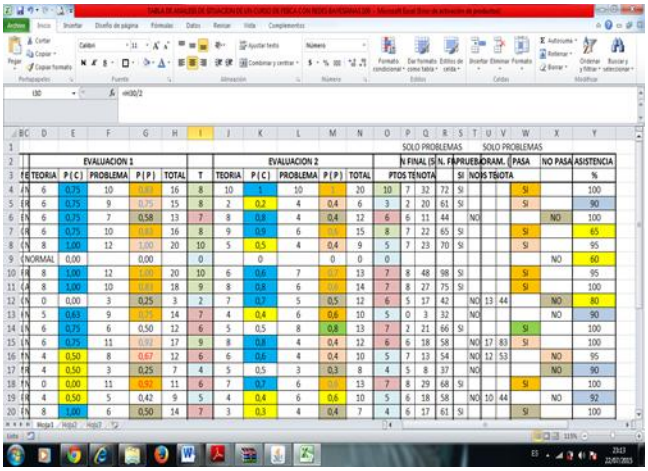

FIGURA 1. Hoja de Excel para la tabulación y evaluación de los datos.

Se tomaron algunas consideraciones para el análisis de los datos, y para poder trabajar con resultados binarios. Es decir, en las evaluaciones el estudiante contesta o no contesta las preguntas conceptuales, y el estudiante resuelve o no resuelve los problemas. Para esto se consideró que, el estudiante no sabe ni contesta las preguntas conceptuales si saca 4 (o menos) sobre 10 puntos; y que el estudiante si sabe resolver problemas si su nota es de 6 (o más) sobre 10 puntos.

Además, para la formación de algunas redes se tomaron consideraciones como: la de los resultados *excelente*, *bueno y malo*, los que se analizaron y se relacionaron con la posibilidad de que el estudiante apruebe o no apruebe la materia. En la Tabla I se dan los rangos para la formación de los grupos.

TABLA I. Formación de grupos por rango de notas.

NOTAS	RANGO
Excelente	8 - 10
Bueno	6 - 7.9
Malo	0 - 5.9

Se denominó *G*, a las notas de la primera evaluación y *H* a las notas de la segunda evaluación. Tomando en cuenta los rangos indicados en la Tabla I, se obtuvo la siguiente clasificación:

TABLA II. Formación de grupos por notas en evaluaciones 1 y 2.

ALUMNOS	GRUPOS	EV-1 G	EV-2 H
Nota alta		9	2
Nota media		8	11
Nota baja		10	14

Se analizaron ocho redes bayesianas en total, cada una buscando la posibilidad de que sea útil a la hora de inferir resultados con ella. Es así que, en la formación de éstas veremos que unas son más importantes que otras, y dan mejores resultados y aseguran que la inferencia sea más confiable. A continuación detallaremos las redes formadas:

A. Red bayesiana 1

Esta red nace de los resultados de la primera evaluación, utilizando dos variables, una dependiente (resolver problemas) "*P1*" y la otra independiente (preguntas conceptuales) "*C1*".

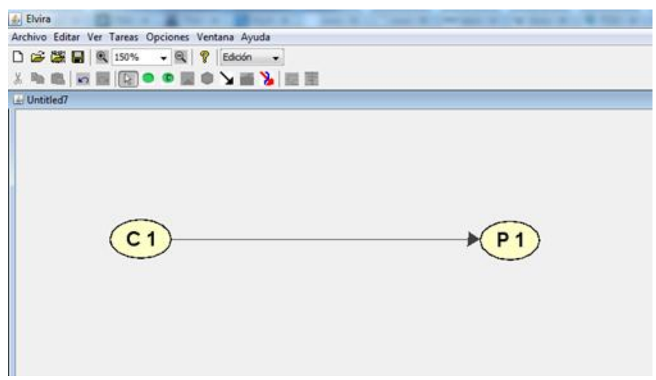

FIGURA 2. Red bayesiana 1-primera evaluación.

- La probabilidad de contestar bien las preguntas conceptuales es: $P(+c) = 0{,}58$.
- La probabilidad condicional de resolver bien los problemas, habiendo contestado bien las preguntas conceptuales es: $P(+p/+c) = 0{,}72$.
- Y la probabilidad condicional de resolver bien los problemas, habiendo contestado mal las preguntas conceptuales es: $P(+p\,/\,-c) = 0{,}55$.

Todos estos datos fueron sacados de los resultados de las evaluaciones en la hoja de Excel de la Figura 1. Y tomando en cuenta las condiciones explicadas anteriormente, de que salir mal en las preguntas conceptuales significaba sacar 4 (o menos), y salir bien en la resolución de problemas significaba sacar de 6 (o más).

Con estos datos podemos calcular:

a) ¿Cuál es la probabilidad *a priori* de que un alumno cualquiera resuelva correctamente los problemas P (+p)?

$$P\ (+p) = P\ (+p/+c)P\ (+p) + P\ (+p/\neg c)P\ (\neg c)$$

$$P\ (+p) = (0,72)(0,58) + (0,55)(0,42)$$

$$P\ (+p) = 0,4176 + 0,231$$

$$P\ (+p) = 0,6486; \quad P(+p) = 0,65$$

La probabilidad de que un alumno resuelva bien los problemas es del 65%.

La probabilidad de que no lo resuelva:

$$P\ (\neg p) = P\ (\neg p/+c)P\ (+p) + P\ (\neg p/\neg c)P\ (\neg c)$$

$$P\ (+p) = (0,28)(0,58) + (0,45)(0,42)$$

$$P\ (\neg c) = 0,3514$$

b) La probabilidad *a posteriori* dada una evidencia observada "*e*".

P* (c) = P (c/e); suponer que la evidencia observada es que cierto alumno ha resuelto correctamente los problemas.

¿Qué probabilidad hay ahora de que conozca las preguntas conceptuales C?

$$P*\ (c) = P\ (+c/+p) =$$

$$\frac{P\ (+c)P\ (+p/+c)}{P\ (+p)} =$$

$$\frac{(0,58)(0,72)}{(0,65)} =$$

$$P\ (+c/+p) = 0,642$$

Y calculando ahora la probabilidad de que el alumno haya resuelto bien el problema y no conozca las preguntas conceptuales sería:

$$P*\ (\neg c) = P\ (\neg c/+p) = 0,358$$

Que como vemos sería la probabilidad complementaria a la anterior.

Ahora vamos a suponer de que la evidencia encontrada es que el alumno no resuelve los problemas, ¿qué probabilidad hay de que conozca las preguntas conceptuales, habiendo resuelto mal los problemas?

$$P*\ (c) = P\ (+c/\neg p) =$$

$$\frac{P\ (+c)P\ (\neg p/+c)}{P\ (\neg p)} = \frac{(0,65)(0,28)}{(0,35)}$$

P (+c/ ¬ p) = 0,52

Y la probabilidad complementaria sería:

$$P*\ (\neg c) = P\ (\neg c/\neg p) = 0,48$$

Que corresponde a una probabilidad de que un alumno no conozca las preguntas conceptuales habiendo resuelto mal los problemas.

Con el software "ELVIRA" ponemos los mismos datos de probabilidad de la variable sin padre, y las probabilidades condicionales de las variables con padres, *P (+c) = 0,58*, *P (+p/+c) = 0,72 y P (+p/ ¬ c) = 0,55*. Y obtenemos los gráficos de las redes bayesianas con los mismos resultados calculados en el problema anterior de la red bayesiana 1 (RB 1).

A continuación mostramos los gráficos que nos genera el programa con los resultados de las probabilidades.

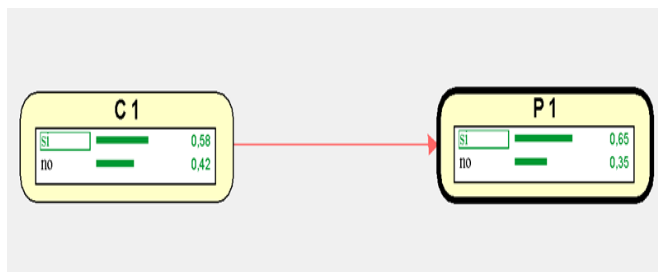

FIGURA 3. Probabilidad *a priori* de red bayesiana 1, calculado por programa ELVIRA.

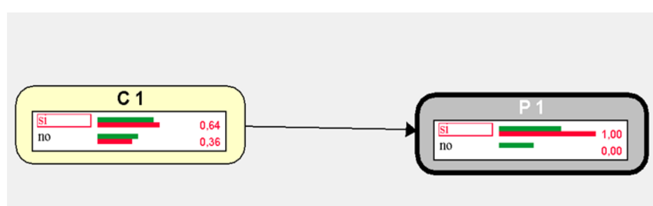

FIGURA 4. Probabilidad *a posteriori* de red bayesiana 1.

Tal como se puede observar, existe concordancia entre los resultados obtenidos por las ecuaciones de Bayes, y los obtenidos en el programa "ELVIRA".

B. Red Bayesiana 2

Esta red nace de los resultados de la segunda evaluación, utilizando dos variables: una dependiente (resolución de problemas) “*P2*” y la otra independiente (preguntas conceptuales) “*C2*.

Utilizando el programa “ELVIRA”, obtenemos la siguiente red bayesiana y sus probabilidades:

FIGURA 5. Red bayesiana 2 – segunda evaluación.

Luego se obtuvo con el programa la probabilidad *a priori*:

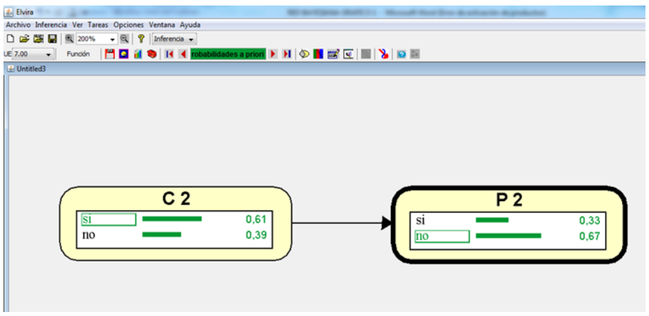

FIGURA 6. Probabilidad *a priori* de red bayesiana 2.

Luego el programa calcula la probabilidad *a posteriori* obteniendo:

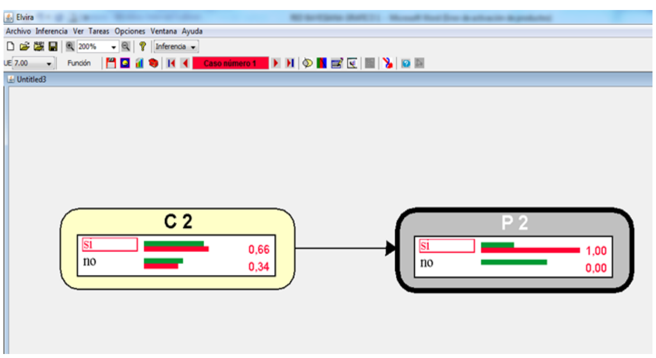

FIGURA 7. Probabilidad *a posteriori* – segunda evaluación.

Tenemos como resultados de esta red:

P (+p2) = 0,33, que sería la probabilidad *a priori*, que indica la probabilidad de resolver correctamente los problemas de la segunda evaluación.

Y *p (+c2/+p2) = 0,66*, que sería la probabilidad *a posteriori*.

P (c) = P (c/e)*; que supone la probabilidad que conozca *C2* (preguntas conceptuales), habiendo la evidencia observada de que, cierto alumno haya resuelto correctamente los problemas de la evaluación 2.

C. Red bayesiana 3

En esta red, ampliamos el modelo de las dos redes anteriores, aquí identificamos la relación de los dos resultados de las evaluaciones, con la posibilidad de aprobar o reprobar la materia.

Utilizando el programa “ELVIRA”, obtenemos la siguiente red bayesiana y sus probabilidades.

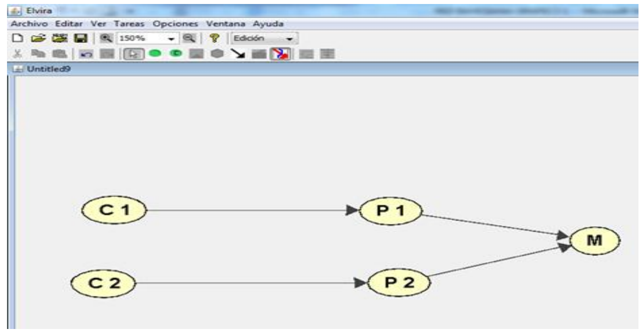

FIGURA 8. Red bayesiana 3.

Luego, se calculó en el programa la probabilidad *a priori*:

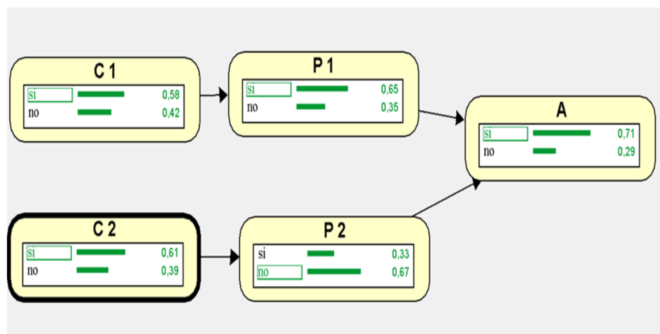

FIGURA 9. Probabilidad *a priori* de red bayesiana 3.

Luego, el programa calcula la probabilidad *a posteriori* obteniendo:

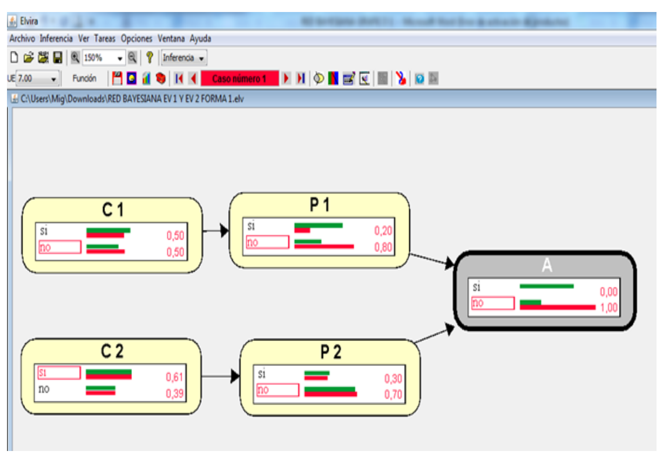

FIGURA 10. Probabilidad *a posteriori* – RB 3 – P(+a).

Tenemos como resultados de esta red:

- *P (+a) = 0,71*, que sería la probabilidad *a priori*, que indica la probabilidad de aprobar la materia con la combinación de las dos evaluaciones.
- *Y p (+c2/+a) = 0,61* y *p (+c1/+a) = 0,58*, que sería la probabilidad *a posteriori*, *P* (c) = P (c/e)*; que supone la probabilidad que conozca *C1* y *C2*, habiendo la evidencia observada de que cierto alumno haya aprobado la materia.

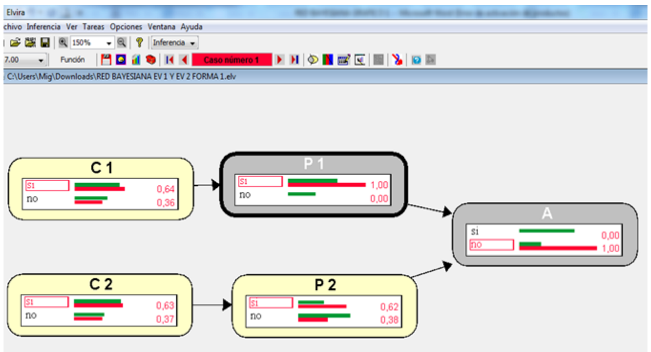

FIGURA 11. Probabilidad *a posteriori* – RB 3 – P (¬a).

Y en la Figura 11, la probabilidad a *posteriori*, si la evidencia que el alumno reprobó la materia.

D. Red bayesiana 4

Es una relación donde se utiliza las dos evaluaciones y el resultado de que si aprueba o no la materia. La variable A significa que el alumno aprueba o no la materia.

De acuerdo a los cálculos, se obtuvo la probabilidad *a priori* y posteriori, de todas las variables, y la probabilidad *a posteriori* de la variable A en este caso, se toma una evidencia futura de que un alumno apruebe o repruebe la materia y se analiza relacionándola con las demás variables.

Utilizando el programa "ELVIRA", obtenemos la siguiente Red bayesiana y sus probabilidades.

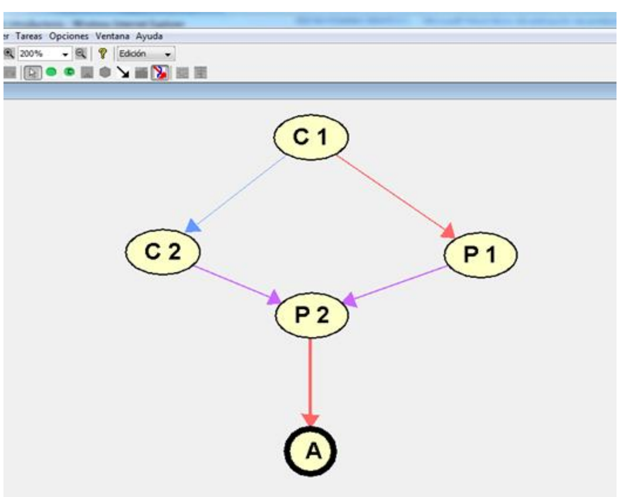

FIGURA 12. Red bayesiana 4.

Luego se calculó en el programa la probabilidad *a priori*:

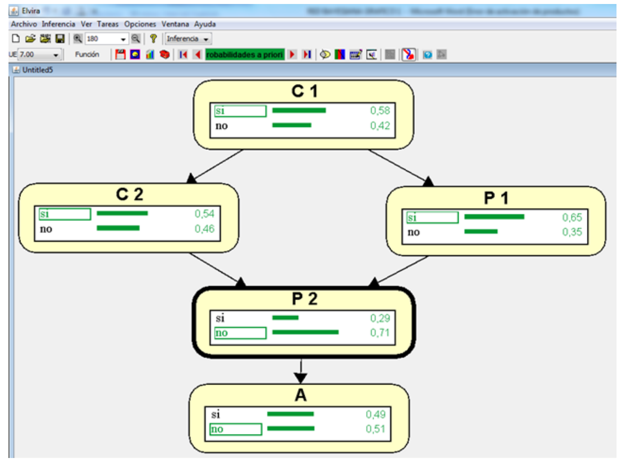

FIGURA 13. Probabilidad *a priori* de red bayesiana 4.

Luego el programa calcula la probabilidad *a posteriori* obteniendo:

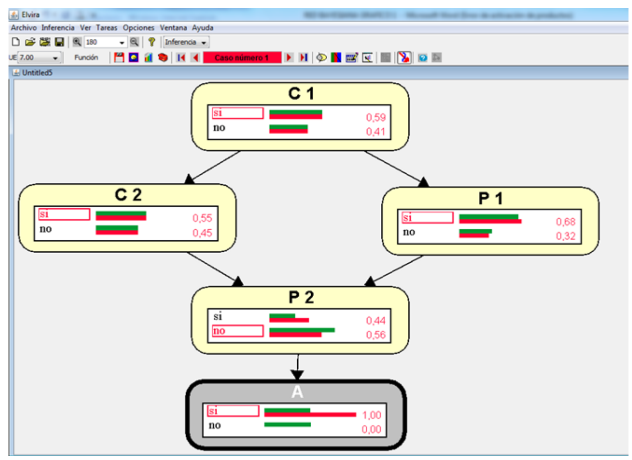

FIGURA 14. Probabilidad *a posteriori* – RB 4 – P (+a).

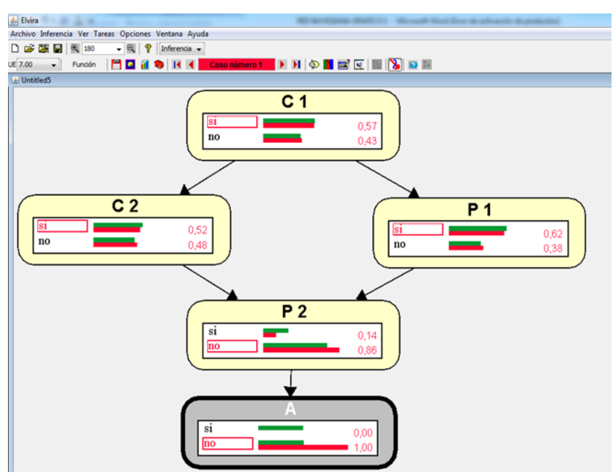

FIGURA 15. Probabilidad *a posteriori* – RB 4 – P (¬ a).

Las probabilidades *a priori*:

- P (+a) = 0,49; es decir la probabilidad de que apruebe la materia el alumno con este arreglo de red es del 49 %
- P (¬a) = 0,51; y la probabilidad de que no apruebe la materia es del 51%
- P (+c2) = 0,54; es decir que hay un 54% de probabilidad que realice las preguntas conceptuales de la evaluación 2
- P (+p2) = 0,29; es decir que hay un 29% de probabilidad que resuelva bien los problemas de la evaluación 2.
- P (+p1) = 0,65; es decir que hay un 65% de probabilidad que resuelva bien los problemas de la evaluación 1.

Las probabilidades *a posteriori*:

- P (+p2/+a) = 0,44; es decir la probabilidad de que un alumno resuelva bien los problemas de la evaluación 2, teniendo como evidencia que este haya aprobado ya la materia.
- P (+p1/+a) = 0,68; es decir la probabilidad de que un alumno resuelva bien los problemas de la evaluación 1, teniendo como evidencia que este haya aprobado ya la materia.
- P (+c2/+a) = 0,55; es decir la probabilidad de que un alumno haya resuelto bien las preguntas conceptuales de la evaluación 2, teniendo como evidencia que este haya aprobado ya la materia.
- P (+c1/+a) = 0,59; es decir la probabilidad de que un alumno haya resuelto bien las preguntas conceptuales de la evaluación 1, teniendo como evidencia que este haya aprobado ya la materia.

Si la evidencia es que el alumno no aprobó la materia tenemos las siguientes probabilidades:

- P (+p2/¬ a) = 0,14; es decir la probabilidad de que un alumno haya resuelto bien los problemas de la evaluación 2, teniendo como evidencia que este no haya aprobado ya la materia.
- P (+p1/¬ a) = 0,62; es decir la probabilidad de que un alumno haya resuelto bien los problemas de la evaluación 1, teniendo como evidencia que este no haya aprobado ya la materia.
- P (+c2/¬ a) = 0,52; es decir la probabilidad de que un alumno haya resuelto bien las preguntas conceptuales de la evaluación 2, teniendo como evidencia que este no haya aprobado ya la materia.
- P (+c1/¬ a) = 0,57; es decir la probabilidad de que un alumno haya resuelto bien las preguntas conceptuales de la evaluación 1, teniendo como evidencia que este no haya aprobado ya la materia.

E. Red bayesiana 5

Es una red formada con las agrupaciones de los estudiantes de acuerdo a su desempeño de la primera evaluación y relacionadas con la variable del desempeño de la segunda evaluación, en este caso son las variables *G* y *H* respectivamente. Anteriormente se explicó la forma en que se los agrupo a los estudiantes, por lo que cada variable tiene 3 posibilidades que son en el programa alto, medio y bajo, que significan agrupaciones con excelentes, buenas y malas calificaciones.

Resolviendo con el programa "ELVIRA", obtenemos la siguiente Red Bayesiana y sus probabilidades.

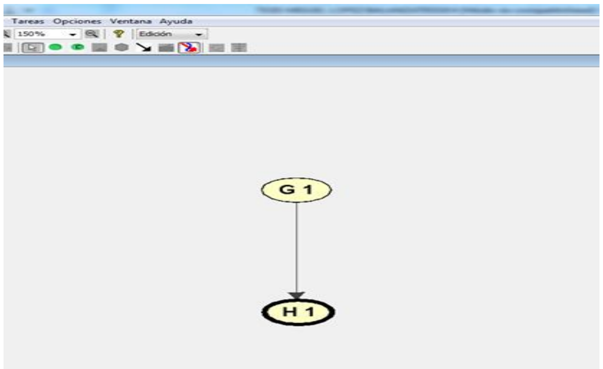

FIGURA 16. Red bayesiana 5.

Luego se calculó en el programa la probabilidad *a priori*:

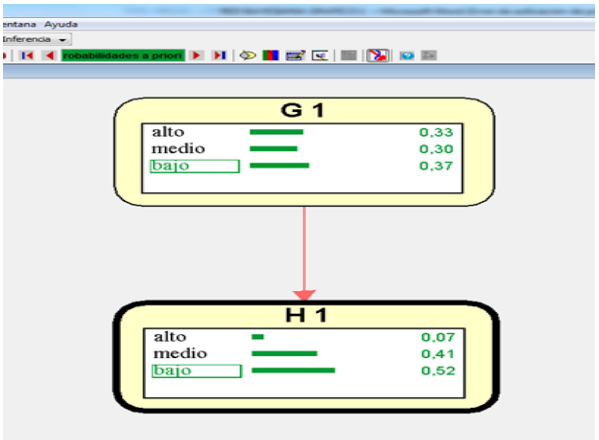

FIGURA 17. Probabilidad *a priori* de red bayesiana 5.

Luego el programa calcula la probabilidad *a posteriori* obteniendo:

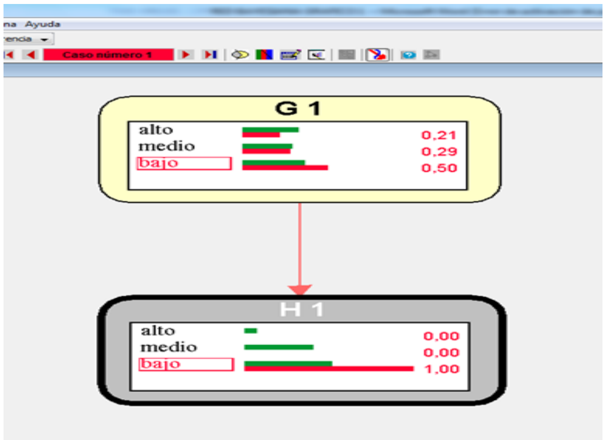

FIGURA 18. Probabilidad *a posteriori* – RB 5 – P(+h1bajo).

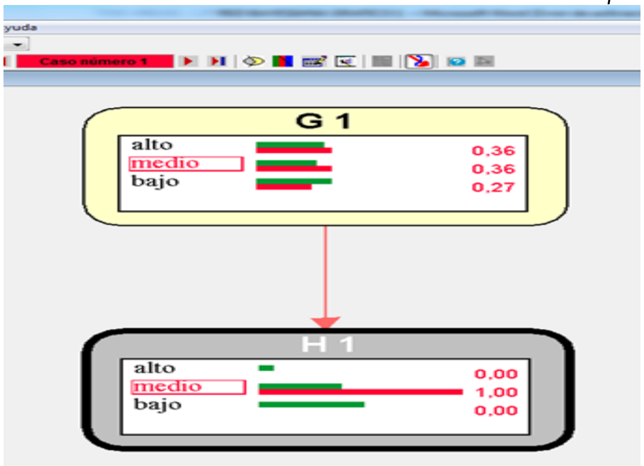

FIGURA 19. Probabilidad *a posteriori* – RB – P (+h1medio).

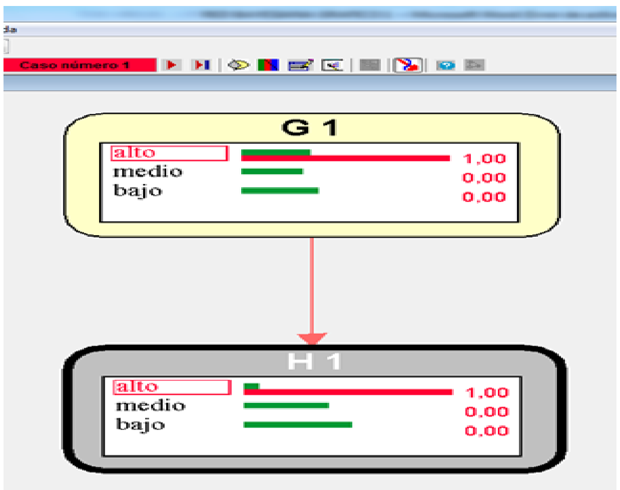

FIGURA 20. Probabilidad *a posteriori* – RB 5 – P (+h1 alto).

Las probabilidades *a priori* del grupo H son:

P (+h1 alto) = 0,07
P (+h1 medio) = 0,41
P (+h1 bajo) = 0,52

Las probabilidades *a posteriori* son:

P (+g1 alto/h1 bajo) = 0,21
P (+g1 medio/h1 bajo) = 0,29
P (+g1 bajo/h1 bajo) = 0,50
P (+g1 alto/h1 medio) = 0,36
P (+g1 medio/h1 medio) = 0,36
P (+g1 bajo/h1 medio) = 0,27
P (+g1 alto/h1 alto) = 1
P (+g1 medio/h1 alto) = 0
P (+g1 bajo/h1 alto) = 0

F. Red bayesiana 6

Red formada por los dos grupos de las evaluaciones y relacionadas con que el estudiante aprueba o no la materia.

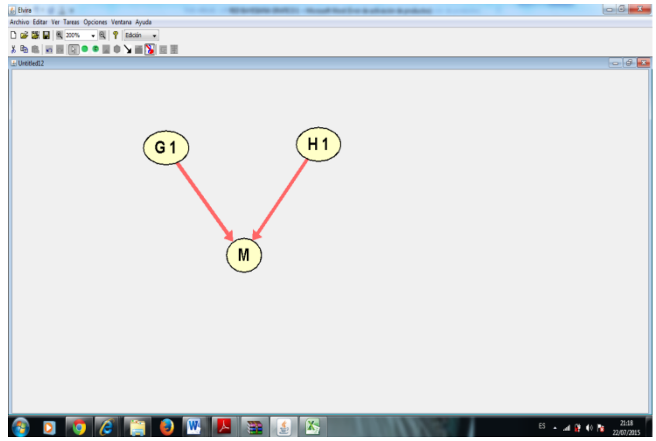

FIGURA 21. Red bayesiana 6.

Resolviendo con el programa "ELVIRA", obtenemos la siguiente red bayesiana y sus probabilidades.

Luego se calculó en el programa la probabilidad *a priori*:

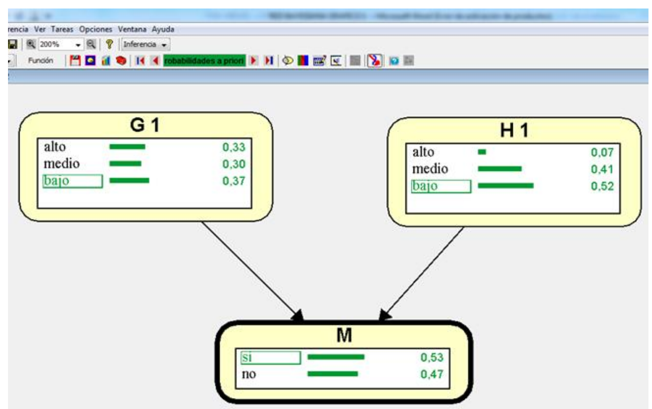

FIGURA 22. Probabilidad *a priori* de red bayesiana 6.

Luego el programa calcula la probabilidad *a posteriori* obteniendo:

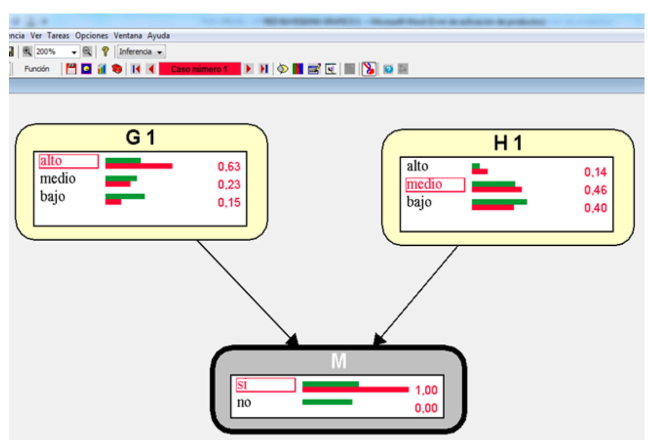

FIGURA 23. Probabilidad *a posteriori* – RB 6 – P (+a).

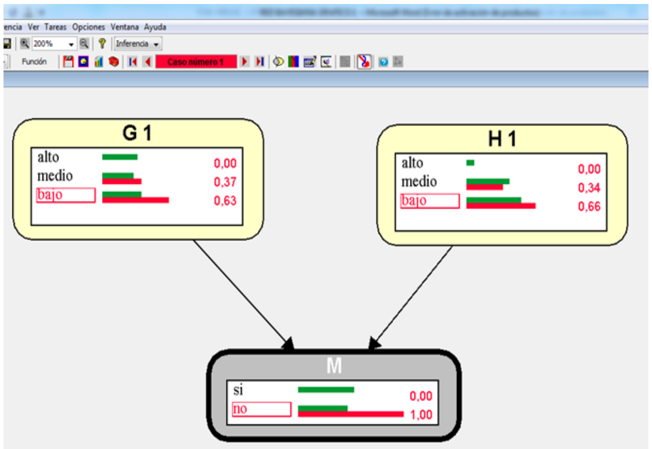

FIGURA 24. Probabilidad *a posteriori* – RB 6 – P (¬ a).

Las probabilidades *a priori* son:

P (+a) = 0,53; es la probabilidad de que el estudiante apruebe la materia, tomando en cuenta el resultado de las notas en los dos grupos formados.

Las probabilidades *a posteriori* son:

P (+g1 alto/+a) = 0,63
P (+h1 alto/+a) = 0,14
P (+g1 alto/¬ a) = 0
P (+h1 alto/¬ a) = 0
P (+g1 medio/+a) = 0,23
P (+h1 medio/+a) = 0,46
P (+g1 medio/¬ a) = 0,37
P (+h1 medio/¬ a) = 0,34
P (+g1 bajo/+a) = 0,15
P (+h1 bajo/+a) = 0,40
P (+g1 bajo/¬ a) = 0,63
P (+h1 bajo/¬ a) = 0,66

G. Red bayesiana 7 y 8

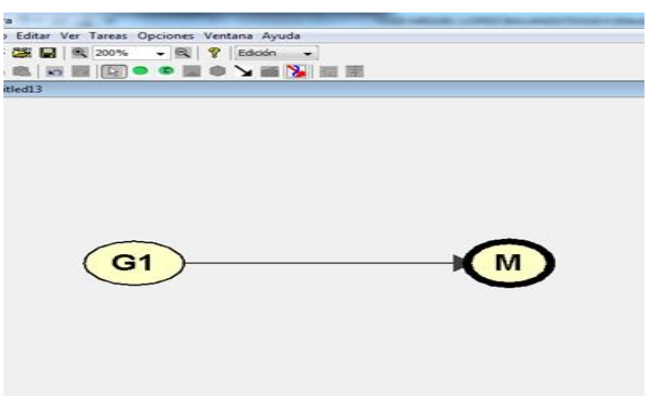

FIGURA 25. Red bayesiana 7.

Estas redes están formadas para relacionar los grupos *G* y *H*, que son los resultados de las evaluaciones 1 y 2 respectivamente, con el resultado final (variable *M*), si aprobó (+*m*) o si no aprobó (¬*m*) la materia; cabe indicar que se relacionarán con sus tres posibilidades de resultados, alto, medio y bajo.

En la Figura 25 puede observarse lo que con el programa “ELVIRA” se pudo calcular.

Las probabilidades *a priori*:

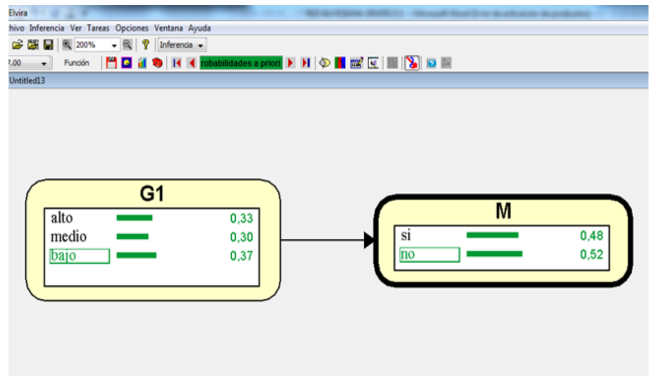

FIGURA 26. Probabilidad *a priori* de red bayesiana 7.

Las probabilidades *a posteriori*:

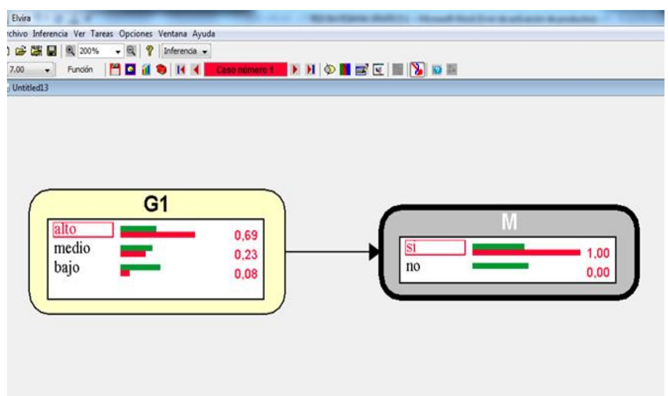

FIGURA 27. Probabilidad *a posteriori* – RB 7 – P (+m).

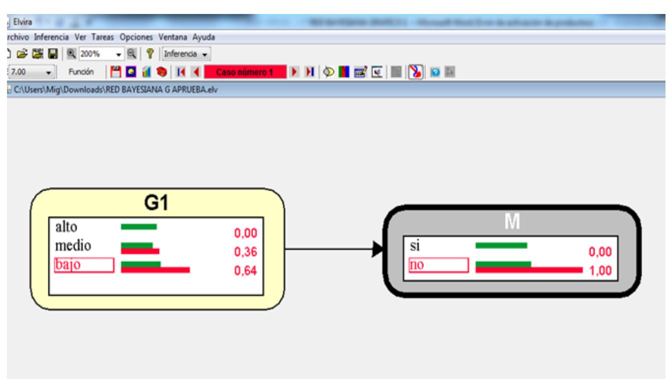

FIGURA 28. Probabilidad *a posteriori* – RB 7 – P (¬ m).

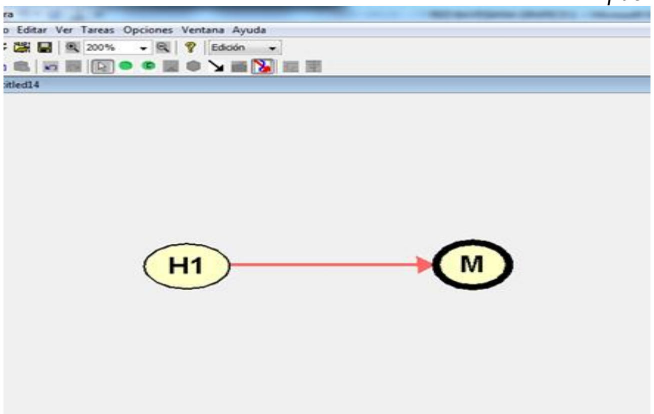

FIGURA 29. Red bayesiana 8.

Las probabilidades *a priori*:

Con el grupo G (RB-7):

P (+m) = 0,48

Con el grupo H (RB-8):

P (+m) = 0,48

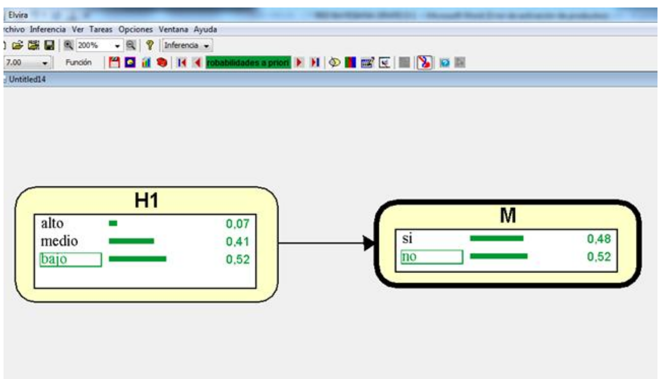

FIGURA 30. Probabilidad *a priori* de red bayesiana 8.

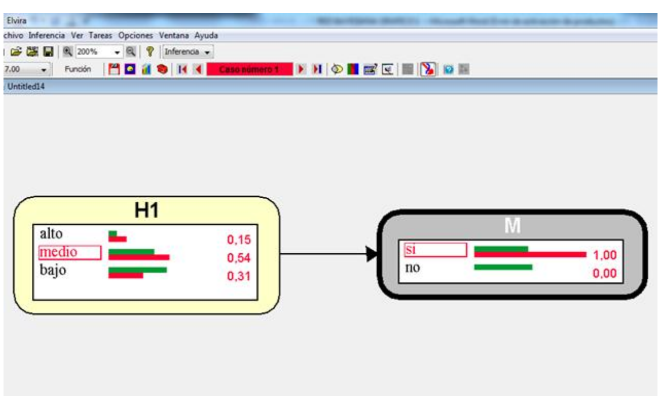

FIGURA 31. Probabilidad *a posteriori* – RB 8 – P (+m).

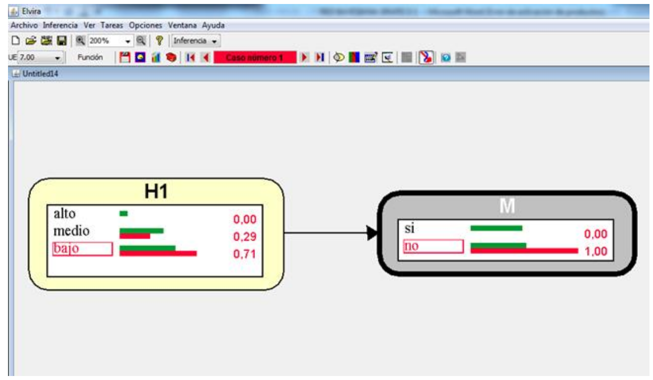

FIGURA 32. Probabilidad *a posteriori* – RB 8 – P (¬ m).

Probabilidades *a posteriori* (RB-7) y (RB-8):

P (+g1 alto/+m) = 0,69
P (+h1 alto/+m) = 0,15
P (+g1 alto/¬ m) = 0
P (+h1 alto/¬ m) = 0
P (+g1 medio/+m) = 0,23
P (+h1 medio/+m) = 0,54
P (+g1 medio/¬ m) = 0,36
P (+h1 medio/¬ m) = 0,29
P (+g1 bajo/+m) = 0,08
P (+h1 bajo/+m) = 0,31
P (+g1 bajo/¬ m) = 0,64
P (+h1 bajo/¬ m) = 0,71

V. CONCLUSIONES

En este trabajo, hemos realizado una metodología para poder analizar el rendimiento académico de los estudiantes de una forma muy sencilla, utilizando el software para redes Bayesianas "ELVIRA", el cual está disponible para cualquier persona en la web.

Cabe indicar que, las ocho redes bayesianas diseñadas y analizadas pueden ser incrementadas y mejoradas, dependiendo de la cantidad de información o datos recopilados.

En la RB 4 se utilizan las variables de tal forma que, se relacionan los conocimientos concatenados entre las dos evaluaciones, siendo el resultado final, la aprobación o no de la materia.

Cabe demostrar que no es toda la información completa para relacionarla con el objetivo final, pero sí infiere con los resultados y se llega a algo muy cercano a la realidad, tal es así, que la probabilidad de aprobar la materia nos da el 49%, información muy cercana a la realidad final que fue del 48%.

En los resultados de esta red se observa que, si un alumno aprobó la materia, la probabilidad de conocer las preguntas conceptuales de la evaluación 2 y la evaluación 1 es respectivamente del 55% al 59%, y de resolver bien los problemas está entre el 44% y 68%.

Con los resultados anteriores queda claro que, el estudiante puede aprobar la materia con un promedio medio de conocimientos y habilidades en la resolución de problemas. Que con un incremento en la dificultad de los problemas planteados se incrementa su influencia sobre el resultado final en el conocimiento de las preguntas conceptuales.

Se recomienda que de ser utilizado este trabajo, se lo realice en las primeras evaluaciones del año, de forma individual y colectiva para poder lograr el verdadero objetivo, que es mejorar el desempeño de sus alumnos. De esta manera, se les clasifica y se les agrupa para trabajar en pares o grupos, donde lidere alguien que haya podido demostrar un mejor desempeño, y sea capaz de transmitir sus conocimientos, y que sus resultados fueron más positivos que sus inferencias.

En lo que respecta a las redes diseñadas, sería mejor si se aplicaran tres evaluaciones, para tener datos más confiables y un verdadero comportamiento del aprovechamiento de los estudiantes.

En los resultados obtenidos en las dos primeras redes, demostramos que, en la materia de Física es muy importante conocer la teoría, para poder aplicarla en la resolución de problemas.

Las redes bayesianas son de gran utilidad para analizar datos aplicados a la educación y se convierten en una gran herramienta de ayuda para tomar decisiones acertadas que podrían mejorar el desempeño de los estudiantes, es necesario continuar con las investigaciones del tema, ya que tiene muchas aplicaciones y despierta el interés investigativo, a pesar de ser extenso. Se espera que este trabajo sea el comienzo de más investigaciones para un mayor aprendizaje.

REFERENCIAS

[1] Almond, R. & Shute, V., *Bayesian networks: A teacher´s view*, International Journal of Approximate Reasoning **50**, 450-460 (2008).

[2] Marín, Á., *Sistemas expertos, redes bayesianas y sus aplicaciones*, Semana ESIDE, Facultad de Ingeniería de la Universidad de Deusto, Bilbao, España (2005).

[3] Valenzuela-Rendón, M., *Redes bayesianas,* (Centro de Sistemas Inteligentes, Tecnológico de Monterrey, México, 2008).

[4] Díez, F. J., *Introducción al razonamiento aproximado*, (Universidad Nacional de Educación a Distancia UNED, Madrid, 2005).

[5] Cortez Bohigas, Ma. del Mar., *Diccionario de las Ciencias de La Educación,* (Mar de Plata, *ibídem*, 2008).

[6] Tejedor, F. y García-Valcárcel, A., *Causas del bajo rendimiento del estudiante universitario (en opinión de los profesores y alumnos). Propuesta de mejora en el marco del EEES*, Revista de Educación **342**, 443-473 (2007).

[7] Chain, R., Cruz Ramírez, N., Martínez Morales, M. y Jácome, N., *Examen de selección y probabilidades de éxito escolar en estudios superiores. Estudio en una universidad pública estatal mexicana*, Revista Electrónica de Investigación Educativa **5**, 99-116 (2003)
. http://redie.uabc.mx/vol5no1/contenido-chain.html.