

Modelo Bayesiano para la Optimización de Aseguradoras de Autos

Francisco Javier Tapia Moreno
Héctor Antonio Villa Martínez

Universidad de Sonora
e-mail: ftapia@mat.uson.mx, hvilla@mat.uson.mx

Resumen

La toma de decisiones en los distintos niveles de las organizaciones cada vez es de mayor complejidad, dadas las crecientes restricciones de disponibilidad de todo tipo de recursos. En una compañía de seguros para automóviles, los gerentes se ven en la necesidad de decidir si otorgar, renovar, cancelar o rechazar una póliza de seguros a un cliente. Existen grandes cantidades de datos almacenados en distintos formatos en tales compañías, las cuales pueden ser usadas para ayudarse a decidir si otorgar o no la póliza de seguros a un cliente específico. Es decir, dichas aseguradoras pueden explotar íntegramente los datos acerca de los patrones de compra y comportamiento de sus clientes, entender mejor las motivaciones de los mismos, ayudando con ello a la reducción de los costos y del fraude, optimizando así los recursos de la compañía.

Este artículo presenta el diseño de un modelo bayesiano que integra las características principales del conductor y de su auto, y que sirve para predecir si se otorga o no la póliza de seguros al cliente. Se usan redes bayesianas para prever la probabilidad de un accidente automovilístico, basándose en los datos del cliente disponibles en el tiempo de la solicitud de compra de la póliza. Las redes bayesianas son una herramienta útil en la estimación de probabilidades ante incertidumbres y nuevas evidencias. Este modelo puede optimizar la utilidad de una compañía de seguros de automóviles, porque éste sugiere la autorización o no, de la póliza mencionada. El modelo se valida mediante una muestra aleatoria de 1,200 automovilistas con sus respectivos automóviles.

1 Introducción

En el rubro de seguros y salud privada, la minería de datos puede ser un apoyo fundamental para obtener ventajas competitivas. Las compañías aseguradoras pueden explotar al máximo los datos acerca de los patrones de compra y comportamiento de sus clientes, logrando así entender mejor las motivaciones de los mismos y colaborando con ello a la reducción del fraude. Los análisis realizados mediante las técnicas de la minería de datos pueden ayudar a las compañías mencionadas a reducir costos, aumentar las utilidades, conservar a los clientes actuales y atraer nueva clientela.

En los últimos años las tecnologías de la información, aplicada al sector de seguros y salud privada han evolucionado considerablemente. Innovaciones tecnológicas tales como la minería de datos [6] y el data warehouse [2], han reducido los costos de almacenamiento, acceso y procesamiento de datos permitiendo que las compañías aseguradoras puedan llevar a cabo una gestión personalizada de cada cliente, adaptando sus primas y servicios a las necesidades de cada uno de ellos, y mediante su historial correlacionar, en un futuro, las

primas con los comportamientos de los clientes que aumentan el riesgo. Uno de los problemas a que se enfrentan las compañías aseguradoras, es otorgar o no la póliza al cliente.

Este artículo presenta el diseño de un modelo bayesiano que integra las características principales del conductor y de su auto, y que sirve para predecir si se otorga o no la póliza de seguros al cliente. Se usan redes bayesianas para prever la probabilidad de un accidente automovilístico, basándose en los datos del cliente disponibles en el tiempo de la solicitud de compra de la póliza. Las redes bayesianas son una herramienta útil en la estimación de probabilidades ante incertidumbres y nuevas evidencias. Este modelo puede optimizar la utilidad de una compañía de seguros de automóviles, porque éste sugiere la autorización o no de la mencionadas póliza. El resto del documento está compuesto como sigue. La sección 2 comenta lo referente a los seguros de automóviles. La sección 3 muestra la metodología usada para integrar el modelo. La sección 4 presenta el modelo y las variables que lo conforman. En la sección 5 se valida el modelo mediante una muestra aleatoria de 1,200 clientes. Por último, la sección 6 expone las conclusiones y las pesquisas futuras.

2 Seguro de automóviles

De acuerdo con el sitio dmv.org [3], el seguro de automóvil está diseñado para proteger a los conductores y pasajeros de cualquier dificultad financiera después de un accidente u otro incidente relacionado con vehículos. Sin seguro de automóvil, las víctimas de un accidente de auto por lo general no pueden obtener ningún tipo de compensación. En el caso de que la víctima demande al conductor culpable, éste puede a menudo enfrentarse a gastos enormes.

En algunos países, como Estados Unidos, es obligatorio contar con un seguro de automóvil. De esta forma, el seguro ayuda al gobierno a proteger a los conductores de los enormes costos que se pueden generar con un accidente. En particular, el estado se asegura de que tanto el conductor culpable pueda cubrir los costos de las lesiones o daños a la propiedad resultantes del accidente, sin perder grandes sumas de dinero en un pleito, y de que las víctimas del accidente puedan buscar la atención médica requerida, pagar los gastos funerarios, y/o arreglar su propiedad sin tener que pagar nada de su bolsillo.

El tipo más común de seguro es el seguro de responsabilidad civil o de daños a terceros, que cubre los daños a la otra persona cuando el conductor asegurado tiene la culpa. Sin embargo, existen muchos otros tipos de coberturas que ofrecen una protección adicional respecto a los costos de las lesiones y daños a la propiedad para el conductor y sus pasajeros. Estas coberturas incluyen:

1. Cobertura amplia. Paga por daños al automóvil del conductor (menos un deducible) aunque éste haya sido culpable.
2. Cobertura de no seguro / seguro insuficiente en lesiones corporales del conductor. Ayuda a pagar los gastos del conductor asegurado cuando ha sido golpeado por otro conductor sin seguro o con seguro insuficiente.

3. Cobertura de no seguro / seguro insuficiente en propiedad. Ayuda a cubrir los daños al automóvil del conductor asegurado si ha sido golpeado por otro conductor sin seguro o con seguro insuficiente.
4. Cobertura de pagos de gastos médicos. Ayuda a pagar los gastos médicos del conductor asegurado y de sus pasajeros, independientemente de quien fue encontrado culpable.

Al igual que todos los seguros, las tasas de seguros de automóviles se evalúan en función del riesgo. Esto significa que el proveedor de seguros calcula la probabilidad de hacer una reclamación y basa su prima en dicho nivel de riesgo. Las compañías de seguros de automóvil han encontrado que ciertos factores afectan estadísticamente las probabilidades de tener un accidente, por lo que usan esos factores para determinar sus primas. Estos factores incluyen (pero no se limitan a):

1. El historial de conducción.
2. Número de multas.
3. Número de accidentes.
4. Marca del vehículo, modelo y año.
5. Edad, sexo y estado civil del conductor.
6. La ocupación del conductor.
7. Historial de crédito (en algunos países).
8. Historial de reclamaciones.
9. Cantidad de kilómetros que el conductor maneja al mes o al año.

Esto implica que un conductor de bajo riesgo obtendrá una prima más baja que un conductor de alto riesgo.

3 Redes bayesianas

Las redes bayesianas (RBs), también conocidas como redes de creencia, pertenecen a la familia de los modelos gráficos probabilísticos (MGs). Estas estructuras gráficas se usan para representar conocimiento acerca de un dominio incierto. En particular, cada nodo del gráfico representa una variable aleatoria, mientras que las aristas entre nodos representan dependencias probabilísticas entre las variables aleatorias. Estas dependencias condicionales en la gráfica son a menudo estimadas por medio de métodos estadísticos y computacionales. Por lo tanto, las RBs combinan principios de teoría de gráficas, ciencias computacionales y estadística. Las MGs con aristas no dirigidas generalmente se les denominan campos aleatorios de Markov o redes de Markov. Estas redes proporcionan una definición simple de

independencia entre cualesquiera dos nodos distintos, basado en el concepto de una cubierta de Markov. Las redes de Markov son conocidas en áreas del conocimiento tales como la física estadística y visión por computadora [9, 8]. Las RBs pertenecen a otra estructura de MG conocida como gráfica acíclica dirigida (GAD) la cual es muy conocida en el área de la estadística, en el aprendizaje de máquina y en las sociedades de inteligencia artificial. Las RBs son matemáticamente rigurosas e intuitivamente entendibles y llevan a cabo una representación efectiva y el cálculo de la distribución de probabilidad conjunta (DPC) sobre un conjunto de variables aleatorias [11].

La estructura de una GAD se define por dos conjuntos: el conjunto de nodos (vértices) y el conjunto de aristas dirigidas. Los nodos representan variables aleatorias y se indican por medio de círculos etiquetados con los nombres de las variables. Las aristas representan la dependencia directa entre las variables y se señalan con flechas entre los nodos. En particular, una arista del nodo X_i al nodo X_j representa una dependencia estadística entre las variables correspondientes. Por lo tanto, la flecha indica que el valor que toma la variable X_j depende del valor que tome la variable X_i , o en términos generales, significa que la variable X_i "influye" en X_j . En este caso, al nodo X_i se le denomina el *padre* de X_j y de manera similar, a X_j se le llama el *hijo* de X_i . Una extensión de dichos términos genealógicos se utiliza a menudo para definir los conjuntos de "descendientes" (el conjunto de nodos a los que se puede llegar por una ruta directa desde el nodo antecesor) o los nodos "antecesores" (el conjunto de nodos desde los cuáles el nodo puede ser alcanzado en una ruta directa) [5]. La estructura de la gráfica acíclica garantiza que no hay ningún nodo que pueda ser su propio antepasado o su propio descendiente. Tal condición es de vital importancia para la factorización de la probabilidad conjunta de una colección de nodos que se observa en la relación 1. Hay que tener en cuenta que aunque las flechas representan conexiones causales directas entre las variables, el *proceso de razonamiento* puede operar en redes bayesianas o RBs mediante la propagación de la información en cualquier dirección [12].

Una RB refleja una afirmación de independencia condicional simple. Es decir, que cada variable es independiente de sus no descendientes en la gráfica, dado el estado de sus padres. Esta propiedad se utiliza para reducir, a veces de manera significativa, el número de parámetros que se requieren para caracterizar la densidad de probabilidad conjunta (DPC) de las variables. Esta reducción proporciona una solución eficiente para calcular las probabilidades a posteriori, dada la evidencia [13, 4, 1].

Además de la estructura GAD, que a menudo se considera como la parte "cualitativa" del modelo, es necesario especificar los parámetros "cuantitativos" del mismo. Los parámetros se describen de una manera consistente con una propiedad Markoviana, donde la distribución de probabilidad condicional (DPC) en cada nodo sólo depende de sus padres. Para las variables aleatorias discretas, esta probabilidad condicional es, a menudo, representada por una tabla que enumera la probabilidad local de que un nodo hijo tome cada uno de los valores posibles (para cada combinación de los valores de sus padres). La distribución conjunta de una colección de variables se puede determinar únicamente por estas tablas de probabilidad condicional locales (TPCs).

Una red bayesiana B es una gráfica acíclica con anotaciones que representa una DPC

sobre un conjunto de variables aleatorias V . La red se define como $B = \langle G, \Theta \rangle$, donde G es la GAD cuyos nodos X_1, X_2, \dots, X_n representan variables aleatorias y cuyas aristas representan las dependencias directas entre esas variables. La gráfica G codifica las hipótesis de independencia, por medio de cada variable X_i , la cual es independiente de sus descendientes dados sus padres en G . La segunda componente (Θ) indica el conjunto de parámetros de la red. Este conjunto contiene los parámetros $\Theta_{x_i/\pi_i} = P_B(x_i|\pi_i)$ para cada intento x_i de X_i condicionado a π_i , el conjunto de padres de X_i en G . En consecuencia, B indica una única DPC sobre V que es:

$$P_B(X_1, X_2, \dots, X_n) = \prod_{i=1}^n P_B(X_i|\pi_i) = \prod_{i=1}^n \theta_{X_i|\pi_i} \quad (1)$$

Si X_i no tiene padres, su distribución de probabilidad local se denomina incondicional, de otra manera será condicional. Si la variable representada por un nodo es analizada, entonces el nodo es un nodo evidencia, de otro modo al nodo se le considera como perdido o latente [12, 7].

4 Construcción del modelo

Para construir una RB para un dominio de aplicación específico, se requieren cubrir los tres siguientes pasos: 1) Identificar cada una de las variables importantes en el dominio de interés, 2) Determinar el conjunto de los posibles resultados o estados. Es decir, construir la parte cualitativa que consiste en identificar las independencias entre las variables involucradas y expresarlas en una GAD que codifique las afirmaciones de las independencias condicionales, 3) Llenar el modelo con probabilidades numéricas.

Las variables aleatorias participantes son: a) Información sobre cliente: edad, sexo, problemas (físicos, visuales, etc.), antecedentes (alcoholismo, drogas, penales, etc.), b) Información sobre el automóvil: modelo, promedio de kilómetros recorridos al año, historial (participación en siniestros, reporte delictivo, documentación falsa, etc.) y c) Status de la póliza: decisión del sistema de cómputo. La tabla 1 muestra las variables del modelo y sus estados, cumpliendo así con el primer paso de los tres requeridos.

Para cubrir el paso 2, en la figura 1 se muestra la GAD donde se identifican las independencias entre las variables mencionadas en la tabla 1, y que codifica las afirmaciones de las independencias condicionales. La estructura de la GAD de la figura 1 y la relación 1 permiten calcular:

- i Probabilidades de perfiles del conductor tales como: $P(\text{conductor con problemas visuales})$, $P(\text{menor de 27 años; mujer; con problemas visuales, físicos o mentales; con alcoholismo o drogas, conductor no apto})$.
- ii Probabilidades sobre perfiles del automóvil tales como: $P(\text{auto ya siniestrado})$; $P(\text{más de 20,000 kilómetros al año})$ y $P(\text{modelo reciente; sin antecedentes de siniestro; más de 20,000 kilómetros por año; automóvil admitido})$.

Tabla 1: Variables del modelo y sus estados

VARIABLES DEL CONDUCTOR	ESTADOS DE LAS VARIABLES
1. Edad	18-27 (JO); 27 o más (AM)
2. Sexo	Hombre (HO); Mujer (MU)
3. Problemas físicos, visuales, etc.	Con problemas (CP); Sin problemas (SP)
4. Antecedentes (alcoholismo, drogas, penales, etc.)	Con antecedentes (AP); Sin antecedentes (SA)
5. Situación del conductor	Aprobado (CA); No aprobado (CN)
VARIABLES DEL AUTOMÓVIL	ESTADOS DE LAS VARIABLES
6. Modelo del automóvil	Modelo reciente (MR) (2012 a 2016); Otro modelo (OM)
7. Promedio de kilómetros recorridos al año	20,000 o menos (IV); Más de 20,000 (SV)
8. Historial (participación en siniestros, acciones delictivas, etc.)	Antecedentes (HP); Sin antecedentes (HN)
9. Situación del automóvil	Admitido (AD); No admitido (NA)
VARIABLE DE LA SOLICITUD DE PÓLIZA	ESTADOS DE LA VARIABLE
10. Decisión del sistema de cómputo	Aprobada (PA); Rechazada (PR)

- iii Probabilidades conjuntas de perfiles conductor-automóvil-estado de la póliza tales como: $P(\text{menor de 27 años; mujer; con problemas visuales, físicos o mentales; con alcoholismo o drogas, conductor no apto; modelo reciente, sin antecedentes de siniestro; más de 20,000 kilómetros por año; automóvil admitido; póliza rechazada})$.
- iv Probabilidades condicionales tales como: $P(\text{aceptar póliza} \mid \text{más de 27 años; mujer; sin problemas; sin antecedentes})$ y $P(\text{aceptar póliza} \mid \text{menor de 27 años; hombre; con problemas visuales; con problemas de alcoholismo, conductor apto; modelo reciente, sin antecedentes de siniestro; más de 20,000 kilómetros por año; automóvil admitido})$ y,
- v Probabilidades a posteriori tales como: $P(\text{aceptar póliza})$ o $P(\text{conductor no apto})$.

Una vez determinado el perfil del conductor y del automóvil, el sistema de cómputo conjunta ambas probabilidades y este cálculo, determina la credibilidad del sistema que un conductor y un automóvil específico pertenezcan a cierto perfil, y en base a este perfil decide rechazar o aceptar la póliza al cliente.

El tercer y último paso, consiste en llenar el modelo con probabilidades numéricas. Para cubrir este paso, se consideran las estadísticas obtenidas de una muestra de 1,200 conductores, con sus respectivos automóviles, de acuerdo a la siguiente metodología: las funciones de probabilidad asociadas a cada uno de los nodos independientes se generan gradualmente mediante los resultados que se obtienen de las apreciaciones realizadas a los conductores y a sus respectivos automóviles. Se asignaron a los conductores y automóviles, características



Figura 1: Gráfica acíclica dirigida que recopila las independencias entre las variables de la tabla 1

aleatorias de las señaladas en la tabla 1. Se calcularon las frecuencias relativas y se propusieron como las probabilidades previas asociadas a cada uno de los nodos padres y nodos hijos, las cuales representan las funciones de distribución de probabilidad de cada una de las variables y que son mostradas en la tabla 2. Estas funciones son las encargadas de generar las métricas para lograr pronosticar el perfil del conductor y de su automóvil, y así decidir la aprobación o rechazo de la póliza de seguro.

5 Resultados en la implementación

Hemos utilizado el software MSBNX [10] para obtener los datos que se presentan en la figura 2. Es decir, las probabilidades posteriores de cada una de las variables involucradas en el modelo. Con estos resultados, podemos obtener las probabilidades de ocurrencia de cada uno de los 1,024 perfiles de conductor-auto-póliza posibles. Para ilustrar la funcionalidad del modelo se presentan algunos ejemplos.

Si se desea saber la probabilidad que conductor tenga problemas visuales, el modelo nos indica que $P(\text{conductor con problemas visuales}) = 0.43$ (ver figura 2).

Por otro lado, la probabilidad de que un conductor sea mujer menor de 27 años, con problemas visuales, físicos o mentales alcoholismo o drogas y que sea considerado un conductor apto de acuerdo a los resultados de la figura 2 y usando la nomenclatura de la tabla 1 es:

$$P(JO, MU, CP, AP, CN) = 0.1600 \times 0.5100 \times 0.4300 \times 0.4300 \times 0.6291 = 0.00949$$

Es decir, existe aproximadamente un 1% de posibilidades de que un conductor presente las características arriba mencionadas.

Tabla 2: Funciones de distribución de probabilidad de cada variable

(a)		(b)	
Sexo	Probabilidad	Modelo del auto	Probabilidad
Femenino	0.20	Reciente	0.617
Masculino	0.80	No reciente	0.383

(c)		(d)	
Edad	Probabilidad	Kilometraje por año	Probabilidad
18-27	0.61	Más de 10,000	0.646
Más de 27	0.39	Menos de 10,000	0.354

(e)		(f)	
Problemas (alcohol o drogas)	Probabilidad	Antecedentes	Probabilidad
Si	0.4775	Si	0.38
No	0.5225	No	0.62

(g)	
Problemas (físicos o visuales)	Probabilidad
Si	0.28
No	0.72

Ahora, si se desea saber la probabilidad de aceptar la póliza a un cliente sabiendo que es hombre, menor de 27 años, con problemas de alcoholismo, con un modelo reciente, sin antecedentes de siniestro, con más de 20 mil kilómetros por año, con automóvil no admitido, por la definición de probabilidad condicional, ésta es:

$$\begin{aligned} P(PA|(JO, HO, CP, SA, CA, AP, HN, SV, AD)) &= \\ &= \frac{0.3954 \times 0.5100 \times 0.8400 \times 0.4300 \times 0.4300 \times 0.6291 \times 0.6100 \times 0.5800 \times 0.5700 \times 0.6291}{0.3954} \\ &= 0.0063 \end{aligned}$$

Es decir, la probabilidad que una póliza sea aceptada a un solicitante con las características antes mencionadas es del 0.63%. De manera similar, se pueden calcular las probabilidades mencionadas en los puntos ii, iii y v de la sección 4

También, este modelo es útil para realizar predicciones en caso de incertidumbre ante datos perdidos. Por ejemplo, se puede calcular la probabilidad de que un solicitante tenga las características siguientes: menor de 27 años, hombre, con problemas visuales, con problemas de alcoholismo, conductor apto, modelo reciente, sin antecedentes de siniestro, más de 20,000 kilómetros por año, automóvil admitido dado que se aceptó su póliza. Este cálculo se hace usando el teorema de Bayes [14][p. 153] y se obtiene:

$$P((JO, HO, CP, SA, CA, AP, HN, SV, AD)|PA) = \frac{0.0063 \times 0.0063}{0.3954} = 0.0001$$

De acuerdo con el modelo, hay una probabilidad del 0.01% que un solicitante al que se le aceptó la póliza tenga las características antes señaladas.

Otro caso es cuando se tienen datos incompletos del solicitante. Por ejemplo, sólo se sabe que es hombre mayor de 27 años y sin antecedentes penales y se tiene la duda de aceptar o no la póliza. La probabilidad que ofrece el modelo de aceptar la solicitud y otorgar dicha póliza, dadas las características anteriores es:

$$P(PA/(AM, HO, SP)) = \frac{0.3954 \times 0.39 \times 0.62}{0.57 \times 0.49 \times 0.57} = 0.6005$$

Este resultado ofrece el beneficio de la duda de aceptar la solicitud al cliente y entregarle su póliza de seguros.

6 Conclusiones

Se ha construido un modelo que es útil para pronosticar si se otorga o no la póliza de seguros para su automóvil al cliente. Esto, a consideración nuestra, puede optimizar las utilidades de las compañías aseguradoras de automóviles. El modelo bayesiano fue validado usando los resultados obtenidos de una muestra de 1,200 automovilistas con sus respectivos

Node Name	State 0	State 1
Antecedentes penales.	CP	SP
	0,4300	0,5700
Edad del Conductor; menos de 27; 27 o más.	JO	AM
	0,5100	0,4900
Participación del auto en siniestros o en acciones delictuosas.	HP	HN
	0,5800	0,4200
Kilometraje promedio por año. 20 mil o menos; más de 20 mil.	IV	SV
	0,4300	0,5700
Años de vida del automóvil (5 o menos; más de 5)	MR	DM
	0,6100	0,3900
Decision del sistema.	PA	PR
	0,3954	0,6046
Con o Sin Problemas visuales, físicos, alcohol, drogas etc.	CP	SP
	0,4300	0,5700
Sexo del Conductor	HO	MU
	0,8400	0,1600
Resultado de Revisión mecánica o de circulación al automovil.	AD	NA
	0,8481	0,1519
Aptitud del conductor.	CA	CN
	0,6291	0,3709

Figura 2: Probabilidades a posteriori que arroja el modelo.

automóviles. Se utilizó el software MSBNx para la realización de los cálculos de las probabilidades posteriores. El diseño e implementación del modelo Bayesiano dejó los siguientes resultados:

- Un grupo de relaciones de causa-efecto entre características personales del solicitante, características del automóvil a asegurar y la decisión del sistema electrónico usando el modelo.
- Un conjunto de tablas con las probabilidades previas de las variables que componen nuestro modelo que se generaron de una muestra aleatoria de 1,200 solicitantes con sus respectivos automóviles y que pueden ser utilizados para inicializar otros modelos y realizar inferencias.
- Una colección de probabilidades previas y posteriores de cada una de las variables del modelo Bayesiano, tablas de distribución de probabilidades previas de cada uno de los nodos padre y tablas de distribución de probabilidades condicionales de cada uno de los nodos hijos, todas usadas para inicializar el modelo.
- Propusimos un conjunto de funciones de distribución de probabilidades locales, que generarán las métricas de aprendizaje para los estados de las variables que conforman el modelo Bayesiano y cuyos parámetros serán determinados gradualmente a medida que las redes bayesianas sean alimentadas con los datos recientes de los solicitantes.

Este modelo puede usarse total o parcialmente para realizar diagnósticos, ya sea acerca del perfil del solicitante o del perfil del automóvil o del perfil de ambos, y para resolver casos de incertidumbre o de datos perdidos acerca de las características personales del solicitante o de su automóvil.

Como trabajo futuro se pueden agregar más variables al modelo tales como el estrato del conductor (rural o urbano), si el automóvil se encuentra protegido o no por las noches y si es conducido por más de un conductor.

Referencias

- [1] Ben-Gal, Irad, Bayesian Networks, Encyclopedia of Statistics in Quality and Reliability 1, 1800 páginas, Wiley, Hoboken, NJ, USA (2008).
- [2] Corr, Lawrence y Stagnitto, Jim, Agile Data Warehouse Design: Collaborative Dimensional Modeling, from Whiteboard to Star Schema, 328 páginas, DecisionOne Press, Devon, PA, USA (2011).
- [3] DMV.org <http://www.dmv.org> (2016). Accesado 12 de octubre de 2016.
- [4] Friedman, Nir, Geiger, Dan y Goldszmidt, Moises, Bayesian Network Classifiers, Machine Learning 29, 131-163 (1997).
- [5] Griffiths, Thomas L. y Yuille, Alan, A primer on probabilistic inference, Trends in Cognitive Sciences Supplement to special issue on Probabilistic Models of Cognition, 10(7), 1–11, (2006).
- [6] Han, Jiawei, Data Mining: Concepts and Techniques, 2nd edition, Morgan Kaufmann Publishers, Burlington, MA, USA (2006).
- [7] Henrion, Max, Some Practical Issues in Constructing Belief Networks, In Uncertainty in Artificial Intelligence 3, Kanl, L. N., Levitt, T. S., Lemmer, J. F. editores, Elsevier Science, Amsterdam, Holanda (1989).
- [8] Jensen, Finn Verner y Nielsen, Thomas Dyhre, Bayesian Networks and Decision Graphs, 2nd edition, 448 páginas, Springer, New York, NY, USA (2007).
- [9] Jordan, Michael Irwin, Learning in Graphical Models, 644 páginas, MIT Press, Cambridge, MA, USA (1999).
- [10] MSBNx – Microsoft Bayesian Network Editor, <http://research.microsoft.com/en-us/um/redmond/groups/adapt/msbnx/> (2016)
- [11] Pearl, Judea, Probabilistic Reasoning in Intelligent Systems, 552 páginas, Morgan Kaufmann, San Francisco, CA, USA (1988).
- [12] Pearl, Judea y Russell, Stuart, Bayesian Networks, In Handbook of Brain Theory and Neural Networks pp. 157-160, M. Arbib, editor, MIT Press, Cambridge, MA, USA (2001).
- [13] Spirtes, Peter, Glymour, Clay y Schienes, Richard, Causation Prediction and Search, 529 páginas, Springer-Verlag, New York, NY, USA (1993).

- [14] Triola, Mario F., Estadística, novena edición, 872 páginas, Pearson Addison Wesley, México (2004)