Un modelo de red bayesiana de la informalidad laboral en Veracruz orientado a una simulación social basada en agentes

Jean Christian Díaz-Preciado, Alejandro Guerra-Hernández, Nicandro Cruz-Ramírez

Universidad Veracruzana, Centro de Investigación en Inteligencia Artificial, Xalapa, Ver., Mexico

jean_christian12@msn.com, aguerra@uv.mx, ncruz@uv.mx

Resumen. La informalidad en el empleo en México es un fenómeno social de interés, ya que cerca de un 60 % de los trabajadores se desempeña en este sector. En este trabajo se propone un análisis de la informalidad en el empleo con la creación de un modelo de red bayesiana a partir de la base de datos generada de la Encuesta Nacional de Ocupación y Empleo, obtenida mediante el Instituto Nacional de Estadística y Geografía, con el propósito de utilizarla posteriormente en una simulación social basada en agentes y artefactos, en la cual los agentes obtengan algunas de sus propiedades de la base de datos y otras por inferencias a la red bayesiana generando datos artificiales. Se hace una comparación estadística de los datos generados en el sistema con los datos reales, lo cual se utilizará en un futuro para la validación de la simulación social bajo un paradigma lógico-estadístico.

Keywords: Base de datos, encuesta, redes bayesianas, agentes, validación estadística.

A Bayesian Network Model of Labour Informality in Veracruz Oriented to Agent-based Social Simulation

Abstract. In Mexico, informal employment is a social phenomenon of interest, given that about 60 % of the workers are in this situation. This paper presents an analysis of informal employment based a bayesian network model obtained from the data from the National Survey of Occupation and Employment, obtained by the National Institute of Statistics and Geography. The model is intended to be used in an agent-based social simulation, where agents get some properties directly from the data base and some others through bayesian inference, generating in this way artificial data. A statistical comparison of the data generated in the

system and the real data will be used in the future for validation of social simulation under a logical-statistical paradigm.

Keywords: Data base, survey, Bayesian network, agents, statistical validation.

1. Introducción

En México, el Instituto Nacional de Estadística y Geografía (INEGI) es el encargado de recabar información de personas, hogares y empresas, mediante la aplicación de encuestas y censos en periodos determinados, que tienen como objetivo principal proveer información para la generación de estadistícas que sirven como parámetro para la toma de decisiones en la implementación de políticas públicas. Estas encuestas y censos tienen diferentes temáticas, en este trabajo nos enfocamos en la Encuesta Nacional de Ocupación y Empleo (ENOE), que proporciona información sobre la ocupación de las personas, es decir, si tienen empleo o no; y en el caso de los empleados, que características tienen sus empleos. Los resultados publicados por el INEGI acerca de la informalidad laboral son presentados mediante indicadores y estadísticas con una cobertura geográfica nacional y estatal, desglosadas por sexo, edad y sector de la actividad [6].

El objetivo de este trabajo es construir un modelo que nos permita clasificar la situación laboral de una persona como formal o informal, dadas las prestaciones que proporciona su empleo y el contexto individual del trabajador. Se decidió que el modelo tomase la forma de una red bayesiana y se adoptó una aproximación de minería de datos basada en agentes para su construcción. Aunque esto no es mandatorio, nuestro interés por usar el modelo más adelante, en una simulación social basada en agentes, justifica la decisión. Hemos adoptado una aproximación de minería de datos basada en Agentes y Artefactos [8], muy similar a la usada en la herramienta JaCa-DDM [7]: Se provee una serie de artefactos basados en Weka [10], para almacenar datos, generar el modelo y evaluarlo. Los agentes usan estos artefactos en el proceso de aprendizaje. Puesto que el modelo y los datos son accesibles a los agentes, vía estos artefactos, los agentes pueden obtener algunas de sus propiedades de la base de datos y otras mediante inferencias bayesianas a partir del modelo. Los agentes reportan sus actividades generando una base de datos artificial, la cual es comparada estadísticamente con la base de datos real. Eventualmente nos gustaría modelar como afecta la implementación de políticas públicas la decisión laboral de los agentes, es decir, si optan por una situación formal o informal.

El artículo está organizado de la siguiente manera: En el capítulo 2 se muestran las características de la base de datos ENOE y la descripción de las variables utilizadas en este trabajo. En el capítulo 3 se hace la descripción del sistema que genera el modelo y los datos artificiales. Posteriormente, el capítulo 4 describe el diseño experimental y el capítulo 5 los resultados obtenidos del mismo. Finalmente se presentan las conclusiones y trabajo futuro, en los capítulos 6 y 7, respectivamente.

2. Base de datos ENOE

La base de datos ENOE está construída a partir de entrevistas realizadas en 120,000 viviendas repartidas en todo México, recabando información de alrededor de 800,000 personas. La información de la ENOE, puede estudiarse a nivel de vivienda, hogar y persona, la base de datos con la que se realizó el modelo fue obtenida de la tabla sociodemográfica (SDEMT110 [5]) del primer trimestre del año 2010. Los periodos posteriores podrán usarse para la validación de los datos artificiales generados.

El INEGI clasifica a la población en edad de trabajar legalmente, mayores de 15 años, en dos grandes grupos: Población Económicamente Activa (PEA), que es la que ejerce presión en el mercado laboral; y Población No Económicamente Activa (PNEA). Como podemos observar en la figura 1, la PEA se divide a su vez en Población Ocupada y Desocupada, según se tenga o no. Dentro de la población ocupada hay población sub ocupada refiriéndose a las personas que tienen un empleo pero continúan en busca de otro. La PNEA se divide en las personas disponibles, que aunque no están ocupadas ni buscando empleo al momento de la encuesta, bajo ciertas circunstancias podrían decidir incorporarse al mercado laboral; y las personas no disponibles, que son las que se encuentran bajo un contexto que les impide laborar.

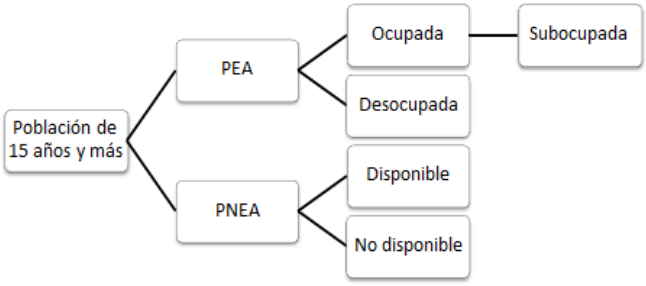


Fig. 1. Clasificación de la población [4].

Con la idea de atender la problemática regional, en este trabajo se considera la población sub ocupada que reside en el estado de Veracruz.

2.1. Descripción de variables

El cuadro 1 describe las variables consideradas para este trabajo, las cuales dividimos en generales y laborales. Las variables generales incluyen sexo, edad, escolaridad y si las personas estudian al momento de la encuesta; como se mencionó, nos enfocamos en las personas sub ocupadas, por ello se incluyen las variables que nos indican si se encuentran en busqueda de un nuevo empleo, y el motivo de la busqueda. Las variables laborales corresponden a algunas de

las características de los empleos, como es el nivel de ingreso, la duración de la jornada y la formalidad o informalidad del empleo. La variable de clasificación de empleo nos indica la infomalidad o formalidad del mismo, por ello es considerada como variable clase, ya que se quiere observar que tipo de empleo puede tener una persona dadas sus características.

Tabla 1. Variables seleccionadas para la creación del modelo [5].

Atributos generales			**Atributos laborales**		
Atributo	**Variable**	**Descripción**	**Atributo**	**Variable**	**Descripción**
Sexo	1	Hombre	Ingreso	1	Hasta 1 salario mínimo
	2	Mujer		2	De 1 a 2 salarios mínimos
Edad	1	14 a 24 años		3	De 2 a 3 salarios mínimos
	2	25 a 44 años		4	Des 3 a 5 salarios mínimos
	3	45 a 64 años		5	Mas de 5 salarios mínimos
	4	65 años y mas	Jornada	1	Ausente temporales
Escolaridad	1	Preescolar		2	Menos de 15 horas
	2	Primaria		3	De 15 a 34 horas
	3	Secundaria		4	De 35 a 48 horas
	4	Bachillerato		5	Mas de 48 horas
	5	Normal	Clasificación de empleos	1	Empleo informal
	6	Técnica		2	Empleo formal
	7	Profesional			
	8	Maestría			
	9	Doctorado			
Estudia actualmente	1	Sí			
	2	No			
Busca otro empleo	1	Sí			
	2	No			
Motivo de busqueda	1	Para tener otro empleo			
	2	Para cambiarse de empleo			
	3	No buscó			

Los nombres de las variables de la fuente original fueron cambiados para dar mayor claridad. Es importante especificar la relación que existe entre las variables generales y su representación como propiedades de los agentes que definiremos en nuestro sistema como trabajadores; así como la relación entre las variables laborales y su representación como propiedades de los artefactos que definiremos en nuestro sistema como empresas. El siguiente capítulo presenta la descripción detallada del sistema de Agentes y Artefactos propuesto.

3. Sistema multiagente para crear el modelo

Los agentes del sistema están basados en redes probabilísticas para el manejo de la incertidumbre. Se implementó una red bayesiana, la cual se define como un modelo probabilísitico representado mediante un grafo acíclico dirigido (GAD), en el cual los nodos representan las variables del fenómeno y las dependencias probabilistas que existan entre ellas se encuentran en la estructura del grafo. Asociada a cada nodo de la red hay una distribución de probabilidad condicional (TPC), dependiente de los nodos padre [9].

La razón de utilizar redes bayesianas es facilitar la interpretación del modelo mediante el GAD y que nos permite observar la probabilidad que tiene una persona de tener un empleo formal o informal, y de las características del empleo, según su edad, sexo y escolaridad. Esto se realiza por medio de inferencias al modelo enviando como evidencia las variables generales [3].

Dado que el sistema está basado en Agentes y Artefactos, es deseable que la red bayesiana sea accesible a los agentes, ya sea como parte de ellos o como parte del algún artefacto. Para construir el modelo, hemos extendido JaCa-DMM [7], definiendo una nueva estrategia de aprendizaje basada en redes bayesianas y agregando un artefacto que encapsula SamIam [1], para realizar las inferencias. La figura 2 muestra el diagrama general del sistema, que tiene como entrada una base de datos y una red bayesiana que puede ser ingresada manualmente o generada por el sistema a partir de los datos. A continuación se describen a detalle las tareas realizadas por los artefactos y agentes en el sistema, en la figura 3 podemos observar el diagrama de los procesos realizados por estos.

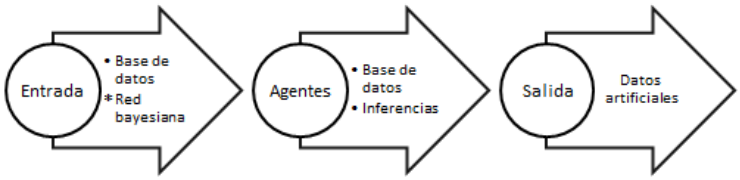


Fig. 2. Diagrama general del sistema multiagente.

3.1. Artefactos

El sistema cuenta con 3 artefactos que realizan tareas correspondientes a la minería de datos, los cuales se describen a continuación:

InstancesBase Este artefacto se adopta directamente de la herramienta JaCa-DDM, que a su vez la adopta de Weka. Se trata de un repositorio de ejemplos de entrenamiento que puede cargar archivos ARFF para su uso en clasificadores, evaluadores y demás herramientas Weka. Al tomar la forma de artefacto, los ejemplos y sus atributos son accesibles a los agentes del sistema y a otros artefactos.

BayesNet Este artefacto encapsula los métodos de construcción de redes bayesianas implementados en Weka. Su tarea principal es construir modelos a partir de datos y ponerlos a disposición del artefacto basado en SamIam. Dependiendo de las entradas del sistema realiza las siguientes tareas:

- Leer una red bayesiana generada previamente en formato XMLBIF

- Generar una red bayesiana a partir de los datos del artefacto InstancesBase y crear un modelo con los parámetros descritos en el capítulo 4 (Diseño experimental). Estos parámetros son fijos. Una vez generado el modelo, éste se guarda en formato XMLBIF.

SamIam Es el artefacto que los agentes usan para hacer inferencias basadas en una red bayesiana, la generada con el artefacto BayesNet. Por medio de las herramientas de SamIam, envía los parámetros para tener acceso a las tablas TPC; Recibe las evidencias de los agentes trabajadores para hacer la inferencia de las variables no conocidas, poniendo a disposición de los agentes las TPC obtenidas.

3.2. Agentes

El sistema cuenta con tres clases de agentes, dos se utilizan para crear agentes que controlan los experimentos y una para crear agentes que representan trabajadores. A continuación se describen las tareas que realiza cada agente:

Agente learning El sistema inicia su ejecución con este agente, el cual tiene como meta realizar las tareas correspondientes a la minería de datos por medio de los artefactos, para que los demás agentes tengan acceso a la base de datos y al modelo. Para poder concretar su meta, crea los artefactos y establece las ligas necesarias entre ellos. Su plan sigue esta secuencia:

- Leer base de datos (IntancesBase),
- Generar modelo (BayesNet),
- Preparar modelo para inferencias (SamIam).

El agente está diseñado para iniciar sus acciones ya sea recibiendo la base de datos y crear la red bayesiana; O recibir un modelo generado previamente. Al concluir sus tareas, envía un mensaje al agente control para que este comience sus actividades.

Agente control La meta de este agente es crear a los agentes que representan a los trabajadores a partir de los datos almacenados en el artefacto InstancesBase. Su única creencia es el número de personas que debe crear. Al momento de crear un nuevo agente le proporciona un nombre, que corresponde al número de caso de la base de datos, para que el nuevo agente obtenga sus propiedades generales de éste. También es el encargado de controlar el acceso al artefacto SamIam al momento de las inferencias de los agentes que representan los trabajadores.

Agente persona Esta clase de agente se usa para representar trabajadores, por lo que su meta principal es instanciar sus propiedades generales y laborales. Para las propiedades generales, recupera sus datos almacenados en el artefacto

InstancesBase a partir de su nombre. Las propiedades laborales son inferidas en el artefacto SamIAm, con base en la evidencia que proveen las propiedades generales del agente y el modelo almacenado en el artefacto BayesNet. Puesto que las TPC generadas tienen 2 o más variables, se ejecuta una acción interna que representa la ruleta propuesta por DeJong [2] y según su probabilidad condicional, se determina la propiedad laboral del agente. Una vez obtenidas todas sus propiedades, estas son almacenadas en un archivo de texto.

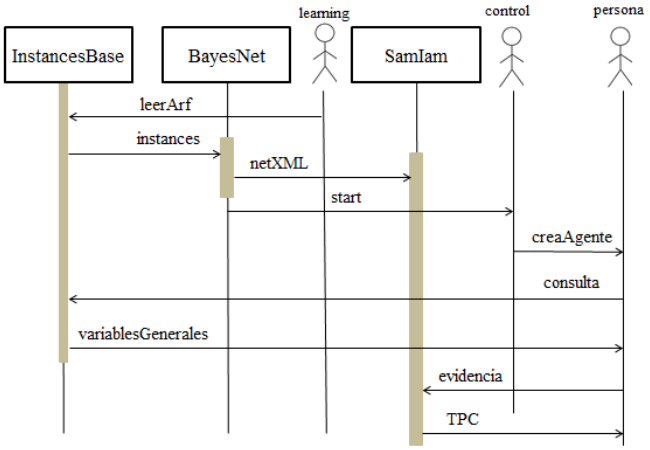


Fig. 3. Diagrama de procesos del sistema multiagente.

4. Diseño experimental

El sistema multi-agente propuesto, proporciona como salida una base de datos artificial y, si se desea, la red bayesiana generada de la base de datos. Como se mencionó, se determinó construir este modelo a partir de las personas sub ocupadas con residencia en el estado de Veracruz, las cuales en la ENOE del primer trimestre del 2010 constituyen un universo de 595 casos.

El cuadro 2, muestra los parámetros implementados para la generación de la red bayesiana, los cuales se determinaron en base a la observación de las redes generadas en experimentos utilizando Weka. En estos experimentos se observó que los parámetros que modificaban significativamente al modelo fueron el iniciarlo con un modelo Naive y la implementación de la manta de Markov, ya que sin ella, dos variables consideradas como relevantes, la edad y si el trabajador está estudiando, quedaban fuera de la red. Se llevó a cabo una validación cruzada del model con diez pliegues.

Dado que las variables generales son obtenidas directamente de la ENOE, la validación de los datos artificiales se realiza comparando la distribución de las variables laborales. Posteriormente se generó una red bayesiana con la base de

Tabla 2. Parámetros para generar el modelo en Weka.

Estimador	Simple Estimador	A 0.5
Algoritmo de Busqueda	HillClimber	
Parámetros del Algoritmo de busqueda	initNaiveBayes	False
	MarkovBlanket Classifier	True
	mxNrOfParents	10,000
	scoreTYPE	MDL
	useArcReversal	True

datos artificial la cual se comparó con la base de datos generada por el sistema para complementar la validación.

5. Resultados

El modelo generado por el sistema a partir de los datos reales, es muy similar al generado por Weka con los mismos datos. Con los datos artificiales se generó un modelo en Weka usando los mismo parámetros. El modelo obtenido con los datos de la ENOE se muestra en la figura 4(a), y el modelo generado por los datos artificiales en la figura 4(b). Los resultados estadísticos calculados por Weka se pueden observar en el cuadro 3.

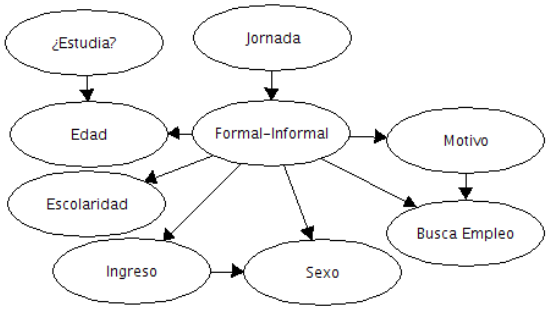
Tabla 3. Comparación de estadísticas de la generación del modelo.

	Datos reales	**Datos Artificiales**
Porcentaje de clasificados correctamente	74.11 %	63.80 %
Desviación Estándar	5.46	6.13
Área bajo la curva ROC	0.8180	0.6383

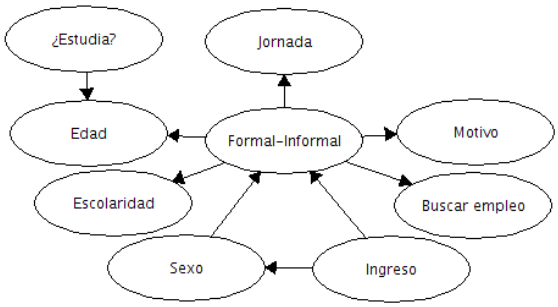
Se generaron las gráficas de curva ROC, con umbral de 0.5, para cada variable de la clase, con ambas bases de datos. Las gráficas ROC para la variable de informalidad en el empleo se muestran en las Figura 5(a) para los datos reales y en la Figura 5(b) para los artificiales, y para la variable de formalidad en el empleo en las Figuras 5(c) y 5(d).

También se realizó una comparación de las distribuciones de las variables que se obtuvieron en las inferencias al modelo. Dado que se generó el mismo número de agentes que de casos de la base de datos, la distribución de las variables generales son las mismas. En el Cuadro 4 se muestran las estadísticas de las variables laborales.

La Figura 6 muestra la comparación gráfica de la distribución de las variables. Los datos generados por medio de las inferencias y la aplicación de la ruleta, nos proporcionan resultados muy parecidos a los reales.

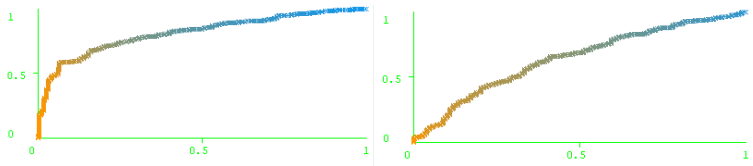


(a)

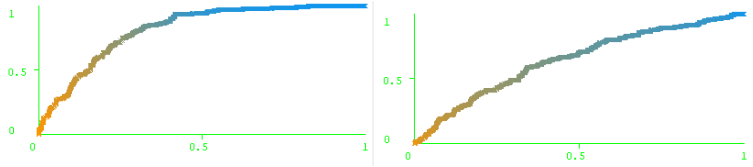


(b)

Fig. 4. Modelos de red bayesiana obtenidos: (a) Datos reales y (b) Datos artificiales.



(a) Datos reales (b) Datos artificiales



(c) Datos reales (d) Datos artificiales

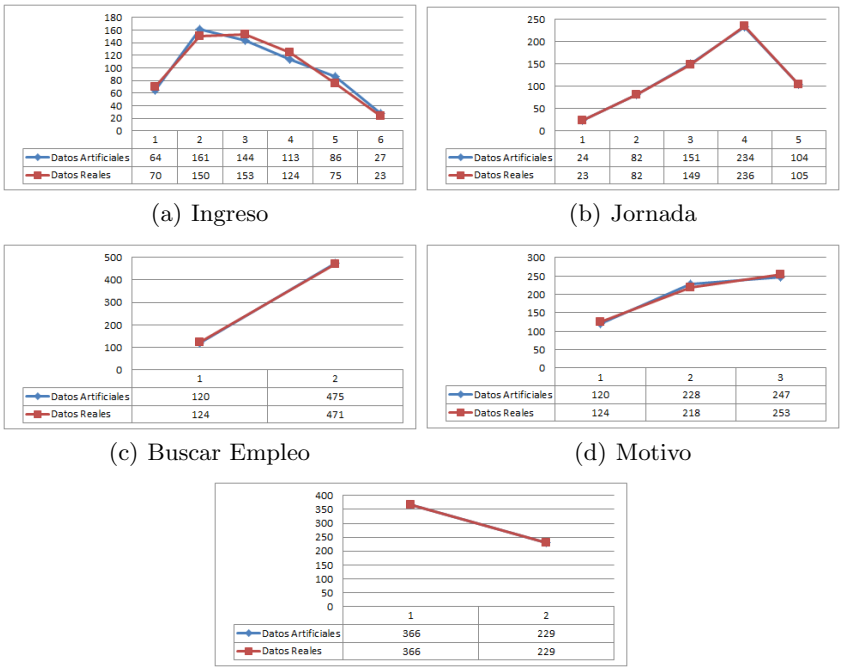
Fig. 5. Curvas ROC de los modelos generados.

5.1. Informalidad en el Estado de Veracruz

Como se mencionó, uno de los principales objetivos del trabajo del INEGI es dotar de estadísticas e información a los órganos encargados de la generación de

Tabla 4. Comparación de estadísticas.

Variable	Media		EE Media		DesvEst		Varianza		Asimetria		Curtosis	
	Real	Artificial	Real	Artificial	Real	Artificial	Real	Artificial	Real	Artificial	Real	Artificial
Ingreso	3.0891	3.1294	0.0547	0.0560	1.3335	1.3652	1.7782	1.8637	0.25	0.28	-0.71	-0.80
Jornada	3.5345	3.5244	0.0432	0.0434	1.0541	1.0576	1.1112	1.1185	-0.49	-0.48	-0.38	-0.39
Buscar	1.7916	1.7983	0.0167	0.165	0.4065	0.4616	0.1652	0.1613	-1.44	-1.49	0.07	0.22
Motivo	2.2168	3.2134	0.0314	0.0310	0.7665	0.7564	0.5876	0.5722	-0.39	-0.38	-1.21	-1.17
Formal/Informal	1.3849	1.3849	0.0200	0.0200	0.4870	0.4870	0.2371	0.2371	0.47	-1.78	-1.78	-1.78



(a) Ingreso (b) Jornada

(c) Buscar Empleo (d) Motivo

(e) Informalidad

Fig. 6. Comparación de distribución de variables.

políticas públicas. Estas estadísticas se presentan como gráficas con 2 variables a observar, por ejemplo la cantidad de personas por sexo que tienen un empleo formal o informal.

En este trabajo se realizó el análisis de los datos por medio de consultas al modelo, observando la propagación de las probabilidades dados ciertos valores de las variables que se observaron. Es importante mencionar que para este experimento, no se consideraron todas las variables que provee la base de datos ENOE y solo se trabajó con las personas subocupadas del Estado de Veracruz. A continuación se presentan tres consultas generadas:

- La variable clasificiación de empleo es dependiente del número de horas que trabaja una persona, se observó que solo en el rango de 35 a 48 horas laborales, hay más personas con empleos formales que informales, figura 7.

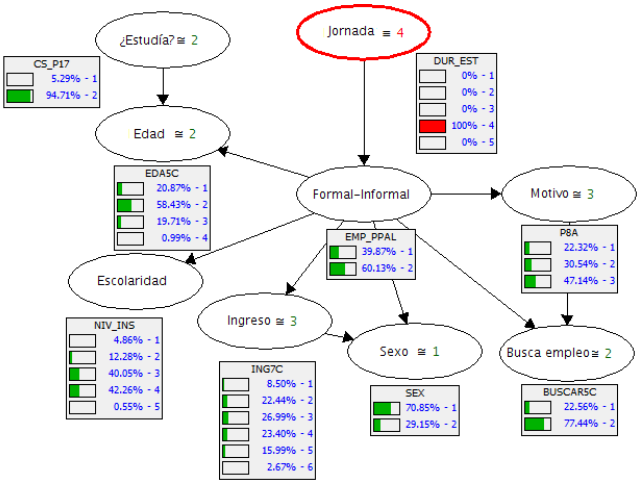


Fig. 7. Propagación por Jornada Laboral [1].

- Para los empleos informales se observó que las personas cuentan con un menor nivel escolar, los sueldos son menores y hay más personas en busca de un nuevo empleo para dejar el actual, Figura 8.

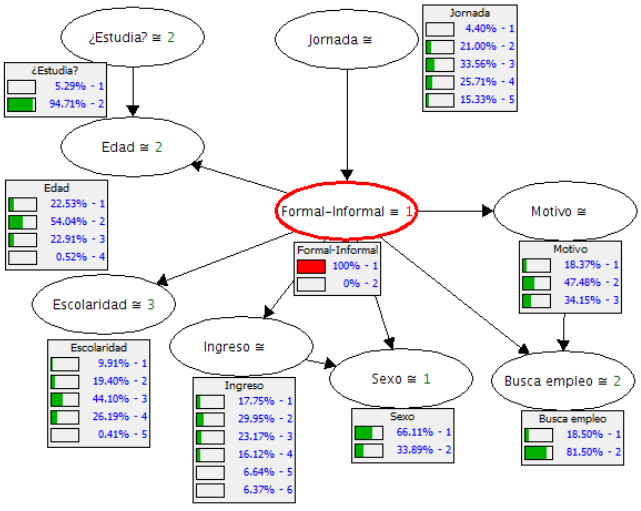


Fig. 8. Propagación por Informalidad en el Empleo [1].

- Para los empleos formales se observó que las personas tienen un mayor nivel escolar, los sueldos se sitúan en rangos más altos y es menor el número de

persona en busca de otro empleo, sin embargo las que están en busca de otro empleo es en su mayoría por que desean dejar su empleo actual, figura 9.

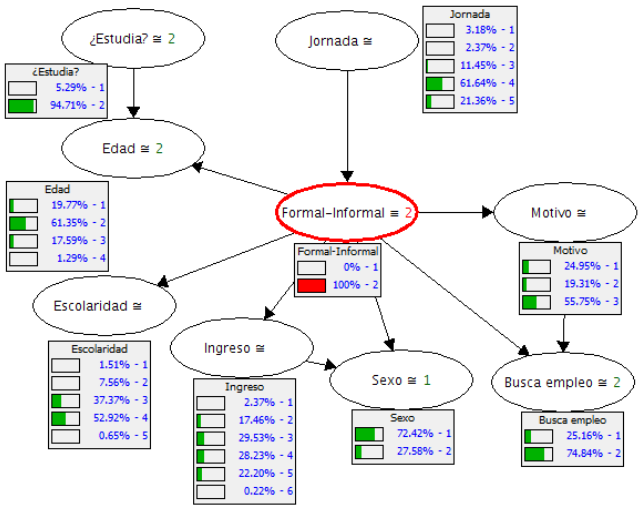


Fig. 9. Propagación por Formalidad en el Empleo [1].

6. Conclusiones

Podemos observar como la red generada con los datos artificiales, es muy similar a la generada con los datos reales, considerando que el valor de las variables que adquieren los agentes como propiedades laborales, está determinado por una ruleta, se da oportunidad de adquirir valores que no tengan la probabilidad más grande. Las diferencias en los GAD de los modelos se describen en el cuadro 5.

Tabla 5. Diferencia entre modelos.

Datos reales	Datos artificiales
El nodo Jornada es independiente	El nodo Jornada es dependiente del nodo de clasificación del empleo
El nodo de clasificación del empleo es dependiente del nodo Jornada	El nodo de clasificación del empleo es dependiente de los nodos Ingreso y Sexo

A pesar que los datos con los que se genera el modelo son solo 595, y el número de nodos de la red son 9, la distribución de las variables que se obtienen

por medio de inferencias tienen una gran similitud con las reales. Consideramos que el porcentaje de casos clasificados correctamente con el modelo generado es bueno teniendo 74.11 % con una desviación estándar de 5.46, y a pesar de que las distribuciones de las variables de los datos artificiales son similares, el porcentaje que se obtuvo es de 63.80 %, con una desviación estándar de 6.13, evidentemente el porcentaje disminuye más de un 10 %, sin embargo la desviación estándar no es muy alta, por lo que se espera tener mejoras experimentando con bases de datos que contemplen un número mayor de casos.

Se generó el área bajo la curva ROC, para visualizar como los datos artificiales tienden a disminuir la generalidad del modelo, y con esto a cometer errores en la clasificación. Estos resultados nos proporcionan un panorama del número de agentes que debe tener nuestra simulación así como la complejidad de la red bayesiana que se implementará, para mantener la consistencia que nos proporciona el modelo en nuestros datos artificiales. Las comparaciones estadísticas que observamos en el cuadro 4 nos muestran las similitudes entre las bases de datos. Observamos que el atributo correspondiente al ingreso es el que tiene mayor diferencia en la distribución de los datos, esto se debe al número de variables que tiene y la distribución de las mismas, que observamos en la desviación estándar y la varianza, como complemento se calculó la curtosis que es de -0.71 y el valor de asimetria de 0.25; el atributo jornada también tiene un número mayor de variables, pero no se presentan cambios significativos en los datos ya que tiene una desviación estándar menor, y tiene una curtosis de -0.39 con un valor de asimetría de -0.48. Los atributos que solo tienen dos o tres variables presentaron las menores diferencias en la comparación, teniendo una distribución igual en la variable correspondiente a la formalidad o informalidad del empleo. El error cuádratico medio (EMC) calculado es de 0.81, con lo cual se determinó que los resultados son aceptables para el caso de estudio.

La implementación de la minería de datos en un entorno de agentes y artefactos, es una herramienta útil que nos facilita observar el flujo de trabajo que se realiza. Los artefactos nos proporcionan la distribución de las tareas que se requieran efectuar, ya sea desde la lectura de la base de datos para la generación del modelo o partiendo de un modelo ya generado. Los agentes se utilizaron para definir el orden en que se deben efectuar los procesos. Los resultados de la predicción mediante inferencias a redes bayesianas nos proporcionan datos aproximados a los datos con los que se genera el modelo. La determinación de analizar una base de datos obtenida de encuestas a personas con redes bayesianas, nos proporciona información sobre la causalidad del fenómeno a observar.

7. Trabajo futuro

En un futuro se prentede utilizar las inferencias que realizan los agentes al modelo en una simulación social, en donde el valor de los atributos que obtengan influyan en su toma de decisiones. Dicho esto se probará las inferencias con bases de datos con más casos, y con periodos de tiempo definidos, por ejemplo, observar el comportamiento de las variables en un periodo de un año equivalente a cuatro

encuestas de la ENOE. Como se mencionó la ENOE puede observarse a nivel hogar, por lo que se pretende hacer experimentos a este nivel, representando los hogares por medio de artefactos, con el fin que los agentes que pertenezcan a un mismo hogar tengan acceso a la información de los integrantes de su familia. Se pretende implementar por medio de artefactos, una abstracción de políticas públicas, es decir, de qué forma afectan a un hogar o individuo, y agregar comportamientos de los agentes con respecto a los cambios con los que se enfrentará.

Agradecimientos. El primer autor cuenta con el apoyo de la beca CONACyT número 633473.

Referencias

1. Darwiche, A.: Modeling and Reasoning with Bayesian Networks. Cambridge University Pres, New York (2009)
2. De Jong, K.A.: Analysis of the behavior of a class of genetic adaptive systems. Tech. Rep. 185, The University of Michigan (1975)
3. Glymour, C.: Discovering Causal Structure. Academic Press, Orlando, Florida (1987)
4. INEGI: Conociendo la base de datos de la ENOE. Datos ajustados a proyecciones de población 2010. INEGI (2010)
5. INEGI: ENOE. Descripción de Archivos. INEGI (2010)
6. INEGI: México: Nuevas estadísticas de informalidad laboral. INEGI (2013)
7. Limón, X., Guerra-Hernández, A., Cruz-Ramírez, N., Grimaldo, F.: An agents and artifacts approach to distributed data mining. In: Advances in Soft Computing and Its Applications, pp. 338–349. Springer (2013)
8. Omicini, A., Ricci, A., Viroli, M.: Artifacts in the A&A meta-model for multi-agent systems. Autonomous Agents and Multi-Agent Systems 17(3), 432–456 (2008)
9. Pearl, J.: Probabilistic Reasoning in Intelligence Systems. Morgan Kauffman, San Mateo, CA (1988)
10. Witten, I.H., Frank, E.: Data mining, Practical Machine Learning Tools and Techniques. Morgan Kauffman, San Francisco, CA (2011)