

Satisfacción del Alumnado mediante Técnicas de Descubrimiento basadas en Lógica Fuzzy

Susana Beatriz Ruiz¹ ; Rafael Espin Andrade²; Myriam Beatriz Herrera¹

¹Facultad de Ciencias Exactas, Físicas y Naturales, Universidad Nacional de San Juan, Argentina.

² Universidad Autónoma de Coahuila, Blvd Revolución y Donato Guerra, s/n

sbruizr@yahoo.com.ar, rafaelalejandroespinandrade@gmail.com, myriamhrrr@gmail.com

Resumen

Al igual que la mayoría de las organizaciones, las instituciones educativas deben evaluar la calidad de los servicios prestados. Aquí, el concepto de "calidad" se entiende como la forma en que esta organización satisface las expectativas de sus estudiantes para contribuir a la imagen que los estudiantes tienen de ella. Aunque la calidad percibida en el servicio recibido está fuertemente relacionada con las necesidades de los estudiantes, es posible extraer indicadores objetivos como su satisfacción como un factor esencial para abordar la calidad de la educación superior. Este artículo presenta los resultados de un estudio exploratorio sobre la satisfacción de 112 estudiantes de la Facultad de Ciencias Exactas, Naturales y Físicas de la Universidad Nacional de San Juan durante el año 2017. Mediante el uso de técnicas de descubrimiento de conocimiento basadas en lógicas multivalentes, descubrimos relaciones entre los estados lingüísticos de un conjunto de datos (predicados difusos) que cumplen la forma $P \rightarrow$ Satisfacción. Consideramos que la información encontrada es valiosa para tomar decisiones institucionales con el fin de mejorar la calidad educativa.

Palabras clave: Satisfacción, Descubrimiento, Fuzzy, Eureka-Universe

Introducción

En la actualidad, la mayoría de las instituciones, empresas u organizaciones analizan la calidad de sus productos y/o servicios. De igual forma

las instituciones educativas se ven obligadas a valorar la calidad educativa. La calidad se debe entender como "un concepto abstracto tan amplio en definición y aplicación que cada organización puede entenderlo desde sus propios intereses" [1]. El concepto abarca el modo en que la empresa satisface todas las especificaciones de sus clientes y todo contacto con ellos, ya que contribuye a la imagen que se forman los clientes en sus mentes sobre la empresa [2]. La percepción de la calidad en el servicio debe estar respaldada por una estructura y gestión organizacional, de tal forma que se cree una filosofía de servicio que busque superar las expectativas de los clientes [3]. Cabe destacar que la calidad percibida está muy relacionada con la adecuación de las características del objeto a las necesidades del individuo. A pesar de esto, el contacto del individuo con el servicio que recibe y con los agentes que se lo ofrecen, pueden brindar datos objetivos [4].

Las instituciones educativas universitarias no deben estar ajenas a esta mirada. Los estudios sobre satisfacción estudiantil en las universidades como indicadores para evaluar la calidad educativa son particularmente importantes, porque la satisfacción estudiantil mejora el rendimiento académico, reduce la deserción, el cambio de carrera de los inscritos y es un requisito para el éxito en el aprendizaje. A su vez, una comprobada calidad educativa fortalece la imagen y el prestigio de la institución [5]. Valorar la satisfacción estudiantil y determinar cuáles son las variables asociadas en este aspecto, contribuye a la toma de decisiones acertadas en la gestión orientada a la calidad universitaria.

La disponibilidad de grandes volúmenes de información y el uso generalizado de herramientas informáticas, hoy en día, ha transformado el análisis de datos orientándolo hacia determinadas técnicas especializadas englobadas bajo el nombre de Minería de Datos o Data Mining (DM). Las técnicas de MD persiguen el descubrimiento automático del conocimiento contenido en la información almacenada de modo ordenado en grandes bases de datos. Tiene como el objetivo de detectar patrones de comportamiento consistentes o relaciones entre las diferentes variables para aplicarlos a nuevos conjuntos de datos [6]. Para realizar este proceso no trivial de identificación válida, novedosa, potencialmente útil y entendible de patrones comprensibles que se encuentran ocultos en los datos se utilizan técnicas estadísticas que permiten reunir y analizar los datos, como también técnicas provenientes de la inteligencia artificial así como redes neuronales, algoritmos genéticos, técnicas de Lógica Fuzzy y otros tipos de heurísticas aplicados a los procesos de extracción de patrones ocultos o información útil de los conjuntos de datos. Tanto las técnicas estadísticas como las técnicas de inteligencia artificial son bastante poderosas; en algunos casos, estas son tan solo dos enfoques o alternativas diferentes para la solución de un mismo problema, en otras ocasiones son técnicas complementarias porque resuelven problemas de naturaleza diferentes [7].

La Lógica Fuzzy (LF) es un método de razonamiento aproximado no probabilista, que puede definirse como una extensión de la Lógica Multivaluada que facilita enormemente el modelado de información cualitativa de forma aproximada. Su éxito se debe principalmente a la posibilidad de resolver problemas de una gran complejidad y poco definidos que, mediante métodos tradicionales, son difíciles de solucionar.

La LF considera el concepto de conjunto fuzzy bajo el cual reside la idea de que los elementos sobre los cuales se construye el pensamiento humano no son números sino etiquetas lingüísticas. La LF nos permite representar el conocimiento común (que es en su mayoría

lingüístico cualitativo y no necesariamente cuantitativo) en un lenguaje matemático a través de la teoría de conjuntos fuzzy y las funciones características o de pertenencia asociadas a ellos. Se puede ver un ejemplo de su aplicación en expresiones como "el alumno está muy satisfecho con la carrera" o "el alumno tiene buena relación con compañeros". Sin embargo, no es tan fácil definir qué se entiende por "muy satisfecho" y "buena relación" porque es difícil especificar una medida a partir del cual se puede considerar muy satisfecho un alumno o que tiene buena relación un alumno con sus compañeros.

La LF funciona con conjuntos de datos donde no hay límites bien definidos. Usa expresiones que no son completamente verdaderas ni completamente falsas, se aplica a conceptos que pueden tomar cualquier valor de precisión dentro de un conjunto de valores que se extienden entre dos extremos, la verdad absoluta y la falsedad completa. La idea general es que las cosas no son negras o blancas, pero hay infinitos tonos de gris.

Una forma de implementar el "principio de gradualismo", propiedad esencial de la LF, es la definición de lógicas donde los predicados son funciones del universo X en el intervalo $[0,1]$ y las operaciones de conjunción, disyunción, negación e implicación, se definen de manera que cuando se restringe al dominio [8] se obtiene la Lógica Booleana. Las diferentes formas de definir las operaciones y sus propiedades determinan diferentes lógicas multivalentes que forman parte del Paradigma de la Lógica Fuzzy [9].

Los estados lingüísticos son aquellas variables cuyos valores son palabras o etiquetas en vez de números, que pueden representar elementos que sean difíciles de definir o describir en términos numéricos. Representan conjuntos fuzzy. Los valores que pueden tomar los estados lingüísticos son llamados etiquetas lingüísticas que son definidas como conjuntos fuzzy. El concepto de variable lingüística juega un rol importante en la representación imprecisa del conocimiento. La composición de múltiples variables lingüísticas a menudo

constituye un sistema y una descripción completa de algunos conocimientos. Por ejemplo, una persona puede considerarse una variable lingüística que se puede componer con otra variable lingüística “Edad” “Altura”, “Peso”, “Apariencia”, etc. Los valores de una variable lingüística pueden consistir en términos primarios tales como "joven", "viejo" dentro de la "Edad" manejable y aquellos que se construyen a partir del uso de modificadores o coberturas lingüísticas como "muy", "ligeramente", "más o menos", "bastante", "extremadamente", etc. y los conectivos lógicos "no", “o” e "y" [10].

La conversión de un valor a un estado lingüístico se le llama “Fuzzyficación”. Mientras que el proceso inverso es llamado “defuzzyficación”.

Una función de membresía o de pertenencia $f_A(x)$ en un espacio X caracteriza a un conjunto fuzzy A , donde cada valor x es asociado a un número real dentro del intervalo $[0,1]$, el cual representa qué tanto pertenece x al conjunto A . La función de pertenencia depende del contexto en el que se trabaje, del problema a resolver o la aplicación a construir, e incluso del experto que modela el problema y del usuario que toma la decisión. Las funciones de pertenencia pueden ser continuas o discretas, dependiendo del universo de discurso.

Un predicado es lo que se afirma o niega de un objeto. Los predicados clásicos dan lugar a una clara división del universo A de acuerdo con el cumplimiento del predicado $P(x)$. Pero las personas utilizan cotidianamente predicados que no se pueden reducir a esta división, son los predicados difusos o vagos. Los predicados son estructuras flexibles que facilitan la extracción de conocimiento, están compuestos por los operadores y los estados o variables lingüísticas.

Por medio de los tipos de lógica se calcula el valor de verdad para los operadores fuzzy: conjunción, disyunción, negación y orden.

La Lógica Fuzzy Compensatoria (LC) constituye una rama de la Lógica Fuzzy. Es un reciente sistema multivalente que rompe con la

axiomática tradicional de este tipo de sistemas para lograr un comportamiento semánticamente mejor a los sistemas clásicos. En general los modelos basados en esta lógica combinan la experiencia y el conocimiento con datos numéricos, por lo que puede ser visto como una "caja gris". Fue creada por el grupo científico multidisciplinario Gestión Empresarial en la Incertidumbre: Investigación y Servicios (GEMINIS).

Se puede afirmar que, en general, un indicador directo de la calidad educativa es la satisfacción del alumno por la carrera [11][12][13][14]. La satisfacción del estudiantado universitario constituye un factor imprescindible en el abordaje del tema de la calidad de la educación superior. Vista la importancia del tema, en este trabajo se muestran resultados de un estudio exploratorio sobre la satisfacción de 112 estudiantes que cursan carreras en la Facultad de Ciencias Exactas, Naturales y Físicas de la Universidad Nacional de San Juan (Argentina), en el año 2017, con el propósito de determinar las variables asociadas que mejor caracterizan la satisfacción del alumnado por carreras que cursan, , aplicando técnicas de descubrimiento de conocimiento basados en lógicas multivalentes. Los análisis se centran en el descubrimiento de relaciones de la forma $P \rightarrow$ Satisfacción, entre estados lingüísticos a partir de un conjunto de datos proveniente de una encuesta.

Metodología

Se quiere caracterizar la satisfacción del alumnado por carreras de la FCEFyN de la UNSJ (Argentina) mediante la determinación de variables asociadas, a partir de datos provenientes de una encuesta en la que participaron 112 alumnos en el año 2017. La encuesta trata sobre Factores de Riesgo y Calidad de Vida y fue elaborada con la herramienta web de encuestas online EncuestaFácil.com, y cuenta con varias secciones. Las variables consideradas se pueden agrupar en: variables que caracterizan

la Facultad, Universidad y la carrera/s que cursa el estudiante; variables que representan características personales del estudiante y de su familia (edad, sexo, su relación con pares, etc.). Variables asociadas al rendimiento (rendimiento, promedio, etc.); y variables que representan el esfuerzo y motivación del estudiante (Ej. hs. de estudio, asistencia a la universidad, la percepción sobre la carrera en cuanto a que mejorará las expectativas en el futuro, el grado de satisfacción por la carrera, etc.). Se puede consultar la encuesta en: <https://www.encuestafacil.com/RespWeb/Qn.aspx?EID=2197195>. Las preguntas consideradas no son mutuamente excluyentes entre sí. Por lo que cada pregunta es una variable en sí misma; las respuestas alternativas a las preguntas de cada sección sí son mutuamente excluyentes y cada una de estas respuestas es una modalidad (ítems, categoría o etiqueta lingüística) de las variables cualitativas a la que pertenece. La información resultante es pre procesada convenientemente, utilizando el software Excel, para la aplicación de procedimientos en el análisis. En este estudio, a partir del cuestionario, se consideraron un total de 16 variables lingüísticas y 56 estados lingüísticos, o categorías, para realizar los estudios.

Dada la importancia de la temática a abordar, con el propósito de descubrir relaciones relevantes entre variables lingüísticas de la forma lógica $P \rightarrow$ Satisfacción, este trabajo tiene en cuenta la naturaleza borrosa de las variables lingüísticas involucradas en la encuesta. Se aplican algoritmos de descubrimiento de conocimiento basado en LF, tales que permitan estimar un conjunto de parámetros para maximizar la veracidad de las implicaciones en forma universal. En este trabajo, dada la complejidad de la búsqueda, se emplea un sistema recientemente diseñado, denominado Eureka-Universe 2.4.6, tanto para la tarea de “descubrimiento” de predicados fuzzy, la estimación de parámetros óptimos y evaluación de la veracidad de los predicados, Se considera LC, que es un paradigma de una Teoría Lógica interpretable [15].

La LC [16] en su definición tiene en cuenta los siguientes axiomas.

Sean en lo sucesivo $x = (x_1, x_2, \dots, x_i, \dots, x_j, \dots, x_n)$ e $y = (y_1, y_2, \dots, y_i, \dots, y_j, \dots, x_n)$ elementos cualesquiera del producto cartesiano en el dominio $[0,1]^n$. Una cuarteta de operadores continuos: conjunción, disyunción, negación y orden (c,d,n,o) constituyen una LC si cumplen con los siguientes axiomas:

- Compensación:

$$\min(x_1, x_2, \dots, x_i, \dots, x_j, \dots, x_n) \leq c(x_1, x_2, \dots, x_i, \dots, x_j, \dots, x_n) \leq \max(x_1, x_2, \dots, x_i, \dots, x_j, \dots, x_n)$$
- Conmutatividad:

$$c(x_1, x_2, \dots, x_i, \dots, x_j, \dots, x_n) = c(x_1, x_2, \dots, x_j, \dots, x_i, \dots, x_n)$$
- Crecimiento estricto: Si $x_1 = y_1, x_2 = y_2, \dots, x_{i-1} = y_{i-1}, x_{i+1} = y_{i+1}, \dots, x_n = y_n$ excepto por $x_i > y_i$ entonces

$$c(x_1, x_2, \dots, x_i, \dots, x_j, \dots, x_n) > c(y_1, y_2, \dots, y_i, \dots, y_j, \dots, x_n).$$
- Axioma del veto: Si $x_i = 0$ para algún i , entonces $c(x)=0$.

La Lógica Fuzzy Compensatoria basada en la Media Geométrica (GMBCL) [17] es tal que:

$$c_1(x_1, x_2, \dots, x_n) = (x_1 \cdot x_2 \cdot \dots \cdot x_n)^{\frac{1}{n}}$$

$$d_1(x_1, x_2, \dots, x_i, \dots, x_j, \dots, x_n) = 1 - [(1 - x_1) \cdot (1 - x_2) \cdot \dots \cdot (1 - x_n)]^{\frac{1}{n}}$$

$$o_1[x, y] = 0.5[c_1(x) - c_1(y)] + 0.5$$

$$n(x_i) = 1 - x_i.$$

Los operadores universal y existencial se definen para la GMBCL de la siguiente forma (en el caso discreto) [15]:

$$\forall_{x \in U} p(x) = \bigwedge_{x \in U} p(x) = \sqrt[n]{\prod_{x \in U} p(x)}$$

$$= \begin{cases} \exp\left(\frac{1}{n} \sum_{x \in U} \ln(p(x))\right) & , \text{for } x: p(x) \neq 0 \\ 0 & , \text{in any other case} \end{cases}$$

$$\exists_{x \in U} p(x) = \bigvee_{x \in U} p(x) = 1 - \sqrt[n]{\prod_{x \in U} (1 - p(x))}$$

$$= \begin{cases} 1 - \exp\left(\frac{1}{n} \sum_{x \in U} \ln(1 - p(x))\right) & , \text{for } x: p(x) \neq 1 \\ 1 & , \text{in any other case} \end{cases}$$

Las fórmulas en el caso continuo se definen de la siguiente forma:

$$\forall x p(x) = \begin{cases} \frac{\int_X \ln(p(x)) dx}{\int_X dx} & , \text{if } p(x) > 0 \text{ for any } x \in X \\ 0 & , \text{in any other case} \end{cases}$$

$$\exists x p(x) = \begin{cases} 1 - e^{-\frac{\int_X \ln(1-p(x)) dx}{\int_X dx}} & , \text{if } p(x) < 1 \text{ for any } x \in X \\ 1 & , \text{in any other case} \end{cases}$$

La Lógica Fuzzy Compensatoria basada en la Media Aritmética (AMBCL) [18] (en el caso discreto de variables):

$$c_2(x_1, x_2, \dots, x_n) = \left[\min(x_1, x_2, \dots, x_n) \frac{\sum_{i=1}^n x_i}{n} \right]^{\frac{1}{2}}$$

$$d_2(x_1, x_2, \dots, x_n) = 1 - \left[\min(1 - x_1, 1 - x_2, \dots, 1 - x_n) \frac{\sum_{i=1}^n (1 - x_i)}{n} \right]^{\frac{1}{2}}$$

mientras que en el caso continuo:

$$\forall x p(x) = \left[\min_X(p(x)) \frac{\int_X p(x) dx}{\int_X dx} \right]^{\frac{1}{2}}$$

$$\exists x p(x) = 1 - \left[\min_X(1 - p(x)) \frac{\int_X (1 - p(x)) dx}{\int_X dx} \right]^{\frac{1}{2}}$$

Un modelo de un sistema de LC incluye: datos de entrada, un módulo “fuzzyficador”, un núcleo con una base de conocimiento y un motor de inferencia que genera conocimiento a partir de reglas, un módulo “defuzzyficador” y salida/s de datos.

A través de un sistema LC se puede evaluar predicados, descubrir conocimiento y realizar inferencias. El software Eureka-Universe es un sistema único que permite resolver tareas y combinarlas en soluciones para la toma de decisiones en forma sencilla. Puede resolver tareas de: evaluación, descubrimiento e inferencias. En las tareas de evaluación, calcula los valores de verdad de un predicado difuso para un set de datos. Construye un predicado seleccionando las variables lingüísticas y los operadores. Evalúa cada registro en el set de datos y obtiene un valor de verdad, se calcula también el operador existencial y el universal. En el descubrimiento se busca relaciones entre los estados lingüísticos de un set de datos

(predicados difusos) que cumplan con las especificaciones del usuario. En la búsqueda de predicados difusos se utilizan algoritmos genéticos y el ajuste de los parámetros de las funciones de pertenencia definidas en los estados lingüísticos.

Eureka-Universe está programado en Java. Permite el trabajo con los distintos tipos de lógicas, entre ellas la Lógica Compensatoria Basada en la Media Geométrica y la Lógica Compensatoria basada en la Media Aritmética. La Figura 1 proporciona una imagen inicial del sistema para su aplicación donde, al activar la ventana de ayuda, muestra el grupo de científicos e investigadores que participaron en el desarrollo de la versión 2.4.6.

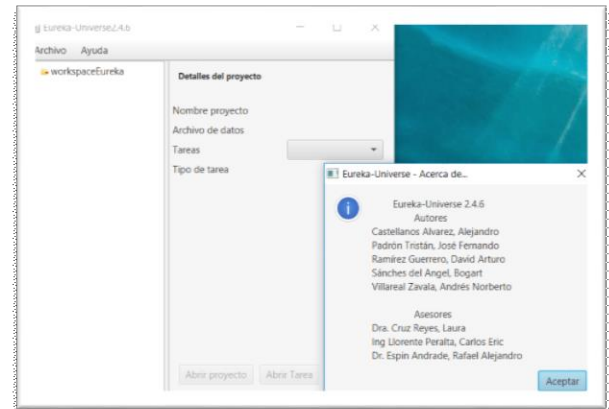


Figura 1: Ventana inicial de Eureka-Universe

Los estados lingüísticos en Eureka-Universe son una asociación entre el nombre de una columna del set de datos, una función de pertenencia y un nombre de estado o etiqueta. En las tareas de Descubrimiento se definen los "estados lingüísticos" sin especificar la función de pertenencia de dichos estados, para que el sistema utilice Funciones de Pertenencias Generalizadas (FPGs) con parámetros variables.

Para el procesamiento de datos de encuesta, en este trabajo, como paso inicial, se divide en forma aleatoria en una muestra para descubrimiento de relaciones entre variables constituida por las respuestas de 78 alumnos (70% de alumnos encuestados) y una muestra test (30% de alumnos) para evaluar resultados. Se aplica el sistema Eureka-Universe 2.4.6.

para la búsqueda de relaciones relevantes entre estados lingüísticos. En este trabajo se considera una relación “P→Satisfacción” relevante si la veracidad del predicado universal $(\forall x)(P(x) \rightarrow \text{Satisfacción}(x))$ resulta mayor o igual a 0.9. A partir de ello se siguen los siguientes pasos:

1) Paso inicial: Lectura de la base de datos para descubrimiento. Creación de Proyectos de Descubrimiento para buscar relaciones relevantes entre estados lingüísticos, para predicados P simples (incluye sólo un estado lingüístico). Las lógicas que se emplean son GMBCL y AMBCL. Las funciones de pertenencias de los estados lingüísticos no se especifican en ningún proyecto.

2) Ejecución de tareas de descubrimiento definidos en los Proyectos creados en la etapa anterior. Análisis y evaluación de resultados. Selección de relaciones relevantes.

3) Evaluación de resultados considerando la muestra test. Análisis y comparación de resultados.

4) Planteos de hipótesis. Definición y ejecución de nuevos proyectos de descubrimiento y evaluación, bajo supuestos considerados, para predicados P compuestos, considerando la muestra de descubrimiento. Selección, análisis y comparación de resultados considerando además la muestra test.

Resultados

A continuación, se presentan los resultados obtenidos, según la metodología de trabajo detallada en la sección anterior. Los resultados se analizan desde las perspectivas específicas de estudio y teniendo en cuenta el objetivo planteado en el presente trabajo.

A partir de la lectura de la base de datos en Excel, que se visualiza en la Figura 2, se definen distintos proyectos para descubrir relaciones entre variables lingüísticas difusas, de interés en el estudio, utilizando el sistema Eureka-Universe. La primera fila de la matriz de datos de la encuesta contiene una lista de variables lingüísticas, entre ellas las que se trabajaron en la primera etapa del trabajo.

Mientras que las restantes filas (112 filas) representan las respuestas de los alumnos en la encuesta. Las puntuaciones para cada respuesta, son transformadas a valores continuos, de tal forma de facilitar el procesamiento de datos en el sistema.

La Tabla 1 muestra resultados que permiten caracterizar los descubrimientos obtenidos al ejecutar el sistema 10 veces, para proyectos con distintas LC y utilizando la muestra de descubrimiento. En cada ejecución, por defecto, el sistema realiza los descubrimientos empleando algoritmos genéticos mediante 50 iteraciones. Se puede observar que, entre el conjunto de predicados descubiertos, el máximo valor de veracidad hallado es aplicando la GMBCL. Este valor óptimo se logra estableciendo una relación entre las variables “Dinero para Alimento” y “Satisfacción por Carrera” mediante la estimación de parámetros de funciones de pertenencia generalizada. También con la GMBCL se descubren, en términos medios, un mayor número de predicados relevantes por ejecución, donde intervienen las variables: “Tiempo dedicado a práctica”, “Relación con compañeros”, “Dinero para alimento”, “Calidad del material de estudio”, “Edad”, “Nivel de exigencia de la carrera”, “Efecto de los aplazos”, “Rendimiento según materias”, “Año que cursa”, “Dinero para estudio”, “Relación con docentes”, “Genero” y “Asistencia por semana”. Mientras que con la LDCMA se descubren dos variables lingüísticas relacionadas con la satisfacción: “Tiempo dedicado a practica” y “Tiempo dedicado a teoría”.

Del análisis de los resultados a partir de la Tabla 2 se tiene en cuenta los valores de veracidad de predicados universales $GP(x)$ definidos en la primera fila de la Tabla, tanto para la muestra para descubrimiento como para la muestra test utilizando GMBCL. Para ambas muestras, los valores de veracidad hallados mediante tareas de evaluación de los predicados universales correspondientes a las relaciones R1, R2, R3 y R4, son todos mayores a 0.90. Para cada relación R_i descubierta se tiene además que las

veracidades halladas para la muestra de descubrimiento y para la muestra test difieren levemente en cada caso, lo que es un indicador de buena calidad de las mismas. Las implicaciones descubiertas expresan que las variables lingüísticas “Dinero para alimento”, “Asistencia por semana”, “Relación con compañeros” y “Carrera en relación a mejoras en el futuro” están relacionadas con la variable lingüística “Satisfacción por Carrera”.

Figura 2: Base de datos procesados en Excel

La primera columna de la Tabla 2 muestra reglas relevantes, tanto para la muestra de descubrimiento como para la muestra test, descubiertas con el sistema Eureka-Universe utilizando GMBCL.

Para interpretar las relaciones R_i de la Tabla 2 se realizan representaciones gráficas de las respectivas FPG estimadas en cada caso, con ayuda del sistema Eureka-Universe, donde se tienen en cuenta los estados de las variables lingüísticas involucrados. Así por ejemplo para interpretar la relación R_1 se analizan las gráficas de las funciones de pertenencias estimadas dadas en las Figuras 3 y 4. Se consideran los estados de las variables lingüísticas “Dinero alimento” y “Satisfacción por Carreras” definidas en el cuestionario. Los estados para la variable “Dinero alimento” son: “1: suficiente”, “2: más o menos” y “3: insuficiente”; mientras que los estados de “Satisfacción por Carreras” son “1: satisfecho”, “2: duda” y “3: insatisfecho”.

Predicados Universales	GMBCL	AMBCL
Máximo valor de verdad	0.979428	0.908960
Promedio de valores máximos de veracidad	0.908542	0.873402
Número máximo de predicados relevantes descubiertos por ejecución.	5	1
Mediana de predicados relevantes descubiertos por ejecución.	1.5	0
Total de variables relevantes descubiertas en diez ejecuciones.	13	2

Tabla 1: Cuadro comparativo de resultados con distintas lógicas fuzzy compensatorias.

$G_P(x): '(\forall x)(P(x) \rightarrow \text{Satisfacción}(x))$		
$R_i : P \rightarrow \text{Satisfacción}$	Muestra descubrimiento	Muestra Test
	Valor Verdad $G_P(x)$	Valor Verdad $G_P(x)$
R1 (IMP "Dinero_Alimento", "Satisfacion-carrera")	0.979151	0.983254
R2 (IMP "Asistencia-semana" "Satisfacion-carrera")	0.963075	0.961204
R3 (IMP "Relacion_Comp" "Satisfacion-carrera")	0.97095	0.934230
R4 (IMP "CarreraenRelacionFuturo""Satisfacion-carrera")	0.940794	0.931273

Tabla 2: Cuadro comparativo de valores de verdad para muestra de descubrimiento y test.

La relación R1 se puede interpretar aproximadamente: “Si un alumno expresa que el dinero que dispone para alimento es más o menos suficiente, entonces (el alumno) está satisfecho o duda sobre su satisfacción por la carrera”. Procediendo en forma similar, para el resto de las relaciones relevantes de la Tabla 2, estas se pudieron interpretar: R2: “Si un alumno/a no asiste a clases en la semana entonces puede estar satisfecho o dudar sobre su carrera.”; R3: “Si un alumno tiene una relación mala o no se relaciona con sus compañeros entonces puede estar satisfecho, insatisfecho o dudar sobre la carrera”; y R4: “Si un alumno expresa que la carrera en relación a futuro no mejorará su calidad de vida o duda de ello entonces puede estar satisfecho o dudar sobre su carrera”.

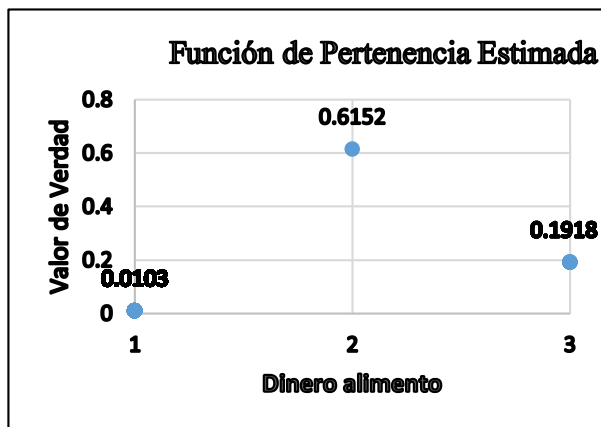


Figura 3: Gráfica de la función de pertenencia estimada para la variable “Dinero para alimento”

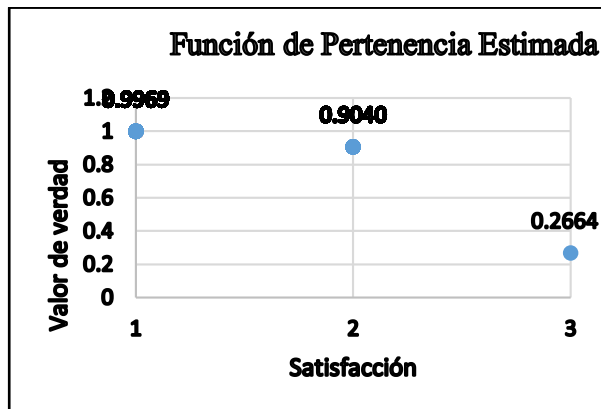


Figura 4: Gráfica de una función de pertenencia estimada para la variable “Satisfacción por la Carrera”, utilizando el sistema Eureka-Universe

Asumiendo los parámetros de la FPG de la tesis, para la R1, como verdaderos, se procede a reanudar los análisis para nuevos descubrimientos de relaciones empleando Eureka-Universe. En las nuevas instancias se considera la posibilidad de conjunciones en los antecedentes de las implicaciones. Como se visualiza en la Figura 5 se descubren más de 10 relaciones relevantes nuevas entre estados lingüísticos. A continuación, se selecciona una de ellas, la relación que se denomina R5 definida en la Tabla 3.

Figura 5: Salidas de relaciones relevantes descubiertas aplicando Eureka-Universe

$G_P(x): (\forall x)(P(x) \rightarrow \text{Satisfacción}(x))$		
$R_i : P \rightarrow \text{Satisfacción}$	Muestra descubrimiento	Muestra descubrimiento
	Valor Verdad GP(x)	Valor Verdad GP(x)
R5 (IMP (AND "Material_Biblioteca", "Edad", "Genero", "Actualizacion_Material" "Estudio_Padre")) "Satisfaccion-carrera")	0.924598	0.903798

Tabla 3: Cuadro comparativo de valores de verdad del predicado universal en estudio, para una relación descubierta, considerando las muestras de descubrimiento y test.

Teniendo en cuenta la información proporcionadas por las representaciones de las funciones de pertenencias generalizadas estimadas y las codificaciones establecidas en el cuestionario para los estados lingüísticos, R5 se pudo interpretar aproximadamente: “Si un alumno/a considera que el Material de Biblioteca es más o menos suficiente a suficiente, tiene 25 años o más, es mujer, expresa que el material educativo está más o menos actualizado a actualizado, el nivel de estudio alcanzado por el padre es terciario o universitario incompleto o nivel superior, entonces el alumno está satisfecho o duda de la carrera”.

Conclusiones

Este trabajo muestra un informe que resulta de un análisis exploratorio de datos proveniente de una encuesta, que resulta de aplicar técnicas de descubrimiento de conocimiento basadas en la Lógica Multivalente. Con el propósito de caracterizar al alumnado, en cuanto a la satisfacción por las carreras que cursan, se aplican procedimientos de descubrimiento de relaciones entre variables lingüísticas, basadas en predicados de la forma $P \rightarrow$ Satisfacción, utilizando el paradigma de la LC.

Desde el empleo del sistema Eureka-Universe, utilizando la GMBCL y AMBCL, se realiza un estudio de relaciones relevantes. Se pudo descubrir relaciones fuzzy que permiten identificar variables lingüísticas relacionadas con la satisfacción del alumno por sus carreras, como también definir nuevas variables subyacentes, como conjunción de variables lingüísticas consideradas en la encuesta, para la caracterización.

El mayor número de relaciones relevantes se obtienen al aplicar GMBCL. Estas al ser interpretadas muestran que la valoración de la calidad educativa en las carreras impartidas, desde la perspectiva del alumnado, están vinculadas con aspectos personales, actitudinales, institucionales, económicos, sociales y culturales

En este estudio el sistema Eureka-Universe resultó una herramienta novedosa, simple de utilizar y muy valiosa por su gran potencial en el descubrimiento de conocimiento útil, a partir de la información de naturaleza fuzzy utilizando distintos paradigmas de la LC.

Consideramos que la información extraída en este estudio, mediante la aplicación de técnicas basadas en Lógica Multivalente, resulta valiosa a la hora de la toma de decisiones institucional y para la mejora de la calidad educativa.

Queda como trabajo a futuro: el estudio relaciones entre variables subyacentes y análisis complementarios para la caracterización del alumnado, como es, por ejemplo, la definición de modelos de clasificación utilizando modelos de regresión apropiados, dentro de Minería de Datos. También el análisis del rol de las funciones de pertenencia generalizadas estimadas y sus posibles combinaciones en la definición e interpretación de las relaciones fuzzy. Consideramos que puede resultar de gran utilidad para definir patrones característicos en torno a la temática de estudio en este trabajo, como también para estudios similares.

Bibliografía

- [1] Traba, L., Barletta, M. y Velázquez, J. (2010). *Teoría y práctica de las organizaciones: herramienta para la gestión de la calidad*. Santa Fe: Universidad Nacional del Litoral.
- [2] Fernández Ziegler, R. (2003). *Planificación y Control de Gestión*. Buenos Aires: Editorial de la Universidad de Quilmes.
- [3] Botero, M.M. y Peña, P. (2006). *Calidad en el servicio: el cliente incógnito*. Suma Psicológica, 13(2), 217-228.
- [4] C. Salvador, C. ; Pozo C.; Alonso, E. (2008)- *Percepción del cliente de los predictores de calidad en el sector servicios*- Boletín de Psicología, No. 94, Noviembre 2008, 69-84 To link to this article:

<https://www.uv.es/seoane/boletin/previos/N94-5.pdf>

[5] Surdez, E. G., Sandoval, M del C. y Lamoyi, C. L. (2018). *Satisfacción estudiantil en la valoración de la calidad educativa universitaria*. *Educación y Educadores*, 21(1), 9-26. DOI: 10.5294/edu.2018.21.1.1

Satisfacción estudiantil en la valoración de la calidad educativa universitaria]

[6] Marín Llanes Luis A., Carro Cartaya Juan Carlos "La Minería de Datos como Herramienta en el Proceso de Inteligencia Competitiva" Consultoría Biomundi, Dirección de Inteligencia Corporativa, Instituto de Información Científica y Tecnológica (IDICT), CUBA, Taller Nacional sobre Inteligencia Empresarial IntEmpres'2000 [Online] available:

<http://www.redciencia.cu/empres/index.htm>

[7] ROJAS DIAZ, J.; CHAVARRO PORRAS, J.; MORENO LAVERDE, R(2009). *Técnicas de lógica difusa aplicadas a la minería de datos*. *Scientia et technica*, [S.l.], v. 3, n. 40, ISSN 2344-7214. [Online] available: <<http://revistas.utp.edu.co/index.php/revistaciencia/article/view/3095/1695>>.

doi:<http://dx.doi.org/10.22517/23447214.3095>.

[8] MARIN ORTEGA, P. G. Á. L. 2009. *Contribución a la modelación de una arquitectura empresarial, para soluciones de inteligencia de negocios*.

[9] DUBOIS D. Y PRADE, H. 1980. *Fuzzy Sets and Systems: Theory and Applications*.

[10] DUBOIS D. Y PRADE, H. 2009. *Fuzzy Inference and Fuzzy Control*. p.190.

[11] ALVES, H. y RAPOSO, M. (2004). *La medición de la satisfacción en la enseñanza universitaria: El ejemplo de la Universidad de Da Beira Interior*. *International Review on Public and Nonprofit Marketing*, Portugal, 1(1): 73-88.

[12] SALINAS A. y MARTÍNEZ, P. (2007). Principales factores de satisfacción entre los estudiantes universitarios. *Revista Internacional de Ciencias Sociales y Humanidades*. México, XVII (001):163-192.

[13] SALINAS, A. (2007). Satisfacción del estudiante y calidad Universitaria: Un análisis explicatorio

en la Unidad Académica Multidisciplinaria Agronomía y Ciencias de la Universidad Autónoma de Tamaulipas. Tesis Doctoral, Universidad de Sevilla.

[14] FERNÁNDEZ, J.; FERNÁNDEZ, S; ÁLVAREZ, A.; MARTÍNEZ, P. (2007). Éxito Académico y Satisfacción de los Estudiantes con la Enseñanza Universitaria. *Revista Electrónica de Investigación y evaluación Educativa*. España, 13(2):203-214.]

[15] Rafael A. ESPIN-ANDRADE, Erick GONZALEZ, Witold PEDRYCZ & Eduardo FERNANDEZ (2016) An Interpretable Logical Theory: The case of Compensatory Fuzzy Logic, *International Journal of Computational Intelligence Systems*, 9:4, 612-626. To link to this article:

<http://dx.doi.org/10.1080/18756891.2016.1204111>

[16] R. ESPÍN ANDRADE, J. MARX GÓMEZ, G. MAZCORRO TÉLLEZ, and E. FERNÁNDEZ GONZÁLEZ, (2004). Compensatory Logic: A Fuzzy Approach to Decision Making. In *Proceedings of 4th International Symposium on Engineering of Intelligent Systems (EIS' 2004)*, Madeira, Portugal.

[17] R. ESPIN ANDRADE and E. FERNÁNDEZ GONZÁLEZ, (2009) "La Lógica Difusa Compensatoria: Una Plataforma para el Razonamiento y la Representación del Conocimiento en un Ambiente de Decisión Multicriterio". *Análisis Multicriterio para la Toma de Decisiones: Métodos y Aplicaciones*. Coedición: Editorial Plaza y Valdes / Editorial Universidad de Occidente.

[18] BOUCHET, A; PASTORE, J.; BRUN, M.; BALLARIN, V. ,(2010). *Logica Difusa Compensatoria basada en la media aritmética y su aplicación en la Morfología Matemática Difusa*.