

Análisis de la incertidumbre en la medición del nivel cognitivo usando Lógica Difusa

Uncertainty analysis in cognitive level measurement using Fuzzy Logic

Constanza Raquel Huapaya¹, Francisco Ángel José Lizarralde¹, Marcela Paula González², Delia Esther Benchoff¹

¹ Universidad Nacional de Mar del Plata, Facultad de Ingeniería, Mar del Plata, Argentina

² Universidad Nacional de Mar del Plata, Facultad de Psicología, Mar del Plata, Argentina

constanza.huapaya@gmail.com, lizarra@fi.mdp.edu.ar, gonzalezmarcelapaula@gmail.com, ebenchoff@fi.mdp.edu.ar

Recibido: 21/12/2021 | Corregido: 28/04/2022 | Aceptado: 11/06/2022

Cita sugerida: C. R. Huapaya, F. Á. J. Lizarralde, M. P. González, D. E. Benchoff, "Análisis de la incertidumbre en la medición del nivel cognitivo usando Lógica Difusa," *Revista Iberoamericana de Tecnología en Educación y Educación en Tecnología*, no. 33, pp. 43-53, 2022. doi: 10.24215/18509959.33.e5

Esta obra se distribuye bajo **Licencia Creative Commons CC-BY-NC 4.0**

Resumen

El objetivo del presente artículo es formular un enfoque del tratamiento de la incertidumbre en la medición del estado cognitivo. Esta incertidumbre se origina en la apreciación subjetiva del evaluador de las acciones del estudiante, sujeta a su experiencia y sensibilidad. Por tal motivo, se utiliza la Lógica Difusa como base de un diseño del modelo del diagnóstico. En el modelo propuesto se identificaron elementos que agregan información relevante a la evaluación si se la compara con la realizada con los métodos tradicionales. Dichos elementos son:

- Variables lingüísticas que agregan información sobre el esfuerzo individual en el aprendizaje a lo largo de un período académico, arrojando información sobre su nivel final de desempeño. Se obtiene un perfil individual.
- Niveles cognitivos basados en la taxonomía revisada de Bloom. A partir de ellos se obtienen perfiles grupales. Asimismo, se ha medido la incertidumbre total de cada grupo de estudiantes.

Se presenta un ejemplo del modelo donde se comparan los valores observados contra los inferidos por el sistema. Asimismo, se presentan valores sobre la performance del modelo.

Palabras clave: Incertidumbre; Nivel cognitivo; Lógica difusa.

Abstract

The aim of the present article is to develop an approach on the uncertainty treatment in the cognitive state measurement. This uncertainty originates in the evaluator's subjective assessment of the student, which is subject to their experience and sensitivity. Therefore, Fuzzy Logic is used as the basis diagnostic model design. In the proposed model, elements that add relevant information to the evaluation have been identified if it is compared to the one carried out with traditional methods. Those elements are as follows:

- Linguistic variables that add information about the individual effort in learning throughout an academic period, providing information about their final performance level. An individual profile is obtained.
- Cognitive levels based on Bloom's revised taxonomy. From those levels, group profiles are obtained. Moreover, the total uncertainty from each group of students has been measured.

An example model is presented in which the observed values are compared to those inferred by the system. In addition, the model's performance values are presented.

Keywords: Uncertainty; Cognitive level; Fuzzy logic.

1. Introducción

El diagnóstico cognitivo es el proceso de inferir el nivel cognitivo a partir de observaciones del desempeño del estudiante [1]. Este diagnóstico es la actividad central de un sistema computacional educativo cuyo objetivo sea construir un modelo dinámico y adaptativo de la actividad del estudiante. El diagnóstico cognitivo de los estudiantes busca identificar sus errores y ayudarlos a alcanzar el conocimiento. El reconocimiento de los errores que poseen los estudiantes es crítico y valioso tanto para los profesores como para los alumnos. Los errores descubiertos permiten un intercambio docente-alumno de gran valor pedagógico: el alumno puede mejorar su técnica de estudio y los maestros sus estrategias de enseñanza.

En un modelo de evaluación del conocimiento, donde no hay comunicación directa entre el docente y el alumno, la información que adquiere el sistema difiere de la que se obtiene de la comunicación cara a cara, donde la interacción es más rica y sin interferencias.

Uno de los principales obstáculos en el proceso de un diagnóstico es la información imperfecta. Esta imperfección se origina en la naturaleza subjetiva de la cognición humana (p.e. la interpretación del docente sobre el rendimiento del estudiante, medido generalmente con métodos estadísticos y aritméticos) y en el modo que el docente comunica sus preferencias y hábitos al sistema.

Además, la adquisición de esta información produce datos sin elaborar, i.e., son datos que pierden precisión debido a la cuantificación posterior. La información registrada por el sistema incorpora estas inexactitudes. Posteriormente, los sistemas deben construir un método de decisión basado en datos imperfectos.

Algunas de estas barreras fueron listadas por [2]: a) el proceso de modelización de un estudiante posee gran cantidad de incertidumbre debido a la naturaleza interpretativa de las observaciones y de las suposiciones realizadas; b) cómo se realiza la interpretación de los datos de la interacción y c) el proceder imprevisto de los estudiantes. El presente artículo explora un posible modelo de solución al problema

A fin de examinar conceptos comparables inmersos en esta temática, vamos a intentar plantear la relación entre imprecisión, inconsistencia e incertidumbre (existen otros términos lingüísticos que no veremos aquí como incompletitud, inexactitud), ver tabla 1. No existe acuerdo generalizado sobre el uso de estos términos [3].

Tabla 1. Información imperfecta

Información imperfecta	Imprecisión	Varios mundos satisfacen la sentencia	
	Inconsistencia	Conclusión incoherente	
	Incertidumbre	vaguedad (=borrosidad)	
		ambigüedad	no-especificidad

1.1. Información imperfecta. Incertidumbre

Generalmente las representaciones del mundo real son imperfectas. Es fácil comprobar que cuando se trabaja con modelos del mundo real, generalmente se hace con aproximaciones. Por el contrario, puede considerarse que la información es perfecta cuando es precisa, consistente y cierta.

El término imperfección de los datos se usa aquí como el concepto más general, el cual engloba a los demás (inconsistencia, imprecisión e incertidumbre)[4]. En la figura 1 se aprecia una taxonomía de la información imperfecta. La imprecisión y la inconsistencia son propiedades asociadas al contenido de una sentencia (sin importar si algún mundo real es compatible con la información) mientras que la incertidumbre nace de la falta de información del mundo real para decidir si la sentencia que constituye la información es verdadera o falsa.

Cuando existe más de un modo de representación del mundo real estamos en presencia de **información imprecisa**. En este caso, lo único que se sabe es que la verdadera representación del mundo real es algún elemento del conjunto M de modelos del mundo.

Cuando se combinan varias sentencias puede aparecer la **inconsistencia**, generando en este caso, algún tipo de error. El conflicto en los datos lleva a conclusiones incoherentes.

La **incertidumbre** es una meta-propiedad que expresa la opinión de un agente sobre cuál de los mundos es el mundo real, esto es, la incertidumbre expresa la opinión de un usuario (o agente en sentido general) sobre cuál de los mundos posibles en M representa mejor al mundo real. La incertidumbre se relaciona con el estado de nuestro conocimiento sobre la relación entre el mundo real y las sentencias que describen ese mundo. Las sentencias pueden ser verdaderas o falsas, pero la incertidumbre aparece cuando el conocimiento que se posee acerca del mundo real no nos permite decidir si la sentencia es verdadera o falsa. La incertidumbre es una propiedad de la relación entre la información y nuestro conocimiento de la realidad.

Un tipo de incertidumbre es la **ambigüedad**. En términos muy generales, la ambigüedad se presenta cuando con información deficiente debe elegirse una de entre una colección de alternativas claramente definidas y mutuamente excluyentes (la exclusión en este contexto significa que solo puede elegirse una de las alternativas). Para posibilitar su tratamiento matemático, entiéndase ambigüedad como la dificultad para determinar el subconjunto al cual pertenece un elemento del cual sólo se dispone de una pobre evidencia o información. La investigación sobre ambigüedad ha permitido establecer dos subclases: el conflicto y la no-especificidad. La evidencia es conflictiva cuando apunta a varios subconjuntos, esto es, expresa discordancia entre conjuntos de alternativas. La no-especificidad alude al tamaño de los subconjuntos, cuanto más elementos

contengan los subconjuntos, menos específica es la evidencia, es decir, expresa la discrepancia entre los tamaños (o cardinalidad) de varios conjuntos de alternativas, el más grande de los conjuntos es el menos específico. Así pues, una evidencia perfecta (no conflictiva y específica) señalará un único subconjunto unitario.

Y finalmente, otro tipo de incertidumbre es la **vaguedad**. La vaguedad acepta más de una interpretación, pero no tiene fundamentos para decidir cuál de ellas debería ser aceptada. En la vaguedad [5,6], la incertidumbre se origina por una deficiente especificación de cada alternativa. En el tratamiento formal de la vaguedad, las alternativas se representan por medio de conjuntos borrosos, los cuales resaltan la dificultad de discriminación.

1.2. Incertidumbre en el diagnóstico cognitivo

El diagnóstico cognitivo puede concebirse como una instancia del diagnóstico general. El diagnóstico cognitivo es una actividad fundamental en sistemas que construyen modelos de usuarios dinámicos. Este proceso suele ser denominado modelado del estudiante.

Valorar el conocimiento del estudiante no es tan fácil como medir la altura y peso de una persona. La medición se hace sobre atributos especiales como es la interpretación de las representaciones mentales y procesos de resolución de problemas, los cuales no son visibles exteriormente. La incertidumbre está presente en este proceso y debe ser considerada [7]. Este es uno de los grandes desafíos de la medición del nivel cognitivo.

Un enfoque para investigar el conocimiento que posee un estudiante, de gran desarrollo actual, es la Evaluación Diagnóstica Cognitiva EDC [8]. La evaluación diagnóstica cognitiva mide las estructuras específicas de conocimiento y habilidades de procesamiento de los estudiantes para brindar información sobre sus fortalezas y debilidades cognitivas, así como para mejorar sus oportunidades de aprendizaje.

Otro enfoque del diseño y construcción de tests es el Diseño Centrado en la Evidencia [9]. El DCE es un conjunto de prácticas que sirven para clarificar las inferencias que se pretenden hacer acerca de los estudiantes en base a las notas de los tests.

Los dos enfoques (EDC y DCE) parten sabiendo que la información utilizada es limitada y conciben a la evaluación como un proceso de razonamiento sobre los datos de la actividad de los alumnos para llegar a afirmaciones acerca de lo que saben.

En los sistemas computacionales de aprendizaje, las interacciones de los estudiantes capturadas y registradas se consideran evidencias para el estado cognitivo del estudiante. Cada interacción individual relacionada a un concepto específico puede considerarse evidencia para determinar el nivel de conocimiento.

La información con incertidumbre nos impulsa a recurrir a técnicas de razonamiento apropiadas para aproximar el diagnóstico cognitivo. Los dos métodos más utilizados en este contexto son Las Redes Bayesianas y la Lógica Difusa (LD). Aquí exploraremos LD por considerarla muy apropiada para tratar la vaguedad.

1.3. Lógica difusa (LD)

El concepto de LD fue propuesto por Lofti A. Zadeh en la Universidad de Berkeley, quien se mostró disconforme con los conjuntos clásicos (llamados conjuntos críps o nítidos o precisos) que solo permiten dos valores de verdad. Zadeh presentó a la LD como una forma de procesar información permitiendo pertenencias parciales a unos conjuntos, que en contraposición a los tradicionales, los llamó conjuntos difusos o borrosos (conjuntos fuzzy). Los conjuntos difusos de Zadeh no pretenden ser una contribución a la filosofía de la vaguedad, sino que se originaron en la necesidad de una representación computacional para términos lingüísticos. La característica principal de los conjuntos difusos es la gradualidad vista como muchas categorías expresadas en lengua natural incluyendo la verdad. Las extensiones de un predicado graduado forman un conjunto difuso, un conjunto donde la transición entre pertenecer y no pertenecer es más bien gradual en lugar de abrupto [10,11].

2. Medición del nivel cognitivo basada en LD

En el campo educativo, los logros en el aprendizaje se consideran como desarrollos suaves y graduales de habilidades cognitivas y por ello se requiere que la evaluación ofrezca información útil acerca del cambio en la organización y la estructura del conocimiento.

Cuando se desea considerar y manipular la incertidumbre asociada con la evaluación subjetiva de los docentes/evaluadores, los modelos que conviene usar son los basados en la Lógica Difusa [12,13,14]. Una de las características principales de la Lógica Difusa es la obtención de descripciones de propiedades mediante valores que cambian suavemente y en forma continua y asocia las particiones semánticas de estos valores a etiquetas lingüísticas. Por ejemplo, la medición borrosa de competencia en un área particular, en el marco de una lista de criterios de medición acordados, es determinar cuán bien un valor particular asociado con estos criterios refleja el ideal semántico del concepto competencia. Con respecto a la descripción del comportamiento del estudiante, la Lógica Difusa puede ser usada para proporcionar una descripción lingüística de las características cognitivas basadas en las opiniones de los docentes y manipular la incertidumbre asociada con la evaluación subjetiva [15,16].

2.1. Propósito del diagnóstico cognitivo

En general, el diagnóstico cognitivo está asociado a las fallas en el conocimiento. Una herramienta de diagnóstico investiga la naturaleza de las fallas, asumiendo que concierne a un componente que no trabaja apropiadamente. El diagnóstico cognitivo trata de encontrar diversas fallas estructurales del conocimiento. El análisis teórico del diagnóstico es de mayor complejidad cuando la estructura del conocimiento no es fija.

Una de las definiciones más difundidas pertenece a K. VanLehn quien sostiene que el proceso de adquisición y análisis de los datos individuales de un alumno se denomina diagnóstico del estudiante [17]. El diagnóstico del estudiante es un método para estudiar efectivamente las mediciones de la actividad de los estudiantes y hacer estimaciones sobre sus características internas. La incertidumbre de los datos es un componente que incide en el diagnóstico del estado cognitivo del estudiante provocando errores de difícil manejo.

El propósito del diagnóstico no considera solamente el rendimiento o conducta final del alumno en una tarea determinada sino que estima la evaluación paso a paso del proceso que genera tal actuación. Esta nueva forma de abordar el diagnóstico es denominado [18,19] *Cognitive diagnosis*.

3. Modelo propuesto para medir el nivel cognitivo

La concepción del modelo propuesto intenta facilitar a los usuarios el traslado de su experiencia docente al diseño del diagnóstico. Por tal motivo se favorece la transferencia de la visión epistemológica personal sobre el aprendizaje y la enseñanza [20]. La epistemología personal refleja las visiones individuales sobre qué es el conocimiento, cómo obtener el conocimiento, y el grado de certidumbre con el cual se puede obtener [21].

El objetivo principal del modelo consiste en mejorar la precisión de un sistema de diagnóstico bajo el presupuesto que contamos con evidencia afectada por la incertidumbre.

La metodología propuesta para el desarrollo del modelo facilita el diseño combinando resultados de pruebas tradicionales y proponiendo reglas de inferencia a fin de arribar a una evaluación final del nivel cognitivo, tanto individual como grupal (p.e. un enfoque EDC). Por tal motivo, se ha procurado alcanzar un diseño con baja dificultad. El docente puede utilizar el contenido específico de su asignatura y los resultados de las pruebas habituales. Esto es, se espera que su práctica educativa pueda ser mejorada mediante el uso de este modelo [22].

Para la presentación del problema del diagnóstico cognitivo se ha elegido un contexto específico en el dominio y en el tiempo. La actividad del estudiante estará relacionada con una porción del dominio del conocimiento, como por ejemplo, Análisis Numérico. Para

la implementación del prototipo del sistema se ha previsto las siguientes acciones previas:

[1] Análisis del dominio del conocimiento: este análisis se hará con la cooperación de docentes expertos en el dominio. Se recortará una sección del dominio (p.e. un tema de una asignatura) que se considere importante para ser evaluada. Por ejemplo, métodos numéricos para ecuaciones diferenciales ordinarias

[2] Descomposición del dominio: el tema será dividido en secciones tan pequeñas (granularidad) como el experto lo considere apropiado. Por ejemplo, dividir a los métodos numéricos para ecuaciones diferenciales ordinarias según los algoritmos basados en la metodología de Runge-Kutta.

[3] Desarrollo de una base de datos que contenga preguntas y ejercicios: la base de datos almacenará diferentes categorías de pruebas a fin de ser presentadas al estudiante de acuerdo a la etapa de la evaluación en que se encuentre (diagnóstica, formativa o sumativa). Las preguntas tendrán asociadas las respuestas correctas.

La definición del modelo de diagnóstico consta de la siguiente 4-upla:

Diagnóstico = $\langle \text{Tema, Comportamiento, Interpretación, NivelFinal} \rangle$

Tema: es el área del conocimiento que se desea evaluar y sobre el que se analiza el diagnóstico cognitivo.

Comportamiento: es un conjunto que guarda los resultados de las pruebas. Se define, considerando por cada tema, un conjunto de n pruebas p_i y el resultado de la evaluación E_i , $i = 1, \dots, n$.

Comportamiento = $(\text{Tema}, (\text{pruebas}(p_1, p_2, \dots, p_n), (E_1, E_2, \dots, E_n)))$

Interpretación es la función que mapea la evidencia del comportamiento del estudiante en su perfil cognitivo individual. La información de entrada es la evidencia mensurable del comportamiento del estudiante y será utilizada para estimar ciertas características cognitivas no evidentes de los estudiantes (ver figura 2).

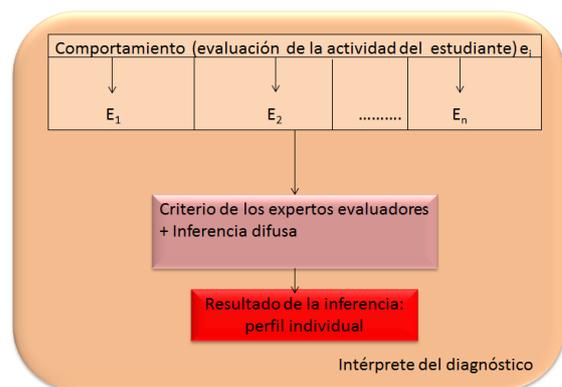


Figura 2. Intérprete

La epistemología personal de cada docente proporciona la base de las reglas difusas que crearán el nivel del conocimiento del estudiante.

3.1. Perfiles cognitivos

La información que se encuentra en el diagnóstico depende de la naturaleza del sistema. Esta información puede variar desde las respuestas a preguntas en un cuestionario, hasta la resolución de problemas hechos por el estudiante. Asimismo, el historial académico del estudiante puede ser un componente de diagnóstico. La salida del módulo de diagnóstico es difícil de definir. Como objetivo primario, un modelo de diagnóstico intenta reflejar con cierta precisión el estado de conocimiento del alumno y puede entregar un perfil cognitivo del estudiante.

El término perfil se deriva de la psicología y se entiende como el conjunto de medidas o características de una persona o grupo de personas, cada una de las cuales se expresa en la misma unidad de medición. Si nos referimos al perfil de un usuario, en general, puede definirse como una colección de información personal. La información se almacena sin el agregado de la interpretación de esos datos. Estas propiedades son almacenadas luego de asignarle valores. Estos valores pueden ser los finales o cambiar a lo largo del tiempo.

Si tratamos con el perfil cognitivo [23] de un estudiante la información se refiere al conjunto de habilidades y estrategias de aprendizaje que utiliza el estudiante. Estos perfiles representan habilidades cognitivas, competencias intelectuales, intenciones, estilos de aprendizaje, preferencias e interacciones con el sistema. De manera similar, [24] considera un estado del conocimiento que comprende el nivel del conocimiento, errores y malos entendidos del estudiante.

En la figura 3 se presenta un esquema general del diagnóstico donde la entrada es la información registrada de las actividades del estudiante y la salida son los perfiles cognitivos.



Figura 3. Modelo del diagnóstico

Como toda evaluación de los docentes/evaluadores existe el fenómeno de la vaguedad cuando se trata de analizar el nivel de adquisición del conocimiento por parte de los alumnos (es difícil precisar ese nivel de adquisición). El perfil individual toma la evidencia de cada estudiante a lo

largo de un ciclo lectivo y muestra una variable lingüística sobre aspectos del aprendizaje.

Además, se examinan algunas características cognitivas comunes de los alumnos. A tal fin se han definido perfiles de aprendizaje de grupos de estudiantes. En este caso se presenta el fenómeno de la ambigüedad (no es específico a qué perfil pertenece cada alumno) [25, 26].

Los perfiles están basados en la taxonomía revisada de Bloom [27]. Luego se mide la incertidumbre presente en cada grupo.

3.1.1. Perfiles individuales

Los docentes experimentados definen una descripción subjetiva de las respuestas del estudiante a través de los variables lingüísticas E_1, E_2, \dots, E_n sobre las evidencias mensurables (como las notas de exámenes). Por ejemplo, E_1 puede representar el nivel de aprobación. Cada variable puede tomar un número diferente de valores lingüísticos.

El conjunto $P(E_i) = \{V_{i1}, V_{i2}, \dots, V_{im}\}$ es el conjunto de valores de E_i . Continuando con el ejemplo, V_{i1} : bajo, V_{i2} : medio, V_{i3} : alto. $P(E_i)=$ medio indica que el estudiante alcanzó un nivel medio en su aprobación (ver figura 4).

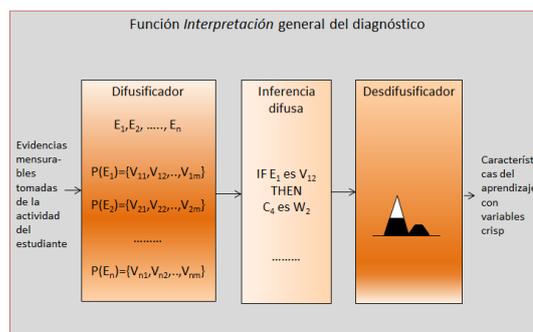


Figura 4. Esquema de la función Interpretación

La inferencia difusa representa el razonamiento de los docentes cuando categorizan cualitativamente a los estudiantes de acuerdo a sus habilidades cognitivas. La descripción cualitativa de las características de los estudiantes se representa con las variables lingüísticas $C_k, k = 1, \dots, p$. Cada una de ellas puede tomar cantidad diferente de valores lingüísticos: $VL(C_k) = \{W_{k1}, W_{k2}, \dots, W_{kl}\}$.

Por ejemplo, un experto puede decidir que la aprobación de ciertas pruebas guiará el nivel final de aprobación. A continuación se presentan 2 reglas que podrían formar parte de la inferencia difusa:

Regla 1. Si (nivel de aprobación de las pruebas es alto) y (nota final respecto a la media está por arriba de la media) y (progresión de notas es creciente) Entonces (nivel de conocimiento es alto)

Regla 2. Si (nivel de aprobación de las pruebas es medio) y (nota final respecto a la media está alrededor de la media) y (progresión de notas es estable) Entonces (nivel de conocimiento es medio)

La desdifusificación (o defuzificación) emula la decisión final de los docentes cuando clasifica a los estudiantes en uno de los valores $W_{k1}, W_{k2}, \dots, W_{kd}$.

Por último, puede llegarse a un ResultadoFinal donde se calculan los valores críps. Puede recurrirse a diversos métodos como el procedimiento del centro de gravedad: $\text{ResultadoFinal}(W_{ki}) = \{x \text{ tal que } 0 < x \leq 10, x \in \mathbb{R}\}$.

3.1.2. Perfiles grupales

El modelo para medir la ambigüedad en la adquisición del conocimiento de un grupo de estudiantes [28] parte de perfiles cognitivos construidos sobre una teoría de aprendizaje.

Este artículo ha tomado la taxonomía de Bloom. La taxonomía de Objetivos Educativos se ha transformado en una herramienta de gran importancia para estructurar y entender el proceso de aprendizaje. Según el autor, el aprendizaje comprende: el dominio cognitivo (procesamiento de la información, conocimiento y habilidades mentales); el dominio afectivo (actitudes y sentimientos) y el dominio psicomotor (habilidades físicas, manuales y de manipulación). La idea central de esta taxonomía es definir objetivos educativos, es decir, aquellos que los educadores buscan que sus alumnos sepan. Tienen una estructura jerárquica que va de lo más simple a lo más complejo o elaborado. Cuando los docentes programan sus clases pueden considerar estos niveles y mediante las diferentes actividades, avanzar hacia un nivel superior.

La taxonomía revisada no trata sobre herramientas y tecnologías sino que esta es sólo el medio para alcanzar las seis acciones que se aprecian en la figura 5. Uno de los aspectos clave de esta revisión es el cambio de los sustantivos de la propuesta original a verbos, para significar las acciones correspondientes a cada categoría. Otro aspecto fue considerar la síntesis con un criterio más amplio y relacionarla con crear (considerando que toda síntesis es en sí misma una creación), además, se modificó la secuencia en que se presentan las distintas categorías.

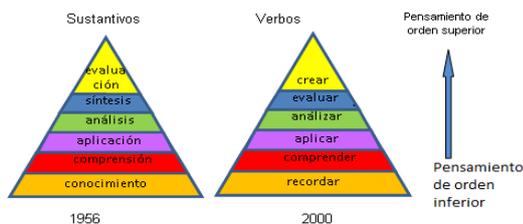


Figura 5: Taxonomía de Bloom revisada

Para definir los niveles cognitivos, se parte considerando un grupo de N estudiantes durante su proceso de aprendizaje de un tópico específico. Tomando como base los seis niveles de la taxonomía revisada de Bloom, se denominará X_1, X_2 y X_3 a los niveles cognitivos según su complejidad creciente: X_1 recordar y comprender; X_2 : aplicar y analizar, y X_3 : evaluar y crear.

Cada uno de estos tres estados, a su vez, será asociado a una de las siguientes etiquetas lingüísticas: muy bajo

(MB), bajo (B), medio (M), alto (A) y muy alto (MA). De este modo, se dispondrá del conjunto $U = \{MB, B, M, A, MA\}$.

Se denota como $n_{X_i,MB}, n_{X_i,B}, n_{X_i,M}, n_{X_i,A}, n_{X_i,MA}$ al número de estudiantes del grupo que han alcanzado un nivel de adquisición de conocimiento muy bajo, bajo, medio, alto o muy alto respectivamente en cada estado X_i . A fin de simplificar la notación se escribirá más abreviadamente:

$n_{i,MB}, n_{i,B}, n_{i,M}, n_{i,A}, n_{i,MA}$ Por ejemplo, $n_{1,M}$ es la cantidad de alumnos que lograron el nivel medio en el estado de recordar y comprender.

Definimos la función de pertenencia $m_{X_i}(t)$ ($i=1,2,3$) para t de U del siguiente modo (N es la cantidad de estudiantes del grupo):

$$m_{X_i}(t) = \begin{cases} 1 & \text{si } \frac{4N}{5} \in n_{i,t} \in N \\ 0,75 & \text{si } \frac{3N}{5} \in n_{i,t} \in \frac{4N}{5} \\ 0,5 & \text{si } \frac{2N}{5} \in n_{i,t} \in \frac{3N}{5} \\ 0,25 & \text{si } \frac{N}{5} \in n_{i,t} \in \frac{2N}{5} \\ 0 & \text{si } 0 \in n_{i,t} \in \frac{N}{5} \end{cases}$$

Por ejemplo, si $m_{X_1}(B) = 0,5$ indica que el nivel B (bajo) en el nivel X_1 tiene un valor de pertenencia de 0,5.

Para caracterizar a los niveles cognitivos X_i , a cada uno se le asocia el siguiente conjunto:

$$P_i = \{(t, m_{X_i}(t)) : t \in U\}, i = 1, 2, 3$$

El ejemplo, $P_1 = \{(MB, 0), (B, 0,75), (M, 0,25), (A, 0), (MA, 0)\}$ representa una instancia del nivel X_1 .

Se define como **perfil de un estudiante** a la terna $w = (a, b, c)$ donde a, b , y c son elementos de U . Además, se determina que la terna está bien ordenada si a corresponde a un grado adquisición de conocimiento igual o mayor que b , y b corresponde a un grado de adquisición mayor o igual a c . Por ejemplo la terna (MA,M,B) está bien ordenada mientras que la terna (MB,A,M) está desordenada.

Se define el grado de pertenencia de un perfil w del siguiente modo:

$$m_R(w) = \begin{cases} m_{X_1}(a)m_{X_2}(b)m_{X_3}(c) & \text{si } w \text{ es una terna bien ordenada} \\ 0 & \text{en todo otro caso} \end{cases}$$

La función m_R vale cero cuando la terna no está bien ordenada a fin de evitar casos inadmisibles como por ejemplo perfiles donde un estudiante no comprenda los conceptos fundamentales de un tópico específico pueda, luego, crear nuevo conocimiento a partir de ellos. Un caso de estos perfiles es (M,MB,A). A fin de representar todos los posibles perfiles de un estudiante bajo la consideración de los tres estados presentados se define la siguiente relación difusa:

$$R = \{f(w, m_R(w)) : w = (a, b, c) \in U^3\}$$

La cantidad total de posibles perfiles es 125 (5^3).

La posibilidad r_w de un perfil es definida del siguiente modo, donde $\max(m_R(w))$ denota el valor máximo de $m_R(w)$:

$$r_w = m_R(w) / \max(m_R(w))$$

4. Ilustración del modelo

Se presenta un ejemplo concreto para exponer el modelo. En este caso se hará una estimación del nivel de conocimiento alcanzado por el estudiante en el momento de la terminación del dictado de una asignatura (ver figura 6). Esto es, se intenta capturar un indicador que integre el logro académico alcanzado basándose en información que se genera en nuestras cátedras universitarias. Se analiza la información adquirida por el sistema sobre las notas de exámenes parciales, finales y pruebas computacionales específicas.

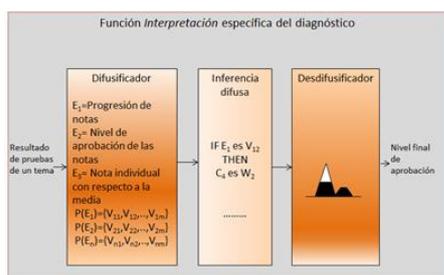


Figura 6. Función interpretación

La implementación del modelo ha sido desarrollada con FisPro [29]. FisPro es una herramienta para diseño y optimización de sistemas de inferencia difusa.

FisPro permite diseñar sistemas difusos a partir del conocimiento experto perteneciente a una determinada área de conocimiento, o, proyectar un sistema de inferencia difusa desde datos del problema a solucionar.

La cooperación entre el conocimiento y los datos genera mejores sistemas. Si bien los sistemas pueden ser diseñados sólo a partir de datos, el experto puede visualizar las variables más importantes del sistema así como describir su comportamiento usando reglas lingüísticas. Las reglas expertas se basan en la experiencia y presentan un alto nivel de generalización.

Los datos, una forma extensiva de conocimiento inducido, son básicamente piezas de evidencia positiva. Pero, a su vez, son incompletos por naturaleza. Una muestra no puede cubrir todas las posibles situaciones (figura 7).



Figura 7. Esquema del intérprete

Tanto el diseño de reglas expertas como la inducción automática pueden ser combinados para crear sistemas más completos y de mejor rendimiento. Este es el caso seguido en el diseño de Diagnóstico. La colaboración entre conocimiento y datos se puede dar en los diferentes pasos del diseño. El experto puede intervenir en todos los pasos del diseño: en la partición difusa y en el diseño de las reglas, donde el conocimiento puede complementar el diseño agregando funciones de pertenencia (FP) y reglas donde no haya datos disponibles, optimización de los parámetros de FP y validación del sistema. En este último paso, se puede controlar tanto la exactitud numérica como el análisis del conocimiento inducido.

Asimismo, hay herramientas que ilustran el mecanismo de razonamiento y otras que miden el rendimiento del sistema sobre los conjuntos de datos. FisPro posee como punto destacado la opción de comprender fácilmente cómo opera el sistema manipulando las restricciones impuestas a los algoritmos para lograr reglas de razonamientos fáciles de interpretar.

4.1. Comportamiento

En los siguientes apartados se expone la partición difusa de las variables de entrada y salida. Asimismo, se presenta la base de reglas de inferencia creadas con ellas.

El modelo desarrollado para estimar la variable lingüística de salida **NivelFinal** usa tres variables lingüísticas de entrada: **Progresión de notas**, **Nivel de aprobación de las pruebas** y **Nota final respecto a la media del curso**. A continuación, se presenta una breve descripción de cada una de ellas.

Progresión de notas: se considera importante la evolución de las notas a lo largo del período académico. Un crecimiento importante de las notas indica un mejoramiento en el aprendizaje, mientras que un decrecimiento sostenido implica lo contrario. Los otros casos pueden ser considerados como no significativos (en el sentido de que no ascienden o descienden en forma monótona) y no inciden sobre el cálculo del nivel del conocimiento. Sus términos son tres conjuntos difusos: *creciente*, *estable* y *decreciente*. Las funciones de pertenencia son de tipo gaussiano. Considerando que en el ciclo académico se registran n notas, el cómputo de estos valores se inicia calculando la recta resultante de la aproximación por cuadrados mínimos, siendo la variable x el tiempo. Por ejemplo, si el valor de la pendiente es mayor a -1 y menor que 1 , la variable **Progresión de notas** toma el valor *estable*.

Nivel de Aprobación de las pruebas: es la información más relevante para decidir el nivel de conocimiento de cada estudiante. Los resultados de las pruebas (incluyendo exámenes) son almacenados en el sistema. La variable toma 5 términos: *no satisfactorio*, *poco satisfactorio*, *regular*, *satisfactorio* y *muy satisfactorio*. Las funciones de pertenencia son de tipo gaussiano (ver figura 8).

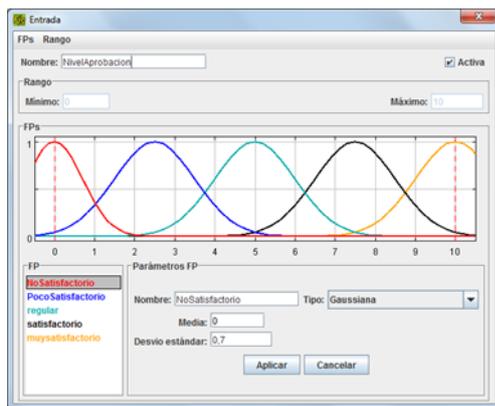


Figura 8. Variable lingüística Nivel de Aprobación

Nota final respecto a la media del curso: aquí se compara el rendimiento individual de cada alumno a través de su nota Nivel de Aprobación con respecto al desempeño promedio de todos los estudiantes del curso. Se han determinado los siguientes 3 términos: *debajo de la media*, *alrededor de la media* y *arriba de la media*. Las funciones de pertenencia son de tipo gaussiano.

La variable de salida **Nivel Final** de conocimiento toma 5 términos: *deficiente*, *regular*, *medio*, *bueno*, y *muy bueno* indicando la categoría alcanzada por el estudiante según la información analizada en la progresión de notas, nivel de aprobación de las pruebas y nota final respecto a la media del curso (Ver figura 9). El nivel estimado provee una guía al evaluador, asistiéndole en su toma de decisión.



Figura 9. Variable de salida

4.2. Creación de reglas

Las reglas de inferencia fueron diseñadas completamente por expertos del área de Análisis Numérico. Para inferir los valores del nivel de conocimiento, se ha considerado que Nivel de aprobación de las pruebas es la variable de entrada que guiará fundamentalmente la categorización y que las otras dos variables consideradas (progresión de notas y nota respecto a la media), contribuirán en menor medida. Los expertos involucrados en el desarrollo propusieron 30 reglas en lenguaje natural, dos de las cuales se muestran a continuación:

Regla 1. Si (Nota final respecto a la media está por debajo de la media) y (Progresión de notas es decreciente) y (Nivel de aprobación de las pruebas es no satisfactorio) Entonces (Nivel Final es deficiente).

Regla 15. Si (Nota final respecto a la media está alrededor de la media) y (Progresión de notas es estable) y (Nivel de aprobación de las pruebas es satisfactorio) Entonces (Nivel de conocimiento es bueno).

En la figura 10 se aprecia la inferencia difusa representada por 30 reglas difusas.

Regla	Peso	Activa	SI RespectoMedia	Y Progresion	Y NivelAprobacion	ENTONCES Nivel...
1	1	✓	decremento	decreciente	NoSatisfactorio	deficiente
2	1	✓	decremento	decreciente	PocoSatisfactorio	deficiente
3	1	✓	decremento	decreciente	satisfactorio	regular
4	1	✓	decremento	estable	NoSatisfactorio	deficiente
5	1	✓	decremento	estable	PocoSatisfactorio	deficiente
6	1	✓	decremento	estable	regular	medio
7	1	✓	decremento	creciente	NoSatisfactorio	deficiente
8	1	✓	decremento	creciente	PocoSatisfactorio	deficiente
9	1	✓	decremento	creciente	regular	regular
10	2	✓	ArededorMedia	decreciente	regular	regular
11	1	✓	ArededorMedia	decreciente	muySatisfactorio	medio
12	1	✓	ArededorMedia	estable	NoSatisfactorio	deficiente
13	1	✓	ArededorMedia	estable	PocoSatisfactorio	regular
14	1	✓	ArededorMedia	estable	regular	medio
15	1	✓	ArededorMedia	estable	satisfactorio	bueno
16	1	✓	ArededorMedia	creciente	regular	medio
17	1	✓	ArededorMedia	creciente	satisfactorio	bueno
18	1	✓	ArribaMedia	decreciente	regular	regular
19	1	✓	ArribaMedia	decreciente	satisfactorio	bueno
20	1	✓	ArribaMedia	decreciente	muySatisfactorio	bueno
21	1	✓	ArribaMedia	estable	regular	bueno
22	1	✓	ArribaMedia	estable	satisfactorio	bueno
23	1	✓	ArribaMedia	estable	muySatisfactorio	muybueno
24	1	✓	ArribaMedia	creciente	regular	bueno
25	1	✓	ArribaMedia	creciente	satisfactorio	bueno
26	1	✓	ArribaMedia	creciente	muySatisfactorio	muybueno
27	1	✓	ArribaMedia	estable	muySatisfactorio	muybueno
28	1	✓	decremento	creciente	regular	regular
29	1	✓	decremento	creciente	satisfactorio	medio
30	1	✓	decremento	creciente	muySatisfactorio	bueno

Figura 10. 30 reglas difusas

4.3. Comportamiento del sistema

El comportamiento del sistema puede ser observado a través de herramientas gráficas. La visualización puede ser en 2D y 3D. En particular, en la figura 11, puede analizarse la influencia de la variable **Nivel de Aprobación** en la variable **Nivel Final** modificando las otras dos variables de entrada.

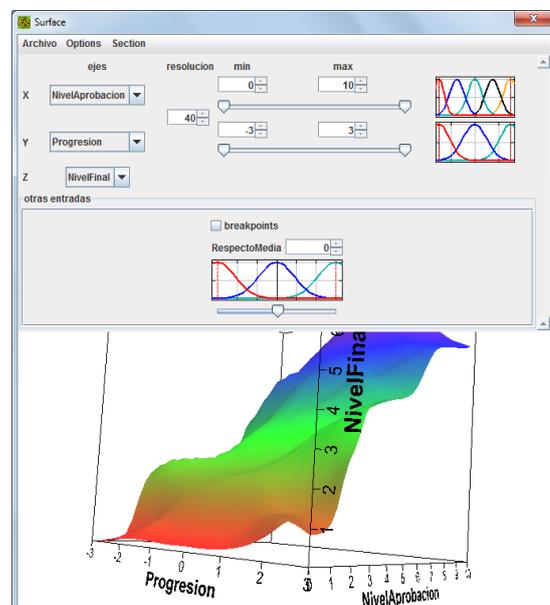


Figura 11. Inferencia del sistema

La inferencia en el sistema es observada moviendo el cursor en cada rango de las variables, apreciando el comportamiento de cada regla y agregación. En este caso, se verifica la preponderancia de la variable de entrada

Nivel de Aprobación al modificar los valores de entrada. El color de la superficie cambia del rojo (mínimo) al púrpura (máximo).

4.3.1. Validación de la inferencia

La evaluación del sistema se realizó con los datos de 54 estudiantes. Si se desea analizar la inferencia para un caso particular, se tiene que hacer doble click en la fila correspondiente en la tabla de datos y se abre la pantalla de inferencia con el valor inferido (ver figura 12).

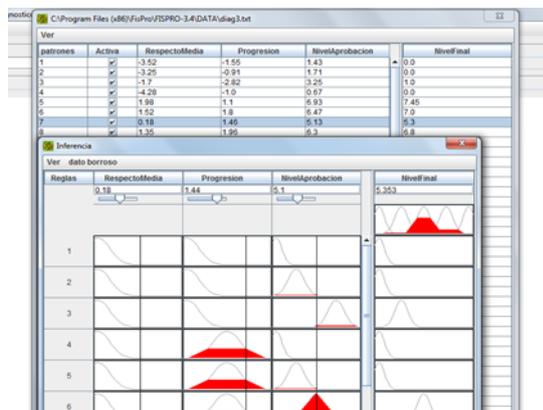


Figura 12. Inferencia de un caso

El valor inferido es 5.353 mientras el valor de la evaluación de docentes es 5.3. Otra modalidad visual de la relación de los datos se aprecia en la figura 13 donde se ve la correlación entre los datos observados y los inferidos.

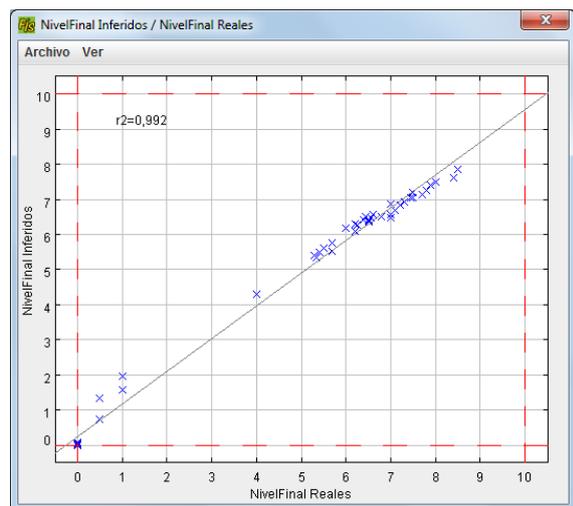


Figura 13. Correlación entre datos observados e inferidos

La evaluación de la inferencia se hace con la función Performance. Dados los datos, la función estima los siguientes tres índices:

Error máximo: valor absoluto máximo de las diferencias entre los valores originales e inferidos de la variable de salida (Nivel Final)

Índice de cobertura: algunos datos se pueden corresponder a las reglas en un nivel bajo. Un dato se considera como blanco si su grado de correspondencia

acumulado es menor que el umbral consignado (en este caso 0.1)

Índice de performance: para la salida borrosa que se ha elegido se usa la función SetErrorIndex y toma tres valores: PI, RMS (Root Mean Squared Error) y MAE (Mean Absolute Error).

Los resultados se aprecian en la figura 14.

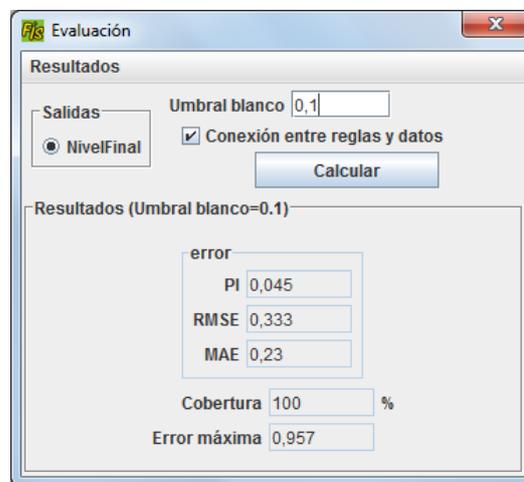


Figura 14. Evaluación del sistema

4.3.2. Perfiles grupales

A modo de ejemplo se definió los siguientes tres niveles cognitivos para una asignatura de programación:

Nivel P₁: expresar la comprensión de conceptos de programación básica. Por ejemplo, variable, tipo de datos, etc.

Nivel P₂: aplicar y analizar las estructuras de datos y estructuras de control de ejecución. Por ejemplo, arreglos, ciclos, etc.

Nivel P₃: evaluar y crear conocimiento mediante el desarrollo de programas en computadora. Por ejemplo, programas, algoritmos para evaluar un polinomio, etc.

Un perfil específico, a partir de este ejemplo de niveles cognitivos, es (A,M,B) indicando que este alumno posee un nivel Alto en la comprensión de conceptos de la programación básica; un nivel Medio en el análisis de las estructuras de datos y ejecución, y finalmente, un nivel Bajo en el desarrollo de programas. En la figura 15 se aprecia un ejemplo de perfiles grupales.

	Perfiles			Grado de pertenencia m_A	Posibilidad r_A
	P ₁	P ₂	P ₃		
1	Muy alto	Muy alto	Bajo	0.0938	1.0000
2	Medio	Medio	Bajo	0.0625	0.6667
3	Muy alto	Muy alto	Muy bajo	0.0469	0.5000
4	Muy alto	Muy alto	Medio	0.0469	0.5000
5	Medio	Medio	Muy bajo	0.0313	0.3333
6	Medio	Medio	Medio	0.0313	0.3333
7	Muy alto	Medio	Bajo	0.0313	0.3333
8	Muy alto	Medio	Muy bajo	0.0156	0.1667
9	Muy alto	Medio	Medio	0.0156	0.1667
Conflicto: 0.4918					
No especificidad: 1.4962					
Incertidumbre total: 1.9980					

Figura 15. Perfiles con función de pertenencia distinta de cero y sus posibilidades pertenecientes al grupo 1 con 54 estudiantes (N=54)

Conclusiones

El punto central del presente artículo es el análisis de la incertidumbre involucrada en la evaluación del rendimiento de estudiantes universitarios. La propuesta emplea información tradicional obtenida de la actividad de

los estudiantes a lo largo de un período académico. El modelo propuesto calcula la ambigüedad que se manifiesta cuando se desea asignar un perfil cognitivo específico a grupos de estudiantes. La experiencia mostró un camino para obtener información adicional a partir de las notas tradicionales y proponer al evaluador otra faceta de la actividad del estudiante, como son los perfiles cognitivos.

La comparación de niveles cognitivos y la incertidumbre total de los grupos conduce a un mejoramiento de estrategias docentes. De la experiencia se desprende que la metodología propuesta puede ser usada en otras áreas del conocimiento.

Se ha usado la Lógica Difusa como fundamento de un modelo de diagnóstico. Una de las principales ventajas de este modelo es que permite una representación interpretable del conocimiento. Esto es, el conocimiento está basado en reglas tanto para el razonamiento bien definido como para el razonamiento intuitivo fruto de experiencia.

Un aporte interesante de esta investigación es la flexibilidad en el uso de los criterios cualitativos y cuantitativos propuestos por los docentes para la evaluación de los estudiantes. Esta característica permite un fácil mejoramiento tanto de las variables lingüísticas como las reglas difusas. Por ejemplo, en el modelo presentado se ha llegado a la conclusión que cuatro de las 30 reglas pueden ser descartadas siguiendo la opinión de los docentes experimentados.

Como futura línea de investigación se agregarán nuevas evidencias al modelo de diagnóstico como la fecha de las evaluaciones, nivel de dificultad de los tópicos y se incorporarán nuevas variables lingüísticas, como nivel de interés y motivación.

Referencias

- [1] S. Ohlsson, P. Langley, "Psychological evaluation of path hypotheses in cognitive diagnosis," in *Learning issues for intelligent tutoring systems*, H. Mandl y A. Lesgold Ed. New York: Springer Verlag, 1988, pp. 42-62.
- [2] R. Mazza, C. Milani, "Exploring usage analysis in learning systems: Gaining insights from visualizations," in the Workshop on Usage Analysis in Learning Systems at the 12th International Conference on Artificial Intelligence in Education, New York, 2005, pp. 1-6.
- [3] F. Dowlathahi, L. J. Kohout, "Intentional possibilistic approach to dealing with incompleteness, vagueness and

uncertainty," *Fuzzy sets and Systems*, vol. 25, pp. 277-295, 1988

[4] P. Smets, "Imperfect Information: Imprecision and Uncertainty," in *Uncertainty Management in Information Systems*, Motroy Smets, Ed. Springer, 1997, pp. 225-254.

[5] B. Russell. "Vagueness," *The Australasian Journal of Psychology and Philosophy*, vol. 1, no. 2, pp. 84-92, 1923.

[6] L. Godo H. Prade D. Dubois, F. Esteva., "An information-based discussion of vagueness: six scenarios leading to vagueness" in *Handbook of Categorization in Cognitive Science*, Ed. Elsevier, Amsterdam, 2005, Chapter 40

[7] C. Huapaya, F. Lizarralde, G. Arona, J. Vivas. "Análisis del aprendizaje usando el nivel de incertidumbre en la adquisición de conocimiento", in Proceedings of the WICC 2013, Entre Ríos, 2013, pp. 1126-1130.

[8] J. P. Leighton, M.J. Gierl, "Why cognitive diagnostic assessment?," in *Cognitive diagnostic assessment for education: Theory and applications Rds.*, J. P. Leighton, M. J. Gierl, Ed. Cambridge University, 2007, pp. 3-18.

[9] M. J. Zieky, "An introduction to the use of evidence-centered design in test development," *Psicología educativa*, vol. 20, pp. 79-87, 2014

[10] L.A. Zadeh, "Fuzzy sets," *Information and Control*, vol. 8, pp. 338-353, 1965.

[11] L. A. Zadeh, "Fuzzy logic, neural networks and soft computing," *Communication of the ACM*, vol. 37, pp. 77-84, 1993

[12] S. Ibrahim, K. Seong-in, "A fuzzy system for evaluating students' learning achievement," *Expert Systems with Applications*, vol. 36, pp. 6236-6243, 2009

[13] B. Alibek, S. Sharipbay, U. Ulyukova, S. Talgat, K. Batyrkhan, "Student's performance evaluation by fuzzy logic," *Procedia Computer Science*, vol. 102, pp. 98-105, 2016

[14] K. Chrysafiadi, M. Virvou, "Evaluating the integration of fuzzy logic into the student model of a web-based learning environment," *Expert System with Applications*, vol. 39, pp. 13127-13134, 2012

[15] P. Petts, G. Swift, S. Nightingale, "Evaluation of Student Understanding of Uncertainty in Level 1 Undergraduate Physics Laboratories" in Proceedings of the 10th International Conference New Perspectives in Science Education, Florence,, 2021, pp. 413-419.

[16] G. Gokmena, T. Akincib, M. Tektaúc, N. Onatc, G. Kocyigita, N. Tektaú, "Evaluation of student performance in laboratory applications using fuzzy logic," *Procedia Social and Behavioral Sciences*, vol. 2, pp. 902-909, 2010

[17] K. VanLehn, "Student modelling," in *Foundations of Intelligent Tutoring Systems*, M. Polson Ed., Hillsdale. N.J Lawrence Erlbaum Associates, 1988, pp. 55-78.

[18] P. Langley, J. Wogulis, S. Ohlsson, "Rules and principles in cognitive diagnosis," in *Diagnostic*

Monitoring of Skill and Knowledge Acquisition, N. Frederiksen, R. Glaser, A. Lesgold and M. G. Shafto, Ed. Hillsdale, NJ: Lawrence Erlbaum Associates, 1990, pp. 217–250.

[19] P. D. Nichols, S. F. Chipman, R. L. Brennan, *Cognitively Diagnostic Assessment*. New York: Routledge, 1995.

[20] B. K. Hofer, L. D. Bendixen. "Personal epistemology: Theory, research, and future directions," in *APA educational psychology handbook, Vol. 1. Theories, constructs, and critical issues*, K. R. Harris, S. Graham, T. Urduan, C. B. McCormick, G. M. Sinatra, J. Sweller Ed. American Psychological Association. 2012, pp. 227–256.

[21] A. S. Eryasar, A. Kilinc, "The Coherence Between Epistemologies and SSI Teaching", *Science and Education*, vol. 31, no. 1, pp. 123–147, 2022

[22] A. Rieu, T. Leuders, K. Loibl, "Teachers' diagnostic judgments on tasks as information processing—The role of pedagogical content knowledge for task diagnosis," *Teaching and Teacher Education*. vol. 111, 103621, 2022.

[23] E. Rich, "Users are individuals :- individualizing user models," *Int. J. Man-Machine Studies*, vol. 18, pp. 199–214, 1983

[24] E. Wenger, *Artificial Intelligence and Tutoring Systems: Computational and Cognitive Approaches to the Communication of Knowledge*, Morgan Kaufmann, 1998.

[25] C. Law, "Using fuzzy numbers in educational grading systems," *Fuzzy sets and systems*, vol. 83, no. 3 pp. 311–323, 1996

[26] C. Sun, D. Chang, "Fuzzy assessment of learning performance of junior high school students" in *Proceedings of the First National Symposium on Fuzzy Theory and Applications*, Hsinchu, Taiwan, 1993, pp. 1–10.

[27] L. D. Anderson, B. Krathwohl, *A Taxonomy for Learning, Teaching and Assessing: a Revision of Bloom's Taxonomy of Educational Objectives*, New York Longman, 2001.

[28] M. Voskoglou, "Transition across levels in the process of learning: a fuzzy model", *Journal of Mathematical Modeling and Application*, vol 1 no. 1 pp. 37–44, 2009.

[29] S. Guillaume, B. Charnomordic, "Fuzzy inference systems: an integrated modelling environment for collaboration between expert knowledge and data using Fispro," *Expert Systems with Applications*, vol. 39, no. 10, pp. 8744–8755, 2012

Información de Contacto de los Autores:

Constanza Raquel Huapaya
Castelli 698
Mar del Plata
Argentina

constanza.huapaya@gmail.com
ORCID: <https://orcid.org/0000-0001-9219-0291>

Francisco A. J. Lizarralde
Av. Juan B. Justo 4302
Mar del Plata
Argentina

flizarra@fi.mdp.edu.ar
ORCID: <https://orcid.org/0000-0002-5401-6791>

Marcela Paula González
Arenales 3292
Mar del Plata
Argentina

gonzalezmarcelapaula@gmail.com
ORCID: <https://orcid.org/0000-0001-7176-7194>

Delia Esther Benchoff
Av. Juan B. Justo 4302
Mar del Plata
Argentina

ebenchoff@fi.mdp.edu.ar
ORCID: <https://orcid.org/0000-0001-5467-579X>

Constanza Raquel Huapaya

Especialista en Tecnología Informática Aplicada en Educación. (UNLP) Investigadora en Ambientes Inteligentes de Aprendizaje. Personalización.

Francisco A. J. Lizarralde

Magister en Tecnología Informática Aplicada en Educación (UNLP). Especialista en Tecnología Informática Aplicada en Educación (UNLP). Investigador en Análisis de Redes Semánticas - Entornos Virtuales 3D de Aprendizaje (UNMdP).

Marcela Paula González

Magister en Psicología Social (UNMdP). Investigadora en Personalización y Adaptación del Aprendizaje en Ambientes Virtuales.

Delia Esther Benchoff

Magister en Tecnología Informática Aplicada en Educación (UNLP). Investigadora en Personalización y Adaptación del Aprendizaje en Ambientes Virtuales.