
Un modelo difuso para el diagnóstico de alumnos universitarios con bajo rendimiento

A fuzzy model for the diagnosis of low-achieving college students

Ramón De la Cruz Lázaro, Maritza Espinoza Limaylla, Roque Mizuno Inafuko,
Marcos Sotelo Bedón, María Vidaurre Gastelú.

Universidad Nacional Mayor de San Marcos, Campus Universitario s/n, Lima, Perú

rmdlazaro@gmail.com, meslim3@gmail.com, msperu21@gmail.com, roque.mizuno@gmail.com,
evidaurreg@yahoo.com.

RESUMEN

Una de las tareas más complejas de un docente es la evaluación fidedigna del progreso de sus estudiantes. Los criterios individuales en la evaluación del nivel de competencia es un tema ampliamente debatido. el objetivo del presente trabajo es mostrar un procedimiento para el diagnóstico de los estudiantes cuyo conocimiento es insuficiente para aprobar. Se han diseñado dos metodologías para tal fin. Las reglas lógicas para llegar a una recomendación o decisión fueron elaboradas sobre criterios subjetivos de docentes expertos. el tratamiento de estas reglas fue planteado con lógica difusa y con lógica booleana. Otro objetivo del trabajo es hacer una comparación de ambas metodologías para poder decir cual de ellas es más conveniente usar. se ha comprobado que el modelo difuso ofrece mejores resultados que el modelo booleano.

El razonamiento de los docentes se describió en conjuntos y reglas difusas. el trabajo se basó en tres variables etiquetadas verbalmente: nota de examen parcial, nota de prácticas o dinámicas grupales y promedio

Palabras claves: Lógica Difusa, evaluación de alumnos, variables lingüísticas, lógica booleana, sistemas de inferencia, reglas de inferencia.

ABSTRACT

One of the most complex tasks of a teacher is the accurate assessment of progress their students. The individual criteria in assessing the level of competition is a widely debated topic. The aim of this paper is to show a procedure for the diagnosis of students whose knowledge is insufficient to pass. It Two methodologies have been designed for this purpose. The logical rules to arrive at a recommendation or decisions were made on subjective criteria of expert teachers. The Treatment of these rules was raised with fuzzy logic and Boolean logic. Other study aims to compare the two methodologies to say which one is more convenient to use. It has been found that the fuzzy model provides better results than the Boolean model.

The reasoning of teachers described in sets and fuzzy rules. The work is based on three variables labeled verbally: note the partial examination, practical note or group dynamics and average.

Keywords: Fuzzy Logic, student assessment, linguistic variables, logic Boolean inference systems, rules of inference.

I. INTRODUCCIÓN

el proceso de formación profesional universitaria, ilustrado en la Figura 1, muestra que debe desarrollarse los aspectos: Académico, investigación y responsabilidad social. sin embargo, en el país prácticamente sólo se aborda el académico, y con una serie de deficiencias, lo cual se refleja en el alto índice de *desaprobación*, que en muchos casos lleva a la *deserción*.

La debilidad a tratar en esta ocasión es el alto índice de desaprobación, que en ciertas materias supera el 80% de los participantes.

II. SISTEMAS DE LÓGICA DIFUSA O BORROSA

Los Sistemas de Lógica Difusa (SLD) [1], junto con las redes neuronales artificiales (RNA), constituyen las principales herramientas de modelado de las técnicas de Soft Computing. En la lógica difusa se define brevemente una "herramienta para transformar el conocimiento estructurado humano a algoritmos ejecutables". el campo de la lógica difusa es muy amplio y cubre muchos conceptos matemáticos y lógicos.

El concepto de la lógica difusa (Fuzzy Logic) se aplica con diferentes sentidos. Así, la lógica difusa puede considerarse como un sistema lógico que persigue modelar el razonamiento humano cuando éste es aproximado en lugar de exacto. sin embargo, en un sentido más amplio, la lógica difusa puede entenderse como un conjunto difuso de teoría de clases cuyos límites no están definidos o son difusos o borrosos. Los métodos basados en la lógica difusa pueden utilizarse para diseñar sistemas inteligentes sobre la base del conocimiento expresado en un lenguaje común. en la actualidad casi no existen áreas de actividad humana en las que la lógica difusa no haya sido empleada.

La principal razón de que la lógica difusa sea tan versátil es que permite tanto el procesado de información expresada de forma *numérica* como de información expresada de manera *simbólica*. Frecuentemente, los sistemas diseñados y desarrollados empleando lógica difusa han resultado ser más eficientes que aquellos basados en aproximaciones convencionales.

en el modelado mediante lógica difusa se sigue un modelo conceptual de caja blanca, donde se tiene cierto conocimiento sobre la solución, siendo el conocimiento mismo lo que se modela. La lógica difusa puede considerarse una herramienta muy eficaz para transformar el conocimiento estructurado humano en útiles algoritmos. Al igual que en el proceso de razonamiento e inferencia humanos, cada afirmación, cada medida u observación, posee cierto grado de certeza. este grado se expresa a través de funciones de pertenencia que miden el grado de pertenencia de ciertas entradas a los subconjuntos difusos dados.

III. PROBLEMÁTICA

La evaluación tradicional de los estudiantes los mecaniza a solamente prepararse para los parciales, prestan poca importancia al dictado de clases por diversas razones: estudian otra carrera, trabajan, participan en algún taller estudiantil. Toda su energía se dirige hacia estas actividades y olvidan los objetivos iniciales con los que ingresaron a la universidad. es importante valorar el trabajo académico diario que se desarrolla en cada sesión, que los estudiantes reflexionen sobre su rol como actores que generan su propio conocimiento. Establecer una tarea específica para cada sesión implica desarrollar un plan de trabajo: el establecimiento de un objetivo final que puede ser considerado como producto de la sesión y establecer las etapas que se desa-



Figura 1. Proceso de formación profesional universitaria

rollaran para el logro de ese objetivo. Las actividades a desarrollar deben ser significativas y apropiadas para desarrollar la interacción y fortalecer el aprendizaje colaborativo. Como toda actividad pedagógica debe ser evaluada, esta evaluación permitirá al estudiante reflexionar sobre su propio aprendizaje, desarrollar una meta cognición y ser consciente sobre los procesos cognitivos que desarrolló para lograr el objetivo final.

Es costumbre clasificar a los estudiantes de acuerdo a un puntaje promedio de sus capacidades cognitivas, y todavía persiste la actitud de orientar a los que tienen bajo promedio hacia saberes técnico-operacionales porque se considera inútil intentar que aprendan ciencias y/o humanidades. el surgimiento de nuevos paradigmas educativos como el de inteligencias múltiples (Gardner, 1983) [2] que propone que la inteligencia no es inmutable ni unitaria, sino que es diversa de acuerdo a cada persona permite explicitar que estas inteligencias son flexibles, distintas e independientes. Esto implica que un científico es tan inteligente como un artista o un deportista, simplemente sus inteligencias pertenecen a diferentes campos.

Vigotsky [3] plantea que el desarrollo no es un proceso estático, sino que supone la idea de potencialidad lo que otorga importancia a la intervención del docente como mediador entre la cultura y el individuo. el mero contacto con los objetos no garantiza el aprendizaje.

La Zona de Desarrollo potencial es la distancia entre el nivel de resolución de una tarea que una persona

puede alcanzar actuando independientemente y el nivel que puede alcanzar con la ayuda de un compañero experto en esa tarea. Vigotsky denomina Zona de Desarrollo Próximo (ZDP) al espacio en que gracias a la interacción y la ayuda de otros, una persona puede trabajar y resolver un problema o realizar una tarea con un nivel superior al que no sería capaz de resolver individualmente. en cada alumno y para cada contenido de aprendizaje existe una zona que esta próxima a desarrollarse y otra que en ese momento está fuera de su alcance. En la ZDP es en donde deben situarse los procesos de enseñanza y de aprendizaje. En la ZDP es donde se desencadena el proceso de construcción de conocimiento del alumno y se avanza en el desarrollo. No tendría sentido intervenir en lo que los alumnos pueden hacer solos.

IV. MODELO DE EVALUACION DE RENDIMIENTO

La propuesta consiste en definir un modelo booleano y otro difuso (basado en [5]), usando las mismas reglas de inferencia y los mismos indicadores lingüísticos. Ambos modelos permiten establecer una sugerencia o decisión (**variable de salida**) dirigida a los alumnos con nota desaproboratoria en el curso (Promedio < 11). Ambos modelos necesitan como datos de entrada: **indicadores (variables de entrada)** y **reglas de inferencia**.

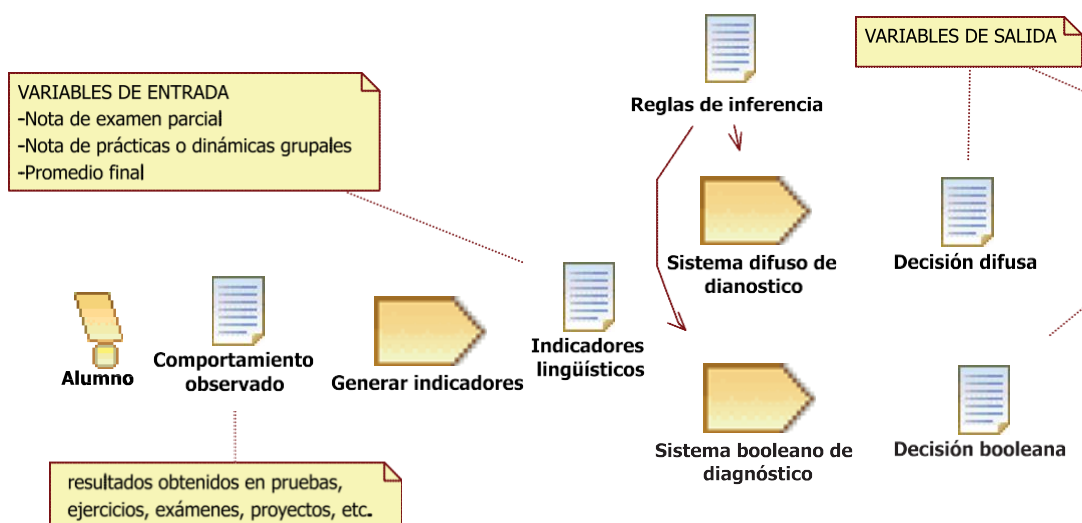


Figura 2. Comparación de los resultados obtenidos por los 2 modelos propuestos

Variables de entrada

La propuesta contempla realizar un análisis de rendimiento a la mitad del semestre, toda vez que aquí se puede disponer del 50% del puntaje (que va de 0 a 20), el cual usualmente se obtiene a partir de las notas de: Examen parcial y dinámicas (prácticas laboratorios en equipo), por medio de una fórmula como esta:

$$\text{Promedio} = 0.6 * \text{Examen Parcial} + 0.4 * \text{Dinámicas} \dots (1)$$

es por eso que ambos modelos usan los 3 indicadores lingüísticos siguientes como entradas:

- **Nota de examen parcial (EP).** esta nota varía entre 0 y 20. Además es múltiplo de 0.25 (0, 0.25, 0.5, 1, ..., 19, 19.25, 19.5, 19.75, 20)
- **Nota de dinámicas (DI).** es múltiplo de 0.25 y varía entre los siguientes valores:

$$\max(-0.6 * EP / 0.4, 0) \leq DI \leq \min((11 - 0.6 * EP) / 0.4 - 0.001, 20)$$

Simplificando se obtiene:

$$0 \leq DI \leq \min((11 - 0.6 * EP) / 0.4 - 0.001, 20) \dots (2)$$

- **Promedio (PR).** Se obtiene con la fórmula (1). Se considera alumnos de bajo rendimiento, a aquellos cuyo promedio esta en el rango:

$$0 \leq PR < 11 \dots (3)$$

Cada indicador puede tomar 3 valores lingüísticos: Pésimo (P o 1), Muy Malo (MM o 2) y Malo (M o 3).

Variable de salida

- **Decisión (D).** Toma 3 valores, los cuales recomiendan que el alumno debe: Reforzar algunos aspectos (R), recibir tutoría (T) o retirarse del curso (E).

Reglas de inferencia

Son once (11) reglas establecidas con el juicio experto de dos (2) docentes (Tabla 1).

El modelo booleano

en el modelo booleano los indicadores toman valores de acuerdo al rango en que se encuentra su valor numérico. A continuación se resume la relación entre los valores numéricos y lingüísticos de los indicadores.

Tabla1. Reglas definidas por docentes expertos.

PR	EP	DI	D
1	0	0	excluir
2	1	1	excluir
2	1	2	excluir
2	1	3	Tutoría
2	2	1	Tutoría
2	2	2	Tutoría
2	2	3	Tutoría
2	3	1	Recom
2	3	2	Recom
2	3	3	Recom
3	0	0	Recom

Tabla 2. Rangos para hallar los valores de los indicadores

	1	2	3
	Pésimo (<)	M.Malo	Malo (>=)
EP	5	(5, 7)	7.00
DI	4	(4,8)	8.00
PR	3	(3, 9)	9.00

La variable de salida: Decisión (D), sólo toma 3 valores discretos:

Tabla 3. Valores de la variable de salida

	R	T	e
	Reforzar	Dar tutoría	Retirarlo
D	1	2	3

es decir, si D vale 3, el estudiante será invitado a retirarse de la materia; si tiene 1, se le dará una serie de pautas para reforzar su aprendizaje; y si D vale 2, el estudiante será asistido por un tutor.

El modelo difuso

Para llevar a cabo el análisis del bajo rendimiento (promedio menor a 11) se hace uso de lógica difusa, asignando una función de pertenencia a cada una de las 3 variables lingüísticas de entrada y la variable de salida:

Tabla 4. Funciones de pertenencia de las diferentes variables lingüísticas.

Variables	Tipo	Valores lingüísticos		
Promedio final (PR)	entrada	Pésimo $e(0, 1)$	Muy malo $e(6, 1)$	Malo $e(10.75, 1)$
Nota de examen parcial (EP)	entrada	Pésimo $e(0, 1.67)$	Muy malo $e(6, 0.3)$	Malo $e(10.75, 1.3)$
Nota de dinámicas (DI)	entrada	Pésimo $e(0, 1.3)$	Muy malo $e(6, 0.69)$	Malo $e(10.75, 1)$
Decisión (D)	salida	Necesita Reforzar $t(0, 0.5, 1)$	Necesita tutoría $t(1, 1.5, 2)$	Debe ser retirado $t(2, 2.5, 3)$

Donde:

$$e(m, k) = e^{-\frac{(x-m)^2}{2k^2}}$$

$$t(a, m, b) =$$

$$\begin{cases} 0, & x \leq a \\ \frac{x-a}{m-a}, & a < x \leq m \\ \frac{b-x}{b-m}, & m < x < b \\ 0, & x \geq b \end{cases}$$

V. RESULTADOS

el sistema fue implementado en una hoja de cálculo en excel. se usó el método de inferencia difusa de Mamdani [4].

Las pruebas (quince) se realizaron con datos aleatorios

de grupos de 20 alumnos, obteniéndose el resultado consignado en la Tabla 5. en cada prueba se trató un grupo diferente, haciendo el comparativo de ambos modelos. Para la defusificación de la variable de salida, se empleó el método del centroide [4].

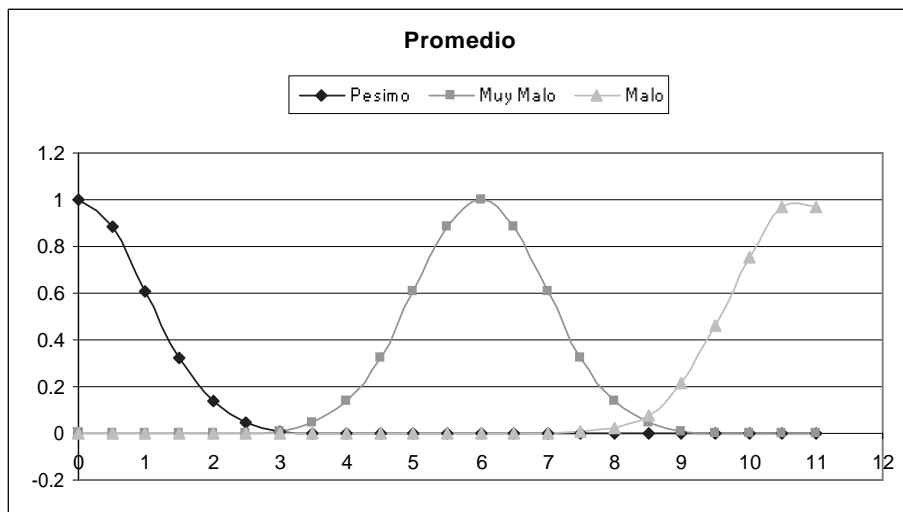


Figura3. Variable promedio final y sus 3 funciones de pertenencia

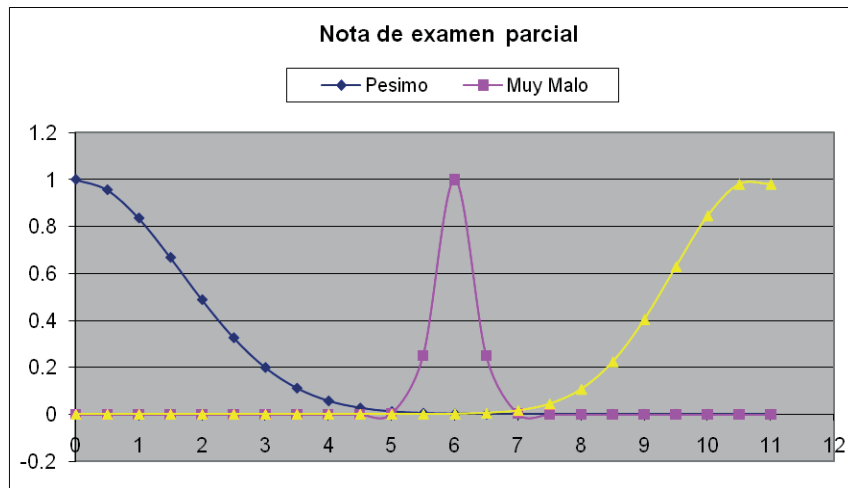


Figura 6. La variable "decisión" y sus 3 funciones de pertenencia

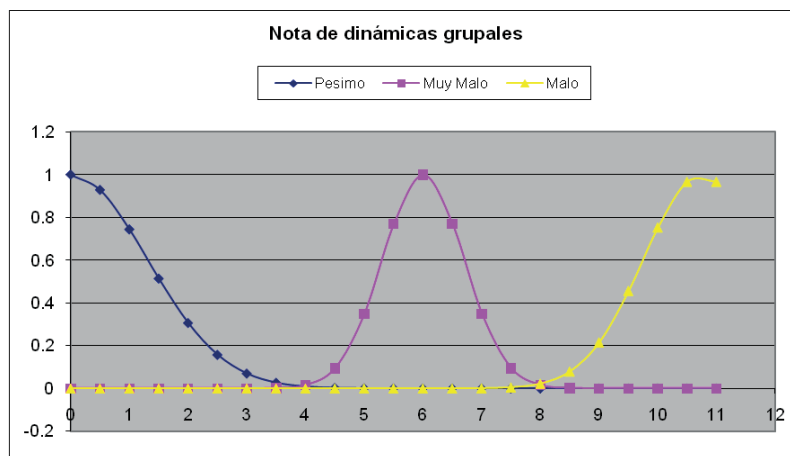


Figura 4. La variable "examen parcial" y sus 3 funciones de pertenencia

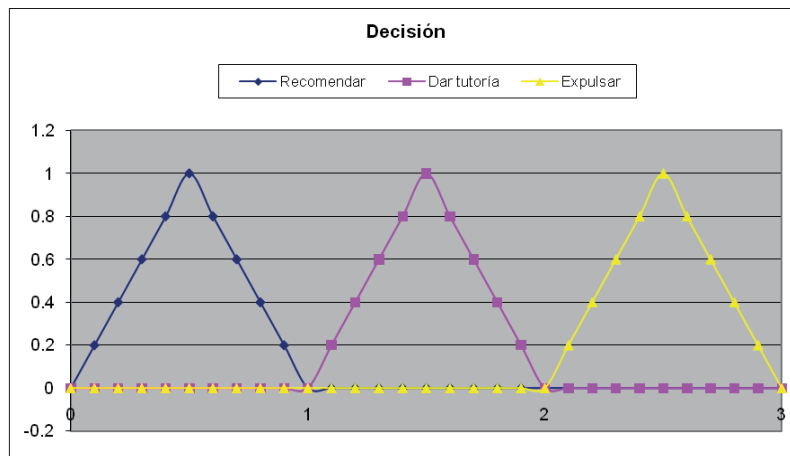


Figura 5. La variable "dinámicas" y sus 3 funciones de pertenencia

Tabla 5. Resultados de la evaluación para ambos modelos.

Nº de ejecución	Nº de Coincidencias	Nº de veces que Difuso fue más flexible que Booleano	Nº de veces que Difuso fue menos flexible que Booleano
1	16	4	0
2	16	3	1
3	18	2	0
4	16	3	1
5	13	6	1
6	15	5	0
7	15	3	2
8	15	5	0
9	18	2	0
10	16	3	1
11	13	7	0
12	18	1	1
13	16	3	1
14	14	6	0
15	13	5	2
suma	232	58	10
Promedio	15.467	3.867	0.667
% promedio	77.33%	19.33%	3.33%

VI. CONCLUSIONES y RECOMENDACIONES

en promedio, los modelos difuso y el booleano propuestos presentan una coincidencia de un 77.33% en los resultados. Además el modelo difuso es 19.33% más flexible que el modelo booleano, y esto se da porque existen casos donde por ejemplo, si el modelo booleano considera que un alumno debe ser retirado, el modelo difuso considera que el alumno no debe ser retirado, sino que se le debe dar tutoría. También se observa que el modelo booleano fue más flexible que el difuso en un 3.33% de alumnos evaluados.

entonces se recomienda usar un modelo difuso debido a que es más flexible que el modelo booleano y por lo tanto permite dar mayor oportunidad a los alumnos para que mejoren su rendimiento antes de sugerirles que se retiren del curso.

el sistema se puede mejorar considerando la nota de un examen de entrada al inicio del semestre. Luego se puede crear un nuevo indicador que considere este dato.

en un trabajo a futuro se puede hacer que el sistema provea las acciones correctivas a partir de una base de conocimientos.

Algunas ventajas son:

- Poder ofrecer al estudiante a mitad de semestre una retroalimentación sobre su desempeño el docente puede comparar su evaluación tradicional con las recomendaciones del sistema.
- se puede agrupar a los alumnos para darles una enseñanza personalizada.
- Ahorro en costos debido a evaluaciones adicionales con personal.

se puede considerar como innovación:

- Contar con una herramienta basada en el modelo de lógica difusa que apoye a la evaluación tradicional, permitiendo identificar a los alumnos de bajo rendimiento, con posibilidades de aprobar, de tal manera que se apliquen las acciones correctivas necesarias.

VII. REFERENCIAS

1. [soledad, 2010] soledad Valero Cubas, Tesis doctoral "Arquitectura de Búsqueda Basada en Técnicas soft Computing para la Resolución de Problemas Combinatorios en Diferentes Dominios de Aplicación" (2010).
2. [Gardner, 1983] Irving B. Weine: Handbook of Psychology, educational psychology, Artículo "Multiple Intelligence Perspective on Talent and Giftedness Gardner (1983) postulates the existence of eight ... Gardner's criteria (1983)..."
3. [Lev, 1995] Lev Vygotsky, Pensamiento y lenguaje (1995).
4. http://www.dma.fi.upm.es/java/fuzzy/fuzzyinf/main_en.htm; sanjay Krishnankutty Alonso; *eMathTeacher: Mamdani's Fuzzy Inference Method*, accedido el 16 de octubre del 2011
5. Constanza Huapaya | Francisco Lizarralde | Graciela Arona | Jorge Vivas; Sesgos en la evaluación del estudiante con bajo rendimiento. Un sistema de mejoramiento diagnóstico basado en Lógica Difusa, 2010