

Sesgos en la evaluación del estudiante con bajo rendimiento. Un sistema de mejoramiento diagnóstico basado en Lógica Difusa

Constanza Huapaya¹ | Francisco Lizarralde² | Graciela Arona³ | Jorge Vivas⁴

Resumen

Una de las tareas más complejas de un docente es la evaluación fidedigna del progreso de sus estudiantes. Los sesgos personales en la evaluación del nivel de competencia es un fenómeno ampliamente debatido. El objetivo del presente trabajo es poner a consideración de los docentes un procedimiento para el diagnóstico de los estudiantes cuyo conocimiento es insuficiente para aprobar. Se ha diseñado una metodología para mejorar los procedimientos estadísticos tradicionales, la cual se aplicó sobre alumnos de Ingeniería. La recomendación fue elaborada sobre criterios subjetivos de docentes expertos en el área de las asignaturas tratadas y el tratamiento de estas recomendaciones se hizo con lógica difusa (LD). El razonamiento de los docentes se describió en conjuntos y reglas difusas. El trabajo se basó en cuatro variables etiquetadas verbalmente: *progresión decreciente de las notas, pobreza en la resolución de problemas en equipo, nota con respecto a la media y nota de examen*

Summary

One of the teacher's most complex tasks is the reliable evaluation of the students' progress. Personal bias in the students' competence evaluation is a highly debated phenomenon. The objective of this article is to offer the teachers a procedure for the diagnosis of students whose knowledge is insufficient to pass. A methodology has been designed to improve traditional statistical procedures. Such methodology was applied on Engineering students. The recommendation was elaborated on the subjective criteria of expert teachers in the subjects of the Engineering course of studies. The treatment of such recommendation was carried out with fuzzy logic (FL). The reasoning of the teachers was described in sets and fuzzy rules. This work was based on four verbally labeled variables: *decreasing progression of grades, poor group problem-solving, grade with respect to the average and grade of the mid-term exam with respect to the passing grade*. The output verbal

parcial con respecto a la aprobación. La variable verbal de salida *recomendación* sugiere a los docentes si el nivel de logro de un estudiante particular puede considerarse como poco insuficiente o fuertemente insuficiente. Se ha comprobado el buen desempeño del sistema con respecto a las decisiones de los docentes.

Palabras clave: sesgos en la evaluación - lógica difusa - diagnóstico del estudiante basado en computadora.

variable *recommendation* suggests to the teachers whether a given student's achievement level can be considered barely insufficient or strongly insufficient. The good performance of the system has been proven in comparison to the teachers' decisions.

Key words: evaluation biases - fuzzy logic - student diagnosis based on computer.

Fecha de recepción: 09/11/09
Fecha de aceptación: 10/12/09

Introducción

El proceso de inferencia de las modificaciones internas del estudiante a partir de su comportamiento observado se llama *diagnóstico cognitivo del estudiante*. En este proceso complejo participan docentes y estudiantes, y da lugar a planteos sobre el estado cognitivo individual de cada alumno y del grupo de la clase. El diagnóstico cognitivo se construye sobre la interpretación de los resultados obtenidos en pruebas, ejercicios, exámenes, proyectos, etc. Como resultado del proceso de inferencia se construye la caracterización del estudiante sobre su estado cognitivo. El diagnóstico involucra la consideración de los siguientes aspectos:

1. Registrar el comportamiento observable del estudiante en términos de medidas específicas.
2. Inferir las características internas del conocimiento del estudiante en base a la información recogida, y además, que sean sustantivas para el aprendizaje.
3. Desarrollar un método para extraer esta información a través del monitoreo y seguimiento del estudiante.

El punto central de los desarrollos para solucionar la problemática presentada es investigar un método que pueda analizar efectivamente las mediciones de la actividad de los estudiantes y hacer estimaciones sobre sus características internas.

Los sistemas computacionales para el diagnóstico del estudiante utilizan un *modelo del estudiante* (estructura que manipula las características individuales de los alumnos) (Sison, 1988) (Hartley, 1995) para guiar el comportamiento del sistema sobre los consejos para el evaluador y la orientación del estudiante en su aprendizaje.

Existen razones importantes para buscar nuevos métodos de evaluación del conocimiento del alumno, a saber:

- *Puntuación imprecisa*: los docentes asignan, al trabajo de los estudiantes, una nota dentro de un esquema pre-determinado. Por sesgos personales o contextuales, esta nota puede diferir de un evaluador a otro, dependiendo de la experiencia y sensibilidad de cada uno de ellos. En general, la nota que se le asigna a un estudiante es una aproximación.

Otra forma de evaluar sus aptitudes, competencias y habilidades es calificarlos con términos verbales como excelente, regular, etc. Se considera que aptitud, competencia y habilidad son conceptos imprecisos, como también lo son los términos excelente y regular. Puede presuponerse que etiquetas verbales de carácter difuso puedan ser capturadas mejor usando técnicas difusas.

- *Control de la calificación*: usualmente los docentes evalúan colocando una nota al logro de su estudiante. Casi nunca se usa un método de evalua-

ción alternativo para verificar esa nota final del curso, con la cual se da por aprobado o desaprobado. Un nuevo método eficiente de evaluación puede resultar de gran utilidad para confirmar o refutar la decisión tomada con métodos tradicionales.

- *Uso de la lengua natural en la evaluación académica:* las pruebas pueden ser evaluadas usando términos lingüísticos y con ellos lograr mayor flexibilidad en el juicio sobre el logro del estudiante.

De aquí se infiere que uno de los principales obstáculos en el proceso de diagnóstico por medio de un sistema computacional es la *incertidumbre*. Esta incertidumbre se origina en la naturaleza abstracta 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 al sistema. Además, la adquisición de esta información produce datos sin elaborar, esto es, son datos que pierden precisión debido a la cuantificación posterior. La información registrada por el sistema incorpora estas inexactitudes. Luego, los sistemas deben construir un método de decisión basado en datos imprecisos. Una de las metodologías que tratan esta problemática es la lógica difusa.

Breve introducción a la lógica difusa

La lógica difusa (Zadeh, 1965, 1988, 1994) puede verse desde dos puntos de vista: el primero (más restringido) propone un sistema lógico cuyo objetivo es formalizar el razonamiento aproximado. Si bien la lógica difusa es una lógica multivaluada, no lo es en el sentido tradicional de los sistemas lógicos multivaluados. Los conceptos centrales son variables lingüísticas, formas canónicas, reglas lógicas difusas, cuantificadores difusos y modos de razonamiento tales como razonamiento silogístico y razonamiento interpolatorio.

El segundo enfoque, más amplio, es casi sinónimo de teoría de los conjuntos difusos. Esta es una teoría de clases con bordes blandos y es más amplia que la lógica difusa en su sentido estricto; contiene, además, aritmética difusa, programación matemática difusa, teoría de grafos difusos, etc.

Los conjuntos difusos utilizan funciones, definidas en un universo de objetos (X) en el intervalo $[0, 1]$. Así la función $m: X \rightarrow [0, 1]$ indica el grado de pertenencia de cada elemento al conjunto difuso. Así, $x \in A$ se indica con el par $(x, 0)$ y $x \in A$ con el par $(x, 1)$. Nótese que si solo se establecen estos dos casos se obtiene la teoría tradicional de conjuntos, pero con la lógica difusa es posible expresar el grado de pertenencia de x al conjunto A con todos los valores intermedios entre 0 y 1. La función m se la conoce como *función de pertenencia*.

Es importante notar que la especificación de la función de pertenencia es subjetiva, es decir, que los valores pueden diferir de una persona a otra; esta subjetividad proviene de las diferencias individuales en la percepción o expresión de conceptos abstractos. Este es el punto que diferencia la probabilidad (medida matemática de fenómenos aleatorios) de la posibilidad (subjetividad, i.e. un continuo psicológico).

La lógica difusa en la evaluación del conocimiento del estudiante

La teoría de la lógica difusa busca proveer una notación matemática que permita expresar la incertidumbre asociada con los procesos cognitivos humanos. La lógica difusa es particularmente relevante para tratar fenómenos con incertidumbre y que involucran un cambio suave y progresivo. El razonamiento seguido durante un proceso de evaluación se basa en un continuo cuyos extremos pueden denominarse, por ejemplo, «definitivamente competente» y «definitivamente incompetente» si se desea mensurar el nivel de pericia alcanzado.

Un alumno puede posicionarse, según algún criterio seguido de evaluación, en algún punto del continuo. El evaluador puede proponer distintas escalas para tomar decisiones, según el aspecto que desea medir. Por ejemplo, si se desea saber la «velocidad en la resolución de problemas» de un estudiante, puede pensarse en una escala cuyos extremos sean «rápido» y «lento».

La evaluación humana del logro involucra, generalmente, etiquetas verbales como excelente, bueno o insatisfactorio, las cuales están fuertemente imbuidas de subjetividad (estas etiquetas verbales se denominan variables lingüísticas en la lógica difusa). Las etiquetas verbales surgen a partir de diversos componentes de la evaluación como cuestionarios, planillas especializadas, etc. El evaluador propone estas etiquetas con las cuales el sistema llega al diagnóstico. Las funciones de pertenencia $m(x)$ sobre el logro de los estudiantes pueden definirse, por ejemplo en el aprendizaje del lenguaje Francés, como se ve en la figura 1. Aquí se han definido 5 etiquetas verbales (insatisfactorio, satisfactorio, bueno, muy bueno, excelente) sobre diversos intervalos de notas. Por ejemplo, notas en el intervalo (0,4] tienen un valor de pertenencia a insatisfactorio (línea punteada) de 1, mientras las notas en el intervalo (4,5) sus valores de pertenencia son decrecientes.

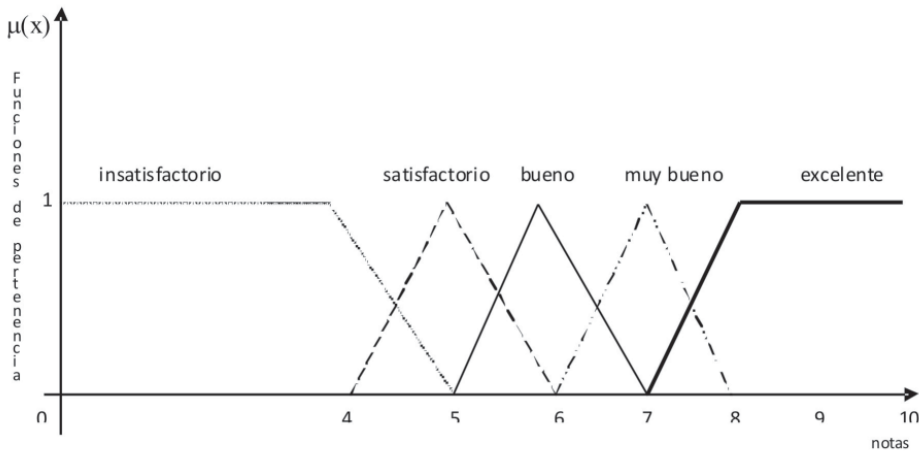


Figura 1: cinco funciones de pertenencia $m(x)$ para el logro en Francés

La subjetividad se aprecia cuando se trata de definir el logro de los estudiantes. El rendimiento en el aprendizaje del Francés puede ser apreciado de diferente modo que en el aprendizaje de otra asignatura. Por ejemplo, la etiqueta verbal excelente (traducido en una función $m(x)$ con línea resaltada para notas superiores a 7) en Francés puede conformar a la función de la figura 1, pero el excelente en otra área del conocimiento puede asignarse a notas por arriba o por debajo de 7.

Los modelos difusos pueden ‘imitar’ el modo de adquirir conocimiento por parte de los docentes cuando evalúan las características de aprendizaje de sus alumnos, tales como actitudes, nivel de conocimiento, competencias, motivación y estilo de aprendizaje. En particular, la lógica difusa (Dubois, 1996) y (Zadeh, 1975) es usada para acceder a un modo de razonamiento cualitativo cercano al

razonamiento humano cuando toma decisiones, principalmente el manejo de la *imprecisión* y *vaguedad*. La forma de lograrlo es a través de la combinación de hechos difusos y relaciones difusas.

En la literatura del área, existen varios desarrollos para la evaluación del estudiante basados en lógica difusa. Estos estudios pueden categorizarse en cuatro grupos:

1. **Similitud difusa** (Bitwas, 1995): calcula la similitud entre los términos lingüísticos *excelente*, *muy bueno*, *bueno*, *satisfactorio* e *insatisfactorio* y las notas difusas asignadas a cada pregunta de las pruebas de los alumnos. Esta metodología provee una evaluación muy detallada.
2. **Valores esperados difusos** (Law, 1996): creación de un modelo estructural de notas usando el enfo-

que difuso. Se calculan los valores esperados para cada nota u otra modalidad de evaluación. Luego se calcula la nueva calificación usando un método para obtener un valor nítido a partir de valores difusos (o defuzzificación). Finalmente se calcula la nota final como un promedio ponderado con diversos pesos. Esta metodología involucra un costoso proceso computacional.

3. Uso de funciones de pertenencia en la teoría estadística

(Fourali, 1994), (Chen y Lee, 1999), (Weon y Kim, 2001): la característica común en estos desarrollos es el uso de varios criterios de evaluación para el logro de un estudiante (p.e. un portfolio). Estos métodos llevan las evaluaciones a valores difusos y luego calculan las notas finales con algún proceso estadístico, como promedios ponderados. Son métodos con cálculos simples.

4. Basado en reglas simples difusas

(Shimizu y Yamashita, 2000): se usan reglas difusas cuando la evaluación se basa en la experticia y sensibilidad de docentes experimentados. Por ejemplo, los autores mencionados aplicaron el razonamiento difuso en la evaluación de la caligrafía. Este enfoque usa inferencia difusa muy simple, basada en sólo 9 reglas de producción difusas.

La evaluación del logro académico, en general, se basa actualmente en técnicas aritmético-estadísticas y son ampliamente aceptadas por la comunidad educativa. Las técnicas difusas aparecieron a mediados de la década de los '90 con el objetivo de ser una alternativa en aquellos casos donde intervienen diferentes evaluadores con criterios distintos, como lo muestran Fourali y Weon. Actualmente, los enfoques se amplían hacia la captura del razonamiento seguido por los docentes cuando evalúan a sus alumnos. Los métodos presentados de Fourali y Chen y Lee son bastantes simples y no usan mecanismos de inferencia difusa. Asimismo, los métodos de Law y Biswas presentan cálculos más complejos, pero tampoco poseen inferencia difusa. Shimizu y Yamashita han usado inferencia simple en áreas del conocimiento que necesitan fuertemente los criterios subjetivos del evaluador, como las artes.

Una inferencia muy simple puede construirse en base a datos sobre cierta cantidad de pruebas. Para implementar la inferencia se usan funciones predictivas. Por ejemplo, una función muy simple puede calcular el promedio de todos los alumnos en cada una de las pruebas. Con esta función se pueden establecer dos grupos: el de los alumnos que *aprueban el curso* (los que sobrepasan los valores promedios de todas las pruebas) y el de los alumnos de *pobre desempeño* (los que no sobrepasan ningún valor promedio y necesitarán apoyo en diversos niveles para terminar el curso). Hay otros subconjuntos de estudiantes. Los resultados de este tipo de categorización forman las **descripciones difusas** de los alumnos.

A continuación se presenta un modelo para el diagnóstico del estudiante basado en notas registradas de una asignatura, usando inferencia difusa.

El modelo propuesto

El objetivo del modelo de diagnóstico desarrollado es medir el nivel de conocimiento de los alumnos que no aprueban una asignatura a fin de implementar, posteriormente, acciones remediales apropiadas ajustadas al logro individual alcanzado. Para alcanzar tal objetivo contempla las siguientes características:

1. Usar las evaluaciones originales (notas, conceptos, etc.) como entradas y transformarlas en etiquetas verbales en lugar de pedir al evaluador la asignación de valores difusos.
2. Usar mecanismos de inferencia basados en la opinión de docentes expertos además de manipular números difusos.
3. Poder usar diferentes niveles de evaluación sin necesidad de desarrollar otro método.

Las fuentes de evidencia mensurable que usa el modelo comprenden pruebas individuales escritas sobre conceptos teóricos, pruebas individuales de resolución de problemas en computadora y un proyecto integrador, en equipo, de resolución de problemas en computadora.

Los resultados de estas pruebas se registran a mediados y al final de un ciclo académico. Por ejemplo, para el curso de

Análisis Numérico para Ingeniería se guardan 6 atributos (2 pruebas sobre conceptos teóricos, 2 pruebas individuales de resolución de problemas, 1 proyecto colectivo y 1 nota final).

En base a los resultados de las pruebas se crean las descripciones difusas de los estudiantes usando conjuntos difusos, con las correspondientes funciones de pertenencia. Para su creación se recurre a la opinión del docente-evaluador sobre el nivel considerado. Por ejemplo la respuesta a la pregunta: ¿si un estudiante ha respondido correctamente el 48 % de las pruebas, cuál considera Ud. qué es el nivel de suficiencia alcanzado? Las respuestas sirven para construir el conjunto difuso y la función $m(x)$.

Si el evaluador docente desea brindar ayuda sobre un tópico, conviene recurrir a los conjuntos difusos que representan los niveles «insuficiencia» caracterizando a los alumnos que todavía no han llegado al nivel mínimo exigido para aprobar. Por razones prácticas la información solicitada puede ser discreta adquiriendo para las notas los siguientes 11 valores de $x = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$ mientras que para la función de pertenencia se obtienen los siguientes valores para $m(x) = \{1, 0.9, 0.8, 0.6, 0.3, 0.15, 0, 0, 0, 0, 0\}$. En la figura 2 se aprecia los niveles de «insuficiencia» con una representación continua.

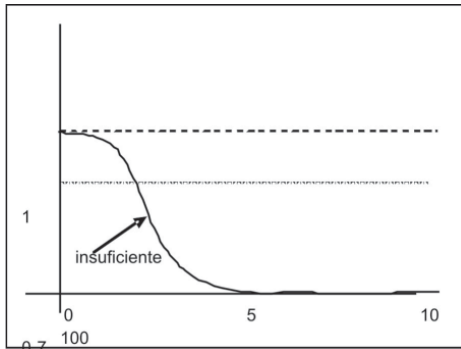


Figura 2: conjunto difuso para insuficiente

En este ejemplo, el docente puede decidir que si el grado de 'insuficiencia es mayor que 0.7' se debe brindar ayuda inmediata a ese alumno.

Las descripciones difusas de los estudiantes dan las bases para interpretar las reglas lingüísticas difusas. Para construir estas reglas se necesita definir las expresiones lingüísticas difusas. La estructura de estas expresiones es:

modificador lingüístico difuso	concepto lingüístico
--------------------------------	----------------------

Por ejemplo, el concepto de desempeño puede representar a alumnos que deberán ser atendidos especialmente por el docente. Para declarar el concepto *moderadamente insuficiente* use utiliza la siguiente expresión difusa:

moderadamente	insuficiente
---------------	--------------

Pero el concepto lingüístico *insuficiente* puede poseer diversos modificadores lingüísticos además de *moderadamente*. Por ejemplo, puede definirse una secuencia de constructores de este tipo: *extremadamente, fuertemente, bastante, moderadamente y poco*.

El modelo del diagnóstico del estudiante

Nuestro modelo predictivo permite establecer una *recomendación* difusa en el universo de los estudiantes que no alcanzaron el nivel de suficiencia necesario para aprobar la asignatura. A partir de las calificaciones de un ciclo lectivo (en nuestro caso 2007) se creó un modelo heurístico.

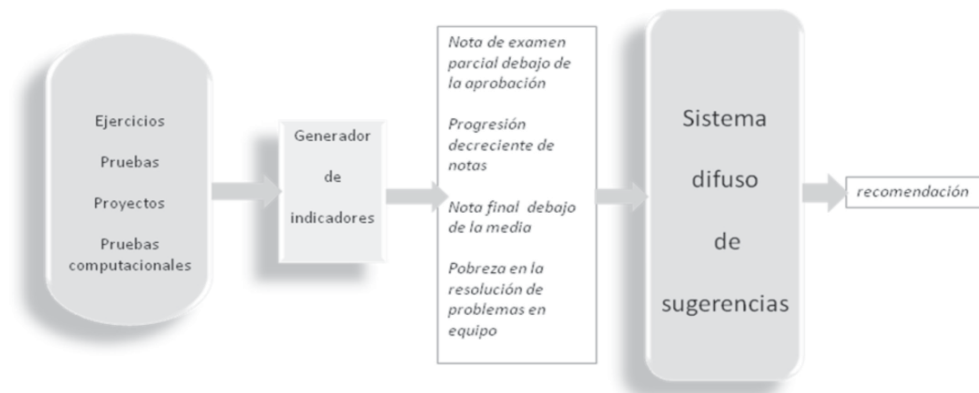


Figura 3: modelo difuso para emitir recomendaciones sobre alumnos con nivel insuficiente

Este modelo está compuesto por indicadores (que serán variables verbales) basados en criterios subjetivos de los docentes expertos del área. En nuestra aplicación se establecieron cuatro de ellas: *Progresión decreciente de notas*, *Nota de examen parcial debajo de la aprobación*, *Nota final con respecto a la media* y *Pobreza en la resolución de problemas en equipo* (ver figura 3). Estas cuatro variables verbales alimentan al Sistema de sugerencias compuesto por reglas lingüísticas difusas, las cuales finalmente calculan el valor de la *recomendación*. Las variables verbales de entrada que nuestra aplicación utiliza para el modelo predictivo son las siguientes:

- En *Progresión decreciente de notas* es calculada en base a las notas registradas de cada alumno considerando su cronología: se compara la nota del primer parcial con la del segundo y de este con la nota del proyecto.

Progresión decreciente de notas toma los siguientes valores: decrece poco (la diferencia entre las tres notas es decreciente y su diferencia es menor o igual a 1), decrece moderadamente (la diferencia entre las tres notas es decreciente y su diferencia es menor o igual a 2 y mayor que 1) y decrece fuertemente (la diferencia entre las tres notas es decreciente y su diferencia es mayor que 2).

- *Nota de examen parcial debajo de la aprobación*: la nota 5 es el límite inferior de la aprobación, por tal motivo los alumnos que se encuentran por debajo son considerados como posibles candidatos a la recuperación de contenidos. La variable toma los siguientes valores: un poco debajo de la aprobación, moderadamente debajo de la aprobación y fuertemente debajo de la aprobación.

- *Nota final debajo de la media:* se compara el rendimiento individual de cada alumno con respecto al desempeño promedio de todos los estudiantes del curso. Se toman los siguientes valores: un poco debajo de la media, moderadamente debajo de la media y fuertemente debajo de la media. Para Análisis Numérico la media del año 2007 es 4.75 sobre 106 alumnos.
- *Pobreza en la resolución de problemas en equipo:* en las asignaturas donde es importante resolución de problemas, como Análisis Numérico, se evalúa el logro en la adquisición de criterios de aplicación apropiada de los métodos numéricos. Los proyectos tienen como objetivo que el alumno integre conocimientos de métodos numéricos, programación de computadoras y problemas propios de las especialidades de ingeniería. Estos proyectos se desarrollan en equipos de 3 estudiantes. Los valores que toma esta variable considera a los alumnos que se encuentran debajo del límite de aprobación: poco pobre, moderadamente pobre, y fuertemente pobre. Para el caso particular de la asignatura Análisis Numérico la nota mínima de aprobación es 5.

A fin de visualizar una de las variables lingüísticas, veamos su representación en la figura 4, donde *Nota final debajo de la media* y los valores que toma representados como conjuntos difusos: un poco debajo de la media, moderadamente debajo de la media y fuertemente debajo de la media.

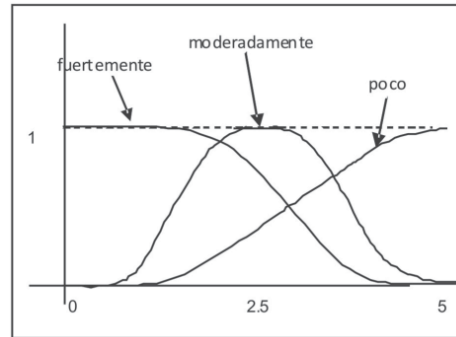


Figura 4: variable lingüística *Nota final debajo de la media*

La variable de salida, *recomendación* toma los siguientes valores: *poco insuficiente* (PI), *moderadamente insuficiente* (MI), *bastante insuficiente* (BI), *fuertemente insuficiente* (FI), y *extremadamente insuficiente* (EI).

Para alcanzar los valores de la *recomendación*, que depende de las cuatro variables lingüísticas de entrada recién vistas, se definen reglas de producción difusa como las siguientes:

1. SI la *Progresión decreciente de notas* decrece fuertemente ENTONCES *recomendación* es FI.
2. SI la *progresión decreciente de notas* decrece moderadamente Y *nota de examen parcial debajo de la aprobación* es moderadamente debajo de la aprobación Y *nota final debajo de la media* es un poco debajo de la media Y *Pobreza en la resolución de problemas en equipo* es moderadamente pobre ENTONCES *recomendación* es MI.

La regla número 2 posee un antecedente compuesto por valores de las cuatro variables verbales de entrada, mientras que la número 1 es un ejemplo simple donde sólo se ha considerado una variable en el antecedente. Nuestro sistema se compone de reglas como la número 2.

La cantidad de reglas que contienen el razonamiento de los docentes expertos para calcular el valor de la *recomendación* son 81 (cada una de las 4 variables de entrada toma 3 valores, por lo tanto son 3⁴ combinaciones de esos valores), las cuales figuran en la tabla 1.

Pobreza en la resolución de problemas en equipo	Progresión decreciente de notas	Notas de parcial con respecto a la aprobación	Nota final con respecto a la media	Recomendación
Poco pobre	Decrece poco	Poco por debajo	Poco por debajo	PI
Poco pobre	Decrece poco	Poco por debajo	Moderad. por debajo	PI
Poco pobre	Decrece poco	Poco por debajo	Fuertem. por debajo	MI
Poco pobre	Decrece poco	Moderad. por debajo	Poco por debajo	PI
Poco pobre	Decrece poco	Moderad. por debajo	Moderad. por debajo	PI
Poco pobre	Decrece poco	Moderad. por debajo	Fuertem. por debajo	MI
Poco pobre	Decrece poco	Fuertem. por debajo	Poco por debajo	MI
Poco pobre	Decrece poco	Fuertem. por debajo	Moderad. por debajo	MI
Poco pobre	Decrece poco	Fuertem. por debajo	Fuertem. por debajo	BI
Poco pobre	Decrece moderadam.	Poco por debajo	Poco por debajo	PI
Poco pobre	Decrece moderadam.	Poco por debajo	Moderad. por debajo	PI
Poco pobre	Decrece moderadam.	Poco por debajo	Fuertem. por debajo	MI
Poco pobre	Decrece moderadam.	Moderad. por debajo	Poco por debajo	MI
Poco pobre	Decrece moderadam.	Moderad. por debajo	Moderad. por debajo	MI
Poco pobre	Decrece moderadam.	Moderad. por debajo	Fuertem. por debajo	MI
Poco pobre	Decrece moderadam.	Fuertem. por debajo	Poco por debajo	MI
Poco pobre	Decrece moderadam.	Fuertem. por debajo	Moderad. por debajo	MI
Poco pobre	Decrece moderadam.	Fuertem. por debajo	Fuertem. por debajo	BI
Poco pobre	Decrece fuertem.	Poco por debajo	Poco por debajo	MI
Poco pobre	Decrece fuertem.	Poco por debajo	Moderad. por debajo	MI
Poco pobre	Decrece fuertem.	Poco por debajo	Fuertem. por debajo	BI
Poco pobre	Decrece fuertem.	Moderad. por debajo	Poco por debajo	BI
Poco pobre	Decrece fuertem.	Moderad. por debajo	Moderad. por debajo	MI
Poco pobre	Decrece fuertem.	Moderad. por debajo	Fuertem. por debajo	BI
Poco pobre	Decrece fuertem.	Fuertem. por debajo	Poco por debajo	MI
Poco pobre	Decrece fuertem.	Fuertem. por debajo	Moderad. por debajo	MI
Poco pobre	Decrece fuertem.	Fuertem. por debajo	Fuertem. por debajo	BI
Poco pobre	Decrece fuertem.	Fuertem. por debajo	Fuertem. por debajo	BI
Moderadamente pobre	Decrece poco	Poco por debajo	Poco por debajo	PI
Moderadamente pobre	Decrece poco	Poco por debajo	Moderad. por debajo	PI
Moderadamente pobre	Decrece poco	Poco por debajo	Fuertem. por debajo	MI
Moderadamente pobre	Decrece poco	Moderad. por debajo	Poco por debajo	PI
Moderadamente pobre	Decrece poco	Moderad. por debajo	Moderad. por debajo	MI
Moderadamente pobre	Decrece poco	Moderad. por debajo	Fuertem. por debajo	MI
Moderadamente pobre	Decrece poco	Fuertem. por debajo	Poco por debajo	MI
Moderadamente pobre	Decrece poco	Fuertem. por debajo	Moderad. por debajo	MI
Moderadamente pobre	Decrece poco	Fuertem. por debajo	Fuertem. por debajo	BI
Moderadamente pobre	Decrece moderadam.	Poco por debajo	Poco por debajo	MI
Moderadamente pobre	Decrece moderadam.	Poco por debajo	Moderad. por debajo	MI
Moderadamente pobre	Decrece moderadam.	Poco por debajo	Fuertem. por debajo	MI
Moderadamente pobre	Decrece moderadam.	Moderad. por debajo	Poco por debajo	MI
Moderadamente pobre	Decrece moderadam.	Moderad. por debajo	Moderad. por debajo	MI
Moderadamente pobre	Decrece moderadam.	Moderad. por debajo	Fuertem. por debajo	MI
Moderadamente pobre	Decrece moderadam.	Fuertem. por debajo	Poco por debajo	BI
Moderadamente pobre	Decrece moderadam.	Fuertem. por debajo	Moderad. por debajo	MI
Moderadamente pobre	Decrece moderadam.	Fuertem. por debajo	Fuertem. por debajo	BI
Moderadamente pobre	Decrece fuertem.	Poco por debajo	Poco por debajo	MI
Moderadamente pobre	Decrece fuertem.	Poco por debajo	Moderad. por debajo	MI
Moderadamente pobre	Decrece fuertem.	Poco por debajo	Fuertem. por debajo	BI
Moderadamente pobre	Decrece fuertem.	Fuertem. por debajo	Poco por debajo	BI
Moderadamente pobre	Decrece fuertem.	Fuertem. por debajo	Fuertem. por debajo	BI

Moderadamente pobre	Decrece fuertem.	Moderad. por debajo	Poco por debajo	MI
Moderadamente pobre	Decrece fuertem.	Moderad. por debajo	Moderad. por debajo	MI
Moderadamente pobre	Decrece fuertem.	Moderad. por debajo	Fuertem. por debajo	BI
Moderadamente pobre	Decrece fuertem.	Fuertem. por debajo	Poco por debajo	MI
Moderadamente pobre	Decrece fuertem.	Fuertem. por debajo	Moderad. por debajo	BI
Fuertemente pobre	Decrece poco	Poco por debajo	Moderad. por debajo	BI
Fuertemente pobre	Decrece poco	Poco por debajo	Fuertem. por debajo	BI
Fuertemente pobre	Decrece poco	Moderad. por debajo	Poco por debajo	BI
Fuertemente pobre	Decrece poco	Moderad. por debajo	Moderad. por debajo	MI
Fuertemente pobre	Decrece poco	Moderad. por debajo	Fuertem. por debajo	BI
Fuertemente pobre	Decrece poco	Fuertem. por debajo	Poco por debajo	MI
Fuertemente pobre	Decrece poco	Fuertem. por debajo	Moderad. por debajo	BI
Fuertemente pobre	Decrece poco	Fuertem. por debajo	Fuertem. por debajo	FI
Fuertemente pobre	Decrece moderadam.	Poco por debajo	Poco por debajo	MI
Fuertemente pobre	Decrece moderadam.	Poco por debajo	Moderad. por debajo	BI
Fuertemente pobre	Decrece moderadam.	Poco por debajo	Fuertem. por debajo	BI
Fuertemente pobre	Decrece moderadam.	Moderad. por debajo	Poco por debajo	FI
Fuertemente pobre	Decrece moderadam.	Moderad. por debajo	Moderad. por debajo	BI
Fuertemente pobre	Decrece moderadam.	Moderad. por debajo	Fuertem. por debajo	BI
Fuertemente pobre	Decrece moderadam.	Fuertem. por debajo	Poco por debajo	BI
Fuertemente pobre	Decrece moderadam.	Fuertem. por debajo	Moderad. por debajo	BI
Fuertemente pobre	Decrece moderadam.	Fuertem. por debajo	Fuertem. por debajo	FI
Fuertemente pobre	Decrece fuertem.	Poco por debajo	Poco por debajo	BI
Fuertemente pobre	Decrece fuertem.	Poco por debajo	Moderad. por debajo	FI
Fuertemente pobre	Decrece fuertem.	Poco por debajo	Fuertem. por debajo	BI
Fuertemente pobre	Decrece fuertem.	Moderad. por debajo	Poco por debajo	FI
Fuertemente pobre	Decrece fuertem.	Moderad. por debajo	Poco por debajo	FI
Fuertemente pobre	Decrece fuertem.	Moderad. por debajo	Moderad. por debajo	FI
Fuertemente pobre	Decrece fuertem.	Moderad. por debajo	Fuertem. por debajo	EI
Fuertemente pobre	Decrece fuertem.	Fuertem. por debajo	Poco por debajo	FI
Fuertemente pobre	Decrece fuertem.	Fuertem. por debajo	Moderad. por debajo	EI

Tabla 1: las 81 reglas de producción del sistema de recomendación

Resultados

El sistema fue implementado en el ambiente de Matlab, particularmente con el toolbox de Lógica Difusa. Se usó el método de inferencia difusa de Mamdani (Mamdani 1975, 1976, 1977).

En la figura 5 se aprecia la definición de una de las variables verbales y sus funciones de pertenencia que representan a los conjuntos difusos definidos.

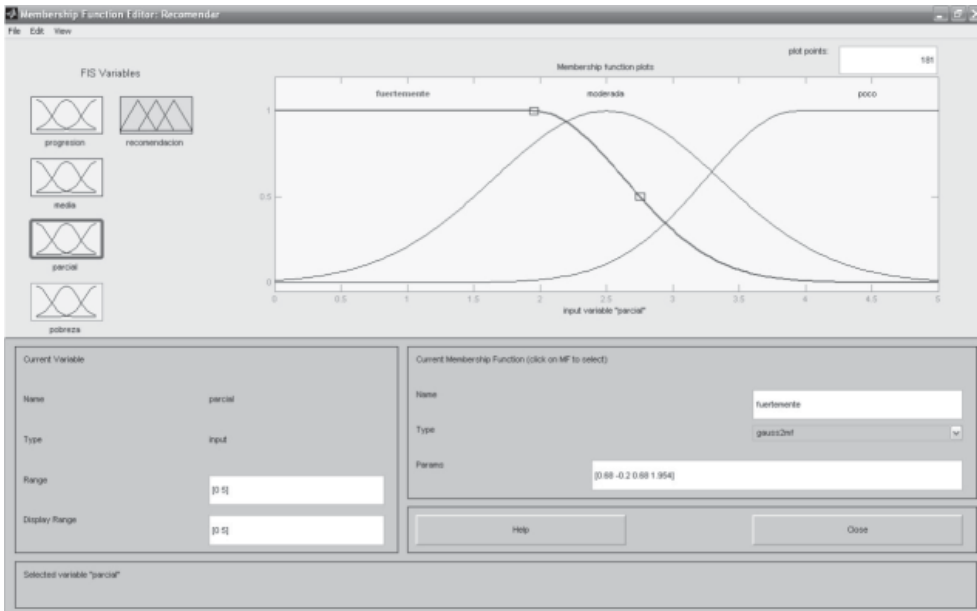


Figura 5: edición de la variable *Nota de examen parcial debajo de la aprobación* y sus 3 funciones de pertenencia

Por razones prácticas, se muestra, en la figura 6, 27 de las reglas de producción que capturan el razonamiento de los docentes de la asignatura.

En ventanas como la de la figura 6 puede visualizarse el proceso de inferencia completo para una entrada particular. Se aprecia cómo la forma de las funcio-

nes de pertenencia influyen en el resultado.

Las variables y sus valores actuales se muestran en la cabecera de las columnas. Los valores de las variables de entrada pueden modificarse en el panel de Input o arrastrando la línea de control que las atraviesa. El resultado, en la variable de salida *recomendación* se visualiza en la última columna.

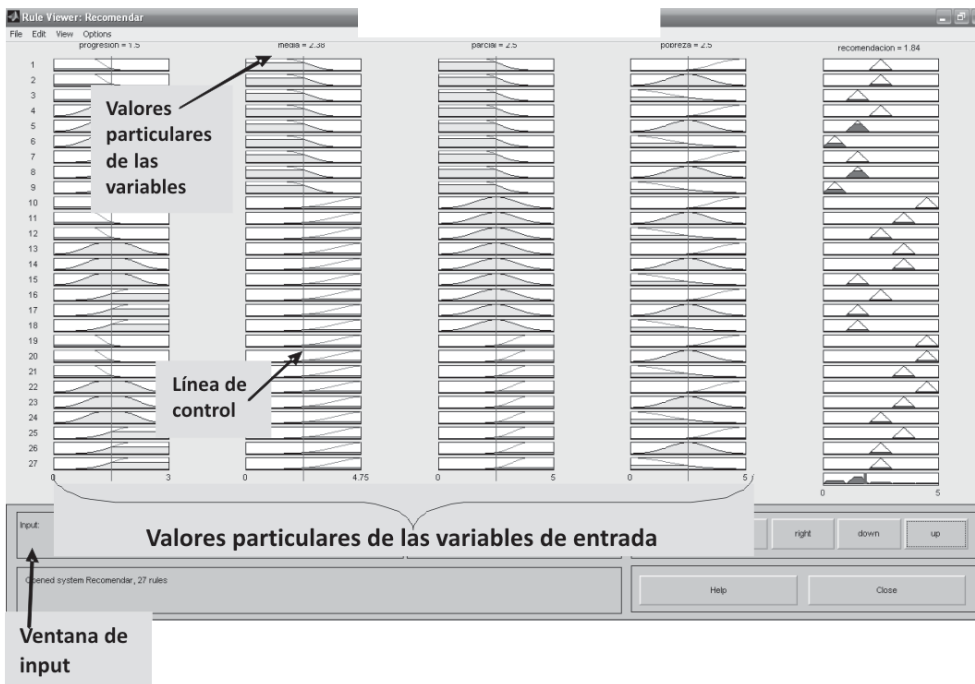


Figura 6: visualización de 27 reglas de producción

Los valores de la *recomendación* pueden convertirse a valores difusos siguiendo los intervalos que figuran en la figura 7

donde se aprecia que el valor 2.5 corresponde al valor *bastante insuficiente* (BI):

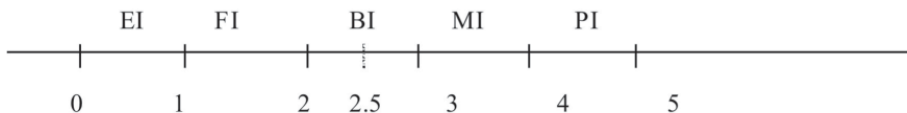


Figura 7: fuzzificación de los valores de la recomendación dados por MatLab

Las pruebas preliminares se hicieron en la asignatura Análisis Numérico para Ingeniería con datos del ciclo lectivo 2008. Sobre un total de 130 alumnos con actividad académica, 25 no alcanzaron la aprobación. Las acciones de recuperación se

hacen actualmente con un examen integrador de todos los temas de la asignatura, aplicado de la misma manera a todos los estudiantes. El sistema desarrollado puede ser usado para personalizar no fue decreciente, aunque con notas

muy bajas. El modelo debe ser probado con otros casos de estudios a fin de mejorar el rendimiento. Está previsto su mejoramiento con los resultados del ciclo lectivo 2009.

Conclusión

La propuesta emplea un modelo lingüístico difuso y da lugar a un mecanismo de inferencia similar a la forma en que los humanos toman decisiones, característica considerada importante en el diseño de un mecanismo de inferencia difusa. Para desarrollar tal mecanismo se ha utilizado como datos no difusos las notas de los estudiantes y como datos difusos el criterio subjetivo de los docentes.

Si bien los resultados alcanzados pueden considerarse satisfactorios, es necesario mejorar el conocimiento experto del sistema, contenido en la base de reglas difusas, a fin de ampliar la cobertura y exactitud del sistema. Como futuro trabajo sobre el sistema difuso está previsto agregar y mejorar las variables lingüísticas, así como enriquecer el sistema de reglas. Referido al sistema de diagnóstico del estudiante, está previsto agregar conocimiento sobre las posibles acciones pedagógicas reparadoras basadas en los resultados de la variable *recomendación*.

Notas

CIMEPB (Centro de Investigación en Procesos Básicos, Metodologías y Educación) Facultad Ingeniería-Psicología. UNMDP.

¹ Especialista en Tecnología Informática Aplicada a Educación y Calculista Científico, Profesor Titular de Computación y Análisis Numérico, Juan B. Justo 4302, 0223-4816600 int. 259, huapaya@fi.mdp.edu.ar. Facultad de Ingeniería (UNMDP).

² Especialista en Tecnología Informática Aplicada a Educación e Ingeniero Electrónico, Jefe de Trabajos Prácticos de Análisis Numérico y Computación. Juan B. Justo 4302, 0223-4816600 int. 259, flizarra@fi.mdp.edu.ar. Facultad de Ingeniería (UNMDP).

³ Computador Científico, Jefe de Trabajos Prácticos de Computación e Introducción a la Inteligencia Artificial. Juan B. Justo 4302, 0223-4816600 int 259, grarona@fi.mdp.edu.ar. Facultad de Ingeniería (UNMDP).

⁴ Doctor en Psicología. Director del (CIMEPB) Centro de Investigación en Procesos Básicos, Metodologías y Educación. Funes 3280, 0223-4752266 int 110, jvivas@mdp.edu.ar. Facultad de Psicología (UNMDP).

Bibliografía

BITWAS, R. (1995). «An application of fuzzy sets in students' evaluation». *Fuzzy sets and systems*. Pp 187-194. ELSEVIER.

CHEN, S. M., LEE, C.H. (1999). New methods for Students' Evaluation using fuzzy sets». *Fuzzy sets and systems*. 104:209:218.

DUBOIS, D. y PRADE, H. (1996) «What are fuzzy rules and how to use them». *Fuzzy sets and systems*, 84:169-185.

FOURALI, C. (1994) «Fuzzy logic and the Quality of Assessment of Portfolios». *Fuzzy sets and systems*. 68:123-139.

HARTLEY, R., PAIVA, A., SELF, J. (1995) «Externalizing Learner Models» Proceedings de *International Conference on Artificial Intelligence in Education*, editado por J. Greer, pp. 509-516. Washington. AACE.

LAW, C.(1996). «Using fuzzy numbers in educational grading systems». En *Fuzzy sets and systems*. Pp 311-323. ELSEVIER.

MAMDANI, E. y ASSILIAN, S. (1975), «An experiment in linguistic synthesis with a fuzzy logic controller,» *International Journal of Man-Machine Studies*, Vol. 7, No. 1, pp. 1-13.

MAMDANI, E.H (1976) «Advances in the linguistic synthesis of fuzzy controllers,» *International Journal of Man-Machine Studies*, Vol. 8, pp. 669-678.

MAMDANI, E.H.,(1977) «Applications of fuzzy logic to approximate reasoning using linguistic synthesis,» *IEEE Transactions on Computers*, Vol. 26, No. 12, pp. 1182-1191.

- SISON, R. y SHIMURA, M. (1988), «Student modeling and machine learning». *International Journal of Artificial Intelligence in Education*. 9, 128-158.
- ZADEH, L. (1965). «Fuzzy sets». *Inf. and control*. Vol 8, 338-353.
- ZADEH, L. «Fuzzy logic». *IEEE Computer*. Vol 21(4), 83-93. 1988.
- ZADEH, L. (1994). «Fuzzy logic, neural networks and soft computing». *Communications of the ACM*. Vol 37(3), 77-84, 1994.
- ZADEH, L. (1975). «The concept of a linguistic variable and its application to approximate Reasoning» . *I-III Information Sci* 8, , 199-250 y 301-357.



Cómo citar este artículo: Huapaya, C., Lizarralde, F., Arona, G., & Vivas, J. (2010). "Sesgos en la evaluación del estudiante con bajo rendimiento. Un sistema de mejoramiento diagnóstico basado en Lógica Difusa". *Revista de Educación* [en línea], 1 [citado AAAA-MM-DD]. Disponible en Internet: http://200.16.240.69/ojs/index.php/r_educ/article/view/13. ISSN 1853–1326.