



Modelo de lógica borrosa aplicada para valorar rúbricas de evaluación

Césari Matilde

Diplomatura en Métodos de Explotación Inteligente de Datos, Centro de Investigación CeReCoN, UTN - FRM
matilde.cesari@frm.utn.edu.ar

Césari Ricardo

Diplomatura en Métodos de Explotación Inteligente de Datos, Centro de Investigación CeReCoN, UTN - FRM
ricardocesari@ymail.com

Resumen

Se propone como aporte en el contexto universitario y el medio social. la implementación de rubricas de evaluación haciendo uso de un enfoque basado en la lógica difusa (LD). Los datos imprecisos, contiene incertezas que no pueden ser caracterizadas usando estadística o incluso razonamiento lógico tradicional. La Imprecisión propia en la mayoría de palabras humanas que son usadas para evaluar conceptos y derivar conclusiones, se logra minimizar su efecto de valoraciones subjetivas con el uso de abstracción y analogías; y unas pocas oraciones que describen contextos complejos. En este artículo se presenta como contribución metodológica la aplicación de la lógica difusa en un sistema que tiene como entradas las valoraciones realizadas a un conjunto de criterios pertenecientes a la rúbrica y como salida la calificación obtenida a partir de una serie de reglas de inferencia que relacionan los diversos niveles de los criterios de la rúbrica con una escala cualitativa. La Rubrica Difusa es aplicada sobre el modelo de rúbricas analíticas diseñado y presentado por Neil Carlo. Textualmente, el modelo expresa que los resultados de aplicación asociados a un nivel de dominio, son evaluados mediante la rúbrica analítica. En este modelo, la rúbrica analítica, está compuesta por uno o más criterios de evaluación, asociados a uno o más descriptores detallados por un contexto, mediante la aritmética borrosa es posible incorporar al modelo reglas lógicas que representan relaciones entre criterios y obtener una calificación ponderada por nivel y en general más objetivo a partir de valoraciones subjetivas.

Palabras clave: la evaluación, rubrica difusa, teoría de la posibilidad



Abstract

It is proposed as a contribution in the university context and the social environment. the implementation of evaluation rubrics using an approach based on fuzzy logic (LD). Imprecise data contains uncertainties that cannot be characterized using statistics or even traditional logical reasoning. The own Imprecision in the majority of human words that are used to evaluate concepts and derive conclusions, it is possible to minimize its effect of subjective evaluations with the use of abstraction and analogies; and a few sentences describing complex contexts. This article presents as a methodological contribution the application of fuzzy logic in a system that has as inputs the evaluations made to a set of criteria belonging to the rubric and as output the qualification obtained from a series of inference rules that relate the various levels of rubric criteria with a qualitative scale. The Fuzzy Rubric is applied to the analytical rubric model designed and presented by Neil Carlo. Textually, the model expresses that the application results associated with a mastery level are evaluated through the analytical rubric. In this model, the analytical rubric is made up of one or more evaluation criteria, associated with one or more descriptors detailed by a context. Using fuzzy arithmetic, it is possible to incorporate logical rules into the model that represent relationships between criteria and obtain a weighted rating. by level and in general more objective from subjective assessments.

Key words: evaluation, fuzzy rubric, possibility theory.

Introducción:

La evaluación es un proceso cognitivo complejo que implica diversos mecanismos en los cuales es necesario identificar los elementos que van a ser evaluados, fijar el marco en el que se va a realizar la evaluación, recopilar la información y finalmente obtener una valoración de los elementos evaluados. Nos encontramos a menudo con la necesidad de evaluar a partir de apreciaciones provenientes del ser humano, por ejemplo, para calificar a un grupo de alumnos, evaluar procedimientos en una auditoria, o para determinar la viabilidad de proyectos de desarrollo en ingeniería, incluso cualquier caso en que necesitamos obtener una nota numérica a partir de notas verbales de calificaciones de diversas unidades temáticas [1]

La información que manejamos diariamente presenta imperfecciones, generalmente no es de tipo precisa. La incerteza es aquella que surge de procesos humanos, como ser la sensación, la percepción, la experiencia cognoscitiva, el razonamiento y el pensamiento. Específicamente, la Incerteza Léxica, trata la imprecisión propia en la mayoría de palabras humanas, usadas para evaluar conceptos y derivar conclusiones; con el uso de abstracción y analogías, unas pocas oraciones describirán contextos complejos que son difíciles de modelar con precisión matemática.



En procesos de evaluación la información proporcionada por las personas implica siempre incertidumbre e imprecisión. [2;3]

La lógica difusa corresponde a una herramienta matemática que permite llegar a una conclusión específica, partiendo de una información de entrada, la cual puede ser indefinida o inexacta [4]. La importancia de la lógica difusa radica en el hecho de posibilitar la generación de resultados inteligibles que relacionan datos numéricos con términos lingüísticos que resultan más cercanos al lenguaje natural [5;6]. A partir de lo anterior, la lógica fuzzy puede ser de utilidad en cuanto a la obtención de niveles cualitativos dentro del proceso de evaluación.

La Rúbrica Difusa será aplicada sobre el modelo de rúbricas analíticas diseñado y presentado por Carlos Neil [7], en la Revista Electrónica de Tecnología Educativa número 80; utilizando el paradigma orientado a objetos, que, mediante un diagrama de clases, se detallan todos los componentes y sus relaciones para su potencial implementación. Textualmente, el modelo expresa que los resultados de aplicación asociados a un nivel de dominio, son evaluados mediante la rúbrica analítica. En este modelo, la rúbrica analítica, está compuesta por uno o más criterios de evaluación, asociados a uno o más descriptores detallados por un contexto que, a su vez, está formado por dos criterios analíticos, comentarios y propuestas de solución. Por otro lado, de cada descriptor se obtiene un nivel de logro. [7].

Determinar valores numéricos únicos en muchas ocasiones impide aceptar que este tipo de procesos no es por naturaleza exacto. Por lo que, el uso de la aproximación lingüística difusa para modelar esta información puede ayudar; ofrece un marco de trabajo natural donde las cantidades borrosas son interpretadas como distribuciones de posibilidad, describiendo los valores del conocimiento impreciso en conjuntos borrosos.

La teoría de la posibilidad, constituyen modelos, que resultan especialmente útiles para tratar con la imprecisión de manera más "natural" y más "humana". Emplean una terminología particular: Fuzzy (Difuso o Borroso); Crisp (Nítido); Fuzzification (Borrosificación: convertir un conjunto nítido en borroso); y Defuzzification (Desborrosificación: convertir un conjunto borroso en un valor —Crisp).[1]

Uno de los propósitos de la rúbrica borrosa propuesta por Diana Patricia Bedoya y Ruiza, Carlos Mario Sierra Duqueb [8] es sustentar el procesamiento formal, con miras a superar el proceder intuitivo- de valoraciones cuantitativas y cualitativas; en tal sentido, se acudió al estudio y aplicación del marco teórico de los sistemas difusos.

El instrumento de valoración se enfoca en el uso de sistemas de inferencia difusa o borrosa, siguiendo la metodología de análisis de datos imprecisos con lógica borrosa [9;10].

Un sistema de inferencia difusa se entiende como una articulación de elementos que pretenden realizar un proceso de razonamiento; el sistema se activa cuando de un conjunto de proposiciones borrosas requiere inferirse una conclusión que es, generalmente, difusa.

Para el caso que nos atañe una proposición borrosa hace referencia a la valoración cualitativa o cuantitativa efectuada por uno de los tipos de actores, respecto del nivel de cumplimiento en una de las dimensiones de formación, para un criterio de desempeño asociado con una competencia específica. La ambigüedad existente en expresiones lingüísticas se expresa en el concepto de grado de pertenencia o también denominada función de pertenencia.

Matemáticamente, los conceptos de sí/no, verdadero/falso están representados por medio del concepto clásico de conjunto, pero necesitamos extenderlo para poder representar este tipo de información más difusa.

La comprensión de este tipo de sistemas requiere una fundamentación en la noción de conjunto difuso. Un conjunto difuso permite a sus elementos tener un grado de pertenencia. Si el valor 1 se asigna a los elementos que están completamente en el conjunto, y 0 a los que están completamente fuera, entonces los objetos que están parcialmente en el conjunto tendrán un valor de pertenencia estrictamente entre 0 y 1 (Figura 1).

Un conjunto difuso A en un universo de discurso X se determina por medio de una función de pertenencia μ_A , que toma valores en el intervalo $[0,1]$; $\mu(x)$ es el grado de pertenencia del elemento x al conjunto A

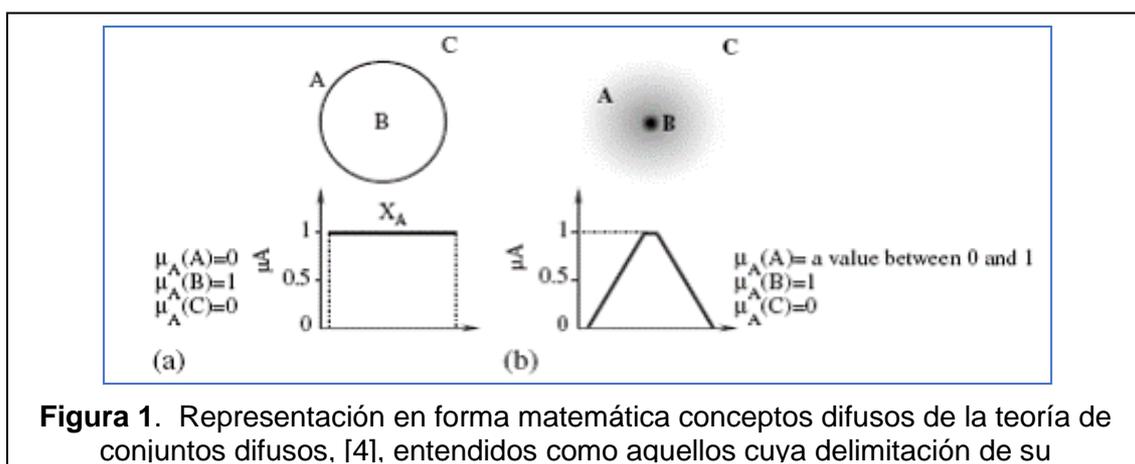


Figura 1. Representación en forma matemática conceptos difusos de la teoría de conjuntos difusos, [4], entendidos como aquellos cuya delimitación de su

Metodología

Para el desarrollo del instrumento de valoración empleando un sistema borroso, al que llamaremos Rúbrica borrosa, es apropiado exponer algunas guías o lineamientos que establecen el desenvolvimiento que debe seguir la rúbrica en un proceso de valoración

El estado inicial del proceso de valoración se establece con el ingreso de un valor numérico– uno por cada criterio A y B+ de cada contexto por cada criterio establecido para una competencia o rúbrica. Cada criterio de evaluación, producto de la descomposición del tema/problema a evaluar, tiene asociada una importancia relativa expresada en términos porcentuales, se denominan porcentaje de valoración general (Figura 2).

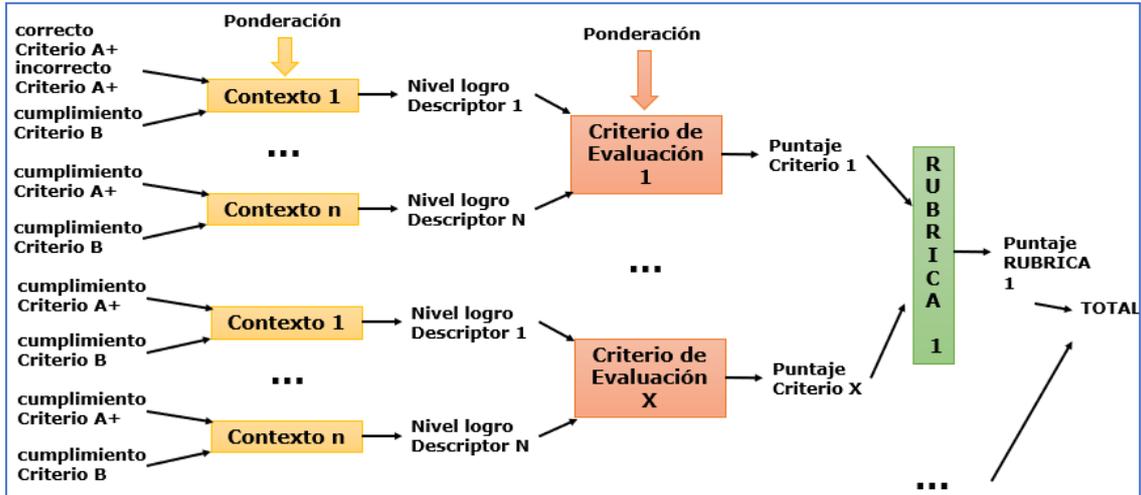


Figura 2. Descomposición esquemática de los criterios de evaluación en contextos el diseño de rúbricas analíticas de [7]. Fuente Elaboración propia

De forma análoga, los contextos vinculados a los criterios de evaluación poseen una ponderación porcentual asociada y tiene como objetivo establecer un matiz en la consideración del valor numérico obtenido, se describen como porcentaje de valoración parcial [7].

La elección de la escala y la determinación de los pesos son colocados según criterio del evaluador en un principio tomamos todos los pesos como uno, es decir todo tiene la misma importancia, luego según parámetros o reglamentación relacionadas al sistema de calificación se definen los pesos por contexto y por criterio de evaluación, tanto estáticos como aquellos que pueden cambiar de manera dinámica durante la evaluación.

A partir de acá se define lo siguiente:

El sistema debe posibilitar la especificación, en la escala de valoración, de valores cualitativos mediante conjuntos difusos (Figura 3 y 4). Cada institución educativa tiene la facultad de definir: la cantidad de conjuntos difusos (granularidad) por cada criterio, la función de pertenencia correspondiente en la escala de valoración numérica, así como la cantidad de niveles de la escala numérica de valoración (su resolución).

Las funciones de pertenencia de los conjuntos a representar son usadas durante todo el proceso, es decir, las diferentes valoraciones y los resultados obtenidos son dados en la misma escala.

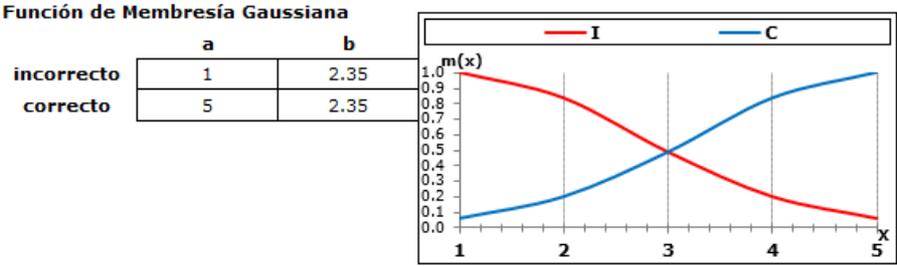


Figura 3. Variable lingüística para representar el cumplimiento de los criterios A+ y B. Dos conjuntos borrosos definidos con funciones gaussianas con parámetro a=media y b=desvío. Fuente Elaboración propia

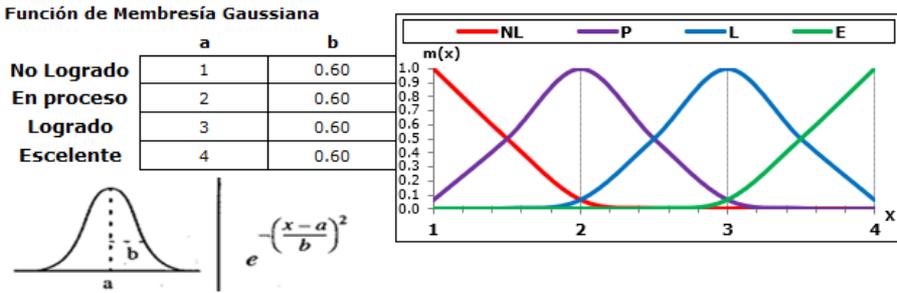


Figura 4. Variable lingüística para representar el Nivel de logro de los contextos. Dos cuatro conjuntos borrosos definidos con funciones gaussianas con parámetro a=media y b=desvío. Fuente Elaboración propia

Cada criterio de evaluación se puede subdividir en uno o más descriptores. Cada uno de ellos, a su vez, estará compuesto por un contexto vinculado y dos criterios analíticos A+ y B, que podrán ser correctos o incorrectos [7].

Los niveles de logro, permite establecer los niveles con los que se evaluarán los criterios de evaluación. Deben expresarse en forma gradual, considerando desde un nivel mínimo hasta un nivel óptimo: excelente, el estudiante manifiesta poseer un nivel de competencia excelente; logrado, demuestra un buen dominio de la competencia; en proceso, desempeña los criterios de calidad de forma aceptable y no logrado, no posee un nivel de dominio suficiente [7]

Los descriptores permiten detallar las características de cada nivel de logro según el grado en que los criterios de evaluación hayan sido alcanzados. Constituyen, en el diseño de las rúbricas, el elemento más complejo que requieren, por parte de los docentes, mucho esfuerzo. [7]



Este proceso requiere la definición previa de las escalas de valoración, y la definición de las bases de reglas difusas correspondientes a cada módulo. Para llevar a cabo el proceso de inferencia se debe:

Paso 1: Seleccionar la base de reglas con la cual se desea hacer la valoración [7].

Paso 2: Ingresar la valoración, numérica o lingüística, para cada criterio. Cuando un valor asignado es numérico, el módulo lo borrosifica, es decir, se le establece el grado de pertenencia en el conjunto difuso estipulado para el criterio de desempeño en cada regla borrosa

En la ecuación de la Figura 5, x son los valores originales de la escala continua de 1 a 5 para evaluar si un criterio A+ y B es correcto o incorrecto, los parámetros a y b corresponden a la función gaussiana definida para cada uno de los dos conjuntos definidos para evaluar el cumplimiento del criterio.

Borrosificación

$$\mu_{a,b}(x) = \exp[-b \cdot (x-a)^2]$$

Figura 5. Ecuación para calcular a partir de un valor x la pertenencia al conjunto difuso, en forma de campana con media a

Paso 3: Evaluar cada regla una vez efectuada la fusificación: se identifica si para todos los criterios involucrados en el antecedente de la regla, el valor de pertenencia del elemento fusificado es mayor a cero; de ser así, se aplica el operador de implicación. El valor resultante se denominará valor de activación (o grado de cumplimiento)

Obtener la salida borrosa de cada regla. El grado de pertenencia de cada valor de la escala numérica del consecuente al nuevo conjunto borroso es el grado original cuando no supere el valor de activación de la regla, o es este último en caso contrario.

No logrado = MIN (incorrecto A+ y incorrecto B)
en proceso = MIN(incorrecto A+ y correcto B)
logrado = MIN(correcto A+ y incorrecto B)
excelente = MIN(correcto A+ y correcto B)

Figura 6. Ecuación para calcular la implicación lógica de las reglas

En la **Figura 6** se plantean las funciones para evaluar las reglas de inferencia, MIN, representa la función de obtener el mínimo (AND lógico) de los valores de pertenencia de los conjuntos incorrecto y correcto de los criterios A y B+.

Paso 4: Agregar las salidas producidas por las reglas. Se estipula para cada elemento en la escala de valoración definido en salida, el máximo grado de



pertenencia en los conjuntos borrosos resultantes de cada regla; este valor será el grado de pertenencia del elemento en el conjunto borroso final o agregado. La obtención de un **Valor Borroso agregado** por cada criterio de evaluación se realiza mediante una métrica difusa. El método es de tipo métrico, usa ponderaciones, como métrica utiliza la media armónica. Los valores V_{ik} (valor borroso de la característica k en el nivel i) se ponderan por el peso P_{ik} , asimismo se dividen por el mismo, estos cálculos permiten aplicar la media armónica [1]

Paso 5. Calcular un valor concreto del conjunto borroso obtenido en el paso anterior.

Por medio de los **métodos de desborrosificación**, a partir del conjunto de valores borrosos agregados, se obtiene información precisa expresada mediante un valor crisp. Cada método utiliza una fórmula matemática para calcular la salida final. La bibliografía sugiere diversas alternativas de operadores de desborrosificación, entre ellos, los denominados clásicos como la teoría —mean of maximum (MOM) basada en el núcleo del número borroso; teoría —center of areall (COA), basada en el concepto del centro del área y la teoría llamada media borrosa —fuzzy mean (FM) [1]

De manera genérica se propone para la desborrosificación el método modificado de centro máximo y se obtiene un valor crisp por criterio de evaluación, a través de la media ponderada de los criterios podemos obtener una valoración total para la rúbrica evaluada.

The equation is presented in a box with a light orange background. The title "DesBorrosificación" is at the top. The formula is:
$$V_{CRISP} = \frac{\sum_{i=1}^k \left(\frac{t_i^2}{b^2} \cdot a_i \right)}{\sum_{i=1}^k \left(\frac{t_i^2}{b^2} \right)}$$

Figura 7. Ecuación para calcular a partir de un número borroso constituido por valores de pertenencia a cada conjunto difuso de la variable lingüística, en un valor cuantitativo

En la ecuación de la **Figura 7**, t son los valores agregados para cada conjunto del nivel de logro de un criterio y los parámetros a y b corresponden a la función gaussiana definida para cada uno de esos conjuntos borrosos.

La calificación general permite hacer un seguimiento en el tiempo, mediante la construcción de un panel de control que permite ver la evolución de la evaluación y estimar las calificaciones con las valoraciones analizadas en el momento.

Conclusiones

Se plantea una tecnología actual que ofrece un marco flexible de evaluación donde se pueda expresar preferencias y valoraciones subjetivas en conjuntos de etiquetas lingüísticas distintos, puede mejorar los resultados de la evaluación.

Esta contribución propone un modelo de evaluación, basado en los test de viabilidad de proyectos de desarrollo de sistemas expertos, con un marco de evaluación que define un contexto lingüístico modelado mediante jerarquías lingüísticas para mejorar la precisión de los procesos de computación con palabras.

Como evidencian los resultados, es una propuesta genérica que se puede aplicar a todo tipo de evaluación; en particular destaca las ventajas como herramienta de evaluación para el docente, permitiendo mejorar la subjetividad del método de evaluación actual permitiendo medir el rendimiento continuo del alumno de manera más humana.

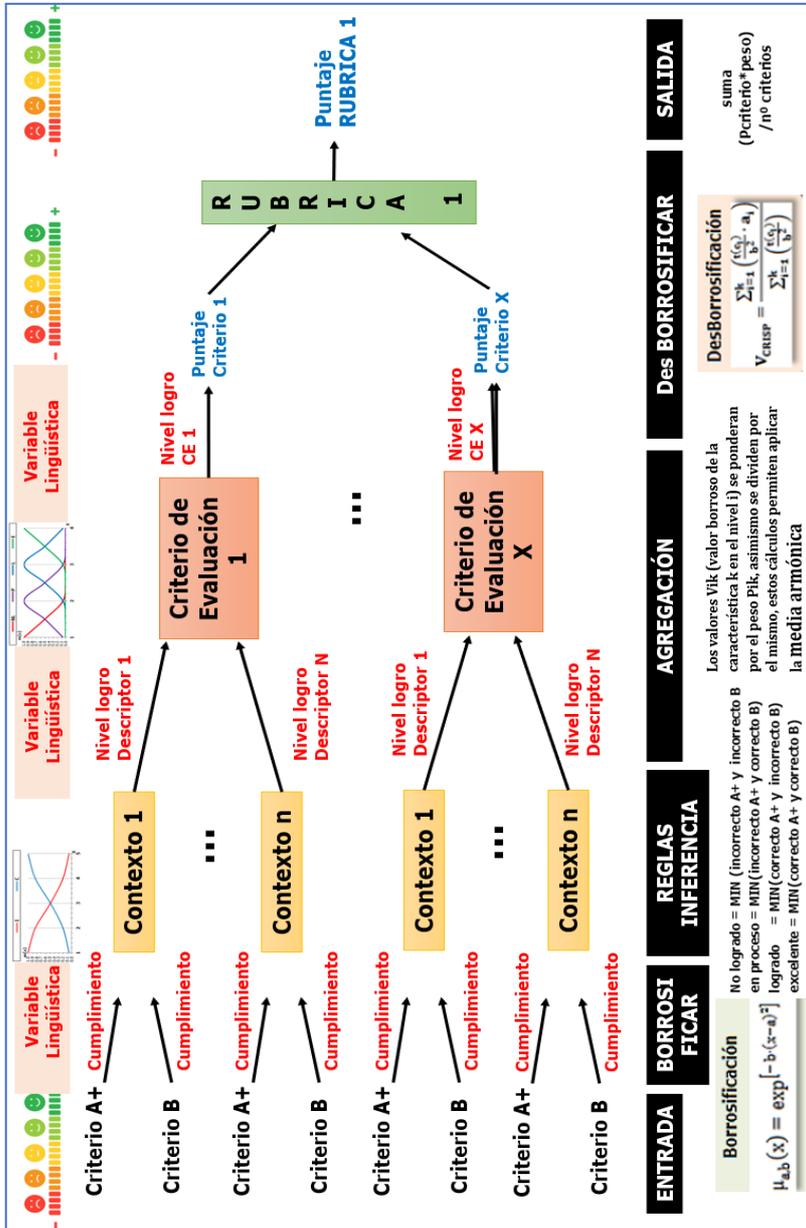


Figura 8. Modelo de rúbrica borrosa propuesto. Fuente Elaboración propia



Citas

- [1] Césari, M., & Césari, R. (2016). Sistema de Calificación con Lógica difusa. Noveno Encuentro de Investigadores y Docentes de Ingeniería-EnIDI 2017.
- [2] Chanchí, G. G. E., Sierra, M. L., & Campo, M. Y. (2021). Aplicación de la lógica difusa en la implementación de rúbricas de evaluación en el contexto universitario. *Revista Ibérica de Sistemas e Tecnologias de Informação*, (E42), 174-187.
- [3] Chanchi, G. E., Ospina, M. A., & Monroy, M. E. (2020). Chanchí, G. G. E., Sierra, M. L., & Campo, M. Y. (2021). Aplicación de la lógica difusa en la implementación de rúbricas de evaluación en el contexto universitario. *Revista Ibérica de Sistemas e Tecnologias de Informação*, (E42), 174-187. en el análisis de inspecciones heurísticas de usabilidad. *Revista ESPACIOS*. ISSN, 798, 1015.
- [4] Zadeh, L. A., Klir, G. J., & Yuan, B. (1996). *Fuzzy sets, fuzzy logic, and fuzzy systems: selected papers (Vol. 6)*. World Scientific.
- [5] Colino, E. A., & Gracia, J. R. (2007). Conceptos y aplicaciones de la lógica borrosa. *Técnica industrial*, (269), 58-62.
- [6] Ramírez, N. V., & Laguna, M. (2012). La lógica borrosa: conjuntos borrosos, razonamiento aproximado y control borroso. *Pistas Educativas*, 100, 55-65.
- [7] Neil, C., Battaglia, N., & Zemborain, M. E. D. V. (2022). Marco metodológico para el diseño de rúbricas analíticas. *EduTec. Revista Electrónica De Tecnología Educativa*, (80).
- [8] Bedoya Ruiz, D. P., & Sierra Duque, C. M. (2016, July). Evaluación por competencias utilizando lógica difusa. In *In-Red 2016. II Congreso nacional de innovación educativa y docencia en red*. Editorial Universitat Politècnica de València.
- [9] Césari, M. I., Ventrera, N. B., & Gámbaro, A. (2018). Análisis de datos sensoriales de tomate triturado con lógica difusa y técnicas multivariadas. *Revista de la Facultad de Ciencias Agrarias. Universidad Nacional de Cuyo*, 50(1), 233-248.
- [10] Césari, I. (2016). Metodología de Análisis Sensoriométrico con Lógica Borrosa e Ingeniería de Modelos Borrosos para evaluar Calidad Sensorial en Alimentos. Universidad Tecnológica Nacional Regional Mendoza
- [11] Neil, C., Battaglia, N., & Zemborain, M. E. D. V. (2022). Marco metodológico para el diseño de rúbricas analíticas. *EduTec. Revista Electrónica De Tecnología Educativa*, (80).