

# Sistema de Recomendación de Asignaturas en el Proceso de Registro y Matrícula de Estudiantes Universitarios

*Recommendation System of Subjects in the Registration and Enrollment Process of University Students*

## ARTICLE HISTORY

Received 16 April 2020  
Accepted 06 June 2020

### **Milton Mariduenaa-Arroyave**

Facultad de Ciencias Matemáticas y Físicas  
Universidad de Guayaquil  
Guayaquil, Ecuador  
milton.mariduenaa@ug.edu.ec

### **Lorenzo Cevallos-Torres**

Facultad de Ciencias Matemáticas y Físicas  
Universidad de Guayaquil  
Guayaquil, Ecuador  
lorenzo.cevallost@ug.edu.ec

### **Miguel Botto-Tobar**

Facultad de Ciencias Matemáticas y Físicas  
Universidad de Guayaquil  
Guayaquil, Ecuador  
miguel.bottot@ug.edu.ec

# Sistema de Recomendación de Asignaturas en el Proceso de Registro y Matrícula de Estudiantes Universitarios

## *Recommendation System of Subjects in the Registration and Enrollment Process of University Students*

**Milton Mariduenaa-Arroyave**

Facultad de Ciencias  
Matemáticas y Físicas  
Universidad de Guayaquil  
Guayaquil, Ecuador  
milton.mariduenaa@ug.edu.ec

**Lorenzo Cevallos-Torres**

Facultad de Ciencias  
Matemáticas y Físicas  
Universidad de Guayaquil  
Guayaquil, Ecuador  
lorenzo.cevallost@ug.edu.ec

**Miguel Botto-Tobar**

Facultad de Ciencias  
Matemáticas y Físicas  
Universidad de Guayaquil  
Guayaquil, Ecuador  
miguel.bottot@ug.edu.ec

**Resumen** — Actualmente, los sistemas de recomendación son ampliamente utilizados para analizar preferencias de usuarios y sugerirles ítems afines. En el ámbito universitario, el momento en el cual una asignatura es elegida por un estudiante para su siguiente etapa educativa, es monitoreado por un Consejero Académico, quien de acuerdo con el récord estudiantil, y comparando perfiles similares a lo largo de su carrera, debe recomendar cuáles asignaturas, podrían contribuir al rendimiento y aprendizaje del estudiante. El presente trabajo representó un esfuerzo por diseñar un recomendador que se apoya en el modelado de las relaciones causales existentes entre las asignaturas del pénsum curricular de una carrera universitaria, empleando mapas cognitivos difusos y operadores de agregación OWA. Se aplicó el flujo de trabajo del modelo propuesto y su implementación, a través de la herramienta informática (FCM-Decision). Se desarrolló un estudio de caso con los récords estudiantiles de una universidad en Guayaquil, y además, se realizó un experimento para probar los resultados de interpretabilidad con otros modelos existentes. Entre los principales resultados están la fiabilidad de las métricas para el análisis estático de mapas difusos, la similitud respecto a un estudiante objetivo, y la importancia que cada asignatura representa en un nuevo registro.

**Palabras clave** — Sistemas de Recomendación, Asignaturas, Récord\_estudiantil, Mapas Cognitivos Difusos.

**Abstract** — *Currently, recommendation systems are widely used to analyze user preferences and suggest related items. At the university level, the moment in which a subject is chosen by a student for his next educational stage, is monitored by an Academic Counselor, who according to the student record, and comparing similar profiles throughout his career, must recommend which subjects could contribute to student performance and learning. The present work represented an effort to design a recommender that is based on the modeling of the causal relationships existing between the subjects of the curricular curriculum of a university career, using fuzzy cognitive maps and OWA aggregation operators. The workflow of the proposed model and its implementation were applied through the computer tool (FCM-Decision). A case study with the student records of a university in Guayaquil was developed, and an experiment was also carried out to test interpretability results with other existing models. Among the main results are the reliability of the metrics for the static analysis of fuzzy maps, the similarity with respect to a target student, and the importance that each subject represents in a new record.*

**Keywords** — Recommendation Systems, Subjects, Student\_Record, Fuzzy Cognitive Maps.

## 1 INTRODUCCIÓN

En la actualidad la toma de decisiones sobre temas o aspectos de la vida real, resulta ser compleja para las personas. Con el progreso en las tecnologías de punta y el incremento de los sistemas de gestión y control de procesos, se han desarrollado herramientas de asistencia a la toma de decisiones, partiendo de información previa registrada. Dentro de los elementos principales que juegan un papel fundamental a la hora de tomar una decisión se manifiestan: (i) El grado de madurez de las personas; (ii) El nivel de conocimientos; y, (iii) la información disponible relacionada con la decisión que se desea tomar.

Estos elementos en el contexto universitario no son la excepción, también pueden considerarse por el estudiante al momento de tomar una decisión, que implique mejorar su rendimiento académico futuro. La elección de una asignatura, o un grupo de ellas asociadas a un perfil académico, son eventos que se presentan en las diferentes etapas de formación de un estudiante y que resulta difícil, para aquellos con falta de experiencia o porque cuentan con deficiente información, para tomar una mejor decisión, frente a un proceso de registro o matriculación en un nuevo ciclo lectivo.

En el proceso de toma de decisiones, es posible modelar la información de un suceso mediante un grafo, que permita identificar y analizar la relación que existe entre los datos que la componen. El grafo se representa como una Matriz de Decisión, con información existente que puede ser incierta, ambigua, imprecisa y borrosa [1]. En la toma de decisiones bajo incertidumbre, los modelos causales pueden ser empleados como una alternativa para comprender e implementar sistemas complejos [2] [3].

Como parte de las técnicas empleadas para la inferencia causal, se destacan las Redes Bayesianas (RB) y los Mapas Cognitivos Difusos (MCD). Ambas técnicas son representadas en términos computacionales como grafos dirigidos [4].

Los MCD han sido empleados en diferentes dominios de la vida real [5]. Su empleo en la toma de decisiones [6] y el desarrollo de sistemas de ayuda para este fin [7] se destacan como áreas de uso frecuente. Su representación en forma de grafos dirigidos permite esquematizar la información y la forma cómo se encuentran relacionados los elementos que componen el problema. Esta técnica permite

visualizar y estudiar la dinámica de cambios de un estado determinado a otro, los mismos que son analizados en diferentes situaciones [8]. Por otra parte, los sistemas de recomendación, también se han convertido en herramientas muy útiles al proporcionar ayuda a la hora de tomar decisiones en diferentes entornos de la sociedad [9].

En ciertos sistemas educativos universitarios, un consejero académico es aquel funcionario de gestión, por lo general un docente encargado de orientar a los alumnos en educación superior. Su función principal, basada en su experiencia, es brindar soporte a los estudiantes en su visión respecto a la decisión acerca de temas importantes, como es la selección de las asignaturas de la malla curricular de su carrera profesional, en época de registros académicos o procesos de matriculación de inicio de semestre, entre otros temas. Además, constituyen un apoyo a los estudiantes con problemas en su rendimiento académico, ayudando a los mismos a encontrar posibles soluciones para su mejora.

Manejar de forma eficiente toda la información académica referente a los estudiantes es un proceso complejo para estos profesionales. El gran volumen de información a analizar para la selección de criterios y consejos a brindar, es un proceso complejo en el cual es posible obviar información relevante del perfil del estudiante. Colateralmente a esta problemática se tiene la carencia de una plataforma virtual para la gestión de toda la información, que se genera durante el proceso y que es de vital importancia para la toma de decisiones al momento de aconsejar al estudiante basado en experiencias y perfiles similares, así como ser considerado como punto de partida para aquellos nuevos especialistas que se incorporen en el proceso docente de consejería estudiantil.

Como objetivo general de la investigación, está desarrollar un sistema inteligente de recomendación de asignaturas, apoyado en el uso de Mapas Cognitivos Difusos (MCD) a partir del análisis del Perfil Académico de un estudiante universitario; y como objetivos específicos: (i) estructurar los componentes del sistema; (ii) describir el Flujo de Trabajo; y, (iii) validar el modelo del sistema. La novedad científica del trabajo se refleja en la propia concepción del modelo de recomendación y la fundamentación teórica de flujo de trabajo basado en inferencias causales.

El aporte práctico de la investigación se manifiesta en una herramienta informática (FCM-Decision) para la generación de recomendaciones de asignaturas.

La estructura del artículo consta de cuatro apartados: En la Sección II se desarrolla la fundamentación y el marco teórico conceptual de los Modelos de recomendación basados en MCD. En la Sección III, se desarrolla el análisis y diseño del Sistema de Recomendación propuesto para la ayuda a la toma de

decisiones en la elección de asignaturas durante el registro de un nuevo ciclo lectivo, estimando las calificaciones obtenidas y que podría obtener, así como el esfuerzo a realizar para mejorar el aprovechamiento académico. En la Sección IV, se expone la validación de los algoritmos, la presentación de resultados, su análisis y discusión. Finalmente, se presentan las Conclusiones y Recomendaciones resultantes de la investigación en la Sección V, acompañadas de las respectivas Referencias Bibliográficas.

## 2 MARCO TEÓRICO PARA EL DESARROLLO DEL SISTEMA

### 2.1 Toma de decisiones

La toma de decisiones es el proceso de aprendizaje natural o estructurado, mediante el cual se elige entre dos o más alternativas, opciones o formas para resolver diferentes situaciones o conflictos de la vida, la familia, la empresa u organización. Cada día se toman decisiones, las cuales no necesariamente son la mejor opción. Cuando se toman decisiones es necesario estar conscientes, que cada una conlleva consecuencias positivas o negativas y también implica un riesgo. La toma de decisiones a nivel individual está caracterizada por el uso del razonamiento y pensamiento de una persona, para elegir una decisión ante un problema o conflicto que le presente la vida.

En [10] proponen un procedimiento para la toma de decisiones, el cual consta de las siguientes actividades:

1. Definir el problema de toma de decisiones.
2. Analizar el problema y la información.
3. Identificar las alternativas de solución.
4. Establecer criterios o expertos por los cuales ellas puedan ser evaluadas.
5. Evaluar alternativas y seleccionar la mejor.

En el análisis de los procesos de toma de decisiones, se inicia configurando la matriz de decisión, en la Fig. 1. Se detalla una representación de los elementos que la componen:

	$X_1$	$X_2$	...	$X_j$	...	$X_n$
$A_1$	$x_{11}$	$x_{12}$	...	$x_{1j}$	...	$x_{1n}$
$A_2$	$x_{21}$	$x_{22}$	...	$x_{2j}$	...	$x_{2n}$
...	...	...	...	...	...	...
$A_i$	$x_{i1}$	$x_{i2}$	...	$x_{ij}$	...	$x_{in}$
...	...	...	...	...	...	...
$A_m$	$x_{m1}$	$x_{m2}$	...	$x_{mj}$	...	$x_{mn}$

Fig. 1. Matriz de decisiones

Los elementos presentes en la matriz de decisión son los siguientes:

$A_1, \dots, A_m$ : Los diferentes estados de la naturaleza que se pueden presentar.

$X_1, \dots, X_n$ : Las diferentes decisiones o estrategias que puede adoptar el decisor.

$x_{11}, \dots, x_{mn}$ : Los diferentes resultados, que surgen de combinar una estrategia con los

diferentes estados de la naturaleza. Éstos pueden ser exactos o imprecisos, según la información disponible para el problema.

Los ambientes de decisión se pueden clasificar de la siguiente forma [1]:

— Certidumbre: La decisión es inmediata y se conoce cuál es la mejor decisión posible.

— Riesgo: Se desconoce cuál es la mejor solución posible, pero sí se puede asociar un nivel de probabilidad asociado a las posibles situaciones futuras.

— Incertidumbre: Se determina cuáles son los posibles resultados futuros, pero se desconoce el nivel de probabilidad asociada a ellos.

Los Métodos de Decisión Multicriterio (MDMC) [11] constituyen un conjunto de metodologías que brindan soporte a las circunstancias, donde criterios en conflicto, metas, objetivos, y puntos de vista, deben ser tomados en consideración a la hora de la toma de decisiones. Esta acción bajo múltiples criterios es una tarea subjetiva que depende de las preferencias de un grupo de decisores. Se relaciona con la estructuración del problema, el modelado de las preferencias, la construcción y caracterización de diferentes funciones de agregación de criterios, así como de la construcción de herramientas, que les brinden soporte a los modelos. En la actualidad la decisión multicriterio es considerada como la parte más importante de la teoría de la decisión [12].

### 2.2 Sistemas de recomendación

El objetivo de los sistemas de recomendación (RS) es facilitar el proceso de toma de decisiones, en el instante de elegir entre un conjunto de posibilidades. La mayoría de los sistemas de recomendación están catalogados en base a la fuente de conocimiento: así, pueden distinguirse entre los sistemas basados en contenido y los sistemas colaborativos.

Por lo general, los sistemas de recomendación se encargan de proporcionar a los usuarios consejos e información personalizada sobre productos o servicios que puedan ser de interés a la hora de tomar una decisión. Este proceso, en el que el sistema guía al usuario a la hora de realizar una elección, puede proporcionar resultados que sean de gran utilidad, ya sea ahorrando tiempo, proporcionando datos relevantes de forma cómoda y fácil, e incluso obteniendo información que permite valorar opciones que de otra forma antes no se habrían contemplado, algo muy apreciado por la mayoría de los usuarios.

Simplificando, se puede decir que un sistema de recomendación reduce su problemática al predecir la puntuación que el usuario daría a una serie de ítems que todavía no ha calificado. Esta predicción, de la forma más intuitiva, podría basarse en las calificaciones basadas en puntajes, que otros usuarios asignaron en el pasado (sistemas colaborativos o sistemas de filtrado colaborativo), aunque también puede obtenerse usando otra clase de información distinta a las valoraciones mencionadas, como pueden ser las características concretas de los productos, los perfiles de preferencias de los distintos usuarios a la hora de tomar decisiones, información demográfica sobre ellos, heurísticas basadas en comportamientos humanos, entre otros; y por supuesto, usar hibridaciones de diversas aproximaciones, entre las existentes.

Aunque los RS tienen el mismo objetivo, ayudar al usuario generando una serie de recomendaciones, de manera que se minimice al máximo la búsqueda que el usuario debe realizar, existen diversas fuentes de las que se nutren de la información requerida y diversas técnicas que mediante estos sistemas construyen, evalúan y proporcionan sus resultados [13]: (i) Basados en contenido [14]; (ii) Demográficos [15]; (iii) Basados en conocimiento [16]; (iv) Basados en utilidad [17]; (v) Sistemas de recomendación Híbridos [18]; y, (vi) los Sistemas de recomendación Basados en Filtrado Colaborativo.

Estos últimos RS se aplican actualmente en muchos sectores del entorno web. Los usuarios de un sistema colaborativo comparten sus valoraciones y opiniones con respecto a los ítems que conocen de forma, en cambio que otros usuarios del sistema, puedan decidir qué elección realizar. Así mismo, otros sistemas proporcionan recomendaciones personalizadas, para aquellos elementos que resultan atractivos para el usuario.

Es importante recalcar, que en el filtrado colaborativo son los usuarios, quienes determinan la relevancia, calidad, prioridad e interés de los ítems, por lo que se puede realizar el filtrado sobre elementos difíciles de analizar mediante computación. Esto se consigue gracias a que el problema no se intenta solucionar analizando los elementos que se trata de recomendar, sino manejando directamente las valoraciones, objetivas y/o subjetivas, que sobre esos elementos se han realizado.

Además, el filtrado colaborativo tiene la competencia de discernir cómo se adapta un ítem a las necesidades o intereses de los

usuarios, basándose en la propia capacidad de los humanos de analizar en términos de calidad o gusto, algo difícil de realizar por procesos computacionales.

Por otro lado, se puede aportar al usuario ítems fortuitos por casualidad o accidente, haciendo recomendaciones de ítems importantes que dicho usuario no esperaba encontrar, que de otro modo posiblemente jamás habría considerado. Este fenómeno es conocido con el término en inglés serendipity, y hace referencia a un hallazgo casual beneficioso, afortunado y no esperado [19]. Si bien esta técnica de recomendación deja la elección del sistema de recomendación al usuario del sistema hay que tener en cuenta que se necesita una alta capacidad de cómputo para que este sistema sea usable. El sistema debe estar configurado para poder calcular recomendaciones de forma paralela y estar preparado para situaciones de alta carga en las horas pico del sistema.

Como se aprecia, el potencial del filtrado colaborativo es significativo, cuando se debe discriminar y tener en cuenta comportamientos subjetivos. Sin embargo, según [20], para alcanzar el potencial completo, es casi seguro que debe hibridarse con tecnologías de otros sistemas, basados en contenido.

### 2.3 Algoritmos de Agrupamiento

Con el objetivo de identificar elementos pertenecientes a un mismo grupo y distinguir las diversas formas de distribución de la información en un dominio de estudio, el empleo de algoritmos de agrupamiento juega un papel muy importante para la obtención de mejores resultados.

Los algoritmos de recomendación colaborativos pueden ser agrupados en dos clases generales [21]:

- Algoritmos basados en memoria (o en vecindad, o heurísticos)  
Son esencialmente implementaciones heurísticas, que realizan predicciones basadas en una colección completa de ítems valorados previamente por el usuario. Es decir, el valor de una puntuación no conocida  $v$  para un usuario  $u$  sobre un ítem  $i$  se calcula como un agregado de las valoraciones de otros usuarios (los  $K$  más parecidos) para el mismo ítem  $i$ .
- Algoritmos basados en modelos:  
Utilizan una colección de valoraciones para aprender un patrón. Mejora el rendimiento en cuanto a la predicción

porque da un fundamento más intuitivo, el cual será utilizado a la hora de realizar las futuras predicciones. Algunos algoritmos colaborativos basados en modelos se fundamentan en:

- Uso de redes bayesianas [22] y sus derivados [23].
- Métodos de clustering para filtrado colaborativo [24].
- Análisis iterativo de la componente principal [25].
- Otro método interesante es utilizar modelos de aprendizaje de manera, que se pueda tratar el filtrado colaborativo como un problema de clasificación. [26].

Teniendo en cuenta las características de los algoritmos de filtrado, que fueron argumentados y los datos que serán manejados en nuestro problema, por la simplicidad e intuición de entendimiento, su estructura flexible, adaptable al razonamiento causal, la propuesta estará centrada en los algoritmos basados en modelos. Con la implementación de un modelo de Mapa Cognitivo Difuso (MCD) para su simulación, se realizarán las predicciones de asignaturas del sistema a desarrollar para dar solución al problema.

### 2.4 Distancia euclidiana

En álgebra, geometría y más específicamente, en análisis real, análisis complejo y geometría analítica, se trata de una función no negativa usada en diversos contextos para calcular la distancia entre dos puntos, primero en el plano y luego en el espacio. También sirve para definir la distancia entre dos puntos en otros tipos de espacios o más dimensiones. Sus bases se encuentran en la aplicación del Teorema de Pitágoras sobre triángulos rectángulos, donde la distancia euclidiana viene a ser por lo general, la longitud de la hipotenusa, conformado por cada punto y los vectores proyectados sobre los ejes directores al nivel de la hipotenusa.

En  $\mathbb{R}^2$ , sean los puntos  $A = (x_A; y_A)$ ,  $B = (x_B; y_B)$  se define la distancia euclidiana entre dichos puntos por:

$$d_E(A, B) = \sqrt{(x_A - x_B)^2 + (y_A - y_B)^2} \quad (1)$$

En  $\mathbb{R}^3$ , sean los puntos  $A = (x_A; y_A; z_A)$  y  $B = (x_B; y_B; z_B)$  se define la distancia euclidiana mediante la expresión:

$$d_E(A, B) = \sqrt{(x_A - x_B)^2 + (y_A - y_B)^2 + (z_A - z_B)^2} \quad (2)$$

Y de manera más general en un espacio  $\mathbb{R}^N$ , la distancia euclidiana entre dos puntos  $N$ -dimensionales:  $A = (a_1; a_2; \dots; a_N)$  y  $B = (b_1; b_2; \dots; b_N)$  se define como:

$$d_E(A, B) = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + \dots + (a_N - b_N)^2} \quad (3)$$

De manera general, la métrica euclidiana entre dos puntos se define como la longitud del segmento de recta que une a dichos puntos.

La distancia euclidiana por su sencillez en su cálculo, será en consecuencia, recomendable también utilizarla cuando las variables sean homogéneas y estén medidas en unidades similares y/o cuando se desconozca la matriz de varianzas.

### 2.5 Mapas cognitivos difusos

Los mapas cognitivos [2] y [3] representan el conocimiento causal sin indicar el grado de fortaleza de la relación causa-efecto. Definen y simulan la dinámica de sistemas por medio de conocimiento cualitativo [27].

Los Mapas Cognitivos Difusos MCD, son una técnica introducida por [28] como una extensión de los mapas cognitivos [20]. Los MCD mejoran los mapas cognitivos describiendo la fortaleza de la relación mediante el empleo de valores borrosos en el intervalo  $[-1,1]$ . Los nodos son conceptos causales y pueden modelar eventos, acciones, valores, metas o procesos [29]. Constituyen una estructura de grafo difuso con retroalimentación para representar causalidad. Combinan herramientas teóricas de los mapas cognitivos, la lógica difusa, las redes neuronales, las redes semánticas, los sistemas expertos, y los sistemas dinámicos no lineales [4].

Un MCD elemental puede ser representado a través de un dígrafo, en el cual los nodos representan conceptos y los arcos indican su relación causal [27]. La matriz de adyacencia es obtenida a partir de estos valores asignados a los arcos, tal como se muestra en la Fig. 2.

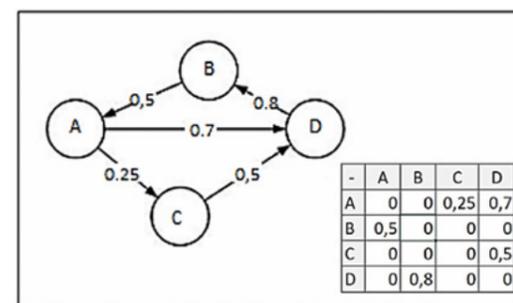


Fig. 2. Mapa cognitivo difuso y su correspondiente matriz de adyacencia.

Entre los elementos que permiten una representación más realista del conocimiento se encuentra la posibilidad de representar retroalimentación, la influencia del tiempo, la vaguedad y la ambigüedad. Presentan una

mayor usabilidad para obtener conocimiento de los expertos. Actualmente no existe una amplia disponibilidad de herramientas tanto comerciales como libres que den soporte a los MCD.

En los MCD existen tres posibles tipos de relaciones causales entre conceptos:

- Causalidad positiva ( $W_{ij} > 0$ ): Indica una causalidad positiva entre los conceptos  $C_i$  y  $C_j$ , es decir, el incremento/disminución en el valor de  $C_i$  lleva al incremento/disminución en el valor de  $C_j$ .
- Causalidad negativa ( $W_{ij} < 0$ ): Indica una causalidad negativa entre los conceptos  $C_i$  y  $C_j$ , es decir, el incremento/disminución en el valor de  $C_i$  lleva a la disminución/incremento en el valor de  $C_j$ .
- No existencia de relaciones ( $W_{ij} = 0$ ): Indica la no existencia de relación causal entre  $C_i$  y  $C_j$ .

Los valores de los conceptos son calculados en cada paso de la simulación, determinando la influencia de los conceptos interconectados al concepto específico de acuerdo a la siguiente regla:

$$A_i^{(t+1)} = f(A_i^{(t)} + \sum_{j=1, j \neq i}^n A_j^{(t)} \cdot W_{ji}) \quad (4)$$

Donde  $A_i^{(t+1)}$  es el valor del concepto  $C_i$  en el paso  $(t+1)$  de la simulación,  $A_i^{(t)}$  y  $A_j^{(t)}$  son los valores del concepto  $C_i$  y  $C_j$ , respectivamente en el paso  $t$  de la simulación,  $W_{ji}$  es el peso de la conexión que va del concepto  $C_j$  al concepto  $C_i$ , y  $f(x)$  es la función de activación [30].

La causalidad juega un papel fundamental en la cognición humana especialmente en la toma de decisiones [31]. La definición sistémica plantea la causalidad como las interacciones que involucran un conjunto de elementos (las entradas o causas) afectando el estado futuro de otros (con las salidas o efectos) [29].

Para definir una relación como causal esta debe cumplir tres condiciones [12]:

- Asimetría: la causa siempre precede al efecto.
- Linealidad: la causa es seguida por un efecto.
- Transitividad: si A causa B y B causa C, entonces A causa C.

Los enlaces entre causa y efecto son frecuentemente imprecisos, imperfectos y con presencia de incertidumbre por naturaleza [31]

[32]. Para considerar la causalidad imperfecta desde un punto de vista computacional, se requiere la obtención de modelos causales imprecisos.

Surge entonces la necesidad de plantear la causalidad en términos de lógica difusa ofreciendo esta un marco adecuado para manejar casos de causalidad imperfecta. La teoría de los conjuntos difusos o borrosos fue introducida por Zadeh [33]. Está basada en la teoría clásica de conjuntos, añadiendo una función de pertenencia [34].

Una función de pertenencia o inclusión  $\mu_A(t)$  indica el grado  $\mu$  en que la variable  $t$  está incluida en el concepto representado por la etiqueta  $A$  [35]. Para la definición de estas funciones de pertenencia se utilizan convenientemente ciertas familias, por coincidir con el significado lingüístico de las etiquetas más utilizadas. Las más frecuentes son: la triangular, trapezoidal y gaussiana Fig. 3.

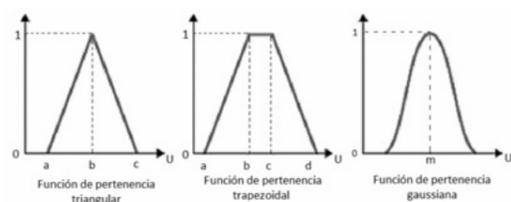


Fig. 2. Mapa cognitivo difuso y su correspondiente matriz de adyacencia.

Los nodos de los MCD están representados por los conceptos a modelar y los arcos o aristas, determinan las relaciones causales entre estos. Se emplea el término granularidad del conjunto de términos lingüísticos para expresar el grado o nivel de causalidad, o intensidad de conexión entre los nodos [36].

La granularidad presente en las relaciones causales [37] hace adecuada su representación mediante variables lingüísticas [38, 39] El empleo de estas variables lingüísticas implica la necesidad de realizar procesos de computación con palabras [40].

La computación con palabras es una metodología que permite realizar un proceso de razonamiento utilizando palabras pertenecientes a un lenguaje en lugar de números. Dicha metodología permite crear y enriquecer modelos de decisión en los cuales la información vaga e imprecisa [41] es modelada a través de variables lingüísticas.

El modelo de representación lingüística de 2-tuplas, permite realizar procesos de

computación con palabras sin pérdida de información, basándose en el concepto de traslación simbólica. En los cinco últimos años, se propuso un modelo para combinar la computación con palabras mediante el modelo de representación de las 2-tuplas lingüísticas y los mapas cognitivos difusos [42].

Debido a esta característica y a su relativa facilidad de uso, el autor considera este modelo como el más adecuado para la computación con palabras en MCD, especialmente cuando estos se obtienen a partir de juicio de expertos.

Sea  $S=\{s_0, s_1, \dots, s_g\}$  un conjunto de términos lingüísticos y  $\beta \in [0, g]$  un valor en el intervalo de granularidad de  $S$ .

Definición 1: [43] La Traslación Simbólica de  $\beta$  en un término lingüístico  $S=(s_i)$ , es un número  $\alpha$  valorado en el intervalo  $[-.5, .5)$  que expresa la diferencia de información entre una cantidad de información expresada por el valor  $\beta \in [0, g]$ , obtenido en una operación simbólica y el valor entero más próximo, donde  $i \in \{0, \dots, g\}$  que indica el índice de la etiqueta lingüística  $s_i$  más cercana en  $S$ .

A partir de este concepto se desarrolló un nuevo modelo de representación de la información lingüística el cual hace uso de un par de valores, también denominado 2-tuplas. Este modelo de representación define un conjunto de funciones que facilitan las operaciones sobre 2-tuplas.

Definición 2: [43] Sea  $S=\{s_0, s_1, \dots, s_g\}$  un conjunto de términos lingüísticos y  $\beta \in [0, g]$  un valor que representa el resultado de una operación simbólica, entonces la 2-tupla lingüística que expresa la información equivalente a  $\beta$ , se obtiene usando la siguiente función:

$$\Delta: [0, g] \rightarrow S \times [-.5, .5) \quad (5)$$

$$\Delta(\beta) = (s_i, \alpha), \begin{cases} s_i, & i = \text{round}(\beta) \\ \alpha = \beta - i, & \alpha \in [-.5, .5) \end{cases} \quad (6)$$

Donde round es el operador usual de redondeo para  $(s_i)$ , que es la etiqueta lingüística con índice  $i$  más cercano a  $\beta$  y  $\alpha$  es el valor de la traslación simbólica.

Siendo  $\langle S \rangle = S \times [-.5, .5)$ , es importante resaltar que la inversa  $\Delta^{-1}: \langle S \rangle \rightarrow [0, g]$  está definida como  $\Delta^{-1}(s_i, \alpha) = i + \alpha$ . De este modo, una 2-tupla lingüística  $\langle S \rangle$  queda identificada con su valor numérico en  $[0, g]$ .

### 3 ANÁLISIS Y DISEÑO DEL SISTEMA DE RECOMENDACIÓN

El modelo se basa en el historial de calificaciones del estudiante. Adicionalmente se toman en cuenta otros aspectos, tales como: asignaturas aprobadas del nivel de estudio en que se encuentre y las asignaturas pendientes de cursar.

Se utiliza un enfoque basado en la toma de decisiones bajo incertidumbre [29] para determinar el nivel de importancia de las asignaturas, considerando la utilización de técnicas de softcomputing [8], específicamente la computación con palabras mediante el modelo de las 2-tuplas lingüísticas evitando la pérdida de información [44].

Se propone un modelo de recomendación y ayuda a la toma de decisiones basado en la relación de los resultados de la integración del algoritmo de cálculo de distancia, la modelación y el análisis de MCD.

Se puede manejar información lingüística y numérica [45] en la modelación del MCD. Se da preferencia al manejo de información numérica, sin embargo, los expertos pueden manejar ambos dominios de información.

A continuación, se muestra la Tabla I con la relación de las etiquetas que se utilizan en el modelado:

TABLA I RELACIÓN DE ETIQUETAS LINGÜÍSTICAS.

Etiqueta	Valor
Muy alta (VH)	1.0
Alta (H)	0.8
Normal (Z)	0.5
Baja (L)	0.4
Muy baja (VL)	0.2

Como se observa, la relación de los conceptos que serán modelados en el MCD se expresará en 5 criterios esencialmente. Desde una relación más débil con un valor de 0.2 hasta una relación total con un valor de 1.0, se utilizan para la representación de las relaciones del mapa. El valor 0 no se incluye porque su

equivalencia es la NO relación entre conceptos.

#### 3.1 Estructura general del modelo

En la Fig. 4, se muestra una representación gráfica del modelo que rige la investigación:

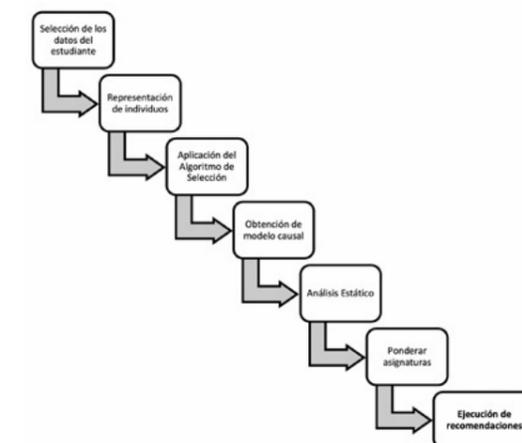


Fig. 4. Etapas del Modelo

#### Proceso de selección de datos de estudiantes

En el modelo propuesto se presenta como información para la investigación, el historial de calificaciones de los estudiantes de la carrera de Ingeniería de Sistemas Computacionales de una universidad de la ciudad de Guayaquil, Ecuador.

Para el estudio se tomarán los datos de los estudiantes de los 2 primeros niveles académicos de la carrera en los períodos lectivos 2017 CI, CII y 2018 CII. A partir de esta información se aplicará la selección de asignaturas y estudiantes, para generar como salida del sistema, una recomendación a un estudiante específico.

Como primer paso para iniciar el proceso se debe seleccionar el estudiante al cual se quiere brindar consejería académica. Para ello, se escogerá un estudiante, en este caso, un estudiante del primer nivel que se denominará estudiante objetivo.

El segundo paso estará planteado por la selección de los estudiantes de la misma

carrera, pero un nivel superior al estudiante objetivo. Serán válidos para el proceso de selección, aquellos estudiantes que obtuvieron calificaciones con cierto grado de similitud, en las asignaturas del nivel académico del estudiante objetivo.

El proceso de selección de los individuos semejantes al estudiante objetivo se obtendrá de la ejecución del algoritmo de distancia euclidiana en función de las calificaciones de los mismos. Como resultado de la ejecución se obtiene el estudiante con mayor similitud al estudiante objetivo.

• **Representación de individuos**

La función de distancia euclidiana, con base en el teorema de Pitágoras como se hace referencia en el capítulo anterior, es el cálculo de la distancia entre dos puntos en un espacio cartesiano.

Los estudiantes seleccionados serán representados en una matriz bidimensional de estudiantes x calificaciones. Adicionalmente en la matriz se incluye el estudiante objetivo con sus calificaciones, actuando como centro sobre el cual se realizarán los cálculos.

• **Aplicación de algoritmo de selección**

La aplicación del algoritmo se basa en la función  $d(A,B)=\sqrt{(\sum_{i=1}^N(b_i-a_i)^2)}$  en cada una de las calificaciones por asignaturas, donde siempre estará representado por las calificaciones del punto A el estudiante objetivo. El menor promedio de la suma de las distancias desde (A) hasta cada uno de los estudiantes (B) de la matriz representa al individuo con mayor similitud respecto al estudiante objetivo (A). Teniendo este resultado, se recomienda al estudiante (A) el valor aproximado de la calificación por asignatura que puede alcanzar en el próximo nivel de estudio, manteniendo el mismo rigor académico que tiene en la actualidad.

Con el objetivo de representar la relación de las asignaturas de una determinada carrera se diseña la estructura general de un modelo basado en el uso de los MCD como técnica de inferencia causal para fusionar el conocimiento de los expertos y evaluar las distintas alternativas que estos nos ofrecen para su análisis.

• **Obtención de Modelo Causal**

El modelo está soportado en una herramienta informática desarrollada por (Reimar, 2012) denominada FCM-Decision.

A continuación, se presenta el flujo de trabajo del modelo en FCM Fig. 5.

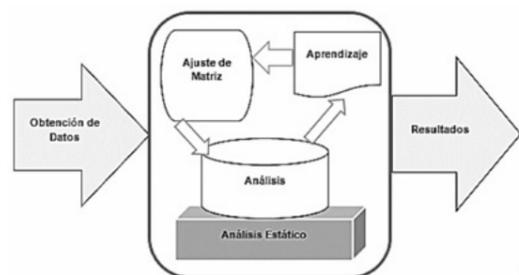


Fig. 5. Flujo de trabajo del modelo.

Las actividades principales representadas en el flujo de trabajo de la figura anterior son: la obtención de datos como pasos iniciales del proceso, el ajuste de la matriz de adyacencia que se obtiene producto de la representación de los datos y la relación que existe entre ellos, el análisis estático y el aprendizaje como paso importante para el ajuste de valores causales por los expertos para la obtención de un resultado más representativo.

El flujo de trabajo a seguir para la obtención del MCD que representará el conocimiento causal que los expertos tienen sobre el dominio, está basado en un conjunto de pasos que permiten la agregación de los mapas individuales. La participación de múltiples expertos ayuda a la comprensión conjunta del problema, así como a la búsqueda de una solución por consenso.

A continuación, se describen los pasos del flujo de trabajo en la Fig. 6:

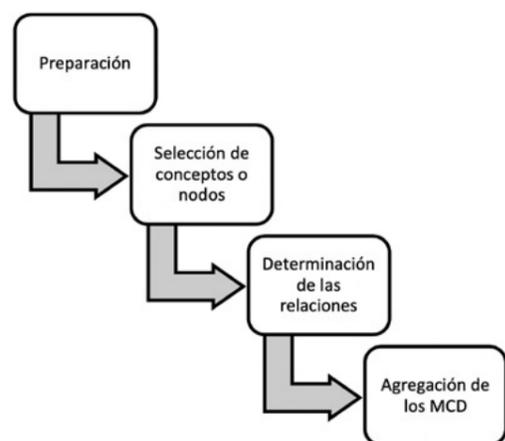


Fig.6. Flujo del trabajo de obtención del modelo causal

— Preparación: Se identifican las fuentes de información que serán analizadas en el estudio. Para la obtención de datos, es aconsejable la participación de varios expertos con diferentes criterios. La información que se modelará, al igual que en el flujo de selección de estudiantes, se extrae del historial de calificaciones de estos en una universidad de la ciudad de Guayaquil. Además, se cuenta con los datos ofrecidos por expertos donde cuantitativamente exponen la relación que existe entre las diferentes asignaturas de la carrera de Ingeniería en Sistemas Computacionales de esta universidad.

— Selección de conceptos o nodos: Son determinados los conceptos que serán representados en el mapa. Cada uno de los conceptos representará una asignatura de la carrera, las cuales deben ser asignaturas del ciclo lectivo que se encuentre cursando el estudiante objetivo.

— Determinación de las relaciones: Se determina el tipo de causalidad existente entre conceptos. La causalidad puede tomar valores negativos y positivos, en dependencia del área de estudio que se esté modelando. En el modelo que se analiza, solo se trabajará con valores positivos, representando la relación que existe entre las asignaturas de un ciclo lectivo.

El peso de la conexión que va del concepto  $C_j$  al concepto  $C_i$  es representado mediante 2-tuplas lingüísticas del siguiente modo:

$$W_{ij} = (s_{ij}^w, a_{ij}^w) = \beta_{ij}^w \quad (7)$$

Dado que el modelo lingüístico basado en 2-tuplas permite realizar una transformación entre una 2-tupla lingüística y un valor numérico en el intervalo de granularidad  $[0,g-1]$  y los MCD trabajan sobre valoraciones numéricas expresadas en el intervalo  $[-1,1]$ , es necesario realizar una transformación de una 2-tupla a un valor numérico en el intervalo  $[-1,1]$  que nos permita trabajar con MCD.

— Agregación de los MCD: Como último paso del presente flujo se realiza la agregación del mapa modelado por los expertos. El análisis de un MCD se puede hacer mediante un análisis estático y/o análisis dinámico. Para el modelo que se presenta solo se realiza el análisis estático, considerando que para la obtención de los resultados es suficiente su aplicación. Mediante la agregación de los mapas individuales el modelo colectivo se hace menos susceptible a creencias potencialmente erróneas de un único experto y sirviendo como base de conocimiento colectivo [46]. La Fig.

7 muestra el flujo de trabajo para la obtención del modelo colectivo:

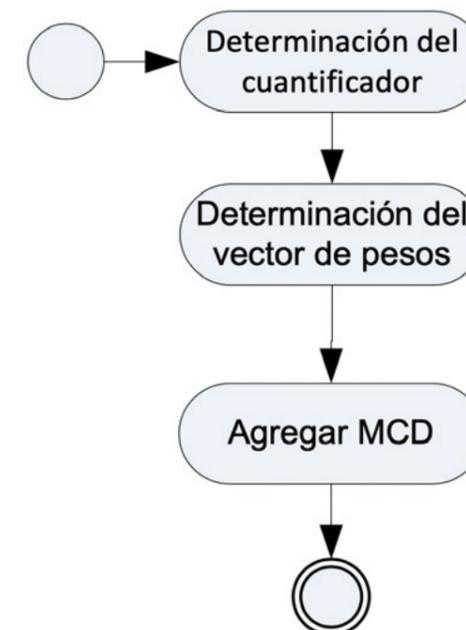


Fig.7. Flujo de trabajo para la agregación de los modelos individuales.

Se propone en este trabajo, el empleo de la familia de operadores OWA y en especial el cálculo del vector de pesos mediante cuantificadores lingüísticos. La Familia de operadores *Ordered Weighted Averaging OWA* -o traducido al español, media ponderada ordenada [47]. Este método unifica los criterios clásicos de decisión con incertidumbre en un solo modelo. Es decir, esta unificación abarca los criterios optimista, el pesimista, el de Laplace y el de Hurwicz en una sola expresión [1].

Este operador puede ser definido de la forma siguiente:

Definición 3: Un operador OWA es una función  $F: \mathbb{R}^n \rightarrow \mathbb{R}$  de dimensión  $n$  si tiene un vector asociado  $W$  de dimensión  $n$  con  $w_j \in [0,1]$  y  $\sum_{j=1}^n w_j = 1$ , de forma tal que:

$$F(a_1, a_2, \dots, a_n) = \sum_{j=1}^n w_j b_j \quad (8)$$

Donde  $b_j$  es el  $j$ -ésimo más grande de los  $a_j$ .

Es decir si  $W = \{w_1 = \frac{6}{11}, w_2 = \frac{2}{11}, w_3 = \frac{3}{11}\}$  y  $A = \{3,4,5\}$ , entonces:

$$F(3,4,5) = \sum_{j=1}^n w_j b_j = (\frac{6}{11})(5) + (\frac{2}{11})(4) + (\frac{3}{11})(3) = \frac{47}{11}$$

El vector de pesos  $w_j$  asociado al operador OWA puede ser determinado de modo funcional [48]. Por otra parte, este vector es empleado para indicar el nivel de compensación entre los criterios y el nivel de optimismo del decisor [23]. Dentro de este enfoque se destaca el empleo de los cuantificadores regulares no decrecientes  $Q$  [49]. Esto evita la necesidad de determinar parámetros externos al modelo, aunque en principio el modelo es flexible tanto para admitir otras formas del cálculo del vector de pesos, como otros operadores de agregación.

A continuación, se describe el flujo de trabajo de agregación:

- Determinación del cuantificador: Los cuantificadores lingüísticos difusos son empleados para representar cuantificadores del quehacer diario como por ejemplo: casi todos, muchos, al menos la mitad, entre otros. [49]. El primer paso consiste en la selección del cuantificador a emplear.
- Determinación del vector de pesos. En este caso se sugiere que el vector de pesos asociado al operador OWA sea determinada por un cuantificador regular no decreciente  $Q$  [48].

Un Cuantificador  $Q$  es un conjunto difuso  $[0,1]$  donde dada una proporción  $r \in [0,1]$ ,  $Q(r)$  representa el grado en que  $r$  representa la semántica definida en  $Q$ , con regla de correspondencia:

$$Q(r) = \begin{cases} 0 & \text{si } r \leq \alpha \\ \frac{r-\alpha}{\beta-\alpha} & \text{si } \alpha < r \leq \beta \\ 1 & \text{si } r > \beta \end{cases} \quad (9)$$

Si  $Q$  es un Cuantificador Incremental Monótono Regular (RIM por sus siglas en inglés) entonces el valor agregado de la alternativa  $x=(a_1, \dots, a_n)$  esta dado por  $F_q(a_1, \dots, a_n)$ , donde  $F_q$  es un operador OWA derivado de  $Q$ . Los cuantificadores RIM pueden ser empleados para generar los pesos del = operador OWA utilizando la siguiente expresión,

$$w_i Q\left(\frac{i}{n}\right) - Q\left(\frac{i-1}{n}\right) \quad (10)$$

— Agregar MCD: En caso de participar más de un experto y de obtenerse los modelos de manera independiente, se realiza la agregación. Esta actividad consiste en la fusión de los MCD individuales obtenidos.

• **Análisis estático**

Para determinar la importancia y selección de los nodos representados en un MCD se utiliza el análisis estático, como se muestra en la Fig. 8.



Fig. 8. Análisis estático.

- Selección de las medidas: Se selecciona el aspecto del MCD o la combinación que se desea analizar. En el presente estudio, se decidió establecer el nivel de fuerza de la conexión entre los nodos (asignaturas), para determinar su importancia dentro del grafo.
- Cálculo de centralidad: Se calculan los valores de entrada y salida de los nodos (Indegree y Outdegree, respectivamente) para determinar la centralidad. De emplearse más de una medida de centralidad se determina un valor compuesto de centralidad mediante la agregación de nuevos valores.

En este trabajo, como propuesta inicial, se propone la transformación del valor numérico equivalente de una 2-tupla  $\beta$  a un valor numérico en el intervalo  $[-1,1]$  del siguiente modo:

$$\gamma: [0, g - 1] \rightarrow [-1,1] \quad (11)$$

$$\gamma(v) = \frac{2v}{g-1} - 1 \quad (12)$$

$$\gamma^{-1}: [-1,1] \rightarrow [0, g - 1] \quad (13)$$

$$\gamma^{-1}(v) = \frac{(v+1)(g-1)}{2} \quad (14)$$

Donde  $g$  es la granularidad del conjunto de términos  $S$ .

Las funciones anteriores junto a las funciones asociadas con la 2-tupla, permiten la representación de relaciones causales tanto positiva como negativa y la agregación de estos valores para la realización del proceso de

inferencia causal utilizando las 2-tuplas.

- Ordenamiento: En esta actividad se ordenan los nodos de acuerdo a su importancia en el modelo. Se representa el MCD como un grafo dirigido ponderado  $(V,E)$ , donde  $V$  es el conjunto de nodos y  $E$  es el conjunto de conexiones entre esos nodos.

Para priorizar los nodos más importantes se determina la centralidad del factor  $(C_i)$  a partir de su Outdegree  $(od_i)$  e Indegree  $(id_i)$ , teniendo en cuenta la magnitud de los pesos  $c_{ij}$  de la siguiente manera:

• *Outdegree*:  $od(v_j)$  es la suma de las filas en la matriz de adyacencia y refleja la fortaleza de las relaciones  $(c_{ij})$  salientes de la variable.

$$od(v_j) = \sum_{i=1}^N c_{ij} \quad (15)$$

• *Indegree*:  $id(v_j)$  es la suma de las columnas y refleja la intensidad de las relaciones  $(c_{ji})$  entrantes de la variable.

$$id(v_j) = \sum_{i=1}^N c_{ji} \quad (16)$$

• Grado de Centralidad  $(C_i)$  se calcula a partir de la suma de su grado de entrada  $(id_i)$  y grado de salida  $(od_i)$ , tal como se expresa en la fórmula siguiente:

$$C_i = id_i + od_i \quad (17)$$

La centralidad en un MCD indica que tan fuertemente está relacionado un nodo con otros, a partir de sus conexiones directas. Los nodos se clasifican de acuerdo con las siguientes reglas:

- Las variables transmisoras tienen *Outdegree* positivo o indeterminado, y cero *Indegree*.
- Las variables receptoras en cambio, tienen una *Indegree* positivo o indeterminado, y cero *Outdegree*.
- Las variables ordinarias: Tienen un grado de *Indegree* y *Outdegree* distinto de cero.

Los Mapas de Decisión Difusa (FDM por sus siglas en inglés) se propusieron en el

año 2006 para resolver el problema de toma de decisión multicriterio teniendo en cuenta la dependencia y la retroalimentación [49]. Para superar la deficiencia en los métodos AHP y ANP. El FDM incorpora el método de valores propios (eigenvalues), y la ecuación de ponderación [50].

Las ventajas de FDM [50, 51] es superar los inconvenientes de los métodos: Analytic Hierarchy Process AHP y el Analytic Network Process ANP, empleando las diferentes funciones de umbral para indicar los diversos tipos de relación entre los criterios. La capacidad para resolver tanto el compuesto como los efectos de interacción y tratar con influencias directas e indirectas es otra área de ventaja.

Los FDM están basados en los mapas cognitivos difusos (MCD) e indican el grado de influencia entre los criterios [52]. Siendo además una opción adecuada para el modelado de interdependencia [53].

La importancia entre los nodos se compara para derivar el vector de ponderación local utilizando el enfoque de valor propio [50]. Esto se debe hacer por los expertos de dominio de acuerdo con la escala de preferencia como se muestra en la Tabla II.

TABLA II. ESCALA PARA LA RELACIÓN ENTRE LOS CRITERIOS.

Descripción	Valor Numérico
Igual Importancia	1
Importancia Moderada	2
Importancia Fuerte	3
Importancia Muy Fuerte	4
Importancia Extrema	5

Se desarrolló un mapa cognitivo difuso para indicar la influencia entre los criterios por el experto. Interdependencias causales son modeladas de esta manera. Esta etapa consiste en la formación de FDM de sub-factores, como nodos según los expertos.

El cálculo de la ecuación de actualización para obtener la matriz de estado estacionario es la siguiente:

$$C^{(t+1)} = f(C^{(t)} \cdot E), C^0 = I_{n \times n} \quad (18)$$

Donde  $I_{n \times n}$  denota la matriz de identidad,  $E=[W_{ij}]$  es una matriz de  $n \times n$ , que reúne los valores de peso entre los conceptos  $C_i$  y  $C_j$ ,  $C^{(t+1)}$  y  $C^{(t)}$  son las matrices de estado en iteraciones  $(t+1)$  y  $(t)$ , respectivamente,  $C^0$  es la matriz inicial, y  $f$  es la función de transformación de umbral.

Se debe normalizar primero el vector de Peso Local  $V$  y la matriz de estado estacionario  $M$ , como sigue:

$$V_n = \frac{1}{k}V, M_n = \frac{1}{c}M \quad (19)$$

Donde  $V_n$  es el vector normalizado de peso local, y  $M_n$  es la matriz normalizada de estado estacionario. Por último, se normaliza el peso global de  $W$

Donde  $k$  es el elemento más grande de  $V$  y  $c$  es la mayor suma de la fila de  $M$ . Entonces, el vector de peso global de  $W$ , puede ser calculado como sigue:

## 4 VALIDACIÓN DEL SISTEMA DE RECOMENDACIÓN BASADO EN MAPAS COGNITIVOS DIFUSOS

La aplicación del proceso de ayuda a la toma de decisiones se torna compleja cuando se tiene un conjunto de datos imprecisos y no se cuenta con herramientas que automaticen el proceso. En este capítulo se presenta un caso de estudio basado en los récords estudiantiles de la carrera de Ingeniería en Sistemas Computacionales de la Facultad de Ciencias

$$W = V_n + V_n M_n \quad (20)$$

### • Ponderar asignaturas

En esta actividad se filtran los cursos de acuerdo al perfil del usuario para encontrar cuáles son las más adecuadas para este.

$$NP_i = C_i N_i \quad (21)$$

Donde  $NP_i$  representa la nota ponderada,  $C_i$  el valor del peso obtenido por el método FDM y  $N_i$  la nota promedio de los  $N$  vecinos más cercanos

### • Ejecutar recomendaciones

Una vez calculada la ponderación de las notas de las asignaturas y el valor de centralidad, representado por el siguiente vector:

$$NP_i = (np_i, \dots, np_n) \quad (22)$$

Los mejores serán aquellos, que mejor satisfagan las necesidades del perfil del usuario.

Matemáticas y Físicas de una universidad de la ciudad de Guayaquil, la implementación de los sistemas informáticos para el cálculo de distancia entre estudiantes, el FCM-Decision, como una de las herramientas utilizadas para el modelado y análisis de Mapas Cognitivos Difusos (MCD).

### 4.1 Aplicaciones informáticas utilizadas.

Para el estudio se desarrolla un pequeño sistema que nos permite realizar la selección y cálculo de métricas necesarias para obtener los resultados del modelo. La pantalla del sistema para el cálculo de distancias en la Fig. 9, es una aplicación de escritorio desarrollada en Visual Studio .Net C# sobre una base de datos en Sqlite donde se guardan todos los registros de los estudiantes.

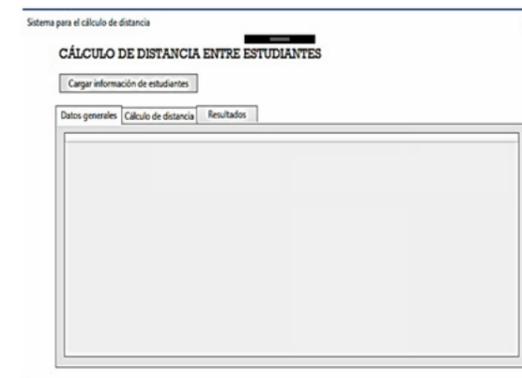


Fig.9. Pantalla del Sistema Informático desarrollado para el cálculo de distancia

El diseño del MCD donde se modelan las asignaturas se realiza con la herramienta FCM-Decision [54]. El desarrollo de la aplicación fue llevado a cabo utilizando la metodología de desarrollo RUP, UML para el modelado del sistema, como lenguaje de programación Java y XML como lenguaje de etiquetado para el tratamiento de ficheros.

La aplicación consta de un panel central donde se encuentra el área de trabajo para la obtención visual del MCD. A la izquierda se muestra la paleta con todos los componentes que brinda la aplicación para el diseño y a la derecha se representa de forma jerárquica los nodos que se encuentran insertados en el área de trabajo. En la parte superior se encuentra ubicada la barra de herramientas la cual contiene funcionalidades básicas de la aplicación. En la Fig. 10, se muestra la vista de las principales funcionalidades y del análisis estático del sistema.

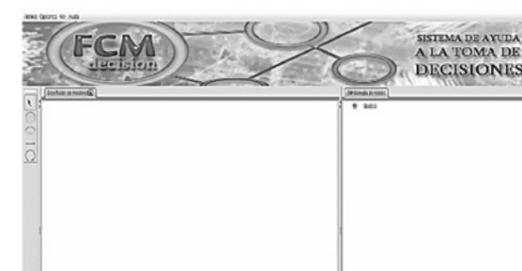


Fig.10. Vista principal del sistema.

El software implementado ha tenido dos alternativas de aplicación: desarrollar sistemas basados en el conocimiento, para resolver problemas concretos, permitiendo simplificar el proceso de ingeniería del conocimiento causal. También ha sido utilizada como medio de enseñanza para el desarrollo de habilidades prácticas en la construcción de sistemas inteligentes y de la ingeniería de software.

### 4.2 Caso de estudio

El caso de estudio para el análisis se divide en dos partes esenciales. En una primera parte se realiza la selección de los estudiantes con evaluaciones similares (perfiles semejantes) al estudiante que se desea ayudar a tomar una decisión y/o mostrar las alternativas de estudio - asignaturas recomendadas - para mejorar sus resultados académicos. En la segunda parte del estudio se realiza el modelado de la relación que existe entre las diferentes asignaturas que conforman el nivel a cursar por el estudiante en cuestión, así como determinar el grado de importancia que presenta cada una de ellas.

### 4.3 Validación de la propuesta.

Se procede a validar el modelo propuesto atendiendo a las etapas enunciadas en el epígrafe A, de la Sección III:

Para la selección de los estudiantes que serán analizados respecto a un estudiante objetivo, se debe tener en cuenta que se tiene que trabajar con dos ciclos consecutivos de cualquiera de las carreras, en dicho caso, de la facultad y carrera del estudiante a dar una recomendación.

Se cuenta con el registro de calificaciones de los estudiantes de la carrera de Ingeniería de Sistemas Computacionales de una universidad de la ciudad de Guayaquil. En el caso de estudio se tomará el registro de notas de los estudiantes del ciclo 2016-2017 CII del nivel 1 y ciclo 2015-2016 CII del nivel 1. Del primer ciclo que se hace referencia, se seleccionan los estudiantes a los cuales se quiere dar una recomendación de estudio referente al siguiente nivel a cursar y en el segundo ciclo se aplica el cálculo de la distancia con aquellos estudiantes con similares calificaciones respecto al estudiante seleccionado.

El algoritmo se divide en varios pasos descritos a continuación:

Se selecciona un estudiante A del primer ciclo que consta con las notas que en la Tabla III, se describen:

TABLA III. RELACIÓN DE NOTAS DEL ESTUDIANTE A.

Asignatura	Nota
Matemáticas I	9 puntos
Matemáticas Discretas	7 puntos
Programación I	7 puntos
Introducción Computacional	7 puntos
Circuitos Eléctricos	7 puntos
Idioma Español	9 puntos

Se seleccionan los estudiantes del segundo ciclo referido para el estudio y en el mismo nivel del estudiante A. Con la relación seleccionada se construye una matriz bidimensional M(x,y), donde y representa las asignaturas del ciclo y el x la lista de estudiantes. El valor M(x,y) contiene la nota de cada estudiante, respecto a la asignatura en la posición (x,y) de acuerdo a la Tabla IV, que sigue:

TABLA III. RELACIÓN DE NOTAS DEL ESTUDIANTE A.

Estudiante	6.0	7.0	8.0	9.0	10.0
Estudiante 1	6.0	7.0	8.0	9.0	10.0
Estudiante 2	10.0	10.0	9.0	9.0	10.0
Estudiante 3	7.0	8.0	7.0	8.0	9.0
Estudiante 4	8.0	8.0	6.0	10.0	9.0
Estudiante N	8.0	7.0	7.0	8.0	9.0

Teniendo seleccionado el escenario, se procede a la ejecución del algoritmo de cálculo de distancia. Para ello se incluye como registro adicional a la matriz M de calificaciones del estudiante A. El cálculo se realiza tomando como referencia el vector de notas del estudiante A y la distancia que existe entre cada una de las calificaciones de los estudiantes representados en la matriz.

Haciendo referencia a la ecuación (3), se propone una forma simplificada:  $d(X,Y) = \sqrt{(\sum_{i=1}^N (y_i - x_i)^2)}$ , donde X representa el estudiante a recomendar y Y, al resto representados en la matriz M.

En cada iteración del algoritmo son normalizados los valores de las notas de los estudiantes. Se divide el valor de la nota por el mayor valor representado en la matriz, o lo que es lo mismo, por el valor de la nota máxima que se puede alcanzar que es 10.0 puntos.

Realizado el cálculo se seleccionarán aquellos estudiantes donde el valor de distancia obtenido se encuentre dentro del 85% de similitud respecto al estudiante A. Como resultado se tienen 36 registros dentro del rango de similitud especificado.

Los estudiantes comprendidos en los resultados del cálculo de la distancia, se toman como muestra para determinar el promedio de

notas a alcanzar en el próximo ciclo a cursar por el estudiante A. Este promedio que se obtiene de las asignaturas del presente ciclo terminado por los estudiantes resultantes representa la calificación aproximada que puede alcanzar el estudiante A si mantiene un sistema de estudio similar al del curso terminado.

La relación de asignaturas a cursar el próximo ciclo, así como las notas estimadas que el estudiante puede alcanzar manteniendo un comportamiento estable respecto al ciclo concluido y en referencia a la similitud que presenta con los 36 estudiantes que se obtuvieron como resultado de la ejecución del algoritmo, se evidencia en la Tabla V.

TABLA V. VALORES APROXIMADOS DE NOTAS DE PRÓXIMO CICLO.

Asignatura	Nota Estimada
Matemática II	6.56 puntos
Álgebra Lineal	7.03 puntos
Programación II	5.93 puntos
Circuitos electrónicos	7.61 puntos
Contabilidad general	7.08 puntos
Comunicación oral y escrita	7.72 puntos

#### 4.4 Mapa Cognitivo Difuso de Asignaturas

Haciendo uso de la herramienta FCM-Decision, se registran los nodos y relaciones que serán representados en el mapa. Cada uno de los nodos representa una asignatura y la relación que exista entre cada una de ellas estará representada mediante arcos con un peso que indica cuan fuerte están relacionadas. Los valores de los arcos están comprendidos en el siguiente conjunto de términos lingüísticos, véase Tabla IV, donde a mayor valor, mayor es la relación que existe entre nodos que se relaciona.

TABLA VI. CONJUNTO DE TÉRMINOS LINGÜÍSTICOS.

No	Etiqueta	Número Difuso Triangular
s <sub>0</sub>	Negativamente Muy Alto (NVVH)	(-1,-1,-0.66)
s <sub>1</sub>	Negativamente Moderado (NH)	(-1,-0.667,-0.333)
s <sub>2</sub>	Ligeramente Negativo (NL)	(-0.667,-0.333, 0)
s <sub>3</sub>	Ninguno (Z)	(-0.333,0,0.333)
s <sub>4</sub>	Ligeramente Positivo (PL)	(0,0,0.333,0.667)
s <sub>5</sub>	Positivamente Moderado (PH)	(0.333,0.667,1)
s <sub>6</sub>	Positivamente Muy Alto (PVVH)	(0.667,1,1)

Para el caso de estudio se toman las asignaturas del 2do.nivel de la Carrera de Ingeniería en Sistemas Computacionales y haciendo uso del criterio de expertos se establece la influencia que tiene cada una de ellas sobre el resto.

Se asignan los valores de forma lingüística y luego se obtiene el valor numérico correspondiente. Por ejemplo, para el caso de la etiqueta positivamente muy alto, mediante la aplicación de (9), se tiene:

$$\gamma(6) = \frac{2 * 6}{7 - 1} - 1 = 2 - 1 = 1$$

TABLA VII. NODOS DEFINIDOS PARA DETERMINAR LA IMPORTANCIA DE LAS ASIGNATURAS.

Nodos	Descripción
N-1	Matemática II
N-2	Álgebra Lineal
N-3	Programación II
N-4	Circuitos Electrónicos
N-5	Comunicación Oral y Escrita
N-6	Contabilidad General

Partiendo del registro de nodos en la herramienta, en la Fig. 11, se muestra el diseño automático del mapa que se analizará y la matriz de adyacencia resultante sobre la cual se aplica el análisis para la obtención de las alternativas de solución.

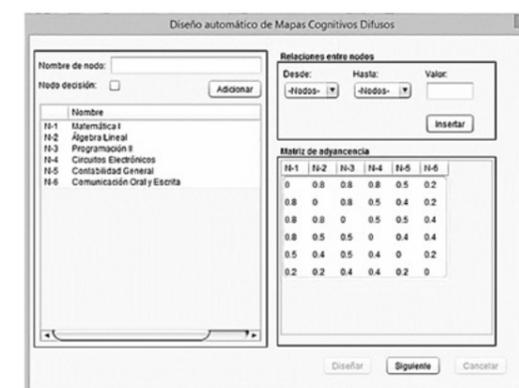


Fig. 11. Diseño automático del Mapa

Con los valores registrados se diseña el mapa, obteniendo un grafo resultante de las asignaturas y la intensidad de influencia de una asignatura sobre otra, como sigue en la Fig. 12:

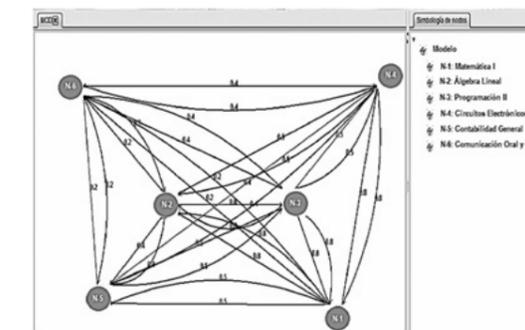


Fig. 12. Diseño automático de Grafo MCD

Una vez diseñado el MCD, donde los nodos son las asignaturas y las relaciones representan la influencia que tiene cada una sobre las demás, se procede a realizar el análisis estático del MCD.

Como sigue en la Fig. 13, se muestran los resultados luego de ejecutar el análisis estático del mapa:

Factores	Outdegree	Indegree	Sumatoria	Centralidad
Matemática I	3.1	3.1	6.2	0.209
Álgebra Lineal	2.7	2.7	5.4	0.182
Programación II	3.0	3.0	6.0	0.203
Circuitos Electrónicos	2.6	2.6	5.2	0.176
Contabilidad General	2.0	2.0	4.0	0.135
Comunicación Oral y Escrita	1.4	1.4	2.8	0.095

Fig. 13. Resultados de ejecución

Los valores de centralidad calculados en la simulación del análisis estático del MCD representa el grado de influencia que tiene cada nodo en el espacio de estudio. Estos resultados se ordenan descendientemente y se obtienen los conceptos:  $N1 > N3 > N2 > N4 > N5 > N6$ . Esta sucesión representa la importancia que estas asignaturas tienen en el diseño.

A continuación, se obtiene la ponderación de cada una de las asignaturas.

Para la elección de la mejor alternativa, se debe derivar los puntajes de influencia de

cada criterio y calcular las puntuaciones de influencia de cada factor.

La matriz de puntuación (E) se obtuvo como sigue:

$$\begin{bmatrix} [1. & 0.25 & 0.2 & 0.2 & 0.33 & 3. & ] \\ [4. & 1. & 1. & 1. & 2. & 1. & ] \\ [5. & 1. & 1. & 1. & 4. & 2. & ] \\ [5. & 1. & 1. & 1. & 2. & 3. & ] \\ [3. & 0.5 & 0.25 & 0.5 & 1. & 3. & ] \\ [0.33 & 1. & 0.33 & 0.33 & 0.33 & 1. & ] \end{bmatrix}$$

Desde la matriz anterior, se pueden calcular los pesos locales por medio del método de valor propio Tabla VIII.

TABLA VIII. PESOS LOCALES

Nodo	Pesos locales
N1	0.11866388
N2	0.30634986
N3	0.39097931
N4	0.11866388
N5	0.19835631
N6	0.1272004

Matriz de adyacencia de la FCM se muestra a continuación:

$$\begin{bmatrix} [0. & 0.8 & 0.8 & 0.8 & 0.5 & 0.2] \\ [0. & 0.8 & 0. & 0.5 & 0.4 & 0.2] \\ [0.8 & 0.8 & 0. & 0.5 & 0.5 & 0.4] \\ [0.8 & 0.5 & 0.5 & 0. & 0.4 & 0.4] \\ [0.5 & 0.4 & 0.5 & 0.4 & 0. & 0.2] \\ [0.2 & 0.2 & 0.4 & 0.4 & 0.2 & 0. ] \end{bmatrix}$$

Por último, mediante el uso de (15), se pueden obtener los Pesos globales como se muestra en la Tabla IX:

TABLA IX. PESOS GLOBALES

Nodo	Pesos Globales
N1	0.14164641
N2	0.16535974
N3	0.23737714
N4	0.21740704
N5	0.14194677
N6	0.0962629

### 4.5 Resultados Experimentales del estudio.

Realizados los flujos de selección de estudiantes según similitud en cuanto a calificaciones alcanzadas y el análisis de importancia de las asignaturas haciendo uso de Mapas Cognitivos Difusos se obtiene la siguiente tabla de relación:

TABLA X. RELACIÓN DE RESULTADOS

Asignatura	Nota Estimada	Valor De Importancia	Nota*Importancia
MATEMÁTICA II	6,56	0,14164641	0,9292
ALGEBRA LINEAL	7,03	0,16535974	1,162479
PROGRAMACION II	5,93	0,23737714	1,407646
CIRCUITOS ELECTRONICOS	7,61	0,21740704	1,654468
CONTABILIDAD GENERAL	7,08	0,14194677	1,004983
COMUNICACION ORAL Y ESCRITA	7,72	0,0962629	0,74315

En la Tabla X, se puede observar cómo influye una asignatura en las calificaciones de los estudiantes, y a su vez, ayuda a tomar una decisión de cómo enfrentar una asignatura con el objetivo de mejorar sus resultados académicos. En el ciclo de estudio escogido para la investigación se tiene que en las asignaturas de mayor rigor es donde los estudiantes presentan más bajas calificaciones, por lo que son estas asignaturas donde ellos deben hacer un mayor esfuerzo con vistas a mejorar sus resultados.

Según el algoritmo se le recomienda cursar las siguientes asignaturas en este orden:

Circuitos electrónicos>programación II>álgebra lineal>contabilidad general>matemática>comunicación oral y Escrita.

### 4.6 Resultados experimentales en la agregación de MCD

Se diseñó un experimento para probar las mejoras en cuanto a interpretabilidad y fiabilidad de los MCD agregados mediante cuantificadores.

#### 1. Datos empleados

En la literatura se reportan distintos estudios de casos donde se integra conocimiento causal de múltiples expertos sin embargo existen pocos datos disponibles para realizar experimentos. Para este trabajo se empleó la base de datos propuesta de Leyva [55] que incluye siete modelos reales obtenidos por el autor y disponibles en: <https://sites.google.com/>

site/mleyvazshomepage/causal-knowledge-datasets. Estos son: factores críticos de éxito de los proyectos de integración de datos (FCE) [56], análisis del caso de negocio para software biomédico [55] (ACN), el modelado de las relaciones entre requisitos no funcionales (RNF), dos modelos relacionados con la Ley de Brook, (Brook1, Brook2) y con los factores que intervienen en la motivación de los ingenieros de software (Mot1), así como de la motivación de los estudiantes de la carrera de Ingeniería en Ciencias Informáticas (Mot2).

#### 2. Variables e indicadores de evaluación

En la presente investigación se ha definido como variable dependiente la efectividad en la recomendación de asignaturas, así como la preservación de las propiedades, tanto estructurales como de comportamiento, de los modelos individuales en el modelo colectivo. Los resultados del análisis estático del MCD agregado con respecto a los MCD de entrada deben ser consistentes [57, 58].

La evaluación del modelo propuesto se realizó teniendo en cuenta aspectos estructurales (densidad y error de la matriz) y de comportamiento (error del estado estable).

En este caso la Densidad del grafo se toma como una medida de la complejidad del modelo e indica la facilidad para interacciones posteriores con el mismo, por parte de los expertos [59]. Esta se define de la siguiente forma:

$$Densidad = \frac{m}{n^2} \quad (23)$$

Donde m es el número de conexiones y n es la cantidad de nodos.

La fiabilidad está relacionada con reflejar la opinión cercana al experto o grupo de expertos. En el aspecto estructural se analiza el Error de Matriz (EM), definido de la siguiente forma [55] [59]:

$$EM = \frac{1}{K \cdot N \cdot N} \sum_{i=1}^N \sum_{j=1}^N \sum_{k=1}^K |e_{ij} - \hat{e}_{ij}^k| \quad (24)$$

Donde  $e_{ij}$  es el valor de la relación existente entre los nodos  $C_i$  y  $C_j$  en el mapa agregado y  $\hat{e}_{ij}^k$  es el peso de la relación existente entre los nodos  $C_i$  y  $C_j$  para el experto k-ésimo. N representa el número de conceptos y K el número de expertos.

### 3. Modelos de agregación empleados en la comparación

La comparación se realiza entre el modelo propuesto guiado por cuantificadores utilizando el cuantificador "muchos" recomendado en el modelo para representar cierto nivel de consenso Fig. 14.

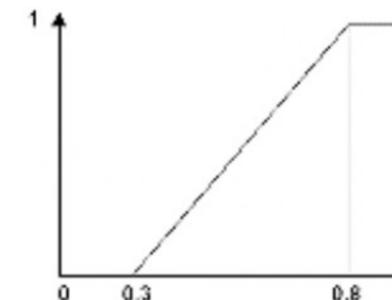


Fig. 14. Cuantificador lingüístico "Muchos"

y la agregación realizada por los siguientes modelos:

- El operador OWA dependiente (D-OWA) (Xu, 2006), siempre que los valores a agregar sean distintos entre sí, sino se devuelve el valor único.
- Media aritmética ( $\mu$ ).
- El OWA centrado (C-OWA). En este caso para el cálculo de los pesos se emplea la propuesta de Yager (Yager, 2007), para  $j \leq \frac{n+1}{2}$  se tiene:

$$w_j = \frac{2(2j-1)}{n^2} \quad (25)$$

#### 4. Análisis de los resultados

Se muestran los datos de los MCD (número de expertos, número de conceptos) y los resultados de la aplicación de los distintos modelos de agregación en la Tabla XI. En negrita se muestran los mejores resultados para cada uno de los casos.

Con la propuesta se obtiene en cuanto a la densidad los mejores resultados 7/7 coincidiendo una vez con la  $Q_{\text{muchos}}$ . En cuanto al EM se obtiene también los mejores en 5/7 casos. En el experimento para la comparación de las muestras se utilizó: Kruskal-Wallis Test o el Mann-Whitney Test según se especifica. En ambos casos para el nivel de significación se aplicó el método de Monte Carlo con intervalos de confianza del 99%. Se considera significativa una significación menor de 0.05.

Durante la aplicación de los test, en los casos que se encontraron diferencias significativas entre los modelos, se formaron diferentes grupos. Los grupos formados cumplen con la siguiente propiedad: los modelos de un mismo grupo no tienen diferencias significativas entre ellos. Los grupos cumplen la siguiente propiedad Grupo1 < Grupo2 <..., con los modelos en los grupos menores reportando los mejores resultados.

TABLA VI. CONJUNTO DE TÉRMINOS LINGÜÍSTICOS.

Modelo	Datos		Operador	Medida	
	Número de conceptos	Número de expertos		Densidad	EM
FCE	10	12	$Q_{muchos}$	0.41	0.231
			D-OWA	0.89	0.251
			$\mu$	0.89	0.255
			C-OWA	0.89	0.238
ACN	8	9	$Q_{muchos}$	0.425	0.291
			D-OWA	0.891	0.298
			$\mu$	0.891	0.302
			C-OWA	0.891	0.290
RNF	10	8	$Q_{muchos}$	0.427	0.239
			D-OWA	0.9	0.245
			$\mu$	0.9	0.250
			C-OWA	0.9	0.238
Brook1	9	8	$Q_{muchos}$	0.154	0.086
			D-OWA	0.481	0.097
			$\mu$	0.481	0.104
			C-OWA	0.481	0.090
Mot1	10	9	$Q_{muchos}$	0.098	0.047
			D-OWA	0.32	0.057
			$\mu$	0.32	0.061
			C-OWA	0.32	0.051
Mot2	7	6	$Q_{muchos}$	0.1344	0.053
			D-OWA	0.388	0.063
			$\mu$	0.388	0.068
			C-OWA	0.388	0.059
Brook2	4	7	$Q_{muchos}$	0.187	0.131
			D-OWA	0.625	0.140
			$\mu$	0.625	0.144
			C-OWA	0.625	0.136

## 5 CONCLUSIONES Y RECOMENDACIONES

### 5.1 CONCLUSIONES

1. En esta investigación se realizó un análisis preliminar de los sistemas de recomendación en general como herramientas capaces de ahorrar tiempo y proporcionar ayuda a la hora de tomar decisiones de diversa índole, para continuar con un estudio más profundo de los Mapas Cognitivos Difusos (MCD) en el campo de la Educación Superior.

Se aplicó el Kruskal-Wallis Test comparando las muestras obtenidas de aplicar los resultados de los modelos respecto a las variables densidad y EM. Como resultado se encontraron diferencias significativas respecto a la variable densidad y no se encontraron diferencias significativas respecto a las variables EM.

Se aplicó Mann-Whitney Test convenientemente para comparar muestras dos a dos determinando finalmente los siguientes grupos respecto a variable EM:

Grupo1:  $Q_{muchos}$

Grupo2: C-OWA, D-OWA,  $\mu$

2. Los MCD son capaces de realizar recomendaciones de asignaturas al estudiante basándose en la importancia que estas representan en los ciclos a los que corresponden, así como de forma general. La representación de asignaturas en un MCD permite recomendar que asignaturas debe el estudiante enfrentar con mayor o menor rigor teniendo en cuenta la importancia que dicha asignatura representa.

- Se resolvió un problema académico para comprobar la efectividad de los MCD y el cálculo de similitud entre estudiantes tomando sus calificaciones como información, permitiendo realizar recomendaciones personalizadas a estudiantes de una universidad de la ciudad de Guayaquil al momento de enfrentar asignaturas con mayores dificultades de aprendizaje o necesidades académicas de refuerzo.
- Para el modelado de los MCD, se trabajó con datos que no provienen de la información previa utilizada en la investigación y que tampoco son recogidos automáticamente por el sistema, sino que son aportados por expertos.
- Haciendo uso de los datos históricos de las calificaciones de alumnos universitarios y su representación en MCD de las asignaturas por ciclo lectivo se desarrolló un sistema para determinar la similitud entre estudiantes de ciclos anteriores respecto a un estudiante objetivo y además se utilizó un sistema para modelar

las asignaturas en un MCD y determinar la importancia que cada una representa, lo que permite orientar al alumnado en la elección de asignaturas y predicción de su desempeño en futuras asignaturas a cursar.

### 5.2 RECOMENDACIONES

- Estudiar el proceso de simulación dinámica de los Mapas Cognitivos Difusos para establecer modelos combinados de valores cuantitativos y cualitativos para obtener mejores resultados.
- Debido a que se trabaja con datos reales de estudiantes de una universidad de la ciudad de Guayaquil, es importante utilizar el mayor número de medidas para garantizar la privacidad de la información.
- Desarrollar un sistema web único integrado para la ejecución del modelo propuesto que permita el acceso al personal docente de la universidad e implicados en el proceso de aprendizaje de la universidad

## REFERENCIAS

- J. Merigó, "New extensions to the OWA operators and its application in decision making," PhD Thesis, Department of Business Administration, University of Barcelona, 2008.
- A.M. Sharif and Z. Irani, "Applying a fuzzy-morphological approach to complexity within management decision making," Emerald Group Publishing Limited. pp. 930-961, 2006.
- M. Glykas and P. Groumos, "Fuzzy Cognitive Maps: Basic Theories and Their Application to Complex Systems," in Fuzzy Cognitive Maps, Berlin: Springer, pp. 1-22, 2010.
- C.W. Ping, "A Methodology for Constructing Causal Knowledge Model from Fuzzy Cognitive Map to Bayesian Belief Network," in Department of Computer Science. Chonnam National University, 2009.
- G. Pajares, J. Sánchez-Lladó and C. López-Martínez, "Fuzzy Cognitive Maps Applied to Synthetic Aperture Radar Image Classifications Advances Concepts for Intelligent Vision Systems," Berlin: Springer, pp. 103-114, 2011.
- J. Carvalho, "Rule Based Fuzzy Cognitive Maps in Humanities Social Sciences and Economics," Berlin: Springer, pp. 289-300, 2012.
- D.k. Iakovidis and E. Papageorgiou, "Intuitionistic Fuzzy Cognitive Maps for Medical Decision Making," Information Technology in Biomedicine, IEEE Transactions on. 15(1): pp. 100-107, 2011.
- L. Curia and A. Lavallo, "Estrategias de decisión en sistemas dinámicos: aplicando mapas cognitivos difusos aplicación a un ejemplo socio-económico," Revista de Gestão da Tecnologia e Sistemas de Informação. 8(3): pp. 663-680, 2011.
- R. Srivastava, M. Buche and T. Roberts, "Belief Function Approach to Evidential Reasoning in Causal Maps," in Causal Mapping for Research in Information Technology, Idea Group Pub., 2005.
- F. Herrera, S. Alons, F. Chiclana and E. Herrera-Viedma, "Computing with words in decision making: foundations, trends and prospects," Fuzzy Optimization and Decision Making. 8(4), pp. 337-364, 2009.

- [11] M. Doumpos and C. Zopounidis, "Preference disaggregation and statistical learning for multicriteria decision support: A review," *European Journal of Operational Research*. 209(3): pp. 203-214, 2010.
- [12] G.F Barberis and M.C.E. Ródenas, "La Ayuda a la Decisión Multicriterio: orígenes, evolución y situación actual," VI Congreso Internacional de Historia de la Estadística y de la Probabilidad. Valencia, 2011.
- [13] G. Adomavicius and A. Tuzhilin, "Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions," *IEEE Trans. Knowl. Data Eng.*, vol. 17, no. 6, pp. 734-749, 2016.
- [14] M.J. Pazzani, "A framework for collaborative, content-based and demographic filtering," *Artif. Intell. Rev.*, vol. 13, no. 5-6, pp. 393-408, 2015.
- [15] M.j. Pazzani and D. Billsus, *Learning and Revising User Profiles: The Identification of Interesting Web Sites*, Kluwer Academic Publishers, pp. 313-331, 1997.
- [16] J.D. Ullman, "Principles of database and knowledge-base systems," vol. 1, Computer Science Press, Inc., pp. 631, 2017.
- [17] R. Burke, K. J. Hammond, and B. C.Young, "Knowledge-based Navigation of Complex Information Spaces," *Proceedings of the Thirteenth National Conference on Artificial Intelligence*, AAAI Press/MIT Press, 2014.
- [18] B. Bezerra and F.D. Carvalho, "A symbolic hybrid approach to face the new user problem in recommender systems," *Ai 2004: Advances in Artificial Intelligence*, *Proceedings, Lecture Notes in Artificial Intelligence 3339*, Berlin: Springer, pp. 1011-1016, 2016.
- [19] T. Murakami, K. Mori and R. Orihara, "Metrics for evaluating the serendipity of recommendation lists," *New Frontiers in Artificial Intelligence*, pp. 40-46, 2008.
- [20] B. Kosko, "Fuzzy cognitive maps," *International Journal of Man-Machine Studies*. 24(1): pp. 65-75, 1986.
- [21] B. Sarwar, G. Karypis, J. Konstan and J. Riedl, "Analysis of recommendation algorithms for e-commerce," *ACM Press*, pp. 158-167, 2000.
- [22] J. Breese, D. Heckerman and C. Kadie, "Empirical Analysis of Predictive Algorithms for Collaborative Filtering," *Proc. Uncertainty in Artificial Intelligence. Proceedings of the Fourteenth Conference*, pp. 18, 2016.
- [23] M. Condliff, D. Madigan, D. Lewis and C. Posse, "Bayesian Mixed-Effects Models for Recommender Systems," *Proceedings of the ACM SIGIR Workshop on Recommender Systems: Algorithms and Evaluation*. 22nd Intl. Conf. on Research and Development in Information Retrieval, 2016.
- [24] L. Ungar and D. Foster, "Clustering Methods for Collaborative Filtering," *Proceedings of the Workshop on Recommendation Systems*, AAAI Press, Menlo Park California, 1998.
- [25] D. Kim and B. Yum, "Collaborative filtering based on iterative principal component analysis," *Expert Syst. Appl.*, vol. 28, no. 4, pp. 823-830, 2015.
- [26] D. Billsus and M. Pazzani, "Learning Collaborative Information Filters," *Proc. Proceedings of the 15th International Conference on Machine Learning*, Morgan Kaufmann Publishers Inc., San Francisco, CA, pp. 46-54, 2017.
- [27] R.M. Axelrod, *Structure of decision: The cognitive maps of political elites*. Princeton University Press Princeton, NJ, 1976.
- [28] J.L Salmeron, *Supporting decision makers with Fuzzy Cognitive Maps*. Industrial Research Institute, Inc. pp. 53-59, 2009.
- [29] J. Salmeron and E. Papageorgiou, "A Fuzzy Grey Cognitive Maps-based Decision Support System for radiotherapy treatment planning," *Knowledge-Based Systems*, pp. 151-160, 2012.
- [30] S. Bueno and J. Salmeron, *Benchmarking main activation functions in fuzzy cognitive maps*. *Expert Systems with Applications*, pp. 5221-5229, 2009.
- [31] C. Puente-Agueda, *Causality in Science*. *Pensamiento Matemático*. pp. 12, 2011.
- [32] R. Burke, *Book Knowledge-based Recommender Systems*, vol. 69, Supplement 32, Marcel Dekker, 2014.
- [33] L. A. Zadeh, "Fuzzy sets. Information and Control," pp.338-353, 1965.
- [34] B. M. Brio and A. S. Molina, *Redes Neuronales y Sistemas Borosos*, 2th ed., Alfaomega, 2001.
- [35] G. J. Klir and B. Yuan, *Fuzzy sets and fuzzy logic*, New Jersey: Prentice Hall, 1995.
- [36] L. Cevallos, M. Leyva, M. Peña, E. Santos

- and A. Guijarro, "The Extended Hierarchical Linguistic Model in Fuzzy Cognitive Maps," *CITI 2016*, Guayaquil: SPRINGER, pp. 39-50, 2016.
- [37] L. J. Mazlack, "Causal modeling approximations in the medical domain," *IEEE International Conference on Fuzzy systems*, June 2011.
- [38] M. R. Berthold and D. J. Hand, *Intelligent Data Analysis: An Introduction*, Springer, 2010.
- [39] M. S. Garcia-Cascales and M. T. Lamata, "Nueva aproximación al método tópsis difuso con etiquetas lingüísticas," *ESTYLF*, Huelva, 2010.
- [40] M. Espinilla-Estévez, "Nuevos modelos de evaluación sensorial con información lingüística," (DEA), Universidad de Jaén, Jaen, 2009.
- [41] F. Herrera, S. Alonso, F. Chiclana and E. Herrera-Viedma, "Computing with words in decision making: foundations, trends and prospects," *Fuzzy Optimization and Decision Making*, pp. 337-364, 2009.
- [42] K. Pérez-Teruel, M. Leyva-Vázquez, M. Espinilla and V. Estrada-Sentí, "Computación con palabras en la toma de decisiones mediante mapas cognitivos difusos," *Revista Cubana de Ciencias Informáticas*, pp. 19-34, 2014.
- [43] F. Herrera and L. Martínez, "A 2-tuple fuzzy linguistic representation model for computing with words Fuzzy Systems," *IEEE Transactions on*, pp. 746-752, 2000.
- [44] A. Altay and G. Kayakutlu, "Fuzzy cognitive mapping in factor elimination: A case study for innovative power and risks," *Procedia Computer Science*, pp. 1111-1119, 2011.
- [45] S. Samarasinghea and G. Strickert, "A New Method for Identifying the Central Nodes in Fuzzy Cognitive Maps using Consensus Centrality Measure," *19th International Congress on Modelling and Simulation*. Perth, Australia, 2011.
- [46] M. Y. Leyva-Vázquez, K. Pérez Teurel, A. Febles Estrada and J. Gulín-González, "Modelo para el análisis de escenarios basado en mapas cognitivos difusos," *Ingeniería y Universidad*, 2013.
- [47] R. Yager, "Quantifier guided aggregation using OWA operators," *International Journal of Intelligent Systems*, pp. 49-73, 1996.
- [48] R. Yager, "Centered OWA operators," in *Soft Computing* pp. 631-639, 2007.
- [49] J. M. Doña-Fernández, "Modelado de los procesos de toma de decisión en entornos sociales mediante operadores de agregación OWA," PhD Thesis, Universidad de Málaga, 2008.
- [50] B. M. Elomda, H. A. Hefny and H. A. Hassan, "An extension of fuzzy decision maps for multi-criteria decision-making," *Egyptian Informatics Journal*, pp. 147-155, 2013.
- [51] G. H. Tzeng, W.H. Chen, R. Yu and M. L. Shih, "Fuzzy decision maps: a generalization of the DEMATEL methods," in *Soft Computing*, pp. 1141-1150, 2010.
- [52] M. Leyva, J. Hechavarría, N. Batista, J. A. Alarcon and O. Gomez, "A framework for PEST analysis based on fuzzy decision maps," *Revista ESPACIOS*, 2018.
- [53] A. Betancourt-Vázquez, K. Pérez-Teruel and M. Leyva-Vázquez, "Modeling and analyzing non-functional requirements interdependencies with neutrosophic logic," *Neutrosophic Sets and Systems*, 2015.
- [54] O. Reimar, M. Leyva and Y. Barroso, "Herramienta para la simulación y análisis de mapas cognitivos difusos," VI Taller de Inteligencia Artificial. UCIENCIA, Habana, Cuba, 2012.
- [55] M. Leyva-Vázquez, "Modelo de Ayuda a la Toma de Decisiones Basado en Mapas Cognitivos Difusos," UCI, La Habana, 2013.
- [56] M. Y. Leyva-Vázquez, R. Rosado-Rosello and A. Febles-Estrada, "Modelado y análisis de los factores críticos de éxito de los proyectos de software mediante mapas cognitivos difusos," in *Ciencias de la Información*, pp. 41-46, 2012.
- [57] W. Stach, "Learning and aggregation of fuzzy cognitive maps-An evolutionary approach," PhD Thesis, University of Alberta, 2011.
- [58] M. Stajdohar and J. Demsar, "Interactive Network Exploration with Orange," *Journal of Statistical Software*, pp. 1-24, 2013.
- [59] W. Stach, W. Pedrycz and L. A. Kurgan, "Learning of fuzzy cognitive maps using density estimate," *Systems, Man, and Cybernetics, Part B: Cybernetics*, pp. 900-912, 2012.

# AUTORES



MILTON RAFAEL  
MARIDUEÑA  
ARROYAVE

Ingeniero en Computación, Máster en Docencia Universitaria e Investigación Educativa, Máster en Investigación Matemática, PhD en Ciencias Pedagógicas y, PhD en Ciencias Técnicas. Docente investigador de la Facultad de Ciencias Matemáticas y Físicas de la Universidad de Guayaquil y en la Facultad de Ciencias Naturales y Matemáticas de la ESPOL. Docente de Posgrado UTEG-UPSE, autor de libros y artículos publicados en revistas indizadas, ha sido Director de Investigación de la UG y Editor Asociado Revista NEUTROSOPHIC COMPUTING and MACHINE LEARNING.

LORENZO JEOVANY  
CEVALLOS TORRES

Ingeniero en Estadística, Máster en Gestión de la productividad y la calidad, Máster en Modelado computacional en Ingeniería, Docente investigador de la Universidad de Guayaquil UG, miembro del grupo de Investigación de Inteligencia Artificial e Información Tecnológica UG, autor de libros de Simulación, Probabilidad & estadística educativa y artículos en revistas indizadas, ha ocupado cargos de Dirección y Subdirección en la carrera de Ingeniería en Sistemas Computacionales así como Coordinador de Investigación en la Facultad de Ciencias Matemáticas y Físicas.



MIGUEL ANGEL  
BOTTO TOBAR

Miguel Botto-Tobar es Ingeniero en Sistemas Computacionales, M.Sc. en Ingeniería de Software. Actualmente, cursa un Doctorado en Ciencias de la Computación en la Universidad Tecnológica de Eindhoven, Países Bajos. Es autor/co-autor de 30 artículos, y libros publicados por Springer Verlag. Es editor de Ecuadorian Journal of Science, Research and Innovation, y editor regional de JOIV: International Journal on Informatics Visualization. Sus áreas de investigación incluyen Ingeniería de Software, Ingeniería de Software Empírica, y Aspectos Sociales en la Ingeniería de Software.