

Predictor Basado en Prototipos Difusos y Clasificación No-supervisada

Anibal Vásquez, Enrique Peláez, Xavier Ochoa

Facultad de Ingeniería en Electricidad y Computación, Centro de Tecnologías de Información
Escuela Superior Politécnica del Litoral, ESPOL
Guayaquil, Ecuador
{anibal.vasquez, epelaez, xavier}@cti.espol.edu.ec

Resumen— La construcción de prototipos difusos es un método que permite describir a los elementos más representativos de un clúster, a través de su tipicidad. Los prototipos, como los datos más representativos de cada clúster, pueden ser usados en un proceso de clasificación como datos de entrenamiento. Estos prototipos y los clusters pueden ser construidos mediante algoritmos de clustering difuso; los clusters representados por los prototipos poseen variables descriptivas y atributos que pueden ser asociados a nuevos datos. El siguiente trabajo propone una arquitectura que utiliza herramientas de clustering y prototipado difuso, para clasificación no-supervisada y predicción a través de la extracción de variables descriptivas. El desarrollo de un caso de estudio permitió validar el modelo de clasificación para predecir el riesgo de falla en el rendimiento académico de estudiantes, basado en su carga semestral y rendimiento académico, en la selección de cursos antes de registrarse, con un porcentaje de certeza significativo.

Palabras Clave—Inteligencia Artificial, Prototipos Difusos, Lógica Difusa, Clasificación No-supervisada, Predicción.

I. INTRODUCCIÓN

La deserción académica es un problema actual en el sistema de educación superior, esto ocurre cuando un estudiante abandona su programa de estudios y en consecuencia no culmina su carrera de manera exitosa, este problema no solo afecta a los estudiantes sino también a las universidades en sí, pues al ofrecer educación incompleta a sus estudiantes, no se usan de forma eficiente los recursos como infraestructura y profesores que imparten las materias [1]. La predicción del rendimiento académico en estudiantes universitarios ha sido objeto de estudio en años recientes, pues es fundamental evaluar y monitorear el desempeño académico en el sistema educativo; muchas metodologías desarrolladas han hecho uso de técnicas de minería de datos e inteligencia computacional, para así dar un pronóstico a los estudiantes antes de tomar determinados cursos en el semestre, este pronóstico es soportado generalmente por datos académicos, socioeconómicos y otras variables demográficas; Bhardwaj y Pal [2] usaron un modelo basado en teoría de Bayes para predecir el rendimiento mediante clasificación; Pandey y Taruna [3] propusieron un modelo de clasificación basado en árboles de decisión para la predicción de rendimiento, además Oladokun et al. [4] y Naser et al. [5] propusieron el uso de redes neuronales para la predicción de rendimiento basada en clasificación supervisada; el uso de estas técnicas generalmente poseen un nivel de precisión mayor al 70%, esto permite ayudar a los estudiantes en el proceso de decisión al tomar materias en un nuevo periodo académico y así contribuir a mejorar su desempeño.

Este trabajo tiene como objetivo proponer un modelo para estimar el riesgo de reprobación de un estudiante en un periodo académico, dicho modelo es soportado por datos académicos históricos, los cuales son usados para establecer similitudes entre los estudiantes, a través de la extracción de las características que describen su desempeño académico y la carga académica que implican las materias que debe cursar, para así, mediante el uso de variables descriptivas, asociar un valor de riesgo de reprobación dependiendo de su similitud con otros y la carga académica que representa en el periodo en el que la tomará.

El modelo propuesto utiliza herramientas de clustering y prototipado difuso para la clasificación no-supervisada y la predicción mediante asociación de variables descriptivas.

El clustering es una técnica de minería de datos que permite realizar sub-agrupaciones denominadas clusters dentro de un conjunto, permitiendo clasificar los elementos de este conjunto en base a una similitud entre ellos; definido formalmente como [6]:

$$C_f = \{x_1, x_2, \dots, x_m\}; x_i \in X, i = 1, 2, \dots, m \quad (1)$$

$$X = \bigcup_{f=1}^m C_f \quad (2)$$

Donde C_f es el cluster subconjunto de X , tal que $C_f \subseteq X$, y $X \subseteq \mathbb{R}^n$.

Existe una generalización de esta técnica para conjuntos difusos, esta generalización establece que cada cluster puede ser representado como un conjunto difuso [7], permitiendo definir una pertenencia parcial de los elementos a un cluster, e incluso pertenecer de forma simultánea a más de un cluster, así cada uno de estos elementos característicos, llamados prototipos, pueden describir a los demás elementos de una categoría.

Este trabajo se encuentra organizado de la siguiente manera: la sección 2 describe la creación de prototipos, y los algoritmos usados para clustering difuso y su validación; la sección 3 describe las variables extraídas en el contexto académico; la sección 4 describe la arquitectura planteada para la estimación de riesgo de reprobación, basada en prototipos difusos y clasificación no-supervisada, mediante técnicas de clustering; la sección 5 presenta los resultados obtenidos, se realiza un análisis de estos y se presenta un contraste con una técnica de minería de datos convencional, como la regresión logística, y finalmente en la sección 6 se describen las conclusiones.

II. CONSTRUCCIÓN DE PROTOTIPOS DIFUSOS MEDIANTE MÉTODOS DE CLUSTERING Y VALIDACIÓN DEL PROCESO

La construcción de prototipos difusos se basa en una noción de tipicidad, la cual es una medida que se asocia a un elemento perteneciente a un determinado cluster, además esta medida permite determinar qué tan bien puede caracterizar o no a los demás elementos del cluster, esta medida se obtiene mediante dos parámetros que describen: qué tan semejante es un elemento a los demás de su propia categoría; y, qué tan diferente o disímil es con respecto a los otros elementos que no pertenecen a su categoría; Rifqi [8] propone la agregación de estos parámetros de semejanza y disimilitud para estimar el grado de tipicidad de un elemento, definido formalmente como:

$$t_{ir} = \begin{cases} \Phi(S_r(x_i), D_r(x_i)), & x_i \in C_r \\ 0, & x_i \notin C_r \end{cases} \quad (3)$$

Donde:

$$S_r(x_i) = \frac{1}{|C_r|} \sum_{y \in C_r} \rho(x_i, y) \quad (4)$$

$$D_r(x_i) = \frac{1}{|X/C_r|} \sum_{z \in X/C_r} \delta(x_i, z) \quad (5)$$

S_r en (4) describe la medida de semejanza de un elemento x_i dentro de la categoría C_r , y esta se define por las distancias ρ entre x_i y los demás elementos y , que pertenecen a la misma categoría C_r .

D_r en (5) describe la medida de disimilitud de un elemento x_i dentro de la categoría C_r , esta se define por las distancias δ entre x_i y los otros elementos y que no pertenecen a su misma categoría C_r .

Las distancias ρ y δ son definidas a través de funciones decrecientes [9], estas son transformaciones lineales que parten de productos escalares o medidas de disimilitud [10], las cuales utilizan distancias normalizadas euclidianas o no-euclidianas para describir la semejanza entre dos elementos; entre las transformaciones lineales usadas, generalmente para describir funciones decrecientes, se encuentran: la transformada de Laplace, la transformada generalizada de Gauss, y funciones sigmoideas. Siendo ρ una distancia que describe semejanza, definida por una función decreciente en el rango de $[0,1]$, por lo tanto es posible definir la distancia δ , que describe disimilitud, mediante el complemento a uno; esto es [11]:

$$\delta = 1 - \rho \quad (6)$$

δ es también una función que se encuentra dentro del rango $[0,1]$.

Una vez determinadas las medidas de tipicidad de cada elemento, el prototipo difuso para la categoría C_r puede ser construido mediante una operación de agregación sobre los elementos que cumplen con una tipicidad mayor a una condición de umbral definida, así el prototipo difuso lo podemos expresar formalmente como [12]:

$$w_r = \Psi(\{x_i/t_{ir} > \tau\}) \quad (7)$$

Donde τ es el umbral para las medidas de tipicidad en los elementos que construirán el prototipo para la categoría C_r .

Existen varios algoritmos de clustering difuso que permiten construir prototipos; Fuzzy C-Means (FCM), basado en un

proceso de media iterativa donde los grados de membresía asociados a cada elemento describen el grado de pertenencia de un elemento a un cluster, los valores de membresía de cada elemento a cada cluster son definidos por la matriz de membresía $U_{n \times c} = [u_{ir}]$, donde cada elemento u_{ir} es definido formalmente como [13]:

$$u_{ir} = \left(\sum_{s=1}^c \left(\frac{d_{is}}{d_{is}^m} \right)^{\frac{2}{m-1}} \right)^{-1} \quad (8)$$

Donde x_i es un elemento del cluster C_r ; d_{ir} es la distancia entre x_i y el prototipo w_r del cluster C_r y definido por la norma A tal que, $d_{ir} = \|x_i - w_r\|_A$; de igual manera d_{is} está definido por la norma A , pero esta describe la distancia entre x_i y el prototipo w_s de cada cluster C_s tal que, $d_{is} = \|x_i - w_s\|_A$; m es el exponente de ponderación, el cual determina cuan difusos son los límites de cada cluster, por la forma de la curvatura de la función exponencial en u_{ir} ; el exponente de ponderación m y el número de clusters c son parámetros definidos por el usuario. FCM puede ser considerado como una optimización de la semejanza ya que mediante el complemento de los valores de membresía se puede determinar la disimilitud asociada a cada elemento.

En los procesos de clustering difuso existen medidas de validación que permiten estimar el número de clusters c que describen un conjunto de datos en varias categorías, de esta forma podemos definir un proceso de categorización no-supervisada [14], estas medidas surgen como una necesidad de la validación de clusters en conjuntos multidimensionales. Una de estas medidas es el coeficiente de validación de Dave [15], a partir del coeficiente de partición de Bezdeck [13], definido formalmente como:

$$V_{FDC} = 1 - \frac{c}{c-1} (1 - V_{FC}) \quad (9)$$

Donde c es el número de clusters a validar y V_{FC} es el coeficiente de partición de Bezdeck, este último definido como:

$$V_{FC} = \frac{1}{n} \sum_{r=1}^c \sum_{i=1}^n (u_{ir})^2 \quad (10)$$

El objetivo del coeficiente de Dave es la optimización mediante minimización para el número de clusters c , tal como [14]:

$$\min(V_{FDC}) \quad (11)$$

El número de clusters c obtenido en el proceso de validación, permite definir las categorías en el proceso de clustering, para así construir los prototipos difusos que pueden ser usados como datos de entrenamiento por una Máquina de Soporte de Vectores (SVM) [16] o técnicas de defusificación [17], mediante las cuales podemos clasificar nuevos datos y predecir su comportamiento mediante variables descriptivas asociadas a cada categoría.

III. EXTRACCIÓN DE VARIABLES EN EL CONTEXTO ACADÉMICO

Las medidas usadas en el proceso de clustering son extraídas de la historia académica de estudiantes, estas variables extraídas son consideradas como características que permiten la abstracción de la carga académica semestral y el rendimiento de un estudiante.

La carga académica semestral es descrita mediante medidas asociadas a cada materia que definen qué tan difícil es aprobarlas, y con qué rigurosidad son calificadas; Caulkins et. al [18] introdujeron las medidas de dificultad y rigurosidad, estas medidas están definidas a su vez por: el promedio académico de los estudiantes y la calificación que obtuvo cada uno de ellos en un curso.

La medida de dificultad es definida formalmente como:

$$\alpha_j = \frac{\sum_i \alpha_i \eta_{ij}}{\sum_i (\alpha_i - GPA_i)} \quad (12)$$

Y la rigurosidad es:

$$\beta_j = \frac{\sum_i (GPA_i - \mu_j)}{N_j} \quad (13)$$

Donde GPA_i es el promedio académico de un estudiante i , η_{ij} es la calificación del estudiante i en la materia j , y N_j es el número de estudiantes que han cursado la materia j .

Además, Mendez et. al [19] introdujeron la medida de distribución de calificaciones, la cual es una medida asociada igualmente a una materia, y describe la asimetría de la distribución estadística de la diferencia entre el promedio académico de cada estudiante, que ha cursado la materia a medir, y la calificación obtenida por cada estudiante en dicha materia, esta es definida formalmente como:

$$Sk_j = skewness_j(GPA_i - \eta_{ij}) \quad (14)$$

Estas medidas son usadas como vector de características, este vector formado por las tres medidas antes descritas define la carga académica semestral, basada en la dificultad y rigurosidad con que es calificada una materia.

A través de esta medida es posible describir el rendimiento de un estudiante mediante su historia académica, y extraer las variables que describan las habilidades que adquieren al aprobar cada materia, independientemente del nivel en el que el estudiante se encuentre, Mendez et. al [19] propuso el uso de un Análisis Exploratorio de Factores [20] para develar dichas estructuras en la malla curricular de la carrera de Ingeniería en Ciencias Computacionales de la Universidad ESPO, este análisis develó 5 factores que componen el núcleo de 26 materias en la malla curricular, estos factores son:

1. Factor de preparación en ingeniería básica
2. Factor de temas avanzados de Ciencias Computacionales
3. Factor de interacción con el cliente
4. Factor de programación
5. Factor de cursos no estrechamente relacionados con las Ciencias Computacionales

El Factor de preparación en ingeniería básica está conformado por materias que están relacionadas con las bases generales para las carreras de ingeniería como: cálculo, física, estadísticas y química, estas materias permiten el desarrollo de la habilidad para entender conceptos abstractos e interpretar fenómenos; el Factor de temas avanzados de Ciencias Computacionales agrupa materias relacionadas a tópicos actuales de las Ciencias Computacionales como: Inteligencia Artificial, Sistemas Operativos, Interacción Humano-computador, Ingeniería de Software, etc., estas materias permiten el desarrollo de la habilidad para identificar los componentes de un sistema y su relación; el Factor de interacción con el cliente está conformado por materias que

permiten expresarse de forma apropiada, verbal y escrita, estas permiten desarrollar la habilidad para comunicarse con usuarios finales y clientes, parte del proceso de desarrollo de software; El Factor de programación agrupa materias como Lenguajes de Programación y Fundamentos de Programación, las cuales permiten desarrollar la habilidad para programar, necesaria en el proceso de desarrollo de software para manejar conceptos, estrategias y patrones de diseño; finalmente el Factor de cursos no estrechamente relacionados con las Ciencias Computacionales está formado por materias generalmente relacionadas a la ingeniería eléctrica.

Con base en estas agrupaciones de materias interrelacionadas que definen los factores en la malla curricular, se propone una medida que describe el desempeño en un estudiante para las materias dentro de dichas agrupaciones, la medida propuesta se puede definir como:

$$f_{F_k, i} = \frac{n_{a_{F_k}}}{n_{F_k}} \cdot \sum_{j \in F_k} \frac{\eta_{ij}}{n_{F_k}} \quad (15)$$

Donde F_k es el grupo de materias del factor k ; $k = 1, \dots, 5$; $n_{a_{F_k}}$ es el número de materias en el factor F_k que ha aprobado un estudiante i , n_{F_k} es el número de materias en el factor F_k que ha tomado un estudiante i , y η_{ij} es la calificación del estudiante i en la materia j dentro de F_k . Esta medida está formada por dos componentes en una multiplicación, el primer componente describe la eficiencia del estudiante para aprobar las materias en F_k , y el segundo es una analogía al promedio académico, pero solo para las materias en el factor F_k .

De esta forma el rendimiento de un estudiante puede ser definido por un vector de características compuesto por las medidas dentro de cada uno de los cinco factores.

Estos vectores asociados a carga académica semestral y rendimiento académico pueden ser usados en un proceso de clustering, permitiendo clasificar de forma no-supervisada a estudiantes en combinación con semestres, y así predecir su riesgo de reprobación basado en la tasa de reprobación de estudiantes con características similares.

IV. DISEÑO DEL PREDICTOR DE RIESGO DE REPROBACIÓN

El prototipo diseñado está conformado por tres componentes principales: un componente de pre-procesamiento, un componente de clustering y un componente para clasificar y estimar el riesgo de reprobación, en la figura 1 se puede apreciar el diagrama de componentes del software de estimación de riesgo de reprobación.

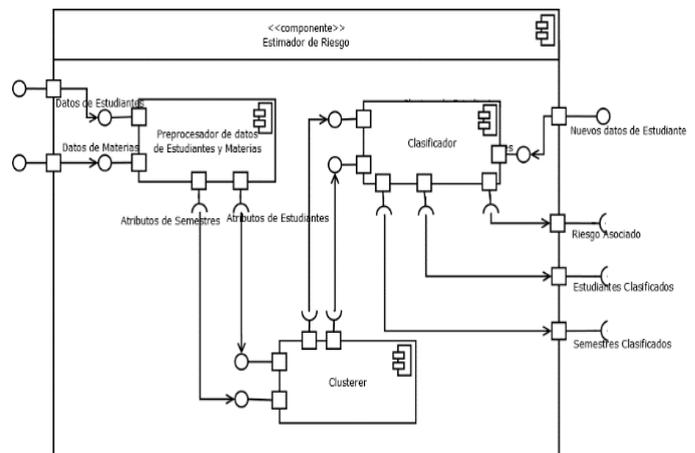


Fig. 1. Diagrama de componentes del software de estimación de riesgo de reprobación

El componente de pre-procesamiento tiene la función de seleccionar los atributos y extraer las características en la historia académica de los estudiantes; características que son usadas por el componente de clustering, el cual construye los prototipos y determina los clusters, mediante un proceso de validación, para generar las categorías de estudiantes y semestres; las categorías obtenidas son usadas luego, en el componente de estimación y clasificación, para estimar la razón de reprobación asociada a cada grupo de estudiante-semestre, el cual también clasifica nuevos datos de estudiantes mediante defusificación y asociarlos, según su valor de membresía, dada por la similitud a una categoría, a un riesgo de reprobación; y así predecir la reprobación en al menos una materia durante el semestre.

V. RESULTADOS

Varios experimentos fueron realizados para el análisis de la precisión del modelo propuesto, para lo cual se realizaron pruebas usando los datos académicos históricos de estudiantes de Ingeniería en Ciencias Computacionales registrados entre los años 1978 y 2013, todas las calificaciones fueron convalidadas y representadas con el núcleo de la malla curricular de la carrera, el cual está compuesto por 26 materias como se menciona en [19].

Se realizó una validación del número de clusters obtenidos en el proceso de clustering mediante la optimización del índice de Dave, para este proceso se usaron las cinco medidas que describen rendimiento con el objetivo de categorizar los estudiantes; en esta prueba se obtuvo un número de clusters $c = 5$ con índice de $V_{index} = 0.850776$ siendo el máximo valor, en la figura 2 se puede observar el gráfico de radar para las medidas en cada uno de los clusters obtenidos.

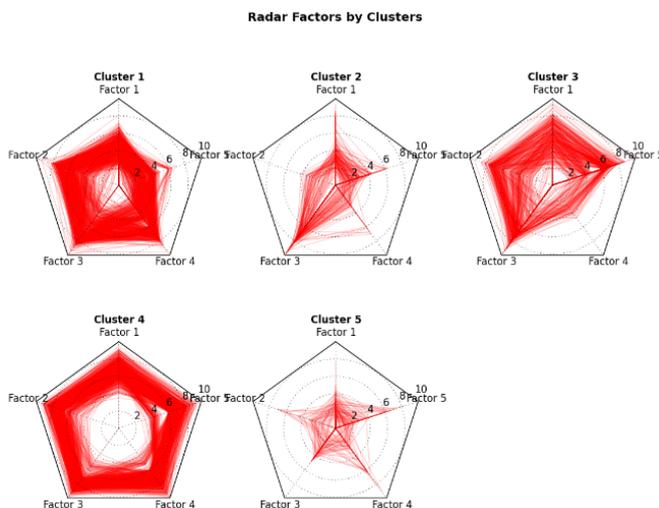


Fig. 2. Gráficos de radar en cada cluster de estudiantes, las cinco medidas de habilidades

Como se puede observar en el gráfico, el primer cluster está formado por estudiantes que poseen habilidades en los cinco factores relativamente equilibradas a excepción del factor 5, ligado a los cursos no estrechamente relacionados a Ciencias Computacionales; el segundo cluster está formado por estudiantes con medidas mayormente bajas en todos los factores a excepción del factor 3, el cual está relacionado a materias con un fuerte componente de interacción con el cliente; el tercer cluster está conformado por estudiantes con una medida relativamente baja en el factor 4, factor relacionado a las habilidades de programación, y en los demás factores posee una medida relativamente alta; en el cuarto cluster

podemos encontrar estudiantes con medidas relativamente altas y equilibradas; finalmente en el quinto cluster encontramos estudiantes con medidas generalmente bajas. En este experimento se pudo identificar los clusters que serán usados en el contexto de categorías de estudiantes.

Para el análisis del nivel de certeza que posee el modelo, se realizó una verificación con el Score de Brier [21], el cual permite medir la precisión en un modelo de predicción, esto se realiza con el contraste entre la probabilidad predicha y la frecuencia relativa observada, para este caso la verificación fue realizada en el pronóstico del riesgo de reprobación. El Score de Brier puede ser descompuesto en tres términos que describen [22]: la incertidumbre de que ocurra o no el evento; la confiabilidad dada por la cercanía entre la probabilidad pronosticada y la probabilidad real; y la resolución que define la diferencia entre la probabilidad de los pronósticos y el promedio observado; con estos términos es posible determinar el score de skill de Brier el cual mide la diferencia entre el score de la predicción y el score de un modelo no cualificado; para valores negativos del score de skill se puede interpretar al modelo como no cualificado para predicción.

Para el análisis con el primer término del año 2013 se encontró un score de Brier $bs = 0.2164$, cercano a cero, un score de skill positivo $ss = 0.1304$; con una **confiabilidad** = 0.0101 cercana a cero, una **resolución** = 0.04256 igualmente cercana a cero y además un porcentaje de incertidumbre de 24.88%, el riesgo pronosticado vs. la frecuencia relativa observada se puede observar en el gráfico de confiabilidad en la figura 3.

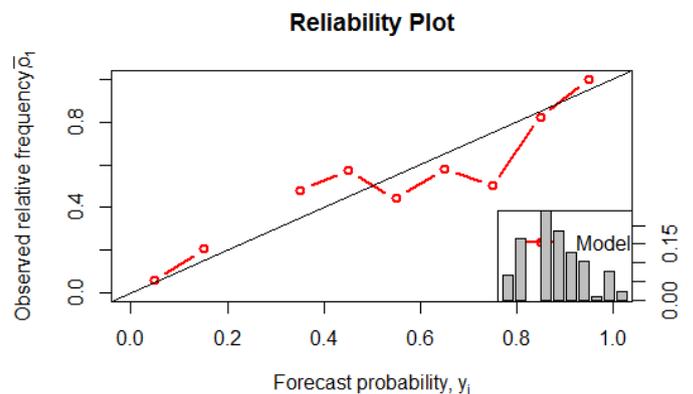


Fig. 3. Confiabilidad para la estimación de riesgo de falla en estudiantes de Ciencias Computacionales que tomaron materias en el término académico 2013-I usando el modelo propuesto.

En el análisis con el segundo término del 2013 se encontró un score de Brier de $bs = 0.2422$, aproximado al score del primer semestre, un score de skill $ss = 0.03113$ también positivo, con una **confiabilidad** = 0.01486 muy cercana a cero, una **resolución** = 0.02264 también cercana a cero, y una incertidumbre de 24.99%, en la Fig. 4 se puede observar el gráfico de confiabilidad para este caso.

Se realizó un contraste entre el modelo y una técnica comúnmente usada en minería de datos, Regresión Logística [23], para predecir el riesgo de reprobación utilizando el mismo conjunto de datos. Como resultado de esta verificación se obtuvo un score de Brier $bs = 0.3405$, para el primer semestre del 2013 y un score de skill $ss = -0.3683$, con una **confiabilidad** = 0.1164; además, una **resolución** = 0.0247, con una incertidumbre de 24.88%; en la figura 5, se puede observar el gráfico de confiabilidad, el

cual muestra la diferencia marcada entre la frecuencia relativa observada y el riesgo predicho.

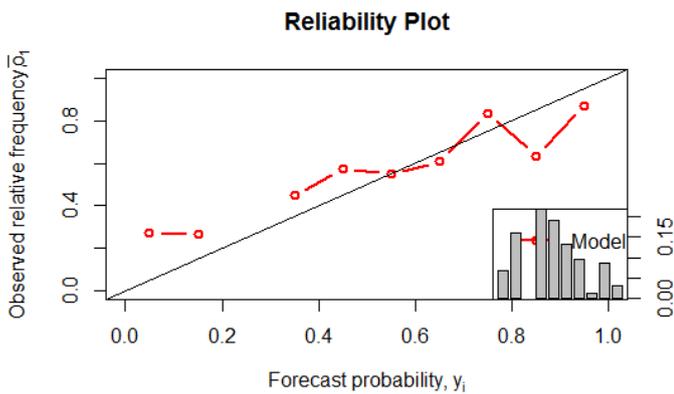


Fig. 4. Confiabilidad para la estimación de riesgo de falla en estudiantes de Ciencias Computacionales que tomaron materias en el término académico 2013-II usando el modelo propuesto.

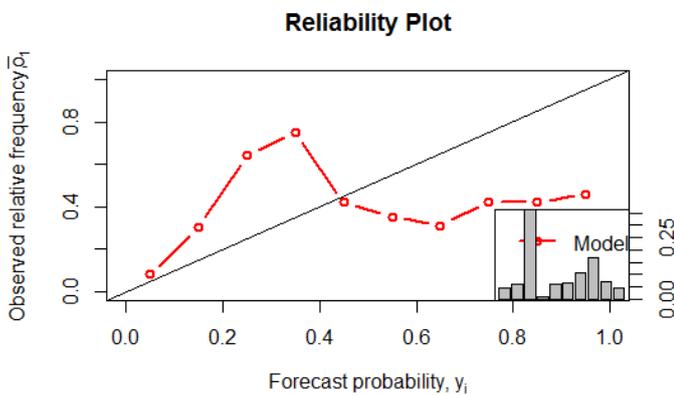


Fig. 5. Confiabilidad para la estimación de riesgo de falla, usando regresión logística, en estudiantes de Ciencias Computacionales que tomaron materias en el término académico 2013-I.

Para el segundo semestre del 2013, en la predicción mediante Regresión Logística, se pudo observar un score de Brier $bs = 0.3217$, un score de skill $ss = -0.2869$, con una confiabilidad = 0.08261, una resolución = 0.01089, y una incertidumbre de 24.99%; en la figura 6, se puede apreciar la diferencia entre frecuencia relativa observada y el riesgo predicho.

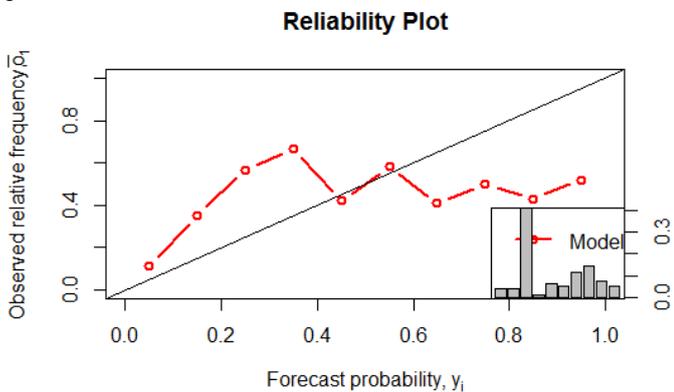


Fig. 6. Confiabilidad para la estimación de riesgo de falla, usando regresión logística, en estudiantes de Ciencias Computacionales que tomaron materias en el término académico 2013-II.

VI. CONCLUSIONES

El uso de prototipos difusos propuesto como clasificador no-supervisado para la predicción del rendimiento en estudiantes de pre-grado, mediante el pronóstico de riesgo de

reprobación, con base en la historia académica de estudiantes similares, utiliza algoritmos de clustering difuso para la construcción de prototipos dentro de un conjunto de datos, mediante la semejanza y disimilitud entre los elementos de dicho conjunto, este proceso permite identificar los elementos más típicos del conjunto o prototipos, los cuales pueden ser usados en un clasificador no-supervisado para definir la similitud con nuevos datos y asociar un valor de riesgo.

La validación en el proceso de clustering permite estimar el número de clusters, en los casos en que no se conoce el número de categorías dentro de un conjunto de datos, y sea difícil visualizar la distribución de los datos en cada cluster.

Cada una de las categorías, definidas en el proceso de validación, depende de las características o atributos que son seleccionados, o extraídos de cada elemento del conjunto de datos, puesto que estos además de aportar a la semántica de los prototipos construidos, también describe la distribución de los elementos dentro de cada cluster.

En este caso los atributos usados permitieron describir el rendimiento de los estudiantes y la carga académica a la que se enfrentaban, y así definir la similitud entre estudiantes en el proceso de clustering, así como asociar cada cluster a una tasa de reprobación para generalizarla a nuevos datos clasificados, como un valor de riesgo de reprobación.

Con la arquitectura propuesta se diseñó un prototipo, el cual permitió predecir el riesgo de falla en el rendimiento de estudiantes de pre-grado en la carrera de Ingeniería en Ciencias Computacionales, esto con una certeza superior al 75% en cada uno de los términos académicos del año 2013, en el que fue probado, permitiendo a los estudiantes tomar decisiones más educadas en el proceso de selección de materias a registrarse, con base en la predicción del riesgo de reprobación.

Como estrategia a futuro, se espera incorporar otros factores como la carga extra-universitaria, información demográfica y socioeconómica, para mejorar la precisión del modelo de predicción; además la medición de las habilidades ganadas, que definen el rendimiento académico, está en función de las materias en la malla curricular de los estudiantes cuyos datos fueron usados en el caso de estudio, por lo que otra estrategia a futuro sería la descripción del rendimiento, mediante extracción de variables, de manera independiente a la carrera en la que el estudiante se encuentre.

REFERENCIAS

- [1] V. McGaha, and J. Fitzpatrick, "Personal and social contributors to dropout risk for undergraduate students", *College Student Journal*, vol. 39 no. 2, 2005, pp. 287.
- [2] B. K. Bhardwaj, and S. Pal, "Data Mining: A prediction for performance improvement using classification", *arXiv preprint arXiv: 1201.3418*, 2012.
- [3] M. Pandey and S. Taruna. "A Comparative Study of Ensemble Methods for Students' Performance Modeling." *International Journal of Computer Applications*, vol. 103 no. 8, 2014.
- [4] V. Oladokun, A. Adebajo, and O. Charles-Owaba. "Predicting students' academic performance using artificial neural network: A case study of an engineering course." *The Pacific Journal of Science and Technology*, vol. 9 no. 1, 2008, pp. 72-79.
- [5] S. A. Naser, I. Zaqout, M. A. Ghosh, R. Atallah and E. Alajrami, "Predicting Student Performance Using Artificial Neural Network: in the Faculty of Engineering and Information Technology", *International Journal of Hybrid Information Technology*, vol. 8 no. 2, 2015, pp. 221-228.
- [6] L. Kaufman, and P. J. Rousseeuw. "Finding groups in data: an introduction to cluster analysis", vol. 344, John Wiley & Sons, 2009.

- [7] J. C. Bezdek, "Pattern recognition with fuzzy objective function algorithms", Kluwer Academic Publishers, 1981.
- [8] M. Rifqi, "Constructing prototypes from large databases", In International conference on Information Processing and Management of Uncertainty in knowledge-based systems, IPMU'96, 1996, pp. 301-306.
- [9] M. Rifqi, V. Berger, and B. Bouchon-Meunier, "Discrimination power of measures of comparison". Fuzzy sets and system 110 no. 2, 2000, pp. 189-196.
- [10] A. J. Smola and B. Schölkopf, "Learning with kernels". GMD-Forschungszentrum Informationstechnik, 1998, pp. 210.
- [11] M. J. Lesot, "Similarity, typicality and fuzzy prototypes for numerical data", In 6th European Congress on Systems Science, Workshop Similarity and resemblance, vol. 94, 2005, pp. 95.
- [12] M. J. Lesot, M. Rifqi, and B. Bouchon-Meunier, "Fuzzy prototypes: From a cognitive view to a machine learning principle", In Fuzzy Sets and Their Extensions: Representation, Aggregation and Models, Springer Berlin Heidelberg, 2008, pp. 431-452.
- [13] J. C. Bezdek, R. Ehrlich, and W. Full, "FCM: The fuzzy c-means clustering algorithm", Computers & Geosciences 10 no. 2, 1984, pp. 191-203.
- [14] W. Wang, and Y. Zhang, "On fuzzy cluster validity indices", Fuzzy sets and systems 158 no. 19, 2007, pp. 2095-2117.
- [15] R. N. Dave, "Validating fuzzy partitions obtained through c-shells clustering", Pattern Recognition Letters 17, no. 6, 1996, pp. 613-623.
- [16] J. A. Suykens and J. Vandewalle, "Least squares support vector machine classifiers", Neural processing letters, 1999, pp. 293-300.
- [17] T. Runkler, "Selection of appropriate defuzzification methods using application specific properties", Fuzzy Systems, IEEE Transactions on, vol. 5 no. 1, 1997, pp. 72-79.
- [18] J. P. Caulkins, P. D. Larkey, and J. Wei, "Adjusting GPA to reflect course difficulty", 1996.
- [19] G. Méndez, X. Ochoa and K. Chiluíza, "Techniques for data-driven curriculum analysis", In Proceedins of the Fourth International Conference on Learning Analytics and Knowledge. ACM, 2014, pp. 148-157.
- [20] L. R. Fabrigar and D. T. Wegener, "Exploratory factor analysis", Oxford University Press, 2011.
- [21] G. W. Brier, "Verification of forecasts expressed in terms of probability". Monthly weather review, vol.78 no. 1, 1950, pp. 1-3.
- [22] A. H. Murphy, "A new vector partition of the probability score". Journal of Applied Meteorology, vol. 12 no. 4, 1973, pp. 595-600.
- [23] Jr, D. W. Hosmer, S. Lemeshow and R. X. Sturdivant, "Applied logistic regression", vol. 398, John Wiley & Sons, 2013.
- [24] M. J. Lesot, Mouillet, and B. Bouchon-Meunier, "Fuzzy prototypes based on typicality degrees. In Computational Intelligence, Theory and Applications". Springer Berlin Heidelberg, 2005, pp. 125-138.
- [25] R. Krishnapuram, and J. M. Keller, "A possibilistic approach to clustering". Fuzzy Systems, IEEE Transactions, 1993, pp. 98-110.



Anibal Vásquez. Recibió la Ingeniería en Ciencias Computacionales en la ESPOL en 2015. Actualmente es Ayudante de investigación en el programa de Tecnologías para la Enseñanza y el Aprendizaje en el Centro de Tecnologías de Información en la ESPOL.



Enrique Peláez. Recibió la Ingeniería en Electricidad en la ESPOL en 1989, M.S. in Electrical and Computer Engineering en University of South Carolina en 1991 y el título de PhD. in Electrical and Computer Engineering en University of South Carolina en 1994. Actualmente es Director del Centro de investigación en Tecnologías de Información en la ESPOL y profesor de postgrado en el área de Inteligencia Computacional de la Facultad de Ingeniería en Electricidad y Computación de la ESPOL.



Xavier Ochoa. Recibió la Ingeniería en Computación en la ESPOL en 2000, M.S. in Applied Computer Science en Free University of Brussels en 2002 y el título de PhD. in Engineering por la Catholic University of Leuvenolina en 2008. Actualmente es Director del programa de Tecnologías para la Enseñanza y el Aprendizaje en el Centro de Tecnologías de Información en la ESPOL y profesor de postgrado en la Facultad de Ingeniería en Electricidad y Computación.