

FORTHCOMING 03B50-22-01-03

RECONOCIMIENTO DE PATRONES DE APRENDIZAJE PARA ELEVAR LA CALIDAD EDUCATIVA EN LA EDUCACIÓN SUPERIOR MEDIDO EN UNA ESCALA NEUTROSÓFICA INDETERMINADA DE LIKERT

Joffre Ruperto Paladines Rodríguez^{1*}, Ricardo Sánchez Casanova^{**} y Maikel Yelandi Leyva Vázquez^{***}^{*}Universidad Católica de Santiago de Guayaquil, Ecuador^{**}Universidad de la Habana, Cuba^{***}Facultad de Ciencias Matemáticas y Físicas, Universidad de Guayaquil, Ecuador

RESUMEN

La utilización de las tecnologías en el desarrollo de actividades áulicas contribuye al mejoramiento de la calidad educativa en la Educación Superior, permitiendo elevar los resultados de aprendizaje. El problema científico radica en cómo contribuir a la identificación de patrones de aprendizaje generados a través de la utilización de la plataforma Moodle. Los métodos analítico-sintético, inductivo-deductivo, histórico-lógico y sistematización posibilitaron identificar técnicas apropiadas que permiten reconocer patrones almacenados en bases de datos. El objetivo del presente estudio se basa en el análisis de los datos obtenidos en la plataforma Moodle y su procesamiento en la herramienta Orange, mediante la metodología CRIP-DM. Se emplearon redes neuronales para predecir la evaluación final y el índice de SHAP para explicar el modelo. Se puede apreciar que los 3 atributos más importantes están relacionados con los conocimientos adquiridos y el nivel de esfuerzo. Se aplicaron encuestas a estudiantes, quienes respondieron usando una escala de Likert Indeterminada, que consiste en evaluar cada pregunta en cinco componentes en lugar de una sola, lo cual otorga a las respuestas mayor indeterminación y precisión.

PALABRAS CLAVE: Reconocimiento de patrones, tecnologías, calidad educativa, índice SHAP, escala de Likert Indeterminada, Neutrosofía.

MSC: 03B50, 68T30, 68T35, 68T37

ABSTRACT

The use of technologies in the development of classroom activities contributes to improved educational quality in Higher Education, allowing raising learning results. The scientific problem lies in how to contribute to the identification of learning patterns generated through the Moodle platform. The analytical-synthetic, inductive-deductive, historical-logical, and systematization methods made it possible to identify techniques that allow us to recognize patterns stored in databases. The objective of this study is based on the analysis of the data obtained in the Moodle platform and its processing in Orange tool, using the CRIP-DM methodology. Neural networks were used to predict the final evaluation and the SHAP index to explain the model. It can be seen that the three most important attributes are related to the knowledge acquired and the level of effort. Questionnaires were applied to students, which answered using an indeterminate Likert scale that consists in the assessment of every question with five components instead of only one of them, so, responses are more indeterminate and accurate.

KEYWORDS: Pattern recognition, technologies, educational quality, SHAP index, indeterminate Likert scale, Neutrosophy.

1. INTRODUCCIÓN

Los *avances tecnológicos* han sido parte fundamental en el desarrollo durante todas las etapas de la historia de la humanidad y han permitido cambios significativos en diversos procesos y múltiples campos y áreas de la ciencia y la tecnología, lo que ha hecho evolucionar a las personas en pensamiento, criterios y actitudes socio-culturales.

Dichos avances han involucrado a las comunicaciones, la microelectrónica, la cibernética y la informática aportando con mayores contribuciones al campo educativo concretando el desarrollo de nuevos modelos

¹ Email: joffre_paladines@hotmail.com

pedagógicos dentro del proceso de enseñanza-aprendizaje tanto en el proceso presencial, a distancia y mixta.

Las Tecnologías de la Informática y las Comunicaciones (TIC) aportan un reto al sistema educativo relacionado con la necesidad de superar el modelo pedagógico de la enseñanza tradicional, centrada en el profesor, en la transmisión de contenidos por parte de este y en el carácter reproductivo, memorístico y receptivo del aprendizaje por parte del alumno ([24]). La utilización de las TIC involucradas en los procesos de enseñanza-aprendizaje ha contribuido a mejorar la calidad educativa y elevar el nivel de aprendizaje por parte de los estudiantes y en los profesores, lográndose un cambio significativo al impartir las clases ampliando los tipos de recursos tecnológicos involucrados en este proceso educativo. En este sentido, en Iberoamérica se han desarrollado diversos proyectos de innovación cuyo objetivo central ha sido la introducción de tecnologías en las instituciones de Educación Superior donde se involucran los programas tradicionales con la utilización constante de las TIC en los procesos de educación en línea.

Tomando en consideración lo antes mencionado, en el *Sistema Educativo Superior* actual se crean y se generan la simulación de manejo de laboratorios, aulas, espacios, foros virtuales, el desarrollo de sistemas expertos, se mejoran las herramientas de trabajo que se utilizan en la educación de ambiente web. Las nuevas situaciones de enseñanza y aprendizaje basadas en las tecnologías promueven la participación en experiencias educativas altamente interactivas.

La Universidad Católica de Santiago de Guayaquil no está al margen de la utilización de la tecnología involucrada en el proceso constante y diario de la enseñanza-aprendizaje, que considera utilizar la plataforma Moodle como su herramienta principal para la interacción de docentes y estudiantes en el apoyo y fortalecimiento de las actividades académicas para la formación de futuros profesionales. De hecho, Moodle proporciona un poderoso conjunto de herramientas centradas en el estudiante y un entorno de aprendizaje colaborativo para mejorar la capacidad de enseñanza y aprendizaje, ya son más de 10 años que contribuye al desarrollo de la pedagogía de constructivismo social, proporcionando herramientas colaborativas como foros, blogs, chats, wikis entre otras, permitiendo facilitar el aprendizaje mixto (*blended learning*).

En la carrera de Animación Digital de la Facultad de Artes y Humanidades de la Universidad Católica de Santiago de Guayaquil se detecta que hay poco control y seguimiento continuo a las actividades desarrolladas por los estudiantes en la Plataforma Moodle, por lo cual se considera la valoración, creación, implantación y evaluación de una aplicación o de un modelo de análisis de datos que permita detectar el comportamiento generado junto al grado de aprendizaje adquirido a través de las actividades planificadas en la plataforma.

Esto conlleva a que no se pueden obtener métricas sobre los resultados generados y que permitan mostrar estadísticas u otros tipos de análisis que identifiquen patrones tanto de rendimiento, como de aprendizaje, ni mucho menos identificar qué elementos del proceso de aprendizaje no se están aprovechando al máximo.

En este artículo se estudian 259 estudiantes que son encuestados y cuyos resultados se procesan mediante herramientas de programación como Orange. Para ganar en exactitud se les pide que las respuestas sean basadas en una Escala de Likert Indeterminada ([18]). Con la utilización de esta escala, los encuestados pueden dar cinco evaluaciones a una sola pregunta, en cinco aspectos diferentes, que son: Pertenencia negativa, Indeterminación tendiente hacia pertenencia negativa, Pertenencia indeterminada, Indeterminación tendiente hacia pertenencia positiva, Pertenencia positiva. De esta manera se capturan las posibles contradicciones de sentimientos y pensamiento de los encuestados.

Esta herramienta es una generalización de la escala de Likert clásica ([20]) en el campo de la Neutrosofía. La Neutrosofía es la rama de la filosofía que estudia las indeterminaciones debido a las contradicciones, las paradojas, las incoherencias, las inconsistencias, entre otros aspectos del conocimiento ([28]).

Otra herramienta neutrosófica que generaliza una clásica es el método Delphi neutrosófico donde se incorpora la indeterminación en la evaluación de un grupo de expertos. Este se puede consultar en más detalles en [30]. No obstante los autores prefirieron utilizar una encuesta para estudiantes quienes evalúan mediante una escala de Likert. Una aproximación a las ciencias pedagógicas desde la neutrosofía se puede encontrar en [7] y de Investigación Operacional aplicada en pedagogía se puede consultar en [21].

2. LAS TECNOLOGÍAS DE LA INFORMÁTICA Y LAS COMUNICACIONES EN LOS PROCESOS EDUCATIVOS DE LAS INSTITUCIONES DE EDUCACIÓN SUPERIOR

La integración de las TIC en los procesos educativos de las Instituciones de Educación Superior (IES), corresponde a los objetivos de la eficiencia y eficacia en la calidad de los procesos educativos ([13]) estos autores coinciden en que el uso y empleo de las TIC permite mejorar los procesos educativos.

Por ende, la innovación tecnológica de las universidades hacia la transformación digital son retos y cambios en los paradigmas ya establecidos como estándares a mantener, sin duda alguna es inevitable la

modificación de los modelos educativos considerando la reconfiguración y reorientación de la enseñanza tradicional. Valverde ([31]) menciona que se deben generar políticas educativas, tecnologías educativas más utilizadas, modelos y prácticas educativas digitales encaminadas hacia la incorporación de la digitalización de la docencia universitaria.

A partir de la década de los noventa empezó la revolución de las *plataformas educativas*, teniendo como función principal facilitar el proceso de aprendizaje en las universidades, permitiendo en muchos casos mantener una educación dual, combinando la educación presencial con la modalidad *online* estableciendo la modalidad mixta o también conocida como *blended learning*, muchas veces mencionada por autores como educación semipresencial ([4]).

El estudio realizado por Área, Hernández y Sosa ([3]), identifican modelos de integración didáctica que permiten establecer la funcionalidad que tienen las tecnologías en la educación y sus procesos educativos, como son: la frecuencia con la que se utilizan las TIC y diversidad de las actividades didácticas.

Enfatizando básicamente en las actividades de aprendizaje que se aplican en el aula, tales como la socialización de los contenidos, la búsqueda de información con el apoyo del Internet para el desarrollo de actividades, tareas o ejercicios por medio de alguna plataforma en línea. La evolución de la web se ha convertido en un punto de concentración global y enfocada al mejoramiento de la calidad educativa considerando a los procesos de enseñanza, aprendizaje, metodologías, didácticas y tecnologías educativas, las cuales tienen una relación estrechamente ligada a la aplicación y uso de las tecnologías.

Basados en estos principios, según Marín, Inciarte, Hernández y Pitre ([22]) se deben establecer estrategias que sean consideradas y utilizadas por las IES en Ecuador que les permita ser más competitivas y sostenibles orientadas a la innovación y la intervención significativa del ámbito social y educativo. En consecuencia según Mora ([23]), las IES deben aplicar una contextualización en los cambios de los modelos educativos aplicados a partir de la era del conocimiento puntualizando el valor económico de la educación y su desarrollo tecnológico y así se establezcan y diseñen políticas educativas respondiendo directamente a los requerimientos y necesidades del medio, evaluando y mejorando continuamente los procesos y gestión de las TIC.

3. ANÁLISIS DE DATOS

La información en sus diversas modalidades expresivas, ha existido en todos los tiempos y en múltiples circunstancias sociales y culturales, pero solo a partir del siglo XX comienza a ser considerada como un fenómeno de gradual importancia en todos los terrenos, razón por la cual la sociedad necesita hombres capaces de consultar un gran volumen de información y utilizar este caudal de conocimientos en la solución de forma creadora de los problemas que se plantean.

El *Análisis de Datos* (AD), es la acción de someterlos a operaciones con la finalidad de obtener conclusiones precisas que permiten alcanzar objetivos que solo se pueden mientras exista una recolección de datos a los cuales se puedan establecer dichos análisis. En la década de los años 90 empiezan a crearse los primeros programas para el análisis de datos, es así como Richard West, Richard Young, Judith Johnson y Christine Macfarlane de la Universidad Estatal de Utah incorporan reglas de decisiones que son aplicados a los datos obtenidos en múltiples registros ([6]).

Un programa llamado *Sheri* demostró que se puede incorporar gráficos y generar informes, otro programa capaz de realizar los mismos análisis y procesos es *Eventlog* desarrollado por Robert Henderson para una empresa de educación, *Conduit Educational Software*, con el propósito de que lo utilizaran investigadores que registraban datos de observaciones.

El investigador Bakeman ([5]) estableció el ajuste a los tipos de datos sobre el cálculo de coeficientes de control de calidad del dato que incorporaban la tecnología de IBM, luego en la Universidad de Kent State, se construye por Robert Zuckerman, Lyle Barton y Harold Johnson el programa DA+A, el cual aplicaba técnicas analíticas que eran consideradas adecuadas, tales como estadísticos descriptivos, análisis de ciclos, coeficientes de concordancia, detección de patrones, generando una tendencia de desarrollo de programas informáticos para el registro de datos observacionales que estaban siendo utilizados.

Son los autores Coll, Engel y Niño ([11]) quienes establecen tres requerimientos para aplicar las analíticas:

1. Análisis de datos de alto nivel que permita definir e identificar perfiles diferenciados de las actividades de los estudiantes a partir de combinaciones de indicadores e índices de distintos tipos.
2. Permitir presentar la información de los estudiantes en formato de visualización sencillo.
3. Debe ser flexible para su aplicación en una amplia gama de entornos de aprendizaje en línea en distintos tipos de actividades con objetivos y contenidos educativos distintos y con diferentes grupos de estudiantes.

Sin duda alguna la analítica de aprendizaje juega un papel importante en los procesos educativos, los autores Amo y Santiago ([1]) establecen una relación entre la analítica de aprendizaje, las metodologías educativas, las tecnologías innovadoras y la utilización de entornos virtuales de aprendizaje, lo cual posibilita el seguimiento y evaluación del aprendizaje, relacionan la analítica de aprendizaje con los diferentes estilos de aprendizajes y niveles cognitivos de aprendizaje considerando a la taxonomía de Bloom como el principal referente en sus estudios, apoyándose en Anderson y Krathwohl ([2]) y del proceso cognitivo de Gagné ([15]) y su taxonomía, complementándose con el análisis de las inteligencias múltiples.

En consecuencia, las analíticas de aprendizaje en la actualidad se convierten en un recurso para la docencia y el mejoramiento constante del aprendizaje, es decir, que son sistemas dinámicos que se encuentran en constante cambios por las múltiples variables que se descubren en este proceso, esto se torna complicado y difícil al momento de realizar un análisis y generar un modelo cuantitativo. Bossolasco y otros ([8]) manifiestan:

“Sería conveniente trabajar a partir de una adecuada combinación entre los datos que arrojan las analíticas y que ofrecen ideas en escala, el análisis de dichos datos desde marcos conceptuales pertinentes y la triangulación de estos datos cuantitativos con otro grupo de datos densos, de origen cualitativos”.

La **minería de datos** (MD) o *Data Mining*, es la exploración de datos, consiste en la detección de información durante el procesamiento de grandes cantidades de datos, es la aplicación de un conjunto de técnicas que se aplican para la extracción de datos que sean válidos, permite el análisis de información estructurada, exploración y clasificación, realizando búsquedas de patrones con apoyo del *Big Data*.

En la actualidad se utiliza la MD en diferentes campos y áreas de la ciencia, como en las finanzas, banca, estudios de mercados y comercios, en la industria, salud, medicina, bioingeniería, educación entre otras, siendo una tendencia a nivel mundial la aplicación de este proceso utilizando programas que permiten obtener resultados precisos y de forma inmediata.

El Descubrimiento de Conocimiento en Bases de Datos, traducido del idioma inglés (KDD- *Knowledge Discovery in Databases*), es un campo de la informática que permite descubrir nueva información desde grandes bases de datos. Se identifican dos fases importantes en el análisis de los datos almacenados como son: preprocesamiento de datos derivando otros subprocesos (preparación de datos, mecanismo de captura de datos, organización y tratamiento de datos) y posprocesamiento de los resultados obtenidos luego del procesamiento de datos, Fayyad, Piatetsky, Smyth y Uthurusamy ([14]) mencionan que: “KDD es el proceso no trivial de identificar patrones, a partir de datos, que son válidos, nuevos, potencialmente útiles y comprensibles”. Esto concuerda con Costa y otros ([12]), quienes mencionan como un proceso general de descubrimiento de conocimiento compuesto por tres etapas consideradas como estándares, los cuales deben ser nuevos, comprensibles y útiles, permitiendo aportar un nuevo beneficio hasta llegar a ser entendido rápidamente por los usuarios para la correcta toma de decisiones.

3.1. Analíticas de Aprendizaje

Es importante establecer que el proceso de aprendizaje es un factor primordial dentro del accionar educativo en todas las IES, por eso la necesidad de implementar estrategias y modelos que van orientados a mejorar y elevar el nivel y calidad educativa, considerando que la educación ha tenido una constante adaptación a las tecnologías educativas, por ende, la necesidad de establecer investigaciones que permitan conocer la realidad de la educación superior.

Muchas investigaciones van orientadas a descubrir las fallas en el sistema educativo y para determinar qué información es relevante para mejorar la educación dentro de las aulas. Las Analíticas de Aprendizaje (*Learning Analytics*) permiten comprender las necesidades de estudiantes y docentes en el ámbito educacional aplicadas en investigaciones por Johnson y otros ([17]) permitiendo desarrollar estrategias que mejoren la calidad del proceso, como lo menciona Rojas ([25]) quien realiza una investigación bibliográfica sobre la analítica de aprendizaje convirtiéndose en un referente en este campo teniendo información en español, donde se establece la definición sobre *Learning Analytics* (LA) que se define como la medición, recopilación, análisis y presentación de datos sobre los estudiantes y sus contextos, para comprender y optimizar el aprendizaje y los entornos en que se produce.

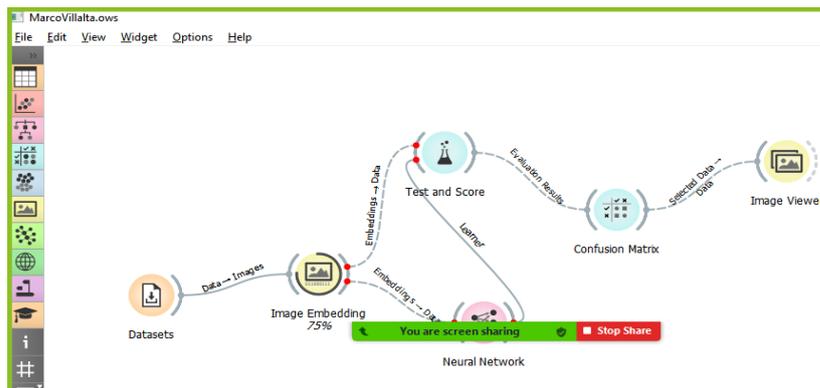
Para entender claramente sobre las analíticas de aprendizaje Slade y Prinsloo ([27]) la definen como: “recopilación, el análisis, el uso y la difusión adecuada de los datos procesables generados por los estudiantes con el fin de crear un apoyo cognitivo, administrativo y eficaz adecuado para los alumnos”. Estos datos son generados por medio de la utilización de plataformas educativas, registros de calificaciones, registros de asistencias, porcentajes de graduados, interacciones en chat, blog entre otras. La analítica de aprendizaje es un nuevo campo de las ciencias de la educación, de donde se extraen datos de las actividades que generan estudiantes y docentes, aprovechando dicha información dinámica que se

obtiene de los entornos virtuales de aprendizaje, estos son analizados e integrados a los resultados de predicciones permitiendo mejorar los procesos de enseñanza-aprendizaje. El informe *Horizon* del *The New Media Consortium* y de la Universidad de Oberta de Cataluña proponen su implementación de cuatro a cinco años, definiendo que las analíticas de aprendizaje consisten en la interpretación de un amplio rango de datos producidos y recogidos acerca de los estudiantes para orientar su progresión académica, predecir actuaciones futuras e identificar elementos problemáticos.

El objetivo de la recolección, registro, análisis y presentación de datos posibilita que los profesores puedan adaptar de manera rápida y eficaz las estrategias educativas al nivel de necesidad y capacidad de cada alumno, las analíticas de aprendizaje responden a la necesidad de llevar a cabo el seguimiento y control de la actividad en el campus para la toma de decisiones estratégicas. Por otro lado, pretenden aprovechar la gran cantidad de datos producidos por los estudiantes en actividades académicas.

Buckingham y Ferguson ([9]) establecen de acuerdo al informe de la UNESCO que el nivel micro posibilita recoger datos, realizar análisis posibilitando intervenir oportunamente sobre los resultados obtenidos, así se podrá tener una visión sobre los hábitos educativos y de aprendizaje de los estudiantes proporcionando recomendaciones para sus mejoras. Todas las actividades desarrolladas en las diversas plataformas educativas se denominan huellas digitales, como la interacción en un foro, desarrollo de tareas, aplicación de evaluaciones en línea, ver videos, la descarga de materiales, el ingreso a las aulas virtuales.

Al referirse como huella digital, Siemens ([26]) indica que son senderos digitales todo aquello que se genera dentro del proceso de aprendizaje a través de una plataforma educativa, el estudiante genera puntos de datos cuando navega, estableciendo patrones o hábitos de lectura, escritura, concluyendo que los senderos de datos permiten explorar al aprendizaje desde varias perspectivas. Johnson, Adams y Cummins ([16]) mencionan sobre los beneficios de EDM para estudiantes, profesores, administradores investigadores y sociedad en general, afianzando sobre el concepto de EDM, Calvet y Juan ([10]) afirman que: “aprendizaje automático y de minería de datos para lograr mejores desempeño a través de decisiones basadas en datos históricos”.



Se determina utilizar la herramienta tecnológica Orange la cual permite encontrar y visualizar datos y aprendizaje automático, la principal ventaja de este programa consiste en que es de código abierto,

Figura 1. Programación Visual en Orange.

permite crear flujos de trabajo de análisis de datos de forma visual, con una caja de herramientas amplia y diversa. Orange permite explorar datos interactivamente para un análisis cualitativo rápido con visualizaciones limpias. La interfaz gráfica de usuario le permite concentrarse en el análisis exploratorio de datos en lugar de la codificación, mientras que los valores predeterminados inteligentes hacen que la creación rápida de prototipos de flujo de trabajo de análisis de datos.

3.2. Programación visual

La fuerza educativa de Orange proviene de la combinación de programación visual y visualizaciones interactivas. También los autores han diseñado algunos *widgets* educativos que se han creado explícitamente para apoyar la enseñanza. Orange es un paquete de software de programación visual basado en componentes para visualización de datos, aprendizaje automático, minería de datos y análisis de datos. Los componentes de Orange se denominan *widgets* y van desde la visualización de datos, la selección de subconjuntos y el preprocesamiento hasta la evaluación empírica de algoritmos de aprendizaje y el modelado predictivo.

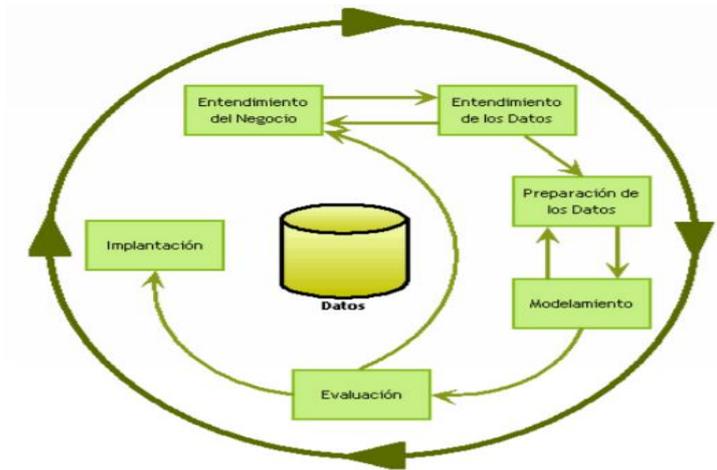
La programación visual se implementa a mediante una interfaz en la que los flujos de trabajo (Figura 1) se crean vinculando *widgets* predefinidos o diseñados por el usuario, mientras que los usuarios avanzados pueden usar Orange como una biblioteca de Python para la manipulación de datos y la alteración de *widgets* (software).

4. MATERIALES Y MÉTODOS

4.1. Detalles técnicos de programación

Se utiliza la metodología CRIP-DM. CRISP-DM fue creada por el grupo de empresas SPSS, NCR y Daimler Chrysler en el año 2000, es actualmente la guía de referencia más utilizada en el desarrollo de proyectos de *Data Mining*. La guía estructura el proceso en seis fases que son:

1. Comprensión del negocio.
2. Comprensión de los datos.
3. Preparación de los datos.
4. Modelado.
5. Evaluación.
6. Implantación.



Para el análisis de los datos a través del programa Orange se utilizó la recopilación de datos por medio de una encuesta desarrollada en formulario de *GoogleForm* de 259 estudiantes de diferentes paralelos del Primer Ciclo en dos periodos diferentes 2-2020 y 1-2021.

Se emplea un molde de red neuronal para predecir la evaluación final del estudiante.

Las **redes neuronales** (NN por sus siglas en inglés) son un

Figura 2. Guía de referencia CRISP-DM. **Fuente:** Inteligencia artificial. **Autor:** elaboración propia.

conjunto de neuronas conectadas entre sí, con el objetivo de lograr el aprendizaje automático de datos por los potentes métodos que estas utilizan para la abstracción; al extraer la información las NN tienen la capacidad de clasificar las características, agruparlas o dar reconocimiento de análisis estadísticos en las muestras tomadas.

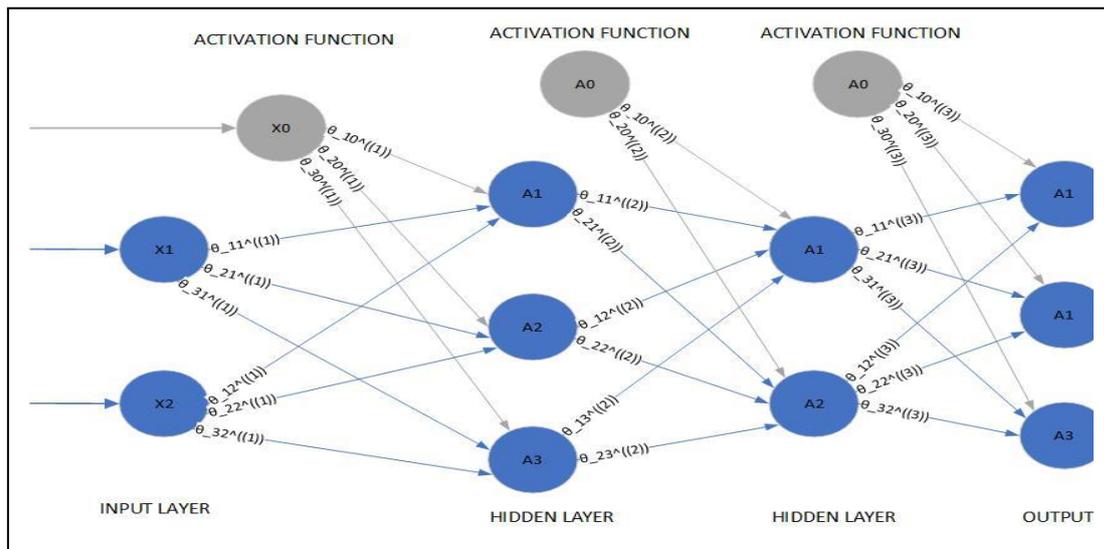


Figura 3. Arquitectura de una red neuronal. Una red neuronal con 4 capas: Capa de entrada, dos capas ocultas y una capa de salida.

Las neuronas azules X representan la capa entrada de la red con una notación $\theta_n^{(n)}$ que representa los pesos que conectan a cada neurona en la siguiente capa, las capas ocultas a_n . La salida de las capas ocultas por cada neurona se obtiene al aplicar la función de activación que se defina en el aprendizaje.

Para comprender la red neuronal se empleó el índice de SHAP. Dicho índice especifica la explicación como:

$$\mathbf{g}(\mathbf{z}'_1) = \boldsymbol{\phi}_0 + \sum_{j=1}^M \boldsymbol{\phi}_j \mathbf{z}'_1 \quad (1)$$

Donde g es el *modelo de explicación*, $\mathbf{z}' \in \{0,1\}^M$ es el *vector de coalición*, M es el tamaño máximo de la coalición y $\boldsymbol{\phi}_j \in \mathbb{R}$ es la atribución de características.

4.2. La escala Likert Indeterminada

Esta subsección contiene los conceptos principales para diseñar una escala Likert neutrosófica. Este tema se introduce porque como se explicó en la subsección anterior se aplicaron cuestionarios a un grupo de estudiantes. Los investigadores decidieron utilizar una escala lo más exacta posible donde los estudiantes expresaran sus opiniones al costo de una mayor indeterminación.

Definición 1: ([28]) Un *Conjunto Neutrosófico de Valor Único* (CNVU) sobre U es $A = \{ \langle x; T_A(x), I_A(x), F_A(x) \rangle : x \in U \}$, donde $T_A: U \rightarrow [0, 1]$, $I_A: U \rightarrow [0, 1]$, y $F_A: U \rightarrow [0, 1]$, $0 \leq T_A(x) + I_A(x) + F_A(x) \leq 3$.

Definición 2: ([19][29]) Una *lógica neutrosófica refinada* se define como un valor de veracidad T que se divide en varios tipos de valores de veracidad: T_1, T_2, \dots, T_p , I en varias indeterminaciones: I_1, I_2, \dots, I_r y F en varias falsedades F_1, F_2, \dots, F_s , donde $p, r, s \geq 1$ son enteros y $p + r + s = n$.

Definición 3: ([19]) Un *conjunto neutrosófico indeterminado triple refinado* (CNITR) A en X se caracteriza por funciones de pertenencia positiva $P_A(x)$, indeterminada $I_A(x)$, negativa $N_A(x)$, indeterminada positiva $I_{P_A}(x)$ e indeterminada negativa $I_{N_A}(x)$. Cada una de ellas tiene un peso $w_m \in [0, 1]$ asociado. Para cada $x \in X$, hay $P_A(x), I_{P_A}(x), I_A(x), I_{N_A}(x), N_A(x) \in [0, 1]$, $w_P^m(P_A(x)), w_{I_P}^m(I_{P_A}(x)), w_I^m(I_A(x)), w_{I_N}^m(I_{N_A}(x)), w_N^m(N_A(x)) \in [0, 1]$ y $0 \leq P_A(x) + I_{P_A}(x) + I_A(x) + I_{N_A}(x) + N_A(x) \leq 5$. Por tanto, un CNITR A se puede representar por $A = \{ \langle x; P_A(x), I_{P_A}(x), I_A(x), I_{N_A}(x), N_A(x) \rangle | x \in X \}$.

Sean A y B dos CNITR de un universo de discurso finito, $X = \{x_1, x_2, \dots, x_n\}$, que se denotan por:

$$A = \{ \langle x; P_A(x), I_{P_A}(x), I_A(x), I_{N_A}(x), N_A(x) \rangle | x \in X \} \text{ y } B = \{ \langle x; P_B(x), I_{P_B}(x), I_B(x), I_{N_B}(x), N_B(x) \rangle | x \in X \},$$

Donde $P_A(x_i), I_{P_A}(x_i), I_A(x_i), I_{N_A}(x_i), N_A(x_i), P_B(x_i), I_{P_B}(x_i), I_B(x_i), I_{N_B}(x_i), N_B(x_i) \in [0, 1]$, para cada $x_i \in X$. Sea w_i ($i = 1, 2, \dots, n$) los pesos de los elementos x_i ($i = 1, 2, \dots, n$), con $w_i \geq 0$ ($i = 1, 2, \dots, n$) y $\sum_{i=1}^n w_i = 1$

La *distancia ponderada generalizada del CNITR* se define como sigue, ([18]):

$$d_\lambda(A, B) = \left\{ \frac{1}{5} \sum_{i=1}^n w_i \left[|P_A(x_i) - P_B(x_i)|^\lambda + |I_{P_A}(x_i) - I_{P_B}(x_i)|^\lambda + |I_A(x_i) - I_B(x_i)|^\lambda + |I_{N_A}(x_i) - I_{N_B}(x_i)|^\lambda + |N_A(x_i) - N_B(x_i)|^\lambda \right] \right\}^{1/\lambda} \quad (2)$$

Donde $\lambda > 0$.

La *Escala de Likert Indeterminada* consiste en los siguientes cinco elementos ([18]):

- Pertenencia negativa,
- Indeterminación tendiente hacia pertenencia negativa,
- Pertenencia indeterminada,
- Indeterminación tendiente hacia pertenencia positiva,
- Pertenencia positiva.

Por tanto, en esta escala la persona que responde debe dar cinco valoraciones en lugar de una, por cada una de las cinco anteriores. Para facilitar esto se utiliza una escala de 1 a 10, que luego es dividida por 10. Por ejemplo una opinión (8, 4, 2, 0,0) significa que el encuestado considera que lo que valora no lo hace como negativo, ni con indeterminación hacia lo negativo. Por otra parte considera que es indeterminado en un valor de 2, es indeterminado hacia lo positivo con un valor de 4 y es positivo en un valor de 8. Estos valores se dividen por 10 para llevarlos a una escala en $[0, 1]$ y se obtiene $\left(\frac{8}{10}, \frac{4}{10}, \frac{2}{10}, 0,0 \right) = \left(\frac{4}{5}, \frac{2}{5}, \frac{1}{5}, 0,0 \right)$.

Estos resultados se convierten en un escalar calculando 1 menos la distancia del valor anterior al valor ideal (1,0,0,0,0). Por ejemplo, $(\frac{4}{5}, \frac{2}{5}, \frac{1}{5}, 0, 0)$ se sustituye por el valor escalar $1 -$

$$d_1 \left(\left(\frac{4}{5}, \frac{2}{5}, \frac{1}{5}, 0, 0 \right), (1, 0, 0, 0, 0) \right) = 1 - \left(\left| 1 - \frac{4}{5} \right| + \left| \frac{2}{5} - 0 \right| + \left| \frac{1}{5} - 0 \right| \right) = 1 - \left(\frac{1}{5} + \frac{2}{5} + \frac{1}{5} \right) = \frac{1}{5}.$$

5. RESULTADOS.

Se analizaron los datos de la encuesta previa al curso y los resultados docentes durante el curso. Se utilizó el paradigma de la programación visual basada en Orange. Se Diseñó el siguiente *pipeline*.

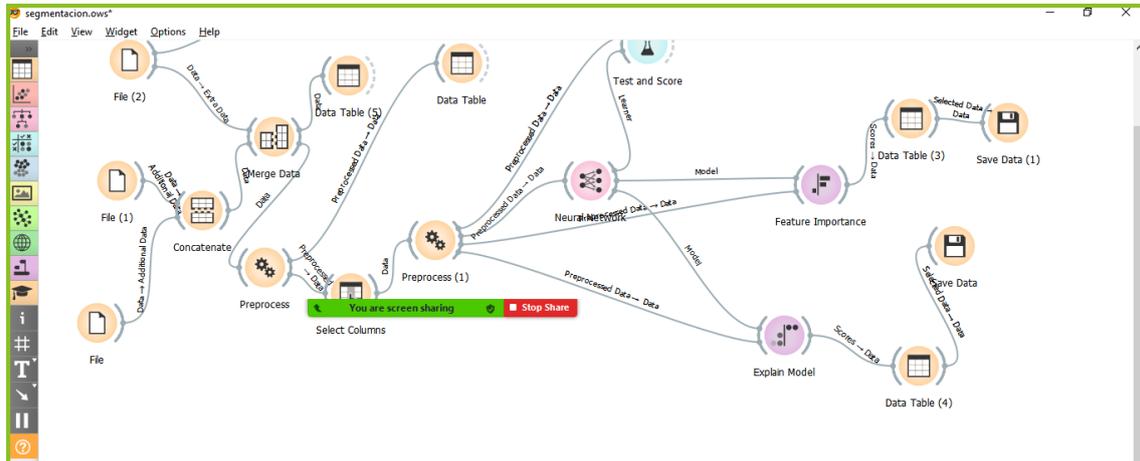


Figura 4. Pipeline Orange.

A continuación, se muestran los valores de las variables más relevantes en la predicción de la nota final del estudiante en el curso.

Característica	Puntuación
Habilidades y dedicación del profesor [El profesor era un formador eficaz]=Totalmente de acuerdo	0,095661247
Conocimientos adquiridos [Nivel de habilidades o conocimientos al principio del curso]=Muy bueno	0,081769594
Conocimientos adquiridos [Nivel de habilidades o conocimientos al principio del curso]=Satisfactorio	0,076621481
Habilidades y dedicación del profesor [El profesor era un formador eficaz]=De acuerdo	0,06153057
Conocimientos adquiridos [Nivel de habilidades o conocimientos al final del curso]=Excelente	0,041246531

Tabla 1. Resultados del índice de SHAP

A continuación, se pueden apreciar gráficamente los resultados:

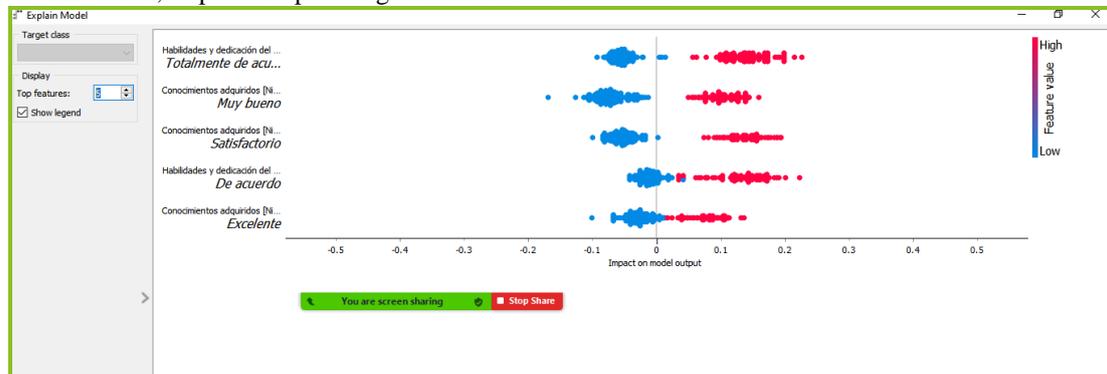


Figura 5. Resultados del índice de SHAP

Los resultados muestran que los factores están correlacionados positivamente con la nota final. Adicionalmente se analiza la importancia de los atributos.

Características	Media
Conocimientos adquiridos [Nivel de habilidades o conocimientos al final del curso]	0,035737
Conocimientos adquiridos [Nivel de habilidades o conocimientos necesarios para completar el curso]	0,0222
Nivel de esfuerzo que usted aplicó al curso [Nivel de esfuerzo que has dedicado al curso]	0,02084
Conocimientos adquiridos [Nivel de habilidades o conocimientos al principio del curso]	0,019961
Conocimientos adquiridos [¿En qué medida ha contribuido el curso a mejorar tus habilidades o conocimientos?]	0,018715

Tabla 2. Importancia de los atributos.

Los resultados de la Tabla 2 se muestran gráficamente en la Figura 3.

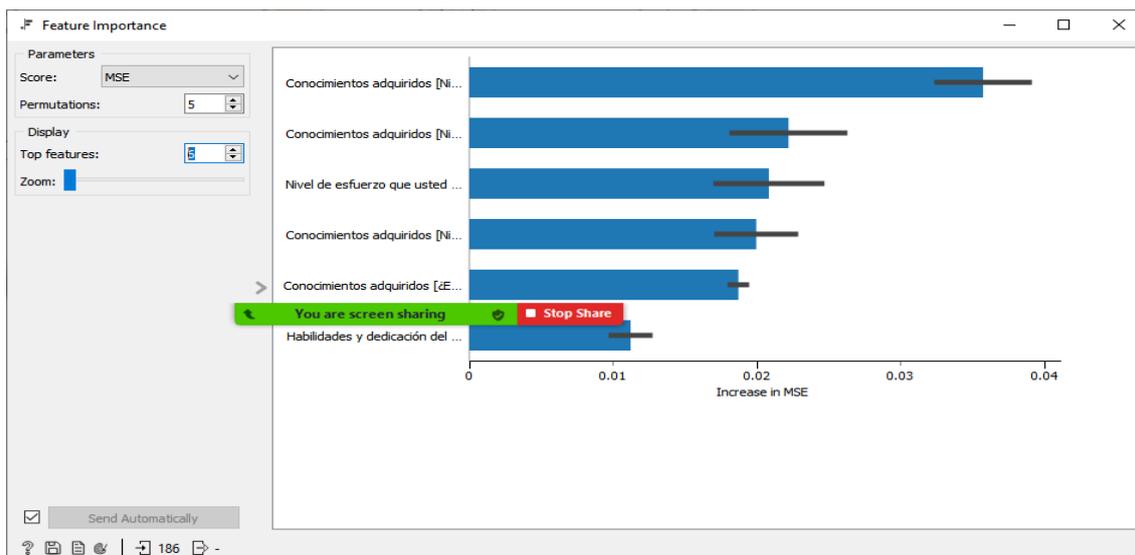


Figura 6. Importancia de los atributos.

Se puede apreciar que los 3 atributos más importantes están relacionados con los conocimientos adquiridos y el nivel de esfuerzo.

6. CONCLUSIONES

Resulta importante para el proceso de aprendizaje la implantación del análisis del aprendizaje para toma de decisiones en materia educativa. El objetivo del presente estudio se basa en el análisis de los datos obtenidos en la plataforma Moodle y su procesamiento en la herramienta Orange. La utilización del método CRIPS junto al uso del índice SHAP permitió descubrir los patrones que influyen en la evaluación final de los estudiantes. Se utilizó redes neuronales para predecir la evaluación final de los estudiantes, se puede apreciar que los 3 atributos más importantes están relacionados con los conocimientos adquiridos y el nivel de esfuerzo. Las encuestas aplicadas se evaluaron mediante una escala de Likert indeterminada que permitió capturar con mayor exactitud las opiniones de los estudiantes, incluyéndose las posibles inconsistencias. Como trabajos futuros se plantea a la utilización de técnicas de agrupamiento para segmentar los estudiantes.

RECEIVED: JULY, 2021.
REVISED: FEBRUARY, 2022.

REFERENCIAS

- [1] AMO, D. y SANTIAGO, R. (2017). Learning Analytics. La narración del aprendizaje a través de los datos. *Revista de Medios y Educación*, 52, 249-250.
- [2] ANDERSON, L. W. y KRATHWOHL, D. R. (2001). *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*. Allyn & Bacon. Boston, MA (Pearson Education Group).
- [3] ÁREA-MOREIRA, M., HERNÁNDEZ-RIVERO, V. y SOSA-ALONSO, J. J. (2016). Models of educational integration of ICTs in the classroom. *Comunicar. Media Education Research Journal*, 24, 79-87.

- [4] AVELLO MARTÍNEZ, R., & DUART, J. M. (2016). Nuevas tendencias de aprendizaje colaborativo en e-learning: Claves para su implementación efectiva. **Estudios pedagógicos (Valdivia)**, 42, 271-282.
- [5] BAKEMAN, R. (1978). Untangling streams of behavior: Sequential analysis of observation data. **Observing behavior**, 2, 63-78.
- [6] BARTON, L. E. y JOHNSON, H. A. (1990). Observational technology: An update. En **Ecobehavioral analysis and developmental disabilities** (pp. 201-227). Springer, New York.
- [7] BATISTA HERNÁNDEZ, N., LEYVA VÁZQUEZ, M.Y., GONZÁLEZ CABALLERO, E., VALENCIA CRUZATY, L.E., ORTEGA CHÁVEZ W. y SMARANDACHE, F. (2021). A new method to assess entrepreneurship competence in university students using based on plithogenic numbers and SWOT analysis. **International Journal of Fuzzy Logic and Intelligent Systems**, 21, 280-292.
- [8] BOSSOLASCO, M. L., CASANOVA, B. A., ENRICO, E. E., SANTOS, D. A. D. y ENRICO, R. J. (2018). Perfiles de apropiación de TIC y desempeño académico en estudiantes universitarios. Un estudio de caso utilizando analítica del aprendizaje. **Virtualidad, Educación y Ciencia**, 9, 44-61.
- [9] BUCKINGHAM, S. y FERGUSON, R. (2012). Social learning analytics. **Journal of educational technology & society**, 15, 3-26.
- [10] CALVET LIÑÁN, L. y JUAN PÉREZ, Á. A. (2015). Educational Data Mining and Learning Analytics: Differences, similarities, and time evolution. **International Journal of Educational Technology in Higher Education**, 12, 98-112.
- [11] COLL, C., ENGEL, A. y NIÑO, S. (2017). La actividad de los participantes como fuente de información para promover la colaboración. Una analítica del aprendizaje basada en el modelo de Influencia Educativa Distribuida. **Revista de Educación a Distancia (RED)**, 53, 1-36.
- [12] COSTA, E., BAKER, R., AMORIM, L., MAGALHÃES, J. y MARINHO, T. (2013). Mineração de dados educacionais: Conceitos, técnicas, ferramentas e aplicações. **Jornada de Atualização em Informática na Educação**, 1, 1-29.
- [13] CRUZ PÉREZ, M. A. (2020). Scientific content in research training through ICT in university students. **E-Ciencias de la Información**, 10, 136-158.
- [14] FAYYAD, U. M., PIATETSKY-SHAPIO, G., SMYTH, P. y UTHURUSAMY, R. (1996). **Advances in knowledge discovery and data mining**. Cambridge: AAAI/MITPress.
- [15] GAGNÉ, R. M. (1975). Principios básicos del aprendizaje para la instrucción. México D.F. : Diana.
- [16] JOHNSON, L., ADAMS, S. y CUMMINS, M. (2012). **The NMC horizon report 2012 higher education**. Austin, TX: The New Media Consortium.
- [17] JOHNSON, L., BECKER, S. A., CUMMINS, M., ESTRADA, V., FREEMAN, A. y HALL, C. (2016). **NMC horizon report: 2016 higher education edition**. Austin, TX: The New Media Consortium.
- [18] KANDASAMY, I., KANDASAMY, W. B. V., OBBINENI, J.M. y SMARANDACHE, F. (2020) Indeterminate Likert scale: feedback based on neutrosophy, its distance measures and clustering algorithm. **Soft Computing**, 24, 7459-7468.
- [19] KANDASAMY, I. y SMARANDACHE, F. (2016) Triple refined indeterminate neutrosophic sets for personality classification. En: **2016 IEEE Symposium Series on Computational Intelligence (SSCI)**. IEEE. 2016.
- [20] LIKERT, R. (1932). A technique for the measurement of attitudes. **Archives of psychology**, 22, 5-55.
- [21] MANTURANO PÉREZ, T.U., ORTEGA CHÁVEZ, W., VÁSQUEZ PÉREZ, J.K., SANGAMA SÁNCHEZ, J.L. (2021). Determinación de los criterios a medir para efectuar la gestión del talento de liderazgo en las instituciones peruanas de Educación Superior. **Revista Investigación Operacional**, 42, 425-432.
- [22] MARÍN, F. V., INCIARTE, A. DE J., HERNÁNDEZ, H. G. y PITRE, R. C. (2017). Estrategias de las Instituciones de Educación Superior para la Integración de las Tecnología de la Información y la Comunicación y de la Innovación en los Procesos de Enseñanza. Un Estudio en el Distrito de Barranquilla, Colombia. **Formación universitaria**, 10, 29-38.
- [23] MORA, J. G. (2004). La necesidad del cambio educativo para la sociedad del conocimiento. **Revista Iberoamericana de educación**, 35, 13-37.
- [24] RODRÍGUEZ, E. (2009). Ventajas e inconvenientes de las TIC en el aula. **Cuadernos de Educación y Desarrollo**, 1, 45-61.
- [25] ROJAS, P. (2017). Learning Analytics: Una revisión de la literatura. **Educación y Educadores**, 20, 406-127.
- [26] SIEMENS, G. (2013). Learning analytics: The emergence of a discipline. **American Behavioral Scientist**, 57, 1380-1400.
- [27] SLADE, S. y PRINSLOO, P. (2013). Learning analytics: Ethical issues and dilemmas. **American Behavioral Scientist**, 57, 1510-1529.
- [28] SMARANDACHE, F. (2002). **Neutrosophy, a new Branch of Philosophy**. Craiova: Infinite Study.
- [29] SMARANDACHE, F. (2013) **n-valued refined neutrosophic logic and its applications in physics**. Craiova: Infinite Study.
- [30] SMARANDACHE F., ESTUPIÑÁN RICARDO, J., GONZÁLEZ CABALLERO, E., LEYVA VÁZQUEZ M. Y. and BATISTA HERNÁNDEZ, N. (2020). Delphi method for evaluating scientific research proposals in a neutrosophic environment. **Neutrosophic Sets and Systems**, 34, 204-213.

[31] VALVERDE BERROCOSO, J. (2018). **Campus digitales en la Educación Superior. Experiencias e investigaciones.** Universidad de Extremadura, Servicio de Publicaciones.