

Artículo de Investigación

Factores de la modalidad en línea-emergente que inciden en el sistema de educación superior del Ecuador

Factors of the emerging online learning modality affecting Ecuador's higher education system.

María Isabel Uvidia-Fassler¹, Andrés Cisneros-Barahona¹, Pablo Méndez-Naranjo¹, Henry Villa-Yáñez¹

¹Universidad Nacional de Chimborazo, Riobamba, Ecuador, 060108;
ascisneros@unach.edu.ec; pmendez@unach.edu.ec; hvilla@unach.edu.ec

*Correspondencia: muvidia@unach.edu.ec

Citación: Uvidia-Fassler, M., Cisneros-Barahona, A., Méndez-Naranjo, P. & Villa-Yáñez, H., (2024). Factores de la modalidad en línea-emergente que inciden en el sistema de educación superior del Ecuador. *Novasinerгия*. 7(2). 87-114.

<https://doi.org/10.37135/ns.01.14.06>

Recibido: 11 abril 2024

Aceptado: 3 junio 2024

Publicado: 03 julio 2024

Novasinerгия
ISSN: 2631-2654



Copyright: 2024 derechos otorgados por los autores a Novasinerгия.

Este es un artículo de acceso abierto distribuido bajo los términos y condiciones de una licencia de Creative Commons Attribution (CC BY NC). (<http://creativecommons.org/licenses/by/4.0/>).

Resumen: La pandemia por COVID-19 influyó en el proceso de educación superior generando una transición entre la modalidad presencial a la modalidad en línea-emergente en la que se identificaron varias dificultades para el acceso debido a la inequidad, problemas tecnológicos, crisis de salud y económica a nivel mundial, y en Ecuador a nivel de docentes y estudiantes. El descubrimiento de conocimiento en bases de datos de KDD – *Educacional*, el análisis de técnicas de minería de datos de clasificación y agrupación, y el complemento de análisis de datos con la inteligencia de negocios determinaron los factores desde la perspectiva de la institución, del docente y de los estudiantes que inciden en el proceso de educación en línea, donde se demuestra además, la importancia del modelo de educación, la metodología del proceso de enseñanza aprendizaje, el entorno de aprendizaje y el bienestar de los docentes y estudiantes que garantizan una educación superior en línea de calidad con igualdad de oportunidades en el acceso.

Palabras clave: COVID-19, Descubrimiento de conocimiento en bases de datos, Educación superior, Factores en educación en línea, Minería de datos educacional.

Abstract: *The COVID-19 pandemic significantly impacted higher education, forcing a rapid transition from face-to-face modality to the emerging online learning modality. This transition exposed several challenges, including, inequitable access to technology, global health and economic crisis, and specific difficulties faced by both teachers and students in Ecuador. Knowledge discovery in KDD – Educational databases, the analysis of classification and clustering data mining techniques and the complement of data analysis with business intelligence determined the factors from the perspective of the institution, teachers and students that influence the online education process. In addition, the importance of the education model, the teaching-learning process methodology, the learning environment, and the well-being of teachers and students that guarantee quality online higher education with equal access opportunities is demonstrated.*

Keywords: COVID-19, Knowledge discovery in databases, Higher education, Educational factors, Educational data mining.

1. Introducción

La base de la sociedad es la educación y, además, es el arma más poderosa que puedes usar para cambiar el mundo (Mandela, 2014). Casi la mitad de los estudiantes del mundo han sido afectados por el cierre parcial o total de las escuelas, y más de 100 millones de niños no podrán alcanzar el nivel mínimo en lectura por la crisis sanitaria (UNESCO, 2021).

La aparición de la expansión de la enfermedad de coronavirus (COVID-19) apareció por primera vez en Wuhan (China) el 31 de diciembre de 2019 y fue notificado por la Organización Mundial de la Salud (OMS) en marzo de 2020. La OMS rápidamente comunicó la información científica sobre el nuevo virus solicitando al mundo tomar las medidas de protección y prevención para evitar la propagación del brote y precautelar la salud (Organización Mundial de la Salud, 2020). Hasta el 20 de abril la UNESCO determinó el cierre del 91.3% de escuelas afectando a la población estudiantil mundial, es decir, 1,575,270,054 millones de estudiantes (UNESCO y IESALC, 2020a).

Las universidades como cualquier otra institución social han tenido que enfrentarse a diferentes epidemias que han interrumpido su funcionamiento cotidiano. La Universidad de Cambridge en 1665 cerró sus puertas por la peste negra que golpeó a Inglaterra, por tal motivo Isaac Newton tuvo que volver a su hogar en Woolsthorpe Manor, y sentado desde su jardín, vio caer la manzana que le permitió generar la teoría de la gravitación universal, o por lo menos, eso es lo que narra William Stukeley. De esta manera, se demuestra que las instituciones de educación superior pueden dejar de funcionar presencialmente, pero las actividades continúan con la gente comprometida con la ciencia y la formación, siendo así que Cambridge cierra por segunda vez sus puertas en el año 2020 (UNESCO y IESALC, 2020b).

La educación tiene varios retos como el crecimiento sin calidad, inequidades en el acceso y los logros y la pérdida progresiva del financiamiento público, que sumado a la pandemia por COVID-19 ha generado mayores problemas a la educación superior de América Latina y el Caribe (UNESCO y IESALC, 2020c). Pero esta crisis también ha generado oportunidades para la educación virtual, creando un incremento inesperado, permitiendo un avance importante para el XXI, en la era de la universalización de los conocimientos de educación superior en el mundo (UNESCO y IESALC, 2020d).

En Ecuador, debido al estado de excepción por la pandemia por COVID-19, el Consejo de Educación Superior determinó la Normativa transitoria para el desarrollo de actividades académicas en las Instituciones de Educación Superior, mencionando que “las disposiciones que tienen por objeto garantizar el derecho a la educación de los estudiantes y la consecuente ejecución de la oferta académica vigente de todas las instituciones de educación superior (IES), debido al estado de excepción decretado por la emergencia sanitaria que rige en el territorio nacional” (CES, 2020).

La Universidad Nacional de Chimborazo (UNACH), universidad pública ubicada en la provincia de Chimborazo, con 12.000 estudiantes distribuidos en 31 carreras presenciales de pregrado, mediante el Consejo Universitario en marzo de 2020 aprueba las “Directrices para la aplicación en la Universidad Nacional de Chimborazo de la normativa transitoria emitida por el CES, para el desarrollo de actividades académicas en las instituciones de educación

superior, debido al estado de excepción decretado por la emergencia sanitaria ocasionada por la pandemia de COVID-19" (UNACH, 2020), permitiendo que docentes y estudiantes puedan continuar con las actividades académicas para garantizar una educación superior de calidad.

La crisis económica y sanitaria por COVID-19 hizo que obligatoriamente se cree una modalidad en línea en muchas universidades que contaban con carreras únicamente en modalidad presencial, por tal motivo el autor Roberto Bustamante plantea que la estrategia temporal se convierte en permanente, generándose una Educación Remota de Emergencia (Bustamante, 2020).

En cada país se vivieron realidades diferentes, por ejemplo en Ecuador, 700 mil estudiantes universitarios pasaron de tener clases presenciales, a recibirlas a través del internet (Calvopiña, 2020), el 53,2% de los hogares ecuatorianos si tiene acceso a internet, el 70,7% de las personas utilizan internet y el 81,8% utiliza un teléfono celular inteligente, según la encuesta TIC-2020 INEC (Instituto Nacional de Estadística y Censos, 2021), siendo datos importantes para la realidad de la educación en todos los niveles.

"El Descubrimiento de Conocimiento en Bases de Datos (*KDD*), es el proceso no trivial de identificar patrones en los datos que sean válidos, novedosos, potencialmente útiles y, por último, comprensibles" (Fayyad et al., 1996), además es un análisis automático, exploratorio y de modelado de los repositorios de datos de gran tamaño, que los autores Maimon y Rokach (2015) definen a *KDD* como un proceso organizado de identificación válida, novedosa y útil que genera patrones comprensibles de los conjuntos de datos grandes y complejos, es iterativo e interactivo y se define en nueve pasos.

Los sistemas de minería de datos (*DM*) como redes neuronales, sistemas expertos, programación genética y árboles de decisión se desarrollan bajo lenguajes de última generación basados en Inteligencia Artificial y utilizan métodos matemáticos (Bernabeu, 2010), adicionalmente, es vital considerar que las herramientas OLAP (*Online Analytical Processing*: procesamiento analítico en línea) y en los reportes y consultas, el análisis parte de una pregunta o hipótesis generada por los usuarios, mientras que *DM* genera estas hipótesis.

Las herramientas de *DM* generalmente se integran a bases ya existentes como *Data Warehouse* (*DW*), permitiendo incrementar el valor de las fuentes de datos para que incluso puedan integrarse a otros productos como OLAP, siendo la limpieza de datos y la integración partes importantes de esta construcción (Bernabeu, 2010).

Los métodos de clasificación como rough sets, árboles de decisión, redes neuronales, redes bayesianas, algoritmos genéticos entre otros encuentran propiedades comunes entre los objetos de una base de datos. El modelo de árboles de decisión es de los más utilizados, popular por su simplicidad y facilidad para entender, además, la agrupación de objetos físicos o abstractos se llama segmentación, *clustering* o clasificación no supervisada y comparten un número de propiedades por lo que son considerados homogéneos (Jiménez y Timarán, 2015).

Un complemento importante es la Inteligencia de Negocios (*BI*, *Business Intelligence*) que permite almacenar y procesar gran cantidad de datos, con el objetivo de transformarlos en

conocimiento y en decisiones en tiempo real, a través de un sencillo análisis y exploración (Bernabeu, 2010).

A partir de los datos generados desde la modalidad presencial y su transición por la pandemia por COVID-19, durante este estudio se determinan los factores que influyen en el modelo de educación superior en línea.

2. Metodología

El alcance de la investigación abarca la educación superior durante el periodo de emergencia sanitaria causada por la pandemia de COVID-19, con un enfoque específico en Ecuador. Examina las desigualdades en el acceso a la educación en línea, problemas tecnológicos, y crisis de salud y económica. Se incluye el análisis de datos educativos utilizando técnicas de minería de datos y la inteligencia de negocios para determinar los factores clave que afectan la educación en línea.

La investigación se enfoca en el impacto de la pandemia de COVID-19 en la educación superior, específicamente en la transición de la modalidad presencial a la modalidad en línea. Se centra en identificar y analizar las dificultades y factores que influyen en el acceso y la calidad de la educación en línea desde múltiples perspectivas: institucional, docente y estudiantil.

El diseño de la investigación es de tipo descriptivo y analítico, utilizando un enfoque cuantitativo para el descubrimiento de conocimiento en bases de datos (*KDD*) en el contexto educativo. Se emplean técnicas de minería de datos, como la clasificación y la agrupación, para analizar los datos. Además, se complementa con análisis de inteligencia de negocios para identificar los factores determinantes que inciden en el proceso de educación en línea. El estudio considera diversas variables relacionadas con el modelo educativo, la metodología del proceso de enseñanza-aprendizaje, el entorno de aprendizaje y el bienestar de los docentes y estudiantes.

La investigación se desarrolló mediante una metodología mixta: exploración cuantitativa, además del análisis cualitativo de las experiencias y percepciones de los actores del proceso de educación reflejada en los patrones obtenidos de los datos. La metodología aplicada para *KDD* en este trabajo es el resultado de investigaciones previas como de los autores Moine et al. (2011), que proponen metodologías de *KDD* con una secuencia de pasos como: *CRISP-DM*, *SEMMA* y *Catalyst*.

La metodología *CRISP-DM* es la más usada y de fácil implementación, mientras que *SEMMA* tiene enfoques técnicos que no incluye al análisis del negocio, siendo una parte muy importante para definir la muestra de datos con lo que se va a trabajar, y *Catalyst* (*P3TQ*) no explica en alto nivel de detalle las tareas a realizarse.

Por tal motivo, con el análisis de las metodologías y con la base de las fases genéricas del proceso de *KDD*, la metodología "Descubrimiento de Conocimiento en Base de Datos Educativos – *KDD – Educativa*" (Figura 1) es la metodología producto de investigaciones de los autores Uvidia et al. (2018) y Cisneros et al. (2021).

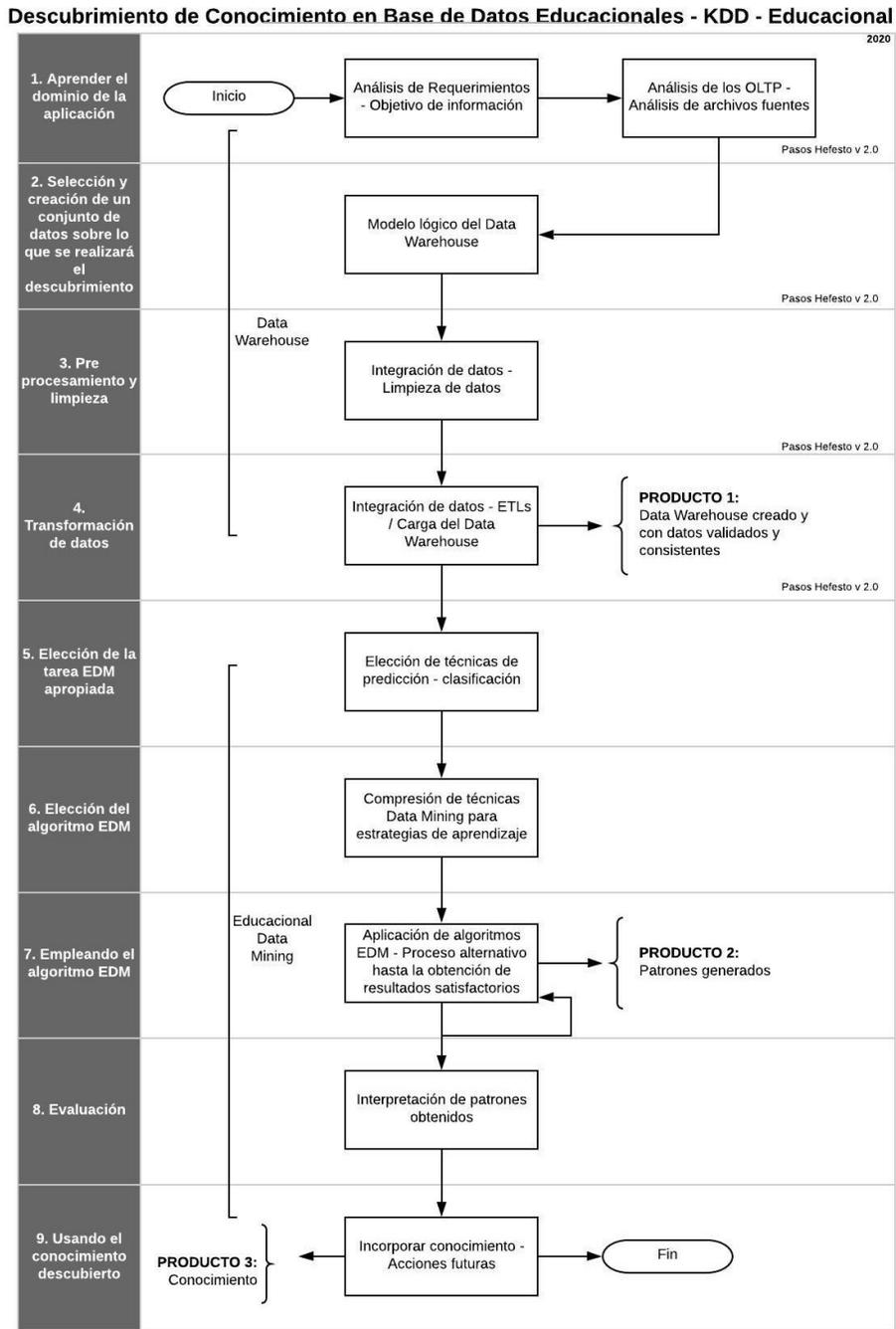


Figura 1: Metodología KDD – Educativa

La Figura 1 muestra la metodología de Descubrimiento de Conocimiento en Base de Datos Educativas - KDD – Educativa con 9 pasos que identifican el objetivo del problema, define las fuentes de datos a utilizar, selecciona las técnicas de minería de datos de predicción y clasificación, genera patrones y principalmente conocimiento para la aplicación de acciones y estrategias.

“Aprender el dominio de la aplicación” es el primer paso en el que se definen todos los requerimientos de información, identificando el área de aplicación o escenario en el que se

tomarán decisiones, además, se define el objetivo, las características de los datos, fuentes y atributos. La siguiente matriz recopila los datos como se muestra en la Tabla 1. Esta tabla contiene la información de los archivos OLTP (*OnLine Transaction Processing*: procesamiento de transacciones en línea).

Tabla 1: Matriz Aprender el dominio de la aplicación

PROCESO:				
N°	Requerimiento	Descripción	Características	Fuente

Según Uvidia et al. (2018) y Cisneros et al. (2021), es vital el primer paso ya que determina los datos que van a generar conocimiento.

“Selección y creación de un conjunto de datos para el descubrimiento” es la segunda fase que permite completar la información adicional para crear el *Data Warehouse*, *Data Mart* o *dataset* de datos. En este paso se define el Modelo Lógico de DW, tipo de esquema, dimensiones, tablas de hechos (*Fact*) y las relaciones entre estas. La matriz de la tabla 2, permite determinar todas las características de los datos fuente: atributos, tipos de datos, columnas, claves principales de las tablas y continuar con el diseño y su modelo conceptual si es un *DW* o *Data Mart*.

Tabla 2: Matriz Selección y Creación DW

PROCESO:						
N°	Requerimiento	Descripción	Características	Dataset	Tipo de Dato	Fuente

“Pre procesamiento y limpieza” es la tercera fase que permite asegurar la consistencia y confiabilidad de la información que se almacena en el DW, siendo el principal objetivo recabar, validar y analizar datos que pueden ser atípicos y faltantes, ya que los datos deben ser correctos y confiables para reflejar la realidad de la empresa y se determine una línea base de información consistente y adecuada para la toma de decisiones.

“Transformación de datos” es la cuarta fase de la metodología y genera el primer producto que es el DW. Se aplica la técnica de Integración de Información: ETL (Extracción, Transformación y Carga), mediante la cual se realiza la carga la información desde las fuentes al DW, con datos correctos y tipos de datos adecuados (transformación).

Al iniciar el subproceso DM, en la quinta fase “Elección de la tarea de DM apropiada”, se selecciona el tipo de algoritmo DM a aplicar, ya sea clasificación, regresión, agrupación o predicción. La selección del algoritmo depende de los objetivos *KDD*, y sus dos objetivos principales son: la predicción y la descripción. La predicción son técnicas supervisadas, mientras que la descripción son técnicas no supervisadas (Maimon y Rokach, 2015).

En la sexta fase “Elección del algoritmo de DM”, se escoge el algoritmo de acuerdo con la precisión o porcentaje de confiabilidad del algoritmo frente a la comprensión de los resultados. Es importante incluir desde los algoritmos más simples hasta los más complejos como: árboles de decisión, redes bayesianas, regresiones lineales, optimización de secuencia mínima, u otros, para interpretar patrones y generar conocimiento.

“Empleando el algoritmo de DM” es la séptima fase en la que se emplea la herramienta de DM, como R, WEKA, RapidMiner, Knime, entre otras. Este es un proceso iterativo ya que se repite hasta encontrar resultados satisfactorios que cumpla con los objetivos planteados en la primera fase. Además, se obtiene el segundo producto que son los patrones para evaluar.

“Evaluación” es la octava fase en la que se valoran e interpretan los patrones obtenidos (reglas, fiabilidad), se revisan los resultados de DM, su comprensión y la utilidad del modelo frente al conocimiento descubierto. Los patrones obtenidos se almacenan para posteriores análisis.

En la última fase “Usando el conocimiento descubierto”, se incorpora el conocimiento para acciones futuras, permita tener un sentido activo para medir efectos y aplicar soluciones, determinando el éxito del proceso KDD. El conocimiento es el tercer producto generado, que constituye la base para que se implementen planes de mejora, guías o proyectos, y asegure la oportuna toma de decisiones.

2.1. Escenario del estudio

La Universidad Nacional de Chimborazo con 12.000 estudiantes aproximadamente, tiene varias herramientas tecnológicas que permiten desarrollar las actividades académicas que constituyen el escenario que se muestra en la Figura 2.



Figura 2: Sistema Académico SICOA y las aulas virtuales Moodle

En la Figura 2 se muestran las principales herramientas de la UNACH en las cuales se desarrollan la modalidad presencial y en línea.

Sistema Académico SICOA: Es el sistema informático de control académico que contiene y registra la información de los estudiantes como: calificaciones y matrículas, además, información de docentes como: sílabos, evaluación integral, distributivos, capacitación, avances académicos, horarios de clases y diferentes encuestas dirigidas a la comunidad universitaria. SICOA es un sistema robusto con datos académicos de la UNACH.

Aulas virtuales Moodle: Es una plataforma de aprendizaje con ambientes personalizados diseñada para docentes y estudiantes (administrado por la UNACH).

2.2. Periodos académicos de análisis

En la UNACH, cada año se desarrollan dos periodos académicos semestrales. Esta investigación realiza el análisis desde el año 2018 hasta el 2019 con actividades en modalidad presencial, y los periodos académicos a partir del año 2020 hasta el 2021 con modalidad en línea, en la Tabla 3 se muestran los periodos académicos considerados.

Tabla 3: Descripción de los periodos académicos de análisis

AÑO	PERIODO ACADÉMICO	DESCRIPCIÓN
2018	2018-1S	Período académico comprendido entre marzo – agosto 2018
2018	2018-2S	Período académico comprendido entre septiembre 2018 – febrero 2019
2019	2019-1S	Período académico comprendido entre marzo – agosto 2019
2019	2019-2S	Período académico comprendido entre septiembre 2019 – febrero 2020
2020	2020-1S	Período académico comprendido entre marzo – agosto 2020
2020	2020-2S	Período académico comprendido entre septiembre 2020 – febrero 2021
2021	2021-1S	Período académico comprendido entre marzo – agosto 2021

Durante la investigación se encontraba en curso el periodo académico 2021-1S, por lo tanto, algunos de los resultados de este año no se muestran ya que varios procesos alcanzan sus resultados al cierre del periodo académico.

Las herramientas utilizadas para esta investigación son:

- *Weka*: es un software de código abierto emitido bajo la Licencia Pública General *GNU*, es una colección de algoritmos de aprendizaje automático para tareas de minería de datos que contiene herramientas para la preparación de datos, clasificación, regresión, agrupamiento, minería de reglas de asociación y visualización (University of Waikato, 2021).
- *RapidMiner*: es un software que tiene capacidades de mejora de la productividad para la preparación automatizada de datos y el desarrollo de modelos en el mercado, así como herramientas visuales integrales para generar datos y canalizaciones de aprendizaje automático. *RapidMiner* es fácil de usar y tiene el rigor de la ciencia de datos y ha sido líder en *The Forrester Wave* (*Rapidminer*, 2021).
- *Microsoft Power BI*: es la plataforma enfocada en la Inteligencia de Negocios (BI) que es fácil de usar y permite obtener conocimientos sobre los datos mediante informes de análisis para la empresa (Microsoft, 2021).

3. Resultados

Posterior a la aplicación de la metodología en la UNACH, se muestran los resultados obtenidos en cada fase.

3.1. Fase 1. Aprender el dominio de la aplicación

La Tabla 4 muestra los resultados obtenidos.

Tabla 4: Matriz Aprender el dominio de la aplicación – resultados

PROCESO: Desarrollo académico de la UNACH				
N°	Requerimiento	Descripción	Características	Fuente
1	Notas de los estudiantes por componente de evaluación	Los datos analizados por periodo académico pertenecen a los períodos de las modalidades presencial y en línea debido a la pandemia	Datos desde el año 2018 al 2020	SICOA
2	Resultados de la evaluación integral docente	Los datos analizados por periodo académico pertenecen a los períodos de las modalidades presencial y en línea debido a la pandemia	Datos desde el año 2018 al 2020	SICOA
3	Desarrollo de las actividades académicas en las aulas virtuales	Los datos analizados de las conexiones a las aulas virtuales por periodo académico pertenecen a los períodos de la modalidad en línea debido a la pandemia	Datos desde el año 2020 al 2021	MOODLE
4	Resultados de la Encuesta para estudiantes sobre la situación actual durante la emergencia sanitaria	Los datos analizados por periodo académico pertenecen a los períodos de la modalidad en línea debido a la pandemia	Datos desde el año 2020 al 2021	SICOA

Según Uvidia et al. (2018) y Cisneros et al. (2021), el aprender el dominio de la aplicación identifica los requerimientos de la información y las fuentes de donde se obtienen datos para la UNACH.

3.2. Fase 2. Selección y creación de un conjunto de datos para el descubrimiento

La Tabla 5 muestra el resultado de la segunda fase con los tipos de datos a manejar.

Tabla 5: Matriz Selección y Creación DW – resultados

PROCESO: Desarrollo académico de la UNACH						
	Requerimiento	Descripción	Características	Dataset	Tipo de Dato	Fuente
1	Notas de los estudiantes por componente de evaluación	Los datos analizados por periodo académico pertenecen a los períodos de las modalidades presencial y en línea debido a la pandemia	Datos desde el año 2018 al 2020	NOTAS	Numéricas	SICOA
2	Resultados de la evaluación integral docente	Los datos analizados por periodo académico pertenecen a los períodos de las modalidades presencial y en línea debido a la pandemia	Datos desde el año 2018 al 2020	RESULTADOS DE LA EVALUACIÓN	Numéricas	SICOA
3	Desarrollo de las actividades académicas en las aulas virtuales	Los datos analizados de las conexiones a las aulas virtuales por periodo académico pertenecen a los períodos de modalidad en línea debido a la pandemia	Datos desde el año 2020 al 2021	RESULTADOS DEL USO DE AULAS VIRTUALES	Numéricas, categóricas	MOODLE
4	Resultados de la Encuesta para estudiantes sobre la situación actual durante la emergencia sanitaria	Los datos analizados por periodo académico pertenecen a los períodos de modalidad en línea debido a la pandemia	Datos desde el año 2020 al 2021	RESULTADOS DE LA ENCUESTA	Numéricas, categóricas	SICOA

Según Uvidia et al. (2018) y Cisneros et al. (2021), la selección y creación de un conjunto de datos para el descubrimiento identifica de manera puntual los elementos del repositorio para el análisis de la UNACH.

3.3. Fase 3. Pre procesamiento y limpieza

La tercera fase, asegura la extracción de datos de sistemas confiables con integridad de la información, y se convierte en una especie de auditoría ya que se analizan los datos que están siendo generados, y para la UNACH no existieron datos atípicos, erróneos o incompletos, es decir, la UNACH maneja datos consistentes. Finalmente, no fue necesario realizar limpieza de datos en las fuentes seleccionadas.

3.4. Fase 4. Transformación de datos

Como resultado de esta fase se obtuvieron para el análisis los siguientes conjuntos de datos que se muestran en la Tabla 6.

Tabla 6: *Datasets* obtenidos – productos

No.	DATASET	DESCRIPCIÓN	PROTECCIÓN DE DATOS	CANTIDAD DE REGISTROS
1	Notas	<i>Dataset</i> resultante que contiene las notas de los estudiantes desde el año 2018 hasta el año 2020	No existen datos personales, sino únicamente información de las asignaturas, asistencias y notas. No se utilizaron técnicas de anonimización.	13871
2	Evaluación integral docente	<i>Dataset</i> resultante que contiene los resultados de la evaluación integral docente desde el año 2018 hasta el año 2020	No existen datos personales de los docentes, sino únicamente información consolidada por unidades académicas (facultades y nivelación). No se utilizaron técnicas de anonimización.	100
3	Resultados de encuesta	<i>Dataset</i> resultante que contiene los resultados de la encuesta aplicada por emergencia sanitaria desde el año 2020 hasta el año 2021	No existen datos personales de los estudiantes encuestados, sino únicamente información consolidada por pregunta. No se utilizaron técnicas de anonimización.	562.708

Los *datasets* resultantes tienen datos de acuerdo con objetivo definido en la primera fase para el caso de la UNACH.

3.5. Fase 5. Elección de la tarea de DM apropiada

Con los *datasets* de análisis definidos se determinaron las técnicas de minería de datos adecuadas para obtener los patrones que serán analizados mediante Weka y *RapidMiner*, las técnicas seleccionadas son:

- Algoritmos de Clasificación: J48, Redes Bayesianas, *Random Tree*, *Random Forest* y *Hoeffding Tree*.
- Algoritmos de Agrupación: *kMeans*, *FarthestFirst*.

El análisis de los datos para este estudio se centra en el *dataset* Notas por ser la parte esencial de la universidad en el área académica y por estar completa desde el año 2018 al 2020.

En la Tabla 7 se muestran los datos para los algoritmos de minería de datos seleccionados.

Tabla 7: *Dataset* Notas y sus campos

COLUMNA	TIPO DE DATO	VALORES
---------	--------------	---------

Año	Cuantitativa - Discreta	2018 2019 2020
Facultad	Cualitativa – Nominal	EDUCACIÓN INGENIERÍA NIVELACIÓN POLIÍTICAS Y ADMINISTRATIVAS SALUD
Promedio Asignatura	Cuantitativa - Continua	0,1,2,3,4,5,6,7,8,9,10

El *dataset* de análisis contiene los campos para clasificar el rendimiento académico de los estudiantes de la UNACH.

3.6. Fase 6. Elección del algoritmo de DM

En la sexta fase, se desarrolla cada algoritmo escogido:

3.6.1. Algoritmos de Clasificación

Para determinar el rendimiento académico de los estudiantes durante los periodos de las modalidades presencial y en línea se realizaron los algoritmos de clasificación mediante árboles. Se analizaron varios algoritmos de clasificación, para identificar los mejores resultados.

En la Figura 3 se muestra la ejecución de todos los algoritmos de clasificación.



Figura 3: Algoritmos de clasificación ejecutados

Mediante la ejecución de varios algoritmos de clasificación se determina el mejor.

3.6.2. Algoritmos de Agrupación

Mediante los algoritmos de agrupación los grupos de notas de los estudiantes para conocer el rendimiento académico durante los periodos de las modalidades presencial y en línea. Se analizaron dos algoritmos de clasificación, con el objetivo de escoger el mejor resultado.

En la Figura 4 se puede observar la ejecución de los dos algoritmos de agrupación.

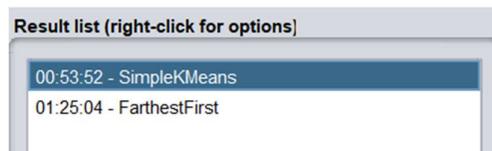


Figura 4: Algoritmos de agrupación ejecutados

Mediante la ejecución de los dos algoritmos de agrupación se determina el mejor.

3.7. Fase 7. Empleando el algoritmo de DM

La ejecución de los algoritmos de clasificación y agrupación en el *dataset* de notas con *Weka* permitió obtener el segundo producto que son los patrones para evaluar en la siguiente fase para alcanzar el conocimiento.

3.7.1. Algoritmos de Clasificación

En la Figura 5 se observa la ejecución del algoritmo de árbol de decisión J48.

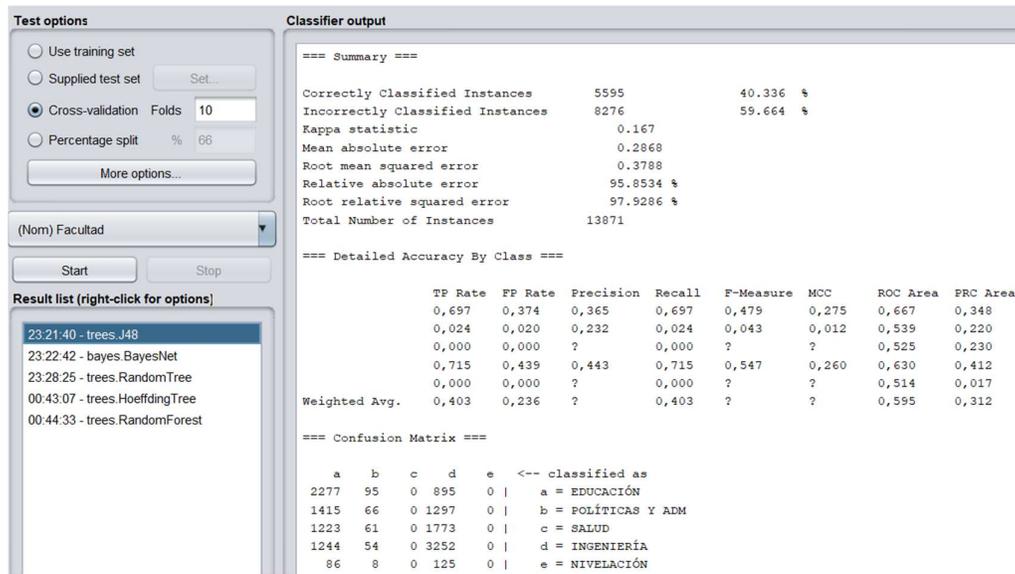


Figura 5: Algoritmo de clasificación J48 ejecutado

El algoritmo J48 al ser ejecutado tiene el 40,34% de instancias correctamente clasificadas.

En la Figura 6 se muestra la ejecución del algoritmo de redes Bayesianas.

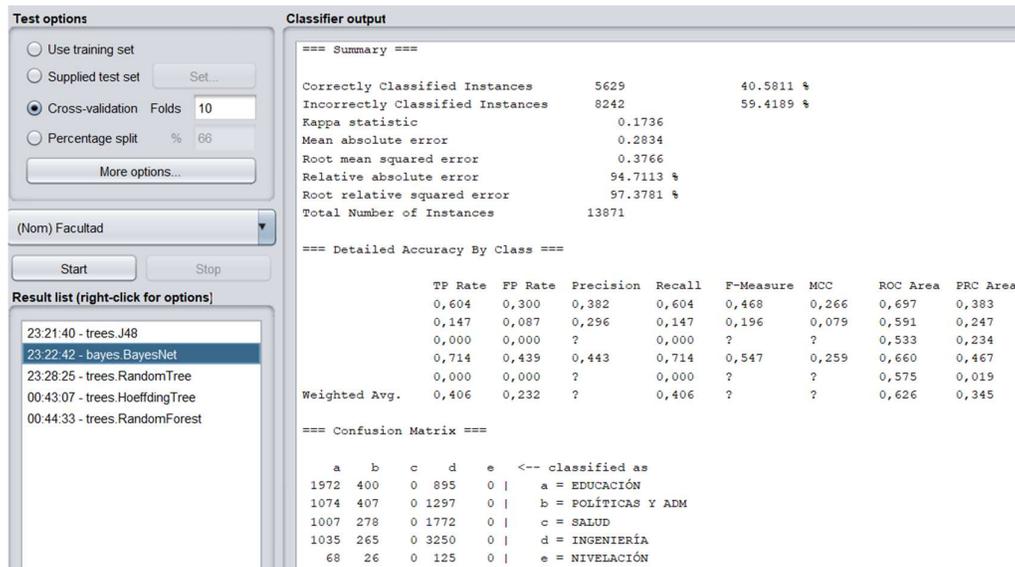


Figura 6: Algoritmo de clasificación Redes Bayesianas ejecutado

El algoritmo de redes Bayesianas al ser ejecutado tiene el 40,58% de instancias correctamente clasificadas.

En la Figura 7 se muestra la ejecución del algoritmo de *Random Tree*.

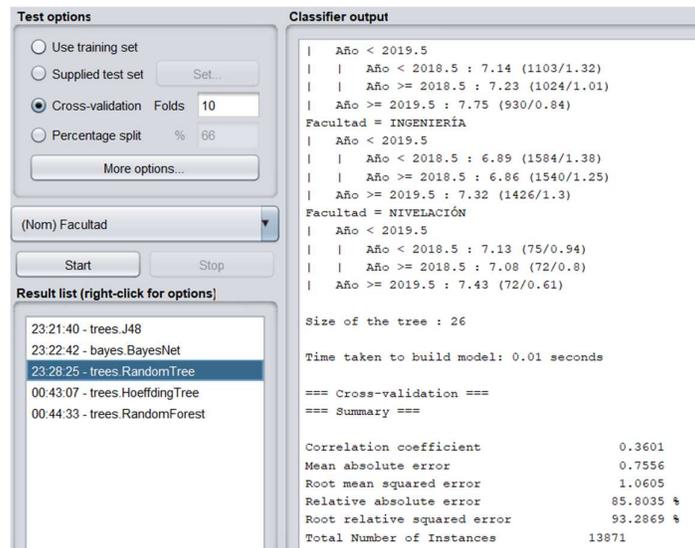


Figura 7: Algoritmo de clasificación Random Tree ejecutado

El algoritmo de *Random Tree* al ser ejecutado tiene un coeficiente de correlación de 0.361 que determina una correlación positiva ya que se encuentre entre 0 y 1.

En la Figura 8 se muestra la ejecución del algoritmo de *Hoeffding Tree*.

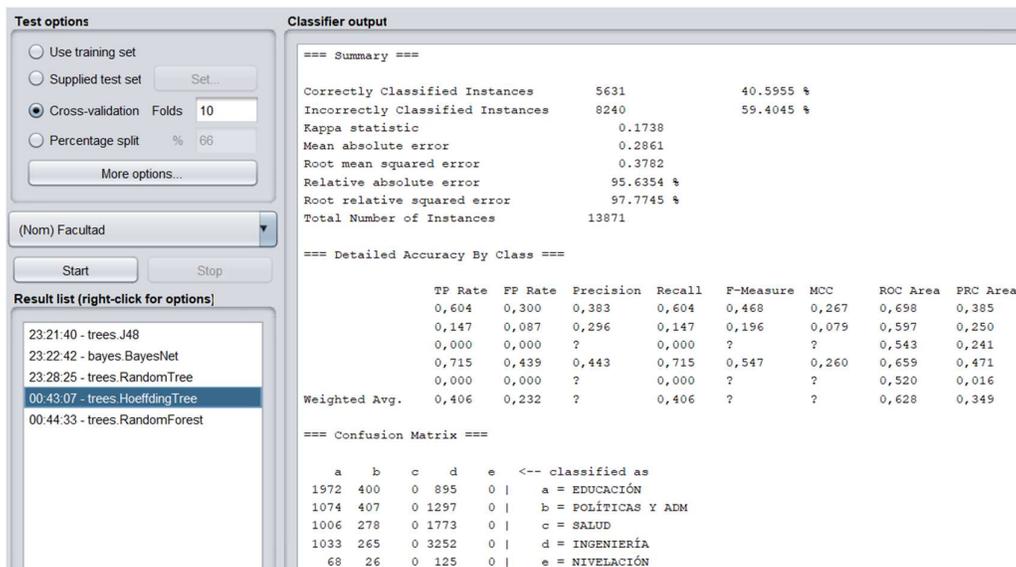


Figura 8: Algoritmo de clasificación Hoeffding Tree ejecutado

El algoritmo de *Hoeffding Tree* al ser ejecutado tiene el 40,60% de instancias correctamente clasificadas.

En la Figura 9 se muestra la ejecución del algoritmo de *Random Forest*.

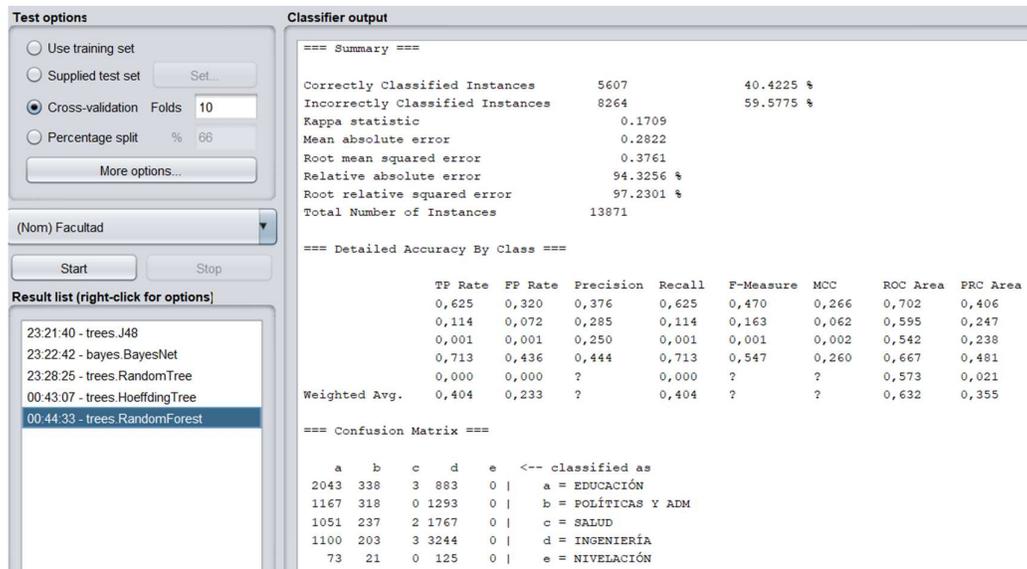


Figura 9: Algoritmo de clasificación *Random Forest* ejecutado

El algoritmo de *Random Forest* al ser ejecutado tiene el 40,42% de instancias correctamente clasificadas.

La ejecución de los algoritmos de clasificación permite tener toda la información necesaria para tomar la mejor opción en la siguiente fase.

3.7.2. Algoritmos de Agrupación

En la Figura 10 se muestra la ejecución del algoritmo simple *kMeans* en *Weka*.

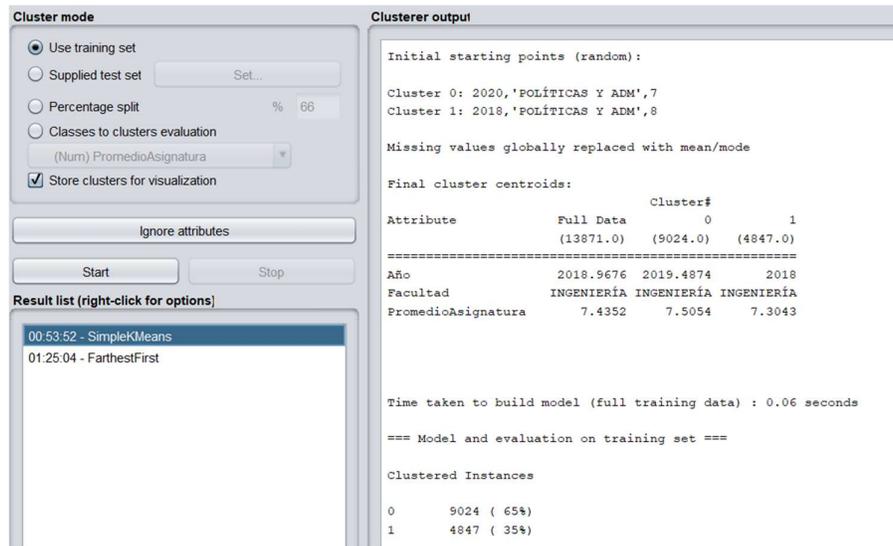


Figura 10: Algoritmo de agrupación simple *kMeans* ejecutado - *Weka*

El algoritmo simple *kMeans* al ser ejecutado muestra la agrupación de dos *clústers* o grupos, en un 65% y 35% respectivamente.

En la Figura 11 se muestra la ejecución del algoritmo simple *kMeans* en *RapidMiner*.



Figura 11: Algoritmo de agrupación simple *kMeans* ejecutado - *RapidMiner*

El algoritmo presenta dos grupos, el primer grupo, *cluster_0*, tiene calificaciones superiores a 6,5 puntos y tiene el 86,26% de la concentración de los datos, es decir, 11965 registros, mientras que, en el segundo grupo, *cluster_1*, existe el 13,74%, es decir, 1906 registros. Se evidencia que la mayor cantidad de estudiantes universitarios aprueban sus materias con notas superiores a los 6,5 puntos.

En la Figura 12 se muestra la ejecución del algoritmo de *FarthestFirst*.

```

Cluster mode
- [x] Use training set
- [ ] Supplied test set (Set...)
- [ ] Percentage split (% 66)
- [ ] Classes to clusters evaluation (Num) PromedioAsignatura
- [x] Store clusters for visualization
- Ignore attributes
- Start
- Stop
- Result list (right-click for options)
  00:53:52 - SimpleKMeans
  01:25:04 - FarthestFirst

Cluster output
- Facultad
- PromedioAsignatura
- Test mode: evaluate on training data
- === Clustering model (full training set) ===
- FarthestFirst
- =====
- Cluster centroids:
- Cluster 0
  2018.0 POLÍTICAS Y ADM 7.0
- Cluster 1
  2020.0 EDUCACIÓN 0.0
- Time taken to build model (full training data) : 0.02 seconds
- === Model and evaluation on training set ===
- Clustered Instances
  0      9138 ( 66%)
  1      4733 ( 34%)
    
```

Figura 12: Algoritmo de agrupación *FarthestFirst* ejecutado

El algoritmo de *FarthestFirst* al ser ejecutado muestra la agrupación de dos *clústers* o grupos, en un 66% y 34% respectivamente.

La ejecución de los algoritmos de agrupación permite tener toda la información necesaria, para tomar la mejor opción en la siguiente fase.

3.8. Fase 8. Evaluación

Con la generación de los patrones en fases previas, se interpretan los resultados de cada uno de los algoritmos de clasificación y agrupación, que se muestran a continuación:

3.8.1. Algoritmos de Clasificación

El árbol de decisión *Random Tree* generado evidencia de mejor manera la clasificación de los datos del rendimiento académico de los estudiantes durante las modalidades presencial y en línea desde el año 2018 hasta el año 2020, como se muestra en la figura 13.

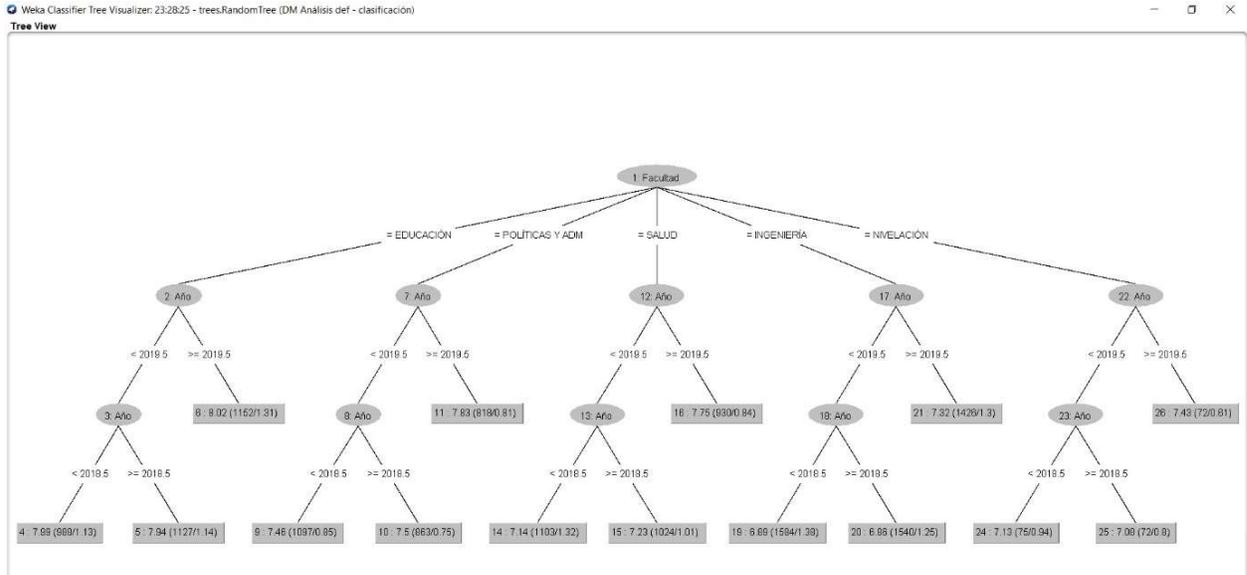


Figura 13: Algoritmo de Clasificación – *Random Tree*

Mediante el algoritmo de clasificación *Random Tree* se muestra claramente el conjunto de árboles que hacen referencia a la evolución del rendimiento académico por facultades en los tres años de análisis. Las hojas muestran que la mayoría de los datos están correctamente clasificados, y que la nota promedio de las asignaturas es más alta en el último año y en la modalidad en línea.

3.8.2. Algoritmos de Agrupación

El algoritmo simple *kMeans* evidencia de mejor manera la gráfica con los datos del rendimiento académico de los estudiantes durante las modalidades presencial y en línea desde el año 2018 al año 2020 y los agrupa en dos grupos como se muestra en la figura 14.

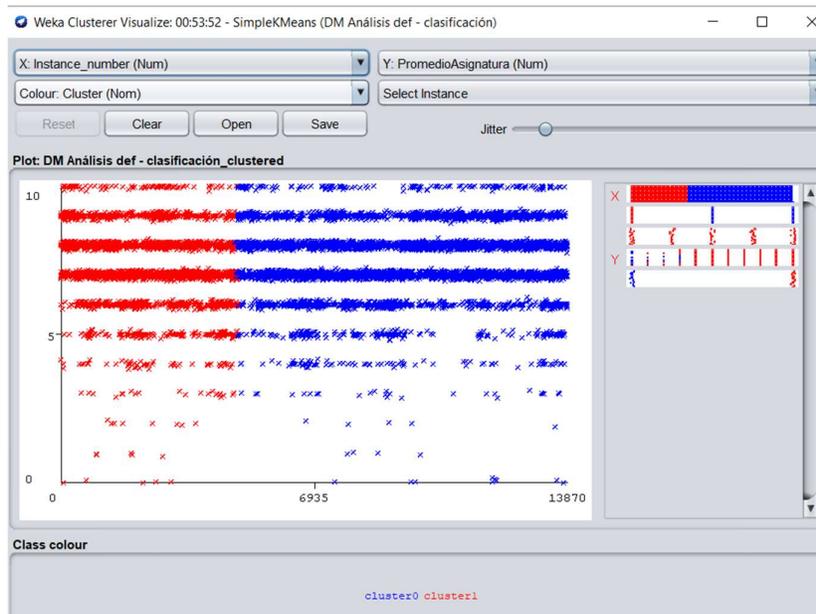


Figura 14: Algoritmo de Agrupación – Simple *kMeans*

El algoritmo de agrupación Simple *kMeans* muestra los dos grupos de datos homogéneos, el grupo 0 tiene la mayor concentración de datos con un 65% y son valores superiores a 5 puntos.

3.9. Fase 9. Usando el conocimiento descubierto

En esta última fase se obtiene el conocimiento sobre los factores que inciden en la educación superior en línea, en la Figura 15 y Figura 16 mediante *Weka* y *RapidMiner* se obtuvo información complementaria para la generación de conocimiento.

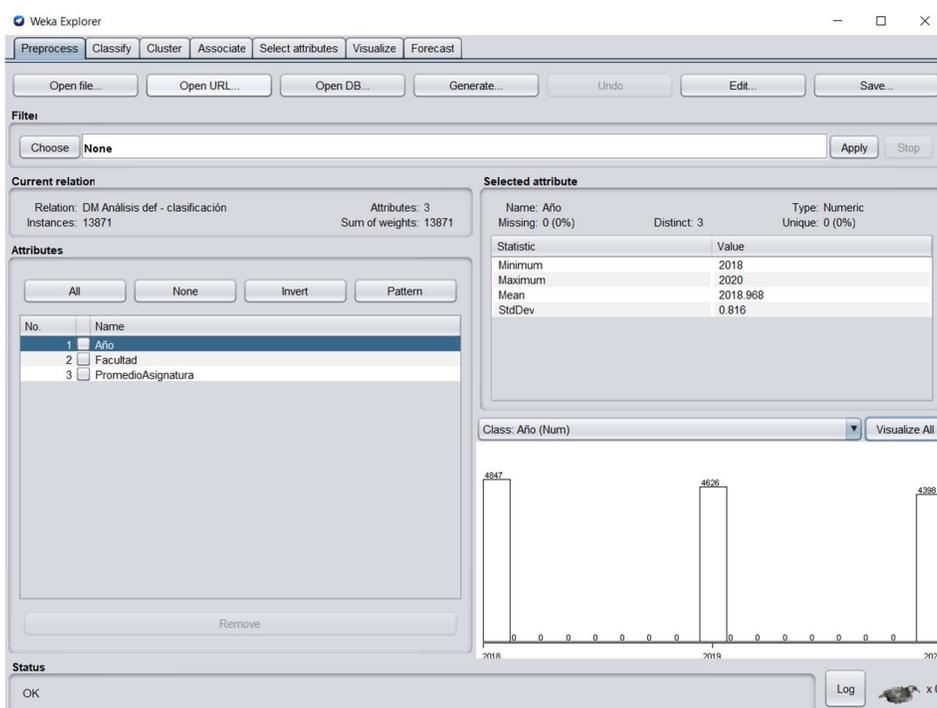


Figura 15: Distribución de los datos por años – *Weka*

La distribución de los registros por año es mayor en el año 2018 y el año 2020 es el que tiene menor cantidad de datos.

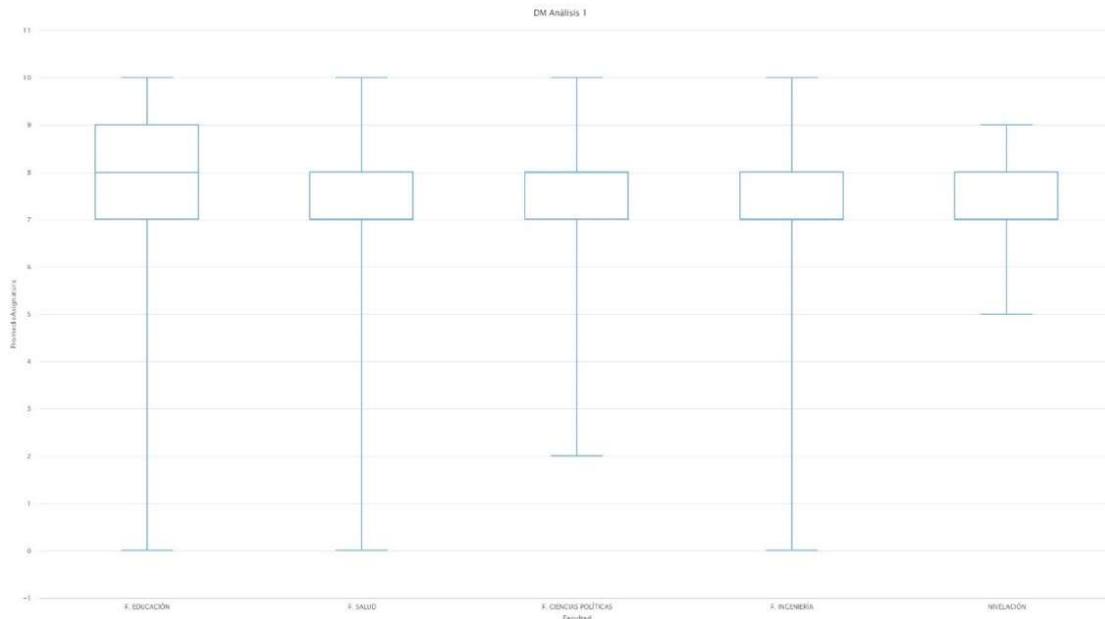


Figura16: Gráfica de caja y bigotes – *RapidMiner*

Mediante la gráfica de caja y bigotes se muestra que la Facultad de Educación es la que tiene las notas más altas y que dentro de la concentración de las notas, 7 puntos es la nota mínima.

Lista de conocimientos

Se generan los siguientes conocimientos:

- El 86,26% de los estudiantes tienen notas superiores a 6,5 puntos durante los periodos académicos desde el año 2018 hasta el 2021, es decir no existe cambio en el rendimiento académico de los estudiantes en las modalidades presencial y en línea.
- Durante la modalidad en línea, el porcentaje de asistencia a los encuentros síncronos es superior al 70%, por lo que los estudiantes aprueban en todas las asignaturas.
- El año 2020, que es el año de pandemia, concentra las mejores calificaciones en todas las facultades. El puntaje más alto lo tiene la Facultad de Educación con 8.02 puntos; el puntaje más bajo lo tiene la Facultad de Ingeniería con 7.32 puntos; y el año con mayor cantidad de estudiantes es el año 2019, a excepción de la Facultad de Educación que el año con mayor cantidad de estudiantes es el año 2020, el año de pandemia por COVID-19.

- 10 puntos es la nota máxima para las cuatro Facultades, mientras que para Nivelación es de 9 puntos. 8 puntos es la mediana de nota para la Facultad de Educación y la Facultad de Ciencias Políticas y Administrativas, y de 7 puntos para la Facultad de Salud, Facultad de Ingeniería y Nivelación.

3.9.1. Reportes de Inteligencia de Negocios

Se identificaron aspectos importantes de análisis mediante reportes de Inteligencia de Negocios con *Power BI*, como complemento al análisis que lo menciona la Metodología *KDD – Educacional*.

Análisis de las notas de los estudiantes por componente de evaluación

En la Tabla 8 se puede observar la evolución de los resultados de las calificaciones de los estudiantes.

Tabla 8: Tabla de la evolución de las notas – resultados del año 2018 al 2020

Año	Facultad	Promedio de Docencia	Promedio de Experimentación	Promedio de Autónomo	Promedio de Promedio Asignatura	Promedio de Asistencia
2018	FACULTAD DE CIENCIAS DE LA EDUCACIÓN, HUMANAS Y TECNOLOGÍAS	3.24	2.63	2.54	7.99	93.13
2018	FACULTAD DE CIENCIAS DE LA SALUD	2.72	2.30	2.32	7.14	92.50
2018	FACULTAD DE CIENCIAS POLÍTICAS Y ADMINISTRATIVAS	2.83	2.42	2.40	7.46	91.24
2018	FACULTAD DE INGENIERÍA	2.51	2.24	2.24	6.89	92.74
2018	UNIDAD DE NIVELACIÓN Y ADMISIÓN	2.71	2.45	2.47	7.13	95.47
2019	FACULTAD DE CIENCIAS DE LA EDUCACIÓN, HUMANAS Y TECNOLOGÍAS	3.21	2.59	2.51	7.94	93.39
2019	FACULTAD DE CIENCIAS DE LA SALUD	2.77	2.33	2.33	7.23	94.52
2019	FACULTAD DE CIENCIAS POLÍTICAS Y ADMINISTRATIVAS	2.83	2.45	2.42	7.50	91.78
2019	FACULTAD DE INGENIERÍA	2.46	2.21	2.25	6.86	92.92
2019	UNIDAD DE NIVELACIÓN Y ADMISIÓN	2.65	2.49	2.40	7.08	96.19
2020	FACULTAD DE CIENCIAS DE LA EDUCACIÓN, HUMANAS Y TECNOLOGÍAS	3.24	2.67	2.60	8.02	95.65
2020	FACULTAD DE CIENCIAS DE LA SALUD	3.07	2.55	2.52	7.75	97.99
2020	FACULTAD DE CIENCIAS POLÍTICAS Y ADMINISTRATIVAS	3.05	2.59	2.56	7.83	95.87
2020	FACULTAD DE INGENIERÍA	2.74	2.38	2.40	7.32	97.12
2020	UNIDAD DE NIVELACIÓN Y ADMISIÓN	2.86	2.54	2.49	7.43	97.22
Total		2,85	2,43	2,41	7,44	94,06

Durante los periodos académicos de la pandemia por COVID-19 que se desarrolla en modalidad en línea, no existe cambio en los componentes de la evaluación de los estudiantes.

El mayor análisis dentro del área académica es la evolución de las notas de los estudiantes, en la Figura 17 se evidencia que durante la pandemia por COVID-19 las notas mejoraron.

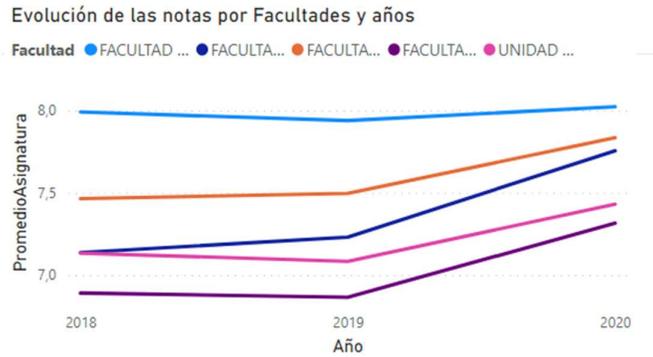


Figura 17: Resultado de la evolución de las notas en los años 2018, 2019 y 2020 de la UNACH

La mejora de las notas en el año 2020 en modalidad en línea, se observan con las líneas de evolución ascendente, sin embargo, esto no significa que el proceso de enseñanza aprendizaje sea mejor o superior, sino más bien muestra la importancia del uso de instrumentos de evaluación adecuadas y acordes a la nueva modalidad, con el respaldo de normativa y guías pertinentes.

Análisis de los resultados de evaluación docente

Mediante la evaluación integral docente, los estudiantes califican el trabajo de los docentes en la heteroevaluación, siendo importante que el docente realice su autoevaluación y que los directivos evalúen el trabajo de los docentes también, los tres componentes forman parte de la evaluación integral docente. Durante los últimos 3 años, este resultado ha sido excelente como se muestra en la Figura 18, sin embargo, en el año 2020 de pandemia por COVID-19, este puntaje obtuvo los valores históricos más altos, es decir, en la modalidad en línea.

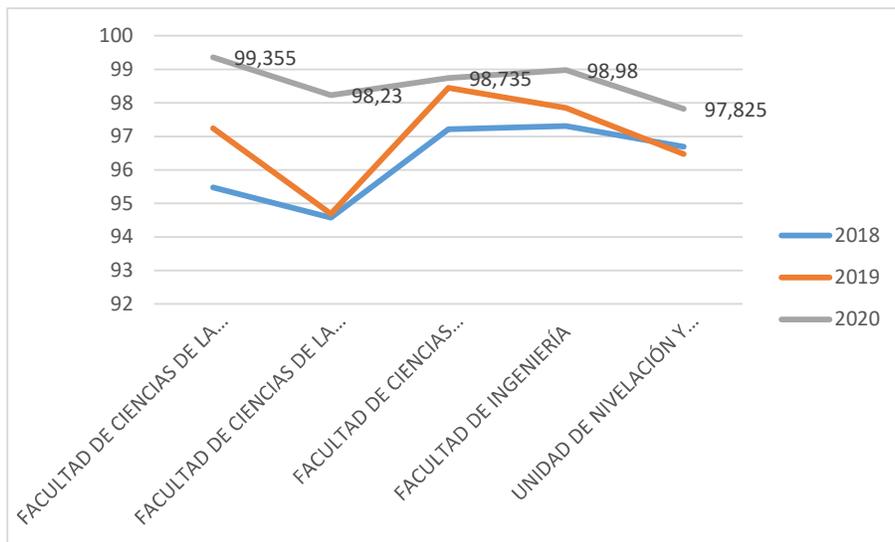


Figura 18: Evolución de la evaluación integral docente

En la figura 18 se demuestra que las habilidades y competencias de los docentes son adecuados, que las capacitaciones recibidas han sido de gran apoyo y evolución, pero también al ser superior el puntaje, se puede analizar que el instrumento de evaluación

integral docente requiere de una revisión de sus componentes cuando existe una modalidad en línea.

Análisis de los resultados de las conexiones a las aulas virtuales durante la modalidad en línea

Moodle ha sido la plataforma primordial para la modalidad en línea, ya que las aulas virtuales han permitido almacenar los contenidos y las actividades académicas para generar conocimiento, desde mayo del año 2020 hasta el año 2021 han mantenido una tendencia promedio de 3000 conexiones como se muestra en la Figura 19.

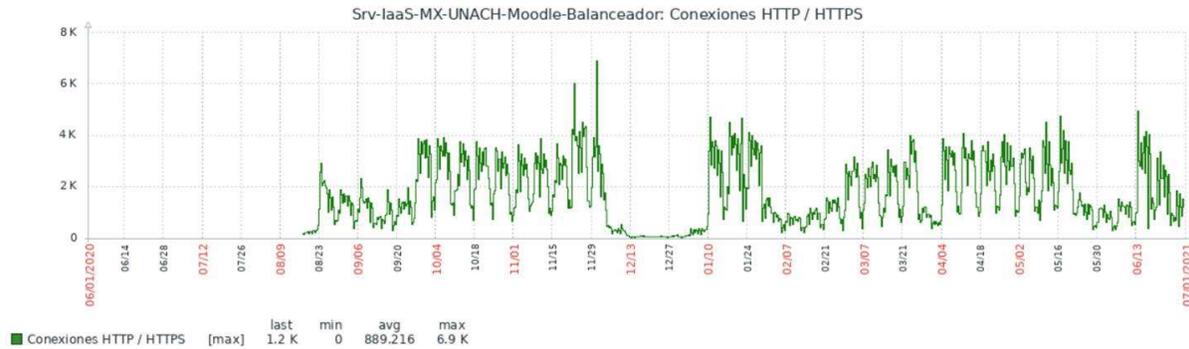


Figura 19: Resultado de las conexiones a las aulas virtuales Moodle

Los resultados de registros constan desde agosto de 2020 hasta julio de 2021, manteniendo el promedio de conexiones de usuarios recurrentes.

Las conexiones en Moodle alcanzan un valor máximo de 6.900 conexiones, demostrando la importancia del entorno de aprendizaje para la modalidad en línea, además identifica el aumento de las conexiones concurrentes de usuarios y el uso de recursos tecnológicos. Finalmente, en diciembre de cada año existe el receso académico, por tal motivo existe una reducción de conexiones.

Análisis de la Encuesta para estudiantes sobre la situación actual durante la emergencia sanitaria

La UNACH adicionalmente, realizó encuestas a los estudiantes durante los años 2020 y 2021, en las cuales se obtuvieron los siguientes resultados más importantes:

El 23.30% de estudiantes durante el año 2020 de pandemia por COVID-19 pensó en suspender sus actividades académicas como muestra la Figura 20.

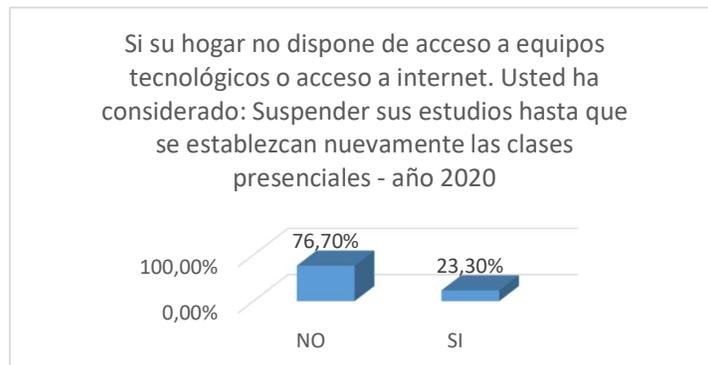


Figura 20: Resultado de suspensión de clases 2020 UNACH

El 54,95% de la población estudiantil, en el año 2020 ha identificado la necesidad de una ayuda económica para continuar con las actividades académicas en la modalidad en línea como muestra la Figura 21.

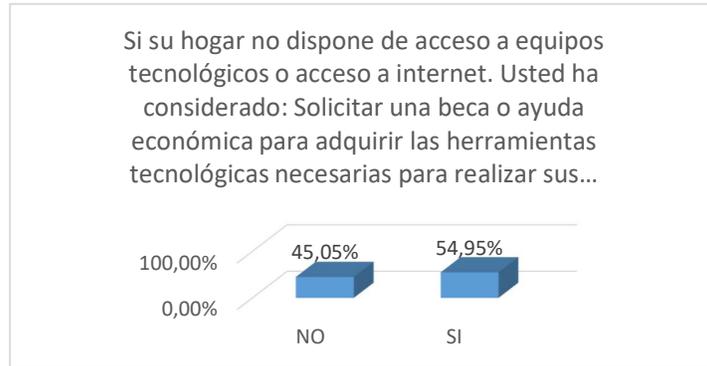


Figura 21: Resultado de solicitud de becas o ayudas económicas 2020 UNACH

Las condiciones de acceso a dispositivos con internet mejoraron en el año 2021, es así como más del 25% tiene la opción de más de 8 horas para desarrollar sus actividades académicas, como indica la Figura 22.

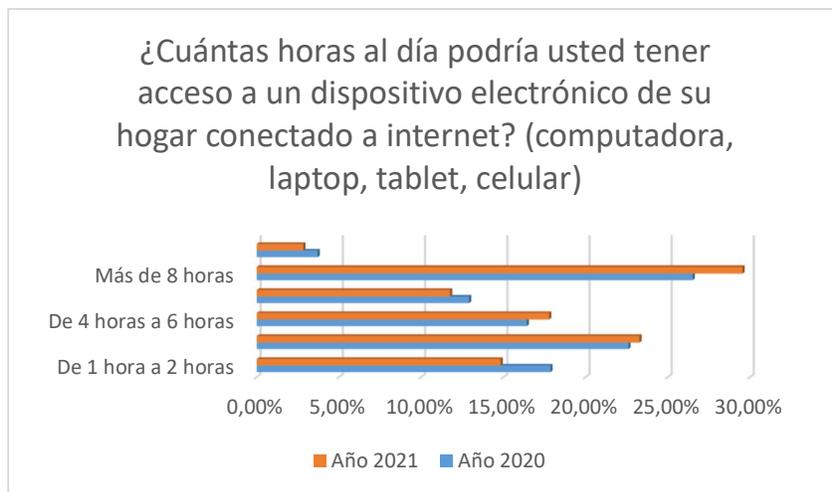


Figura 22: Resultado de horas de acceso a dispositivo con internet 2020 y 2021 UNACH

La dinámica de los años de pandemia por COVID-19 permitieron frente a la necesidad de la comunidad universitaria que busquen las formas de mejorar sus condiciones y continuar sus estudios.

3.9.2. Factores que inciden en la educación superior en línea

En este importante análisis de datos mediante el desarrollo de la Metodología *KDD-Educacional*, mediante la investigación cualitativa, se analizó el contenido y se interpretó los resultados de la investigación cuantitativa, que identificó patrones y significados que fueron clasificados por perspectivas de los diferentes actores del proceso de enseñanza-aprendizaje, a continuación, se muestran los factores que inciden en la educación superior en línea desde la modalidad en línea-emergente, donde la enseñanza tradicional pasó al

aprendizaje en línea durante las crisis. Las perspectivas son: institución, docentes y estudiantes. La Tabla 9 muestra los factores desde la institución.

Tabla 9: Factores que inciden en la modalidad en línea desde la perspectiva de la institución

FACTOR	DESCRIPCIÓN
Económico	Las instituciones de educación superior requieren estabilidad económica para garantizar la calidad de la educación superior. La crisis económica mundial afecta completamente su desarrollo.
Igualdad de oportunidades en el acceso	La universidad debe implementar medidas de acción afirmativa que garanticen el acceso a la educación superior, teniendo en cuenta que la situación de los estudiantes de las instituciones públicas es compleja.
Política de la modalidad de estudio	Las regulaciones para la educación en línea deben ser claras, precisas y garantizar su desarrollo y crecimiento.
Recursos pedagógicos	El liderazgo pedagógico en la educación superior en línea debe garantizar el proceso de enseñanza-aprendizaje, es necesario fortalecerlo.
Oferta Académica	Los estudios de pertinencia y demanda respaldan la presentación de una oferta académica adecuada que garantice la igualdad de oportunidades en el acceso.
Infraestructura tecnológica	La tecnología es el pilar de la educación en línea, por lo tanto, debe haber planes de ejecución y mantenimiento.
Sistemas y plataformas informáticas	Es importante automatizar por completo los procesos académicos y administrativos en la modalidad en línea.
Modelo de educación en línea	El modelo de educación en línea debe ser el pilar del desarrollo académico de la institución, con instrumentos claros que reflejen la realidad de esta nueva modalidad.

Los factores que determinan la educación superior en línea desde la perspectiva de la institución garantizan su adecuado desarrollo permiten su fortalecimiento.

En la Tabla 10 se muestran los factores desde la perspectiva del docente.

Tabla 10: Factores que inciden en la modalidad en línea desde la perspectiva del docente

FACTOR	DESCRIPCIÓN
Capacitación	Los docentes requieren recibir el conocimiento técnico y pedagógico que se desarrolla en la modalidad en línea, generando competencias administrativas y técnicas enfocadas en los contenidos de la tecnología educativa.
Metodología de enseñanza-aprendizaje	La metodología empleada en la modalidad en línea es uno de los factores más importantes que inciden en su desarrollo, siendo imprescindible que los docentes la dominen.
Actitud	La actitud del docente frente a una modalidad diferente debe inspirar la evolución, crecimiento y adaptación de las nuevas generaciones, siendo la creatividad una de las más importantes.
Sensibilidad a las diferencias de los estudiantes	El docente debe tener la sensibilidad para entender las diferencias de los estudiantes, especialmente en este proceso de pandemia por COVID-19
Tutorías	El docente debe dedicar un tiempo importante al seguimiento académico y a las tutorías con el objetivo de acompañar a los estudiantes en el proceso

Los factores que determinan la educación superior en línea desde la perspectiva del docente fortalecen el desarrollo de la educación en línea por la evolución de las competencias y habilidades de los docentes.

En la Tabla 11 se muestran los factores desde la perspectiva del estudiante.

Tabla 11: Factores que inciden en la modalidad en línea desde la perspectiva del estudiante

FACTOR	DESCRIPCIÓN
Organización de tiempos y tareas	Los estudiantes deben estar conscientes de la importancia del manejo del tiempo durante la modalidad en línea, que no exige horarios, pero sí cumplimiento de actividades.
Actitud	El estudiante debe tener una actitud positiva y de crecimiento frente a la necesidad académica.
Dominio del uso de herramientas tecnológicas	Los estudiantes deben dominar el uso de las herramientas en las cuales desarrollarán los procesos de enseñanza aprendizaje. Es importante que existan capacitaciones e inducciones

Los factores que determinan la educación superior en línea desde la perspectiva del estudiante determinarán su éxito y el cumplimiento del objetivo de formación profesional.

4. Discusión

En la presente investigación existen varios aportes significativos, entre estos, la metodología de “Descubrimiento de Conocimiento en Base de Datos Educativos - *KDD – Educativa*” que acopla los pasos a la UNACH como institución de educación superior, como un aporte vital que marca el inicio de los análisis respectivos de los datos en el área de educación, frente a autores como Moine et al. (2011), que proponen modelos o metodologías de *KDD* como: *CRISP-DM*, *SEMMA* y *Catalyst*, respectivamente de manera general.

La metodología *KDD-Educativa* dentro de los productos más importantes tiene el conocimiento que se generó y que en comparación con los estudios de la UNESCO e IESALC, en este estudio se definen los factores de la modalidad en línea-emergente que inciden en el Modelo Educativo en Línea del Sistema de Educación Superior en Ecuador, desde la perspectiva de la institución, del docente y de los estudiantes. Esto permite, a partir de su identificación, encontrar las mejores estrategias para mejorar el proceso de enseñanza-aprendizaje en nuevas modalidades como una oportunidad que se generó post pandemia, y que asegurará su desarrollo ya que también se incluye el análisis comparativo entre dos modalidades, tanto la presencial como la modalidad en línea-emergente.

Además, según Rochina-Chisag y Tabuena (2022) los métodos de enseñanza tradicionales siguen prevaleciendo en las instituciones de educación superior y la pandemia por COVID-19 ha generado la transición al aprendizaje en línea en todo el mundo, además, según Henríquez et al. (2022) la transición de clases presenciales a clases virtuales se generó a partir del COVID-19, en donde existieron desafíos, que requirieron estrategias, existiendo oportunidades que experimentaron los profesores y los estudiantes. El presente estudio permitió además de analizar datos, determinar mediante la interpretación de los patrones, los factores que inciden en la educación en modalidad en línea, realizados a partir de los datos de la educación presencial y en línea-emergente.

Según Giler-Valverde et al. (2023), el modelo de educación en línea es influenciado por las innovaciones tecnológicas en Ecuador y las universidades ofrecen mejores servicios que se adaptan a los cambios tecnológicos. El presente estudio demuestra con la base de los datos que los factores más importantes desde la perspectiva de la institución son la infraestructura tecnológica, sistemas y plataformas informáticas que son el pilar para el desarrollo de la educación en línea, además, desde la perspectiva del estudiante el dominio del uso de herramientas tecnológicas se convierte en un requisito, ya que, desde la modalidad en línea-emergente la tecnología se fortaleció y consta como aporte al modelo educativo del sistema de educación superior en Ecuador.

Otro aporte significativo dentro del análisis de los datos, es que las calificaciones de los estudiantes en la modalidad en línea-emergente sobre todo en el año 2020 fueron mejores en todas las facultades de la UNACH, lo que refleja la importancia de la herramienta de evaluación para garantizar el proceso de enseñanza-aprendizaje, es por esto que los factores de política de la modalidad de estudio, recursos pedagógicos y el modelo de educación en línea desde la perspectiva de la institución, y la metodología de enseñanza aprendizaje

desde la perspectiva del docente, se determinan para garantizar educación de calidad que forme profesionales competentes y con conocimientos.

Los factores económico, igualdad de oportunidades en el acceso, y la oferta académica desde la perspectiva de la institución, son factores que fueron determinados de acuerdo a la situación establecida por los estudiantes en las encuestas desde la recolección de los datos, ratificando el sentido de la universidad pública en el sistema de educación superior del país y de lo importante que es conocer la realidad y situación de los estudiantes que escogen carreras con pertinencia requiriendo la igualdad de oportunidades en el acceso para su proceso de formación como base en una modalidad en línea.

Los factores de organización de tiempo y actitud desde la perspectiva del estudiante, y, actitud y sensibilidad a las diferencias de los estudiantes desde la perspectiva del docente, son los factores que también se determinaron a partir de los datos analizados en las encuestas, que, desde la interpretación de los resultados como parte de la investigación cualitativa, se consideraron puntos importantes.

Desde la perspectiva del docente la capacitación constituye la base de evolución del profesional para garantizar el proceso de enseñanza-aprendizaje con calidad, de acuerdo con los datos analizados, los resultados de la evaluación docente tuvieron los puntajes más altos durante pandemia, por tal motivo este factor es necesario, además porque el docente en educación en línea debe destinar tiempo adecuado al factor tutoría que requiere capacitación para mejorar las competencias del mismo.

Finalmente, desde la perspectiva de la institución el factor más importante es la política de la modalidad de estudio que asegura las bases del proceso de enseñanza-aprendizaje. Desde la perspectiva del docente, el factor más significativo son las tutorías que se convierten en el espacio para dar seguimiento y acompañamiento al estudiante para fortalecer sus competencias. Y desde la perspectiva del estudiante, los factores más relevantes son la actitud con su necesidad académica y el dominio del uso de herramientas tecnológicas, que le brinda los medios para alcanzar el objetivo de formación profesional.

5. Conclusiones

La generación de datos producidos durante la pandemia por COVID-19 en las instituciones de educación superior, específicamente en la UNACH, fueron una oportunidad para aplicar la investigación cuantitativa mediante el proceso de descubrimiento de conocimiento en bases de datos *KDD – Educacional*, y seleccionar las mejores técnicas de minería de datos de clasificación y agrupación siendo *random tree* y *simple kMeans* las más precisas. Además, mediante la inteligencia de negocios se complementó el conocimiento, y mediante la investigación cualitativa, desde la interpretación de los datos, se generaron los factores que desde la modalidad presencial y en línea-emergente inciden en el modelo de educación superior en línea.

El principal pilar identificado para la educación en línea son las plataformas informáticas e infraestructura tecnológica, ya que permite el desarrollo de las actividades académicas, genera gran cantidad de datos que pueden ser analizados para plantear acciones de mejora, sobre todo frente a las conexiones concurrentes de los usuarios, el uso de las plataformas

adecuadas para el proceso de enseñanza-aprendizaje y además el conocimiento sobre el rendimiento de los estudiantes, por tal motivo, es importante considerar la planificación de presupuestos si se desea sostener una nueva modalidad en línea con educación de calidad que garantice igualdad de oportunidades.

Mediante la identificación de los factores que inciden en la educación superior en línea desde la perspectiva de la institución, los docentes y los estudiantes, se pueden detectar los factores que se deben trabajar y fortalecer especialmente en los procesos de evaluación de los estudiantes y docentes, ya que durante el año 2020 los estudiantes mejoraron su rendimiento académico, manteniendo como promedio más alto a la Facultad de Educación con 8.02 puntos, y el más bajo a la Facultad de Ingeniería con 7.32 puntos, y los docentes mejoraron su evaluación integral docente con el promedio más alto a nivel histórico de la UNACH, siendo el 98.825%; dejando a la interpretación si el puntaje refleja una mejora de las habilidades y competencias de los dos actores del proceso de educación durante la pandemia por COVID-19, pero además, se abre la opción de considerar más y nuevas variables para el desarrollo académico de calidad.

Las encuestas de la UNACH durante los años 2020 y 2021 permitieron evidenciar el escenario complejo que mantuvo la educación en línea-emergente, ya que los estudiantes de las universidades públicas no están preparados para asumirla inmediatamente, sino más bien que el éxito de una modalidad en línea en primera instancia está dado por la necesidad del estudiante en el proceso de formación, lo que muestra el 55% de los estudiantes en las encuestas al identificar que requieren ayudas económicas si la modalidad de estudios es en línea. También se evidencia el trabajo de la UNACH al conseguir datos que puedan evidenciar la realidad de los estudiantes para plantear estrategias y tomar decisiones, cumpliendo el objetivo de un proceso *KDD* y la minería de datos educacional.

Finalmente, las futuras investigaciones deben analizar el impacto de la pandemia de COVID-19 en la educación superior y la transición a la modalidad en línea enfocadas en varias áreas clave, siendo fundamental evaluar a largo plazo el efecto en el rendimiento académico y el desarrollo profesional de los estudiantes, incluyendo su empleabilidad y éxito post-graduación. Además, es importante investigar la adaptación y efectividad de diversas metodologías pedagógicas y enfoques de enseñanza en entornos virtuales, especialmente con el uso de tecnologías emergentes como la realidad aumentada y virtual, que también estén acompañadas de las políticas y estrategias implementadas para reducir la brecha digital y promover la equidad en el acceso a la educación en línea, evaluando su efectividad en diferentes contextos socioeconómicos.

Contribuciones de los autores

En concordancia con la taxonomía establecida internacionalmente para la asignación de créditos a autores de artículos científicos (<https://casrai.org/credit/>). Los autores declaran sus contribuciones en la siguiente matriz:

	Uvidia-Fassler, M.	Cisneros-Barahona, A.	Méndez-Naranjo, P.	Villa-Yáñez, H.
Conceptualización				
Análisis formal				
Investigación				
Metodología				
Validación				
Redacción – revisión y edición				

Conflicto de Interés

Los autores declaran que no existen conflictos de interés de naturaleza alguna con la presente investigación.

Referencias

Bernabeu, R. (19 de julio de 2010). *Data Warehousing y metodología Hefesto*. <https://www.dataprix.com/es/data-warehousing-y-metodologia-hefesto/data-warehousing-y-metodologia-hefesto>

Bustamante, R. (2020). Educación en cuarentena: cuando la emergencia se vuelve permanente. *Aportes para el diálogo y las acción* 1-12. <https://www.grade.org.pe/creer/recurso/educacion-en-cuarentena-cuando-la-emergencia-se-vuelve-permanente/>

Calvopiña, V. (4 de Octubre de 2020). *Estudiar : un privilegio en pandemia Jóvenes entre la crisis y la exclusión*. <https://wambra.ec/estudiar-en-pandemia-un-privilegio-crisis-exclusion/>

Cisneros, A.S., Uvidia, M.I., Samaniego, G.N., Dumancela, G.J., y Casignia, B.A. (2021). Complementary Admission Processes Implemented by Ecuadorian Public Universities Promote Equal Opportunities in Access: An Analysis Through Knowledge Discovery in Databases. In: Botto-Tobar, M., Zamora, W., Larrea Plúa, J., Bazurto Roldan, J., Santamaría Philco, A. (eds) *Systems and Information Sciences. ICCIS 2020. Advances in Intelligent Systems and Computing*, vol 1273. Springer, Cham. https://doi.org/10.1007/978-3-030-59194-6_18

Consejo de Educación Superior [CES]. RPC-SE-03-No.046-2020. <https://www.ces.gob.ec/wp-content/uploads/2020/05/CODIFICADO-RPC-SE-03-No.046-2020.pdf>

Fayyad, U., Piatetsky-Shapiro, G., y Smyth, P. (1996). From data mining to knowledge discovery in databases. *AI Magazine*, 17(3), 37. <https://doi.org/10.1609/aimag.v17i3.1230>

Giler-Valverde, G. P., Luy-Navarrete, W. A., y Merino-Velásquez, J. (2023). Efectividad del modelo online en la graduación de estudiantes de la Universidad de Guayaquil, Ecuador. *Conciencia Digital*, 6(1), 42-57 <https://doi.org/10.33262/concienciadigital.v6i1.2434>

Henríquez, E. F., Salazar, Y., Santín-Picoita, F., Morales, J. C., y Torres, Y. (2022). The Impact of Covid19 on Traditional Higher Education in Ecuador. *Journal of Higher Education, Theory, and Practice*, 22(4), 30-40 <https://doi.org/10.33423/jhetp.v22i4.5126>

Instituto Nacional de Estadística y Censos. (2021). *Tecnologías de la Información y Comunicación-TIC*. <https://www.ecuadorencifras.gob.ec/tecnologias-de-la-informacion-y-comunicacion-tic/>.

- Jiménez, J. A., y Timarán, S. (2015). Caracterización de la deserción estudiantil en educación superior con minería de datos. *Revista Tecnológica - ESPOL*, 28(5) 447–463. <https://rte.espol.edu.ec/index.php/tecnologica/article/view/453>
- Maimon, O., y Rokach, L. (2006). *Data Mining and Knowledge Discovery Handbook*, Springer New York, NY. <https://doi.org/10.1007/b107408>.
- Mandela, N. (2014). *Nelson Mandela por sí mismo*. Plataforma Editorial. <https://www.plataformaeditorial.com/libro/3948-nelson-mandela-por-si-mismo>
- Microsoft. (2021). *Microsoft*. <https://www.microsoft.com/es-es/power-platform/products/power-bi>
- Moine, M., Haedo, S., y Gordillo, D. (2011). *Estudio comparativo de metodologías para minería de datos*. CACIC 2011 - XVII Congreso Argentino De Ciencias De La Computación. https://sedici.unlp.edu.ar/bitstream/handle/10915/20034/Documento_completo.pdf?sequence=1
- Organización Mundial de la Salud. (2020). *Coronavirus (COVID-19)*. <https://www.who.int/es/emergencias/diseases/novel-coronavirus-2019>
- Rapidminer. (2021). *RapidMiner*. 1–7. <https://altair.com/altair-rapidminer>
- Rochina-Chisag, Á. G., y Tabuena, A. (2022). Online Learning as an Alternative Learning Modality in Ecuador's Education Institutions Amidst Crises and Outbreaks: A SWOT Analysis. *Journal of Learning for Development*. <https://doi.org/10.56059/jl4d.v9i3.665>
- UNACH. (2020). *Directrices para la aplicación en la universidad nacional de chimborazo, de la normativa transitoria emitida por el ces, para el desarrollo de actividades académicas en las instituciones de educación superior, debido al estado de excepción decretado por la emergencia sanitaria ocasionada por la pandemia de covid19*.
- UNESCO. (2021). *Education: From COVID-19 school closures to recovery*. <https://es.unesco.org/covid19/educationresponse>
- UNESCO, & IESALC. (2020a). *¿Cómo las universidades garantizan la continuidad pedagógica y planifican el futuro? Aportes emergentes del primer foro de rectores ante la pandemia del COVID-19*. <https://bit.ly/4cvNqB9>
- UNESCO, y IESALC. (2020b). *COVID-19 y educación superior. De los efectos inmediatos al día después*. <https://www.iesalc.unesco.org/wp-content/uploads/2020/04/COVID-19-070420-ES-2-1.pdf>
- UNESCO, y IESALC. (2020c). *El Coronavirus COVID-19 y la educación superior: impacto y recomendaciones*. <https://bit.ly/3KQMcF3>
- UNESCO, & IESALC. (2020d). *Informe del IESALC analiza los impactos del #COVID19 y ofrece recomendaciones a gobiernos e instituciones de educación superior*. <https://bit.ly/3ze1K2W>
- Uvidia Fassler, M.I., Cisneros Barahona, A.S., Ávila-Pesántez, D.F., Rodríguez Flores, I.E. (2018). Moving Towards a Methodology Employing Knowledge Discovery in Databases to Assist in Decision Making Regarding Academic Placement and Student Admissions for Universities. In: Botto-Tobar, M., Esparza-Cruz, N., León-Acurio, J., Crespo-Torres, N., Beltrán-Mora, M. (eds) *Technology Trends. CITT 2017. Communications in Computer and Information Science*, vol 798. Springer, Cham. https://doi.org/10.1007/978-3-319-72727-1_16
- University of Waikato. (2021). *The Weka Workbench*. <http://www.cms.waikato.ac.nz/ml/weka/>