UNIVERSIDAD ESTATAL DEL SUR DE MANABÍ

INSTITUTO DE POSGRADO
MAESTRÍA EN TECNOLOGÍAS DE LA INFORMACIÓN Y LA COMUNICACIÓN

Proyecto de Titulación Modalidad Artículos Científicos Profesionales De Alto Nivel

TEMA

Técnica de minería de datos para procesos educativos en estudiantes con necesidades educativas especiales basado en un modelo predictivo

Proyecto de Titulación previo a la obtención de título de Magister en Tecnologías de la Información y la Comunicación

AUTORA:

Ing. Hilda Brigitte Quijije Quiroz

TUTORA:

Lcda. Kirenia Maldonado Zuñiga. Mg. CE.

JIPIJAPA - MANABÍ – ECUADOR

2023

UNIVERSIDAD ESTATAL DEL SUR DE MANABÍ

INSTITUTO DE POSGRADO
MAESTRÍA EN TECNOLOGÍAS DE LA INFORMACIÓN Y LA COMUNICACIÓN

Proyecto De Titulación Modalidad Artículos Científicos Profesionales De Alto Nivel

TEMA:

Técnica de minería de datos para procesos educativos en estudiantes con necesidades educativas especiales basado en un modelo predictivo

Proyecto de Titulación previo a la obtención de título de Magister en Tecnologías de la Información y la Comunicación

AUTORA: Hilda Brigitte Quijije Quiroz

Ingeniera en Tecnologías de la Información

quijije-hilda2868@unesum.edu.ec

TUTORA: Kirenia Maldonado Zuñiga

Doctora en Tecnologías de la Información y Comunicación

kirenia.maldonado@unesum.edu.ec

Línea de investigación aprobada en el Posgrado/ Grupo de Investigación/ Programa de Investigación: Tecnologías de la Información y Comunicación

JIPIJAPA – MANABÍ – ECUADOR

2023

UNIVERSIDAD ESTATAL DEL SUR DE MANABÍ
Creada mediante Ley Publicada en el registro Oficial 261 del 7 de febrero de 2001
INSTITUTO DE POSGRADO
MAESTRÍA EN TECNOLOGÍAS DE LA INFORMACIÓN Y LA COMUNICACIÓN

APROBACIÓN DEL TUTOR

En mi calidad de Tutor del Trabajo de Titulación: **"Técnica de minería de datos para procesos educativos en estudiantes con necesidades educativas especiales basado en un modelo predictivo"** presentado por, **Hilda Brigitte Quijije Quiroz**, para optar por el título Magister en Tecnologías de la Información y la Comunicación, certifico que dicho Trabajo de Titulación ha sido revisado en todas sus partes y se considera que reúne los requisitos y méritos suficientes para ser sometido a la presentación para la valoración por parte del tribunal que se designe y su exposición y defensa pública.

Asimismo, le hago entrega de: *Carta de aceptación del artículo, Carta de publicación, Artículo y del Proyecto* cedido por el maestrante **Hilda Brigitte Quijije Quiroz**, para el debido análisis de los miembros del tribunal que se designe para continuar con el proceso de titulación.

Particular que comunico para fines pertinentes.

Jipijapa, 01 de julio del 2023

Lcda. Kirenia Maldonado Zuñiga, PhD.
C.I. 1756936884
TUTORA

APROBACIÓN TRIBUNAL DE TITULACIÓN

El proyecto de Titulación: "**Técnica de minería de datos para procesos educativos en estudiantes con necesidades educativas especiales basado en un modelo predictivo**" ha sido revisado, aprobado y autorizada su impresión y empastado, previo a la obtención de título de Magister en Tecnologías de la Información y la Comunicación; dado que el presente trabajo reúne los requisitos de contenido y forma.

Jipijapa, 07 de julio del 2023

.....

Ing. Lenin Jonatan Pin García, PHD.

CI. 1307678860

Presidente del tribunal

.....

Ing. Vicente Fray Romero Castro, Mg

CI. 1309181079

Miembro del tribunal

.....

Ing. Franklin Jhimmy Toala Arias, Mg

CI. 1307688232

Miembro del tribunal

RESPONSABILIDAD DE AUTORÍA

La responsabilidad de investigación, teorías. ideas, análisis, resultados, conclusiones, recomendaciones y propuestas planteadas en el presente Proyecto de Titulación en modalidad artículos profesionales de alto nivel son de exclusiva responsabilidad de su autor.

Así mismo, autorizo a la **UNIVERSIDAD ESTATAL DEL SUR DE MANABÍ** para que realice la digitalización y publicación de este trabajo de investigación en el repositorio virtual, de conformidad ando dispuesto en el Art. 144 de la Ley Orgánica de Educación Superior.

Jipijapa, 28 de junio del 2023

.....

Ing. Hilda Brigitte Quijije Quiroz

CI. 1350842868

MAESTRANDO

RENUNCIA DE DERECHOS

Quien suscribe, cede los derechos de autor intelectual total y parcial del presente proyecto de titulación a la Universidad Estatal del Sur de Manabí

Jipijapa 28 de junio del 2023

Ing. Hilda Brigitte Quijije Quiroz

CI. 1350842868

MAESTRANDO

AVAL DEL VEEDOR

Quién suscribe, declara que el presente Proyecto De Titulación: **"Técnica de minería de datos para procesos educativos en estudiantes con necesidades educativas especiales basado en un modelo predictivo"**, contiene las observaciones realizadas por los lectores en sesión científica del tribunal.

Jipijapa, 07 de julio del 2023

Ing. Lenin Jonatan Pin García, PHD.

CI. 1307678860

Presidente del tribunal

AGRADECIMIENTO

Gracias a Dios por la fortaleza espiritual que me brinda a diario, a mis padres por el amor, esfuerzo y apoyo en cada una de mis decisiones y metas que establezco, a mis hermanos y sobrino por hacer mayormente feliz todos los días y por confiar en mis capacidades.

A mi tutora Lcda. Kirenia Maldonado Zuñiga, gracias por la paciencia, enseñanzas y dedicación en cada uno de los procesos dentro de este trabajo, a los docentes que con sus aprendizajes en cada módulos aportaron y potencializaron mis conocimientos, gracias por permitirme destacar mis habilidades y aplicarlas en mis actividades laborales y personales.

Un agradecimiento rotundo a las autoridades de la Unidad Educativa Plan Internacional por la disposición de la información brindada a esta investigación y por la aprobación rotunda de los resultados finales.

Gracias a todos los que me apoyaron mientras cursaba mi etapa de preparación profesional, sin duda forman parte de este maravilloso logro.

Hilda Brigitte Quijije Quiroz

DEDICATORIA

Dedico este trabajo a Dios mi ser supremo, a mis padres Sr. Genaro Quijije y Sra. Rosario Quiroz personajes principales a lo largo de mi vida, mi motivación, superación y apoyo en lo personal y laboral, a mis hermanos Karen, John y Kristine pilares fundamentales en la realización de este proyecto, a mi pequeño sobrino Saúl quien me ayuda a desestresarme con sus juegos y travesuras, a mis tíos y abuelitos.

A mis fuentes de inspiración por ayúdame a salir de mi zona de confort y reconfortar el amor por lo que disfruto hacer y a lo que me dedicaré en el futuro, a mis docentes quienes con sus enseñanzas han permitido la culminación de este trabajo, a los amigos que aún no conozco y a los que ya tengo, a mis conocidos y demás personas que depositaron su confianza en mi capacidad como profesional y maestrante.

Hilda Brigitte Quijije Quiroz

UNIVERSIDAD ESTATAL DEL SUR DE MANABÍ

INSTITUTO DE POSGRADO

MAESTRÍA EN TECNOLOGÍAS DE LA INFORMACIÓN Y LA COMUNICACIÓN

TEMA:

Técnica de minería de datos para procesos educativos en estudiantes con necesidades educativas especiales basado en un modelo predictivo

Autor:

Ing. Hilda Brigitte Quijije Quiroz

Tutora:

Lcda. Kirenia Maldonado Zuñiga. Mg. CE.

ÍNDICE DE CONTENIDOS

1. INTRODUCCIÓN 1

2. Planteamiento del problema de investigación 2

2.1. Formulación del problema 3

2.2. Definición del objeto de estudio 3

3. OBJETIVOS 3

3.1. Objetivo general 3

3.2. Objetivos específicos 3

3.3. El campo de acción 3

4. JUSTIFICACIÓN 4

5. ANTECEDENTES 5

5.1. Bases teóricas 8

5.1.1. Minería de datos 8

5.1.2. Técnicas de minería de datos 9

5.1.3. Arboles de decisión 10

5.1.4. WEKA 12

5.1.5. Proceso de enseñanza 14

5.1.6. Proceso de enseñanza con estudiantes con necesidades educativas especiales 16

6. METODOLOGÍA 18

7. RESULTADOS Y DISCUSIÓN 21

8. CONCLUSIONES 33

9. RECOMENDACIONES 34

10. REFERENCIAS BIBLIOGRÁFICAS 35

11. ANEXOS 3

Índice de tablas

Tabla 1: *Atributos con sus respectivos valores posibles dentro de la investigación* 26

Tabla 2: *Algoritmos clasificadores de WEKA* 37

Tabla 3: *Aplicaciones para la predicción* 38

Índice de figuras

Fig. 1: *WEKA GUI Chooser* 22
Fig. 2: *Metodología basada de proyectos de minería de datos* 30
Fig. 3: *Datos de Excel (XML)* 31
Fig. 4: *Carga de archivo en WEKA* 31
Fig. 5: *Modelo filtrado del procesamiento y análisis de datos en WEKA* 33
Fig. 6: *Clasificación del procesamiento de los datos* 34
Fig. 7: *Clasificación de datos en WEKA por el algoritmo J48* 34
Fig. 8: *Composición del algoritmo J48* 35
Fig. 9: *Aplicativo de datos en WEKA por el algoritmo J48* 35
Fig. 10: *Modelo árbol de decisión de los datos en WEKA (Datos Generales)* 36
Fig. 11: *Modelo árbol de decisión de los datos en WEKA (Datos clasificados)* 36
Figura 12. *Modelo de datos clasificados en el algoritmo BayesNet* 37
Figura 13. *Modelo de datos clasificados en el algoritmo OneR -B 6* 37

UNIVERSIDAD ESTATAL DEL SUR DE MANABÍ

INSTITUTO DE POSGRADOS

MAESTRÍA EN TECNOLOGÍAS DE LA INFORMACIÓN Y LA COMUNICACIÓN

Título: Técnica de minería de datos para procesos educativos en estudiantes con necesidades educativas especiales basado en un modelo predictivo

Autora: Hilda Brigitte Quijije Quiroz

Tutora: Lcda. Kirenia Maldonado Zuñiga. Mg. CE.

Resumen

En el contexto educativo, la minería de datos puede ser utilizada para analizar los datos enfocados en el proceso de enseñanza de los estudiantes y los procesos educativos que ayuden a identificar los factores que influyen desde diversos aspectos a partir de la retención del aprendizaje, rendimiento académico, la asistencia y la participación en actividades extracurriculares. Esto puede permitir a los educadores tomar medidas para mejorar los temas y habilidades en los que los estudiantes con necesidades educativas especiales y adaptar el plan de estudios en consecuencia. Esta investigación tuvo el objetivo de diseñar un modelo predictivo mediante técnicas de minería de datos en procesos educativos aplicados a estudiantes con necesidades educativas especiales en la Unidad Educativa “Plan Internacional” cuyo desarrollo se realizó mediante el uso de WEKA como herramienta para el análisis de datos y modelado predictivo presentando un aporte en la búsqueda de estrategias que permitan mejorar el proceso de enseñanza con los estudiantes con NNE. Se utilizaron métodos de investigación científica para sustentar el proceso de caracterización del estudio basado en un enfoque cuantitativo a nivel descriptivo y de campo.

Palabras clave: estrategias; herramientas; procesos educativos**.**

UNIVERSIDAD ESTATAL DEL SUR DE MANABÍ

INSTITUTO DE POSGRADOS

MAESTRÍA EN TECNOLOGÍAS DE LA INFORMACIÓN Y LA COMUNICACIÓN

Title: Data mining technique for educational processes in students with special educational needs based on a predictive model

Author: Hilda Brigitte Quijije Quiroz

Tutora: Lcda. Kirenia Maldonado Zuñiga. Mg. CE.

Abstract

In the educational context, data mining can be used to analyze data focused on the process of teaching students and educational processes to help identify the factors that influence various aspects from learning retention, academic performance, attendance and participation in extracurricular activities. This may enable educators to take steps to improve the subjects and skills in which students with special educational needs and adapt the curriculum accordingly. This research had the objective of designing a predictive model through data mining techniques in educational processes applied to students with special educational needs in the Educational Unit "Plan Internacional" whose development was carried out with WEKA as a tool for data analysis and predictive modeling presenting a contribution in the search for strategies to improve the teaching process with students with NNE. Scientific research methods were used to support the research characterization process based on a quantitative approach at a descriptive and field level.

Keywords: educational processes; strategies; tools.

UNIVERSIDAD ESTATAL DEL SUR DE MANABI
Creada el 7 de febrero del 2001, según Registro Oficial # 261
CENTRO DE IDIOMAS

CERTIFICADO No. 647

Lic.
Kirenia Maldonado Zúñiga
COORDINADOR DE LA MAESTRIA EN TECNOLOGÍA DE LA INFORMACIÓN Y LA COMUNICACIÓN– POSTGRADO - UNESUM
En su despacho.-

De mi consideración:

Por medio de la presente me permito CERTIFICAR que fue corregido el Summary, correspondiente a la Tesis de Grado **"TÉCNICA DE MINERÍA DE DATOS PARA PROCESOS EDUCATIVOS EN ESTUDIANTES CON NECESIDADES EDUCATIVAS ESPECIALES BASADO EN UN MODELO PREDICTIVO".** Previo a la obtención del título de Magister en Tecnología de la Información y la Comunicación al maestrante **Hilda Brigitte Quijije Quiroz,** mismo que fue corregido por la Lic. Liliam Rosalía Sánchez Choéz, Mg, Eii.

Particular que hago extensivo para los fines consiguientes.

Jipijapa, 3 de Julio del 2023.

Atentamente,

UNIVERSIDAD ESTATAL DEL SUR DE MANABI
UNESUM
CENTRO DE IDIOMAS
JIPIJAPA - MANABI

Lic. Paola Yadira Moreira Aguayo, Mg. Eii.
COORDINADORA DEL CENTRO DE IDIOMAS

Cc. Archivo
PYMA/jt

1. INTRODUCCIÓN

Las Tecnologías de la Información y comunicación (TIC) aportan al campo de la educación aspectos innovadores que suponen una mejora cualitativa en las formas de enseñar y aprender con un desarrollo multidisciplinario en el que intervienen varias ciencias, entre ellas la pedagogía y la computación. Su utilización contribuye a que los procesos de enseñanza-aprendizaje sean más colaborativos, interactivos y flexibles a partir de la aplicabilidad que se alcanza en la personalización de la enseñanza.

Hasta el actual momento el uso de las tecnologías enfocadas en la minería de datos representa una alternativa eficaz para el análisis de datos y la vigente necesidad de mejorar la calidad educativa presenta múltiples pautas para transformar la visión activa de enseñanza de forma espiral, la forma autónoma de trabajo en grupos seleccionados se efectúa en base a la referencia del efecto mateo y su contextualización muchas veces se basa en las ventajas acumulativas que se presentan de forma indirecta, sin embargo desde la mirada en forma espiral la perspectiva incide en la correcta integración de procesos que conlleven la utilización de nuevas metodologías y la oportunidad de clasificar, comparar y evaluar pautas mediante softwares que determinen un análisis evaluativo a través de técnicas basadas en minería de datos por medio de la extracción de registros en etapas iterativas que permitan obtener conocimiento nuevo.

Desde los factores la parte más denotable se basa en el rendimiento académico de estudiantes con necesidades educativas especiales que puede variar dependiendo de la naturaleza y el grado de la discapacidad, así como de los recursos disponibles para el estudiante. Sin embargo, con los apoyos adecuados, muchos estudiantes con pueden alcanzar un rendimiento académico comparable al de sus compañeros. Por ello, es importante que se les brinden los apoyos y recursos adecuados, incluyendo la adaptación

de los materiales de enseñanza y evaluación, el acceso a tecnologías asistivas, el uso de intérpretes y servicios de apoyo emocional.

También es importante que los docentes y el personal de la escuela estén capacitados en cómo trabajar con estudiantes con discapacidad y cómo proporcionar los apoyos y recursos necesarios. La colaboración entre los padres, los docentes y otros miembros del equipo de apoyo también puede ser fundamental para mejorar el rendimiento académico de los estudiantes con diversas necesidades y que estén apegadas a las adaptaciones curriculares según los lineamientos del ministerio de educación.

2. Planteamiento del problema de investigación

Las TIC en áreas de conocimientos es notable en los actuales momentos, por tanto, el uso de herramientas que permitan mejorar estos procesos es importante en cada campo, no obstante, la validación de los modelos generados para la extracción de nuevos lineamientos que aborden metodologías basadas y filtradas por técnicas de minerías de datos que potencien estos procesos, en cuanto a lo determinado existe un punto de inflexión entre dos líneas que parten en base a la educación y las tecnologías sobre un grupo en específico con el cual se trabajará dentro de este proyecto.

Dentro de lo abordado se plantea la aplicación de minería de datos para determinar roles respecto a la usabilidad de herramientas tecnológicas con estudiantes con necesidades educativas especiales en la Unidad Educativa “Plan Internacional” perteneciente al Cantón Daule, la clasificación, errores y redundancias presentes en los datos a obtener, así como recodificar algunas variables que permitan contrastar el rendimiento académico de los estudiantes en base a nuevos modelos en la enseñanza de este grupo seleccionado.

2.1. Formulación del problema

¿Qué aporte brindara aplicar técnicas de minería de datos con el uso del WEKA en el proceso de enseñanza - aprendizaje de estudiantes con necesidades educativas especiales de la Unidad Educativa Plan Internacional?

2.2. Definición del objeto de estudio

Minería de datos

3. OBJETIVOS

3.1. Objetivo general

Desarrollar un modelo predictivo mediante técnicas de minería de datos en procesos educativos aplicados a estudiantes con necesidades educativas especiales en la Unidad Educativa Plan Internacional.

3.2. Objetivos específicos

Efectuar el levantamiento de información histórica del rendimiento académico de los estudiantes con necesidades educativas especiales.

Depurar la información más relevante en el proceso de análisis de los datos obtenidos mediante el software WEKA con la aplicación de la técnica de árboles de decisión utilizando como base al algoritmo J48.

Determinar el modelo predictivo del rendimiento académico mediante WEKA con el uso de técnicas de minería de datos en procesos educativos en estudiantes con necesidades educativa en la Unidad Educativa Plan Internacional.

3.3. El campo de acción

Procesos educativos en la Unidad Educativa Plan Internacional.

4. JUSTIFICACIÓN

El presente trabajo presenta se justifica debido a dos aspectos esenciales que abordan el determinar la tendencia y aplicabilidad de las TIC en estudiantes con necesidades educativas especiales y el contraste significativo de un modelo predictivo mediante el uso de técnicas de minería de datos y algoritmos de identificación, se trabajara con plataformas que mediaran la aceptabilidad, progreso y niveles de dificultad de dichas plataformas en este grupo de estudiantes lo que ayuda a mejorar la calidad de la educación al predecir las dificultades encontradas al trabajar con estos estudiantes . A través de este artículo, también se investiga la precisión de algunas técnicas de clasificación como enfoque principal en la selección de datos,

En cuanto a la investigación es preciso mencionar el uso de un árbol de decisión, cada nodo representa una variable de entrada y cada rama representa una posible decisión o acción. El objetivo de este se centra en la clasificación los datos de entrada en categorías o valores de salida basados en las decisiones tomadas en los nodos mediante el uso del algoritmo J48. Donde pueden ser utilizados en muchos campos son utilizados para tomar decisiones basadas en datos de forma asertiva así mismo la determinación del área y factibilidades de la Unidad Educativa pues al ser un área rural es difícil el acceso a tecnologías, por ende, el porqué de esta investigación es presentar de forma representativa las áreas de potencialización de dichos estudiantes y las recomendaciones propias resultantes de la investigación.

En base a ello es preciso mencionar que el grupo seleccionado son estudiantes con necesidades educativas especiales, por lo tanto se referencia el rendimiento de ellos en las asignaturas de matemática y lengua y literatura como conocer el intervalo de ellas, así mismo es factible que en base a ello se busquen alternativas que mejoren y potencialicen

el trabajo con este grupo basado en actividades lúdicas en estas áreas como la introducción de herramientas tecnológicas en base a el trabajo con ellos.

5. ANTECEDENTES

(Jiménez Toledo & Timarán Pereira, 2015) realizaron la investigación titulada, Caracterización de la deserción estudiantil en educación superior con minería de datos. Teniendo como objetivo principal detectar patrones de deserción estudiantil a partir de los datos socioeconómicos, académicos, disciplinares e institucionales de los estudiantes de los programas de pregrado de la Universidad de Nariño e Institución Universitaria CESMAG. La investigación aplicada fue de tipo descriptivo bajo el enfoque cuantitativo, aplicando un diseño no experimental, la ruta es cuali-cuantitativa (mixta). Las técnicas de minería de datos aplicadas fueron clasificación basada en árboles de decisión, asociación y clustering, utilizando la herramienta libre de minería de datos WEKA. Obteniendo como resultado la detección de patrones de deserción estudiantil confiables utilizando árboles de decisión se generaron 35 árboles variando el factor de confianza C de 0.1 hasta 0.5 con un incremento de 0.1 y el número de instancias por nodo M de 10 en 10 iniciando en 10 hasta 70. Llegando a la conclusión que los perfiles de deserción estudiantil obtenidos a través de las técnicas de minería de datos: clasificación, asociación y agrupamiento indican que éstas son capaces de generar modelos consistentes con la realidad observada y el respaldo teórico, basándose únicamente en los datos que se encontraron almacenados en las bases de datos.

(Díaz Landa et al., 2021) en el articulo científico con la temática, Rendimiento académico de estudiantes en educación superior: predicciones de factores influyentes a partir de árboles de decisión, mostrando como objetivo principal predecir el rendimiento académico de estudiantes de maestrías en educación, mediante la introducción de la

investigación aplicada y el uso de técnicas de clasificación de minería de datos utilizando el algoritmo J48 aportaron mediante la predicción la clasificación el análisis de la base de entrenamiento entre los atributos de la predicción respectiva correspondiendo a la exactitud del modelo, así mismo los resultados obtenidos pone en evidencia que es posible aplicar la metodología de minería de datos empleando árboles de decisión para generar el modelo de predicción del rendimiento académico, se demuestra la asociación en la predicción del rendimiento académico empleando árbol de decisión conjuntamente con J48 algoritmo bayesiano del software WEKA.

(Espinoza Mina, 2018) en la revisión sistemática de su artículo titulado, WEKA, áreas de aplicación y sus algoritmos: una revisión sistemática de literatura partiendo desde el objetivo compilar una lista de los campos o áreas de actividad en los que se utiliza WEKA como una herramienta de análisis de datos y los algoritmos que se emplean. Ajustando criterios de inclusión y exclusión de la selección de material científico mediante el diagrama prisma. al revisar los artículos científicos primarios se encontró que nueve de ellos están aplicados a la ciencia de la informática, por cuanto demuestran la efectividad de WEKA como herramienta computacional de aprendizaje automático y minería de datos además muestra al algoritmo Naïve Bayes como el más utilizado en el aprendizaje inductivo de estos procesos. Se debe diferenciar el uso de las herramientas como WEKA y la programación de algoritmos específicos para determinados problemas; la programación tiene que ver principalmente con el desarrollo de una solución a un problema específico, que herramientas ya desarrolladas no pueden resolverlos de forma óptima y esto lo que buscan los algoritmos encontrados en los artículos revisados.

(Quintana Bajaña & Yagual Tomalá, 2017) en la investigación, Propuesta de aplicación predictiva de aprobación de una asignatura con flujo previo a través de algoritmos basados en software WEKA para estudiantes del último semestre de la Carrera de Ingeniería en

Sistemas Computacionales de la Universidad de Guayaquil. Para obtener el título profesional de Ingeniero En Sistemas Computacionales, cuyo objetivo se enfocó desarrollar un análisis evaluativo a través de técnicas basadas en minería de datos por medio de la extracción en registros de calificaciones mediante la aplicación WEKA. La modalidad empleada en esta propuesta definió varios métodos y procesos los cuáles nos permitieron obtener la información, obteniendo una muestra al azar de registros de las calificaciones de estudiantes arribando a el resultado esperado para el proceso de clasificación de los datos. La minería de datos es un factor muy importante que permite tomar decisiones sobre el comportamiento de los datos. No existe ninguna herramienta que asegure una confianza del 100% en lo que respecta al ámbito de la minería de datos para ejercer predicciones sobre registros ya existentes que son estimados para una determinada evaluación.

(Rodríguez Rodríguez et al., 2021) en el artículo, Implementación de algoritmos de Inteligencia Artificial en la predicción de nuevos conocimientos mediante enseñanza constructivista, describen describe una solución al escenario antes planteado a partir de la implementación de algoritmos computaciones de Inteligencia Artificial en la predicción de nuevos conocimientos mediante una enseñanza constructivista, así mismo, Se aplicó en un escenario real con datos personales, institucionales y las interacciones de los estudiantes. La gestión fue guiada por la metodología CRISP - DM que permita desarrollar el proyecto, se hizo uso de la metodología de minería de datos entre las cuales se ostentan las siguientes fases, Compresión del Negocio, Objetivos del Negocio, Comprensión de los Datos, Recolección de Datos iníciales, Preparación de los Datos, Estructura de los Datos, A partir del estudio realizado, se implementó en la presente investigación la metodología de minería de datos CRISP - DM para el análisis y procesamiento de las informaciones. La implementación de algoritmos computacionales

sustentados en Inteligencia Artificial permitió la predicción de nuevos conocimientos mediante un enfoque de enseñanza constructivista.

5.1. Bases teóricas

5.1.1. Minería de datos

La minería de datos consiste en extraer datos implícitos, previamente desconocidos, y el valor potencial para el conocimiento de la toma de decisiones y reglas a partir de una gran cantidad de datos. Con la continua expansión del mercado de minería de datos, el estudio relacionado con la minería de datos es penetrada constantemente. La minería de datos es una ciencia aplicada, y una tecnología práctica desde el mismo principio, por lo que es muy importante seleccionar el apropiado herramientas de minería de datos en análisis de inteligencia de minería de datos en cuenta de su fuerte carácter experimental. (Shen, Liu, & Shen, 2010).

Así mismo diversos autores destacan a la minería de datos como el proceso de descubrir nuevas e importantes relaciones, patrones y tendencias utilizando diversas tareas en grandes volúmenes de datos. Estas tareas tienen como objetivo encontrar patrones, perfiles y tendencias a través del análisis de datos utilizando clasificación, agrupamiento, patrones de secuencia y relaciones, entre otros (Timarán Pereira et al., 2020).

En este contexto se enfoca una correlación de las áreas aplicadas en la minería de datos entre las cuales se deduce la siguiente investigación:

Suénaga (2015) indica que la minería de datos en la educación (MDE) no es un concepto nuevo, su estudio y aplicación ha tomado mayor relevancia en los últimos años. La utilización de las técnicas de MD permite deducir fenómenos dentro del ámbito educativo; de esta forma, es posible determinar la probabilidad de desertar o continuar con sus estudios de los estudiantes, así como el desempeño de estos durante el cursado.

El producto final de los modelos beneficia a estudiantes, docentes, padres y gestores de la educación, no sólo para informar sobre la situación de los estudiantes cuyo desempeño podría estar asociado con una característica particular (positiva o negativa), sino también como asesoramiento para la toma de decisiones. Dicho de otra manera, se pretende que estos modelos finales faciliten la reflexión y la autorregulación durante los estudios.

La minería de datos estudia algoritmos y paradigmas computacionales que permitir que las computadoras descubran la estructura en las bases de datos, realicen predicción y pronóstico, y en general mejorar su rendimiento a través de la interacción con los datos. El aprendizaje automático es preocupado por la construcción de sistemas informáticos que tienen la capacidad de mejorar su desempeño en un dominio dado a través de la experiencia. El aprendizaje automático y la minería de datos son cada vez más áreas importantes de la ingeniería y la informática y han aplicado con éxito a una amplia gama de problemas en la ciencia e ingeniería. Recientemente, reconociendo la importancia de estas áreas en informática e ingeniería (Markov & Russell, 2006).

La aplicación de la minería de datos educacional en diversas investigaciones plantea una solución para la mejora del rendimiento académico, pero el análisis de grandes cantidades de datos y la experimentación de diferentes métodos y técnicas para encontrar un modelo que sea fiable suponen de tiempo y esfuerzo. Existen técnicas de investigación documental que facilitan el análisis de toda la información existente acerca de un tema en común, aplicando técnicas que pueden determinar la concordancia de los resultados y comparar las observaciones de los diferentes autores para llegar a una conclusión acertada (Santiesteban et al., 2017).

5.1.2. Técnicas de minería de datos

Las técnicas de minería de datos son pasos estructurados en el tiempo y se derivan de elementos de inteligencia artificial y estadísticas, son algoritmos avanzados aplicados a

la estructura. Los datos se pueden utilizar para producir ciertos resultados. Las técnicas conocidas según diversos autores son:

Redes neuronales. - Son un paradigma de aprendizaje y procesamiento automático inspirado en la forma en que funciona el sistema nervioso de los animales. Se trata de un sistema de interconexión de neuronas en una red que colabora para producir un estímulo de salida. Algunos ejemplos de red neuronal son: el perceptrón y perceptrón multicapa.

Regresión lineal. - Es la más utilizada para formar relaciones entre datos. Rápida y eficaz pero insuficiente en espacios multidimensionales donde puedan relacionarse más de 2 variables

Árboles de decisión. - Un árbol de decisión es un modelo de predicción utilizado en el ámbito de la inteligencia artificial, dada una base de datos se construyen estos diagramas de construcciones lógicas, muy similares a los sistemas de predicción basados en reglas, que sirven para representar y categorizar una serie de condiciones que suceden de forma sucesiva, para la resolución de un problema (Diaz, 2021).

5.1.3. Arboles de decisión

Los árboles de decisión son una técnica utilizada en la minería de datos para la clasificación y la predicción, utiliza una estructura de árbol para representar una serie de decisiones y sus posibles resultados.

Dada una base de datos se construyen diagramas de construcciones lógicas, muy similares a los sistemas de predicción basados en reglas, que sirven para representar y categorizar una serie de condiciones que ocurren de forma sucesiva, para la resolución de un problema. suele contener nodos internos, nodos de probabilidad, nodos hojas y arcos. Un nodo interno contiene una prueba sobre algún valor de una de las propiedades. Un nodo de probabilidad indica que debe ocurrir un evento aleatorio de acuerdo con la

naturaleza del problema, este tipo de nodos es redondo, los demás son cuadrados. Un nodo hoja representa el valor que devolverá el árbol de decisión y finalmente las ramas brindan los posibles caminos que se tienen de acuerdo con la decisión tomada (Diaz, 2021).

La tarea de aprendizaje para la cual los árboles de decisión se adecuan mejor es la clasificación. Por tanto, dentro de la redacción se plantean diversas posturas que indican, clasificar es determinar de entre varias clases a qué clase pertenece un objeto; la estructura de condición y ramificación de un árbol de decisión es idónea para resolver diversos problemas. Una de las ventajas de este tipo de modelos, en relación, por ejemplo, a las redes neuronales o los vectores de soporte, es que el resultado que se obtiene es inteligible para los seres humanos (y también para sistemas semi-automáticos que procesen reglas) (Miranda & Guzmán, 2017).

Los árboles de decisión son útiles en la minería de datos porque son fáciles de entender e interpretar, y pueden manejar datos categóricos y numéricos. Sin embargo, pueden ser propensos al sobreajuste y la complejidad, lo que puede reducir su precisión. Por lo tanto, es importante equilibrar la complejidad del árbol con la precisión de las predicciones mediante técnicas como la poda de árboles y la selección de características.

Los árboles de decisión construyen modelos de clasificación de forma ramificada. Su implementación se basa en la inducción de reglas, y la iteración bajo el enfoque de división y segmentación. La aplicación de árboles de decisión con la técnica J48, o también conocida como C4.5, se presenta como un algoritmo que desarrolla estructura en forma de reglas basado en subconjuntos de datos extraídos del conglomerado total de datos de entrenamiento. De esta forma el subconjunto de reglas requiere ser evaluado en su bondad y precisión para garantizar la efectividad en sus predicciones. El árbol de

decisión construye modelos de clasificación en forma de árboles. Cada nodo interior en estos árboles representa una de las variables de entrada, y tiene un número de ramas igual al número de valores posibles de esa variable de entrada. Cada nodo hoja tiene un valor del atributo de destino. El nodo hoja representa la decisión tomada en función de los valores de las variables de entrada desde la raíz hasta la hoja (Ospina-Mateus & Quintana Jiménez, 2019).

5.1.4. WEKA

WEKA es una herramienta de plataforma de minería de datos para el aprendizaje automático, las primeras pruebas se ejecutaron en la Universidad de Waikato, en Nueva Zelanda en el año 1997. WEKA es un conjunto de Algoritmos de Machine Learning y Data Mining. Esta diseñado bajo un software programado en lenguaje JAVA el cual tiene una interfaz GUI para interactuar con archivos y conjuntos de datos predefinidos con 49 herramientas de preprocesamiento para ellos, contiene herramientas bajo 76 algoritmos de clasificación, 15 atributos evaluadores y 10 algoritmos de búsqueda para la selección de características.

Existen 3 algoritmos para encontrar varias reglas de asociación, también tiene 3 interfaces gráficas de usuario: "The Explorer", "The Experimenter" y "The Knowledge Flow". El formato de archivo o la extensión utilizada para almacenar datos en WEKA es ARFF, formato de archivo de relación de atributos. También incluye herramientas para la visualización. WEKA tiene la capacidad de expandirse y contener los nuevos algoritmos; estos algoritmos expandidos pueden aplicarse directamente al conjunto de datos para Machine Learning (Meena & Choudhary, 2017).

Existen varios enfoques de algoritmos de agrupamiento, debido a que, para un mismo conjunto de datos, aplicando diferentes algoritmos de agrupamiento se pueden obtener

resultados diferentes. Por lo que surge la necesidad de evaluar las particiones obtenidas y poder determinar la calidad de los resultados alcanzados. Existen medidas que permiten realizar una evaluación de la estructura resultante de cada algoritmo de agrupamiento, obteniendo de manera cuantitativa un valor de calidad de estas. Estas medidas de calidad se conocen bajo el nombre de índices de validación de grupo (Castillo et al., 2020).

Así mismo se denota dentro de su interfaz algunas opciones que permiten realizar trabajos de minería de datos, entre las cuales se destacan:

Explorer: Es la interfaz visual de WEKA para trabajar de manera gráfica de una manera sencilla. Este modo permite procesar, clasificar, asociar y visualizar datos de una manera fácil e intuitiva sobre un sólo archivo de datos.

Experimenter: es un modo útil para aplicar uno o varios métodos de clasificación de manera automática. Con esta ventana se facilita la realización de experimentos a gran escala.

KnowledgeFlow: es la interfaz gráfica, y es utilizado para desarrollar proyectos a través de flujos de información.

Simple CLI: se le conoce como interfaz línea de comandos, y se usa para llamar directamente a los paquetes de Java que WEKA incorpora (Navas, 2016).

Figura1. *WEKA GUI Chooser*

Según la perspectiva de Charalampopoulos & Anagnostopoulos (2011) en el estudio baso en "*Un estudio comparable que emplea algoritmos de agrupamiento/clasificación WEKA para la clasificación de páginas web*" ostenta que, la clasificación de datos es un problema en Ciencias de la Información pues describe la difícil tarea de asignar un documento a uno o más categorías según su contenido. Muchos algoritmos se han creado con el uso de métodos con sus aplicaciones que varían de publicidad, indexación de bases de datos, motores de búsqueda (como mencionado) y muchos más. La mayoría de los algoritmos utilizados denotan ciertas inconsistencias en la búsqueda del conocimiento científico que pone en riesgo la veracidad de integración y búsqueda del conocimiento mediante estos datos e información.

5.1.5. Proceso de enseñanza

Las sociedades mundiales consideran importante la necesidad de enriquecer la formación cultural del hombre, cuya preparación lo coloca a la cabeza del desarrollo del mundo moderno; un hombre civilizado que comprende los problemas de su contexto en su origen y desarrollo que entrar con firmeza en la batalla de las ideas para adoptar una actitud transformadora encaminada a la consecución de los ideales sociales de nuestro tiempo (Barcia & Carvajal, 2015).

El proceso de enseñanza es un conjunto de actividades que se llevan a cabo con el fin de transmitir conocimientos, habilidades y valores de un individuo a otro. Este proceso implica la participación de un docente y un estudiante, quienes trabajan juntos para alcanzar los objetivos de aprendizaje.

Desde esta perspectiva analítica, las metas tienen un significado fundamental, no tanto enunciados fijos e inalterables a lograr a toda costa, sino pautas que guíen el proceso educativo respecto a la enseñanza. No se puede suponer que las metas u objetivos fijados por el educador deban ser necesariamente traducido directamente en resultados tangibles, observables y medibles. Por un lado, porque los estudiantes se diferencian sustancialmente entre sí, comienzan su aprendizaje con niveles de desarrollo diferentes en sus conocimientos y habilidades, provienen de diferentes familias y entornos sociales, historia de vida específica y rasgos de personalidad muy diferentes. Las experiencias de formación son experienciales y están relacionadas con diferentes planes de referencia. Por otra parte, y con base en lo anterior, la vida de cada estudiante transcurre bien durante el proceso mismo (Bravo & Cáceres, 2016).

El proceso de enseñanza se divide en varias etapas, que incluyen la planificación, la implementación y la evaluación. En la etapa de planificación, el docente identifica los objetivos de aprendizaje y selecciona los métodos y materiales que se utilizarán para alcanzarlos. Durante la implementación, el docente utiliza diferentes estrategias de enseñanza, como la lectura, la escritura, la discusión en grupo, la resolución de problemas y la práctica.

La evaluación es otra parte importante del proceso de enseñanza, ya que permite al docente medir el progreso de los estudiantes y determinar si se han alcanzado los

objetivos de aprendizaje. La evaluación puede ser formativa, es decir, se realiza durante el proceso de enseñanza, o sumativa, que se realiza al final del proceso.

El proceso de enseñanza puede verse afectado por varios factores, incluidos el ambiente de aprendizaje, la motivación del estudiante, el estilo de enseñanza del docente y el nivel de preparación del estudiante. Por lo tanto, es importante que los docentes consideren estos factores al planificar y llevar a cabo el proceso de enseñanza.

5.1.6. Proceso de enseñanza con estudiantes con necesidades educativas especiales

El proceso de enseñanza dirigido a estudiantes con necesidades educativas especiales es similar al proceso educativo común, pero requiere consideraciones especiales para garantizar que se satisfagan las necesidades únicas de los estudiantes con discapacidades.

Es importante que los maestros sean conscientes de las discapacidades específicas de los estudiantes y de sus fortalezas y limitaciones individuales. Esto les permite adaptar el aprendizaje a las necesidades de cada alumno, desde esta perspectiva (Yenchong & Barcia, 2020), indican desde la mirada de otros autores que el proceso de desarrollar y reconocer las capacidades que tienen los estudiantes con NEE, se desarrollan estrategias de adaptación cuyas funciones cumplen con el sistema educativo obteniéndose ganancias en los objetivos que se pretenden alcanzar y las prácticas inclusivas que ayuden a mejorar la pertenencia de los estudiantes con sus respectivas instituciones; pero no solo se refieren a las oportunidades de aprender, sino también de participar en la vida social escolar, sentir aceptación por parte de sus compañeros y el acompañamiento docente adecuado.

La selección de estrategias de enseñanza apropiadas para estudiantes con discapacidades. Por ejemplo, los estudiantes con discapacidades visuales pueden necesitar materiales de aprendizaje en Braille o en letra grande, mientras que los

estudiantes con discapacidades auditivas pueden necesitar intérpretes de lenguaje de señas o materiales visuales para apoyar la comprensión de los conceptos. Tercero, los maestros deben ser sensibles a las necesidades emocionales de los estudiantes con discapacidades. Pueden experimentar frustración, estrés y ansiedad debido a las limitaciones de aprendizaje, por lo que es importante brindarles apoyo emocional y recursos y servicios de apoyo para ayudarlos a lidiar con estos sentimientos.

Es la educación uno de los espacios más demandados pues esta resulta vital para lograr la inserción de las personas con discapacidad a una sociedad en la que deberán desenvolverse como actores activos de la misma y para ello deben ser preparados, a la vez, el resto de la sociedad debe prepararse para convivir con ellos donde se haga con respeto, aceptación y de forma activa (Santiesteban et al., 2017)

En definitiva, se puede afirmar que el proceso de aprendizaje de los alumnos con necesidades educativas especiales requiere una atención especial a las necesidades individuales de cada alumno, así como a la elección de estrategias didácticas y la creación de un entorno de aprendizaje seguro y sin barreras. Esto permite a los maestros ayudar a los estudiantes con discapacidades a alcanzar su máximo potencial y alcanzar el éxito académico, en base a ello (Rojas Salgado, 2021), menciona que sin duda alguna, las estrategias didácticas que emprendan los docentes deben sustentarse en la diversidad de características que presentan los estudiantes, especialmente aquellos con NEE y, a su vez, aprovechar los beneficios que brindan las tecnologías para inclusión educativa de las personas en situación de discapacidad. La variabilidad y flexibilidad son aspectos clave de estas, por ello simbolizan un gran potencial para integrarlas, en particular, en el proceso educativo de los estudiantes con NEE. En ese sentido, incorporar en las estrategias didácticas, tanto materiales físicos como virtuales que demanden de tecnología, permitirá abordar esa diversidad de características de la población estudiantil.

6. METODOLOGÍA

La asunción del presente trabajo presenta una investigación aplicada, cualitativa y su fundamentación se basa a partir de un enfoque interpretativo, analítico con el entorno y la búsqueda de soluciones mediante la investigación abierta de los conceptos y la experimentación aplicable de técnicas y métodos científicos en pro al objeto de la investigación.

Así mismo el empleo de métodos de recolección de datos con el uso de recursos estadísticos para la exposición de datos que provienen de un cálculo o medición.

En su contexto se hace uso de los métodos teóricos, empíricos y matemáticos los cales ostentan el trabajo bajo una revisión bibliográfica de los acontecimientos e investigaciones realizadas con aportes significativos que aportan en el desarrollo de la aplicabilidad de este trabajo, así mismo a nivel práctico de hace uso de la metodología CRISP-DM (Cross Industry Standard Process for Data Mining) compuesta de seis etapas: análisis del problema, análisis de los datos, preparación de los datos, modelado, evaluación e implementación.

Métodos teóricos:

Histórico lógico: se utilizó para el estudio explícito de los antecedentes y bases teóricas que fundamentan la información de hechos aplicados, verídicos, de relevancia para la investigación.

Análisis-síntesis: se utilizó en el análisis de los componentes de la investigación, grupos identificados y seleccionados en el tratamiento de datos que posibilita la sistematización de las estrategias relacionadas en el desarrollo del trabajo de investigación.

Inducción–deducción: se hizo uso de este método en base al conocimiento enfocado en la naturaleza del problema de la investigación, compilar datos esenciales que permitan llegar a soluciones concretas a partir del objeto y campo de la investigación.

Sistémico-estructural-funcional: Desde la concepción particular de la investigación, basada en la determinación y definición de los componentes y relaciones demostrativas, en la fundamentación basada en modelo predictivo, se utilizó a partir de la identificación de la estructura, funcionalidad y niveles de relaciones jerárquicas entre los componentes del modelo así mismo los algoritmos utilizados durante la ejecución de minería de datos.

Métodos empíricos:

Observación: Método empleado en el análisis de fenómenos registrados en el campo de la investigación focalizando principales datos que permitan el ordenamiento de piezas que deán significado y alcance de la investigación.

Entrevista: Se utilizó en la recolección de orientaciones a partir de la interacción de autoridades de la Unidad educativa tomada para la realización de la investigación, así mismo la sistematización de adecuadas acciones y compromisos para favorecer a los grupos identificados en la investigación.

Método comparativo

Dada la investigación el método comparativo se utilizó como estrategia que para analizar y comprender las similitudes y diferencias entre dos o más elementos, fenómenos, grupos o situaciones en el procesamiento de los datos en WEKA.

Para ello se establecieron los algoritmos comparativos propios de WEKA:

Algoritmo J48:

Desde la concepción de la investigación, este algoritmo se utilizó para la clasificación de individuos o entidades netas del modelo predictivo basado en arboles de decisión dados por el programa WEKA, la identificación de estancias correctamente clasificadas y por ende la visualización del árbol de decisión en donde se identifican los periodos relativos al proceso de enseñanza aprendizaje de las asignaturas del tronco común (Matemática y Lengua y Literatura).

Algoritmo Bayesnet

Algoritmo utilizado en la comparativa con enfoque de representación de las relaciones probabilísticas entre variables e inferencias sobre los datos recopilados y procesados.

Algoritmo OneR -B 6

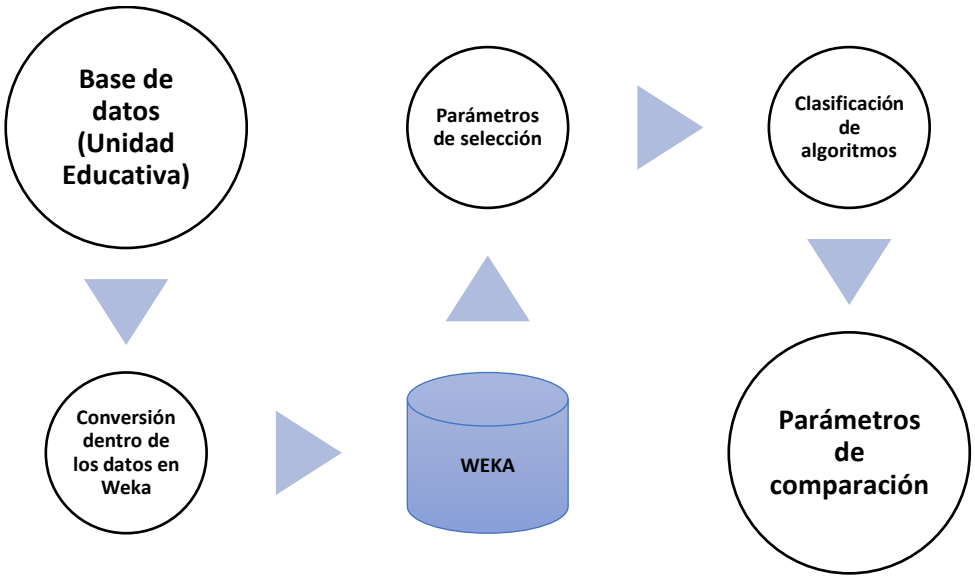
Se utilizó como algoritmo comparativo basado en un modelo de reglas de clasificación de una única variable, siendo así un algoritmo simple pero efectivo hacia la selección de la variable que mejor predice la clase y establece una regla basada en esa variable.

Método práctico de la investigación

Arquitectura de la propuesta metodológica

Este método se utilizó como guía del modelo de procesos de minería de datos basado en las fases del proyecto de investigación, cuya finalidad estableció que cada proceso genere una validez en los atributos seleccionados, algoritmos de clasificación y selección de modelos.

Figura 2: *Metodología basada de proyectos de minería de datos*



Fuente: Elaboración propia

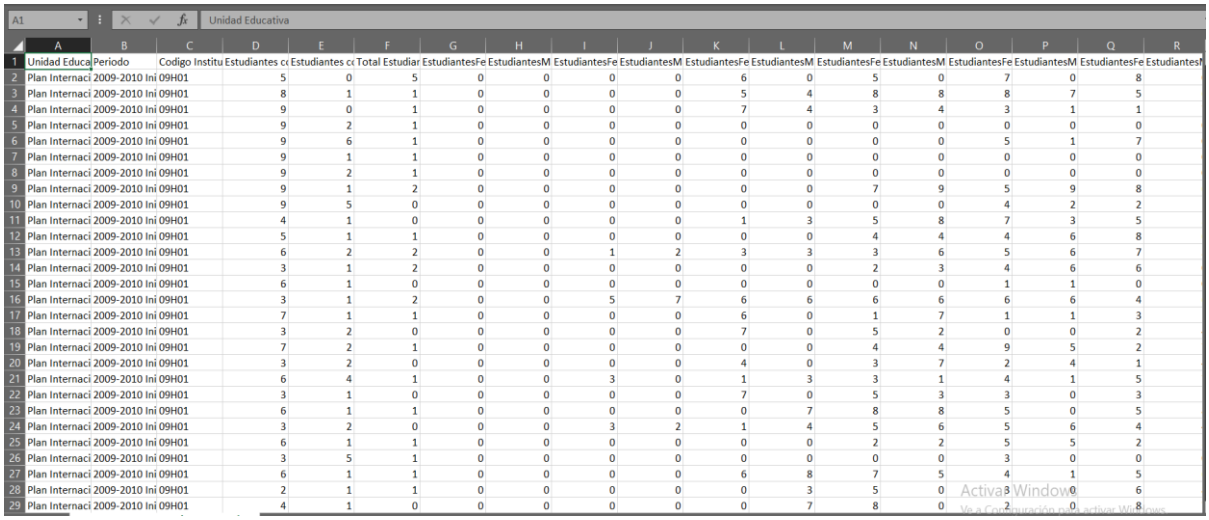
7. RESULTADOS Y DISCUSIÓN

Recopilación de datos

La investigación se encuentra basada por aspectos significativos, desde la instancia se enmarca el proceso de ETL, el alojamiento de los datos se atribuye al formato XML con datos nominales y ordinales.

Desde este punto se establecen las clases y atributos para ser procesados, por ellos WEKA permite realizar proceso de ETL partiendo desde la extracción de datos, desde el sistema original, en la transformación de datos de forma clara, intangibles y de utilidad, de carga de los datos de acuerdo con los requerimientos y dirección de la investigación.

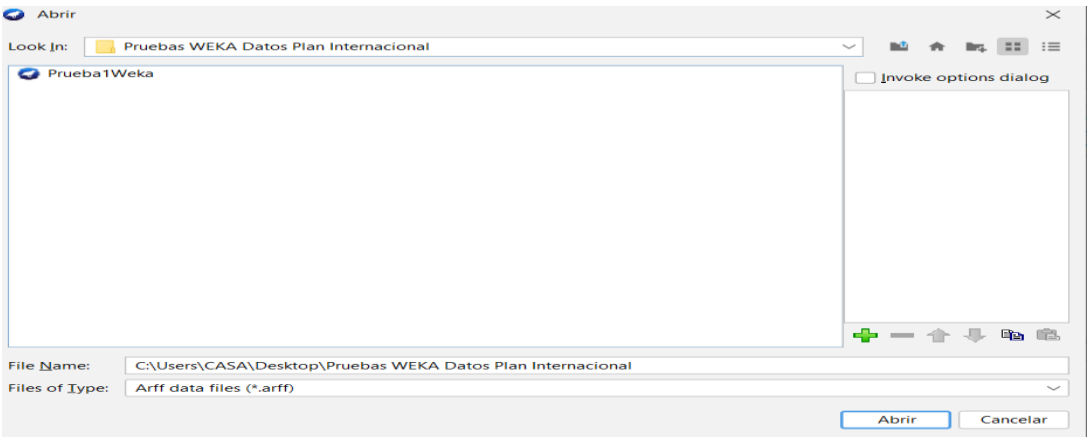
Figura 3. *Datos de Excel (XML)*



Unidad Educa	Periodo	Codigo Institu	Estudiantes c	Estudiantes c	Total Estudiar	EstudiantesFe	EstudiantesM	EstudiantesFe	EstudiantesM	EstudiantesFe	EstudiantesM	EstudiantesFe	EstudiantesM	EstudiantesFe	EstudiantesM	EstudiantesFe	EstudiantesM	EstudiantesFe
Plan Internaci	2009-2010 Ini	09H01	5	0	5	0	0	0	0	6	0	5	0	7	0	8		
Plan Internaci	2009-2010 Ini	09H01	8	1	1	0	0	0	0	5	4	8	8	8	7	5		
Plan Internaci	2009-2010 Ini	09H01	9	0	1	0	0	0	0	7	4	3	4	3	1	1		
Plan Internaci	2009-2010 Ini	09H01	9	2	1	0	0	0	0	0	0	0	0	0	0	0		
Plan Internaci	2009-2010 Ini	09H01	9	6	1	0	0	0	0	0	0	0	0	5	1	7		
Plan Internaci	2009-2010 Ini	09H01	9	1	1	0	0	0	0	0	0	0	0	0	0	0		
Plan Internaci	2009-2010 Ini	09H01	9	2	1	0	0	0	0	0	0	0	0	0	0	0		
Plan Internaci	2009-2010 Ini	09H01	9	1	2	0	0	0	0	0	0	7	9	5	9	8		
Plan Internaci	2009-2010 Ini	09H01	9	5	0	0	0	0	0	0	0	0	0	4	2	2		
Plan Internaci	2009-2010 Ini	09H01	4	1	0	0	0	0	0	1	3	5	8	7	3	5		
Plan Internaci	2009-2010 Ini	09H01	5	1	1	0	0	0	0	0	0	4	4	4	6	8		
Plan Internaci	2009-2010 Ini	09H01	6	2	2	0	0	1	2	3	3	3	6	5	6	7		
Plan Internaci	2009-2010 Ini	09H01	3	1	2	0	0	0	0	0	0	2	3	4	6	6		
Plan Internaci	2009-2010 Ini	09H01	6	1	0	0	0	0	0	0	0	0	0	1	1	0		
Plan Internaci	2009-2010 Ini	09H01	3	1	2	0	0	5	7	6	6	6	6	6	6	4		
Plan Internaci	2009-2010 Ini	09H01	7	1	1	0	0	0	0	6	0	1	7	1	1	3		
Plan Internaci	2009-2010 Ini	09H01	3	2	0	0	0	0	0	7	0	5	2	0	0	2		
Plan Internaci	2009-2010 Ini	09H01	7	2	1	0	0	0	0	0	0	4	4	9	5	2		
Plan Internaci	2009-2010 Ini	09H01	3	2	0	0	0	0	0	4	0	3	7	2	4	1		
Plan Internaci	2009-2010 Ini	09H01	6	4	1	0	0	3	0	1	3	3	1	4	1	5		
Plan Internaci	2009-2010 Ini	09H01	3	1	0	0	0	0	0	7	0	5	3	3	0	3		
Plan Internaci	2009-2010 Ini	09H01	6	1	1	0	0	0	0	0	7	8	8	5	0	5		
Plan Internaci	2009-2010 Ini	09H01	3	2	0	0	0	3	2	1	4	5	6	5	6	4		
Plan Internaci	2009-2010 Ini	09H01	6	1	1	0	0	0	0	0	0	2	2	5	5	2		
Plan Internaci	2009-2010 Ini	09H01	3	5	1	0	0	0	0	0	0	0	0	3	0	0		
Plan Internaci	2009-2010 Ini	09H01	6	1	1	0	0	0	0	6	8	7	5	4	1	5		
Plan Internaci	2009-2010 Ini	09H01	2	1	1	0	0	0	0	0	3	5	0	3	0	6		
Plan Internaci	2009-2010 Ini	09H01	4	1	0	0	0	0	0	0	7	8	0	2	0	8		

La instancia del requerimiento de datos son procesados de acuerdo con las necesidades, se parte desde una primera evaluación de los datos significativos y valorativos presentes para la continuidad del proceso, además se recalca que el formato dentro del procesamiento de WEKA es fundamental, desde este aspecto es importante asegurar el formato denominado ARFF por defectos, sin embargo presenta otras opciones de trabajo indicando CSV que son Archivos separados por comas o tabuladores en donde la primera línea se establecen los atributos, C4.5 archivos codificados y agrupados al igual que un fichero.

Figura 4. *Carga de archivo en WEKA*



La muestra de datos presenta la correspondencia de estudiantes con necesidades educativas especiales en la Unidad Educativa Plan Internacional. Desde este aspecto la información recopilada fue el periodo académico dentro del proceso educativo correspondiente, la cantidad de estudiantes en dichos periodos matriculados por paralelo, balance de genero (femenino y masculino) de los estudiantes con NNE, promedio de calificación en las Asignaturas de Matemática y Lengua y Literatura.

Tabla 1: *Atributos con sus respectivos valores posibles dentro de la investigación*

Atributos	**Valores posibles**
Periodo académico	**2010 – 2022 (NM)**
Cantidad de estudiantes con NNE (Histórica)	**575 (CC)**
Género Femenino y Masculino	**F= 200; M=375 (CC)**
Promedio general de asignatura de Matemática	**P= 9≅2 (CC)**
Promedio general de asignatura de Lengua y Literatura	**P=9≅2 (CC)**

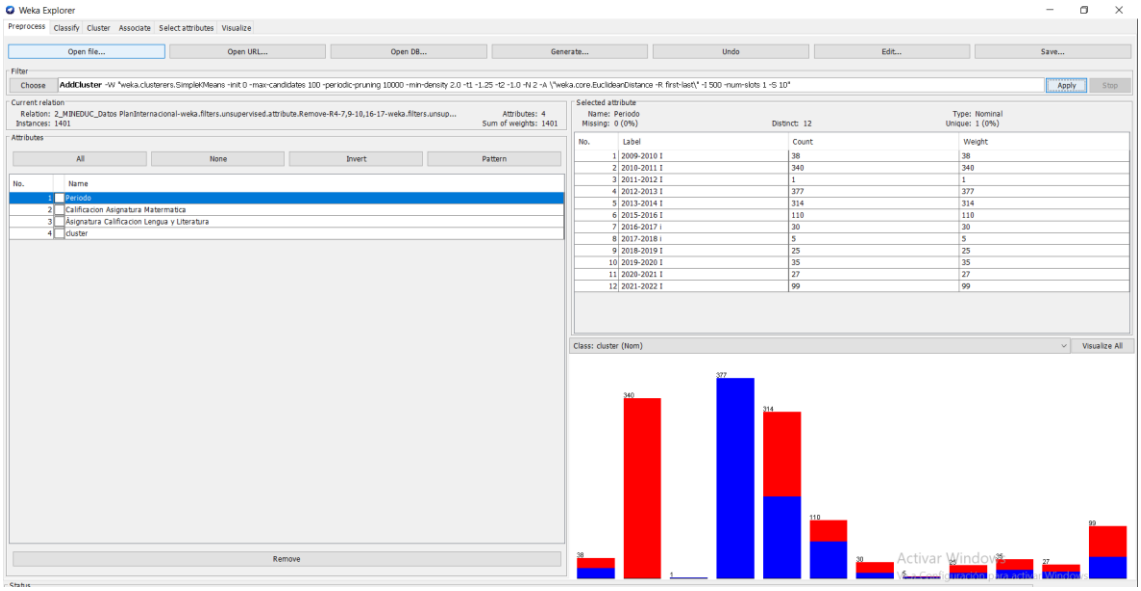
Fuente: Construcción propia

Dentro del proceso de valoración se indica el análisis de los datos obtenidos en WEKA de forma general y depura de los datos relevantes que ostentan la dirección de la investigación.

La interpretación selecciona los siguientes datos:

- ✓ Periodo académico
- ✓ Total, de estudiantes
- ✓ Total, de estudiantes con NNE
- ✓ Calificación de asignatura de Matemática
- ✓ Calificación de asignatura de Lengua y Literatura

Figura 5. *Modelo filtrado del procesamiento y análisis de datos en WEKA*

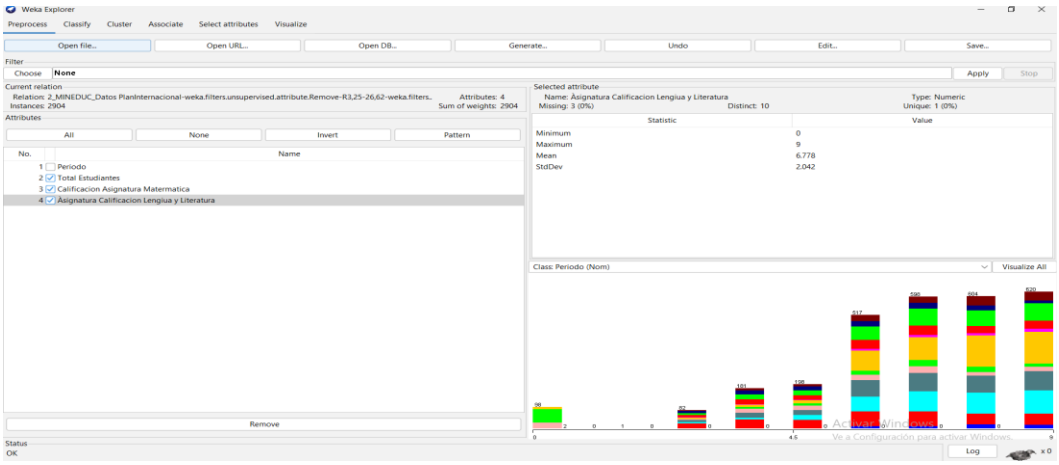


Modelo J48

La implementación del algoritmo J48 de WEKA está basado en el árbol de decisión C4.5. Los árboles de decisión son parte de la taxonomía supervisada. Es decir, contiene variables dependiente o denominadas clases, basado en objetivos de clasificación para determinar el valor de esa o distintas clases. El proceso de creación de árboles comienza con el nodo raíz, conectados al nodo de las muestras y casos de entrenamiento. El primer paso es elegir la variable o el atributo que divide las muestras de entrenamiento originales (nodos raíz) de modo que haya una variación mínima de clase dentro de los subconjuntos generados. Por tanto, una vez determinada la variable que da mayor uniformidad con respecto a la clase de nodos hijos, se vuelve a realizar el análisis para cada nodo hijo. Si todos los nodos hoja contuvieran instancias de la misma clase, este proceso estaría acotado dentro de ciertos límites, pero este valor extremo se alcanzaría si se implementaran los métodos de poda previa y posterior del árbol.

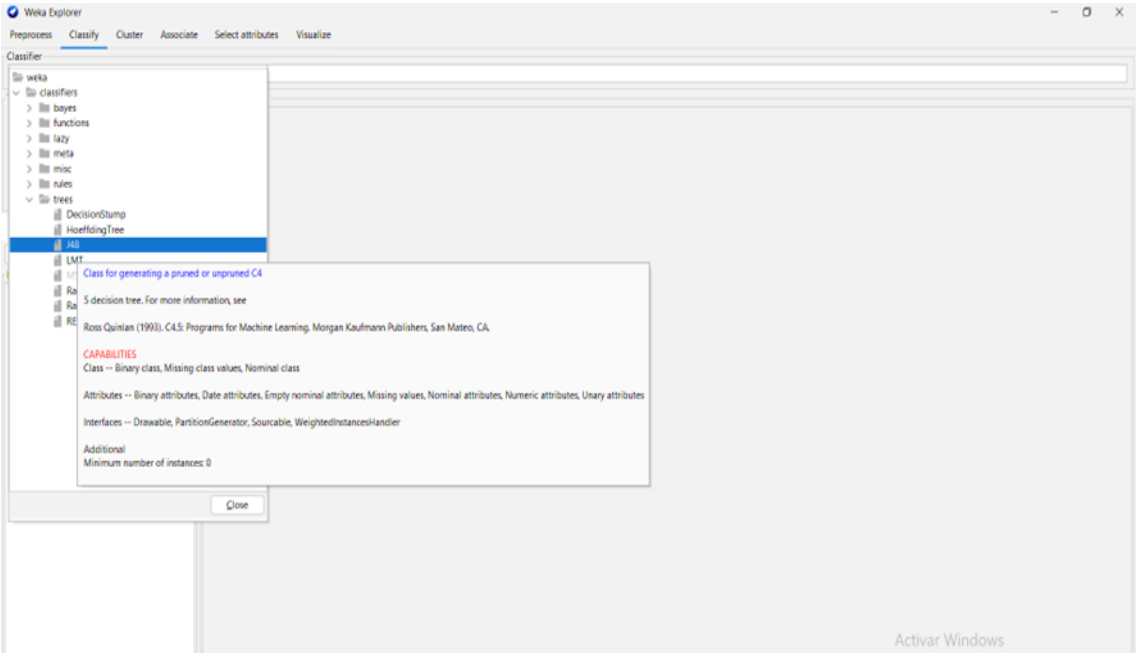
Se describen las instancias, atributos y clases a procesar:

Figura 6. *Clasificación del procesamiento de los datos*



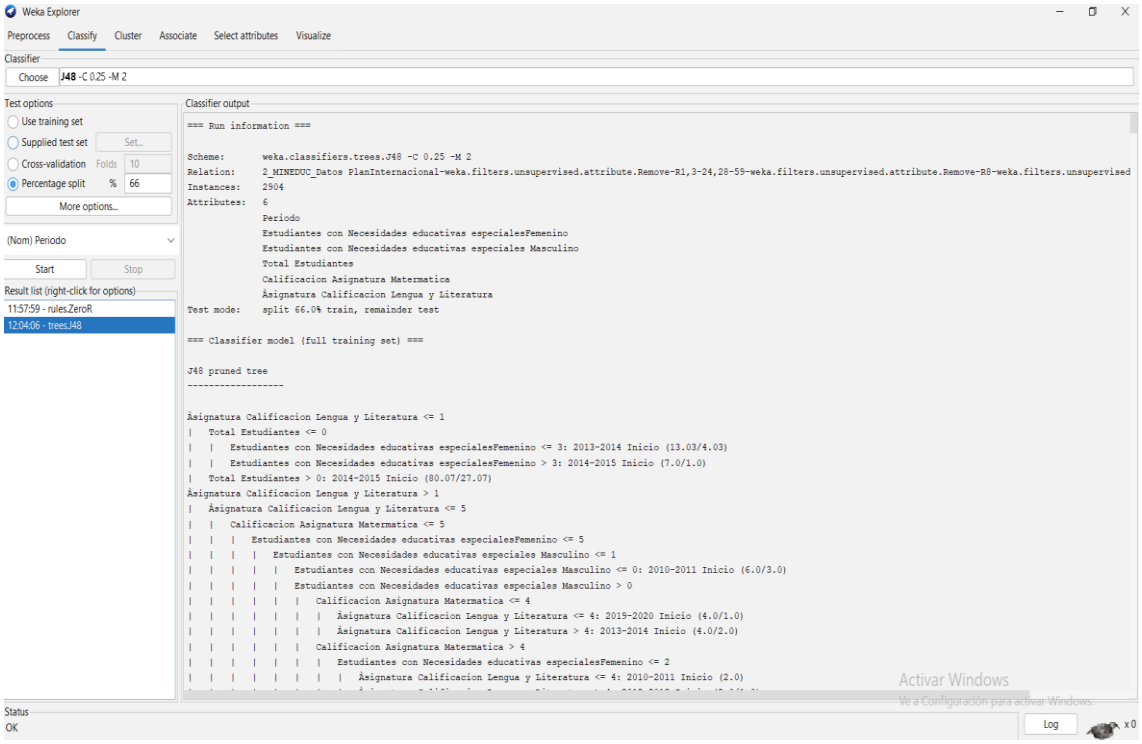
Proceso de clasificación en WEKA después de la clasificación de clases y atributos.

Figura 7. *Clasificación de datos en WEKA por el algoritmo J48*



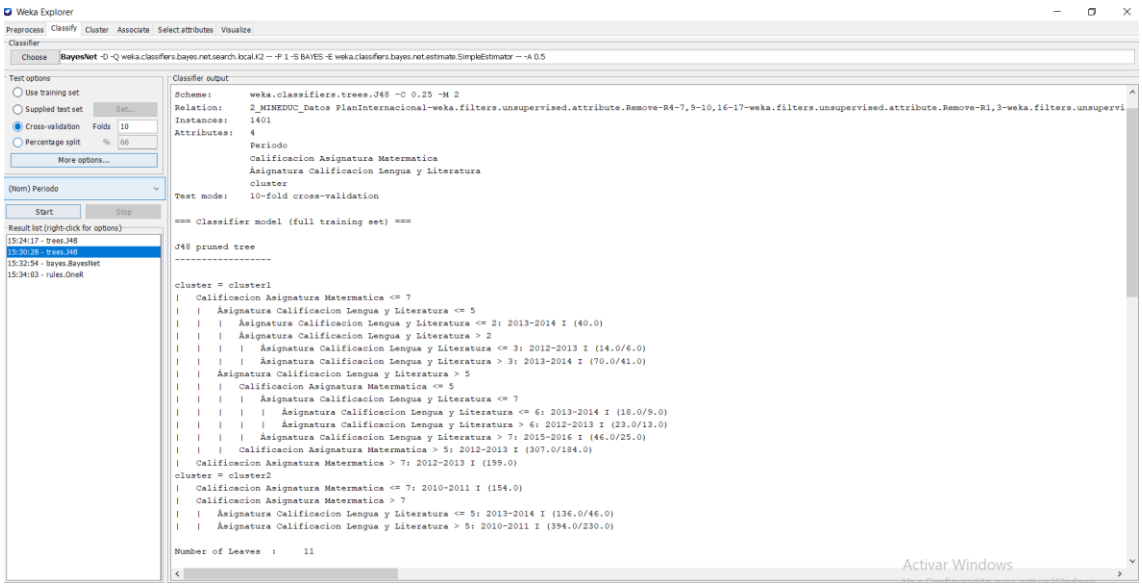
Para el proceso de clasificación en WEKA es importante clasificar el algoritmo a utilizar.

Figura 8. *Composición del algoritmo J48*



Bajo este parámetro se muestran los atributos (6), el algoritmo utilizado y los datos del árbol previamente no grafica.

Figura 9. *Aplicativo de datos en WEKA por el algoritmo J48*



De acuerdo con la clase a interpretar se genera un árbol de decisión de los datos obtenidos para su verificación y análisis, en este caso el análisis de las asignaturas de matemática y lengua y literatura.

Para ello es determinante el identificar la raíz, los nodos y las hojas.

Figura 10. *Modelo de árbol de decisión de los datos en WEKA (Datos Generales)*

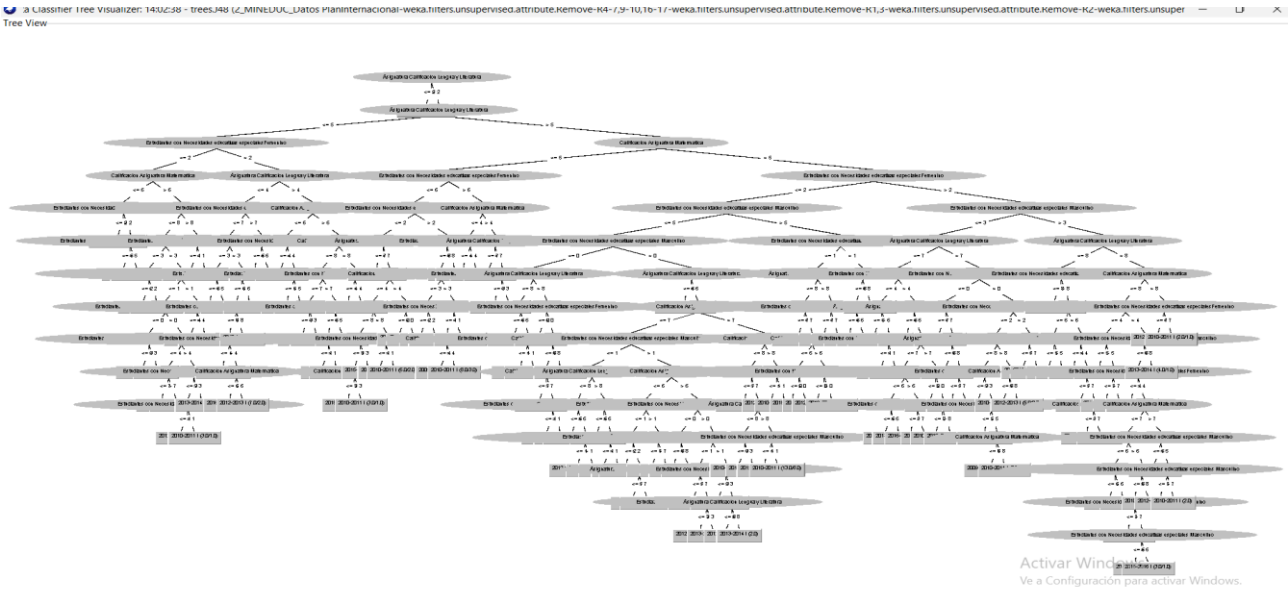
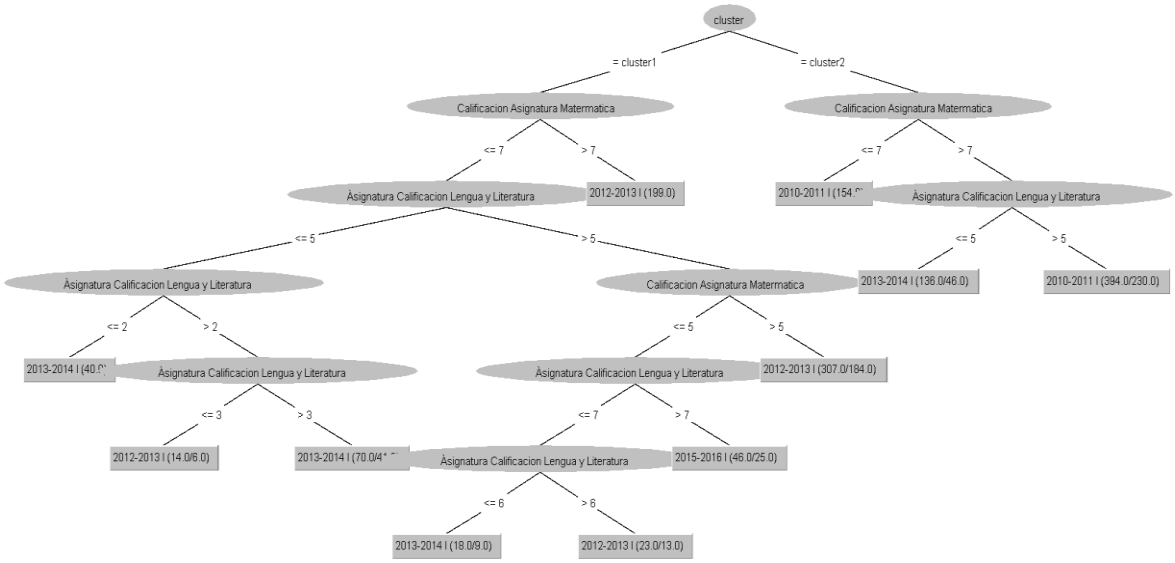


Figura 11. *Modelo de árbol de decisión de los datos en WEKA (Datos clasificados)*



En la figura 11 se presenta los datos valorados mediante el árbol de decisión de los atributos basado en las asignaturas de Lengua y Literatura y Matemática, dando como respuesta un intervalo de mejoramiento del rendimiento de los estudiantes de acuerdo con el periodo académico, en ella podemos especificar que en el periodo académico el balance de mayor calificación es de 8 en la asignatura de Lengua y Literatura y matemática con calificación <2, en este caso 833 de instancias correctamente clasificadas e instancias clasificadas incorrectamente 568 asentado en la suma total de 1401 instancias dado por los datos.

Por otro la se muestra, la comparativa y datos clasificados del algoritmo BayesNet:

Figura 12. *Modelo de datos clasificados en el algoritmo BayesNet*

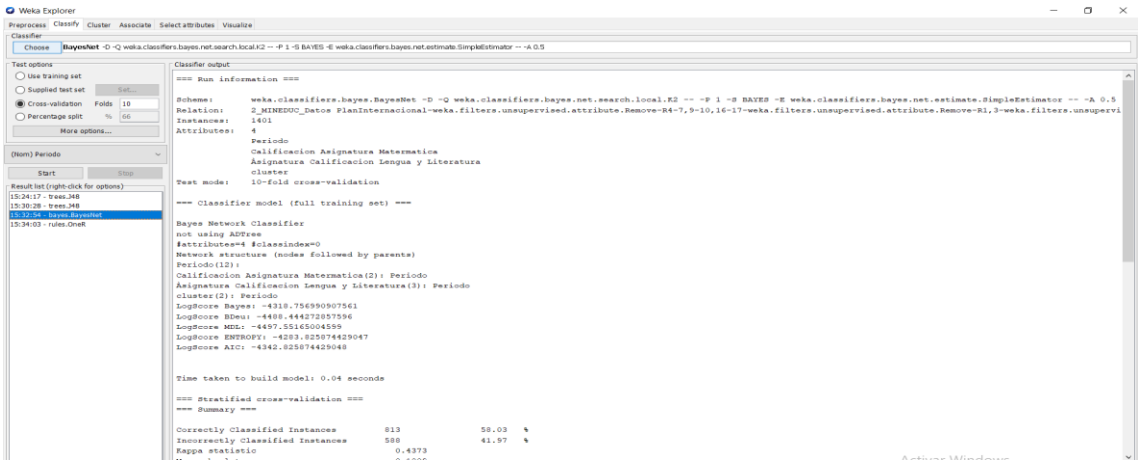
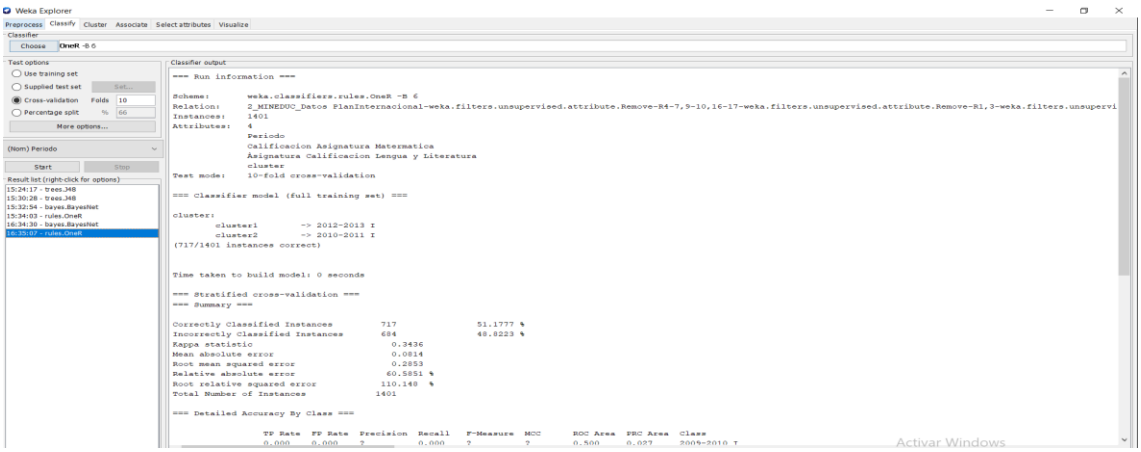


Figura 13. *Modelo de datos clasificados en el algoritmo OneR -B 6*



Evaluaciones de Resultados de los Algoritmos de Clasificación en WEKA

Los resultados de las herramientas se presentan en forma de esquema y gráficamente de acuerdo con la utilización de cada uno de los algoritmos:

Tabla 2: *Algoritmos clasificadores de WEKA*

Algoritmos	**Asignaturas**	**ICC**
J48 (C4.5)	Lengua y literatura	59.4575 %
	Matemática	
BayesNet	Lengua y literatura	58.03 %
	Matemática	
OneR	Lengua y literatura	51.1777 %
	Matemática	

Fuente: Construcción propia

La Tabla 2 representa la evaluación de resultados obtenidos a partir de tres algoritmos de WEKA (J48, BayesNet y OneR). Se aprecia en primera estancia los algoritmos utilizados, los porcentajes de las instancias clasificadas correctamente (ICC). Se observa que en el algoritmo J48 obtuvo un porcentaje de ICC (59.4575 %) denotando el mayor porcentaje de instancias correctamente clasificadas, para BayesNet (58.03 %) y para OneR (51.1777 %) por ende, se denota de acuerdo con los datos valores en gran medida a la robustez de los algoritmos, sin embargo, es importar recalcar una desvarianza entre el algoritmo seleccionado (J48) con mayor precisión hacia el algoritmo BayesNet como un algoritmo robusto.

Comparativa de aplicaciones para la predicción de datos

Tabla 3: *Aplicaciones para la predicción*

Plataformas	Precisión	Utilización
WEKA	51%	43%
Power BI (PI)	61%	82%
TABLEAU PUBLIC	49%	49%

Fuente: Construcción propia

Se han realizado pruebas tomando como ejes WEKA, Power BI (PI) y Tableau Public las cuales han aportado a la investigación como entes de comparativos para la predicción de los datos.

Desde ello se destaca, resultados en la visualización de los datos considerando a Weka con un porcentaje del 51% de precisión de los datos en base a los algoritmos utilizados, Powe BI con un 61% de acuerdo con sus múltiples funciones para procesar datos y así mismo para su visualización y a Tableau Public 49% por su procesamiento de datos y servicios en la nube.

Discusión

Basado en el estudio (Mhetre & Nagar, 2017) muestra que la minería de datos es considerado más útil en el campo educativo, puesto que su postura presenta aplicabilidad en diversos factores nuevos para mejorar el desempeño del estudiante y los resultados se pueden obtener basado en datos sabios anuales. Además, más minería de datos técnicas como k-means, algoritmos de agrupamiento y otros algoritmos de clasificación se pueden aplicar en estos datos, en base a ello la presente investigación determino la utilidad de los procesos y técnicas de minería de datos arrojan que de forma relativa que la usabilidad WEKA y sus algoritmos ayudan a la generación de predicciones de procesos sin embargo,

Al aplicar el algoritmo de J48 se pueden utilizar las tres herramientas de estudio, así como lo menciona (Fernández & Lujan-Mora, 2017), en donde según su investigación basado en procesos educativos con WEKA, se observa datos distintos de acuerdo al criterio de cada uno, así mismo presentan datos similares para el centroide de clústeres, se muestra además que Weka muestra diferentes valores que están dentro de un rango aceptable.

Este estudio empleó la validación cruzada de los atributos según el periodo académico y las asignaturas seleccionadas (Matemática y Lengua y literatura) en los procesos académicos de estudiantes con NNE para aumentar la diferencia en la evaluación y obtener resultados más confiables en las métricas de los modelos. Además, produjo una menor cantidad de atributos y se emplearon más métricas de evaluación para caracterizar de mejor manera los modelos de predicción. Es importante destacar que a medida que aumenta la instancias clasificadas correctamente mayor es la exactitud de las predicciones por ende el modelo se vuelve más útil en entornos reales para predecir el rendimiento académico por periodo académico destacando así el algoritmo J48 como un algoritmo robusto.

En última instancia, es relevante señalar que las características utilizadas en este estudio pueden ser aplicadas en la construcción de modelos predictivos del rendimiento académico en otros cursos. Es decir, la metodología desarrollada puede ser replicada para diferentes cursos, y los atributos utilizados pueden ser recopilados como datos históricos con relevancia para la implementación de estrategias y adaptaciones pedagógicas destinadas a las asignaturas del tronco común.

Esto muestra la validación de los algoritmos utilizados y la manera muy similar en términos de precisión cuando se aplica un determinado algoritmo de clasificación los cuales permiten generar un modelo. Basado en la literatura es importante considerar que

WEKA presenta el mayor número de algoritmos implementados de forma nativa, seguido por RapidMiner y finalmente Knime. Estas dos últimas herramientas utilizan un número significativo de algoritmos importados de WEKA en sus base de datos.

8. CONCLUSIONES

El uso de técnicas de minería de datos direccionada al ámbito educativo está en continua evolución de acuerdo con los requerimientos. Desde este aspecto la obtención de datos que permitan validad procesos y retentivas del conocimiento de los establecimientos educativos es relevante, queda mucho por mejorar en cuanto al análisis y aproximaciones al estudio de prácticas de minería de datos sobre todo en aplicaciones prácticas y las tareas de predicción y descripción que permitan tomar decisiones de los procesos que se llevan dentro del régimen educativo, por ello el énfasis de las asignaturas del tronco común como matemática y lengua y literatura son los ejes principales en la valoración del rendimiento académico de los estudiantes con necesidades educativas especiales por medio de la valoración de forma cuantitativa y cualitativa y los ejes a fortalecer en estas dos áreas.

En consecuencia, a la investigación se recalca que la minería de datos es un proceso de extracción de información útil y significativa de grandes conjuntos de datos, basado a ello la importancia de la extracción de datos ayuda a conseguir una forma predictiva de análisis del aprendizaje automático para descubrir patrones y tendencias en los datos, el uso de los clasificadores de árboles de decisión de forma simple, sin embargo en la minería de datos existen diversos algoritmos y aplicaciones robustas que potencian predicciones de datos, así mismo en el ámbito de WEKA tales como NaiveBayes, OneR los cuales sirven para mejorar o determinar predicciones de forma precisa.

A través de WEKA, se ha realizado el proceso de ETL, es decir el análisis y aplicación de forma estructurada, practica y constructiva de los métodos que se encuentran implementados. Con el filtrado de datos se encontró en instancias gran masa de datos históricos obtenidos y se han creado un modelo de minería de datos de muestra, considerando los dos atributos a asemejar con respecto a las asignaturas en donde un

mayor déficit se presenta en el área de matemática, por lo que la falta de potencialización de dichos conocimientos en estudiantes con necesidades educativas especiales representa un desbalance en la adaptación de contenidos.

9. RECOMENDACIONES

La recopilación exacta de datos de relevancia, sobre todo registros sobre la aplicación de factores cualitativos que permitan gestionar e identificar una predicción mayormente completa y asertiva, además que de acuerdo con esto se logre en la Institución educativa la mejora en la potenciación de habilidades del área de Matemática y Lengua y Literatura mediante adaptaciones pedagógicas adecuadas, propias para cada grado y necesidades educativas de los estudiantes.

La comparación de los datos obtenidos y procesados mediante otros algoritmos y plataformas que fomenten criterios evaluativos y predictivos pues es importante recalcar que, aunque WEKA es una popular plataforma de software de aprendizaje automático y minería de datos que ofrece una amplia gama de algoritmos para realizar tareas de análisis y modelado de dato existen en la actualidad otras plataformas como el caso de Power BI, Tableau Public que de forma sencilla ofrecen predicción y visualización de datos para la toma de decisiones.

Continuar con los procesos evaluativos que fomenten la interpretación de datos anual propio, que permitan conocer desfaces o avances en los procesos educativos de los estudiantes con NNE y a su vez identificar mejoras para su adaptación educativa de acuerdo que potencialicen su proceso de enseñanza-aprendizaje por parte del cuerpo docente.

10.REFERENCIAS BIBLIOGRÁFICAS

Barcia, M. J. J., & Carvajal, Z. B. T. (2015). El proceso de enseñanza aprendizaje en la educación superior. *Revista Electrónica Formación y Calidad Educativa (REFCalE)*, *3*, 139–154.

Bravo, L. G., & Cáceres, M. M. (2016). El proceso de enseñanza-aprendizaje desde una perspectiva comunicativa. *Revista Iberoamericana de Educación*, *1*(7).

Castillo, Y. E., Laborí De La Nuez, B., & Sifonte, R. S. (2020). *Extensión de índices de validación de grupo en la herramienta WEKA para la evaluación de algoritmos de agrupamiento Extension of group validation indices in the WEKA tool for the evaluation of clustering algorithms*. *13*(9), 179–187. http://publicaciones.uci.cu

Diaz, L. B. N. (2021). *Predicción del rendimiento académico utilizando la técnica de árboles de decisión en los programas de maestría de educación en la Escuela de Posgrado de la Universidad Nacional José Faustino Sánchez Carrión* [Universidad Nacional José Faustino Sánchez Carrión]. http://repositorio.unjfsc.edu.pe/bitstream/handle/20.500.14067/4395/Brenda%20Nicole%20Diaz%20Landa.pdf?sequence=1&isAllowed=y

Díaz Landa, B., Meleán Romero, R., & Marín Rodríguez, W. (2021). Rendimiento académico de estudiantes en Educación Superior: predicciones de factores influyentes a partir de árboles de decisión. *Telos Revista de Estudios Interdisciplinarios En Ciencias Sociales*, *23*(3), 616–639. https://doi.org/10.36390/telos233.08

Espinoza Mina, M. A. (2018). Weka, áreas de aplicación y sus algoritmos: una revisión sistemática de literatura. *REVISTA CIENTÍFICA ECOCIENCIA*, *5*, 1–26. https://doi.org/10.21855/ecociencia.50.153

Fernandez, D. B., & Lujan-Mora, S. (2017). Comparison of applications for educational data mining in Engineering Education. *2017 IEEE World Engineering Education Conference (EDUNINE)*, 81–85. https://doi.org/10.1109/EDUNINE.2017.7918187

Jiménez Toledo, J. A., & Timarán Pereira, S. R. (2015). Caracterización de la deserción estudiantil en educación superior con minería de datos. *Revista Tecnológica - ESPOL*, *28*(5).

Markov, Z., & Russell, I. (2006). An introduction to the WEKA data mining system. *Proceedings of the 11th Annual SIGCSE Conference on Innovation and Technology in Computer Science Education*, 367–368. https://doi.org/10.1145/1140124.1140127

Meena, G., & Choudhary, R. R. (2017). A review paper on IDS classification using KDD 99 and NSL KDD dataset in WEKA. *2017 International Conference on Computer, Communications and Electronics (Comptelix)*, 553–558. https://doi.org/10.1109/COMPTELIX.2017.8004032

Mhetre, V., & Nagar, M. (2017). Classification based data mining algorithms to predict slow, average and fast learners in educational system using WEKA. *2017 International Conference on Computing Methodologies and Communication (ICCMC)*, 475–479. https://doi.org/10.1109/ICCMC.2017.8282735

Miranda, M. A., & Guzmán, J. (2017). Análisis de la Deserción de Estudiantes Universitarios usando Técnicas de Minería de Datos. *Formación Universitaria*, *10*(3), 61–68. https://doi.org/10.4067/S0718-50062017000300007

Navas, M. F. (2016). *Introducción a la minería de datos con Weka: Aplicación a un problema económico*. Universidad de Jaén.

Ospina-Mateus, H., & Quintana Jiménez, L. A. (2019). Predicción de accidentes viales en Cartagena, Colombia, con árboles de decisión y reglas de asociación. *Economía & Región*, *13*(2), 83–115. https://doi.org/10.32397/er.vol13.n2.3

Quintana Bajaña, V. E., & Yagual Tomala, S. A. (2017). *Propuesta de aplicación predictiva de aprobación de una asignatura con flujo previo a través de algoritmos basados en software WEKA para estudiantes del último semestre de la Carrera de Ingeniería en Sistemas Computacionales de la Universidad de Guayaquil*. Universidad de Ecuador.

Rodríguez Rodríguez, A., Rodríguez González, A. del C., Pino Tarragó, J. C., & Domínguez Gálvez, D. L. (2021). Implementación de algoritmos de Inteligencia Artificial en la predicción de nuevos conocimientos mediante enseñanza constructivista. *Serie Científica de La Universidad de Las Ciencias Informáticas*, *14(3)*, 131–141.

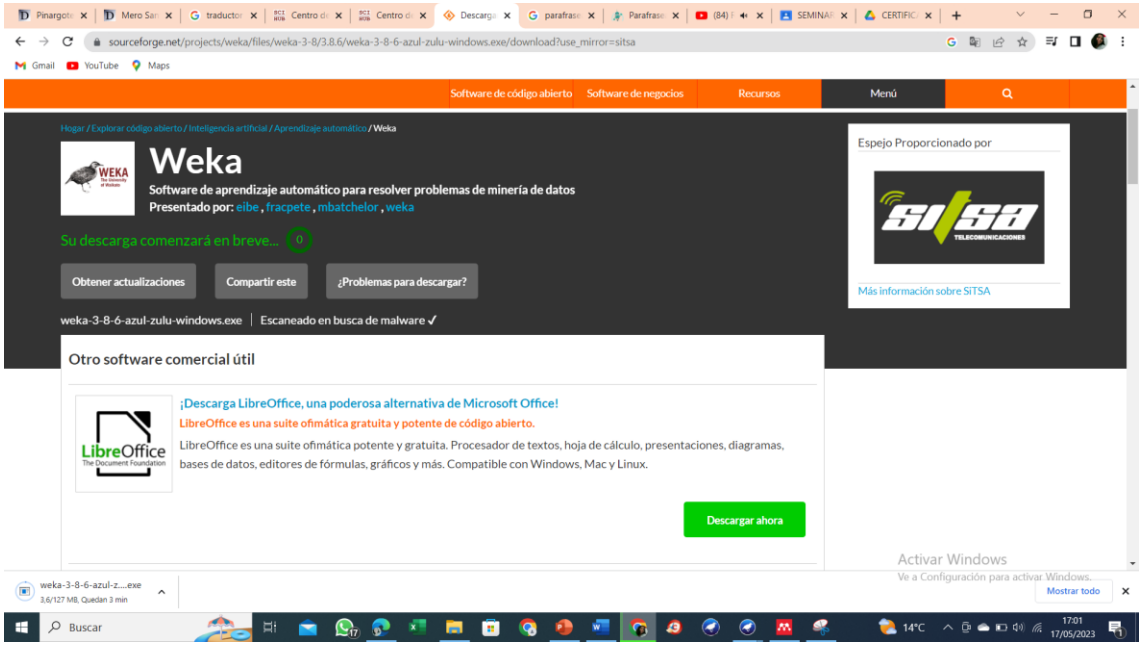
Rojas Salgado, M. E. (2021). Los retos de una educación virtual para estudiantes con necesidades educativas especiales. *HAMUT'AY*, *8*(1), 9. https://doi.org/10.21503/hamu.v8i1.2232

Santiesteban, S. I., Barba, A. J. V., & Álvarez, D. F. (2017). Inclusión de estudiantes con necesidades educativas especiales en la Universidad Técnica del Norte del Ecuador. *Revista Universidad y Sociedad*, *9*(3), 162–167.

Timarán Pereira, R., Hidalgo Troya, A., & Caicedo Zambrano, J. (2020). Factores asociados al desempeño académico en Lectura Crítica en las pruebas Saber 11° con árboles de decisión. *Investigación e Innovación En Ingenierías*, *8*(3), 29–37. https://doi.org/10.17081/invinno.8.3.4701

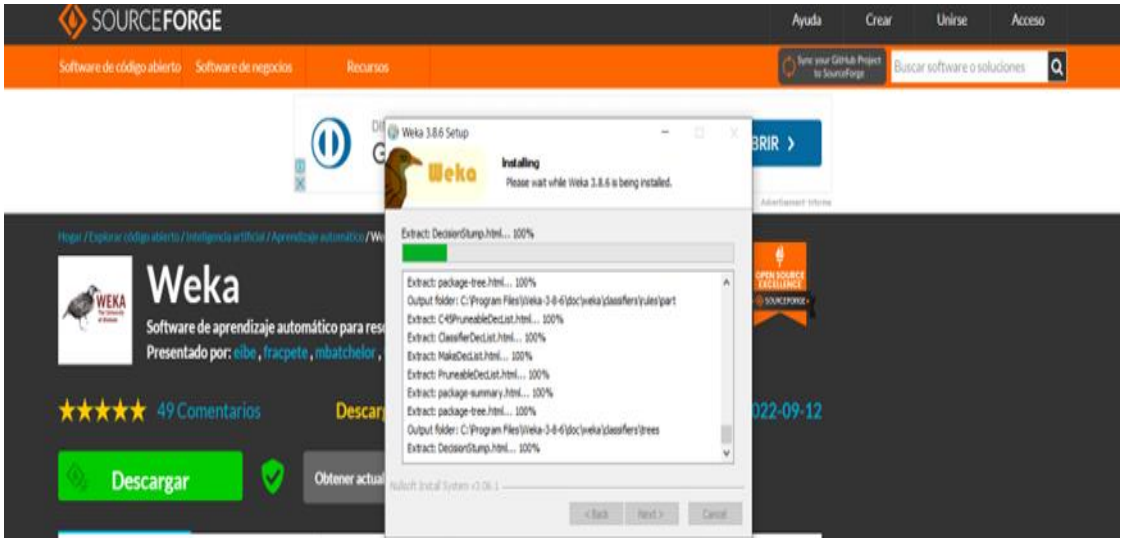
Yenchong, M. W. E., & Barcia, B. M. F. (2020). Formación integral para estudiantes con necesidades educativas especiales Formación integral para estudiantes con necesidades educativas especiales Comprehensive training for students with Special Educational Needs Treinamento abrangente para alunos com necessidades educacionais especiais. *Revista Científica Dominio de Las Ciencias*, *6*, 361–377. https://doi.org/10.23857/dc.v6i3.1222

11.ANEXOS

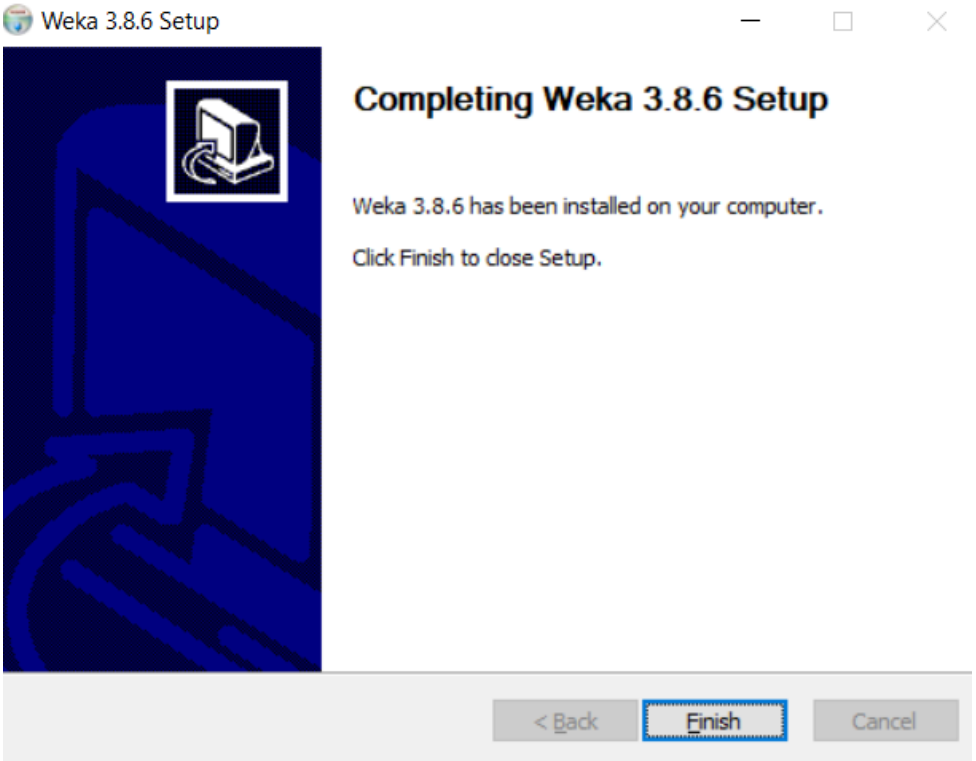


Descarga de WEKA

Instalación de Weka

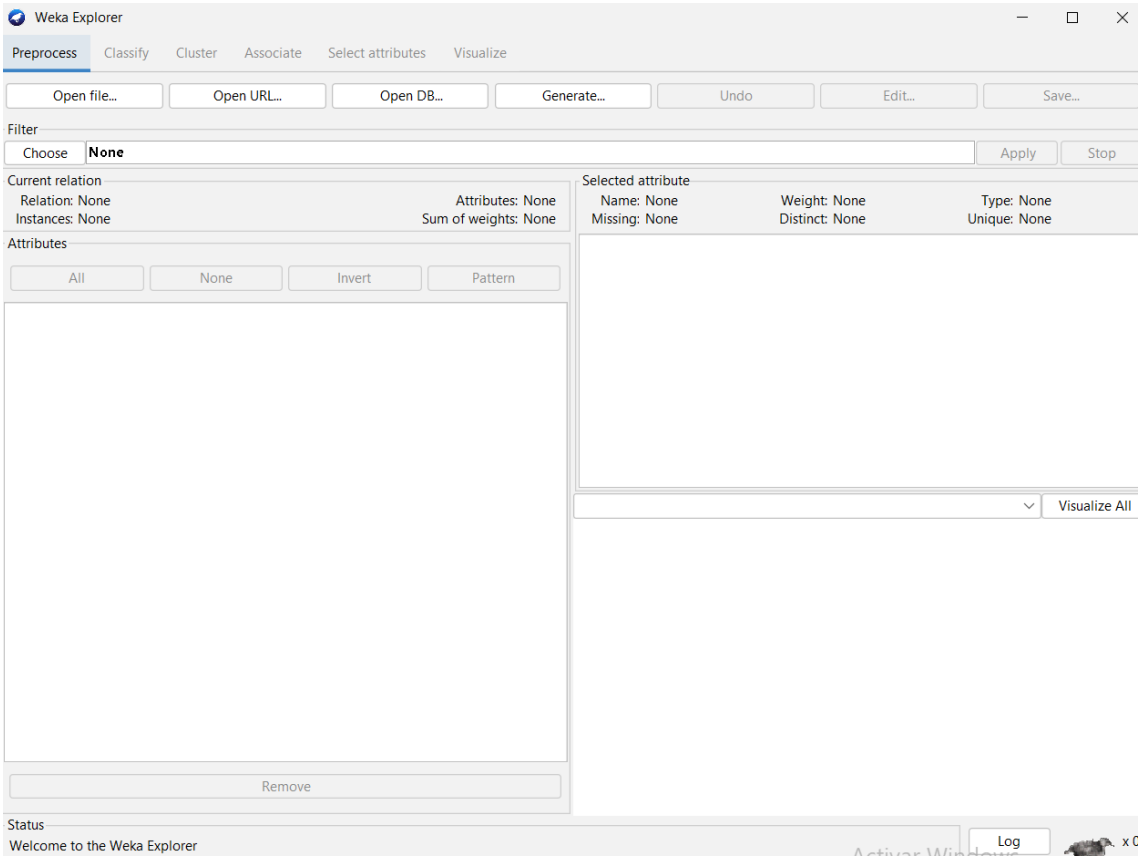


Instalación de Weka

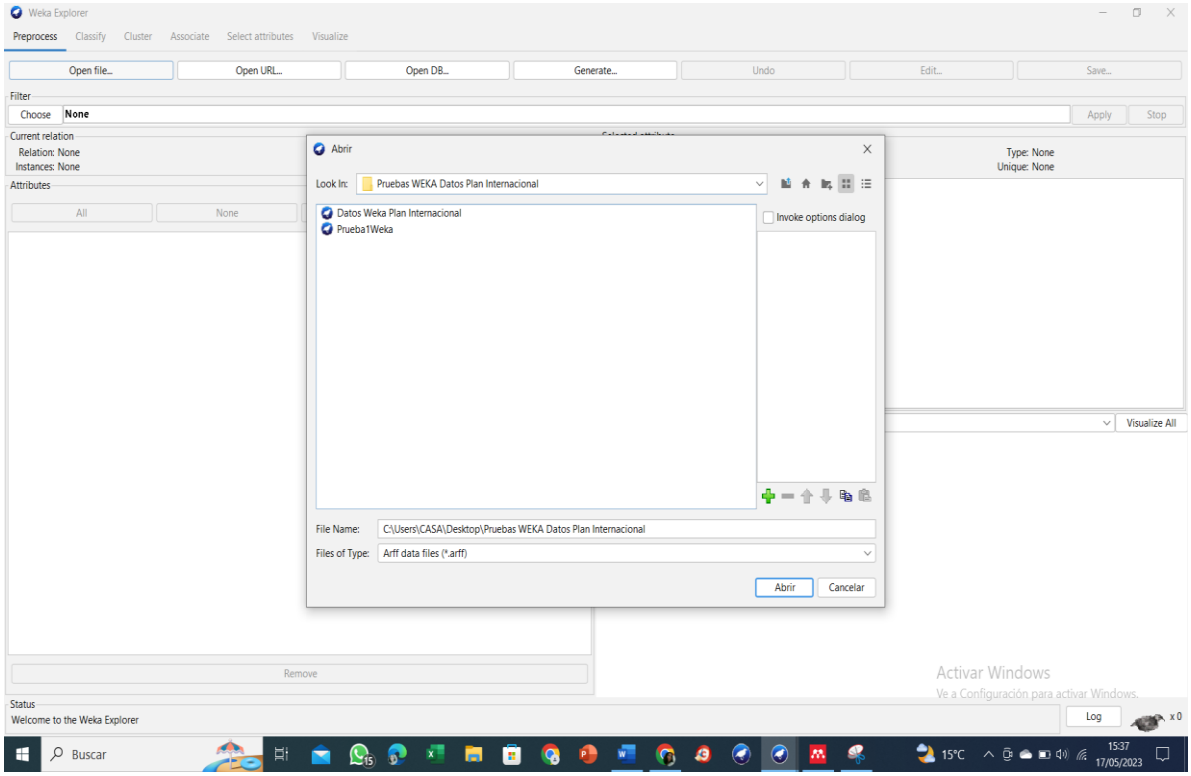


Instalación de Weka

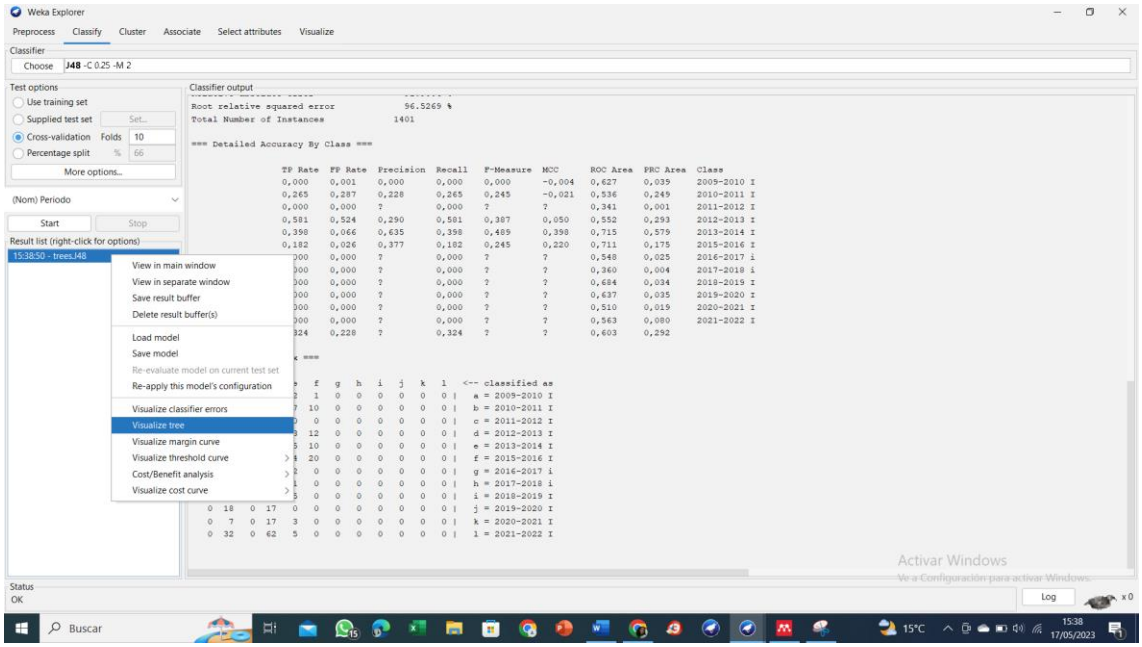
Ambiente WEKA



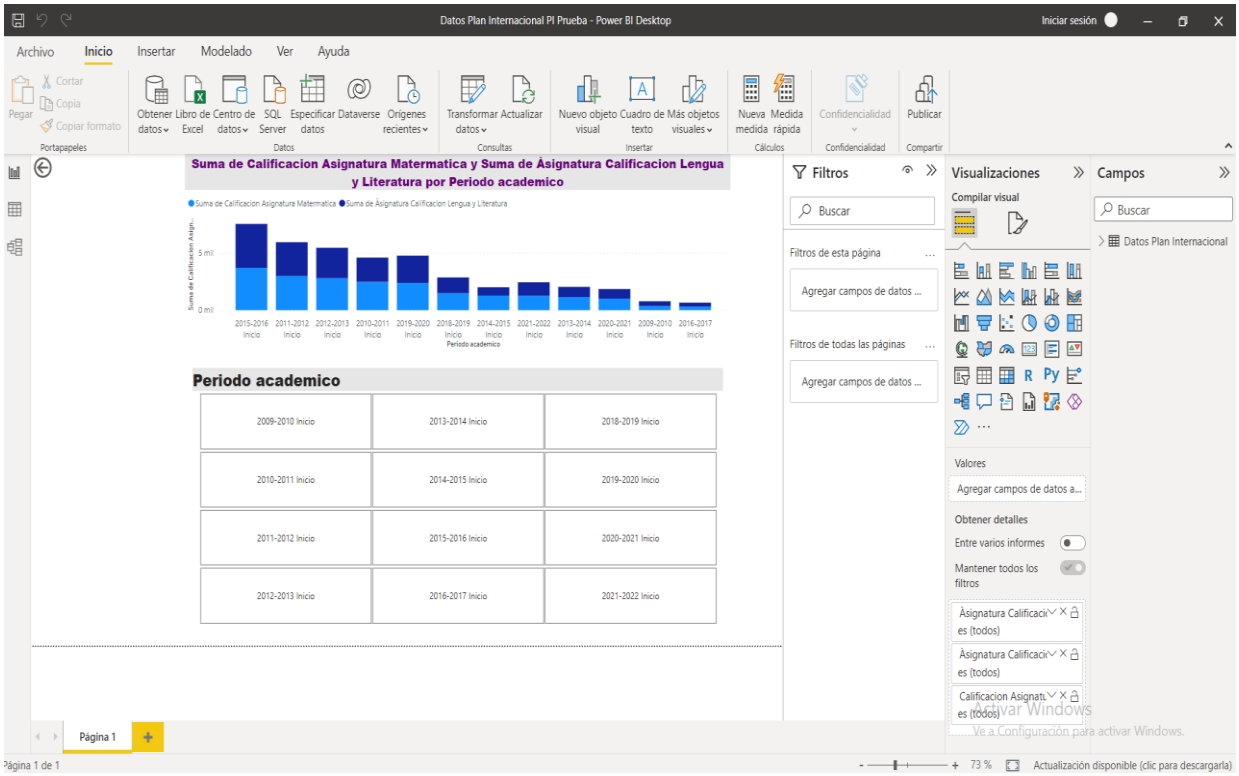
Interfaz WEKA



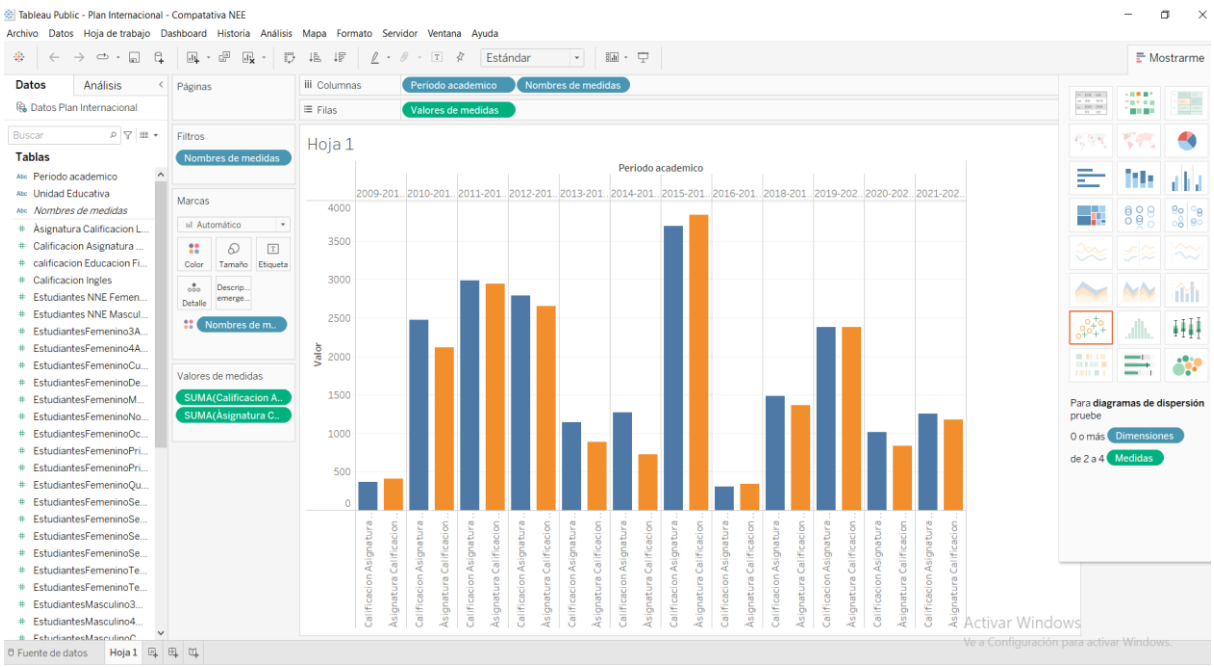
Carga de datos en WEKA



Visualización de árbol de decisiones en WEKA

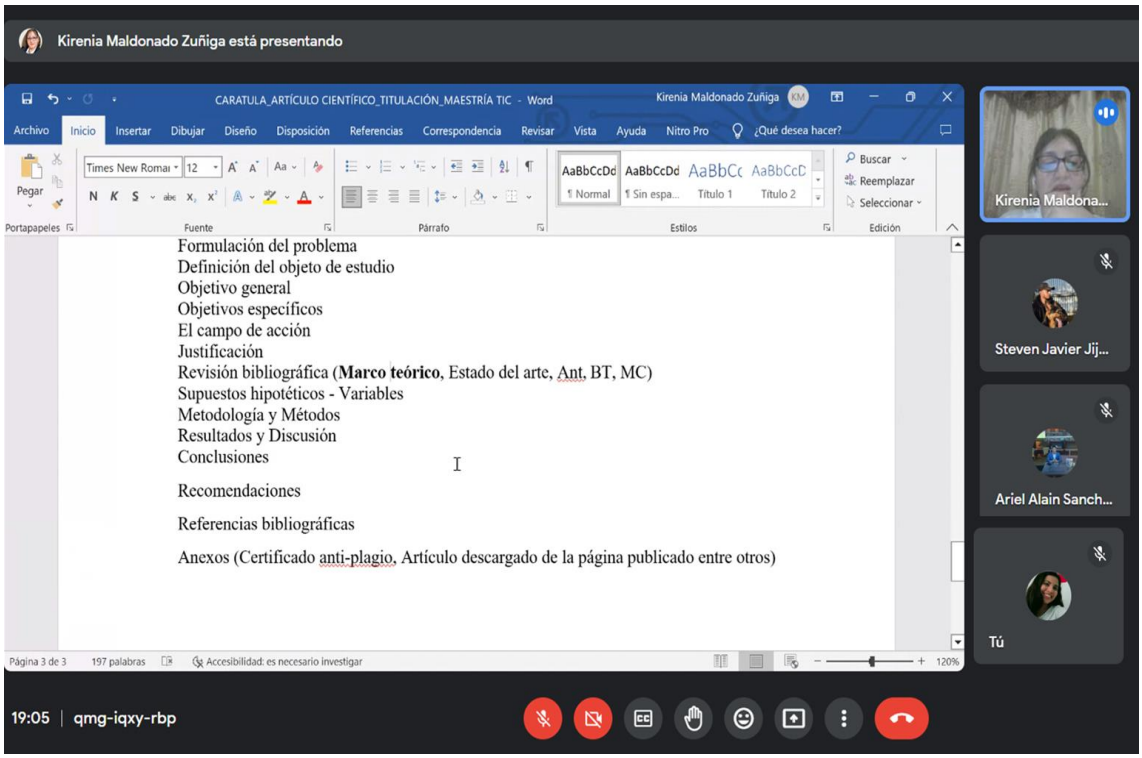


Visualización de datos en Power BI



Visualización de datos en Tableau Public

Reunión con comisión pedagógica de la información solicitada a la UE



Reunión con la tutora designada en el trabajo de investigación

CERTIFICADO DE ANÁLISIS
magister

Tesina Hilda Quijije Quiroz Maestria (1)

4% Similitudes

< 1% Texto entre comillas
0% similitudes entre comillas
2% Idioma no reconocido

Nombre del documento: Tesina Hilda Quijije Quiroz Maestria (1).docx
ID del documento: 983bcbbad28cf31963cd98115dd028446bd07008
Tamaño del documento original: 5,75 MB

Depositante: Maldonado Kirenia
Fecha de depósito: 13/7/2023
Tipo de carga: interface
fecha de fin de análisis: 13/7/2023

Número de palabras: 10.113
Número de caracteres: 67.443

Ubicación de las similitudes en el documento:

Fuentes principales detectadas

Nº	Descripciones	Similitudes	Ubicaciones	Datos adicionales
1	Documento de otro usuario #fadf83 El documento proviene de otro grupo 2 fuentes similares	< 1%		Palabras idénticas : < 1% (25 palabras)
2	Documento de otro usuario #085553 El documento proviene de otro grupo 1 fuente similar	< 1%		Palabras idénticas : < 1% (24 palabras)
3	dialnet.unirioja.es | Ricardo Timarán Pereira - Dialnet https://dialnet.unirioja.es/servlet/autor?codigo=3016196 1 fuente similar	< 1%		Palabras idénticas : < 1% (22 palabras)

Fuentes con similitudes fortuitas

Nº	Descripciones	Similitudes	Ubicaciones	Datos adicionales
1	Freddy Villafuerte Tesis_OK.docx | Freddy Villafuerte Tesis_OK #ac00e4 El documento proviene de mi grupo	< 1%		Palabras idénticas : < 1% (38 palabras)
2	editorialalema.org | Aplicación de herramientas tecnológicas para la evaluación del ... http://editorialalema.org/index.php/pentaciencias/article/view/301	< 1%		Palabras idénticas : < 1% (40 palabras)
3	www.gob.ec https://www.gob.ec/sites/default/files/regulations/2020-02/Reglamento de Régimen Académico.pdf	< 1%		Palabras idénticas : < 1% (29 palabras)
4	Documento de otro usuario #df91c7 El documento proviene de otro grupo	< 1%		Palabras idénticas : < 1% (32 palabras)
5	Documento de otro usuario #2c99b6 El documento proviene de otro grupo	< 1%		Palabras idénticas : < 1% (17 palabras)

Fuentes ignoradas

Estas fuentes han sido retiradas del cálculo del porcentaje de similitud por el propietario del documento.

Nº	Descripciones	Similitudes	Ubicaciones	Datos adicionales
1	3. PROYECTO DE TITULACIÓN_Maestría_TIC-Javier(1).docx | 3. PROYECTO D... #a87de7 El documento proviene de mi grupo	4%		Palabras idénticas : 4% (381 palabras)
2	repositorio.unjfsc.edu.pe | Predicción del rendimiento académico utilizando la técni... http://repositorio.unjfsc.edu.pe/handle/UNJFSC/4395	3%		Palabras idénticas : 3% (292 palabras)
3	repositorio.unjfsc.edu.pe | Predicción de la deserción estudiantil utilizando la técnic... https://repositorio.unjfsc.edu.pe/handle/20.500.14067/5576	3%		Palabras idénticas : 3% (289 palabras)
4	TESINA FINAL - RODRÍGUEZ SINISTERRA_15_abril.docx | TESINA FINAL - RO... #0f9a19 El documento proviene de mi grupo	2%		Palabras idénticas : 2% (250 palabras)
5	repositorio.unesum.edu.ec https://repositorio.unesum.edu.ec/bitstream/53000/3559/1/Quijije Quiroz Hilda Brigitte_PDF.pdf	2%		Palabras idénticas : 2% (267 palabras)
6	TESINA FINAL - RODRÍGUEZ SINISTERRA_15_abril.docx | TESINA FINAL - RO... #300402 El documento proviene de mi grupo	2%		Palabras idénticas : 2% (223 palabras)
7	TESINA FINAL - RODRÍGUEZ SINISTERRA_18_abril.docx | TESINA FINAL - RO... #8dddd1 El documento proviene de mi grupo	2%		Palabras idénticas : 2% (210 palabras)
8	repositorio.ug.edu.ec http://repositorio.ug.edu.ec/bitstream/redug/19883/1/B-CISC-PTG.1296 Quintana Bajaña Vanessa Este...	2%		Palabras idénticas : 2% (207 palabras)
9	www.redalyc.org | Rendimiento académico de estudiantes en Educación Superior: p... https://www.redalyc.org/journal/993/99368569008/	2%		Palabras idénticas : 2% (189 palabras)

Nº		Descripciones	Similitudes	Ubicaciones	Datos adicionales
10		**www.redalyc.org** | Rendimiento académico de estudiantes en Educación Superior: p... https://www.redalyc.org/journal/993/99368569008/99368569008.pdf	2%		Palabras idénticas : 2% (187 palabras)
11		**www.redalyc.org** | Rendimiento académico de estudiantes en Educación Superior: p... https://www.redalyc.org/journal/993/99368569008/html/	2%		Palabras idénticas : 2% (181 palabras)
12		**dialnet.unirioja.es** https://dialnet.unirioja.es/descarga/articulo/8133982.pdf	2%		Palabras idénticas : 2% (175 palabras)
13		**dialnet.unirioja.es** https://dialnet.unirioja.es/descarga/articulo/7971394.pdf	1%		Palabras idénticas : 1% (136 palabras)
14		**hdl.handle.net** | Implementación de una Red Neuronal para mejorar el pronóstico d... https://hdl.handle.net/20.500.12672/11163	1%		Palabras idénticas : 1% (150 palabras)
15		**editorialalema.org** | Revista Científica Arbitrada Multidisciplinaria PENTACIENCIAS http://editorialalema.org/index.php/pentaciencias	1%		Palabras idénticas : 1% (133 palabras)
16		**rte.espol.edu.ec** http://rte.espol.edu.ec/index.php/tecnologica/article/download/453/318	1%		Palabras idénticas : 1% (139 palabras)
17		**dialnet.unirioja.es** https://dialnet.unirioja.es/descarga/articulo/8590452.pdf	1%		Palabras idénticas : 1% (125 palabras)
18		**www.itca.edu.sv** https://www.itca.edu.sv/wp-content/themes/elaniin-itca/docs/2011-Diseno-de-un-modelo-basado-en-te...	1%		Palabras idénticas : 1% (147 palabras)
19		**hdl.handle.net** | Construcción de un modelo predictivo de tipo Gradient boosting pa... http://hdl.handle.net/10810/15852	1%		Palabras idénticas : 1% (143 palabras)
20		**hdl.handle.net** | Implementación de un sistema de predicción del rendimiento acad... https://hdl.handle.net/20.500.14168/590	1%		Palabras idénticas : 1% (137 palabras)

« ‹ 1 2 3 4 › »

Fuentes mencionadas (sin similitudes detectadas)

Estas fuentes han sido citadas en el documento sin encontrar similitudes.

1 http://publicaciones.uci.cu

2 https://doi.org/10.1109/EDUNINE.2017.7918187

3 https://doi.org/10.1145/1140124.1140127

4 https://doi.org/10.1109/COMPTELIX.2017.8004032

5 https://doi.org/10.1109/ICCMC.2017.8282735

TÉCNICA DE MINERÍA DE DATOS PARA PROCESOS EDUCATIVOS EN ESTUDIANTES CON NECESIDADES EDUCATIVAS ESPECIALES BASADO EN UN MODELO PREDICTIVO

DATA MINING TECHNIQUE FOR EDUCATIONAL PROCESSES IN STUDENTS WITH SPECIAL EDUCATIONAL NEEDS BASED ON A PREDICTIVE MODEL

Hilda Brigitte Quijije Quiroz [1*]

[1] Ingeniera en Tecnologías de la Información. Estudiante de la Maestría en Tecnologías de la Información y la Comunicación en el Instituto de Posgrado de la Universidad Estatal del Sur de Manabí. Jipijapa, Manabí, Ecuador. Docente de la Unidad Educativa Plan Internacional, Daule, Guayas, Ecuador. ORCID: https://orcid.org/0000-0002-6696-3075. Correo: quijije-hilda2868@unesum.edu.ec

Kirenia Maldonado Zuñiga [2]

[2] Doctorando en Tecnologías de la Información y Comunicación, Universidad Nacional de Piura, Perú. Magister en Ciencias de la Educación, Licenciada en Educación Informática. Docente en la Carrera Tecnologías de la Información. Coordinadora y catedrática de la Maestría en Tecnologías de la Información y la Comunicación en el Instituto de Posgrado de la Universidad Estatal del Sur de Manabí. Jipijapa, Manabí, Ecuador. ORCID: https://orcid.org/0000-0002-3764-5633. Correo: kirenia.maldonado@unesum.edu.ec

*** Autor para correspondencia: quijije-hilda2868@unesum.edu.ec**

Resumen

En el contexto educativo, la minería de datos puede ser utilizada para analizar los datos enfocados en el proceso de enseñanza de los estudiantes y los procesos educativos que ayuden a identificar los factores que influyen desde diversos aspectos a partir de la retención del aprendizaje, rendimiento académico, la asistencia y la participación en actividades extracurriculares. Esto puede permitir a los educadores tomar medidas para mejorar los temas y habilidades en los que los estudiantes con necesidades educativas especiales y adaptar el plan de estudios en consecuencia. Esta investigación tuvo el objetivo de diseñar un modelo predictivo mediante técnicas de minería de datos en procesos educativos aplicados a estudiantes con necesidades educativas especiales en la Unidad Educativa "Plan Internacional" cuyo desarrollo se realizó mediante el uso de WEKA como herramienta para el análisis de datos y modelado predictivo presentando un aporte en la búsqueda de estrategias que permitan mejorar el proceso de enseñanza con los estudiantes con NNE. Se

Esta obra está bajo una licencia *Creative Commons* de tipo (CC-BY-NC-SA).
Grupo Editorial "ALEMA-Pentaciencias" E-mail: alema.pentaciencias@gmail.com

utilizaron métodos de investigación científica para sustentar el proceso de caracterización del estudio basado en un enfoque cuantitativo a nivel descriptivo y de campo. Este estudio es un aporte que se relaciona al proyecto del grupo de investigación de la maestría en Tecnologías de la Información y la Comunicación, titulado "Desarrollo e Innovación en Tecnologías de la Información y las Comunicaciones".

Palabras clave: estrategias; herramientas; procesos educativos

Abstract

In the educational context, data mining can be used to analyze data focused on the teaching process of students and educational processes that help identify the factors that influence from various aspects from learning retention, academic performance, attendance and participation in extracurricular activities. This can allow educators to take steps to improve the subjects and skills in which students with special educational needs and adapt the curriculum accordingly. This research had the objective of designing a predictive model through data mining techniques in educational processes applied to students with special educational needs in the Educational Unit "Plan International" whose development was carried out through the use of WEKA as a tool for data analysis and predictive modeling presenting a contribution in the search for strategies to improve the teaching process with students with NNE. Scientific research methods were used to support the characterization process of the study based on a quantitative approach at the descriptive and field level. This study is a contribution that is related to the project of the research group of the master's degree in Information and Communication Technologies, entitled "Development and Innovation in Information and Communication Technologies".

Keywords: *educational processes; strategies; tool*

Fecha de recibido: 04/05/2023
Fecha de aceptado: 14/06/2023
Fecha de publicado: 10/07/2023

Introducción

Las Tecnologías de la Información y comunicación (TIC) aportan al campo de la educación aspectos innovadores que suponen una mejora cualitativa en las formas de enseñar y aprender con un desarrollo multidisciplinario en el que intervienen varias ciencias, entre ellas la pedagogía y la computación. Su utilización contribuye a que los procesos de enseñanza-aprendizaje sean más colaborativos, interactivos y flexibles a partir de la aplicabilidad que se alcanza en la personalización de la enseñanza.

Hasta el actual momento el uso de las tecnologías enfocadas en la minería de datos representa una alternativa eficaz para el análisis de datos y la vigente necesidad de mejorar la calidad educativa presenta múltiples pautas para transformar la visión activa de enseñanza de forma espiral, la forma autónoma de trabajo en grupos seleccionados se efectúa en base a la referencia del efecto mateo y su contextualización muchas veces se basa

Esta obra está bajo una licencia *Creative Commons* de tipo (CC-BY-NC-SA).
Grupo Editorial "ALEMA-Pentaciencias" E-mail: alema.pentaciencias@gmail.com

en las ventajas acumulativas que se presentan de forma indirecta, sin embargo desde la mirada en forma espiral la perspectiva incide en la correcta integración de procesos que conlleven la utilización de nuevas metodologías y la oportunidad de clasificar, comparar y evaluar pautas mediante softwares que determinen un análisis evaluativo a través de técnicas basadas en minería de datos por medio de la extracción de registros en etapas iterativas que permitan obtener conocimiento nuevo.

La aplicación de la minería de datos educacional en diversas investigaciones plantea una solución para la mejora del rendimiento académico, pero el análisis de grandes cantidades de datos y la experimentación de diferentes métodos y técnicas para encontrar un modelo que sea fiable suponen de tiempo y esfuerzo. Existen técnicas de investigación documental que facilitan el análisis de toda la información existente acerca de un tema en común, aplicando técnicas que pueden determinar la concordancia de los resultados y comparar las observaciones de los diferentes autores para llegar a una conclusión acertada (Santiesteban et al., 2017).

Existen varios enfoques de algoritmos de agrupamiento, debido a que, para un mismo conjunto de datos, aplicando diferentes algoritmos de agrupamiento se pueden obtener resultados diferentes. Por lo que surge la necesidad de evaluar las particiones obtenidas y poder determinar la calidad de los resultados alcanzados. Existen medidas que permiten realizar una evaluación de la estructura resultante de cada algoritmo de agrupamiento, obteniendo de manera cuantitativa un valor de calidad de estas. Estas medidas de calidad se conocen bajo el nombre de índices de validación de grupo (Castillo et al., 2020).

Desde los factores la parte más notable se basa en el rendimiento académico de estudiantes con necesidades educativas especiales que puede variar dependiendo de la naturaleza y el grado de la discapacidad, así como de los recursos disponibles para el estudiante. Sin embargo, con los apoyos adecuados, muchos estudiantes con pueden alcanzar un rendimiento académico comparable al de sus compañeros. Por ello, es importante que se les brinden los apoyos y recursos adecuados, incluyendo la adaptación de los materiales de enseñanza y evaluación, el acceso a tecnologías asistivas, el uso de intérpretes y servicios de apoyo emocional. En la investigación se aplicó teorías relacionadas al proyecto de investigación titulado "La enseñanza constructivista sustentada en la Inteligencia Artificial"

La gestión fue guiada por la metodología CRISP - DM que permita desarrollar el proyecto, se hizo uso de la metodología de minería de datos entre las cuales se ostentan las siguientes fases, Compresión del Negocio, Objetivos del Negocio, Comprensión de los Datos, Recolección de Datos iniciales, Preparación de los Datos, Estructura de los Datos, A partir del estudio realizado, se implementó en la presente investigación la metodología de minería de datos CRISP - DM para el análisis y procesamiento de las informaciones. La implementación de algoritmos computacionales sustentados en Inteligencia Artificial permitió la predicción de nuevos conocimientos mediante un enfoque de enseñanza constructivista (Rodríguez Rodríguez et al., 2021).

También es importante que los docentes y el personal de la escuela estén capacitados en cómo trabajar con estudiantes con discapacidad y cómo proporcionar los apoyos y recursos necesarios. La colaboración entre los padres, los docentes y otros miembros del equipo de apoyo también puede ser fundamental para mejorar el rendimiento académico de los estudiantes con diversas necesidades y que estén apegadas a las adaptaciones curriculares según los lineamientos del ministerio de educación.

Esta obra está bajo una licencia *Creative Commons* de tipo (CC-BY-NC-SA).
Grupo Editorial "ALEMA-Pentaciencias" E-mail: alema.pentaciencias@gmail.com

(Rojas Salgado, 2021), menciona que, sin duda alguna, las estrategias didácticas que emprendan los docentes deben sustentarse en la diversidad de características que presentan los estudiantes, especialmente aquellos con NEE y, a su vez, aprovechar los beneficios que brindan las tecnologías para inclusión educativa de las personas en situación de discapacidad. La variabilidad y flexibilidad son aspectos clave de estas, por ello simbolizan un gran potencial para integrarlas, en particular, en el proceso educativo de los estudiantes con NEE. En ese sentido, incorporar en las estrategias didácticas, tanto materiales físicos como virtuales que demanden de tecnología, permitirá abordar esa diversidad de características de la población estudiantil.

Dentro de lo abordado se aplicó el uso de minería de datos para determinar roles respecto a la usabilidad de herramientas tecnológicas con estudiantes con necesidades educativas especiales en la Unidad Educativa "Plan Internacional" perteneciente al Cantón Daule, la clasificación, errores y redundancias presentes en los datos a obtener, así como recodificar algunas variables que permitan contrastar el rendimiento académico de los estudiantes en base a nuevos modelos en la enseñanza de este grupo seleccionado. Esta investigación se relaciona con el proyecto de investigación titulado "Metodología para auditoría automática de peligros y puntos críticos de control aplicando minería de proceso".

Las sociedades mundiales consideran importante la necesidad de enriquecer la formación cultural del hombre, cuya preparación lo coloca a la cabeza del desarrollo del mundo moderno; un hombre civilizado que comprende los problemas de su contexto en su origen y desarrollo que entrar con firmeza en la batalla de las ideas para adoptar una actitud transformadora encaminada a la consecución de los ideales sociales de nuestro tiempo (Barcia & Carvajal, 2015).

En base a ello es preciso mencionar que el grupo seleccionado son estudiantes con necesidades educativas especiales, por lo tanto se referencia el rendimiento de ellos en las asignaturas de matemática y lengua y literatura como conocer el intervalo de ellas, así mismo es factible que en base a ello se busquen alternativas que mejoren y potencialicen el trabajo con este grupo basado en actividades lúdicas en estas áreas como la introducción de herramientas tecnológicas en base a el trabajo con ellos.

Materiales y métodos

WEKA es una herramienta de plataforma de minería de datos para el aprendizaje automático, las primeras pruebas se ejecutaron en la Universidad de Waikato, en Nueva Zelanda en el año 1997. WEKA es un conjunto de Algoritmos de *Machine Learning* y *Data Mining*. Está diseñado bajo un software programado en lenguaje JAVA el cual tiene una interfaz GUI para interactuar con archivos y conjuntos de datos predefinidos con 49 herramientas de preprocesamiento para ellos, contiene herramientas bajo 76 algoritmos de clasificación, 15 atributos evaluadores y 10 algoritmos de búsqueda para la selección de características.

Según la perspectiva de Charalampopoulos & Anagnostopoulos (2011) baso "Un estudio comparable que emplea algoritmos de agrupamiento/clasificación WEKA para la clasificación de páginas web" ostenta que, la clasificación de datos es un problema en Ciencias de la Información pues describe la difícil tarea de asignar un documento a uno o más categorías según su contenido. Muchos algoritmos se han creado con el uso de métodos con sus aplicaciones que varían de publicidad, indexación de bases de datos, motores de búsqueda (como mencionado) y muchos más. La mayoría de los algoritmos utilizados denotan ciertas inconsistencias

Esta obra está bajo una licencia *Creative Commons* de tipo (CC-BY-NC-SA).
Grupo Editorial "ALEMA-Pentaciencias" E-mail: alema.pentaciencias@gmail.com

en la búsqueda del conocimiento científico que pone en riesgo la veracidad de integración y búsqueda del conocimiento mediante estos datos e información.

Dirigido a estos enfoques, este estudio presentó una investigación aplicada, cualitativa y su fundamentación se basó a partir de un enfoque interpretativo, analítico con el entorno y la búsqueda de soluciones mediante la investigación abierta de los conceptos y la experimentación aplicable de técnicas y métodos científicos en pro al objeto de la investigación.

Así mismo se empleó métodos de recolección de datos con el uso de recursos estadísticos para la exposición de datos que provienen de un cálculo o medición.

En su contexto los métodos utilizados son: teóricos, empíricos y matemáticos los cales ostentan el trabajo bajo una revisión bibliográfica de los acontecimientos e investigaciones realizadas con aportes significativos que aportan en el desarrollo de la aplicabilidad de este trabajo, así mismo a nivel práctico de hace uso de la metodología CRISP-DM (*Cross Industry Standard Process for Data Mining*) compuesta de seis etapas: análisis del problema, análisis de los datos, preparación de los datos, modelado, evaluación e implementación.

Métodos de Investigación

Histórico lógico: se utilizó para el estudio explícito de los antecedentes y bases teóricas que fundamentan la información de hechos aplicados, verídicos, de relevancia para la investigación.

Análisis-síntesis: se utilizó en el análisis de los componentes de la investigación, grupos identificados y seleccionados en el tratamiento de datos que posibilita la sistematización de las estrategias relacionadas en el desarrollo del trabajo de investigación.

Inducción–deducción: se hizo uso de este método en base al conocimiento enfocado en la naturaleza del problema de la investigación, compilar datos esenciales que permitan llegar a soluciones concretas a partir del objeto y campo de la investigación.

Sistémico-estructural-funcional: Desde la concepción particular de la investigación, basada en la determinación y definición de los componentes y relaciones demostrativas, en la fundamentación basada en modelo predictivo, se utilizó a partir de la identificación de la estructura, funcionalidad y niveles de relaciones jerárquicas entre los componentes del modelo así mismo los algoritmos utilizados durante la ejecución de minería de datos.

Métodos empíricos:

Observación: Método empleado en el análisis de fenómenos registrados en el campo de la investigación focalizando principales datos que permitan el ordenamiento de piezas que deán significado y alcance de la investigación.

Entrevista: Se utilizó en la recolección de orientaciones a partir de la interacción de autoridades de la Unidad educativa tomada para la realización de la investigación, así mismo la sistematización de adecuadas acciones y compromisos para favorecer a los grupos identificados en la investigación.

Método comparativo

Esta obra está bajo una licencia *Creative Commons* de tipo (CC-BY-NC-SA).
Grupo Editorial "ALEMA-Pentaciencias" E-mail: alema.pentaciencias@gmail.com

Dada la investigación el método comparativo se utilizó como estrategia que para analizar y comprender las similitudes y diferencias entre dos o más elementos, fenómenos, grupos o situaciones en el procesamiento de los datos en WEKA.

Para ello se establecieron los algoritmos comparativos propios de WEKA:

Algoritmo J48:

Desde la concepción de la investigación, este algoritmo se utilizó para la clasificación de individuos o entidades netas del modelo predictivo basado en arboles de decisión dados por el programa WEKA, la identificación de estancias correctamente clasificadas y por ende la visualización del árbol de decisión en donde se identifican los periodos relativos al proceso de enseñanza aprendizaje de las asignaturas del tronco común (Matemática y Lengua y Literatura).

Algoritmo Bayesnet

Algoritmo utilizado en la comparativa con enfoque de representación de las relaciones probabilísticas entre variables e inferencias sobre los datos recopilados y procesados.

Algoritmo OneR -B 6

Se utilizó como algoritmo comparativo basado en un modelo de reglas de clasificación de una única variable, siendo así un algoritmo simple pero efectivo hacia la selección de la variable que mejor predice la clase y establece una regla basada en esa variable.

Método práctico de la investigación

Arquitectura de la propuesta metodológica

Este método se utilizó como guía del modelo de procesos de minería de datos basado en las fases del proyecto de investigación, cuya finalidad estableció que cada proceso genere una validez en los atributos seleccionados, algoritmos de clasificación y selección de modelos.

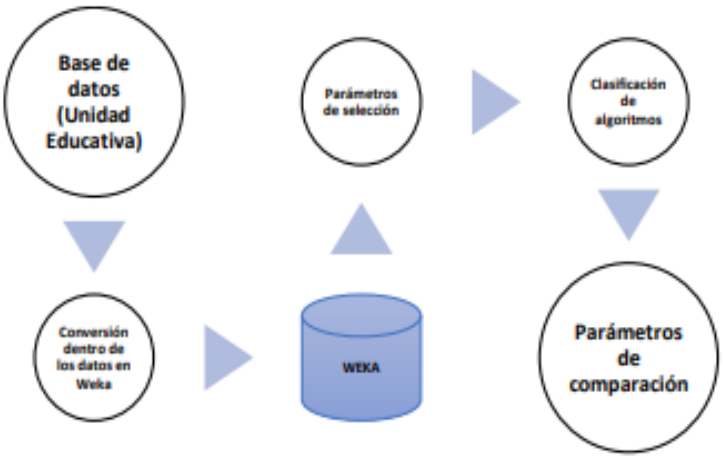


Figura 1: Metodología basada de proyectos de minería de datos
Fuente: Elaboración propia.

Esta obra está bajo una licencia *Creative Commons* de tipo (CC-BY-NC-SA).
Grupo Editorial "ALEMA-Pentaciencias" E-mail: alema.pentaciencias@gmail.com

Resultados y discusión

La investigación se encuentra basada por aspectos significativos, desde la instancia se enmarca el proceso de ETL, el alojamiento de los datos se atribuye al formato XML con datos nominales y ordinales. Desde este punto se establecen las clases y atributos para ser procesados, por ellos WEKA permite realizar proceso de ETL partiendo desde la extracción de datos, desde el sistema original, en la transformación de datos de forma clara, intangibles y de utilidad, de carga de los datos de acuerdo con los requerimientos y dirección de la investigación.

La muestra de datos presenta la correspondencia de estudiantes con necesidades educativas especiales en la Unidad Educativa Plan Internacional. Desde este aspecto la información recopilada fue el periodo académico dentro del proceso educativo correspondiente, la cantidad de estudiantes en dichos periodos matriculados por paralelo, balance de género (femenino y masculino) de los estudiantes con NNE, promedio de calificación en las Asignaturas de Matemática y Lengua y Literatura.

Tabla 1: Atributos con sus respectivos valores posibles dentro de la investigación.

Atributos	Valores posibles
Periodo académico	**2010 – 2022 (NM)**
Cantidad de estudiantes con NNE (Histórica)	**575 (CC)**
Género Femenino y Masculino	**F= 200; M=375 (CC)**
Promedio general de asignatura de Matemática	**P= 9≅2 (CC)**
Promedio general de asignatura de Lengua y Literatura	**P=9≅2 (CC)**

Fuente: Construcción propia.

Dentro del proceso de valoración se indica el análisis de los datos obtenidos en WEKA de forma general y depura de los datos relevantes que ostentan la dirección de la investigación.

La interpretación selecciona los siguientes datos:

- Periodo académico
- Total, de estudiantes
- Total, de estudiantes con NNE
- Calificación de asignatura de Matemática
- Calificación de asignatura de Lengua y Literatura

Evaluaciones de Resultados de los Algoritmos de Clasificación en WEKA

Los resultados de las herramientas se presentan en forma de esquema y gráficamente de acuerdo con la utilización de cada uno de los algoritmos:

Tabla 2: Algoritmos clasificadores de WEKA

Algoritmos	Asignaturas	ICC
J48 (C4.5)	Lengua y literatura	59.4575 %
	Matemática	
BayesNet	Lengua y literatura	58.03 %

Esta obra está bajo una licencia *Creative Commons* de tipo (CC-BY-NC-SA).
Grupo Editorial "ALEMA-Pentaciencias" E-mail: alema.pentaciencias@gmail.com

	Matemática	
OneR	Lengua y literatura	51.1777 %
	Matemática	

Fuente: Construcción propia.

La Tabla 2 representa la evaluación de resultados obtenidos a partir de tres algoritmos de WEKA (J48, BayesNet y OneR). Se aprecia en primera estancia los algoritmos utilizados, los porcentajes de las instancias clasificadas correctamente (ICC). Se observa que en el algoritmo J48 obtuvo un porcentaje de ICC (59.4575 %) denotando el mayor porcentaje de instancias correctamente clasificadas, para BayesNet (58.03 %) y para OneR (51.1777 %) por ende, se denota de acuerdo con los datos valores en gran medida a la robustez de los algoritmos, sin embargo, es importar recalcar una desvarianza entre el algoritmo seleccionado (J48) con mayor precisión hacia el algoritmo BayesNet como un algoritmo robusto.

Modelo J48

La implementación del algoritmo J48 de WEKA está basado en el árbol de decisión C4.5. Los árboles de decisión son parte de la taxonomía supervisada. Es decir, contiene variables dependientes o denominadas clases, basado en objetivos de clasificación para determinar el valor de esa o distintas clases. El proceso de creación de árboles comienza con el nodo raíz, conectados al nodo de las muestras y casos de entrenamiento. El primer paso es elegir la variable o el atributo que divide las muestras de entrenamiento originales (nodos raíz) de modo que haya una variación mínima de clase dentro de los subconjuntos generados. Por tanto, una vez determinada la variable que da mayor uniformidad con respecto a la clase de nodos hijos, se vuelve a realizar el análisis para cada nodo hijo. Si todos los nodos hoja contuvieran instancias de la misma clase, este proceso estaría acotado dentro de ciertos límites, pero este valor extremo se alcanzaría si se implementaran los métodos de poda previa y posterior del árbol.

Esta obra está bajo una licencia *Creative Commons* de tipo (CC-BY-NC-SA).
Grupo Editorial "ALEMA-Pentaciencias" E-mail: alema.pentaciencias@gmail.com

Figura 2. Aplicativo de datos en WEKA por el algoritmo J48.

Por otro la se muestra, la comparativa y datos clasificados del algoritmo BayesNet:

Figura 3. Modelo de datos clasificados en el algoritmo BayesNet.

Esta obra está bajo una licencia *Creative Commons* de tipo (CC-BY-NC-SA).

Grupo Editorial "ALEMA-Pentaciencias" E-mail: alema.pentaciencias@gmail.com

Datos relacionados al algoritmo *OneR -B 6*

Figura 4. Modelo de datos clasificados en el algoritmo OneR -B 6.

De acuerdo con la clase a interpretar se genera un árbol de decisión de los datos obtenidos para su verificación y análisis, en este caso el análisis de las asignaturas de matemática y lengua y literatura. Para ello es determinante el identificar la raíz, los nodos y las hojas.

Figura 5. Modelo de árbol de decisión de los datos en WEKA (Datos clasificados).

Esta obra está bajo una licencia *Creative Commons* de tipo (CC-BY-NC-SA).
Grupo Editorial "ALEMA-Pentaciencias" E-mail: alema.pentaciencias@gmail.com

En la figura 5 se presenta los datos valorados mediante el árbol de decisión de los atributos basado en las asignaturas de Lengua y Literatura y Matemática, dando como respuesta un intervalo de mejoramiento del rendimiento de los estudiantes de acuerdo con el periodo académico, en ella podemos especificar que en el periodo académico el balance de mayor calificación es de 8 en la asignatura de Lengua y Literatura y matemática con calificación <2, en este caso 833 de instancias correctamente clasificadas e instancias clasificadas incorrectamente 568 asentado en la suma total de 1401 instancias dado por los datos.

Comparativa de aplicaciones para la predicción de datos

Tabla 3: Aplicaciones para la predicción

Plataformas	Precisión	Utilización
WEKA	51%	43%
Power BI (PI)	61%	82%
TABLEAU PUBLIC	49%	49%

Fuente: Construcción propia

Se han realizado pruebas tomando como ejes WEKA, Power BI (PI) y Tableau Public las cuales han aportado a la investigación como entes de comparativos para la predicción de los datos.

Desde ello se destaca, resultados en la visualización de los datos considerando a Weka con un porcentaje del 51% de precisión de los datos en base a los algoritmos utilizados, Powe BI con un 61% de acuerdo con sus múltiples funciones para procesar datos y así mismo para su visualización y a Tableau Public 49% por su procesamiento de datos y servicios en la nube.

Discusión

Basado en el estudio (Mhetre & Nagar, 2017) muestra que la minería de datos es considerado más útil en el campo educativo, puesto que su postura presenta aplicabilidad en diversos factores nuevos para mejorar el desempeño del estudiante y los resultados se pueden obtener basado en datos sabios anuales. Además, más minería de datos técnicas como k-means, algoritmos de agrupamiento y otros algoritmos de clasificación se pueden aplicar en estos datos, en base a ello la presente investigación determino la utilidad de los procesos y técnicas de minería de datos arrojan que de forma relativa que la usabilidad WEKA y sus algoritmos ayudan a la generación de predicciones de procesos sin embargo, Al aplicar el algoritmo de J48 se pueden utilizar las tres herramientas de estudio, así como lo menciona (Fernández & Lujan-Mora, 2017), en donde según su investigación basado en procesos educativos con WEKA, se observa datos distintos de acuerdo al criterio de cada uno, así mismo presentan datos similares para el centroide de clústeres, se muestra además que WEKA muestra diferentes valores que están dentro de un rango aceptable.

Este estudio empleó la validación cruzada de los atributos según el periodo académico y las asignaturas seleccionadas (Matemática y Lengua y literatura) en los procesos académicos de estudiantes con NNE para aumentar la diferencia en la evaluación y obtener resultados más confiables en las métricas de los modelos. Además, produjo una menor cantidad de atributos y se emplearon más métricas de evaluación para caracterizar de mejor manera los modelos de predicción. Es importante destacar que a medida que aumenta

Esta obra está bajo una licencia *Creative Commons* de tipo (CC-BY-NC-SA).
Grupo Editorial "ALEMA-Pentaciencias" E-mail: alema.pentaciencias@gmail.com

la instancias clasificadas correctamente mayor es la exactitud de las predicciones por ende el modelo se vuelve más útil en entornos reales para predecir el rendimiento académico por periodo académico destacando así el algoritmo J48 como un algoritmo robusto.

En última instancia, es relevante señalar que las características utilizadas en este estudio pueden ser aplicadas en la construcción de modelos predictivos del rendimiento académico en otros cursos. Es decir, la metodología desarrollada puede ser replicada para diferentes cursos, y los atributos utilizados pueden ser recopilados como datos históricos con relevancia para la implementación de estrategias y adaptaciones pedagógicas destinadas a las asignaturas del tronco común.

Esto muestra la validación de los algoritmos utilizados y la manera muy similar en términos de precisión cuando se aplica un determinado algoritmo de clasificación los cuales permiten generar un modelo. Basado en la literatura es importante considerar que WEKA presenta el mayor número de algoritmos implementados de forma nativa, seguido por RapidMiner y finalmente Knime. Estas dos últimas herramientas utilizan un número significativo de algoritmos importados de WEKA en sus bases de datos.

Conclusiones

El uso de técnicas de minería de datos direccionada al ámbito educativo está en continua evolución de acuerdo con los requerimientos. Desde este aspecto la obtención de datos que permitan validad procesos y retentivas del conocimiento de los establecimientos educativos es relevante, queda mucho por mejorar en cuanto al análisis y aproximaciones al estudio de prácticas de minería de datos sobre todo en aplicaciones prácticas y las tareas de predicción y descripción que permitan tomar decisiones de los procesos que se llevan dentro del régimen educativo, por ello el énfasis de las asignaturas del tronco común como matemática y lengua y literatura son los ejes principales en la valoración del rendimiento académico de los estudiantes con necesidades educativas especiales por medio de la valoración de forma cuantitativa y cualitativa y los ejes a fortalecer en estas dos áreas.

En consecuencia, a la investigación se recalca que la minería de datos es un proceso de extracción de información útil y significativa de grandes conjuntos de datos, basado a ello la importancia de la extracción de datos ayuda a conseguir una forma predictiva de análisis del aprendizaje automático para descubrir patrones y tendencias en los datos, el uso de los clasificadores de árboles de decisión de forma simple, sin embargo en la minería de datos existen diversos algoritmos y aplicaciones robustas que potencian predicciones de datos, así mismo en el ámbito de WEKA tales como NaiveBayes, OneR los cuales sirven para mejorar o determinar predicciones de forma precisa.

A través de WEKA, se ha realizado el proceso de ETL, es decir el análisis y aplicación de forma estructurada, práctica y constructiva de los métodos que se encuentran implementados. Con el filtrado de datos se encontró en instancias gran masa de datos históricos obtenidos y se han creado un modelo de minería de datos de muestra, considerando los dos atributos a asemejar con respecto a las asignaturas en donde un mayor déficit se presenta en el área de matemática, por lo que la falta de potencialización de dichos conocimientos en estudiantes con necesidades educativas especiales representa un desbalance en la adaptación de contenidos.

Esta obra está bajo una licencia *Creative Commons* de tipo (CC-BY-NC-SA).

Grupo Editorial "ALEMA-Pentaciencias" E-mail: alema.pentaciencias@gmail.com

Agradecimientos

Un agradecimiento rotundo a las autoridades de la Unidad Educativa Plan Internacional por la disposición de la información brindada a esta investigación y por la aprobación rotunda de los resultados finales.
Rectora: Msc. Tania Alvarado A.
Vicerrector: Msc. Alex Mindiolaza M.
Responsable comisión Pedagógica: Karina Manzaba Ch.

Referencias

Barcia, M. J. J., & Carvajal, Z. B. T. (2015). El proceso de enseñanza aprendizaje en la educación superior. *Revista Electrónica Formación y Calidad Educativa (REFCalE)*, *3*, 139–154.

Castillo, Y. E., Labori De La Nuez, B., & Sifonte, R. S. (2020). *Extensión de índices de validación de grupo en la herramienta WEKA para la evaluación de algoritmos de agrupamiento Extension of group validation indices in the WEKA tool for the evaluation of clustering algorithms*. *13*(9), 179–187. http://publicaciones.uci.cu

Fernández, D. B., & Lujan-Mora, S. (2017). Comparison of applications for educational data mining in Engineering Education. *2017 IEEE World Engineering Education Conference (EDUNINE)*, 81–85. https://doi.org/10.1109/EDUNINE.2017.7918187

Mhetre, V., & Nagar, M. (2017). Classification based data mining algorithms to predict slow, average and fast learners in educational system using WEKA. *2017 International Conference on Computing Methodologies and Communication (ICCMC)*, 475–479. https://doi.org/10.1109/ICCMC.2017.8282735

Rodríguez Rodríguez, A., Rodríguez González, A. del C., Pino Tarragó, J. C., & Domínguez Gálvez, D. L. (2021). Implementación de algoritmos de Inteligencia Artificial en la predicción de nuevos conocimientos mediante enseñanza constructivista. *Serie Científica de La Universidad de Las Ciencias Informáticas*, *14(3)*, 131–141.

Rojas Salgado, M. E. (2021). Los retos de una educación virtual para estudiantes con necesidades educativas especiales. *HAMUT'AY*, *8*(1), 9. https://doi.org/10.21503/hamu.v8i1.2232

Santiesteban, S. I., Barba, A. J. V., & Álvarez, D. F. (2017). Inclusión de estudiantes con necesidades educativas especiales en la Universidad Técnica del Norte del Ecuador. *Revista Universidad y Sociedad*, *9*(3), 162–167.

Esta obra está bajo una licencia *Creative Commons* de tipo (CC-BY-NC-SA).
Grupo Editorial "ALEMA-Pentaciencias" E-mail: alema.pentaciencias@gmail.com

Revista Científica Arbitrada Multidisciplinaria
PENTACIENCIAS
e-ISSN: 2806-5794
https://editorialalema.org/index.php/pentaciencias

CARTA DE ACEPTACIÓN PARA PUBLICACIÓN DE ARTÍCULO

Dra. Leonor de la Concepción Moreno Suárez
Directora del Comité Editorial-Jefe
alema.pentaciencias@gmail.com

Ecuador, 14 de junio del 2023

EDITORIAL ALEMA INTERNACIONAL ORG

Estimados colegas:

Hilda Brigitte Quijije Quiroz [1], Kirenia Maldonado Zuñiga [2]

[1] Ingeniera en Tecnologías de la Información. Estudiante de la Maestría en Tecnologías de la Información y la Comunicación en el Instituto de Posgrado de la Universidad Estatal del Sur de Manabí. Jipijapa, Manabí, Ecuador. Docente de la Unidad Educativa Plan Internacional, Daule, Guayas, Ecuador. Correo: quijije-hilda2868@unesum.edu.ec
[2] Doctorando en Tecnologías de la Información y Comunicación, Universidad Nacional de Piura, Perú. Magister en Ciencias de la Educación, Licenciada en Educación Informática. Docente en la Carrera Tecnologías de la Información. Coordinadora y catedrática de la Maestría en Tecnologías de la Información y la Comunicación en el Instituto de Posgrado de la Universidad Estatal del Sur de Manabí. Jipijapa, Manabí, Ecuador. Correo: kirenia.maldonado@unesum.edu.ec

Me complace informarle que después del proceso de revisión por pares, el artículo **"Técnica de minería de datos para procesos educativos en estudiantes con necesidades educativas especiales basado en un modelo predictivo"** ha sido **ACEPTADO** para ser publicado por la Revista Científica Arbitrada Multidisciplinaria **PENTACIENCIAS** en su Vol. 5, No.5 (Julio Septiembre), 2023. e-ISSN: 2806-5794.

Saludos cordiales

Firmado electrónicamente por:
LEONOR DE LA CONCEPCION MORENO SUAREZ

Dra. Leonor de la Concepción Moreno Suárez
Revista Científica Arbitrada Multidisciplinaria PENTACIENCIAS
Directora del Comité Editorial-Jefe

Revista Científica Arbitrada Multidisciplinaria PENTACIENCIAS
Ecuador
EDITORIAL ALEMA INTERNACIONAL
Indexada en:

Revista Científica Arbitrada Multidisciplinaria
PENTACIENCIAS
e-ISSN: 2806-5794
https://editorialalema.org/index.php/pentaciencias

CERTIFICADO DE PUBLICACIÓN

Dra. Leonor de la Concepción Moreno Suárez
Directora del Comité Editorial-Jefe
alema.pentaciencias@gmail.com

Ecuador, 10 de Julio del 2023

EDITORIAL ALEMA INTERNACIONAL ORG

Después del proceso de revisión por pares, el artículo **"Técnica de minería de datos para procesos educativos en estudiantes con necesidades educativas especiales basado en un modelo predictivo",** ha sido **PUBLICADO** por la Revista Científica Arbitrada Multidisciplinaria **PENTACIENCIAS** en su Vol. 5, No.5 (Julio Septiembre), 2023. e-ISSN: 2806-5794.

De los autores:

Hilda Brigitte Quijije Quiroz [1], Kirenia Maldonado Zuñiga [2]

[1] Ingeniera en Tecnologías de la Información. Estudiante de la Maestría en Tecnologías de la Información y la Comunicación en el Instituto de Posgrado de la Universidad Estatal del Sur de Manabí. Jipijapa, Manabí, Ecuador. Docente de la Unidad Educativa Plan Internacional, Daule, Guayas, Ecuador. Correo: quijije-hilda2868@unesum.edu.ec
[2] Doctorando en Tecnologías de la Información y Comunicación, Universidad Nacional de Piura, Perú. Magister en Ciencias de la Educación, Licenciada en Educación Informática. Docente en la Carrera Tecnologías de la Información. Coordinadora y catedrática de la Maestría en Tecnologías de la Información y la Comunicación en el Instituto de Posgrado de la Universidad Estatal del Sur de Manabí. Jipijapa, Manabí, Ecuador. Correo: kirenia.maldonado@unesum.edu.ec

Firmado electrónicamente por:
LEONOR DE LA CONCEPCION MORENO SUAREZ

Dra. Leonor de la Concepción Moreno Suárez
Revista Científica Arbitrada Multidisciplinaria PENTACIENCIAS
Directora del Comité Editorial-Jefe

Revista Científica Arbitrada Multidisciplinaria PENTACIENCIAS
Ecuador
EDITORIAL ALEMA INTERNACIONAL
Indexada en: