UNIVERSIDAD TECNOLÓGICA DE LOS ANDES

FACULTAD DE INGENIERÍA

ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS E INFORMÁTICA

Tesis

Minería de datos y la ciberseguridad en la Universidad Tecnológica de los Andes, Abancay-2025

Asesor:

Dr. Baptista Velásquez, Adolfo Rafael

Autor:

Huamaní Martínez, Mary Carmen

Para optar el título profesional de:

Ingeniero de Sistemas e Informática

Abancay- Apurímac - Perú

2025

Acta De Sustentación

Universidad Tecnológica de los Andes

Transformando vidas

ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS E INFORMÁTICA

ACTA DE SUSTENTACIÓN DE TÍTULO PROFESIONAL Acta N°:001-2025

En la ciudad de **Abancay,** a los **24** días del mes de **noviembre** del 2025, siendo las **11:00** horas, se reunieron los integrantes del Jurado designado por Resolución Directoral **N° 063-2025-UTEA-FI-EPIS,** de la Escuela Profesional de **Ingeniería de Sistemas e Informática,** Facultad de **Ingeniería.**

Presidente:	Mg. Soria Donaires Fredy
Dictaminante:	Mg. Maruri Malpartida Nilton
Replicante:	Mg. Ugarte Warthon Katerine

Para evaluar la sustentación, en la modalidad de:

(X) Tesis ( ) Trabajo de suficiencia profesional

Titulado:

Minería de datos y la Ciberseguridad en la Universidad Tecnológica de los Andes Abancay -2025

Desarrollado por el (la) Bachiller:

Br. Huamani Martínez Mary Carmen

(Apellidos y Nombres)

Para optar el Título Profesional de:

Ingeniero de Sistemas e Informática

(Denominación del Título)

Concluido el acto, el Jurado dictaminó que el (la) mencionado(a) bachiller fue:
APROBADO(S) (X)

Por: **Unanimidad**

Emitiéndose la calificación final de:

Bachiller (Apellidos y Nombres)	**Calificación (**)**
Huamani Martínez Mary Carmen	Aprobado

Siendo las **12:15** horas concluyó la sesión, firmando los integrantes del Jurado.

Presidente: Mg. Soria Donaires Fredy Firma

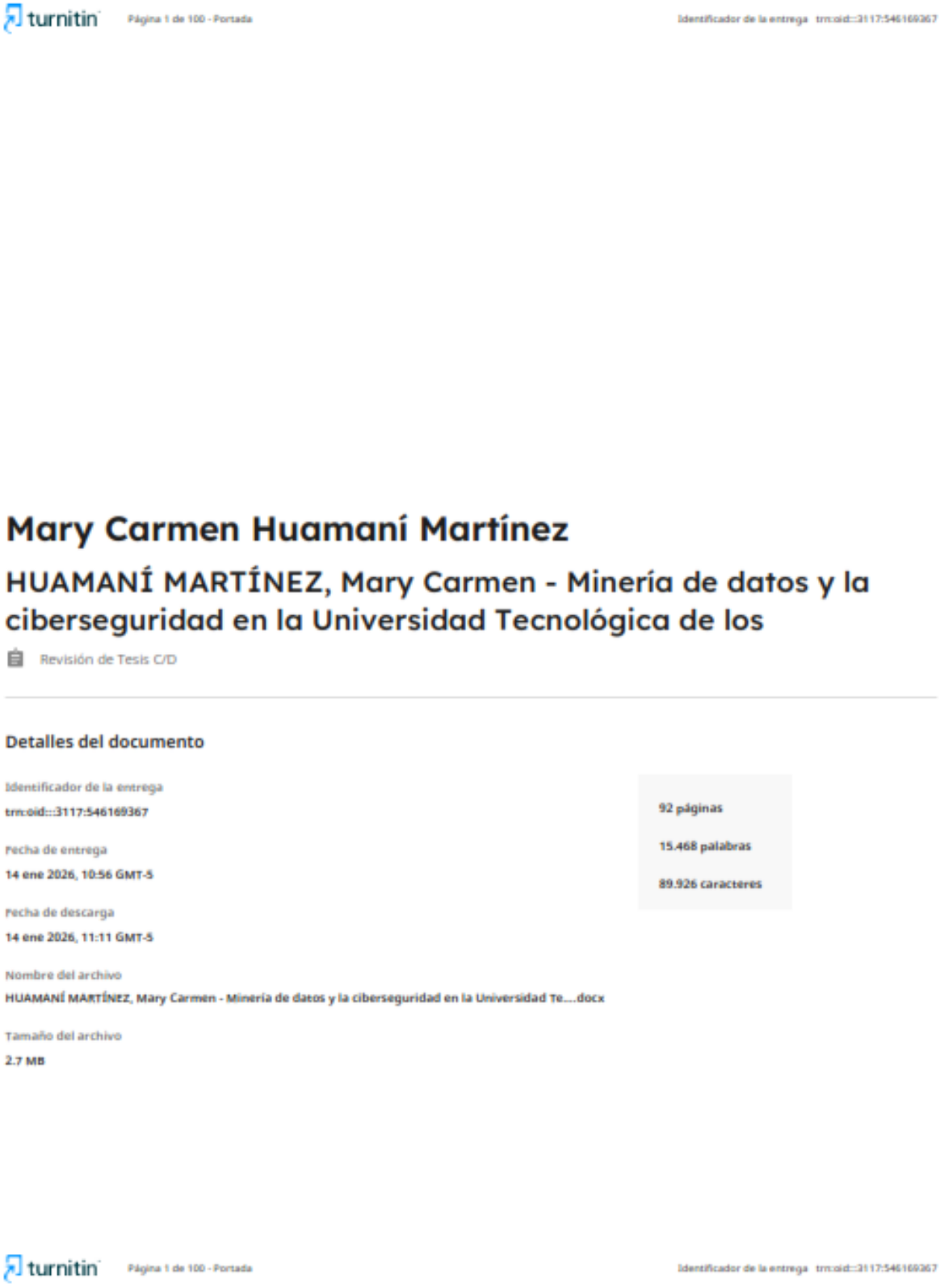
Dictaminante: Mg. Maruri Malpartida Nilton Firma

Replicante: Mg. Ugarte Warthon Katerine Firma

(*): Mayoría: Dos Integrantes del Jurado aprueban o desaprueban; Unanimidad: Todos los integrantes del jurado aprueban y desaprueban
(**): 0 a 10: Desaprobado, 11 a 15: Aprobado, 16 a 18: Aprobado Notable, 19 y 20: Aprobado con Distinción, Art. A8 RGGAT.

UNIVERSIDAD TECNOLÓGICA DE LOS ANDES
Ciudad Universitaria Av. Perú N° 700, Abancay, Central Telefónica 051 (083) 321559
Filial Cusco, Av. Grau N° 516, Teléfono (084) 251565
Filial Andahuaylas, Av. Juan Antonio Trelles N° 513 Teléfono (083) 421752
www.utea.edu.pe

Reporte De Similitud



24% Similitud general

El total combinado de todas las coincidencias, incluidas las fuentes superpuestas, para ca...

Filtrado desde el informe

- Bibliografía
- Texto citado
- Texto mencionado
- Coincidencias menores (menos de 9 palabras)

Exclusiones

- N.º de fuente excluida

Fuentes principales

21% Fuentes de Internet
4% Publicaciones
20% Trabajos entregados (trabajos del estudiante)

Marcas de integridad

N.º de alertas de integridad para revisión

Los algoritmos de nuestro sistema analizan un documento e
buscar inconsistencias que permitirían distinguirlo de una e
advertimos algo extraño, lo marcamos como una alerta para

Una marca de alerta no es necesariamente un indicador de
recomendamos que preste atención y la revise.

Metadatos

Datos de los Autores		
Apellidos y Nombres	:	Br. Huamaní Martínez, Mary Carmen
Tipo de Documento de Identidad	:	DNI: 46743045
URL ORCID	:	------
Datos de los asesores		
Apellidos y Nombres	:	Dr. Baptista Velasquez, Adolfo Rafael
Tipo de documento de Identidad	:	DNI:
Número de Documento de Identidad	:	45970028
URL ORCID	:	https://orcid.org/0000-0002-0475-0867
Datos de la Investigación		
Facultad	:	Ingeniería
Escuela Profesional	:	Ingeniería de Sistemas e Informática
Línea de Investigación	:	Informática, sociedad y gestión del conocimiento
Rango de años en que se realizó la investigación	:	Diciembre 2024 hasta agosto de 2025.
Fuente de financiamiento	:	Todo el recurso fue financiado por la investigadora.
Porcentaje de similitud	:	Índice de similitud 24%
URL de OCDE	:	https://purl.org/pe-repo/ocde/ford# 2.02.04

Agradecimientos

A Dios, por darme la vida, fortaleza y sabiduría para seguir adelante en los momentos más difíciles de este proceso.

A mis padres, Francisco y Graciela, por su amor, sacrificio y apoyo incondicional, quienes con mucha disciplina lograron alcanzar esta meta tan deseada.

A mis hermanos, César, Edith e Iris, por ser mis amigos y compañeros de vida que siempre estuvieron a mi lado, alentándome a seguir luchando por mis sueños.

A mi princesa Sulay, la cual desde que llegó a mi vida cambió por completo; me enseñó que, a pesar de las dificultades, uno puede llegar a cumplir las metas más anheladas.

A mi asesor, Dr. Adolfo Rafael Baptista Velásquez, por su orientación, compromiso y útiles recomendaciones, que resultaron esenciales para el desarrollo de la investigación.

Dedicatoria

Dedico mi Tesis a mis padres Francisco y Graciela, que son mis mayores inspiraciones en la vida, que gracias a su sabiduría puedo llegar a alcanzar esta meta tan anhelada, gracias a sus consejos y palabras de aliento que llegaron a ser el principal motivo para seguir luchando por mis metas y sueños.

Resumen

En la investigación se consideró el objetivo “existe un nivel de incidencia evidente de la minería de datos en la ciberseguridad de la Universidad Tecnológica de los Andes, Abancay 2025”. Fundamentado en la metodología tipo dogmática, nivel explicativo y diseño no experimental transeccional correlacional causal, bajo una población y muestra de 14 unidades de análisis establecida por el método no probabilístico, aplicando la técnica de la encuesta y el instrumento del cuestionario para lograr datos sustanciales. Alcanzando resultados, donde el 92.86% de los colaboradores universitarios afirmaron regular el manejo básico y cotidiano de la minería de datos orientado al tratamiento de la información diaria universitaria, incidiendo en la seguridad; toda vez que el 92.86% señalaron regular la ciberseguridad de la información, frente a las amenazas digitales. Concluyendo que, la minería de datos incide evidentemente en la ciberseguridad de la Universidad Tecnológica de los Andes, Abancay 2025, fijada por Wilcoxon con sig. $0.001<0.05$; en vista que la minería de datos como técnica de análisis predictivo y descriptivo de procesamiento de grandes volúmenes de datos no sólo llegará a fortalecer la detección, prevención y respuesta a las amenazas digitales, sino que también permitirá identificar vulnerabilidades críticas de los sistemas informáticos, la reducción de incidentes como accesos no autorizados, los ataques a la red académica, la filtración de datos sensibles y el fortalecimiento de la infraestructura de la ciberseguridad universitaria.

Palabras clave: Minería de datos, ciberseguridad, principio de seguridad, estado de la información y contramedidas.

Abstract

The research objective was considered "there is an evident level of incidence of data mining in the cybersecurity of the Technological University of the Andes, Abancay 2025". Based on the dogmatic type methodology, explanatory level and non-experimental correlational causal design, under a population and sample of 14 analysis units established by the non-probabilistic method, applying the survey technique and the questionnaire instrument to obtain substantial data. Achieving results, where 92.86% of university collaborators stated that they regulate the basic and daily management of data mining aimed at the treatment of daily university information, influencing security; since 92.86% indicated that they regulate the cybersecurity of information, against digital threats. Concluding that, data mining evidently affects the cybersecurity of the Technological University of the Andes, Abancay 2025, set by Wilcoxon with sig. 0.001 <0.05; Given that data mining, as a predictive and descriptive analysis technique for processing large volumes of data, will not only strengthen the detection, prevention, and response to digital threats, but will also identify critical vulnerabilities in computer systems, reduce incidents such as unauthorized access, attacks on the academic network, the leakage of sensitive data, and strengthen the university cybersecurity infrastructure.

Key words: Data mining, cybersecurity, security principle, information status, and countermeasures.

Índice

Índice De Tablas

Índice De Figuras

Índice De Anexos

I. Introducción

La información en las organizaciones a nivel mundial crece sin parar, algunas estimaciones apuntan a que el 90% de los datos en el mundo se ha creado en los últimos dos años y se predice un crecimiento de un 40 % anual, en este contexto, la Minería de datos (MD) o data mining (DM) se presenta como una práctica estratégica relevante para las organizaciones que emplean la inteligencia empresarial (business intelligence) (Grupo Iberdrola [GI], 2024). De donde, en la MD o DM, es crucial distinguir la información relevante de la irrelevante, permitiendo así a las organizaciones enfocar sus recursos de manera más efectiva (Sánchez, 2024). De donde la realidad de las organizaciones latinoamericanas ha acelerado el paso hacia la transformación digital, trayendo no solo ventajas, sino también nuevos riesgos y medidas de seguridad necesarias para que las empresas puedan proteger la información y documentos digitales de los ciberdelincuentes (Guzmán, 2022). En esa línea, la seguridad en la minería de datos protege la información al evitar el acceso, la alteración o la divulgación no autorizados durante el proceso del tratamiento de datos, siendo crucial la ciberseguridad en la minería de datos cuando se llegan a proteger datos sensibles o confidenciales, para el cual se llegan a utilizar técnicas de seguridad como la autenticación de usuarios, el cifrado de datos, los sistemas de detección de intrusiones y la computación segura de múltiples partes, hasta lograr la confidencialidad e integridad de los datos y evitar el acceso no autorizado (Brook, 2024).

El escenario en la que se encuentran muchas organizaciones peruanas y especialmente la Universidad Tecnológica de los Andes, Abancay; es que la toma de sus decisiones la dan de forma intuitiva o por conocimientos históricos que posee la entidad, por situaciones que han atravesado previamente; sin embargo, deben emplear diversas técnicas y metodologías para

proteger los datos y contra posibles amenazas durante el proceso de minería de datos en la organización (Dongo y Silva, 2020).

Realidades sostenidas en líneas anteriores, que facultaron la organización del estudio en diferentes capítulos; detallando al capítulo uno; donde se específica la introducción de la investigación; capítulo dos: problema de investigación; que comprende la descripción de la realidad problemática, el establecimiento de los problemas, los objetivos, las respectivas justificaciones e importancia, así como las hipótesis del estudio, la variables problemáticas, la operacionalización de variables; capitulo tres: marco teórico; con los pertinentes antecedentes, las bases teóricas, y definición de términos; capitulo cuatro; el diseño metodológico; partiendo del tipo y alcance de estudio, así como el diseño de investigación, la población y muestra, además de las técnicas e instrumentos para el logro de datos, y al final las técnicas de procesamiento y análisis de datos; capítulo quinto: resultados y discusiones; presentando los resultados descriptivos e inferenciales de la investigación, así como las coherentes discusiones; capitulo seis: las conclusiones del estudio; capítulo siete: las recomendaciones oportunas; capítulo ocho las fuentes bibliográficas consultadas, y al final el capítulo nueve: con los anexos de la investigación.

II. Problema De Investigación

2.1. Descripción de la realidad problemática

En un mundo globalizado donde el acceso a la información es casi ilimitado, la capacidad de analizar y extraer datos relevantes marca la diferencia entre las empresas que prosperan y las que quedan atrás, por cuanto las organizaciones a nivel mundial, se encuentran analizando a fondo los datos que fueron recolectando sin prestarles demasiada atención, y con gran sorpresa, descubren que dentro de los respectivos datos existe una riqueza de información no aprovechada, tales como: patrones de comportamiento de clientes, tendencias de compra, preferencias de productos y mucho más, encontrando oro, no en una mina, sino en tu base de datos, es aquí donde la minería de datos (MD) o data mining (DM) entra en juego como una herramienta clave para transformar esos datos crudos en conocimiento procesable, que pueda aplicarse directamente a las decisiones organizacionales (Base Sur Digital [BSD], 2024).

En esa línea, en el mundo del business, la minería de datos (MD) o data mining (DM), es una práctica cada vez más extendida, convirtiéndose en una de las prioridades principales de los CIO's (responsables de tecnologías de información), y son herramientas fundamentales para captar e integrar un conjunto aún más amplio de datos como parte de los procesos de toma de decisiones, y que forman parte del proceso de definición de estrategias de negocio y toma de decisiones empresariales (Bismart, 2024). De donde, en la MD o DM, es crucial distinguir la información relevante de la irrelevante, permitiendo así a las organizaciones enfocar sus recursos de manera más efectiva (Sánchez, 2024). La información en las organizaciones a nivel mundial crece sin parar, algunas estimaciones apuntan a que el 90% de los datos en el mundo se ha creado en los últimos dos años y se predice un crecimiento de un 40 % anual, en este contexto, la MD o DM se presenta como

una práctica estratégica relevante para las empresas que emplean la inteligencia empresarial (business intelligence) (Grupo Iberdrola [GI], 2024). La realidad de las empresas latinoamericanas ha acelerado el paso hacia la transformación digital, trayendo no solo ventajas, sino también nuevos riesgos y medidas de seguridad necesarias para que las empresas puedan proteger la información y documentos digitales de los ciberdelincuentes, sin importar si se tratan de organizaciones pequeñas, medianas o grandes, señalando que en el 2020 América Latina sufrió al menos 91 billones de intentos de ciberataques a organizaciones y al menos el 45,3% de ellas fueron víctimas, por lo que la Ciberseguridad ha tomado un papel más protagónico (Guzmán, 2022). La seguridad en la minería de datos protege la información al evitar el acceso, la alteración o la divulgación no autorizados durante el proceso del tratamiento de datos, incluyendo la protección del almacenamiento de datos y las bases de datos, los algoritmos de minería de datos, el proceso de transmisión de datos y los resultados del análisis de minería de datos; siendo crucial la ciberseguridad en la minería de datos cuando se llegan a proteger datos sensibles o confidenciales, para el cual se llegan a utilizar técnicas de seguridad como la autenticación de usuarios, el cifrado de datos, los sistemas de detección de intrusiones y la computación segura de múltiples partes, hasta lograr la confidencialidad e integridad de los datos y evitar el acceso no autorizado (Brook, 2024).

El escenario en la que se encuentran muchas organizaciones peruanas es que la toma de sus decisiones la dan de forma intuitiva o por conocimientos históricos que posee la entidad, por situaciones que han atravesado previamente; sin embargo, los contextos cambian y con ellos los clientes o usuarios, y sus requerimientos, es por ello que las organizaciones y entidades educativas superiores deben ir de la mano con la tecnología disponible, como ser del data mining cuya capacidad radica en la transformación de datos

crudos en información estratégica, lo que facilita la creación de ventajas competitivas y la identificación de nuevas oportunidades de mercado y que estén concatenadas a la seguridad en la minería de datos, empleando diversas técnicas y metodologías para proteger los datos y contra posibles amenazas durante el proceso de minería de datos en la organización (Dongo y Silva, 2020).

En la Universidad Tecnológica de los Andes, sede Abancay, las tecnologías digitales que maneja le permiten almacenar grandes volúmenes de datos académicos y financieros, donde todo ese flujo de información recopilada han logrado satisfacer las necesidades diarias de la entidad universitaria, pero con preocupación se observa que en la actualidad se considera que presentan problemas inherentes a las capacidades humanas para analizar, transformar y la seguridad de la información para generar un conocimiento útil y relevante que apoye a la toma de decisiones de las autoridades y la comunidad uteina, en vista que los respectivos datos almacenados guardan información valiosa, pero por tratarse de una condensación masiva de datos no recurren a la aplicación de técnicas y herramientas de Minería de Datos (MD) o Data Mining (DM).

Ambiente que exigió el desarrollo de la investigación, donde se pudo observar la realidad del comportamiento de la minería de datos y de la ciberseguridad en la entidad universitaria, para responder interrogantes múltiples y cada vez más compuestas sobre lo que saben los responsables del tratamiento de datos, autoridades y administrativos, o si se encuentran comprometidos con la ciberseguridad de las grandes cantidades de información procesadas automáticamente por la MD como un habilitador de innovación y comprender la importancia de proteger los sistemas que se manejan en el control de datos e información interna, para concurrir a un conocimiento útil para los usuarios y/o clientes, y satisfacer sus propósitos relacionados al entendimiento a detalle de un sistema

de información sobre el contexto y la calidad educativa, que permita mejorar los procesos de los servicios educativos universitarios actuales.

2.1.1 Problema General

¿Cuál es el nivel de incidencia de la minería de datos en la ciberseguridad de la Universidad Tecnológica de los Andes, ¿Abancay 2025?

2.1.2 Problemas Específicos

- ¿Cuál es el nivel de incidencia de la minería de datos en los principios de seguridad de la ciberseguridad de la Universidad Tecnológica de los Andes, Abancay 2025?
- ¿Cuál es el nivel de incidencia de la minería de datos en los estados de la información de la ciberseguridad de la Universidad Tecnológica de los Andes, Abancay 2025?
- ¿Cuál es el nivel de incidencia de la minería de datos en las contramedidas de la ciberseguridad de la Universidad Tecnológica de los Andes, Abancay 2025?

2.2 Objetivos

2.2.1 Objetivo General

Establecer el nivel de incidencia de la minería de datos en la ciberseguridad de la Universidad Tecnológica de los Andes, Abancay 2025.

2.2.2 Objetivos Específicos

- Identificar el nivel de incidencia de la minería de datos en los principios de seguridad de la ciberseguridad de la Universidad Tecnológica de los Andes, Abancay 2025.
- Identificar el nivel de incidencia de la minería de datos en los estados de la información de la ciberseguridad de la Universidad Tecnológica de los Andes, Abancay 2025.
- Identificar el nivel de incidencia de la minería de datos en las contramedidas de la ciberseguridad de la Universidad Tecnológica de los Andes, Abancay 2025.

2.3 Justificación e importancia

2.3.1 Justificación

Las razones que justifican el desarrollo de la investigación, parte de la significativa importancia que presenta ella, para analizar y observar de qué manera se viene desarrollando el tratamiento y manejo de los datos, que son generados del entorno educativo universitario para visualizar los patrones de comportamiento por intermedio de la minería de datos, a partir de la exploración de los tipos únicos de datos académicos y administrativos almacenados, y el conjunto de prácticas que se están realizando para medir la seguridad y proteger todos los sistemas tecnológicos físicos y lógicos, así como los sistemas que se encuentran implantados en la institución educativa superior para el control de datos e información interna, frente a un posible ataque cibernético que puedan llegar a secuestrar la información organizacional, a partir de las cuales llegar a proponer estrategias de mejora continua e innovación, con la finalidad de diseñar acciones para poder resolver y mejorar los procesos de análisis de información para la intensión de descubrir patrones y correlaciones seguras que permitan garantizar y mejorar los aprendizajes educativos de forma automatizada, oportuna y segura para la toma de decisiones y la satisfacción de los usuarios de la primera Casa Superior de Estudios de Abancay, Apurímac.

2.3.2 Importancia

El estudio sobre la minería de datos y la incidencia que pueda generar en la ciberseguridad en la zona de influencia es de gran significancia porque permite comprender y observar la gestión de los volúmenes de información que se encuentran generando en la universidad, además de la comprensión de cómo las técnicas de análisis masivo de la información impactan en la protección, gestión y resiliencia de los sistemas informáticos y

datos académicos-administrativos universitarios, a partir de los resultados a ser alcanzados la investigación fortalecerá tanto la seguridad digital e información, como la capacidad de gestión de datos de la UTEA, convirtiéndose en un aporte al conocimiento de las variables problemáticas, así como para el desarrollo académico, científico y tecnológico digital.

2.4 Hipótesis

2.4.1 Hipótesis general

Existe un nivel de incidencia evidente de la minería de datos en la ciberseguridad de la Universidad Tecnológica de los Andes, Abancay 2025.

2.4.2 Hipótesis específicas

- Existe un nivel de incidencia evidente de la minería de datos en los principios de seguridad de la ciberseguridad de la Universidad Tecnológica de los Andes, Abancay 2025.
- Existe un nivel de incidencia evidente de la minería de datos en los estados de la información de la ciberseguridad de la Universidad Tecnológica de los Andes, Abancay 2025.
- Existe un nivel de incidencia evidente de la minería de datos en las contramedidas de la ciberseguridad de la Universidad Tecnológica de los Andes, Abancay 2025.

2.5 Variables

2.5.1 Variable Independiente:

Minería de datos.

2.5.2 Variable dependiente:

Ciberseguridad

2.5.3 Operacionalización de Variables

Variable	Definición conceptual	Definición operacional	Dimensiones	Indicadores
Variable independiente: Minería de datos	"Es una técnica asistida por computadora que se utiliza en los análisis para procesar y explorar grandes conjuntos de datos" (Amazon Web Services [AWS], 2023).	Son procedimientos para la operación, análisis y exploración descriptiva, predictiva y prescriptiva de inmensas cantidades de datos almacenados en la organización.	Descriptivas	Asociación. Funciones Agrupamiento. Eventos. Relevancia.
			Predictivas	Clasificación de datos. Estrategias. Categorías. Regresión. Detección de anomalías.
			Prescriptivas	Automatización. Reglas. Análisis de información. Optimización. Simulaciones. Respuestas.

Variable	Definición conceptual	Definición operacional	Dimensiones	Indicadores
Variable dependiente: Ciberseguridad	"Es la práctica de proteger equipos, redes, aplicaciones de software, sistemas críticos y datos de posibles amenazas digitales" (Amazon Web Services [AWS], 2024).	Comprende los procedimientos para brindar seguridad a los sistemas y tecnologías de información y comunicación, así como la información considerando el principio de seguridad, el estado de la información y las contramedidas que aplica la organización.	Principio de la seguridad.	Confidencialidad. Integridad. Disponibilidad
			Estado de la información.	Identificación. Transito Almacenado. Proceso.
			Contramedidas	Habilidades profesionales. Tecnologías Dispositivos. Productos. Políticas Prácticas Actualización.

III. Marco Teórico

3.1 Antecedentes del problema

3.1.1 Internacional

A partir de las consultas de estudios y artículos científicos, se tiene antecedentes internacionales como la investigación de Matilde (2023), bajo el objetivo de identificar qué tipo de conocimiento puede estar faltando en los planes de estudio actuales, así como medir tendencias y proponer líneas de acción. Con una metodología de enfoque mixto y de tipo exploratorio, descriptivo, no experimental, transversal. Cuyos resultados más relevantes se tiene que existe una leve diferencia entre el nivel de conocimientos en ciberseguridad solicitados por las empresas y el que realmente tienen los egresados de carreras afines a la seguridad cibernética. Concluyendo que, la ciberseguridad no es considerada de manera proactiva dentro de las organizaciones, y que es necesario seguir investigando por qué existe un déficit tan grande de profesionales especialistas en ciberseguridad.

En el estudio de Vallejo y Tenelanda (2022), cuyo objetivo fue; mostrar el aporte a la seguridad de la información de la minería de datos en el contexto de la detección de intrusos. Manejo la metodología CRoss-Industry Standard Process for Data Mining CRISP-DM. Concluyendo que, la minería de datos es el proceso de ahondar en los datos para detectar patrones y relaciones ocultos; emplea una orientación empresarial clara y potentes tecnologías analíticas para explorar rápida y concienzudamente montañas de datos y extraer de ellas la información útil y aplicable que se necesita.

Pozo (2022), en la investigación donde el objetivo fue; conocer las estrategias que utiliza el gobierno ecuatoriano para el cuidado y protección de los sistemas cibernéticos. Utilizando una metodología cualitativa, explicativa y descriptiva. Concluyendo que, no

existe un modelo de ciberseguridad específico que pueda dar seguridad al entorno actual, sino que utiliza un conjunto de estrategias, lineamientos y objetivos de carácter político para proteger el ciberespacio, donde el entorno digital ecuatoriano no se enfoca en garantizar la identificación, protección y detección de los ataques debido a que no se cuenta con un modelo oportuno para la ciberseguridad y se limita al cumplimiento de un plan estratégico de políticas de ciberseguridad.

Por otro lado, Torres (2022), en su investigación señala como objetivo; aplicar minería de datos para determinar los factores más influyentes en la ocurrencia de siniestros de tránsito en Ecuador en el año 2020. Con una metodología cuantitativa, explicativa y las herramientas OpenRefine y RStudio. Llegando a resultados de que el algoritmo CHAID Exhaustivo fue el que obtuvo los mejores resultados con un porcentaje de clasificación correcta del 58,38% y 44,60% de precisión, con el cual se identificó los patrones más importantes en los datos y se evaluó las posibles asociaciones entre las variables recogidas. Concluyendo que, el factor humano es el factor más influyente con una probabilidad de ocurrencia del 69,64%.

3.1.2 Nacional

En la investigación de Correa (2022), donde el objetivo fue; determinar la incidencia de la ciberseguridad en el tratamiento de datos personales en una Municipalidad distrital de Lima Sur, 2021. Aplicando una metodología tipo aplicada, diseño no experimental, corte transversal, correlacional-causal. De cuyos resultados llega a concluir que, la ciberseguridad incide significativamente en el tratamiento de datos personales en la entidad de estudio.

Por su parte y considerando a Arizaca (2022), quién desarrolló la investigación basada en el objetivo; evaluar las transferencias económicas a las municipalidades, y su

impacto en el desarrollo humano y la reducción de la pobreza de los distritos mineros del Perú aplicando data mining. La metodología cuantitativa y descriptiva. De los resultados alcanzados concluye que, las correlaciones de la variable de interés con los indicadores de desarrollo como IDH, pobreza y NBI son prácticamente nulas, en cambio las transferencias de Foncomun y del programa de vaso de leche tienen una mejor correlación, variando entre 0.26 y 0.30 respectivamente que a pesar de ser baja demostraron que hay una influencia de estos recursos sobre los indicadores de pobreza.

Partiendo de la investigación de Jimenez (2022), basado en el objetivo; establecer un modelo de detección de amenazas digitales para mitigar los riesgos de ciberseguridad en las organizaciones mediante la evaluación de la influencia de la infraestructura de ciberseguridad en el índice global de ciberseguridad. La Metodología de estudio cuantitativa, no experimental, longitudinal, descriptiva correlacional. Arribando a la conclusión que, los diferentes escenarios donde las organizaciones se desarrollan consideran aspectos para tener en cuenta en la construcción de plataformas seguras, donde cada escenario debe ser considerado al momento de planificar la estrategia de prevención o defensa de los servicios digitales, considerando que cada tecnología demanda servicios de protección diferente y gestión especializada.

En consideración al estudio de Chero (2020), donde el objetivo fue; identificar las técnicas de minería de datos para mejorar el análisis de la gestión de recursos humanos del proyecto especial Altomayo. Siendo la metodología descriptiva, cuasi experimental. Llegando a la conclusión que, la calidad de datos y tipo de datos pertenecen a variables categóricas de tipo binomial y polinomial que permiten aplicar los algoritmos clasificadores, estableciendo que los algoritmos de árbol de clasificación y redes bayesianas son los que se adecuan a los datos a examinar para generar el aprendizaje en

la fase de entrenamiento, donde el algoritmo de árbol de decisiones tiene un menor rendimiento sobre la red bayesiana en cuanto al propósito de predicción.

3.1.3 A Regional y/o locales

Dongo y Silva (2020), en la investigación cuyo objetivo fue, identificar y mostrar cómo aplican algunas empresas del sector retail la minería de datos en sus operaciones. Con una metodología cualitativa. Quienes llegan a la conclusión que, la minería de datos es una herramienta que ayuda a mejorar la toma de decisiones en empresas del sector retail, más no es una herramienta generadora de alternativas de solución a un determinado problema que presente la empresa, siendo necesario que un personal encargado que analice los datos procesados por la herramienta, vea su proyecciones y los posibles efectos en la empresa, después del análisis correspondiente, se puede tomar una decisión que permitirá mejorar la empresa.

Finalmente, Hernández (2020), en su estudio cuyo objetivo fue, identificar vulnerabilidades informáticas mediante la aplicación de las herramientas de detección de vulnerabilidades en el portal Web de la Universidad Andina del Cusco. Con una metodología experimental, descriptivo. Donde los resultados basados en las pruebas de vulnerabilidad de los programas informáticos Kali-Linux y Acunetix Web Vulnerability Scanner y existen vulnerabilidades en el portal Web. Concluyendo que, el portal Web de la Universidad Andina del Cusco presenta vulnerabilidades de nivel medio/alto, tanto en vulnerabilidades de diseño, implementación y uso.

3.2 Bases teóricas

3.2.1 Minería de datos

La minería de datos “es una técnica asistida por computadora que se utiliza en los análisis para procesar y explorar grandes conjuntos de datos” (Amazon Web Services

[AWS], 2023).

De acuerdo a las manifestaciones de Holdsworth (2024), la minería de datos "es el uso del machine learning y el análisis estadístico para descubrir patrones y otra información valiosa a partir de grandes conjuntos de datos".

Además, Business Technology Platform (BTP 2023), afirman que la minería de datos

"es el proceso de utilizar herramientas analíticas avanzadas para extraer información útil proveniente de una acumulación de datos".

3.2.1.1 Características de la minería de datos

Según el Instituto Europeo de Posgrado (IEP 2020), la minería de datos o Data Mining para llegar a manejar inmensas cantidades de datos para y tomar las mejores decisiones en la organización, contiene determinadas características, siendo:

- Tendencias: caracterizado en conseguir y hallar tendencias entre grandes volúmenes de datos, por intermedio de patrones o reglas que se repiten y que pueden llegar a explicar o detectar ciertos comportamientos.
- Ayuda a la toma de decisiones: una cantidad significativa de datos no sirve de mucho si no se analizan y extraen conclusiones, por cuanto debe bucear dentro de esos datos para quedarse con los relevantes y arribar a decisiones informadas.
- Previsión: ayuda a predecir y, por lo tanto, puede ser un complemento para que las empresas se anticipen a posibles escenarios.
- Descubrir conocimiento: descubre conocimientos en las bases de datos, no siendo importante tener una inmensa base de datos, sino dar con la clave de patrones que permitan generar ese conocimiento.

- Tecnología: sus técnicas van mejorando conforme lo hace su tecnología y, actualmente, se sirve de disciplinas como la estadística (estudia la variabilidad), inteligencia artificial (inteligencia llevada a cabo por máquinas) y machine learning (aprendizaje automático de máquinas).

Diferentes ámbitos: constituye en el análisis predictivo que se puede realizar en cualquier sector, aunque los más recurrentes son el marketing, los comercios o la banca.

3.2.1.2 Funcionamiento de la minería de datos

Kaspersky (2025), señala que, la minería de datos "implica la evaluación y el análisis de grandes volúmenes de información para encontrar patrones y tendencias importantes

(párr. 1), determinando que el proceso presenta los siguientes pasos:

- Define la meta: determina el objetivo claro al comienzo del proceso de minería de datos, respondiendo a las siguientes interrogantes: ¿quieres obtener más información sobre el comportamiento de los clientes?, ¿Quieres reducir los costos o aumentar las ganancias?, ¿Quieres identificar el fraude?, etc.
- Recopila los datos: por lo general, las organizaciones tienen datos almacenados en varias bases de datos.
- Depura los datos: una vez seleccionado los datos, es necesario depurarlos, reformatearlos y validarlos.
- Interroga los datos: los analistas se familiarizan con los datos mediante la ejecución de análisis estadísticos y la creación de figuras y tablas visuales.
- Desarrolla un modelo: el objetivo es encontrar un enfoque a la minería de datos que produzca los resultados más útiles.

- Valida los resultados: se examinan los resultados para verificar si los hallazgos son precisos. Si no lo son, se debe rehacer el modelo y volver a intentarlo.
- Implementa el modelo: las apreciaciones que se descubrieron pueden utilizarse para cumplir la meta que se definió al comienzo del proceso (Kaspersky, 2025).

3.2.1.3 Objetivos de la minería de datos

De acuerdo a las manifestaciones de IBM (2025), cuando se trabaja con datos para definir una solución técnica al problema comercial, y busca lograr los objetivos que deben ser concretos:

- Describir el tipo de problema de minería datos, como clúster, predicción o clasificación.
- Documentar objetivos técnicos, utilizando unidades específicas de tiempo, como predicciones con una validez de tiempos de ejecución.
- Proporcionar datos reales para resultados deseados. (IBM, 2025, párr. 1).

3.2.1.4 Fases del proceso de minería de datos

Al analizar las distintas fases o procesos estándares para la minería de datos, los equipos de datos se trasladan de una fase a otra de acuerdo a la necesidad (Amazon Web Services [AWS], 2023), pudiendo realizar algunas de estas tareas o apoyarlas:

- Comprensión del negocio: identifica los objetivos y el alcance del proyecto, con las partes interesadas de la organización para identificar cierta información:
  - Problemas que se deben abordar
  - Restricciones o limitaciones del proyecto
  - El impacto empresarial de las posibles soluciones

- Comprensión de los datos: luego de comprender el problema organizacional, se realiza un análisis preliminar de los datos, recopilando conjuntos de datos de diversos orígenes, obtienen los derechos de acceso y elaboran un informe de descripción de datos, al final evalúan la calidad de los datos y eligen un conjunto de datos final para la siguiente fase.
- Preparación de los datos: el software de minería de datos requiere datos de alta calidad, por cuanto recopilan y almacenan datos por razones distintas a la minería.
- La preparación de los datos implica los siguientes procesos.
    - Limpiar los datos; gestionar los datos que faltan, los errores de datos, los valores predeterminados y las correcciones de datos.
    - Integrar los datos: combinar dos conjuntos de datos dispares para obtener el conjunto de datos objetivo final.
    - Dar formato a los datos; convertir los tipos de datos o configurar los datos para la tecnología de minería específica que se utiliza.
- Modelado de datos: introducir datos preparados en el software de minería de datos y estudian los resultados, además se deben escribir pruebas para evaluar la calidad de los resultados de la minería de datos. Para modelar los datos, se cuenta con las siguientes opciones:
    - Entrenar los modelos de machine learning (ML) a partir de conjuntos de datos más pequeños con resultados conocidos
    - Utilizar el modelo para analizar más a fondo conjuntos de datos desconocidos

- Ajustar y volver a configurar el software de minería de datos hasta que los resultados sean satisfactorios

- Evaluación: luego de creados los modelos, se comienza a medirlos con respecto a los objetivos empresariales originales, compartiendo los resultados con los analistas de negocio y obtienen comentarios.
- Implementación: en la implementación, otras partes interesadas utilizan el modelo de trabajo para generar inteligencia empresarial, parten de la planificación del proceso de implementación, que incluye instruir a otros sobre las funciones del modelo, realizar un seguimiento continuo y mantener la aplicación de minería de datos (Amazon Web Services [AWS], 2023).

3.2.1.5 Desafíos de la minería de datos

La minería de datos contiene desafíos para las empresas, que le permita detectar los datos redundantes, el error humano, los problemas de seguridad y la falta de profesionales especializados (Zendesk, 2023), entre ellas se tiene:

- Datos redundantes: las organizaciones señalan que la duplicación de datos es uno de los principales problemas del equipo de inteligencia, sobre todo por la falta de una herramienta compartida entre diferentes unidades y un estándar para ingresar datos.
- Error humano: significa que alguien ha recopilado y almacenado los respectivos datos antes, y es posible que no se haya hecho de acuerdo con las mejores prácticas.
- Seguridad y privacidad: la ciberseguridad es clave para cualquier empresa, por cuanto la privacidad de los datos ya no es un requisito de unos pocos clientes, sino una responsabilidad comercial según la ley.

- Falta de profesionales especializados: La tecnología es una gran aliada, pero no hace milagros, por más sofisticadas que sean las técnicas de procesamiento de datos, un profesional especializado siempre es necesario (Zendesk, 2023).

3.2.1.6 Técnicas de minería de datos

Holdsworth (2024), afirma que existen varias técnicas de minería de datos (párr. 1), sosteniendo que las que se detallan a continuación son algunos de los tipos más populares:

- Reglas de la asociación: método basado en reglas si/entonces para encontrar relaciones entre variables en un conjunto de datos. La fuerza de las relaciones se mide por el apoyo y la confianza. El nivel de confianza se basa en la frecuencia con la que las sentencias si o entonces son verdaderas.
- Clasificación: se predefinen clases de objetos, de acuerdo a los requerimientos de la empresa, con definiciones de los atributos que los objetos tienen en común, permitiendo agrupar los datos subyacentes para agilizar su análisis.
- Agrupación: asociada con la clasificación, la agrupación responde a similitudes, pero también proporciona más agrupaciones basadas en diferencias.
- Árbol de decisión: utiliza el análisis de clasificación o regresión para clasificar o predecir posibles resultados en función de un conjunto de decisiones.
- Vecino K más cercano (KNN): conocido como algoritmo KNN, es algoritmo no paramétrico que clasifica los puntos de datos según su proximidad y asociación con otros datos disponibles.

- Redes neuronales: aplicadas en algoritmos de deep learning, redes neuronales procesan datos por medio de los algoritmos de conectividad, donde cada nodo se compone de entradas, ponderaciones, un sesgo y una salida.
- Análisis predictivo: combinar la minería de datos con técnicas de modelado estadístico y machine learning, se pueden analizar datos históricos utilizando el análisis predictivo para crear modelos gráficos o matemáticos destinados a identificar patrones, prever acontecimientos y resultados futuros e identificar riesgos y oportunidades.
- Análisis de regresión: descubre relaciones en los datos mediante la predicción de resultados basados en variables predeterminadas, pudiendo incluir árboles de decisión y regresión lineal multivariada (Holdsworth, 2024).

3.2.1.7 Seguridad de la minería de datos

La seguridad de la minería de datos de acuerdo a Brook (2024), presenta diversas técnicas para el tratamiento y almacenamiento de datos en una organización (párr. 1), incluyendo a:

- La anonimizarían de datos: asegura que los datos permanezcan confidenciales al hacerlo anónimo, por medio de la eliminación de toda la información identificable.
- La encriptación de datos: El cifrado es un método en el que los datos se convierten en un código para evitar el acceso no autorizado.
- Enmascaramiento de datos: los datos están enmascarados u ocultos, donde sólo las personas autorizadas pueden acceder a los datos originales´, y los datos reales se almacenan en forma o formato enmascarado.

- Infraestructura de seguridad: cuando el sistema de minería de datos es inherentemente seguro, puede añadir una capa de seguridad, presentando características de seguridad para el control de acceso basado en roles y transmisión segura de datos.
- Control de acceso: Esta técnica impide el acceso no autorizado a la base de datos, solamente los usuarios autorizados pueden realizarlo de acuerdo con la política de control de acceso.
- Rutas de Auditoría: Consiste en el monitoreo y mantenimiento de un registro de todos los accesos de datos, ayudan a rastrear quién accedió a los datos, cuándo y qué cambios se hicieron.
- Sanización de Datos: elimina la información sensible del conjunto de datos antes de compartir o publicar.
- Privacidad-Preservación; oculta datos sensibles o protege la privacidad mediante la introducción de ruidos, agregación, intercambio de datos o la generación de datos sintéticos.
- Detección de Amenaza Temprano: utiliza para encontrar patrones y anomalías dentro del tráfico de redes, detectando oportunamente las amenazas, potencialmente deteniéndolas antes de causar daños mayores.
- Vigilancia de la Red Precis: ayuda a identificar patrones de actividad o comportamientos inusuales que los métodos tradicionales de detección de amenazas podrían perder, como múltiples intentos de inicio de sesión de una dirección IP inusual.

- Reducción de falsos positivos: mejora la precisión de las detecciones de amenazas aprendiendo de grandes conjuntos de datos, reduciendo así el número de falsos positivos y liberando recursos.
- Evaluación Integral del Riesgo: analizan grandes cantidades de datos, proporcionando una instantánea completa de vulnerabilidades y riesgos potenciales en una red.
- Análisis predictivo: detectan amenazas actuales y predecir futuras en base a patrones identificados.
- Detecte de amenazas internas: detectan comportamientos o actividades anormales dentro de la red, indicando posibles amenazas interna o violación de datos.
- Detecte fraude: siendo eficaz para identificar patrones típicos del fraude y, por lo tanto, puede ser utilizado para detectar intentos de fraude.
- Costo-Efectivo: puede automatizar el proceso de análisis y detección, reduciendo la necesidad de mano de obra adicional y los costos asociados (Brook, 2024).

3.2.1.8 Dimensiones de la minería de datos

Según Coppola (2023), las técnicas de minería de datos dependen sustancialmente a las necesidades de cada organización, las mismas que se pueden categorizar en tres dimensiones significativas, tales como: las dimensiones o técnicas descriptivas, las predictivas y las prescriptivas (párr. 1).

- Las descriptivas: en las que se pueden agrupar las siguientes técnicas:
  - Técnicas de asociación; basadas en funciones donde se llega a conocer cuáles son los productos o servicios que más salen en una determinada

época del año, por medio de la búsqueda de nuevos eventos o atributos relevantes comparados con los ya existentes dentro del negocio

- Técnicas de agrupamiento; los algoritmos detectan una regularidad en los datos y pueden asociarlos, siendo sencillo llevar a cabo un proceso de agrupamiento que englobe todos los resultados con el mismo comportamiento, valor o relevancia.

- Las predictivas; comprenden las siguientes técnicas:
  - Técnicas de clasificación; permite proyectar cómo puede verse en el futuro cierta información y hacer predicciones comerciales o estratégicas con base en ello, recopilando varios atributos en categorías significativas para la organización.
  - Técnicas de regresión; sirve para ubicar relaciones y calcular probabilidades con base en datos.
  - Técnicas de detección de anomalías; siendo capaz de detectar valores atípicos a través del rastreo o clasificación de datos, detectando irregularidades y predecir su resultado o las consecuencias, gracias al aprendizaje obtenido de otros casos similares.
  - Las prescriptivas; constituido por las técnicas:
  - Técnicas de automatización; establecen reglas o comandos dependiendo de los resultados del análisis de la información.
  - Técnicas de optimización; generan simulaciones para la toma de decisión frente al resultado de una analítica de los datos, por lo tanto, obtienen una mejor respuesta basada en casos anteriores.

3.2.2 Ciberseguridad

Para Kaspersky (2025), la ciberseguridad “es la práctica de defender los ordenadores, servidores, dispositivos móviles, sistemas electrónicos, redes y los datos de ataques maliciosos” (párr. 1).

De la misma manera la ciberseguridad “es la práctica de proteger equipos, redes, aplicaciones de software, sistemas críticos y datos de posibles amenazas digitales” (Amazon Web Services [AWS], 2024).

3.2.2.1 Principios de la ciberseguridad.

Los principios de la ciberseguridad “son un conjunto de protocolos, medidas y operaciones destinadas a reducir los riesgos ante los problemas informáticos, a detectar y prevenir amenazas y a garantizar la recuperación del sistema” (UNIR Formación Profesional [UNIRFP], 2023), para el cual es necesario aplicar los siguientes principios:

- Confidencialidad; la información solo debe ser conocida por las personas autorizadas para ello, donde ciertos datos o software solo pueden ser accesibles para las personas autorizadas. También se lo identifica como principio de privacidad.
- Integridad; se refiere a la preservación de los datos, se debe garantizar que los datos se guardan en un sitio seguro y que no deben ser manipulados por nadie.
- Disponibilidad; garantiza que los datos van a estar disponible para las personas autorizadas en cualquier momento que estas la precisen, así mismo implica que, en caso de producirse algún incidente informático, la información puede ser recuperada, pero siempre bajo condiciones de seguridad y de recuperabilidad ante un ciberataque.

- Autenticidad: controla que la información de la que dispone la organización debe ser legítima, siendo pilar de la ciberseguridad que garantiza que el autor de la información o del documento es auténtico (UNIR Formación Profesional [UNIRFP], 2023).

3.2.2.2 Objetivos de la ciberseguridad.

De acuerdo a las afirmaciones de Ortega (2025), los objetivos de la ciberseguridad son fundamentales para proteger la información en un mundo cada vez más digitalizado (párr. 1), de donde los objetivos que persigue son:

- Prevención ante amenazas y ataques: siendo la prevención es la primera línea de defensa en ciberseguridad y su objetivo principal es evitar que las amenazas y ataques cibernéticos ocurran, por intermedio de medidas proactivas que dificultan el acceso no autorizado y protegen los datos sensibles, tales como los firewalls y sistemas de control de acceso, la encriptación de datos y el desarrollo y ejecución de políticas de seguridad.
- Detección temprana de ciberamenazas: cuando las medidas preventivas no son suficientes y que algunas amenazas logren superar las defensas iniciales, es donde entra a jugar la detección; cuya finalidad de la detección es identificar y reconocer rápidamente cualquier actividad sospechosa o anómala que pueda indicar la presencia de un ataque, entre las cuales se encuentra el sistema de detección de intrusiones (IDS), el análisis de logs y monitoreo continuo, y las Alertas y notificaciones clave (Ortega, 2025).
- Recuperación: se centra en restablecer las operaciones normales de la organización tras un ataque cibernético, llegando a minimizar las interrupciones y garantizar que la organización pueda continuar funcionando, incluso después de un incidente de seguridad. Las estrategias de recuperación son los planes de contingencia y recuperación

ante desastres, hacer backups regulares y realizar evaluaciones post-incidentes (Ortega, 2025).

- Confidencialidad de la información: asegurar que los datos sensibles solo sean accesibles para las personas o sistemas autorizados. La confidencialidad se protege mediante el uso de técnicas como la encriptación, el control de acceso y la autenticación.
- Integridad de los datos: es asegurarse de que la información no sea alterada de manera no autorizada. Esto implica proteger los datos contra modificaciones maliciosas o accidentales que puedan comprometer su exactitud y veracidad.
- Disponibilidad: asegura que los sistemas y datos estén accesibles para los usuarios autorizados cuando los necesiten, de donde la disponibilidad se ve amenazada por ataques como DDoS (Distributed Denial of Service), que buscan interrumpir el acceso a servicios y sistemas.
- Autenticidad: se refiere a la comprobación de que los usuarios, dispositivos, y datos son genuinos, asegurando que las identidades digitales sean verdaderas y que la información provenga de fuentes legítimas (Ortega, 2025).

3.2.2.3 Importancia de la ciberseguridad.

De las consideraciones de Toledo (2025), las organizaciones, en mayor o menor medida, se encuentran inmersas en procesos de transformación digital, donde todos comparten el mismo medio de transporte: Internet, y eso facilita la labor a los ciberdelincuentes, de donde las redes y sistemas son más vulnerables y se encuentren menos protegidos; incluso hay quien afirma que “todo es hackeable” (párr. 1), por cuando la ciberseguridad es sustancial para:

- Preservar los datos.

- Proteger datos ante manipulaciones.
- Proteger el acceso a ellos.
- Proteger la operatividad de sistemas y su integridad.
- Evitar la instalación de espías o roben (datos, sonidos e imágenes).
- Evitar caballos de troya, o la activación de puertas traseras, que permitan tomar el control de nuestros sistemas.
- Protección de dispositivos personales (portátiles, móviles) ante pérdidas y robos.
- Seguridad de la infraestructura TIC.
- Cuidar el valor de la información que publica en la red.
- Garantizar el uso de la Internet (Toledo, 2025).

Existen cuatro principios fundamentales de la ciberseguridad (Ortega, 2024), siendo:

La integridad, la confidencialidad, la disponibilidad y la autenticación, las mismas permiten:

- Promover en los usuarios la capacidad de prevención, detección, reacción, análisis, recuperación, respuesta, investigación y coordinación ante ciberataques de cualquier tipo.
- Impulsar la seguridad de los sistemas de información tanto de usuarios independientes como de empresas.
- Informar a los usuarios sobre los riesgos del ciberespacio.
- Participar en la mejora de la ciberseguridad a nivel internacional.
- Compartir y mantener los conocimientos, las habilidades y las capacidades tecnológicas que se requieren para ejercer la ciberseguridad (Ortega, 2024).

3.2.2.4 Tipos de ciberseguridad.

Zendesk (2024), indica que la protección informática abarca diversos enfoques y especialidades (párr. 1), existiendo diferentes tipos de ciberseguridad que desempeña un papel crucial en la protección digital y defensa integral de la empresa:

- Ciberseguridad de aplicaciones: engloba vulnerabilidades de software y garantiza una autenticación segura, así como contra ataques de scripts maliciosos o inyecciones de código, que insertan información con el objetivo de comprometer el funcionamiento normal de las apps.
- Ciberseguridad de la nube: brinda seguridad de los entornos de cloud computing, protegiendo datos y aplicaciones almacenadas en forma remota, así como la seguridad de las conexiones entre sistemas locales y la nube.
- Ciberseguridad móvil: protección de dispositivos como teléfonos inteligentes y tabletas contra malware, ataques de aplicaciones móviles y vulnerabilidades específicas de este tipo de plataformas.
- Ciberseguridad de infraestructura crítica: seguridad de sistemas que son fundamentales para el funcionamiento de una sociedad, como redes eléctricas, sistemas de agua y de transporte.
- Ciberseguridad de internet de las cosas: protege dispositivos conectados a internet, como cámaras de seguridad, electrodomésticos inteligentes y dispositivos médicos.
- Ciberseguridad financiera: enfocado a brindar seguridad a las digitalizaciones de las operaciones bancarias y financieras, protegiendo contra el robo de datos financieros y fraudes.

- Ciberseguridad de red: proteger la confidencialidad de las redes informáticas, contra amenazas como intrusiones, vulnerabilidades y ataques de denegación de servicio (Zendesk, 2024).

3.2.2.5 Aplicaciones de la ciberseguridad.

La ciberseguridad se aplica en diferentes contextos, desde los negocios hasta la informática móvil (Kaspersky, 2025), y presenta algunas categorías comunes:

- La seguridad de red: protege una red informática de los intrusos, ya sean atacantes dirigidos o malware oportunista.
- La seguridad de las aplicaciones: mantiene el software y los dispositivos libres de amenazas, de donde la seguridad eficaz comienza en la etapa de diseño, mucho antes de la implementación de un software o dispositivo
- La seguridad de la información: protege la integridad y la privacidad de los datos, tanto en el almacenamiento como en el tránsito.
- La seguridad operativa: conforma los procesos y decisiones para manejar y proteger los recursos de datos, detectando los permisos que tienen los usuarios para acceder a una red y los procedimientos que determinan cómo y dónde pueden almacenarse o compartirse los datos.
- La recuperación ante desastres y la continuidad del negocio: sostiene a la forma en que una empresa responde a un incidente de ciberseguridad o a cualquier otro evento que cause que se detengan sus operaciones o se pierdan datos.
- La capacitación del usuario final: aborda a las personas, observando si se incumplen las buenas prácticas de seguridad, cualquier persona puede introducir

accidentalmente un virus en un sistema que de otro modo sería seguro (Kaspersky, 2025).

3.2.2.6 Medidas preventivas correctivas en ciberseguridad.

Para Zendesk (2024), las medidas preventivas correctivas en ciberseguridad, adopta un enfoque integral, para abordar los diferentes riesgos de ataques cibernéticos (párr. 1), que incluye el uso de:

- Firewalls: proporciona barreras de seguridad que controlan el tráfico entre redes y permiten o bloquean la comunicación según reglas predefinidas, siendo la primera defensa contra accesos no autorizados.
- Software antivirus: identifican, previenen y eliminan programas maliciosos, pudiendo escanear archivos y actividades en tiempo real para defender de forma proactiva contra amenazas.
- Herramientas de detección de intrusiones: monitorean y analizan el tráfico de red en busca de patrones anómalos, llegando a detectar y alertar sobre actividades sospechosas.
- Cifrado: garantiza la confidencialidad al convertir los datos en un formato ilegible para quienes no tienen la clave, siendo esencial en la transmisión para evitar la exposición a interceptaciones.
- Políticas de seguridad: determinan reglas de acceso, contraseñas seguras y protocolos para el manejo de información, así mismo orientan el comportamiento de usuarios y aseguran consistencia en medidas de seguridad.

- Concientización y la capacitación: aspectos críticos para prevenir ataques como el phishing, donde la participación activa de las personas puede ser la primera línea de defensa (Zendesk, 2024).

3.2.2.7 Dimensiones de la ciberseguridad.

De las manifestaciones realizadas por Loranca (2021), se rescata, que la ciberseguridad es algo que se ha vuelto recurrente y que muchas veces no se profundiza en cuanto a su significancia (párr. 1), señalando que se debe considerar tres dimensiones para establecer y evaluar la información de seguridad:

- Principio de la seguridad: constituida por la triada: a) confidencialidad, que busca prevenir la divulgación no autorizada de información, b) integridad; busca la precisión, uniformidad y confiabilidad de la información, y c) disponibilidad; garantiza el acceso a la información siempre que se lo requiera.
- Estados de la información: sirve para identificar los estados de información de datos, tomando en cuenta los tres estados de los datos; en tránsito, almacenado y en proceso.
- Contramedidas: permite identificar las habilidades de los profesionales de las tecnologías de la información, debiendo definir las estrategias y aplicar las herramientas que se utilizaran, tales como: a) tecnologías, dispositivos y productos que ayuden a proteger, b) las políticas a ser establecidas, procedimientos, y continuidad de las prácticas adecuadas, y c) actualización de forma permanente del conocimiento exigido para hacer frente a nuevas amenazas (Loranca (2021).

3.3 Definición de términos

Minería de datos

"Es una técnica asistida por computadora que se utiliza en los análisis para procesar y explorar grandes conjuntos de datos" (Amazon Web Services [AWS], 2023).

Ciberseguridad

"Es la práctica de proteger equipos, redes, aplicaciones de software, sistemas críticos y datos de posibles amenazas digitales" (Amazon Web Services [AWS], 2024).

Anonimizarían de los datos

"Es un proceso clave para la protección de la privacidad y consiste en convertir los datos en anónimos mediante técnicas que reduzcan el riesgo de identificación de las personas" (Sociedad Mercantil Estatal para la Gestión de la Innovación y las Tecnologías Turísticas [SEGITTUR], 2023).

Aprendizaje automático o machine learning

"Es una función de la inteligencia artificial que se vale del procesamiento de los datos y del empleo de algoritmos para posibilitar a los equipos informáticos aprender de forma automática y similar a como lo hace un humano" (SEGITTUR, 2023).

Fuga de datos

"Es la pérdida de confidencialidad de una información, bien sea por una brecha de seguridad interna, por un error o descuido humano o por el efecto de un ataque informático" (SEGITTUR, 2023).

Inteligencia artificial

"Es la habilidad de una máquina para imitar el funcionamiento de la mente humana con acciones como el razonamiento, el aprendizaje, la creatividad o la planificación" (SEGITTUR, 2023).

IoT o internet de las cosas

"Se conoce así a la conexión de los objetos físicos a Internet para la transmisión e intercambio de datos" (SEGITTUR, 2023).

Algoritmos de minería de datos

"Son conjuntos de cálculos y heurísticas que permiten crear modelos a partir de datos".

Data Warehouse (Almacén de datos)

"Un almacén de datos es un sistema que se utiliza para hacer un análisis rápido de las tendencias comerciales utilizando datos de muchas fuentes" (Kharkovyna, 2020).

Data Science (Ciencia de los datos)

"Es las manipulaciones con datos y convertir datos desordenados y dispares en material comprensible" (Kharkovyna, 2020).

Selección de Datos

"Identificar y seleccionar un conjunto de datos adecuado para el análisis" (Tecnologiabi, 2025)

Pre-procesamiento

"Limpiar los datos eliminando inconsistencias, datos faltantes o irrelevantes" (Tecnologiabi, 2025).

Transformación

"Convertir los datos en un formato adecuado para la minería de datos".

Evaluación

"Validar y evaluar los patrones y resultados obtenidos".

Presentación de Resultados

Interpretar y comunicar los hallazgos de manera comprensible (Tecnologiabi, 2025)

Amenaza

"Circunstancia desfavorable que puede ocurrir y que cuando sucede tiene consecuencias negativas sobre los activos provocando su indisponibilidad, funcionamiento incorrecto o pérdida de valor" (Instituto Nacional de Ciberseguridad [INCIBE], 2021).

Análisis de riesgos

"Es un proceso que comprende la identificación de activos de información, sus vulnerabilidades y las amenazas a los que se encuentran expuestos, así como la probabilidad de ocurrencia y el impacto de las mismas, a fin de determinar los controles adecuados para tratar el riesgo" (INCIBE], 2021).

Análisis de vulnerabilidades

"Búsqueda y documentación de fallos, carencias o debilidades físicas (inundaciones, incendios, controles de acceso…) y lógicas (configuraciones, actualizaciones...) en un sistema informático, que puedan ser empleados por terceros con fines ilícitos, suponiendo un riesgo para la organización y los propios sistemas" (INCIBE], 2021).

Antivirus

"Software de protección para evitar que ejecutemos algún tipo de software malicioso en nuestro equipo que infecte al equipo" (INCIBE], 2021).

Ataque activo

"Tipo de ataque detectable que se caracteriza por la modificación del contenido de la información, así como de los recursos o funcionamiento del sistema, pudiendo causar daños a dicho sistema" (INCIBE], 2021).

Brecha de seguridad

"Violaciones de la seguridad que ocasionan la destrucción, pérdida o alteración accidental o deliberada de datos personales cuando están siendo transmitidos, están almacenados o son objeto de otros tratamientos" (INCIBE], 2021).

Ciberataque

"Intento deliberado de un ciberdelincuente de obtener acceso al sistema informático sin autorización sirviéndose de diferentes técnicas y vulnerabilidades para la realización de actividades con fines maliciosos, como el robo de información, extorsión del propietario o simplemente daños al sistema" (INCIBE], 2021).

Ciberdelincuente

"Persona que realiza actividades delictivas en la red contra personas o sistemas informáticos, provocando daños económicos o reputacionales mediante robo, filtrado de información, deterioro de software o hardware, fraude y extorsión" (INCIBE], 2021).

Contraseña

"Forma de autentificación de un usuario, a través de una clave secreta, para controlar el acceso a algún recurso o herramienta" (INCIBE], 2021).

Control de acceso

"Sistema de verificación que permite el acceso a un determinado recurso si la persona o entidad tiene los derechos necesarios para solicitarlo" (INCIBE], 2021).

Copia de seguridad

"Proceso mediante el cual se duplica la información existente de un soporte a otro, con el fin de poder recuperar los datos contenidos en caso de fallo del primer soporte de alojamiento" (INCIBE], 2021).

IV. Metodología

4.1 Tipo y nivel de investigación

4.1.1 Tipo de investigación

Por las características de las variables, el estudio es tipo básica, donde se pudo lograr datos de los fenómenos problemáticos, para fortalecer los conocimientos existentes y ampliar el manejo científico de la minería de datos y de la ciberseguridad en la organización educativa en estudio.

De donde la investigación tipo dogmática, de acuerdo a las manifestaciones de Muntané (2010), se originan en un marco teórico y permanece, a partir del cual se incrementa los conocimientos científicos, pero sin contrastar con ningún escenario práctico"

4.1.2 Nivel de investigación

Enmarcada en un nivel explicativo, por cuanto del dato logrado de la realidad se pudo determinar las posibles causas generadas por la minería de datos en la ciberseguridad, es decir, se buscó indagar la relación de causa y efecto entre las variables problemáticas en la zona objeto de investigación

Señalando que los estudios de nivel explicativo "pretenden establecer las causas de los sucesos o fenómenos que se estudian" (Hernández y colaboradores, 2014, p. 95).

4.2 Diseño de investigación

Parte de un diseño no experimental transeccional correlacional causal, en donde los datos de las respectivas variables no tuvieron ningún tipo de manipulación deliberada, fueron manejadas en su contexto real, logrando obtenerlas en un momento único y

establecer la relación de causalidad entre la minería de datos que genera en la ciberseguridad de la Casa Superior de Estudios objeto de investigación.

De donde el diseño no experimental en un estudio “son los que se ejecutan donde no existe presencia de manipulación de las variables, siendo observables para analizar su comportamiento” (Hernández y colaboradores, 2014)

Por otro lado, son transeccionales las investigaciones “cuando los datos se obtienen en un solo momento del ambiente en la que se encuentran las variables” (Hernández y colaboradores, 2014).

En la misma línea, Hernández et al. (2014), asientan que el estudio correlacional causal “llegan a describir las relaciones entre dos o más variables en un momento determinado en función de la relación causa-efecto” (p. 158).

4.3 Ámbito temporal y espacial

4.3.1 Ámbito temporal

Se puso en marcha el estudio a partir del mes de diciembre 2024 y culminó en mayo de 2025.

4.3.2 Ámbito espacial

El espacio geográfico para su ejecución fue la Universidad Tecnológica de los Andes de Abancay.

4.4 Población y muestra

4.4.1 Población

Para Hernández et al. (2014), señalan que la población “es un conjunto de sujetos con atributos particulares, encontrándose en un contexto específico de los cuales se obtendrán datos de forma general” (Hernández y colaboradores, 2014).

La población de estudio estuvo establecida por los talentos humanos que cumplen sus funciones en la dirección general de administración, la oficina de tecnologías de información, el centro de cómputo e informática, y la oficina de servicios académicos de la Universidad Tecnológica de los Andes, sede Abancay. Contando con 14 unidades muéstrales.

4.4.2 Muestra

Una muestra "es el subconjunto de casos del universo de los cuales se obtendrá información y representará la percepción de la población" (Hernández y colaboradores, 2014).

Para la estimación de la muestra se efectuado por el método no probabilístico, en vista que la población es pequeña, no se aplicó ninguna prueba probabilística para determinar el tamaño de la muestra. Además, para la selección de las unidades muestrales que participaran en el estudio, se aplicó el muestreo por conveniencia, en razón de las características específicas y comunes que contaban los colaboradores de las unidades orgánicas consideradas para la investigación. Por cuanto la muestra está conformada por 14 unidades muéstrales.

4.5 Técnicas e instrumentos para la recolección de datos

4.5.1 Técnica

Partiendo de la encuesta, la misma fue el medio por el cual se lograron los datos de las respectivas variables en estudio, sosteniendo de manera sistemática la medición de las dimensiones e indicadores de la minería de datos y la ciberseguridad, que permitieron identificar la causa y efecto entre las variables en la Universidad Tecnológica de los Andes, sede Abancay.

La encuesta permite “la recolección de datos que implica la formulación y administración de un conjunto de preguntas a una muestra representativa de individuos con el fin de recopilar información sobre una problemática en particular” (Anguita et al., 2003 citado por Blanchar y Martinez, 2025).

4.5.2 Instrumentos

El cuestionario, la misma permitió acopiar los datos a partir de las dimensiones e indicadores de las variables objeto de investigación, las que se aplicaron a los colaboradores de la universidad y que contaron con preguntas cerradas bajo una escala de medición de opción múltiple tipo Likert.

El cuestionario para la variable minería de datos, se aplicó las dimensiones: descriptivas, predictivas y prescriptivas con una escala de medición: 1: Muy inadecuado, 2: Inadecuado, 3: Regular, 4: Adecuado, y 5: Muy adecuado. Mientras para la variable ciberseguridad se manejó las dimensiones: principio de la seguridad, estados de la información y contramedidas bajo una escala de medición: 1: Muy inadecuado, 2: Inadecuado, 3: Regular, 4: Adecuado, y 5: Muy adecuado.

De acuerdo a Hernández et al. (2014), el cuestionario “es un conjunto de ítems orientadas a un, o más eventos que se desea medir”.

4.6 Validación y confiabilidad de los instrumentos

4.6.1 Validación

Cada instrumento fue validado por juicio de expertos, quienes determinaron la coherencia, pertinencia, claridad, etc. de las respectivas preguntas que contenían el cuestionario, antes de su aplicación.

4.6.2 Confiabilidad

Los instrumentos estuvieron sometidos a su fiabilidad una vez obtenida la información, la misma que se realizó por medio de la prueba alfa de Cronbach, la que determinó la consistencia de los datos alcanzados.

Tabla 1

Confiabilidad de los instrumentos

Variable	Alfa de Cronbach
Minería de datos	0.817
Ciberseguridad	0.753
Confiabilidad de variables	0.785

Nota: Nivel de la fiabilidad de los instrumentos

La presente tabla 1, contiene valores de alfa de Cronbach, donde el rango de confiabilidad de las variables es 0.785 encontrándose por encima del rango de 0.7 y muy cercano a la unidad (1), estableciendo una confiabilidad adecuada y que los datos alcanzados son significativos, fiables y consistentes para llegar a lograr los objetivos formulados y resolver el problema de la investigación.

4.7. Métodos y técnicas para la presentación y análisis de datos

La información obtenida se llegó a procesar por medio de la estadística descriptiva e inferencial, logrando construir las respectivas bases de datos de cada variable a partir de las cuales construir las tablas con las pertinentes frecuencias y porcentajes, así como las figuras de las dimensiones de la minería de datos y de la ciberseguridad, para luego efectuar el análisis, la interpretación y las discusiones de los resultados de forma efectiva y contrastarlos con investigaciones previas al estudio. Al final se efectuó la comprobación de las hipótesis, previamente se determinó la prueba de normalidad de los datos para observar el comportamiento del supuesto de normalidad o de no normalidad de los

mismos, para luego considerar el uso de los estadísticos paramétricos o no paramétricos para contrastación de la hipótesis de investigación. Todo el tratamiento de la investigación se realizará por el software SPSS v. 29 y Microsoft Excel.

V. Resultados Y Discusiones

5.1 Resultados descriptivos

5.1.1 Variable independiente: Minería de datos

Tabla 2

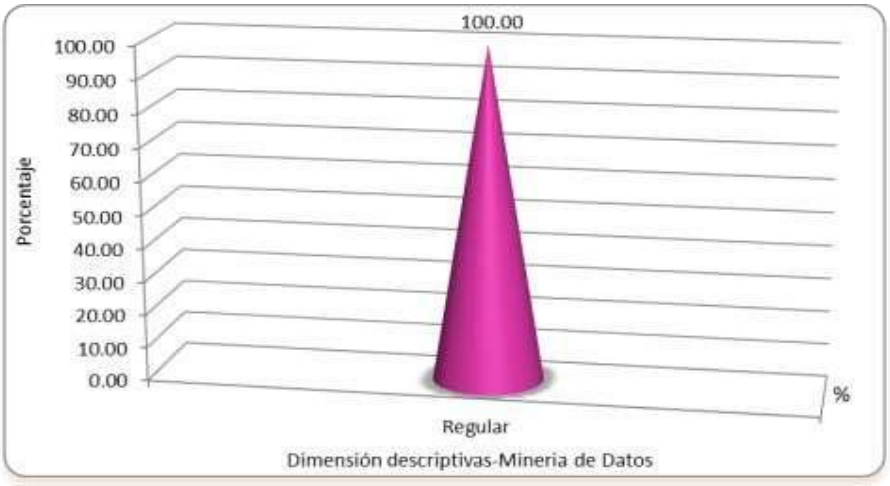
Situación descriptiva de la minería de datos

Afirmación	f	%	Porcentaje Acumulado
Regular	14	100.00	100
Total	14	100	

Nota: Naturaleza descriptiva de la minería de datos

Figura 1 Proporción descriptiva de la minería de datos

Proporción descriptiva de la minería de datos



Nota: Situación porcentual descriptiva de la minería de datos

Análisis e interpretación:

En los datos visualizados en la tabla 2 y figura 1, donde el 100% de los colaboradores participantes en el estudio manifestaron regular la situación descriptiva para convertir los datos en información entendible y concisa que facilite un análisis más avanzado para la toma de decisiones.

Tabla 3

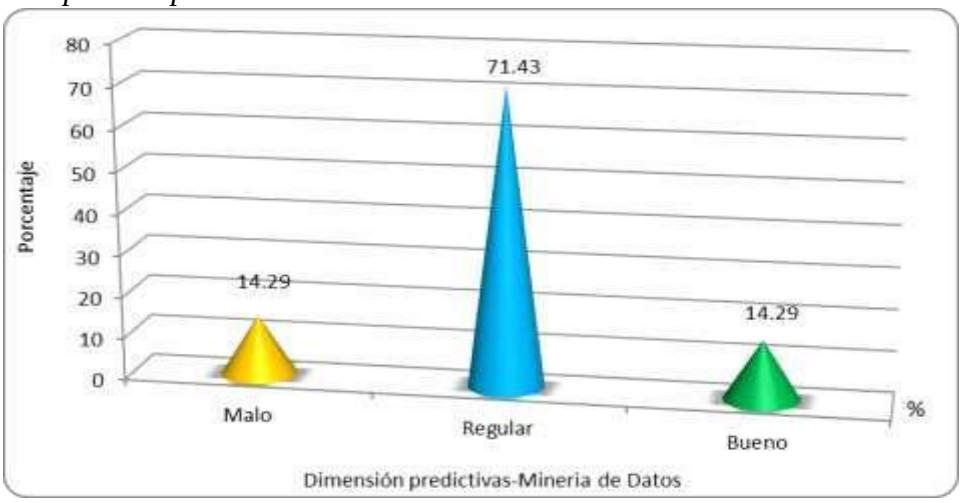
Situación predictiva de la minería de datos

Afirmación	f	%	Porcentaje Acumulado
Malo	2	14.29	14.29
Regular	10	71.43	85.72
Bueno	2	14.29	100
Total	14	100	

Nota: Naturaleza prescriptiva de la minería de datos

Figura 2

Proporción predictiva de la minería de datos



Nota: Situación porcentual predictiva de la minería de datos

Análisis e interpretación:

Las unidades de análisis seleccionadas para el estudio proporcionaron información y que se encuentran en la tabla y figura precedente, donde el 71.43% indicaron regular, seguido del 14.29% que afirmaron bueno y malo respectivamente los procesos predictivos en el manejo de los datos históricos y actuales académicos y administrativos de la universidad para anticiparse a los sucesos futuros o comportamientos probables que puedan presentar.

Tabla 4

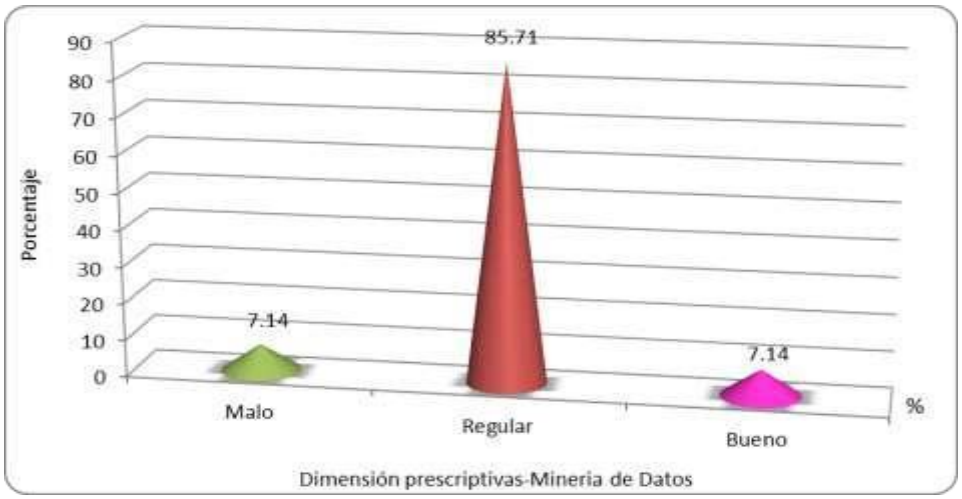
Situación prescriptiva de la minería de datos

Afirmación	f	%	Porcentaje Acumulado
Malo	1	7.14	7.14
Regular	12	85.71	92.85
Bueno	1	7.14	100
Total	14	99	

Nota: Naturaleza prescriptiva de la minería de datos

Figura 3

Proporción prescriptiva de la minería de datos



Nota: Situación porcentual prescriptiva de la minería de datos

Análisis e interpretación:

Distinguiendo la tabla 4 y figura que antecede, se aprecia que el 85.71% de las unidades de análisis sostuvieron regular, además del 7.14% que argumentaron bueno y malo respectivamente las fases prescriptivas para optimizar resultados, decisiones y estrategias concretas basados en los datos y la optimización de los mismos para evitar riesgos futuros en la universidad.

Tabla 5

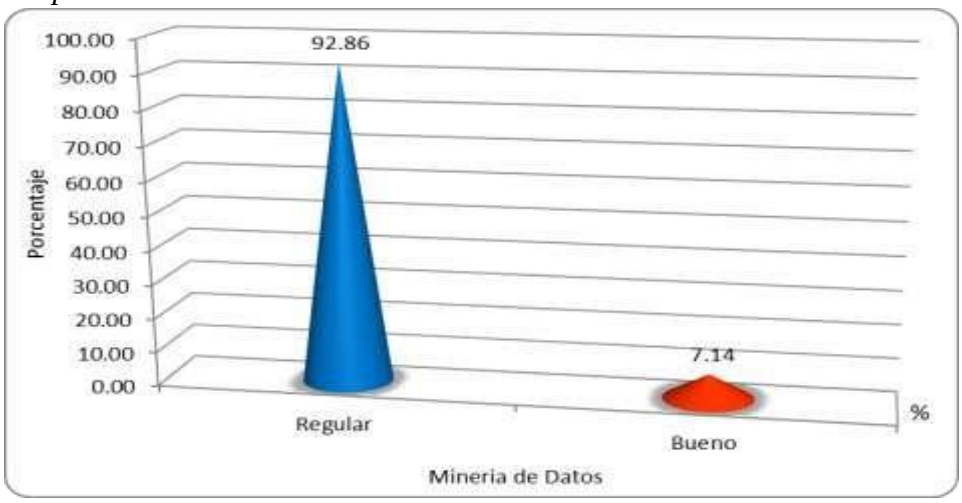
Situación de la minería de datos

Afirmación	f	%	Porcentaje Acumulado
Regular	13	92.86	92.86
Bueno	1	7.14	100
Total	14	100	

Nota: Naturaleza de la minería de datos

Figura 4

Proporción de la minería de datos



Nota: Situación porcentual de la minería de datos

Análisis e interpretación:

Al ver la información de la tabla precedente y figura 4, se emite que un 92.86% del talento humano de las unidades universitarias dijeron regular y sólo el 7.14% asentaron bueno el tratamiento básico y cotidiano de la minería de datos, donde la universidad recolecta, limpia, organiza y consulta los datos o información para apoyar sus operaciones académicas y administrativas diarias, sin que exista modelos avanzados de análisis u optimización de los mismos.

5.1.2 Variable ciberseguridad (Variable dependiente)

Tabla 6

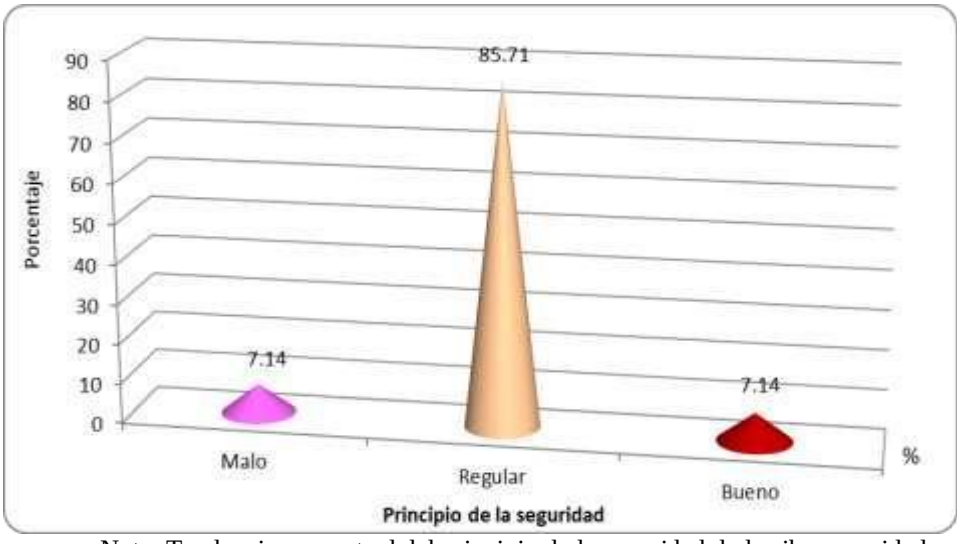
Condiciones del principio de la seguridad de la ciberseguridad

Afirmación	f	%	Porcentaje Acumulado
Malo	1	7.14	7.14
Regular	12	85.71	92.85
Bueno	1	7.14	100
Total	14	100	

Nota: Constitución del principio de la seguridad de la ciberseguridad

Figura 5

Disposición porcentual del principio de la seguridad de la ciberseguridad



Nota: Tendencia porcentual del principio de la seguridad de la ciberseguridad

Análisis e interpretación:

En la tabla 6 y figura 5, se aprecia que el 85.71% de las unidades de análisis indicaron regular, además del 7.14% dijeron bueno y malo de manera respectiva los principios de la seguridad para garantizar que los sistemas, datos y redes tecnológicas estén protegidos frente a amenazas intrínsecas y extrínsecas, y asegurando que la información pueda ser accedida, utilizada y compartida de manera segura y confiable.

Tabla 7

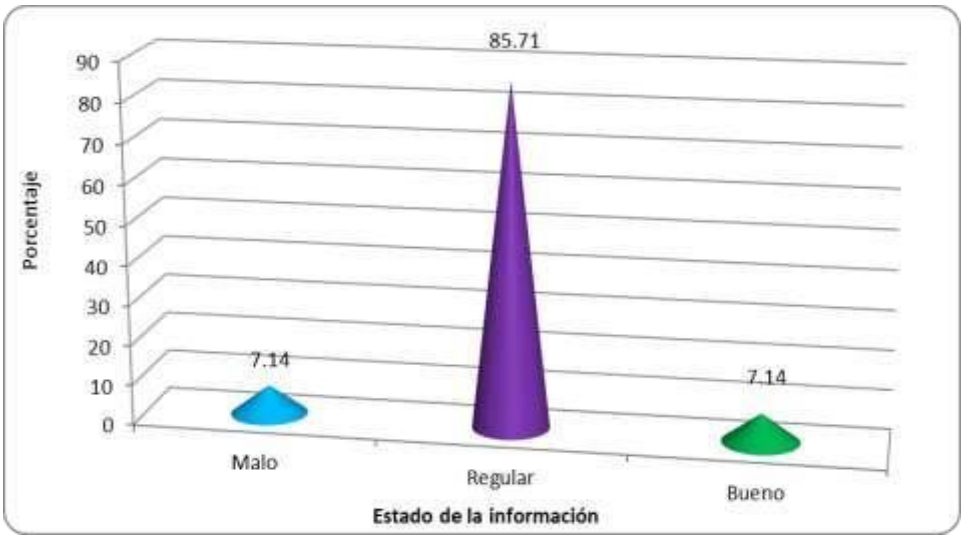
Condiciones del estado de la información de la ciberseguridad

Afirmación	f	%	Porcentaje Acumulado
Malo	1	7.14	7.14
Regular	12	85.71	92.85
Bueno	1	7.14	100
Total	14	100	

Nota: Constitución del estado de la información de la ciberseguridad

Figura 6

Disposición porcentual del estado de la información de la ciberseguridad



Nota: Tendencia porcentual del estado de la información de la ciberseguridad

Análisis e interpretación:

En la presenta tabla 7 y figura que antecede se visualiza que el 85.71% de los responsables de tecnología de información señalaron regular, luego un 7.14% indicaron bueno y malo respectivamente, el estado de la información actualmente relacionada a la protección, gestión y exposición del entorno digital frente a posibles riesgos y amenazas.

Tabla 8

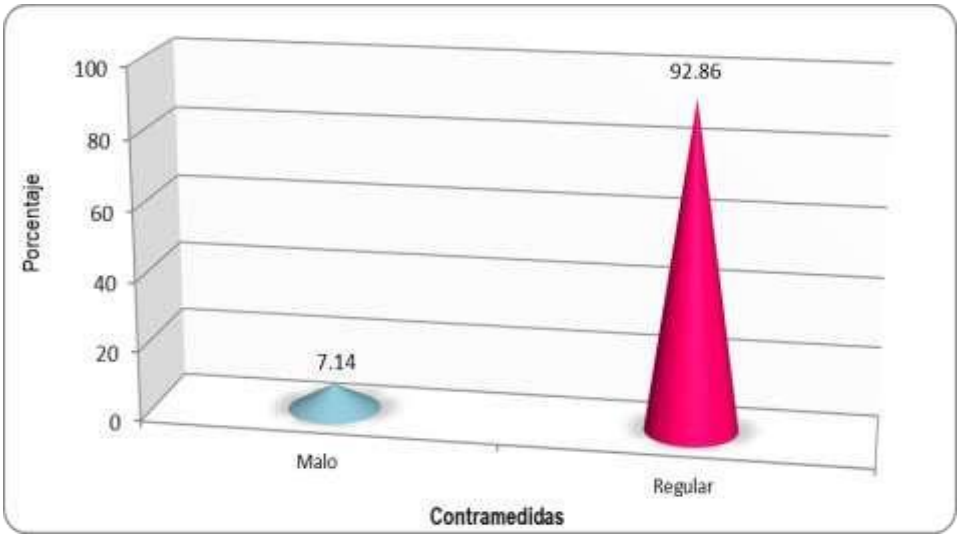
Condiciones de las contramedidas de la ciberseguridad

Afirmación	f	%	Porcentaje Acumulado
Malo	1	7.14	7.14
Regular	13	92.86	100
Total	14	99	

Nota: Constitución de las contramedidas de la ciberseguridad

Figura 7

Disposición porcentual de las contramedidas de la ciberseguridad



Nota: Tendencia porcentual de las contramedidas de la ciberseguridad

Análisis e interpretación:

Al observar la tabla y figura anterior se desprende que el 92.86% de las unidades de análisis subrayaron regular y tan sólo el 7.14% revelaron de malo el estado actual de las contramedidas basadas en las herramientas, estrategias y prácticas manejadas para prevenir, detectar, responder y mitigar ciberataques en la Casa Superior de Estudios objeto de investigación.

Tabla 9

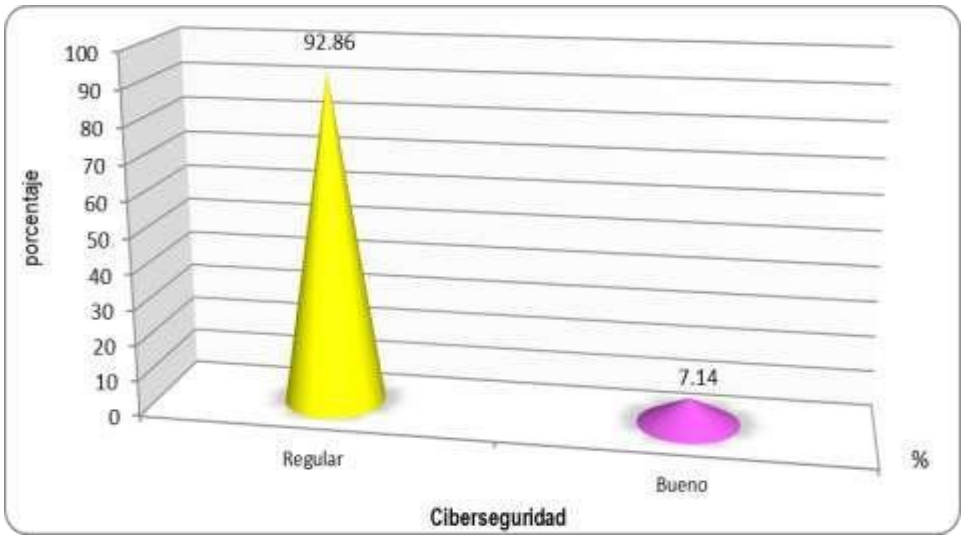
Condiciones de la ciberseguridad

Afirmación	f	%	Porcentaje Acumulado
Regular	13	92.86	92.86
Bueno	1	7.14	100
Total	14	100	

Nota: Constitución de la ciberseguridad

Figura 8 Disposición porcentual de la ciberseguridad

Disposición porcentual de la ciberseguridad



Nota: Tendencia porcentual de la ciberseguridad

Análisis e interpretación:

Partiendo de los datos visualizados en la tabla 9 y la respectiva figura 8 se aprecia que el 92.86% de colaboradores universitarios especificaron regular y al final un 7.14% puntualizaron de bueno la ciberseguridad de los sistemas e información frente a las amenazas digitales, para los cuales no se cuenta con sistemas avanzados de monitoreo, ni cifrado robustos o Zero Trust (modelo de seguridad en red).

5.2 Contrastación de hipótesis de estudio

a)Prueba de normalidad de los datos

Tabla 10

Estadístico de Shapiro-Wilk

Variables	Estadístico	gl	Sig.
Minería de datos	.843	14	.018
Ciberseguridad	.896	14	.100

Nota. Nivel de distribución de normalidad de los datos obtenidos

Análisis e interpretación:

La tabla 10 presenta información de la prueba de normalidad de los datos por ShapiroWilk, en razón que la muestra es menor a 50 unidades de análisis, donde la variable minería de datos presenta un sig. 0.018 siendo inferior al error 0.05 que señala que los datos presentan un supuesto de distribución no normal, mientras la variable ciberseguridad muestra un sig. 0.100 que es superior al error 0.05 determinando que los datos presentan un supuesto de distribución normal, es así que existiendo diferencias de comportamiento entre las variables se considera para la validación de las hipótesis de investigación la prueba no paramétrica de Wilcoxon.

b) Contrastación de hipótesis general

Tabla 11

Incidencia de la minería de datos en la ciberseguridad

Hipótesis nula	Prueba	Sig.	Decisión
La mediana de las diferencias entre Minería de datos y Ciberseguridad es igual a 0.	Prueba de Wilcoxon de los rangos con signo para muestras relacionadas	,001	Rechace la hipótesis nula.

Nota: Validación de la hipótesis general

Tabla 12

Análisis e interpretación

Hipótesis estadística	**Ho**: No existe un nivel de incidencia evidente de la minería de datos en la ciberseguridad de la Universidad Tecnológica de los Andes, Abancay 2025. **Ha**: Sí existe un nivel de incidencia evidente de la minería de datos en la ciberseguridad de la Universidad Tecnológica de los Andes, Abancay 2025.
Nivel de significancia	$\alpha = 0.05$
Valor p calculado	$p = 0.001$
Conclusión	Demostrando que $p < 0.05$, se rechaza hipótesis nula (Ho) y se acepta la hipótesis alterna (Ha), llegando a concluir que, sí existe un nivel de incidencia evidente de la minería de datos en la ciberseguridad de la Universidad Tecnológica de los Andes, Abancay 2025.

Nota: Incidencia de la minería de datos en la ciberseguridad

c) Contrastación de hipótesis específicas

Tabla 13

Incidencia de la minería de datos en los principios de la seguridad

Hipótesis nula	Prueba	Sig.	Decisión
La mediana de las diferencias entre Minería de datos y Principio de la seguridad es igual a 0.	Prueba de Wilcoxon de los rangos con signo para muestras relacionadas	,001	Rechace la hipótesis nula.

Nota: Validación de la primera hipótesis específica

Tabla 14

Análisis e interpretación

Hipótesis estadística	**Ho**: No existe un nivel de incidencia evidente de la minería de datos en los principios de seguridad de la ciberseguridad de la Universidad Tecnológica de los Andes, Abancay 2025. **Ha**: Sí existe un nivel de incidencia evidente de la minería de datos en los principios de seguridad de la ciberseguridad de la Universidad Tecnológica de los Andes, Abancay 2025.
Nivel de significancia	$\alpha = 0.05$

Valor p calculado	p = 0.001
Conclusión	Puntualizando que p > 0.05, se rechaza hipótesis nula (Ho) y se acepta la hipótesis alterna (Ha), llegando a concluir que, sí existe un nivel de incidencia evidente de la minería de datos en los principios de seguridad de la ciberseguridad de la Universidad Tecnológica de los Andes, Abancay 2025.

Nota: Incidencia de la minería de datos en los principios de la seguridad

Tabla 15

Incidencia de la minería de datos en el estado de la información

Hipótesis nula	Prueba	Sig.	Decisión
La mediana de las diferencias entre Minería de datos y Estado de la información es igual a 0.	Prueba de Wilcoxon de los rangos con signo para muestras relacionadas	,001	Rechace la hipótesis nula.

Nota: Validación de la segunda hipótesis específica

Tabla 16

Análisis e interpretación

Hipótesis estadística	**Ho**: No existe un nivel de incidencia evidente de la minería de datos en los estados de la información de la ciberseguridad de la Universidad Tecnológica de los Andes, Abancay 2025. **Ha**: Sí existe un nivel de incidencia evidente de la minería de datos en los estados de la información de la ciberseguridad de la Universidad Tecnológica de los Andes, Abancay 2025.
Nivel de significancia	$\alpha = 0.05$
Valor p calculado	p = 0.001
Conclusión	Puntualizando que $p > 0.05$, se rechaza hipótesis nula (Ho) y se acepta la hipótesis alterna (Ha), llegando a concluir que, sí existe un nivel de incidencia evidente de la minería de datos en los estados de la información de la ciberseguridad de la Universidad Tecnológica de los Andes, Abancay 2025.

Nota: Incidencia de la minería de datos en los estados de la información

Tabla 17

Incidencia de la minería de datos en las contramedidas

Hipótesis nula	Prueba	Sig.	Decisión
La mediana de las diferencias entre Minería de datos y Contramedidas es igual a 0.	Prueba de Wilcoxon de los rangos con signo para muestras relacionadas	,001	Rechace la hipótesis nula.

Nota: Validación de la tercera hipótesis específica

Tabla 18

Análisis e interpretación

Hipótesis estadística	**Ho:** No existe un nivel de incidencia evidente de la minería de datos en las contramedidas de la ciberseguridad de la Universidad Tecnológica de los Andes, Abancay 2025. **Ha:** Sí existe un nivel de incidencia evidente de la minería de datos en las contramedidas de la ciberseguridad de la Universidad Tecnológica de los Andes, Abancay 2025.
Nivel de significancia	$\alpha = 0.05$
Valor p calculado	$p = 0.001$
Conclusión	Puntualizando que $p > 0.05$, se rechaza hipótesis nula (Ho) y se acepta la hipótesis alterna (Ha), llegando a concluir que, sí existe un nivel de incidencia evidente de la minería de datos en las contramedidas de la ciberseguridad de la Universidad Tecnológica de los Andes, Abancay 2025.

Nota: Incidencia de la minería de datos en las contramedidas

5.3 Discusiones

Los hallazgos del estudio permiten desarrollar las pertinentes discusiones con resultados de estudio previos, logrando contrastar la hipótesis de investigación en base a la prueba no paramétrica de Wilcoxon cuyo *p*-valor alcanzado 0.001 es menor a la significancia 0.05 que puntualiza de la existencia de un nivel de incidencia evidente de la minería de datos en la ciberseguridad de la Universidad Tecnológica de los Andes, Abancay 2025, realidad esgrimida por el 92.86% del talento humano de las unidades académicas y administrativas de la universidad que sostuvieron regular el tratamiento básico y cotidiano de la minería de datos,

asentados en la recolecta, limpia, organiza y consulta de información para apoyar las distintas operaciones académicas y administrativas diarias; ambiente que impacta en el 92.86% que informaron de regular la ciberseguridad de los datos, evidenciando que la tecnología puede impactar de manera significativa en la detección, prevención y respuesta ante amenazas digitales, toda vez que la minería de datos en virtud a los modelos descriptivos, predictivos y prescriptivos incide de forma evidente en vista que fortalece la capacidad de la UTEA para proteger sus sistemas e información, detectar ataques y gestionar riesgos, convirtiéndose en una herramienta clave dentro de la estrategia de manejo de los principios de seguridad, el estado de la información y las contramedidas de la ciberseguridad, para que los sistemas de información de la universidad mantengan su confiabilidad, funcionalidad y rendimiento, y garanticen la integridad, precisión, coherencia y exactitud de los datos, así como la puesta en práctica de las habilidades, capacidades y profesionalismo del talento humano que se dedica a la detección de amenazas, de patrones no autorizados, comportamientos y respuestas a incidentes que proceden de ataques internos y/o externos dentro de sus estrategias de ciberseguridad universitario.

Resultados logrados en el estudio donde se encuentran concordantes a los logrados por Vallejo y Tenelanda (2022), quienes indicaron que, la minería de datos es el proceso de ahondar en los datos para detectar patrones y relaciones ocultos para la seguridad de la información. Además, a los hallazgos de Correa (2022), afirmando que, la ciberseguridad incide significativamente en el tratamiento de datos personales en las entidades. De la misma manera con los resultados de Dongo y Silva (2020), donde asientan que, la minería de datos es una herramienta que ayuda a mejorar la toma de decisiones en empresas del sector.

En el mismo contexto los hallazgos de la investigación también se encuentran asociados a los resultados logrados por Hernández (2020), en donde puntualiza que, el portal Web de la Universidad Andina del Cusco presenta vulnerabilidades de nivel medio/alto, tanto en

vulnerabilidades de diseño, implementación y uso. Así como a los resultados de Matilde (2023), en donde se puntualiza que, la ciberseguridad no es considerada de manera proactiva dentro de las organizaciones, y que es necesario seguir investigando por qué existe un déficit tan grande de profesionales especialistas en ciberseguridad. Al igual a los hallazgos de Torres (2022), quién confirma que, el talento humano es el factor más influyente en el tratamiento de minería de datos para detectar la ocurrencia de siniestros de tránsito. Además, se distingue que los resultados del estudio también se encuentran concordantes con los de Pozo (2022), señalando que, no existe un modelo de ciberseguridad específica que pueda dar seguridad a los datos del entorno actual, debiendo utilizarse un conjunto de estrategias, lineamientos y objetivos de carácter político para proteger el ciberespacio. Finalmente, las realidades presentadas están anidadas a los hallazgos de Jimenez (2022), manifestando que, en virtud a los diferentes escenarios donde las organizaciones se desarrollan, teniendo en cuenta la construcción de plataformas seguras al momento de planificar la estrategia de prevención o defensa de los servicios digitales, donde cada tecnología demanda servicios de protección diferente y gestión es especializada.

VI. Conclusiones

Primera: Se concluye que, la minería de datos incide evidentemente en la ciberseguridad de la Universidad Tecnológica de los Andes, Abancay 2025, asentada por Wilcoxon con un pvalor logrado 0.001 siendo menor al error 0.05; puntualizando que la minería de datos como técnica de análisis predictivo y descriptivo de procesamiento de grandes volúmenes de datos no sólo llegara a fortalecer la detección, prevención y respuesta de las amenazas digitales, sino que también permitirá identificar vulnerabilidades críticas de los sistemas informáticos, la reducción de incidentes como accesos no autorizados, los ataques a la red académica, la filtración de datos sensibles y el fortalecimiento de la infraestructura de la ciberseguridad universitaria.

Segunda: Así mismo se concluye que, existe un nivel de incidencia evidente de la minería de datos con los principios de seguridad de la ciberseguridad de la Universidad Tecnológica de los Andes, Abancay 2025, realidad esgrimida por Wilcoxon bajo un *p*-calculado 0.001 estando por debajo del error 0.05; toda vez que la minería de datos como técnica no sólo potencia la protección de confidencialidad, integridad, disponibilidad y trazabilidad de la información académica-administrativa universitaria, sino que se transforma en un elemento estratégico para cumplir de forma efectiva con los principios de la ciberseguridad de los activos tecnológicos y académicos de la UTEA.

Tercera: De otra parte se concluye que, sí existe un nivel de incidencia evidente de la minería de datos en los estados de la información de la ciberseguridad de la Universidad Tecnológica de los Andes, Abancay 2025, contexto señalado por Wilcoxon cuyo *p*-alcanzado 0.001 encontrándose por debajo del error 0.05; ya que la aplicación de la tecnología de la minería de datos permite identificar vulnerabilidades, evaluar riesgos y monitorear de forma continua el

estado de los datos de la UTEA, para optimizar la detección de amenazas, fortalecer los mecanismos de protección y facilitar la toma de decisiones en la mantención de un estado de ciberseguridad sólido.

Cuarta: Al final se concluye que, existe un nivel de incidencia evidente de la minería de datos en las contramedidas de la ciberseguridad de la Universidad Tecnológica de los Andes, Abancay 2025, basado a la prueba de Wilcoxon con un *p*-valor 0.001 que es inferior al error 0.05; donde la naturaleza de la técnica de la minería de datos permite identificar patrones de ataque, prever vulnerabilidades y mejorar las medidas preventivas y correctivas de seguridad de datos, consolidándose como un recurso significativo para robustecer las estrategias defensivas y garantizar un ambiente tecnológico digital más seguro y resiliente académico y administrativo universitario.

VII. Recomendaciones

Primera: A las autoridades y los responsables de tecnologías de información de la UTEA, deberán diseñar estrategias con miras a fortalecer la gestión de datos y seguridad, implementando procedimientos y políticas que normalicen el uso de las técnicas de minería de datos que garanticen el cumplimiento de los principios de confidencialidad, integridad y disponibilidad de la información académica y administrativa uteina, así como la detección temprana de patrones extraños o posibles amenazas, creando alertas anticipadas que permitirán advertir vulnerabilidades en los sistemas informáticos y contar con una eficiente ciberseguridad.

Segunda: A los responsables académicos y administrativos, así como a los de la oficina de tecnologías de información de la universidad, les compete implantar acciones base de capacitación orientadas tanto al talento humano técnico y administrativo, para los cuales deben desarrollar programas de formación en ciberseguridad y manejo responsable de datos, así como la adopción de herramientas de minería de datos con enfoque seguro y estrictos protocolos de anonimización, cifrado y que eviten filtraciones de datos e información sensible de docentes, discentes y personal no docente.

Tercera: Al talento humano de la oficina de tecnologías de información y centro de cómputo de la universidad, deben diseñar programas integrales de herramientas avanzadas de análisis predictivo e incorporar modelos de minería de datos que no solo permitan identificar amenazas actuales, sino que también anticipen posibles ataques cibernéticos (Malware, Phishing, Ransomware, Ataques de denegación de servicio “DDoS”, etc.) y brechas de seguridad, así como desarrollar mecanismos que registren quién o quiénes,

cómo y cuándo se accede o procesa información, reduciendo riesgos de manipulación o pérdida de datos de la UTEA.

Cuarta: Al ápice estratégico y responsables de tecnologías de información de la universidad, así como a los investigadores imbuidos en los fenómenos abordados, deberán impulsar, inducir, promover estudios y proyectos universitarios que utilicen minería de datos para fortalecer estrategias defensivas y de seguridad de datos, así como el rediseño de los protocolos internos de la universidad con el fin de integrar los resultados de minería de datos como insumo, de priorización de riesgos y fortalecer controles de acceso, autenticación y monitorio, reduciendo dependencia de soluciones externas y creando capacidades propias.

VIII. Referencias

Amazon Web Services (2024). ¿Qué es la ciberseguridad?. Internet, consultado el 18/01/2025 y disponible en: https://.amazon.com/es/what-is/cybersecurity/

Amazon Web Services (2023). ¿Qué es la minería de datos?. Internet, consultado el 12/01/2025 y disponible en: https://aws.amazon.com/es/what-is/data-mining/

Arizaca-Avalos, A. (2022). Aplicación de data mining para predecir el impacto de las transferencias económicas mineras en el desarrollo humano y pobreza del Perú. Internet, consultado el 22/12/2024 y disponible en: https://repositorio.unap.edu.pe/handle/20.500.14082/18339

Base Sur Digital (2024). De Datos a Dólares: Cómo la Minería de Datos Impulsará el Crecimiento Empresarial en 2024. Internet, consultado el 16/12/2024 y disponible en: https://basesur.ar/blog/de-datos-a-dolares-como-la-mineria-de-datos-impulsara elcrecimiento-empresarial-en-2024

Bismart (2024). ¿Qué es la minería de datos? Lo que todo empresario debería saber sobre data mining. Internet, consultado el 16/12/2024 y disponible en: https://blog.bismart.com/que-es-la-mineria-de-datos-para-que-sirve-negocios

Business Technology Platform (2023). ¿Qué es la minería de datos?. Internet, consultado el 14/01/2025 y disponible en: https://www.sap.com/latinamerica/products/technologyplatform/hana/what-is-datamining.html

Blanchar-Martinez, T. C. y Martinez-Trujillo, N. E. (2025). ¿Entrevista o encuesta?: Una diferencia necesaria. Internet, consultado el 23/01/2025 y disponible en: https://nuevaepoca.revistalatinacs.org/index.php/revista/article/view/2339

Brook, Chr. (2024). ¿Qué es la seguridad en la minería de datos? Destacamos sus amenazas, oportunidades y beneficios. Internet, consultado el 15/01/2025 y disponible en: https://www.digitalguardian.com/blog/what-data-mining-security-highlighting itsthreatsopportunities-and-benefits

Coppola, M. (2023). Qué es la minería de datos: conceptos, técnicas y ejemplos. Internet, consultado el 18/01/2025 y disponible en: https://blog.hubspot.es/marketing/mineriadatos#herramientas

Correa-Coronel, M. M. (2022). Ciberseguridad y su incidencia en el tratamiento de datos personales en una Municipalidad distrital de Lima Sur, 2021. Internet, consultado el 22/12/2024 y disponible en: https://repositorio.ucv.edu.pe/handle/20.500.12692/85975

Chero-Montalbán, J. G. (2020). Técnicas de minería de datos en el diseño de aplicaciones para mejorar el análisis de la gestión de recursos humanos del proyecto especial Altomayo Internet, consultado el 23/12/2024 y disponible en: https://repositorio.uss.edu.pe/handle/20.500.12802/6788

Dongo-Pozo, A. Fr. y Silva-Cama, X. P. (2020). Análisis de la minería de datos aplicada en empresas del sector retail. Internet, consultado el 23/12/2024 y disponible en: https://repositorio.ucsp.edu.pe/backend/api/core/bitstreams/e548e3c5-9cab-4dcdb55e9923c7a9e80b/content

Guzmán, A. (2022). La importancia de la Ciberseguridad en las empresas. Internet, consultado el 17/01/2025 y disponible en: https://welcome.atlasgov.com/es/blog/ciberseguridad/laimportancia-de-laciberseguridad-en-las-empresas/

Grupo Iberdrola (2024). Data mining': definición, ejemplos y aplicaciones: Descubre cómo el 'data mining' predecirá nuestro comportamiento. Internet, consultado el 02/01/2025 y disponible el: https://www.iberdrola.com/innovacion/data-mining-definicion-ejemplosyaplicaciones

Hernández-Sampieri, R., Fernández-Collado, C. y Baptista-Lucio, P. (2014). Metodología de la investigación, sexta edición. Editorial McGraw-Hill Interamericana. México.

Hernández-Mechate, E. J. (2020). Vulnerabilidades informáticas en el portal web de la Universidad Andina del Cusco. Internet, consultado el 23/12/2024 y disponible en: https://repositorio.uandina.edu.pe/item/607cc99f-1dd9-4dd2-a0a3-bfe8967d2739

Holdsworth, J. (2024). ¿Qué es la minería de datos?. Internet, consultado el 14/01/2025 y disponible en: https://www.ibm.com/es-es/topics/data-mining

IBM (2025). Objetivos de la minería de datos. Internet, consultado el 16/01/2025 y disponible en: https://www.ibm.com/docs/es/spss-modeler/18.5.0?topic=goals-data-mining

Instituto Europeo de Posgrado (2020). Minería de Datos: ¿Qué es y en qué consiste?. Internet, consultado el 16/01/2025 y disponible en: https://www.iep-edu.com.co/mineria-dedatos/ Instituto Nacional de Ciberseguridad (2021). Glosario de términos de ciberseguridad Una guía de aproximación para el empresario. Internet, consultado el 20/01/2025 y disponible en: https://www.incibe.es/sites/default/files/contenidos/guias/doc/guia_glosario_ciberseg ur idad_2021.pdf

Jimenez-Chuque, F. E. (2022). Modelo de detección de amenazas digitales para mitigar los riesgos de ciberseguridad en las organizaciones, 2019. Internet, consultado el 22/12/2024 y disponible en:

Kaspersky (2025), ¿Qué es la minería de datos y por qué es importante?. Internet, consultado el 15/01/2025 y disponible en: https://latam.kaspersky.com/resourcecenter/definitions/data-mining?srsltid=AfmBOooQ4YHoHgKBkwfvuUQvARKtiViWheHceEaLg6Lg8NB8oxdOz 2do

Kaspersky (2025a). ¿Qué es la ciberseguridad? Internet, consultado el 17/01/2025 y disponible en: https://latam.kaspersky.com/resource-center/definitions/what-iscybersecurity?srsltid=AfmBOoqPjGXk92t1nY0lX8ITyDWd9P_KZoCqCLlIf2Urh859C5e Po43 c

Kharkovyna, O. (2020). El Glosario Definitivo De La Ciencia De Los Datos. Internet, consultado el 20/01/2025 y disponible en: https://www.datasource.ai/es/data-sciencearticles/elglosario-definitivo-de-la-ciencia-de-los-datos

Loranca. M. (2021). El cubo de la ciberseguridad. Internet, consultado el 17/01/2025 y disponible en: https://i-networks.com.mx/el-cubo-de-la-ciberseguridad/

Matilde-Espino, Y. (2023). Análisis de Inclusión de la Ciberseguridad en la Educación y su Relevancia Empresarial. Internet, consultado el 20/12/2024 y disponible en: https://ring.uaq.mx/handle/123456789/9429

Muntané-Relat, J. (2010). Introducción a la investigación básica. Internet, consultado el 22/01/2025 y disponible en: https://www.sapd.es/revista/2010/33/3/03/pdf

Ortega, K. (2025). ¿Cuáles son los objetivos de la ciberseguridad?. Internet, consultado el 18/01/2025 y disponible en: https://worldcampus.saintleo.edu/blog/estudiarciberseguridad-en-linea-objetivos-de-laciberseguridad

Pozo-Acosta, L. (2022). Ciberseguridad y medidas de protección de la información adoptadas por el estado ecuatoriano. Internet, consultado el 20/12/2024 y disponible en: https://repositorio.iaen.edu.ec/handle/24000/6103

Sánchez, E. (2025). Data mining: Definición y aplicaciones esenciales 2025. Internet, consultado el 04/01/2025 y disponible en: https://blog.mercately.com/marketing/datamining

Sociedad Mercantil Estatal para la Gestión de la Innovación y las Tecnologías Turísticas (2023). Glosario básico para no perderse entre datos: 1ª parte, de la "A" a la "F". Internet, consultado el 20/01/2025 y disponible en: https://www.dataestur.es/blog/glosario-significados-terminos-ciencia-datos/

Tecnologiabi (2025). Minería de Datos. Internet, consultado el 20/01/2025 y disponible en: https://tecnologiabi.com/glosario-bi/mineria-de-datos/

Toledo, R. (2025). Por qué es importante la ciberseguridad. Internet, consultado el 19/01/2025 y disponible en: https://www.grupocibernos.com/blog/por-que-es-importantelaciberseguridad

Torres-Quezada, Y. S. (2022). Minería de datos para determinar los factores más influyentes en la ocurrencia de Siniestros de Tránsito en Ecuador en el año 2020. Internet, consultado el 20/12/2024 y disponible en: https://dspace.unl.edu.ec/jspui/bitstream/123456789/24502/1/YulissaStefania_Torres Quezada.pdf.

UNIR Formación Profesional (2023). Los 4 principios de la seguridad informática y su implementación. Internet, consultado el 18/01/2025 y disponible en: https://unirfp.unir.net/revista/ingenieria-y-tecnologia/principios-seguridad-informatica/

Vallejo-P., D. y Tenelanda-V., G. (2022) Minería de datos aplicada en detección de intrusos. Internet, consultado el 18/01/2025 y disponible en: https://dialnet.unirioja.es/descarga/articulo/4694116.pdf

Zendesk (2024). ¿Qué es la ciberseguridad y cuál es su relación con la IA?. Internet, consultado el 19/01/2025 y disponible en: https://www.zendesk.com.mx/blog/ciberseguridad/

Zendesk (2023). Data mining: 6 etapas para obtener información valiosa. Internet, consultado el 16/01/2025 y disponible en: https://www.zendesk.com.mx/blog/data-mining-que-es/

Los anexos, panel fotográfico y otros documentos están resguardados en la oficina de repositorio digital institucional en la Biblioteca Central de la Universidad Tecnológica de los Andes