

UNIVERSIDAD NACIONAL DANIEL ALCIDES CARRIÓN

FACULTAD DE INGENIERÍA

ESCUELA DE FORMACIÓN PROFESIONAL DE SISTEMAS Y

COMPUTACIÓN



TESIS

Aplicación de técnicas de minería de datos para medir la calidad del servicio Educativo y lectoescritura en estudiantes de primaria del módulo educación de la ENAHO Trimestre 4, 2021

Para optar el título profesional de:

Ingeniero de Sistemas y Computación

Autor: Bach. Victor Antony HUANCA MAYHUASCA

Asesor: Mg. Pit Frank ALANIA RICALDI

Cerro de Pasco – Perú – 2023

UNIVERSIDAD NACIONAL DANIEL ALCIDES CARRIÓN

FACULTAD DE INGENIERÍA

ESCUELA DE FORMACIÓN PROFESIONAL DE SISTEMAS Y

COMPUTACIÓN



TESIS

Aplicación de técnicas de minería de datos para medir la calidad del servicio Educativo y lectoescritura en estudiantes de primaria del módulo educación de la ENAHO Trimestre 4, 2021

Sustentada y aprobada ante los miembros del jurado:

Mg. Oscar Cleворio CAMPOS SALVATIERRA
PRESIDENTE

Mg. Lisbeth Gisela NEGRETE CARHUARICRA
MIEMBRO

Ing. Arturo Melquiades TRINIDAD MALPARTIDA
MIEMBRO

DEDICATORIA.

Esta tesis está dedicada a:

A mis padres Victor y Jubana quienes con su amor, paciencia y esfuerzo me han permitido llegar a cumplir hoy un sueño más, gracias por inculcar en mi el ejemplo de esfuerzo y valentía.

A mis hermanas Katherin y Lizzet por su cariño y apoyo incondicional, durante todo este proceso, por estar conmigo en todo momento, gracias.

AGRADECIMIENTO

A mi familia, por haberme dado la oportunidad de formarme como profesional y haber sido mi apoyo durante todo este tiempo.

De manera especial a mi tutor de tesis, por haberme guiado, no solo en la elaboración de este trabajo de titulación, sino a lo largo de mi carrera universitaria y haberme brindado el apoyo para desarrollarme profesionalmente.

A la Universidad Daniel Alcides Carrión, por haberme brindado tantas oportunidades y enriquecerme en conocimiento.

RESUMEN

El trabajo de investigación que realice se titula: Aplicación de técnicas de minería de datos para medir la calidad del servicio Educativo y lectoescritura en estudiantes de primaria del módulo educación de la ENAHO Trimestre 4, 2021”. El objetivo principal es Aplicar las técnicas de minería de datos ayudarán a medir la calidad del servicio educativo y lectoescritura en estudiantes de primaria del módulo educación de la ENAHO Trimestre 4, 2021. El diseño utilizado es apropiado. El propósito de este tipo de investigación es medir el grado de asociación que existe entre dos o más conceptos o variables. Los residentes son residentes de casas particulares en las ciudades y zonas rurales del país, estudiantes de primaria. El tamaño de la muestra fue de Conformada por 26589 encuestas de la ENAHO del cuarto trimestre del año 2021. Para la evaluación usamos el instrumento parte de la encuesta nacional de hogares – ENAHO, se tuvo como resultado La alfabetización es necesaria para percibir adecuadamente la calidad de los servicios educativos que brinda el gobierno nacional y las instituciones privadas. Por lo tanto, en el marco de la Encuesta Nacional de Hogares (ENAHO), la información anterior se recopila con fines de investigación y análisis. Los resultados se confirmaron con consistencia interna mediante el análisis factorial confirmatorio, que concluyó que su percepción de la calidad de los servicios educativos recibidos en la escuela no estaba relacionada con las capacidades de los niños. Lee y escribe.

Palabras Clave. Técnicas de minería de datos, Medir la calidad del servicio educativo y lectoescritura.

ABSTRACT

The research work that I carry out is entitled: Application of data mining techniques to measure the quality of the Educational service and literacy in primary students of the ENAHO education module Quarter 4, 2021 ". The main objective is to Apply the data mining techniques will help measure the quality of the educational service and literacy in primary students of the education module of the ENAHO Quarter 4, 2021. The design used is appropriate. The purpose of this type of research is to measure the degree of association that exists between two or more concepts or variables. The residents are residents of private homes in the cities and rural areas of the country, primary school students. The sample size was made up of 26,589 ENAHO surveys from the fourth quarter of 2021. For the evaluation we used the instrument part of the national household survey - ENAHO, the result was Literacy is necessary to adequately perceive the quality educational services provided by the national government and private institutions. Therefore, within the framework of the National Household Survey (ENAHO), the above information is collected for research and analysis purposes. The results were confirmed with internal consistency through confirmatory factor analysis, which concluded that their perception of the quality of educational services received at school was not related to the abilities of the children. Read and write.

Keywords. Data mining techniques, Measuring the quality of the educational and literacy service.

INTRODUCCIÓN

Trabajar en este trabajo es el resultado final de un gran compromiso y trabajo duro.

Filtrado, conexión a bases de datos y procesamiento de datos seguido de análisis mediante minería de datos, análisis factorial confirmatorio y árboles de decisión.

Como parte de un estudio que tiene como objetivo comprender la relación entre las percepciones de los estudiantes de primaria sobre la calidad de los servicios educativos en el Perú y los niveles de alfabetización en el 4to trimestre del módulo educativo de la ENAHO 2021, utilizando minería de datos.

En vista de sus actividades de promoción y difusión de la investigación estadística, el Instituto Nacional de Estadística e Informática (INEI) pone a disposición un sistema de microdatos. El sistema brinda acceso a bases de datos y documentos obtenidos de estudios e investigaciones realizadas por el INEI en los últimos años, garantizando la confidencialidad estadística de la información de ley de seguridad.

La presente está compuesta por los siguientes capítulos y apartados mencionados a continuación:

Capítulo I “el cual incluye los siguientes apartados: Identificación y determinación del problema, delimitación de investigación, formulación del problema, formulación de objetivos, justificaciones de la investigación y limitaciones de la investigación”.

Capitulo II “el cual incluye los siguientes apartados: Bases teóricas – científicas, definición de términos, identificación de hipótesis y variables y finalmente la definición operacional”.

Capitulo III el cual abarca: “Tipo de investigación, métodos de investigación, diseño de investigación, población y muestra, técnicas e instrumentos de recolección de datos, técnicas de procesamiento y análisis de datos, tratamiento estadístico de datos, selección, validación y confiabilidad de los instrumentos de investigación y orientación ética”.

Capitulo IV el cual consta de: “Resultados y discusión y abarca los siguientes apartados: descripción del trabajo, análisis e interpretación de resultados, prueba de hipótesis y discusión de resultados”

Finalmente, “se muestran las conclusiones, recomendaciones, referencias bibliográficas y anexos”

El autor.

ÍNDICE

DEDICATORIA

AGRADECIMIENTO

RESUMEN

ABSTRACT

INTRODUCCIÓN

ÍNDICE

CAPITULO I

PROBLEMA DE INVESTIGACIÓN

1.1.	Identificación y determinación del problema.....	1
1.2.	Delimitación de la investigación.....	2
1.3.	Formulación del problema	3
1.3.1.	Problema general	3
1.3.2.	Problemas específicos:	3
1.4.	Formulación de objetivos.....	3
1.4.1.	Objetivo General.	3
1.4.2.	Objetivos específicos.	3
1.5.	Justificación de la investigación.....	4
1.6.	Limitaciones de la investigación	4

CAPITULO II

MARCO TEÓRICO

2.1.	Antecedentes de estudio.	5
2.1.1.	A nivel Internacional.....	5

2.1.2.	A nivel Nacional	6
2.1.3.	A nivel Local.....	8
2.2.	Bases teóricas – científicas.	8
2.2.1.	Técnicas de minería de datos.	8
2.2.2.	Medir la calidad del servicio educativo y lectoescritura	26
2.2.3.	Análisis Factorial	31
2.2.4.	Análisis Factorial Confirmatorio.....	31
2.3.	Definición de términos básicos.	31
2.4.	Formulación de Hipótesis	32
2.4.1.	Hipótesis General.....	32
2.4.2.	Hipótesis Específicas	32
2.5.	Identificación de Variables.....	33
2.6.	Definición Operacional de variables e indicadores	33

CAPITULO III

METODOLOGÍA Y TECNICAS DE INVESTIGACIÓN

3.1.	Tipo de investigación	34
3.2.	Nivel de investigación	34
3.3.	Métodos de investigación	35
3.4.	Diseño de investigación.....	35
3.5.	Población y muestra	35
3.5.1.	Población.....	35
3.5.2.	Muestra.....	35
3.6.	Técnicas e instrumentos de recolección de datos	35

3.7. Selección, validación y confiabilidad de los instrumentos de investigación.....	36
3.8. Técnicas de procesamiento y análisis de datos.....	39
3.9. Tratamiento Estadístico.....	39
3.10. Orientación ética filosófica y epistémica.....	39

CAPITULO IV

RESULTADOS Y DISCUSIÓN

4.1. Descripción del trabajo de campo.....	40
4.2. Presentación, análisis e interpretación de resultados.....	41
4.3. Prueba de Hipótesis	62
4.4. Discusión de resultados.....	64

CONCLUSIONES

RECOMENDACIONES

REFERENCIAS BIBLIOGRÁFICAS

ANEXOS

ÍNDICE DE TABLAS

Tabla 1. Operacionalización de variables.....	33
Tabla 2. Módulo de Estudio	36
Tabla 3. Módulo Educación ENAHO Trimestre 4, 2021.....	42
Tabla 4. Módulo Educación Idioma ENAHO Trimestre 4, 2021.....	42
Tabla 5. Módulo Educación Centro de Estudios ENAHO Trimestre 4, 2021.....	43
Tabla 6. Módulo Educación Resultado ENAHO Trimestre 4, 2021	43
Tabla 7. Módulo Educación Grado de Estudios ENAHO Trimestre 4, 2021	44
Tabla 8. Módulo Educación según sexo ENAHO Trimestre 4, 2021	44
Tabla 9. Módulo Educación según edad ENAHO Trimestre 4, 2021.....	45
Tabla 10. Módulo Educación según región ENAHO Trimestre 4, 2021	46
Tabla 11. Módulo Educación según infraestructura ENAHO Trimestre 4, 2021	46
Tabla 12. Módulo Educación según equipamiento ENAHO Trimestre 4, 2021.....	47
Tabla 13. Módulo Educación según enseñanza ENAHO Trimestre 4, 2021.....	47

Tabla 14. Módulo Educación según dotación de materiales ENAHO Trimestre 4, 2021
..... **48**

Tabla 15. Módulo Educación según apoyo de los padres ENAHO Trimestre 4, 2021. **49**

ÍNDICE DE FIGURAS

Figura 1.....	14
Modelo de proceso de la Metodología CRISP–DM.....	14
Figura 2.....	14
Fase de comprensión del negocio o problema.....	14
Figura 3.....	16
Fase de comprensión de los datos.....	16
Figura 4.....	18
Fase de preparación de los datos.....	18
Figura 5.....	20
Fase de Modelado.....	20
Figura 6.....	22
Fase de evaluación	22
Figura 7.....	24
Fase de implementación	24
Figura 8.....	40
Código de importación de base de datos.	40
Figura 9.....	41
Base de Datos R-Project.....	41
Figura 10. Modelos de Análisis Factorial en Amos SPSS	51
Figura 11. Análisis Factorial de Servicio Educativo	51
Figura 12. Diagrama de ruta de factor no estandarizado.....	52
Figura 13. Diagrama de ruta de factor estandarizado.....	52
Figura 14. Árbol de decisión en SPSS	59
Figura 15. Ganancia para nodos	60

CAPITULO I

PROBLEMA DE INVESTIGACIÓN

1.1. Identificación y determinación del problema

En vista de sus actividades de promoción y difusión de la investigación estadística, el Instituto Nacional de Estadística e Informática (INEI) pone a disposición un sistema de microdatos. El sistema brinda acceso a las bases de datos y documentos obtenidos como resultado de investigaciones e investigaciones realizadas por el INEI en los últimos años, garantizando la confidencialidad estadística de la información de las garantías reales.

Por lo tanto, utilizando la base de datos de la Encuesta Nacional de Hogares (ENAHOG) del 4to trimestre de 2021, se hace necesaria la minería de datos para reflejar y comunicar la realidad del Perú.

Por otro lado, la calidad de los servicios educativos en el Perú depende directamente del Ministerio de Educación, el cual ha estado en crisis en todos los niveles educativos en las instituciones educativas públicas, traducándose en la ausencia de un sistema de calidad educativa en los servicios educativos. . . , teniendo en cuenta la falta de infraestructura tecnológica moderna, la falta de equipamiento de aula en carpetas, laboratorios, computadores, etc., la falta de

métodos de enseñanza de los docentes, la falta de material didáctico como libros, láminas, maquetas, entre otros., y sobre todo la carencia de una escuela de padres donde se ve necesario el apoyo y orientación de sus hijos.

El Comercio, en las entrevistas realizada a Hugo Díaz vicepresidente del Consejo Nacional de Educación titulada “los tres problemas que existen en el sector educación del Perú” considera como uno de dichos problemas es la falta de calidad y equidad, además indico que somos un país demasiado inequitativo en términos de calidad de los aprendizajes. (Díaz, 2014)

“La inteligencia del del niño por comprender el lenguaje escrito, genera la creación de hipótesis infantiles sobre cómo funciona este instrumento cultural. Al cuestionarse continuamente por esta actividad se hace consciente de los signos escritos, acercándose cada vez más a su dominio. La importancia de permitir que el niño se apropie de los conocimientos es evidente. Esta apropiación debe darse en la escuela y en el hogar como un proceso natural e implícito en todas las actividades cotidianas del niño. Promover la lectura expresando constantemente la funcionalidad de ésta en la cultura, aumenta la conciencia del manejo cotidiano del lenguaje escrito. Además, la escuela debe planear la instrucción con base en el nivel conceptual desarrollado espontáneamente por el niño y debe llevar a cabo actividades de enseñanza donde se promueva el auto-descubrimiento”. (Montealegre & Forero, 2006)

1.2. Delimitación de la investigación.

1.2.1. Espacial

Desarrollaré el trabajo de investigación con los datos del módulo educación de la ENAHO Trimestre 4, 2021.

1.2.2. Temporal

Los datos ha usar para la interpretación y el procesamiento son los del año 2021.

1.2.3. Conceptual

Se desarrolla la búsqueda teórica de la aplicación de las técnicas de minería de datos y medir la calidad de servicio Educativo y lectoescritura en estudiantes de primaria.

1.3. Formulación del problema

1.3.1. Problema general

¿Las técnicas de minería de datos ayudará a predecir la calidad del servicio educativo y lectoescritura en estudiantes de primaria del módulo educación de la ENAHO Trimestre 4, 2021?

1.3.2. Problemas específicos:

¿Existe un análisis factorial confirmatorio que predice la calidad del servicio educativo y lectoescritura en estudiantes de primaria del módulo educación de la ENAHO Trimestre 4, 2021?

¿Se podrá realizar un análisis mediante los arboles de decisión para predecir la calidad del servicio educativo y lectoescritura en estudiantes de primaria del módulo educación de la ENAHO Trimestre 4, 2021?

1.4. Formulación de objetivos

1.4.1. Objetivo General.

Aplicar las técnicas de minería de datos ayudarán a predecir la calidad del servicio educativo y lectoescritura en estudiantes de primaria del módulo educación de la ENAHO Trimestre 4, 2021.

1.4.2. Objetivos específicos.

Realizar el análisis factorial confirmatorio que predice la calidad del servicio educativo y lectoescritura en estudiantes de primaria del módulo educación de la ENAHO Trimestre 4, 2021.

Realizar el análisis mediante los árboles de decisión para predecir la calidad del servicio educativo y lectoescritura en estudiantes de primaria del módulo educación de la ENAHO Trimestre 4, 2021.

1.5. Justificación de la investigación

Este trabajo de investigación es relevante para la industria de la educación pública, especialmente en la región de Pasco, porque a través de la minería de datos se descubrieron posibles correlaciones entre las percepciones de los estudiantes de primaria sobre la calidad del servicio. La educación peruana que recibieron en la institución a la que asistieron estuvo ligada a su capacidad para leer y escribir correctamente.

Por otro lado, es muy importante para los investigadores en el campo de la educación, porque quieren tener al alcance de la mano la información más actualizada sobre la educación primaria en el Perú, y también es muy importante para los estadísticos, porque se aprovecha. para minería de datos para explorar la base de datos de ENAHO - módulo de capacitación.

1.6. Limitaciones de la investigación

La minería de datos tiene varias técnicas, esta vez para la investigación solo usaremos el uso del Análisis Factorial Confirmatorio y los Árboles de Decisión, tal como están en nuestros objetivos.

CAPITULO II

MARCO TEÓRICO

2.1. Antecedentes de estudio.

2.1.1. A nivel Internacional

- Pio Alvarez (2018) Tesis intitulada “Minería de calidad de datos: Aplicación de técnicas de minería de datos para la Evaluación de la calidad de los Datos” El objetivo de este estudio es explorar el estado del arte en calidad de datos, desde su definición y los diversos conceptos existentes hasta las técnicas comunes para tratarla, y analizar la posibilidad de integrar estos dos campos. Se han sugerido algunas técnicas específicas para aplicar técnicas de minería de datos para resolver problemas de calidad de datos. Este trabajo explora en profundidad los dos campos presentados, para luego analizar los mecanismos que permiten conectarlos, y finalmente introducir métodos que nos permitan aproximarnos a la calidad de los conjuntos de datos analizados utilizando las posibilidades inherentes a la minería de datos. Este trabajo dio como resultado dos nuevas propuestas para el uso de métodos de minería de datos para la evaluación de la calidad de los datos, que se presentaron en dos eventos

internacionales especiales. Uno se enfoca en evaluar si el conjunto de datos está lo suficientemente actualizado y el otro se enfoca en analizar los datos faltantes. Además, se plantea una tercera propuesta, que aún está en desarrollo, para evaluar su aplicabilidad en función de las características del conjunto de datos.

- Montealegre y Fore (2006) Tesis intitulada “Desarrollo de la lectoescritura: adquisición y dominio” Se concluyó que la adquisición del lenguaje escrito es inequívoca desde que se aprende a leer hasta que se aprende a leer. La etapa de aprender a escribir de los niños y niñas debe presentarse como un paso hacia su integración cultural, por lo que se enfatiza su función, uso cotidiano y relevancia para la comprensión de otros saberes. La escritura es un tema valioso, porque la habilidad para escribir es la primera evidencia de la adquisición gradual de los principios del alfabeto y un ejemplo convincente de alfabetización superior posterior. Ser capaz de transmitir conceptos e ideas a través de un texto escrito significa usar múltiples campos y ser lo suficientemente claro para transcribir las ideas en sí mismas de una manera estructurada, clara y coherente. El contexto afecta el procesamiento del lenguaje escrito, porque la cosmovisión, la cultura, la interacción social, las situaciones sociales y escolares, etc. determina la estructura mental del sujeto en el que se integra la nueva información Los factores personales, sociales y ambientales deben ser entendidos antes de construir procedimientos metodológicos. en la página de instrucciones.

2.1.2. A nivel Nacional

- Piscocya Ordoñez (2016) “Aplicación de técnicas de minería de datos para predecir la deserción estudiantil en la Educación Básica regular en la Región de Lambayeque” El objetivo es desarrollar una herramienta utilizando técnicas de minería de datos que brinde a los usuarios acceso a información precisa

que prediga los estudiantes que se matricularán en los próximos años. Los diseños utilizados son técnicos y experimentales – cuasi-experimentales. La población es un elemento de investigación que identifica registros entre 2006 y 2015 en la región Lambayeque, la cual consta de tres Ugels: Chiclayo, Lambayeque y Ferreñafe, la muestra es Ugel Chiclayo. Las herramientas utilizadas son mapas que nos ayudan a resumir las predicciones que obtendremos, recopilación de datos, resultados, comparación de dos técnicas, redes neuronales y ETS. Se observó que la red neuronal alcanzó el mayor nivel de confianza en comparación con ERS, lo que se representa en el valor obtenido mediante el cálculo del coeficiente. Teniendo en cuenta la última métrica de la sección Disponibilidad de los sistemas de red, el tiempo promedio de generación de cálculos es de 30,6 segundos. En el nivel secundario, me tomó un promedio de 29,5 segundos generar estimaciones de disponibilidad del sistema. Finalmente, se creó una aplicación web para evaluar los resultados obtenidos. El sistema está desarrollado en php, el cual proporciona una interfaz capaz de interactuar con el servidor para ejecutar modelos, reales o simulados, donde se recupera de un almacén de datos el historial de cada escuela analizada para que el usuario realice pruebas de correlación.

- Anaya Carrillo y Bedón Trujillo (2018) Tesis intitulada “Minería de datos sobre la opinión de la calidad del servicio educativo peruano y la lectoescritura en estudiantes de Educación Primaria del módulo Educación de la ENAHO Trimestre 2, 2016” El objetivo fue determinar mediante minería de datos si existe una relación entre las actitudes hacia la calidad de los servicios educativos en el Perú y la alfabetización de los estudiantes de primaria del segundo trimestre de 2016 del módulo de educación de la ENAHO. El diseño utilizado no es experimental. La población son residentes de casas particulares en las ciudades y el interior del país, son estudiantes, la muestra está conformada por 29.585 encuestas ENAHO en el segundo trimestre del 2016.

Además, la encuesta también incluyó investigación de características, limpieza y selección de productos para satisfacer. procesamiento de datos; es decir, los encuestados eran estudiantes de primaria. Instrumentos utilizados Se recopiló a partir de la encuesta ENAHO 2do trimestre 2016 y la Encuesta Nacional de Hogares INEI. La alfabetización es necesaria para percibir adecuadamente la calidad de los servicios educativos que brinda el gobierno nacional y las instituciones privadas. Por lo tanto, en el marco de la Encuesta Nacional de Hogares (ENAHO), la información anterior se recopila con fines de investigación y análisis. Los resultados se validan con coherencia interna mediante el análisis factorial confirmatorio, mientras que los estudiantes proporcionan perfiles de sus habilidades de alfabetización mediante árboles de decisión. En resumen, las principales características de la alfabetización fueron que la edad de 7 años se asoció significativamente con la pertenencia a la costa y la sierra, la edad de 8 años se asoció con la pertenencia a la costa, la sierra y la selva, y las edades de 9, 10 y más de 11 años se asociaron significativamente con el nivel del conocimiento. como leer y escribir. De igual forma, los niños de 11 años que saben leer y escribir y que expresan opiniones importantes sobre la calidad de los servicios educativos en cuanto a equipamiento (carpetas, laboratorios, computadoras, etc.) se distinguen de los niños que no saben leer ni escribir. cómo. saber leer y escribir hasta los 6 años con una vinculación significativa a la pertenencia a la costa, sierra o selva.

2.1.3. A nivel Local

- No se encontró evidencia o literatura de contenido similar en la biblioteca de la Universidad Nacional Daniel Alcides Carrión.

2.2. Bases teóricas – científicas.

2.2.1. Técnicas de minería de datos.

2.2.1.1. Minería de datos.

La minería de datos o minería de datos (que es la fase analítica de las "bases de datos de descubrimiento de conocimiento" o KDD) es un campo de la estadística y la informática que se refiere al proceso de tratar de descubrir patrones en grandes conjuntos de datos..

Según Morales (2003) "La Minería de Datos es un proceso completo de descubrimiento de conocimiento que involucra varios pasos":

- Comprender el dominio de la aplicación, el conocimiento utilizable relacionado y los objetivos del usuario.
- Seleccionar conjuntos de datos para realizar el proceso de registro.
- Limpieza y preprocesamiento de datos, desarrollando estrategias adecuadas para eliminar ruido, valores incompletos, valores fuera de rango, valores inconsistentes, etc.
- Seleccionar la tarea de descubrimiento a realizar, como clasificación, agrupación o clustering, reglas de asociación, etc.
- Seleccione el algoritmo a utilizar.
- Transformar los datos a un formato requerido por algoritmos de minería de datos específicos, encontrar atributos útiles, reducir la dimensionalidad de los datos, etc.
- Realizar un proceso de minería de datos para descubrir patrones interesantes.
- Evaluar patrones descubiertos y su presentación usando técnicas de visualización. Puede ser necesario eliminar patrones redundantes o poco interesantes, o repetir algunos de los pasos anteriores con otros datos, otros algoritmos, otros objetivos u otras estrategias.
- Utilizar los conocimientos descubiertos incorporándolos a un sistema o simplemente almacenándolos e informándolos a las partes interesadas.

“Es muy importante la etapa del pre-procesamiento de los datos y su transformación al formato requerido por el algoritmo, ya que dependiendo de cómo se realicen estas tareas, va a depender la calidad final de los patrones descubiertos. Un patrón es interesante si es fácilmente entendible por las personas, potencialmente útil, novedoso o valida alguna hipótesis que el usuario busca confirmar. Un patrón interesante representa conocimiento como menciona Ale” (2005).

Las principales técnicas de minería de datos se suelen clasificar según su tarea de descubrimiento en:

1. Agrupación de datos:

Chen y Han (1996) “La agrupación o clustering consiste en agrupar un conjunto de datos basándose en la similitud de los valores de sus atributos. El clustering identifica regiones densamente pobladas, denominadas clusters, de acuerdo a alguna medida de distancia establecida”.

De esta forma, tratamos de maximizar la similitud de casos dentro de cada grupo y minimizar la similitud entre grupos. Las técnicas de agrupamiento se han estudiado en estadísticas, aprendizaje automático, bases de datos espaciales y minería de datos. Dos de los algoritmos de agrupamiento más utilizados son los mapas autoorganizados (SOM) y K-means. SOM, también conocida como Kohonen Network, fue fundada en 1982 por Teuvo Kohonen. Es un modelo de red neuronal que mapea funciones similares a las que suceden en el cerebro. SOM se basa en el aprendizaje competitivo y no supervisado, lo que significa que el proceso de aprendizaje no requiere intervención humana y no requiere el conocimiento de las propiedades de la información de entrada. SOM proporciona un mapa topológico de los datos representados en diferentes dimensiones

utilizando unidades de mapa (neuronas) para simplificar la representación. Las neuronas suelen crear un mapa bidimensional, por lo que el mapa transforma el problema de múltiples dimensiones en el espacio a un plano. La propiedad de conservación de la topología significa que el mapa conserva las distancias relativas entre los puntos. Los puntos que están cerca uno del otro en el espacio de entrada original se asigna a las neuronas vecinas en el SOM. Por esta razón, SOM es útil como herramienta para analizar clases de datos multidimensionales y también tiene una capacidad de generalización, lo que significa que la red puede reconocer o caracterizar entradas que nunca antes había encontrado. K-means es un método iterativo que tiene como objetivo crear k clústeres, donde k está predeterminado antes de que comience el proceso. K significa que primero divide los datos en k subconjuntos no vacíos, calcula el centroide de cada partición como el centro del grupo y asigna cada elemento de datos al grupo más cercano al centroide. Luego, los datos se vuelven a particionar iterativamente hasta que no haya más datos que no cambien el clúster de una iteración a la siguiente.

2. Clasificación de datos:

Chen y Han (1996) "La clasificación se utiliza para clasificar un conjunto de datos basado en los valores de sus atributos. Por ejemplo, se podría clasificar a distintas personas para la otorgación de un préstamo en riesgo bajo, medio y alto, teniendo en cuenta información histórica de las mismas"

La clasificación encuentra atributos comunes entre un conjunto de objetos y los clasifica en diferentes clases de acuerdo con un modelo de clasificación. Para construir este modelo se utiliza un conjunto de entrenamiento, donde cada instancia contiene un conjunto de atributos

y el valor de la clase a la que pertenece. El objetivo de la clasificación es analizar los datos de entrenamiento y, usando métodos supervisados, desarrollar una descripción o modelo para cada clase usando las características disponibles en los datos. Esta descripción o modelo nos permite clasificar otros casos donde la clase es desconocida. El método se llama supervisado porque el conjunto de entrenamiento tiene categorías de membresía conocidas y se le dice al modelo si la clasificación que hizo fue correcta. La construcción del modelo se basa en esta retroalimentación instructiva del conductor.

Los algoritmos mayormente utilizados para las tareas de clasificación son los algoritmos de inducción. En la actualidad existen numerosos enfoques de algoritmos

de inducción y variedad en cada enfoque, el presente trabajo hará hincapié en aquellos orientados a generar árboles de decisión.

La clasificación basada en árboles de decisión es un método de aprendizaje supervisado que construye árboles de decisión a partir de un conjunto de entrenamiento. Un sistema típico de construcción de árboles de decisión es ID3, que utiliza la teoría de la información para reducir la cantidad de pruebas requeridas para clasificar objetos. Usando heurística, ID3 garantiza un árbol simple, pero no necesariamente el más simple. Una extensión de ID3 es C4.5, que amplía el dominio de la taxonomía de atributos categóricos a atributos numéricos. Un paso importante en la construcción de árboles de decisión es la poda, que elimina las ramas innecesarias, lo que acelera la clasificación y mejora la precisión de la clasificación de los datos

3. Asociación:

La minería de reglas de asociación consiste en encontrar reglas de la forma $(A_1yA_2y...yA_m) \Rightarrow (B_1yB_2y...yB_n)$, donde A_i y B_j son valores de

atributos del conjunto de datos (Chen & Han, 1996) Por ejemplo, se podría encontrar en un gran repositorio de datos de compras en un supermercado, la regla de asociación correspondiente, si un cliente compra leche, entonces compra pan. Una regla de asociación es una sentencia probabilística acerca de la co-ocurrencia de ciertos eventos en una base de datos, y es particularmente aplicable a grandes conjuntos de datos.

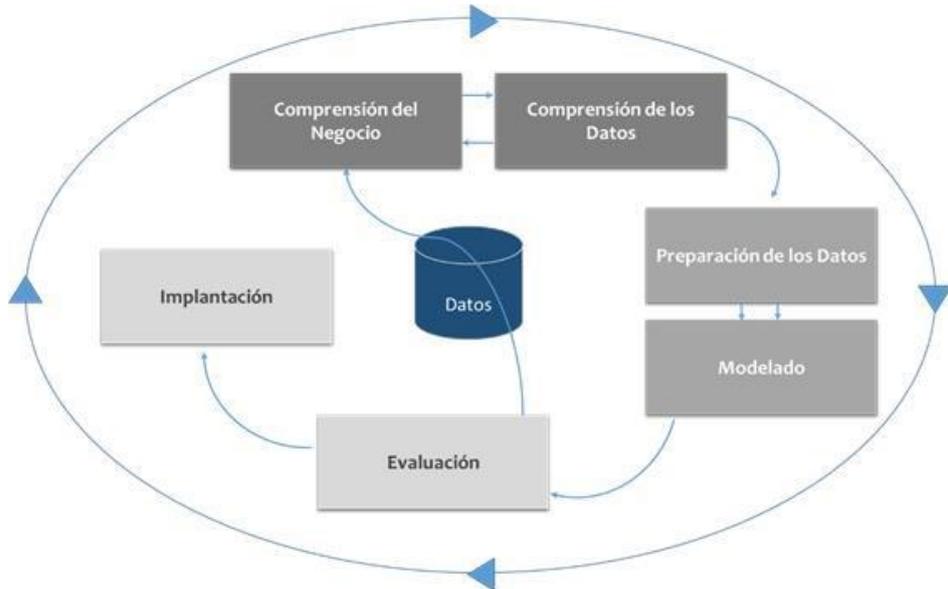
Existen varios algoritmos que realizan el descubrimiento de reglas de asociación, uno de los más utilizados es A priori.

2.2.1.2. CRISP-DM (Cross Industry Standard Process for Data Mining).

IBM (2012), Se considera que CRISP-DM (Cross-Industry Data Mining Standard Process) proporciona una descripción estandarizada del ciclo de vida de un proyecto de análisis de datos estándar, similar al ciclo de vida del desarrollo de software, que se lleva a cabo en el desarrollo de software utilizando un modelo de ciclo de datos. . El modelo CRISPDM cubre las fases del proyecto, sus respectivas tareas y las relaciones entre estas tareas. No es posible identificar todas las relaciones en este nivel de descripción; puede haber relaciones entre cualquier tarea según el propósito, el contexto y el interés del usuario en los datos.

El enfoque CRISP-DM trata el proceso de análisis de datos como un proyecto profesional, creando así un entorno más rico que influye en el desarrollo del modelo. Este contexto tiene en cuenta la existencia de clientes que no forman parte del equipo de desarrollo y el hecho de que el proyecto no termina solo después de encontrar el modelo ideal (ya que luego debe implementarse y mantenerse), sino que está vinculado a otros proyectos y debe estar completamente documentado para poder usar y trabajar a partir del conocimiento adquirido por otros equipos de desarrollo.

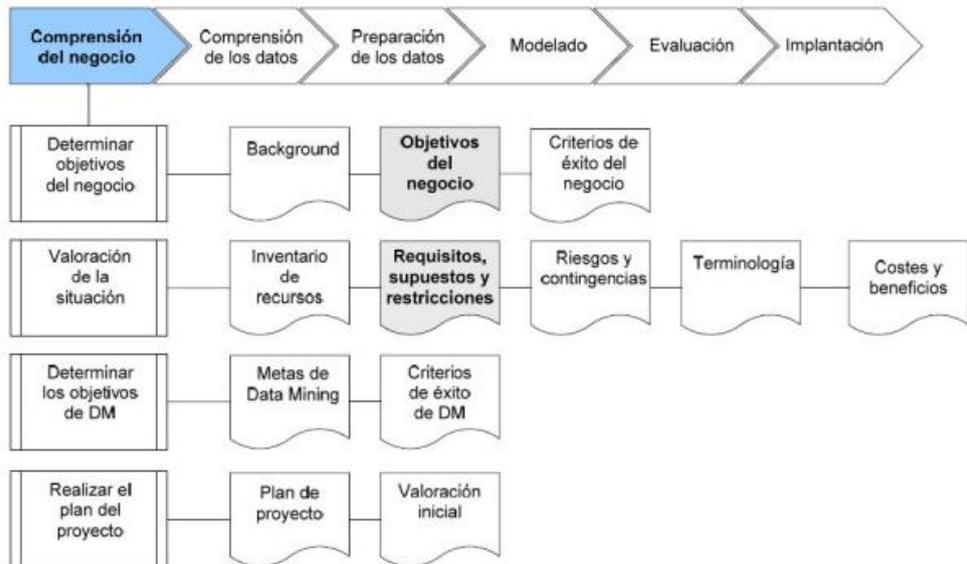
Figura 1.
Modelo de proceso de la Metodología CRISP-DM



Fases del CRISP-DM:

Figura 2.

Fase de comprensión del negocio o problema.



La primera fase del Manual de referencia del método CRISP-DM se denomina fase de comprensión del negocio o del problema.

Quizás lo más importante es que combina la tarea de comprender los objetivos y requisitos del proyecto desde una perspectiva comercial o institucional a medida que los traduce en objetivos técnicos y planes de proyecto. Sin comprender estos objetivos, ningún algoritmo, por sofisticado que sea, puede lograr resultados fiables. Por lo tanto, para aprovechar al máximo la minería de datos, es necesario comprender el problema en cuestión para que se puedan recopilar los datos correctos y los resultados se interpreten correctamente. A continuación se muestra una descripción de cada tarea principal que conforma esta fase:

1. Determinar los objetivos del negocio.

Esta es la primera tarea a desarrollar con el objetivo de identificar el problema a resolver, por qué se necesita la minería de datos y definir criterios de éxito. Los problemas pueden ser tan diversos como detectar fraudes con tarjetas de crédito, detectar intentos no autorizados de ingresar al sistema, asegurar el éxito de una determinada campaña publicitaria, etc.

2. Evaluación de la situación.

En este ejercicio, antes de iniciar el proceso de DM, se debe definir el estado de la situación, teniendo en cuenta:

¿Cuál es el conocimiento previo del problema? ¿Tiene suficientes datos para resolver el problema? ¿Cuál es la relación costo-beneficio de una aplicación de DM? En esta fase se definen los requerimientos del problema desde una perspectiva de negocio y minería de datos.

3. Determinación de los objetivos de DM.

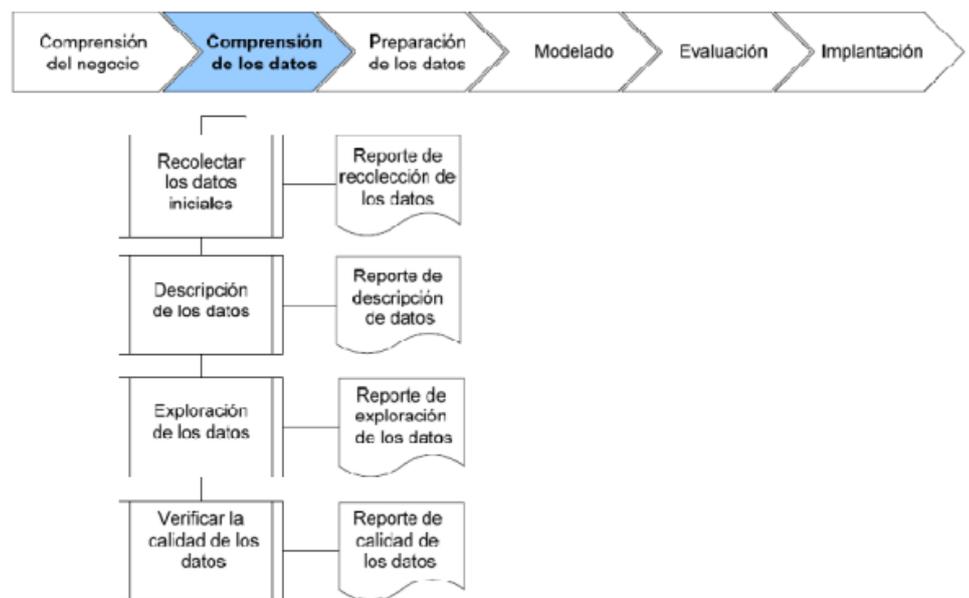
Este ejercicio pretende reflejar el objetivo comercial en relación con el objetivo del proyecto DM, si el objetivo comercial es realizar una campaña publicitaria para aumentar el crédito hipotecario, el objetivo

DM será, por ejemplo, determinar el perfil de capacidad de endeudamiento del cliente.

4. Producción de un plan del proyecto.

Finalmente, la tarea final de CRISP-DM Fase 1 tiene como objetivo desarrollar un plan de proyecto, incluidos los pasos a seguir y los métodos a utilizar.

Figura 3.
Fase de comprensión de los datos.



Consiste en la recogida de datos iniciales con el objetivo de realizar una primera conexión con el problema, conocerlo, identificar sus características e identificar las relaciones más evidentes para definir las primeras hipótesis.

Esta fase y las dos siguientes son las que requieren más esfuerzo y tiempo en un proyecto de DM. Generalmente, si una organización tiene una base de datos empresarial, debe crear una nueva base de datos específicamente para proyectos de DM.

Esto crea muchos problemas debido a los frecuentes y grandes accesos a la base de datos que se pueden generar durante el desarrollo del proyecto para consultas y posibles cambios.

1. Recolección de datos iniciales

La primera tarea en esta segunda fase, es la recolección de los datos iniciales y su adecuación para el futuro procesamiento.

El objetivo de este ejercicio es, en primer lugar:

- Elaboración de informes con listados de datos recopilados
- Tu ubicación
- los métodos utilizados en su montaje
- Problemas y soluciones específicos del proceso

2. Descripción de los datos.

Después de obtener los datos iniciales, deben ser descritos. Este proceso incluye la creación de conjuntos de datos, su identificación, el significado de cada campo y la descripción del formato original

3. Exploración de datos.

Luego continúa con la exploración que tiene como objetivo encontrar la estructura general de los datos.

Implica la aplicación de pruebas estadísticas básicas que revelan las características de los datos, la creación de tablas de frecuencia y la creación de gráficos de distribución.

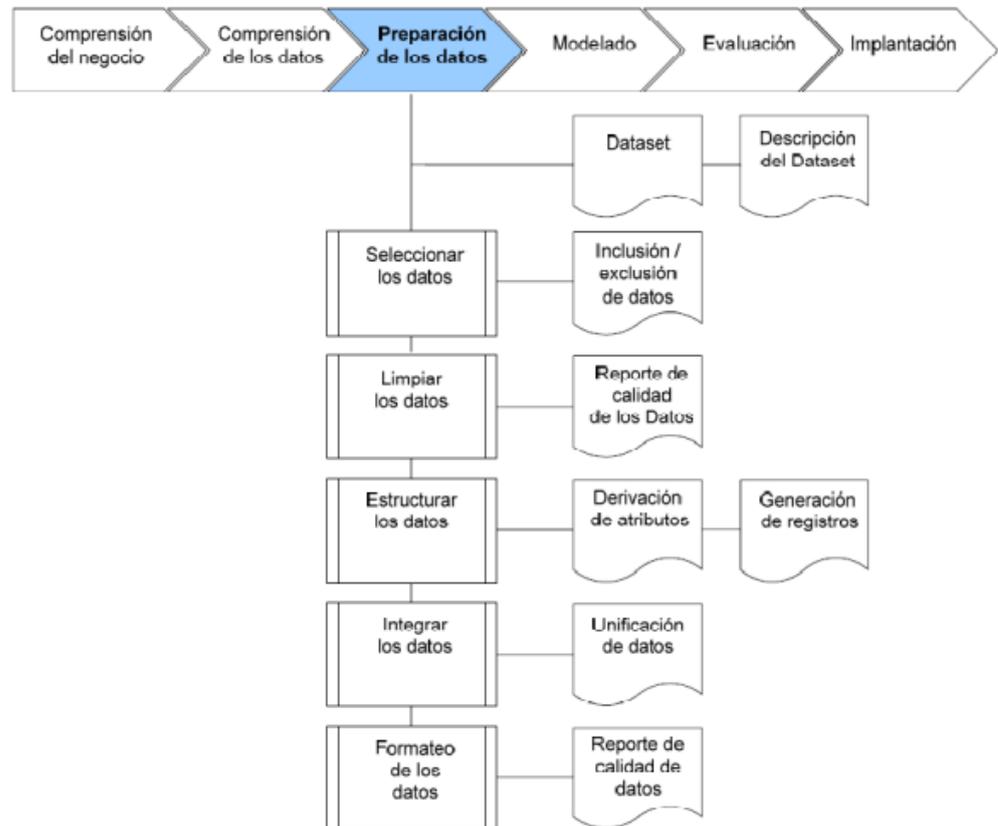
El resultado de este ejercicio es un informe de exploración de datos.

4. Verificación de la calidad de los datos.

Este ejercicio examina los datos, determina la consistencia de los valores individuales de los campos, encuentra el número y la distribución de valores nulos y valores fuera de rango, lo que puede crear un proceso ruidoso.

En conclusión, la idea en este punto, es asegurar la completitud y corrección de los datos

Figura 4.
Fase de preparación de los datos



En esta etapa, una vez que se realiza la recopilación inicial de datos, se puede adaptar para el uso posterior de técnicas de minería de datos, como técnicas de visualización de datos, búsqueda de relaciones entre variables u otras actividades de exploración de datos.

La preparación de datos incluye tareas generales como la selección de datos para la aplicación de métodos de modelado específicos, limpieza de datos, generación de variables adicionales, integración de diferentes fuentes de datos y cambios de formato.

1. Selección de datos

En esta fase se selecciona un subconjunto de los datos obtenidos en la fase anterior según los criterios previamente definidos en la fase anterior:

- Calidad de datos versus integridad
- Corrección de limitaciones de volumen o tipo de datos relacionadas con la tecnología DM seleccionada.

2. Limpieza de los datos

Esta tarea complementa la anterior y es una de las que más tiempo consume, ya que se pueden utilizar diferentes métodos para optimizar la calidad de los datos en preparación para la fase de modelado. Algunos métodos se utilizan para este propósito:

- Normalización de datos.
- Discretización de campos numéricos
- Manejo de valores faltantes
- Reducir la cantidad de datos

3. Estructuración de los datos

Esta tarea incluye pasos de preparación de datos tales como:
Generar nuevos atributos a partir de atributos existentes
Integrar nuevos registros o convertir valores de atributos existentes.

4. Integración de los datos

La integración de datos implica la creación de nuevas estructuras a partir de datos seleccionados, como:

Generación de nuevos dominios a partir de dominios existentes
establecer un nuevo récord

Combinar campos de tablas o tablas nuevas donde las características de varios registros u otros campos se recopilan en una nueva tabla de resumen.

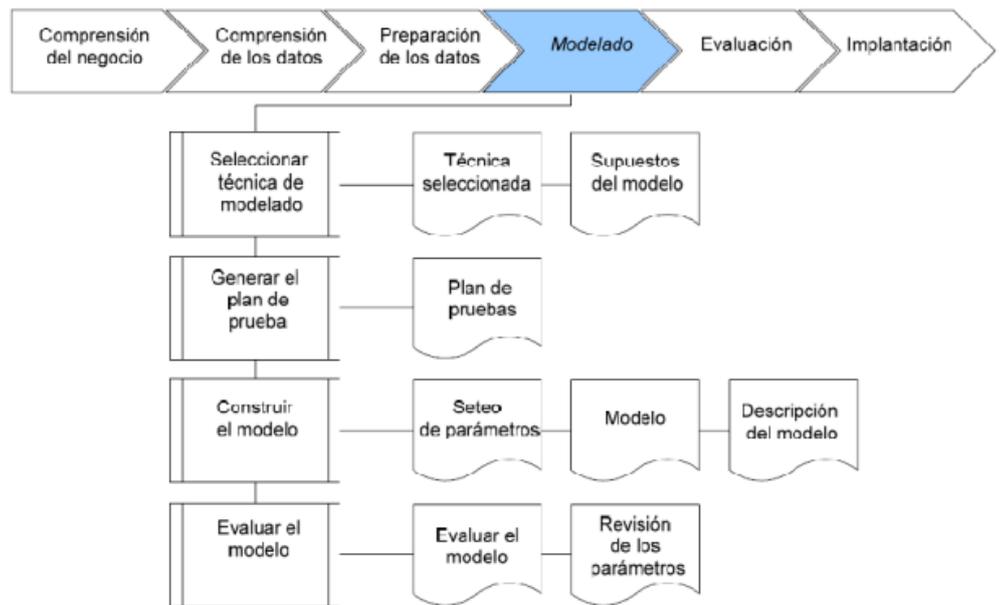
5. Formateo de los datos

Esta tarea consiste principalmente en transformar sintácticamente los datos sin cambiar su significado para permitir o facilitar el uso de técnicas específicas de DM, tales como:

Reorganizar campos y/o registros en una tabla o ajustar los valores de los campos para que se ajusten a las restricciones de la herramienta de modelado:

Elimine comas, tabulaciones, caracteres especiales, valores de cadena máximos y mínimos, etc.

Figura 5.
Fase de Modelado



La tecnología a utilizar en esta etapa se selecciona en base a los siguientes criterios:

- preguntas relevantes
- hay suficientes datos
- Completar los requisitos de la pregunta.
- Tiempo suficiente para comprar modelos.
- conocimientos técnicos

Antes de modelar los datos, se debe determinar un método para evaluar el modelo para determinar qué tan bueno es.

Después de completar estas tareas generales, pasamos a la generación y evaluación del modelo.

Los parámetros utilizados para generar el modelo dependen de las características de los datos y de las características de precisión deseadas del modelo. Se describen las tareas y los resultados de esta fase. Las principales tareas en esta etapa se describen a continuación:

1. Selección de la técnica de modelado

Esta tarea consiste en la selección de la técnica de DM más apropiada al tipo de problema a resolver. Para esta selección, se debe considerar el objetivo principal del proyecto y la relación con las herramientas de DM existentes.

Por ejemplo, si el problema es de clasificación, se podrá elegir de entre árboles de decisión, k-nearest neighbour o razonamiento basado en casos (CBR); si el problema es de predicción, análisis de regresión, redes neuronales; o si el problema es de segmentación, redes neuronales, técnicas de visualización, etc.

2. Generación del plan de prueba

Una vez creado el modelo, se debe generar un programa para probar su calidad y validez. Por ejemplo, en tareas supervisadas de DM, como la clasificación, la tasa de error se usa a menudo como una medida de calidad. Luego, los datos generalmente se dividen en dos conjuntos, uno para entrenamiento y otro para prueba, y se crea un modelo en el conjunto de entrenamiento y la calidad del modelo resultante se mide usando el conjunto de prueba.

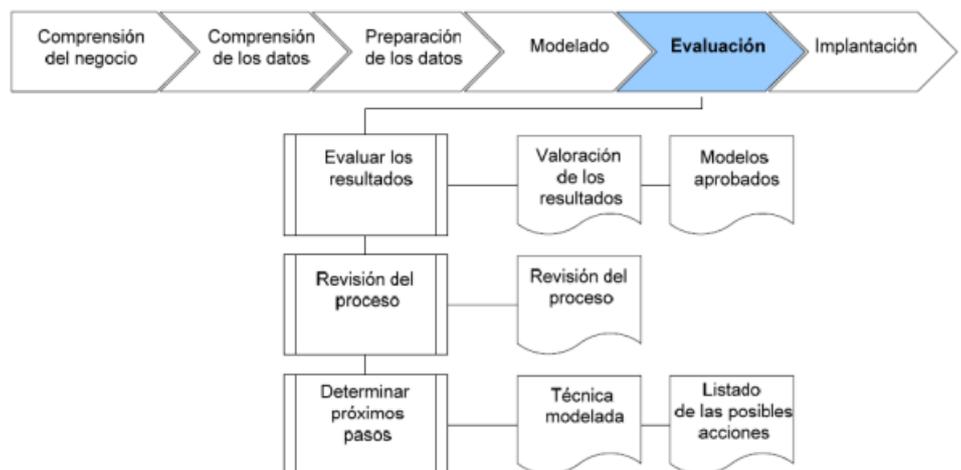
3. Construcción del Modelo

Una vez que se selecciona un método, se ejecuta en datos preparados previamente para generar uno o más modelos. Todos los métodos de modelado tienen un conjunto de parámetros que definen las propiedades del modelo generado. La elección de los mejores parámetros es un proceso iterativo basado únicamente en los resultados obtenidos. Deben interpretarse y justificarse su aplicación.

4. Evaluación del modelo

En este ejercicio, los ingenieros de DM interpretan los modelos en términos de conocimiento de dominio preexistente y criterios de éxito predeterminados. Los expertos en el dominio del problema evalúan el modelo en el contexto del dominio, los expertos en minería de datos aplican sus propios criterios (confianza en los conjuntos de prueba, pérdidas o ganancias de la tabla, etc.).

Figura 6.
Fase de evaluación



En esta fase se evalúa el modelo, teniendo en cuenta el cumplimiento de los criterios de éxito del problema.

Debe considerarse, además, que la fiabilidad calculada para el modelo se aplica solamente para los datos sobre los que se realizó el análisis.

Es preciso revisar el proceso, teniendo en cuenta los resultados obtenidos, para poder repetir algún paso anterior, en el que se haya posiblemente cometido algún error.

Considerar que se pueden emplear múltiples herramientas para la interpretación de los resultados.

Las matrices de confusión Edelstein, 1999 “son muy empleadas en problemas de clasificación y consisten en una tabla que indica cuantas clasificaciones se han hecho para cada tipo, la diagonal de la tabla representa las clasificaciones correctas”.

Si el modelo generado es válido en función de los criterios de éxito establecidos en la fase anterior, se procede a la explotación del modelo. Detalla las tareas que componen esta fase y los resultados que se deben obtener:

1. Evaluación de los resultados.

Factores como la precisión y la generalización del modelo generativo se consideraron en la fase de evaluación anterior. Este ejercicio implica evaluar el modelo frente a los objetivos comerciales y tratar de determinar si existe un caso comercial para el modelo o si es recomendable probar el modelo frente a un problema real si el tiempo y las limitaciones lo permiten. Además de los resultados directamente relacionados con los objetivos del proyecto

¿Se recomienda evaluar el modelo contra otras métricas distintas a la original? Puede revelar más.

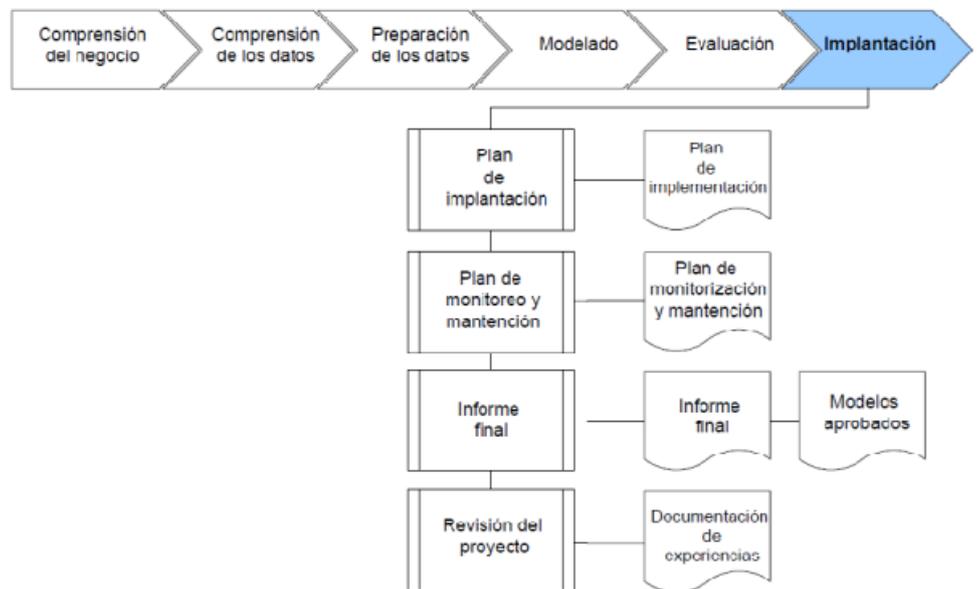
2. Revisión del proceso

El proceso de revisión se refiere a calificar todo el proceso de DM para identificar elementos que se pueden mejorar.

3. Determinación de futuras fases

Si se encuentra que los pasos anteriores han dado resultados satisfactorios, se puede continuar con el siguiente paso, de lo contrario, se puede decidir hacer otra iteración con diferentes parámetros de la etapa de preparación de datos o modelado. Tal vez incluso decida comenzar un nuevo proyecto de DM desde cero en este punto.

Figura 7.
Fase de implementación



Una vez que se crea y valida un modelo, el conocimiento resultante se traduce en acciones en el proceso comercial, ya sea que los analistas recomienden acciones en función de las observaciones del modelo y sus resultados, o aplicando el modelo a diferentes conjuntos de datos o como parte de un proceso.

Por ejemplo, en análisis de riesgo de crédito, detección de fraude, etc. Por lo general, un proyecto de minería de datos no termina con la implementación del modelo, ya que los resultados deben registrarse y presentarse de manera comprensible para los usuarios para garantizar el crecimiento del conocimiento.

Por otro lado, durante la etapa de desarrollo se debe asegurar el mantenimiento de la aplicación y la posible distribución de resultados:

1. Plan de implementación

Para implementar los resultados de DM en la organización, esta tarea toma los resultados de la evaluación y desarrolla una estrategia para implementarlos. Si existe un procedimiento general para crear el modelo, este procedimiento debe documentarse para su futura implementación.

2. Monitorización y mantenimiento

Si se implementan modelos de procesos de minería de datos en el dominio del problema como parte del trabajo diario, se recomienda preparar estrategias de monitoreo y mantenimiento para aplicar a los modelos.

Los comentarios de la supervisión y el mantenimiento pueden indicar si el modelo se está utilizando correctamente.

3. Informe Final

Esta es la conclusión del proyecto DM completado. Según el plan de implementación, este informe puede ser simplemente un resumen de los puntos destacados del proyecto y las lecciones aprendidas, o puede ser un informe final que contiene y explica los resultados alcanzados por el proyecto.

4. Revisión del proyecto

En este punto se evalúa qué fue lo correcto y qué lo incorrecto, qué es lo que se hizo bien y qué es lo que se requiere mejorar.

2.2.2. Medir la calidad del servicio educativo y lectoescritura

2.2.2.1. Lectoescritura

“La lectoescritura constituye un pilar básico en la vida de todo ser humano. No solo se trata de una herramienta indispensable para acceder a los objetivos y contenidos educativos si no que es uno de los principales vehículos de la cultura existente. La importancia de la lectura y escritura radica en que se trata de un agente imprescindible en la trasmisión cultural entre generaciones y personas en general, siendo el vehículo que permite adquirir información, cultura y aprendizajes de manera universal” (Lucas Griñan, 2014).

La lectura y la escritura pueden desarrollar no solo las habilidades lectoras, sino también la estrecha relación entre la lectura y otras habilidades básicas, especialmente las habilidades de estudio, las habilidades de procesamiento de información y las habilidades digitales, de modo que se pueda adquirir la alfabetización mediática.

“La lectoescritura constituye una estrategia y un proceso en nuestros alumnos; como estrategia en el proceso de enseñanza-aprendizaje, y como proceso para acercarnos a la comprensión del texto. La lectura y la escritura son elementos inseparables de un mismo proceso mental” (Ferreiro, 2011)

La Real Academia de España (RAE) define la lectura como el acto de leer, que consiste en mirar una página escrita o impresa y comprender el significado de los signos utilizados.

Álvarez (2005) “distingue tres tipos de lectura: oral, en voz alta y visual. En la lectura oral el principal destinatario es nuestro propio cerebro. Su finalidad es que el lector las escuche para su decodificación”.

La Real Academia Española (RAE) define la escritura como la representación de palabras o ideas dibujadas sobre papel u otras superficies con letras u otros símbolos.

El profesor Robert Reel define la escritura como una actividad sensoriomotora en la que la coordinación mano-ojo y la participación de los músculos y articulaciones responsables de la motricidad fina de la mano juegan un papel fundamental. Por tanto, toda ortografía requiere del movimiento de las manos para formar letras y palabras, y la percepción visual requiere de la intervención de distintos grupos musculares, por lo que se define como una actividad sensoriomotora.

En la lectura la comprensión está estrechamente relacionada a la necesidad del autor y del lector en hablar la misma lengua, con el objetivo de llegar juntos al vasto campo del entendimiento; ya que este es el resultado de una compleja actividad en la que el lector debe ejercitar las operaciones y las habilidades relativas a la comprensión lectora y recurrir a muchos tipos de conocimientos, como aclara Colomer (2002)

En el diálogo educativo existen innumerables divergencias teóricas, pero ninguna niega la importancia de aprender a leer y ni la de todas las implicaciones para garantizar el desarrollo cultural del alumno. Por eso, uno de los principales roles de la escuela es producir buenos lectores, y si esto no sucediera, la institución ya no valoraría su papel representativo para penetrar los horizontes culturales, porque "la mayor parte del conocimiento es que la lectura puede entrar". sus vidas. también fuera de la escuela"

Ver la lectura en la perspectiva del desarrollo del acto de leer es descubrir el contexto vivido en el desarrollo de la práctica de la enseñanza de la lectura, lo que obliga a los docentes a poner en el foco de la discusión conceptos arraigados en la enseñanza tradicional. Que argumenta que la centralización del maestro es una forma de promover y controlar la comprensión de los niños sobre el sistema de lectura como una disciplina intelectual que busca activamente el conocimiento sobre el mundo que les rodea.

Para algunos docentes, el propósito de enseñar a leer es permitir que los estudiantes aprendan a leer, explicar, prevenir, responder preguntas después de leer o verificar hechos, en este caso, la enseñanza de la lectura es fragmentada. y de la manera correcta. Al comienzo del proceso de aprendizaje, que se considera habilidad más que conocimiento, se programan actividades visuales, motrices o auditivas en forma de reproducción, imitación y preparación.

2.2.2.2. Calidad de los servicios educativos

La calidad del diseño es una medida de si un producto o servicio cumple con los requisitos del cliente. La silla no es tan buena como su material, pero sí su idoneidad para ver la televisión o trabajar en la oficina. La diferencia radica en el diseño. Las lecciones de matemáticas (planificación para el desempeño futuro) pueden ser efectivas (en términos de calidad) para estudiantes con estilos cognitivos convergentes, pero no suficientes para estudiantes diversos.

Edwards Deming (2000) "calidad es todo aquello que le hiciera ahorrar a la empresa y a su vez cumplir a tiempo con la entrega del producto al cliente, donde esto cambió totalmente la manera de pensar a todos los gerentes al decirles que era más barato hacer un producto nuevo que corregirlo ya que esto conlleva a la pérdida de tiempo y esfuerzo,

también Deming recalcó en observar las líneas de producción para observar donde se estaba dando el problema y por qué y así eliminar lo que estaba fallando de raíz y esto nos ayudaba a conocer cada vez mejor el sistema y aplicarlo a una forma de mejora continua y contribuir a la satisfacción del cliente”.

La finalidad de la educación es desarrollar a las personas en los aspectos técnicos y humanísticos. Esto significa tratar no solo de desarrollar una persona que sea capaz de enfrentar problemas reales con inteligencia y brindarles soluciones efectivas y concretas utilizando o desarrollando la tecnología más adecuada a las necesidades, sino también que la persona esté preparada para comprender a sus pares. Vivir en sociedad y actuar sobre la base de los más altos principios morales. En el mismo sentido, la Comunidad Europea afirma que "...la educación y la formación, adoptando valores comunes, tienen funciones esenciales en relación con la integración social y el desarrollo personal; la transmisión del patrimonio cultural y el aprendizaje autónomo". Esto significa que la formación integral de una persona.

La eficacia según García Hoz (1982) “se refiere a la necesidad de que todos los componentes cumplan adecuadamente su función, de tal forma que el ser humano desarrolle todas sus potencialidades y compense sus limitaciones. La eficacia no refiere, en este caso, a la producción abundante de algo material, refiere más bien a la indicación de que una actividad ha alcanzado el fin que persigue. La eficacia opera sobre la aceptación previa de una finalidad que se deben concretar en objetivos que han de ser alcanzados utilizando los medios que se puedan disponer. La eficacia es, entonces, una relación entre objetivos (extensión, nivel, dificultad) que han de ser alcanzados; los medios de que se dispone (recursos materiales, humanos, tecnológicos) y la organización y métodos

utilizados en la actividad misma de educar y los resultados, que son las metas finales a alcanzar”.

La educación de calidad asegura un mejor nivel de vida para quienes la reciben, y un crecimiento rápido y sostenible en los países que la brindan, porque una mejor educación aumenta la tasa de innovación e invención, pero con mejores métodos de producción y tecnología, asegura la productividad.

Entre los principales factores que determinan la calidad de la educación podemos mencionar los siguientes:

- Recursos físicos disponibles: aulas, salas de recursos, biblioteca, laboratorios, patios, instalaciones deportivas, mobiliario y recursos educativos.
- Recursos humanos: estándares académicos y pedagógicos del personal, experiencia general y actitud del personal, habilidades de trabajo en equipo, proporción alumno-docente, tiempo asignado. Los servicios y actividades que prestan las personas determinan la calidad de cualquier organización. En este sentido, su participación y compromiso es muy importante.
- Dirección, administración y gestión profesional del centro: trabajo de dirección, organización, actividad de servicio, relaciones interpersonales, coordinación y control.
- Docencia: PEI (Programa de Educación Institucional), POA (Programa de Actividades Anuales), evaluación inicial de los estudiantes, adecuación de objetivos y contenidos, atención a la diversidad, métodos de enseñanza, uso de recursos educativos, evaluación, capacitación, implementación de los objetivos del programa.

2.2.3. Análisis Factorial

“Es un método multivariante que pretende expresar p variables observables como una combinación lineal de m variables hipotéticas o latentes, denominadas factores. Tiene una formulación parecida al Análisis de Componentes Principales, pero el modelo que relaciona variables y factores es diferente en Análisis Factorial. Si la matriz de correlaciones existe, las componentes principales también existen, mientras que el modelo factorial podría ser aceptado o no mediante un test estadístico”. (Rokach & Maimon, 2008)

2.2.4. Análisis Factorial Confirmatorio

Esta técnica permite contrastar con un modelo previamente establecido, en el que el investigador determina de antemano el conjunto global de relaciones entre los elementos que lo componen. La factorial confirmatoria asume que el investigador ya puede usar un modelo estructural, datos, preferiblemente basados en una teoría bien establecida, y simplemente necesita confirmar que esta estructura también se puede lograr empíricamente. persiguen un objetivo: “explicar las covarianzas o correlaciones entre un conjunto de variables observadas o medidas a través de un conjunto reducido de variables latentes o factores. (Bollen, 1989)

2.3. Definición de términos básicos.

• Aprendizaje supervisado:

Los datos del conjunto de entrenamiento tienen dos tipos de atributos: los datos en sí mismos y alguna información relacionada con la resolución de problemas.

• Aprendizaje no supervisado:

Los datos del conjunto de entrenamiento solo contienen información de los ejemplos, sin información para guiar el proceso de aprendizaje. En este caso, ningún profesor freelance decide aprender.

• CHAID:

Detección automática de interacciones mediante chi-cuadrado

- **Calidad de servicio:**

Está formado por todos los atributos o cualidades que lo componen y le dan valor. Debido a la intangibilidad, indivisibilidad, heterogeneidad y vida útil de los productos, la calidad del producto es fácil de medir, mientras que la calidad del servicio es difícil de medir.

- **ENAHO:**

Encuesta Nacional de Hogares.

- **Estandarización:**

Es un proceso que facilita la comparación de dos variables en la misma escala de medida con media cero y varianza uno.

- **Servicio educativo:**

Son estructuras de apoyo permanente al profesorado, especialmente a la labor docente del Centro Educativo.

2.4. Formulación de Hipótesis

2.4.1. Hipótesis General

La aplicación de las técnicas de minería de datos ayudará a predecir la calidad del servicio educativo y lectoescritura en estudiantes de primaria del módulo educación de la ENAHO Trimestre 4, 2021.

2.4.2. Hipótesis Específicas

El análisis factorial confirmatorio predecirá la calidad del servicio educativo y lectoescritura en estudiantes de primaria del módulo educación de la ENAHO Trimestre 4, 2021.

El análisis mediante los árboles de decisión predecirá la calidad del servicio educativo y lectoescritura en estudiantes de primaria del módulo educación de la ENAHO Trimestre 4, 2021.

2.5. Identificación de Variables

2.5.1. Variables independientes

Técnicas de minería de datos.

2.5.2. Variables dependientes

Predecir la calidad del servicio educativo y lectoescritura.

2.6. Definición Operacional de variables e indicadores

Tabla 1. Operacionalización de variables

VARIABLE	DIMENSIONES	INDICADORES
INDEPENDIENTE	Análisis factorial confirmatorio	Infraestructura
Técnicas de minería de datos	Análisis mediante los árboles de decisión para ver la calidad de la educación	Equipamiento Enseñanza docente Materiales Apoyo de padres
DEPENDIENTE	Desarrollo	Lectura
Medir la calidad del servicio educativo y lectoescritura	Habilidades Capacidad de expresión	Escritura

CAPITULO III

METODOLOGÍA Y TECNICAS DE INVESTIGACIÓN

3.1. Tipo de investigación

“El tipo de investigación es aplicada guarda íntima relación con la básica, pues depende de los descubrimientos y avances de la investigación básica y se enriquece con ellos, pero se caracteriza por su interés en la aplicación, utilización y consecuencias prácticas de los conocimientos. La investigación aplicada busca el conocer para hacer, para actuar, para construir o para modificar” (Hernández Sampieri et al., 2014)

3.2. Nivel de investigación

“La presente investigación es de nivel explicativo van más allá de la descripción. Están dirigidos a responder a las causas de los eventos físicos o sociales. Se centran en explicar por qué ocurre un fenómeno y en qué condiciones se da éste. Las investigaciones explicativas son más estructuradas”.(Hernández Sampieri, 2014)

3.3. Métodos de investigación

La investigación que realizare es de método analítico-inductivo utiliza el razonamiento para obtener conclusiones que parten de hechos aceptados como válidos, para llegar a conclusiones.(Hernández Sampieri & Mendoza Torres, 2018)

3.4. Diseño de investigación

El presente trabajo de tesis es correlacional este tipo de estudios tienen como propósito medir el grado de relación que exista entre dos o más conceptos o variables, miden cada una de ellas y después, cuantifican y analizan la vinculación. Tales correlaciones se sustentan en hipótesis sometidas a prueba. (Hernández Sampieri & Mendoza Torres, 2018)

3.5. Población y muestra

3.5.1. Población

Residentes de viviendas particulares del área urbana y rural del país que son estudiantes del nivel primario.

3.5.2. Muestra

Consta de 26589 encuestas ENAHO en 4° trimestre 2021, consta de encuestas que cumplen con las características de investigación, limpieza y selección de productos de minería de datos; es decir. los encuestados eran estudiantes de primaria.

3.6. Técnicas e instrumentos de recolección de datos

Para el presente trabajo de tesis se utiliza la técnica de búsqueda de información. Estos instrumentos se utilizan para recopilar información para el desarrollo de la investigación.

Técnicas:

- Base de datos de la ENAHO del cuarto trimestre del año 2021.

Instrumentos:

- Encuesta nacional de hogares del INEI.

3.7. Selección, validación y confiabilidad de los instrumentos de investigación.

Selección de datos.

“Uno de los problemas a solucionar fue la fiabilidad del instrumento de medición de la calidad del servicio educativo peruano desarrollado por el INEI, como parte de la encuesta nacional de hogares – ENAHO, así como, el perfil de dichos indicadores de calidad asociados a la lectoescritura de los estudiantes de educación primaria” (Anaya Carrillo & Bedón Trujillo, 2018)

Confiabilidad de datos

Los datos fueron recolectados de la página del INEI, en el siguiente enlace presentado de microdatos (<http://inei.inei.gob.pe/microdatos/index.htm>) con la base de datos de la ENAHO – Metodología actualizada -Condiciones de Vida y Pobreza del año 2021 del trimestre 4, se cuenta con 8 módulos:

Tabla 2. Módulo de Estudio

Código Módulo	Módulo
1	Características de la Vivienda y del Hogar
2	Características de los Miembros del Hogar
3	Educación
4	Salud
5	Empleo e Ingresos
7	Gastos en Alimentos y Bebidas (Módulo 601)
37	Programas Sociales (Miembros del Hogar)
85	Gobernabilidad, Democracia y Transparencia

De los mencionados, para la investigación de la minería de datos se utilizará los siguientes módulos.

Código Módulo	Módulo
2	Características de los Miembros del Hogar
3	Educación

Donde:

Características de los Miembros del Hogar:

Compuesta por la base de datos de 39 campos de variable y 29344 datos.

Educación:

Compuesta por la base de datos de 508 campos de variables y 26589 datos.

Variables que seleccionaremos son:

- Sabe leer y escribir
- Calidad del servicio educativo en infraestructura
- Calidad del servicio educativo en equipamiento
- Calidad del servicio educativo en enseñanza de los maestros
- Calidad del servicio educativo en dotación de materiales educativos
- Calidad del servicio educativo en apoyo a la participación de los padres de familias
- Idioma que habla
- Tipo de centro de estudios
- Resultados académicos del año pasado
- Grado matriculado
- Uso de internet
- Sexo del estudiante
- Edad del estudiante
- Región donde habita

Procedimientos de recolección de datos

Durante el proceso de datos, los tres archivos anteriores se combinaron para luego filtrar y seleccionar las variables a estudiar. Se eliminaron las entradas con campos en blanco, así como aquellas que no llegaban a la edad establecida para los estudiantes de primaria. Después de la limpieza de datos, la base de datos consta de 15 campos y la confiabilidad de los datos indica 4062 casos.

Para el objetivo de validación de la calidad del servicio educativo peruano, mediante el análisis factorial confirmatorio por medio del programa AMOS se requirió las siguientes variables formateadas a la misma escala de medición de la ENAHO 2016:

- Infraestructura (paredes, techos, pisos)
- Equipamiento (Carpetas, Laboratorios, Computadoras, etc.)
- Enseñanza de los maestros
- Dotación de materiales educativos (Libros, Láminas, maquetas, etc.)
- Apoyo a la participación de los padres de familias (Promueven la participación de los padres de familia, escuela de padre)

Mientras que para los árboles de decisión se hicieron uso de las variables en su forma original observada mediante la ENAHO 2016:

- Sabe leer y escribir
- Idioma o lengua materna que aprendió en su niñez
- El resultado que obtuvo en el 2015
- Grado de Estudios en el que está matriculado
- Sexo del estudiante
- Edad en años cumplidos del estudiante
- Región donde habita el estudiante
- Tipo de Centro de Estudios al que asistió en el 2015
- Uso del servicio de Internet por cabina pública

3.8. Técnicas de procesamiento y análisis de datos.

Para la técnica de procesamiento de datos realizaremos mediante las encuestas y el software SPSS y R-Project.

3.9. Tratamiento Estadístico.

El análisis e interpretación estadística de los datos de este estudio se realizó utilizando medidas de tendencia central y dispersión, presentadas en forma de diagramas estadísticos y ordenadas en tablas estadísticas. Además, se utilizó el análisis de minería de datos.

3.10. Orientación ética filosófica y epistémica.

La información que la empresa ha otorgado fue usada solamente para fines de esta investigación. Los investigadores se comprometen a respetar la corrección de los resultados y la fiabilidad de los datos presentados por la Encuesta nacional de hogares del INEI.

CAPITULO IV

RESULTADOS Y DISCUSIÓN

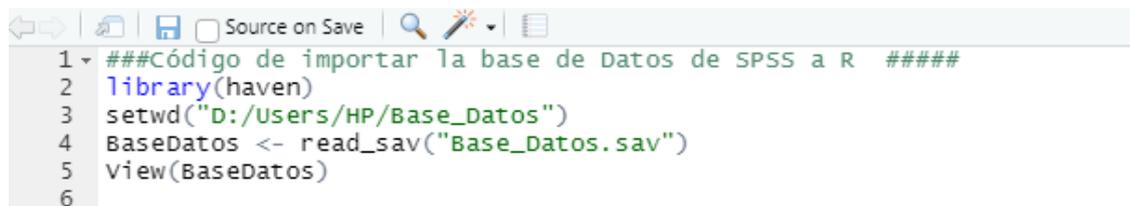
4.1. Descripción del trabajo de campo

4.1.1. Modelado

La minería de los datos de estudio se realizó por medio del paquete R-Project version 4.2.2, los datos preprocesados se almacenaron en SPSS v26, cuyos registros se limpiaron, normalizaron y exportaron a RProject..

El código se importó desde la base de datos de SPSS a R-Project.

Figura 8.
Código de importación de base de datos.



```
1  ###Código de importar la base de Datos de SPSS a R  #####
2  library(haven)
3  setwd("D:/Users/HP/Base_Datos")
4  BaseDatos <- read_sav("Base_Datos.sav")
5  View(BaseDatos)
6
```

Obteniendo la Base de Datos en R-Project.

Figura 9.
Base de Datos R-Project

	AÑO	MES	CONGLOME	VIVIENDA	HOGAR	UBIGEO	DOMINIO	ESTRATO	CODPERSO
	Año de la encuesta	Mes de ejecución de la encuesta	Número de conglomerado	Número de selección de la vivienda	Número secuencial del Hogar	Ubicación geográfica	Dominio geográfico	Estrato geográfico	Número de orden de la persona
1	2021	10	005119	045	11	010705	7	7	05
2	2021	10	005119	048	11	010705	7	7	04
3	2021	10	005119	048	11	010705	7	7	05
4	2021	10	005119	049	11	010705	7	7	05
5	2021	10	005119	100	11	010705	7	7	03
6	2021	10	005119	101	11	010705	7	7	03
7	2021	10	005132	068	11	010512	4	7	03
8	2021	10	005132	160	11	010512	4	7	03
9	2021	10	005132	161	11	010512	4	7	03
10	2021	10	005137	009	11	010513	4	8	03
11	2021	10	005137	012	11	010513	4	8	04
12	2021	10	005137	056	11	010513	4	8	05
13	2021	10	005137	060	11	010513	4	8	05
14	2021	10	005196	002	11	021809	2	2	03
15	2021	10	005196	033	11	021809	2	2	03
16	2021	10	005196	033	11	021809	2	2	04
17	2021	10	005196	033	11	021809	2	2	05
18	2021	10	005196	097	11	021809	2	2	04

Del total de variables incluidas en la base de datos, sólo haremos uso a aquellas relacionadas con los objetivos de estudio, tales como:

- 302. ¿Sabe leer y escribir?
- 308B1. Infraestructura (paredes, techos, pisos)
- 308B2. Equipamiento (Carpetas, Laboratorios, Computadoras, etc.)
- 308B3. Enseñanza de los maestros
- 308B4. Dotación de materiales educativos (Libros, Láminas, maquetas, etc.)
- 308B5. Apoyo a la participación de los padres de familia (Promueven la participación de los padres de familia, escuela de padres, etc.)

4.2. Presentación, análisis e interpretación de resultados

A continuación, se presentará los resultados obtenidos por R-Project de las tablas de la Base de Datos ENAHO Trimestre 4, 2021.

1. ¿Sabe leer y escribir? Respuestas por los estudiantes de educación primaria del Módulo Educación ENAHO Trimestre 4, 2021.

¿Sabe leer y escribir?	n	%
Si	2747	67.63
No	1315	32.37
Total	4062	100

Tabla 3. Módulo Educación ENAHO Trimestre 4, 2021

Interpretación: De la muestra de estudio Tabla N°3 Según las observaciones, el 67.63% de los alumnos de primaria saben leer y escribir, mientras que el 32.37% restante de los alumnos de primaria no saben leer ni escribir

2. ¿Sabe leer y escribir? Respuestas por los estudiantes de educación primaria según idioma o lenguas materna que aprendió en su niñez del Módulo Educación ENAHO Trimestre 4, 2021.

Idioma o lengua materna que aprendió en su niñez	¿Sabe leer y escribir?			
	Si		No	
	n	%	n	%
Quechua	139	5.06	66	5.02
Aimara	5	0.18	2	0.15
Otra lengua nativa	26	0.95	23	1.75
Castellano	2510	91.37	1180	89.73
Portugués	3	0.11	1	0.08
Otra lengua extranjera	1	0.04	1	0.08
No escucha/no habla	2	0.07	7	0.53
Lengua de señas peruanas	0	0.00	2	0.15
Ashaninka	13	0.47	11	0.84
Awajún/Aguarun	37	1.35	14	1.06
Shipibo – Konibo	10	0.36	5	0.38
Shawi / Chayahuita	1	0.04	3	0.23
Matsigenka / Machiguenga	0	0.00	0	0.00
Achuar	0	0.00	0	0.00
Total	2747	100	1315	100

Tabla 4. Módulo Educación Idioma ENAHO Trimestre 4, 2021

Interpretación: De la muestra de estudio Tabla N°4 Se ha observado que el 91.37% de los alumnos de primaria que saben leer y escribir han aprendido español o su lengua materna de niños, mientras que el

89.73% de los alumnos de primaria que no saben leer no han aprendido español o su lengua materna de niños.

3. ¿Sabe leer y escribir? Respuestas por los estudiantes de educación primaria según centro de estudios del Módulo Educación ENAHO Trimestre 4, 2021.

Centro de estudios	¿Sabe leer y escribir?			
	Si		No	
	n	%	n	%
Estatal	2388	86.96	1102	90.55
No estatal	358	13.04	115	9.45
Total	2746	100	1217	100

Tabla 5. Módulo Educación Centro de Estudios ENAHO Trimestre 4, 2021

Interpretación: De la muestra de estudio Tabla N°5 Se observó que el 86.96% de los alumnos de primaria que saben leer y escribir estudiaron en centros educativos gubernamentales, mientras que el 90.55% de los alumnos de primaria que no saben leer y escribir estudiaron en centros educativos no gubernamentales.

4. ¿Sabe leer y escribir? Respuestas por los estudiantes de educación primaria según resultado que obtuvo el año pasado del Módulo Educación ENAHO Trimestre 4, 2021.

Resultado que obtuvo el año pasado	¿Sabe leer y escribir?			
	Si		No	
	n	%	n	%
Aprobado	2487	91.47	264	22.06
Desaprobado	7	0.00	10	0.84
Retirado	5	0.18	7	0.58
Otro	220	8.09	916	76.52
No aprueba, ni desaprueba (inicial)	0	0.00	0	0.00
Total	2719	100	1197	100

Tabla 6. Módulo Educación Resultado ENAHO Trimestre 4, 2021

Interpretación: De la muestra de estudio Tabla N°6 Se observó que el año pasado el 91.47% de los alumnos de primaria que sabían leer y

escribir aprobaron sus estudios, mientras que el 22.06% de los alumnos de primaria que no sabían leer y escribir aprobaron sus estudios.

5. ¿Sabe leer y escribir? Respuestas por los estudiantes de educación primaria según resultado que obtuvo el año pasado según su grado del Módulo Educación ENAHO Trimestre 4, 2021.

Grado de Estudios	¿Sabe leer y escribir?			
	Si		No	
	n	%	n	%
Primero	177	6.62	309	53.09
Segundo	358	13.39	158	27.15
Tercero	507	18.97	65	11.17
Cuarto	523	19.57	32	5.50
Quinto	539	20.16	18	3.09
Sexto	569	21.29	0	0.00
Total	2673	100	582	100

Tabla 7. Módulo Educación Grado de Estudios ENAHO Trimestre 4, 2021

Interpretación: De la muestra de estudio Tabla N°7 según las observaciones, el 21.29% de los alumnos de primaria que saben leer y escribir están estudiando sexto grado, mientras que el 53.09% de los alumnos de primaria que no saben leer y escribir están estudiando primer grado.

6. ¿Sabe leer y escribir? Respuestas por los estudiantes de educación primaria según sexo del Módulo Educación ENAHO Trimestre 4, 2021.

Sexo	¿Sabe leer y escribir?			
	Si		No	
	n	%	n	%
Hombre	1375	50.05	706	53.69
Mujer	1372	49.95	609	46.31
Total	2747	100	1315	100

Tabla 8. Módulo Educación según sexo ENAHO Trimestre 4, 2021

Interpretación: De la muestra de estudio Tabla N°8 Según las observaciones, el 50.95% de los alumnos de primaria que saben leer

y escribir son hombres, mientras que el 53.09% de los alumnos de primaria que no saben leer ni escribir son hombres.

7. ¿Sabe leer y escribir? Respuestas por los estudiantes de educación primaria según edad del Módulo Educación ENAHO Trimestre 4, 2021.

Edad	¿Sabe leer y escribir?			
	Si		No	
	n	%	n	%
5 años	31	1.13	429	32.62
6 años	81	2.95	402	30.57
7 años	229	8.34	242	18.40
8 años	422	15.36	119	9.05
9 años	468	17.04	51	3.88
10 años	507	18.46	31	2.36
11 años	531	19.33	22	1.67
12 años	383	13.94	8	0.61
13 años	58	2.11	2	0.15
14 años	26	0.95	6	0.46
15 años	11	0.40	3	0.23
Total	2747	100	1315	100

Tabla 9. Módulo Educación según edad ENAHO Trimestre 4, 2021

Interpretación: De la muestra de estudio Tabla N°9 se observa que el 18.46% de los estudiantes de educación primaria que saben leer y escribir tienen 10 años de edad, mientras que el 30.57% de los estudiantes de educación primaria que no saben leer y ni escribir tienen 6 años de edad.

8. ¿Sabe leer y escribir? Respuestas por los estudiantes de educación primaria según región del Módulo Educación ENAHO Trimestre 4, 2021.

Región	¿Sabe leer y escribir?			
	Si		No	
	n	%	n	%
Costa Norte	440	16.02	165	12.55
Costa Centro	239	8.70	108	8.21
Costa Sur	175	6.37	51	3.88
Sierra Norte	174	6.33	87	6.62
Sierra Centro	390	14.20	229	17.41
Sierra Sur	302	10.99	134	10.19
Selva	737	26.83	440	33.46
Lima Metropolitana	290	10.56	101	7.68
Total	2747	100	1315	100

Tabla 10. Módulo Educación según región ENAHO Trimestre 4, 2021

Interpretación: De la muestra de estudio Tabla N°10 Se observó que el 31.53% de los alumnos de primaria que sabían leer y escribir eran de la zona de Sera mientras que el 33.46% de los alumnos de primaria que no sabían leer ni escribir eran de la zona de la selva.

9. ¿Sabe leer y escribir? Respuestas por los estudiantes de educación primaria según servicio educativo: Infraestructura en el Módulo Educación ENAHO Trimestre 4, 2021.

Infraestructura (paredes, techos, pisos)	¿Sabe leer y escribir?			
	Si		No	
	n	%	n	%
Muy malo	24	0.89	11	0.89
Malo	335	12.44	174	14.12
Bueno	2110	78.35	941	76.38
Muy bueno	114	4.23	40	3.25
No sabe	110	4.08	66	5.36
Total	2693	100	1232	100

Tabla 11. Módulo Educación según infraestructura ENAHO Trimestre 4, 2021

Interpretación: De la muestra de estudio Tabla N°11 según las observaciones, el 78.35% de los estudiantes de primaria con habilidades de lectura y escritura tienen una buena calificación de los servicios educativos relacionados con la infraestructura, y el 76.38% de los estudiantes de primaria con habilidades de lectura y escritura tienen una buena calificación de los servicios educativos relacionados con la infraestructura.

10. ¿Sabe leer y escribir? Respuestas por los estudiantes de educación primaria según servicio educativo: Equipamiento en el Módulo Educación ENAHO Trimestre 4, 2021.

Equipamiento (Carpetas, Laboratorios, Computadoras, etc.)	¿Sabe leer y escribir?			
	Si		No	
	n	%	n	%
Muy malo	27	1.00	9	0.73
Malo	402	14.93	202	16.40
Bueno	1939	72.00	837	67.94
Muy bueno	68	2.53	30	2.44
No sabe	257	9.54	154	12.50
Total	2693	100	1232	100

Tabla 12. Módulo Educación según equipamiento ENAHO Trimestre 4, 2021

Interpretación: De la muestra de estudio Tabla N°12 se observa que el 72% de los estudiantes de educación primaria que saben leer y escribir califican el servicio educativo en el equipamiento como bueno, mientras que el 67.94% de los estudiantes de educación primaria que no saben leer y ni escribir califican el servicio educativo en el equipamiento como bueno.

11. ¿Sabe leer y escribir? Respuestas por los estudiantes de educación primaria según servicio educativo: Enseñanza de los Maestros en el Módulo Educación ENAHO Trimestre 4, 2021.

Enseñanza de los maestros	¿Sabe leer y escribir?			
	Si		No	
	n	%	n	%
Muy malo	27	1.00	19	1.54
Malo	268	9.95	173	14.04
Bueno	2170	80.58	972	78.90
Muy bueno	214	7.95	57	4.63
No sabe	14	0.52	11	0.89
Total	2693	100	1232	100

Tabla 13. Módulo Educación según enseñanza ENAHO Trimestre 4, 2021

Interpretación: De la muestra de estudio Tabla N°13 se observa que el 80.58% de los estudiantes de educación primaria que saben leer y escribir “califican el servicio educativo en la enseñanza de los maestros

como bueno, mientras que el 78.90% de los estudiantes de educación primaria que no saben leer y ni escribir califican el servicio educativo en la enseñanza de los maestros como bueno”.

12. ¿Sabe leer y escribir? Respuestas por los estudiantes de educación primaria según servicio educativo: dotación de materiales educativos en el Módulo Educación ENAHO Trimestre 4, 2021.

Dotación de materiales educativos (Libros, Láminas, maquetas, etc.)	¿Sabe leer y escribir?			
	Si		No	
	n	%	n	%
Muy malo	9	0.33	3	0.24
Malo	187	6.94	118	9.58
Bueno	2309	85.74	1005	81.57
Muy bueno	101	3.75	33	2.68
No sabe	87	3.23	73	5.93
Total	2693	100.00	1232	100

Tabla 14. Módulo Educación según dotación de materiales ENAHO Trimestre 4, 2021

Interpretación: De la muestra de estudio Tabla N°14 se observa que el 85.74% de los estudiantes de educación primaria que saben leer y escribir los servicios educativos que apoyan la participación de los padres se califican como buenos, y el 73,21% de los estudiantes de primaria que no saben leer ni escribir califican los servicios educativos que apoyan la participación de los padres como buenos.

13. ¿Sabe leer y escribir? Respuestas por los estudiantes de educación primaria según servicio educativo: Apoyo a la participación de los padres de familia en el Módulo Educación ENAHO Trimestre 4, 2021.

Apoyo a la participación de los padres de familia (Promueven la participación de los padres de familia, escuela de padres, etc)	¿Sabe leer y escribir?			
	Si		No	
	n	%	n	%
Muy malo	11	0.41	7	0.57
Malo	291	10.81	163	13.23
Bueno	2056	76.35	902	73.21
Muy bueno	112	4.16	28	2.27
No sabe	223	8.28	132	10.71
Total	2693	100.00	1232	100

Tabla 15. Módulo Educación según apoyo de los padres ENAHO Trimestre 4, 2021

Interpretación: De la muestra de estudio Tabla N°15 se observa que el 76.35% de los estudiantes de educación primaria que saben leer y escribir los servicios educativos que apoyan la participación de los padres se califican como buenos, y el 73.21% de los estudiantes de primaria que no saben leer ni escribir califican los servicios educativos que apoyan la participación de los padres como buenos.

14. Análisis factorial confirmatorio respecto a la calidad del servicio educativo

El siguiente análisis tiene como propósito confirmar que los 5 puntos de medición explican la calidad de los servicios educativos, la cual estará determinada por las siguientes características iniciales del modelo.

Variables a observar:

- Infraestructura (P308B1)
- Equipamiento (P308B2)
- Enseñanza de maestros (P308B3)
- Dotación de materiales educativos (P308B4)
- Apoyo a la participación de padres de familia (P308B5)

Variable latente: factor común

- Servicio educativo (ξ).

Variable latente: factores específicos

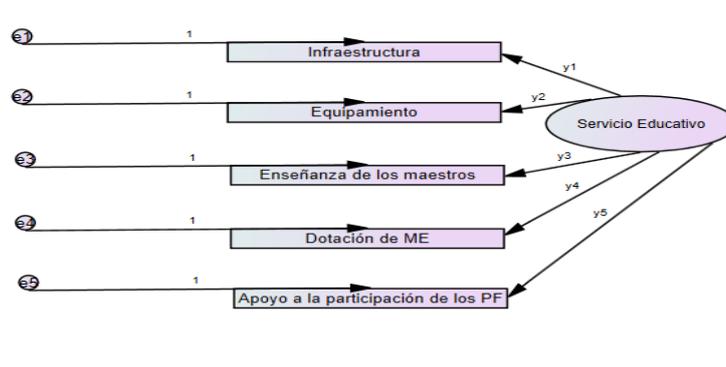
- Factor Especifico de infraestructura (δ_1).
- Factor Especifico de equipamiento (δ_2).
- Factor Especifico de enseñanza docente(δ_3).
- Factor Especifico de dotación de materiales educativos(δ_4).
- Factor Especifico de apoyo a la participación de padres de familia (δ_5).

Intensidad de relación (carga factorial) entre el servicio educativo y:

- La infraestructura (λ_1).
- El equipamiento (λ_2).
- La enseñanza docente (λ_3).
- La dotación de materiales educativos (λ_4).
- La participación de padres de familia (λ_5).

Modelo de Análisis factorial confirmatorio.

Figura 10. Modelos de Análisis Factorial en Amos SPSS



Modelo estructural se realizó mediante la estimación de distribución libre asintótica. Los resultados se muestran a continuación para lo cual se trabajo con las variables mencionadas:

Figura 11. Análisis Factorial de Servicio Educativo

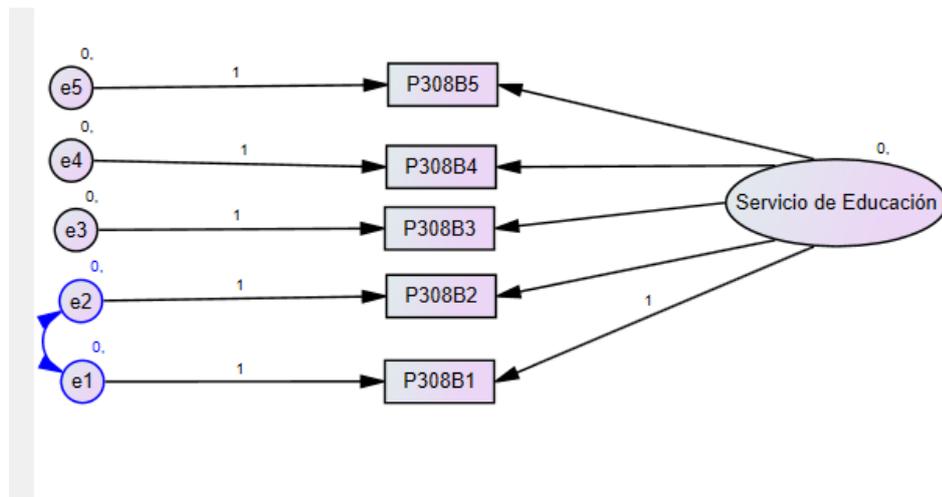


Diagrama de ruta del modelo de un factor (No estandarizado) lo presentamos a continuación:

Figura 12. Diagrama de ruta de factor no estandarizado

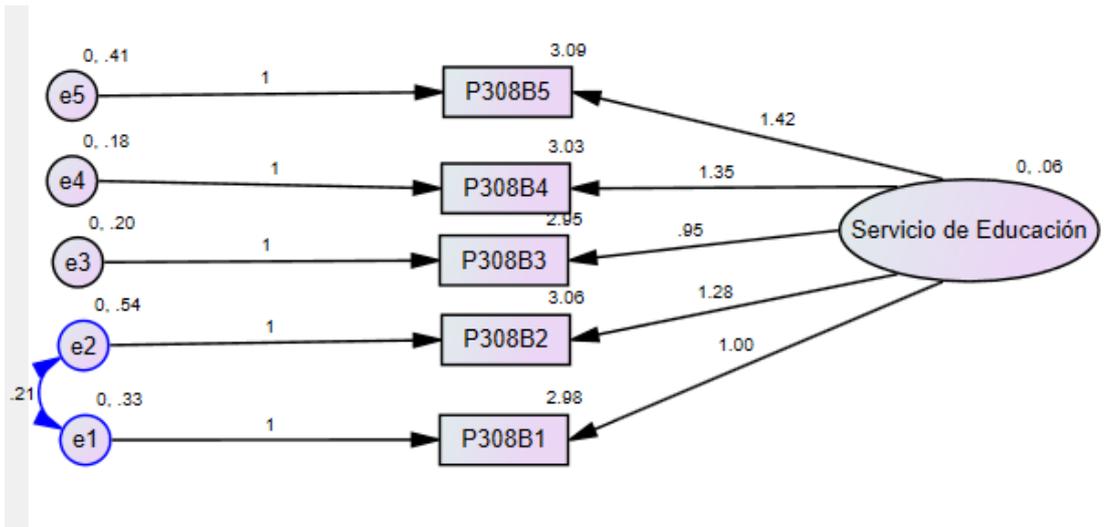
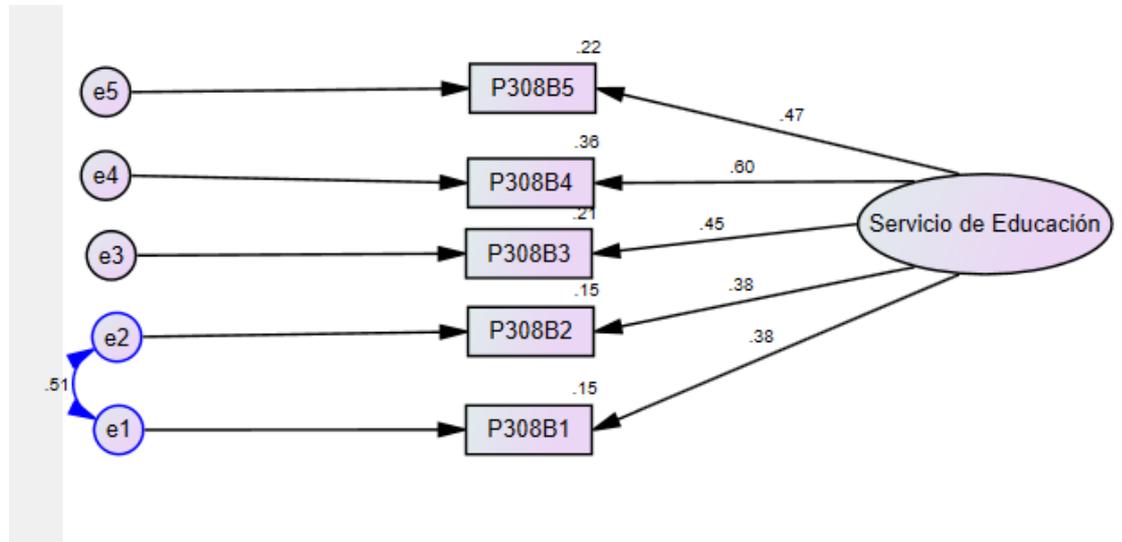


Diagrama de ruta del modelo de un factor (Estandarizado) lo presentamos a continuación:

Figura 13. Diagrama de ruta de factor estandarizado



Así los modelos quedan descritos de la siguiente manera:

- Infraestructura = 1*ServicioEducación+0.33
- Equipamiento =1.28*ServicioEducación+0.54
- EnseñanzaMaestros =0.95*ServicioEducación+0.20
- MaterialesE =1.35*ServicioEducación+0.18

- ParticipaciónPF = 1.42*ServicioEducación+0.41

Estimación del modelo:

Número de parámetros conocidos	:20
Número de parámetros desconocidos	:16
Diferencia	:4

Por lo que se dice que el modelo se encuentra sobre identificado.

Grados de libertad (g.l.) = 4

Índice descriptivo de bondad de ajuste:

$$\chi^2 = 11.159, sig(p) = 0.025$$

$$H_0: S = \Sigma$$

$$H_1: S \neq \Sigma$$

Por lo que el índice descriptivo de bondad de ajuste indica que el modelo reproduce el comportamiento de los datos observados de un modo razonablemente bueno.

Índice comparativo $\chi^2/g.l.$:

Como el valor es igual a 2.79, es mayor o igual que 2, entonces significa que el modelo presenta un buen ajuste.

Error cuadrático medio de aproximación (RMSEA):

RMSEA = 0.021, representando un buen ajuste al modelo que se requiere.

Índice de ajuste normado (NFI):

NFI= 0.996 Esto representa un buen ajuste del modelo y una reducción proporcional de la función ajustada de cero al modelo propuesto.

Índice de ajuste comparativo (CFI):

CFI: 0.997, representando un buen ajuste al modelo que se requiere

Pesos de la estimación:

	Estimate	S.E.	C.R.	P	Label
P308B1 <--- r	1.000				
P308B2 <--- r	1.283	.067	19.119	***	par_1
P308B3 <--- r	.955	.069	13.807	***	par_2
P308B4 <--- r	1.351	.095	14.203	***	par_3
P308B5 <--- r	1.420	.102	13.942	***	par_4

-----Los coeficientes estimados son significativos -----

Pesos de la estimación estandarizado

	Estimate
P308B1 <--- r	.384
P308B2 <--- r	.385
P308B3 <--- r	.454
P308B4 <--- r	.601
P308B5 <--- r	.468

Como se observa, la variable (P308B4) "dotación de materiales educativos" se encuentra mejor ponderado frente a las demás variables.

	Estimate	S.E.	C.R.	P	Label
e1 <--> e2	.214	.009	23.673	***	par_5

Los modelos de covarianza significativa entre e1 y e2 tienen un coeficiente no estandarizado de 0.214, mientras que la misma relación de covarianza tiene un coeficiente estandarizado igual a 0.511.

	Estimate
e1 <--> e2	.511

Asimismo, la varianza del modelo de validación es significativa incluso al 99.99% de confianza.

4.2.1. Árbol de clasificación respecto a los componentes de la calidad del servicio educativo frente a la lectoescritura.

Desarrollando el árbol de clasificación que tendrán las siguientes características:

	Estimate	S.E.	C.R.	P	Label
r	.057	.006	8.953	***	par_11
e1	.327	.009	37.994	***	par_12
e2	.537	.014	37.984	***	par_13
e3	.199	.006	35.100	***	par_14
e4	.183	.007	24.407	***	par_15
e5	.409	.012	34.317	***	par_16

Variables explicativas:

- Infraestructura (P308B1)
- Equipamiento (P308B2)
- Enseñanza del docente (P308B3)
- Dotación de materiales educativos (P308B4)
- Apoyo a la participación de padres de familia (P308B5)
- Edad (P208A)
- Sexo (P207)
- Región (DOMINIO)

Variable Dependiente:

- ¿Sabe leer y escribir? (P302): con respuestas Si = 1 y No = 2.

Para la elaboración de los árboles de clasificación se hizo uso del método CHAID, que a continuación se detalla:

Resumen del modelo

Especificaciones	Método de crecimiento	CHAID
	Variable dependiente	¿Sabe leer y escribir? - Respuesta espontánea
	Variables independientes	En su opinión, ¿El servicio educativo que brinda el centro de enseñanza donde asiste actualmente en relación a: Infraestructura(paredes, techos, pisos) es?, En su opinión, ¿El servicio educativo que brinda el centro de enseñanza donde asiste actualmente en relación a: Equipamiento (Carpetas, Laboratorios, Computadoras, etc.) es?, En su opinión, ¿El servicio educativo que brinda el centro de enseñanza donde asiste actualmente en relación a: Enseñanza de los maestros es?, En su opinión, ¿El servicio educativo que brinda el centro de enseñanza donde asiste actualmente en relación a: Dotación de materiales educativos (Libros, Láminas, maquetas, etc.) es?, En su opinión, ¿El servicio educativo que brinda el centro de enseñanza donde asiste actualmente en relación a: Apoyo a la participación de los padres de familiar (Promueven participación de los padres de familia, escuela de padres) es?, Sexo, Dominio geográfico, ¿Qué edad tiene en años cumplidos ? (En años)
	Validación	Muestra de división
	Máxima profundidad del árbol	3
Resultados	Casos mínimos en nodo padre	100
	Casos mínimos en nodo hijo	50
	Variables independientes incluidas	¿Qué edad tiene en años cumplidos ? (En años), Dominio geográfico, Sexo, En su opinión, ¿El servicio educativo que brinda el centro de enseñanza donde asiste actualmente en relación a: Apoyo a la participación de los padres de familiar (Promueven participación de los padres de familia, escuela de padres) es?, En su opinión, ¿El servicio educativo que brinda el centro de enseñanza donde asiste actualmente en relación a: Equipamiento (Carpetas, Laboratorios, Computadoras, etc.) es?
	Número de nodos	26
	Número de nodos terminales	17
	Profundidad	3

La muestra de división son 80% y 20% para el entrenamiento y la comprobación respectivamente.

Costes de clasificación errónea

Observado	Pronosticado	
	Si	No
Si	,000	1,000
No	1,000	,000

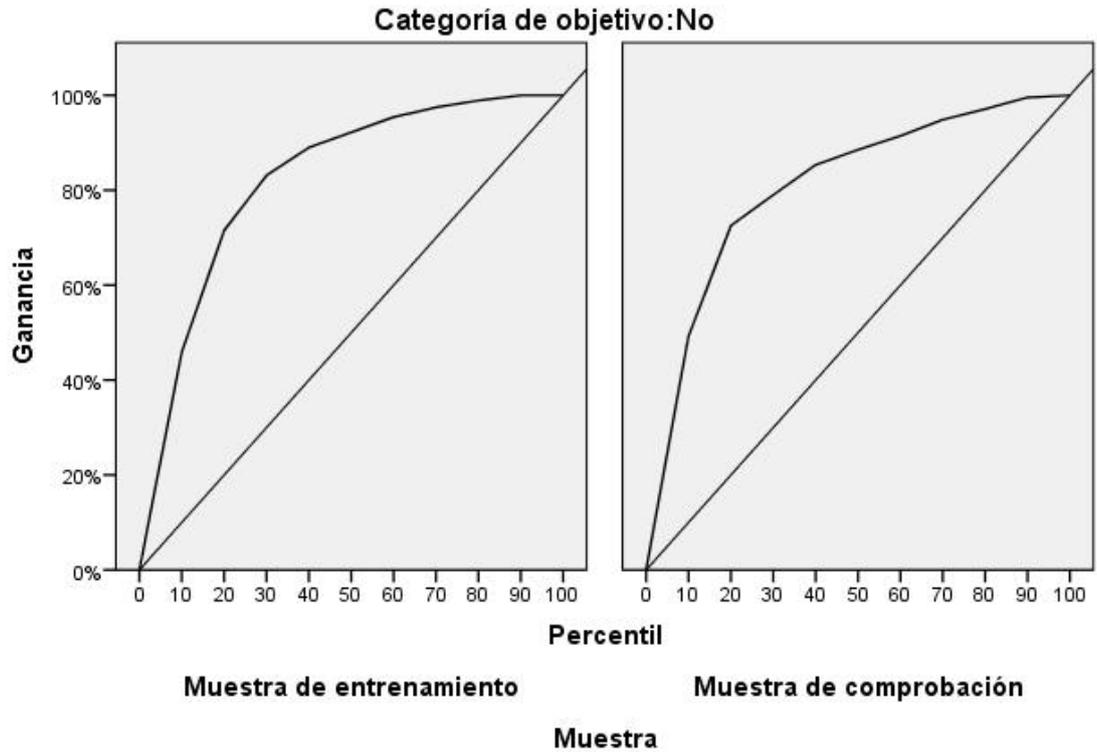
Variable dependiente: ¿Sabe leer y escribir? - Respuesta espontánea

La recodificación del saber leer y escribir: Si = 0 y No = 1.

Los resultados de la prueba de aprendizaje se muestran en la siguiente figura: Los niños que no saben leer y escribir se caracterizan por ser menores de seis años y pertenecer a una región (costa, sierra y selva), mientras que los niños que saben escribir se caracterizan por: La lectura y escritura muestra indicios de que los mayores de 8 años muestran puntajes altos en la dimensión de instalaciones del servicio educativo.

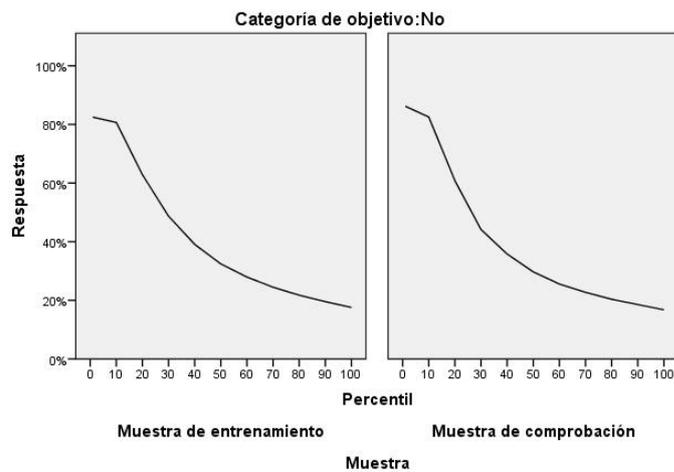
Ejemplo	Nodo	Si		No		Total		Categoría pronosticada	Nodo padre	Variable	Variable independiente primaria			Valores de división
		N	Porcentaje	N	Porcentaje	N	Porcentaje				Sig.*	Chi-cuadrado	gl	
Entrenamiento	0	2719	69,4%	1197	30,6%	3916	100,0%	SI						
	1	30	7,4%	375	92,6%	405	10,3%	No	0	¿Que edad tiene en años cumplidos ? (En años)	,000	2065,013	6	<= 5.0
	2	79	17,5%	373	82,5%	452	11,5%	No	0	¿Que edad tiene en años cumplidos ? (En años)	,000	2065,013	6	(5.0, 6.0]
	3	227	49,2%	234	50,8%	461	11,8%	No	0	¿Que edad tiene en años cumplidos ? (En años)	,000	2065,013	6	(6.0, 7.0]
	4	420	78,9%	112	21,1%	532	13,6%	SI	0	¿Que edad tiene en años cumplidos ? (En años)	,000	2065,013	6	(7.0, 8.0]
	5	467	90,7%	48	9,3%	515	13,2%	SI	0	¿Que edad tiene en años cumplidos ? (En años)	,000	2065,013	6	(8.0, 9.0]
	6	1033	95,5%	49	4,5%	1082	27,6%	SI	0	¿Que edad tiene en años cumplidos ? (En años)	,000	2065,013	6	(9.0, 11.0]
	7	463	98,7%	6	1,3%	469	12,0%	SI	0	¿Que edad tiene en años cumplidos ? (En años)	,000	2065,013	6	> 11.0
	8	8	3,3%	234	96,7%	242	6,2%	No	1	Dominio geográfico	,016	14,749	1	Selva; Sierra Norte; Sierra Centro; Sierra Sur
	9	22	13,5%	141	86,5%	163	4,2%	No	1	Dominio geográfico	,016	14,749	1	Costa Centro; Costa Sur; Lima Metropolitana; Costa Norte
	10	36	12,2%	259	87,8%	295	7,5%	No	2	Dominio geográfico	,007	16,382	1	Selva; Sierra Norte; Sierra Centro; Costa Norte
	11	43	27,4%	114	72,6%	157	4,0%	No	2	Dominio geográfico	,007	16,382	1	Costa Centro; Sierra Sur; Lima Metropolitana; Costa Norte
	12	81	37,7%	134	62,3%	215	5,5%	No	3	Dominio geográfico	,001	26,777	2	Selva; Sierra Norte; Sierra Centro
	13	92	54,4%	77	45,6%	169	4,3%	SI	3	Dominio geográfico	,001	26,777	2	Costa Centro; Sierra Sur; Costa Norte
	14	54	70,1%	23	29,9%	77	2,0%	SI	3	Dominio geográfico	,001	26,777	2	Costa Sur; Lima Metropolitana
	15	183	74,4%	63	25,6%	246	6,3%	SI	4	Dominio geográfico	,000	38,002	2	Selva; Sierra Norte; Sierra Sur
	16	186	91,6%	17	8,4%	203	5,2%	SI	4	Dominio geográfico	,000	38,002	2	Costa Centro; Costa Sur; Lima Metropolitana; Costa Norte
	17	51	61,4%	32	38,6%	83	2,1%	SI	4	Dominio geográfico	,000	38,002	2	Sierra Centro
	18	246	98,1%	10	3,9%	256	6,5%	SI	5	Sexo	,000	17,655	1	Mujer
	19	221	85,3%	38	14,7%	259	6,6%	SI	5	Sexo	,000	17,655	1	Hombre
20	140	88,7%	16	10,3%	156	4,0%	SI	6	En su opinión, ¿El servicio educativo que brinda el centro de enseñanza donde asiste actualmente en relación a: Apoyo a la participación de los padres de familia (Promueven participación de los padres de familia, escuela de padres) es?	,006	13,831	1	Malo, -missing-	
21	893	96,4%	33	3,6%	926	23,6%	SI	6	En su opinión, ¿El servicio educativo que brinda el centro de enseñanza donde asiste actualmente en relación a: Apoyo a la participación de los padres de familia (Promueven participación de los padres de familia, escuela de padres) es?	,006	13,831	1	Bueno; No sabe; Muy malo; Muy malo	
22	77	95,1%	4	4,9%	81	2,1%	SI	7	En su opinión, ¿El servicio educativo que brinda el centro de enseñanza donde asiste actualmente en relación a: Equipamiento (Carpetas, Laboratorios, Computadoras, etc.) es?	,040	10,379	1	Malo, -missing-	
23	388	99,5%	2	0,5%	388	9,9%	SI	7	En su opinión, ¿El servicio educativo que brinda el centro de enseñanza donde asiste actualmente en relación a: Equipamiento (Carpetas, Laboratorios, Computadoras, etc.) es?	,040	10,379	1	Bueno; No sabe; Muy malo; Muy bueno	
24	50	70,4%	21	29,6%	71	1,8%	SI	19	Dominio geográfico	,004	17,359	1	Selva	
25	171	91,0%	17	9,0%	188	4,8%	SI	19	Dominio geográfico	,004	17,359	1	Sierra Norte; Costa Centro; Sierra Centro; Sierra Sur; Costa Sur; Lima Metropolitana; Costa Norte	
Prueba	0	28	19,2%	118	80,8%	146	100,0%	SI						
	1	1	1,8%	54	98,2%	55	37,7%	No	0	¿Que edad tiene en años cumplidos ? (En años)	,000	2065,013	6	<= 5.0
	2	2	6,5%	29	93,5%	31	21,2%	No	0	¿Que edad tiene en años cumplidos ? (En años)	,000	2065,013	6	(5.0, 6.0]
	3	2	20,0%	8	80,0%	10	6,8%	No	0	¿Que edad tiene en años cumplidos ? (En años)	,000	2065,013	6	(6.0, 7.0]
	4	2	22,2%	7	77,8%	9	6,2%	SI	0	¿Que edad tiene en años cumplidos ? (En años)	,000	2065,013	6	(7.0, 8.0]
	5	1	25,0%	3	75,0%	4	2,7%	SI	0	¿Que edad tiene en años cumplidos ? (En años)	,000	2065,013	6	(8.0, 9.0]
	6	5	55,6%	4	44,4%	9	6,2%	SI	0	¿Que edad tiene en años cumplidos ? (En años)	,000	2065,013	6	(9.0, 11.0]
	7	15	63,6%	13	46,4%	28	19,2%	SI	0	¿Que edad tiene en años cumplidos ? (En años)	,000	2065,013	6	> 11.0
8	0	0,0%	37	100,0%	37	25,3%	No	1	Dominio geográfico	,016	14,749	1	Selva; Sierra Norte; Sierra Centro; Sierra Sur	

Figura 15. Ganancia para nodos



Método de crecimiento: CHAID
Variable dependiente: Sabe leer y escribir

Se observa que el objeto no sabe leer ni escribir si presenta poder de discriminación tanto para la muestra de entrenamiento y la muestra de comprobación.



Método de crecimiento: CHAID
Variable dependiente: Sabe leer y escribir

Riesgo

Ejemplo	Estimación	Desv. Error
Entrenamiento	,129	,005
Prueba	,226	,035

Método de crecimiento: CHAID

Variable dependiente: ¿Sabe leer y escribir?

- Respuesta espontánea

Los niveles de riesgo son adecuados tanto para la muestra de entrenamiento como de la muestra de prueba.

Clasificación

Ejemplo	Observado	Pronosticado		Porcentaje correcto
		Si	No	
Entrenamiento	Si	2529	190	93,0%
	No	315	882	73,7%
	Porcentaje global	72,6%	27,4%	87,1%
Prueba	Si	25	3	89,3%
	No	30	88	74,6%
	Porcentaje global	37,7%	62,3%	77,4%

Método de crecimiento: CHAID

Variable dependiente: ¿Sabe leer y escribir? - Respuesta espontánea

Finalmente, el porcentaje de clasificación correcta para la muestra de entrenamiento es de 87.10%, mientras que para la muestra de prueba el porcentaje de clasificación correcta es de 77.40%, por lo que son indicadores adecuados de la clasificación.

4.3. Prueba de Hipótesis

H₀: La aplicación de las técnicas de minería de datos ayudará a medir la calidad del servicio educativo y lectoescritura en estudiantes de primaria del módulo educación de la ENAHO Trimestre 4, 2021.

H_a: La aplicación de las técnicas de minería de datos no ayudará a medir la calidad del servicio educativo y lectoescritura en estudiantes de primaria del módulo educación de la ENAHO Trimestre 4, 2021.

$$\chi^2 = 11.159, sig(p) = 0.025$$

Donde:

$$\text{sig}(p) = 0.025$$

$0.05 > 0.025$ por lo tanto se rechaza la H_a **aceptando la Hipótesis nula o aseveración original.**

4.3.1. Evaluación.

“Respecto al análisis factorial confirmatorio se elaboró el modelado con un 80% (3916 registros) de los datos elegidos aleatoriamente, posteriormente se realizó el modelado del 109 análisis factorial confirmatorio incluyendo el 20% (146 registros) de los datos restantes, en la que se observó que en el primer modelamiento un comportamiento semejante cuando se incluyó el 100% de los datos”.

“De igual manera, para el análisis de los árboles de decisión se trabajó con el 80% (3916 registros) para el modelo de entrenamiento y el 20% (146 registros) para predicción, donde se predijo un 77,40% de clasificación correcta”

4.3.2. Implementación

En cuanto a la herramienta para medir la calidad de los servicios educativos en el Perú, cabe señalar que la herramienta consta de 5 pilares: infraestructura, equipamiento, formación docente, dotación de material didáctico y apoyo a la participación de los padres en forma de familia, referente a el pilar básico de la calidad de los servicios educativos en el Perú, destacando la subdimensión oculta entre equipamiento e infraestructura, como un activo fijo propiedad de cada institución, capaz de activar los restantes tres indicadores de calidad de los servicios educativos.

Con respecto al análisis del árbol de decisiones, se observó que la alfabetización no está relacionada con el cumplimiento de la calidad de los servicios educativos en el Perú, sino que está influenciada por la edad del niño Factor de madurez cognitiva.

4.4. Discusión de resultados

La alfabetización es necesaria para percibir adecuadamente la calidad de los servicios educativos que brinda el gobierno nacional y las instituciones privadas. Por lo tanto, en el marco de la Encuesta Nacional de Hogares (ENAH), la información anterior se recopila con fines de investigación y análisis. Los resultados se validan con consistencia interna mediante análisis factorial confirmatorio, mientras que los estudiantes proporcionan un perfil de sus habilidades de lectura mediante un árbol de decisión.

Esta calidad tal como lo señala Municio (2000) es una medida de ajuste del producto o servicio a los requisitos demandados por los clientes y Crosby (1980) donde señala que la calidad de un bien o servicio no es otra cosa que la conformidad del mismo con los requerimientos de los clientes, así, las necesidades de los estudiantes peruanos respecto a infraestructura, equipamiento y dotación de materiales hace que el estado peruano actúe de manera directa en la mejora y/o adecuación de las mismas, tal como lo indica García Hoz (1982) que los medios que se disponen con influyentes en la calidad de los servicios. Por otro lado, el Consejo Nacional de Educación (2007) “la mejora de los aprendizajes en la educación básica depende de un conjunto de factores, que tienen que ver con el equipamiento de las instituciones, la gestión escolar, la formación de los profesores, nuevas reglas de juego para el ejercicio de la docencia, evaluación periódica de los progresos en el rendimiento de los alumnos, una educación inicial con mayor cobertura y calidad y programas que compensen las desventajas iniciales de los alumnos a nivel de su salud y nutrición. Estos últimos respecto a la enseñanza de los docentes y la participación de los padres hace que sea activa dicha participación de los actores principales como son los docentes dentro de las aulas de enseñanza aprendizaje, así como de los padres dentro de los ambientes del hogar, dinamizando el contexto educativo de los estudiantes, por ende, la calidad de servicio educativo peruano

medido por intermedio de los cinco indicadores considerados son factibles, ya que la corroboración mediante el análisis factorial confirmatorio resaltan su buen ajuste de manera lineal con el servicio educativo, con indicadores de ajuste del modelo adecuados y altos. Además, se observa que los errores aleatorios de la infraestructura se encuentran covariados con los errores aleatorios del equipamiento en los centros de enseñanza que asisten los niños”.

En segundo lugar, los niños analfabetos se definen como niños menores de seis años, lo cual es consistente con el hecho de que cada niño tiene su propio ritmo de aprendizaje. Para aprender a leer y escribir, primero debe lograr un buen nivel de habla y una correcta coordinación de movimientos. Por lo tanto, la edad ideal para leer y escribir se estima alrededor de los 6 años, lo que significa que estos niños recién comienzan a aprender a leer y escribir, mientras que los niños mayores de 8 años ya saben leer y escribir. Inglés. En la mayoría de los casos, entendibles.

“Sin embargo, por problemas de pobreza y dificultad de acceso a una educación adecuada se cuenta con un 32.37% de niños que aún no saben leer ni escribir, pese a que dichos niños debieron haber pasado por la educación inicial y al momento de la transición a la educación primaria ya se tiene noción de leer y escribir de manera silábica, sin embargo, ello no sucede en los lugares desfavorecidos de la costa, sierra y selva del país”. Lo que concuerda con lo indicado por Ferreiro en Zorzi (2003) en que la mayor parte del conocimiento que alguien puede obtener en su vida será mediante la lectura, aunque sea fuera de la escuela, entendiéndose como el paso del tiempo y el proceso propio de maduración y la necesidad de comunicación tanto verbal como escrita de las personas, tal como lo indica Braslavsky (2005) donde las actividades que se realizan en familia relacionadas con la lectura y escritura en edades tempranas asientan las bases para que los niños aprendan a leer y escribir en la escuela, donde posteriormente, será el tiempo de ocio y el placer por la lectura lo que hará

que se perfeccionen y aumenten las habilidades lectoras. El ambiente en el que crecen los niños debe estar relacionado con la lectura, es decir, hay libros, revistas, documentos en su entorno, y la lectura es una actividad cotidiana agradable en su vida

“La opinión de la calidad del servicio educativo que le brinda la institución educativa al cual asiste no se encuentra relacionado ni es influenciado por la capacidad de saber leer ni escribir, ya que dicha capacidad no influye en su satisfacción por los servicios que recibe, de la cual opina libremente en la forma como lo percibe tal como lo indica” Deming (2000) donde indica que la calidad es todo aquello que contribuye a la satisfacción del cliente.

CONCLUSIONES

Para la aplicación de técnicas de minería de datos para medir la calidad del servicio Educativo y lectoescritura en estudiantes de primaria del módulo educación de la ENAHO Trimestre 4, 2021 se concluye:

Su percepción de la calidad de los servicios de educación escolar no estaba relacionada con la alfabetización de los niños.

Los indicadores infraestructura, equipamiento, capacitación docente, disponibilidad de materiales educativos y apoyo a la participación de los padres explican la calidad de los servicios educativos, excepto infraestructura y equipamiento, que es una covariable que se ajusta bien al modelo del evaluador.

Las características sobresalientes de la alfabetización fueron que los 7 años se asociaron significativamente con la pertenencia a la costa y la sierra, los 8 años se asociaron con la pertenencia a la costa, la sierra y la selva, y los 9, 10 y mayores de once años se asociaron significativamente con el conocimiento, según Read y escribe. De igual forma, los niños de 11 años que saben leer y escribir y que expresan opiniones importantes sobre la calidad de los servicios educativos en cuanto a equipamiento (carpetas, laboratorios, computadoras, etc.) se distinguen de los niños que no saben leer ni escribir. cómo. saber leer y escribir hasta los 6 años con una vinculación significativa a la pertenencia a la costa, sierra o selva.

Se realizó el análisis factorial confirmatorio que mide la calidad del servicio educativo y lectoescritura en estudiantes de primaria del módulo educación de la ENAHO Trimestre 4, 2021 teniendo un análisis satisfactorio ya presentado.

Se realizó el análisis mediante los árboles de decisión para medir la calidad del servicio educativo y lectoescritura en estudiantes de primaria del módulo educación de la ENAHO Trimestre 4, 2021 se cumplió satisfactoriamente los objetivos.

La aplicación de las técnicas de minería de datos ayudó a medir la calidad del servicio educativo y lectoescritura en estudiantes de primaria del módulo educación de la ENAHO Trimestre 4, 2021.

RECOMENDACIONES

- Al Ministerio de Educación y a las instituciones trabajar en el sector de la educación; Incidir en la mejora de la infraestructura, equipamiento, formación docente, dotación de material didáctico y apoyo a la participación de los padres de familia para mejorar y garantizar la calidad de los servicios educativos en el Perú.
- A los directivos de la escuela profesional, Implementar cursos de minería de datos para equipar a los estudiantes con las herramientas adecuadas y profundizar el conocimiento en minería de datos.
- Finalmente se ha demostrado que el proyecto ha sido viable, se recomienda seguir la investigación.

REFERENCIAS BIBLIOGRÁFICAS

- Ale, J. (2005). *Introducción a Data Mining*.
- Álvarez, A. I. (2005). *Leer en español. Ed. Nobel. Oviedo*.
- Anaya Carrillo, M. A., & Bedón Trujillo, J. L. (2018). Minería de datos sobre la opinión de la calidad del servicio educativo peruano y la lectoescritura en estudiantes de educación primaria del módulo Educación de la ENAHO Trimestre 2,2016 [SANTIAGO ANTÚNEZ DE MAYOLO]. In *Ciencia e Investigación* (Vol. 4, Issue 1). <https://www.mendeley.com/catalogue/santiago-antunez-mayolo/>
- Bollen, K. A. (1989). *Structural equations with latent variables*. New York: John Wiley y Sons. <https://doi.org/10.1002/9781118619179>
- Chen, M., & Han, J. (1996). *Data mining: An overview from database perspective*. *IEEE Transactions on Knowledge and Data Eng.*
- Díaz, H. (2014). Obtenido de Los tres problemas que existen en el sector educación del Perú. *El Comercio*. <http://elcomercio.pe/economia/peru/tresproblemas-%0Aque-existen-sector-educacion-peru-noticia-1714189>
- Edwards Deming. (2000). *Out of the Crisis 2nd ed. Cambridge*.
- Ferreiro, E. (2011). *LA LECTO ESCRITURA*. <http://lasferrerianas5.blogspot.com.es/>
- García Hoz, V. (1982). *Calidad de educación, trabajo y libertad*.
- Hernández Sampieri, R. (2014). Metodología de la investigación. In S. A. D. C. . McGRAW-HILL / INTERAMERICANA EDITORES (Ed.), *Journal of Chemical Information and Modeling* (Sexta, Vol. 53, Issue 9).
- Hernández Sampieri, R., Fernández Collado, C., & Baptista Lucio, P. (2014). *Metodología de la Investigación* (S. A. D. C. . McGRAW-HILL / INTERAMERICANA EDITORES (ed.); Sexta).
- Hernández Sampieri, R., & Mendoza Torres, C. (2018). *Metodología de la Investigación*.
- Lucas Griñan, V. (2014). La lectoescritura en la Etapa de educación Primaria. In *Tesis de Pregrado*.
- Montealegre, R., & Forero, L. A. (2006). *Desarrollo de la lectoescritura: adquisición y dominio*.
- Morales, E. (2003). *Descubrimiento de Conocimiento en Bases de Datos*.
- Pio Alvarez, S. (2018). *Minería de Calidad de Datos: Aplicación de técnicas de minería de datos para la Evaluación de la Calidad de los Datos*. UNIVERSIDAD DE LA REPÚBLICA MONTEVIDEO – URUGUAY.
- Piscocoya Ordoñez, L. (2016). “*Aplicación de técnicas de minería de datos para predecir la deserción estudiantil en la educación básica regular en la región de Lambayeque*.” UNIVERSIDAD SEÑOR DE SIPÁN.
- Rokach, L., & Maimon, O. (2008). *DATA MINING WITH DECISION TREES Theory and Applications*.

ANEXOS

Anexo 01: Instrumento De Recolección de Datos

Encuesta

300. EDUCACIÓN (Para las personas de 3 años y más de edad)																																																																													
Persona N°	Nombre :	Informante N°																																																																											
Para personas de 15 años y más de edad																																																																													
300A. ¿CUÁL ES EL IDIOMA O LENGUA MATERNA QUE APRENDIÓ EN SU NIÑEZ: Quechua? 1 Aymara? 2 Asháninka? 10 Awajún/Aguaruna? 11 Shipibo – Konibo? 12 Shawi / Chayahuita? 13 Matsigenka / Mochiguanga? 14 Achuar? 15 Otra lengua nativa? (Especifique) 3 Castellano? 4 Portuguesa? 6 Otra lengua extranjera? (Especifique) 7 NO ESCUCHA/NO HABLA 8 LENGUA DE SEÑAS PERUANAS 9		302A. EN LOS ÚLTIMOS 12 MESES, ¿RECIBIÓ: <table border="1" style="width: 100%; border-collapse: collapse; font-size: x-small;"> <tr> <th style="width: 10%;"></th> <th style="width: 10%;">Sí</th> <th style="width: 10%;">No</th> <th style="width: 10%;"></th> <th style="width: 10%;"></th> <th style="width: 10%;"></th> </tr> <tr> <td></td> <td></td> <td></td> <td style="text-align: center;">PASE A 302</td> <td>¿Quitin la día?</td> <td></td> </tr> <tr> <td></td> <td></td> <td></td> <td></td> <td>PRONAMADUALFA ...</td> <td>1</td> </tr> <tr> <td></td> <td></td> <td></td> <td></td> <td>Ugala</td> <td>2</td> </tr> <tr> <td></td> <td></td> <td></td> <td></td> <td>ONG</td> <td>3</td> </tr> <tr> <td></td> <td></td> <td></td> <td></td> <td>Otro (especificar)</td> <td>4</td> </tr> </table> 1. Programa de alfabetización? ... <table border="1" style="display: inline-table; font-size: x-small;"> <tr><td>1</td><td>2</td></tr> </table>			Sí	No							PASE A 302	¿Quitin la día?						PRONAMADUALFA ...	1					Ugala	2					ONG	3					Otro (especificar)	4	1	2																																				
	Sí	No																																																																											
			PASE A 302	¿Quitin la día?																																																																									
				PRONAMADUALFA ...	1																																																																								
				Ugala	2																																																																								
				ONG	3																																																																								
				Otro (especificar)	4																																																																								
1	2																																																																												
301. ¿CUÁL ES EL ÚLTIMO AÑO O GRADO DE ESTUDIOS Y NIVEL QUE APROBÓ? <table border="1" style="width: 100%; border-collapse: collapse; font-size: x-small;"> <thead> <tr> <th rowspan="2"></th> <th rowspan="2">Año</th> <th rowspan="2">Grado</th> <th colspan="2">Centro de Estudios</th> <th rowspan="2"></th> </tr> <tr> <th>Estatal</th> <th>No Estatal</th> </tr> </thead> <tbody> <tr><td>Sin nivel..... 1</td><td></td><td></td><td></td><td></td><td rowspan="2">PASE A 302</td></tr> <tr><td>Educación inicial..... 2</td><td></td><td></td><td>1</td><td>2</td></tr> <tr><td>Primaria incompleta..... 3</td><td></td><td></td><td>1</td><td>2</td><td rowspan="2">PASE A 303</td></tr> <tr><td>Primaria completa..... 4</td><td></td><td></td><td>1</td><td>2</td></tr> <tr><td>Secund. incompleta..... 5</td><td></td><td></td><td>1</td><td>2</td><td rowspan="2">PASE A 302</td></tr> <tr><td>Secund. completa..... 6</td><td></td><td></td><td>1</td><td>2</td></tr> <tr><td>Básica especial..... 7</td><td></td><td></td><td>1</td><td>2</td><td rowspan="2">PASE A 302</td></tr> <tr><td>Sup. no universitaria incompleta..... 7</td><td></td><td></td><td>1</td><td>2</td></tr> <tr><td>Sup. no universitaria completa..... 8</td><td></td><td></td><td>1</td><td>2</td><td rowspan="2">PASE A 302</td></tr> <tr><td>Sup. universitaria incompleta..... 9</td><td></td><td></td><td>1</td><td>2</td></tr> <tr><td>Sup. universitaria completa..... 10</td><td></td><td></td><td>1</td><td>2</td><td rowspan="2">PASE A 302</td></tr> <tr><td>Maestría/Doctorado..... 11</td><td></td><td></td><td>1</td><td>2</td></tr> </tbody> </table>			Año	Grado	Centro de Estudios			Estatal	No Estatal	Sin nivel..... 1					PASE A 302	Educación inicial..... 2			1	2	Primaria incompleta..... 3			1	2	PASE A 303	Primaria completa..... 4			1	2	Secund. incompleta..... 5			1	2	PASE A 302	Secund. completa..... 6			1	2	Básica especial..... 7			1	2	PASE A 302	Sup. no universitaria incompleta..... 7			1	2	Sup. no universitaria completa..... 8			1	2	PASE A 302	Sup. universitaria incompleta..... 9			1	2	Sup. universitaria completa..... 10			1	2	PASE A 302	Maestría/Doctorado..... 11			1	2	303. EL AÑO PASADO (.....), ¿ESTUVO MATRICULADO EN ALGÚN CENTRO O PROGRAMA DE EDUCACIÓN BÁSICA O SUPERIOR? Sí 1 No 2 → PASE A 306	
	Año				Grado	Centro de Estudios																																																																							
		Estatal	No Estatal																																																																										
Sin nivel..... 1					PASE A 302																																																																								
Educación inicial..... 2			1	2																																																																									
Primaria incompleta..... 3			1	2	PASE A 303																																																																								
Primaria completa..... 4			1	2																																																																									
Secund. incompleta..... 5			1	2	PASE A 302																																																																								
Secund. completa..... 6			1	2																																																																									
Básica especial..... 7			1	2	PASE A 302																																																																								
Sup. no universitaria incompleta..... 7			1	2																																																																									
Sup. no universitaria completa..... 8			1	2	PASE A 302																																																																								
Sup. universitaria incompleta..... 9			1	2																																																																									
Sup. universitaria completa..... 10			1	2	PASE A 302																																																																								
Maestría/Doctorado..... 11			1	2																																																																									
301A. ¿CUÁL ES LA CARRERA SUPERIOR UNIVERSITARIA O NO UNIVERSITARIA QUE UD. ESTUDIA O HA ESTUDIADO? (Especifique) NINGUNA 1 → PASE A 303 NO SABE 2		304. ¿CUÁL ES EL GRADO O AÑO DE ESTUDIOS AL QUE ASISTIÓ EL AÑO PASADO (.....)? <table border="1" style="width: 100%; border-collapse: collapse; font-size: x-small;"> <thead> <tr> <th rowspan="2"></th> <th rowspan="2">Año</th> <th rowspan="2">Grado</th> <th colspan="2">Centro de Estudios</th> </tr> <tr> <th>Estatal</th> <th>No Estatal</th> </tr> </thead> <tbody> <tr><td>Educación inicial..... 1</td><td></td><td></td><td>1</td><td>2</td></tr> <tr><td>Primaria..... 2</td><td></td><td></td><td>1</td><td>2</td></tr> <tr><td>Secundaria..... 3</td><td></td><td></td><td>1</td><td>2</td></tr> <tr><td>Básica especial..... 7</td><td></td><td></td><td>1</td><td>2</td></tr> <tr><td>Sup. no universitaria..... 4</td><td></td><td></td><td>1</td><td>2</td></tr> <tr><td>Sup. universitaria..... 5</td><td></td><td></td><td>1</td><td>2</td></tr> <tr><td>Maestría/Doctorado..... 6</td><td></td><td></td><td>1</td><td>2</td></tr> </tbody> </table>			Año	Grado	Centro de Estudios		Estatal	No Estatal	Educación inicial..... 1			1	2	Primaria..... 2			1	2	Secundaria..... 3			1	2	Básica especial..... 7			1	2	Sup. no universitaria..... 4			1	2	Sup. universitaria..... 5			1	2	Maestría/Doctorado..... 6			1	2																																
	Año	Grado	Centro de Estudios																																																																										
			Estatal	No Estatal																																																																									
Educación inicial..... 1			1	2																																																																									
Primaria..... 2			1	2																																																																									
Secundaria..... 3			1	2																																																																									
Básica especial..... 7			1	2																																																																									
Sup. no universitaria..... 4			1	2																																																																									
Sup. universitaria..... 5			1	2																																																																									
Maestría/Doctorado..... 6			1	2																																																																									
301B. ¿CUÁL ES EL NOMBRE DEL CENTRO DE ESTUDIOS DONDE ESTUDIA O HA ESTUDIADO LA CARRERA SUPERIOR UNIVERSITARIA O NO UNIVERSITARIA? (Especifique) DEPARTAMENTO: NO SABINO RECUERDA..... 1		305. EL RESULTADO QUE OBTUVO EL AÑO PASADO (.....) FUE: ¿Aprobado? Año 1 ¿Desaprobado? 2 ¿Reñido? 3 ¿Otro? 4 (Especifique) ¿No aprueba, ni desaprueba (inicia)? 5																																																																											
302. ¿SABE LEER Y ESCRIBIR? <table border="1" style="width: 100%; border-collapse: collapse; font-size: x-small;"> <tr> <th colspan="2">Respuesta Espontánea</th> <th colspan="2">Respuesta con Cartilla de Lectura</th> <th rowspan="2">No se aplicó cartilla</th> </tr> <tr> <th>Sí</th> <th>No</th> <th>Sí</th> <th>No</th> </tr> <tr> <td>1</td> <td>2</td> <td>1</td> <td>2</td> <td>3</td> </tr> </table>		Respuesta Espontánea		Respuesta con Cartilla de Lectura		No se aplicó cartilla	Sí	No	Sí	No	1	2	1	2	3	306. ESTE AÑO, ¿ESTÁ MATRICULADO EN ALGÚN CENTRO O PROGRAMA DE EDUCACIÓN BÁSICA O SUPERIOR? Sí 1 No 2 → PASE A 310A																																																													
Respuesta Espontánea		Respuesta con Cartilla de Lectura		No se aplicó cartilla																																																																									
Sí	No	Sí	No																																																																										
1	2	1	2	3																																																																									
ENCUESTADOR(A): Luego de la respuesta espontánea aplique la Cartilla de lectura a todas las personas de 15 años y más de edad.		307. ACTUALMENTE, ¿ASISTE A ALGÚN CENTRO O PROGRAMA DE EDUCACIÓN BÁSICA O SUPERIOR BAJO LA MODALIDAD DE EDUCACIÓN A DISTANCIA? Sí 1 No 2 → PASE A 310A																																																																											

307A. A TRAVÉS DE QUE MEDIOS ACCEDA AL CONTENIDO DE LAS CLASES DE DISTANCIA:
(Acepte una o más alternativas)

¿Televisión? 1
 ¿Radio/equipo de sonido? 2
 ¿Plataforma virtual/página web? (*) 3
 (*) Computadora o Tablet o celular
 ¿Otro? 4
 (Especifique)

307B. EL DESARROLLO DE LAS CLASES A DISTANCIA LAS RECIBE CON:
(Acepte una o más alternativas)

¿Interacción con el profesor? 1
 ¿Videos? 2
 ¿Documentos (textos, cuadros, etc.)? 3
 ¿Otro? 4
 (Especifique)

ENCUESTADOR(A): Sólo si en pregunta 307 está circulado el código 1 "SI", continúe con la pregunta 308B. De lo contrario pase a la pregunta 310A.

308B. EN SU OPINIÓN, ¿EL SERVICIO EDUCATIVO QUE BRINDA EL CENTRO DE ENSEÑANZA DONDE ASISTE ACTUALMENTE EN RELACIÓN A:

	Muy mal	Mal	Bueno	Muy bueno	No sabe	No responde
Infraestructura (paredes, techos, pisos) es 1	1	2	3	4	5	
Equipamiento (Carpetas, Laboratorio, Computadoras, etc.) es 2	1	2	3	4	5	
Enseñanza de los maestros es 3	1	2	3	4	5	
Dotación de materiales educativos (Libros, Láminas, maquetas, etc.) es 4	1	2	3	4	5	6
Apoyo a la participación de los padres de familia/Promueven la participación de los padres de familia, escuela de padres, etc.) es 5	1	2	3	4	5	6

308C. LA INSTITUCIÓN EDUCATIVA O COLEGIO, INSTITUTO SUPERIOR O UNIVERSIDAD A LA QUE ASISTE, ¿ESTA UBICADA:

Aquí en este Distrito? 1

En otro Distrito?

Distrito: _____

Provincia: _____

Departamento: _____

310A. EN LOS ÚLTIMOS 12 MESES, DE.....
 A....., ¿RECIBIÓ ENSEÑANZA EN ALGÚN CENTRO O PROGRAMA DE ESTUDIOS CUYA DURACIÓN SEA MENOR A 3 AÑOS (GETPRO, ACADEMIA U OTROS)?

Si 1
 No 2

Para personas de 14 años y más de edad

310-B1 EN EL PRESENTE AÑO O EN AÑOS ANTERIORES, ¿RECIBIÓ ENSEÑANZA DE NIVEL AUXILIAR TÉCNICO O TÉCNICO, EN ALGÚN CENTRO O PROGRAMA DE ESTUDIOS CUYA DURACIÓN SEA MENOR A 3 AÑOS?

Si 1
 No 2 → **Pase a RECUADRO (A)**

310-C1 ¿CUÁL ES EL NOMBRE DE LA ÚLTIMA CARRERA DE NIVEL AUXILIAR TÉCNICO O TÉCNICO, QUE USTED ESTUDIA O HA ESTUDIADO?

(Especifique)

NO SABE 1

310D. ¿CUÁL ES EL TIEMPO DE DURACIÓN DE LA CARRERA DE NIVEL AUXILIAR TÉCNICO O TÉCNICO, QUE USTED ESTUDIA O HA ESTUDIADO?

AÑOS	_____
MESES	_____

310E. ¿CUÁL ES EL NOMBRE DEL CENTRO DE ESTUDIOS DONDE ESTUDIA O HA ESTUDIADO LA CARRERA DE NIVEL AUXILIAR TÉCNICO O TÉCNICO?

(Especifique)

DEPARTAMENTO: _____

NO SABE/NO RECUERDA 1

(A) ENCUESTADOR(A): Verifique las siguientes preguntas:

Si - 303 = 2
 - 308 = 2 y
 - 310A = 2 } **Pase a RECUADRO (B)**

y
 De lo contrario, continúe con Págs. 311

311. EN LOS ÚLTIMOS 12 MESES, DE..... A....., ¿GASTÓ, OBTUVO, CONSIGUIÓ O LE REGALARON:

¿Cuántas veces realizó el gasto durante los últimos 12 meses? →

	A. ¿Cómo obtuvo el/los?							B. ¿Cuántas veces al mes/día de.....?	C. ¿Dónde compró el/los?..... PASÉ A "1"	D. ¿Cuánto cree que le costaría el tenerlo que pagar por este producto o servicio?..... SI EN 311-A CIRCULÓ EL CÓDIGO 2, 3, 4, 5, 6 y 7	E. Número de veces								
	1. Comprado?.....	2. Autorizado?.....	3. Pago en especie?.....	4. Otros logros?.....	5. Programa escolar?.....	6. Otro (especificar).....													
	SI	NO	Código					SI	Código	MONTOS:									
			1	2	3	4	5	6	7			1	2	3	4	5	6	7	
1. Uniforme escolar?.....	1	2	1	2	3	4	5	6	7										
2. Calzado escolar?.....	1	2	1	2	3	4	5	6	7										
3. Libros y textos?.....	1	2	1	2	3	4	5	6	7										
4. Utiles escolares?.....	1	2	1	2	3	4	5	6	7										
5. Matrícula?.....	1	2	1	2	3	4	5	6	7										
6. APAFA?.....	1	2	1	2	3	4	5	6	7										
7. Otros (fotocopias, cuotas extraordinarias, etc.?..... <i>(Especifique)</i>	1	2	1	2	3	4	5	6	7										
8. Laptop del Programa "Una Laptop por Niño".....	1	2						6											
9. Tablet - MINEDU	1	2						6											
TOTAL																			

1. Autorizado
2. Librado
3. Centro de estudio
4. Paga
5. Pasa de mano

6. Mue
7. Rociado, Areca, etc.
8. Tende expeditamente al por mayor
9. Tende expeditamente al por menor
10. Otro (especificar)

312. EN EL MES ANTERIOR, GASTÓ EN:

¿Cuántas veces realizó el gasto durante los últimos 12 meses? →

	A. ¿Cómo obtiene el/los?							B. ¿Cuántas veces al mes/día de.....?	C. ¿Cuánto cree que le costaría el tenerlo que pagar?.....	D. Número de veces									
	1. Comprado?.....	2. Autorizado?.....	3. Pago en especie?.....	4. Otros logros?.....	5. Programa escolar?.....	6. Otro (especificar).....													
	SI	NO	Código					SI	Código	MONTOS:									
			1	2	3	4	5	6	7			1	2	3	4	5	6	7	
1. Pasaje mensual escolar?.....	1	2	1	2	3	4	5	6	7										
2. Movilidad particular al centro de enseñanza?.....	1	2	1	2	3	4	5	6	7										
TOTAL																			

B. ENCUESTADOR(A):
Transcribe la edad de pregunta 208, capítulo 200.

Menores de 25 años:
Transcribe el código de pgs. 208 y 207.

Si alguno de ellos es igual a 2 continúe con 213A.

De 25 años o más → PASÉ A 214A

313A. ¿CUÁL ES LA PRINCIPAL RAZÓN POR LA QUE NO ESTÁ MATRICULADO O NO ASISTE A ALGÚN CENTRO O PROGRAMA DE EDUCACIÓN BÁSICA O SUPERIOR?

- | | |
|--|--|
| Problemas económicos.....1 | No existe centro de educación básica o superior en el centro poblado.....7 |
| Estoy trabajando.....2 | Asiste a un centro de Educación Técnico-Productivo.....8 |
| Terminó sus estudios: secundarios/ superiores/Asiste a academia preuniversitaria.....3 | No me interesa/no me gusta el estudio.....9 |
| No tiene la edad suficiente (para el grupo 3-5 años).....4 | Se dedica a los quehaceres del hogar.....10 |
| Problemas familiares.....5 | Otra razón.....11 |
| De vacaciones.....6 | |

OBSERVACIONES: (Especifique)

SERVICIO DE INTERNET (Para las personas de 6 años y más de edad)

314A. EN EL MES ANTERIOR, ¿UD. HIZO USO DEL SERVICIO DE INTERNET?

Si..... 1
 No..... 2 →

PASE A 315A

314B. EN EL MES ANTERIOR, ¿USÓ UD. EL SERVICIO DE INTERNET EN:
 (Acepte una o más alternativas)

- El hogar? 1
- El trabajo? 2
- Un establecimiento educativo? 3
- Una cabina pública? 4
- En casa de otra persona? 5
- Acceso móvil a internet? 7
- Otro? 6
 (Especifique)

314B1. EN EL MES ANTERIOR, ¿EL SERVICIO DE INTERNET LO USÓ A TRAVÉS DE UNA/UN:
 (Acepte una o más alternativas)

- Computadora? 1
- Laptop? 2
- Tablet? 6
- Celular sin plan de datos? 8
- Celular con plan de datos? 9
- Otro? 7
 (Especifique)

314D. ¿UD. USA INTERNET AL MENOS:
 (Referido al uso de mayor frecuencia)

- Una vez al día? 1
- Una vez a la semana? 2
- Una vez al mes? 3
- Cada 2 meses o más? 4

ENCUESTADOR(A): Si en la pregunta 314B alguna de las alternativas que circuló corresponde al código 4 "Cabina Pública", continúe con la pregunta 315. De lo contrario pase a la pregunta 316.

315. EN EL MES ANTERIOR ¿OBTUVO EL SERVICIO DE INTERNET EN CABINA PÚBLICA: Comprado? 1 Autosuministró? 2 Pago en especie? 3 Otras hogares? 4 Programa social? 5 Otro? (especifique) 6	315A. ¿CUÁNTO FUE EL MONTO DEL SERVICIO EN EL MES ANTERIOR?	315B. ¿CUÁNTO CREE QUE LE COSTARÍA SI TUVIERA QUE PAGAR POR ESTE SERVICIO QUE USÓ EL MES ANTERIOR?			
		MONTO \$1.			
Código	\$1.	210-0 Bolívano	210-2 Bolívano	210-4 Bolívano	210-6 Bolívano

316. ¿USÓ EL INTERNET PARA:
 (Acepte una o más alternativas)

	Si	No
1. Obtener información (sobre bienes y servicios, salud, organizaciones gubernamentales)?	1	2
2. Comunicarse (e-mail, chat, llamadas por Skype, WhatsApp, Facebook, Twitter, etc.)?	1	2
3. Comprar productos y/o servicios?	1	2
4. Operaciones de banco electrónica y otros servicios financieros?	1	2
5. Educación formal y actividades de capacitación?	1	2
6. Transacciones (interactivas) con organizaciones estatales/autoridades públicas?	1	2
7. Actividades de entretenimiento (juegos de video, obtener películas, música, videos, escuchar radio, leer periódico, etc.)?	1	2
8. Vender productos y/o servicios (Mercado Libre, OLX, Facebook, etc.)?	1	2
9. Descarga de antivirus/aplicativos/software(programas)?	1	2
10. Otros?(Especifique)	1	2
11. Otros?(Especifique)	1	2

TELEFONÍA CELULAR

318A. ¿EN EL MES ANTERIOR, UD. UTILIZÓ....
 (Acepte una o más alternativas)

- Teléfono celular propio? 1
- Teléfono celular de un familiar o amigo/a? 2
- Teléfono celular de su centro de trabajo? 3
- Teléfono celular alquilado? 6
- Otro? 4
 (Especifique)
- NO UTILIZA 5

Para personas de 14 años y más de edad

300T. ¿CUÁL ES EL NÚMERO DE SU CELULAR?

- El número es..... 1 → **N° de Celular**
- No quiere dar el número de teléfono..... 2
- No tiene teléfono celular 3
- No Sabe 4

OBSERVACIONES

Anexo 02: Procedimiento de Validación y Confiabilidad

Validación de Instrumentos



Universidad Nacional Daniel Alcides Carrión

FACULTAD DE INGENIERIA

ESCUELA DE FORMACION PROFESIONAL DE INGENIERIA DE SISTEMAS

FICHA DE VALIDACIÓN DEL INSTRUMENTO “JUICIO DE EXPERTOS”

I. DATOS PERSONALES.

- APELLIDOS Y NOMBRES DEL EXPERTO:** PAREDES LOPEZ, ELVIS JESUS
- GRADO ACADÉMICO:** INGENIERO DE SISTEMAS Y COMPUTACION
- CARGO E INSTITUCIÓN DONDE LABORA:** UNDAC
- TÍTULO DE LA INVESTIGACIÓN:** Aplicación de técnicas de minería de datos para medir la calidad del servicio Educativo y lectoescritura en estudiantes de primaria del módulo educación de la ENAHO Trimestre 4, 2021.
- AUTOR DEL INSTRUMENTO:** Bach. Victor Antony HUANCA MAYHUASCA
- NOMBRE DEL INSTRUMENTO:** FICHAS DE REGISTRO

II. ASPECTOS DE EVALUACIÓN.

Después de haber leído las matrices de consistencia y de contrastación de variables; y analizado los ítems del instrumento correspondiente lea Ud. Las siguientes preguntas, dándole un puntaje para su validación marcando los números de puntaje del cuadro según considere (1. Completamente en desacuerdo. 2. En desacuerdo. 3. De acuerdo. 4. Completamente de acuerdo)

N°	Indicadores / Criterios: Preguntas	1	2	3	4	Observaciones
1	Claridad: Está formulado con lenguaje apropiado				X	
2	Objetividad: Está expresado en conductas observadas				X	
3	Actualidad: ¿El instrumento de recolección de datos mide correctamente los indicadores?				X	
4	Organización: ¿Existe una organización lógica entre (variables e indicadores)?				X	
5	Suficiencia: ¿Los instrumentos son suficientes para las mediciones de todos los indicadores?				X	
6	Intencionalidad: Es adecuado para valorar aspectos sobre la comprensión espacial en relación a las capacidades de define, identifica, señala y ubica.				X	
7	Consistencia: ¿Los objetivos y variables están formulados de forma que puedan ser medibles y comprobados?				X	
8	Coherencia: ¿Hay coherencia entre las variables, dimensiones e indicadores?				X	
9	Metodología: ¿La estrategia responde al propósito de la investigación?				X	
10	actualidad: ¿Es adecuado el avance de la ciencia y tecnología y la experiencia del tesista?				X	
	TOTAL				40	
	TOTAL GENERAL				40	

Opinión de aplicabilidad: Ninguno


.....
PAREDES LOPEZ ELVIS JESUS



Universidad Nacional Daniel Alcides Carrión

FACULTAD DE INGENIERIA

ESCUELA DE FORMACION PROFESIONAL DE INGENIERIA DE SISTEMAS

FICHA DE VALIDACIÓN DEL INSTRUMENTO “JUICIO DE EXPERTOS”

I. DATOS PERSONALES.

- a. **APELLIDOS Y NOMBRES DEL EXPERTO:** VICENTE CRISTOBAL, JOHANNES AVILIO
- b. **GRADO ACADÉMICO:** INGENIERO DE SISTEMAS Y COMPUTACION
- c. **CARGO E INSTITUCIÓN DONDE LABORA:** INGENIERO DE SEGURIDAD INFORMATICA – GOBIERNO REGIONAL PASCO
- d. **TITULO DE LA INVESTIGACIÓN:** Aplicación de técnicas de minería de datos para medir la calidad del servicio Educativo y lectoescritura en estudiantes de primaria del módulo educación de la ENAHO Trimestre 4, 2021.
- e. **AUTOR DEL INSTRUMENTO:** Bach. Victor Antony HUANCA MAYHUASCA
- f. **NOMBRE DEL INSTRUMENTO:** FICHAS DE REGISTRO

II. ASPECTOS DE EVALUACIÓN.

Después de haber leído las matrices de consistencia y de contrastación de variables; y analizado los ítems del instrumento correspondiente lea Ud. Las siguientes preguntas, dándole un puntaje para su validación marcando los números de puntaje del cuadro según considere (1. Completamente en desacuerdo. 2. En desacuerdo. 3. De acuerdo. 4. Completamente de acuerdo)

N°	Indicadores / Criterios: Preguntas	1	2	3	4	Observaciones
1	Claridad: Está formulado con lenguaje apropiado				X	
2	Objetividad: Está expresado en conductas observadas				X	
3	Actualidad: ¿El instrumento de recolección de datos mide correctamente los indicadores?				X	
4	Organización: ¿Existe una organización lógica entre (variables e indicadores)?				X	
5	Suficiencia: ¿Los instrumentos son suficientes para las mediciones de todos los indicadores?				X	
6	Intencionalidad: Es adecuado para valorar aspectos sobre la comprensión espacial en relación a las capacidades de define, identifica, señala y ubica.				X	
7	Consistencia: ¿Los objetivos y variables están formulados de forma que puedan ser medibles y comprobados?				X	
8	Coherencia: ¿Hay coherencia entre las variables, dimensiones e indicadores?				X	
9	Metodología: ¿La estrategia responde al propósito de la investigación?				X	
10	actualidad: ¿Es adecuado el avance de la ciencia y tecnología y la experiencia del tesista?				X	
	TOTAL				40	
	TOTAL GENERAL				40	

Opinión de aplicabilidad: Ninguno


Firma del Experto
DNI: 72647100
CIP: 258778
N° Telefónico: 935784094



Universidad Nacional Daniel Alcides Carrión

FACULTAD DE INGENIERIA

ESCUELA DE FORMACION PROFESIONAL DE INGENIERIA DE SISTEMAS

FICHA DE VALIDACIÓN DEL INSTRUMENTO “JUICIO DE EXPERTOS”

I. DATOS PERSONALES.

- a. **APELLIDOS Y NOMBRES DEL EXPERTO:** RAMON VICENTE, LILIANA MADELEINE
- b. **GRADO ACADÉMICO:** INGENIERO DE SISTEMAS Y COMPUTACION
- c. **CARGO E INSTITUCIÓN DONDE LABORA:** INDEPENDIENTE
- d. **TITULO DE LA INVESTIGACIÓN:** Aplicación de técnicas de minería de datos para medir la calidad del servicio Educativo y lectoescritura en estudiantes de primaria del módulo educación de la ENAHO Trimestre 4, 2021.
- e. **AUTOR DEL INSTRUMENTO:** Bach. Víctor Antony HUANCA MAYHUASCA
- f. **NOMBRE DEL INSTRUMENTO:** FICHAS DE REGISTRO

II. ASPECTOS DE EVALUACIÓN.

Después de haber leído las matrices de consistencia y de contrastación de variables; y analizado los ítems del instrumento correspondiente lea Ud. Las siguientes preguntas, dándole un puntaje para su validación marcando los números de puntaje del cuadro según considere (1. Completamente en desacuerdo. 2. En desacuerdo. 3. De acuerdo. 4. Completamente de acuerdo)

N°	Indicadores / Criterios: Preguntas	1	2	3	4	Observaciones
1	Claridad: Está formulado con lenguaje apropiado				X	
2	Objetividad: Está expresado en conductas observadas				X	
3	Actualidad: ¿El instrumento de recolección de datos mide correctamente los indicadores?				X	
4	Organización: ¿Existe una organización lógica entre (variables e indicadores)?				X	
5	Suficiencia: ¿Los instrumentos son suficientes para las mediciones de todos los indicadores?				X	
6	Intencionalidad: Es adecuado para valorar aspectos sobre la comprensión espacial en relación a las capacidades de define, identifica, señala y ubica.				X	
7	Consistencia: ¿Los objetivos y variables están formulados de forma que puedan ser medibles y comprobados?				X	
8	Coherencia: ¿Hay coherencia entre las variables, dimensiones e indicadores?				X	
9	Metodología: ¿La estrategia responde al propósito de la investigación?				X	
10	actualidad: ¿Es adecuado el avance de la ciencia y tecnología y la experiencia del tesista?				X	
	TOTAL				40	
	TOTAL GENERAL				40	

Opinión de aplicabilidad: Ninguno

RAMON VICENTE, Liliana M.