



UNIVERSIDAD DEL AZUAY

DEPARTAMENTO DE POSGRADOS

Magister en Sistemas de Información, mención Inteligencia de
Negocios

FACTORES PERSONALES QUE IMPACTAN EN EL DESEMPEÑO DE UN USUARIO FRENTE A UN JUEGO SERIO ORIENTADO AL REFUERZO DE LA MEMORIA

Autor:

Jefferson Hernan Arias Ochoa

Director:

Irene Priscila Cedillo Orellana Ph. D.

Cuenca – Ecuador

2022

Dedicatoria

A Dios por enseñarme a nunca rendirme y siempre perseguir mis sueños con esfuerzo y dedicación.

A mis padres Piedad y Hernan, quienes siempre me han apoyado en todas las decisiones tomadas a lo largo de mi vida y me han sabido guiar por el camino del bien, por medio de su paciencia y cariño.

A mi hermana Dayanna, la cual es mi motivación a seguir adelante día a día y no decaer ante ningún obstáculo.

A mis abuelitas María (†) y Rebeca, quienes siempre han estado pendientes de mí en todo momento brindándome su apoyo y amor de forma incondicional.

Agradecimientos

A mi directora, Ing. Priscila Cedillo, una gran persona y profesional, quien una vez más, me supo guiar y transmitir los conocimientos necesarios para la ejecución de este trabajo, mismos que han sido transmitidos con paciencia y dedicación.

A la Ing. María Inés Acosta que, gracias a sus enseñanzas, ayuda y aliento, permitieron desarrollar y culminar este trabajo.

A la Ing. Alexandra Bermeo que, gracias a su apoyo y guía fue posible culminar y presentar el presente trabajo.

A Emily, Esteban, Miguel Ángel, Paola y Paúl quienes me demostraron que una valiosa amistad puede perdurar con el tiempo.

A Vanessa, por su amistad, empatía, cariño y apoyo durante el desarrollo y culminación de este trabajo.

A mis amigos y compañeros de trabajo Catalina, Yolanda, María Esperanza y Christian, por demostrarme que el calor y cariño de familia existe dentro de un ambiente laboral.

A mi familia por el cariño y el apoyo brindado durante toda mi vida.

A todos ustedes mil gracias.

Resumen

El uso de los juegos serios, con el pasar del tiempo, se ha extendido a varios campos; siendo uno de ellos la psicología. Cuyas prácticas permiten reforzar distintas habilidades cognitivas, en este caso la memoria. Estos juegos generan datos, mismos que permiten la medición de variables durante la interacción del usuario con los mismos. Por ejemplo: demográficas por parte del jugador e indicadores de rendimiento dentro del juego.

En este trabajo de titulación, se ha adaptado un juego serio, orientado al refuerzo de la memoria. El mismo que, a través de sus datos, permite la medición de variables enmarcadas con las funciones ejecutivas que promueven una buena salud mental, para su posterior recolección y análisis de datos; además, se ha aplicado la metodología Crisp-DM para obtener un modelo de minería de datos que permita describir el comportamiento de los datos y generar un conjunto de inferencias.

Finalmente, para mostrar la precisión de los resultados obtenidos utilizando el modelo de minería de datos propuesto, se ha realizado un caso de estudio. La propuesta fue evaluada con cuatro psicólogos clínicos, de quienes se obtuvieron resultados que validan la solución obtenida y proporcionan información relacionada al modelo.

Palabras clave. Juegos serios, Minería de Datos, Funciones Cognitivas, Análisis de Datos.

IRENE
PRISCILA
CEDILLO
ORELLANA

Digitally signed by
IRENE PRISCILA
CEDILLO ORELLANA
Date: 2022.09.22
14:52:50 -05'00'

Ing. Priscila Cedillo Orellana Ph.D
Director del trabajo de titulación



Ing. Jefferson Arias Ochoa
Autor

Abstract

Over time, the use of serious games has spread to various areas; psychology being one of them. Whose practices allows the reinforcement of different cognitive skills, in this case, memory. These games enable the compilation of variables during the interaction with the user. For example, demographic data per player and performance indicators within the game.

During this work, a serious game has been adapted, aimed at strengthening memory. Also, it includes appropriate variables within the spectrum of mental health, for later data collection and analysis. In addition, the Crisp-DM methodology has been applied to obtain a data mining model, which allows describing the behavior of the data and generating a set of inferences.

Finally, to obtain the accuracy of the results gathered from the data mining model, a study case had been carried out. On top of that, the proposal was tested by four clinical psychologists, from which results obtained validated the solution found and provide information related to the model.

Keywords. Serious games, Data mining, Cognitive functions, Data analysis.

Translated by:

IRENE PRISCILA
CEDILLO
ORELLANA

Digitally signed by
IRENE PRISCILA
CEDILLO ORELLANA
Date: 2022.09.22
14:53:10 -05'00'

Ing. Priscila Cedillo Orellana Ph.D
Final Degree Project Director



Ing. Jefferson Arias Ochoa
Author



Índice de Contenido

Dedicatoria.....	ii
Agradecimientos	iii
Resumen	iv
Abstract.....	v
1. Introducción	1
1.1 Problemática	1
1.2 Preguntas de Investigación.....	1
1.3 Objetivo General	2
1.4 Objetivos Específicos.....	2
1.5 Metodología de la Investigación.	2
1.6 Estructura del Trabajo de Titulación.	4
2. Marco Tecnológico y Estado del Arte.....	5
2.1 Marco Tecnológico.....	5
2.1.1 Salud Mental y Funciones Cognitivas	5
2.1.2 Aplicaciones para el Refuerzo Cognitivo y Minería de Datos	6
2.1.3 Análisis de Datos.....	8
2.2 Estado del Arte	9
3. Metodología Aplicada	12
3.1 Actores	12
3.2 Pre-Procesar los Datos.....	14
3.3 Transformar los Datos	14
3.4 Conseguir los Modelos	15
3.5 Evaluar los Modelos.....	16
4. Instanciación de la Metodología	18
4.1 Adaptación del juego serio.....	18
4.2 Fase de pre-procesar los datos	20
4.2.1 Pre-procesar los datos	20
4.3 Fase de transformar los datos	24
4.3.1 Transformar los datos	24
4.4 Fase de Conseguir los Modelos	28

4.4.1 Conseguir los modelos.....	29
4.5 Fase de evaluación.....	30
4.5.1 Evaluar los modelos.....	30
5. Evaluación.....	33
5.1 Diseñar el caso de estudio.....	33
5.5.1 Definición de objetivos, hipótesis y variables.....	34
5.2 Ejecución y análisis del caso de estudio.....	35
5.2.1 Selección de los sujetos.....	35
5.2.2 Diseño y ejecución del caso de estudio.....	35
5.2.3 Análisis de los resultados.....	37
5.3 Amenazas a la validez.....	39
5.3.1 Validez de conclusión.....	39
5.3.2 Validez interna.....	39
5.3.3 Validez de constructo.....	39
5.3.4 Validez externa.....	39
6. Conclusiones.....	39
7. Referencias.....	42
8. Anexos.....	46

Índice de Figuras

Figura 1. Metodología de Gorscheck aplicada a este trabajo de titulación.	3
Figura 2. Estructura del trabajo de titulación alineada a los objetivos y metodología de la investigación.....	4
Figura 3. Diagrama SPEM de la metodología Crisp-DM adaptada a este trabajo de titulación.	12
Figura 4. Diagrama de casos de uso, respecto a los actores que participan en la metodología.....	13
Figura 5. Fase de pre-procesamiento de los datos.....	14
Figura 6. Fase de transformación de los datos.	15
Figura 7. Fase de obtención del modelo.	16
Figura 8. Fase de evaluación.	16
Figura 9. Inclusión de las variables dentro del juego serio desplegado.....	18
Figura 10. Variables agregadas dentro del formulario de registro.	19
Figura 11. Juego serio desplegado.	19
Figura 12. Fase de pre-procesamiento de los datos del juego serio.	20
Figura 13. Datos que deben ser llevados a un proceso de limpieza.	22
Figura 14. Herramienta RapidMiner usada para la limpieza de datos.	23
Figura 15. Estadística de las variables a utilizar.	23
Figura 16. Nueva fuente de datos.	24
Figura 17. Fase de transformación de los datos.	24
Figura 18. Dataset sin realizar la transformación de los datos.	26
Figura 19. Proceso de etiquetado de los datos.	26
Figura 20. Nuevo conjunto de datos.	28
Figura 21. Fase de obtención del modelo de minería de datos del juego serio.	29
Figura 22. Conectores para generar el árbol de decisión.	29
Figura 23. Árbol de decisión generado.	30
Figura 24. Fase de evaluación de los modelos obtenidos.	30
Figura 25. Página de inicio.	36
Figura 26. Página del caso de estudio.	37
Figura 27. Página de la encuesta.....	37

Índice de Tablas

Tabla 1. Tabla de estudios relacionados.....	10
Tabla 2. Variables recopiladas al hacer uso del juego serio.....	20
Tabla 3. Variables a aplicar la limpieza de datos.....	22
Tabla 4. Variables a realizar el etiquetado/transformación de datos.....	25
Tabla 5. Variables de interés.....	27
Tabla 6. Inferencias obtenidas.....	31
Tabla 7. Plantilla para la definición del Goal-Question-Metric.....	33
Tabla 8. Definición del Goal-Question-Metric.....	33
Tabla 9. Cuestionario para medir la precisión de los resultados del modelo.....	34
Tabla 10. Pregunta cerrada (Sí/No).....	35
Tabla 11. Percepciones del usuario con respecto a las tareas del ejercicio aplicado...	37
Tabla 12. Resultados de las hipótesis planteadas.....	38

Índice de Anexos

Anexo 1 Formulario de evaluación.....	46
Anexo 2 Evidencia de evaluación.....	48
Anexo 3 Resultados de la encuesta.....	49
Anexo 4 Adaptación del juego serio.....	52

1. Introducción

Este capítulo tiene como objetivo presentar la problemática, preguntas de investigación, objetivos generales y específicos, la metodología de investigación utilizada. Finalmente, se detalla la estructura que posee este trabajo.

1.1 Problemática

En el área de la Psicología, existen varios test que permiten evaluar diferentes funciones cognitivas, tales como: atención, memoria, lenguaje, entre otras. El test de Mini-Mental State Examination (MMSE), constituye una breve prueba de deterioro cognitivo, utilizada para detectar demencia (Galea & Woodward, 2005). Otro test orientado a evaluar problemas en la memoria, es el test conductual de Rivermead (RBMT), (Alonso & Prieto, 2004). Dichas pruebas son aplicadas de diversas formas, tomando en cuenta ciertas variables asociadas a cada paciente; los resultados de las mismas, proveen al experto una referencia útil al momento de brindar un diagnóstico sobre las áreas evaluadas.

Por otra parte, existen varias metodologías para hacer análisis o minería de datos ya sea: Crisp-DM, KDD y SEMMA (Azevedo & Santos, 2008), las cuales permiten generar conocimiento sobre cierto conjunto de datos en un determinado contexto, es decir, dada una cierta problemática ya sea en el ámbito de la industria o la academia; al aplicar estas metodologías, se permite ir más allá de solo recolectar y almacenar los datos; pudiendo llegar a explorar, y visualizar los datos de tal manera que se pueda descubrir patrones. Éstos patrones permiten explicar el comportamiento de los datos a lo largo del tiempo. Una pieza clave durante este proceso son las técnicas de minería de datos que, con el pasar del tiempo, se han extendido a varios campos (Jothi, 2015). Una de las disciplinas tecnológicas, creadas para la obtención de estos patrones, son los juegos serios (Corrigan, 2015). Según Lopes et al. (2018), los juegos serios consisten en pruebas o test mentales que utilizan la diversión para el entrenamiento de ciertas capacidades con fines políticos, educativos, salud o comunicación. Estos juegos recopilan variables demográficas asociadas al usuario y al juego; en este sentido, se busca determinar, cuáles son los factores personales que impactan en el desempeño del usuario frente al juego aplicado, haciendo uso de las técnicas de minería de datos adecuadas, para obtener un modelo idóneo para su posterior evaluación y validación.

1.2 Preguntas de Investigación

Las preguntas de investigación planteadas durante el desarrollo de este estudio, sirven para dirigir el mismo y obtener ideas claras. En este contexto, es importante indagar en ciertas variables idóneas, dentro del marco de la salud mental, para su posterior inclusión en un juego serio. Es importante, obtener varios estudios relacionados al área de minería de datos y juegos serios, con el fin de determinar brechas de investigación, y, finalmente, determinar las técnicas

de minería de datos que permitan obtener modelos, para describir el comportamiento de los datos. Así, se han planteado las siguientes preguntas de investigación:

- ¿Qué variables demográficas dentro del marco de la salud mental, se deben incluir en un juego serio orientado al refuerzo de la memoria?
- ¿Cómo abordan la industria y la academia, la aplicación de minería de datos en los juegos serios, para generar conocimiento?
- ¿Qué técnicas de minería de datos permiten obtener modelos idóneos de un juego serio orientado al refuerzo de la memoria?

1.3 Objetivo General

Como objetivo general se ha visto necesario establecer que técnicas de minería de datos son las idóneas para generar conocimiento sobre un conjunto de datos de un juego serio, es por ello, que el objetivo general planteado es el siguiente:

- Aplicar técnicas de minería de datos que permitan obtener un modelo para determinar los factores personales, que inciden en el desempeño del jugador frente a un juego serio orientado a la memoria.

1.4 Objetivos Específicos

Con el fin de cumplir el objetivo general, se han propuesto los siguientes objetivos específicos:

- Evidenciar el estado actual de la investigación relacionada con la salud mental, juegos serios y minería de datos, a través del análisis del estado tecnológico de la investigación en el área.
- Adaptar un juego serio, mismo que permita registrar las variables apropiadas, para analizar aspectos relacionados con el desempeño del jugador dentro del marco de la salud mental, para su posterior recolección de datos.
- Aplicar la Metodología Crisp-DM, para la obtención de un modelo de minería de datos, que permita determinar los factores relevantes que inciden en el desempeño del jugador frente al juego serio.
- Evaluar el modelo obtenido a través de un caso de estudio, con la finalidad de dar validez a la solución.

1.5 Metodología de la Investigación.

Para la elaboración de este trabajo de titulación, se aplicará la metodología de Gorschek et al. (2006), haciendo uso de cuatro de los siete pasos propuestos mostrados en la Figura 1, esto es debido al alcance de este trabajo de titulación y a la necesidad futura de generar un medio de

adopción tecnológica de la solución en contextos reales. Los pasos a ser empleados son descritos a continuación:

- **Análisis del problema:** Consiste en comprender qué variables dentro del marco de la salud mental, son importantes dentro de un juego serio orientado al refuerzo de la memoria, para su posterior adaptación, despliegue, recolección y análisis de datos.
- **Formulación del problema:** Una vez analizado el problema, se lo formula de forma clara y precisa, en esta parte se describe el planteamiento del problema, justificación del estudio, las preguntas de investigación, objetivos que esta investigación debe alcanzar.
 - o **Estudio del estado del arte:** Se realiza la obtención de estudios relacionados con juegos serios, salud mental y minería de datos, con el fin de determinar el estado actual de la investigación, revisando soluciones existentes que permitan identificar brechas que la investigación desea abordar.
- **Solución candidata:** Se propone una solución al problema establecido, por medio de un modelo de minería de datos, el cual permite describir el comportamiento de los usuarios que hicieron uso del juego serio.
- **Validación en la academia:** La validación de la solución propuesta se la realiza por medio de un caso de estudio. Misma que permite validar o rechazar la solución candidata desde el punto de vista de un psicólogo clínico o investigador.

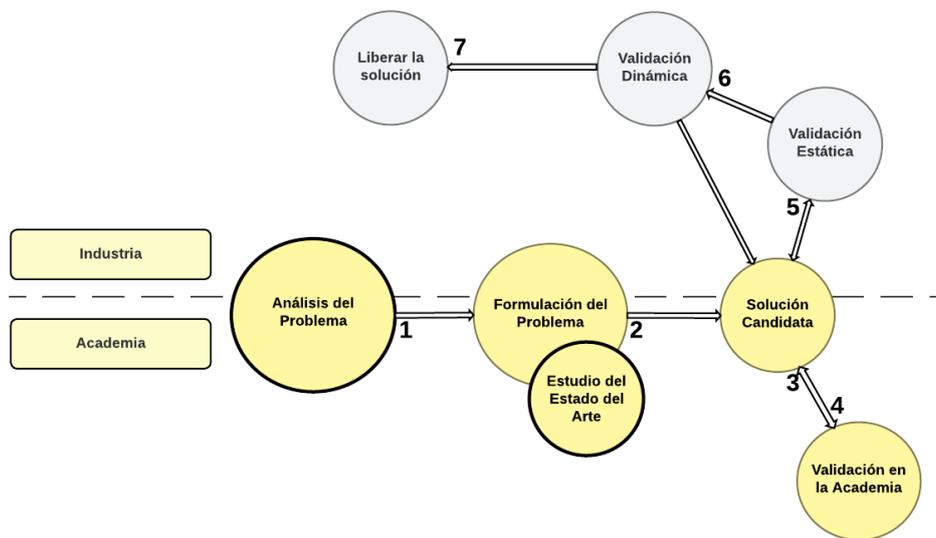


Figura 1. Metodología de Gorschek aplicada a este trabajo de titulación.

1.6 Estructura del Trabajo de Titulación.

La estructura de este trabajo de titulación se alinea con las actividades del modelo de transferencia tecnológica propuesto por Gorschek et al. (2006). A continuación, se describen brevemente los capítulos presentados a lo largo de este documento.

Capítulo 2: Se presenta el estado actual de la investigación; además, los conceptos teóricos claves asociados a este trabajo de titulación, que sirven de ayuda al lector para un mejor entendimiento.

Capítulo 3: Se describe a detalle la metodología utilizada, describiendo cada actividad, sus tareas, roles, entradas y salidas.

Capítulo 4: Se presentan los resultados obtenidos, los cuales satisfacen los objetivos planteados en este trabajo de titulación.

Capítulo 5: Se presenta la evaluación del modelo obtenido mediante un caso de estudio, en el cual se describen los modelos obtenidos durante la ejecución de este trabajo y las métricas definidas para la obtención del modelo final.

Capítulo 6: Se describe la información más valiosa y relevante obtenida durante el desarrollo de este trabajo.

En la Figura 2, se muestra la estructura del trabajo de titulación, en donde se puede observar que cada uno de los entregables (capítulos) se encuentran alineados a los objetivos y la metodología propuesta para el desarrollo de este trabajo.

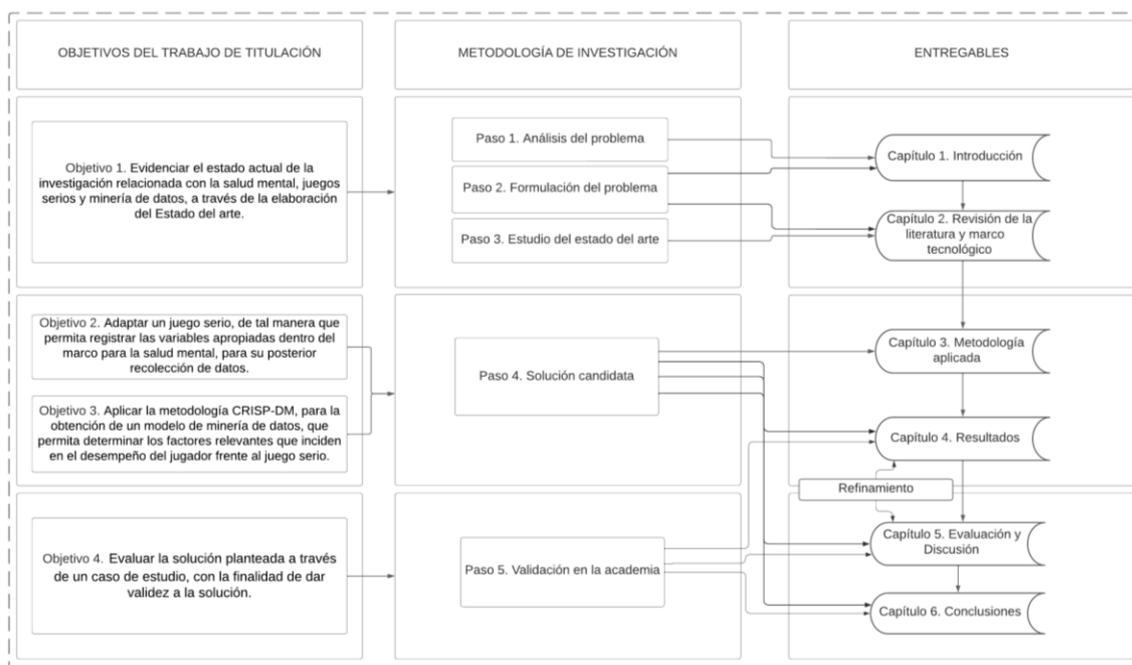


Figura 2. Estructura del trabajo de titulación alineada a los objetivos y metodología de la investigación.

2. Marco Tecnológico y Estado del Arte

En este capítulo se aborda el conjunto de conceptos que hacen de este trabajo, un producto autocontenido y consistente, que permita al lector una idea cabal de todas las definiciones y tecnologías que serán descritas y mencionadas a lo largo del estudio.

2.1 Marco Tecnológico

En esta sección se proporciona al lector, los conceptos fundamentales para un mejor entendimiento de este trabajo de titulación; entre los cuales tenemos la salud mental y funciones cognitivas, haciendo énfasis en la memoria. Las aplicaciones más comunes usadas para el para refuerzo cognitivo, ya sea con o sin uso de la tecnología, así como también, las técnicas de minería de datos usadas para generar conocimiento sobre un conjunto de datos. Finalmente, se habla del uso del análisis de datos, y sus distintos métodos que pueden ser utilizados para obtener conclusiones.

2.1.1 Salud Mental y Funciones Cognitivas

La Organización Mundial de la Salud (OMS), define a la salud mental como: “un estado de completo bienestar, físico, mental y social; y no solamente, la ausencia de afecciones o enfermedades”. Con base en esta definición, la salud mental es un componente fundamental de salud, que permite a las personas hacer frente al estrés del día a día, trabajar productivamente y realizar sus actividades cotidianas (Etienne, 2018).

Una enfermedad mental es una alteración de tipo emocional, cognitivo y/o del comportamiento, en la que, se ven afectados procesos psicológicos básicos, tales como: emoción, motivación, cognición, conciencia, conducta, percepción, aprendizaje y lenguaje (Control, 2022). Para una persona que sufre de la alteración de alguno de estos factores, se dificulta la adaptación al entorno cultural y social en el que vive; dicha persona, padece alguna forma de malestar subjetivo; sin embargo, es importante saber que, este tipo de problemas de salud mental pueden ser tratados, en busca de la recuperación del paciente, para que éste pueda tener una mejor calidad de vida (Control, 2022). En este sentido, existen varios factores que pueden afectar a la salud mental, los cuales son: condiciones económicas, profesión, educación, calidad de vivienda, género, edad, etnia, entre otros. De estos factores, el género y el estado económico, son los factores con una puntuación mayor para afecciones de salud mental (Raphael & Mikkonen, 2010).

Otros tipos de factores son los de clase biológica, de los cuales los antecedentes familiares pueden aumentar la probabilidad de afecciones mentales; ya que, ciertos genes y variantes exponen una persona a un riesgo mayor. Tener un gen vinculado con un trastorno mental, como la depresión o esquizofrenia, no garantiza que, el paciente desarrolle dicha afección. Las afecciones de salud mental, tales como el estrés, la depresión y la ansiedad,

pueden desarrollarse debido a problemas de salud física subyacentes que cambian la vida, como el cáncer, la diabetes y el dolor crónico. (Control, 2022)

El funcionamiento cognitivo de un individuo está regido por varios procesos o habilidades mentales; éstas incluyen el aprendizaje, el razonamiento, la memoria, el lenguaje, la atención, entre otras (Fisher, 2019). Cada una de estas funciones forman parte de las habilidades cognitivas, las cuales nos permiten recibir, procesar y elaborar información. Estas habilidades se desarrollan a lo largo de los años, comenzando desde etapas primarias del desarrollo, hasta aproximadamente la segunda década de vida (Klein, 2015); posteriormente, éstas empiezan a disminuir, siendo curiosamente ésta, la época en la que se empieza generalmente a trabajar (Klein, 2015).

Para un mejor entendimiento sobre las funciones cognitivas y su funcionamiento, se explica a continuación las más relevantes, haciendo énfasis en la memoria.

Memoria: La memoria hace referencia a los procesos que son usados para adquirir, almacenar, retener y recuperar información. Según Cherry et al. (2019), la memoria puede ser clasificada en tres tipos:

- **Memoria sensorial:** En este tipo de memoria, la información se mantiene almacenada durante un periodo muy breve de tiempo, generalmente durante no más de medio segundo.
- **Memoria a corto plazo:** En este tipo de memoria, se encuentra la información que actualmente conocemos o en la que estamos pensando. Esta memoria se mantendrá durante aproximadamente 20 a 30 segundos.
- **Memoria a largo plazo:** Se refiere al almacenamiento continuo de información. Esta información está en gran medida fuera de nuestro conocimiento, pero se puede llamar a la memoria de trabajo para usarla cuando sea necesario.

Razonamiento: Esta función hace referencia a los procesos cognitivos mediante los cuales las personas sacan conclusiones sobre problemas significativos que comprenden u observan. Existen tres tipos de razonamiento: deducción, inducción y abducción. Las deducciones son inferencias verdaderas en aquellos casos en que las premisas son verdaderas. Las inducciones son una forma de razonamiento que consiste en el establecimiento de una premisa general a partir de la observación de aspectos particulares. Y, las abducciones son tipos especiales de inducciones que producen hipótesis explicativas (Khemlani, 2018).

Lenguaje: El lenguaje es una propiedad que se basa en un sistema de comunicación basado en palabras y la combinación de éstas en oraciones. La comunicación por medio del lenguaje se puede denominar comunicación lingüística; además, existen otras formas de comunicación como reír, sonreír, gritar, etc.; las cuales forman parte de una comunicación no lingüística (Santana, 2016).

2.1.2 Aplicaciones para el Refuerzo Cognitivo y Minería de Datos

Existen varias aplicaciones para realizar refuerzo cognitivo, ya sea desde los métodos tradicionales aplicados por los profesionales de la salud; así como también, software o aplicativos desarrollados para ser usados en un teléfono móvil o computador personal. A continuación, son detallados algunos de los métodos más comunes dentro del refuerzo cognitivo.

Juegos Serios: Según Lopes et al. (2018), un juego serio es una prueba mental con reglas específicas; en éstos, se utiliza la diversión como una forma de capacitación corporativa. Las aplicaciones de juegos serios relacionados con la salud y el cuidado de la salud mental son cada vez más comunes (Fleming, 2014). En la actualidad, existe una gran cantidad de juegos serios con distintos enfoques, siendo los más comunes los juegos orientados al refuerzo o entrenamiento de lenguaje, memoria y atención; además, las ventajas que ofrecen las aplicaciones informáticas como recurso terapéutico son: i) Soporte a la mejora del rendimiento del usuario frente a la ejecución de una tarea, ii) el tamaño de los estímulos y iii) la modalidad de presentación. Los juegos serios proporcionan una retroalimentación en un paciente que resulta más motivadora y competitiva, ya que los mismos proporcionan al usuario participe, otra forma de experiencia de juego, frente a los ejercicios tradicionales en papel y lápiz (Portellano, 2005).

Compensación y sustitución: Por sustitución se entiende la construcción de un método nuevo que reemplace la respuesta tras el daño producido por una lesión cerebral, es decir, es aplicada cuando se produce la pérdida completa de un sistema funcional. Este tipo de rehabilitación cognitiva, pone énfasis en realizar actividades con un objetivo funcional mediante ayudas externas que eliminen o reduzcan la necesidad de requisitos cognitivos. Por ejemplo: domótica, sistemas de audio para vida asistida, libros sonoros para no videntes, entre otros (Portellano, 2005).

Terapia en Grupo: La terapia en grupo trabaja en la necesidad emocional de entender los problemas y compartirlos con otros pacientes que estén en las mismas condiciones. La utilización de grupos mejora la conducta social que fue alterada tras algún daño, tiene un menor coste económico, además, facilita la adquisición de conciencia del problema del cual es frecuentemente afectado el paciente (Portellano, 2005).

Haciendo énfasis en los juegos serios, los cuales generan una cantidad considerable de datos asociados a los usuarios que son participes de estos, existen varios métodos de minería de datos que pueden ser utilizados para generar conocimiento. Según Riquelme et al. (2006), la minería de datos consiste en la búsqueda en grandes volúmenes de datos (bases de datos) para encontrar información de provecho que sirva para la toma de decisiones. En este sentido, las técnicas más comunes de minería de datos son descritas a continuación:

Clasificación: Hace referencia a realizar clasificación de datos dentro de las clases categóricas predefinidas. Se permite responder a los siguientes tipos de preguntas: ¿Este cliente puede optar por un crédito?, ¿Qué tipo de enfermedad posee un paciente?, entre otras.

Clustering: Al hablar del término clúster se hace referencia a agrupar registros, observaciones u objetos similares, es decir, un clúster es una colección de registros que son similares entre sí, y diferentes a los registros de otro clúster.

Regresión: Es un campo de estudio, la cual hace énfasis en la relación estadística entre dos variables continuas, conocidas como respuesta y variables de predicción. Existen casos que pueden haber más de una variable a predecir, convirtiéndose el problema en una regresión lineal múltiple.

Generación de reglas: Esta técnica, permite generar reglas que hacen referencia al descubrimiento de relaciones de asociación y dependencias funcionales entre los diferentes atributos analizados.

2.1.3 Análisis de Datos

Según Loh et al. (2015), los objetivos principales de analizar uno o varios juegos serios son i) obtener información útil y valiosa para mejorar el juego o el diseño de aprendizaje y ii) mejorar las habilidades y el rendimiento de los aprendices de juego para convencer mejor a las partes interesadas de la efectividad del juego. Los métodos de recopilación de datos utilizados para el análisis de juegos serios se pueden separar en dos categorías: in situ y ex situ. La recopilación in situ se produce en el propio juego (por ejemplo, registrar eventos de juego, tiempos y variables demográficas del usuario), mientras que ex situ son datos recopilados fuera del juego. Las discusiones de grupos focales y las encuestas posteriores a la prueba son ejemplos de datos ex situ. Los métodos de recopilación in situ son preferidos por los investigadores frente a los ex situ, esto es debido a que se elimina una gran cantidad de datos subjetivos, los cuales no cumplen con la investigación de alto nivel (Quellmalz, 2012).

Existen diferentes tipos de datos que contienen un tipo diferente de información, que puede ayudar a las partes interesadas dentro del campo de e-health, telemedicina, entre otros, a generar conocimiento sobre los mismos. Según Rashid et al. (2020), estos datos pueden ser caracterizados en tres categorías, las cuales son:

Datos cualitativos: Los datos cualitativos se ocupan de números y variables que puedan ser medidas de forma objetiva, es decir, en el escenario que un paciente visite un consultorio médico, la información proporcionada (edad, género, estatura) es almacenada en una base de datos, para su posterior procesamiento, análisis y difusión.

Datos cuantitativos: Estos datos al contrario de los cualitativos, se ocupan de características que son se pueden medir fácilmente pero que se pueden observar subjetivamente. Por ejemplo: Encuestas relacionadas a cierta temática, las cuales están orientadas a evaluar o generar estadísticas.

Datos multimedia: Este tipo de datos incluye las imágenes médicas como tomografías, resonancias magnéticas, rayos X, entre otros.

Finalmente, existen varias técnicas en el campo del análisis de datos (Rashid, 2020), las cuales permiten generar conocimiento, entre los métodos más comunes se listan los siguientes:

Aprendizaje profundo: En el campo de la medicina, las imágenes médicas suelen ser interpretadas por un especialista, lo cual está limitado debido a su subjetividad, complejidad de la imagen y variación de opinión entre los distintos intérpretes. Por este motivo, aparece este tipo de técnicas llamadas aprendizaje profundo (Deep Learning), que consisten en entrenar un procesador para que realice tareas tales como: reconocimiento de voz, reconocimiento de imágenes y predicciones, configurando parámetros básicos acerca de los datos para su posterior procesamiento.

Redes bayesianas: Las redes bayesianas son un modelo gráfico probabilístico, que representa las variables aleatorias del modelo y sus influencias entre ellas por medio de un grafo acíclico dirigido. Este grafo, cada borde (arista) corresponde a una dependencia condicional y cada nodo corresponde a una variable aleatoria única. Dentro del campo de salud, una aplicación para esta técnica es la representación de las relaciones estadísticas entre las enfermedades y síntomas.

Redes neuronales: Dentro del campo de la salud mental, la predicción de enfermedades es muy importante para proteger a los pacientes de enfermedades graves. El aprendizaje automático y la inteligencia artificial se han ido desarrollando de tal manera que permite abordar este tipo de problemas, por lo que, una red neuronal permite realizar tareas de predicción sobre cierto dataset o información obtenida.

2.2 Estado del Arte

En el campo de la Telemedicina, la minería de datos proporciona la generación de conocimiento necesaria para la toma de decisiones por parte del experto del dominio. Wong et al. (2006) presenta una arquitectura para el intercambio de información entre departamentos de un hospital, por medio de un servidor local; el cual, se convierte en un repositorio central de datos para compartir los registros de los pacientes. El objetivo de ese estudio fue centralizar la información en un servidor local, para en un futuro realizar minería de datos sobre la información obtenida de cada uno de los pacientes. Por otra parte, Gheorgue et al. (2014) presenta la importancia de integrar las técnicas de minería de datos dentro de los sistemas de telemedicina, para mejorar los servicios de atención médica brindados a los pacientes, para lograr esto se implementó un almacén de datos, que recopila los datos más relevantes de fuentes internas y externas, para su posterior tratamiento y generación de conocimiento. Así, dicho estudio aplica técnicas clasificación, reglas de asociación y agrupamiento. Finalmente, Khanapi et al. (2015) presenta la implementación de un almacén de datos orientado a las aplicaciones de telemedicina, dicho almacén de datos fue evaluado utilizando la técnica de casos de prueba, los resultados obtenidos mostraron que dicho insumo, proporciona los elementos más importantes de información al sistema de telemedicina, y es útil al momento de usarlo durante las consultas a pacientes.

Por otra parte, existen varios estudios relacionados con el campo de los juegos serios, Orellana et al. (2022) presenta la aplicación de minería de datos en un juego serio orientado al refuerzo de la atención, en esta investigación se aplica la técnica de clustering (K-means), en este estudio, lo que se buscó es describir patrones entre jugadores, los atributos analizados fueron el nivel de juego y el tiempo de respuesta. Los resultados de esta investigación mostraron un patrón oculto basado en el género entre los jugadores y su comportamiento en los diferentes niveles. Por otra parte, Halim et al. (2019) propone el uso de técnicas de minería de datos para correlacionar información con los cinco principales rasgos de personalidad (franqueza, escrupulosidad, extraversión, simpatía y neuroticismo), utilizando tres conjuntos de datos de juegos online (starcraft, world of warcraft y age of empires), para ello se aplicó las siguientes técnicas de clustering: k-means, k-medias, fuzzy c-mean y agrupamiento jerárquico. En cambio, Braun et al. (2017) presenta un algoritmo de minería de datos para agrupar y visualizar datos del juego Overwatch, el cual utiliza la propagación de afinidad para realizar la agrupación y gráficos bidimensionales para visualizar los datos. En este sentido, los datos del juego son usados para generar conocimiento sobre los jugadores actuales. Mismo que permite el análisis de grupos individuales, y proporciona estadísticas que tienen una alta correlación con las estrategias de los jugadores ganadores, estas estadísticas permiten a los nuevos jugadores conocer cómo se juega un personaje y, por lo tanto, pueden ayudar a conocer sus prioridades sobre cada héroe personaje. En otras palabras, el algoritmo ayuda a analizar los datos de juego en línea, obtener información sobre la agrupación o grupos de jugadores y ofrecer sugerencias a los nuevos jugadores del juego. Por otro lado, Slimani et al. (2018) presenta un juego serio llamado Elisa, orientado a la formación médica aplicado a un grupo de estudiantes de biología. Este juego consta de tres pasos: i) Preparación de las muestras, ii) Ejecución del juego y iii) Evaluación de los resultados. Los resultados de este juego se utilizaron como insumo para la aplicación de técnicas de clustering, permitiendo identificar cinco clústeres asociados a la efectividad del aprendizaje. Finalmente, Robles-Bykbaev et al. (2019), presenta una plataforma interactiva dirigida a la cultura Cañari del Ecuador, dicha aplicación tiene como objetivo, proporcionar un entorno en donde los niños puedan jugar juegos serios educativos accediendo al contenido multimedia de esta cultura; además, ésta plataforma posee un módulo de minería de datos para analizar automáticamente la actividad de los usuarios para sugerir áreas educativas que deben reforzarse.

En la Tabla 1, se muestran los estudios obtenidos, si bien existen varias aplicaciones de minería de datos, ya sea dentro del área cognitiva, educación, videojuegos, entre otros. No existen estudios que analicen los datos de juegos serios orientados al refuerzo de la memoria. Es por esto que, este trabajo de investigación busca aplicar técnicas de minería de datos, para determinar los factores personales que impactan en el desempeño de un jugador frente a un juego serio orientado al refuerzo de la memoria.

Tabla 1. Tabla de estudios relacionados.

<i>Estudio</i>	<i>Principales características</i>	<i>Área de aplicación</i>
----------------	------------------------------------	---------------------------

Gheorgue et al. (2014)	Centralizar los datos de pacientes en un almacén de datos, para su posterior aplicación de minería de datos.	Telemedicina
Khanapi et al. (2015)	Implementación de un almacén de datos orientado a las aplicaciones de telemedicina.	Telemedicina
Orellana et al. (2022)	Aplicación de minería de datos en un juego serio orientado al refuerzo de la atención.	Juegos serios
Halim et al. (2019)	Uso de técnicas de minería de datos para correlacionar información con los cinco principales rasgos de personalidad.	Videojuegos
Braun et al. (2017)	Algoritmo de minería de datos para agrupar y visualizar datos del juego Overwatch	Videojuegos
Slimani et al. (2018)	Presenta un juego serio llamado Elisa, orientado a la formación médica aplicado a un grupo de estudiantes de biología.	Juegos serios
Robles-Bykbaev et al. (2019)	Plataforma interactiva dirigida a la cultura Cañari del Ecuador	Educación

3. Metodología Aplicada

La metodología guía utilizada en este trabajo de titulación es Crisp-DM (Wirth & Hipp, 2000), dicha metodología fue adaptada de tal manera que permita cumplir con los objetivos planteados en este trabajo de titulación. Para ello, se utilizó el lenguaje de modelado SPEM 2.0 y sus especificaciones, con el fin de representar de manera gráfica la metodología como un método de minería de datos como se muestra en la Figura 3.

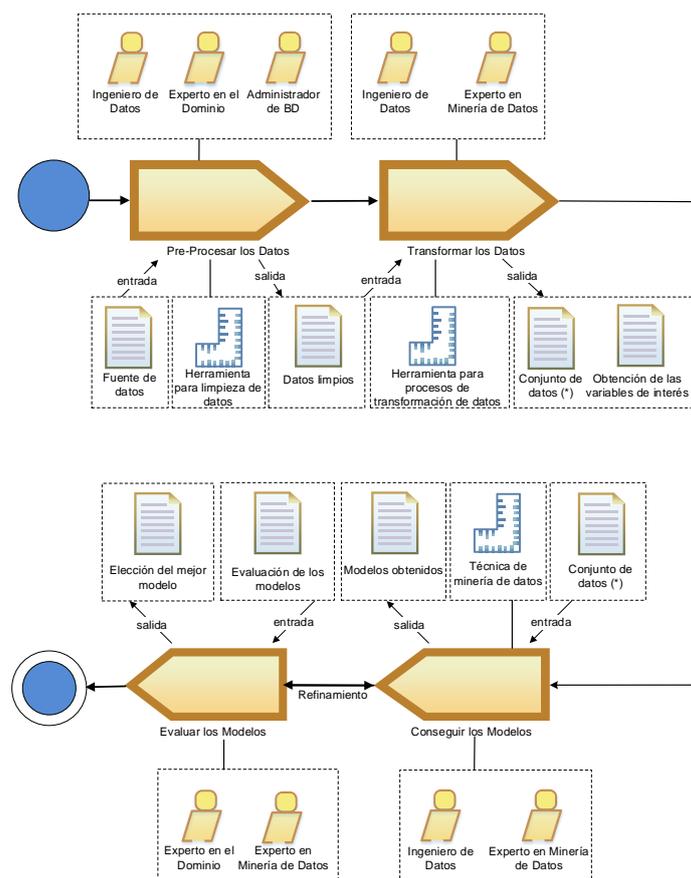


Figura 3. Diagrama SPEM de la metodología Crisp-DM adaptada a este trabajo de titulación.

3.1 Actores

En la Figura 4, se muestran los actores involucrados a lo largo del desarrollo y ejecución de la metodología. Posteriormente se detalla de manera general cada uno de ellos, con la finalidad de dar a conocer sus características y tareas.

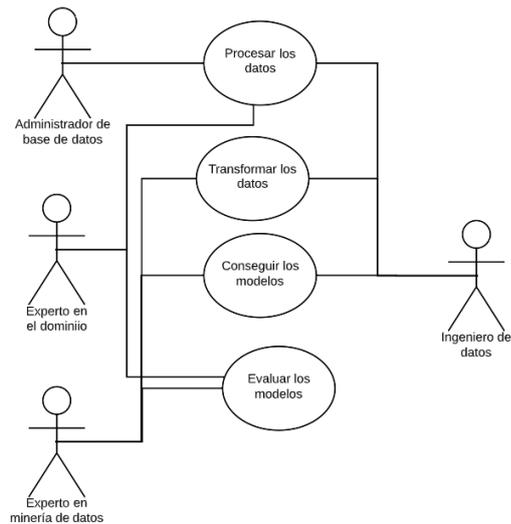


Figura 4. Diagrama de casos de uso, respecto a los actores que participan en la metodología.

Experto en minería de datos

Un experto en minería de datos es comúnmente conocido como un minero o explorador de datos, por medio del descubrimiento de patrones en medio de enormes cantidades de datos. Su intención es la de aportar información valiosa a las organizaciones, para la toma de decisiones futuras. (García, 2019)

Experto en el dominio

Un experto en el dominio es una persona que posee amplios conocimientos en un tema determinado, en este caso, debe poseer estudios en psicología y juegos serios. Según Campbell et al. (2013), un experto en el dominio posee las siguientes características:

- Es una persona que posee los conocimientos o habilidades dentro de un área en particular.
- Es un término que es usado con frecuencia en el desarrollo de software, el cual, hace referencia al término del dominio dentro de un área en específico, más no al dominio de software.

Ingeniero de datos

Un ingeniero de datos es el encargado de la gestión y almacenamiento de los datos con el objetivo de detectar tendencias. De esta persona depende como se procesan los datos, es decir, su objetivo es transformar la materia prima de tal manera, que sea la adecuada para que sea convertida en productos. (UNIR, 2020)

Administrador de base de datos

El administrador de base de datos (DBA), es el responsable de coordinar y asegurar que las tareas asociadas al mantenimiento de un entorno de base datos sean las adecuadas. Un DBA debe asegurar que las bases de datos de una entidad y sus programas relacionados

funcionen de manera correcta en los distintos entornos ya sea producción, pruebas, producción, entre otros. (ComputerWeekly, 2021)

3.2 Pre-Procesar los Datos

La primera fase de este método, permite realizar la limpieza de la fuente de datos, es por esto que, es de vital importancia elegir una herramienta adecuada que permita alcanzar este objetivo, para que de esta manera se obtenga una fuente de datos lo más limpia posible para su posterior tratamiento, para su mejor comprensión de esta fase se la puede observar en la Figura 5.

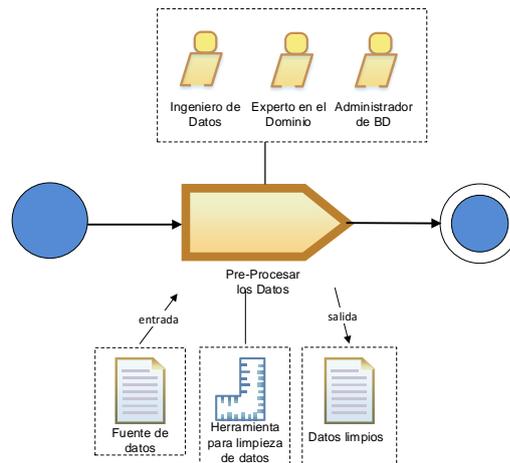


Figura 5. Fase de pre-procesamiento de los datos.

La fase de pre-procesar los datos se compone de una tarea: generación de la nueva fuente de datos limpia; de la cual, se puede visualizar los distintos roles, productos y guías que se obtienen al finalizar esta tarea.

El pre-procesamiento de los datos está soportado por tres roles, los cuales son: 1) Ingeniero de datos, 2) Experto en el dominio, y 3) Administrador de base de datos; quienes tienen como entrada la fuente de datos y como salida de esta actividad los datos limpios para su posterior tratamiento en la siguiente fase.

3.3 Transformar los Datos

La segunda fase de este método, permite tratar los datos a conveniencia según sea necesario, es decir, realizar procesos de etiquetado, transformación, entre otros. Tareas necesarias, las cuales permitan aplicar técnicas de minería de datos en la siguiente fase de este modelo, para su mejor comprensión de esta fase se la puede observar en la Figura 6.

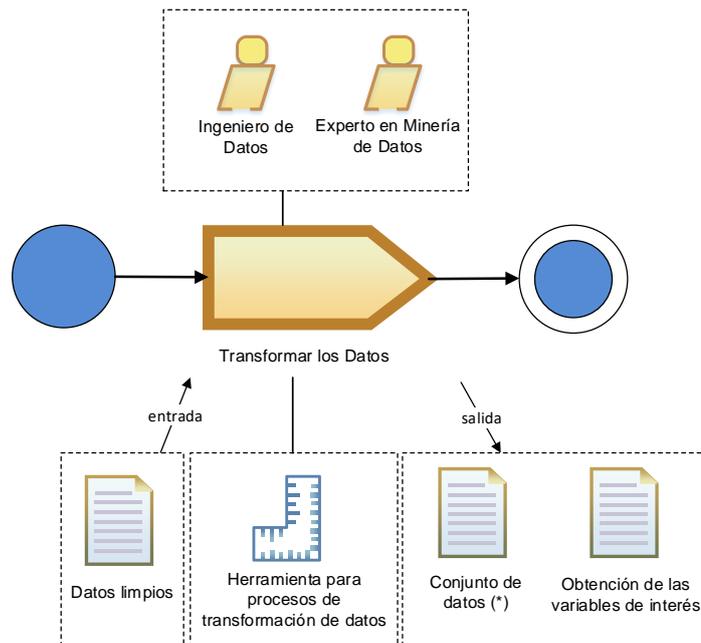


Figura 6. Fase de transformación de los datos.

La fase de transformar los datos se compone de una tarea: obtener las variables más relevantes, para su posterior generación de un modelo de minería de datos; de esta fase se pueden visualizar los roles, productos y guías necesarios.

La transformación de datos está soportada por dos roles, los cuales son: 1) Ingeniero de datos, y 2) Experto en minería de datos; quienes tienen como entrada los datos limpios y la herramienta para procesos de transformación de datos, y como salida de esta actividad el conjunto de datos y las variables de interés.

3.4 Conseguir los Modelos

En esta fase, se obtienen varios modelos de minería de datos en función del nuevo conjunto de datos obtenido en la fase anterior, estos modelos son utilizados en la fase final para determinar el modelo que cumplan con los objetivos planteados en fases iniciales, para su mejor comprensión de esta fase se la puede observar en la Figura 7.

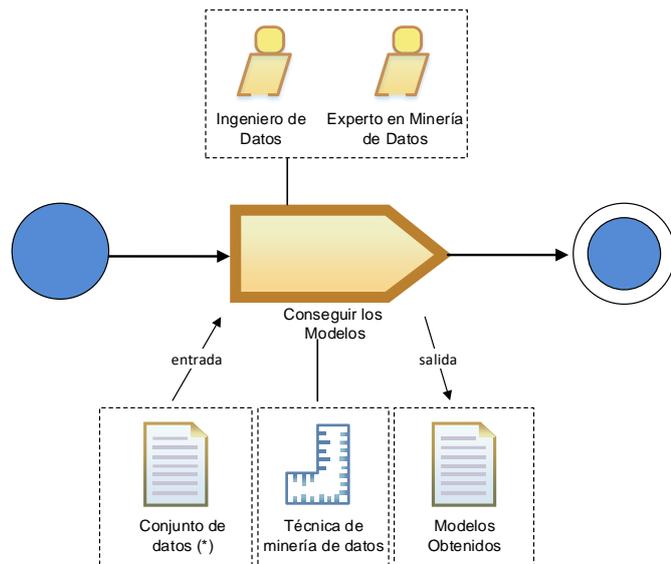


Figura 7. Fase de obtención del modelo.

La fase de conseguir los modelos se compone de una tarea: generar varios modelos de minería de datos; del cual, se puede visualizar los distintos roles, productos y guías que se obtienen al finalizar esta tarea.

Esta fase está soportada por dos roles: 1) Ingeniero de datos, y 2) Experto en minería de datos; quienes tienen como entrada el dataset y como salida de esta actividad se generan varios modelos de minería de datos para su posterior evaluación.

3.5 Evaluar los Modelos

En la fase final de este método, se realiza la evaluación de cada uno de los modelos obtenidos, de los cuales se definen métricas para determinar el mejor modelo. Mismo que debe generar conocimiento, y permitir explicar el comportamiento de los datos tratados en fases anteriores, para su mejor comprensión de esta fase se la puede observar en la Figura 8.

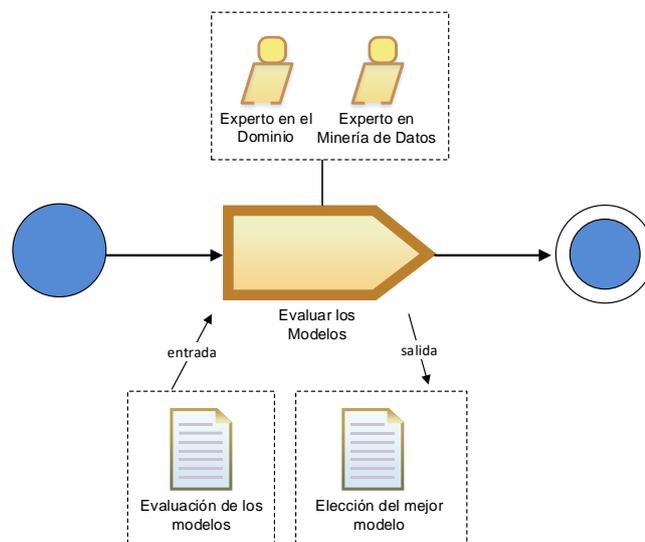


Figura 8. Fase de evaluación.

Esta fase se compone de una tarea: elegir el modelo idóneo de minería de datos; del cual, se puede visualizar los distintos roles y productos que se obtienen al finalizar esta tarea.

Finalmente, esta fase está soportada por dos roles: 1) Experto en el dominio, y 2) Experto en minería de datos; quienes tienen como entrada el conjunto de modelos de minería de datos y como salida de esta actividad se obtiene el modelo adecuado que permita cumplir con los objetivos planteados en la fase inicial de este método.

4. Instanciación de la Metodología

En este capítulo, se detalla la adaptación de un juego serio para recoger información; además, se realiza la instanciación de la metodología, explicando de manera detallada la ejecución de cada uno de los pasos, guías, tareas y entregables obtenidos. Finalmente, se obtiene un modelo de minería de datos, para su posterior evaluación y validación mediante un caso de estudio.

4.1 Adaptación del juego serio

Antes de instanciar la metodología propuesta, como parte de los objetivos de este trabajo de titulación se realizó la adaptación de un juego serio orientado al refuerzo de la memoria, el proceso realizado para lograr esto es descrito a continuación.

Juego serio adaptado

Las variables incluidas dentro del juego serio fueron las siguientes: 1) Nivel de ingresos social, 2) Carrera Universitaria, 3) Nivel de instrucción primaria, 4) Tipo de colegio, 5) Si tuvo o no Covid-19, 6) Antecedentes neurodegenerativos, 7) Tiempo de uso estimado de una computadora al día, 8) Cantidad de horas que realiza deporte, 9) Peso y 10) Altura.

Estas variables fueron agregadas según Organización Mundial de la Salud (OMS) y se encuentran descritas en la sección 2.1.1 Salud Mental y Funciones Cognitivas, el juego serio desplegado se muestra en la Figura 9, Figura 10 y Figura 11.

← Formulario de ingreso, los campos con (*) son obligatorios.
La información proporcionada será anonimizada, para fines académicos y de investigación.

Nombres y Apellidos (*)

Cédula (*)

Fecha de Nacimiento (*)

Ciudad (*)

Teléfono (*)

Peso aproximado en KILOGRAMOS (*) (Ejemplo: 50.4)

Altura aproximada en METROS (*) (Ejemplo: 1.50)

Cantidad en horas que realiza deporte al día (*) (Ejemplo 5)

Empezar

Figura 9. Inclusión de las variables dentro del juego serio desplegado.

Cantidad en horas que utiliza una computadora o celular al día (*) (Ejemplo 1)

Seleccione su tipo de colegio: (*) Ninguno ↓

Seleccione su carrera: (*) Ninguno ↓

Seleccione su nivel de instrucción actual (*): Primaria ↓

Se ha contagiado de Covid? (*): No ↓

Posee antecedentes neurodegenerativos? (*): No ↓

Nivel de ingresos económicos (*): 0-1000 ↓

Seleccione su género (*): Masculino ↓

¿Posee alguna discapacidad?

Empezar

Figura 10. Variables agregadas dentro del formulario de registro.

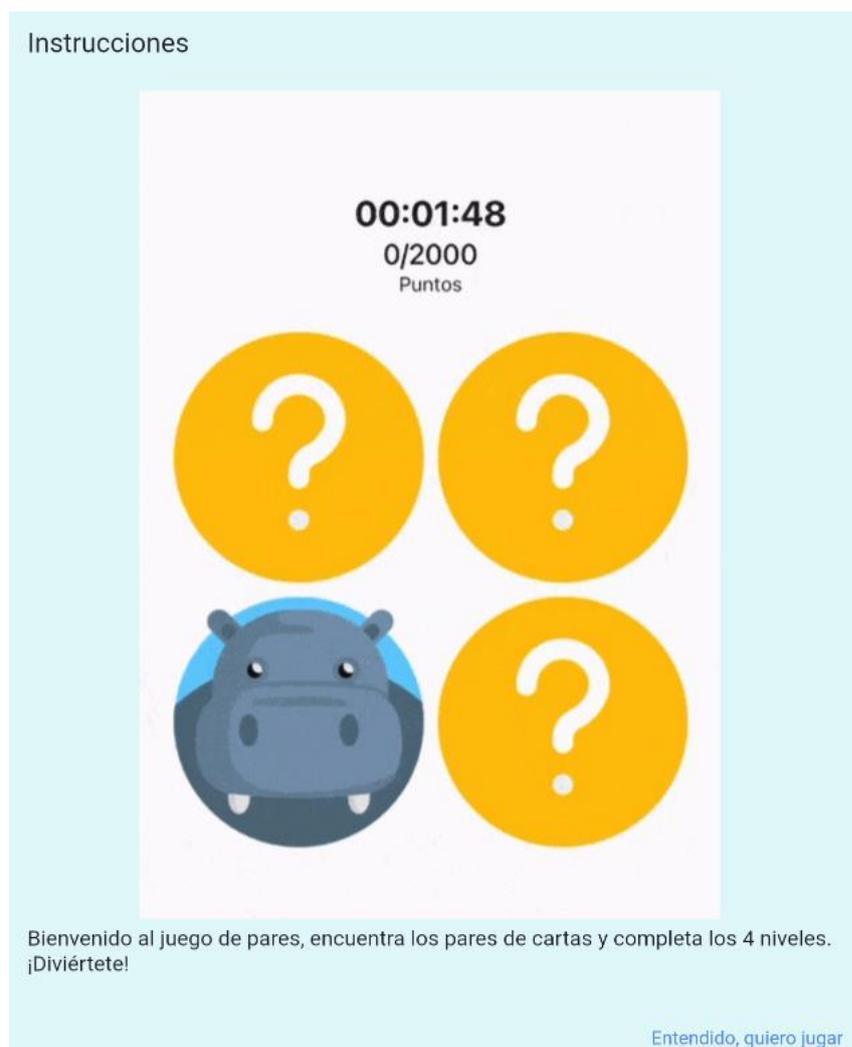


Figura 11. Juego serio desplegado.

El link de acceso del juego serio es: <https://jserionew-8e818.web.app/#/>, el cual se encuentra funcional; además, puede ser accedido por cualquier usuario para su uso y refuerzo de la memoria, ya sea desde una computadora como también un celular.

4.2 Fase de pre-procesar los datos

Esta fase comprende una tarea la cual es la de pre-procesamiento de los datos; además, se describen en este apartado cada una de las tareas, productos de trabajos y guías que fueron utilizadas para obtener los entregables de esta fase, dichos elementos pueden ser visualizados en la Figura 12.

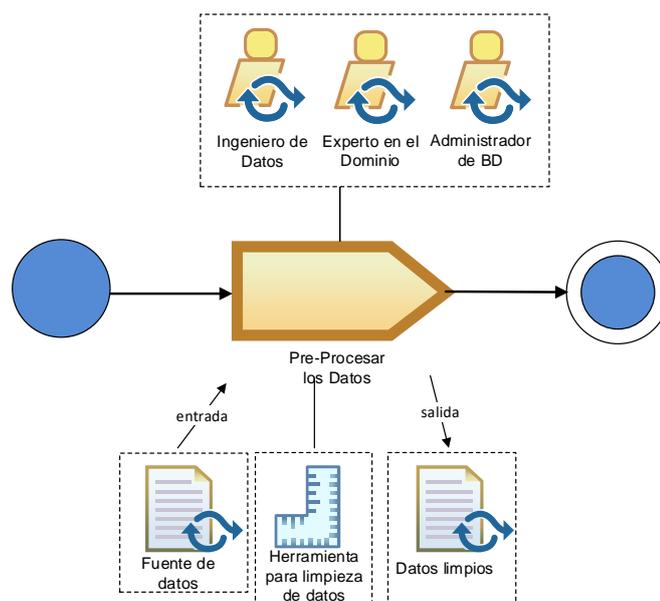


Figura 12. Fase de pre-procesamiento de los datos del juego serio.

4.2.1 Pre-procesar los datos

En esta tarea, se hace uso de la fuente de datos obtenida del juego serio, para ello se debe definir una herramienta para realizar la limpieza de datos, con el fin de tener información sin ruido para el posterior uso, dicha herramienta y fuente de datos son descritas a continuación.

Fuente de datos

Las variables recopiladas al hacer uso del juego serio, son descritas en la Tabla 2.

Tabla 2. Variables recopiladas al hacer uso del juego serio.

Nombre Variable	Tipo	Descripción
FechaJuego	Date	Se almacena la fecha que le usuario realizó el intento.
TiempoFinal	String	Contiene el tiempo empleado en jugar todos los niveles.
TiemposNivel4_tiempolnicio	String	Contiene el tiempo al iniciar el nivel 4 del juego.

TiemposNivel4_tiempoFinal	String	Contiene el tiempo al finalizar el nivel 4 del juego.
TiemposNivel3_tiempoInicio	String	Contiene el tiempo al iniciar el nivel 3 del juego.
TiemposNivel3_tiempoFinal	String	Contiene el tiempo al finalizar el nivel 3 del juego.
TiemposNivel2_tiempoInicio	String	Contiene el tiempo al iniciar el nivel 2 del juego.
TiemposNivel2_tiempoFinal	String	Contiene el tiempo al finalizar el nivel 2 del juego.
TiemposNivel1_tiempoInicio	String	Contiene el tiempo al iniciar el nivel 1 del juego.
TiemposNivel1_tiempoFinal	String	Contiene el tiempo al finalizar el nivel 1 del juego.
Cédula	String	Contiene la cédula del usuario.
Peso	String	Contiene el Peso del jugador.
Puntos	String	Hace referencia al puntaje obtenido del usuario.
Edad	Int	La edad del jugador.
ComputadorHoras	Int	La cantidad de horas que el jugador emplea en una computadora o celular.
Carrera	String	En el caso que posea una carrera universitaria.
EnfermedadCovid	Boleano	Si el usuario se enfermó o no de Covid.
AntecedentesNeuro	Boleano	Si el usuario posee antecedentes neurodegenerativos.
TipoColegio	String	El tipo de colegio del cual proviene el usuario.
FechaNacimiento	Date	La fecha de nacimiento.
Discapacidad	Boleano	Si posee alguna discapacidad.
DeporteHoras	String	La cantidad de horas que emplea haciendo deporte.
Teléfono	String	Un número de contacto.
NivelInstruccion	String	Nivel de instrucción más alta del jugador.
IngresosEconomicos	String	Los ingresos económicos del usuario en dólares.
Nombre	String	Nombre del jugador.
Altura	String	Altura del jugador.
TieneDiscapacidad	Boleano	Si el usuario posee o no discapacidad.
TiempoNivel1	String	El tiempo empleado en el nivel 1.
TiempoNivel2	String	El tiempo empleado en el nivel 2.
TiempoNivel3	String	El tiempo empleado en el nivel 3.
TiempoNivel4	String	El tiempo empleado en el nivel 4.

Adicionalmente, el dataset está formado por 33 variables, de las cuales las columnas cédula y nombres han sido eliminadas debido a que se busca anonimizar la información. Sin

embargo, existen variables que deben pasar por un proceso de limpieza, debido a que se necesita cierta información asociada a dicha columna, es por esto que, en la Tabla 3 se muestran las variables consideradas para realizar este proceso.

Tabla 3. Variables a aplicar la limpieza de datos.

<i>Variable</i>	<i>Limpieza</i>
Peso	Debe contener solo números y se muestra en pocas ocasiones datos del tipo "33.5 kg".
ComputadorHoras	Debe contener solo números y se muestra en pocas ocasiones datos del tipo "5 horas".
deporteHoras	Debe contener solo números y se muestra en pocas ocasiones datos del tipo "1 h".
Carrera	Posee caracteres basura, esto es debido a las tildes.
Altura	Debe contener solo números y se muestra en pocas ocasiones datos del tipo "1.69 metros".
Ciudad	Posee caracteres basura, esto es debido a las tildes.

En la Figura 13 se muestra los datos recopilados que poseen caracteres basura, o que pueden causar conflicto al momento de realizar el proceso de obtención del modelo de minería de datos.

computadorHoras	carrera
	7 Ciencias de la Computaci ³ n
	4 Otros
	4 Estudios Internacionales
	10 Ciencias de la Computaci ³ n
	10 Ciencias de la Computaci ³ n
	10 Ciencias de la Computaci ³ n
	10 Ciencias de la Computaci ³ n
	10 Ciencias de la Computaci ³ n
6h	Medicina
6h	Medicina
6h	Medicina

Figura 13. Datos que deben ser llevados a un proceso de limpieza.

Herramienta para la limpieza de datos

La herramienta utilizada para realizar el proceso descrito anteriormente es RapidMiner (Lee & Mierswam, 2022), la cual permite realizar ya sea procesos de minería de datos, como también limpieza de datos, en la Figura 21 se muestran los conectores usados dentro de dicha herramienta.

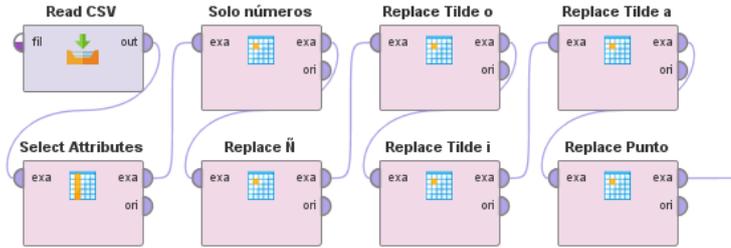


Figura 14. Herramienta RapidMiner usada para la limpieza de datos.

Una vez finalizado la limpieza de datos, se realizó un análisis estadístico de las variables, con el fin de determinar si existen columnas con datos faltantes, para su posterior tratamiento, este proceso puede ser visualizado en la Figura 15.

Name	Type	Missing	Statistics
peso	Polynomial	0	Least: 97.5 (1) Most: 150 (126)
computadorHoras	Polynomial	0	Least: 2.30 (1) Most: 6 (195)
deporteHoras	Polynomial	0	Least: 5 (1) Most: 0 (396)
altura	Polynomial	0	Least: 164 (1) Most: 156 (123)
carrera	Polynomial	0	Least: Marketing (1) Most: Otros (237)
ciudad	Polynomial	0	Least: Pasaje-El Oro (1) Most: Cuenca (595)
fechaJuego	Polynomial	0	Least: 2022-06- [...] 00:00 (1) Most: 2022-05- [...] 0:00 (11)
tiempoFinal	Polynomial	0	Least: 05:45.2 (1) Most: 02:07.7 (19)
tiemposNivel4_tiempolnicio	Polynomial	0	Least: 03:21.4 (1) Most: 00:26.6 (16)
tiemposNivel4_tiempoFinal	Polynomial	0	Least: 05:45.2 (1) Most: 02:07.7 (19)
tiemposNivel1_tiempolnicio	Polynomial	0	Least: 00:19.0 (1) Most: 00:01.9 (73)
tiemposNivel1_tiempoFinal	Polynomial	0	Least: 00:58.9 (1) Most: 00:04.0 (55)
tiemposNivel3_tiempolnicio	Polynomial	0	Least: 01:44.1 (1) Most: 00:11.9 (19)

Figura 15. Estadística de las variables a utilizar.

Como se puede observar en la figura anterior, no existen columnas con datos faltantes o nulos, esto es debido a que, al momento de realizar la programación y adaptación del juego, se realizaron las validaciones respectivas en el ingreso de los datos, con el fin de optimizar el tiempo y omitir este proceso.

Datos limpios

La nueva fuente de datos obtenida al utilizar el proceso de limpieza de datos, puede ser visualizada en la Figura 16, la cual sirve de entrada para los siguientes fases del método propuesto.

A	B	C	D	E	F	G	H	I	J	K	L	M
computadorHoras	deporteHoras	carrera	puntos	tiempoFinal	antecedentesNeuro	tipoColegio	nivelInstruccion	enfermadoCovid	genero	ingresosEconomicos	edad	IMC
0	0	Ninguno	1700	01:50.8	No	Ninguno	Primaria	No	Masculino	0-1000	56	27.94
0	0	Ninguno	1575	03:46.8	No	Ninguno	Primaria	No	Masculino	0-1000	56	27.94
8	0	Otros	1700	01:32.2	No	Urbano	Tercer Nivel	No	Masculino	0-1000	46	26.71
2	0	Otros	1675	01:11.3	No	Urbano	Tercer Nivel	No	Femenino	0-1000	42	28.58
6	1	Otros	1800	01:02.9	No	Urbano	Tercer Nivel	No	Masculino	0-1000	58	26.42
3	0	Otros	1575	01:57.6	No	Ninguno	Primaria	Si	Femenino	0-1000	58	30.43
4	0	Ninguno	1650	01:57.2	No	Urbano	Secundaria	No	Femenino	0-1000	53	19.83
10	2	Otros	1625	01:17.2	No	Ninguno	Tercer Nivel	No	Femenino	1000-2000	50	31.11
12	1	Ninguno	1450	02:54.2	No	Ninguno	Cuarto Nivel	Si	Masculino	1000-2000	55	26.56
10	0	Otros	1725	01:17.3	No	Urbano	Cuarto Nivel	Si	Femenino	2000-3000	35	26.22
1	0	Otros	1725	01:15.9	No	Ninguno	Tercer Nivel	Si	Masculino	3000-4000	54	23.45
2	2	Ninguno	1150	05:45.2	No	Ninguno	Primaria	No	Femenino	0-1000	69	26.67
8	1	Derecho	1700	01:15.8	No	Ninguno	Tercer Nivel	Si	Masculino	0-1000	43	25.95
1	1	Ninguno	1725	01:15.1	No	Ninguno	Tercer Nivel	Si	Femenino	2000-3000	40	20.28
8	0	Contabilidad	1625	01:36.9	No	Urbano	Tercer Nivel	Si	Femenino	0-1000	41	25.81
1	1	Derecho	1775	01:18.8	No	Urbano	Tercer Nivel	Si	Masculino	0-1000	38	25.21
3	1	Psicologia Educativa	1850	01:05.0	No	Rural	Tercer Nivel	No	Masculino	0-1000	37	27.18
4	1	Psicologia Clinica	1675	01:13.6	No	Ninguno	Secundaria	No	Femenino	0-1000	45	23.05
4	0	Psicologia Educativa	1900	00:42.4	No	Urbano	Cuarto Nivel	No	Femenino	1000-2000	42	26.67
6	1	Otros	1725	01:15.1	No	Urbano	Cuarto Nivel	Si	Femenino	1000-2000	37	30.76
6	2	Estudios Internacionales	1725	01:33.7	Si	Urbano	Secundaria	Si	Masculino	0-1000	24	23.94
10	1	Ciencias de la Computacion	1825	01:55.5	No	Urbano	Cuarto Nivel	No	Femenino	1000-2000	39	24.22
6	0	Ciencias de la Computacion	1850	00:49.8	No	Ninguno	Secundaria	Si	Masculino	0-1000	25	27.34
12	0	Otros	1850	00:49.8	No	Urbano	Cuarto Nivel	No	Femenino	1000-2000	38	21.88
3	1	Estudios Internacionales	1775	01:42.9	No	Urbano	Tercer Nivel	No	Femenino	0-1000	23	19.33
4	1	Administracion	1750	01:09.1	No	Urbano	Tercer Nivel	Si	Masculino	1000-2000	39	27.78
12	0.5	Estudios Internacionales	1525	02:08.1	No	Urbano	Secundaria	Si	Femenino	0-1000	23	21.93
9	0	Contabilidad	1800	00:56.1	No	Urbano	Tercer Nivel	Si	Femenino	0-1000	36	26.71
10	1	Marketing	1800	01:26.9	Si	Urbano	Tercer Nivel	Si	Femenino	0-1000	33	25.24
9	1	Otros	1800	01:46.4	No	Urbano	Tercer Nivel	No	Femenino	0-1000	32	24.44
6.5	1	Estudios Internacionales	1825	01:02.2	No	Urbano	Tercer Nivel	Si	Masculino	1000-2000	26	21.22
7	1	Ninguno	1725	01:42.2	No	Ninguno	Primaria	No	Masculino	0-1000	34	29.41
3	5	Ciencias de la Computacion	1575	01:28.4	No	Urbano	Secundaria	No	Masculino	0-1000	23	24.97
10	1	Ninguno	1325	01:47.4	No	Ninguno	Cuarto Nivel	Si	Femenino	3000-4000	35	20.31
1	1	Medicina	1675	01:15.6	No	Urbano	Tercer Nivel	No	Femenino	0-1000	22	17.99

Figura 16. Nueva fuente de datos.

4.3 Fase de transformar los datos

Esta fase comprende una tarea la cual es la de transformación de los datos; además, se describen en este apartado cada una de las tareas, productos de trabajos y guías que fueron utilizadas para obtener los entregables de esta fase, dichos elementos pueden ser visualizados en la Figura 17.

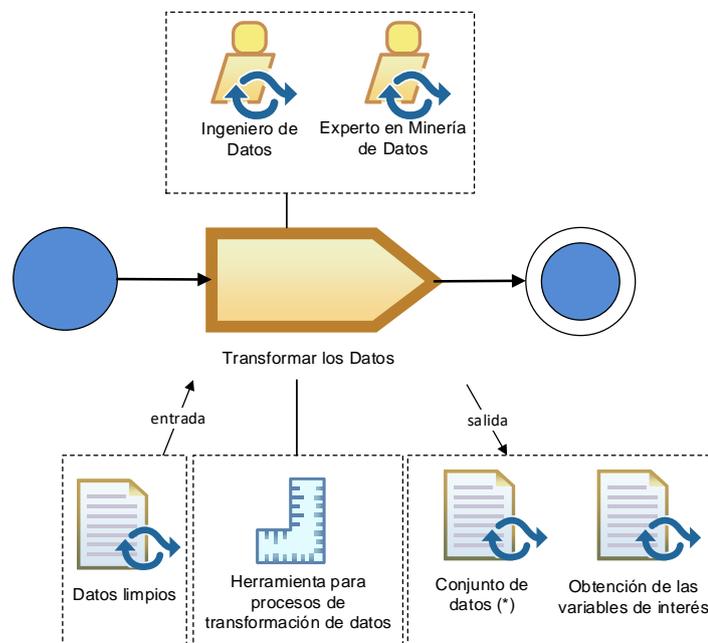


Figura 17. Fase de transformación de los datos.

4.3.1 Transformar los datos

A partir de la nueva fuente de datos, se debe determinar las variables más relevantes; dichas variables nos permitirán obtener un modelo de minería de datos, que permita generar conocimiento o patrones de comportamiento sobre los datos, para ello se utilizaron las siguientes herramientas descritas a continuación.

Herramienta para procesos de transformación

La herramienta seleccionada para realizar el proceso de transformación de datos fue RapidMiner, este software a más de realizar tareas de minería de datos, también permite realizar procesos de etiquetado y transformación de datos; en este sentido, las variables consideradas para el proceso de transformación son mostrada en la Tabla 4.

Para realizar el etiquetado de la variable **edad**, se tomó en cuenta el estudio propuesto por Leuw et al. (1973), en donde propone clasificar a la edad en 6 grupos: **lactancia** desde el nacimiento hasta los 2 años de edad; **infancia** desde los 3 años hasta los 12; **adolescencia** desde los 13 años hasta los 20 años; **adultez inicial** desde los 21 hasta los 40 años; **adultez madura** desde los 41 años hasta los 60 años y, finalmente **vejez** desde los 61 años en adelante.

En cuanto al **índice de masa corporal (IMC)**, según Navarrete et al. (2016) realiza la siguiente clasificación: **peso bajo** cuando el índice es menor a 18.5; **peso normal** cuando el índice está comprendido entre 18.5 a 24.9; **sobrepeso** si el índice está comprendido entre 25 a 29.9 y **obesidad** si el índice está comprendido entre 30 a 34.9, cabe resaltar que existen muchas más divisiones para esta variable, pero, en este caso, todos usuarios que utilizaron el juego estaban comprendidos entre estos grupos, por lo tanto, para evitar tener varias divisiones se hizo uso de estos grupos.

Finalmente, el etiquetado de la variable **puntaje** fue realizado en base al estudio de Scapin et al. (2010), en donde se propone la siguiente clasificación: **excelente** cuando el puntaje obtenido es desde 10 a 9; **muy bueno** cuando el puntaje obtenido es desde 8.9 a 8; **bueno** cuando el puntaje obtenido es desde 7.9 a 7; **regular** cuando el puntaje obtenido es desde 6.9 a 6 y, **bajo** cuando el puntaje obtenido es menor a 5.9. Cabe resaltar que el puntaje máximo es de 2000 y el mínimo 0, en este caso se adaptaron estas etiquetas, con el fin de satisfacer la escala de los puntajes, de forma resumida en la Tabla 4 se muestra los umbrales descritos anteriormente.

Tabla 4. Variables a realizar el etiquetado/transformación de datos.

Variable	Criterio de transformación/división
Altura y Peso	Transformar en índice de masa corporal (IMC). $IMC = \frac{Peso(kg)}{(Altura(m))^2}$
IMC	Menos de 18.5 > Inferior. 18.5 – 24.9 > Normal. 25.0 – 29.9 > Superior.

Edad

Mayor a 30 > Obesidad.

Entre 12 a 19 años > Adolescencia.

Entre 20 a 40 años > Adulthood inicial.

Entre 41 a 60 años > Adulthood madura.

Mayor a 61 años > Vejez

Puntaje

2000 – 1900 > Excelente.

1899 – 1800 > Muy Bueno.

1799 – 1700 > Bueno.

1699 - 1600 > Regular.

Menor a 1599 > Bajo.

Previo a realizar la transformación de los datos, en la Figura 18 se muestra el dataset con los datos sin realizar el proceso descrito en la Tabla 4.

A	B	C	D	E	F	G	H	I	J	K	L	M
computadori	deporte	Horas carrera	puntajeEtiqu	puntos	antecede	tipoColegio	nivelinstruccion	edadEtiqueta	IMCEtiqueta	etiquetaTiempo	enfermadoCovid	genero
4	2	Medicina	Excelente	1975	No	Urbano	Tercer Nivel	Adulthood inicial	Superior	Rapido	Si	Masculino
6	0	Medicina	Excelente	1950	No	Urbano	Tercer Nivel	Adolescencia	Normal	Medio	No	Femenino
4	1.5	Estudios Interna	Excelente	1925	Si	Ninguno	Secundaria	Adulthood inicial	Normal	Rapido	No	Masculino
6	2	Otros	Excelente	1925	No	Urbano	Tercer Nivel Tecnológico	Adulthood inicial	Normal	Rapido	No	Masculino
6	2	Medicina	Excelente	1925	No	Urbano	Tercer Nivel	Adulthood inicial	Normal	Rapido	Si	Masculino
18	0	Medicina	Excelente	1925	No	Urbano	Tercer Nivel	Adolescencia	Normal	Rapido	Si	Femenino
2	0	Ninguno	Excelente	1925	No	Urbano	Secundaria	adulthood madura	Obesidad	Rapido	No	Femenino
6	2	Medicina	Excelente	1925	No	Urbano	Secundaria	Adolescencia	Normal	Rapido	Si	Masculino
4	0	Psicologia Educa	Excelente	1900	No	Urbano	Cuarto Nivel	adulthood madura	Superior	Rapido	No	Femenino
17	0	Medicina	Excelente	1900	No	Urbano	Tercer Nivel	Adolescencia	Normal	Rapido	Si	Femenino
6	1	Medicina	Excelente	1900	No	Urbano	Tercer Nivel	Adulthood inicial	Normal	Rapido	Si	Femenino
6	1	Medicina	Excelente	1900	No	Urbano	Tercer Nivel	Adolescencia	Normal	Rapido	Si	Femenino
8	2	Estudios Interna	Excelente	1900	No	Urbano	Tercer Nivel	Adulthood inicial	Normal	Rapido	Si	Femenino
5	2	Ninguno	Excelente	1900	No	Urbano	Secundaria	Adolescencia	Normal	Rapido	Si	Masculino
6	1	Estudios Interna	Excelente	1900	No	Urbano	Secundaria	Adulthood inicial	Superior	Rapido	No	Femenino
8	1	Estudios Interna	Excelente	1900	No	Urbano	Tercer Nivel	Adulthood inicial	Normal	Rapido	Si	Femenino
8	1	Estudios Interna	Excelente	1900	No	Urbano	Cuarto Nivel	Adulthood inicial	Normal	Rapido	No	Femenino
13	0	Administracion	Muy bueno	1875	No	Urbano	Tercer Nivel Tecnológico	Adulthood inicial	Normal	Rapido	No	Femenino
3	3	Otros	Muy bueno	1875	No	Urbano	Secundaria	Adolescencia	Normal	Rapido	Si	Masculino
3	1	Estudios Interna	Muy bueno	1875	No	Urbano	Tercer Nivel	Adulthood inicial	Normal	Rapido	No	Femenino
5	3	Estudios Interna	Muy bueno	1875	No	Urbano	Tercer Nivel	Adulthood inicial	Obesidad	Rapido	Si	Masculino
8	2	Derecho	Muy bueno	1875	No	Urbano	Cuarto Nivel	Adulthood inicial	Normal	Rapido	Si	Masculino
3	1	Psicologia Educa	Muy bueno	1850	No	Rural	Tercer Nivel	Adulthood inicial	Superior	Rapido	No	Masculino
6	0	Ciencias de la Cc	Muy bueno	1850	No	Ninguno	Secundaria	Adulthood inicial	Superior	Rapido	Si	Masculino
12	0	Otros	Muy bueno	1850	No	Urbano	Cuarto Nivel	Adulthood inicial	Normal	Rapido	No	Femenino
4	3	Ciencias de la Cc	Muy bueno	1850	No	Urbano	Tercer Nivel Tecnológico	Adulthood inicial	Normal	Rapido	No	Masculino
4	2	Medicina	Muy bueno	1850	No	Urbano	Secundaria	Adulthood inicial	Normal	Rapido	Si	Masculino
6	2	Otros	Muy bueno	1850	No	Urbano	Secundaria	Adulthood inicial	Normal	Rapido	Si	Femenino
4	2	Estudios Interna	Muy bueno	1850	No	Urbano	Tercer Nivel	Adulthood inicial	Normal	Rapido	No	Masculino
4	2	Estudios Interna	Muy bueno	1850	No	Urbano	Tercer Nivel	Adulthood inicial	Superior	Rapido	No	Femenino
8	0	Otros	Muy bueno	1850	No	Rural	Secundaria	Adulthood inicial	Superior	Rapido	Si	Femenino
5	0	Otros	Muy bueno	1850	No	Urbano	Secundaria	Adulthood inicial	Normal	Rapido	No	Masculino
9	0	Medicina	Muy bueno	1850	No	Urbano	Tercer Nivel	Adolescencia	Normal	Rapido	Si	Femenino
20	1	Estudios Interna	Muy bueno	1850	No	Urbano	Tercer Nivel	Adulthood inicial	Normal	Rapido	Si	Femenino
10	1	Ciencias de la Cc	Muy bueno	1825	No	Urbano	Cuarto Nivel	Adulthood inicial	Normal	Medio	No	Femenino

Figura 18. Dataset sin realizar la transformación de los datos.

El proceso realizado dentro del software RapidMiner para realizar la transformación de los datos propuesto en la Tabla 4, es mostrado en la Figura 19, en el cual se mapean los valores propuestos y se genera un nuevo dataset para su posterior uso.



Figura 19. Proceso de etiquetado de los datos.

Obtención de las variables de interés

En la Tabla 5, se muestran las variables de interés asociadas al juego serio, las cuales servirán para obtener las inferencias necesarias durante el proceso de obtención de los modelos de minería de datos, para su posterior evaluación y validación por medio de un caso de estudio.

Tabla 5. Variables de interés.

Variable	Descripción
Género	Permite conocer si las personas de género femenino o masculino fueron los más eficientes frente al juego serio.
Puntaje	Determina en función de las demás variables, si existe algún tipo de comportamiento de los usuarios.
Enfermado Covid	De igual manera que la variable anterior, permite conocer si una persona que tuvo covid tiene un mejor rendimiento frente una persona que aún no pasa por la enfermedad.
Edad	Se busca determinar si existe algún patrón de comportamiento en función de la edad.
IMC	Se puede determinar si el desempeño del usuario es influido por su índice de masa corporal.

Conjunto de datos (*)

Los conectores utilizados en la herramienta de transformación de datos (Figura 19), permitieron realizar los cambios adecuados al dataset, de tal manera que se permita realizar las tareas necesarias para generar los modelos adecuados, cabe resaltar que, se realizó un proceso enfocado a etiquetado de datos, esto es debido a que se busca generar un conjunto de reglas de asociación, con el fin de generar inferencias con respecto a las variables analizadas, para su posterior validación con un experto en el dominio. Finalmente, el nuevo dataset puede ser visualizado en la Figura 20.

A	B	C	D	E	F	G	H	I
puntajeEtiquetado	antecedentesNeuro	tipoColegio	nivelnInstruccion	edadEtiqueta	IMCEtiqueta	enfermadoC	genero	ingresosEcor
Excelente	No	Urbano	Tercer Nivel	Adultez inicial	Superior	Si	Masculino	1000-2000
Excelente	No	Urbano	Tercer Nivel	Adolescencia	Normal	No	Femenino	0-1000
Excelente	Si	Ninguno	Secundaria	Adultez inicial	Normal	No	Masculino	0-1000
Excelente	No	Urbano	Tercer Nivel Tecnológico	Adultez inicial	Normal	No	Masculino	0-1000
Excelente	No	Urbano	Tercer Nivel	Adultez inicial	Normal	Si	Masculino	0-1000
Excelente	No	Urbano	Tercer Nivel	Adolescencia	Normal	Si	Femenino	0-1000
Excelente	No	Urbano	Secundaria	adultez madura	Obesidad	No	Femenino	0-1000
Excelente	No	Urbano	Secundaria	Adolescencia	Normal	Si	Masculino	2000-3000
Excelente	No	Urbano	Cuarto Nivel	adultez madura	Superior	No	Femenino	1000-2000
Excelente	No	Urbano	Tercer Nivel	Adolescencia	Normal	Si	Femenino	0-1000
Excelente	No	Urbano	Tercer Nivel	Adultez inicial	Normal	Si	Femenino	0-1000
Excelente	No	Urbano	Tercer Nivel	Adolescencia	Normal	Si	Femenino	0-1000
Excelente	No	Urbano	Tercer Nivel	Adultez inicial	Normal	Si	Femenino	0-1000
Excelente	No	Urbano	Secundaria	Adolescencia	Normal	Si	Masculino	0-1000
Excelente	No	Urbano	Secundaria	Adultez inicial	Superior	No	Femenino	0-1000
Excelente	No	Urbano	Tercer Nivel	Adultez inicial	Normal	Si	Femenino	0-1000
Excelente	No	Urbano	Cuarto Nivel	Adultez inicial	Normal	No	Femenino	0-1000
Muy bueno	No	Urbano	Tercer Nivel Tecnológico	Adultez inicial	Normal	No	Femenino	0-1000
Muy bueno	No	Urbano	Secundaria	Adolescencia	Normal	Si	Masculino	0-1000
Muy bueno	No	Urbano	Tercer Nivel	Adultez inicial	Normal	No	Femenino	0-1000
Muy bueno	No	Urbano	Tercer Nivel	Adultez inicial	Obesidad	Si	Masculino	0-1000
Muy bueno	No	Urbano	Cuarto Nivel	Adultez inicial	Normal	Si	Masculino	0-1000
Muy bueno	No	Rural	Tercer Nivel	Adultez inicial	Superior	No	Masculino	0-1000
Muy bueno	No	Ninguno	Secundaria	Adultez inicial	Superior	Si	Masculino	0-1000
Muy bueno	No	Urbano	Cuarto Nivel	Adultez inicial	Normal	No	Femenino	1000-2000
Muy bueno	No	Urbano	Tercer Nivel Tecnológico	Adultez inicial	Normal	No	Masculino	0-1000
Muy bueno	No	Urbano	Secundaria	Adultez inicial	Normal	Si	Masculino	0-1000
Muy bueno	No	Urbano	Secundaria	Adultez inicial	Normal	Si	Femenino	0-1000
Muy bueno	No	Urbano	Tercer Nivel	Adultez inicial	Normal	No	Masculino	0-1000
Muy bueno	No	Urbano	Tercer Nivel	Adultez inicial	Superior	No	Femenino	0-1000
Muy bueno	No	Rural	Secundaria	Adultez inicial	Superior	Si	Femenino	0-1000
Muy bueno	No	Urbano	Secundaria	Adultez inicial	Normal	No	Masculino	0-1000
Muy bueno	No	Urbano	Tercer Nivel	Adolescencia	Normal	Si	Femenino	0-1000
Muy bueno	No	Urbano	Tercer Nivel	Adultez inicial	Normal	Si	Femenino	0-1000
Muy bueno	No	Urbano	Cuarto Nivel	Adultez inicial	Normal	No	Femenino	1000-2000
Muy bueno	No	Urbano	Tercer Nivel	Adultez inicial	Normal	Si	Masculino	1000-2000
Muy bueno	No	Ninguno	Tercer Nivel	Adultez inicial	Normal	Si	Femenino	0-1000

Figura 20. Nuevo conjunto de datos.

4.4 Fase de Conseguir los Modelos

Esta fase comprende una tarea, la cual es la de conseguir el o los modelos de minería de datos; además, se describen en este apartado cada una de las tareas, productos de trabajos y guías que fueron utilizadas para obtener los entregables de esta fase, dichos elementos pueden ser visualizados en la Figura 21.

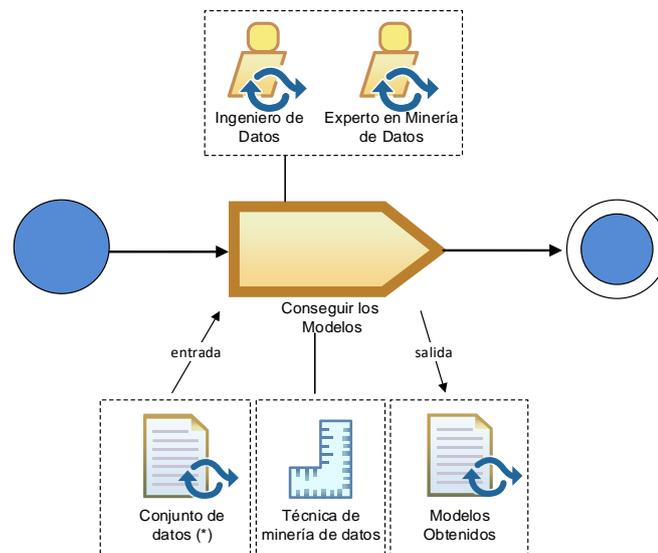


Figura 21. Fase de obtención del modelo de minería de datos del juego serio.

4.4.1 Conseguir los modelos

A partir del conjunto de datos de la actividad anterior, se debe obtener un modelo de minería de datos que permita generar inferencias sobre el dataset; además, se debe determinar una técnica de minería de datos dependiendo del tipo de tarea, siendo en este caso, establecer el comportamiento de los usuarios en función del puntaje obtenido durante su interacción con el juego, para que de esta manera se permita responder a uno de los objetivos planteados en este trabajo de titulación, el cual es “*Determinar los factores relevantes que inciden en el desempeño del jugador frente al juego serio*”.

Técnica de minería de datos

La técnica de minería de datos elegida fue un árbol de decisión, esta técnica permite explicar el comportamiento de los usuarios en función del puntaje obtenido durante su interacción con el juego, es decir, en base a su puntaje se obtienen un conjunto de ramas que son las variables mostradas en la Tabla 5, permitiendo explicar el comportamiento de los usuarios y realizar inferencias en función al número de muestras obtenidas, para su posterior evaluación con los expertos.

Modelo Obtenido

En la Figura 22 y Figura 23, se muestra el proceso y el árbol de decisión obtenido en RapidMiner, el mismo que permite generar cierta cantidad de inferencias en función de las muestras obtenidas del conjunto de datos.

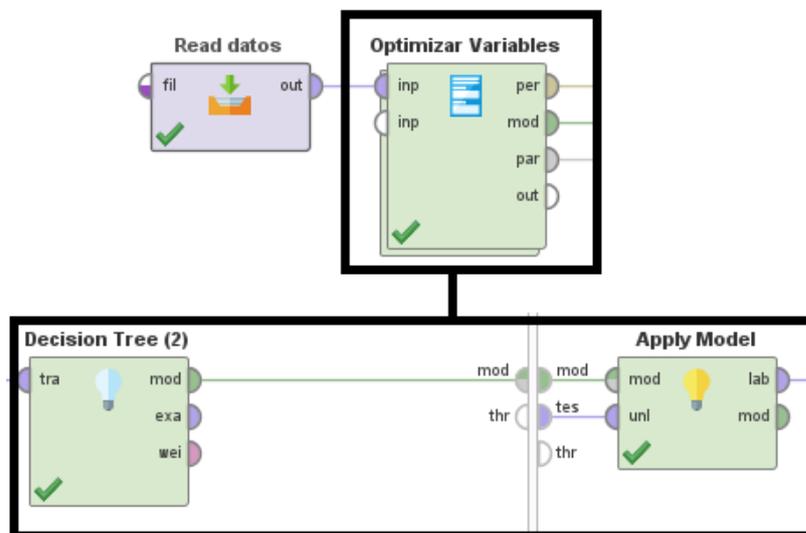


Figura 22. Conectores para generar el árbol de decisión.

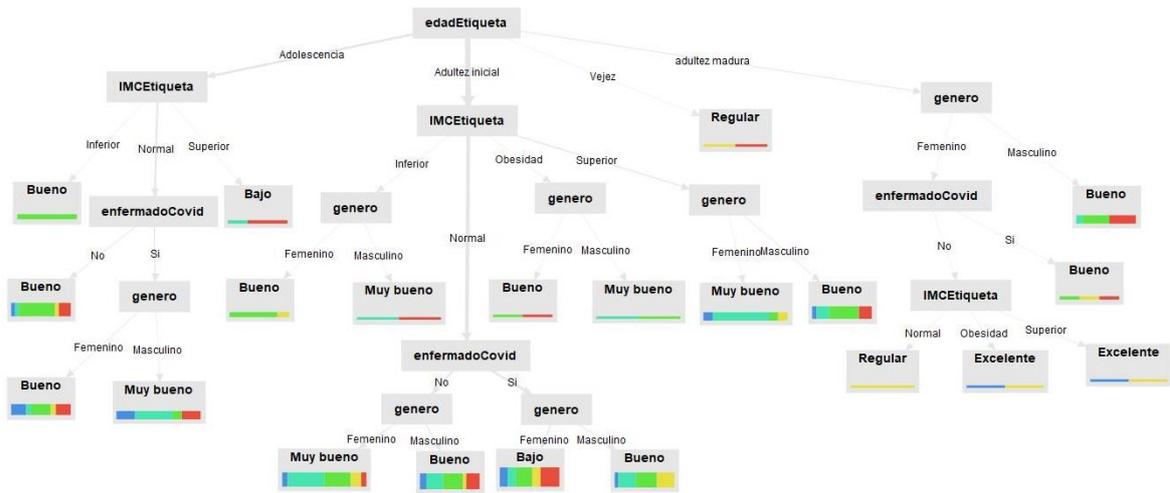


Figura 23. Árbol de decisión generado.

4.5 Fase de evaluación

Esta fase comprende la tarea final de este modelo que es la evaluación del modelo obtenido de minería de datos; además, se describen en este apartado cada una de las tareas y productos de trabajo que fueron utilizadas para obtener los entregables de esta fase, dichos elementos pueden ser visualizados en la Figura 24.

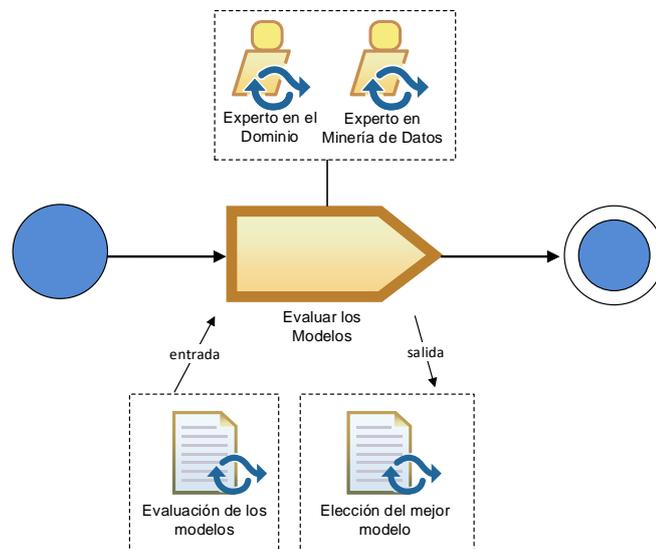


Figura 24. Fase de evaluación de los modelos obtenidos.

4.5.1 Evaluar los modelos

Finalmente, en esta tarea se busca determinar el mejor modelo de minería de datos, por lo tanto, al ser un árbol de decisión, se busca que la precisión del modelo sea buena, para obtener una clasificación lo suficientemente robusta y que permita explicar el comportamiento de los datos.

Evaluación de los modelos

Para evaluar de manera idónea los modelos de minería obtenidos, se determinó que la precisión obtenida en los modelos debe ser lo suficientemente confiable, es decir, el modelo con mejor precisión será el que describa de manera adecuada el comportamiento de los datos.

Elección del mejor modelo

El mejor modelo generado por el árbol de decisión es el propuesto en la Figura 23, del cual se obtuvieron las siguientes inferencias que son descritas en base a la Tabla 4 y son listadas a continuación en la Tabla 6:

Tabla 6. Inferencias obtenidas.

<i>Número</i>	<i>Inferencia</i>
1	Los participantes de sexo masculino en la edad de adultez madura (41 a 60 años) obtuvieron una puntuación baja (menor a 1599 puntos) en el juego de memoria.
2	Los participantes de sexo femenino en la edad de adultez madura (41 a 60 años) que se enfermaron de covid, obtuvieron una puntuación baja (menor a 1599 puntos) en el juego de memoria.
3	Los participantes en la edad de vejez (mayor a 61 años), obtuvieron una puntuación baja en el juego de memoria.
4	Los participantes en la edad de la adultez inicial (20 a 40 años), con un índice corporal superior al normal (25 a 29.9 IMC) de género masculino y femenino, obtuvieron una puntuación muy buena (1899 a 1800 puntos) en el juego de memoria.
5	Los participantes de sexo femenino en la edad de la adultez inicial (20 a 40 años), con un índice de masa corporal normal (18.5 a 24.9 IMC), que se enfermaron de covid, obtuvieron un puntaje bajo (menor a 1599 puntos) en el juego de memoria.
6	Los participantes de sexo masculino en la edad de la adultez inicial (20 a 40 años), con un índice de masa corporal normal (18.5 a 24.9 IMC), que se enfermaron de covid, obtuvieron un puntaje muy bueno (1899 a 1800 puntos) en el juego de memoria.
7	Los participantes de sexo femenino y masculino en la edad de la adultez inicial (20 a 40 años), con un índice de masa corporal normal (18.5 a 24.9 IMC), que NO se enfermaron de covid, obtuvieron un puntaje muy bueno (1899 a 1800 puntos) en el juego de memoria.
8	Los participantes de sexo femenino y masculino en la edad de la adolescencia (12 a 19 años), con un índice de masa corporal normal (18.5 a 24.9 IMC), que se enfermaron de covid, obtuvieron un puntaje excelente (2000 a 1900 puntos) en el juego de memoria.

Los participantes en la edad de la adolescencia (12 a 19 años), con un índice de masa corporal inferior al normal (menor a 18.5 IMC), obtuvieron un puntaje bueno (1899 a 1800 puntos) en el juego de memoria.

5. Evaluación

En este capítulo se detalla la evaluación del modelo de minería de datos generado con respecto a la precisión de los resultados obtenidos. Por otra parte, se medirá la eficiencia y efectividad cuando el modelo es utilizado por psicólogos clínicos o profesionales en formación en su último año. Para ello, se aplicará un caso de estudio, el mismo que se define como “Una investigación empírica que investiga un fenómeno contemporáneo en su contexto real, donde los límites entre el fenómeno y el contexto no se muestran de forma precisa, y en el que múltiples fuentes de evidencia son utilizadas” (Runeson & Höst, 2009). Éste ha sido aplicado siguiendo las guías del autor, con el fin de obtener resultados confiables, evitando sesgos y problemas de validez.

5.1 Diseñar el caso de estudio

Para definir el alcance del caso de estudio, se ha utilizado la plantilla propuesta por Basili, Caldiera, & Rombach (1994) llamado Goal-Question-Metric (GQM), misma que se puede observar en la Tabla 7.

Tabla 7. Plantilla para la definición del Goal-Question-Metric.

Analizar	<Objeto(s) de Estudio> - ¿Qué es lo que se analiza?
Con el propósito de	<Propósito> - ¿Qué intención tiene el estudio?
Con respecto a	<Enfoque de Calidad> - ¿Cuál es el efecto estudiado?
Desde el punto de vista de	<Perspectiva> - ¿Quién se ve afectado?
En el contexto de	<Contexto> - ¿Dónde tiene lugar el estudio, sobre qué artefactos y con qué tipo de participantes?

Para el caso de estudio orientado a evaluar el modelo de minería de datos aplicado a un juego serio, se ha definido el esquema de Basili et al. (1994) de la siguiente manera:

Tabla 8. Definición del Goal-Question-Metric.

Evaluar	El modelo de minería de datos aplicado a un juego serio
Con el propósito de	De evaluar la precisión de los resultados del modelo
Desde el punto de vista de	Psicólogo Clínico/Investigador.

5.5.1 Definición de objetivos, hipótesis y variables

La evaluación a realizar está enfocada en evaluar la precisión de los resultados del modelo obtenido. En este contexto, se realizará las evaluaciones del modelo orientado a evaluar un conjunto de inferencias, dentro del cual surgen los siguientes objetivos: i) Puntuar las variables de interés relacionadas a un juego serio orientado al refuerzo de la memoria, ii) Interactuar con un juego serio, iii) Ordenar un conjunto de inferencias relacionadas al juego y iv) Puntuar las inferencias obtenidas del modelo de minería de datos.

Una vez definido los objetivos, es necesario tomar en consideración lo siguiente:

- Medir la precisión de los resultados obtenidos del modelo y su calidad en cuanto a los resultados obtenidos.
- La eficiencia actual, la cual es el esfuerzo requerido por parte del usuario para entender y evaluar el modelo. La medida común para cuantificar este parámetro es el tiempo
- La efectividad, la cual es la calidad del modelo obtenido, en varios de los casos esta medida puede ser determinada por la capacidad del evaluador al realizar cierta tarea, teniendo en cuenta si fue realizada con éxito o fracaso.

Finalmente, la posibilidad de los resultados obtenidos del modelo de minería de datos sea aceptado en la práctica para la toma de decisiones, se propone la siguiente hipótesis:

- H_{10} : Los resultados del modelo obtenido se perciben como precisos, $H_{10}=\neg H_{11}$
- H_{20} : No existe una diferencia significativa en la efectividad de los sujetos al realizar tareas de puntuación de inferencias, cuando usan un modelo de minería de datos, $H_{21}=\neg H_{20}$
- H_{30} : No existe una diferencia significativa en la eficiencia de los sujetos al realizar tareas de puntuación de inferencias, cuando usan un modelo de minería de datos, $H_{31}=\neg H_{30}$

En la Tabla 9 se muestran las preguntas que fueron utilizadas, para medir la variable de precisión de los resultados del modelo.

Tabla 9. Cuestionario para medir la precisión de los resultados del modelo.

Pregunta	Declaración positiva cinco puntos
PR1	Los resultados del modelo son cercanos a la realidad
PR2	Los resultados del modelo son de fácil interpretación.

PR3	Los resultados del modelo pueden ayudar en la toma de decisiones en tiempo real a un profesional.
PR4	Los modelos de minería de datos pueden describir adecuadamente el comportamiento de los datos.
PR5	Los resultados del modelo permiten generar inferencias cercanas a la realidad.
PR6	Las inferencias obtenidas del modelo son de fácil entendimiento e interpretación.
PR7	Los resultados obtenidos del modelo son precisos.
PR8	Los árboles de decisiones generan inferencias cercanas a la realidad.
PR9	La precisión de los resultados es fiable y confiable.
PR10	Los resultados obtenidos del modelo son confiables.

Este cuestionario fue aplicado después de que el sujeto haya terminado de ejecutar el cuasi experimento, sobre la precisión de los resultados obtenidos del modelo de minería de datos. La escala para las preguntas fue una escala de Likert de 5 puntos, con preguntas opuestas.

En la Tabla 10 se muestra una pregunta cerrada (PC), que se realizó a más de las preguntas planteadas anteriormente, la misma que brinda información relacionada, con respecto al uso del modelo en la práctica profesional, para la toma de decisiones.

Tabla 10. Pregunta cerrada (Sí/No).

Código	Pregunta
PC1	Considera que los resultados del modelo obtenido son útiles para la toma de decisiones en un juego serio orientado al refuerzo de la memoria.

5.2 Ejecución y análisis del caso de estudio

5.2.1 Selección de los sujetos

Para la ejecución de este caso de estudio, se contó con la participación de 4 psicólogos correspondientes al área social. Mismos que fueron seleccionados, de acuerdo a las habilidades y conocimientos relacionados para evaluar el modelo de minería de datos.

5.2.2 Diseño y ejecución del caso de estudio

Para obtener los datos, se realizó una página web, la cual contiene los documentos y guías necesarias para la ejecución del caso de estudio. Posteriormente se realizó la invitación formal a cada uno de los participantes mediante correo electrónico, en donde se indicaba las pautas necesarias para realizar el mismo. De esta manera, se recolectaron los datos del tiempo que les tomó a los participantes ejecutar las tareas propuestas. Posteriormente se les proporcionó un cuestionario que puede ser visualizado en el Anexo 1; este cuestionario consta de diez preguntas cerradas, que fueron necesarias para analizar la variable subjetiva y una pregunta cerrada, para determinar el uso de los participantes con respecto al modelo.

La página web consta de tres páginas, las mismas que son explicadas a continuación:

- **Página de inicio:** En esta página, se muestra el objetivo del modelo generado, el mismo que también se encuentra explicado en el documento de la plantilla del experimento.



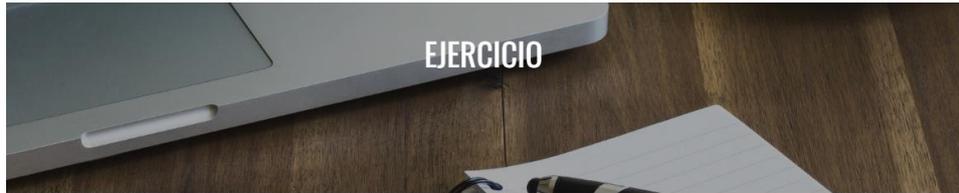
Explicación del Método

El presente modelo describe el comportamiento de 172 participantes que hicieron uso de un juego serio orientado a la memoria, se aplicó la metodología de minería de datos Crisp-DM, la misma que permite generar conocimiento sobre el conjunto de datos obtenido. Finalmente, el resultado de este proceso fue un conjunto de inferencias, que describen el comportamiento de los usuarios partícipes del juego.



Figura 25. Página de inicio.

- **Página del caso de estudio:** En esta página, se explican los documentos necesarios, ya sea la plantilla y anexos necesarios para realizar la ejecución del experimento. Mismos que pueden ser descargados al pulsar el botón “Descargar”.



Ejecución del Ejercicio

En esta sección encontrará el ejercicio asociado a la experimentación con el nombre "**Plantilla del Ejercicio**", del cual se debe realizar las indicaciones propuestas en ese documento; además, encontrará los anexos necesarios para la ejecución del mismo.



Figura 26. Página del caso de estudio.

- **Página de la encuesta:** En esta página, se muestra la encuesta que debe ser respondida por cada uno de los participantes que participaron en el experimento. La matriz de respuestas servirá para medir la precisión de los resultados.

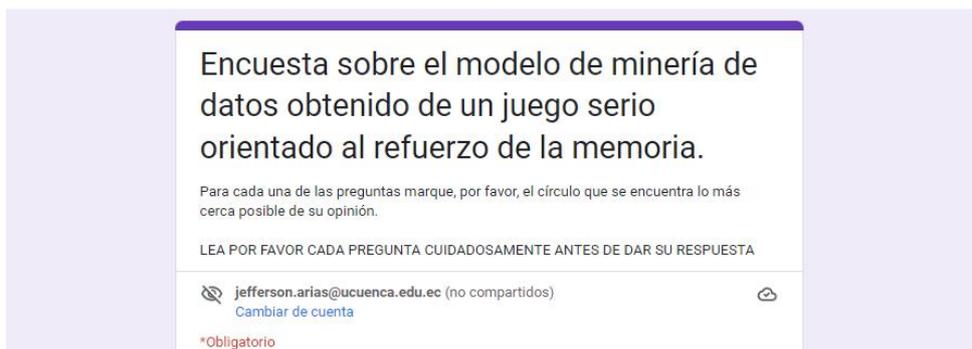


Figura 27. Página de la encuesta.

5.2.3 Análisis de los resultados

Una vez recolectados los datos del ejercicio aplicado, se obtuvieron los siguientes resultados descritos en la Tabla 11.

Tabla 11. Percepciones del usuario con respecto a las tareas del ejercicio aplicado.

Sujeto	Tarea 1	Tarea 2	Tarea 3	Tarea 4	Conclusión
1	Éxito	Éxito	Éxito	Éxito	El usuario concluye de manera satisfactoria el ejercicio propuesto.
2	Éxito	Éxito	Éxito	Éxito con explicación oportuna	El usuario tuvo dificultades al realizar la tarea final, pero con ayuda oportuna concluye el ejercicio exitosamente.
3	Éxito con explicación oportuna	Éxito	Éxito	Éxito	El usuario tuvo dificultades al realizar la tarea final, pero con ayuda oportuna concluye el ejercicio exitosamente.
4	Éxito	Éxito	Éxito	Éxito	El usuario concluye de manera satisfactoria el ejercicio propuesto.

En base a la tabla anterior, se percibe que la Tarea 1, que consistía en obtener varias relevantes de incluir en un juego serio, y, la Tarea 4, la cual se debía puntuar un conjunto de inferencias generadas en base al modelo de minería de datos, necesitaron una explicación más a fondo con respecto a los sujetos 2 y 3. Con respecto a los sujetos 1 y 4, no existió mayor inconveniente al momento de realizar el ejercicio propuesto.

Finalmente, se aplicó la encuesta para medir la percepción de los sujetos, con respecto a la precisión de los resultados del modelo de minería de datos. Las respuestas de los participantes, pueden ser visualizadas en el Anexo 3; además, para aceptar o rechazar las hipótesis propuestas, se triangularon los resultados obtenidos del ejercicio y la encuesta, los mismo que son mostrados en la Tabla 12.

Tabla 12. Resultados de las hipótesis planteadas.

Hipótesis	Acción	Resultado
H ₁₀	Aceptar	Los resultados del modelo obtenido se perciben como precisos
H ₂₀	Rechazar	Existe una diferencia significativa en la efectividad de los sujetos al realizar tareas de puntuación de inferencias, cuando usan un modelo de minería de datos
H ₃₀	Aceptar	No existe una diferencia significativa en la eficiencia de los sujetos al realizar tareas de puntuación de inferencias, cuando usan un modelo de minería de datos.

5.3 Amenazas a la validez

A continuación, se detallan los problemas que pueden poner en peligro la validez del experimento aplicado, Cook y Campell (1979) consideran cuatro tipos de validez: de conclusión, interna, constructo y externa.

5.3.1 Validez de conclusión

Con respecto a la validez de conclusión, se escogieron los test estadísticos adecuados según la naturaleza de los datos. Además, se debe tener en cuenta que el experimento fue realizado por cuatro participantes; sin embargo, los resultados obtenidos son satisfactorios, ya que, los sujetos lograron realizar de manera adecuada, cada una de las tareas planteadas.

5.3.2 Validez interna

Para mitigar la validez interna, se seleccionaron sujetos que posean cierta experiencia con respecto a juegos serios y salud mental; además, se realizó el acompañamiento respectivo durante la ejecución del cuasi experimento, de tal manera que, en el caso que existieran dudas al momento de realizar el ejercicio, puedan ser solventadas de manera inmediata.

5.3.3 Validez de constructo

La validez de constructo es la confiabilidad del cuestionario aplicado. Para aumentar esta medida, se realizó la prueba de alfa de Cronbach con respecto a la variable precisión de los resultados (PR). Siendo el umbral mínimo aceptado $\alpha = 0.70$, entonces $PR = 0.86$. Dando validez al estudio realizado.

5.3.4 Validez externa

Al no tener profesionales de la psicología clínica a disposición, la generalización de los resultados, se puede ver afectada, por utilizar estudiantes que se encuentran finalizando la carrera, como participantes dentro del experimento. Por ello, se consideran que los resultados obtenidos son válidos en un ámbito de psicólogos clínicos principiantes, siguiendo las recomendaciones propuestas por Kitchenham et al. (2002).

6. Conclusiones

Con respecto al primer objetivo específico planteado: “*Evidenciar el estado actual de la investigación relacionada con la salud mental, juegos serios y minería de datos, a través de la elaboración del estado del arte*”, se obtuvieron las siguientes conclusiones:

- Existen pocos juegos serios orientados al refuerzo de la memoria. Los mismos que

no son explotados por medio de la aplicación de técnicas de minería de datos, siendo un área de potencial explotación, al momento de generar datos para realizar tareas de minería de datos.

- Existen varias soluciones de minería de datos en juegos online, las mismas que describen el comportamiento de sus jugadores, por ejemplo: rasgos psicológicos, personalidad, etc. Ya sean usuarios antiguos, como también nuevos dentro del mundo de los videojuegos.

El segundo objetivo específico fue: *“Adaptar un juego serio, mismo que permita registrar las variables apropiadas, para analizar aspectos relacionados con el desempeño del jugador dentro del marco de la salud mental, para su posterior recolección de datos”*, se obtuvieron las siguientes conclusiones:

- Es necesario tener un experto en el dominio para la obtención de estas variables, debido a que, al ser un trabajo realizado por un profesional ajeno al área de psicología, la obtención de los resultados finales, puede verse afectados al obtener variables que no sean relevantes dentro del área.
- La falta de experiencia en el lenguaje de programación y despliegue en la nube, puede ser un factor en contra, que puede conllevar a que el tiempo de ejecución del trabajo propuesto sea elevado.

El tercer objetivo específico propuesto fue: *“Aplicar la Metodología Crisp-DM, para la obtención de un modelo de minería de datos, que permita determinar los factores relevantes que inciden en el desempeño del jugador frente al juego serio.”*, se obtuvieron las siguientes conclusiones:

- Se deben realizar las validaciones necesarias con respecto al registro de datos, esto puede afectar a los procesos de limpieza y procesamiento de los datos.
- Existen varias herramientas o lenguajes de programación a más de RapidMiner, estas herramientas pueden ser elegidas según los objetivos planteados y experticia.
- La elección de la técnica de minería de datos, debe ser la acorde a los objetivos propuestos para realizar la aplicación de la metodología Crisp-DM.
- La participación del experto en el dominio, en las fases iniciales de la metodología, es de vital importancia, para la obtención de un modelo de minería de datos idóneo.

El último objetivo específico fue: *“Evaluar el modelo obtenido a través de un caso de estudio, con la finalidad de dar validez a la solución.”*, se obtuvieron las siguientes conclusiones:

- Los participantes consideraron que los resultados del modelo fueron precisos y tienen presente su uso para la toma de decisiones en este tipo de contextos.
- Existe una diferencia significativa en la efectividad de los participantes al realizar tareas de puntuación de inferencias, cuando usan un modelo de minería de datos.

- No existe una diferencia significativa en la eficiencia de los participantes al realizar tareas de puntuación de inferencias, cuando usan un modelo de minería de datos.

7. Referencias

- Alonso, M. A., & Prieto, P. (2004). Validación de la versión en español del Test Conductual de Memoria de Rivermead (RBMT) para población mayor de 70 años. *Psicothema*, 16(2), 325–328.
- Azevedo, A., & Santos, M. F. (2008). KDD, semma and CRISP-DM: A parallel overview. *MCCSIS'08 - IADIS Multi Conference on Computer Science and Information Systems; Proceedings of Informatics 2008 and Data Mining 2008, January 2008*, 182–185.
- Basili, V. R., Caldiera, G., & Rombach, H. D. (1994). The goal question metric approach. *Encyclopedia of Software Engineering*, 2, 528–532. <https://doi.org/10.1.1.104.8626>
- Braun, P., Cuzzocrea, A., Keding, T. D., Leung, C. K., Padzor, A. G. M., & Sayson, D. (2017). Game Data Mining: Clustering and Visualization of Online Game Data in Cyber-Physical Worlds. *Procedia Computer Science*, 112, 2259–2268. <https://doi.org/10.1016/j.procs.2017.08.141>
- Campbell, G. (2013). *Domain Expert*. <https://www.spiekerpoint.com/2013/02/26/what-is-a-domain-expert/>
- Cherry, K. (2019). *How Human Memory Works*. VeryWellMind.
- ComputerWeekly. (2021). *Administrador de base de datos DBA*. <https://www.computerweekly.com/es/definicion/Administrador-de-base-de-datos-o-DBA>
- Control, C. for D. (2022). *NCHHSTP Social Determinants of Health*. <https://www.cdc.gov/nchhstp/socialdeterminants>
- Corrigan, S., Zon, G. D. R., Maij, A., McDonald, N., & Mårtensson, L. (2015). An approach to collaborative learning and the serious game development. *Cognition, Technology and Work*, 17(2), 269–278. <https://doi.org/10.1007/s10111-014-0289-8>
- Etienne, C. F. (2018). Salud mental como componente de la salud universal. *Revista Panamericana de Salud Pública*, 42, 1–2. <https://doi.org/10.26633/rpsp.2018.140>
- Fisher, G. G., Chacon, M., & Chaffee, D. S. (2019). Theories of Cognitive Aging and Work. In *Work Across the Lifespan*. Elsevier Inc. <https://doi.org/10.1016/b978-0-12-812756-8.00002-5>
- Galea, M., & Woodward, M. (2005). Mini-Mental State Examination (MMSE): Commentary. *Australian Journal of Physiotherapy*, 51(3), 198. [https://doi.org/10.1016/S0004-9514\(05\)70034-9](https://doi.org/10.1016/S0004-9514(05)70034-9)
- Garcia, Y. (2019). *¿Qué es el minado de Datos o Data Mininig?* <https://www.iebschool.com/blog/data-mining-mineria-datos-big-data/>
- GHEORGHE, M., & PETRE, R. (2014). Integrating Data Mining Techniques into Telemedicine Systems. *Informatica Economica*, 18(1/2014), 120–130.

<https://doi.org/10.12948/issn14531305/18.1.2014.11>

- Gorschek, T., Garre, P., Larsson, S., & Wohlin, C. (2006). A model for technology transfer in practice. *IEEE Software*, 23(6), 88–95. <https://doi.org/10.1109/MS.2006.147>
- Halim, Z., Atif, M., Rashid, A., & Edwin, C. A. (2019). Profiling Players Using Real-World Datasets: Clustering the Data and Correlating the Results with the Big-Five Personality Traits. *IEEE Transactions on Affective Computing*, 10(4), 568–584. <https://doi.org/10.1109/TAFFC.2017.2751602>
- Jothi, N., Rashid, N. A., & Husain, W. (2015). Data Mining in Healthcare - A Review. *Procedia Computer Science*, 72, 306–313. <https://doi.org/10.1016/j.procs.2015.12.145>
- Khanapi, M., Ghani, A., Jaber, M. M., & Suryana, N. (2015). *Telemedicine supported by data warehouse architecture*. 10(2), 2006–2008.
- Khemlani, S. S. (2018). Reasoning. In *Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience* (Issue February). <https://doi.org/10.1002/9781119170174.epcn311>
- Kitchenham, B. A., Kitchenham, B. A., Society, I. C., Pfleeger, S. L., Pickard, L. M., Jones, P. W., Hoaglin, D. C., & El Emam, K. (2002). Preliminary Guidelines for Empirical Research in Software Engineering. *Ieee Transactions on Software Engineering*, 28(8), 721–734.
- Klein, R. M., Dilchert, S., Ones, D. S., & Dages, K. D. (2015). Cognitive predictors and age-based adverse impact among business executives. *Journal of Applied Psychology*, 100(5), 1497–1510. <https://doi.org/10.1037/a0038991>
- Lee, P., & Mierswam, I. (2022). *RapidMiner*. <https://rapidminer.com>
- Leuw, J. R. de, Tella, G. di, & Zymelman, M. (1973). Las etapas del desarrollo económico argentino. *Revista Española de La Opinión Pública*, 34, 529. <https://doi.org/10.2307/40199236>
- Loh, C. S., Sheng, Y., & Ifenthaler, D. (2015). Serious games analytics: Methodologies for performance measurement, assessment, and improvement. In *Serious Games Analytics: Methodologies for Performance Measurement, Assessment, and Improvement* (Issue November). <https://doi.org/10.1007/978-3-319-05834-4>
- Lopes, S., Magalhães, P., Pereira, A., Martins, J., Magalhães, C., Chaleta, E., & Rosário, P. (2018). Games used with serious purposes: A systematic review of interventions in patients with cerebral palsy. *Frontiers in Psychology*, 9(SEP). <https://doi.org/10.3389/fpsyg.2018.01712>
- Navarrete Mejía, P. J., Loayza Alarico, M. J., Velasco Guerrero, J. C., Huatuco Collantes, Z. A., & Abregú Meza, R. A. (2016). Índice De Masa Corporal Y Niveles Séricos De Lípidos. *Horizonte Médico (Lima)*, 16(2), 13–18. <https://doi.org/10.24265/horizmed.2016.v16n2.03>
- Orellana, M., Lima, J., & Urigüen, M. A. (2022). Data Mining Applied to a Serious Game. *Communications in Computer and Information Science*, 1, 1–11.

https://doi.org/https://doi.org/10.1007/978-3-030-99170-8_5

Portellano, J. A. (2005). *INTRODUCCION A LA NEUROPSICOLOGIA - PORTELLANO* (S. A. U. McGRAW-HILL/INTERAMERICANA DE ESPAÑA (ed.)). McGRAW-HILL.

Quellmalz, E. S., Timms, M. J., Silberglitt, M. D., & Buckley, B. C. (2012). Science assessments for all: Integrating science simulations into balanced state science assessment systems. *Journal of Research in Science Teaching*, 49(3), 363–393. <https://doi.org/10.1002/tea.21005>

Raphael, D., & Mikkonen, J. (2010). Social Determinants of Health: The Canadian Facts Juha Mikkonen. In *York University School of Health Policy and Management*. <http://www.thecanadianfacts.org/>

Rashid, T. A., Chakraborty, C., & Fraser, K. (2020). Advances in telemedicine for health monitoring: Technologies, design and applications. *Advances in Telemedicine for Health Monitoring, January 2021*, 1–297. <https://doi.org/10.1049/PBHE023E>

Riquelme, J. C., Ruiz, R., & Gilbert, K. (2006). Minería de datos: Conceptos y tendencias. *Inteligencia Artificial*, 10(29), 11–18.

Robles-Bykbaev, Y., Galan-Montesdeoca, J., Segarra-Vanegas, V., Robles-Bykbaev, V., Pesantez-Aviles, F., & Vinanzaca-Padilla, E. (2019). An interactive educational platform based on data mining and serious games to contribute to preservation and learning of the Cañari indigenous cultural heritage in Ecuador. *2018 IEEE Biennial Congress of Argentina, ARGENCON 2018*. <https://doi.org/10.1109/ARGENCON.2018.8646260>

Runeson, P., & Höst, M. (2009). Guidelines for conducting and reporting case study research in software engineering. *Empirical Software Engineering*, 14(2), 131–164. <https://doi.org/10.1007/s10664-008-9102-8>

Santana, C. (2016). What Is Language? *Ergo, an Open Access Journal of Philosophy*, 3(20200313), 1–21. <https://doi.org/10.3998/ergo.12405314.0003.019>

Scapin, C. L. (2010). *Tabla de conversión de puntajes. 2005*, 1–12.

Slimani, A., Elouaai, F., Elaachak, L., Yedri, O. B., & Bouhorma, M. (2018). Learning analytics through serious games: Data mining algorithms for performance measurement and improvement purposes. *International Journal of Emerging Technologies in Learning*, 13(1), 46–64. <https://doi.org/10.3991/ijet.v13i01.7518>

UNIR. (2020). *El ingeniero de datos*. <https://www.unir.net/ingenieria/revista/ingeniero-de-datos/>

Wirth, R., & Hipp, J. (2000). CRISP-DM: towards a standard process model for data mining. Proceedings of the Fourth International Conference on the Practical Application of Knowledge Discovery and Data Mining, 29-39. *Proceedings of the Fourth International Conference on the Practical Application of Knowledge Discovery and Data Mining*, 24959, 29–39. https://www.researchgate.net/publication/239585378_CRISP-

DM_Towards_a_standard_process_model_for_data_mining

Wong Kok Seng, Bin Besar, R., & Abas, F. S. (2006). *Collaborative Support for Medical Data Mining in Telemedicine*. 1894–1899. <https://doi.org/10.1109/ictta.2006.1684678>

8. Anexos

Anexo 1 Formulario de evaluación.

Encuesta sobre el modelo de minería de datos obtenido de un juego serio orientado al refuerzo de la memoria.

Para cada una de las preguntas marque, por favor, el círculo que se encuentra lo más cerca posible de su opinión.

Debe tener en cuenta que todas estas preguntas están planteadas en función del Anexo 2 y el experimento que fue realizado anteriormente.

LEA POR FAVOR CADA PREGUNTA CUIDADOSAMENTE ANTES DE DAR SU RESPUESTA

1. Considero que los resultados del modelo son cercanos a la realidad. *

1 2 3 4 5
Totalmente en Desacuerdo Totalmente en Acuerdo

2. Creo que los resultados del modelo son difíciles de interpretar. *

1 2 3 4 5
Totalmente en Acuerdo Totalmente en Desacuerdo

3. Considero que los resultados del modelo, pueden ayudar en la toma de decisiones en tiempo real a un profesional de la salud. *

1 2 3 4 5
Totalmente en Desacuerdo Totalmente de Acuerdo

4. Pienso que los modelos de minería de datos NO pueden describir adecuadamente el comportamiento de los datos. *

1 2 3 4 5
Totalmente de Acuerdo Totalmente en Desacuerdo

5. Considero que los resultados del modelo permiten generar inferencias cercanas a la realidad. *

1 2 3 4 5
Totalmente en Desacuerdo Totalmente de Acuerdo

6. Creo que las inferencias obtenidas del modelo son difíciles de entender e interpretar. *

	1	2	3	4	5	
Totalmente de Acuerdo	<input type="radio"/>	Totalmente en Desacuerdo				

7. Considero que los resultados obtenidos del modelo son precisos. *

	1	2	3	4	5	
Totalmente en Desacuerdo	<input type="radio"/>	Totalmente de Acuerdo				

8. Creo que los árboles de decisiones NO generan inferencias cercanas a la realidad. *

	1	2	3	4	5	
Totalmente de Acuerdo	<input type="radio"/>	Totalmente en Desacuerdo				

9. Pienso que la precisión de los resultados es fiable y confiable. *

	1	2	3	4	5	
Totalmente en Desacuerdo	<input type="radio"/>	Totalmente de Acuerdo				

10. Creo que los resultados obtenidos del modelo NO son confiables. *

	1	2	3	4	5	
Totalmente de Acuerdo	<input type="radio"/>	Totalmente en Desacuerdo				

Utilizando el Anexo 2 (Tome en cuenta que las hojas finales de la figura son la puntuación que * obtiene el jugador) asociado al experimento, por favor responder:

Considera que los resultados del modelo obtenido son útiles para la toma de decisiones en un juego serio orientado al refuerzo de la memoria.

Sí

No

Anexo 2 Evidencia de evaluación.

Página web:

<https://sites.google.com/ucuenca.edu.ec/experimentomodelomineria/experimento>



Ejecución del Ejercicio

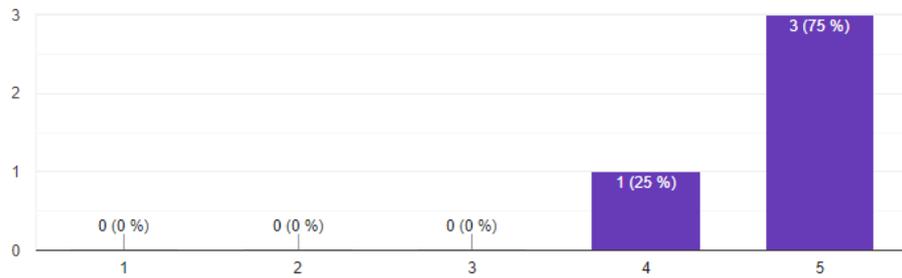
En esta sección encontrará el ejercicio asociado a la experimentación con el nombre "**Plantilla del Experimento**", del cual se debe realizar las indicaciones propuestas en ese documento; además, encontrará los anexos necesarios para la ejecución del mismo.



Anexo 3 Resultados de la encuesta.

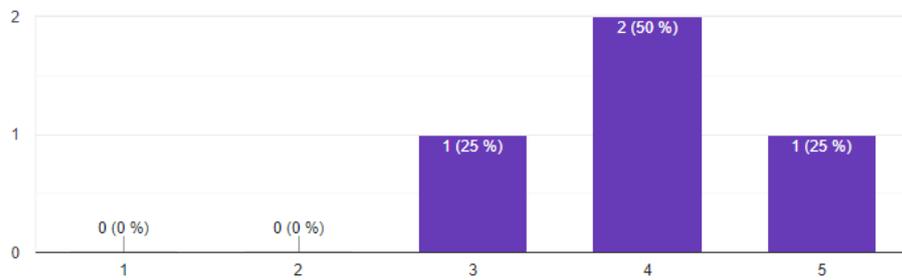
1. Considero que los resultados del modelo son cercanos a la realidad.

4 respuestas



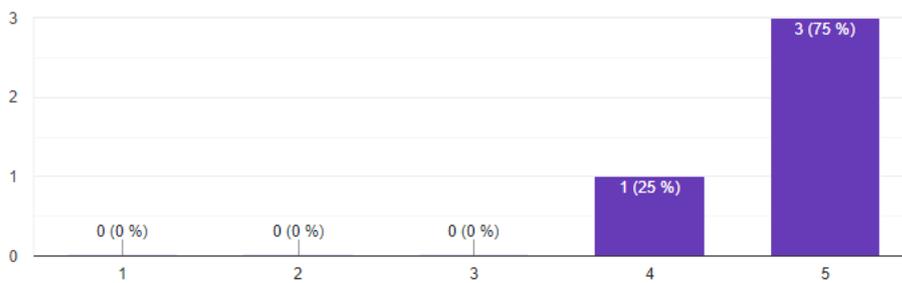
2. Creo que los resultados del modelo son difíciles de interpretar.

4 respuestas



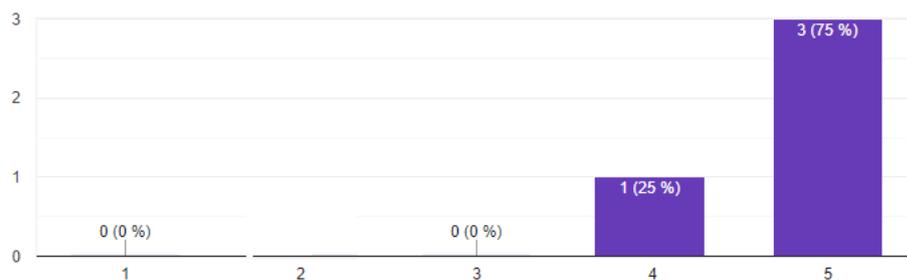
3. Considero que los resultados del modelo, pueden ayudar en la toma de decisiones en tiempo real a un profesional de la salud.

4 respuestas



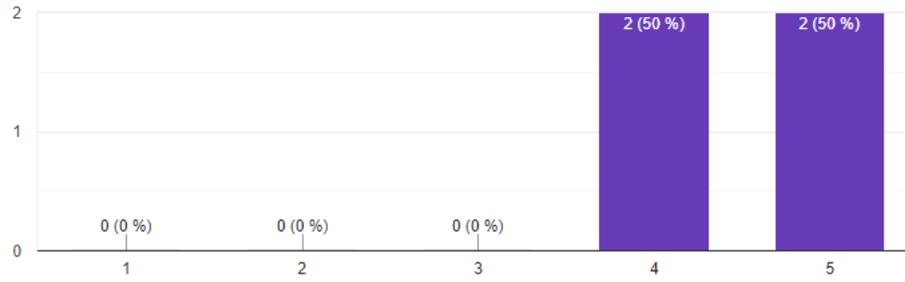
4. Pienso que los modelos de minería de datos NO pueden describir adecuadamente el comportamiento de los datos.

4 respuestas



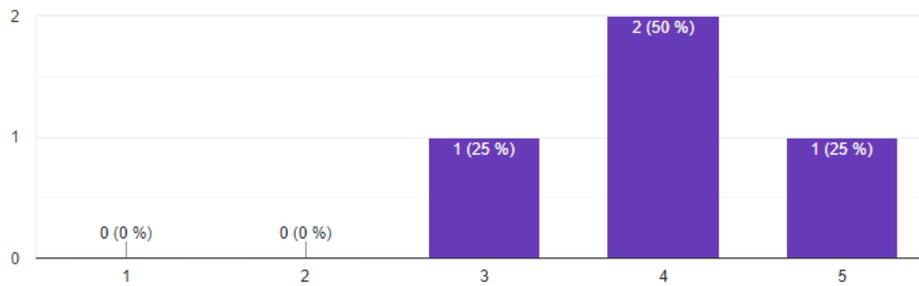
5. Considero que los resultados del modelo permiten generar inferencias cercanas a la realidad.

4 respuestas



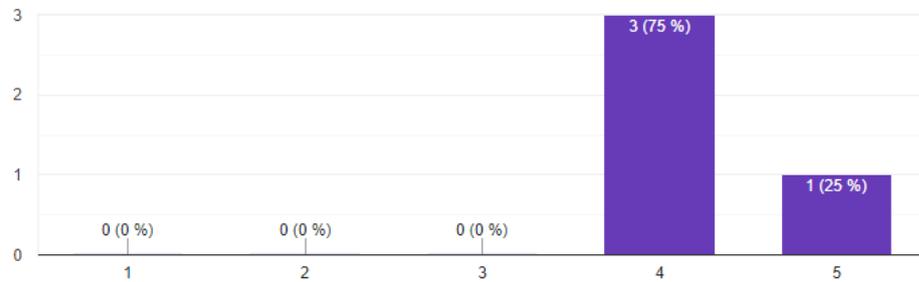
6. Creo que las inferencias obtenidas del modelo son difíciles de entender e interpretar.

4 respuestas



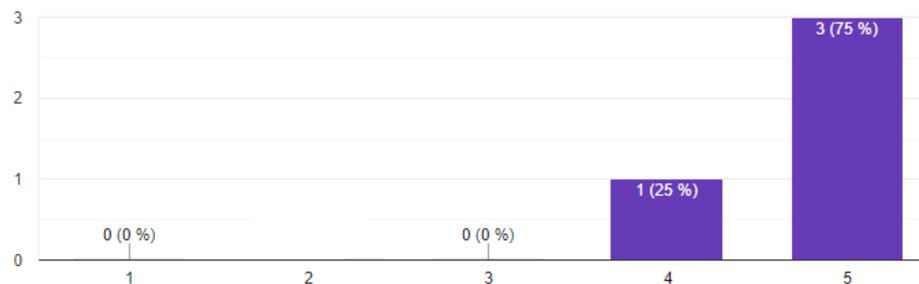
7. Considero que los resultados obtenidos del modelo son precisos.

4 respuestas



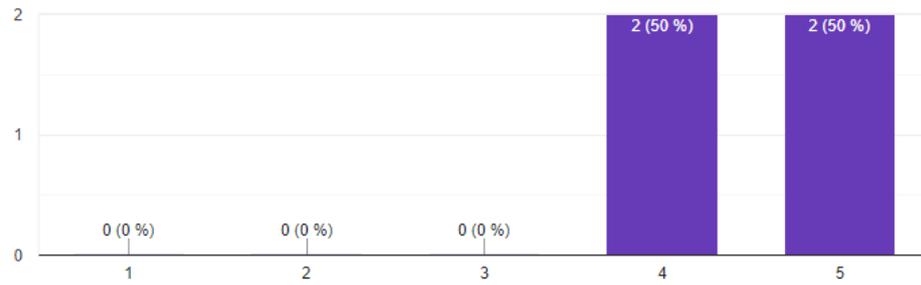
8. Creo que los árboles de decisiones NO generan inferencias cercanas a la realidad.

4 respuestas



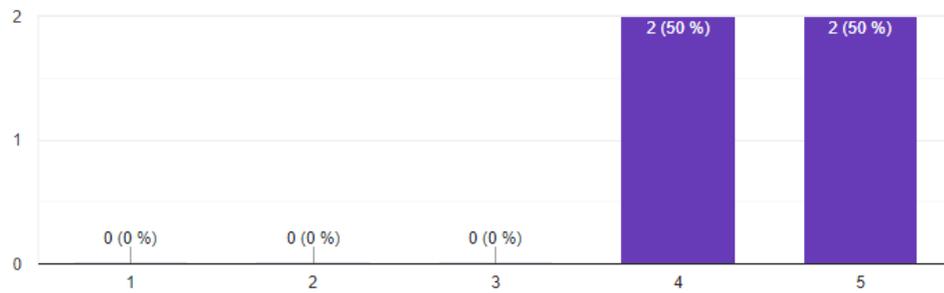
9. Pienso que la precisión de los resultados es fiable y confiable.

4 respuestas



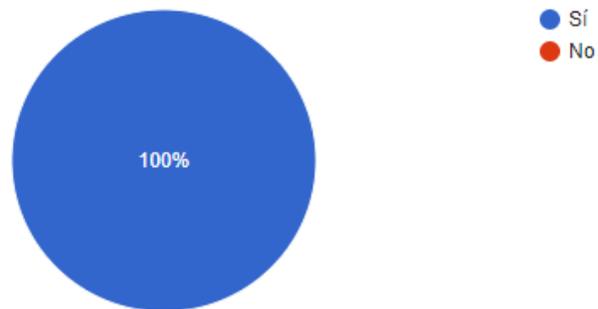
10. Creo que los resultados obtenidos del modelo NO son confiables.

4 respuestas



Utilizando el Anexo 2 (Tome en cuenta que las hojas finales de la figura son la puntuación que obtiene el jugador) asociado al experimento, por favor responder: Considera que los resultados del modelo obtenido son útiles para la toma de decisiones en un juego serio orientado al refuerzo de la memoria.

4 respuestas



Anexo 4 Adaptación del juego serio.

```
import 'package:cloud_firestore/cloud_firestore.dart';

class Persona {
  String id;
  String _nombre;
  String _cedula;
  String _carrera;
  String _telefono;
  String _ciudad;
  String _fechaNacimiento;
  String _sexo;
  String _tipoColegio;
  String _discapacidad;
  String _nivelInstruccion;
  String _poseeCovid;
  String _antecedentesNeuro;
  String _ingresosEconomicos;
  String _peso;
  String _altura;
  String _deporte;
  String _computador;
  int _edad;
  DocumentReference persona;
  bool _tieneDiscapacidad;
}
```

```
Persona.fromMap(Map snapshot, String id, DocumentReference persona)
  : id = id ?? '',
    _nombre = snapshot['nombre'] ?? '',
    _cedula = snapshot['cedula'] ?? '',
    _carrera = snapshot['carrera'] ?? '',
    _telefono = snapshot['telefono'] ?? '',
    _ciudad = snapshot['ciudad'] ?? '',
    _edad = snapshot['edad'] ?? 0,
    _fechaNacimiento = snapshot['fechaNacimiento'] ?? '',
    _sexo = snapshot['genero'] ?? '',
    _tipoColegio = snapshot['tipoColegio'] ?? '',
    _nivelInstruccion = snapshot['nivelInstruccion'] ?? '',
    _poseeCovid = snapshot['enfermadoCovid'] ?? '',
    _antecedentesNeuro = snapshot['antecedentesNeuro'] ?? '',
    _ingresosEconomicos = snapshot['ingresosEconomicos'] ?? '',
    _peso = snapshot['peso'] ?? '',
    _altura = snapshot['altura'] ?? '',
    _deporte = snapshot['deporteHoras'] ?? '',
    _computador = snapshot['computadorHoras'] ?? '',
    _tieneDiscapacidad = snapshot['tieneDiscapacidad'] ?? false,
    _discapacidad = snapshot['discapacidad'] ?? '',
    persona = persona ?? null;
```

```

GlobalKey<FormState> formKey;
File image;
String dropdownValue = 'Auditiva';
String dropdownValueSexo = 'Masculino';
String dropdownValueColegio = 'Ninguno';
String dropdownValueNivelInstruccion = 'Primaria';
String dropdownValueCovid = 'No';
String dropdownValueAntecedentes = 'No';
String dropdownValueIngresos = '0-1000';
String dropdownValueCarrera = 'Ninguno';

set setFormKey(formKeyV) {
  formKey = formKeyV;
}

```

```

class _RegistroPageState extends State<RegistroPage> {
  bool _opciones = false;
  TextEditingController _nombreInputController;
  TextEditingController _pesoInputController;
  TextEditingController _nombreCarreraInputController;
  TextEditingController _ciudadInputController;
  TextEditingController _telefonoInputController;
  TextEditingController _cedulaInputController;
  TextEditingController _deporteInputController;
  TextEditingController _alturaInputController;
  TextEditingController _fechaNacimientoInputController;
  TextEditingController _computadoraInputController;
  DateTime _fechaNacimiento;
}

```

```

),
divider(),
Container(
  child: TextFormField(
    controller: _pesoInputController,
    validator: (value) => value.isEmpty
      ? 'Ingrese su peso aproximado en kg '
      : null,
    textCapitalization: TextCapitalization.sentences,
    style: TextStyle(
      fontSize: 17.0, color: Colors.black),
    keyboardType: TextInputType.number,
    inputFormatters: [
      FilteringTextInputFormatter.allow(
        RegExp(r'[0-9|.]')),
    ],
    decoration: InputDecoration(
      icon: Icon(FontAwesomeIcons.solidIdCard),
      hoverColor: Color.fromRGBO(77, 135, 250, 1),
      labelText:
        'Peso aproximado en KILOGRAMOS (*) (Ejemplo: 50.4)'),
  ),
),
divider(),
Container(
  child: TextFormField(
    controller: _alturaInputController,
    validator: (value) => value.isEmpty
      ? 'Ingrese su altura aproximada en metros'
      : null,
    textCapitalization: TextCapitalization.sentences,
    style: TextStyle(
      fontSize: 17.0, color: Colors.black),
    keyboardType: TextInputType.number,
    inputFormatters: [
      FilteringTextInputFormatter.allow(
        RegExp(r'[0-9|.]')),
    ],
    decoration: InputDecoration(
      icon: Icon(FontAwesomeIcons.solidIdCard),
      hoverColor: Color.fromRGBO(77, 135, 250, 1),
      labelText:
        'Altura aproximada en METROS (*) (Ejemplo: 1.50)'),
  ),
),

```