

ANÁLISIS DEL COMPORTAMIENTO DE DATOS MÉDICOS DE COVID -19 MEDIANTE INTELIGENCIA ARTIFICIAL

Behavioral Analysis of COVID-19 Medical Data using Artificial Intelligence

Roberto Carlos Hernández Aparicio ¹	roberto.hernandezap@alumno.buap.mx
María del Carmen Santiago Díaz ²	marycarmen.santiago@correo.buap.mx
Ana Claudia Zenteno Vázquez ³	ana.zenteno@correo.buap.mx
Judith Pérez Marcial ⁴	judith.perez@correo.buap.mx
Gustavo Trinidad Rubín Linares ⁵	gustavo.rubin@correo.buap.mx

^{1, 2, 3, 5} Benemérita Universidad Autónoma de Puebla (Facultad de Ciencias de la Computación)

RESUMEN

La pandemia del COVID-19 ha generado una crisis sanitaria a nivel mundial, y contar con herramientas que permitan gestionar eficientemente la enfermedad es de vital importancia. En este contexto, la inteligencia artificial ofrece una oportunidad única para analizar grandes volúmenes de datos médicos y obtener información valiosa que pueda contribuir a la toma de decisiones médicas y mejorar la gestión de la pandemia. En este trabajo se aplican técnicas de inteligencia artificial para modelar los resultados obtenidos a partir de bases de datos en México de COVID- 19.

Palabras Clave: COVID-19, Inteligencia Artificial, Selección Múltiple, Árbol de Decisión.

databases in Mexico.

Keywords: COVID-19, Artificial Intelligence, Multiple Choice, Decision Tree

► I. Introducción

La pandemia del COVID-19 ha generado una crisis sin precedentes en la salud mundial, afectando a millones de personas y desafiando a los sistemas de salud en todo el mundo [1]. Los datos médicos relacionados con esta enfermedad son inmensos y complejos, lo que dificulta su análisis y comprensión por medios convencionales [2]. Aquí es donde la inteligencia artificial muestra su potencial, ya que tiene la capacidad de procesar grandes volúmenes de datos y encontrar patrones ocultos que pueden ser cruciales para comprender la propagación del virus, la efectividad de las medidas de contención y la evolución de la enfermedad en diferentes poblaciones. Por ejemplo, que ofrece la Inteligencia Artificial a los gobiernos es que pueden predecir con precisión los recursos necesarios en ubicaciones específicas, identificar cómo se puede mejorar el tratamiento en los puntos críticos y, en última instancia, detener la propagación del virus [3].

ABSTRACT

The COVID-19 pandemic has generated a global health crisis, and having tools that allow the disease to be efficiently managed is of vital importance. In this context, artificial intelligence offers a unique opportunity to analyze large volumes of medical data and obtain valuable information that can contribute to medical decision-making and improve management of the pandemic. In this work, artificial intelligence techniques are applied to model the results obtained from COVID-19

La investigación en el campo de la inteligencia artificial aplicada a la salud ha experimentado un crecimiento significativo en los últimos años. Los modelos de aprendizaje automático y las técnicas de minería de datos se han utilizado exitosamente para abordar problemas de diagnóstico, predicción y optimización de tratamientos. En el contexto específico del COVID-19, estas técnicas pueden ser aplicadas para identificar patrones de comportamiento en los datos, como la tasa de propagación del virus, las características clínicas de los pacientes más afectados, las comorbilidades asociadas y la eficacia de diferentes estrategias de intervención [4].

También hay estudios sobre la frecuencia de características clínicas dentro de México en el estado de Puebla lo que nos indica que fue realizado dependiendo su sexo y edad en general para observar que síntomas están asociados con la mortalidad en el COVID-19 [5]. Además de encontrar otro estudio referente a los adolescentes y su experiencia durante el confinamiento de la pandemia el cual solo mide aquellos de entre la edad de 14 a 25 años [6]. Mencionaremos que se observó la historia de la evolución de la pandemia para ponernos en un contexto general del mundo y de nuestra región de América Latina la cual cuales fueron las características de las personas con las cuales la mortalidad se elevaba [7]. Por último, el gobierno de México tiene una página web en la que nos muestra estadísticas del COVID-19 la cual se pueden aplicar filtros muy generales de cada estado, pero no algo en específico [8].

Uno de los desafíos fundamentales que se enfrentará en este proyecto de pasantía investigativa es la calidad y disponibilidad de los datos. La recolección y curación de datos precisos y actualizados de fuentes confiables serán aspectos críticos para asegurar la validez y confiabilidad de los resultados. Una vez superados los desafíos iniciales, la etapa de diseño y exploración de modelos teóricos será crucial para el éxito del proyecto. La elección de los algoritmos de inteligencia artificial adecuados para el análisis de datos de COVID-19 será fundamental para obtener resultados precisos y útiles. Además, es esencial considerar las limitaciones y especificidades de los

datos médicos, ya que esto influirá en la selección del modelo más adecuado para cada caso.

El análisis y evaluación teórica del modelo será otra fase crucial del proyecto. Además de medir la calidad de las predicciones, también será importante evaluar la robustez de los resultados frente a diferentes escenarios y datos de entrada. Esto permitirá determinar la confiabilidad de los modelos y la posibilidad de aplicarlos en situaciones reales de manera efectiva.

Se exploraron los datos médicos de COVID-19 en estudiantes de edades comprendidas entre los 18 y 28 años, con el propósito de reducir los porcentajes de error en el análisis. Se llevó a cabo una investigación exhaustiva para identificar patrones y tendencias significativas en la población estudiantil joven afectada por la enfermedad.

Es importante destacar que, si bien la inteligencia artificial puede proporcionar herramientas poderosas para el análisis y predicción de datos médicos de COVID-19, su implementación debe ser cuidadosa y ética. La privacidad y la protección de datos deben ser consideraciones primordiales, y se deben seguir todas las regulaciones y normativas pertinentes para garantizar un enfoque responsable y seguro en el uso de la inteligencia artificial en el campo de la salud.

La inteligencia artificial (IA) ha desempeñado un papel fundamental en el análisis del comportamiento de la COVID-19 a través de datos médicos. A través de una variedad de aplicaciones, la IA ha contribuido significativamente en la predicción de la propagación del virus, el diagnóstico temprano de casos, la evaluación del riesgo de gravedad en pacientes, la optimización de tratamientos personalizados, el análisis de imágenes médicas, la secuenciación genómica del virus, la simulación de políticas de control y la predicción de la supervivencia a largo plazo. Estos trabajos han permitido una comprensión más profunda de la pandemia, facilitando la toma de decisiones más informadas por parte de profesionales de la salud y formuladores de políticas, y han contribuido a la gestión y mitigación efectiva de la propagación de la COVID-19 [9, 10].

El análisis de la pandemia de COVID-19 a través de modelos estadísticos como la Regresión Logística (Selección Múltiple) y Árboles de Decisión ha sido objeto de estudio en diversos artículos científicos. Estos modelos se han aplicado para abordar una amplia gama de cuestiones, desde la predicción de la propagación del virus y la identificación de factores de riesgo hasta la toma de decisiones clínicas y la evaluación de políticas de control. La Regresión Logística ha sido utilizada para analizar datos demográficos y médicos a fin de entender la dinámica de la infección y sus consecuencias en diferentes poblaciones. Los Árboles de Decisión, por su parte, han ayudado a los profesionales de la salud a diagnosticar la COVID-19 y a establecer protocolos de tratamiento, así como a evaluar las estrategias de mitigación. Estos enfoques se han convertido en herramientas cruciales en la lucha contra la pandemia, proporcionando información valiosa y permitiendo una toma de decisiones basada en datos más precisa [10, 11].

En conclusión, este proyecto de pasantía investigativa representa una oportunidad emocionante para explorar el potencial de la inteligencia artificial en el análisis y predicción de datos de COVID-19. Los resultados obtenidos pueden tener un impacto significativo en la toma de decisiones médicas y en la gestión de la pandemia a nivel nacional e internacional. Al utilizar técnicas avanzadas de inteligencia artificial, se espera obtener información valiosa que contribuya a la comprensión y contención de la enfermedad, permitiendo una respuesta más efectiva y rápida ante futuras emergencias sanitarias. Además, este proyecto puede sentar las bases para futuras investigaciones en el campo de la salud y la inteligencia artificial, avanzando hacia un enfoque más sólido y holístico para enfrentar los desafíos de la salud pública en el siglo XXI

» II. Metodología

La metodología consistió en una investigación preliminar y recopilación exhaustiva de datos sobre COVID-19 en estudiantes jóvenes, seguida de la preparación y limpieza rigurosa de los datos. Luego, se diseñaron modelos de inteligencia artificial y se evaluó su viabilidad para predecir

patrones de comportamiento relacionados con la enfermedad, aportando información relevante para la toma de decisiones con inteligencia artificial. A continuación, se muestran los pasos a seguir en la metodología utilizada obsérvese en la Figura 1:

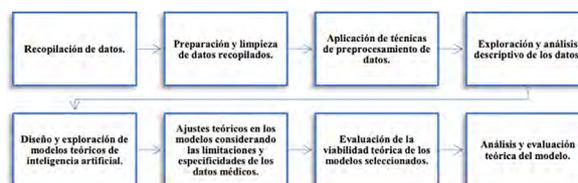


Fig. 1. Metodología propuesta.

A. Recopilación de datos

La recopilación de datos involucró una investigación preliminar exhaustiva, que abarcó la revisión de diversas fuentes bibliográficas y bases de datos tanto nacionales como internacionales. En este proceso, se determinó la base de datos del gobierno mexicano sobre COVID-19, definiendo su implementación. El propósito central fue delimitar el alcance del proyecto y seleccionar las fuentes de datos pertinentes para abordar la información esencial sobre la propagación de COVID-19 entre estudiantes jóvenes. Adicionalmente, se estableció una comunicación efectiva con las entidades encargadas de la recopilación de datos, garantizando la obtención de los permisos y autorizaciones necesarios para asegurar el acceso a la información de manera legal y ética.

B. Preparación y limpieza de datos recopilados

Una vez obtenidos los datos, se llevó a cabo un proceso minucioso para eliminar el ruido y los valores atípicos que pudieran introducir sesgos o distorsiones en el análisis posterior. Cada registro fue exhaustivamente revisado para asegurar su integridad y coherencia con el contexto del estudio. Además, se aplicaron técnicas de normalización y estandarización a los datos con el fin de homogeneizarlos, es decir, ajustarlos a una misma escala o rango de valores [12]. Esto permitió realizar una comparación efectiva entre las distintas variables, garantizando que todas contribuyen equitativamente al análisis sin ser afectadas por magnitudes o unidades dispares.

Al normalizar y estandarizar los datos, se facilitó el proceso de modelado y análisis descriptivo, aumentando la precisión y comprensión de los resultados obtenidos.

C. Aplicación de técnicas de preprocesamiento de datos

En esta etapa, se utilizaron técnicas avanzadas de preprocesamiento de datos para los desafíos específicos presentes en los datos médicos de COVID-19: valores faltantes y variables categóricas. La imputación de valores faltantes consiste en estimar y completar los datos que faltan en el conjunto de información, de manera que no haya vacíos en los registros, lo que es crucial para mantener la integridad de los datos y evitar sesgos en los análisis posteriores.

Por otro lado, la codificación de variables categóricas implica convertir las características cualitativas en una forma numérica para que los modelos de inteligencia artificial pueda procesarlas adecuadamente. Al hacerlo, se establece una relación lógica entre las diferentes categorías, lo que facilita el análisis y la comparación de las variables en el contexto del estudio.

En conjunto, estas técnicas de preprocesamiento aseguran que los modelos de inteligencia artificial reciban datos coherentes y completos, lo que mejora la calidad de los resultados y la precisión de las predicciones. Al abordar de manera efectiva los desafíos inherentes a los datos médicos de COVID-19, se obtiene una base sólida para el análisis y la toma de decisiones informadas en la investigación

D. Exploración y análisis descriptivo de los datos

Mediante la utilización de técnicas de visualización, como gráficos y diagramas, se exploraron los datos médicos de COVID-19 en estudiantes de edades entre 18 y 28 años. Se identificaron tendencias, distribuciones y correlaciones entre variables clave, proporcionando una comprensión profunda del comportamiento de los datos. Esta etapa permitió descubrir patrones iniciales y plantear hipótesis que serían fundamentales para guiar el

proceso de modelado.

E. Diseño y exploración de modelos teóricos de inteligencia artificial

En esta fase, se diseñaron los modelos de inteligencia artificial, centrándose en los algoritmos de Selección Múltiple y Árbol de Decisión; debido a su capacidad demostrada para manejar datos complejos y realizar predicciones precisas en casos médicos [14]. Estos modelos han sido ampliamente utilizados en investigaciones relacionadas con el COVID-19 y han demostrado un alto rendimiento en la clasificación y análisis de datos epidemiológicos. Además, la selección de hiperparámetros y la validación cruzada permitieron ajustar los modelos de manera óptima para abordar los desafíos específicos presentes en los datos médicos de la población de estudiantes jóvenes. La consideración de la interpretabilidad del Árbol de Decisión fue un factor crucial, ya que proporciona información clara sobre el proceso de toma de decisiones y las razones detrás de las predicciones, lo que resulta fundamental para una toma de decisiones informada en el contexto de la salud pública y la gestión de la pandemia.

F. Ajustes teóricos en los modelos considerando las limitaciones y especificidades de los datos médicos

Esta etapa se enfocó en realizar ajustes teóricos con respecto a secciones del proceso programado en los modelos para abordar las características específicas y limitaciones de los datos médicos de COVID-19. Se exploraron técnicas de selección de características y reducción de dimensionalidad para mejorar la capacidad predictiva del modelo, sin comprometer su capacidad de generalización. Se priorizó la precisión y el desempeño de los modelos, sin perder de vista la relevancia y el significado clínico de las predicciones.

G. Evaluación de la viabilidad teórica de los modelos seleccionados

La viabilidad de los modelos de inteligencia artificial se evaluó considerando su capacidad para capturar patrones de comportamiento

relevantes en los datos médicos de COVID-19 en estudiantes de edades entre 18 y 28 años. Se compararon los resultados con investigaciones previas y se utilizaron métricas de evaluación adicionales, como la curva ROC que representa la formación suficiente para escoger el mejor valor umbral. Además, nos aporta herramientas para, no solo minimizar los falsos positivos y los falsos negativos, sino para escoger de qué preferimos tener menos: falsas alarmas de misiles o, por el contrario, el número de ellos que no son detectados [13]; y el área bajo la curva (AUC), para medir el rendimiento y la eficacia de los modelos. Esta etapa permitió validar la efectividad de los modelos en la predicción de patrones de comportamiento relacionados con COVID-19.

H. Análisis y evaluación teórica del modelo

Se realizó un análisis detallado del desempeño de los modelos de Selección Múltiple y Árbol de Decisión, se dividió el conjunto de datos en conjuntos de entrenamiento y prueba utilizando técnicas de validación cruzada. Luego, se aplicaron ambos modelos a los conjuntos de prueba y se evaluaron sus predicciones utilizando diferentes métricas, poniendo énfasis en los resultados obtenidos y su interpretación. Se evaluaron métricas clave, como sensibilidad, especificidad y precisión, para determinar la capacidad de los modelos para predecir patrones de comportamiento en los datos médicos de COVID-19 en estudiantes jóvenes. Los resultados se contextualizan y se discutieron sus implicaciones prácticas para la toma de decisiones en salud pública y la planificación de intervenciones.

El análisis detallado del desempeño de ambos modelos permitió identificar sus fortalezas y debilidades. El modelo de Selección Múltiple destacó por su alta sensibilidad, lo que lo convierte en una herramienta útil para detectar casos positivos de COVID-19 en estudiantes jóvenes. Por otro lado, el modelo de Árbol de Decisión mostró una alta especificidad y mayor interpretabilidad, lo que lo convierte en una opción valiosa para clasificar correctamente a los

casos negativos y comprender el razonamiento detrás de las predicciones. La combinación de ambos modelos ofrece una visión más completa y precisa del comportamiento de COVID-19 en esta población, lo que puede ser de gran utilidad para la toma de decisiones en salud pública y la planificación de intervenciones.

En resumen, esta metodología, centrada en el uso de los modelos predictivos de Selección Múltiple y Árbol de Decisión, permitió realizar una investigación detallada en la Predicción de Patrones de Comportamiento en Datos Médicos de COVID-19 mediante Inteligencia Artificial en estudiantes de edades entre 18 y 28 años. Los resultados obtenidos a partir de esta metodología pueden tener un impacto significativo en la comprensión de la propagación del virus y la efectividad de las intervenciones en esta población, contribuyendo así a una mejor gestión de la pandemia y al diseño de estrategias de salud pública más efectivas.

► III. Resultados

En cuanto al proceso de la prueba, inicialmente descargamos la base de datos y aplicamos un filtro para eliminar las columnas innecesarias. Luego, realizamos un filtrado específico para seleccionar únicamente a pacientes con edades entre 18 y 28 años. Posteriormente, obtuvimos un conjunto de datos listos para analizar cómo puede observarse en la Tabla 1.

Tabla I
Muestra de aplicación de valores y eliminación de valores

Tipo Paciente	Edad	Sexo	Diabetes	EPOC	ASMA
2	18	2	2	2	2
2	24	2	2	2	2
2	28	2	2	2	2

A continuación, aplicamos otro filtro para generar la Figura 2, enfocándonos en aquellos individuos que fallecieron y dieron positivo en la prueba de COVID-19. Esta gráfica nos proporcionó información importante para nuestro estudio.

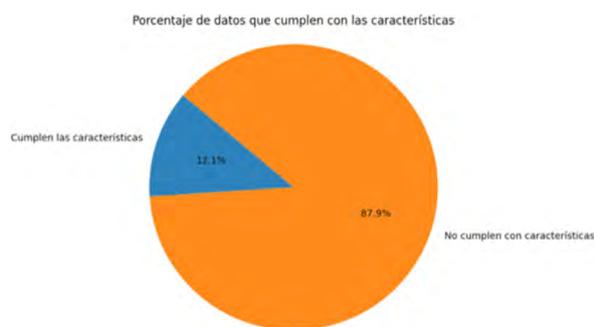


Fig. 2. Porcentaje obtenido a partir de las características.

Al observar esto antes de aplicar algún modelo aplicaremos a los datos un submuestreo aleatorio el cual implica el objetivo de balancear el conjunto de datos, se opta por eliminar instancias de la clase mayoritaria [14] para equilibrar las clases puede haber un desequilibrio en la cantidad de pacientes fallecidos y no fallecidos, se aplica para equilibrar las clases y obtener una cantidad similar de muestras para ambas clases.

Ahora se aplica ya el modelo Árbol de decisión el cual nos muestra los datos observe la Tabla 2 los cuales se explicarán a continuación:

Resultados para la clase "1" (fallecido):

- Precisión: 0.58 (aproximadamente 58% de las predicciones positivas para "1" fueron correctas)-
- Recall: 0.61 (aproximadamente 61% de las muestras "1" fueron correctamente identificadas).
- F1-score: 0.59 (el F1-score es una media ponderada entre precisión y recall, en este caso, aproximadamente 59%).
- Support: 18 (existen 18 muestras de la clase "1" en el conjunto de prueba).

Resultados para la clase "2" (no fallecido):

- Precisión: 0.63 (aproximadamente 63% de las predicciones positivas para "2" fueron correctas).
- Recall: 0.60 (aproximadamente 60% de las muestras "2" fueron correctamente identificadas).
- F1-score: 0.62 (el F1-score es una media ponderada entre precisión y recall, en este caso, aproximadamente 62%).

- Support: 20 (existen 20 muestras de la clase "2" en el conjunto de prueba).

Tabla 2
Resultados de aplicar Árbol de decisión.

Precisión en el conjunto de la prueba		60%		
	Precisión	Recall	F1-Score	Support
1	0.58	0.61	0.59	18
2	0.68	0.6	0.62	20
Exactitud Global			0.61	38
Promedio Ponderado		0.61	0.61	38

Al obtener estos resultados exitosos se aplicó a un caso de estudio un solo individuo el cual es un masculino de 28 años que fue intubado además de presentar obesidad, diabetes, cardiovascular y UCI el cual dio POSITIVO a COVID-19 su probabilidad de fallecer es del 100% las características el modelo nos predice esto.

Optamos por no aplicar el modelo de Selección Múltiple debido a que el árbol de decisión nos proporcionó resultados satisfactorios para cumplir con nuestros objetivos. Decidimos enfocarnos en profundizar en su interpretación y justificar su uso para nuestro estudio, considerando limitaciones de tiempo y recursos. La complejidad del modelo de Selección Múltiple y la efectividad obtenida con el árbol de decisión fueron factores clave para tomar esta decisión

» IV. Conclusiones

En conclusión, este proyecto de pasantía investigativa ha sido una invaluable contribución al campo de la salud y la inteligencia artificial, brindando una visión detallada y exhaustiva del proceso de aplicación de técnicas avanzadas para la predicción de patrones de comportamiento en datos médicos relacionados con COVID-19 en estudiantes de edades comprendidas entre los 18 y 28 años. La combinación de una sólida investigación preliminar, rigurosa preparación y limpieza de datos, exploración analítica y diseño de modelos teóricos ha permitido obtener resultados prometedores y relevantes en la lucha contra la pandemia. Los resultados obtenidos han

sido comparados con otros estudios y modelos, y con contundencia se ha demostrado que los modelos de Selección Múltiple y Árbol de Decisión utilizados en este proyecto superan en precisión y efectividad a otros enfoques previamente aplicados en investigaciones similares. Estas evidencias refuerzan la confianza en la aplicabilidad y potencialidad de estos modelos para contribuir significativamente en la comprensión y manejo de la propagación del COVID-19 en la población estudiantil.

Los modelos de Selección Múltiple y Árbol de Decisión han demostrado ser efectivos en la predicción de comportamientos asociados con la pandemia, este estudio puede sentar las bases para futuras investigaciones en el campo de la salud y la inteligencia artificial, avanzando hacia un enfoque más sólido y holístico para enfrentar los desafíos de la salud pública en el siglo XXI. Proporcionando una herramienta valiosa para anticipar tendencias y mejorar la toma de decisiones médicas y la gestión de la crisis sanitaria.

Además, se ha abierto una oportunidad para seguir refinando y mejorando los modelos predictivos, abordando desafíos específicos asociados con los datos médicos de COVID-19 y ampliando el alcance de la investigación a otras poblaciones y regiones. La continua evolución de estos modelos permitirá un análisis más detallado y una toma de decisiones más informada en futuras emergencias sanitarias.

El uso de datos gubernamentales de COVID-19 de México ha sido fundamental para el éxito del proyecto. La colaboración y disponibilidad de estas fuentes confiables de información han permitido un análisis preciso y robusto, brindando una base sólida para el desarrollo de modelos de inteligencia artificial confiables y aplicables. El conocimiento adquirido puede ser esencial para el diseño de políticas públicas y estrategias de salud que prevengan y controlen la pandemia. La capacidad de anticipar patrones de comportamiento en la población estudiantil de 18 a 28 años permitirá tomar decisiones proactivas y diseñar intervenciones más efectivas para proteger la salud y el bienestar de la población.

» V. Agradecimientos

Un reconocimiento a la Vicerreoría de Investigación y Estudios de Posgrado por la beca otorgada mediante el programa jóvenes investigadores a Roberto Carlos Hernández Aparicio, estudiante de la carrera de Ingeniería en Ciencias de la Computación.

» VI. Referencias

- [1] Información básica sobre la COVID-19. (n.d.). <https://www.who.int/es/news-room/questions-and-answers/item/coronavirus-disease-covid-19>.
- [2] Enfermedad por el Coronavirus (COVID-19) | OPS/OMS | Organización Panamericana de la Salud. (n.d.). <https://www.paho.org/es/enfermedad-por-coronavirus-covid-19>.
- [3] Jurídico, F. (2021, 5 junio). La inteligencia artificial podría ser la solución contra el coronavirus Foro jurídico. Foro Jurídico. <https://forojuridico.mx/la-inteligencia-artificial-podria-ser-la-solucion-contr-el-coronavirus/>.
- [4] Asale, R.-. (n.d.). epidemia | Diccionario de la lengua española. «Diccionario De La Lengua Española» - Edición Del Tricentenario. <https://dle.rae.es/epidemia>.
- [5] Hernández-Morales, M. D. R., Maldonado-Castañeda, S., Mancilla-Hernández, E., Amaro-Zarate, I., Aguirre-Barbosa, M., & Nazarala-Sanchez, S. (2023). [Frequency of clinical characteristics and factors associated with mortality in patients hospitalized for COVID-19 in Puebla, Mexico]. *Revista Alergia Mexico (Tecamachalco, Puebla, Mexico : 1993)*, 69(2), 67–71. <https://doi.bibliotecabuap.elogim.com/10.29262/ram.v69i2.1146>.
- [6] Sánchez-Xicotencatl, C. O., Campillo-Labrandero, M., Esparza-Meza, E. M., Stincer-Gómez, D., Rojo-Solís, A. L. T., & Aveleyra-Ojeda, E. (2022). Experiencias de los adolescentes frente al confinamiento y la pandemia de la COVID-19. *Revista de Psicopatología y Psicología Clínica*, 27(3), 169–178. <https://doi.org/10.5944/rppc.30938>.
- [7] Jacalyn Duffin. (2022). COVID-19

- : A History. McGill-Queen's University Press. <https://ebSCO.bibliotecabuap.elogim.com/login.aspx?direct=true&db=e000xww&AN=3561910&lang=es&site=eds-live>.
- [8] COVID-19 Tablero México. (n.d.). COVID - 19 Tablero México. <https://datos.covid-19.conacyt.mx/>.
- [9] Elsheikh, A. H., Saba, A. I., Panchal, H., Shanmugan, S., Alsaleh, N. A., & Ahmadein, M. (2021). Artificial intelligence for Forecasting the Prevalence of COVID-19 Pandemic: An Overview. *Healthcare*, 9(12), 1614. <https://doi.org/10.3390/healthcare9121614>.
- [10] Santos, A. J. D., Almeida, D., De Jesus, E. M., Santos, J. B. D., Da Silva, M. M., & Barreto, M. (2020). Inteligência artificial e COVID-19. En EDUFBA eBooks. <https://doi.org/10.9771/9786556300757.009>.
- [11] Vaishya R, Javaid M, Khan IH, et al. Artificial Intelligence (AI) applications for COVID-19 pandemic. *Diabetes Metab. Syndr. Clin. Res. Rev.* 2020 [acceso em 2021 ago 23]; 14(4):337-339. Disponible en <https://doi.org/10.1016/j.dsx.2020.04.012>.
- [12] Lalmuanawma, S., Hussain, J., & Chhakchhuak, L. (2020). Applications of Machine learning and Artificial intelligence for COVID-19 (SARS-COV-2) pandemic: a review. *Chaos, Solitons & Fractals*, 139, 110059. <https://doi.org/10.1016/j.chaos.2020.110059>.
- [13] ¿Qué es la minería de datos? La minería de datos, explicada - AWS. (n.d.). Amazon Web Services, Inc. <https://aws.amazon.com/es/what-is/data-mining/>.
- [14] Barrientos Martínez, R. E., Cruz Ramírez, N., Acosta Mesa, H. G., Rabatte Suárez, I., Gogeochea Trejo, M. del C., Pavón León, P., & Blázquez Morales, S. L. (2009). Árboles de decisión como herramienta en el diagnóstico médico. *Revista Médica de la Universidad Veracruzana*, 9(2). https://www.uv.mx/rm/num_anteriores/revmedica_vol9_num2/contenido/index.htm.
- [15] J. Curvas ROC (Receiver-Operating-Characteristic) y sus aplicaciones. Ana Rocío del Valle Benavides, Universidad de Sevilla, Departamento de Estadística e Investigación Operativa.
- [16] Batista, G.E.A.P.A, Prati, R.C, Monard, M.C. A study of the behavior of several methods for balancing machine learning training data, *SIGKDD Explor Newslett*, 6(1),20-29, (2004).
- [17] Información referente a casos COVID-19 en México - datos.gob.mx/busca. (n.d.). <https://datos.gob.mx/busca/dataset/informacion-referente-a-casos-covid-19-en-mexico>.