

# Logarithmic Loss in Machine Learning Models for Liver Cirrhosis Detection.

Darwin Patiño-Pérez, Ph.D<sup>1</sup>, Miguel Molina-Calderón, MSIG<sup>1</sup>,  
Ángel Ochoa-Flores, MSc<sup>1</sup>, José Castro-Carrasco, MSc<sup>2</sup>,

<sup>1</sup>Universidad de Guayaquil, Facultad de Ciencias Matemáticas y Física, Ecuador, [darwin.patinop@ug.edu.ec](mailto:darwin.patinop@ug.edu.ec),  
[miguel.molinac@ug.edu.ec](mailto:miguel.molinac@ug.edu.ec), [angel.ochoaf@ug.edu.ec](mailto:angel.ochoaf@ug.edu.ec),

<sup>2</sup>Universidad de Guayaquil, Facultad de Ciencias Administrativas, Ecuador, [jose.castroca@ug.edu.ec](mailto:jose.castroca@ug.edu.ec),

*Abstract. – Accurately predicting the survival rate of patients with cirrhosis is very important in healthcare. In this study, we compared five classification models using Mayo Clinic data for primary biliary cirrhosis. The log loss score is used to evaluate the accuracy of the model in predicting survival. RandomForest shows the lowest log loss, followed by LogisticRegression and SVM with consistent prediction accuracy. On the other hand, Naïve Bayes and kNN show more accurate results. K-fold cross-validation verifies the stability of the model. Limitations such as dataset dependency and lack of cirrhosis-specific studies were identified, indicating the need for future external validation and development of more accurate models applicable in clinical settings. In conclusion, RandomForest stands out for its high performance, but it is essential to carefully evaluate other models before clinical implementation to predict survival in cirrhotic patients.*

*Keywords: Liver Cirrhosis, Classification Models, Log Loss, Cross Validation, Machine Learning.*

# Pérdida Logarítmica en Modelos de Aprendizaje Automático para la Detección de Cirrosis Hepática. Logarithmic Loss in Machine Learning Models for Liver Cirrhosis Detection.

Darwin Patiño-Pérez, Ph.D<sup>1</sup>, Miguel Molina-Calderón, MSIG<sup>1</sup>,

Ángel Ochoa-Flores, MSc<sup>1</sup>, José Castro-Carrasco, MSc<sup>2</sup>,

<sup>1</sup>Universidad de Guayaquil, Facultad de Ciencias Matemáticas y Física, Ecuador, darwin.patinop@ug.edu.ec, miguel.molinac@ug.edu.ec, angel.ochoaf@ug.edu.ec,

<sup>2</sup>Universidad de Guayaquil, Facultad de Ciencias Administrativas, Ecuador, [jose.castroca@ug.edu.ec](mailto:jose.castroca@ug.edu.ec),

**Resumen.** – Predecir con precisión la tasa de supervivencia de los pacientes con cirrosis es muy importante en la atención médica. En este estudio, comparamos cinco modelos de clasificación utilizando datos de Mayo Clinic para la cirrosis biliar primaria. La puntuación de pérdida logarítmica se utiliza para evaluar la precisión del modelo a la hora de predecir la supervivencia. RandomForest muestra la pérdida de registros más baja, seguida de LogisticRegression y SVM con una precisión de predicción constante. Por otro lado, Naive Bayes y kNN muestran resultados más precisos. La validación cruzada K-fold verifica la estabilidad del modelo. Se identificaron limitaciones como la dependencia del conjunto de datos y la falta de estudios específicos de cirrosis, lo que indica la necesidad de una futura validación externa y el desarrollo de modelos más precisos aplicables en entornos clínicos. En conclusión, RandomForest destaca por su alto rendimiento, pero es fundamental evaluar cuidadosamente otros modelos antes de su implementación clínica para predecir la supervivencia en pacientes cirróticos.

**Palabras Claves:** Cirrosis Hepática, Modelos de Clasificación, Pérdida Logarítmica, Validación Cruzada, Aprendizaje Automático.

**Abstract.** – Accurately predicting the survival rate of patients with cirrhosis is very important in healthcare. In this study, we compared five classification models using Mayo Clinic data for primary biliary cirrhosis. The log loss score is used to evaluate the accuracy of the model in predicting survival. RandomForest shows the lowest log loss, followed by LogisticRegression and SVM with consistent prediction accuracy. On the other hand, Naive Bayes and kNN show more accurate results. K-fold cross-validation verifies the stability of the model. Limitations such as dataset dependency and lack of cirrhosis-specific studies were identified, indicating the need for future external validation and development of more accurate models applicable in clinical settings. In conclusion, RandomForest stands out for its high performance, but it is essential to carefully evaluate other models before clinical implementation to predict survival in cirrhotic patients.

**Keywords:** Liver Cirrhosis, Classification Models, Log Loss, Cross Validation, Machine Learning.

Digital Object Identifier: (only for full papers, inserted by LACCEI).  
ISSN, ISBN: (to be inserted by LACCEI).

## I. INTRODUCCION

Predecir con precisión la supervivencia en pacientes con cirrosis es un desafío continuo en medicina que tiene implicaciones importantes para la atención clínica y las decisiones de tratamiento. La cirrosis, caracterizada por la formación de tejido cicatricial en el hígado, provoca una disminución gradual de la función hepática, afectando diversos aspectos de la salud [1]. Esta disfunción está estrechamente relacionada con el aumento de la mortalidad por enfermedades hepáticas como la hepatitis B y C, que según la Organización Mundial de la Salud (OMS) son responsables de un número importante de muertes anualmente [2].



Fig.1 Cirrosis Hepática

La cirrosis hepática representa un desafío significativo para la salud pública en la ciudad de Guayaquil, Ecuador. Esta enfermedad crónica del hígado, caracterizada por la cicatrización irreversible del tejido hepático, está estrechamente vinculada a factores de riesgo como el consumo excesivo de alcohol [3], la hepatitis viral y la obesidad.

En Guayaquil, la prevalencia de cirrosis hepática está en aumento, lo que plantea importantes preocupaciones para el sistema de salud local y la calidad de vida de la población. Uno de los principales problemas asociados con la cirrosis hepática en Guayaquil es el diagnóstico tardío. Muchas personas no son conscientes de su condición hasta que experimentan síntomas avanzados, como ictericia, ascitis o sangrado gastrointestinal [4].

Esta falta de detección temprana conduce a un manejo inadecuado de la enfermedad y reduce las opciones de tratamiento efectivo. Además, la disponibilidad limitada de servicios médicos especializados en hepatología en la ciudad puede contribuir a retrasos en la atención y resultados subóptimos para los pacientes con cirrosis hepática.

Los síntomas asociados con la cirrosis incluyen una amplia gama de signos físicos y mentales que tienen un profundo impacto en la calidad de vida de los pacientes, que se refiere a cómo los individuos perciben su salud, teniendo en cuenta factores tanto objetivos como subjetivos [5]. Estos síntomas, que van desde ictericia y retención de líquidos hasta un mayor riesgo de cáncer de hígado, reflejan en parte el impacto devastador que tiene la enfermedad en la salud y la esperanza de vida de los afectados [6]. En este contexto, predecir con precisión la tasa de supervivencia de los pacientes con cirrosis es importante porque puede predecir eventos futuros y aumentar la confianza en el diagnóstico y el tratamiento [7].

La prevención y el manejo de la cirrosis hepática en Guayaquil requieren un enfoque integral que aborde tanto los factores de riesgo como las barreras en el acceso a la atención médica. Las iniciativas de educación pública sobre los peligros del consumo excesivo de alcohol, la importancia de la vacunación contra la hepatitis viral y la promoción de estilos de vida saludables son fundamentales para reducir la incidencia de la enfermedad [8]. Además, es crucial mejorar el acceso a servicios de atención médica especializada y promover la detección temprana mediante programas de detección oportuna y exámenes de rutina. En última instancia, abordar el impacto de la cirrosis hepática en la ciudad de Guayaquil requiere una colaboración coordinada entre el gobierno [9] y los profesionales de la salud además de incluir a la comunidad en general [10].

A medida que el volumen de datos relacionados con la salud crece exponencialmente [11], los métodos tradicionales de extracción manual de información se están volviendo no rentables. En respuesta a esta complejidad y a la necesidad de un procesamiento eficiente, se han logrado avances significativos en la implementación de tecnologías inteligentes en la atención sanitaria [12]. Estos desarrollos han llevado al crecimiento de campos de inteligencia artificial como el aprendizaje automático, que resuelven problemas sin programación directa [13]. Esta técnica genera resultados identificando patrones en un conjunto de datos. También muestra el enorme crecimiento y la importancia de la inteligencia artificial en diversos campos como la medicina, la industria y las finanzas, y muestra su eficacia en el análisis y procesamiento de información [14].

Como resultado, los algoritmos de aprendizaje automático han surgido como una nueva forma de procesar grandes cantidades de datos, especialmente datos complejos y completos en muchas situaciones [15]. Desde esta perspectiva, las herramientas de aprendizaje automático, cuyo objetivo es crear sistemas que puedan aprenderse o mejorar su rendimiento basándose en los datos adquiridos, han demostrado ser invaluable [16].

En particular, como se menciona en [17], son algoritmos del aprendizaje automático que identifican clases o categorías dentro de un conjunto de datos predefinido. Para evaluar la eficacia de estos modelos a la hora de predecir la supervivencia en pacientes cirróticos, era importante utilizar datos de evaluación como la pérdida logarítmica multinomial dado que es un problema de clasificación multiclase. Según lo observado en [18], esta métrica, es particularmente adecuada para evaluar predicciones probabilísticas, se convierte en una parte clave de una evaluación rigurosa del rendimiento del modelo. Las métricas son herramientas utilizadas para evaluar el desempeño de los modelos de aprendizaje automático, y su función es comparar e identificar el modelo que mejor se adapta a las necesidades y objetivos específicos del problema en cuestión [19].

En este estudio, se tomaron datos de [20] para entrenar los modelos y realizar un análisis comparativo integral de diferentes modelos de clasificación aplicados al conjunto de datos "Predicción de supervivencia del paciente con cirrosis" del estudio clínico de cirrosis biliar primaria (CBP). Se utilizó la métrica de pérdida logarítmica como herramienta principal para evaluar la precisión de estos modelos en la predicción de la supervivencia en pacientes con cirrosis hepática.

Los modelos específicos evaluados incluyen regresión logística, máquinas de vectores de soporte, Naive Bayes, bosque aleatorio [21] y K-vecinos más cercanos (kNN). El objetivo principal es identificar los modelos más precisos, así como proporcionar información importante sobre la idoneidad de estos modelos para futuras aplicaciones clínicas para mejorar la práctica clínica y la medicina personalizada.

## II. MATERIALES Y METODOS

### A. Dataset

El conjunto de datos "Cirrhosis Patient Survival Prediction", sobre cirrosis biliar primaria (CBP) se lo uso para el entrenamiento de los modelos de predicción, el mismo contiene información demográfica, características clínicas y marcadores bioquímicos de pacientes con cirrosis hepática, así como su estado de supervivencia y la descripción de cada una de las variables que lo conforman están descritas en la Tabla I.

Antes de iniciar el análisis, se llevó a cabo un riguroso proceso de preprocesamiento de datos, fundamental para garantizar la validez de los resultados en la predicción de la supervivencia de pacientes con cirrosis hepática. El procesamiento de datos implica recopilar y modificar información, ya sea en cantidades pequeñas o grandes. Su propósito es maximizar el valor de los datos obtenidos para contribuir al avance de un proyecto y alcanzar sus metas. Este proceso abarcó diversas etapas esenciales. En primer lugar, se realizó una exhaustiva limpieza de los datos, un paso crucial para corregir posibles errores en la información sin procesar. Esto resulta fundamental, ya que imprecisiones en los modelos de *machine learning* podrían influir significativamente en la precisión de nuestras predicciones clínicas [22].

TABLA 1  
CARACTERÍSTICAS Y ETIQUETA

Variables Independientes	Descripción
N_Days	Días de seguimiento de los pacientes.
Drug	Tratamiento farmacológico aplicado.
Age	Edad de los pacientes.
Sex	Género de los pacientes.
Ascites	Presencia de ascitis, acumulación anormal de líquido en la cavidad abdominal.
Hepatomegaly	Agrandamiento anormal del hígado.
Spiders	Presencia de telangiectasias cutáneas, comúnmente observadas en pacientes con enfermedad hepática.
Edema	Presencia de edema, acumulación anormal de líquido en tejidos.
Bilirubin	Niveles de bilirrubina en la sangre.
Cholesterol	Niveles de colesterol.
Albumin	Niveles de albúmina en la sangre.
Copper	Niveles de cobre en el suero sanguíneo.
Alk_Phos	Actividad de fosfatasa alcalina en suero.
SGOT	Niveles de aspartato aminotransferasa en suero.
Tryglicerides	Niveles de triglicéridos en suero.
Platelets:	Recuento de plaquetas en la sangre.
Prothrombin	Tiempo de protrombina, una medida de la coagulación sanguínea.
Stage	Estadio de la enfermedad hepática.
Variable Objetivo	Descripción
Status	Estado de supervivencia de los pacientes.

### B. Normalización y Estandarización

Posteriormente, se aplicaron técnicas de normalización y estandarización a las variables numéricas. Esta fase fue vital para adaptar y ampliar nuestros datos, asegurando una comparación precisa y minimizando el impacto de valores atípicos inusuales que podrían distorsionar la interpretación de los resultados en el contexto médico [22]. Además, se procedió con la codificación adecuada de las variables categóricas, un aspecto esencial debido a que estos rasgos discretos no se interpretan directamente como valores numéricos. Al codificarse de forma apropiada, se transformaron las características en formatos comprensibles para los algoritmos de modelado, como los utilizados en la predicción de la supervivencia en pacientes con cirrosis hepática [23]. Se implementaron métodos como el one-hot encoding, permitiendo a los algoritmos mejorar sus capacidades predictivas al trabajar con estas características categóricas [24].

### C. Marco Teórico

Se seleccionaron varios modelos de clasificación ampliamente reconocidos y utilizados en la literatura para la predicción de supervivencia en pacientes con diversas enfermedades [25]. En el caso de la regresión logística, como se muestra en la Fig. 2, según [26] es útil cuando se busca prever la presencia o ausencia de una característica o resultado basándose en los valores de un conjunto de predictores.

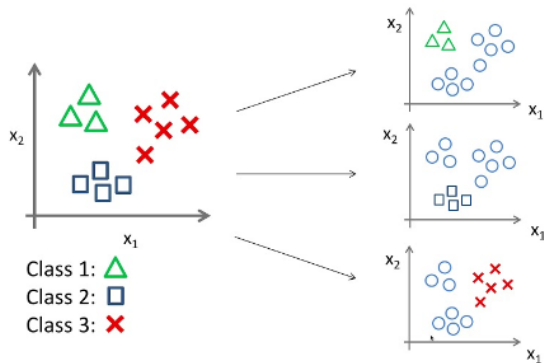


Fig. 2 Regresión Logística

El Support Vector Machine (SVM) según [27], se emplea para clasificar datos en función de dos o más clases etiquetadas, actuando como un clasificador discriminativo definido formalmente a través de un hiperplano óptimo como se muestra en la Fig. 3, donde se aprecia la separación de todas las clases.

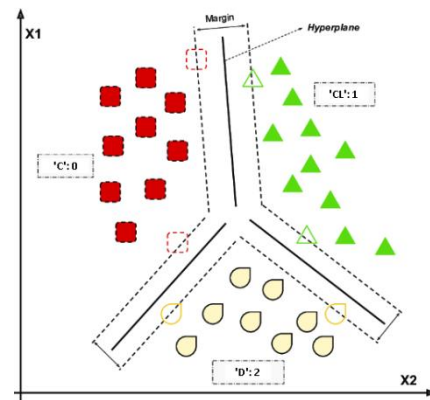


Fig. 3 Support Vector Machine

Mientras que k-Nearest Neighbors (KNN) según [28], se presenta como un clasificador de aprendizaje supervisado no paramétrico que utiliza la proximidad para realizar clasificaciones o predicciones sobre la agrupación de puntos cercanos a un punto de dato individual, como se muestra en la Fig. 4.

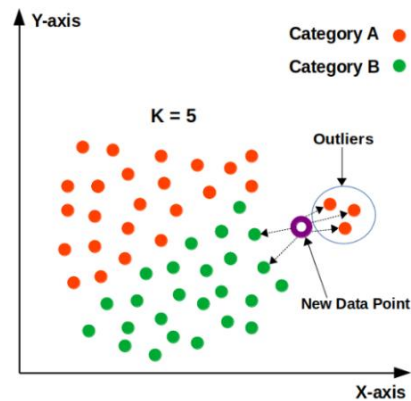


Fig. 4 k-Nearest Neighbors

En cuanto a Random Forest descrito en [29], este se utiliza para resolver problemas de clasificación y regresión, construyendo árboles de decisión a partir de diferentes muestras y tomando su voto mayoritario para decidir la clasificación, así como el promedio en caso de regresión.

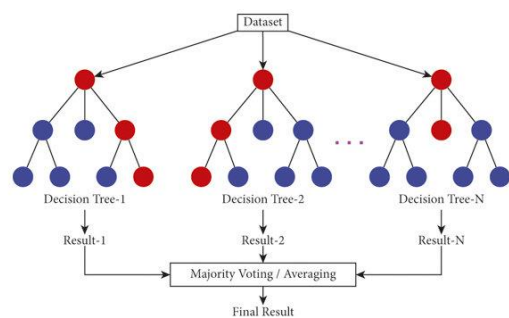


Fig. 5 Random Forest

Finalmente, en relación con el Naive Bayes según [30], se destaca por ofrecer una forma sencilla de construir modelos con un comportamiento bueno, basándose en la simplicidad al asumir que las variables predictoras no dependen unas de otras.

La métrica principal utilizada para evaluar el rendimiento de los modelos fue la pérdida logarítmica (log loss). Aunque interpretar directamente los valores brutos de log-loss es complicado, esta métrica se muestra efectiva para comparar un modelo de aprendizaje automático con otro [31]. Para cada modelo seleccionado, se llevó a cabo un análisis individual, los modelos se entrenaron con los datos de entrenamiento y se calcularon las predicciones probabilísticas sobre el conjunto de prueba para determinar la pérdida logarítmica.

Asimismo, se implementó la validación cruzada k-fold con un número de divisiones igual a 5 para cada modelo. Esta técnica evalúa el rendimiento de un modelo de *machine learning*, con el propósito de identificar de manera ágil el más efectivo. Esto facilita la comprensión y aplicación del modelado predictivo al ser una metodología sencilla de implementar [32].

Todas las tareas de preprocesamiento, modelado y evaluación se llevaron a cabo en el entorno de programación de Google Colab utilizando Python. Se hizo uso de la biblioteca Scikit-learn para acceder a su amplia gama de herramientas y funcionalidades. Scikit-learn ofrece una variedad de algoritmos de aprendizaje automático, técnicas de preprocesamiento de datos, métodos de evaluación de modelos y más, facilitando así la implementación de modelos de clasificación y la evaluación de su rendimiento en este entorno específico.

#### D. Métrica de Evaluación

La **pérdida logarítmica**, también conocida como entropía cruzada binaria o pérdida logística, es una función de pérdida comúnmente utilizada en el entrenamiento de modelos de clasificación en *machine learning*. Puesto que es una métrica que se usa para problemas de clasificación binaria en el contexto de aprendizaje supervisado, se la ha utilizado para medir la discrepancia entre las predicciones de un modelo y las etiquetas reales de los datos. La pérdida logarítmica se puede adaptar y utilizar en problemas de clasificación multiclase. Esta adaptación se conoce como pérdida logarítmica multinomial o pérdida logarítmica categórica. En lugar de trabajar con una sola variable objetivo-binaria (0 o 1), en un problema de clasificación multiclase, se tiene una variable objetivo que puede tomar múltiples valores, correspondientes a las diferentes clases en el problema de clasificación. La fórmula generalizada de la pérdida logarítmica multinomial para múltiples clases es:

$$\text{Pérdida Logarítmica} = -\frac{1}{N} \sum_{i=1}^N \cdot \sum_{j=1}^M y_{ij} \log(p_{ij}) \quad (1)$$

Donde:

N es el número total de instancias en el conjunto de datos.  
M es el número de clases en el problema de clasificación.

En general la pérdida logarítmica penaliza los errores en la predicción de probabilidades de clase en problemas de clasificación, y se utiliza comúnmente como función de pérdida en algoritmos de aprendizaje supervisado como la regresión logística y las redes neuronales cuando se trata de problemas de clasificación. Una pérdida logarítmica más baja indica un mejor rendimiento del modelo.

#### E. Metodología

La metodología usada para la predicción de cirrosis está basada en KDD (*Knowledge Discovery in Databases*) o Descubrimiento de Conocimiento en Bases de Datos que es una metodología en sí misma [33]. Definida como un proceso no trivial, automático y exploratorio que sirve para analizar grandes conjuntos de datos y extraer información útil, novedosa, comprensible y potencialmente útil [34]; para la etapa de minería de datos se evaluarán un grupo de técnicas de *machine learning* de donde finalmente se seleccionarán las técnicas que ofrezcan el mejor de los rendimientos.

### III. RESULTADOS

Se evaluaron cinco modelos de clasificación para predecir la supervivencia en pacientes con cirrosis hepática utilizando el conjunto de datos "Cirrhosis Patient Survival Prediction".

Los modelos base o también conocidos como modelos de referencia actúan como punto inicial para evaluar el rendimiento de modelos más sofisticados, estableciendo un estándar inicial y proporcionando un punto de comparación para medir mejoras y eficiencia. Dichos modelos mostraron diferentes desempeños en la predicción de la supervivencia de los pacientes con cirrosis hepática.

TABLA II  
PÉRDIDA LOGARÍTMICA DE LOS MODELOS DE CLASIFICACIÓN

Modelo	Pérdida Logarítmica Negativa
LogisticRegression	0.632891
SVM	0.689882
kNN	3.148747
RandomForest	0.527007
NaiveBayes	1.682638

La Tabla II, muestra los puntajes de pérdida logarítmica obtenidos por diferentes modelos de clasificación en la predicción de la supervivencia en pacientes con cirrosis hepática.

**LogisticRegression**, con una pérdida logarítmica de 0.632891. Aunque no fue el puntaje más bajo, mostró un desempeño moderado en la predicción de la supervivencia, ubicándose en un rango intermedio en comparación con los otros modelos. **Support Vector Machine (SVM)**, su pérdida alcanzó un puntaje de 0.689882, mostrando un desempeño similar al anterior, su pérdida logarítmica fue un poco más alta, lo que sugiere una capacidad de predicción es aceptable pero ligeramente inferior. **K-Nearest Neighbors (kNN)**, su pérdida alcanzó un puntaje de

3.148747, situándose en el rango más alto de pérdida logarítmica y, por consiguiente, sugiriendo un rendimiento menos preciso en la predicción de la supervivencia en comparación con otros modelos. **Random Forest**, presentó una pérdida logarítmica más baja entre los modelos, con un puntaje de 0.527007. Esta baja pérdida logarítmica sugiere una capacidad más sólida para predecir la supervivencia en pacientes con cirrosis hepática. **Naive Bayes**, presentó un puntaje significativamente más alto de 1.682638, lo que indica que no tiene un óptimo desempeño en la predicción de la supervivencia en comparación con otros modelos. Su pérdida logarítmica considerablemente alta sugiere una menor precisión en las predicciones.

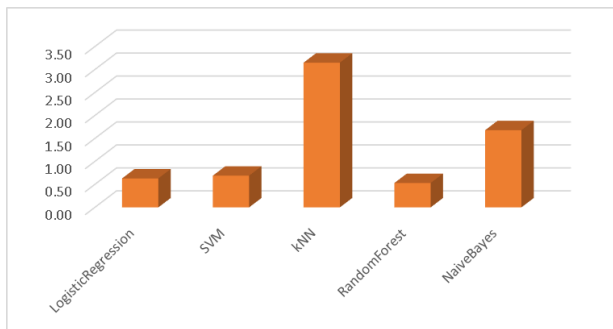


Fig. 6 Pérdida Logarítmica en los Modelos de Clasificación

En la Fig. 6, se ilustra la comparación de la pérdida logarítmica entre varios modelos de clasificación utilizados para predecir la supervivencia en pacientes con cirrosis hepática. Los valores más bajos indican un mejor rendimiento en la predicción de la supervivencia. Además, se realizó una validación cruzada k-fold con cinco divisiones para los modelos Random Forest, SVM y Logistic Regression. Estos modelos fueron seleccionados debido a su destacado desempeño en la evaluación inicial.

Estos resultados reflejan diferencias significativas en la precisión de los modelos para predecir la supervivencia en pacientes con cirrosis hepática, donde Random Forest muestra un desempeño más destacado seguido por Logistic Regression y SVM, mientras que KNN y Naive Bayes presentan puntajes más altos, indicando un rendimiento poco óptimo en la predicción.

TABLA III  
PÉRDIDA LOGARÍTMICA DURANTE VALIDACIÓN CRUZADA

	Random Forest	SVM	Logistic Regression
Prom Neg Log Loss	-0.561	-0.672	-0.63
Min Neg Log Loss	-0.63	-0.704	-0.648
Mas Neg Log Loss	-0.533	-0.653	-0.613

La estrategia de selección permitió una evaluación más profunda de estos modelos en múltiples divisiones del

conjunto de datos, asegurando así una estimación más sólida de su rendimiento general.

Los resultados detallados de la validación cruzada se presentan en la Tabla III, y estos valores reflejan las métricas de pérdida logarítmica promedio, mínima y máxima obtenidas durante la validación cruzada para cada modelo respectivamente, ofreciendo una visión detallada de su rendimiento en múltiples divisiones del conjunto de datos.

Durante la validación cruzada, Random Forest se destacó por mantener una pérdida logarítmica baja y consistente, reafirmando su sólido desempeño en la predicción de la supervivencia en pacientes con cirrosis hepática. Aunque SVM mostró puntajes de pérdida logarítmica más altos en comparación con Random Forest, demostró una consistencia notable en sus predicciones a lo largo de las iteraciones de la validación cruzada. Logistic Regression, por su parte, exhibió resultados consistentes con una pérdida logarítmica similar a la de Random Forest en promedio, y con una variación mínima entre las iteraciones, lo que resalta su estabilidad en la predicción de la supervivencia de estos pacientes.

Es importante tener en cuenta las limitaciones de este estudio, como la dependencia del conjunto de datos utilizado y las características específicas de la cirrosis hepática en el conjunto de pacientes de datos "Cirrhosis Patient Survival Prediction" Mayo Clinic. Además, futuras investigaciones podrían explorar la inclusión de más características clínicas o el uso de técnicas avanzadas de modelado para mejorar aún más la precisión de la predicción de la supervivencia en estos pacientes. A pesar de estas consideraciones, los hallazgos respaldan la idoneidad de Random Forest como modelo preferido debido a su desempeño superior en términos de precisión. No obstante, SVM y Logistic Regression también ofrecen resultados consistentes, lo que sugiere su posible utilidad en aplicaciones clínicas para la predicción de la supervivencia en pacientes con cirrosis hepática.

#### IV. DISCUSION Y CONCLUSIONES

La predicción precisa de la supervivencia en cirrosis hepática es fundamental en la atención clínica. Este análisis implica examinar datos médicos pasados y presentes para mejorar la toma de decisiones de los profesionales de la salud. Se evaluaron cinco modelos de clasificación, destacando que Random Forest mostró el mejor rendimiento con la pérdida logarítmica más baja. Logistic Regression y SVM también fueron consistentes, aunque menos precisos que Random Forest, aún mostraron estabilidad en la predicción, sugiriendo su uso clínico potencial.

Los resultados obtenidos en el estudio destacan la utilidad de diversos modelos de clasificación para predecir la supervivencia en pacientes con cirrosis hepática. Aunque la literatura específica en esta área es limitada, investigaciones anteriores han explorado exitosamente el uso de algoritmos de clasificación en el ámbito médico. Se abordará el aporte de este ámbito desde un aspecto médico y científico

analizando las contribuciones de otros estudios sobre diversas enfermedades, resaltando su éxito en la aplicación de algoritmos similares.

Se realizó un estudio exhaustivo que ha explorado la eficacia de diferentes modelos de clasificación para predecir la supervivencia en pacientes con cirrosis hepática. Entre los modelos evaluados, Random Forest ha demostrado consistentemente la capacidad más sólida para predecir la supervivencia, seguido por Logistic Regression y SVM, que mostraron estabilidad en sus predicciones.

La evaluación se basó en la métrica de pérdida logarítmica, una herramienta esencial para comparar y medir la precisión de los modelos en la predicción de probabilidades. En este contexto, la baja pérdida logarítmica obtenida por Random Forest destaca su precisión para predecir la supervivencia en cirrosis hepática, mientras que los puntajes más altos de Naive Bayes y KNN indican una menor precisión en estas predicciones.

Además, se implementó la validación cruzada k-fold con cinco divisiones para evaluar la robustez de los modelos en diferentes subconjuntos de datos. Esto reafirmó la consistencia de Random Forest en mantener una pérdida logarítmica baja y consistente, respaldando su solidez en la predicción de la supervivencia en pacientes con cirrosis hepática. Mientras que Logistic Regression y SVM, aunque mostraron puntajes ligeramente más altos en pérdida logarítmica, demostraron una consistencia notable en sus predicciones a lo largo de las iteraciones de la validación cruzada.

Estos resultados resaltan la importancia de la selección cuidadosa de métricas de evaluación y técnicas de validación para comprender mejor el rendimiento de los modelos de clasificación en contextos médicos específicos, como la predicción de la supervivencia en cirrosis hepática.

En conclusión, mientras Random Forest emerge como el modelo preferido debido a su desempeño superior en términos de precisión, los resultados de Naive Bayes y KNN subrayan la importancia de una evaluación exhaustiva de los modelos antes de su aplicación clínica en la predicción de la supervivencia en pacientes con cirrosis hepática.

## REFERENCIAS

[1] E. D. Bethea and D. S. Pratt, "Valoración de la función hepática," in *Harrison. Principios de Medicina Interna*, vol. 2, 2023.

[2] Francesc, "CIRROSIS HEPÁTICA Definición," *cirrosis hepatica*, vol. 23, no. 3, 2020.

[3] A. Louvet, "Cirrosis asociada al alcohol," *EMC - Tratado Med.*, vol. 25, no. 2, 2021.

[4] E. F. Daza and E. F. Juan, "Aproximación al diagnóstico de enfermedades hepáticas por el laboratorio clínico," *Med. y Lab.*, vol. 14, no. 11–12, 2008.

[5] G. Sánchez-Visconti, "Función hepática y parámetros analíticos," *Lav*, 2010.

[6] J. E. Prieto O, S. Sánchez P, R. G. Prieto, E. L. Rojas D, L. González, and F. Mendivelso, "Características clínicas y descompensación en pacientes con cirrosis hepática atendidos en dos centros de hepatología en la ciudad de Bogotá D.C., 2010-2014," *Rev. Colomb. Gastroenterol.*, vol. 31, no. 1, 2016.

[7] M. Ferreira Cardoso *et al.*, "The impact and evolution of acute-on-chronic liver failure in decompensated cirrhosis: A Portuguese single-center study," *Gastroenterol. y Hepatol. (English Ed.)*, vol. 42, no. 5, 2019.

[8] L. Salleras Sanmarti, "VACUNACION CONTRA LA HEPATITIS B," *Atencion Primaria*, vol. 9, no. 1, 1992.

[9] D. A. Vásquez, M. V. Cely Mero, O. Muñoz Roca, R. L. Briones Jimenez, and C. Moncayo Valencia, "Prevalencia del hígado graso no alcohólico en estudiantes de la Universidad Católica de Santiago de Guayaquil, periodo 2014," *Medicina (B. Aires)*, vol. 20, no. 1, 2019.

[10] M. F. Delgado-López, J. Y. Chamaidan-Moreno, K. L. Labanda-Jaramillo, D. P. Delgado-Suárez, D. G. Echeverría-Gia, and S. G. Matute-Ortiz, "Estudio de casos sobre la cirrosis hepática y sus complicaciones en el Hospital Universitario de Guayaquil, año 2015," *Dominio las Ciencias*, vol. 5, no. 1, 2019.

[11] J. Alonso-Arévalo and A. Mirón-Canelo, "Aplicaciones móviles en salud: potencial, normativa de seguridad y regulación Mobile health applications: potential, regulation and security," *Scielo*, vol. 28, no. 3, 2017.

[12] A. Beltrán-Ostos, A. M. Urdaneta, and J. A. González, "Estado del arte, ventajas y limitaciones de la inteligencia artificial en epidemiología y salud pública," *Medicina (B. Aires)*, vol. 43, no. 4, 2022.

[13] D. N. Martínez-García, V. M. Dalgo-Flores, J. L. Herrera-López, E. I. Analuisa-Jiménez, and E. F. Velasco-Acurio, "Avances de la inteligencia artificial en salud," *Dominio las Ciencias*, vol. 5, no. 3, 2019.

[14] M. Anaya and C. Rodríguez, "ABC de la inteligencia artificial (IA) aplicada en la salud," *Medicina (B. Aires)*, vol. 43, no. 4, 2022.

[15] M. Learning, "Machine Learning: ¿ Están nuestros datos preparados?," *Miradax*, 2018.

[16] I. M. Álvarez-Bonilla, A. J. Romero-Fernández, G. E. Fernández-Villacrés, and L. R. Freire-Lescano, "Machine learning para la extracción de información biomédica en un laboratorio clínico," *CIENCIAMATRIA*, vol. 8, no. 4, 2022.

[17] I. N. Salamanca Rativa, "Técnicas de aprendizaje automático aplicadas en los sistemas de predicción," *Tecnol. Investig. y Acad.*, vol. 8, no. 1, 2021.

[18] S. Wang and Q. Chen, "The study of multiple classes boosting classification method based on local similarity," *Algorithms*, vol. 14, no. 2, 2021.

[19] T. Santamaría-Lopez, D. Patiño-Perez, V. González-Ruiz, and L. Flores-Carvajal, "IMPLEMENTATION OF MACHINE LEARNING TECHNIQUES AND CREATION OF AN ARTIFICIAL NEURAL NETWORK FOR THE PREDICTION OF THE ACADEMIC PERFORMANCE OF STUDENTS IN UNIVERSITY ENVIRONMENTS THAT USE E-LEARNING AND STREAMING," *Dyna*, vol. 98, no. 3, 2023.

[20] Kaggle, "Kaggle: Your Home for Data Science," *Website*, 2021.

[21] D. Patiño-pérez *et al.*, "Modelos de Machine Learning basados en Aprendizaje Supervisado para la Detección de Diabetes Mellitus en la Ciudad de Guayaquil. Machine Learning Models based in Supervised Learning for the Detection of Diabetes Mellitus in the City of Guayaquil .," *LACCEI*, pp. 1–8, 2022.

[22] C. Russo, H. Ramón, N. Alonso, B. Cicerchia, L. Esnaola, and J. P. Tessore, "Tratamiento Masivo de Datos Utilizando Técnicas de Machine Learning," *XVIII Work. Investig. en Ciencias la Comput.*, 2016.

[23] L. Judith Sandoval, "Algoritmos de aprendizaje automático para análisis y predicción de datos," *Rev. Tecnológica ITCA-FEPADE*, vol. 11, 2018.

[24] A. Y. Hussein, P. Falcarin, and A. T. Sadiq, "Enhancement performance of random forest algorithm via one hot encoding for IoT IDS," *Period. Eng. Nat. Sci.*, vol. 9, no. 3, 2021.

[25] D. Patiño-Pérez, C. Munive-Mora, L. Chóez-Acosta, A. Collantes-Farah, M. Molina-Calderón, and Z. Reyez-Sánchez, "Rapid Detection of Diabetic Retinopathy Through Deep Learning," in *Proceedings of the LACCEI international Multi-conference for Engineering, Education and Technology*, 2023, vol. 2023-July.

[26] IBM, "What is Machine Learning? | IBM," *The International Business Machines Corporation*, 2023.

[27] R. Gandhi, "Support Vector Machine — Introduction to Machine Learning Algorithms," *Towar. Data Sci.*, 2018.

[28] IBM Cloud Education, "¿Qué es machine learning? - México | IBM," *IBM*, 2020.

[29] M. Schonlau and R. Y. Zou, "The random forest algorithm for statistical learning," *Stata J.*, vol. 20, no. 1, 2020.

[30] S. Chen, G. I. Webb, L. Liu, and X. Ma, "A novel selective naïve

- Bayes algorithm.” *Knowledge-Based Syst.*, vol. 192, 2020.
- [31] R. Borja-Robalino, A. Monleón-Getino, and J. Rodellar, “Estandarización de métricas de rendimiento para clasificadores Machine y Deep Learning,” *Rev. Ibérica Sist. y Tecnol. la Inf.*, 2022.
- [32] J.-L. Fernández-Hernández, P. Herranz-Hernández, and L. Segovia-Torres, “Validación cruzada sobre una misma muestra: una práctica sin fundamento,” *R.E.M.A. Rev. electrónica Metodol. Apl.*, vol. 24, no. 1, 2022.
- [33] L. Al-Alawi, J. Al Shaqsi, A. Tarhini, and A. S. Al-Busaidi, “Using machine learning to predict factors affecting academic performance: the case of college students on academic probation,” *Educ. Inf. Technol.*, vol. 28, no. 10, 2023.
- [34] M. Sabhnani and G. Serpen, “Why machine learning algorithms fail in misuse detection on KDD intrusion detection data set,” *Intell. Data Anal.*, vol. 8, no. 4, 2004.