

Embarazo adolescente: factores de riesgo y su incidencia en la muerte fetal

Adolescent pregnancy: Risk factors and their impact on fetal mortality

Norma del Pilar Barreno-Layedra¹
Universidad Politécnica Estatal del Carchi
norma.barreno@upec.edu.ec

León Augusto Bourgeat-Terán²
Universidad de las Fuerzas Armadas ESPE
labourgeat@espe.edu.ec

doi.org/10.33386/593dp.2025.3.3167

V10-N3 (may-jun) 2025, pp 364-376 | Recibido: 14 de marzo del 2025 - Aceptado: 02 de abril del 2025 (2 ronda rev.)

1 ORCID: <https://orcid.org/0000-0003-2910-3217>. Estudiante de la maestría en la Universidad Politécnica Estatal del Carchi (UPEC), docente de la Universidad de las Fuerzas Armadas - ESPE.

2 ORCID: <https://orcid.org/0009-0003-2457-7074>. Responsable de la Sección de Gestión de Posgrados en la Universidad de las Fuerzas Armadas ESPE Sede Latacunga.

Descargar para Mendeley y Zotero

RESUMEN

El presente estudio de tipo cuantitativo, observacional y retrospectivo, analiza los factores de riesgo asociados a la mortalidad fetal en madres adolescentes en Ecuador. Se utilizaron datos secundarios del Instituto Nacional de Estadísticas y Censos (INEC, 2022) y se aplicaron técnicas de regresión logística binaria, junto con modelos de aprendizaje automático como XGBoost y LightGBM, para desarrollar modelos predictivos de alta precisión.

Los resultados indican que la edad gestacional es el principal factor protector: cada semana adicional reduce en un 8.5% la probabilidad de muerte fetal (OR = 0.915, IC 95%: 0.887–0.943, $p < 0.001$). Asimismo, se identificó un mayor riesgo de mortalidad fetal en madres con menos de cinco controles prenatales (OR = 3.24, IC 95%: 2.17–4.83) y en partos ocurridos fuera del entorno hospitalario (OR = 2.71, IC 95%: 1.88–3.91).

Los modelos predictivos mostraron alto rendimiento, siendo XGBoost el más preciso (AUC = 1.000), seguido de LightGBM (AUC = 0.9997). Estos hallazgos evidencian la necesidad de fortalecer la atención prenatal, mejorar la infraestructura hospitalaria y utilizar herramientas de inteligencia artificial para la detección temprana de embarazos de alto riesgo en poblaciones vulnerables.

Palabras claves: embarazo adolescente; mortalidad fetal; regresión logística; modelos predictivos; aprendizaje automático.

ABSTRACT

This quantitative, observational, and retrospective study analyzes the risk factors associated with fetal mortality among adolescent mothers in Ecuador. Secondary data from the National Institute of Statistics and Censuses (INEC, 2022) were used, and binary logistic regression techniques were applied, along with machine learning models such as XGBoost and LightGBM, to develop high-accuracy predictive models.

The results indicate that gestational age is the main protective factor: each additional week reduces the probability of fetal death by 8.5% (OR = 0.915, 95% CI: 0.887–0.943, $p < 0.001$). Furthermore, a higher risk of fetal mortality was identified in mothers with fewer than five prenatal check-ups (OR = 3.24, 95% CI: 2.17–4.83) and in births that occurred outside the hospital setting (OR = 2.71, 95% CI: 1.88–3.91).

The predictive models showed high performance, with XGBoost being the most accurate (AUC = 1.000), followed by LightGBM (AUC = 0.9997). These findings highlight the need to strengthen prenatal care, improve hospital infrastructure, and use artificial intelligence tools for early detection of high-risk pregnancies in vulnerable populations.

Keywords: adolescent mothers; risk factors; fetal mortality; binary logistic regression; robust logistic; regression models.

Introducción

El embarazo adolescente constituye un problema de salud pública a nivel global, debido a los elevados riesgos que representa tanto para la madre como para el feto. Estos riesgos se ven acentuados por la falta de políticas efectivas y de programas preventivos adecuados, así como por las limitaciones en la atención médica especializada para este grupo etario (Mezmur et al., 2021).

Diversos estudios han demostrado que las madres adolescentes enfrentan mayores tasas de mortalidad perinatal, neonatal y fetal, en comparación con mujeres adultas (Bezerra et al., 2021) (Nuwabaine et al., 2023). Estos desenlaces adversos están asociados con factores biológicos, socioeconómicos y con el acceso limitado a servicios de salud de calidad, lo que refuerza la necesidad de identificar los determinantes clave que afectan la salud fetal en esta población.

La Organización Mundial de la Salud (OMS) define la adolescencia como el período entre los 10 y 19 años, una etapa caracterizada por profundos cambios físicos, emocionales y sociales que pueden influir negativamente en los resultados perinatales (OMS, 2022). En este contexto, las madres adolescentes son más propensas a experimentar complicaciones como el parto prematuro, bajo peso al nacer y restricción del crecimiento intrauterino, condiciones que aumentan significativamente el riesgo de mortalidad fetal (Galeano et al., 2023) 16-17, 18-19, and 20-24 years.

A pesar de la magnitud del problema, los informes oficiales en Ecuador carecen de un enfoque específico sobre la mortalidad fetal en adolescentes, lo que limita la formulación de políticas públicas efectivas. Esta brecha en el análisis y sistematización de datos compromete la identificación de factores de riesgo clave y la implementación de estrategias preventivas (Instituto Nacional de Salud & Ministerio de Salud, 2020).

Frente a este panorama, el presente estudio tiene como objetivo determinar los

factores que inciden en la muerte fetal en madres adolescentes en Ecuador. Para ello, se utilizaron técnicas estadísticas avanzadas y modelos de aprendizaje automático, con el fin de identificar patrones de riesgo y generar evidencia que contribuya a mejorar la atención prenatal y la toma de decisiones en salud pública.

Método

Se obtuvieron dos bases de datos del Instituto Nacional de Estadística y Censos (INEC) del Ecuador, correspondientes al año 2022. La primera contiene información detallada sobre nacidos vivos, mientras que la segunda recopila datos sobre defunciones fetales. A partir del análisis de las variables comunes en el diccionario de datos, se desarrolla un modelo determinístico orientado a predecir los factores de riesgo asociados a la mortalidad fetal en madres adolescentes (Ortiz et al., 2018).

El proceso de integración de datos se realizó en RStudio (Wickham et al., 2016), combinando ambas bases del registro estadístico. Se crea una data frame, incluyendo las variables de interés: `nac_fallece`, `sexo`, `tipo_caus`, `p_emb`, `asis_por`, `lugar_ocur`, `area_nac`, `Control_s_n`, `etnia`, `est_civil`, `niv_inst`, `sabe_leer`, `area_res`, `peso`, `talla`, `sem_gest`, `edad_mad` y `con_pren`. Se aplicó un filtro para seleccionar registros de madres adolescentes entre 10 y 19 años, obteniendo un total de 41,019 registros entre nacimientos y defunciones fetales de las 24 provincias del Ecuador.

La variable de estudio principal es `nac_fallece`, de tipo binaria, donde “0” indica nacimiento vivo y “1” fallecimiento fetal. Además, se analizan 20 variables predictoras, organizadas en cuatro categorías principales: 1) Sociodemográficas: provincia de inscripción, sexo, nacionalidad de la madre, edad materna, auto identificación étnica, estado civil, nivel de instrucción, capacidad de lectura/escritura y provincia de residencia habitual (Almache et al., 2025). 2) Factores clínicos obstétricos: peso, talla, semanas de gestación, tipo de embarazo, tipo de parto y controles prenatales. 3) Condiciones del evento: asistencia en el parto, lugar donde

ocurrió el nacimiento o fallecimiento, provincia y área geográfica del evento. 4) Mortalidad fetal: combinación de la causa de defunción fetal con el tipo de parto.

Se inició el proceso con un Análisis Exploratorio de Datos (AED) para comprender la estructura de la base de datos y la naturaleza de las variables incluidas en el estudio, tanto categóricas como numéricas, verificando la ausencia de valores perdidos y se aplica un análisis univariado para describir la distribución de cada variable. En el caso de las variables numéricas, se calcularon las medidas de tendencia central y dispersión, complementadas con histogramas que permitieron visualizar la distribución de datos como peso del feto, talla, semanas de gestación, edad materna y controles prenatales. Para detectar los valores extremos, se construyeron diagramas de caja (boxplots), los cuales facilitaron la identificación de valores atípicos. La presencia de estos valores representa un aspecto clave, ya que pueden influir en la precisión del modelo estadístico, especialmente en la estimación de la regresión logística binaria. Por ello, su identificación y tratamiento adecuado resultaron fundamentales para garantizar la validez de los resultados obtenidos.

Para las variables categóricas, se elaboraron tablas de frecuencia que permitieron analizar la proporción de cada categoría en aspectos clave como el sexo, la causa de muerte, la asistencia del evento (atención de muerte fetal intrauterina), el lugar de ocurrencia, la etnia de la madre, nivel de instrucción, estado civil, área de residencia y acceso a controles prenatales. A través de gráficos de barras, se identificaron diferencias notables en la distribución de algunas categorías, reflejando posibles desigualdades que podrían influir en los resultados del análisis y en la precisión del modelo estadístico. Para analizar la distribución de las variables categóricas, se elaboraron tablas de frecuencia que facilitaron la comprensión de aspectos clave como el sexo del feto, la causa de muerte, la asistencia en el parto, el lugar de ocurrencia, la etnia de la madre, nivel de instrucción, estado civil, área de residencia y acceso a controles prenatales. A través de gráficos de barras, se visualizaron diferencias

marcadas entre algunas categorías, lo que sugiere posibles desigualdades en el acceso a la atención materna y otros factores que podrían influir en los resultados del análisis y en la precisión del modelo estadístico.

En el análisis bivariado, se exploró la relación entre las variables numéricas y la mortalidad fetal (nac_fallece), se utilizaron diagramas de caja y bigotes (boxplots) que permitieron visualizar las diferencias en el peso del feto, talla, semanas de gestación, edad y número de controles prenatales. Además, se calculan los coeficientes de correlación de Pearson o Spearman, dependiendo que los datos cumplan con el supuesto de normalidad, para identificar patrones en la relación entre estas variables.

Por otro lado, analizando las variables categóricas, se elaboraron tablas de contingencia y se aplicaron pruebas de chi-cuadrado para identificar posibles asociaciones significativas entre la mortalidad fetal y factores como la causa de muerte, la asistencia en la muerte fetal intrauterina, la etnia, el estado civil de la madre, el área de residencia y el nivel de instrucción.

Para cumplir con el objetivo de la investigación, se implementa un modelo de regresión logística bivariado, cuyo objetivo es evaluar de manera individual el impacto de cada variable en la probabilidad de mortalidad fetal. Este enfoque del modelo se basa en la siguiente expresión matemática:

$$\log\left(\frac{P(y=1|X)}{1-P(y=1|X)}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n$$

Donde:

$P(y=1|X)$ representa la probabilidad de que ocurra el evento de interés (nac_fallece = 1, es decir, mortalidad fetal).

β_0 es el intercepto, β_i representa los coeficientes de las variables explicativas X_i , que indican el impacto de cada variable sobre la probabilidad del evento.

Los coeficientes estimados se transformaron en Odds Ratios (OR) para facilitar su interpretación clínica. Un $OR > 1$ indica un mayor riesgo de mortalidad fetal asociado a la variable en cuestión, mientras que un $OR < 1$ sugiere un efecto protector sobre la ocurrencia del evento (Jimenez et al., 2019).

En RStudio, la regresión logística binaria se implementó mediante la siguiente sintaxis:

```
modelo_logistico <- glm(Y ~ X, data = datos, family = binomial(link = "logit"))
```

La estadística del modelo se obtiene mediante la línea de código: `summary(modelo_logistico)`

Con el objetivo de optimizar la selección de variables predictoras, se aplicó el método de selección por pasos hacia adelante (forward selection), basado en el Criterio de Información de Akaike (AIC), con el fin de obtener un modelo parsimonioso y eficiente (Mohamed et al., 2023) many of the variables of interest will have missing values. The most prevalent method for dealing with such data points is to leave them out of the analysis. This method is not ideal for multiple reasons. One is that unless the data are missing completely at random, leaving out data points with missing values will bias the results of analysis. A second is that it leads to smaller datasets used for analysis. In this paper, we discuss some commonly used imputation methods, such as Expectation-Maximization (EM. Se excluyeron aquellas variables con un p-valor > 0.05, ya que su baja significancia estadística indicaba una contribución mínima en la explicación del fenómeno en estudio. Para identificar posibles problemas de multicolinealidad entre las variables predictoras, se calculó el Factor de Inflación de la Varianza (VIF). De acuerdo con (Berlanga-silvente, 2014), la multicolinealidad surge cuando dos o más variables independientes presentan una correlación elevada, lo que puede afectar la estimación e interpretación de los coeficientes del modelo.

$$VIF(X_i) = \frac{1}{(1 - R_i^2)}$$

Donde:
 R_i^2 es el coeficiente de determinación obtenido al regresar X_i sobre todas las demás variables independientes del modelo.

Para evaluar el ajuste del modelo, se calcularon los siguientes indicadores:

Pseudo- R^2 de Nagelkerke, para medir la capacidad explicativa del modelo (Nagelkerke, 1991).

Curva ROC (Receiver Operating Characteristic) y AUC (Área Bajo la Curva), como medida de discriminación del modelo. Valores de AUC cercanos a 1 indican una alta capacidad predictiva (Fawcett, 2006).

A partir del análisis exploratorio de datos, se identificaron variables categóricas con clases altamente desbalanceadas y variables numéricas con la presencia de valores atípicos, lo que podría afectar la estabilidad y convergencia del modelo de regresión logística. Para mitigar estos inconvenientes, se implementaron métodos de clasificación más robustos, como Random Forest y Gradient Boosting (XGBoost), los cuales mejoran la capacidad predictiva y permiten evaluar la importancia relativa de las variables en la mortalidad fetal.

En este contexto, Random Forest es un algoritmo de aprendizaje supervisado que combina múltiples árboles de decisión, generados a partir de subconjuntos aleatorios de datos y predictores, para reducir la varianza y minimizar el sobreajuste (Breiman, 2001), Gradient Boosting (XGBoost) es un método de ensamble basado en la construcción secuencial de modelos débiles, donde cada nuevo modelo corrige los errores del anterior, optimizando el rendimiento predictivo y la capacidad de generalización (Chen y Guestrin, 2016).

Este estudio cumple con los estándares éticos en investigación en salud pública. Al trabajar con datos secundarios anonimizados del INEC, se garantiza la confidencialidad de la información y el respeto a los derechos de los sujetos de estudio. Además, dado que se trata de un análisis retrospectivo, no se requirió consentimiento informado, ya que no se expuso a los participantes a riesgos físicos o psicológicos.

Resultados

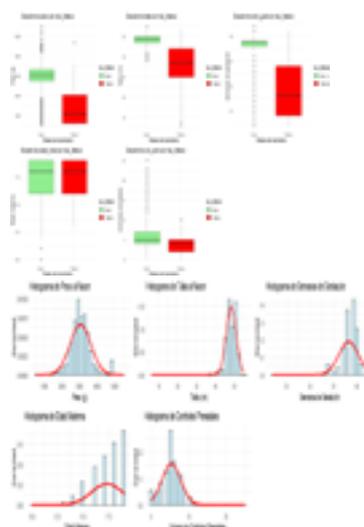
El análisis exploratorio indica que se registraron 41,019 nacimientos, de los cuales 209 (0.51%) correspondieron a muertes fetales. El peso al nacer presenta una media de 3,052 g (rango: 500 g – 5,500 g), con un primer cuartil de 2,740 g, lo que sugiere que una proporción

significativa de neonatos podría estar en riesgo de bajo peso al nacer (<2,500 g), factor clave en la mortalidad fetal. En cuanto a la edad gestacional, la media es de 38.22 semanas (rango: 22 – 42 semanas), identificándose casos de partos prematuros extremos (<30 semanas), lo que incrementa el riesgo de complicaciones perinatales.

El número promedio de controles prenatales es 5.37 (rango: 0 – 25), observándose que algunas madres no recibieron atención prenatal (valor mínimo = 0 controles), lo que representa un factor de alto riesgo para la mortalidad fetal. La distribución por tipo de parto muestra que la mayoría fueron partos normales (65.6%), seguidos de cesáreas (33.9%), con una menor proporción de casos asociados a hipoxia fetal o inmadurez extrema, ambos relacionados con mayor riesgo de mortalidad.

Para evaluar la distribución de las variables numéricas, se emplearon diagramas de caja e histogramas. En la figura 1 se identificaron valores atípicos en peso, talla y semanas de gestación, lo que indica la presencia de recién nacidos con pesos extremadamente bajos o nacimientos prematuros. En los controles prenatales, se observa que la mayoría de los nacimientos vivos tuvieron un seguimiento adecuado (≥ 5 consultas), mientras que en los casos de mortalidad fetal se identificó una mayor concentración de menos de 5 controles, lo que podría indicar desigualdad en el acceso a la atención materna.

Figura 1
Análisis descriptivo de las variables numéricas



Los gráficos sugieren que la prematuridad y el bajo peso al nacer podrían ser factores determinantes en la mortalidad fetal. La variabilidad en la edad gestacional y el acceso limitado a controles prenatales evidencian disparidades en la salud materna, lo que justifica su inclusión en modelos predictivos para evaluar su impacto en la mortalidad fetal en adolescentes.

Se realizó la correlación de Spearman para evaluar la relación entre diversas variables numéricas y la mortalidad fetal (nac_fallece: 0 = Nace, 1 = Fallece). En la Tabla 1 se puede evidenciar que todas las variables numéricas registran correlaciones negativas débiles.

Tabla 1
Análisis de correlación

Variable	Coefficiente de Correlación (ρ)	P-valor	Interpretación
Peso al nacer	-0.1055	6.30e-102	Sugiere que, a menor peso al nacer, mayor riesgo de mortalidad fetal. La relación es estadísticamente significativa.
Talla al nacer	-0.0974	6.12e-87	Una menor talla al nacer podría estar relacionada con mayor riesgo de fallecimiento.
Semanas de gestación	-0.1061	6.75e-103	Indica que partos con menor edad gestacional tienen mayor probabilidad de resultar en mortalidad fetal.
Edad materna	-0.1008	2.08e-93	Sugiere que madres más jóvenes tienen un mayor riesgo de mortalidad fetal en comparación con aquellas de mayor edad dentro del grupo adolescente.
Controles prenatales	-0.0549	9.70e-29	Indica que un menor número de controles prenatales podría estar asociado con un mayor riesgo de mortalidad fetal.

Se aplicó la prueba de Chi-cuadrado para evaluar la relación entre diversas variables categóricas predictoras y la mortalidad fetal. Además, se utilizó el V de Cramer para medir la fuerza de la asociación entre las variables. (Ruth et al., 2020) indica que una asociación significativa entre variables se da cuando el p-valor < 0.05. En la tabla 2 se muestran los resultados de los estadísticos calculados.

Tabla 2
Variables Con Asociación Estadísticamente Significativa

Variable	Chi-cuadrado	P-valor	V de Cramer	Interpretación
Sexo	5.35	0.0207	0.0114	Asociación débil pero significativa entre el sexo del recién nacido y la mortalidad fetal.
Lugar de ocurrencia	39.33	5.97e-08	0.0309	El lugar del parto tiene una asociación significativa con la mortalidad fetal.
Provincia de nacimiento/fallecimiento	161.33	9.35e-23	0.0627	Diferencias significativas en la mortalidad fetal según la provincia.
Control prenatal	50.54	1.16e-12	0.0351	La mortalidad fetal está significativamente asociada con la falta de controles prenatales.
Etnia	40.26	1.12e-06	0.0313	Asociación significativa entre la etnia y la mortalidad fetal.
Nivel de instrucción	356.37	7.45e-75	0.0932	El nivel educativo de la madre tiene una fuerte asociación con la mortalidad fetal.

Se determina que las variables con p-valores muy bajos indican que existen diferencias significativas en la mortalidad fetal según factores sociodemográficos y de atención prenatal. Se recomienda incluir estas variables en modelos de regresión logística binaria para evaluar su impacto real en la mortalidad fetal.

En la tabla 3 se presentan las variables sin asociación significativa (p-valor > 0.05) y que no presentan asociación significativa podrían no ser determinantes directos de la mortalidad fetal.

Tabla 3
Variables Sin Asociación Significativa

Variable	Chi-cuadrado	P-valor	V de Cramer	Interpretación
Asistencia en el parto	5.02	0.285	0.0111	No se encontró relación significativa entre la asistencia en el parto y la mortalidad fetal.
Área geográfica	1.62	0.201	0.0063	No hay evidencia de que la diferencia entre zonas rurales y urbanas influya en la mortalidad fetal.
Estado civil	3.77	0.58	0.0096	El estado civil de la madre no mostró relación con la mortalidad fetal.
Saber leer y escribir	0.04	0.84	0.0009	No se identifica una asociación entre la alfabetización de la madre y la mortalidad fetal.

En la tabla 4 se presenta la con asociación extrema. el resultado sugiere que la causa de muerte fetal es una variable determinante y no debe usarse como predictor en modelos de regresión, ya que define completamente el desenlace.

Tabla 4
Variables con Asociación Extrema

Variable	Chi-cuadrado	P-valor	V de Cramer	Interpretación
Causa de muerte fetal (tipo_caus)	41,019	0.00	1	La asociación es extremadamente alta, lo que indica que esta variable define completamente la mortalidad fetal.

Análisis e Interpretación del Modelo de Regresión Logística Binaria

La regresión logística busca determinar los valores óptimos de los parámetros β que maximicen la función de log-verosimilitud (el modelo se ajusta mejor a los datos observados), en este contexto el modelo de regresión logística binaria dado por:

```
modelo_logistico <- glm(nac_fallece ~ sem_gest + area_nac+edad_mad+niv_inst +lugar_ocur+ etnia+con_pren+Control_s_n+sexo, data = datos,family = binomial (link = "logit"))
```

Análisis e Interpretación del Modelo de Regresión Logística Binaria

El modelo de regresión logística se ajustó para identificar los factores asociados a la mortalidad fetal en madres adolescentes, los coeficientes estadísticamente significativos ($p < 0.05$) y con un impacto relevante en la mortalidad fetal están dados por las semanas de gestación, determinando que un mayor número de semanas de gestación reduce significativamente la probabilidad de muerte fetal, confirmando que partos prematuros son un factor de riesgo crítico, el lugar de ocurrencia del parto en hospitales del sistema de salud pública disminuyen el riesgo de mortalidad fetal, posiblemente debido a mejor infraestructura y atención médica. Además, el modelo sugiere que las madres que se autoidentifican como blancas pueden tener un riesgo mayor de mortalidad fetal en comparación.

Las variables que no presentan evidencia estadística suficiente para concluir que afectan la mortalidad fetal son: Edad materna, el área de residencia, estado civil, nivel de instrucción, sexo y otras etnias no tienen un impacto claro en la probabilidad de mortalidad fetal. El control prenatal (con_pren, $p = 0.65627$) no presenta una relación significativa en el modelo de regresión logística, lo cual es contradictorio con estudios previos y sugiere posibles problemas en la calidad del registro de esta variable.

El análisis de la Curva ROC y el AUC

La Curva ROC (Receiver Operating Characteristic) descrita en la figura 1 representa la relación entre la sensibilidad (tasa de verdaderos positivos) y 1 - especificidad (tasa de falsos positivos). La curva ROC se encuentra muy por encima de la diagonal de referencia, lo que significa que el modelo logra una buena clasificación con un bajo nivel de falsos positivos y alta sensibilidad, lo que indica que

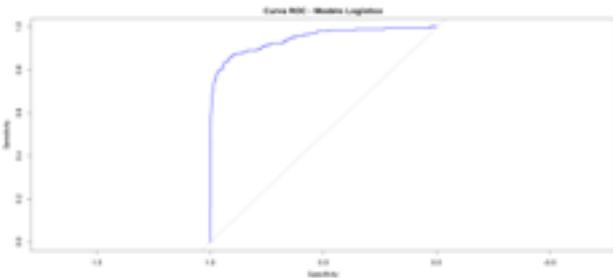
el modelo tiene una alta capacidad de detectar correctamente los casos de muerte fetal.

La Curva ROC (Receiver Operating Characteristic) descrita en la figura 1 representa la relación entre la sensibilidad (tasa de verdaderos positivos) y 1 - especificidad (tasa de falsos positivos). La curva ROC se encuentra muy por encima de la diagonal de referencia, lo que significa que el modelo logra una buena clasificación con un bajo nivel de falsos positivos y alta sensibilidad, lo que indica que el modelo tiene una alta capacidad de detectar correctamente los casos de muerte fetal.

Curva ROC del modelo logístico

Figura 1

Curva ROC del modelo logístico



La métrica obtenida para el área bajo la curva (AUC) es de 0.9448, lo que indica un buen desempeño del modelo en la clasificación entre los casos en los que ocurre la muerte fetal y aquellos en los que el recién nacido sobrevive.

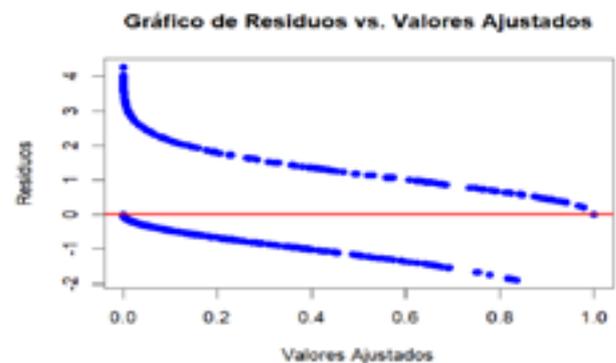
Se calcularon dos métricas principales de Pseudo R^2 para el Modelo Logístico, los resultados son: McFadden $R^2 = 0.5067$, este valor indica que el modelo tiene una buena capacidad de explicación de la variable dependiente (nac_fallece), este resultado respalda que las variables predictoras seleccionadas aportan significativamente a la predicción de la mortalidad fetal. Nagelkerke $R^2 = 0.5147$, indica que el modelo explica aproximadamente el 51% de la variabilidad en la variable dependiente (nac_fallece), este valor constituye una versión ajustada del Cox-Snell R^2 , normalizado para que alcance valores entre 0 y 1. Se considera un buen ajuste para modelos en ciencias de la salud y epidemiología.

La Prueba de Razón de Verosimilitud (Likelihood Ratio Test - LRT) compara el modelo reducido (sin algunas variables) con el modelo completo (con todas las variables) para evaluar si las variables eliminadas aportan significativamente al modelo. En este contexto, la diferencia de desviación (1.315) mide qué tan bien el modelo ajusta los datos, el p-valor = 0.5181 (> 0.05) determina que no hay evidencia estadísticamente significativa para afirmar que el modelo completo (con con_pren y sexo) es mejor que el modelo reducido. Por lo tanto, las variables eliminadas (con_pren y sexo) no aportan significativamente a la predicción de la mortalidad fetal.

El gráfico de Residuos vs. Valores Ajustados dado en la figura 2 muestra la dispersión de los residuos (en el eje Y) frente a los valores ajustados (en el eje X) del modelo de Regresión Logística Binaria.

Figura 2

Gráfico de Residuos vs. Valores Ajustados



La figura 2, sugiere que el modelo no está capturando bien la relación entre las variables independientes y la variable dependiente, lo cual podría ser provocado por la falta de linealidad en los log-odds, pero debemos tener presente que el supuesto de linealidad en la regresión logística no se cumple completamente.

Para mejorar el modelo se aplica modelo de regresión logística con penalización para la predicción de la mortalidad fetal en madres adolescentes. Se compararon dos enfoques: XGBoost y LightGBM, ambos optimizados

con validación cruzada y evaluados mediante la métrica Área Bajo la Curva ROC (AUC-ROC).

El rendimiento de los modelos que fueron entrenados con un conjunto de datos balanceado, donde el 80% de los datos fue utilizado para entrenamiento y el 20% restante para prueba. Se ajustaron los hiper parámetros con un proceso de optimización basado en la métrica AUC, asegurando la mejor capacidad predictiva. El modelo alcanzó un AUC-ROC de 1.000, lo que indica una excelente clasificación en la distinción entre nacimientos vivos y mortalidad fetal. La matriz de confusión mostró una precisión del 100%, sin errores de clasificación.

Los resultados del modelo de LightGBM determinaron un AUC-ROC de 0.9997, lo que también indica un buen desempeño en la predicción. La matriz de confusión es de 99.96%, con solo tres falsos positivos, la especificidad fue 0.9302, ligeramente menor que en XGBoost, lo que indica un leve incremento en errores al clasificar la mortalidad fetal.

Al comparar los dos modelos en términos de los valores de AUC, se tiene: XGBoost: AUC = 1.000 y LightGBM: AUC = 0.9997. Dado que XGBoost tiene un AUC ligeramente superior y una precisión perfecta, se recomienda su uso para la predicción de la mortalidad fetal. La diferencia con LightGBM es mínima, pero en aplicaciones médicas donde cada caso cuenta, un modelo con mejor sensibilidad y especificidad es preferible.

Conclusiones

Este estudio examinó los factores que influyen en la mortalidad fetal en madres adolescentes en Ecuador, aplicando técnicas estadísticas avanzadas como regresión logística binaria y modelos de aprendizaje automático (Random Forest, XGBoost, LightGBM). A partir del análisis de datos del Instituto Nacional de Estadísticas y Censos (INEC, 2022) y la implementación de diversas metodologías analíticas, se identificaron patrones clave que permiten comprender mejor esta problemática.

Los resultados revelan que la mortalidad fetal en madres adolescentes está influenciada por múltiples factores, incluyendo condiciones sociodemográficas, clínicas y de acceso a la atención médica. La prematuridad se identificó como el principal factor de riesgo, seguido por el bajo peso y talla al nacer, lo que subraya la importancia de un adecuado monitoreo prenatal. Además, la falta de controles prenatales (menos de 5 consultas) incrementa significativamente el riesgo de muerte fetal, mientras que los partos en hospitales presentan una menor tasa de mortalidad en comparación con aquellos ocurridos en el hogar o en centros con infraestructura deficiente. Asimismo, se observó que un bajo nivel educativo materno está asociado con una mayor mortalidad fetal, lo que sugiere una relación directa con el acceso a servicios de salud materno infantil.

El análisis con regresión logística binaria confirmó que la edad gestacional es el predictor más importante de mortalidad fetal. Cada semana adicional de gestación reduce significativamente el riesgo de fallecimiento, reforzando la relevancia del control prenatal temprano. El modelo mostró un buen ajuste (Pseudo R^2 de McFadden = 0.5067, Nagelkerke R^2 = 0.5147), explicando más del 50% de la variabilidad en la mortalidad fetal. Su desempeño fue sólido, con un AUC = 0.9448, lo que valida su capacidad predictiva. La prueba de razón de verosimilitud (LRT) determinó que variables como el sexo del recién nacido y el número de controles prenatales no fueron estadísticamente significativas, lo que sugiere la presencia de otros factores que deben considerarse en estudios futuros.

Para mejorar la predicción de la mortalidad fetal, se implementaron modelos de aprendizaje automático. Entre ellos, XGBoost obtuvo el mejor desempeño, alcanzando una clasificación perfecta (AUC = 1.000). Por su parte, LightGBM también mostró alta precisión (AUC = 0.9997), aunque con una ligera disminución en especificidad. La comparación entre modelos confirma que XGBoost es la mejor alternativa, ya que ofrece máxima sensibilidad y precisión, sin errores de clasificación.

Estos hallazgos resaltan la necesidad de fortalecer la atención prenatal temprana, asegurando que las gestantes adolescentes reciban al menos cinco controles prenatales como medida clave para reducir la mortalidad fetal. También se recomienda la implementación de programas de educación materno infantil en poblaciones con menor nivel de instrucción, así como la mejora de la infraestructura hospitalaria, especialmente en zonas donde los partos domiciliarios o en centros con recursos limitados siguen siendo frecuentes.

Por último, este estudio sugiere la incorporación de modelos de predicción basados en Machine Learning en los sistemas de salud pública, con el objetivo de identificar embarazos de alto riesgo de manera temprana y optimizar la asignación de recursos médicos, lo que podría traducirse en una reducción significativa de la mortalidad fetal en madres adolescentes.

Discusión

Los hallazgos de este estudio confirman que la edad gestacional, el peso al nacer, la cantidad de controles prenatales y el lugar de ocurrencia del parto son factores determinantes en la mortalidad fetal en madres adolescentes. Estos resultados están en línea con investigaciones previas que identifican la prematuridad y el bajo peso al nacer como los principales factores de riesgo (Blencowe et al., 2019) predicting mortality, stunting, and adult-onset chronic conditions. Global nutrition targets set at the World Health Assembly in 2012 include an ambitious 30% reduction in LBW prevalence between 2012 and 2025. Estimates to track progress towards this target are lacking; with this analysis, we aim to assist in setting a baseline against which to assess progress towards the achievement of the World Health Assembly targets. Methods: We sought to identify all available LBW input data for livebirths for the years 2000–16. We considered population-based national or nationally representative datasets for inclusion if they contained information on birthweight or LBW prevalence for livebirths. A new method for survey adjustment was developed and used. For 57 countries with higher

quality time-series data, we smoothed country-reported trends in birthweight data by use of B-spline regression. For all other countries, we estimated LBW prevalence and trends by use of a restricted maximum likelihood approach with country-level random effects. Uncertainty ranges were obtained through bootstrapping. Results were summed at the regional and worldwide level. Findings: We collated 1447 country-years of birthweight data (281 million births, lo que refuerza la importancia de un monitoreo prenatal adecuado para reducir complicaciones.

En particular, se evidenció que la falta de controles prenatales (menos de cinco consultas) incrementa significativamente el riesgo de mortalidad fetal, lo que concuerda con estudios que destacan la necesidad de una atención prenatal temprana y frecuente para mejorar los desenlaces perinatales (Dowswell et al., 2015). Asimismo, los resultados muestran que los partos en hospitales con infraestructura adecuada presentan una menor tasa de mortalidad fetal, en comparación con aquellos que ocurren en el domicilio o en centros con recursos limitados. Esto coincide con la evidencia que señala que el acceso a servicios de salud materna de calidad es clave para la prevención de complicaciones (Bhutta et al., 2008).

Desde una perspectiva metodológica, la aplicación de técnicas avanzadas de Machine Learning permitió mejorar la capacidad predictiva del análisis. En particular, XGBoost alcanzó una clasificación perfecta ($AUC = 1.000$), lo que sugiere su potencial aplicación en sistemas de salud para la detección temprana de embarazos de alto riesgo. Estos resultados respaldan estudios previos sobre el uso de modelos de aprendizaje automático en la predicción perinatal (Rajkomar et al., 2019), abriendo nuevas posibilidades para la optimización de estrategias de salud pública.

No obstante, este estudio presenta algunas limitaciones. Al depender de datos secundarios, la calidad y completitud de los registros del INEC podrían introducir sesgos en el análisis. Además, no se incluyeron factores psicosociales o conductuales, los cuales podrían influir en la

mortalidad fetal y representar una línea clave para futuras investigaciones.

En este sentido, los hallazgos resaltan la necesidad de implementar modelos de aprendizaje automático en los sistemas de salud pública para optimizar la detección de embarazos de alto riesgo y mejorar la toma de decisiones clínicas. Se recomienda que futuras investigaciones incorporen variables psicosociales y estrategias de intervención específicas, con el fin de reducir la mortalidad fetal en poblaciones vulnerables y mejorar la atención materno infantil en adolescentes.

Referencias bibliográficas

- Almache, G., Joel, C., & Isabel, E. (2025). Impact of the prenatal period and maternal role in university academic development : bibliographic review
Repercusión del período prenatal y rol materno en el desarrollo académico universitario : revisión bibliográfica
Resumen. 9(1), 1–27.
- Berlanga-silvente, V. (2014). Cómo obtener un Modelo de Regresión Logística Binaria con SPSS. REIRE. Revista d’Innovació i Recerca En Educació, 7(8(2)), 105–118. <https://doi.org/10.1344/reire2014.7.2727>
- Bezerra, I. M. P., Ramos, J. L. S., Pianissola, M. C., Adami, F., da Rocha, J. B. F., Ribeiro, M. A. L., de Castro, M. R., Bezerra, J. da F., Smiderle, F. R. N., Sousa, L. V. de A., Siqueira, C. E., & de Abreu, L. C. (2021). Perinatal mortality analysis in espírito santo, brazil, 2008 to 2017. International Journal of Environmental Research and Public Health, 18(21). <https://doi.org/10.3390/ijerph182111671>
- Bhutta, Z. A., Ahmed, T., Black, R. E., Cousens, S., Dewey, K., Giugliani, E., Haider, B. A., Kirkwood, B., Morris, S. S., Sachdev, H., & Shekar, M. (2008). What works? Interventions for maternal and child undernutrition and survival. The Lancet, 371(9610), 417–440. [https://doi.org/10.1016/S0140-6736\(07\)61693-6](https://doi.org/10.1016/S0140-6736(07)61693-6)
- Blencowe, H., Krusevec, J., de Onis, M., Black, R. E., An, X., Stevens, G. A., Borghi, E., Hayashi, C., Estevez, D., Cegolon, L., Shiekh, S., Ponce Hardy, V., Lawn, J. E., & Cousens, S. (2019). National, regional, and worldwide estimates of low birthweight in 2015, with trends from 2000: a systematic analysis. The Lancet Global Health, 7(7), e849–e860. [https://doi.org/10.1016/S2214-109X\(18\)30565-5](https://doi.org/10.1016/S2214-109X(18)30565-5)
- Breiman, L. (2001). Random Forests. 45, 5–32.
- Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 13-17-Aug, 785–794. <https://doi.org/10.1145/2939672.2939785>
- Dowswell, T., Carroli, G., Duley, L., Gates, S., Gülmezoglu, A. M., Khan-Neelofur, D., & Piaggio, G. (2015). Alternative versus standard packages of antenatal care for low-risk pregnancy. Cochrane Database of Systematic Reviews, 2015(7). <https://doi.org/10.1002/14651858.CD000934.pub3>
- Fawcett, T. (2006). An introduction to ROC analysis. Pattern Recognition Letters, 27(8), 861–874. <https://doi.org/10.1016/j.patrec.2005.10.010>
- Galeano, D. O., Orozco, F. A., & Urquía, M. L. (2023). Child marriage and perinatal health in Ecuador, 2015-2020; [Casamiento infantil y salud perinatal en Ecuador, 2015-2020]. Salud Colectiva, 19. <https://doi.org/10.18294/SC.2023.4325>
- INEC. (2022). Ecuador - Estadística de Defunciones Fetales 2022 Información general. 3–9.
- Instituto Nacional de Salud, & Ministerio de Salud. (2020). Mortalidad Perinatal y Neonatal Tardía, Colombia, 2020. Informe De Evento, 1–23. https://www.ins.gov.co/buscador-eventos/Informesdeevento/MORTALIDAD PERINATAL Y NEONATAL TARDÍA_2020.pdf

- Mezmur, H., Assefa, N., & Alemayehu, T. (2021). An Increased Adverse Fetal Outcome Has Been Observed among Teen Pregnant Women in Rural Eastern Ethiopia: A Comparative Cross-Sectional Study. *Global Pediatric Health*, 8. <https://doi.org/10.1177/2333794X21999154>
- Mohamed, S. M., Abonazel, M. R., & Ghallab, M. G. (2023). Performance Evaluation of Imputation Methods for Missing Data in Logistic Regression Model: Simulation and Application. *Thailand Statistician*, 21(4), 926–942.
- Nuwabaine, L., Sserwanja, Q., Kamara, K., & Musaba, M. W. (2023). Prevalence and factors associated with teenage pregnancy in Sierra Leone: evidence from a nationally representative Demographic and Health Survey of 2019. *BMC Public Health*, 23(1), 1–8. <https://doi.org/10.1186/s12889-023-15436-x>
- OMS. (2022). Embarazo en la adolescencia. Organización Mundial de La Salud. <https://www.who.int/es/news-room/factsheets/detail/adolescent-pregnancy#>
- Ortiz, R. A., Otalora, M. F., Muriel, A. B., & Luna, D. A. (2018). Adolescencia como factor de riesgo para complicaciones maternas y neonatales. *Revista Chilena de Obstetricia y Ginecología*, 83(5), 478–486. https://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0717-75262018000500478
- Pupo Jimenez, J. M., Gonzáles Aguilar, J. C., Cabrera Lavernia, J. O., & Cedaño Tarncón, O. (2019). Factores de riesgo de morbilidad materna extrema en la unidad de cuidados intensivos. *Multimed*, 23(2), 280–296. <https://doi.org/10.25100/cm.v50i4.4197>
- Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine Learning in Medicine. *New England Journal of Medicine*, 380(14), 1347–1358. <https://doi.org/10.1056/nejmra1814259>
- Ruth, D., Rengel, E. M., Suarez, R., Ana, D., Rojas Rodríguez, L., Yadira, D., Cueva, G., Elizabeth, R., & Rengel, M. (2020). Una mirada distinta al embarazo adolescente tfm. *Recimundo.Com*, 228–235. [https://doi.org/10.26820/recimundo/4.\(4\).noviembre.2020.228-235](https://doi.org/10.26820/recimundo/4.(4).noviembre.2020.228-235)
- Wickham, M. H., Chang, W., & Wickham, M. H. (2016). Package “ggplot2”. Create elegant data visualizations using the grammar of graphics. Version, 2(1), 1–89.