



Modelado predictivo de la empleabilidad digital en Ecuador con datos de empleo 2022-2025

Predictive modeling of digital employability in Ecuador with employment data 2022-2025

Ely Stefanie Parra Tauriz¹, Jesús Rafael Hechavarría Hernández², Dayron Rumbaut Rangel³

¹Universidad Bolivariana del Ecuador , Durán Guayas Ecuador , esparrat@ube.edu.ec

²Universidad Bolivariana del Ecuador , Durán Guayas Ecuador jrhechavarria@ube.edu.ec

³Universidad Bolivariana del Ecuador , Durán Guayas Ecuador drumbautr@ube.edu.ec

Resumen. La transformación digital ha cambiado por completo el mundo laboral, donde las habilidades digitales no son opcionales: quien quiera estar en el mercado necesita aprenderlas, dominarlas y utilizarlas de manera estratégica para el mercado laboral. En este estudio se analiza cómo está la empleabilidad digital en Ecuador, con una muestra de 55.118 registros tomados de la Encuesta Nacional de Empleo, Desempleo y Subempleo 2022 -2025, se creó un índice que mide tres dimensiones claves; como el acceso a tecnología, la aplicación de estas herramientas en el trabajo y qué tanto se dominan las habilidades digitales. Clasificando a las personas ocupadas en tres grupos de empleabilidad digital: alta, media y baja. Para entender mejor los perfiles, se utilizó técnicas de ML. El análisis no supervisado mostró dos clusters que están diferenciados por la edad y el nivel de educación. El XGBoost fue el más preciso (0.7791), donde el nivel educativo, el sector económico y el área donde vive cada persona son las variables más importantes para predecir el nivel de empleabilidad digital de un individuo. Con base en los datos muestrales, se identificó que únicamente el 16,8% de la población ocupada analizada pertenecía al grupo con alta empleabilidad digital, mientras que la mayoría (69,5%) se concentraba en el nivel más bajo. Este estudio permite construir un insumo valioso para la acción gubernamental, proporcionando una base para el diseño de políticas públicas enfocadas en reducir la brecha digital y priorizar la equidad y la sostenibilidad en la transformación digital.

Palabras clave: Brecha Digital, Empleo, Mercado De Trabajo, Aprendizaje, Gestión.

Abstract. Digital transformation has fundamentally reshaped the labor market, where digital skills are no longer optional. To participate effectively, individuals must learn, master, and strategically utilize these skills. This study analyzes the state of digital employability in Ecuador. Using a sample of 55,118 records from the National Employment, Unemployment, and Underemployment Survey (ENEMDU) for the period 2022-2025, we constructed an index measuring three key dimensions: access to technology, the application of digital tools in the workplace, and proficiency in digital skills. This index was used to classify the employed population into three groups: high, medium, and low digital employability. To better understand the profiles within these groups, we employed Machine Learning (ML) techniques. Unsupervised analysis revealed two distinct clusters, primarily differentiated by age and educational level. In supervised learning, the XGBoost algorithm demonstrated the highest predictive accuracy (0.7791), identifying educational attainment, economic sector, and geographic area as the most important variables for predicting an individual's level of digital employability. Based on the sample data, the analysis found that only 16.8% of the employed population studied belonged to the high digital employability group, while the majority (69.5%) were concentrated at the lowest level. This study provides a valuable foundation for governmental action, offering an evidence-based rationale for designing public policies aimed at reducing the digital divide and prioritizing equity and sustainability in the nation's digital transformation.

Key words: Digital Divide, Employment, Labour Market, Learning, Management.

1. Introducción

La transformación digital ha redefinido globalmente los requisitos de competencias laborales, estableciendo la empleabilidad digital como un factor crítico para la inserción efectiva en los mercados de trabajo contemporáneos. La transformación tecnológica en América Latina evidencia profundas desigualdades estructurales, reflejadas en brechas digitales persistentes entre distintos grupos sociales, regiones y sectores económicos [19]. Ecuador enfrenta así el desafío estratégico de alinear sus políticas laborales y educativas con la economía digital, haciendo del fortalecimiento de habilidades digitales, el acceso a infraestructura tecnológica y la adaptación organizacional imperativos para la competitividad nacional [3].

Las técnicas de Machine Learning y la analítica predictiva permiten analizar el mercado laboral con base empírica y orienta al futuro. Con estas herramientas, es posible descubrir patrones y comportamientos digitales dentro de la población ocupada, lo que ayuda a medir de forma objetiva aspectos de la empleabilidad que, hasta ahora, eran difíciles de captar [1]. Cuando se integran algoritmos de clasificación y predicción de bibliotecas como scikit-learn y XGBoost, se pueden crear modelos capaces de anticipar perfiles de empleabilidad digital con un nivel de precisión muy alto [5-16].

El objetivo principal de este estudio es desarrollar un modelo predictivo de empleabilidad digital para el mercado laboral del Ecuador, aplicando técnicas de ML a los datos de la Encuesta Nacional de Empleo, Desempleo y Subempleo (ENEMDU) dentro del periodo 2022-2025. Se implementan los algoritmos K-means para análisis no supervisado y Random Forest, XGBoost y Regresión Logística para el aprendizaje supervisado, con el fin de establecer la relación entre variables sociodemográficas y el nivel de empleabilidad digital.

La presente investigación supera las limitaciones de estudios previos en la región, como el de [10] en el análisis laboral ecuatoriano, mediante un enfoque multidimensional que integra acceso tecnológico, aplicación laboral y habilidades avanzadas. Adicionalmente, estudios sectoriales en Ecuador han documentado disparidades tecnológicas significativas que afectan la empleabilidad digital. [6] identificaron brechas críticas en la adopción de Big Data y Cloud Computing entre empresas ecuatorianas, donde solo el 28% utilizaba analytics de datos para la toma de decisiones. Esta brecha se amplía al considerar que el Instituto Nacional de Estadística y Censos (2024) reportó diferencias pronunciadas entre grandes empresas (45% adopción tecnológica) y microempresas (12%), mientras la adopción de comercio electrónico variaba del 15% en retail al 3% en agricultura. Esta transformación digital heterogénea configura demandas diferenciadas de habilidades digitales en el mercado laboral ecuatoriano, ampliando los marcos conceptuales tradicionales [20]. Este abordaje representa una contribución significativa al transformar el análisis tradicional de brecha digital, centrado en indicadores de acceso básico, hacia una evaluación comprehensiva que considera las capacidades efectivas de utilización de herramientas digitales en entornos laborales específicos.

El desafío principal consiste en gestionar eficientemente la información laboral e identificar factores críticos que afectan la empleabilidad digital en la población ecuatoriana, con un impacto potencial en la competitividad nacional y el desarrollo de políticas de capacitación efectivas. Este artículo se compone de cinco secciones: Materiales y métodos, presentación de los resultados, discusión de implicaciones y conclusiones, buscando aportar una base sólida para la toma de decisiones en el contexto de la empleabilidad digital en el Ecuador.

2. Materiales y Métodos

La investigación se desarrolló bajo un enfoque cuantitativo, no experimental, de alcance descriptivo-predictivo y con temporalidad transversal. El objetivo principal fue construir un índice de empleabilidad digital y, con base en él, predecir los perfiles del mercado laboral mediante modelos de machine learning.

Los datos de la Encuesta Nacional de Empleo, Desempleo y Subempleo (ENEMDU) de los años 2022, 2023, 2024 y 2025 componen la base para este estudio, siendo una encuesta de hogares, realizada por el Instituto Nacional de Estadística y Censos (INEC), posee un diseño muestral probabilístico, estratificado y polietápico, que garantiza su representatividad a nivel nacional. La estratificación se realiza por dominios geográficos (provincias y



área urbano-rural), mientras que el diseño polietápico asegura una selección aleatoria y sistemática de viviendas dentro de estos estratos.

Un aspecto metodológico clave a considerar es que el módulo de Tecnologías de la Información y Comunicación (TIC) se administra solo en el mes de julio de cada año dentro de la ENEMDU. Esta característica, lejos de ser una limitante, consolida una serie temporal consistente y comparable. Al analizar el mismo período anualmente, se controlan variaciones estacionales que podrían afectar los indicadores, como los patrones de empleo temporales, asegurando que los cambios observados en el índice de empleabilidad digital reflejen tendencias subyacentes y no fluctuaciones coyunturales.

Esta rigurosa metodología permite realizar inferencias estadísticas robustas sobre la población ecuatoriana y analizar la evolución del mercado laboral con consistencia metodológica durante el período de estudio. Los microdatos anonimizados utilizados en este estudio proceden de la base de personas del módulo de Tecnologías de la Información y Comunicación (TIC). La población de interés comprendió a personas de 15 años o más que declararon estar con empleo al momento de la encuesta (empleo =1). Tras la aplicación de los criterios de inclusión y exclusión, se consolidó una muestra final de 55,118 observaciones para el análisis. La muestra presentó una distribución por sexo del 55.9% hombres y 44.1% mujeres, con una edad media de 42.88 años (DE = 15.55). La distribución por área fue 71.4% urbana y 28.6% rural.

El procesamiento y análisis de datos se realizó en Python 3.10, implementando librerías para la manipulación y limpieza de datos con pandas y NumPy, mientras que el análisis visual empleó matplotlib y Seaborn, para el ML se utilizó scikit-learn para preprocesamiento, clustering y modelos base, complementado con XGBoost para el clasificador de mayor rendimiento. El proceso incluyó desde la transformación de características hasta la evaluación de modelos con validación cruzada (Pedregosa et al., 2011; [5]). La construcción del Índice de Empleabilidad Digital (IED) se basó en tres dimensiones binarias mutuamente excluyentes, diseñadas a partir de variables del módulo TIC de la ENEMDU, según el marco conceptual de [20]. La Dimensión 1 (Acceso a tecnología) consideró el uso de computadora de escritorio o portátil y la conexión diaria a internet. La Dimensión 2 (Aplicación laboral) se cumplió cuando el individuo reportó el uso de internet para fines de trabajo. La Dimensión 3 (Habilidades avanzadas) se determinó cuando el encuestado declaró realizar al menos dos de siete actividades TIC especializadas: enviar correos con archivos adjuntos, utilizar fórmulas aritméticas, conectar e instalar dispositivos, buscar, descargar e instalar software, elaborar presentaciones, transferir archivos entre dispositivos y escribir programas informáticos.

El índice global de cada individuo se obtuvo sumando las tres dimensiones, conforme a la expresión: $IED_i = A_i + L_i + H_i$; donde $IED_i \in \{0,1,2,3\}$ representa el puntaje total. A partir de este valor, se definieron tres perfiles de empleabilidad digital: Alta (cumplimiento de las tres dimensiones, $IED=3$), Media (cumplimiento de dos dimensiones, $IED=2$) y Baja (cumplimiento de una o ninguna dimensión, $IED=0-1$)

El análisis adoptó técnicas descriptivas, no supervisadas y predictivas. La parte descriptiva permitió caracterizar la muestra y analizar la distribución del Índice de Empleabilidad Digital (IED), sentando las bases para los posteriores análisis multivariantes. Para identificar patrones latentes en los datos, se aplicó un análisis no supervisado mediante K-means clustering (Pedregosa et al., 2011), utilizando como variables de entrada los indicadores sociodemográficos y el IED. Para determinar la estructura natural de los datos en cuanto a empleabilidad digital, se aplicaron dos técnicas complementarias: el método del codo y el análisis del coeficiente de silueta, que convergieron en marcar un número óptimo de dos clusters.

Posteriormente, se implementó modelado predictivo supervisado, comparando los algoritmos Random Forest, XGBoost y Regresión Logística Multinomial para predecir el perfil de empleabilidad digital. El conjunto de datos se dividió en 70% para entrenamiento y 30% para prueba, utilizando una validación cruzada de 5 divisiones (5-fold cross-validation) para la evaluación de los modelos. La optimización de hiperparámetros se realizó mediante búsqueda aleatoria, seleccionando el modelo con mayor precisión promedio. La evaluación final incluyó matriz de confusión, reporte de clasificación y análisis de importancia de variables

La investigación cumplió con los estándares éticos para el manejo de información. Este estudio utiliza datos secundarios de la Encuesta Nacional de Empleo, Desempleo y Subempleo (ENEMDU), que por diseño metodológico garantiza el anonimato de los participantes mediante la eliminación de identificadores personales, siguiendo protocolos estrictos de confidencialidad estadística. Su procesamiento se realizó exclusivamente con fines



académicos y de análisis estadístico agregado, respetando los principios de confidencialidad y no exposición a riesgo para los participantes de la encuesta.

3. Resultados

De los 55.118 trabajadores analizados (2022-2025), la distribución de empleabilidad digital, revela una profunda desigualdad en empleabilidad digital. Como muestra la Figura 1, solo 1 de cada 6 trabajadores (16.8%) posee habilidades digitales avanzadas, mientras que 7 de cada 10 (69.5%) se encuentran en el nivel más bajo, exponiendo una brecha digital crítica en el mercado laboral.

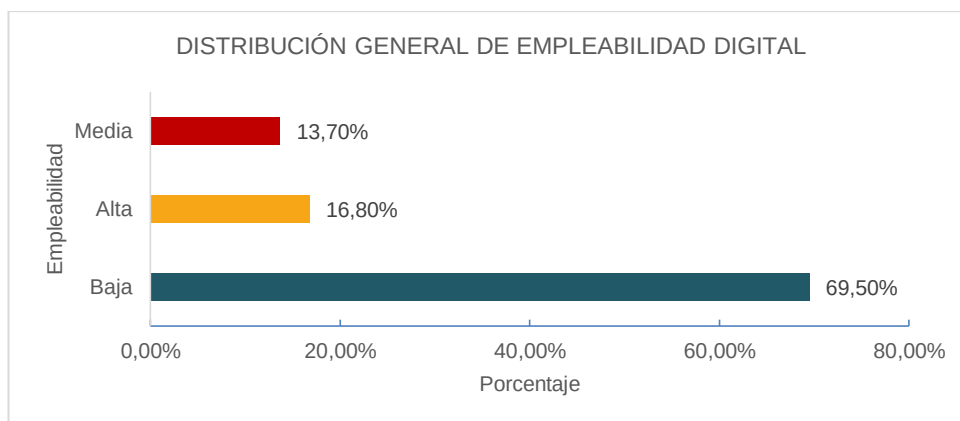


Figura 1: Distribución general de empleabilidad digital en la población ocupada ecuatoriana (2022-2025)

Los datos muestran que la empleabilidad digital no avanzó de forma constante: subió a 17.9% en 2023, bajó a 15.6% en 2024 y volvió a crecer a 17.0% en 2025. Estos altibajos sugieren que las habilidades digitales responden a factores externos, ver figura 2.

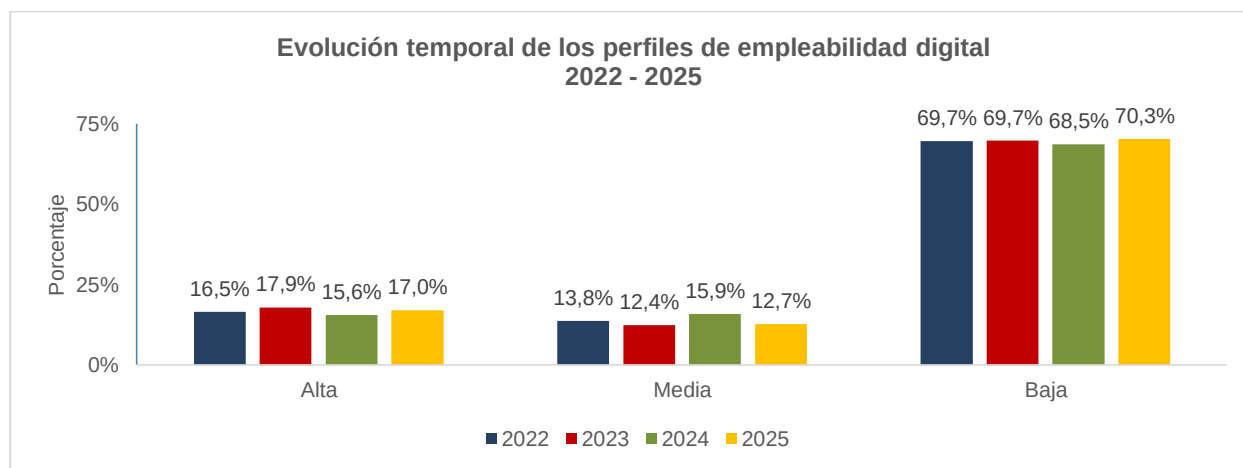


Figura 2: Evolución temporal de los perfiles de empleabilidad digital (2022-2025)

Los resultados evidencian tres patrones de desigualdad en empleabilidad digital. En género, las mujeres presentan ventaja (18.9% vs 15.0% en alta empleabilidad). Territorialmente, la brecha es más pronunciada: las zonas urbanas concentran 21.2% de empleabilidad alta frente a solo 5.6% en rurales. Sectorialmente, mientras Servicios alcanza 28.8% en empleabilidad alta, Agricultura muestra 1.4%, configurando un panorama de exclusión digital que segmenta al mercado laboral de Ecuador. Generacionalmente, en adultos jóvenes (26-35 años) alcanza a 25.4% en empleabilidad alta y decae progresivamente hasta niveles críticos en adultos mayores (65+ años) a 3.4%.

Cabe destacar que el nivel educativo resultó ser el factor más influyente, las personas con educación superior profesional presentan 46.5% de empleabilidad alta, mientras que quienes no tienen educación formal apenas alcanzan 0.1%, una diferencia de 46.4 puntos, ver figura 3.

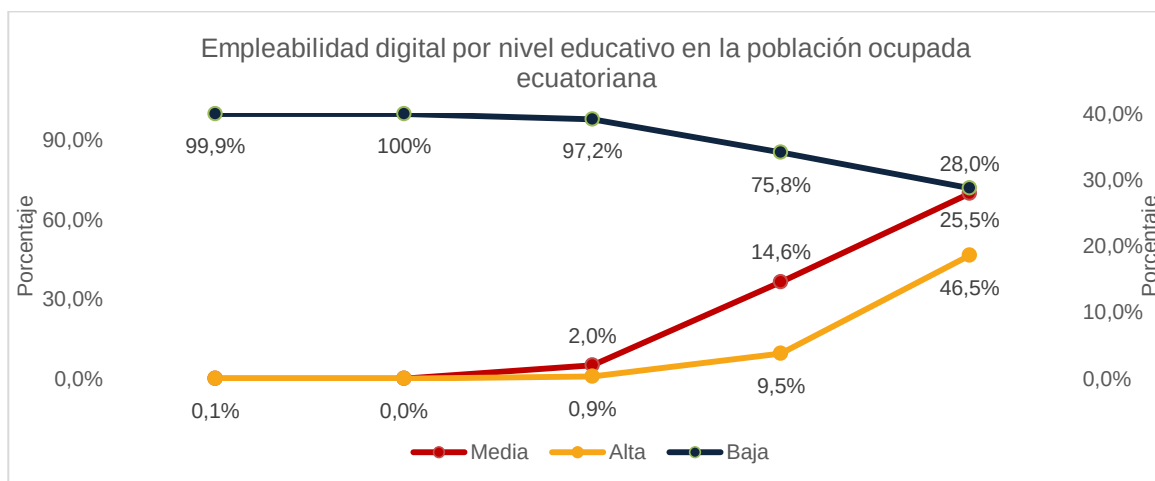


Figura 3: Empleabilidad digital por nivel educativo en la población ocupada ecuatoriana

En cuanto a la distribución sectorial existe profundas desigualdades: el sector servicios lidera con 28.8% de empleabilidad alta, superando en 27.4 puntos al sector agrícola, que apenas alcanza 1.4%. Esta brecha se replica a nivel territorial, donde las zonas urbanas (21.2%) cuadruplican a las rurales (5.6%), con una diferencia de 15.7 puntos, ver figura 4.

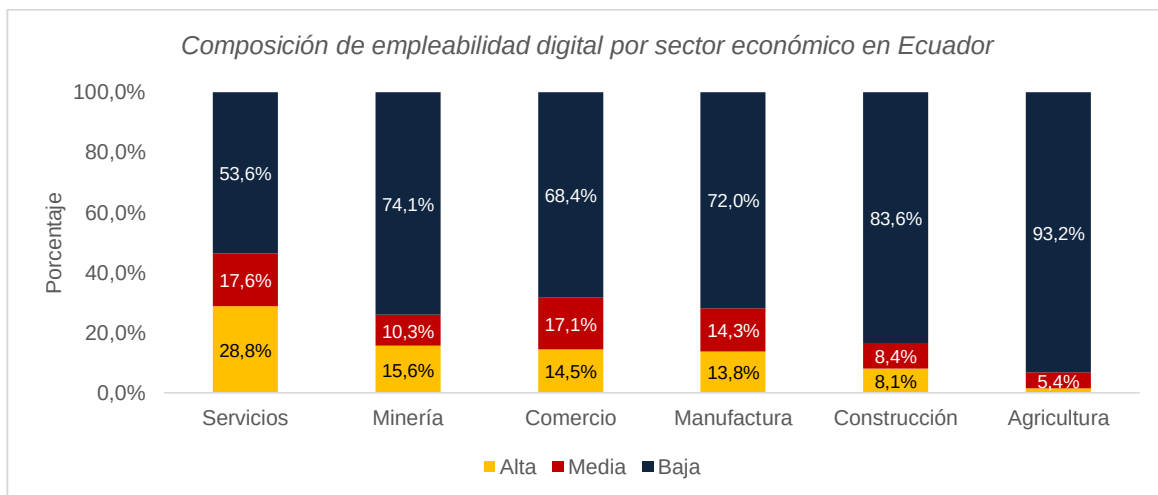


Figura 4: Composición de empleabilidad digital por sector económico en Ecuador

El análisis de clustering no supervisado entificó dos perfiles claramente diferenciados en la fuerza laboral. El primer clustering (55% de los trabajadores) reúne a población joven (31 años en promedio) con alta educación, donde 1 de cada 5 presenta empleabilidad digital avanzada. El segundo clustering (45%) concentra a trabajadores mayores (57 años en promedio) con menor formación, donde solo 1 de cada 10 alcanza alta empleabilidad digital. Esta segmentación sugiere que la brecha digital está profundamente asociada a la edad y el nivel educativo, ver figura 5.

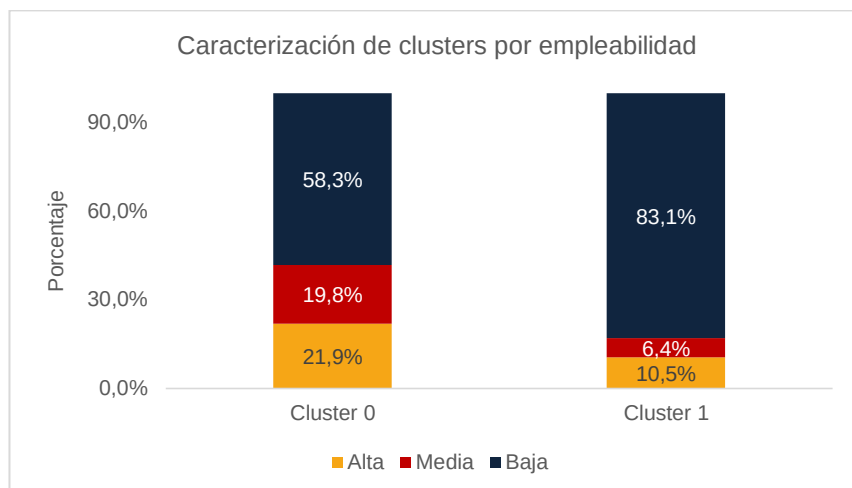


Figura 5: Caracterización de clusters por empleabilidad digital identificados mediante análisis no supervisado

Para predecir los perfiles de empleabilidad digital, se compararon los algoritmos Random Forest, Regresión Logística Multinomial y XGBoost. Este último demostró el mejor rendimiento, confirmándose como el modelo óptimo. Con el fin de maximizar su capacidad predictiva, se sometió a un proceso de optimización de hiperparámetros mediante búsqueda aleatoria y validación cruzada. La configuración óptima encontrada fue: `learning_rate: 0.1`, `max_depth: 3`, `n_estimators: 200` y `subsample: 0.8`.

El modelo XGBoost optimizado se evaluó sobre un conjunto de prueba de 16,536 observaciones, logrando una precisión global (accuracy) del 76.97%. Esta métrica indica que el modelo clasifica correctamente el perfil de empleabilidad digital en aproximadamente 3 de cada 4 casos, lo que constituye un nivel de acierto sólido para un problema de clasificación multiclase con tres categorías desbalanceadas.

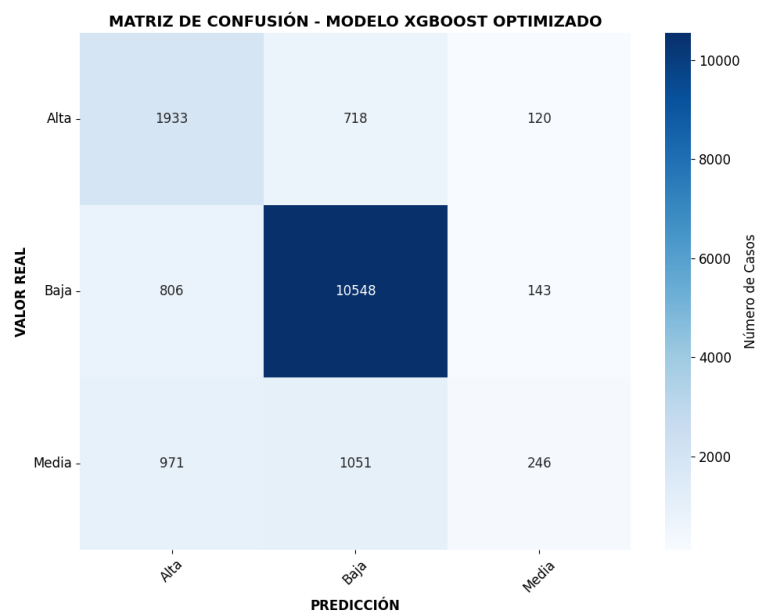
Un análisis detallado de las métricas por clase revela el desempeño diferencial del modelo, ver tabla 1:

- Empleabilidad Baja: El modelo muestra un excelente rendimiento para la clase mayoritaria, con una precisión del 0.86 y un recall del 0.92. Esto significa que es muy efectivo para identificar correctamente a los individuos con baja empleabilidad digital.
- Empleabilidad Alta: El desempeño es moderado, con un recall (0.70) mayor que la precisión (0.52). Esto indica que el modelo es razonablemente bueno para capturar a la mayoría de los individuos con alta empleabilidad (solo se le escapan 3 de cada 10), pero a costa de un número considerable de falsos positivos (clasifica como "Alta" a muchos que no lo son).
- Empleabilidad Media: Esta es la clase más difícil de predecir para el modelo, con una precisión del 0.48 y un recall muy bajo (0.11). El modelo struggle para identificar correctamente a este grupo, confundiéndolo frecuentemente con las clases Alta y Baja.

Tabla 1: Reporte de clasificación del modelo XGBoost optimizado.

Clase	Precisión	Recall	Puntuación F1	Casos de Soporte
Alta	0.52	0.70	0.60	2,771
Baja	0.86	0.92	0.89	11,497
Media	0.48	0.11	0.18	2,268
<i>Accuracy</i>			0.77	16,536
<i>Macro Avg</i>	0.62	0.57	0.55	16,536
<i>Weighted Avg</i>	0.75	0.77	0.74	16,536

La matriz de confusión permite visualizar estos patrones de error, ver figura 6. Los principales errores de clasificación son: 1,051 casos de la clase *Media* clasificados como *Baja* (6.36% del total de predicciones) y 971 casos de *Media* clasificados como *Alta* (5.87%). Asimismo, se observa una confusión bidireccional entre *Alta* y *Baja* (718 y 806 casos, respectivamente). Estos hallazgos sugieren que el modelo encuentra difícil distinguir los perfiles intermedios y, en menor medida, confunde los extremos del espectro de empleabilidad.

**Figura 6:** Matriz de confusión del modelo XGBoost optimizado.

Finalmente, para comprender los factores que impulsan las predicciones, se analizó la importancia de las variables en el modelo XGBoost, identificó que la educación es con gran diferencia, el principal determinante de empleabilidad digital (88.45% de importancia). Los demás factores analizados sector económico (6.61%), zona geográfica (2.49%), edad (1.70%) y género (0.65%) presentaron una influencia notablemente menor en la capacidad predictiva del modelo, ver figura 7

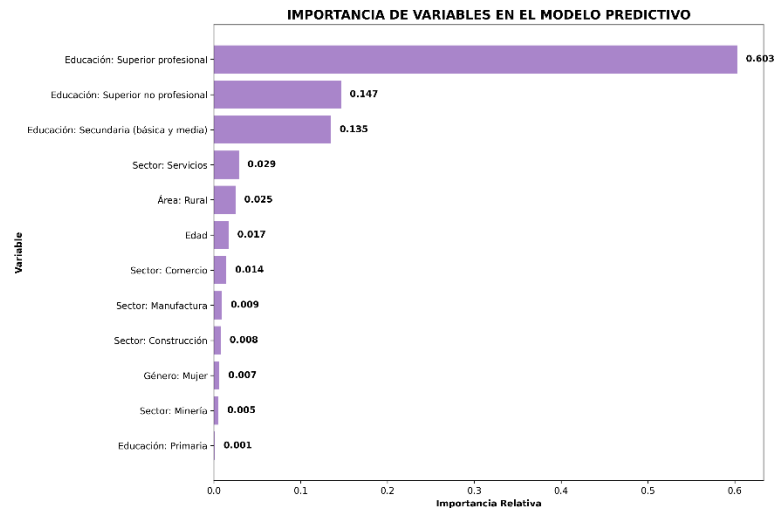


Figura 7: Importancia de variables en el modelo predictivo de empleabilidad digital

Respecto a las relaciones estadísticas, el análisis de correlaciones de Pearson confirmó relaciones significativas: nivel educativo (0.566, moderada positiva), sector económico (0.338, moderada positiva), edad (-0.192, débil negativa), área geográfica (-0.237, débil negativa) y género (0.064, muy débil positiva). Asimismo, se identificaron correlaciones importantes entre predictores como área geográfica y sector económico (-0.526) y nivel educativo con sector económico (0.417), ver figura 8.

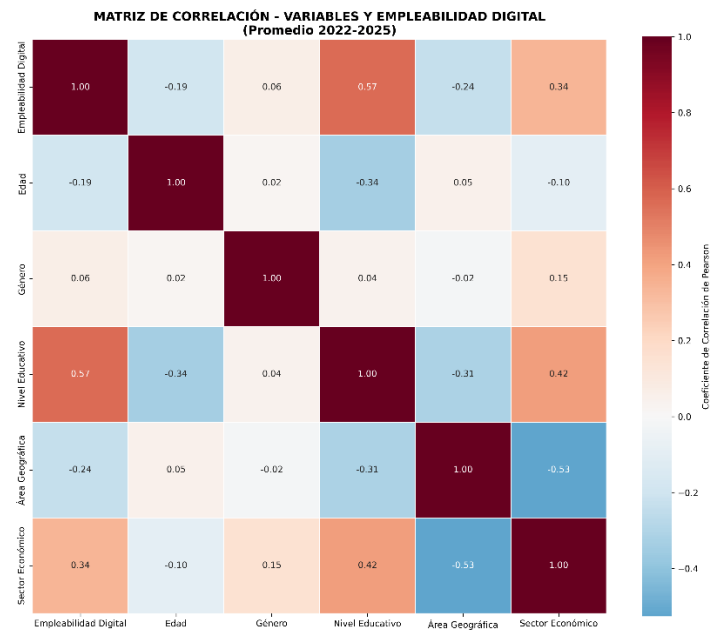


Figura 8: Correlación de variables sociodemográficas con la empleabilidad digital

Finalmente, en cuanto a las competencias específicas, la distribución de habilidades avanzadas mostró que el 67.4% de la población reportó cero habilidades, mientras solo el 1.6% dominó las siete habilidades evaluadas, evidenciando una concentración en competencias básicas. En relación al acceso tecnológico, los patrones mostraron:

internet diario (65.1%), computadora portátil (24.8%) y computadora de escritorio (22.3%), con solo el 33.4% cumpliendo la combinación requerida para la dimensión de acceso.

4. Discusión

El diagnóstico de empleabilidad digital en Ecuador revela un escenario crítico: apenas el 16.8% de la población empleada presenta un perfil competitivo para los empleos digitales, mientras que el 69.5% se encuentra en situación de desventaja. Estos hallazgos coinciden con las advertencias de [19] sobre las profundas desigualdades estructurales en América Latina y con el Banco Interamericano de Desarrollo (2025) respecto a los desafíos regionales en la transición digital. El modelo predictivo desarrollado corrobora esta segmentación, demostrando una alta capacidad para identificar el perfil mayoritario de baja empleabilidad (precisión del 86%, recall del 92%), pero revelando simultáneamente dificultades para clasificar los perfiles intermedios, lo que sugiere la existencia de transiciones graduales y solapamientos en las competencias digitales de la población.

Los resultados obtenidos confirman lo señalado por [20] respecto al papel determinante de la educación en el desarrollo de competencias digitales. En el caso ecuatoriano, coincidiendo con [10], este estudio evidencia que la educación superior profesional actúa como un divisor crítico. El análisis de importancia de variables en el modelo XGBoost refuerza cuantitativamente esta observación, donde el nivel educativo emerge como el predictor abrumadoramente dominante (88.45% de importancia relativa). Este hallazgo valida estadísticamente que el sistema educativo actual opera como el principal mecanismo de estratificación digital, abriendo las puertas a empleos digitales avanzados para quienes acceden a educación superior, mientras marca una frontera infranqueable para quienes carecen de ella, profundizando así la brecha de oportunidades en la economía digital.

Cuando se analiza el género, se muestra un patrón contrastante que difiere de estudios previos, sugiriendo un posible cambio en las dinámicas de apropiación tecnológica por género en Ecuador. Sin embargo, la persistencia de brechas sectoriales y la muy débil influencia del género en el modelo predictivo (0.65% de importancia) indican que los factores estructurales siguen siendo más determinantes que las variables individuales [4]. Esta mínima importancia relativa en el modelo sugiere que, si bien existen diferencias descriptivas en la distribución por género, estas no constituyen factores explicativos significativos cuando se controla por el efecto de otras variables estructurales.

La segmentación por edad evidencia problemas de inclusión digital entre generaciones [4]. El modelo predictivo corrobora esta brecha, aunque con una influencia moderada (1.70% de importancia), superior al género pero muy inferior a la educación. La diferencia de casi 6 años entre los promedios de edad de los perfiles alto y bajo, combinada con la dificultad del modelo para clasificar correctamente a los trabajadores de mayor edad, sugiere la existencia de una brecha generacional digital que requiere políticas diferenciadas y programas de capacitación específicos para trabajadores seniors.

Los resultados sectoriales confirman lo señalado por [21]: la transformación digital no llega a todos por igual. El modelo predictivo identifica al sector económico como el segundo factor en importancia (6.61%), validando empíricamente las profundas disparidades observadas. Mientras el sector servicios avanza con un 28.8% de empleabilidad alta, la agricultura se queda en un 1.4%, mostrando una brecha que refleja una doble vulnerabilidad: no solo es el sector con menor preparación digital, sino que coincide con la población rural, quedando atrapada en un círculo de exclusión tecnológica y territorial. Esta doble vulnerabilidad se manifiesta en el modelo a través de la confusión entre categorías, donde trabajadores del sector agrícola con empleabilidad media son frecuentemente clasificados en la categoría baja.

El desempeño del modelo XGBoost (76.97% de precisión) demuestra la utilidad del machine learning para predecir la empleabilidad digital. La arquitectura optimizada del modelo, con una profundidad limitada de árboles ($\text{max_depth}=3$) y una tasa de aprendizaje conservadora (0.1), sugiere que las relaciones entre variables son reconocibles y estables, aunque con cierta complejidad que requiere múltiples iteraciones (200 estimadores). El análisis de las métricas por clase revela un patrón característico de problemas con datos desbalanceados: mientras el modelo muestra excelente desempeño para la clase mayoritaria (baja empleabilidad), presenta desafíos significativos para identificar la clase media, con un recall críticamente bajo (11%) que indica la dificultad para capturar las transiciones y matices en el desarrollo de competencias digitales.



Es importante reconocer que los resultados están sujetos a limitaciones metodológicas. Por un lado, el uso de datos declarados conlleva el riesgo de que los participantes presenten una imagen favorable de sus habilidades, lo que podría sobrestimar los niveles reales de empleabilidad digital. Por otro lado, el enfoque en variables estructurales dejó fuera elementos psicosociales relevantes, como la autoeficacia o las actitudes hacia la tecnología, que seguramente aportarían matices importantes a la comprensión del fenómeno. Las confusiones observadas en la matriz de confusión, particularmente entre las categorías alta y media, podrían reflejar precisamente la ausencia de estas dimensiones psicosociales en el modelo.

Desde una perspectiva epistémica más amplia, la dificultad del modelo para clasificar la categoría 'Media' (Recall 0.11) sugiere que la empleabilidad digital no se comporta estrictamente bajo la lógica clásica binaria, sino que presenta características propias de la **Lógica Neutrosófica**. Mientras que la empleabilidad 'Alta' y 'Baja' pueden asociarse a los componentes de Verdad (T) y Falsedad (F) respectivamente, el perfil 'Medio' representa un alto grado de **Indeterminación (I)** [18-17].

Los individuos en esta franja poseen competencias fragmentadas (e.g., saben usar redes sociales pero no hojas de cálculo), lo que genera una ambigüedad que los algoritmos tradicionales de *crisp logic* (como XGBoost estándar) fuerzan a una clasificación única, perdiendo matices. Una aproximación futura mediante **conjuntos neutrosóficos de valor único (SVNS)** (Smarandache 2018) permitiría modelar este grupo no como una clase fallida, sino como un subconjunto donde la función de pertenencia a la indeterminación es dominante ($I > T, F$), ofreciendo una representación matemática más fiel de la transición digital incompleta.

5. Conclusiones

La investigación identificó obstáculos críticos en la transición hacia el trabajo digital en Ecuador. Los hallazgos mostraron que factores como la educación, la edad, la ubicación geográfica y el sector económico actuaron como filtros que restringen el acceso a las habilidades necesarias para los empleos del futuro. Los hallazgos demostraron que la mayoría de los trabajadores ecuatorianos se encuentran en situación de desventaja digital, con una distribución marcadamente polarizada donde solo una minoría ha logrado desarrollar las competencias necesarias para insertarse efectivamente en la economía digital.

En cuanto a la educación se mostró como el elemento clave en la formación de las brechas digitales, señalando que el grado de formación es el principal factor que predice la empleabilidad en el ámbito digital. La dramática diferencia entre los extremos del espectro educativo -desde 0.1% en personas sin educación hasta 46.6% en profesionales universitarios evidencia que el sistema educativo actual no está generando las competencias digitales necesarias para la mayoría de la población.

Las brechas territoriales emergieron como otro hallazgo crucial, mostrando una profunda división entre áreas urbanas (21.2% empleabilidad alta) y rurales (5.3%), que refleja desigualdades estructurales en infraestructura, oportunidades económicas y acceso a servicios educativos. El análisis generacional reveló un patrón preocupante de exclusión digital que afecta particularmente a los trabajadores mayores de 55 años (12.3% empleabilidad alta versus 25.4% en jóvenes de 25-34 años).

La polarización sectorial detectada evidencio que los beneficios de la digitalización se están distribuidos de manera desigual, con sectores tradicionales como la agricultura (1.4% empleabilidad alta) mostrando niveles críticos de rezago digital, mientras servicios profesionales alcanzan 61.4%. El modelo predictivo desarrollado validó la efectividad de las técnicas de machine learning para el análisis de empleabilidad digital, mientras que la identificación de clusters naturales sugiere la necesidad de políticas diferenciadas.

Para las políticas públicas, los resultados señalaron la necesidad de acciones que conecten la formación en competencias digitales, el acceso a infraestructura tecnológica y las oportunidades del desarrollo económico. Se recomienda implementar un programa de transformación digital en el que los sectores económicos tradicionales como la agricultura, que combine subsidios del 50% para tecnologías básicas, capacitación sectorial e incentivos fiscales; establecer un sistema nacional de certificación de competencias digitales con meta de 500,000 trabajadores certificados en 5 años; y ejecutar una estrategia de inclusión digital rural que expanda conectividad a 1,000 comunidades, cree centros comunitarios digitales e impulse programas de agricultura de precisión para 50,000 pequeños agricultores. Estas acciones coordinadas entre sector público, privado y académico son esenciales para



construir una fuerza laboral preparada para los requerimientos de la economía digital, evitando que las brechas existentes se traduzcan en desigualdades económicas y sociales permanentes.

Para futuras investigaciones, se recomienda incorporar métodos mixtos que combinen el análisis cuantitativo con aproximaciones cualitativas para profundizar en las barreras culturales y organizacionales identificadas. Igualmente, sería beneficioso llevar a cabo investigaciones longitudinales que faciliten el análisis de las trayectorias relacionadas con la empleabilidad digital, así como la evaluación de la influencia de políticas públicas concretas. La incorporación de variables psicosociales y la aplicación de técnicas de balanceo de datos podrían mejorar significativamente la capacidad predictiva para las clases minoritarias.

Los resultados evidencian la necesidad de políticas que aborden de manera conjunta los espacios de educación y sector económico, priorizando a la población en riesgo de exclusión digital. La capacidad predictiva de los modelos empleados, junto con el análisis.

6. Referencias bibliográficas.

- [1]. Álvarez Calvo, R. F. (2025). La IA y el Machine Learning aplicada a la ingeniería industrial a favor de la mejora en la gestión de operaciones. *Rhombus*, 5(Especial), 1-20. <https://doi.org/10.63058/rhombus.v5iEspecial.343>
- [2]. Angrave, D., Charlwood, A., Kirkpatrick, I., Lawrence, M., & Stuart, M. (2016). HR and analytics: Why HR is set to fail the big data challenge. *Human Resource Management Journal*, 26(1), 1-11. <https://doi.org/10.1111/1748-8583.12090>
- [3]. Banco Interamericano de Desarrollo. (2025). Habilidades para el trabajo en América Latina y el Caribe: impulsar el talento para un futuro sostenible y equitativo. <http://dx.doi.org/10.18235/0013593>
- [4]. Bercovich, N. & Muñoz, M. (2022). Rutas y desafíos para cerrar las brechas de género en materia de habilidades digitales. Comisión Económica para América Latina y el Caribe (CEPAL). <https://www.cepal.org/es/publicaciones/47939-rutas-desafios-cerrar-brechas-genero-materia-habilidades-digitales>
- [5]. Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785-794. <https://doi.org/10.1145/2939672.2939785>
- [6]. Coello Yagual, R. R., & Parrales Ponce, J. D. (2020). Análisis de las ventajas y desventajas del Big Data y el Cloud Computing en el proceso de la toma de decisiones de las empresas que practican comercio electrónico. **Revista Científica Ciencia y Tecnología**, *20*(25), 45-62. <https://doi.org/10.47189/RCCT.V20I25.279>
- [7]. Comisión Económica para América Latina y el Caribe (CEPAL). (2020). Panorama social de América Latina 2020. <https://repositorio.cepal.org/entities/publication/61652648-a144-413a-bab4-1eae639b3ecf>
- [8]. Del Carpio-Mendoza, R. (2024). Predicción del rendimiento académico utilizando modelos de aprendizaje automático: Una revisión sistemática de la literatura. *593 Digital Publisher CEIT*, 9(6), 1038-1054. <https://doi.org/10.33386/593dp.2024.6.2797>
- [9]. Durán Bernardino, M. (2021). Digitalización y empleo: retos del futuro del trabajo desde una perspectiva de género. *Revista de Estudios Jurídicos*, (21), e6761. <https://doi.org/10.17561/rej.n21.6761>
- [10]. Flores Merchán, L. A. (2023). Análisis del empleo adecuado en Ecuador durante el 2023, a través de la integración de modelos econométricos y algoritmos de machine learning (Tesis de pregrado). Universidad de las Américas, Ecuador. <https://dspace.udla.edu.ec/handle/33000/16797>
- [11]. Instituto Nacional de Estadística y Censos. (2024). El Mercado Laboral Ecuatoriano en la Era Digital. Cuadernos de Trabajo - Estadísticas Sociales. https://www.ecuadorencifras.gob.ec/documentos/web-iniec/Bibliotecas/Libros/cuadernos_trabajo/Estadisticas_vacantes_ML_WS.pdf
- [12]. Instituto Nacional de Estadística y Censos. (2025). Tecnologías de la Información y Comunicación - TIC. <https://www.ecuadorencifras.gob.ec/tecnologias-de-la-informacion-y-comunicacion-tic/>
- [13]. Marler, J. H., & Boudreau, J. W. (2017). An evidence-based review of HR analytics. *The International Journal of Human Resource Management*, 28(1), 3-26. <https://doi.org/10.1080/09585192.2016.1244699>
- [14]. Mina, A. (2020). Big data e inteligencia artificial en el futuro manejo de pacientes. ¿Por dónde empezar? ¿En qué punto nos encontramos? *Avances en medicina de laboratorio*, 1(3), 20200052. <https://doi.org/10.1515/almed-2020-0052>



- [15]. OCDE. (2019). Measuring the Digital Transformation: A Roadmap for the Future. OECD Publishing. <https://doi.org/10.1787/9789264311992-en>
- [16]. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12, 2825-2830. <https://jmlr.org/papers/volume12/pedregosa11a/pedregosa11a.pdf>
- [17]. Smarandache, F. (2018). Operadores con conjunto neutrosóficos de valor único Oversets, Undersets y Offset. *Neutrosophic Computing and Machine Learning. ISSN 2574-1101*, 4(4), 03-07.
- [18]. Smarandache, F. (2025). Indeterminación en las Teorías Neutrosóficas y sus Aplicaciones. *Neutrosophic Computing and Machine Learning. ISSN 2574-1101*, 39, 1-7.
- [19]. UNESCO. (2022). Fortaleciendo las sociedades del conocimiento: Marco para la acción. Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura. <https://unesdoc.unesco.org/ark:/48223/pf0000381995>
- [20]. Van Laar, E., Van Deursen, A. J. A. M., Van Dijk, J. A. G. M., & De Haan, J. (2017). The relation between 21st-century skills and digital skills: A systematic literature review. *Computers in Human Behavior*, 72, 577-588. <https://doi.org/10.1016/j.chb.2017.03.010>
- [21]. Varela Bohórquez, F. (2025). El futuro del trabajo en la transformación digital: El reskilling y upskilling como procesos necesarios para satisfacer las nuevas demandas de empleo. *Revista de Economía Digital*, 10(2), 78-95. <http://hdl.handle.net/10017/65545>

