

Uso de aplicación web para la evaluación temprana de conocimientos sobre las TIC's en los egresados del nivel medio del distrito de Zimatlán de Álvarez, Oaxaca

DOI: 10.58299/UTP.261.C877

Autores

Moisés Adrián Hernández Luis

Tecnológico Nacional de México

Instituto Tecnológico de Oaxaca

mahl107@itoaxaca.edu.mx

<https://orcid.org/0009-0004-1583-2923>

Armando Levid Rodríguez Santiago

Universidad Tecnológica

de los Valles Centrales de Oaxaca

levid.rodriguez@gmail.com

<https://orcid.org/0000-0001-5510-9394>



Uso de aplicación web para la evaluación temprana de conocimientos sobre las TIC's en los egresados del nivel medio del distrito de Zimatlán de Álvarez, Oaxaca

Use of a web application for the early assessment of ICT knowledge in high school graduates in the district of Zimatlán de Álvarez, Oaxaca

Resumen

Este estudio presenta el desarrollo de una aplicación web diseñada para evaluar los factores que influyen en el aprendizaje informático de egresados del nivel medio superior y predecir su posible desempeño académico universitario. La plataforma integra tecnologías modernas como MongoDB, Express, React, Python y Flask, mediante las cuales se recolectan datos a través de un formulario estructurado. La información obtenida se procesa con el algoritmo supervisado Random Forest para identificar patrones asociados al riesgo académico. Los resultados muestran que el modelo es funcional para generar predicciones iniciales, aportando un indicador temprano del nivel de preparación del estudiante. La herramienta constituye un apoyo docente para la detección oportuna de casos que requieren acompañamiento académico. Se concluye que la información recolectada es adecuada para alimentar modelos predictivos y que futuras ampliaciones podrían incorporar más variables y otros algoritmos para mejorar su precisión.

Palabras clave: algoritmo; aprendizaje; educación; inteligencia; software.

Abstract

This study presents the development of a web application designed to evaluate factors that influence computer literacy among upper-secondary graduates and to predict their potential academic performance in university-level technology programs. The platform integrates modern technologies such as MongoDB, Express, React, Python, and Flask to collect data through a structured questionnaire. The information gathered is processed using the supervised Random Forest algorithm to identify patterns associated with academic risk. Results indicate that the model is functional for generating preliminary predictions, providing an early indicator of each student's level of preparedness. The tool serves as a valuable support for instructors by enabling the early detection of students who may require academic assistance. The study concludes that the collected data are suitable for predictive modeling and that future work could incorporate additional variables and alternative algorithms to improve accuracy and broaden the application's analytical capabilities.

Keywords: algorithm; education; intelligence; learning; software.

Introducción

El acelerado avance de las tecnologías de la información y la creciente demanda de competencias digitales en el ámbito educativo y profesional han puesto en evidencia la necesidad de comprender, evaluar y fortalecer el nivel de aprendizaje informático adquirido por los estudiantes del nivel medio superior. En el caso particular del estado de Oaxaca, y específicamente del distrito de Zimatlán de Álvarez, los subsistemas de educación media superior presentan condiciones heterogéneas en infraestructura, habilidades docentes y recursos tecnológicos, factores que influyen directamente en la apropiación del conocimiento y en el desempeño académico posterior de los estudiantes que continúan su formación en áreas relacionadas con las tecnologías de la información.

Diversos estudios en educación, cognición y aprendizaje tecnológico han demostrado que las experiencias previas, el acceso a dispositivos, la calidad del acompañamiento docente y las percepciones individuales sobre la informática determinan tanto la construcción de representaciones mentales como la capacidad de aplicación práctica del conocimiento. Asimismo, con la expansión de metodologías pedagógicas orientadas al aprendizaje autónomo apoyado en tecnología, se vuelve imprescindible identificar con mayor precisión los factores que predisponen al éxito o rezago académico en asignaturas de corte informático.

En este contexto, el uso de técnicas de aprendizaje automático ha emergido como una estrategia eficaz para el análisis de datos educativos y la predicción temprana de patrones de riesgo académico. Los modelos supervisados, como Random Forest, permiten procesar grandes volúmenes de información, descubrir tendencias no triviales y generar estimaciones que contribuyen a la toma de decisiones docentes fundamentadas. No obstante, la aplicación de estos modelos en escenarios educativos locales, particularmente en comunidades con desigualdad tecnológica, ha sido limitada.

El presente estudio tiene como propósito desarrollar una aplicación web capaz de recopilar, organizar y analizar información relevante sobre el nivel de conocimientos informáticos de los egresados del nivel medio superior, con el fin de predecir de manera temprana posibles dificultades académicas al ingresar a programas universitarios relacionados con la informática. Mediante la integración de tecnologías web contemporáneas (MongoDB, Express, React, Python y Flask) y la implementación del algoritmo predictivo Random Forest, se busca generar un sistema funcional que ofrezca diagnósticos automatizados y facilite a los docentes la identificación de estudiantes en riesgo, contribuyendo así a estrategias de acompañamiento académico más oportunas y efectivas.

Este proyecto se fundamenta en la hipótesis de que el análisis sistemático de los datos recopilados mediante el formulario permitirá descubrir patrones significativos asociados al aprendizaje informático previo, y que dichos patrones podrán emplearse para anticipar el desempeño futuro de los estudiantes. La aplicación desarrollada no pretende sustituir la complejidad del fenómeno educativo ni abarcar variables psicosociales o personales no cuantificables; sin embargo, sí busca constituirse como una herramienta de valor para la orientación docente y para la mejora continua de la enseñanza de la informática en contextos con limitaciones tecnológicas.

Revisión teórica

Teoría cognitiva de la representación mental

EL cerebro construye estados cognitivos o de conocimiento a través de modelos: primero análogos, lineales, figuras, bordes, etc. Después pasan a un modelo lingüístico que transforma figuras en palabras con valor referencial que cumplen una función de interpretación y describen una realidad (UAEH, 2020).

La representación se da cuando esas imágenes externas que captamos en el ambiente, los transformamos en modelos, que nuestra mente a partir del lenguaje que se utiliza para representar ese objeto se toma la descripción y el significado. En

el lenguaje van inmersos unos conceptos que son: medio, contenido y relación representacional (Moreira, 2002).

Dentro de los cuales la “ semejanza ” es igual a la imagen que representa a un objeto; lo que llama la atención en realidad es el conjunto de propiedades de ese objeto en particular. La causalidad que explica todo lo que aprendemos es por una causa que el objeto provoca por una particular atención que se le pone y se obtiene de él información que se convierte en un conocimiento, esto por sus características y propiedades (Moreira, 2002).

Todo esto lleva a realizar una función representacional, la cual deberá cumplir dentro de nuestro complejo sistema neuronal una función que deberá ser interpretada por nuestro cerebro, almacenada y traducida en una conducta.

En este caso el conocimiento sería el resultado del funcionamiento de una intrincada red de millones de neuronas, comunicadas entre sí por sus dendritas y que van creando caminos y decisiones.

La representación se hallaría distribuida a través de esas redes de unidades y conexiones y los pulsos eléctricos; así pues, la mente, el conocimiento, la percepción y actitud participan como parte fundamental de la comprensión cognitiva de las personas; son operaciones neuronales cuya funcionalidad radica en dar a la persona la capacidad de adaptación (UAEH, 2020). En un sentido estricto, es la imagen de una cosa (incluyendo personas) o proceso en la conciencia humana independiente de la presencia corpórea de lo representado, que queda almacenado en la memoria de un individuo, a partir de una percepción previa.

Tendencia pedagógica tecnológica

En las tendencias pedagógicas son formas del trabajo que se ha de desarrollar en el proceso de enseñanza-aprendizaje, así como el papel que se le asigna tanto al educando como al educador y se plantean así mismo las repercusiones que éstas hayan podido tener en la práctica pedagógica (Bonafé, 1996).

Las tendencias pedagógicas contemporáneas responden al reclamo social de una formación que les permita a los sujetos resolver problemas de diferente índole de forma autónoma y responsable (Bonafé, 1996).

Esto significa, poder enfrentar la búsqueda de soluciones, encontrar una respuesta y tener algún control sobre ésta; los problemas que se presentan implican encontrar respuestas nuevas a preguntas también nuevas.

La tendencia educativa tecnológica es parte de la evolución y constante modernización por la que pasa la humanidad: ya que encamina a los aprendizajes escolares al dominio y manejo de las tecnologías como herramientas fundamentales que acompañan a los estudiantes en las áreas informáticas modernas basadas en tecnología nube, blockchain, entre otros (Digital House, 2024).

Insta a que el maestro enseñe al alumno a aprender por sí mismo. Entre sus tareas debe enseñar a los estudiantes el valor de lo aprendido y ayudarles a relacionar el nuevo aprendizaje en los esquemas conceptuales con los que vive su vida. Se busca que el profesor, se centre en la adaptación de las necesidades educativas actuales.

Cada alumno puede avanzar a su ritmo, lo que le permite desarrollar el conocimiento en el momento. Libera al profesor y al alumno de las aburridas exposiciones de contenido en las que el aprendizaje no puede asegurarse.

Estamos viviendo en la cuarta revolución industrial, donde la mercancía con mayor valor es la información. Esta era está produciendo un grupo de personas que adquieren conocimiento, mediante el uso de la mente más que en el uso del músculo y el conocimiento se consagra con herramientas como la inteligencia artificial, el internet de las cosas y el big data (Digital House, 2024).

Modelo de aprendizaje aplicable a la informática

Uno de los modelos de aprendizaje como lo es el “Procesamiento de la Información”, sería una excelente forma de enseñanza para la informática, pues esta se asemeja al funcionamiento de una computadora; entrada de datos, procesamiento de datos y salida de información (Biblioteca ITAM, 2010).

Este modelo de aprendizaje en donde los elementos más importantes de explicación son las instancias internas (en este caso la operación general de una computadora) tendría la actuación “dinámica” del profesor para que este diseñe la estimulación al alumnado y estos expongan sus respuestas a temas informáticos (Biblioteca ITAM, 2010).

Un aporte personal sería aplicar la metodología de enseñanza conductual dentro de la informática; pues algunos conocimientos son abstractos y se necesita mucha repetición para memorizarlos y después aplicarlos algún caso de uso.

Un ejemplo claro de ello es la materia de la programación, que es una parte de la informática donde el código que se desarrolla sólo funciona integrando las variables que podemos decir que actúan como estímulos que ocasionan o sirven para dar una respuesta correcta o incorrecta.

Son varias críticas y desventajas que se le atribuyen a esta teoría, si bien es cierto Watson no puntualizó muchas ideas que quedaron inconclusas: las aportaciones que esta dan sobre todo a la educación son de importante estudio y reflexión. Pero dada la situación el docente de informática es el único que sabe hasta donde se puede aplicar esta teoría y que consecuencias se tendrán (De la mora, 1990).

La educación general y específica (como lo es la informática) dentro de la escuela es parte de un adiestramiento para el ser humano, donde se le enseña a utilizar primeramente el lenguaje innato que se utiliza dentro de nuestra cultura o sociedad; para lograr así una comunicación con los demás; así como el conocer a los objetos (computadora) que este necesitará emplearlos de tal forma que lo ayuden a solucionar problemas reales de su vida diaria.

Ahora bien, un licenciado o ingeniero en informática (que por lo regular son los que imparten las materias de informática) que no cuenta con imaginación e inventiva, sin conocimientos adecuados o sin amor por lo que hace, echará a perder las aspiraciones de los estudiantes que realmente quieren aprender a usar correctamente la herramienta.

El profesor debe fomentar la creatividad de los estudiantes; deben jugar con la tecnología de punta, que el maestro-alumno realice investigaciones juntos, todo

esto para encontrar nuevas soluciones y estimular la invención de nuevos dispositivos o aplicaciones.

En informática los programas educativos se vuelven obsoletos en muy corto tiempo, por lo que los cursos deben diseñarse de tal manera que puedan incorporar los avances conceptuales y tecnológicos que existen en la actualidad y que en otros países ya ocupan desde hace un buen tiempo.

El aprendizaje automático: machine learning

El aprendizaje automático es una rama de las ciencias computacionales que ha adquirido gran auge en los últimos años sobre todo después de la pandemia mundial del covid19.

Podemos definir el concepto de machine learning de acuerdo con Espinosa (2020), como el campo de estudio que permite que las computadoras tengan la habilidad de aprender mediante la aplicación de algoritmos que simulan los procesos del pensamiento humano.

Este concepto abarca una amplia gama de técnicas que se clasifican en dos grandes tipos:

- Aprendizaje supervisado, donde se tiene una variable objetivo y se entrena a un programa informático con un conjunto de datos para aplicar el resultado en un nuevo conjunto de datos.
- Aprendizaje no supervisado, en el cual no se tiene variable objetivo y el programa informático debe encontrar patrones y relaciones en un conjunto de datos.

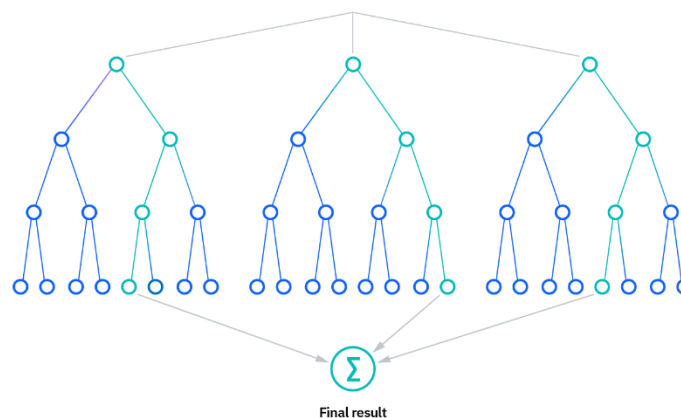
Para este proyecto, se utilizó el algoritmo Random Forest que se muestra en la Figura 1. Este algoritmo es una técnica de aprendizaje supervisado que genera múltiples árboles de decisión sobre un conjunto de datos de entrenamiento: los resultados obtenidos se combinan a fin de obtener un modelo único más robusto en comparación con los resultados de cada árbol por separado (Espinosa, 2020).

Los árboles de decisión son representaciones gráficas, de las posibles soluciones a una decisión basadas en ciertas condiciones que son evaluadas

jerárquicamente, los árboles de decisión, es de los algoritmos de aprendizaje supervisado más utilizados en machine learning (Na8, 2020).

Para obtener el árbol óptimo, hay que determinar entre todos los árboles posibles desde el nodo raíz y los subsiguientes; medir y valorar de alguna manera las predicciones logradas para después compararlas todas y obtener la mejor. La división de nodos continuará hasta que lleguemos a la profundidad máxima posible del árbol ó se limiten los nodos a una cantidad mínima de muestras en cada hoja (Na8, 2020).

Figura 1. Múltiples árboles de decisiones Random Forest



Fuente: IBM, 2020.

De acuerdo a Espinosa (2020), cada árbol se obtiene mediante un proceso de dos etapas:

- Se genera un número considerable de árboles de decisión con el conjunto de datos. Cada árbol contiene un subconjunto aleatorio de variables m (predictores) de forma que $m < M$ (donde M = total de predictores).
- Cada árbol crece hasta su máxima extensión. Cada árbol generado por el algoritmo Random Forest contiene un grupo de observaciones aleatorias elegidas mediante bootstrap, que es una técnica estadística para obtener muestras de una población donde una observación se puede considerar en más de una muestra.

Las observaciones no estimadas en los árboles, también conocidas como “out of the bag”, se utilizan para validar el modelo. Un modelo predictivo es un sistema o herramienta que utiliza datos históricos y técnicas estadísticas para

identificar patrones y predecir resultados o comportamientos futuros (Espinosa 2020).

Sistemas educativos medio superior en el distrito de Zimatlán de Álvarez, Oaxaca

A través de la Subsecretaría de Educación Media Superior (SEMS), encargada de coordinar el trabajo y las actividades de las instituciones educativas que brindan el plan de estudios del Bachillerato General; en Oaxaca coordina los siguientes subsistemas que tienen planteles en las poblaciones pertenecientes al distrito de Zimatlán de Álvarez, Oaxaca.

- Centro de Bachillerato Tecnológico Industrial y de Servicios
Los Bachilleratos Tecnológicos son un subsistema de Educación media Superior que permiten a los estudiantes cursar su bachillerato a la par de una carrera profesional y técnica; planteles en el distrito de Zimatlán de Álvarez: CBTIS 248 - Zimatlán de Álvarez.
- Colegio de Estudios Científicos y Tecnológicos del Estado
Los Bachilleratos Tecnológicos son un subsistema de Educación media Superior creado en 1996 como respuesta al programa nacional para la modernización educativa 1990-1994; planteles en el distrito de Zimatlán de Álvarez: CECYTE 37 - San Gabriel Mixtepec y CECYTE 25 - San Pablo Huixtepec.
- Instituto de Estudios de Bachillerato del Estado de Oaxaca
Subsistema estatal derivado del anterior Telebachillerato creado en el sexenio del Sr José Murat Casab exgobernador de Oaxaca; planteles en el distrito de Zimatlán de Álvarez: IEBO 17 - San Pedro el Alto, IEBO 16 - Santa Gertrudis, IEBO 171 - Santa Ana Tlapacoya y el IEBO 112 – La Ciénega.

Cada uno de estos sistemas y subsistemas educativos medio superior en el estado están enfocados al trabajo productivo, dentro de sus currículos se integran materias de informática y sus egresados en su mayoría tienen la especialidad.

En este sentido, es importante mencionar que según datos de la encuesta nacional sobre disponibilidad y uso de tecnologías de la información en los hogares del INEGI, en 2020 el 43.8% de los hogares mexicanos contaba con una computadora, lo que equivale a 16.9 millones de viviendas. También, el 90.4% de los hogares tenía al menos un televisor, mientras que en el censo se registró una posesión del 88.8% en teléfonos celulares y 56.2% en internet (INEGI,2020). Esto permite deducir el avance significativo de las familias mexicanas por obtener y utilizar la tecnología disponible en el mercado para su uso personal, escolar y esparcimiento.

Metodología

Para desarrollar desde cero esta aplicación web se decidió utilizar la metodología ágil Kanban, debido a su facilidad de implementación y por su eficiente manera de actualizar los avances de las tareas del equipo de trabajo, así como conocer el estado del proyecto en todo tiempo de manera visual mediante un tablero que distribuye las actividades en tres columnas: pendiente, en progreso y terminado.

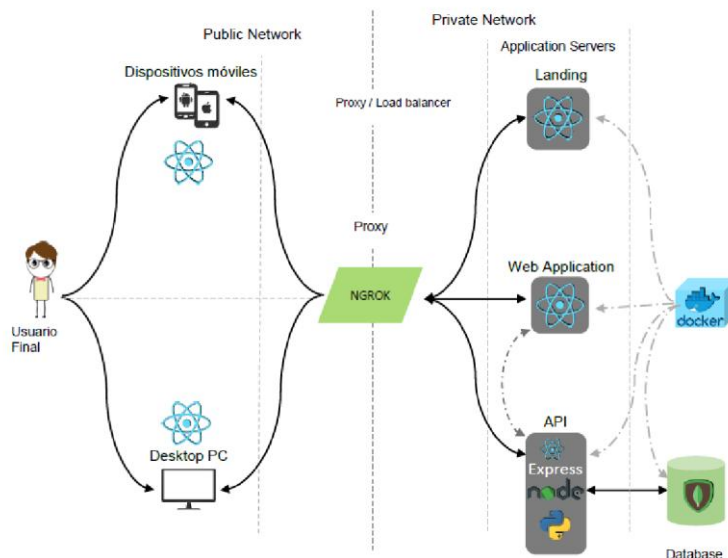
Para ello se definieron las tareas principales que conformarán este desarrollo de la aplicación web y que se detallan a continuación.

Diseño de aplicación web

Implementar una arquitectura sofisticada basada en tecnologías web actuales, que garantice el éxito del proyecto; por tal motivo utilizaremos Mongo como base de datos, en la parte del back end utilizaremos los frameworks Express y Flask.

Para la parte del front end utilizaremos la versatilidad de React y librerías JQuery; y todo esto listo para ejecutarse en un contenedor Docker y salir al internet mediante un enlace Ngrok, tal como se muestra en la figura 2.

Figura 2. Arquitectura de la aplicación web.



Fuente: elaboración propia.

Recolección de datos mediante formulario

PASO 1). El usuario ingresa a la plataforma por medio de un ordenador o dispositivo móvil con conexión a internet.

PASO 2) El usuario rellena el cuestionario con la información que se solicita para la recolección de datos, la encuesta no se podrá enviar hasta que rellene todos los campos, la encuesta se divide en cinco secciones:

- a) Datos generales del usuario
- b) Cuestionario sobre la infraestructura tecnológica del bachillerato de procedencia.
- c) Descripción de la materia de informática del bachillerato de procedencia.
- d) Evaluar el grado de conocimiento de sus antiguos docente.
- e) Evaluación de conocimiento técnico respecto a la carrera de TI.

PASO 3) El usuario validará su información contestando con veracidad cada pregunta, seleccionando un ítem de una escala de Likert de cinco puntos:

- a. Nunca b. Casi nunca c. Regular d. Casi siempre e. Siempre

PASO 4) El usuario envía la información dando clic en el botón enviar, sus datos serán enviados (vía método POST) al servidor y almacenados en una base de datos.

PASO 5) Una vez terminado de responder el cuestionario ya podrá abandonar la página y salir de la aplicación.

- Preparación del dataset para análisis

En la parte del back end, se construirán estructuras de datos unidimensionales (arrays), que contenga los datos de: nombre del alumno, población a la que pertenece, el sistema medio superior al que asistió; así como cada una de las respuestas de los ítems del formulario por secciones.

Cada arreglo de respuestas, tendrán la estructura siguiente:

Array0 = [nombre, población, sistema]

Array1 = [ítem1, ítem2, ítem3, ítem4, ítem5, ítem6, ítem7]

Array2 = [ítem8, ítem9, ítem10, ítem11, ítem12, ítem13, ítem14]

Array3 = [ítem15, ítem16, ítem17, ítem18, ítem19, ítem20, ítem21]

- Ejecución de algoritmo predictivo con Python

Los arreglos 1 al 3 serán introducidos al modelo Random Forest para su tratamiento, el modelo requiere obtener la media aritmética de cada arreglo para su ejecución.

- Presentación de resultados

Una vez entrenado el modelo, se podrá predecir el resultado mediante el siguiente pseudocódigo:

valor = [(ítem1, ítem2,... ítemn, (ítem1 + ítem2 +... ítemn)/n)]

resultado_predicho = función_prediccion(valor)

si (resultado_predicho[0] == 1)

imprimir "En Riesgo académico"

sino

imprimir "Aprueba/Mejora desempeño"

Resultados

La aplicación web fue desarrollada con la tecnología web en el lado de front end: React. Por la otra parte del back end: Ngrok, mismos que se ejecutaron en un ambiente Linux Ubuntu versión 22.04 LTS. Instalando el lenguaje Python en su versión 3.8 como un entorno de trabajo con la librería virtualenv.

Después de desplegar nuestra aplicación, el formulario se contestó con la participación de estudiantes pertenecientes al distrito de Zimatlán de nuevo ingreso al nivel superior en las carreras de tecnologías de la información e ingeniería en sistemas en el año 2023.

A continuación, se describe de manera general; el despliegue de la aplicación web y algunas de las pantallas principales de la aplicación web desarrollada.

- **BACKEND.**

Para ejecutar el backend , desde una terminal se ingresó a la carpeta backend/src/ y se ejecutó lo siguiente:

```
$ python app.py
```

Para llevar a cabo la ejecución de ngrok en una terminal nos dirigimos a la carpeta donde está guardado el archivo ngrok y ejecutamos lo siguiente:

```
$ ngrok http 5000
```

Esto nos genera un link que podemos compartir, el link estará activo por 1 hora.

- **FRONTEND**

Para ejecutar el frontend desde una terminal nos dirigimos a la carpeta frontend y una vez dentro de la carpeta ejecutamos lo siguiente:

```
$ npm start
```

- **LOGIN**

Aquí el usuario debe ingresar a la plataforma con su nombre de usuario y contraseña, mismos que fueron previamente registrados a la plataforma

mediante correo electrónico proporcionado, tal como se muestra en la figura 3.

Figura 3. Pantalla de LOGIN de la aplicación web.

Fuente: elaboración propia.

- **FORMULARIO**

En este apartado del usuario rellenará el cuestionario con su información que se solicita para la recolección de datos generales y de cada uno de los ítems solicitados, tal como se muestra en la figura 4.

Figura 4. Formulario para recabar información de los estudiantes.

Fuente: elaboración propia.

- **DASHBORAD**

Parte primordial de la aplicación web, donde se verán los datos de los estudiantes que se hayan registrado y la información vertida. Para poder ver los resultados del algoritmo Random Forest, se selecciona la casilla de

verificación de la derecha del estudiante, para después dar clic en el botón de abajo con la leyenda Generar Reporte, tal como se muestra en la figura 5.

Figura 5. Dashboard principal de la aplicación web

DATOS DE LOS ESTUDIANTES			EXIT	
Nombre	Email	Población	Acciones	
ARMANDO SANTIAGO RAMIREZ	rando1804@gmail.com	SAN ANTONINO CASTILLO	Borrar	Editar Ver <input type="checkbox"/>
JAVIER HERNANDEZ LOPEZ	javi0209@gmail.com	ZIMATLAN DE ALVAREZ	Borrar	Editar Ver <input type="checkbox"/>

GENERAR REPORTE

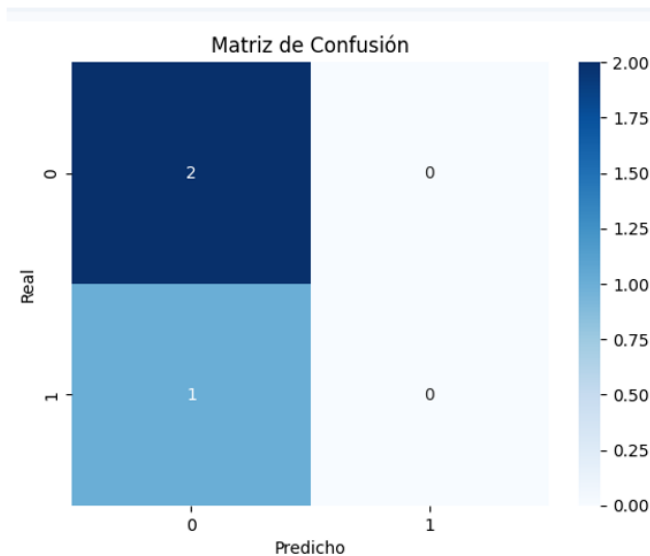
Fuente: elaboración propia.

- **EJECUCION DE ALGORITMO**

En esta matriz de resultado el modelo arrojó dos valores: un valor verdadero positivo y otro valor falso positivo; que se puede interpretar como adecuado para la predicción. Este resultado aparecerá en la parte de consola del lado del back end.

La matriz de confusión de la figura 6, indica que el valor verdadero positivo representado por el color azul fuerte se asemeja a los valores introducidos en los arreglos de respuesta tienden a ser mayores en comparación con el valor falso positivo.

Figura 6. Matriz de resultados.



Fuente: elaboración propia.

- **VENTANA EMERGENTE**

Esta ventana aparecerá del lado del front end, después de dar clic al botón Generar Reporte según sea el caso del estudiante elegido.

Figura 7. Ventana emergente con resultado



Fuente: elaboración propia.

Discusiones

Los resultados obtenidos en este estudio confirman el potencial del aprendizaje automático como herramienta complementaria para la evaluación educativa, particularmente en contextos donde existen brechas tecnológicas y diferencias en la formación previa de los estudiantes. La aplicación web desarrollada permitió integrar en un mismo entorno la recolección, procesamiento y análisis automatizado de datos, demostrando que es posible generar indicadores tempranos de riesgo académico basados en información autoinformada. Esta capacidad predictiva, aunque preliminar, coincide con investigaciones previas que señalan la utilidad de modelos supervisados —como Random Forest— para la identificación de patrones en datos educativos heterogéneos.

No obstante, los resultados también evidencian limitaciones importantes. El tamaño reducido de la muestra afecta la generalización de las predicciones y la estabilidad del modelo, lo cual explica la presencia de falsos positivos y la baja

sensibilidad observada. Asimismo, el formulario recolecta únicamente variables relacionadas con la infraestructura tecnológica, conocimientos previos y percepción de la enseñanza; sin embargo, no incorpora dimensiones psicosociales, motivacionales o contextuales que influyen significativamente en el desempeño académico y que otros estudios han señalado como determinantes críticos.

Otro aspecto relevante es que el modelo utilizado, si bien robusto y adecuado para un primer acercamiento, podría compararse con métodos alternativos como redes neuronales, regresión logística, análisis discriminante o técnicas basadas en boosting, con el fin de validar su superioridad o complementar sus predicciones. Asimismo, la actualización periódica del dataset y la inclusión de más ciclos escolares permitirían mejorar la precisión y reducir la varianza del algoritmo.

En conjunto, los hallazgos sugieren que la herramienta desarrollada es útil como sistema de apoyo para docentes y tutores, pero no debe asumirse como un predictor definitivo del rendimiento académico. Más bien, representa un primer paso hacia la integración de analítica educativa basada en datos, que puede fortalecerse mediante la ampliación del modelo, la incorporación de nuevas características y el análisis comparativo de distintos enfoques de aprendizaje automático. Con ello, se abre la posibilidad de generar estrategias de intervención más informadas y contextualizadas, particularmente en regiones con limitaciones tecnológicas como las observadas en el distrito de Zimatlán de Álvarez, Oaxaca.

Conclusión

Los resultados obtenidos muestran que, aun con un conjunto de datos reducido, el modelo fue capaz de identificar patrones relevantes en las respuestas de los estudiantes, proporcionando indicadores tempranos que pueden apoyar al docente en la toma de decisiones y en la implementación de estrategias de acompañamiento académico. Esto confirma la hipótesis de que la información recolectada es útil para alimentar un modelo predictivo y generar estimaciones iniciales sobre el posible desempeño académico en asignaturas informáticas.

Sin embargo, también se reconoce que la precisión del modelo está condicionada por el tamaño de la muestra, la variabilidad de los datos y la ausencia de variables psicosociales o contextuales que influyen de manera significativa en el rendimiento. Por ello, las predicciones deben interpretarse como un apoyo complementario y no como una determinación concluyente del desempeño académico futuro.

Finalmente, este trabajo constituye un punto de partida para la incorporación de analítica educativa en instituciones con limitaciones tecnológicas, mostrando que la integración de machine learning en procesos de diagnóstico académico es viable, escalable y potencialmente benéfica. Se recomienda ampliar la base de datos, incorporar nuevas características de evaluación y comparar diferentes algoritmos para optimizar la precisión predictiva y fortalecer el impacto pedagógico de la herramienta.

Referencias

- Biblioteca ITAM. (2010). *Reflexiones y recomendaciones sobre educación en informática I*, Estudios ITAM.
<https://estudios.itam.mx/sites/default/files/estudiositammx/files/039/000172755.pdf>
- Bonafé, J. (1996). "Tendencias Educativas Hoy" en *Cuadernos de Pedagogía* (253), 1-12.
- De la mora Ledesma, J. G. (1990). *Psicología del Aprendizaje I*. Editorial Progreso.
- Digital House. (2024). *Digital mindset* [Programa de formación en línea]. Digital House.
- Espinosa, J. (2020). Aplicación de algoritmos Random Forest y XGBoost en una base de solicitudes de tarjetas de crédito. *Ingeniería Investigación y Tecnología*. <https://doi.org/10.22201/fi.25940732e.2020.21.3.022>
- Instituto Nacional de Estadística y Geografía (2020). *Encuesta Nacional sobre Disponibilidad y uso de Tecnologías de la información en los Hogares* (ENDUTIH). <https://www.inegi.org.mx/temas/ticshogares/>
- International Business Machines Corporation (2020,1 enero). *¿Qué es el bosque aleatorio?* IBM Artificial Intelligence <https://www.ibm.com/mx-es/think/topics/random-forest>
- Moreira, M. A. (2002). *Modelos mentales y modelos conceptuales del aprendizaje de la ciencia*. Universidad de Burgos.
- Universidad Autónoma del Estado de Hidalgo. (2020). *Orientaciones para el trabajo en el aula: Representaciones mentales*. Universidad Autónoma del Estado de Hidalgo.
<https://www.uaeh.edu.mx/scige/boletin/prepa4/n8/e3.html>



Certificado de evaluación

La Editorial UTP, con Registro Nacional de Instituciones y Empresas Científicas y Tecnológicas, por la Secretaría de Ciencia, Humanidades, Tecnología e Innovación (SECIHTI) en México; **indexada en catálogos de calidad internacional**. Que, teniendo como **base fundamental el desarrollo del potencial humano**, es líder en el desarrollo y divulgación de producción científica, tecnológica y educativa con altos estándares de calidad en contextos locales, nacionales e internacionales, a través de publicaciones de artículos en revistas, libros, capítulos de libros, recursos educativos, conferencias y congresos.

CERTIFICA

Que el capítulo de libro titulado **“Uso de aplicación web para la evaluación temprana de conocimientos sobre las TIC’s en los egresados del nivel medio del distrito de Zimatlán de Álvarez, Oaxaca”** presentado por los autores **Moisés Adrián Hernández Luis y Armando Levid Rodríguez Santiago** ha sido sometido a un exhaustivo proceso de arbitraje por pares académicos, a través de criterios establecidos para investigaciones de alta calidad, siendo dictaminado como producto de investigación científica, tecnológica y/o educativa de alta calidad. Su publicación en el libro titulado **“Innovación y creatividad: claves para el desarrollo regional”** estará disponible a partir del 9 de diciembre de 2025 en la Biblioteca digital de la Universidad Tecnocientífica del Pacífico.

Se extiende el presente certificado, a los 10 días del mes de noviembre del año 2025.

Transformando con Ciencias
Tepic, Nayarit; México


Dra. Ana Luisa Estrada Esquivel
Directora de la Editorial UTP
Universidad Tecnocientífica del Pacífico


César Alejandro González Guzmán
Coordinador de la Editorial UTP
Universidad Tecnocientífica del Pacífico



Calle 20 de Noviembre, 75 Pte. Col. Mololoa. Tepic, Nayarit, México. C.P. 63050
<https://editorial-utp.com.mx>. Correo electrónico: editorial_utp@tecnocientifica.com.mx. Teléfono: 311 101 01 03