



**UNIVERSIDAD CATÓLICA  
DE SANTIAGO DE GUAYAQUIL  
FACULTAD DE INGENIERÍA  
CARRERA DE COMPUTACIÓN**

**TEMA:**

**Desarrollo de una plataforma interactiva de emprendimientos para la conexión entre emprendedores e inversionistas, aplicando modelos predictivos que permita ponderar el nivel de éxito que puede tener un emprendimiento. Caso de estudio: emprendimientos culinarios.**

**AUTORES:**

**Santacruz Moncayo, Itamar David  
Burgos Calle, Mario Enrique**

**Proyecto de tecnología de información previo a la obtención del título de INGENIERO EN  
CIENCIAS DE LA COMPUTACIÓN**

**TUTOR:**

**Ing. Erazo Ayón, José Miguel. Mgs.**

**Guayaquil, Ecuador  
2 de marzo de 2026**



**UNIVERSIDAD CATÓLICA  
DE SANTIAGO DE GUAYAQUIL  
FACULTAD DE INGENIERÍA  
CARRERA DE COMPUTACIÓN**

**CERTIFICACIÓN**

Certificamos que el presente proyecto de tecnología de información fue realizado en su totalidad por **Santacruz Moncayo, Itamar David** y **Burgos Calle, Mario Enrique**, como requerimiento para la obtención del título de Ingeniero en Sistemas Computacionales.

**TUTOR**

f. \_\_\_\_\_

**Ing. Erazo Ayón, José Miguel. Msg.**

**Guayaquil, a los 2 días del mes de marzo del año 2026**



**UNIVERSIDAD CATÓLICA  
DE SANTIAGO DE GUAYAQUIL  
FACULTAD DE INGENIERÍA  
CARRERA DE COMPUTACIÓN**

**DECLARACIÓN DE RESPONSABILIDAD**

**Yo, Santacruz Moncayo, Itamar David  
Burgos Calle, Mario Enrique**

**DECLARAMOS QUE:**

El presente proyecto de tecnología de información, **“Desarrollo de una plataforma interactiva de emprendimientos para la conexión entre emprendedores e inversionistas, aplicando modelos predictivos que permita ponderar el nivel de éxito que puede tener un emprendimiento. Caso de estudio: emprendimientos culinarios.”**, previo a la obtención del título de **Ingeniero en ciencias de la computación**, ha sido desarrollado respetando derechos intelectuales de terceros conforme las citas que constan en el documento, cuyas fuentes se incorporan en las referencias o bibliografías. Consecuentemente este trabajo es de nuestra total autoría.

En virtud de esta declaración, nos responsabilizamos del contenido, veracidad y alcance del Trabajo de Integración Curricular referido.

**Guayaquil, a los 2 días del mes de marzo del año 2026**

f. \_\_\_\_\_

f. \_\_\_\_\_

**Santacruz Moncayo, Itamar David**

**Burgos Calle, Mario Enrique**



**UNIVERSIDAD CATÓLICA  
DE SANTIAGO DE GUAYAQUIL  
FACULTAD DE INGENIERÍA  
CARRERA DE COMPUTACIÓN**

**AUTORIZACIÓN**

Nosotros, **Santacruz Moncayo, Itamar David**  
**Burgos Calle, Mario Enrique**

Autorizamos a la Universidad Católica de Santiago de Guayaquil a la publicación en la biblioteca de la institución del proyecto de tecnología de información, **“Desarrollo de una plataforma interactiva de emprendimientos para la conexión entre emprendedores e inversionistas, aplicando modelos predictivos que permita ponderar el nivel de éxito que puede tener un emprendimiento. Caso de estudio: emprendimientos culinarios.”** cuyo contenido, ideas y criterios son de nuestra exclusiva responsabilidad y total autoría

**Guayaquil, a los 2 días del mes de marzo del año 2026**

**LOS AUTORES**

f. \_\_\_\_\_

f. \_\_\_\_\_

**Santacruz Moncayo, Itamar David**

**Burgos Calle, Mario Enrique**



UNIVERSIDAD CATÓLICA  
DE SANTIAGO DE GUAYAQUIL  
FACULTAD DE INGENIERÍA  
CARRERA DE COMPUTACIÓN

REPORTE ANTIPLAGIO



Fecha de elaboración: 13 DE FEBRERO DEL 2026



Jose Miguel Erazo Ayon



Nombre del tutor y firma

Ing. José Miguel Erazo Ayón, Mgs.

## **AGRADECIMIENTO:**

Quisiera primeramente expresar mi más profundo agradecimiento a mi papa Walter Santacruz y a mi mama Yadira Moncayo, los pilares de mi vida. Gracias por su amor, por siempre apoyarme en cada decisión, por brindarme lo mejor para mí y por darme la mejor educación posible. Gracias por enseñarme a ser la persona que soy hoy en día y acompañarme a lo largo de este proyecto. A mis hermanos y hermanas por estar siempre presentes.

A mis amigos de facultad por los grandes momentos en estos 4 años de universidad. Sus conocimientos y apoyo mutuo son valores que me han permitido seguir adelante y llegar a este logro juntos.

A mis mejores amigos Nicolas Villena y Ricardo Soriano por su apoyo en los buenos y malos momentos desde el colegio y estar presentes en esta trayectoria.

A mi sobrino Raúl Santacruz por ser un segundo hermano para mí. Su presencia, alegría y apoyo a lo largo de este tiempo fue importante para estar pleno.

A mi pareja Ana Ávila por su constante motivación y apoyo emocional a lo largo de este proceso.

Itamar Santacruz

## AGRADECIMIENTO

Quisiera mandar un agradecimiento a mis padres ya que su constante apoyo, amor y afecto que me han demostrado a lo largo de los distintos aspectos de mi vida, y sé que seguiré recibiendo su amor y comprensión de aquí en adelante, me siento muy afortunado de ser su hijo.

A mis hermanos que están en casa que al igual que mis padres siempre han estado ahí como un par de figuras que yo considero super importante en mi vida y que tenerlos es un regalo que la vida me dio al igual que mis padres

A mis compañeros y amigos de la facultad, ya que su presencia ha representado un aspecto fundamental en lo que fue la carrera, su existencia y apoyo siempre será algo de lo que este agradecido y espero que este no sea el final de nuestro vinculo que se ha formado en este tiempo, ya que su amistad es algo invaluable.

A mis demás amigos que también han estado conmigo y me han apoyado, su existencia siempre es y será agradecida, pues su amistad es para mí algo invaluable.

Mario Burgos

## **DEDICATORIA:**

Este trabajo esta dedicado a nuestras familias, ya que han sido y serán siempre el pilar fundamental de nuestro crecimiento pasado y futuro que hemos tenido, y tendremos. A donde sea que lleguemos en la vida podremos decir con orgullo que nuestra familia estuvo ahí y que su influencia nos marcará hoy y siempre.

También está dedicado a la universidad, que ha sido una pieza importante en nuestra formación estudiantil en estos años que han pasado, y que nos ha enseñado valores que no nosotros consideramos invaluable. Nuestra transición como estudiantes universitarios puede terminar ahí pero nunca hay que olvidar que uno es estudiante toda la vida.

Itamar Santacruz y Mario Burgos.



**UNIVERSIDAD CATÓLICA  
DE SANTIAGO DE GUAYAQUIL  
FACULTAD DE INGENIERÍA  
CARRERA DE COMPUTACIÓN**

**TRIBUNAL DE SUSTENTACIÓN**

f. \_\_\_\_\_

**ING. ANA CAMACHO CORONEL, MGS**

**DIRECTORA DE CARRERA**

f. \_\_\_\_\_

**ING. GALO CORNEJO GOMEZ, MGS**

**DOCENTE DE LA CARRERA**

f. \_\_\_\_\_

**ING. BYRON YONG YONG, MGS**

**OPONENTE**

## INDICE GENERAL

|  |           |
|--|-----------|
| Resumen: .....   | XV        |
| Introducción: .....  | 2         |
| <b>CAPITULO 1</b> .....                                    | <b>4</b>  |
| Planteamiento del problema.....                            | 4         |
| Justificación del problema .....                           | 5         |
| Pregunta de investigación .....                            | 6         |
| Hipótesis .....  | 6         |
| Delimitación.....  | 6         |
| Objetivo General.....                                      | 7         |
| Objetivos Específicos.....                                 | 7         |
| Alcance .....  | 8         |
| <b>CAPITULO II. MARCO TEÓRICO</b> .....                    | <b>9</b>  |
| Marco referencial.....                                     | 9         |
| Marco Conceptual.....                                      | 12        |
| Marco Legal.....   | 51        |
| Ley Orgánica de Protección de Datos Personales.....        | 52        |
| Ley Orgánica de Emprendimiento e Innovación .....          | 53        |
| Registro Nacional de Emprendimiento .....                  | 53        |
| <b>Capitulo III. Metodología de la Investigación</b> ..... | <b>54</b> |
| Tipo de Investigación.....                                 | 54        |
| Enfoque Metodológico.....                                  | 55        |
| Metodología de Desarrollo .....                            | 56        |
| Población y Muestra .....                                  | 58        |
| Técnicas e Instrumentos de Recolección de Datos .....      | 59        |

|   |    |
|---|----|
| Instrumento de validación cualitativa: entrevista semiestructurada..... | 60 |
| Justificación del instrumento .....                                     | 61 |
| Criterios de selección del entrevistado.....                            | 61 |
| Diseño del instrumento .....  | 62 |
| Procedimiento de análisis de la información .....                       | 63 |
| Triangulación metodológica .....  | 64 |
| Capitulo IV. Propuesta Tecnológica .....                                | 65 |
| Objetivo de la Propuesta .....  | 65 |
| Arquitectura del Sistema.....   | 65 |
| Capa de presentación .....  | 67 |
| Capa de lógica de negocio .....   | 67 |
| Capa de acceso de datos.....  | 68 |
| Microservicio de predicción .....                                       | 70 |
| Sistema de Autenticación de perfiles .....                              | 73 |
| Sistema de Mensajería en Tiempo Real.....                               | 74 |
| Desarrollo del modelo predictivo.....                                   | 75 |
| Variables predictoras .....   | 79 |
| Preprocesamiento de Datos.....  | 80 |
| Selección y configuración del algoritmo .....                           | 81 |
| Calibración de probabilidades.....                                      | 82 |
| Entrenamiento y validación .....  | 83 |
| Integración como servicio.....  | 84 |
| Resultados del modelo predictivo .....                                  | 85 |
| Métricas de rendimiento .....   | 85 |
| Importancia de las variables.....                                       | 86 |
| Pruebas de Integración.....   | 88 |

|  |     |
|--|-----|
| Flujo Completo de Usuario.....   | 89  |
| Transparencia Algorítmica.....   | 90  |
| CAPITULO V. CONCLUSIONES Y RECOMENDACIONES .....                         | 96  |
| Conclusiones .....   | 96  |
| Cumplimiento de los Objetivos de Investigación .....                     | 96  |
| Relevancia para el Ecosistema Emprendedor Ecuatoriano .....              | 97  |
| Validación de la arquitectura tecnológica.....                           | 98  |
| Tabla de Trazabilidad .....  | 100 |
| Recomendaciones .....  | 101 |
| Referencias: .....   | 104 |
| Anexos .....   | 110 |
| Anexo 1. Manual de usuario de la plataforma .....                        | 110 |
| Funcionalidades por tipo de usuario (Emprendedor / Inversionista) .....  | 111 |
| Anexo 2. Transcripción de entrevista realizada al Ing. Blas Lorenzo..... | 115 |

## INDICE DE TABLAS

|  |     |
|--|-----|
| Tabla 1: Comparación entre diferentes frameworks de Backend .....            | 40  |
| Tabla 2: Ventajas, Desventajas y Características de herramientas usadas..... | 71  |
| Tabla 3: Distribución de Dataset generado .....                              | 77  |
| Tabla 4: Variables predictorias .....  | 80  |
| Tabla 5: Hiperparámetros del modelo XGBoost.....                             | 82  |
| Tabla 6: Predicciones para perfiles de prueba .....                          | 88  |
| Tabla 7: Resultados de prueba de integración .....                           | 89  |
| Tabla 8: Tabla de Trazabilidad .....   | 100 |

## INDICE DE ILUSTRACIONES

|   |    |
|---|----|
| Ilustración 1: Arquitectura del sistema .....                               | 66 |
| Ilustración 2: Diagrama Entidad-Relación de BDD .....                       | 69 |
| Ilustración 3: Flujo de Predicción .....                                    | 73 |
| Ilustración 4: Diagrama de flujo: Sistema de mensajería en tiempo real..... | 75 |
| Ilustración 5: Matriz de confusión .....                                    | 79 |
| Ilustración 6: Variables Predictoras .....                                  | 87 |

## **Resumen:**

Este trabajo de integración curricular presenta el desarrollo de una plataforma digital interactiva orientada a vincular emprendedores del ámbito culinario con posibles inversionistas en el sector Kennedy de Guayaquil. La herramienta incluye un modelo predictivo basado en inteligencia artificial que calcula el nivel de factibilidad y las posibilidades de éxito de los proyectos culinarios propuestos.

La investigación tiene como fundamento datos específicos del Global Entrepreneurship Monitor Ecuador 2023-2024 (Lasio et al., 2024): aunque la Tasa de Actividad Emprendedora Temprana de Ecuador es del 32,65 %, lo que lo ubica entre las naciones más emprendedoras de la zona, solamente cerca del 18 % de los negocios culinarios en Guayas consiguen financiación formal o inversión privada durante los primeros dos años de funcionamiento. Esta disparidad entre la gran cantidad de iniciativas que surgen y las limitadas oportunidades reales de captar recursos pone de relieve la necesidad urgente de soluciones tecnológicas que fomenten un contacto más objetivo y basado en fundamentos entre inversores y emprendedores.

La metodología utilizada es principalmente cuantitativa, aunque incluye algunas contribuciones cualitativas, y se categoriza como investigación tecnológica aplicada. Para desarrollar el modelo predictivo, se implementó el enfoque CRISP-DM. Se generaron 5,000 registros sintéticos para su entrenamiento, empleando distribuciones de probabilidad ajustadas a datos oficiales del SRI (2023), el INEC (2023) y los reportes GEM elaborados por ESPAE-ESPOL (Lasio et al., 2020, 2021). La ausencia de bases históricas completas y fiables en el ámbito local ha inducido el uso de datos sintéticos. Se eligió el algoritmo XGBoost para la categorización. Durante el desarrollo del modelo surgieron algunos ajustes necesarios. En particular, se detectó

que la variable “existencia de plan de negocios” estaba recibiendo un peso desproporcionado en las predicciones, lo que penalizaba injustamente a emprendedores que, aunque no contaban con ese documento formal, mostraban fortalezas claras en formación académica, validación real del mercado y experiencia previa en el sector.

Para corregir este sesgo se aplicó una calibración isotónica de las probabilidades de salida. Gracias a ese ajuste, el modelo final, evaluado mediante validación cruzada estratificada de 5 pliegues, alcanzó un F1-Score del 86 %, una precisión del 94 % y un ROC-AUC de 88,48 %.

La estructura de la plataforma fue creada usando microservicios. El elemento central fue elaborado con el conjunto tecnológico MERN (MongoDB, Express, React y Node .js), mientras que las proyecciones del modelo se ejecutan en un microservicio independiente codificado en Python con FastAPI .

Para evaluar el funcionamiento del sistema en la realidad, se llevó a cabo una entrevista semiestructurada con un inversionista que está activamente implicado en el ámbito gastronómico de Guayaquil. Su valoración cualitativa se alineó en gran medida con las proyecciones del algoritmo, lo que evidencia la utilidad práctica del modelo .

En resumen, los hallazgos indican que implementar métodos de aprendizaje automático en plataformas de intermediación es un enfoque eficaz para reducir la incertidumbre que enfrentan los inversionistas en el sector gastronómico de Ecuador. Al mismo tiempo, se establecen claramente los límites actuales del sistema y se destacan las áreas clave para mejoras en futuras versiones.

## **Introducción:**

La ciudad de Guayaquil ha experimentado un gran crecimiento en el sector gastronómico en los últimos años, Según el GEM Ecuador 2023-2024, las actividades de alojamiento y servicios de comida se han convertido en la segunda actividad más común entre los emprendedores del país. Sin embargo, los emprendimientos culinarios también enfrentan desafíos que llevan al cierre de los negocios, dos de ellos siendo la falta de rentabilidad y la falta de financiamiento (Redacción La Hora, 2024)

Los cierres anticipados de restaurantes son un fenómeno que no sólo afecta a Guayaquil; Este es un problema que afecta a muchos países. Sin embargo, en nuestro país. En las ciudades, las consecuencias suelen ser más graves porque la mayoría de las decisiones de inversión todavía se basan en criterios subjetivos: la intuición del empresario, la experiencia de un amigo que abrió una tienda o reseñas tibias sin mayor apoyo. Escasean los datos objetivos y actualizados que permitan tener una visión integral de la situación: cómo se comporta realmente el consumidor en la zona, hasta qué punto está saturado el mercado en determinado barrio, cuál es el verdadero peso de los costos fijos y variables en la viabilidad del negocio.

De ahí la importancia de contar con instrumentos que hagan accesible información de calidad para todos los actores involucrados. Se ha identificado la necesidad de una plataforma tecnológica inteligente que sea capaz de unir a emprendedores con propuestas novedosas y a inversionistas dispuestos a apoyar proyectos con verdadero potencial, pero que lo haga siempre sobre un fundamento analítico sólido. Una herramienta que, mediante modelos predictivos, entregue estimaciones probabilísticas razonables —con números claros y bien fundamentados— acerca de las posibilidades reales de éxito de un restaurante antes de que inicie operaciones

Este proyecto propone una plataforma pensada para el ecosistema gastronómico de Guayaquil, que no solo permita el encuentro entre quienes tienen el sueño y quienes tienen el capital, sino que ofrezca una evaluación objetiva de cada propuesta. Un puntaje de viabilidad que ayude a tomar decisiones informadas.

## **CAPITULO 1**

### **Planteamiento del problema**

En Guayaquil, existe un alto nivel de actividad emprendedora y un sector gastronómico en auge, sin embargo, los emprendimientos enfrentan problemas de falta de rentabilidad y financiamiento para llevar a cabo la propuesta.

Los emprendedores, en especial aquellos que están iniciando tienen un acceso limitado a inversionistas. Los emprendedores no siempre conocen las expectativas de los inversionistas y, por otra parte, los inversionistas no tienen acceso a información estructurada de los proyectos propuestos, por lo que muchos emprendimientos culinarios con potencial no llegan concretarse debido a esta falta de visibilidad y a la ausencia de plataformas que permitan este encuentro.

En medio de esta incertidumbre, surge una pregunta importante: ¿cómo podemos mitigar el riesgo en la industria de servicios de alimentos utilizando modelos predictivos? ¿basado en datos? La solución es cambiar la forma tradicional de evaluar y tomar decisiones sobre la financiación de estos proyectos, dejándolos atrás. tomar decisiones casi exclusivamente basándose en la intuición y emitir juicios basados en análisis numéricos y objetivos.

Los modelos de machine learning permiten procesar al mismo tiempo varios factores del negocio que antes se miraban de manera separada y bastante subjetiva. Aspectos como la experiencia previa del emprendedor, si cuenta o no con un plan de negocios bien armado, el nivel de estudios alcanzado, la validación que haya obtenido en el mercado (por ejemplo, a través de un MVP o pruebas reales) y el capital que se necesita para arrancar pueden analizarse de forma integrada para calcular una probabilidad concreta de que el emprendimiento pueda tener éxito.

Este enfoque no busca sustituir el criterio personal de los involucrados, sino apoyarlo con información cuantitativa sólida, de modo que tanto los emprendedores como los inversionistas puedan ver con claridad cuáles son los puntos fuertes del proyecto y en qué aspectos habría que trabajar para mejorar sus probabilidades.

Existe también ausencia de herramientas objetivas que permitan al inversor observar y analizar si un potencial emprendimiento llegaría a tener una rentabilidad considerablemente buena, lo que puede llevar a decisiones de inversión basadas en experiencia personal y en información incompleta. En el sector culinario, el análisis de las múltiples variables existentes puede llegar a ser complejo y toma tiempo; un recurso que las personas consideran esencial. Por lo cual, una herramienta que permita visualizar la tasa de rentabilidad de un emprendimiento será crucial para que inversores puedan apoyarse en decisiones objetivas basadas en casos de estudios anteriores, y tener una comunicación constante con el inversor para analizar que variables de negocio podría mejorar.

### **Justificación del problema**

Este trabajo contribuye al análisis de inteligencia artificial, en particular en la creación de modelos predictivos para la toma de decisiones en el ámbito empresarial. Para los emprendedores, la existencia de una plataforma que permita conectar con inversores a través de una plataforma centralizada y, a su vez mantener una comunicación constante contribuiría positivamente a los problemas de falta de visibilidad y financiamiento en emprendimientos culinarios. Para los inversores, una plataforma que estudie las vastas variables presentes en el sector culinario, y según datos históricos represente visualmente la tasa de éxito de dicho negocio, reduce el tiempo de

estudio y costos con la evaluación de proyectos y permitiría ofrecer decisiones objetivas basadas en datos que complementan su análisis subjetivo.

### **Pregunta de investigación**

¿Se puede crear un modelo predictivo fundamentado en machine learning que analice con una precisión mayor al 80% la viabilidad de proyectos culinarios, y conectarlo a una plataforma digital operativa que facilite el contacto directo entre personas emprendedoras e inversores?

### **Hipótesis**

La puesta en marcha de un modelo predictivo basado en el algoritmo XGBoost, integrado dentro de una plataforma digital de conexión, posibilita que se valore con una exactitud superior al 80% la viabilidad de iniciativas culinarias a través del análisis sistemático de variables empresariales esenciales, disminuyendo de este modo la asimetría informativa entre inversores y emprendedores.

### **Delimitación**

El siguiente proyecto se delimita al desarrollo de una plataforma donde emprendedores e inversionistas puedan conectar y a través de una predicción, observar si la propuesta podría ser exitosa. Sin embargo, el modelado predictivo estará diseñado considerando las variables únicamente del sector culinario, sin tomar en cuenta variables de otro segmento de emprendimientos fuera del rango de estudio, aunque esta pueda ser adaptable a futuro.

Al ser un proyecto netamente académico, en el prototipo se considera las limitaciones de acceso a una base de datos de la ciudad de Guayaquil, por lo cual el proyecto se delimita a datos

simulados dentro de un estudio; es decir las predicciones no tomará en cuenta todas las ubicaciones y sectores de la ciudad de Guayaquil, se desarrollará en un entorno accesible.

Se utilizará frameworks modernos de desarrollo web y la implementación de algoritmos de machine learning supervisado.

### **Objetivo General**

Desarrollar una plataforma digital que funcione de manera interactiva y que sirva para poner en contacto a emprendedores del rubro gastronómico con posibles inversionistas en la ciudad de Guayaquil. La plataforma incluirá un modelo de inteligencia artificial que ayude a calcular y mostrar el nivel de potencial o probabilidad de éxito que tendría cada proyecto culinario que se presente.

### **Objetivos Específicos**

Reconocer y describir cuáles son los requisitos más importantes que necesitan los emprendedores gastronómicos de Guayaquil para conseguir financiamiento, ya sea a través de entidades públicas o de fuentes privadas.

Elaborar e implementar un modelo de aprendizaje supervisado que trabaje con información clave del sector de restaurantes y comida, con el fin de estimar qué tan probable es que un emprendimiento culinario prospere, y comprobar su calidad usando las métricas de evaluación correspondientes.

Desarrollar una página web interactiva (usando herramientas y frameworks actuales) que permita a los emprendedores subir y explicar sus ideas de negocio, y que al mismo tiempo deje a los emprendedores revisar las estimaciones que genera el modelo predictivo.

Probar el modelo en condiciones controladas y con casos simulados que se parezcan a la realidad del sector gastronómico en la ciudad, comparando lo que arroja el modelo con lo que indican expertos del área o lo que se ha publicado en estudios previos.

### **Alcance**

- El modelo predictivo se va a construir principalmente con datos que se generen de forma simulada o con conjuntos de información que son bastante limitados, así que su exactitud va a depender mucho de cómo sean esos datos de partida y seguramente necesitará ajustes o más datos reales en el futuro para usarse de verdad.
- La plataforma que se desarrolle en este proyecto se va a enfocar en facilitar el contacto y la comunicación entre emprendedores e inversionistas dentro de un espacio digital pensado para colaborar y recibir orientación.
- Todo el desarrollo tecnológico va a estar orientado a que la plataforma se pueda usar desde cualquier dispositivo que tenga internet, incluyendo celulares y tabletas.
- En este trabajo no se va a incluir ninguna conexión con sistemas bancarios, notariales o legales que sirvan para formalizar o firmar acuerdos entre las personas que se conecten a través de la plataforma.

- Para validar cómo funciona, se hará una prueba con un grupo pequeño de usuarios piloto, principalmente para revisar si es fácil de usar y si cumple con lo que se espera en términos de funcionalidad.

## **CAPITULO II. MARCO TEÓRICO**

En el presente capítulo, se establecerán las bases conceptuales y teóricas que fundamentan el desarrollo de una plataforma interactiva de emprendimientos, utilizando modelos predictivos basados en inteligencia artificial para ponderar el éxito de un emprendimiento.

### **Marco referencial**

(Obschonka & Audretsch, 2019), en su estudio “Artificial intelligence and big data in entrepreneurship: a new era has begun”, explorar nuevas aplicaciones basadas en inteligencia artificial y aplicar diferentes métodos de aprendizaje automático a diversos aspectos del emprendimiento y el desempeño. han demostrado que los algoritmos de aprendizaje supervisado logran altas tasas de precisión en la predicción del éxito empresarial, incluso superiores a los modelos tradicionales. Concluyeron que la inteligencia artificial mejora las capacidades de previsión.

Villachica et al. (2022), en su estudio “Modelo predictivo basado en machine learning dirigido a PYMES de venta, caso de estudio Bluefields”, evalúan las dificultades en cuanto a control y organización en pymes en la ciudad de Bluefield. Se destaca que; la mejor forma para aprovechar los datos sin usar es mediante Machine Learning. Concluyeron que el modelo de predicción es una herramienta que supone mejoras notorias en cuanto a Pymes en la ciudad

anteriormente mencionada, donde su uso genera información que permite tomar decisiones más acertadas para el éxito de los negocios.

(Díaz, 2023), en su proyecto “PROPUESTA DE MODELO PREDICTIVO DEL PRECIO DEL PAN MEDIANTE USO RNA. “UN CASO DE ESTUDIO SECTOR PANIFICADOR, BOGOTÁ”. Desarrolló un modelo basado en redes neuronales artificiales, dirigidas a pequeñas y medianas empresas de pan para poder predecir el precio del pan y consecuentemente mejorar el manejo de recursos utilizados en estas empresas. Los resultados fueron exitosos, alcanzando el correcto funcionamiento del modelo de predicción.

Salazar, Martínez, Reyes-Nava, López-González (2024), en su proyecto “Plataforma interactiva de exploración estadística para municipios del estado de México: acceso, carga y visualización eficiente” desarrollan una plataforma web interactiva que facilita el acceso a distintos datos estadísticos relacionados al municipio en la cual se puede encontrar, mapas interactivos de todo el país mexicano y sus estados. Con esto crearon una plataforma para el acceso de esta información de manera ordenada y fácil de entender para los usuarios.

Bel Durán, Lejarriaga, Pérez, Martín, López, Javier Sánchez (2019), en su estudio “Acceso a la financiación de proyectos de emprendimiento social A través de plataformas en red y su relación con los objetivos de desarrollo sostenible” habla de cómo a falta de financiación de los emprendimientos del ámbito social, estos están recurriendo a crowdfunding y como se han creado plataformas para unir a inversores y emprendedores.

Jairo Buitrago & Luis Cubillos (2023), en su proyecto “Plan de negocio para una plataforma de intermediación entre inversionistas y emprendimientos ubicados en la ciudad de

Bogotá” mediante una metodología Canva, ellos crean un plataforma la cual tiene como objetivo ayudar a los inversionistas encontrar oportunidades de inversión rápidamente y a los proyectos tener un medio de financiación que los ayude en su crecimiento, y con esto ellos esperan su implementación ayude a fomentar la economía local y apoyar el empleo en la región.

En base a estos distintos casos y distintos estudios se ha tomado la decisión de llevar a cabo nuestro proyecto, ya que se tiene la idea que el mismo puede ayudar tanto a emprendimientos tanto grandes como pequeños, y que su creación serviría como una herramienta que podría ser utilizada como apoyo. El cual puede ser necesario en caso de que los mismos no encuentren una forma de sustento para sus proyectos actuales o futuros.

## **Marco Conceptual**

### **Emprendimiento:**

Según (Redacción Ventana EBC, 2023), el emprendimiento es la capacidad de individuos o un grupo de individuos para desarrollar un negocio con el fin de generar valor monetario o social. La relevancia de este término en nuestra sociedad es vital, puesto que los emprendimientos generan cultura, y un gran impacto social. El emprendimiento está asociado con la “búsqueda de generación de valor, mediante la creación o expansión de actividades económicas, identificando y explotando nuevos productos, procesos o mercados” (Saiz-Alvarez, 1 C.E.)

El emprendimiento “genera competitividad, y favorece la concurrencia, creatividad y también la innovación” (Garcia-Reinoso, 2022, p.7).

### **Ecosistema emprendedor:**

Algo crucial para el crecimiento de los emprendimientos, es comprender que los mismos dependen de un ecosistema emprendedor donde se constituyen por un conjunto determinado de actores empresariales, organizaciones e instituciones que, coordinados, posibilitan el emprendimiento dentro de un establecimiento previamente determinado.

Los actores clave que forman parte del ecosistema empresarial incluyen: fundadores, inversores y entidades.

La colaboración conjunta entre inversionistas y startups crea nuevas puertas de asociación, intercambio de ideas y apoyo mutuo, impulsando a un crecimiento sostenible dentro de los proyectos. (Escobar, n.d.).

### **Este ecosistema emprendedor incluye:**

- **Capital humano:** Los propios emprendedores, su experiencia y capacidades. En el sector gastronómico, esto puede incluir desde chefs, con una formación formal hasta chefs más informales.
- **Recursos financieros:** Disponibilidad de capital en sus diferentes formas: ahorros personales, préstamos bancarios, inversión ángel, capital de riesgo, crowdfunding. Aquí es donde muchos emprendimientos culinarios encuentran su principal barrera (Redacción La Hora, 2024).
- **Mercados:** La existencia de clientes dispuestos a pagar por los productos o servicios. En Guayaquil, con 2.7 millones de habitantes y como centro económico del país, existe un mercado gastronómico dinámico y en crecimiento.
- **Instituciones de apoyo:** Existen distintas instituciones que servirían de apoyo, entre esas están, incubadoras, universidades, y entidades del gobierno que podrían ofrecer capacitación para los nuevos emprendedores.
- **Marco regulatorio:** Existen normas y leyes que, al momento de crear un nuevo emprendimiento, son de gran utilidad o limitante. En el sector gastronómico ecuatoriano, se pueden encontrar regulaciones sanitarias, normativas en el ámbito laboral, y patentes del municipio entre otros.
- **Cultura emprendedora:** la comprensión del riesgo como parte del proceso de creación de una empresa, aceptar el fracaso como un modo de aprendizaje, y la comprensión de las actitudes sociales.

### **Emprendimiento culinario:**

Los emprendimientos culinarios son cualquier tipo de negocio centrado exclusivamente en la producción y venta de alimentos y/o bebidas.

Según el GEM Ecuador 2023-2024, al momento de observar las actividades económicas en el Ecuador, el sector de servicios de comida se ha convertido en uno de los más escogido por los emprendedores ecuatorianos, lo que destaca la importancia de este rubro dentro del ecosistema emprendedor del país.

### **Factores clave para el éxito de los emprendimientos culinarios:**

Factores que influyen en el éxito o fracaso de aquellos emprendimientos fueron utilizados de investigación y análisis durante las últimas décadas. Entender estas variables resulta fundamental y necesario para el diseño de modelos predictivos robustos orientando a emprendedores e inversionistas en la toma de decisiones informadas.

### **Plan de Negocios**

Constituye uno de los factores más sólidos de supervivencia empresarial según la literatura especializada.(Welter et al., 2021) , en un estudio publicado en la revista New England Journal of Entrepreneurship, examinaron 120 emprendimientos ubicados en Estados Unidos con el propósito de explorar cómo se relacionan las metodologías Lean Startup con la planificación tradicional de negocios. Los resultados que nos da el estudio indican que, de la planificación tradicional, el único que mostró una correlación significativa y duradera con los indicadores de desempeño reales fue la misión de la empresa es desarrollar un plan de negocios formal.

Basado en estos resultados, los autores afirman que la oposición entre métodos ágiles y planificación tradicional como opciones mutuamente excluyentes es completamente un error. Por lo tanto, recomiendan a los emprendedores, la integración de técnicas típicas del enfoque Lean Startup al momento redactar planes de negocios con una estructura muy trabajada, y que las ventajas de ambos métodos se utilicen en la gestión de proyectos.

En el sector gastronómico en particular, donde los márgenes de ganancia suelen ser bastante estrechos y la competencia se mantiene muy alta, contar con un plan de negocios bien elaborado cobra una relevancia fundamental. A diferencia de lo que dice el mito tan repetido de que “el 90 % de los restaurantes quiebran en el primer año”, información actualizada de (Datassential, 2025) muestra que la tasa real de cierre en los primeros doce meses se ha mantenido estable por debajo del 6 % en los últimos años, salvo el periodo 2021-2022 que estuvo marcado por los efectos de la pandemia de COVID-19.

Aun así, cuando se observa la tasa de fracaso acumulada a cinco años, en 2018 esta llegaba al 31 %, mientras que para 2025 ha bajado de forma notable hasta situarse por debajo del 1 %. Este descenso se explica principalmente por avances importantes en varios aspectos: una mejor planificación financiera por parte de los emprendedores, mayor acceso a datos comparativos (benchmarking) del sector y el uso cada vez más extendido de herramientas y sistemas de gestión más avanzados y precisos.

### **Nivel Educativo del Emprendedor**

El nivel académico del emprendedor ha mostrado correlación positiva con el éxito empresarial. (Lasio et al., 2023) ha reportado que los emprendedores con educación de tercer nivel

completan presentan tasas de fallo empresarial significativamente menores que aquellos con solo la educación secundaria. Esta comparativa se explica parcialmente por las habilidades de gestión, capacidad de análisis y contactos que han obtenido por medio de sus estudios.

### **Experiencia Previa en el Sector**

Este tema representa también otro factor importante. (Moritz et al., 2022), en un estudio publicado en la revista *Frontiers in Organizational Psychology*, encontraron que la experiencia previa que tiene el emprendedor juega un papel muy importante en varios momentos clave del proceso de crear y manejar un negocio. Entre los aspectos que destacaron se encuentran: la cantidad de emprendimientos que la persona ha fundado con anterioridad, los cargos de gestión o dirección que ocupó en esos proyectos, y también variables personales como la edad, el nivel educativo alcanzado, los años que lleva trabajando en el mundo de los negocios y la experiencia técnica concreta en el sector al que pertenece el nuevo emprendimiento.

Todos estos factores contribuyen a lo que en la teoría se conoce como “aprendizaje por la práctica” (learning by doing). Según esta idea, el conocimiento más valioso y realmente útil para emprender no se obtiene solo de cursos o lecturas, sino que se construye principalmente a través de la participación directa y repetida en actividades empresariales reales.

### **Validación de Mercado**

Este término, popularizado inicialmente por Eric Ries en su libro “*The Lean Startup*”, es el proceso de confirmar que existe demanda real para un producto o servicio antes de arriesgar recursos significativos. Ries expone un ciclo iterativo donde el emprendedor desarrolla un

Producto Mínimo Viable (MVP), para exponerlo a clientes potenciales, y pudiendo ajustar su propuesta a base de retroalimentación.

En otro estudio, (Blank & Dorf, n.d.) resumieron el proceso de validación en etapas progresivas: comenzando con la investigación exploratoria inicial mediante encuestas y entrevistas, pasando por varios prototipos, hasta la validación definitiva con clientes que pagaron por el producto. Los emprendimientos que pasaron por todas estas etapas de validación antes de escalar presentan un índice de supervivencia mayor

### **Capital Inicial y Estructura Financiera**

La cantidad de capital inicial es un factor crítico que frecuentemente es subestimado. Diversos estudios recientes acerca de la viabilidad financiera en el sector de restaurantes coinciden en señalar que la subcapitalización y una mala administración del presupuesto siguen siendo, con mucho, las causas más frecuentes de cierre definitivo de estos negocios (Toteat, 2026). Esta desventaja se debe principalmente al hecho de que los empresarios tienden a subestimar no sólo los costos iniciales de establecer una instalación sino también Se necesita capital de trabajo para sostener las operaciones mientras la empresa tarde (a menudo más de lo esperado) en alcanzar el punto de equilibrio. Como resultado, los recursos se agotan antes de que el restaurante pueda generar un ingreso estable y sostenible.

### **Etapas de Desarrollo del Emprendimiento**

Los comienzos de un emprendimiento condicionan significativamente su perfil de riesgo. La idea inicial de los emprendimientos presenta incertidumbre máxima al carecer de validación

práctica, mientras que aquellos ya tienen práctica han superado barreras iniciales y demostrado cierta viabilidad.

### **Número de Empleados**

La cantidad de empleados o colaboradores con los que inicia un emprendimiento en sus primeras fases ha demostrado tener una relación importante con sus posibilidades de permanecer en el mercado a lo largo del tiempo.

En esta línea, (Kim et al., 2023), en el artículo titulado “How to succeed in the market? Predicting startup success using a machine learning approach”, publicado en la revista *Technological Forecasting and Social Change*, identificaron —a través del análisis de importancia de características— que las variables asociadas al número y características de los empleados representan predictores relevantes para determinar el éxito de las nuevas empresas

### **Años en el Negocio**

Uno de los indicadores más sólidos y consistentes para prever la viabilidad futura de un emprendimiento es el tiempo que este ha logrado mantenerse operativo.

(McCarthy et al., 2023), en su estudio titulado “The impact of founder personalities on startup success” y publicado en la revista *Scientific Reports*, encontraron evidencia de que variables como la ubicación geográfica del startup, el nivel de compromiso y dedicación de los fundadores, así como la antigüedad de la empresa, ejercen una influencia significativa sobre los distintos indicadores de éxito empresarial.

## **Proyecciones de Ingresos**

Las proyecciones de ingresos que elabora un emprendedor para su proyecto ofrecen pistas importantes acerca de cuánto entiende realmente del mercado y del grado de realismo con que enfrenta su iniciativa empresarial.

En este sentido, (Razaghzadeh Bidgoli et al., 2024), en su artículo “Predicting the success of startups using a machine learning approach” publicado en la revista *Journal of Innovation and Entrepreneurship*, mostraron —mediante la aplicación del método SHAP— que el volumen de financiamiento obtenido y el nivel de exposición en medios se posicionan como variables especialmente relevantes a la hora de predecir el éxito de los startups.

## **Características distintivas del sector culinario**

La alta intensidad de capital inicial es un punto de inflexión importante al inicio de un proyecto dentro del sector culinario. Abrir un restaurante requiere inversión en equipamiento de cocina, mobiliario, inventario inicial y capital de trabajo para los primeros meses.

La dependencia de la ubicación también es un factor esencial. La ubicación puede determinar en gran medida el éxito de un establecimiento de comida, dado por algunos factores como el tráfico peatonal, las características demográficas de la ubicación, entre otros.

El sector culinario es sumamente dinámico. Los gustos y preferencias a nuevas tendencias lo hacen uno de los sectores más cambiantes.

## **Variables de Ubicación y Entorno**

La relevancia de la ubicación en el sector gastronómico se encuentra ampliamente documentada y constituye uno de los determinantes más críticos para la supervivencia y el desempeño sostenido de los establecimientos. Diversas características influyen de manera decisiva en el peso que adquiere la localización dentro de la ecuación de éxito del negocio. Entre las principales se destacan:

- **Tráfico peatonal y vehicular:** El volumen de personas y vehículos que transitan por las inmediaciones del local incide directamente en su visibilidad y en la captación de clientes potenciales de forma espontánea.
- **Accesibilidad:** al momento de retener y atraer a la clientela, factores como la disponibilidad de estacionamiento y, buenas rutas para la llegada al establecimiento son cruciales para un establecimiento.
- **Perfil demográfico y socioeconómico del entorno:** características como nivel de ingreso promedio, edades, composición del hogar, y hábitos de consumo, determinan en gran medida que tan viable son ciertas ideas del ámbito gastronómico, por ejemplo, existirán dificultades para que un restaurante de alta cocina crezca, en sectores de ingresos bajos, independientemente de la calidad de lo que ofrecen.
- **Zonificación y normativas municipales:** Las regulaciones tales como los horarios de operación, regulaciones en la venta de bebidas alcohólicas, control de ruido y, el uso de espacio público u otras restricciones pueden presentar

limitaciones a ciertos modelos de negocio, afectando sus planes operativos y económicos.

En el caso particular de Guayaquil, la ciudad existe una marcada distinción en su geografía comercial y empresarial. Una zona la cual en la cual destaca, es el sector norte, especialmente en el sector Kennedy, el mismo concentra una población de mayores niveles de ingreso y presenta una gran densidad de establecimientos gastronómicos orientados a segmentos medio y medio-alto, lo que representa oportunidades y desafíos, específicos para los nuevos emprendimientos que se instalan en esta área.

### **Ecosistema de financiamiento para emprendimientos**

#### **Startups (empresas emergentes)**

Para los emprendimientos, el principal reto consiste en el financiamiento inicial para poner en marcha el proyecto. Las “Rondas Semilla” consiste en las primeras rondas de financiamiento de una compañía, y preceden a la Serie A (La primera ronda de inversión de una empresa emergente).

La Ronda Semilla se usa para llevar a la empresa emergente de una idea inicial hacia la tracción inicial, contratar miembros del equipo y experimentar en el mercado. Su objetivo principal es encontrar como el producto propuesto encaja en el mercado. **(Startupeable, n.d.)**

El capital Semilla es esencial para las empresas emergentes, ya que es el tipo de financiación inicial para el arranque de estas empresas, cubriendo gastos como el alquiler, los equipos y costos relacionados al desarrollo del proyecto. **(Startupeable, n.d.)**

La inversión ángel es una provisión de capital semilla de un inversor a una empresa emergente a cambio de capital social o deuda convertible [Haga clic o pulse aquí para escribir texto.](#) Los inversionistas ángeles son personas con recursos financieros que colocan su dinero en negocios que están en sus primeras etapas.

Un inversionista ángel invertirá porque al invertir en una etapa temprana, pueden obtener un gran retorno si el negocio propuesto tiene éxito. Además, la inversión inicial puede dar paso a que el inversor y el emprendedor formen una buena relación y proporcionar asesoramiento en negocios y contactos.

El “Crowdfunding” es uno de los instrumentos de financiación más usados por startups. Permite captar recursos directamente de usuarios a través de plataformas en línea, sin depender del soporte de algún sistema bancario.

En Ecuador, una investigación muestra resultados favorables en startups con respecto al uso de plataformas de crowdfunding para la generación de capital semilla, ahorrando costos, la creación de redes en inversionistas y beneficios en la comunicación. El estudio concluye que el crowdfunding tiene un gran potencial impulsar la inversión en etapas tempranas de startups. (Jacqueline et al., 2024)

### **Modelos predictivos**

Según (IBM) Los modelos de análisis predictivo son modelos diseñados para evaluar datos históricos descubrir patrones, observar tendencias y utilizar esa información para predecir tendencias futuras. Los resultados obtenidos pueden maximizar la posibilidad de alcanzar los objetivos específicos de un estudio previamente analizado.

La variedad de los modelos predictivos se puede encontrar aplicadas a varios campos, como lo pueden ser: el marketing, la medicina, entre otros. La eficacia de la aplicación de modelos predictivos depende de la calidad de los datos obtenidos, el algoritmo que se desea utilizar y las características de las variables obtenidas.

### **Ventajas de los Modelos predictivos**

Algunas de las ventajas de los modelos predictivos, como lo indica QuestionPro en su página web “Modelos predictivo: de los datos a la toma de decisiones inteligente”, incluyen:

Los modelos predictivos nos proporcionan nuevos conocimientos que son aplicables a la hora de toma de decisiones. Al analizar datos históricos los modelos predictivos encuentran patrones que el usuario final no puede identificar de manera oportuna, agilizando tiempo y proporcionando decisiones basados en hechos (datos).

El modelado predictivo aumenta la precisión a medida que se ingresan datos nuevos e históricos, donde se supera el juicio humano en ciertas tareas y lleva a mejores resultados.

El modelado predictivo puede ser utilizado para la identificación de riesgos potenciales y casos de incertidumbre en distintos escenarios, lo cual permite que las organizaciones estén preparadas para minimizar el impacto negativo que puede haber en las operaciones y para la mitigación de riesgos.

El modelado predictivo se utiliza para el análisis del mercado financiero y para la evaluación de riesgos crediticios, esto resulta en una toma de decisiones más inteligentes por parte de los inversores las mismas que son fundamentadas por hechos.

## Métricas de evaluación de modelos predictivos

Las métricas de rendimiento de modelados predictivos son esenciales para el desarrollo de proyectos de ciencias de datos, que tienen como objetivo estimar la precisión de un modelo sobre datos futuros (Singh, 2023)

Algunas de las principales métricas de evaluación incluyen:

**Exactitud:** Esta métrica es común para problemas de clasificación, y se refiere al número de clasificaciones correctamente predichas. Su fórmula matemática es la siguiente

$$\frac{(TP + TN)}{Total}$$

Donde TP se refiere a los verdaderos positivos de una matriz de confusión, TN los verdaderos negativos sobre el total.

**Precisión:** La precisión mide con qué frecuencia es correcta la clasificación del modelo, se calcula como

$$\frac{Tp}{pred. positivas}$$

**Recall:** El recall es la proporción entre los verdaderos positivos (TP) a todo lo positivo.

$$\frac{TP}{TP + FN}$$

Donde FN representa los Falsos negativos de una matriz de confusión.

**F1-score:** La puntuación F1 combina la precisión y el Recall en un valor único, y se usa para medir el rendimiento del modelo. Alcanza su mejor valor en “1” y el peor en “0”.

Su fórmula es la siguiente:

$$2 \cdot \frac{precision \cdot recall}{precision + recall}$$

## **Generación y Validación de Datos Sintéticos en Investigación**

En el ámbito del aprendizaje automático, la carencia de datos de calidad suele representar el factor más limitante para la construcción de modelos predictivos sólidos. La relevancia de esta problemática aumenta al momento de tomar en consideración áreas específicas, en las cuales la obtención de información representa altos costos, está limitada por normativas de privacidad o escasea disponibilidad de registros históricos. Una estrategia viable para contrarrestar estas

limitaciones es utilizar una alternativa metodológica la cual consiste, en la creación de datos sintéticos, siempre que su aplicación se sustente en criterios de rigor metodológico y transparencia en el proceso.

Los datos sintéticos corresponden a información creada de manera artificial que emula los patrones, las relaciones y las propiedades estadísticas presentes en conjuntos de datos reales. Según (Lu et al., 2023) en una revisión publicada en arXiv, este enfoque permite a los investigadores sortear obstáculos de índole regulatoria y agilizar el avance metodológico sin infringir la confidencialidad de los datos personales. La generación sintética no persigue la réplica literal de registros originales, sino la producción de nuevas observaciones que mantengan intactas las distribuciones, correlaciones y atributos estructurales característicos de lo investigado.

La aplicación de datos sintéticos cuenta con antecedentes consolidados en diversas áreas del conocimiento. En el campo de las ciencias de la salud,(Chen et al., 2021) registraron cómo el personal investigador emplea generadores sintéticos para perfeccionar algoritmos de diagnóstico, evitando al mismo tiempo la divulgación de información clínica confidencial. Plataformas como Synthea, creada por la MITRE Corporation, producen historias clínicas sintéticas integrales que han sido validadas y empleadas en una amplia gama de estudios académicos.

Este método para crear datos sintéticos se vuelve particularmente útil en situaciones como la que vivimos en Ecuador, donde simplemente no hay bases de datos históricas bien organizadas ni completas sobre los emprendimientos del sector gastronómico.

### **Consideraciones para el presente estudio**

En estudios donde la obtención de datos primarios enfrenta limitaciones temporales o de acceso, la creación de datos sintéticos a partir de estadísticas oficiales y literatura especializada se erige como una estrategia metodológicamente sustentable para la elaboración de pruebas de concepto. Organismos como el Instituto Nacional de Estadísticas y Censos de Población (INEC) y la Superintendencia de Empresas (SUPERCIAS), así como estudios de referencia como Global Entrepreneurship Monitor Ecuador el seguimiento de la actividad de startups en Ecuador proporciona atribuciones y parámetros verificables. Los mismos pueden ser utilizados como una base para la creación de conjuntos de datos sintéticos que representa el panorama empresarial del Ecuador.

Al momento de que otros investigadores quieran evaluar la fiabilidad de los resultados obtenidos, una documentación amplia del procedimiento es crucial para la misma, también se necesitaría una documentación específica para esto en la cual se incluye las fuentes utilizadas, las suposiciones hechas y las limitaciones del método. De igual manera, la misma sienta las bases para que, en el futuro, el estudio sea expandido o retomado mediante la recolección directa de información primaria del sector en cuestión.

### **Tipos de Inteligencia Artificial**

La IA estrecha, o IA débil, es un tipo de inteligencia artificial enfocada en realizar primordialmente una tarea específica, es decir no puede realizar actividades para las que no fue programada.

Este tipo de inteligencia artificial se usa para reconocer patrones, realizar predicciones, y clasificar información; ideal para lo que se busca en el proceso de realización de la implementación del proyecto.

### **Machine Learning (aprendizaje automático)**

El Machine Learning es una de las ramas de la Inteligencia Artificial. Permite a los sistemas aprender datos, tomar decisiones y hacer predicciones sin que estas características sean explícitamente programadas para cualquier situación. Se emplea Machine Learning para procesar datos y consecuentemente identificar patrones diferenciadores del uno a otro.

### **Tipos de aprendizaje automático**

#### **Aprendizaje supervisado**

El aprendizaje supervisado consiste en entrenar modelos de machine learning a partir de conjuntos de datos etiquetados, es decir, en los que cada ejemplo de entrada se asocia con una salida conocida o valor objetivo. En el marco de este proyecto, las variables de entrada incluyen atributos tales como la ubicación del emprendimiento, el tipo de cocina ofrecida, el monto de inversión inicial, los ingresos proyectados y otros indicadores relevantes del perfil del negocio.

Este enfoque resulta central en la investigación, ya que se orienta específicamente al desarrollo de análisis predictivos, cuyo propósito es estimar la probabilidad de éxito o supervivencia de los emprendimientos gastronómicos evaluados.

Dentro del aprendizaje supervisado se distinguen principalmente dos categorías de algoritmos, según la naturaleza de la variable de salida que se pretende predecir:

- **Algoritmos de regresión:** Estos modelos se emplean cuando la variable objetivo es continua y numérica. Su objetivo principal consiste en identificar y modelar relaciones (lineales o no lineales) entre las variables de entrada y de salida y el valor numérico a predecir, lo cual permite estimar distintas magnitudes, como los ingresos futuros, tiempo de supervivencia estimado o la viabilidad representada mediante un puntaje cuantitativo.
- **Algoritmos de clasificación:** los mismos se utilizan cuando la variable de salida es categórica o discreta. Tomando esto en consideración, el modelo se entrena con ejemplos que ya fueron etiquetados con anterioridad para aprender a asignar nuevas observaciones a una de las categorías ya definidas (por ejemplo, “éxito” versus “fracaso”, “alto potencial” versus “riesgo elevado”, o “recomendable para inversión” versus “no recomendable”).

### **Aprendizaje no supervisado**

El aprendizaje no supervisado extrae inferencias de datos no etiquetados, permitiendo buscar agrupaciones y patrones. Estos sistemas de Machine Learning no supervisado suelen estar asociados a sistemas de recomendación, como ejemplo: “Los clientes que compraron x, también compraron y”.(*Tipos de Aprendizaje Automático | IBM, n.d.*)

### **Aprendizaje por refuerzo**

El aprendizaje por refuerzo aprende mediante prueba y error, maximizando recompensas. Modelos predictivos e Inteligencia artificial para la evaluación de **emprendimientos**  
Los modelos predictivos son herramientas que, utilizando datos históricos y actuales, permite

analizar y entender patrones y estimar la probabilidad de que algún evento futuro ocurra, o para predecir varios valores continuos en diferentes modelos de negocio.

**Algunos de los algoritmos para aprendizaje supervisado son:**

- **Regresión logística:** Se utiliza para predecir la probabilidad de que un caso pertenezca a una clase, basándose en una combinación lineal de variables de entrada y utilizando una función sigmoidea para modelar los resultados entre “0” y “1”
- **Árbol de decisión:** Estructuran datos en función de varias variables. Su estructura comienza con un nodo raíz, nodos internos y ramas.
- **Random Forest (Bosque Aleatorio):** Random Forest escoge una gran cantidad de árboles de decisión, los cuales han sido entrenados con una muestra diferente de los datos y luego los combina. La predicción final es el promedio o voto mayoritario de todos los árboles. Esto reduce el sobreajuste y generalmente representa una mejora en la precisión, pero representando dificultades en la interpretación.
- **Naive Bayes:** Está basado en el teorema de bayes, una formula fundamental de la probabilidad que permite actualizar la probabilidad de un evento “A”, a medida que se obtiene información reciente (evento “B”)
- **K-Nearest Neighbours:** El algoritmo de K-Nearest Neighbors (KNN) es un método de aprendizaje automático supervisado, donde se da una clasificación o predicción de un nuevo punto de datos mirando a sus "k" vecinos más cercanos en el conjunto de entrenamiento, el mismo parte de la idea de que los puntos

similares tienden a estar cerca los unos de los otros. En tareas de clasificación, decide la categoría del nuevo punto por una votación mayoritaria entre las etiquetas de esos vecinos; en regresión, se hace un cálculo , al del promedio de los valores vecinos más próximos.

- **Redes neuronales:** Las redes neuronales son la base de diversas aplicaciones como la visión artificial, reconocimiento de voz y el procesamiento de lenguaje natural. Este modelo se inspira en el cerebro humano, compuestos por nodos interconectados (neuronas) organizadas en varias capas que procesan datos, aprenden patrones y son capaces de realizar predicciones sin necesidad de programar explícitamente cada parámetro o regla.

Gradient Boosting y XGBoost: Gradient Boosting es un modelo diferente al de Random Forest. El Gradient boosting combina múltiples árboles de decisiones en un solo modelo, donde cada nuevo árbol se enfoca específicamente en corregir los errores cometidos por los árboles anteriores, haciéndolo enormemente efectivo para tareas de clasificación y regresión. (IBM, n.d.) Reconociendo que XGBoost (o eXtreme Gradient Boosting) es una implementación optimizada de Gradient Boosting conocida por su velocidad, eficacia y capacidad para escalar datos grandes o medianos. (IBM, n.d.)

### **Área Bajo la Curva ROC**

Receiver Operating Characteristic (ROC) manifiesta la existencia de una relación entre la proporción de aciertos verdaderos y la proporción de errores falsos creando diferentes comienzos de clasificación. El área bajo esta curva (AUC-ROC) brinda una medida agregada de la capacidad

discriminativa del modelo, utilizando un rango donde 0.5 indica desempeño equivalente al azar, mientras que 1.0 representa una clasificación perfecta.

### **Calibración de Probabilidades en Modelos Predictivos**

Un aspecto que con frecuencia se pasa por alto en el desarrollo de modelos de clasificación es la calibración adecuada de las probabilidades predichas. Es posible obtener valores muy altos en métricas de desempeño convencionales —como la precisión, el recall o el área bajo la curva ROC (ROC-AUC)— y, sin embargo, que las probabilidades de salida no guarden correspondencia fiel con las frecuencias reales observadas en los datos. En términos prácticos, cuando el modelo asigna a un emprendimiento una probabilidad de éxito del 70 %, lo deseable sería que, entre todos los casos a los que se le asignó una probabilidad cercana a ese valor, aproximadamente el 70 % hubiera alcanzado efectivamente el éxito.

Diversos trabajos recientes han confirmado que los algoritmos basados en gradient boosting, entre ellos XGBoost, generan de manera sistemática probabilidades que se encuentran distorsionadas. Esto ocurre principalmente por su forma de funcionar aditiva y por el modo en que se centran en ir corrigiendo los errores de las iteraciones anteriores, lo cual termina produciendo puntuaciones muy extremas —ya sea muy cercanas a 0 o muy cercanas a 1— cuando el modelo clasifica ejemplos con un alto grado de confianza (Apxml, n.d.).

La causa de esta distorsión se explica porque estos algoritmos priorizan la correcta clasificación de los ejemplos más difíciles (aquellos que se encuentran próximos a la frontera de decisión), en detrimento de la precisión probabilística. Como resultado, aunque logran un

excelente desempeño discriminativo, las probabilidades crudas que entregan no constituyen estimaciones fiables de la probabilidad real de pertenencia a la clase positiva.

### **Técnicas de Calibración**

Para rectificar estas alteraciones, la literatura sugiere dos métodos fundamentales de post-procesamiento:

El escalado de Platt (Platt Scaling), introducido por John Platt en 1999 originalmente para máquinas de vectores de soporte, consiste en ajustar las salidas crudas del clasificador mediante una función sigmoidea parametrizada. El escalado asume que la distorsión en las probabilidades sigue una forma sigmoideal, lo cual en el caso de algunos clasificadores puede dar resultados efectivos, pero insuficientes o contraproducentes cuando la desviación real presenta distintos patrones.

Por su parte, la regresión isotónica (Isotonic Regression) constituye un método sin parámetros el cual crea una función monótona escalonada de manera directa a partir de los datos de calibración. Debido a que no impone ninguna suposición previa sobre la forma de la distorsión, la misma ofrece mayor flexibilidad y es muy adecuada para corregir patrones de calibración más complejos o para identificar ciertas irregularidades.

La regresión isotónica ofrece garantías de calibración que no dependen de ninguna distribución en particular y que, por lo tanto, no requieren hacer suposiciones sobre cómo se distribuyen los datos de origen. Esto la diferencia claramente del escalado de Platt, ya que este último sí se basa en suposiciones paramétricas específicas acerca de la forma que tienen las probabilidades subyacentes.

## **Justificación de XGBoost para el contexto del proyecto**

La elección de XGBoost se sustenta por su capacidad de poder manejar pequeños, o medianos conjuntos de datos.

## **Proceso metodológico para la construcción de un modelo predictivo**

Una metodología eficiente para el minado de datos y proyectos de Big Data es la metodología CRISP-DM, propuesta inicialmente por la compañía estadounidense de tecnologías IBM (International Business Machines), y se divide en 6 fases iterativas:

- **Insights de negocio:** definir claramente qué queremos pronosticar y por qué. En nuestro caso: predecir el éxito de los negocios culinarios para ayudar a los inversores y las empresas a que tomen mejores decisiones.
- **Comprender los datos:** recopilar datos relevantes y analizarlos. ¿Qué información hay disponible sobre restaurantes más antiguos? ¿Qué variables están disponibles? ¿Qué es? ¿Cuál es la calidad de estos datos
- **Preparación de los datos:** Limpiar, transformar y estructurar datos para modelado. Esto va a incluir utilizar valores faltantes, codificar variables categóricas, normalizar escalas.
- **Modelado:** Seleccionar y entrenar algoritmos de machine learning. Utilizar técnicas diferentes y ajustar parámetros.
- **Evaluación:** Medir qué tan bien funciona el modelo con datos que no vio durante el entrenamiento. Usar métricas apropiadas.

- **Despliegue:** Complementar el modelo en un sistema utilizable (en este caso, integrarlo en una plataforma web).

## **Plataformas digitales**

Una plataforma digital, es un tipo de plataforma en la cual el usuario tiene una experiencia dinámica con el contenido, y la misma puede ser o una aplicación o una página web o ambos.

Las plataformas online permiten el encuentro entre el emprendedor y el inversionista, al ofrecer:

- **Acceso inmediato:** Permitiendo conectarse con personas de cualquier parte del mundo, eliminando las barreras geográficas
- **Perfil filtrado:** Se analiza el enfoque sectorial que un emprendedor le quiere brindar a su proyecto y la magnitud que un inversionista está dispuesto a ofrecer. (Inversionistas | Plataformas Online (AngelList, Gust, Etc.) Para Emprendedores y Startups, n.d.)

## **Plataformas digitales para conexión de emprendedores e inversionistas**

El desarrollo de plataformas digitales que apoyan a emprendedores ha cobrado relevancia en diversos contextos. Investigaciones en Ecuador muestran que muchos emprendedores requieren apoyo para elegir adecuadamente plataformas y comprender su viabilidad, concluyendo que una guía estructurada es crucial para optimizar procesos, tomar decisiones informadas y maximizar beneficios. (Ordóñez-Viscaino & Rivera-Prado, 2024)

En el ámbito de la micro financiación y financiamiento colectivo, las plataformas suelen diferenciarse por el tipo de proyectos que alojan (sociales, creativos, tecnológicos, científicos, pymes), las condiciones de comisión, los requisitos de admisión de proyectos y el grado de acompañamiento ofrecido a emprendedores

Una de las plataformas internacionales más conocidas en este ámbito es Kickstarter, la cual permite juntar fondos para varios proyectos creativos que incluyan arte, diseño, música, tecnología, periodismo.

Tomando en consideración el contexto de la tesis, se propone una plataforma web interactiva que, no solo se encarga de conectar emprendedores e inversionistas, sino que también incorpore un módulo de evaluación predictiva del potencial de éxito de los emprendimientos culinarios. Si se observan las tendencias descritas sobre el análisis predictivo de, FinTech y plataformas colaborativas, estas se alinean con el enfoque del proyecto, donde los modelos de inteligencia artificial se integran en los sistemas de soporte a decisiones para disminuir la incertidumbre y mejorar el manejo de recursos.

### **Arquitectura de software escalable para desarrollo web**

Una arquitectura de software escalable es uno de los requisitos más importantes para el desarrollo de una plataforma web interactiva, esta debe permitir que exista un buen rendimiento, estabilidad y velocidad a medida que el número de usuarios y el volumen de datos aumenta. Entre los conceptos clave necesarios para desarrollar una arquitectura de software que pueda escalar se halla:

- **Integración y despliegue continuos (CI/CD):** Hacer que los controles de calidad validen al sistema a lo largo de todo su ciclo de funcionamiento, así detectando los errores antes de que estos se vuelvan críticos.
- **Tolerancia a fallos:** Asegurarse que el sistema se mantenga operativo incluso se alguno de sus componentes falla, así manteniendo la fiabilidad incluso cuando se den problemas inesperados.
- **Seguridad del sistema:** métodos sólidos de autenticación y autorización (OAuth, JWT)
- **Arquitectura sin estado (stateless):** Por cada petición al servidor se incluye toda la información necesaria para la misma, de esta manera el trabajo entre los servidores se distribuye de manera más, eficiente.
- **Diseño nativo de la nube:** Gracias a la computación en la nube existe un acceso a la escalabilidad elástica. Esto ayuda a gestionar mejor los servicios, ofrece un mejor control de los costos. Lo cual optimiza el uso de recursos para un crecimiento a largo plazo.

## **Arquitectura Tecnológica de Plataformas Web**

### **Arquitectura de tres capas:**

La capa de presentación (frontend) es lo que los usuarios ven e interaccionan.

#### **Incluye:**

- Diseño visual (layouts, colores, tipografía)

- Interfaz de usuario (formularios, botones, navegación)
- Experiencia de usuario (flujos, interacciones, respuestas)

Tecnologías modernas de presentación como React, Vue.js o Angular permiten crear interfaces dinámicas, responsivas (que se adaptan a diferentes tamaños de pantalla), e interactivas sin recargar páginas.

La capa de lógica de negocio (backend) procesa datos, ejecuta reglas de negocio, se comunica con bases de datos. Aquí residen:

- **Autenticación y autorización:** Quién puede acceder a qué recurso.
- **Procesamiento de transacciones:** Registro de usuarios, publicación de proyectos, inversiones.
- **Lógica de negocio:** Reglas de validación, cálculos, notificaciones.
- **Integración con servicios externos:** Pagos, emails, almacenamiento de archivos Frameworks como Django (Python), Express (Node.js), Laravel (PHP), o Spring (Java) facilitan el desarrollo estructurado del backend.
- **La capa de datos almacena información persistente:**
- **Datos estructurados:** usuarios, proyectos, inversiones (bases de datos relacionales como PostgreSQL, MySQL)
- **Datos no estructurados:** imágenes, documentos, videos (almacenamiento de objetos como Amazon S3)
- **Datos en caché:** información de acceso frecuente para mejorar velocidad (Redis, Memcached)

## **Frameworks de desarrollo web moderno**

Los frameworks de desarrollo web permiten la aceleración de creación de aplicaciones web y móviles, agilizando tiempos y fomentando la escalabilidad.

### **Framework de Backend**

#### **Node.js con Express.js**

Express.js es un framework web para Node.js. Se usa para crear aplicaciones del lado del servidor y API RESTful. Proporciona soporte para el uso de middleware, gestión de rutas y conexión con sistemas de bases de datos, lo que lo posiciona como una alternativa ampliamente utilizada por programadores que desarrollan aplicaciones bajo los entornos MEAN (MongoDB, Express, Angular, Node) y MERN (MongoDB, Express, React, Node).

Express.js se caracteriza por ser un framework sencillo basado en Node.js, que brinda a los programadores una herramienta ágil y adaptable para el desarrollo de aplicaciones web. Gracias a su diseño simple, facilita la creación rápida de proyectos sin comprometer las capacidades esenciales. (Chandran, n.d.)

Express.js también destaca por su integración con bases de datos, compatible con varias opciones como lo pueden ser:

- MongoDB con Mongoose
- PostgreSQL con Sequelize
- MySQL con Knex.js

A continuación, se presentará mediante tableado, una comparación entre Express.js y otros populares frameworks de Backend, justificando la elección del uso de dicho framework de desarrollo:

| <b>Framework</b>     | <b>Lenguaje</b> | <b>Curva de aprendizaje</b> | <b>Performance</b> | <b>Flexibilidad</b> |
|----------------------|-----------------|-----------------------------|--------------------|---------------------|
| <b>Express.js</b>    | Javascript      | Bajo                        | Alto               | Muy Alto            |
| <b>Django</b>        | Python          | Medio                       | Medio              | Medio               |
| <b>Ruby on Rails</b> | Ruby            | Medio                       | Medio              | Medio               |

*Tabla 1: Comparación entre diferentes frameworks de Backend*

## **Framework de Frontend**

### **React.js: Definición y características**

React.js, desarrollado inicialmente por Facebook (ahora Meta) en 2013, se ha consolidado como una de las herramientas más influyentes en el desarrollo de interfaces de usuario modernas, especialmente en aplicaciones de página única (Single Page Applications o SPAs). Su popularidad no radica tanto en ser un framework completo, sino en actuar como una librería enfocada exclusivamente en la capa de vista, siguiendo un paradigma declarativo: el desarrollador describe cómo debería verse la interfaz en un estado dado, y React se encarga de actualizar eficientemente el DOM cuando ese estado cambia. (Chandran, n.d.)

Lo que realmente distingue a React es su arquitectura basada en componentes. En lugar de trabajar con plantillas o vistas monolíticas, las interfaces se construyen como un árbol de componentes independientes, encapsulados y reutilizables. Cada componente gestiona su propia lógica de presentación y, en muchos casos, su estado interno. (Chandran, n.d.) Explica que esta modularidad no solo facilita el mantenimiento del código a gran escala, sino que promueve patrones de diseño más limpios y predecibles, algo especialmente valioso en equipos grandes o proyectos de larga duración.

El modelo de componentes de React representa un cambio de paradigma en el desarrollo web frontend. Al tratar la interfaz como una composición de bloques en la cual cada uno funciona de manera autónoma, refiriéndose a que cada uno posee propio estado, lógica y estilo, esto fomenta una separación de responsabilidades más clara y una mayor reutilización del código. Este enfoque no solo representa una mejora para la mantenibilidad y escalabilidad de las aplicaciones, sino que también alinea el desarrollo con principios de diseño de software consolidados, como la encapsulación.(Chandran, n.d.)

### **APIs RESTful:**

La comunicación entre frontend y backend, o entre diferentes servicios, típicamente usa APIs REST (Representational State Transfer).

El API REST se puede definir como un estilo arquitectónico basado en:

Recursos identificados por URLs (ej: /api/proyectos/123)

Operaciones estándar HTTP (GET para leer, POST para crear, PUT para actualizar, DELETE para eliminar)

**Sin estado:** cada petición contiene toda la información necesaria

Respuestas en formato estándar (típicamente JSON)

Base de datos

La base de datos es un aspecto fundamental para el almacenamiento y manejo de los datos que se utilizan, por lo mismo la opción escogida fue:

### **Mongo DB**

Según IBM, MongoDB es un sistema de gestión de bases de datos (DBMS) abierto y no relacional que utiliza tablas y filas para su procesamiento y almacenamiento de los datos, esta utiliza documentos flexibles.

MongoDB no requiere un sistema de gestión de base de datos relacional (RDBMS), por lo que ofrece un modelo con almacenamiento de datos elástico que le proporciona a los usuarios consultar y almacenar fácilmente distintos tipos de datos. Lo cual ayuda no solo a la simplificación de la administración de bases de datos, si no a su vez genera un entorno muy escalable para aplicaciones y servicios multiplataformas.

### **Casos de uso de Mongo DB**

**Permite analítica en tiempo real**

MongoDB maneja sin esfuerzo la conversión de documentos JSON y similares a JSON, como BSON, en objetos Java, permite que el análisis de información en tiempo real en diversos entornos de desarrollo permita que la lectura y escritura de datos en MongoDB sea rápida y muy eficaz.

#### Lenguajes soportados por MongoDB

- Python
- PHP
- Ruby
- Node.js
- C++
- Scala
- JavaScript
- Entre otros

#### **Inteligencia artificial aplicada a negocios**

La inteligencia artificial aplicada a los negocios es el uso de herramientas de IA para aumentar la productividad y eficiencia operativa de los usuarios. Según IBM, las organizaciones utilizan la inteligencia artificial (IA) para fortalecer el análisis de datos y la toma de decisiones, mejorar la experiencia del cliente, generar contenido, optimizar las operaciones de TI, las ventas, las prácticas de marketing y ciberseguridad, entre otras cosas.

Los modelos predictivos representan una ventaja competitiva gracias al anticipo de resultados, optimización de recursos y las estrategias personalizadas que pueden ofrecer. Las Empresas líderes utilizan análisis predictivo para: la segmentación de mercados, definición de ubicaciones óptimas de puntos de venta, ajuste de inventarios, mejorar estrategias de precios, prevenir la rotación constante de los empleados y la retención de los clientes.

En que lo que respecta a los startups y emprendimientos, estudios recientes realizados en Latinoamérica han utilizado modelos predictivos para la evaluación de la intención emprendedora universitaria la cual fue realizada mediante el uso de la regresión ordinal logística, identificando variables explicativas como el tipo de relación con la empresa familiar, la vinculación con la innovación, la motivación para emprender, la actitud emprendedora y el PIB de cada país. Otro estudio desarrolló un modelo de ecuaciones estructurales para determinar la intención de emprendimiento, basándose en la Teoría del Comportamiento Planificado. (Krauss Delorme et al., 2018)

### **Chat en tiempo real**

El chat en tiempo real es una herramienta integrada que permite el intercambio ya de mensajes, en la cual el usuario recibe unas respuestas casi inmediatas a sus preguntas o intercambios que tiene con este chat esta tecnología está basada en websockets. según IBM los websockets son un protocolo que autoriza que un navegador web o una aplicación cliente y una aplicación servidor web pueda utilizar una conexión dúplex para comunicarse. El mismo utiliza una secuencia solicitud respuesta HTTP estándar para establecer una conexión. Para la integración de esta solución software, se ha optado por el uso de socket.IO

## **Socket.IO**

Se escoge esta librería, debido a que la misma permite comunicación basa en eventos, de baja latencia y bidireccional entre el cliente y el servidor.

La conexión con Socket.IO puede establecerse mediante distintos transportes de bajo nivel:

- HTTP long-Polling (sondeo largo)
- WebSocket
- WebTransport

El mismo seleccionará automáticamente la mejor opción disponible dependiendo de:

### **Las capacidades del navegador web**

La red a la que está conectado, debido a que algunas redes bloquean las conexiones WebSocket o WebTransport

Entre algunas características extras que tiene Socket.Io también están:

### **Reconexión con HTTP Long-polling(sondeo largo)**

En el caso que una conexión WebSocket no se pueda establecer, la conexión optará por el uso de HTTP long-polling. Esto es bastante útil ya que a pesar de que la mayoría de los navegadores hoy en día soportan WebSocket, aún existe un porcentaje de gente que tiene problemas con la implementación de este.

## **Reconexión automática**

En ciertas condiciones particulares, la conexión WebSocket entre el servidor y el cliente puede ser interrumpida, en cuyo caso tanto el cliente como el servidor desconocen el estado de la conexión.

Por esto este incluye un mecanismo de latido (heartbeat), que verifica periódicamente el estado de la conexión. Y cuando se da la desconexión, se reconecta automáticamente utilizando un retardo exponencial (exponential back-off), para no sobrecargar el servidor.

## **Buffering de paquetes**

Los paquetes se almacenan automáticamente en un búfer cuando el cliente está desconectado y se enviarán cuando se restablezca la conexión.

## **Marco Contextual**

### **Contexto geográfico y demográfico del Sector Kennedy**

Dentro de Guayaquil, existen distintas zonas con un potencial en el desarrollo de emprendimientos de distintas áreas, de la cuales se tomó la decisión de enfocarse en el sector Kennedy, ubicado en el norte de Guayaquil. El sector Kennedy, con los años se ha consolidado como una de las zonas de mayor desarrollo tanto en el sector comercial como en el sector gastronómico.

Si se toman en consideración los datos del Instituto Nacional de Estadística y Censos (INEC, 2023), en el sector Kennedy se puede encontrar diariamente una población flotante de

45000 personas, lo cual se refiere a las personas que transitan el sector de manera temporal, este número de personas ya sea residentes o gente que solo estaba de paso, acuden a los centros comerciales, lugares educativos y los restaurantes.

Es importante tener en consideración que dentro de este sector existe una gran concentración de actividad económica, relacionada con el sector de alimentos y bebidas, y una cantidad de establecimientos gastronómicos que superan a la media que se encuentra en la ciudad.

El perfil socioeconómico, que caracteriza la zona corresponde a estratos medio y medio-alto, esto indica que la capacidad de consumo es bastante alta para emprendimientos culinarios que tengan propuestas nuevas.

### **Ecosistema Emprendedor en el sector Kennedy**

Dentro del sector existe una alta concentración de emprendimientos gastronómicos de distintos tipos, tanto restaurante establecido y negocios de comida artesanal. Según el Servicio de Rentas Internas (SRI,2023), en esta zona se puede encontrar un número estimado de 380 establecimientos registrados, los cuales caen en la categoría de “Actividades de servicios de comida y de bebidas”, el mismo representa un 8.3% de la cantidad total de establecimiento gastronómicos formales en la ciudad de Guayaquil.

La Cámara de Comercio de Guayaquil (2023) identifica a el sector Kennedy como un ”hub gastronómico emergente”, caracterizado por:

- **Diversidad de formatos:** en el mismo se encuentra, distintas opciones en lo que se refiere a gastronomía, entre restaurantes tradicionales, hasta cosas menos comunes, y más experimentales.
- **Accesibilidad:** Ubicación estratégica, con una facilidad en el transporte y un estacionamiento accesible.
- **Infraestructura comercial:** hay presencia de centros comerciales como Mall del Sol, San Marino y Plaza Kennedy que ayudan al tránsito de personas por el área.
- **Concentración de servicios complementarios:** Zonas residenciales, instituciones educativas, Bancos que genera una demanda constante.

### **Desafíos y Oportunidades para Emprendimientos Culinarios en el Sector**

El sector Kennedy a pesar de tener un gran potencial para los emprendimientos culinarios, no está exento de retos que han sido documentados por la Superintendencia de Compañías (SUPERCIAS,2023):

#### **Desafíos identificados:**

- **Alta competencia:** Gran cantidad de establecimientos similares, que podría generar conflicto en la rentabilidad de los locales.
- **Costos de arriendo:** los locales más premium de la zona presentan valores de arrendamiento entre 800 y 2500 dólares mensuales.
- **Rotación de clientes:** una necesidad de diferenciación del resto de locales para mantener el interés del consumidor.

- **Acceso a financiamiento:** Limitantes para inversión inicial como para la expansión de los negocios.

### **Oportunidades:**

Mercado en crecimiento: el consumo promedio fuera del hogar en el sector ha aumentado un 23% postpandemia (INEC,2023)

- **Tráfico peatonal constante:** Flujo diario de personas que podrían ser potenciales clientes. Ya sea por motivos educativos o temas de trabajo.
- **Perfil de consumidor:** Gente dispuesta a probar experiencias culinarias nuevas.
- **Infraestructura disponible:** Alta disponibilidad de locales comerciales y una gran facilidad para la logística.

### **Necesidad de Conexión entre Emprendedores e inversionistas**

Cuando se habla de la articulación entre emprendedores gastronómicos con propuestas viables, y los inversionistas interesados en el sector, existe una brecha enorme. El Global Entrepreneurship Monitor Ecuador (GEM Ecuador, 2023) reporta que, solo el 18% de los emprendimientos culinarios en la provincia del Guayas, logran acceder a financiamiento formal o inversiones privadas en sus dos primeros años de estar operativos. Esta desconexión se manifiesta en:

- Falta de plataformas especializadas que eliminen la brecha entre los emprendedores y los inversionistas en la industria gastronómica.

- Falta de herramientas para evaluar la viabilidad de los emprendimientos de manera objetiva, con el contexto único del sector.
- Un conocimiento limitado sobre las expectativas de inversión, al mismo tiempo que los inversionistas no cuentan con un mecanismo fiable para la evaluación de riesgos.

### **Justificación de la Delimitación al Sector Kennedy**

El motivo por el cual se escoge el sector Kennedy como sector de estudio se debe a factores tanto metodológicos como prácticos.

#### **Metodológicos:**

Disponibilidad de datos: Buena cantidad de datos estadísticos de este sector respaldados por fuentes como el SRI, SUPERCIAS, y la Cámara de Comercio.

- **Homogeneidad del contexto:** si se toman en cuenta las variables controladas en las mismas encontramos perfiles socioeconómicos, de infraestructura y condiciones del mercado.
- **Representatividad:** este representa un microcosmos representativo del ecosistema gastronómico de la ciudad.
- **Prácticos:** Alta densidad de emprendimientos: un número muy alto de establecimientos con el propósito de la validación del modelo predictivo.
- **Accesibilidad para la investigación:** recopilación de alta facilidad de datos primarios y de validación con actores del sector.

- **Escalabilidad:** los resultados de esta zona podrán ser replicado en otras.

La plataforma que se ha propuesto en este trabajo de titulación busca la solución a esta problemática en el contexto del sector Kennedy, con la implementación de una herramienta tecnológica que conecte tanto a los emprendedores como a los inversionistas, con la inclusión de análisis predictivos basados en inteligencia artificial que consideran las variables del mercado emprendedor de esta zona de la ciudad.

### **Marco Legal**

El presente marco legal establece el fundamento jurídico que sustenta el desarrollo e implementación de la plataforma digital de emprendimientos culinarios con inteligencia artificial en el Ecuador. Se realiza un análisis de la normativa vigente aplicable al tratamiento de datos personales, emprendimiento e innovación, comercio electrónico, inteligencia artificial y protección al consumidor.

Si se observa la Constitución de la República del Ecuador, aprobada en 2008, esta establece el marco fundamental que garantizaría los derechos relacionados con la protección de datos personales, los servicios de calidad, la tecnología e innovación, la fomentación de la ciencia y el acceso a bienes y servicios de calidad.

El artículo 66, numeral 19, de la Constitución reconoce el derecho a la protección de los datos personales. Este derecho incluye tanto la posibilidad de acceder a la información propia como de decidir sobre su uso, además de exigir que se tomen las medidas necesarias para su resguardo adecuado. En el contexto de la plataforma desarrollada, este precepto constitucional

obliga a implementar tanto acciones organizativas como técnicas que aseguren la protección efectiva de los datos de los usuarios que interactúan con el sistema.

El artículo 52 consagra el derecho de las personas a escoger libremente bienes y servicios que ofrezcan la mejor calidad posible, y a recibir información clara, precisa y no engañosa acerca de sus características y contenido. Aplicado al presente proyecto, esta norma refuerza la obligación de que la plataforma entregue a emprendedores e inversionistas datos veraces y transparentes sobre las evaluaciones realizadas, evitando cualquier tipo de presentación que pueda inducir a error en la toma de decisiones de inversión.

### **Ley Orgánica de Protección de Datos Personales**

La Ley Orgánica de Protección de Datos Personales (LOPDP), publicada en el Registro Oficial Suplemento No. 459 del 26 de mayo de 2021, constituye el instrumento normativo principal para la protección de datos personales en Ecuador. Su Reglamento General fue emitido posteriormente para desarrollar la normativa de aplicación.

La LOPDP tiene como objetivo que se garantice el ejercicio del derecho a la protección de datos personales, con lo cual establece los principios, derechos, obligaciones y mecanismos de tutela. La aplicación de esta ley se extiende a todas las personas naturales o jurídicas, ya sean del sector público o privado, que realicen el tratamiento de datos personales dentro del territorio ecuatoriano o cuando los efectos de dicho tratamiento se produzcan en el país.

En el caso concreto de la plataforma desarrollada para emprendimientos culinarios, resulta plenamente aplicable esta normativa, puesto que el sistema recolecta, almacena y procesa datos de carácter personal tanto de los emprendedores como de los inversionistas. Entre la información que

se maneja se encuentran datos de identificación personal, información de contacto, detalles financieros relacionados con el proyecto y datos sobre el perfil profesional y experiencia de los usuarios.

### **Ley Orgánica de Emprendimiento e Innovación**

La Ley Orgánica de Emprendimiento e Innovación, publicada en el Registro Oficial Suplemento No. 151 del 28 de febrero de 2020, junto con su Reglamento (Decreto Ejecutivo 1113 del 27 de julio de 2020), establece el marco normativo para incentivar y fomentar el emprendimiento, la innovación y el desarrollo tecnológico en Ecuador.

El propósito de esta ley es fijar el marco normativo que estimule y fomente la innovación, el emprendimiento y el progreso tecnológico, promoviendo una cultura emprendedora e introduciendo nuevos tipos de financiamiento y sociedades con el fin de robustecer el ecosistema emprendedor.

El alcance de aplicación abarca todas las acciones, tanto públicas como privadas, que están relacionadas con la creación de empresas y la innovación, en el contexto de los distintos tipos de economía: pública, privada, mixta, popular y solidaria.

### **Registro Nacional de Emprendimiento**

Se establece el Registro Nacional de Emprendimiento (RNE) como instrumento para la caracterización de emprendimientos y acceso a servicios preferenciales. Los emprendimientos registrados pueden acceder a beneficios como descuentos en permisos de funcionamiento,

notificaciones sanitarias, certificaciones de Buenas Prácticas de Manufactura y servicios del SENADI.

La plataforma propuesta permitirá la conexión entre emprendedores culinarios y potenciales inversionistas, alineándose con los objetivos de esta ley de fortalecer el ecosistema emprendedor ecuatoriano.

### **Capítulo III. Metodología de la Investigación**

#### **Tipo de Investigación**

El presente trabajo se enmarca en la investigación aplicada tecnológica, orientada al desarrollo de una solución práctica que aborde problemáticas concretas del ecosistema emprendedor culinario. A diferencia de la investigación básica, cuyo propósito es ampliar el conocimiento teórico sin una aplicación inmediata, la investigación aplicada busca resolver problemas específicos mediante la implementación de soluciones tangibles.

Álvarez et al. (2023) afirman que "la investigación tecnológica se enfoca en poner en práctica los saberes científicos con el objetivo de desarrollar nuevos sistemas, productos y procesos; en otras palabras, tiene como propósito sustituir o mejorar algo que ya existe en la producción de bienes y servicios, o crear algo novedoso: programas, equipos, sistemas o procesos".

Teniendo esto en consideración, el problema que se ha logrado identificar es la carencia de herramientas objetivas que den la posibilidad de evaluar la viabilidad de un proyecto antes de comenzar la búsqueda de una inversión. La naturaleza del proyecto también lo sitúa dentro de la investigación explicativa, debido a que el mismo busca establecer relaciones causales entre

variables predictoras y el éxito empresarial. El modelo predictivo desarrollado intenta explicar qué factores incrementan o disminuyen la probabilidad de que un emprendimiento tenga éxito, proporcionando así una herramienta de toma de decisiones fundamentada en evidencia.

### **Enfoque Metodológico**

El proyecto adopta un enfoque predominantemente cuantitativo con elementos cualitativos adicionales. El uso de esta técnica le brinda un enfoque doble a su trabajo: desarrollar un modelo de predicción basado en aprendizaje automático requiere procesar datos numéricos y utilizar métodos estadísticos; por otro, si se toma en consideración una aproximación cualitativa se debe tomar en consideración la comprensión del contexto emprendedor y la validación de la plataforma.

En lo que respecta al proyecto el componente cuantitativo está presente en distintas partes. Por un lado, se encuentra en la recopilación y análisis de estadísticas oficiales sobre emprendimiento, las características del sector gastronómico del Ecuador y en la supervivencia de las empresas. Por otro, en el modelo predictivo más específicamente en su diseño y aplicación, que transforma variables observables en una puntuación numérica de viabilidad. Finalmente, en la evaluación del modelo a través de métricas como el puntaje F1, el área bajo la curva ROC, la precisión, y la exhaustividad.

El componente cualitativo complementa este enfoque al analizar el ecosistema emprendedor local, identificar las barreras y oportunidades que enfrentan los emprendedores culinarios, y evaluar la usabilidad de la plataforma desarrollada.

## Metodología de Desarrollo

Para el desarrollo de la plataforma web se adoptó una metodología ágil con enfoque iterativo-incremental, seleccionada por su flexibilidad para incorporar cambios durante el proceso de desarrollo de la plataforma y su capacidad para entregar funcionalidades de forma progresiva, revisando y reajustando según las iteraciones. Esta metodología es bastante efectiva en proyectos de desarrollo de software en los que los requisitos de este cambian y aumentan a medida que la comprensión del problema se vuelve más profunda.

Para el desarrollo del modelo de aprendizaje automático se decidió llevar a cabo siguiendo la metodología CRISP-DM (Cross-Industry Standard Process for Data Mining), “un marco estándar ampliamente utilizado en proyectos de ciencia de datos que ofrece una estructura sistemática para el tratamiento de problemas analíticos”. (Chapman et al., 1999) autores de esta metodología, la concibieron como un proceso iterativo y flexible, capaz de adaptarse a diversos contextos y dominios de aplicación.

CRISP-DM organiza el trabajo en seis fases interrelacionadas:

- **Comprensión del negocio:** La primera fase de CRISP-DM se centra en analizar los objetivos del proyecto desde la perspectiva empresarial y traducirlos en un problema definido de ciencia de datos. En el presente trabajo, el objetivo de negocio consistió en desarrollar una herramienta destinada a emprendedores e inversionistas que permitiera evaluar de manera objetiva la probabilidad de éxito de emprendimientos en el sector culinario.

- **Comprensión de los Datos:** La segunda fase involucra la identificación, recopilación y exploración de los datos necesarios para el modelado. En el contexto de esta investigación, se enfrentó un desafío: la escasez de datos históricos sistematizados sobre el contexto del proyecto. Ante esta limitación, se adoptó un enfoque de generación de datos sintéticos fundamentado en estadísticas oficiales y hallazgos de investigaciones previas.
- **Preparación de los Datos:** En la fase de preparación de datos se incluyó la generación de un dataset sintético y su posterior adecuación para las tareas de modelado. Fueron creados 5000 registros que imitan o simulan emprendimientos culinarios los cuales cada uno consta de características distintas, este número de registros contiene una cantidad que se considera suficiente para un entrenamiento del modelo, con una complejidad moderada y en el cual no haya un riesgo de sobreajuste.

La utilización de datos sintéticos en el presente estudio obedece a una limitación estructural del ecosistema emprendedor ecuatoriano que se ha aclarado en el alcance inicial del proyecto: la inexistencia de bases de datos históricas estructuradas, de acceso público y con cobertura suficiente sobre los emprendimientos gastronómicos. En particular, resulta especialmente problemático que dichas fuentes no registren de manera sistemática las variables de resultado (éxito o fracaso del negocio) ni los factores determinantes asociados a dichos desenlaces.

Es preciso subrayar que esta investigación tiene el carácter de una demostración de viabilidad técnica (proof of concept) respecto al empleo de modelos predictivos en la evaluación de proyectos emprendedores culinarios.

- **Modelado:** Durante la fase de modelado se realiza una selección del algoritmo, el entrenamiento del modelo predictivo, y la configuración de los hiperparámetros.
- **Evaluación:** La valoración de la calidad y el desempeño de los modelos generados, son el punto central de esta fase.
- **Despliegue:** La fase Despliegue es la última etapa de CRISP-DM, la misma enfoca en la puesta en producción del modelo, y el desarrollo plan de monitoreo y mantenimiento posterior a su producción. En este proyecto, la serialización e integración fueron implementados en un microservicio implementado con FastAPI, un framework de Python diseñado para el desarrollo de APIs web de alto rendimiento y asíncronas.

### **Población y Muestra**

La población objetivo del proyecto se compone de dos grupos principales dentro del ecosistema emprendedor local. El primer grupo incluye a los empresarios que operan en el sector de la gastronomía, refiriéndose a aquellos que han iniciado o planean iniciar un negocio que tenga relación con productos alimenticios en esta área de estudio. Esto incluye tanto a personas naturales las cuales aún no tienen sus ideas formadas por completo como a propietarios de establecimientos los cuales desean ampliar sus operaciones o tener un acceso a recursos financieros adicionales.

El segundo grupo son los inversores potenciales que pueden definirse como aquellos que tienen el capital y un deseo de financiar el negocio de los productos alimenticios. Aquí se incluyen inversionistas ángeles que actúan de manera individual, participantes en plataformas de crowdfunding y gente interesada en la contribución del crecimiento y fortalecimiento de la base empresarial de la zona.

Para el entrenamiento del modelo predictivo fue generado y utilizado datos sintéticos conformados por 5000 registros que recrean emprendimientos gastronómicos con atributos y trayectorias muy variadas. La decisión utilizar esta metodológica es una respuesta a la falta de bases de datos históricos que muestren una documentación completa del desempeño real que existe en el sector Kennedy en lo que respecta a los emprendimientos, a lo largo de los años.

El empleo de datos sintéticos constituye una práctica ampliamente aceptada en la investigación científica cuando los datos reales son escasos o inexistentes. (Patki et al., 2016) señalan que estos datos resultan particularmente útiles para el desarrollo y la evaluación inicial de sistemas, siempre y cuando se documenten de forma transparente sus limitaciones y se deje claro que los hallazgos obtenidos no pueden considerarse directamente extrapolables sin una posterior validación con información real.

### **Técnicas e Instrumentos de Recolección de Datos**

La recopilación de información secundaria se realizó mediante revisión de fuentes oficiales y académicas. Se consultaron bases de datos del INEC, reportes del SRI e informes del Global Entrepreneurship Monitor.

El dataset sintético fue generado mediante un script desarrollado en Python, el cual implementa distribuciones probabilísticas y relaciones entre variables inferidas a partir de las fuentes previamente consultadas.

En el conjunto cada variable de datos se modeló dirigiendo distribuciones calibradas con base a la evidencia disponible. En este caso, la variable utilizada al número de empleados adopta una distribución en la que el 15 % de los emprendimientos reporta no tener empleados, el 45 % presenta entre 1 y 2 empleados, el 25 % entre 3 y 5, y el 15 % que queda supera los 6 empleados. Estas proporciones se realinean con la estructura empresarial que se observa en Ecuador, según los datos publicados por el Instituto Nacional de Estadística y Censos (INEC).

### **Instrumento de validación cualitativa: entrevista semiestructurada**

Con el fin de dar cumplimiento al cuarto objetivo específico planteado en la investigación (el cual exige evaluar el modelo a través de pruebas controladas y escenarios simulados que reflejen la realidad del sector culinario, contrastando sus resultados con criterios establecidos por expertos), se elaboró una entrevista semiestructurada dirigida a un inversionista con trayectoria activa en el ámbito gastronómico de Guayaquil.

La inclusión de esta herramienta cualitativa cumple con los requisitos de claridad metodológica, a pesar de los tradicionales indicadores cuantitativos de los modelos de medición del desempeño estadístico del algoritmo de manera objetiva, no proporcionan suficiente información objetiva para evaluar si las predicciones que genera son consistentes, realistas y confiables desde la perspectiva de la persona que realmente toma la decisión de inversión. La entrevista pretende llenar este vacío incorporando un mecanismo de verificación externa basado

en la evaluación experta de un actor directamente involucrado en el proceso de toma de decisiones en la industria.

### **Justificación del instrumento**

La entrevista semiestructurada fue elegida herramienta para recabar datos primarios en la etapa de validación cualitativa del modelo predictivo. Esta elección se basa en tres criterios metodológicos esenciales. En primer lugar, la validación tiene una naturaleza exploratoria y necesita una herramienta que sea capaz de captar no solo respuestas directas, sino también los condicionantes, matices y razonamientos que el entrevistado presenta de manera espontánea al argumentar sus juicios.

En segundo lugar, el reducido tamaño de inversionistas activos en el sector gastronómico de Guayaquil que satisfacen los criterios de inclusión determinados hace inviable el uso de herramientas cuantitativas, como las encuestas con muestras representativas.

En tercer lugar, el objetivo específico que se valida requiere una comparación de los resultados algorítmicos con criterios establecidos por expertos. Esto implica un diálogo directo que posibilite plantear escenarios específicos y registrar la estimación del experto antes de exponer la predicción del modelo.

### **Criterios de selección del entrevistado**

El entrevistado fue elegido a través de un muestreo no probabilístico intencionado. Esta técnica es apropiada para investigaciones con un componente cualitativo que necesitan que los

participantes tengan experiencia directa y conocimiento especializado acerca del fenómeno en estudio.

Se establecieron cuatro criterios para la selección del entrevistado:

- Tener una experiencia comprobable en la evaluación y financiación de empresas del sector gastronómico.
- Realizar su actividad inversora en Guayaquil.
- Estar dispuesto a participar en una demostración en vivo de la plataforma durante la entrevista, condición esencial para las preguntas del segundo y tercer eje temático.
- No haber estado involucrado en ninguna fase anterior del desarrollo de la plataforma, asegurando así que su criterio evaluativo sea independiente.

### **Diseño del instrumento**

El instrumento de recolección de información se compone de siete preguntas distribuidas en tres bloques temáticos, cada uno de ellos vinculado directamente a aspectos del proyecto que se pretende validar mediante esta consulta:

- **Primer bloque (preguntas 1 y 2):** Este apartado busca identificar y caracterizar los criterios efectivos que el inversionista aplica en la práctica al momento de analizar y decidir sobre un emprendimiento gastronómico.

- **Segundo bloque (preguntas 3, 4 y 5):** En este artículo se indaga sobre la percepción que tenga el inversionista respecto de la utilidad real y concreta de la plataforma como un apoyo en el proceso de la toma de decisiones, mientras que se exploran las dificultades principales o barreras que en este momento enfrenta para encontrar y evaluar proyectos de emprendedores del sector.
- **Tercer bloque (preguntas 6 y 7):** Este bloque tiene como representación el núcleo de la validación cruzada con criterio experto. La pregunta 6 se orienta a recoger observaciones y sugerencias acerca de las variables o factores que el inversionista vaya a considerar relevantes y que fueron incluidos en la versión actual del modelo predictivo. Por su parte, la pregunta 7 se nos presenta un caso concreto y detallado con el objetivo de que el inversionista estime, en forma de porcentaje, la probabilidad de éxito que el asignaría personalmente a ese perfil emprendedor; dicha estimación se utilizará para posterior compararla de manera directa con el valor generado por el algoritmo para el mismo escenario.

### **Procedimiento de análisis de la información**

El proceso se llevó a cabo en tres etapas: la primera, de familiarización, incluyó la transcripción de las respuestas y una lectura minuciosa para detectar las ideas fundamentales del entrevistado; la segunda, de codificación, abarcó la extracción de los conceptos esenciales de cada respuesta y su clasificación en función de su relación con las variables del modelo, las funcionalidades de la plataforma o los elementos que no se habían tenido en cuenta en el sistema actual; por último, en la tercera etapa, denominada interpretación comparativa, se compararon los

descubrimientos cualitativos con los resultados cuantitativos del modelo, detectando coincidencias y discrepancias así como oportunidades para mejorar.

### **Triangulación metodológica**

La validación del modelo predictivo se sustenta en una estrategia de triangulación metodológica, que se define como la combinación de dos o más métodos de investigación en el mismo estudio, con el fin de fortalecer credibilidad en los hallazgos.

La triangulación funciona a dos niveles en el trabajo actual. En el primer nivel, que es la triangulación de datos cualitativos y cuantitativos, se combinan las métricas estadísticas del modelo (F1-Score del 86 %, precisión del 94 % y ROC-AUC de 88,48 %) con la evaluación cualitativa que hace el inversionista acerca de las limitaciones, la utilidad y la coherencia de las predicciones producidas. La validación cualitativa responde a la cuestión "¿las predicciones son razonables para quien toma decisiones reales?", mientras que las métricas cuantitativas responden a "¿el modelo clasifica de manera correcta?".

En la segunda etapa, de triangulación entre fuentes de información, se contraponen los resultados del modelo (entrenado con datos sintéticos calibrados a partir de las estadísticas oficiales del INEC, SRI y GEM Ecuador) con el criterio experto del inversor, que proviene de su experiencia directa en el mercado. La concordancia entre las dos fuentes en temas como la no descalificación automática por falta de plan de negocios o la clasificación positiva del escenario propuesto en la pregunta 7 fortalece la validez externa del modelo dentro de los límites de su alcance actual.

## **Capítulo IV. Propuesta Tecnológica**

El presente capítulo hace referencia al desarrollo técnico de una plataforma digital para conexión de emprendedores culinarios con inversionistas, incluyendo la implementación de un modelo predictivo basado en inteligencia artificial. Se describe la arquitectura del sistema, las decisiones tecnológicas adoptadas, el proceso de construcción del modelo de aprendizaje automático, y los resultados obtenidos durante las pruebas de validación.

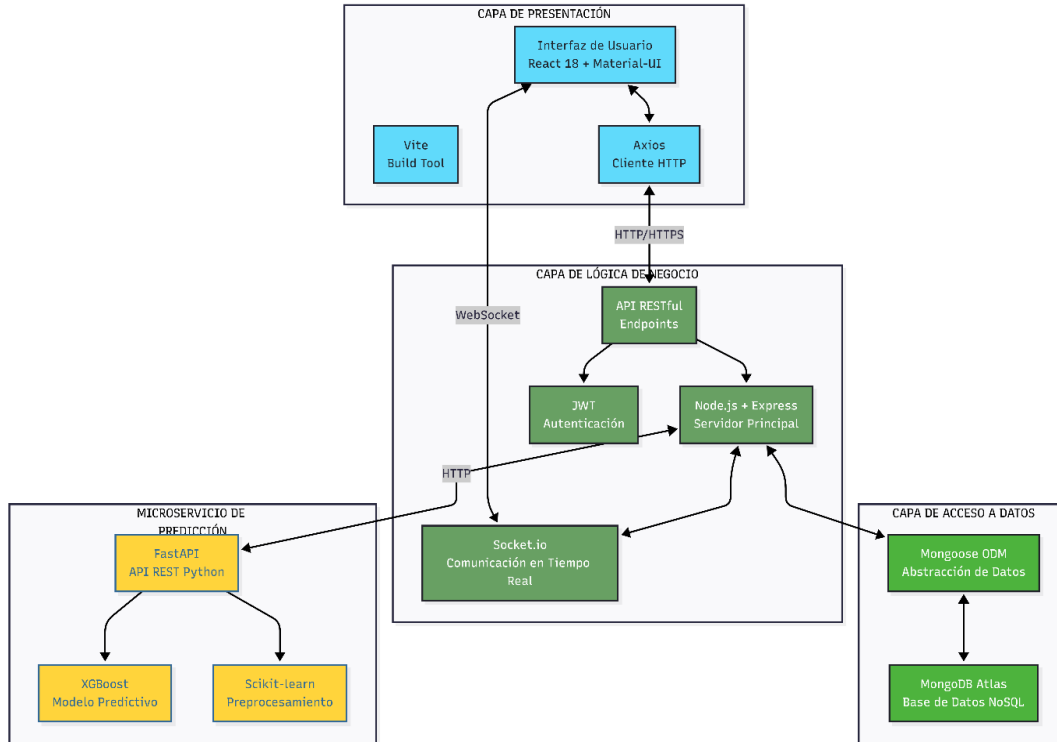
La propuesta tecnológica responde a los objetivos planteados en esta investigación, materializando una solución funcional que integra comunicación en tiempo real, gestión de perfiles diferenciados por tipo de usuario, y un sistema de evaluación de viabilidad empresarial fundamentado en técnicas de machine learning.

### **Objetivo de la Propuesta**

Desarrollar e implementar una plataforma web funcional que permita la interacción entre emprendedores e inversionistas del sector culinario en Guayaquil, incorporando un modelo predictivo capaz de evaluar la probabilidad de éxito de los emprendimientos presentados y generar recomendaciones orientadas a mejorar dicha probabilidad.

### **Arquitectura del Sistema**

La arquitectura del sistema se organiza según el patrón cliente-servidor y se estructura en tres capas principales con separación de responsabilidades: capa de presentación, capa de lógica de negocio y capa de acceso a datos. Además, se incluye un microservicio independiente dedicado específicamente al procesamiento de las predicciones.



*Ilustración 1: Arquitectura del sistema*

Esta decisión de diseño permite el escalado independiente de cada componente y facilita notablemente el mantenimiento del sistema, al mantener aislada la complejidad asociada al modelo predictivo respecto del resto de la aplicación.

La comunicación entre el frontend y el backend se realiza a través de una API RESTful que proporciona endpoints para la autenticación, la gestión de perfiles, la mensajería y las solicitudes de predicción. Para las funcionalidades que requieren comunicación en tiempo real, en particular el sistema de mensajería entre usuarios, se implementó una capa adicional basada en WebSockets.

Previamente, en el capítulo número 2, se estudió diferentes entornos de trabajos, (o frameworks), y la decisión final de elección se fundamenta según los requerimientos del proyecto

y criterio propio de los autores. El stack principal corresponde a lo que en la industria se conoce como stack “MERN” (MongoDB, Express, React, Node.js), complementado con Python para el microservicio de machine learning.

### **Capa de presentación**

Para la interfaz de usuario se utilizó React en su versión 18, librería de JavaScript desarrollada por Meta. El diseño visual se construyó sobre Material-UI, sistema de componentes que implementa las directrices de Material Design de Google. Como herramienta de construcción y servidor de desarrollo se empleó Vite, que ofrece tiempos de recarga significativamente menores que alternativas como Create React App, mejorando la productividad durante el desarrollo. Para la comunicación con el backend se utilizó Axios, cliente HTTP.

### **Capa de lógica de negocio**

El servidor principal se desarrolló con Node.js, entorno de ejecución de JavaScript que permite utilizar el mismo lenguaje en frontend (React.js) y backend. La autenticación se implementó utilizando JSON Web Tokens (JWT), un estándar abierto que permite la transmisión segura de información entre partes mediante un objeto JSON firmado digitalmente.

Los tokens se almacenan en el lado del cliente y se incluyen en cada solicitud a los recursos protegidos. De esta forma se elimina la necesidad de mantener estado de sesión en el servidor.

Para las comunicaciones en tiempo real se incorporó la librería Socket.io, la cual abstrae las complejidades inherentes a la implementación directa de WebSockets y ofrece mecanismos

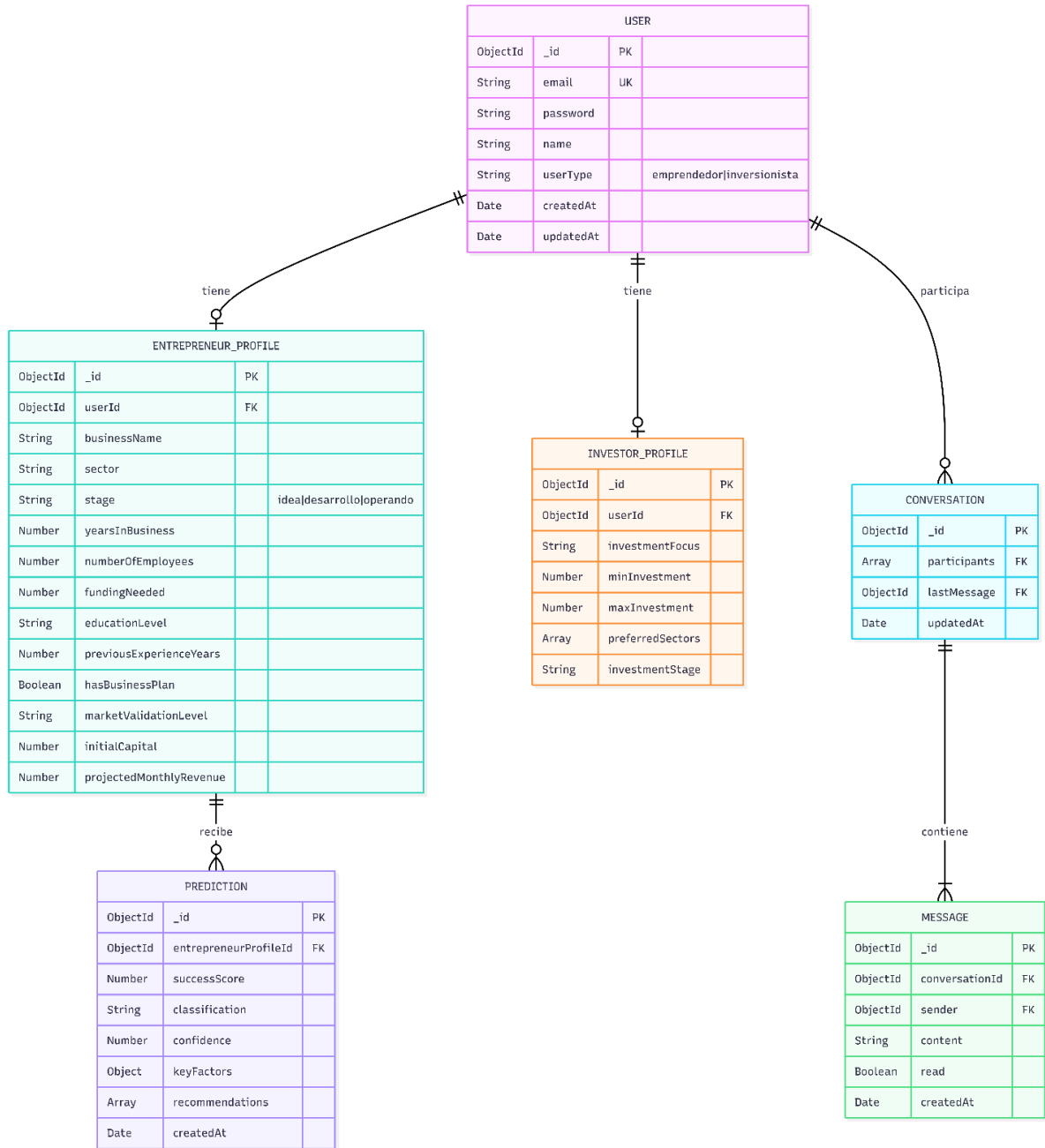
automáticos de fallback para entornos o navegadores que no cuenten con soporte completo de esta tecnología.

El sistema de mensajería entre emprendedores e inversionistas aprovecha estas capacidades para garantizar la entrega instantánea de notificaciones y mensajes

### **Capa de acceso de datos**

MongoDB Atlas provee el servicio de base de datos, ofreciendo una solución documental completamente administrada en la nube. La flexibilidad inherente al modelo de esquema documental resulta especialmente adecuada para este proyecto, dado que los perfiles de emprendedores e inversionistas presentan estructuras de datos heterogéneas y están sujetos a evolución continua a medida que se refinan los requisitos funcionales y de negocio.

Mongoose se emplea como capa de abstracción orientada a objetos entre la aplicación y la base de datos MongoDB



**Ilustración 2: Diagrama Entidad-Relación de BDD**

## **Microservicio de predicción**

Se diseñó e implementó el componente de la inteligencia artificial como un microservicio independiente desarrollado en Python, siendo esta un lenguaje que va a constituir el estándar de facto en el ámbito de la ciencia de datos y el aprendizaje automatizado.

Esta separación arquitectónica permite seleccionar las tecnologías más adecuadas para cada capa del sistema: JavaScript y su ecosistema para el desarrollo web del frontend y backend principal, y Python junto con sus librerías especializadas para las tareas de análisis de datos y modelado predictivo.

En la exposición del modelo mediante una API REST se optó por el framework FastAPI. Este framework, es mucho más actual que otras alternativas consolidadas como Flask o Django REST Framework, una de las características que lo hace destacar es que ofrece una validación de datos automática a través de Pydantic, una generación automática de documentación con la cual se puede interactuar (Swagger/OpenAPI), y un rendimiento superior gracias a su arquitectura asíncrona nativa.

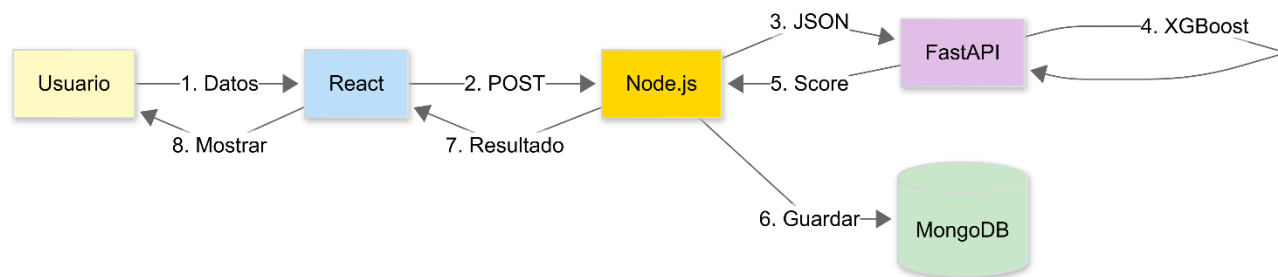
El proceso de creación del modelo predictivo se creó empleando las librerías scikit-learn y XGBoost, ambas opciones muy reconocidas en la industria para las tareas de aprendizaje automático. Scikit-learn generó las herramientas que se necesitaron para el preprocesamiento de datos, la partición de conjuntos de entrenamiento y prueba, así como implementar las métricas de evaluación. Por su parte, XGBoost se utilizó para implementar el algoritmo de clasificación seleccionado, se destaca por su excelente rendimiento tanto en fases de entrenamiento como de inferencia.

| Tecnología | Ventajas  | Desventajas  | Características  |
|------------|---|--|--|
| React      | <ul style="list-style-type: none"> <li>Arquitectura basada en componentes reutilizables</li> <li>Amplio ecosistema de librerías</li> </ul>  | <ul style="list-style-type: none"> <li>Curva de aprendizaje inicial para desarrolladores</li> <li>Requiere librerías adicionales para routing y manejo de estado global</li> </ul>             | <ul style="list-style-type: none"> <li>Flujo de datos unidireccional que facilita el debugging</li> <li>Compatible con Material-UI para interfaces profesionales</li> </ul>                        |
| Node.js    | <ul style="list-style-type: none"> <li>Mismo lenguaje (JavaScript) en frontend y backend</li> <li>Arquitectura no bloqueante ideal para operaciones en tiempo real</li> </ul>     | <ul style="list-style-type: none"> <li>Puede limitar operaciones intensivas de CPU</li> </ul>  | <ul style="list-style-type: none"> <li>Integración nativa con Socket.io para comunicación bidireccional</li> <li>Express proporciona middleware modular para autenticación y validación</li> </ul> |
| MongoDB    | <ul style="list-style-type: none"> <li>Documentos JSON nativos compatibles con el stack desarrollado</li> <li>MongoDB Atlas ofrece escalabilidad automática en la nube</li> </ul> | <ul style="list-style-type: none"> <li>Mayor consumo de almacenamiento comparado con bases relacionales</li> <li>Transacciones multi-documento requieren de configuración adicional</li> </ul> | <ul style="list-style-type: none"> <li>Mongoose ODM facilita validación de esquemas y relaciones</li> <li>Índices de texto completo para búsquedas en la plataforma</li> </ul>                     |
| FastAPI    | <ul style="list-style-type: none"> <li>Validación automática de datos</li> <li>Buen rendimiento</li> <li>Integración sencilla con esquemas de Machine Learning</li> </ul>         | <ul style="list-style-type: none"> <li>Framework bastante reciente</li> <li>Ecosistema de extensiones más reducido que Django</li> </ul>   | <ul style="list-style-type: none"> <li>Inyección de dependencias integrada</li> <li>Soporte nativo para operaciones asíncronas (async/await)</li> </ul>  |

*Tabla 2: Ventajas, Desventajas y Características de herramientas usadas*

Cuando un usuario solicita la evaluación de viabilidad de su emprendimiento, el flujo de datos sigue la secuencia descrita a continuación:

1. El frontend recoge la información del proyecto a través de formularios debidamente validados y envía una solicitud HTTP al backend desarrollado en Node.js.
2. Este backend autentica primero al usuario, realiza una validación adicional de los datos recibidos y, una vez comprobados, reenvía la solicitud al microservicio independiente implementado en Python.
3. En los microservicios, se aplican las transformaciones y el preprocesamiento necesarios a los datos de entrada; Luego se ejecuta el modelo XGBoost previamente entrenado. Los resultados producidos incluyen: una medida estimada de éxito, una clasificación adecuada, identificación de los factores más influyentes y un conjunto de recomendaciones personalizado, adecuado a cada caso concreto.
4. Esto va a retornar al backend principal, donde se persiste lo asociado al perfil del emprendedor. Finalmente, el backend la transmite al frontend para su presentación al usuario.



*Ilustración 3: Flujo de Predicción*

### **Sistema de Autenticación de perfiles**

La plataforma segmenta a los usuarios en dos categorías: emprendedores e inversionistas. Durante el proceso de registro, cada usuario escoge que rol es el más adecuado para él, decisión que muestra una distinción tanto en los campos disponibles en su perfil como las funcionalidades específicas a las que tendrá acceso dentro del sistema.

Los emprendedores configuran perfiles en los cuales podrán incluir información personal básica, una descripción detallada de su proyecto de emprendimiento culinario, en que etapa del proyecto se encuentra, las necesidades específicas para la financiación, y la documentación de soporte respectiva.

Por su parte, los inversionistas completan perfiles que recogen sus preferencias de inversión, los rangos de monto disponibles para invertir, los sectores de interés prioritario y los criterios de evaluación que aplican en su proceso de selección de oportunidades.

## **Sistema de Mensajería en Tiempo Real**

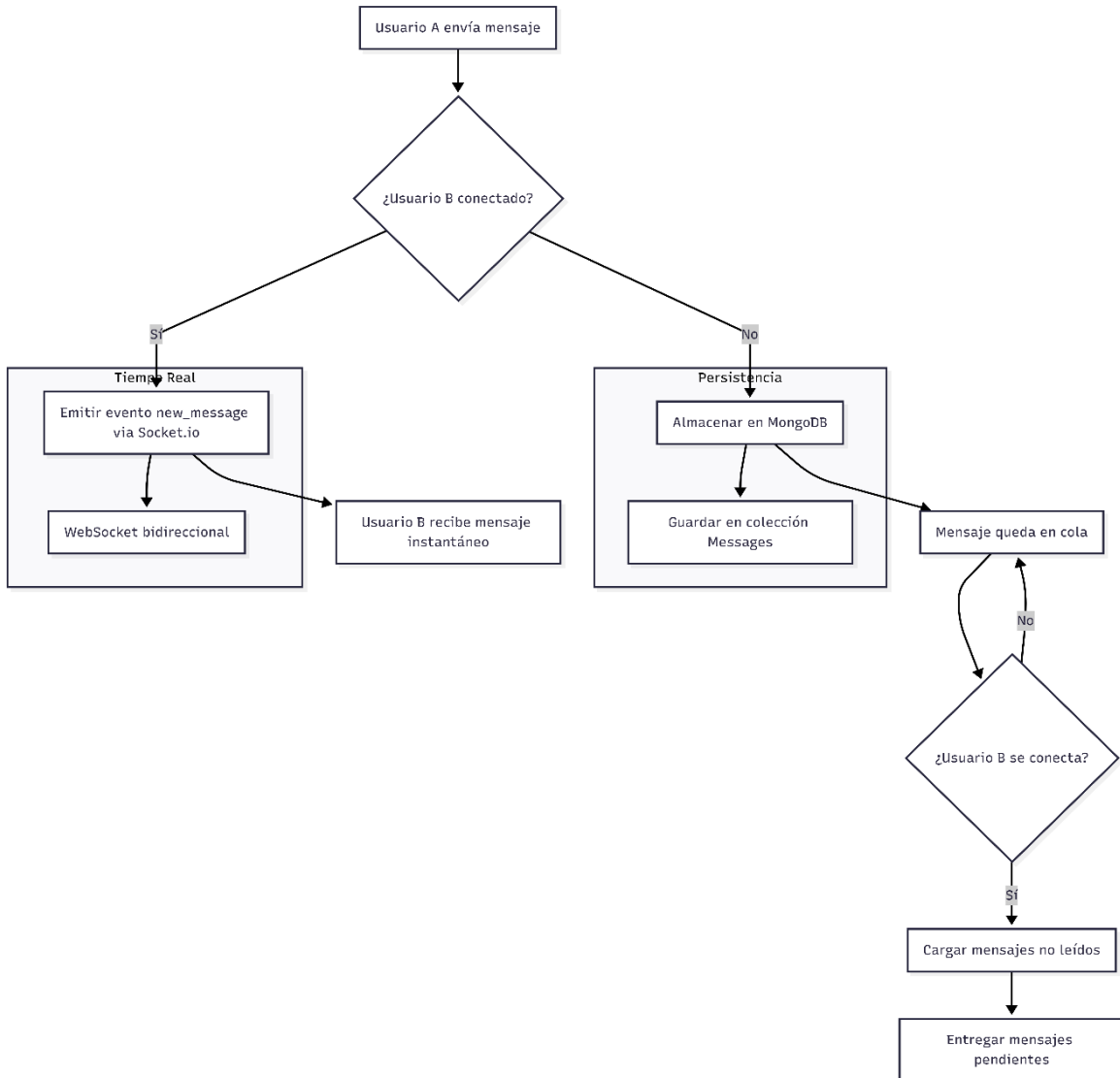
La comunicación entre emprendedores e inversionistas representa una funcionalidad nuclear de la plataforma. El sistema de mensajería permite el intercambio de texto plano entre usuarios, acompañado de notificaciones instantáneas ante la llegada de nuevos mensajes y la persistencia completa del historial de conversaciones.

Socket.io se encarga de gestionar las conexiones bidireccionales en tiempo real, manteniendo un canal persistente y abierto entre cada cliente conectado y el servidor.

Cuando un usuario envía un mensaje, el servidor realiza dos acciones principales de forma inmediata:

1. Almacena el mensaje en la base de datos MongoDB, garantizando su persistencia.
2. Lo transmite en tiempo real al destinatario en caso de que este se encuentre conectado en ese momento.

Si el destinatario no está en línea al momento del envío, el mensaje permanece almacenado en la base de datos y se entrega automáticamente tan pronto como el usuario vuelva a establecer conexión, asegurando así la integridad y continuidad de la conversación.



*Ilustración 4: Diagrama de flujo: Sistema de mensajería en tiempo real*

## Desarrollo del modelo predictivo

### Generación del Dataset

Debido a la falta de bases de datos históricas sobre los emprendimientos culinarios en el sector, se tomó la decisión de generar un dataset sintético basado en estadísticas y hallazgos de investigaciones ya existentes. Este método es válido en el caso que los datos reales no sean accesibles, nos permite demostrar la viabilidad del modelo, mientras se reconoce sus limitaciones.

Se realizó la programación del generador en el sistema Python, que incluyó probabilidades ajustadas basadas en información del Instituto Nacional de Estadística y Censos, el Servicio de Rentas Internas y estudios del Global Entrepreneurship Monitor Ecuador realizados por ESPAE-ESPOL.

Para la función de probabilidad de éxito se consideró algunos factores, basados en literatura académica revisada:

El hecho de que exista un plan de negocios formal aumenta la probabilidad base en 15 puntos porcentuales, esto es consistente con los hallazgos de (Delmar & Shane, 2003) los cuales realizaron una documentación del impacto positivo de la planificación formal para la supervivencia de emprendimientos nuevos.

La experiencia que uno tenga en el sector de la gastronomía añade aproximadamente entre 8 y 15 por ciento de puntos, lo cual se ve reflejado según los metaanálisis de Song et al (2008) sobre los factores de éxito en nuevas empresas.

El nivel de validación de mercado de mercado contribuye progresivamente: un total de 4 puntos son aportados gracias a encuestas, 8 gracias a un producto mínimo viable, y clientes activos añade otros 12. Estos números se fundamentan en los principios de Lean Startup descritos por Ries (2011). La formación educativa del emprendedor modifica la probabilidad en 5 a 8 puntos dependiendo del tipo de formación, sea esta técnica, universitaria, o de postgrado, esto siendo respaldado por las correlaciones documentadas en el GEM Ecuador entre educación y supervivencia empresarial.

La relación entre capital disponible y financiamiento requerido representa un factor positivo cuando hay una ratio favorable (menor a 2:1) y uno negativo cuando el emprendedor depende de manera excesiva de financiamiento externo (mayor a 4:1)

El dataset final se conformo por 5,000 registros de emprendimientos simulados, cada uno caracterizado por 10 variables predictoras identificadas en el marco teórico. La variable objetivo va a indicar si el emprendimiento logró alcanzar los márgenes de éxito (con una supervivencia mayor a 3 años con una rentabilidad positiva) o fracaso. La distribución que quedo resultante refleja aproximadamente 77% de casos exitosos y 23% de fracasos, una proporción consistente con el sesgo de supervivencia observado en datos empresariales reales.

| Característica        | Valor |
|-----------------------|-------|
| Total                 | 5,000 |
| Casos de éxito        | 3850  |
| Casos de fracaso      | 1150  |
| Variables Predictoras | 10    |

*Tabla 3: Distribución de Dataset generado*

La construcción del conjunto de datos se llevó a cabo a través de un procedimiento ordenado desarrollado en Python, el cual integra muestreo aleatorio guiado por distribuciones estadísticas observadas en la realidad con un mecanismo de puntuación ponderada destinado a generar las etiquetas correspondientes a éxito o fracaso de los emprendimientos.

Dicho procedimiento se estructuró en lo que se detalla a continuación:

Cada una de las variables predictoras se construyó mediante funciones de la librería NumPy (específicamente el módulo `numpy.random`), aplicando distribuciones estadísticas ajustadas a datos oficiales y patrones observados en el contexto de la solución software.

Las variables categóricas (etapa en la que se encuentra el proyecto, nivel educativo alcanzado por el emprendedor y existencia de validación de mercado) se generaron utilizando la función `numpy.random.choice()`, asignando probabilidades ponderadas que reproducen las distribuciones reportadas en publicaciones del INEC y del Global Entrepreneurship Monitor para Ecuador.

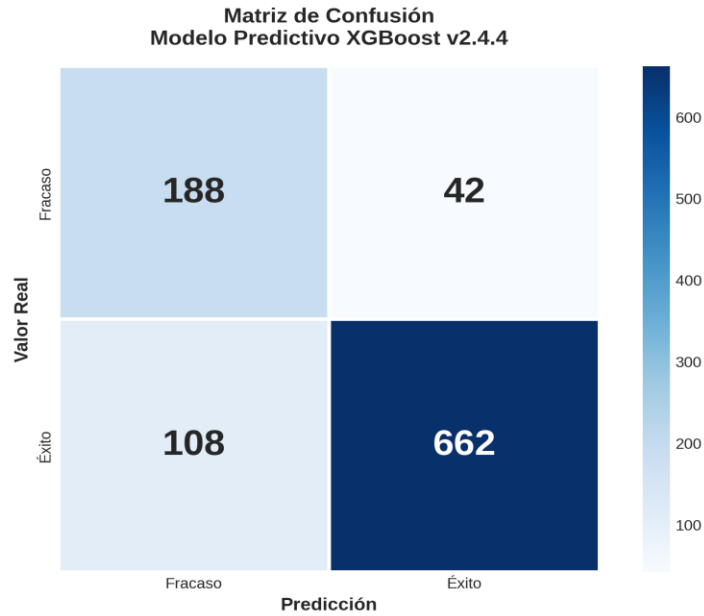
VARIABLES CONTINUAS (años de experiencia empresarial, antigüedad, número de empleados, capital de inversión inicial, ingresos presupuesto esperado y requerido) se toman de una distribución log-normales mediante `numpy.random.lognormal()` recortada al rango real. Estas opciones responde al hecho de que la mayoría de las empresas presentan valores bajos o medios de estos parámetros, mientras que existe una larga tendencia hacia valores altos, Este comportamiento ha sido ampliamente documentado en la literatura sobre distribución empresarial.

El sistema de puntuación se ajustó iterativamente hasta que aproximadamente el 77% de los casos se clasificaron como exitosos y el 23% de los casos se clasificaron como fallidos. Esta relación no se determina al azar sino que se basa en dos premisas principales:

En primer lugar, en las estadísticas publicadas por la Superintendencia de Sociedades Anónimas, Valores y Seguros (SUPERCIAS), se muestra la tasa de supervivencia de las empresas a tres años del 70% al 80% para las empresas nacionales oficialmente registradas.

En segundo lugar, en el sesgo de supervivencia que suele presentarse de forma recurrente en los conjuntos de datos empresariales: los casos de fracaso tienden a estar subrepresentados porque las empresas que cierran en etapas muy tempranas generan pocos o ningún registro histórico accesible, mientras que las que sobreviven acumulan mayor documentación.

## Matriz de confusión del modelo



*Ilustración 5: Matriz de confusión*

La matriz de confusión nos permite visualizar el desempeño del clasificador. De los 1,000 casos de prueba, el modelo clasificó correctamente 662 emprendimientos exitosos (verdaderos positivos) y 188 fracasos (verdaderos negativos). Los 42 falsos positivos representan casos donde el modelo fue optimista, mientras que los 108 falsos negativos indican emprendimientos exitosos que el modelo clasificó conservadoramente como de bajo potencial.

### Variables predictoras

El modelo predictivo se compone de 10 variables predictoras

| Variable           | Tipo       | Descripción               |
|--------------------|------------|---------------------------|
| stage              | Catógórica | Etapas del emprendimiento |
| Years_in_bussiness | Numérica   | Años de operación         |

|                           |          |   |
|---------------------------|----------|---|
| Number_of_employees       | Numérica | Cantidad de empleados                       |
| Funding_needed            | Numérica | Inversión requerida                         |
| Education_level           | Catógica | Nivel educativo del emprendedor             |
| Previous_experience_years | Numérica | Experiencia previa en el sector             |
| Has_bussiness_plan        | Binaria  | Si hay existencia de algún plan de negocios |
| Market_validation_level   | Catógica | Nivel de validación de mercado              |
| Initial_capital           | Numérica | Capital inicial disponible                  |
| Projected_monthly_revenue | Numérica | Ingresos mensuales proyectados              |

*Tabla 4: Variables predictorias*

### **Preprocesamiento de Datos**

Las variables catógicas se codificaron mediante codificación ordinal, asignando valores numéricos que preservan el orden inherente cuando este existe (nivel educativo, validación de mercado y etapa del proyecto).

Esta técnica se adapta mejor con algoritmos basados en árboles como XGBosst que el one-hot ecoding, ya que gracias a que estos pueden aprender divisiones naturales directamente sobre variables numéricas ordenadas, se evita la expansión dimensional innecesaria.

El conjunto de datos tuvo una división en la cual 80% de los registros fueron reservados para entrenamiento y un 20% para la evaluación, con lo cual se mantiene la proporción original de clases mediante estratificación.

Para corregir el desbalance entre clases, se aplicó SMOTE (Synthetic Minority Over-sampling Technique), este algoritmo genera instancias sintéticas de la clase minoritaria mediante interpolación lineal entre ejemplos cercanos en el espacio de características.

(Chawla et al., 2002) demostraron que esta técnica supera al submuestreo aleatorio y al sobremuestreo por replicación directa, al introducir variedad sin perder información relevante. SMOTE se aplicó exclusivamente al conjunto de entrenamiento, logrando un balance perfecto entre las clases de éxito y fracaso antes del ajuste del modelo.

### **Selección y configuración del algoritmo**

Se optó por emplear XGBoost (Extreme Gradient Boosting) como algoritmo principal debido a su desempeño sobresaliente y ampliamente documentado en conjuntos de datos tabulares de tamaño y complejidad moderados. Entre sus ventajas más destacadas se encuentran la capacidad para modelar de forma eficiente relaciones no lineales, el manejo adecuado de variables de tipo mixto (numéricas y categóricas) y su robustez frente a valores faltantes, todo ello sin requerir un preprocesamiento excesivamente intensivo.

Sin embargo, durante la etapa de evaluación y pruebas del modelo se encontró una limitación importante: las probabilidades de salida tenían la tendencia de presentar una polarización marcada, con valores que se acercaban mucho a 0 o a 1 en la mayoría de los casos. Esta distribución no reflejaba fielmente el grado real de incertidumbre asociado a muchos emprendimientos, lo que representaba una reducción considerable de la utilidad práctica de las predicciones para los usuarios finales (emprendedores).

Teniendo en cuenta estas deficiencias se tuvo como propósito de mitigarlas, la configuración de los hiperparámetros se orientó sistemáticamente hacia dos objetivos prioritarios: la corrección progresiva de los sesgos detectados a lo largo del proceso iterativo de refinamiento y ajuste del modelo, y la prevención del sobreajuste.

| <b>Parámetro</b> | <b>Valor</b> |
|------------------|--------------|
| N_estimators     | 200          |
| Max_depth        | 3            |
| Learning_rate    | 0.05         |
| Min_child_weight | 5            |
| gamma            | 0.2          |
| Reg_alpha        | 0.5          |
| Reg_lambda       | 1.0          |
| subsample        | 0.8          |
| Colsample_bytree | 0.8          |

*Tabla 5: Hiperparámetros del modelo XGBoost*

### **Calibración de probabilidades**

En las pruebas preliminares del modelo se detectó un sesgo importante: los emprendedores que no contaban con plan de negocios recibían sistemáticamente probabilidades muy bajas (entre 3 % y 18 %), incluso cuando presentaban otros indicadores claramente positivos, tales como experiencia relevante, validación de mercado mediante producto mínimo viable (MVP) y formación de nivel superior. Este comportamiento si tomamos los propósitos perseguidos por la plataforma, existe una incompatibilidad.

Luego el análisis posterior mostró que XGBoost había aprendido interacciones no lineales en las que la variable `has_business_plan` obtuvo una importancia relativa desproporcionada, alcanzando un 35,24 % de contribución al modelo. Esto representa aproximadamente 2,6 veces el

peso que le correspondía de manera teórica, según la función de puntuación empleada en la generación de los datos sintéticos.

Con el fin de mitigar dicho sesgo, se aplicó una calibración isotónica de las probabilidades mediante la clase `CalibratedClassifierCV` de `scikit-learn`. Esta metodología, propuesta y evaluada por (Niculescu-Mizil & Caruana, n.d.) , ajusta las probabilidades de salida del clasificador para que se alineen más fielmente con las frecuencias reales observadas, empleando regresión isotónica en combinación con validación cruzada de 5 pliegues.

Esta técnica ajusta las probabilidades crudas generadas por el modelo de manera que reflejen con mayor fidelidad la probabilidad real observada en los datos, eliminando la marcada tendencia de XGBoost a producir predicciones extremas e injustificadamente confiadas. El resultado es un sistema de puntuación mucho más interpretable y alineado con las necesidades de apoyo a la toma de decisiones por parte de emprendedores e inversionistas.

### **Entrenamiento y validación**

El entrenamiento del modelo realizó el proceso del conjunto de datos balanceado, ajustando iterativamente los 200 árboles de decisión que componen el ensamble. Cada árbol que le siguió se enfocó en corregir los errores que quedaban de los anteriores, una característica en la cual destaca el gradient boosting.

La validación fue realizada mediante validación cruzada estratificada de 5 pliegues sobre el conjunto el conjunto de entrenamiento, esto ayuda a que las estimaciones del rendimiento de generalización sean más robustas que una partición train-test. Cada que había una iteración 4

pliegues fueron para entrenamiento y el resto para la validación, rotando hasta que todos los pliegues hasta que todos los pliegues aplicaron validación exactamente una vez.

### **Integración como servicio**

El modelo fue serializado mediante joblib y fue desplegado mediante el microservicio FastApi. El modelo se carga en memoria solo una vez, así se evitan problemas de latencia y deserialización en cada petición de predicción realizada.

El endpoint principal de predicción recibe un JSON con las características seleccionadas o descritas del emprendimiento, aplicando transformaciones de preprocesamiento similares o equivalentes a las que se usaron durante el entrenamiento, ejecuta la predicción, y nos da una respuesta que incluye:

- El porcentaje probable de éxito, el cual saca a partir de las probabilidades que XGBoost asigna a cada clase.
- Se da una clasificación basada en la puntuación porcentual, Alto (mayor a 70%), Medio (entre a 40 a 70%), Bajo (menor a 40%).
- Un indicador de confianza que refleja las probabilidades de ambas clases.
- Factores que se indican como positivos o negativos dependiendo de su impacto en la predicción.
- Recomendaciones basadas en el contexto cuando se encuentran carencias en el perfil evaluado.

## Resultados del modelo predictivo

### Métricas de rendimiento

Tras incorporar la calibración isotónica de probabilidades y ampliar significativamente el conjunto de datos de entrenamiento, el modelo alcanzó un nivel de desempeño que justifica su despliegue en un entorno de producción real.

Los resultados obtenidos mediante validación cruzada estratificada de 5 pliegues (5-fold cross-validation) fueron los siguientes:

- **ROC-AUC: 88,48 %:** Esta métrica, que evalúa la capacidad global del modelo para discriminar correctamente entre casos de éxito y fracaso.
- **Precisión: 94 %.** Esto implica que las predicciones positivas tienen una alta fiabilidad: cuando el modelo llega a la conclusión que un emprendimiento se considera viable, la probabilidad que el mismo lo sea es muy alta, reduciendo así el riesgo de que se creen falsas expectativas entre los potenciales inversionistas.
- **F1-Score: 86 %.** Esta medida refleja un equilibrio robusto y consistente entre precisión y exhaustividad (recall), demostrando que el modelo logra identificar una proporción adecuada de casos reales de éxito sin sacrificar excesivamente la calidad de las predicciones.
- **Exactitud (Accuracy): 85 %** en el conjunto de prueba independiente. a pesar que no parezca un porcentaje demasiado elevado, es un resultado más que aceptable y esperable dada la naturaleza conservadora de la calibración

isotónica aplicada, la cual tenía como principal objetivo reducir los falsos positivos y dar probabilidades bien calibradas, lo cual resultaría en una mejor toma de decisiones que una exactitud global inflada.

### **Importancia de las variables**

El análisis de importancia de características, realizado sobre el modelo final evidencia un reajuste significativo en la ponderación relativa de los factores predictivos, consecuencia directa de las correcciones de sesgo aplicadas durante las iteraciones finales:

Plan de Negocios (*has\_business\_plan*): 31,53 %. Continúa siendo el predictor individual de mayor peso, sin embargo, su influencia se redujo de manera sustancial, con el propósito explícito de evitar que la ausencia de un plan formal se traduzca en una penalización excesiva o determinante para el emprendedor.

Etapa del Proyecto (*stage*): 14,34 %. Este resultado refuerza la relevancia crítica del grado de madurez del emprendimiento como indicador clave de probabilidad de supervivencia y éxito a mediano plazo.

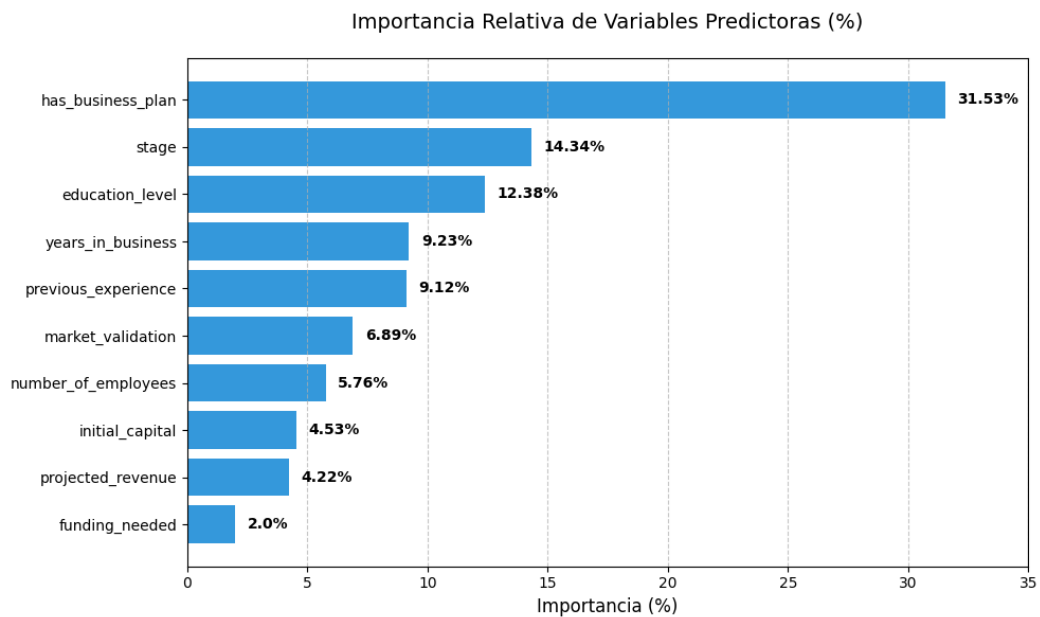
Nivel Educativo (*education\_level*): 12,38 %. Confirma la hipótesis teórica que vincula un mayor nivel de formación académica con capacidades mejoradas de gestión estratégica, financiera y operativa.

- Años de Operación (*years\_in\_business*): 9,23 %.
- Experiencia Previa (*previous\_experience\_years*): 9,12 %.

Ambas variables relacionadas con la trayectoria del emprendedor mantienen una contribución importante y prácticamente equivalente, destacando el valor acumulativo del tiempo de exposición al entorno empresarial.

La existencia de plan de negocios emergió como el factor más determinante con 31.53% de importancia relativa, consistentes en los hallazgos realizados por (Delmar & Shane, 2003) sobre la planificación formal en la supervivencia empresarial.

Por su parte, las variables de carácter estrictamente financiero (como el capital inicial (4,53 %) y los ingresos proyectados (4,22 %)) exhiben una influencia relativamente menor, aunque estable. Estos factores operan principalmente como elementos de ajuste fino que modulan la predicción final una vez establecidos los fundamentos estructurales y de gestión del proyecto.



*Ilustración 6: Variables Predictoras*

| <b>Perfil</b> | <b>Características</b>                                       | <b>Score</b> | <b>Clasificación</b> |
|---------------|--|--------------|----------------------|
| Principiante  | Idea, sin plan, sin experiencia, estudios secundarios        | 28%          | BAJO                 |
| Compensado    | Idea, sin plan, 3 años exp, MVP, postgrado                   | 86.69%       | ALTO                 |
| En desarrollo | Desarrollo, con plan, 2 años exp, MVP, universitario         | 75.77%       | ALTO                 |
| Operando      | Operando, con plan de negocios, 5 años exp, clientes activos | 98.55%       | ALTO                 |

*Tabla 6: Predicciones para perfiles de prueba*

Los resultados obtenidos muestran que el modelo, una vez calibrado, asigna un valor significativo a los factores compensatorios. Así, un emprendedor que carece de plan de negocios formal, pero cuenta con amplia experiencia previa, validación a través de un producto mínimo viable (MVP) y formación de posgrado recibe una evaluación favorable, alcanzando un 86.69 %. En contraste, el perfil carente de cualquier tipo de preparación obtiene una puntuación sustancialmente inferior (27.59 %), lo que refleja adecuadamente el mayor riesgo que conlleva dicha situación.

## **Pruebas de Integración**

### **Comunicación Backend-Microservicio**

Se verificó la comunicación HTTP entre el backend Node.js y el microservicio Python mediante pruebas automatizadas:

| Componente       | Prueba                        | Resultado             |
|------------------|-------------------------------|-----------------------|
| Microservicio ML | Health Check (/health)        | Respuesta correcta    |
| Microservicio ML | Predicción (/predict)         | JSON válido retornado |
| Backend          | Proxy a ML                    | Comunicación exitosa  |
| Backend          | Autenticación JWT             | Tokens válidos        |
| Frontend         | Renderización de predicciones | Componente funcional. |

*Tabla 7: Resultados de prueba de integración*

### **Flujo Completo de Usuario**

Se llevó a cabo una prueba de extremo a extremo del flujo completo de obtención de la predicción, verificando su correcto funcionamiento:

#### **El usuario inicia sesión en la interfaz Front-end mediante su proceso de autenticación.**

1. Completa el formulario correspondiente al perfil del emprendimiento.
2. Solicita la evaluación predictiva del proyecto.
3. El backend valida el token de autenticación y reenvía los datos ingresados al microservicio especializado.
4. El microservicio procesa la información, genera la predicción y devuelve el resultado al backend.
5. El backend registra el resultado en la base de datos y lo remite a la interfaz frontend.

Finalmente, el frontend presenta al usuario una tarjeta de predicción que incluye el puntaje obtenido, la clasificación correspondiente, los factores más influyentes y las recomendaciones personalizadas.

### **Transparencia Algorítmica**

Las predicciones generadas por el modelo deben entenderse como instrumentos de apoyo a la toma de decisiones, y no como determinaciones definitivas o absolutas. Dentro de la interfaz de usuario se refleja esta limitación ya que el mismo indica que los resultados constituyen estimaciones derivadas del análisis de patrones estadísticos, sin que ellos representen resultados 100% reales de los proyectos evaluados.

Asimismo, se muestra información detallada sobre los factores que tuvieron mayor influencia en cada predicción particular, lo que permite al usuario comprender con mayor claridad los elementos que sustentan la evaluación obtenida, por lo tanto tener una valoración más informada y crítica.

### **Validación cualitativa mediante el criterio de un experto**

Tras finalizar la evaluación técnica del modelo predictivo mediante el empleo de métricas estadísticas y las pruebas de integración correspondientes, se llevó a cabo una validación cualitativa complementaria a través de la entrevista semiestructurada planteada en el Capítulo III.

El objetivo principal de esta etapa consistió en confrontar y poner en perspectiva los resultados obtenidos por el algoritmo frente a la apreciación y el criterio práctico de un inversionista con experiencia y actividad actual en el sector gastronómico de Guayaquil.

## **Criterios de evaluación de inversionista con las variables del modelo**

Cuando se le preguntó al entrevistado, el Ing. Lorenzo (inversor del sector culinario de la ciudad de Guayaquil, anteriormente inversor de un restaurante en el sector Kennedy, avenida Delta) cuáles eran los aspectos más relevantes a la hora de decidir si invertir o no en un proyecto gastronómico, señaló tres elementos que considera fundamentales:

- La existencia de procesos estandarizados y de normas operativas claras (lo cual, en su opinión, está estrechamente ligado a contar con un plan de negocios bien estructurado y formalizado).
- La disponibilidad de información financiera histórica que abarque, al menos, los últimos tres años de operación.
- Y, por último, la fortaleza y calidad del equipo humano que dirige y ejecuta el emprendimiento.

Estos criterios coinciden de manera notable con las variables principales incorporadas en la estructura del modelo predictivo desarrollado en esta investigación, exceptuando la disponibilidad de información financiera histórica.

### **El plan de negocios como factor no excluyente**

Al preguntársele si la falta de un plan de negocios formal debería implicar, de manera automática, la exclusión del proyecto como opción de inversión, el Ing. Lorenzo fue tajante: no, no lo descalifica. Explicó que hay numerosos negocios gastronómicos que han alcanzado éxito y rentabilidad sin haber contado inicialmente con estudios o documentos formales previos; muchos surgieron gracias a una buena idea, capital disponible, empuje personal y confianza en el proyecto,

y que la estructura y organización se fueron consolidando sobre la marcha, conforme avanzaba la operación.

Aunque reconoció que disponer de un plan de negocios bien elaborado representa “lo ideal”, ya que proporciona criterios claros y ayuda a reducir riesgos de manera significativa, insistió en que su inexistencia “no invalida ni elimina por completo la posibilidad de invertir en un restaurante”.

### **Percepción de la plataforma como herramienta de apoyo**

En cuanto a la confianza que otorgaría a una herramienta tecnológica capaz de entregar porcentajes estimados de éxito, el Ing. Lorenzo manifestó que los datos que aporta dicha herramienta le proporcionarían “una parte muy importante” del panorama que se requiere para la realización de una primera evaluación del proyecto.

Sin embargo, fue claro con el hecho que su confianza plena dependería de poder complementar esa proyección numérica con una visita directa y una inspección presencial del negocio ya en funcionamiento. En su criterio, “una proyección, por sí sola, nunca resulta determinante para decidir sobre ningún negocio”, lo cual respalda de manera clara la premisa que fue planteada en la sección de Transparencia Algorítmica de este trabajo: las estimaciones que son generadas por el modelo predictivo deben concebirse como herramientas de apoyo en el proceso de toma de decisiones, mas no como veredictos o absolutos.

### **Validación del planteamiento del problema**

En cuanto a la confianza que otorgaría a una herramienta tecnológica capaz de entregar porcentajes estimados de éxito, el Ing. Lorenzo manifestó que los datos aportados por dicha herramienta le proporcionarían “una parte muy importante” del panorama requerido para realizar una primera evaluación del proyecto.

No obstante, precisó que su confianza plena dependería de poder complementar esa proyección numérica con una visita directa y una inspección presencial del negocio ya en funcionamiento. En su criterio, “una proyección, por sí sola, nunca resulta determinante para decidir sobre ningún negocio”, lo cual viene a respaldar de manera clara la premisa planteada en la sección de Transparencia Algorítmica de este trabajo: las estimaciones generadas por el modelo predictivo deben concebirse como herramientas de apoyo en el proceso de toma de decisiones, y no como veredictos o absolutos.

### **Sugerencia de información presentada**

Una vez que el inversionista tuvo oportunidad de revisar una demostración funcional de la plataforma, comentó que los elementos que esta presenta en su análisis de viabilidad —es decir, el porcentaje estimado de éxito, la identificación de factores positivos y negativos, junto con las recomendaciones generadas— no alcanzan por sí solos para tomar una decisión definitiva de inversión.

Explicó que, en su experiencia, es indispensable contar con términos concretos de la negociación, tanto en el plano jurídico como en el financiero. Hizo énfasis, que cualquier inversionista necesita saber con anterioridad y de manera clara que recibe a cambio de su capital: cuál sería su porcentaje de participación accionaria obtendría, cuál sería el esquema de retorno

esperado de la inversión y qué tipo de garantías legales respaldarían la operación. En sus palabras: “Ningún inversionista va a poner dinero únicamente porque le muestren una buena idea o un porcentaje de éxito prometedor”.

Esta apreciación no desvirtúa ni deslegitima el planteamiento de la plataforma; más bien, contribuye a definir con mayor claridad sus límites actuales. Tal como se precisó en la sección de Alcance incluida en el Capítulo I, el desarrollo realizado no incorpora —ni pretende incorporar en esta etapa— funcionalidades que permitan la integración con sistemas financieros, jurídicos o de formalización legal que faciliten o documenten acuerdos entre emprendedores e inversionistas.

La observación del Ing. Lorenzo refuerza precisamente esta delimitación y, al mismo tiempo, señala una línea de evolución lógica y prioritaria para futuras iteraciones del sistema: la incorporación de herramientas que ayuden a estructurar y visualizar los términos básicos de la negociación de inversión.

### **Validación cruzada: Criterio experto frente a predicción algorítmica**

Se expuso al inversionista un caso concreto: un emprendimiento gastronómico que ya cuenta con un producto mínimo viable debidamente validado en el mercado, tres años de trayectoria en el sector, el emprendedor posee formación universitaria, pero no dispone de un plan de negocios formalizado.

El Ing. Lorenzo, apoyándose en su experiencia de que el retorno promedio de la inversión en el rubro gastronómico suele materializarse alrededor de los cinco años, estimó para este perfil una probabilidad de éxito del 66 %. No obstante, precisó que dicha estimación estaría condicionada

a que los registros financieros reales del negocio muestren una recuperación progresiva y proporcional del capital invertido a lo largo de esos tres años de operación.

Con el fin de comparar esta apreciación subjetiva con la salida del modelo predictivo, se procedió a ingresar en la plataforma un perfil equivalente al descrito. El caso más próximo registrado durante las pruebas de validación del modelo es el denominado “Compensado”, que aparece consignado en la Tabla #6 de este capítulo. Este perfil comparte características muy similares (ausencia de plan de negocios formal, tres años de experiencia operativa, validación mediante producto mínimo viable y formación de nivel superior), y el algoritmo le asigna una probabilidad de éxito del 86,69 %.

Resalta una ligera diferencia porcentual dado que el inversionista evalúa el éxito fundamentalmente desde la perspectiva de recuperación financiera de la inversión dentro de un plazo u horizonte temporal definido, según su criterio a 5 años.

Aunque se observa una diferencia notable en los valores cuantitativos obtenidos, ambas valoraciones coinciden en un punto esencial: tanto la apreciación del experto como el resultado del algoritmo consideran al emprendimiento analizado como un proyecto con una probabilidad claramente favorable de éxito.

De esta forma, se descarta la suposición de que la falta de un plan de negocios formal implique, por sí sola, una condena automática al fracaso, siempre y cuando estén presentes otros factores compensatorios de suficiente solidez y consistencia.

## **CAPITULO V. CONCLUSIONES Y RECOMENDACIONES**

En el presente capítulo se exponen las conclusiones principales derivadas del proceso de diseño, desarrollo e implementación de la plataforma digital orientada al apoyo de emprendimientos culinarios en el sector Kennedy de Guayaquil. Asimismo, se formulan recomendaciones específicas dirigidas a fortalecer su mejora continua y a facilitar su eventual escalabilidad.

Los hallazgos que se presentan a continuación responden de manera directa a los objetivos establecidos en el planteamiento de la investigación y se sustentan en la evidencia empírica recopilada a lo largo de las distintas etapas del proyecto: desde el diseño conceptual hasta la evaluación funcional y predictiva del sistema.

### **Conclusiones**

#### **Cumplimiento de los Objetivos de Investigación**

El objetivo general de la investigación, consistente en el desarrollo de una plataforma digital que permita la conexión entre emprendedores culinarios e inversionistas mediante la integración de un modelo predictivo basado en inteligencia artificial, se alcanzó de manera satisfactoria. La solución implementada conforma un ecosistema tecnológico integral que comprende un backend desarrollado en Node.js con el framework Express, un frontend construido en React, una base de datos NoSQL gestionada con MongoDB y un microservicio dedicado al procesamiento de machine learning implementado en Python mediante FastAPI.

En cuanto a los objetivos específicos, se consiguió identificar y formalizar los principales requisitos de financiamiento propios del sector gastronómico al incorporar en el modelo predictivo variables financieras relevantes, tales como el capital inicial requerido, el monto de financiamiento solicitado y los ingresos proyectados.

Por su parte, el modelo predictivo supervisado construido con el algoritmo XGBoost obtuvo un rendimiento de 88,48 % en el indicador ROC-AUC durante la validación cruzada de 5 pliegues, superando con claridad el umbral de desempeño establecido de antemano (80 %). Estos resultados se alinean con hallazgos de investigaciones; por ejemplo, (Razaghzadeh Bidgoli et al., 2024) reportaron valores de precisión comprendidos entre 80 % y 82 % al emplear algoritmos como Random Forest y Gradient Boosting en la predicción del éxito de startups.

### **Relevancia para el Ecosistema Emprendedor Ecuatoriano**

La plataforma desarrollada atiende una necesidad específica y bien documentada del ecosistema emprendedor ecuatoriano. De acuerdo con el informe del Global Entrepreneurship Monitor Ecuador 2023-2024 (Lasio et al., 2023), el país registra una Tasa de Actividad Emprendedora Temprana (TEA) del 32,65 %, lo que lo sitúa como la economía con mayor nivel de emprendimiento. Sin embargo, a pesar de la elevada tasa de creación de nuevos negocios existen importantes dificultades en lo que respecta a la supervivencia empresarial y, grandes dificultades en el acceso a fuentes de financiamiento.

Teniendo esto en consideración, el sector de alimentos y bebidas posee una relevancia e importancia muy grande, ya que representa el 36,9 % del total de establecimientos económicos del país, según datos del Instituto Nacional de Estadística y Censos (INEC, 2023). Esta alta y

significativa proporción justifica la decisión de orientar una solución tecnológica hacia dicho segmento.

La focalización del estudio en el sector Kennedy de Guayaquil permitió recopilar y analizar información contextualizada, directamente vinculada a una realidad local concreta. Esta delimitación geográfica y sectorial contribuyó a aumentar la aplicabilidad práctica de los resultados y evitó caer en generalizaciones que podrían haber comprometido la validez y utilidad del modelo desarrollado.

### **Validación de la arquitectura tecnológica**

La elección de una arquitectura basada en microservicios resultó adecuada para integrar componentes heterogéneos en el sistema desarrollado. Al mantener el modelo de machine learning como un servicio independiente implementado con FastAPI, fue posible avanzar en el desarrollo y en las sucesivas actualizaciones del algoritmo sin generar interrupciones ni dependencias directas en el backend principal.

Gracias a las pruebas de integración que se realizaron se pudo notar que el tiempo de respuesta para las predicciones se mantuvo de manera consistente por debajo de los 500 milisegundos, representando el cumplimiento de los criterios de usabilidad que se definieron en la etapa de diseño.

La selección del algoritmo XGBoost como base del modelo predictivo tuvo una justificación que se ve reflejada plenamente en el contexto del proyecto. Lo que destaca del mismo es su elevada precisión en conjuntos de datos cuyo tamaño se encuentra entre lo pequeño y mediano, también por la capacidad que posee de manejar variables de naturaleza mixta (numéricas

y categóricas) de manera eficiente, y su posibilidad de ofrecer explicaciones claras a través del análisis de importancia de características. Estas cualidades resultaron especialmente relevantes considerando el volumen limitado del conjunto de datos sintéticos empleado (5000 registros).

En relación con el cuarto objetivo específico, la evaluación del modelo predictivo no se restringió exclusivamente al análisis de sus métricas cuantitativas de desempeño, sino que se enriqueció mediante una validación cualitativa basada en una entrevista semiestructurada dirigida a un inversionista con actividad permanente en el sector gastronómico de Guayaquil.

Este enfoque de triangulación permitió confrontar de manera directa las estimaciones generadas por el algoritmo con el juicio práctico y la experiencia de quien, en la realidad del mercado, asume la responsabilidad de decidir sobre inversiones en este rubro.

Los hallazgos obtenidos revelaron coincidencias importantes en puntos clave: tanto el modelo como el inversionista concuerdan en que la falta de un plan de negocios formalizado no debería, por sí sola, actuar como un criterio excluyente o descalificante, siempre que se presenten otros elementos compensatorios de peso significativo, tales como la trayectoria operativa demostrada, la validación efectiva del producto en el mercado y la preparación académica y profesional del emprendedor.

De igual modo, al someterse a consideración un mismo escenario de evaluación, tanto el algoritmo como el experto coincidieron en catalogar el emprendimiento como un proyecto que presenta una probabilidad claramente favorable de alcanzar el éxito.

No obstante, la validación también puso de manifiesto ciertas divergencias constructivas que enriquecen y contribuyen a una comprensión más precisa de las limitaciones presentes en el

modelo actual de la solución software. El inversionista hizo hincapié en la importancia de la disposición de cifras financieras históricas que haya sean debidamente verificables, así como la necesidad de contar con una estructura clara y explícita de los términos de la negociación (entre ellos, el porcentaje de participación accionaria que recibiría el inversionista, las garantías legales que respaldarían la operación y el esquema definido de retorno de la inversión).

Estos aspectos, los cuales no se encuentran incorporados en la versión actual del modelo predictivo, no representan una debilidad insalvable, sino que señalan grados de mejora específicas y de alto valor práctico para las iteraciones futuras del desarrollo.

### Tabla de Trazabilidad

| Objetivo   | Evidencia de Cumplimiento   | Indicador Verificable   |
|--|---|---|
| Desarrollar una plataforma digital que funcione de manera interactiva y que sirva para poner en contacto a emprendedores del rubro gastronómico con posibles inversionistas en la ciudad de Guayaquil. La plataforma incluirá un modelo de inteligencia artificial que ayude a calcular y mostrar el nivel de potencial o probabilidad de éxito que tendría cada proyecto culinario que se presente. | Plataforma web desplegada en producción con arquitectura MERN Stack (MongoDB, Express, React, Node.js) + microservicio ML Python/FastAPI. Sistema de autenticación, perfiles de usuario, sistema de mensajería en tiempo real (Socket.io), y modelo predictivo integrado. | Plataforma final desplegada en el siguiente enlace: <a href="https://sparkling-imagination-production.up.railway.app/">https://sparkling-imagination-production.up.railway.app/</a> |

*Tabla 8: Tabla de Trazabilidad*

## **Recomendaciones**

Se recomienda implementar, de forma progresiva, la recolección de datos reales provenientes de emprendimientos culinarios activos en el sector Kennedy. Este proceso permitiría reentrenar de manera periódica el modelo predictivo, lo cual con el tiempo representaría la sustitución de manera gradual de los datos sintéticos por información empírica obtenida directamente del contexto local.

La incorporación sistemática de datos reales daría una fortaleza significativa de la validez externa de las predicciones, ya habría una reducción de la dependencia de simulaciones y se alinearía de manera más estrecha el comportamiento del modelo con las dinámicas efectivas del ecosistema emprendedor gastronómico de la zona de estudio.

Se da como recomendación la incorporación de un sistema estructurado de retroalimentación que realice un registro de manera sistemática el desempeño real de los emprendimientos que hayan obtenido un financiamiento a través de la plataforma. Este mecanismo de ciclo de retroalimentación da la posibilidad al modelo predictivo de realizar un aprendizaje continuo gracias a la incorporación periódica de nuevos datos empíricos, al tiempo que el mismo podría llevar a cabo una validación prospectiva de su capacidad predictiva. De esta forma, se podría dar una evaluación con mayor rigor de la correspondencia entre las estimaciones generadas por el sistema y los resultados efectivamente observados en el tiempo, dando así un fortalecimiento progresivo de su precisión y confiabilidad en el contexto específico del sector gastronómico local.

La ampliación del estudio hacia otros sectores económicos de Guayaquil surge como una línea de investigación particularmente prometedora. El marco metodológico construido, que

abarca desde la generación de datos sintéticos fundamentada en estadísticas oficiales hasta la arquitectura de integración del componente de machine learning, el mismo tiene una alta capacidad para ser replicado. Y la aplicación de este enfoque podría darse exitosamente en otros ámbitos como el comercio minorista, la manufactura artesanal, los servicios profesionales, entre otros.

En cada caso, se necesitaría identificar y hacer una selección de las variables de predicción más relevantes para el sector en cuestión, así como realizar la correspondiente recalibración del modelo predictivo, haciendo un ajuste en los parámetros y umbrales a las particularidades propias de cada dominio económico.

Otra dirección que podría ser de gran interés para futuras investigaciones es la integración de técnicas de procesamiento de lenguaje natural (NLP) orientadas al análisis de planes de negocio en formato textual. En vez de tratar la existencia del plan de negocios como una variable dicotómica (presente/ausente), sería factible emplear modelos más avanzados capaces de evaluar aspectos cualitativos del documento, tales como su calidad general, grado de completitud, coherencia interna y solidez argumental.

Por último, se propone como línea de trabajo futuro la exploración de modelos de series temporales para modelar y proyectar la evolución del puntaje de éxito a lo largo del ciclo de vida del emprendimiento. Lo cual representa una aproximación de carácter longitudinal que permitiría detectar con mayor precisión los momentos críticos en los que intervenciones puntuales (de carácter financiero, operativo, formativo o estratégico) podrían aumentar de manera significativa la probabilidad, y posibilidad de supervivencia y consolidación del negocio. De esta manera, la plataforma dejaría de limitarse a ofrecer una evaluación estática realizada en un instante determinado para convertirse en un sistema de acompañamiento dinámico y continuo, capaz de

monitorear el trayecto del emprendimiento y de generar recomendaciones adaptadas a cada etapa de su desarrollo.

## Referencias:

- Álvarez, R., Morales, H., & Batista, Y. (2023). La investigación científica o tecnológica y la innovación en los Institutos Superiores Tecnológicos: un debate conceptual. *Universidad y Sociedad*, 15(4), 21-33. [http://scielo.sld.cu/scielo.php?script=sci\\_arttext&pid=S1990-86442023000400021](http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S1990-86442023000400021)
- Angel Investment Network. (n.d.). i. ¿Qué es la inversión ángel? - Red Ecuatoriana de Inversiones. Retrieved December 9, 2025, from <https://www.angelinvestmentnetwork.ec/index.php/aprender/lo-basico/i-que-es-la-inversion-angel>
- Apxml. (n.d.). Advanced Gradient Boosting: XGBoost, LightGBM, CatBoost. Retrieved February 5, 2026, from <https://apxml.com/courses/mastering-gradient-boosting-algorithms>
- Blank, S., & Dorf, B. (n.d.). El manual del emprendedor: La guía paso a paso para crear una gran empresa.
- Buitrago Villalobos, J. A., & Cubillos Naranjo, L. F. (2023). Plan de negocio para una plataforma de intermediación entre inversionistas y emprendimientos ubicados en la ciudad de Bogotá: Caso de estudio: Nómadas digitales en Bogotá (Trabajo de grado, Universidad Distrital Francisco José de Caldas). Repositorio institucional. <https://repository.udistrital.edu.co/server/api/core/bitstreams/ad38e807-b4ba-4d1d-b487-89c1406eda78/content>
- Chandran, M. (n.d.). Los 10 mejores frameworks de aplicaciones web para el desarrollo web. Retrieved January 1, 2026, from [https://www-armia-com.translate.goog/blog/top-10-web-application-frameworks-for-web-development/?\\_x\\_tr\\_sl=en&\\_x\\_tr\\_tl=es&\\_x\\_tr\\_hl=es&\\_x\\_tr\\_pto=tc](https://www-armia-com.translate.goog/blog/top-10-web-application-frameworks-for-web-development/?_x_tr_sl=en&_x_tr_tl=es&_x_tr_hl=es&_x_tr_pto=tc)

- Chapman, P., Clinton, J., Khabaza, T., Reinartz, T., Kerber, R., Shearer, C., & Wirth, R. (1999). CRISP-DM 1.0 Step-by-step data mining guide.
- Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic Minority Over-sampling Technique. *Journal of Artificial Intelligence Research*, 16, 321–357. <https://doi.org/10.1613/JAIR.953>
- Chen, R. J., Lu, M. Y., Chen, T. Y., Williamson, D. F. K., & Mahmood, F. (2021). Synthetic data in machine learning for medicine and healthcare. *Nature Biomedical Engineering* 2021 5:6, 5(6), 493–497. <https://doi.org/10.1038/s41551-021-00751-8>
- Datassential. (2025). Restaurant Failure Rate Plunges in 2025: Datassential Sales Intelligence | Datassential. <https://datassential.com/resource/restaurant-failure-rate/>
- Delmar, F., & Shane, S. (2003). Does business planning facilitate the development of new ventures? *Strategic Management Journal*, 24(12), 1165–1185. <https://doi.org/10.1002/SMJ.349>
- Díaz, D. (2023). PROPUESTA DE MODELO PREDICTIVO DEL PRECIO DEL PAN MEDIANTE USO RNA. “UN CASO DE ESTUDIO SECTOR PANIFICADOR, BOGOTÁ”. <https://repository.uamerica.edu.co/server/api/core/bitstreams/9ad21a1f-ae9f-4b51-83c9-5bbe82dbfe7c/content>
- Durán, B., Lejarriaga Pérez de las Vacas, G., Martín López, S., & Sánchez Espada, J. (2019). Acceso a la financiación de proyectos de emprendimiento social a través de plataformas en red y su relación con los Objetivos de Desarrollo Sostenible (Ponencia). CIRIEC 2019. [https://www.ciriec.uliege.be/wp-content/uploads/2019/06/D7-BEL-DURAN-CIRIEC2019Finalpaper\\_283.pdf](https://www.ciriec.uliege.be/wp-content/uploads/2019/06/D7-BEL-DURAN-CIRIEC2019Finalpaper_283.pdf)

Hernández, R. (2018). Metodología de la investigación. Las rutas cuantitativa, cualitativa y mixta Las rutas Cuantitativa Cualitativa y Mixta. McGRAW-HILL Interamericana Editores S.A. de C.V., 753.  
[https://drive.google.com/file/d/1xJ\\_xF2Zcw20ctWbbqDjsLP5qGyBSk5Vm/view?usp=sharing&usp=embed\\_facebook](https://drive.google.com/file/d/1xJ_xF2Zcw20ctWbbqDjsLP5qGyBSk5Vm/view?usp=sharing&usp=embed_facebook)

Ibm. (2025). MongoDB. What is MongoDB? <https://www.ibm.com/think/topics/mongodb>

IBM. (n.d.). ¿Qué es XGBoost? | IBM. Retrieved January 7, 2026, from <https://www.ibm.com/es-es/think/topics/xgboost>

Introduction | Socket.IO. (2026). <https://socket.io/docs/v4/>

Jacqueline, E., Moran, J., Oswaldo, Á., Saldarriaga, L., Javier, H., De, T., Rosa, L., Xavier, F., Moreno, N., Raúl, J., & Tunja, F. (2024). Crowdfunding como Herramienta para la Generación de Capital Semilla en Start-Ups Ecuatorianas. *Ciencia Latina Revista Científica Multidisciplinar*, 8(1), 545–563. [https://doi.org/10.37811/CL\\_RCM.V8I1.9439](https://doi.org/10.37811/CL_RCM.V8I1.9439)

Kim, J., Kim, H., & Geum, Y. (2023). How to succeed in the market? Predicting startup success using a machine learning approach. *Technological Forecasting and Social Change*, 193, 122614. <https://doi.org/10.1016/J.TECHFORE.2023.122614>

Krauss Delorme, C., Odizzio, A. B., & Volfovicz León, R. (2018). Modelo predictivo de la intención emprendedora universitaria en Latinoamérica. *J. Technol. Manag. Innov*, 13(4). <http://jotmi.org>

Lasio, V., Paula Espinosa, M., Amaya, A., Sarango, P., Dolores Mahauad, M., & Virginia Lasio Adriana Amaya María Paula Espinosa María Dolores Mahauad Paúl Sarango, A. (2023). *Global Entrepreneurship Monitor*. [www.ediloja.com.ec](http://www.ediloja.com.ec)

- Lu, Y., Chen, L., Zhang, Y., Shen, M., Wang, H., Wang, X., van Rechem, C., Fu, T., & Wei, W. (2023). Machine Learning for Synthetic Data Generation: A Review. <https://arxiv.org/pdf/2302.04062>
- McCarthy, P. X., Gong, X., Braesemann, F., Stephany, F., Rizoïu, M. A., & Kern, M. L. (2023). The impact of founder personalities on startup success. *Scientific Reports* 2023 13:1, 13(1), 17200-. <https://doi.org/10.1038/s41598-023-41980-y>
- Moritz, A., Diegel, W., Block, J., & Fisch, C. (2022). VC investors' venture screening: the role of the decision maker's education and experience. *Journal of Business Economics*, 92(1), 27–63. <https://doi.org/10.1007/s11573-021-01042-z>
- Niculescu-Mizil, A., & Caruana, R. (n.d.). Obtaining Calibrated Probabilities from Boosting.
- Obschonka, M., & Audretsch, D. B. (2019). Artificial intelligence and big data in entrepreneurship: a new era has begun. *Small Business Economics*, 55(3), 529-539. <https://doi.org/10.1007/s11187-019-00202-4>
- Ordóñez-Viscaino, V., & Rivera-Prado, M. (2024). Factores fundamentales en la Selección de Plataformas de Marketing de Afiliados para Potenciar el Éxito Empresarial de Emprendedores ecuatorianos en el Marco del Año 2024. 593 *Digital Publisher CEIT*, 9(3), 681–695. <https://doi.org/10.33386/593DP.2024.3.2468>
- Patki, N., Wedge, R., & Veeramachaneni, K. (2016). The Synthetic Data Vault. 2016 IEEE International Conference on Data Science and Advanced Analytics (DSAA), 399–410. <https://doi.org/10.1109/DSAA.2016.49>

Pérez, Yosmaurereen & Cuthbert, Deyvon & Sambola, Dexon. (2022). Modelo predictivo basado en Machine Learning dirigido a PYMES de venta, caso de estudio Bluefields. Ciencia e Interculturalidad. 30. 139-146. 10.5377/rci.v30i01.14267.

Razaghzadeh Bidgoli, M., Raeesi Vanani, I., & Goodarzi, M. (2024). Predicting the success of startups using a machine learning approach. Journal of Innovation and Entrepreneurship, 13(1). <https://doi.org/10.1186/S13731-024-00436-X>

Redacción La Hora. (2024, June 27). Alrededor de 3,3 millones de ecuatorianos han estado involucrados en la creación y puesta en marcha de un emprendimiento. <https://www.lahora.com.ec/archivo/Alrededor-de-33-millones-de-ecuatorianos-han-estado-involucrados-en-la-creacion-y-puesta-en-marcha-de-un-emprendimiento-20240627-0058.html>

Redacción Ventana EBC. (2023, March 31). Emprendimiento: ¿Qué Es y para Qué Sirve? | Ventana EBC. <https://www.ebc.mx/ventana/emprendimiento-que-es-y-para-que-sirve/>

Startupeable. (n.d.). Ronda Semilla | Startupeable. Retrieved December 9, 2025, from <https://startupeable.com/glosario/ronda-semilla/>

Tipos de aprendizaje automático | IBM. (n.d.). Retrieved December 15, 2025, from <https://www.ibm.com/mx-es/think/topics/machine-learning-types>

Toteat. (2026, January 9). Cómo planificar el presupuesto de tu restaurante este 2026 | Toteat Chile. <https://toteat.com/blog/articulo/planificar-presupuesto-restaurante-2026>

WebSphere Application Server. (s. f.). <https://www.ibm.com/docs/es/was/9.0.5?topic=applications-websocket>

Welter, C., Scrimshire, A., Tolonen, D., & Obrimah, E. (2021). The road to entrepreneurial success: business plans, lean startup, or both? *New England Journal of Entrepreneurship*, 24(1), 21–42.  
<https://doi.org/10.1108/NEJE-08-2020-0031>

## Anexos

### Anexo 1. Manual de usuario de la plataforma

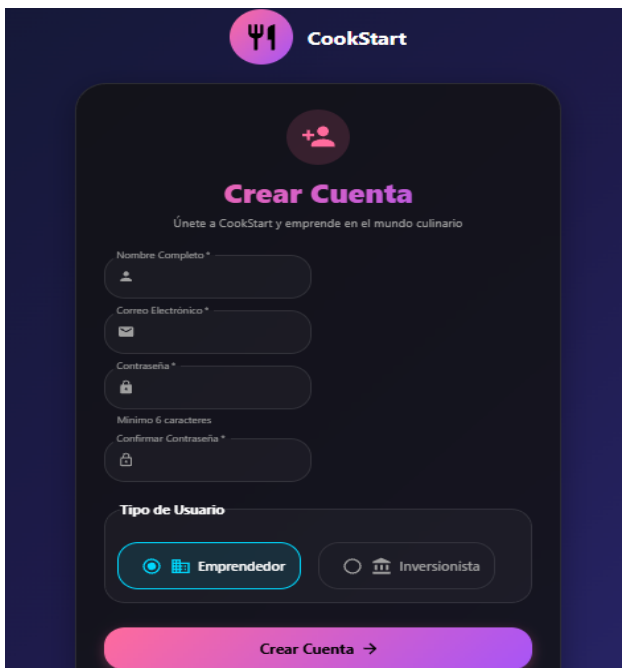
#### Registro de Nuevo Usuario

##### Paso 1: Acceder a la plataforma

Ubicación del botón "Registrarse"

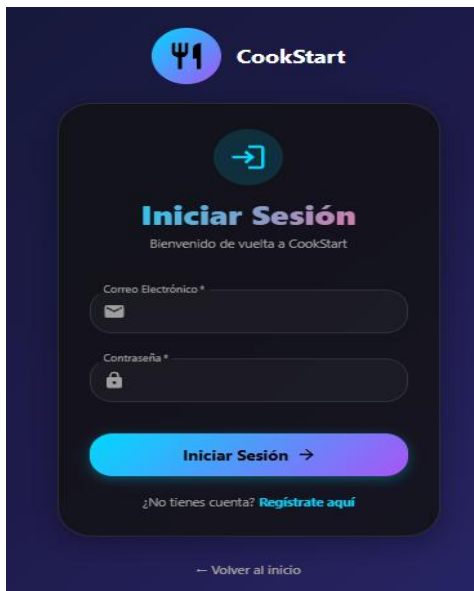


##### Paso 2: Completar formulario de registro

A screenshot of the CookStart registration form. The form is titled "Crear Cuenta" and has the subtitle "Únete a CookStart y emprende en el mundo culinario". It contains the following fields: "Nombre Completo \*" with a person icon, "Correo Electrónico \*" with an envelope icon, "Contraseña \*" with a lock icon, and "Confirmar Contraseña \*" with a lock icon. Below these fields is a section for "Tipo de Usuario" with two radio buttons: "Emprendedor" (selected) and "Inversionista". At the bottom is a large button labeled "Crear Cuenta →".

**Paso 3:** Confirmación automática de registro (redireccionamiento directo al Dashboard principal).

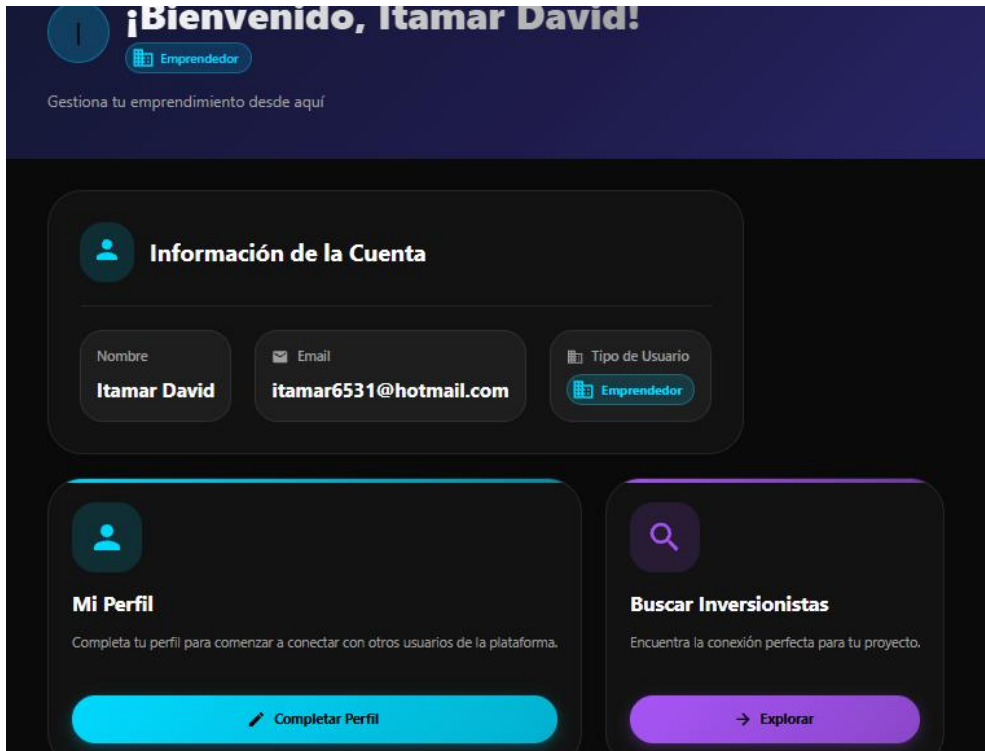
**Paso 4:** Inicio de sesión



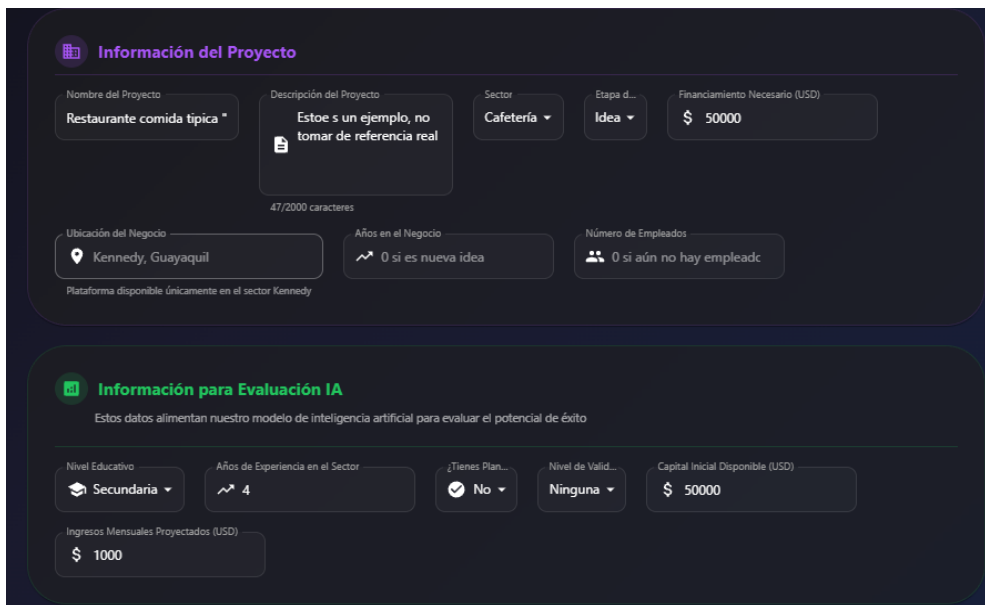
**Funcionalidades por tipo de usuario (Emprendedor / Inversionista)**

**Gestión del Perfil de Emprendedor**

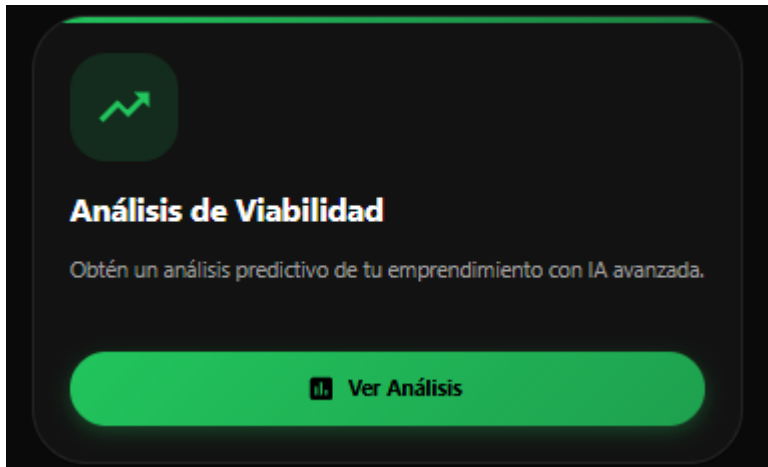
Para completar la información del proyecto, navegar por “Mi perfil” y señalar “Completar Perfil”



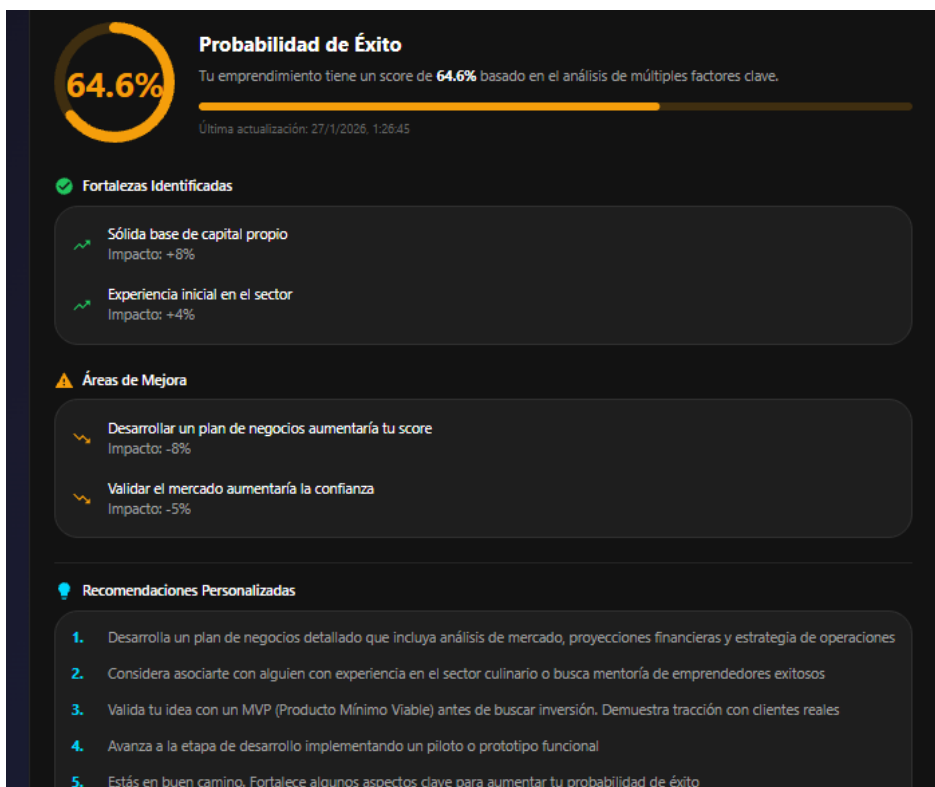
Al entrar en “Completar Perfil” se visualizará su perfil de usuario, donde puede llenar los campos según lo requiera su negocio y, a la vez se encuentra un cuestionario personal que se usará como referencia para el modelo predictivo.



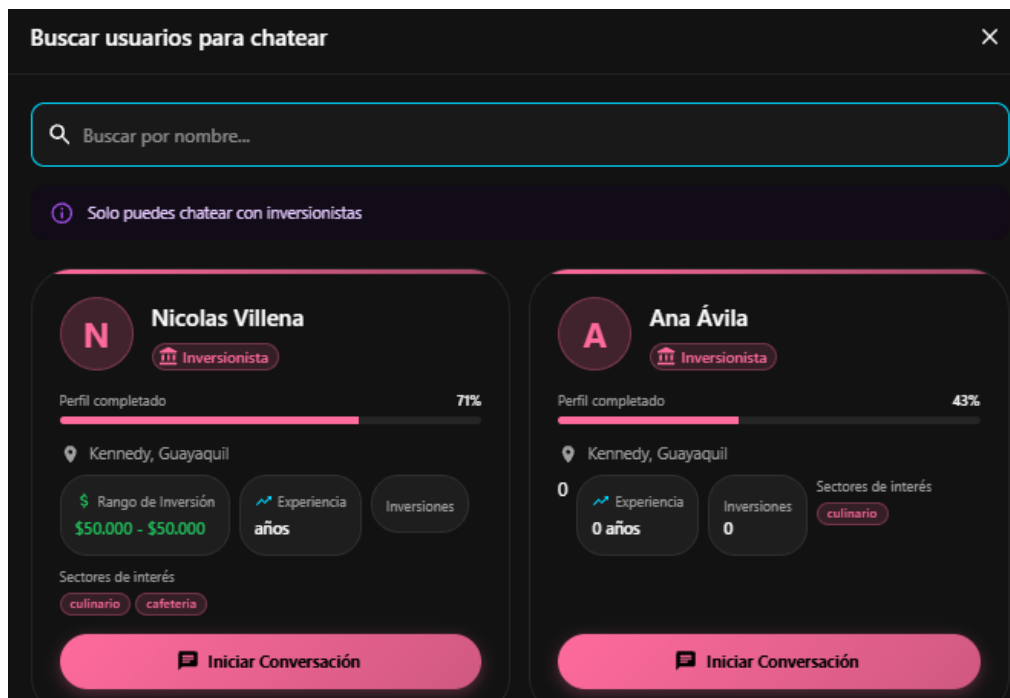
Una vez completado el perfil, clic en “Guardar Perfil” y se podrá visualizar su ponderación de éxito de emprendimiento en el Dashboard, mediante un icono “Análisis de viabilidad”.



Dentro, se visualizará su Score de éxito seguido de recomendaciones que se pueden tomar para poder mejorar dicho Score.



Para la búsqueda de inversionistas, en el Dashboard se encuentra un apartado llamado de nombre “Buscar Inversionistas” con un botón de explorar. Aquí se mostrarán los inversionistas que han creado un perfil en la plataforma y se puede iniciar conversación con aquellos. Se mostrarán en qué tipo de negocio culinarios están interesados (cafés, restaurantes, carretas, etc.) y el rango de inversión que están dispuestos a pagar.



Al iniciar sesión con un inversionista, automáticamente se lo redireccionará al apartado de “Mensajes” donde podrá conversar con el inversor en cuestión, en tiempo real.



## **Anexo 2. Transcripción de entrevista realizada al Ing. Blas Lorenzo.**

**1. Cuando usted evalúa un emprendimiento culinario para una posible inversión, ¿cuáles son los tres factores principales que considera determinantes para decidir si invierte o no?**

"Realmente son varios. Uno de los principales viene a ser los procesos. Si un negocio se ha estructurado con levantamiento de información van a tener estandarizados los procesos y las normas a seguir. Ese es un punto primordial.

Segundo, cifras. Si estamos hablando de un negocio que ya está funcionando, hay que hacer un avalúo de los últimos 3 años. Tú necesitas registros o información histórica de años anteriores para poder promediar o aproximarte a un promedio en saber si esas cifras han sido constantes o no.

Tercero, el equipo humano. El inversionista cree en el producto, pero el equipo humano o las personas que están dentro de ese bote y que van a remar es lo que hacen importante el negocio- ¿Quién se va a hacer cargo de eso? ¿Quién te va a manejar plan y marketing? ¿Quién va a manejar la parte operativa?

**2. ¿Considera usted que la ausencia de un plan de negocios formal debería descalificar automáticamente a un emprendimiento culinario, o existen otros factores que podrían compensar dicha carencia?**

"No considero yo en mi experiencia que lo descalificaría. Tú para poner un negocio lo puedes hacer de dos formas, una a través del plan de negocios, que es lo óptimo. Un plan de

negocios no te garantiza el éxito al 100%, pero lo que hace es que te brinda los parámetros y directrices para minimizar los riesgos.

La otra forma es porque tiene el dinero en el banco o en el colchón, tiene una idea, tiene el tiempo, el ímpetu, las ganas y se lanza con mucha fe. Hay negocios que han empezado así, se han ido haciendo y desarrollando en el camino y son negocios prósperos y exitosos hoy en día.

En lo personal estoy muy seguro de que eso no descalifica la opción de inversión en un restaurante",.

**3. Si una herramienta tecnológica le presentara un porcentaje estimado de éxito de un emprendimiento basado en variables como experiencia del emprendedor, etapa del negocio, nivel educativo, y validación de mercado, ¿en qué medida confiaría en dicha estimación para complementar su análisis personal?**

"A ver, yo con eso tendría gran parte de la información, pero creo yo que habría que aportar algo más que me genere credibilidad. Y la única forma para obtener la credibilidad en coherencia a lo que me estás diciendo de los parámetros, creo yo que podría ser incorporar una visita al establecimiento totalmente libre de horarios, ver el funcionamiento".

**4. Actualmente, ¿qué dificultades enfrenta para identificar y conectar con emprendedores culinarios que buscan financiamiento en Guayaquil?**

"Te diría que es todo lo opuesto [a no encontrar emprendedores]. Emprendedores tú encuentras a diario y en cualquier tipo de plataforma o red social. Lo que hay es una muy limitada

exposición de inversionistas. Hay una sobreoferta de emprendimientos, pero lo que no hay es canales idóneos para contactar con inversionistas o líneas de crédito".

**5. Tras observar una demostración de la plataforma, ¿considera que la información presentada en el análisis de viabilidad (porcentaje de éxito, factores positivos y negativos, recomendaciones) es suficiente y clara para apoyar una decisión de inversión preliminar?**

"Una proyección no es determinante para ningún negocio. Cuando ya se da la negociación final entre el inversionista y el emprendedor, lo que van a manejar es recurso económico. Yo le voy a dar al inversionista \$50,000. ¿Qué obtengo a cambio? Ningún inversionista va a invertir dinero si no le da algo a cambio. Tienes que cubrirte de una serie de garantías porque créeme que ningún inversionista te va a dar dinero para emprender en un negocio si no tiene las garantías legales".

**6. ¿Qué variable o información adicional considera que debería incluirse en el modelo predictivo (o en la plataforma) para que su evaluación sea más completa o realista desde su perspectiva como inversor?**

*El ingeniero sugiere tres adiciones:*

1. **Garantías:** "Incluir un campo en el cual se diga: 'a cambio de los \$50,000 yo doy el 30% de participación accionaria'... o 'los pago en 3 años con una tasa de retorno del 6%'. Ustedes tienen que buscar aproximadamente tres tipos de negociaciones".

2. **Marketing:** "Enfoque en soportar con mayor peso y carga las estrategias de marketing, promoción y publicidad".

3. **Visitas in situ:** "Incorporar un apartado para poder realizar una visita al establecimiento totalmente libre de horarios".

**7. En su experiencia, ¿qué porcentaje aproximado de éxito asignaría usted a un emprendimiento culinario que cuenta con un producto mínimo viable validado, tres años de experiencia en el sector, formación universitaria, pero sin un plan de negocios formal?**

"A ver, en el tema de los negocios hay una norma general que aplica: que tú en 5 años recién recuperas lo invertido. Me hablas de 3 años, entonces estaríamos hablando de aproximadamente el 66%.

Pero yo en esos 3 años, ¿qué es lo que voy a analizar? Si el señor invirtió \$100,000, significa que por año debe tener un retorno de inversión de \$20,000. Si tú me dices 'a ver, sí ha funcionado los 3 años... pero los gastos son altos y han salido tablas' significa que algo no está bien



## DECLARACIÓN Y AUTORIZACIÓN

Nosotros, **Santacruz Moncayo, Itamar David**, con C.C: # **0929006898** y **Burgos Calle, Mario Enrique**, con C.C # **0958178428** autores del proyecto de tecnología de información “**Desarrollo de una plataforma interactiva de emprendimientos para la conexión entre emprendedores e inversionistas, aplicando modelos predictivos que permita ponderar el nivel de éxito que puede tener un emprendimiento. Caso de estudio: emprendimientos culinarios**” requerido para la obtención del título de **Ingeniero en Ciencias de la Computación** en la Universidad Católica de Santiago de Guayaquil.

1.- Declaramos tener pleno conocimiento de la obligación que tienen las instituciones de educación superior, de conformidad con el Artículo 144 de la Ley Orgánica de Educación Superior, de entregar a la SENESCYT en formato digital una copia del referido trabajo de titulación para que sea integrado al Sistema Nacional de Información de la Educación Superior del Ecuador para su difusión pública respetando los derechos de autor.

2.- Autorizamos a la SENESCYT a tener una copia del referido trabajo de titulación, con el propósito de generar un repositorio que democratice la información, respetando las políticas de propiedad intelectual vigentes.

Guayaquil, 2 de marzo del 2026

---

Nombre: **Santacruz Moncayo, Itamar David**

C.C: **0929006898**

---

Nombre: **Burgos Calle, Mario Enrique**

C.C: **0958178428**

# REPOSITORIO NACIONAL EN CIENCIA Y TECNOLOGÍA

## FICHA DE REGISTRO DE TESIS/TRABAJO DE TITULACIÓN

|   |  |  |        |
|---|--|--|--------|
| TEMA Y SUBTEMA:   | Desarrollo de una plataforma interactiva de emprendimientos para la conexión entre emprendedores e inversionistas, aplicando modelos predictivos que permita ponderar el nivel de éxito que puede tener un emprendimiento. Caso de estudio: emprendimientos culinarios.  |  |        |
| AUTOR(ES)   | Santacruz Moncayo, Itamar David<br>Burgos Calle, Mario Enrique   |  |        |
| REVISOR(ES)/TUTOR(ES)                                       | Erazo Ayón, José Miguel  |  |        |
| INSTITUCIÓN:  | Universidad Católica de Santiago de Guayaquil  |  |        |
| FACULTAD:   | Ingeniería   |  |        |
| CARRERA:  | De Computación   |  |        |
| TÍTULO OBTENIDO:  | Ingeniero en Ciencias de la Computación  |  |        |
| FECHA DE PUBLICACIÓN:                                       | 2 de marzo del 2026  | No. DE PÁGINAS:  | 117 p. |
| ÁREAS TEMÁTICAS:  | Inteligencia artificial, Plataformas digitales, Análisis de datos, Sistemas de información, Industria alimentaria, Emprendimiento, Financiación.   |  |        |
| PALABRAS CLAVES/<br>KEYWORDS:                               | Plataforma web, Inteligencia Artificial, Modelos de predicción, Ciencia de datos   |  |        |
| RESUMEN/ABSTRACT:   | <p>El presente trabajo de integración curricular tiene como objetivo principal el desarrollo e implementación de una plataforma digital interactiva orientada a vincular emprendedores del sector culinario con potenciales inversionistas en el sector Kennedy de la ciudad de Guayaquil. La propuesta surge ante la necesidad documentada de reducir la brecha entre la alta tasa de actividad emprendedora registrada en y las limitadas oportunidades de acceso a financiamiento formal que enfrentan los negocios gastronómicos en sus primeras etapas. La solución fue desarrollada sobre una arquitectura MERN (MongoDB, Express.js, React, Node.js) con un microservicio en Python-FastAPI dedicado al procesamiento del modelo de machine learning, el cual emplea el algoritmo XGBoost con calibración isotónica y alcanza un 88,48 % en la métrica ROC-AUC. La plataforma permite a los emprendedores registrar sus propuestas, obtener una evaluación objetiva de viabilidad y comunicarse directamente con inversionistas mediante un sistema de mensajería en tiempo real. Este proyecto es el resultado de la investigación y aplicación de conocimientos adquiridos a lo largo de la Carrera de Ingeniería en Ciencias de la Computación de la Universidad Católica Santiago de Guayaquil, en áreas como ingeniería de software, bases de datos, inteligencia artificial y desarrollo web, las cuales hicieron posible la construcción de esta solución tecnológica orientada al ecosistema emprendedor local.</p> |  |        |
| ADJUNTO PDF:  | <input checked="" type="checkbox"/> SI   | <input type="checkbox"/> NO  |        |
| CONTACTO CON AUTOR/ES:                                      | Teléfono:+593-97-9035134<br>Telefono:+593-96-0541463   | E-mail: <a href="mailto:itamar6531@gmail.com">itamar6531@gmail.com</a> |        |
| CONTACTO CON LA INSTITUCIÓN (COORDINADOR DEL PROCESO UTE):: | Toala Quimí, Edison José<br>Teléfono: +593-990-976776<br>E-mail: edison.toala@cu.ucsg.edu.ec   |  |        |
| <b>SECCIÓN PARA USO DE BIBLIOTECA</b>                       |  |  |        |
| Nº. DE REGISTRO (en base a datos):                          |  |  |        |
| Nº. DE CLASIFICACIÓN:                                       |  |  |        |
| DIRECCIÓN URL (tesis en la web):                            |  |  |        |