

# **Desarrollo de un Agente de Inteligencia Artificial para la Automatización de Reservas de Vuelos y Alojamiento**

## **1. Justificación**

Se ha seleccionado como enfoque de este proyecto, la planificación de viajes ya que es una actividad recurrente en diversos sectores, especialmente en el ámbito empresarial y de asistencia ejecutiva. La búsqueda de vuelos y alojamientos adecuados, así como la optimización de costos y tiempos de desplazamiento, requieren tiempo y análisis detallado. En muchas empresas que ofrecen servicios de asistencia ejecutiva, la gestión de reservas de viaje es una de las tareas que mayor tiempo demanda, representando una oportunidad para la automatización mediante inteligencia artificial (IA).

Este trabajo se enmarca en el desarrollo de un agente de IA diseñado para optimizar la toma de decisiones en la planificación de viajes. Se busca que el agente sea capaz de analizar múltiples opciones de vuelos y alojamientos en función de las necesidades específicas del usuario, considerando variables como costo, duración del viaje, ubicación, y tendencias de demanda.

Para lograr este objetivo, se llevó a cabo un análisis exploratorio de datos (EDA) utilizando conjuntos de datos provenientes de diferentes fuentes que contienen información detallada sobre reservaciones, patrones de viaje y preferencias de los usuarios. A partir de este análisis, se identificaron patrones clave que servirán de base para el desarrollo del modelo de IA.

Este documento presenta la metodología utilizada, los resultados del análisis exploratorio de datos y los siguientes pasos en el desarrollo del sistema. Se sigue la metodología CRISP-

ML(Q), que permite estructurar el proceso de modelado de manera eficiente y garantizar la calidad de los datos antes de la implementación del modelo.

2. Arquitectura del Agente de AI – Enfocado en viajes



Imagen 1:

Componente	Descripción
👤 Usuario	Persona que solicita la búsqueda y reserva de vuelos y hoteles. Proporciona parámetros como destino, fechas, presupuesto y preferencias.
💻 Interfaz de Entrada (Google Colab)	Punto de interacción del usuario con el sistema. En esta fase piloto, el usuario ingresa datos manualmente a través de scripts en Google Colab.
📡 Módulo de Percepción	Recibe y procesa la entrada del usuario. Extrae características clave como fechas, presupuesto y preferencias de viaje.
📊 Fuente de Datos (Datasets Kaggle)	En esta fase piloto, utilizamos datasets preprocesados en lugar de APIs en vivo para alimentar el modelo con información de vuelos, hoteles y tendencias de viajes.
🧠 Motor de IA (Modelo ML)	Componente central que procesa los datos de entrada y genera recomendaciones de vuelos y hoteles basadas en criterios optimizados. Utiliza modelos de Machine Learning como MLP y regresión para evaluar las opciones.
📁 Módulo Cognitivo	Evalúa las opciones disponibles, establece reglas de decisión y filtra los resultados según la mejor opción para el usuario. Puede aplicar ranking basado en factores como costo, escalas y preferencias personales.
📤 Salida de Recomendaciones	Lista de vuelos y alojamientos sugeridos según la preferencia del usuario y las condiciones del mercado disponibles en los datasets.
🔄 Módulo de Acción	Gestiona la entrega de la recomendación final al usuario. En futuras iteraciones, podrá conectarse con APIs de reservas para realizar la compra automáticamente.
📈 Módulo de Aprendizaje	Se retroalimenta con los datos de uso y la selección del usuario para mejorar futuras recomendaciones. En la fase actual, este proceso es manual, pero se puede integrar aprendizaje supervisado en el futuro.
🔍 Modelo en Producción (Pruebas Controladas)	Fase de validación en la que se evalúan los resultados antes de escalar el sistema. Se simulan distintos escenarios para ajustar la precisión del modelo.
🚀 Futuro: APIs Externas	Una vez validado el modelo en entorno controlado, se podrá integrar con APIs de aerolíneas y plataformas de reservas para automatizar completamente el proceso.

Tabla 1: Creación Propia sobre arquitectura del Agente de AI para Turismo

## Explicación de los Módulos

### 1. Módulo de Percepción

**Entrada sensorial:** En este caso, en lugar de sensores físicos, el sistema recibe datos de entrada del usuario mediante un formulario en Google Colab.

**Extracción de características:** Identifica los parámetros clave del usuario, como destino, fecha y presupuesto.

**Reconocimiento de patrones:** Se analizan datos previos de viajes en los datasets para encontrar tendencias relevantes.

### 2. Módulo Cognitivo

**Representación de objetivos:** Se define el objetivo del usuario (por ejemplo, encontrar la opción más económica o con menos escalas).

**Planificación:** Se generan estrategias basadas en reglas predefinidas y machine learning.

**Toma de decisiones:** Se selecciona la mejor combinación de vuelos y hoteles en función de los criterios del usuario y la disponibilidad en los datos.

### 3. Módulo de Acción

**Presentación de resultados:** Se entrega la lista de opciones recomendadas al usuario.

**Ejecución:** En futuras versiones, el módulo podrá conectarse con APIs para realizar la compra automáticamente.

#### 4. Módulo de Aprendizaje

**Aprendizaje supervisado:** Se entrenará con datos históricos de reservas exitosas.

**Aprendizaje por refuerzo:** En iteraciones futuras, el sistema podrá evaluar qué recomendaciones funcionan mejor y ajustar el modelo de forma automática.

#### Flujo del Agente de IA

- El usuario proporciona información sobre su viaje (destino, fechas, presupuesto, preferencias).
- Google Colab recibe los datos y el módulo de percepción los procesa para extraer variables clave.
- El motor de IA (modelo ML) analiza las opciones en los datasets y genera recomendaciones basadas en optimización de costos, tiempos y preferencias.
- El módulo cognitivo filtra y ordena las opciones según reglas predefinidas.
- Se presentan las mejores opciones al usuario en la salida de recomendaciones.
- El módulo de acción entrega las recomendaciones al usuario, quien puede simular cambios o ajustar criterios.
- El módulo de aprendizaje registra los datos de selección del usuario para mejorar futuras sugerencias.
- En la fase de pruebas controladas, se validan las predicciones antes de implementar el sistema en producción.
- A futuro, el sistema podrá conectarse a APIs externas para realizar reservas en aerolíneas y hoteles de manera automática.

#### 3. Análisis Exploratorio de Datos (EDA)

El análisis exploratorio de datos es una etapa crítica para comprender la información disponible y evaluar su calidad antes de proceder con el modelado. Se analizaron tres datasets principales:

- **Traveler Trip Data:** Contiene información sobre destinos, duración del viaje, medios de transporte, costos de alojamiento y perfil de los viajeros.
- **Expedia Travel Dataset:** Incluye registros de búsqueda y reserva de vuelos y hoteles, con detalles sobre ubicación del usuario, fechas y preferencias.
- **Travel Review Rating Dataset:** Proporciona evaluaciones de usuarios sobre hoteles y otros alojamientos, lo que permitirá analizar la percepción de calidad en las reservas.

### 3.1 Estructura y Calidad de los Datos

Se analizaron los principales aspectos estructurales de los datasets:

- **Cantidad de registros y columnas:** Se validó la dimensionalidad de los datos, identificando que la información disponible permite evaluar múltiples factores en la toma de decisiones.
- **Tipos de variables:** Se categorizó cada variable en numérica o categórica, facilitando su posterior procesamiento.
- **Valores faltantes:** Se identificaron registros con valores nulos en campos críticos como “orig\_destination\_distance” y “Transportation cost”, lo que motivó la aplicación de estrategias de imputación.
- **Detección de valores atípicos:** Se encontraron outliers en costos de transporte y alojamiento, lo que llevó a la implementación de técnicas de reducción de sesgo.

### 3.2 Análisis Univariante

Se examinaron las distribuciones de variables clave:

- **Duración de los viajes:** La mayoría de los viajes tienen una duración de entre 7 y 10 días, con algunos casos de estancias más prolongadas.
- **Tipos de alojamiento:** Se observó que la mayoría de los viajeros prefieren hoteles, seguidos por Airbnb y villas.
- **Costos de transporte:** Se evidenció una gran variabilidad en los costos, con vuelos presentando la mayor dispersión.
- **Meses con mayor actividad de viajes:** Se identificaron picos en los meses de mayo y octubre, indicando una tendencia estacional en la demanda.

### 3.3 Análisis Bi/Multivariante

- **Correlaciones:** Se encontró que la duración del viaje tiene una relación moderada con el costo del transporte, y que el tipo de alojamiento influye significativamente en el costo total del viaje.
- **Tendencias Temporales:** Se observaron aumentos en la demanda de viajes en ciertas épocas del año, lo que sugiere la necesidad de considerar la estacionalidad en la recomendación de opciones.
- **Comparación de Costos por Tipo de Alojamiento:** Se confirmó que los hoteles tienen un rango de precios medio, mientras que las villas presentan los costos más elevados.

## 4. Resultados y Hallazgos

El análisis de los datos permitió extraer información clave para la construcción del modelo de IA:

- Los meses de mayor demanda de viajes son mayo y octubre, lo que sugiere la necesidad de considerar factores estacionales.
- El costo del alojamiento está influenciado por la categoría del hospedaje, siendo las villas la opción más costosa.
- Existen valores atípicos en costos de transporte y alojamiento, lo que refuerza la necesidad de un modelo de predicción de precios.
- Los usuarios prefieren medios de transporte más económicos en viajes largos, mientras que en viajes cortos eligen conveniencia sobre precio.

Estos hallazgos serán fundamentales para diseñar el agente de IA y sus mecanismos de recomendación.

## **5. Siguiendo pasos**

Con base en el análisis exploratorio realizado, se plantea el siguiente plan de acción:

1. Definir el modelo de IA para la recomendación de vuelos y alojamientos utilizando aprendizaje supervisado.
2. Entrenar y evaluar el modelo con los datasets obtenidos, optimizando la precisión de las predicciones.
3. Incorporar análisis de predicción de precios mediante modelos de series temporales.
4. Validar la efectividad del sistema con simulaciones basadas en datos reales, asegurando que las recomendaciones sean óptimas.

Aunque la integración con APIs de aerolíneas y hoteles sería ideal en el futuro, para la fase actual del proyecto la validación se realizará mediante simulaciones con los datos recopilados.

## 6. Conclusiones

El análisis exploratorio de datos ha permitido identificar patrones clave en la planificación de viajes, sentando las bases para la automatización de este proceso mediante inteligencia artificial. Se han identificado factores críticos que deben ser considerados en el modelado del sistema, como la variabilidad de precios, la estacionalidad de la demanda y las preferencias de los usuarios.

El desarrollo del agente de IA permitirá optimizar la toma de decisiones en la gestión de reservas de viaje, reduciendo tiempos de búsqueda y costos, y mejorando la experiencia del usuario.

## 7. Referencias

### Metodología

- Visengeriyeva, L., Kammer, A., Bär, I., Kniesz, A., & Plöd, M. (2023). *CRISP-ML(Q). The ML Lifecycle Process*. MLOps. Recuperado de <https://ml-ops.org/content/crisp-ml>

### Lecturas de Referencia

- Kumar Mukhiya, S., & Ahmed, U. (2020). *Hands-On Exploratory Data Analysis with Python*. Packt Publishing. Recuperado de <https://learning.oreilly.com/library/view/hands-on-exploratory-data/9781789537253/>

### Fuentes de Datos

- Kaggle. (2024). *Traveler Trip Data*. Recuperado de <https://www.kaggle.com/datasets/rkiattisak/traveler-trip-data>
- Kaggle. (2024). *Expedia Travel Dataset*. Recuperado de <https://www.kaggle.com/datasets/jacopoferretti/expedia-travel-dataset>
- Kaggle. (2024). *Travel Review Rating Dataset*. Recuperado de <https://www.kaggle.com/datasets/wirachleelakiatiwong/travel-review-rating-dataset>