

## Objetivo

**El objetivo principal de este proyecto es aplicar técnicas de Aprendizaje Automático para abordar un problema de interés utilizando un DataSet relevante**

En este caso vamos a usar el dataset “16\_3\_02\_Oferta\_Hotelera\_por-sector” extraído de la web <https://ipiec.tierradelfuego.gob.ar/estadisticas-economicas-2/>

En Resumen, algunos de los pasos que se aplicaran

Extracción de Datos: Extraer la información y convertirla a un formato estructurado como un DataFrame de pandas.

Preprocesamiento de Datos: Limpiar y preparar los datos para el análisis.

Exploración de Datos: Realizar un análisis exploratorio para entender mejor las características y la distribución de los datos.

Modelo de Aprendizaje Automático: Seleccionar y entrenar modelos adecuados de aprendizaje automático.

Evaluación y Mejora: Evaluar los modelos y mejorar su rendimiento según sea necesario.

El DataFrame creado con los datos. Aquí están los primeros registros de test:

	Periodo	Establecimientos_Hotelero	Plazas_Disponibles_Hotelero
0	2008 enero	43	96224
1	febrero	44	87580
2	marzo	44	96069
3	abril	43	90570
4	mayo	42	90737

Habitaciones_Unidades_Disponibles_Hotelero	Establecimientos_Parahotelero
38936	104
34945	103
38998	103
36480	94
37386	77

Plazas_Disponibles_Parahotelero	Habitaciones_Unidades_Disponibles_Parahotelero
70649	21452
64467	19372
68359	20864
57410	17750
43314	13488

## Siguientes pasos:

**Preprocesamiento de Datos:** Convertir la columna Periodo en un formato de fecha. Rellenar o manejar valores nulos si los hay.

Escalar o normalizar datos según sea necesario.

**Exploración de Datos:** Analizar la distribución de las diferentes características. Identificar patrones y correlaciones.

**Modelado:** Definir un objetivo de modelado claro (por ejemplo, predecir la ocupación hotelera).

Seleccionar modelos apropiados (regresión, clasificación, series temporales, etc.).

Entrenar y evaluar los modelos.

**En esta primera entrega, debe proporcionar una descripción detallada de su proyecto de Aprendizaje Automático.**

Descripción del Proyecto de Aprendizaje Automático: Análisis de la Oferta Hotelera y Parahotelera en Ushuaia

### **Objetivo del Proyecto**

El objetivo principal de este proyecto es aplicar técnicas de Aprendizaje Automático para analizar y predecir la oferta hotelera y parahotelera en Ushuaia, Provincia de Tierra del Fuego, Antártida e Islas del Atlántico Sur, utilizando datos históricos de establecimientos, plazas disponibles y habitaciones/unidades disponibles. Este análisis busca proporcionar insights valiosos para los interesados en la industria del turismo, permitiendo una mejor planificación y toma de decisiones.

### **Descripción del Dataset**

El dataset contiene datos mensuales sobre la oferta hotelera y parahotelera desde enero de 2008 hasta julio de 2010. Las variables incluidas son:

**Periodo:** Mes y año del registro.

**Establecimientos\_Hotelero:** Número de establecimientos hoteleros.

**Plazas\_Disponibles\_Hotelero:** Número de plazas disponibles en el sector hotelero.

**Habitaciones\_Unidades\_Disponibles\_Hotelero:** Número de habitaciones o unidades disponibles en el sector hotelero.

**Establecimientos\_Parahotelero:** Número de establecimientos parahoteleros.

**Plazas\_Disponibles\_Parahotelero:** Número de plazas disponibles en el sector parahotelero.

**Habitaciones\_Unidades\_Disponibles\_Parahotelero:** Número de habitaciones o unidades disponibles en el sector parahotelero.

### **Metodología**

#### **Extracción y Preparación de Datos**

**Extracción de Datos:** Los datos se extrajeron de un archivo y se convirtieron a un DataFrame de pandas.

**Preprocesamiento:** Incluye la conversión de la columna Periodo a un formato de fecha, manejo de valores nulos y escalamiento de los datos si es necesario.

#### **Análisis Exploratorio de Datos (EDA)**

**Visualización de Datos:** Utilización de gráficos para entender la distribución de las variables.

**Estadísticas Descriptivas:** Cálculo de medias, medianas, desviaciones estándar y otras estadísticas relevantes.

**Correlación:** Análisis de la correlación entre las diferentes variables para identificar relaciones importantes.

#### **Modelado Predictivo**

**Selección de Modelos:** Elección de modelos de aprendizaje automático adecuados, como regresión lineal, árboles de decisión o redes neuronales.

**Entrenamiento de Modelos:** Dividir los datos en conjuntos de entrenamiento y prueba, y entrenar los modelos utilizando el conjunto de entrenamiento.

**Evaluación de Modelos:** Medir el rendimiento de los modelos utilizando métricas apropiadas como MAE, MSE, RMSE o  $R^2$ .

#### **Optimización y Validación**

**Optimización de Hiperparámetros:** Ajuste de los hiperparámetros del modelo para mejorar su rendimiento.

**Validación Cruzada:** Uso de técnicas de validación cruzada para asegurar que el modelo generaliza bien a datos no vistos.

### **Interpretación y Comunicación de Resultados**

**Interpretación de Resultados:** Análisis de los resultados obtenidos para extraer conclusiones significativas.

**Visualización de Resultados:** Presentación de los resultados a través de gráficos y visualizaciones claras y comprensibles.

**Informe Final:** Elaboración de un informe detallado que resuma los hallazgos y las recomendaciones basadas en el análisis.

### **Resultados Esperados**

**Patrones de Temporada:** Identificación de patrones estacionales en la oferta hotelera y parahotelera.

**Predicciones de Oferta:** Modelos predictivos que puedan estimar la disponibilidad futura de plazas y habitaciones.

**Insights para la Industria:** Recomendaciones basadas en los datos que puedan ayudar a los gestores hoteleros y planificadores de turismo a mejorar sus estrategias de negocio.

### **Herramientas y Tecnologías**

**Lenguajes de Programación:** Python

**Bibliotecas de Análisis de Datos:** pandas, numpy

**Bibliotecas de Visualización:** matplotlib, seaborn

**Bibliotecas de Aprendizaje Automático:** scikit-learn, tensorflow/keras (opcional)

**Entornos de Desarrollo:** Jupyter Notebook

### **Impacto y Aplicaciones**

Este proyecto tiene el potencial de beneficiar a diversos stakeholders en la industria del turismo en Ushuaia. Los resultados pueden ser utilizados para mejorar la gestión de recursos, optimizar la capacidad hotelera según la demanda estacional, y mejorar la experiencia del cliente al anticipar y satisfacer mejor las necesidades de alojamiento.

Al aplicar técnicas de Aprendizaje Automático, se puede transformar la manera en que se analizan y utilizan los datos históricos, proporcionando una ventaja competitiva en el mercado turístico.

**Formule claramente el objetivo de su trabajo, indicando qué problema específico de su interés abordará con el model**

## **Objetivo del Trabajo**

### **Formulación del Objetivo**

El objetivo de este trabajo es desarrollar un modelo predictivo utilizando técnicas de Aprendizaje Automático para predecir la oferta de plazas disponibles en el sector hotelero y parahotelero en Ushuaia, Provincia de Tierra del Fuego, Antártida e Islas del Atlántico Sur. El modelo se entrenará con datos históricos de establecimientos, plazas disponibles y habitaciones/unidades disponibles con el fin de identificar patrones y tendencias estacionales que permitan anticipar la disponibilidad de plazas y mejorar la planificación y gestión de recursos en la industria del turismo.

### **Problema Específico a Abordar**

**Problema:** La industria del turismo en Ushuaia experimenta fluctuaciones significativas en la oferta y demanda de plazas hoteleras y parahoteleras debido a factores estacionales y otras variables. Esta variabilidad puede llevar a una subutilización de recursos durante períodos de baja demanda y a la falta de disponibilidad durante picos de alta demanda, afectando tanto a los proveedores de alojamiento como a los turistas.

**Específicamente, este trabajo abordará los siguientes aspectos:**

**Predicción de la Disponibilidad de Plazas:** Desarrollar un modelo que prediga la cantidad de plazas disponibles en el sector hotelero y parahotelero para futuros períodos, permitiendo una mejor planificación y gestión de recursos.

**Identificación de Patrones Estacionales:** Analizar los datos históricos para identificar patrones y tendencias estacionales en la oferta de plazas, facilitando la toma de decisiones estratégicas.

**Optimización de Recursos:** Proveer a los gestores hoteleros y planificadores de turismo herramientas predictivas que les permitan optimizar la capacidad de alojamiento, ajustando la oferta según las previsiones de demanda.

### **Relevancia del Problema**

Anticipar la disponibilidad de plazas hoteleras y parahoteleras es crucial para:

**Mejorar la Experiencia del Cliente:** Asegurar que los turistas encuentren alojamiento adecuado en cualquier época del año.

**Optimizar la Utilización de Recursos:** Minimizar los costos asociados con la subutilización o sobrecarga de las instalaciones de alojamiento.

**Planificación Estratégica:** Ayudar a los gestores a desarrollar estrategias de marketing y promociones basadas en previsiones de demanda, incrementando la eficiencia operativa y la rentabilidad.

Al abordar estos problemas mediante el desarrollo de un modelo predictivo, se espera proporcionar a la industria del turismo en Ushuaia herramientas más efectivas para gestionar la oferta de alojamiento, mejorando tanto la satisfacción del cliente como la eficiencia operativa.

**Proporcione detalles sobre el contexto en el que se encuentra el problema y por qué es relevante.**

### **Contexto del Problema**

**Ubicación:** Ushuaia, Provincia de Tierra del Fuego, es un destino turístico clave en Argentina, conocido por su paisaje natural y actividades estacionales.

**Estacionalidad:** El turismo en Ushuaia es altamente estacional, con picos en verano (diciembre a febrero) e invierno (junio a agosto).

**Oferta Diversificada:** Incluye una gama de establecimientos hoteleros y parahoteleros, desde hoteles de lujo hasta cabañas y campings.

### **Relevancia del Problema**

#### **Satisfacción del Cliente:**

**Disponibilidad de Alojamiento:** La falta de plazas durante picos de demanda afecta negativamente la experiencia del turista.

**Calidad del Servicio:** La sobrecarga de establecimientos puede deteriorar la calidad del servicio.

#### **Eficiencia Operativa y Rentabilidad:**

**Gestión de Recursos:** La subutilización en temporada baja y la sobrecarga en temporada alta impactan la eficiencia y rentabilidad.

**Costos Operativos:** Los costos de mantener establecimientos vacíos durante la temporada baja son significativos.

**Planificación Estratégica:**

**Marketing:** Comprender las tendencias de demanda permite desarrollar estrategias de marketing más efectivas.

**Inversión en Infraestructura:** Datos predictivos informan decisiones de inversión, equilibrando oferta y demanda.

**Sostenibilidad del Destino:**

**Impacto Económico:** Un turismo bien gestionado contribuye al desarrollo económico sostenible de la región.

**Preservación Ambiental:** Gestionar el flujo turístico ayuda a minimizar el impacto ambiental.

**Justificación del Proyecto**

Desarrollar un modelo predictivo para la oferta de plazas en Ushuaia es crucial para mejorar la experiencia del turista, la eficiencia operativa de los establecimientos, y la sostenibilidad económica y ambiental de la región. Este proyecto proporciona herramientas basadas en datos que permiten tomar decisiones informadas, abordando los desafíos de la estacionalidad y la gestión de recursos de manera óptima.

**Defina claramente las preguntas de investigación o hipótesis que desea responder a través del aprendizaje automático.**

**Preguntas de Investigación e Hipótesis**

**Preguntas de Investigación**

¿Cuáles son los patrones estacionales en la oferta de plazas disponibles en el sector hotelero y parahotelero de Ushuaia?

Propósito: Identificar los períodos de alta y baja demanda para optimizar la gestión de recursos y la planificación estratégica.

¿Qué variables (por ejemplo, número de establecimientos, habitaciones disponibles) influyen más significativamente en la disponibilidad de plazas en ambos sectores?

Propósito: Determinar los factores clave que afectan la disponibilidad de plazas para mejorar las predicciones y la gestión de la oferta.

¿Es posible predecir la cantidad de plazas disponibles en el sector hotelero y parahotelero para futuros períodos utilizando modelos de aprendizaje automático?

Propósito: Desarrollar un modelo predictivo que ayude a anticipar la demanda y ajustar la oferta de manera proactiva.

¿Cómo varía la precisión de las predicciones al usar diferentes técnicas de aprendizaje automático?

Propósito: Comparar el rendimiento de varios modelos de aprendizaje automático para seleccionar el más adecuado.

¿Cuál es el impacto de eventos específicos (por ejemplo, ferias, festivales, condiciones climáticas) en la oferta y demanda de plazas?

Propósito: Analizar cómo eventos específicos afectan la disponibilidad de plazas y ajustar las estrategias de planificación en consecuencia.

### **Hipótesis**

Hipótesis 1: Existen patrones estacionales claros en la disponibilidad de plazas en Ushuaia, con picos de demanda durante los meses de verano e invierno.

Justificación: La estacionalidad en el turismo es común en regiones con actividades dependientes del clima.

Hipótesis 2: El número de establecimientos y el número de habitaciones disponibles son los factores más influyentes en la disponibilidad de plazas.

Justificación: Estos factores directamente reflejan la capacidad de alojamiento disponible.

Hipótesis 3: Los modelos de aprendizaje automático pueden predecir con precisión la disponibilidad de plazas futuras basándose en datos históricos.

Justificación: Los modelos de aprendizaje automático son capaces de identificar patrones complejos en grandes conjuntos de datos.

Hipótesis 4: Los modelos avanzados, como las redes neuronales o los modelos de bosque aleatorio, ofrecerán una mayor precisión en las predicciones comparados con modelos más simples como la regresión lineal.

Justificación: Los modelos avanzados pueden capturar mejor las no linealidades y las interacciones complejas entre las variables.

Hipótesis 5: Eventos específicos, como festivales o condiciones climáticas extremas, tienen un impacto significativo en la disponibilidad de plazas, causando variaciones inesperadas en la demanda.

Justificación: Los eventos especiales y las condiciones climáticas son factores externos que pueden influir fuertemente en la demanda de alojamiento.

### **Objetivo de las Preguntas e Hipótesis**

Estas preguntas de investigación y hipótesis guiarán el análisis y el desarrollo de modelos predictivos en este proyecto. Al responder a estas preguntas y probar estas hipótesis, el objetivo es proporcionar a la industria del turismo en Ushuaia herramientas basadas en datos para mejorar la planificación, gestión y optimización de la oferta de alojamiento, asegurando una experiencia turística satisfactoria y una operación eficiente y rentable para los establecimientos de alojamiento.