

Challenge: Surf Forecasting and Spot Recommendations

Contexto

Estás desarrollando un sistema avanzado de predicción de olas utilizando Deep Learning para series temporales. Este sistema no solo determinará si será un buen día para surfear en los puntos clave de Playa Grande (Biología, Yacht Club y PG), sino que también evaluará otras playas de Mar del Plata para recomendar el mejor lugar para surfear en los próximos días.

Objetivo

1. Predicción del surf:

- Entrenar un modelo que evalúe si será un buen día de surf en cualquiera de los tres puntos de Playa Grande.
- Basar la predicción en datos históricos de Surfline obtenidos a través de dos métodos: scraping y la biblioteca pysurfline.

2. Recomendación de spots:

- Analizar características de todas las playas de Mar del Plata para asignar un puntaje de calidad (surf score) a cada lugar.
 - Implementar un **modelo de lenguaje (LLM)** que haga recomendaciones personalizadas y justifique su elección.
-

Formato de los datos

1. Archivos pysurfline: Contienen información detallada sobre las condiciones del surf, incluyendo:

- Swells (altura, período, dirección).
- Viento (velocidad y dirección).
- Temperatura, presión atmosférica y probabilidad de buen surf.

2. Archivos scrap: Resumen las condiciones con menos detalle, pero incluyen:

- Rango de altura de olas.
- Clasificación (e.g., "POOR TO FAIR").
- Swells y viento.

Fases del Challenge

Fase 1: Exploración y Preprocesamiento

- Unifica los datos de ambos formatos (pysurfline y scrap).
- Gestiona diferencias en resoluciones y formatos.
- Maneja valores nulos o inconsistencias, asegurando calidad en los datos.

Fase 2: Modelado Predictivo

- Utiliza modelos de Deep Learning para series temporales, como:
 - **LSTMs** o **GRUs** para capturar patrones temporales.
 - **Transformers** si deseas explorar arquitecturas más avanzadas.
- Etiqueta los datos como "Buen día" o "Mal día" basándote en parámetros como altura de olas, dirección del viento, y clasificación de las olas.
- Evalúa el modelo con métricas como **accuracy**, **precision**, y **recall**.

Fase 3: Recomendación de Spots

- Diseña una métrica personalizada para puntuar cada playa en Mar del Plata:
 - Considera swells, viento, temperatura, y estacionalidad.
- Crea un sistema de recomendación basado en la puntuación y características históricas de cada spot.

Fase 4: Implementación de LLM

- Entrena o utiliza un modelo de lenguaje como GPT o Llama para:
 - Generar recomendaciones en lenguaje natural.
 - Justificar la elección del mejor lugar para surfear basándose en las predicciones y análisis.

Fase 5: Evaluación

- Valida las predicciones y recomendaciones con datos reales o históricos.
- Incorpora retroalimentación de surfistas locales para afinar la herramienta.

Entregables

1. **Jupyter Notebook** que incluya:

- Análisis exploratorio de datos.
- Implementación del modelo de predicción.
- Evaluación y visualización de los resultados.

2. **Demo funcional:**

- Herramienta que permita cargar nuevos datos y obtener predicciones.
- Interfaz para visualizar recomendaciones y justificaciones del LLM.