

Guía de Exposición - Predicción de Pasajeros con LSTM + Grid Search

1. Problema de negocio

La dirección del aeropuerto de Madrid-Barajas requiere una herramienta para anticipar el volumen de pasajeros mensual y así ajustar la plantilla necesaria en seguridad, atención, limpieza y operaciones. Un modelo predictivo permite optimizar recursos según la demanda esperada.

2. Enfoque técnico

Se ha implementado un modelo de series temporales multivariadas utilizando LSTM (Long Short-Term Memory), capaz de capturar patrones secuenciales. Para mejorar el rendimiento, se aplicó Grid Search manual sobre hiperparámetros como:

- Número de unidades LSTM
- Dropout
- Learning rate

3. Dataset y preparación

Se utilizaron datos públicos de Eurostat sobre vuelos desde/hacia Madrid. Se filtraron rutas con más de 10 años de datos mensuales y se creó una serie temporal por ruta (fecha - número de pasajeros).

Preprocesado:

- Separación en entrenamiento (<2024) y test (=2024)
- Escalado con MinMaxScaler
- Secuencias de entrada de 12 meses

4. Entrenamiento y selección

Se probaron diferentes combinaciones de hiperparámetros. Se seleccionó la combinación con menor pérdida MSE final. El mejor modelo se reentrenó y se guardó para inferencia.

5. Evaluación y resultados

Se realizaron predicciones para los 12 meses de 2024 y se compararon con los datos reales.

Métricas:

- MAE y RMSE por ruta
- Ranking de las 5 rutas con menor error (mejor predicción)

Visualización:

Comparativa entre predicciones y datos reales para las 5 rutas mejor modeladas.

6. Conclusiones

- El uso de Grid Search mejora notablemente el ajuste del modelo
- El LSTM anticipa tendencias para la mayoría de rutas con alta precisión
- Ideal para planificación operativa en aeropuertos

7. Futuras mejoras

- Incluir variables externas: clima, festivos, eventos
- Ampliar granularidad si se dispone de datos diarios
- Explorar modelos multisalida o attention-based