

Predicción de Precipitaciones Utilizando Datos Históricos y Aprendizaje Profundo

1. Introducción

El pronóstico preciso de las precipitaciones es crucial para diversas aplicaciones, incluyendo la agricultura, la gestión de recursos hídricos, la planificación urbana y la mitigación de desastres naturales. La disponibilidad de datos históricos de lluvia, combinada con técnicas de aprendizaje profundo, ofrece una oportunidad para desarrollar modelos de predicción más precisos y confiables. Este proyecto busca desarrollar un modelo de predicción de precipitaciones para la región de interés de ESPOLETA Tecnologías, utilizando datos históricos proporcionados por La Oficina Nacional de Administración Oceánica y Atmosférica (NOAA, sigla en inglés).

2. Problema a Resolver

ESPOLETA Tecnologías, enfocada en el desarrollo de estaciones meteorológicas, carece actualmente de un predictor de precipitaciones. Este proyecto se centra en desarrollar un modelo de aprendizaje profundo que, utilizando datos históricos de lluvia del NOAA, pueda predecir la probabilidad y cantidad de precipitación en las regiones específicas del país. El objetivo es proporcionar una herramienta valiosa que pueda integrarse en las estaciones meteorológicas de ESPOLETA, mejorando su funcionalidad y ofreciendo un servicio de pronóstico más completo.

3. Datos Disponibles

Se utilizarán datos históricos de lluvia de la región, publicados por el NOAA. Estos datos incluirán información sobre la cantidad de precipitación, la fecha y hora de la medición, y la ubicación geográfica de la estación meteorológica.

- Preprocesamiento:
 - Limpieza de datos para manejar valores faltantes o erróneos.
 - Normalización o estandarización de los datos para mejorar el rendimiento del modelo.
 - Creación de características adicionales (features) como promedios móviles de lluvia, tendencias estacionales, etc.
 - División de los datos en conjuntos de entrenamiento, validación y prueba.

4. Metodología

Dado que estamos tratando con datos de series temporales (datos de lluvia a lo largo del tiempo), utilizaré una Red Neuronal Recurrente (RNN), específicamente una variante llamada Long Short-Term Memory (LSTM). Estas redes son muy buenas para aprender patrones en secuencias de datos.

- Modelo:
 - Arquitectura: Una red LSTM con múltiples capas. Arquitectura sencilla con el menor número de capas posible mientras sea funcional.
 - Capa de Entrada: Recibe la secuencia de datos de lluvia históricos.
 - Capas LSTM: Capturan las dependencias temporales en los datos.
 - Capa Densa: Reduce la dimensionalidad y produce la predicción final.
 - Capa de Salida: Una neurona con función de activación sigmoide (para predecir la probabilidad de lluvia).
 - Implementación: PyTorch.
 - Función de Pérdida Cross: Binary -Entropy.
 -

5. Evaluación del Modelo

- Métricas:
 - Precisión y Recall: Para evaluar la capacidad del modelo para identificar correctamente los días de lluvia y los días sin lluvia.
 - F1-score: Una medida combinada de precisión y recall.
- Visualización:
 - Gráficos de las predicciones del modelo versus los valores reales en el conjunto de prueba.
 - Matrices de confusión para evaluar el rendimiento del modelo en la clasificación de días de lluvia/no lluvia.

6. Conclusiones

Se espera que este proyecto demuestre la viabilidad de utilizar modelos de aprendizaje profundo, específicamente redes LSTM, para predecir precipitaciones utilizando datos históricos, como los del NOAA en este caso. El modelo resultante podría integrarse en las estaciones meteorológicas de ESPOLETA Tecnologías, en un futuro, proporcionando un servicio de pronóstico de precipitaciones valioso y mejorando la competitividad de la empresa.