

Deep Learning y Sistemas Inteligentes - Laboratorio 8 -

Instrucciones:

- Deben unirse a uno de los grupos de Canvas de nombre "Laboratorio 8 #", donde N es un número entre 1 y 15. Los grupos pueden ser de máximo 5 personas.
- Esta actividad debe realizarse en grupos.
- Sólo es necesario que una persona del grupo suba el trabajo a Canvas.
- No se permitirá ni se aceptará cualquier indicio de copia. De presentarse, se procederá según el reglamento correspondiente.

Task - Práctica

Se le proporcionan 5 años de datos de ventas de artículos de una cadena de tiendas y se le pide que prediga 3 meses de ventas para 50 artículos diferentes en 10 tiendas diferentes. Descarguen el dataset de Kaggle [aquí](#) por favor. Para esto considere los siguiente:

1. Preparación de datos:
 - a. Limpieza de datos: Maneje valores faltantes, anomalías o valores atípicos en el conjunto de datos.
 - b. Transformación de datos: Convierta los datos a un formato adecuado para modelar, incluyendo normalización o escalado si es necesario.
2. Preprocesamiento de datos:
 - a. División de series temporales: Divida el conjunto de datos en conjuntos de entrenamiento, validación y prueba. Para datos de series temporales, asegúrese de que el conjunto de prueba contenga los datos más recientes, que abarquen los 3 meses que desea pronosticar.
 - b. Generación de secuencias: Cree secuencias de datos, donde cada secuencia representa una ventana fija de datos históricos de ventas y el objetivo correspondiente (el pronóstico de 3 meses).
3. Selección de modelo:
 - a. Decida si desea utilizar LSTM, GRU o Conv1D (o una combinación de estos) como arquitectura para su modelo. Cada uno tiene sus puntos fuertes, por lo que puede experimentar con los tres y elegir el que funcione mejor. Pero, solo es necesario **uno** de los modelos
4. Arquitectura modelo:
 - a. Diseñe la arquitectura de la red neuronal, que puede incluir:
 - i. Capas LSTM/GRU/Conv1D con unidades o filtros adecuados.
 - ii. Capas adicionales como capas Densas para predicciones finales.
 - iii. Capas de dropout para evitar el overfitting.
 - iv. Otros componentes como la normalización por batches, funciones de activación y funciones avanzadas como mecanismos de atención.
5. Entrenamiento modelo:
 - a. Entrene el modelo utilizando los datos de entrenamiento y válidelos en el conjunto de validación. Puede seguir estos pasos:
 - i. Defina una función de pérdida adecuada (por ejemplo, error cuadrático medio) para tareas de regresión.
 - ii. Elija un optimizador (por ejemplo, Adam o RMSprop).
 - iii. Establezca hiperparámetros como learning rate, tamaño de batch y número de épocas.
 - iv. Supervise el rendimiento del modelo en el conjunto de validación y aplique un early stopping si es necesario.
6. Evaluación del modelo:
 - a. Evalúe el rendimiento del modelo utilizando el conjunto de test. Calcule las métricas de evaluación apropiadas (por ejemplo, MAE, MSE, RMSE) para evaluar la precisión de los pronósticos.
7. Ajuste de hiperparámetros:
 - a. Experimente con diferentes hiperparámetros (por ejemplo, número de capas, unidades, tasas de dropout) para optimizar el rendimiento del modelo.

8. Forecasting:
 - a. Utilice el modelo entrenado para hacer pronósticos para los 50 artículos diferentes en 10 tiendas diferentes durante el período de interés de 3 meses.
9. Visualización:
 - a. Trace las ventas reales contra las previstas para visualizar el rendimiento del modelo.
10. Interpretabilidad del modelo:
 - a. Considere el uso de técnicas como los valores SHAP para interpretar los pronósticos del modelo y comprender qué características o puntos de datos históricos de ventas contribuyen más a las predicciones.
 - i. Si no están familiarizados con los valores SHAP, por favor refiérase a este [link](#) y a este [otro](#), para una lectura introductoria rápida

Task 2 - Teoría

Responda claramente y con una extensión adecuada las siguientes preguntas:

1. ¿Cuál es el problema del gradiente de fuga en las redes LSTM y cómo afecta la efectividad de LSTM para el pronóstico de series temporales?
2. ¿Cómo se aborda la estacionalidad en los datos de series temporales cuando se utilizan LSTM para realizar pronósticos y qué papel juega la diferenciación en el proceso?
3. ¿Cuál es el concepto de "tamaño de ventana" en el pronóstico de series temporales con LSTM y cómo afecta la elección del tamaño de ventana a la capacidad del modelo para capturar patrones a corto y largo plazo?

Entregas en Canvas

1. Jupyter Notebook o el script que usen para resolver el task 1, tanto en PDF como en .ipynb
2. Documento con las respuestas del task 2 en PDF
 - a. Pueden responder en el mismo Jupyter Notebook si así prefieren

Evaluación

1. [4 pts.] Task 1 (4 pts)
2. [1 pts.] Task 2 (0.25 pts pregunta 1,2 y 0.5 pregunta 3)

Total 5 pts.