

Entropía AI Candidates Test

Data Scientist Test: Time Series Forecasting

Instructions

This test is designed to evaluate your knowledge and skills in time series forecasting. It is divided into two sections: theoretical knowledge and a case study.

Section 1: Theory

1. Name two advantages and two disadvantages of using LSTM over ARIMA for time series forecasting.
2. Why is standard k-fold cross-validation not suitable for time series data? What technique would you recommend instead?

Section 2: Case Study

Scenario: Based on the csv send to you, your goal is to forecast monthly demand.

1. Cleaning and preparing the dataset.
2. Choosing and validating forecasting models (e.g., Prophet, XGBoost, LSTM).
3. Evaluating results and communicating your recommendations.

Deliverables:

- Code and explanations in a Jupyter Notebook (Python), saved as an HTML file. The HTML file should include all code, comments, and visualizations. Interactive visualizations (e.g., Plotly) are highly encouraged.
- A presentation (video call) where you explain your findings and insights.
 - 30 minutes total: 20 minutes for your presentation showing your code, and 10 minutes for Q&A.
 - Make sure to walk us through your thought process.
 - Imagine you are presenting to clients.

Please submit your work by the deadline provided. If you have any questions or concerns, please do not hesitate to reach out to us. We wish you the best of luck with the task and thank you for considering a position with entropía AI.