

Peer Review

Review of: "Enhancing Project Performance Forecasting using Machine Learning Techniques"

Jonathan Asante¹

1. Independent researcher

The research article titled "Enhancing Project Performance Forecasting using Machine Learning Techniques" addresses a critical topic in project management. The application of data-driven solutions in this area is indeed timely and has the potential to drive operational excellence and improve productivity. However, there are several aspects where the article requires improvement to enhance its clarity, completeness, and reproducibility. Key concerns include the insufficient description of the techniques employed, lack of transparency regarding the data used, and unclear methodology flow. Reproducibility, a cornerstone of research, is notably missing in its current state.

Below are detailed comments and suggestions for improvement:

Literature Review

- The manuscript briefly touches on ARIMA but does not provide a rationale for its selection over other statistical time series analysis methods such as ARMA, SARIMA, or others. Please elaborate on why ARIMA was chosen and how it compares to alternative techniques in this context.

Methodology

1. Flowchart Inclusion:

- Consider including a workflow or flowchart summarizing the methodology. This will provide readers with a clear visual understanding of the steps taken from data preparation to model evaluation.

2. Data Transparency:

- Specify the total number of data points used in the study.

- Clearly state the time intervals of the data (e.g., hourly, daily, or monthly). This information is crucial for understanding the granularity of the analysis.
- What assumptions were made to simulate the potential impact of external factors on project performance forecasting? These should be explicitly stated.

3. Feature Engineering:

- Explain the rationale behind the creation of the feature-engineered parameters. What was the thought process guiding their development?
- Provide a detailed list of all feature-engineered parameters.

Model Development

ARIMA Model:

- Provide a concise explanation of how the ARIMA model works, including its key components (AR, I, MA).
- Discuss the steps taken to achieve stationarity in the dataset.
- Clearly state the features used for building the ARIMA model.

LSTM Model:

- Similarly, explain the functioning of the LSTM model, focusing on its architecture and how it handles time series data.
- Specify the model parameters:
 - How many hidden layers and neurons were used?
 - Which parameters were included in training the LSTM model?
- Provide details on the dataset split:
 - What percentage of the data was used for training and testing?
- Explain the choice of optimizer used for model training. Why was this particular optimizer selected over others?
- Clarify the sliding window used for the LSTM model. How far back could the model "remember" to inform its predictions?

Addressing these will greatly enhance the clarity, reproducibility, and impact of the research.

Declarations

Potential competing interests: No potential competing interests to declare.