2 January 2025 · CC-BY 4.0

Qeios

Peer Review

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

Sweta Dey[1]

1. Computer Science & Engineering, Indian Institute of Technology Ropar, Ropar, India

Detailed review of the manuscript is stated as follows:

Strengths

1. Relevance and Timeliness:

- The manuscript addresses a critical issue in project management, particularly within the construction industry, where accurate forecasting of project performance metrics is essential. The focus on urban road reconstruction projects is particularly relevant given the increasing complexity and scale of such projects.

2. Innovative Approach:

- The use of machine learning techniques, specifically ARIMA and LSTM networks, to enhance forecasting accuracy represents a significant advancement over traditional methods. This innovative approach highlights the application of modern technology in improving project management practices.

3. Incorporation of External Factors:

- The model's integration of external factors such as weather patterns and resource availability into the forecasting process adds depth to the analysis. This recognition of the dynamic nature of project environments is a notable strength.

4. Comprehensive Methodology:

- The methodology is robust, involving detailed data collection, preprocessing, feature engineering, and model evaluation. The use of k-fold cross-validation and SHAP values for feature importance analysis adds credibility to the findings.

 doi.org/10.32388/WKU9VM

5. Empirical Validation:

- The application of the model to a real-world case study strengthens the research, providing practical insights and demonstrating the model's effectiveness.

Weaknesses

1. Limited Dataset:

- The dataset used in the study is confined to a single urban road reconstruction project. This limitation may affect the generalizability of the findings to other types of projects or different contexts within the construction industry.

2. Assumptions in External Factors Simulation:

- The simulation of external factors is based on reasonable assumptions rather than actual data. This could introduce biases and limit the accuracy of predictions. Real external data could enhance the model's reliability.

3. Complexity of the LSTM Model:

- While the LSTM model is effective in capturing complex patterns, its complexity may pose challenges in terms of interpretability. Practitioners may find it difficult to apply without a strong background in machine learning.

Below are some specific suggestions:

Integration of Real External Data

Weather Data: Instead of relying on simulated weather patterns, the model could integrate actual weather data sourced from meteorological services. This data could include daily precipitation, temperature fluctuations, and extreme weather events, which are known to affect construction timelines and costs. For instance, incorporating historical weather data could help predict delays due to rain or snow, thereby providing more accurate forecasts.

Resource Availability: Real-time data on resource availability, including labor force statistics and equipment usage rates, could enhance the model's accuracy. This could be sourced through project management software or IoT devices that track resource utilization on-site. For example, if a particular subcontractor is delayed due to labor shortages, this information could be fed into the model to adjust forecasts for earned value and cost variance.

doi.org/10.32388/WKU9VM

Traffic Patterns: Data on local traffic conditions could also be integrated, as these can impact project logistics and scheduling. Real-time data from traffic monitoring systems or historical traffic data can help in understanding delays related to material delivery or work zone access.

Regulatory Changes and Local Events: Incorporating data on local events (e.g., festivals, parades) or changes in regulations (e.g., new safety laws) could provide context for potential disruptions in the project timeline.

Techniques and Visualizations for Enhanced Interpretability

SHAP Values: While the article mentions SHAP (SHapley Additive exPlanations) values, elaborating on how to visualize these results could enhance interpretability. Practitioners could benefit from visualizing SHAP values using summary plots that illustrate the impact of various features on predictions. This would allow project managers to easily identify which factors most influence performance metrics.

LSTM Layer Visualization: Techniques such as Grad-CAM (Gradient-weighted Class Activation Mapping) can be adapted for LSTM models to visualize the importance of different time steps in the input sequences. This would help practitioners understand which historical data points are most influential in the model's predictions, enhancing trust in the model's outputs.

Time Series Decomposition: Implementing time series decomposition plots can help visualize the underlying trends, seasonal patterns, and residuals in the data. This can make it easier for practitioners to grasp the temporal dynamics affecting project performance.

Interactive Dashboards: Developing an interactive dashboard that allows users to manipulate input variables and immediately see the effect on forecasts could greatly enhance user engagement and understanding. Tools like Tableau or PowerBI can be utilized to create visualizations that allow for scenario analysis and sensitivity testing.

Feature Importance Ranking: In addition to SHAP, using other feature importance techniques (like permutation importance) and displaying them in a ranked format can help practitioners quickly understand which features to focus on for decision-making.

Incorporating these suggestions can significantly enhance the model's reliability and usability, ultimately aiding project managers in making informed, data-driven decisions.

4. Lack of Comparative Analysis:

- Although the paper compares the performance of the machine learning models with traditional methods like EVM, a more extensive comparative analysis with other modern forecasting techniques could provide a more comprehensive perspective.

Improvements

1. Expand Dataset Scope:

- Future research could involve multiple case studies across different types of construction projects to validate the model's applicability and enhance its generalizability.

2. Utilize Real External Data:

- Incorporating actual data for external factors instead of simulated data would improve the accuracy and reliability of the forecasts.

3. Increase Interpretability:

- Implementing techniques to enhance the interpretability of the LSTM model, such as using simpler models for comparison or providing clearer visualizations of the model's decision-making process, could make the findings more accessible to practitioners.

4. Broaden Comparative Analysis:

- Including comparisons with other advanced forecasting methods, such as Random Forest or Gradient Boosting, could further underscore the strengths and weaknesses of the proposed models.

Applications

1. Project Management:

- The findings can be applied by project managers in the construction industry to enhance decision-making processes, allowing for proactive management of potential deviations from baseline plans.

2. Policy and Planning:

- The research can inform policy-makers in urban planning and infrastructure development by illustrating the importance of incorporating dynamic factors into project forecasting.

3. Training and Development:

- The methodology and findings can serve as a basis for training programs aimed at equipping project managers with the skills needed to utilize machine learning techniques in project performance forecasting.

4. Future Research:

- The paper lays the groundwork for further research into the integration of machine learning in project management, encouraging exploration into different algorithms, data sources, and industries.

Thus, the manuscript presents a valuable contribution to the field of project management by exploring the intersection of machine learning and project performance forecasting. While there are areas for improvement, particularly in the scope of the dataset and the complexity of the models, the research offers important insights and practical applications for enhancing project management practices in the construction industry.

Hence, the decision is accepted with major revision.

Declarations

Potential competing interests: No potential competing interests to declare.