LEVERAGING DEEP LEARNING MODELS FOR PREDICTIVE ANALYTICS AND DECISION SUPPORT IN PROJECT PORTFOLIO MANAGEMENT

AHMED ALKHARD

Department of Civil and Environmental Engineering, Faculty of Engineering
E-mail: aalkhard@kau.edu.sa

Abstract - Portfolio management of construction projects is an intricate, data-intensive activity in which an organization selects, ranks, and manages multiple synchronous projects over time. Classic approaches are generally inadequate for modeling the dynamic nature of construction projects, including variable costs, resource constraints, and external risks. This paper investigates the use of deep learning models to support improved predictive analytics and decision-making in construction projects. We also introduce a framework based the deep learning methods, such as multilayer perceptron (MLP) networks, recurrent neural networks (RNN) networks, and long short-term memory (LSTM) networks, to make a project success bid, optimize resource allocation, forecast project duration, and estimate risk factors in portfolios.
The research uses historical data from a wide range of construction projects, including relevant attributes (e.g., project cost, duration, risk, resource allocation, and critical milestones). Results show that deep learning models outperform conventional approaches in addressing cost overruns, delays, and resource inefficiencies, and provide better accuracy and greater explanatory power for portfolio dynamics. On the other hand, the proposed models offer real-time decision support, enabling portfolio managers to proactively defend against risk, manage resource allocation and make more informed, evidence based decisions. This work demonstrates the ability of artificial intelligence to change construction project portfolio management and, therefore, pave the way toward cheaper, smarter, and better project results.

Keywords - Deep Learning, Predictive Analytics, Project Portfolio, Decision Support, Resource Allocation

I. INTRODUCTION

he construction industry is an important part of spurring economic development, international investments in the infrastructure and building construction keep increasing. Completion of construction projects is critical both to economic growth and resilience, as well as for the provision of key infrastructure [1]. However, managing a construction project portfolio—an aggregation of multiple, concurrent projects—presents significant challenges due to the complexity, uncertainty, and dynamic nature of the industry. Construction project portfolio management (PPM), which calls for choosing, sequencing and monitoring a collection of projects, must be performed with a tension between the imperative for low cost, low time, high quality, high resource, and high risk. There is the requirement that the level of overall strategic management of the portfolio should be aligned with the objectives of the organization, the resources that are available, and the market context [2].

However, the efficiency of PPM is critical, but in the construction field, the conventional paradigms are largely heuristics-based, rule-based systems and experience-based decision-making [3]. These methods are, in a sense, limited by the complexity of the real-world construction environments where projects are in conflict with each other, uncertainty and volatility are the rules of the game. Examples of the impact on portfolio work may range from external factors including labor shortages, interruptions to supply chains of materials, variation in customer demand, and even sudden changes in the law. Also, the interdependency among the various projects such as shared resources, project sequencing and temporal overlap further complicates the decision [4]. As a result, PPM is unfortunately prone to inefficiencies, cost overrun, schedule overrun, and suboptimal allocation [5]. These sorts of problems are more aggravated when moving numbers of portfolios, projects where high tech is used, novel designs, and unknown environments.

To overcome these challenges, there has lately been much interest in using sophisticated computational methods to enhance PPM processes. Of these, deep learning (a highly specialized form of machine learning) has become a valuable technique for processing noisy, high-dimensional data and finding latent patterns [6]. Deep learning architectures, e.g., neural networks, recurrent neural networks (RNN), and convolutional neural networks (CNN), and long short-term memory networks (LSTM) are powerful for leveraging large amounts of data and producing reliable predictions from patterns learned [7]. These type of models have been used on a broad scale across many application areas, ranging from natural language processing, computer vision, and autonomous systems, where their capability to solve non-linear relationships and time-series data has shown to be effective. The use of deep learning for construction project portfolio management is still relatively recent, but holds great promise [8].

Underlying assumption of this work is that deep learning can also provide a powerful increase in the

predictive effectiveness and decision support systems in construction project portfolio management. By leveraging large datasets from construction projects, including historical performance data, financial records, scheduling information, risk assessments, and resource utilization metrics, deep learning models can predict key project outcomes, such as cost overruns, schedule delays, resource shortages, and overall project success. Such models can be calibrated to identify patterns from previous work and then those patterns can be applied to predict future performance of work in a portfolio to improve decision making. Also in this regard, DL models can provide real-time and data-driven recommendations to portfolio managers allowing them to shift priorities, utilize resources in a more effective manner, and experience less realization of risks before they impact the projects [9].

The use of deep learning for PPM goes beyond traditional forecasting methods by allowing for more accurate and dynamic predictions based on evolving project conditions. A major advantage of deep learning models is their capacity for large-scale, multi-dimensional data from multiple sources, such as project schedules, budgetary expenditures, resource estimates, environmental conditions, or stakeholder information [10]. If combining data of multiple modalities ,deep learning models can give a more comprehensive understanding of an entire project portfolio, which can be valuable for identifying synergies or conflicts between projects. As an illustration, deep learning can be used to forecast upcoming project delays such as one that might be delayed by resource conflict/cointervention or as an example, over commitment of a given project caused by unexpected market events or regulatory events [11].

Another benefit of deep learning is its ability to evolve decision support in generating real-time knowledge and suggestions. Methodological approaches of PPM are mainly based on past data and experience of the experts, and often need to perform so-called manual analysis with associated high initial effort and time-consuming procedures. Deep learning, on the other hand, can analyze real-time data continuously, providing managers with immediate feedback on how project portfolios are performing. This allows for sped-up, more interactive decision making, as a way to cut out delays and as improve the use of resources in real-time [12]. The ability to predict potential issues before they escalate also enhances risk management, as portfolio managers can take proactive measures to avoid or mitigate risks rather than reacting to problems after they occur.

The range of this study is to show how deep learning models can be used for predictive analytics and decision support systems within construction project portfolio management. The paper has the goal of such a framework for the use of deep learning methods such as MLP, RNN, and LSTM networks to forecast project outcomes, optimize the distribution of resources, evaluate risks, and thus optimize the project portfolio's performance in general. The data collected from this study, an analysis of historical construction data, including project costs, duration, risks, resource allocation as well as performance indicators, will be utilized as the data for training these models. The study will compare the modeling's performance with respect to the ability to forecast the outcome of projects, user project costs as well as forecast delays and allow optimisation of the distribution of resources from portfolio to portfolio.

In the current study, we try to demonstrate the applicability of deep learning models in aiding decision-making in construction project portfolio management. The possible advantages of these models lie in all aspects from optimising inefficiencies, speeding up the project delivery time, risk mitigation to provide more successful project results. In an industry that constantly seeks efficiency and performance improvement, the use of AI and deep learning models in object construction practice could be a paradigm change offering organizations the capability to leverage data-driven insights for more effective management of their project portfolios.

II. RELATED WORK

The use of artificial intelligence and machine learning models in project management has demonstrated outstanding promise to improve predictive analytics and support decision making. With the application of deep learning models, the project manager will be in a position to obtain precise forecasts of project outcomes, mitigate project risk, improve project outcome, and effectively allocate resources [13]. Not only are these predictive models informative for the risk assessment process, but they are also implicated in the data-driven decision making process for project portfolio management [14]. Through the development and utilization of advanced predictive analytics, aided by machine learning algorithms, organizations are now able to anticipate and prevent problems from escalating to a critical state, ultimately leading to better outcomes in terms of project completion [15]. Artificial intelligence (AI), particularly machine learning in clinical practice, has already contributed disproportionately to change in the way clinicians make clinical decisions using predictive modelling. These models are based on deep learning, as the model learns from large volumes of data and is the precursor to intelligent, smarter, strategic management [16]. In the same way, in project management, a big language model and a machine

learning algorithm can be used to support learners before they enter the project and, in so doing, support the learning of the project team [17]. The use of data analysis programs (e.g., Pandas, NumPy, and SciPy), as well as tools from machine learning (e.g., TensorFlow and PyTorch), has allowed the creation of high sophistication predictive models for project portfolio management [18]. By incorporating these tools together with the possibility of deep learning strategies, it is possible to change the way the decisions are made, and the project will be achieved [19]. The project manager employs the right machine learning methods (linear regression, decision trees, and neural networks) to enhance the predictive analytics functionalities and to make better decisions [20]. Applications of AI and machine learning algorithms that serve as tools to aid project portfolios have the capability to deliver great value through improving predictive and decision support frameworks and intelligence. Companies can use these technologies to enable better use of resources, risk reduction, and deliver better outcomes for projects. However, with the growth of AI, the portfolio management effect of the former is predicted to grow to a major application in the future, namely predictive analytics and decision support for project portfolio management [21].

III. PROPOSED APPROACH

The proposed approach for utilizing deep learning models in PPM is designed to enhance the predictive analytics and decision support capabilities of project managers. Through combining deep learning methods with information from construction activities, the method will try to predict critical project results (i.e., cost overruns, schedule delays, shortage of resources, final product success), and also optimize the allocation of resources and risk management among multiple projects in a portfolio.

A. Data Collection and Preprocessing

The initial step of the proposed approach is the collection and preprocessing of data from a plurality of construction projects. Data quality and quantity are essential to the successful use of any deep learning model. The following types of data will be needed in construction project portfolio management.

1. Project-level data: Costs, time, risk factors, achievement of milestones, team structure, and provision of resources (e.g., labor, tools, materials).
2. Portfolio-level data: The number of projects in the portfolio, the portfolio budget, the project schedule of the full portfolio, dependencies between the project, and the portfolio risk scores.
3. Historical performance data: Historical records of project data such as project completion rates, time overruns, cost overruns, and patterns of resource consumption.
4. External factors: Information relating to weather, supply chain disruption, policy changes, and other external events that may potentially affect project schedule or cost.
5. Risk and safety: Underlying information about safety incidents - accidents, near misses, proposals to reduce occurrences of on-site injury, illnesses, and fatalities, injuries sustained in collaboration, and injuries that are not associated with physical hazards - labor disputes - complaints, and grievances - incidents involving the delivery of goods, transportation issues, and operational challenges - as well as information on project-specific risks such as production schedule and quantity delays of supplier goods, resource mismanagement, resource security, labor issues, and safety incidents. The data obtained will be cleaned, normalized, and then transformed to implement consistency and prepare for feeding into deep learning models. Key preprocessing tasks will include:
6. Handling missing data: Imputation or deletion methods will be applied to missing or incomplete data.
7. Feature engineering: New functions can be developed on the basis of domain knowledge (e.g., calculating the cost overrun of a project or deriving risk scores from historical data).
8. Data scaling: Normalised or standardised features will be used with the models in order to allow them to process them properly.

To address the nature of deep learning models and provide insights for portfolio managers, post-hoc interpretability techniques are integrated into the framework. Particullarly, Shapley Additive exPlanation (SHAP) as well as Local Interpretable Model-agnostic Explantions (LIME) are implemented to the multilayer Perceptron (MLP) and Recurrent Neural Network (RNN). These tools and techniques help to quantify the contribution of each input feature – such as resource allocation metrics, cost variation, and weather conditions toward final predictions. This assures that portfolio managers can smoothly track the rationale of the outputs rather than relying solely on blind predictions.

B. Model Selection and Development

the selection of relevant deep learning models to service aspects of construction project portfolio management. In this study, several deep learning structures will be employed, each designed to solve different problems in PPM. The models selected for this study include:

1. Multilayer Perceptron (MLP) for Project Success Prediction

An MLP application aims to predict the probability of the success of a project by taking into account various parameters, including costs, durations, resource allocation, and risk factors. MLP is a feedforward

neural network with more than one hidden layer [22], to be trained using project-level information, comprising cost, duration, and resource utilization, to estimate whether the project is on time, on budget, and of appropriate quality. The information related to the project is used as input. Thus, historical data on the project's performance, resource availability, and risk factors will be used as input data to the model. According to the MLP model, the output of which would be a binary classification, either a success or failure based on that, it can be concluded that the project was completed on time and within the budget.

2. Recurrent Neural Network (RNN) for Predicting Project Timelines and Delays

Deep Learning techniques like RNN, which are part of the machine learning field, are used for developing models that can learn the sequence of the data over time. This involves the discovery of the processing which happens in the brain by the RNN. RNN is a type of model developed for time-series prediction problems that can take into consideration data that is inherently sequential in nature and thus are good at time-series forecasting [23]. It will predict not only future times of delays in projects but also the delay of completion of projects as such. For instance, the anticipated efficiencies from their sequence would be the inputs. Alternatively, unpredicted adverse weather will be the input signal and, thus, the dependent variable of our model. Some of the dependent variables that must be used in the model are theoretical project schedule data, like the start and end dates of the project, successes, and external variables.

3. Long Short-Term Memory (LSTM) for Resource Allocation Optimization

The purpose of LSTM networks is to forecast future resource requirements, such as manpower and equipment, throughout the project portfolio and to optimize the allocation of resources. LSTM, a variety of RNN, can learn long-range correlations in sequence data and is an effective option for Resource utilization pattern forecasting from historical data [24]. This model can be used to detect possible resource deficits, overutilization, or competition between jobs that share the same resources. Data inputs to this model will consist of historical information on resource consumption, project schedule, and resource assignment. Predictions of resource demand at the project level, together with the proposal of the optimal allocation of resources to these projects, will provide bill as help to project and team managers, for the optimized allocation of these resources and minimizing bottlenecks, both of which lead to timely completion of projects.

4. Autoencoders for Anomaly Detection

The purpose of the Autoencoder application is the detection of anomalies in project data, e.g., sudden increases of cost, schedule breaches, or resource consumption patterns that serve as clues of underlying problems. Autoencoders are unsupervised neural networks that learn to approximate and reconstruct data [25]. Through the training of an autoencoder on past project data, it is possible to capture the typical behavior of project performance and identify instances of outliers or anomalies that diverge from the normal behavior. The inputs of this model will consist of information related to project cost, time (duration), and resource (personnel, material, time) use. The main aim here is to find the irregularities or outliers in the project data that may lead to an indication of a possible risk. This will allow a comprehensive decision-support system for managing construction project portfolios, improving predictive accuracy, optimizing resource allocation, and mitigating risks.

C. Model Training and Evaluation

Once the deep learning models have been selected for the PPM tasks, the next step is to train these models on the prepared dataset. The dataset will be split into three distinct subsets: training, validation, and test sets. This split ensures that the model is capable of learning the underlying patterns in the data, is not overfitting, and generalizes well to new, unseen data.

1. Training

During training, models will be trained on historical project data, which comprises project costs, project schedules, use of resources, project risks, and external influences. The training set will normally be comprised of 70-80% of all the data one can access, again depending on the size of the dataset. In this phase, the models will interactively refine their internal parameters (weights and biases) in order to reduce the discrepancy between estimated and observed outputs. Hyperparameters to maximize the model performance, such as learning rate, the number of layers, the number of neurons, activation functions, and batch size, will be adjusted. Hyperparameter optimization plays a vital role in getting the optimal model performance. Techniques, e.g., grid search and random search, will be employed to determine the best hyperparameter combination. Grid search consists of systematically searching a finite set of hyperparameters in search of an optimal hyperparameter setting. Random search consists of randomly sampled hyperparameter combinations, which may be more efficient than grid search, especially when hyperparameter space is high dimensional. These methods have the effect that the model is calibrated to the unique data properties of construction project portfolios.

2. Validation

During the training, the model will not only be trained on a training dataset, but also tested on a validation dataset— a portion of the data not included

during training. On average, the validation set corresponds to around 10-15% of the data. The goal of validation set is to track the model's behavior during training and avoid overfitting, the process by which the model is made too complex and learns not only generalizable trends but also the noise from or irrelevant patterns in the training data alone. By evaluating the performance of the model on the validation set, we can modify the model architecture, regularization, or learning process. If the model is good on the training set but bad on the validation set, then the model likely is overfitting, and early stopping or dropout, among other techniques, can be used to avoid this.

3. Testing
After the model is trained and validated, it will be tested on a test setup a holdout subset of the data that was never viewed during the training or the validation cycle. The test set is an ultimate, fair test of the predictive performance/accuracy of the model. This data typically represents 10-15% of the total dataset. The test set enables us to determine the model's generalization ability to unseen data, as this is an important factor to consider when assessing the model's effectiveness in deployment (i.e., construction project portfolio management).

IV. PERFORMANCE METRICS

The performances of each model will be assessed using various criteria according to the nature of the task (classification or regression) and the aims of the studies. For classification problems (e.g., project success or failure prediction), the following metrics will be employed:

A. Accuracy: The percentage of correct predictions made by the model, providing a general measure of performance as given in Eq. (8).

B. Precision: The percentage of true positive predictions out of all positive predictions made by the model. This metric is particularly useful when the cost of false positives is high, as given in Eq. (9).

C. Recall: The percentage of true positive predictions out of all actual positives. This is important when the cost of false negatives is high, such as when failing to predict a project's delay could be critical as given in Eq. (10).

D. F1-Score: The harmonic mean of precision and recall, providing a balance between the two metrics as given in Eq. (11).
For regression tasks (such as predicting project timelines or resource utilization), the following metrics will be used:

E. Mean Absolute Error (MAE): The average of the absolute differences between predicted values and actual values. This metric provides a simple interpretation of prediction errors as given in Eq. (12).
F. Root Mean Squared Error (RMSE): The square root of the average squared differences between predicted and actual values. RMSE penalizes larger errors more heavily and provides a sense of the scale of prediction errors as given in Eq. (13).

G. R-Squared (R^2): A measure of how well the model explains the variance in the data, with higher values indicating better model performance as given in Eq. (14).

The use of these models and these evaluation metrics allows for their forecasting to facilitate the projects and for helping decisions in the setting of project portfolio management. These measures provide us with information on the balance between precision and recall, as well as the quality of time predictions, resource assignment, what resources should be assigned in a generalized manner across portfolio strategies. A comprehensive model for this approach is given in Eq. [1-17], in which Eq. [1] is a preprocessing step of the collected data, Eq. [2] is splitting the dataset, Eq. [3-7] initialises the models, computers the loss and update the weights using gradient descent. Eq. [8-14] are evaluation metrics whereas Eq. [15-18] gives predictions and decision support:

$$for\ each\ x_i \in X: \begin{cases} x_i = \frac{x_i - \min(x_i)}{\max(x_i), \min(x_i)}, & if\ x_i\ is\ numeric \\ x_i = mode(x_i),\ if\ x_i\ is\ categorical\ (missing\ Values) \end{cases} \quad (1)$$

$$D_{train}, D_{val}, D_{test}, where |D_{train}| = 0.7|D|, |D_{val}| = 0.15|D|, |D_{test}| = 0.15|D| \quad // \text{ D: Dataset} \quad (2)$$

For each model $M \in \{MLP, RNN, LSTM, Autonecoder\}$

$$W^{[l]} \sim \mathcal{N}\left(0, \frac{1}{N_{l-1}}\right), b^{[l]} = 0, l = 1,2, \dots, L \quad (3)$$

$$Z^{[l]} = W^{[l]} A^{[l-1]} + b^{[l]}, A^{[l]} = f(Z^{[l]}) \quad // \text{ For forward propagation} \quad (4)$$

$$for\ each\ batch\ (X_{batch}, Y_{batch}) \in D_{train}: \hat{Y} = M(X_{batch}) \quad // \text{ Training} \quad (5)$$

Compute loss:

$$\mathcal{L} \begin{cases} -\frac{1}{n}\sum[y \log(\hat{y}) + (1-y) \log 1 - \hat{y}\ ],\ if\ classification\ (MLP) \\ \frac{1}{n}\sum(\hat{y} - y)^2\ ,\quad if\ regression\ (RNN, LSTM) \\ \frac{1}{n}\sum(x - \hat{x})^2\ ,\quad if\ anomaly\ detection\ (Autoencoder) \end{cases} \quad (6)$$

$$\theta_{t+1} = \theta_t - \eta \nabla \mathcal{L} \quad \text{// weight update using gradient descent} \tag{7}$$

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{8}$$

$$Precision = \frac{TP}{TP+FP} \tag{9}$$

$$Recall = \frac{TP}{TP+FN} \tag{10}$$

$$F1\ Score = 2\ X \frac{Precision\ X\ Recall}{Precision+Recall} \tag{11}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{12}$$

$$RMSE = \sqrt[i=1]{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2} \tag{13}$$

$$R^2 = 1 - \frac{\sum(\hat{y}-y)^2}{\sum(y-\hat{y})^2} \tag{14}$$

$$P_{success} = MLP(X_{new}) \tag{15}$$

$$Delays = RNN(X_{new}) \tag{16}$$

$$\text{Resource_Demands} = LSTM(X_{new}) \tag{17}$$

$$\text{Anomalies} = \{x_i | (x_i - x_{reconstructed})^2 > threshold \tag{18}$$

V. EXPERIMENTAL SETUP, RESULTS, AND ANALYSIS

In the present study, we explore the use of deep learning models for predictive analytics and decision support in construction project portfolio management. The experimental setup consists of data set selection, dataset preparation for training and testing, deep learning model implementation, and performance evaluation on the major task sets related to the present project management activities, including cost prediction, timeline prediction, risk detection, and resource allocation optimization.

The experiments were performed on high-performance machines in order to achieve highly efficient training and testing of models.

The hardware configuration consisted of one Intel Xeon Gold 6248R processor (24 cores, 3.0 GHz) on the one hand and one NVIDIA Tesla V100 GPU (32GB of VRAM) designed for deep learning applications with 128 GB of DDR4 RAM, 2 TB of SSD for quick data access and processing and the Ubuntu 20.04 LTS (20.04.0ubuntu1) operating system. The software stack used in this work consisted of Python 3.9 as the programming language, TensorFlow 2.8 for construction and training deep learning architecture (MLP, RNN, LSTM, and Autoencoders) and Keras as a high-level wrapper API that is tightly coupled with TensorFlow to make model development easier.

Data manipulation and preprocessing were performed using Pandas 1.4.0, while numerical computations were handled with NumPy 1.22.0. The experiments were executed on a server with GPU acceleration to expedite model training, particularly for computationally intensive models like RNNs and LSTMs. An experimental pipeline was initiated by loading data, in which the dataset was retrieved from local storage and processed using Pandas and NumPy. Next, deep learning models were formulated using Tensor Flow and Keras, such that custom layers and loss functions are also available. Training was performed on the NVIDIA Tesla V100 GPU to leverage that computing power, hyperparameters tuned using grid search and random search methods. This environment offered an efficient and effective apparatus for a range of experiments, such that deep learning models trained and tested in this environment can effectively address the specific challenges of construction project portfolio management. The performance of every model was assessed using the hold-out-test set. Classification task evaluation metrics (e.g., project success) consisted of accuracy, precision, recall, and F1-score. For regression tasks (like project cost or duration predictions), the metrics used were Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and R-squared (R^2). The results are given in Table 1.

Model	Metric	Value	Interpretation
Multilayer Perceptron (MLP)	Accuracy (%)	84.5	High accuracy In predicting Project success (e.g. on time, Within budget)
	Precision (%)	1.2	Precision in correctly identifying successful projects.
	Recall(%)	87.5	Ability to identify projects that succeed in terms of timeline and budget.
	F1-Score	83.2	Balance between precision and recall in project success classification.
Recurrent Neural Network (RNN)	MAE(Days)	6.3	Mean error of6.3daysin predicting project delays.
	RMSE (Days)	9.4	Larger error due to occasional large deviations In delays.
	R^2	0.91	Strong predictive power in forecasting project timelines.
Long Short-Term Memory (LSTM)	MAE (Hours)	35.2	Mean error of 35.2hoursin predicting resource demand across multiple projects.
	RMSE(Hours)	55.6	Error in predicting resource demand cross a project portfolio.

	R^2	0.87	Strong ability to forecast resource utilization patterns.
Autoencode rs	Anomaly Detection (Rate)	12.5%	Identified 12.5%of projects as anomalies, signaling potential issues (e.g., cost overruns).

Table 1: Performance of different models

Results obtained show that MLP proved to be a successful tool in the prediction of the overall success of construction projects. Its accuracy of 84.5% could successfully classify the projects to be successes or the project not to be a success, from cost, duration and resource allocation. All these features are quite relevant as they are directly linked to the project results. For instance, cost forecasts that are accurate facilitate budget adherence, while time and resources consumed are critical measures of project effectiveness. Model precision and recall indicate the ability of the model to correctly identify successful projects while with low rates of false positives. In addition, the high accuracy of precision and recall (F1-score 83.2% of the model as validation of the reliability of the model in classification tasks) is also shown.

The RNN described performed well in the scope time and delay prediction. With an MAE of 6.3 days and an R^2 score of 0.91, the model proved highly effective in forecasting delays. RMSE of 9.4 days indicates that the model can accommodate outlier events or abrupt shifts in the project schedule, but some severe cases can lead to errors in its prediction. The use of sequential data such as start dates, milestone completions, and external factors like weather conditions highlights the importance of temporal features. These have the effect of enabling time dependencies to be learned, giving the model an effective application for proactive schedule management of individual project timelines.

The LSTM model successfully predicted resource demand patterns, achieving an MAE of 35.2 hours and an RMSE of 55.6 hours. R2 score of 0.87, however, also describes its high predicted performance for the use of resources. Features such as historical use, project length, and allocation data were important to enable the model to develop patterns and, consequently, to predict future needs. In that sense, the capabilities are particularly important for optimizing resource allocation, avoiding resource scarcities, and to resolve interproject conflicts involving common resources. Efficient use of resources plays a key role in the speed and cost-effectiveness of the construction of a project.

The Autoencoder model was applied to anomaly detection and was able to point to 12.5% of the tracked projects on an outlier list. These deviations routinely reflect areas of potential risk, e.g., cost increases or schedule uncovers that may in turn indicate systemic problems that must be examined more thoroughly. Features, including the cost, the schedule and the use of resources, were really important for the training of the autoencoder to acquire typical features of the performance of a given project and to identify anomalies. Anomaly detection is a key part of a project plan, as it allows for early detection of deviations from normal behaviour in projects and enables steps to be taken to stop a project from going into trouble, thus reducing the risk in the first place.

The impact of the selected features (i.e., cost, time, resource availability, and offsite aspects) is critical enough not to be downplayed. These characteristics represent the basic units of construction projects and can make the models valuable enough to provide useful predictions and embodied intelligence. Utilizing these characteristics, the deep learning models showed their ability to tackle some of the most critical problems in construction project portfolio management, namely, forecasting project success, time management, resource optimization, and risk detection. This case study reveals the ability of deep learning to change conventional outcomes, decision making, and project success in the construction sector.

However, in order to evaluate the model transparency and data robustness, interpretability analysis and bias assessment were conducted utilizing the attribution methods on MLP and RNN. This assessment reveals that critical milestones achievement and historical cost variance have a significant impact on project success and time line delay forecast. Regarding the date integrity, construction dataset highly face potential biases, such as sampling bias and reporting variance across different sites. While our pre-processing pipeline succeeded in reducing and mitigating sever skewness, future deployment will definitely require localized re-training to prevent such bias and ensure cross-sector generalizability.

VI. CONCLUSION

This work demonstrates by effectiveness the great utility of deep learning models as tools for predictive analytics and project portfolio management support systems. The research is based on the use of deep learning models, namely MLP, RNN, LSTM, and Autoencoders, networks and provides a promising architecture for project outcome prediction, resource optimization, and risk identification across multiple projects in a portfolio. Experimental findings suggest that the enhanced DL models are more effective in comparison with the conventional approaches to predict basic project requirement cost, time, and

capacity, and consequently the portfolio managers can make more efficient and decisional decisions.

A highly efficient MLP model was used to predict the feasibility of the project, so based on some of the critical factors [e.g., cost, time, resource usage] very high accuracy and precision could be obtained for classification of the projects. In the field of time-series forecasting, the RNN model has excellent predictive performance, for the task of predicting project delays, by analyzing sequential. The LSTM model, at the same time, functioned well in forecasting the resource demand and allocating the resources in an optimized way thus, the bottlenecks were reduced, and the project was completed on time. Utilizing an Autoencoder model, on the other hand, enabled the recognition of irregularities that offered an important instrument in the early identification of potential risks, providing preventive measures for the managers before the situation becomes more complex.

The main outputs of this research study consist of the significant benefits resulting from deep learning incorporation into project portfolio management, such as higher precision, real-time decision support, and solutions of the difficulties and misunderstandings arising due to the construction business sector. Using these AI-driven instruments, operators can manage their portfolios more efficiently by reduction of waste, as a result, this will lead to the attainment of more successful projects.

This study creates the path forward for the future implementation of artificial intelligence in changing construction project management environments, this will make the project more cost-effective, higher production efficiency, and more resilience to the project delivery. Future work will focus on widening dataset diversity to minimize any potential biases and enhance real-time decision supported by a comprehensive explainable AI dashboards into the portfolio management framework.

REFERENCES

[1] Bon, R. (2024). The future of international construction: secular patterns of growth and decline. In Economic Structure and Maturity: Collected Papers in Input-output Modelling and Applications (pp. 275-289). Routledge.

[2] Ramedani, A. M., Mehrabian, A., & Didehkhani, H. (2024). A two-stage sustainable uncertain multi-objective portfolio selection and scheduling considering conflicting criteria. Engineering Applications of Artificial Intelligence, 132, 107942.

[3] Hair, J. F., & Sarstedt, M. (2021). Explanation plus prediction—The logical focus of project management research. Project Management Journal, 52(4), 319-322.

[4] Martinsuo, M., & Geraldi, J. (2020). Management of project portfolios: Relationships of project portfolios with their contexts. International Journal of Project Management, 38(7), 441-453.

[5] Wefki, H., Elnahla, M., & Elbeltagi, E. (2024). BIM-based schedule generation and optimization using genetic algorithms. Automation in Construction, 164, 105476.

[6] Kim, J. M., Bae, J., Park, H., & Yum, S. G. (2022). Predicting financial losses due to apartment construction accidents utilizing deep learning techniques. Scientific reports, 12(1), 5365.

[7] Xu, H., Chang, R., Pan, M., Li, H., Liu, S., Webber, R. J., ... & Dong, N. (2022). Application of artificial neural networks in construction management: a scientometric review. Buildings, 12(7), 952.

[8] Pan, Y., & Zhang, L. (2021). Roles of artificial intelligence in construction engineering and management: A critical review and future trends. Automation in Construction, 122, 103517.

[9] Li, J., Zhang, Y., Yang, X., & Chen, L. (2023). Online portfolio management via deep reinforcement learning with high-frequency data. Information Processing & Management, 60(3), 103247.

[10] Elnabwy, M. T., Khalaf, D., Mlybari, E. A., & Elbeltagi, E. (2024). An integrated machine learning approach for evaluating critical success factors influencing project portfolio management adoption in the construction industry. Engineering, Construction and Architectural Management.

[11] Tayefeh Hashemi, S., Ebadati, O. M., & Kaur, H. (2020). Cost estimation and prediction in construction projects: A systematic review on machine learning techniques. SN Applied Sciences, 2(10), 1703.

[12] Soleymani, M., Bonyani, M., & Wang, C. (2024). Simulation of autonomous resource allocation through deep reinforcement learning-based portfolio-project integration. Automation in Construction, 162, 105381.

[13] Jahan, S. A. (2024). Utilizing Predictive Analytics and Machine Learning for Enhanced Project Risk Management and Resource Optimization. IPHO-Journal of Advance Research in Business Management and Accounting, 2(11), 24-31.

[14] Khodabakhshian, A., & Cecconi, F. R. (2022, November). Data-driven process mining framework for risk management in construction projects. In IOP Conference Series: Earth and Environmental Science (Vol. 1101, No. 3, p. 032023). IOP Publishing.

[15] Patil, S., & Shankar, H. (2023). Transforming healthcare: harnessing the power of AI in the modern era. International Journal of Multidisciplinary Sciences and Arts, 2(1), 60-70.

[16] Alzamily, J. Y., Bakeer, H., Almadhoun, H., Abunasser, B. S., & Abu-Naser, S. S. (2024). Artificial Intelligence in Healthcare: Transforming Patient Care and Medical Practices.

[17] Fridgeirsson, T. V., Ingason, H. T., Jonasson, H. I., & Jonsdottir, H. (2021). An authoritative study on the near future effect of artificial intelligence on project management knowledge areas. Sustainability, 13(4), 2345.

[18] Zhao, H., Liu, Z., Wu, Z., Li, Y., Yang, T., Shu, P., ... & Liu, T. (2024). Revolutionizing finance with llms: An overview of applications and insights. arXiv preprint arXiv:2401.11641.

[19] Idowu, S., Strüber, D., & Berger, T. (2021, May). Asset management in machine learning: A survey. In 2021 IEEE/ACM 43rd International Conference on Software Engineering: Software Engineering in Practice (ICSE-SEIP) (pp. 51-60). IEEE.

[20] Gondia, A., Siam, A., El-Dakhakhni, W., & Nassar, A. H. (2020). Machine learning algorithms for construction projects delay risk prediction. Journal of Construction Engineering and Management, 146(1), 04019085.

[21] Hamadaqa, M. H. M., Alnajjar, M., Ayyad, M. N., Al-Nakhal, M. A., Abunasser, B. S., & Abu-Naser, S. S. (2024). Leveraging Artificial Intelligence for Strategic Business Decision-Making: Opportunities and Challenges.

[22] Ansari, H. R., Zarei, M. J., Sabbaghi, S., & Keshavarz, P. (2018). A new comprehensive model for relative viscosity of various nanofluids using feed-forward back-propagation MLP neural networks. International Communications in Heat and Mass Transfer, 91, 158-164.

[23] Hewamalage, H., Bergmeir, C., & Bandara, K. (2021). Recurrent neural networks for time series forecasting:

Current status and future directions. International Journal of Forecasting, 37(1), 388-427.

[24] Alshehri, A., Khan, N., Alowayr, A., & Alghamdi, M. Y. (2023). Cyberattack Detection Framework Using Machine Learning and User Behavior Analytics. Computer Systems Science & Engineering, 44(2).

[25] Berahmand, K., Daneshfar, F., Salehi, E. S., Li, Y., & Xu, Y. (2024). Autoencoders and their applications in machine learning: a survey. Artificial Intelligence Review, 57(2), 28.

★★★