A Comparative Study on Machine Learning Algorithms for Predicting Construction Projects Delay

Authors
Christian Nnaemeka Egwim[1], Hafiz Alaka[1]

Contact Details
[1]Big Data Technologies and Innovation Laboratory, University of Hertfordshire, Hatfield, United Kingdom.

Abstract

The perpetual occurrence of a global phenomenon – delay in construction industry despite considerable mitigation efforts remains a huge concern to its policy makers. Interestingly, this industry which produces massive amount of data from IoT sensors, building information modelling etc., on most of its projects daily is slow in taking the advantage of contemporary analysis method like machine learning (ML) which best explains factors that can affect a phenomenon like delay based on its predictive capabilities haven been widely adopted across other sectors. In this study therefore, a premise to compare the performance of machine learning algorithms for predicting delay of construction projects was proposed. To begin, a study of the existing body of knowledge on the factors that influence construction project delays was utilised to survey experts in order to obtain quantitative data. The generated dataset was used to train twenty-seven machine learning algorithms in order to develop predictive models. Results from the algorithm evaluation metrics: accuracy, balanced accuracy, Receiver Operating Characteristic Curve (ROC AUC), and f_1-score indeed proved Perceptron model as the top performant model having achieved an accuracy, balanced accuracy, ROC AUC, and f_1-score of 85%, 85%, 0.85 and 085 respectively higher than the rest of the models and unachieved in any previous study in predicting construction projects delay. Ultimately, this model can subsequently be integrated into construction information system to promote evidence-based decision-making, thereby enabling constructive project risk management initiatives in the industry.

Keywords: Algorithms, Machine Learning, Predictive Analytics, Project Delay, Predictive Model

1. INTRODUCTION

The construction industry is a major contributor to the world economy, accounting for 13% of world GDP and projected to grow by 85% to $15.5 billion by 2030, with three major countries – China, the United States, and India – accounting for 57% of worldwide demand (Robinson Graham, 2015). In addition, Woetzel et al. (2017) projects that global infrastructure expenditures will total $3.4 trillion per year between 2013 and 2030, accounting for nearly 4% of total GDP. The industry is also regarded as the major pillar of the economy of any country — representing 3% of the overall economic production of Nigeria, 4.3% of Germany's overall economic production, 6% of the UK's overall economic production, 4.1% and 6.8% of the US and China's entire economic production, etc.

Despite its significance, however, the construction industry continues to underperform. According to Egan (2018), the industry is underperforming, as seen by poor profitability, capital investment, and research and development, mostly caused by delay of its projects, resulting in high customer dissatisfaction with the industry's overall performance. According to several research papers, such as Andrew (2013), Flyvbjerg (2014) and Rhodes (2019), 9 out of 10 worldwide megaprojects encounter delay, which frequently result in cost overruns. Delay is the most important factor in the overall completion of any construction project since it increases excess expenses (Haq et al., 2017).

Several studies have found that construction project delays have a negative impact on the reputation of the construction industry's contribution to the global economy. According to Abdul-Rahman et al., (2011), Shah, (2016), Ji et al., (2018), and Chen et al., (2019), the effects of construction delays can be evaluated in terms of their national footprints, which influence the industry's subsidy to the economy; at an industry level, where delays negatively impact profitability and productivity; and at a project level, where delays foster industry client dissatisfaction with its overall performance, contract termination (s).

Over the years, several statistical methods have been developed to minimise construction project delays. Statistical methods necessitate a set mathematical structure for the relationship between dependent and independent variables. They can also be impacted negatively by missing values and outliers in the dataset. As an example, statistical methods such as Importance Index (II) and Rank Correlation Coefficient (ρ) by Assaf and Al-Hejji (1995), Relative importance index (RII) by Odeh and Battaineh (2002), Relative Importance Weight (RIW) by Ramanathan, Narayanan and Idrus (2012) etc have been employed to mitigate delay, yet delay still strives in the industry.

On the other hand, ML methods are non-parametric technologies that excel at dealing with outliers and missing data. They can choose the most significant explanatory factors while also classifying the dependent variables. Only recently, did the industry began to take advantage of the Machine Learning (ML) methods to predict delay of its projects. More precisely interestingly, only these three literatures at the time of this study: Gondia et al. (2020) used two ML models – Decision Tree and Naïve Bayesian Classifiers (with accuracy value of 74.5% and 78.4% respectively) towards expediting precise project delay risk assessments and forecast in building project in Egypt.

Also, Asadi, Alsubaey & Makatsoris (2015) used two ML approach (with accuracy value of 79.41% and 73.52% for decision tree and Naive Bayes model respectively) to predict delays in construction logistics in Qatar. Furthermore, Yaseen et al. (2020) developed a hybrid artificial intelligence model (a combination of Random Forest and Genetic Algorithm) and achieved an accuracy value of 91.67% for delay problem prediction in Iraq. As a result, these studies have produced a knowledge vacuum that must be filled in other nations. Thus, this study therefore aims to evaluate the application of several ML algorithms to develop predictive models for predicting delay of construction projects in Nigeria. The following objectives will be used to achieve this aim:

1. Carry out literature review toward gathering the most common factors affecting construction projects delay and use it to conduct survey of experts to establish the most applicable factors affecting construction projects delay.
2. Utilize established factors in objective 1 as independent variables for ML algorithms to develop predictive models.

The contribution of this study is therefore to fill the gap in lack of application of ML algorithms in predicting construction project delay in Nigeria. This is novel because the thorough review of existing body of knowledge indicated that this is the first-time robust ML methods are employed to predict delay of construction projects in Nigeria. As a result, several construction sector stakeholders will be able to improve the quality of their decisions and risks on current or future construction projects, fostering trust, increasing productivity and revenue, and, most importantly, yielding timely delivery of construction projects in the industry.

2. LITERATURE REVIEW

The term "delay" is widely used in the construction industry, resulting in a large body of international literature defining the phrase. It is the viewpoints of Assaf & Al-Hejji (2006) that delay can be defined as the time increase beyond the agreed project delivery planned schedule by stakeholders or beyond a legal contract completion date. Furthermore, delay mean different things to different stakeholders (client, contractor, consultant etc.) For client, delay connote loss of revenue or investments at the end of agreed time while to the contractor, a delay can imply an increase in overhead cost (Abdul-Rahman et al., 2009; Akhund et al., 2017; Elawi et al., 2016; Gardezi, Manarvi & Gardezi, 2014; Głuszak & Lešniak, 2015).

Quite several vast bodies of international literature have reviewed the causes of delay in both the developing and developed economies of the world. For instance, Venkatesh & Venkatesan, (2017) reviewed 53 causes of construction delay from different countries categorizing them into two: developing and developed countries. Their results displayed the varying nature of the top 10 causes of delay from country to country. Developing countries: delay in payments by clients; contractor's financial difficulties; deficiencies in planning & scheduling; political instability; etc. Developed countries: weather; delay in drawings, changes & errors in designs; subcontractor & supplier related causes; change orders; slow decision making; delay in approvals; poor site conditions; contractor's financial difficulties; etc.

In the United Kingdom (UK), with reference to (Sullivan & Harris, 1986), unanticipated delay in large construction project can occur due to the following: variation order, design complexity, delay delivery, bad weather, industrial disputes/strikes, pandemics, physical obstructions and significant contractual disputes. They concluded with recommendations for more team building and a greater integration of skills particularly at the early stages of planning a project. An investigation of root causes of delays in highway construction by Ellis & Thomas, (2002) yielded 8 major categories: business practices; procedures; contractors management of scheduling and planning; utilities; differing or unforeseen site conditions; maintenance of traffic; design errors and omissions. Ahmed et al., (2003) produced critical review on the 10 causes of delays in building projects in the Florida region of USA.

According to Mishmish & El-Sayegh, (2018), the most frequent causes of claims in road construction projects in the United Arab Emirates (UAE) are variation; contractor's delay and inadequate site investigation before bidding. Doloi et al., (2012) through questionnaire and interviews analysed factors affecting delays in Indian construction projects. Their report proved the following as the most critical factors: lack of commitment; inefficient site management; poor site coordination; improper planning; lack of clarity in project scope; lack of communication; and substandard contract. Also, Doloi, Sawhney & Iyer, (2012) argued that lack of commitment on contractor's inefficiency, lack of efficient construction planning and client's influence are the major factors affecting delay in Indian construction projects.

A discovery by Chen et al., (2019) detailed five delay causes for grain bin construction projects in China as: shortage of adequate equipment; poor communication among contracting parties; problems with subcontractors; inadequate experience of the design team and frequent change orders by clients. The causes and effects of delay on building construction project delivery time in Nigeria was surveyed by Owalabi et al., (2014) with results showing 15 factors similar to the ones earlier mention by Mansfield, Ugwu & Doran, (1994; Odeyinka, (1997). Furthermore, Ogunde et al., (2017) studied the cause of delay of construction projects in a megacity (Lagos) in Nigeria. They identified 33 major causes and reported the 3 most important ones as: cash flow problems during construction; clients' financial difficulties and poor procurement.

A study on causes of delay in Australia, Malaysia & Ghana construction project by Shah, (2016) revealed that the most important factors in Australia are: planning and scheduling deficiencies; methods of construction; effective monitoring and feedback process, whereas in Ghana: delay in payment certificates; underestimating of project cost; complexity of projects are the most influential factors. However, in Malaysia: contractor's improper planning; poor site management and inadequate contractor experience are the most principal factors. Consequently, these factors/causes identified from these literatures are aggregated as common factors affecting construction projects delay. The next section will detail how these common factors were used to conduct survey of experts to establish the most applicable factors affecting construction projects delay.

3. RESEARCH METHODS

A review of existing literatures on influencing factors of construction projects delay was used to establish the most applicable factors there by fulfilling part of the first objective. Nine applicable factors (see Table 1) were consolidated at the end of the review which was pre-empted as search results became repetitive. These factors were used to design a survey in form of questionnaire to fulfil the remaining part of the first objective.

Table 1
List of Features and Target.

Section ID	Factor ID	Factors
A	F1	Structural design variations
	F2	Changed orders/ discrepancies in contract documents
	F3	Price fluctuation
	F4	Contract management
	F5	Decision making
	F6	Material procurement
	F7	Site conditions
	F8	Political Influence
B	F9	Late delivery of materials by supplier
C	F10	Project delay

The questionnaire was divided into three sections such that each section deals with a specific feature of event under investigation (delay factors). Section A asked the responders to rate how 8 factors affected the duration of the project. Section B enquired what percentage a responder would give to one factor. All these made a total of nine factors as features (independent variables). Also, the responders were asked to rate how long the entire project delayed for in the final section C; this represents the target (dependent variable) for the several

ML model development. In the end, a total of 302 questionnaire were distributed and a total of 120 responses were received.

The questions were designed on a Likert scale with a scale of one to five. Although questions in each section were analysed individually, they were also linked together in such a way that their respective answers accumulatively helped to arrive at a finding (delay). The use of questionnaire research signifies independent observation – implies the questionnaire will be completed in the absence of the researcher, and since one of the objectives of this study is set out to establish the true (most) applicable factor to construction delay makes it a positivist research.

Prior to distribution of the questionnaire, pilot testing was conducted by asking group of experts in construction to comment on the representativeness and suitability of the questions. This was done to ensure thorough understanding of the questions by the responders and to avoid errors when recording data, to assess questions validity and the likely reliability of data to be collected (Saunders et al., 2015: p.425). The responders of the questionnaire were experienced stakeholders from the construction industries in Nigeria.

The collected data received via Google forms were extracted and converted into a comma-separated values file. To achieve the second objectives, this raw data was pre-processed into a clean and an analysable dataset by carrying out data imputation and outlier detection. The resulting clean, and pre-processed dataset was split randomly into two in a ratio of 79% to 21% of training dataset and testing dataset respectively. Several ML algorithms were imported into a running instance of Jupiter Notebook using Scikit-learn - an integral Python programming language module with a broad spectrum of state-of-the-art algorithms for supervised and unsupervised medium-scale problems (Pedregosa et al., 2011).

Since these ML algorithms fit input variables (delay factors) to a known output variable (delay) supervised modelling taxonomy was undoubtedly chosen. All classification algorithms (27 of them) available in scikit-learn version 0.23.1 were used for experimentations in this study. The unseen test dataset (21% of total dataset) was used to fit these 27 ML algorithms in their respective default parameters. As the 27 ML algorithms are all classifiers and the target variable is binary, stratified k-fold a variant of k-fold that returns stratified folds containing about the same proportion of target class as the initial dataset was used for cross-validation where k=10. Finally, accuracy, balanced accuracy, ROC AUC, and f_1-score modelling evaluation metrics were employed to measure the new model performances on the testing dataset.

4. MAIN DISCUSSION

An initial investigation on the data through Exploratory Data Analysis (EDA) showed that the data is a two - dimensional array with 120 rows and 10 columns where the 1st to the 9th columns (F1 – F9 factor IDs) represent the features/ independent variables and the 10th column (F25) represent the target/ dependent variable. Column F10 (see Table 1) were encoded into 0 (no delay) for its ordinal values containing any number <= 3 and 1(delay) for its ordinal values containing 4 and 5 such that on each occurrence, this value (0 or 1) can then help to show if a category (delay) is present or not. Thereafter, the dataset was split using Scikit-learn's train_test_split function at a ratio of 79:21 for training and testing respectively. This pre-processed dataset was used to fit into the following ML algorithms (see Table 2 below) using their respective Scikit-learn's library:

Table 2

Machine Learning Algorithms and Libraries

S/N	Algorithms	Scikit-learn Library
1	Perceptron	Perceptron
2	Linear Support Vector Machine	LinearSVC
3	Linear Discriminant Analysis	LinearDiscriminantAnalysis
4	Support Vector Machine	SVC
5	Ridge (with built-in cross-validation)	RidgeClassifierCV
6	Ridge	RidgeClassifier
7	Random Forest	RandomForestClassifier
8	Passive Aggressive	PassiveAggressiveClassifier
9	Gaussian Naive Bayes	GaussianNB
10	Logistic Regression	LogisticRegression
11	Nearest Centroid	NearestCentroid
12	AdaBoost	AdaBoostClassifier
13	Label Spreading	LabelSpreading
14	Label Propagation	LabelPropagation
15	Bagging	BaggingClassifier
16	Extra-Trees	ExtraTreesClassifier
17	Probability Calibration	CalibratedClassifierCV
18	Bernoulli Naive Bayes	BernoulliNB
19	Extreme Gradient Boosting	XGBClassifier
20	Nu-Support Vector Machine	NuSVC
21	Quadratic Discriminant Analysis	QuadraticDiscriminantAnalysis
22	Light Gradient Boosting Machine	LGBMClassifier
23	K-Nearest Neighbors	KNeighborsClassifier
24	Extremely Randomized Tree	ExtraTreeClassifier
25	Decision Tree	DecisionTreeClassifier
26	Dummy Estimator	DummyClassifier
27	Stochastic Gradient Descent	SGDClassifier

Due to the sample size of the dataset, a stratified10-fold cross validation resampling technique was used to evaluate the performance of machining learning models developed using these 27 ML algorithms in order to mitigate overfitting of the individual models to the test dataset. Table 3 below presents performance evaluation metrics for all the 27 models developed. More precisely, it reveals the accuracy, balanced accuracy, ROC AUC, and f_1-score computed using the stratified10-fold cross validation for the 27 ML algorithms.

Table 3

Performance Evaluation Metrics Report

Model	Accuracy	Balanced Accuracy	ROC AUC	F_1-Score
Perceptron	0.85	0.85	0.85	0.85
Linear Support Vector Machine	0.81	0.81	0.81	0.81
Linear Discriminant Analysis	0.81	0.81	0.81	0.81
Support Vector Machine	0.81	0.81	0.81	0.81
Ridge (with built-in cross-validation)	0.81	0.81	0.81	0.81
Ridge	0.81	0.81	0.81	0.81
Random Forest	0.81	0.81	0.81	0.81
Passive Aggressive	0.81	0.81	0.81	0.80
Gaussian Naive Bayes	0.81	0.81	0.81	0.81
Logistic Regression	0.81	0.81	0.81	0.81
Nearest Centroid	0.77	0.77	0.77	0.77
AdaBoost	0.77	0.77	0.77	0.77
Label Spreading	0.77	0.77	0.77	0.77
Label Propagation	0.77	0.77	0.77	0.77
Bagging	0.77	0.77	0.77	0.77
Extra-Trees	0.77	0.77	0.77	0.77
Probability Calibration	0.77	0.77	0.77	0.77
Bernoulli Naive Bayes	0.77	0.77	0.77	0.77
Extreme Gradient Boosting	0.77	0.77	0.77	0.77
Nu-Support Vector Machine	0.73	0.73	0.73	0.73
Quadratic Discriminant Analysis	0.73	0.73	0.73	0.73
Light Gradient Boosting Machine	0.73	0.73	0.73	0.73
K-Nearest Neighbors	0.69	0.69	0.69	0.69
Extremely Randomized Tree	0.65	0.65	0.65	0.65
Decision Tree	0.65	0.65	0.65	0.65
Dummy Estimator	0.62	0.62	0.62	0.61
Stochastic Gradient Descent	0.50	0.50	0.50	0.43

Comparatively, the Perceptron model became the top performant model having achieved an accuracy, balanced accuracy, ROC AUC, and f_1-score of 85%, 85%, 0.85 and 085 respectively higher than the rest of the models. This is an excellent performance level that is tending towards perfection and unachieved in any previous study, well justifying the need for trailing multiple algorithms when developing forecasting/predictive model.

Gaussian Naive Bayes model also had a great performance with accuracy, balanced accuracy, ROC AUC, and f_1-score of 81%, 81%, 0.81 and 081 respectively higher than the previous studies of Gondia et al. (2020) in Egypt and Yaseen et al. (2020) in Qatar. As Random Forest is built upon Decision Tree, it is no surprising from our research that Random Forest was a

better model than Decision Tree having achieved an accuracy, balanced accuracy, ROC AUC, and f_1-score of 81%, 81%, 0.81 and 081 respectively than it.
The Receiver Operator Characteristic Curve (ROC) is a recall plot of the y-axis against precision of the x-axis. Area under the curve (AUC) is the area under the ROC curve that is generally recognised as the best indicator of the overall performance of a classifier. As the maximum AUC value, indicating excellent accuracy, is 1 and the minimum AUC value, is 0.5, all the 27 algorithms (excluding Stochastic Gradient Descent) implemented in this study appears to have performed pretty well.

5. CONCLUSIONS AND RECOMMENDATIONS

Despite significant mitigating measures, the recurrence of a worldwide issue - construction projects delay – remains a major source of worry for its policymakers. Surprisingly, this industry, which generates massive amounts of data from IoT sensors and building information modelling on a daily basis, has been slow to adopt the most recent analysis method – artificial intelligence/machine learning, which best explains the factors that can affect a phenomenon like delay based on its predictive capabilities, having been widely adopted across other industry. As a result, a foundation for using numerous machine learning algorithms to anticipate construction project delays was architected, implemented, and reported in this study. First, a study of the current body of knowledge on the factors that influence construction project delays was utilised to conduct an expert survey as a method of data collecting and investigation. The resulting dataset applied to 27 ML algorithms was used to develop 27 predictive models. Results from the algorithm evaluation metrics: accuracy, balanced accuracy, ROC AUC, and f_1-score indeed proved Perceptron model as the top performant model having achieved an accuracy, balanced accuracy, ROC AUC, and f_1-score of 85%, 85%, 0.85 and 085 respectively higher than the rest of the models and unachieved in any previous study in predicting construction projects delay. Hence this study contributes to the effort of improving time efficiency of construction projects – a key performance indicator for successful projects. Ultimately, this model can subsequently be integrated into construction information system to promote evidence-based decision-making, thereby enabling constructive project risk management initiatives. Thus, will help improve the quality of decisions and risks to be taken by the policy makers on their present or future construction projects which as a result will foster trust, increase in productivity and revenue and more importantly yield timely delivery of construction projects in the industry. While the proposed contemporary method of analysis is assumed to be applicable in mitigating delay of any construction project within the sector, the unique data transformation employed in this study may not, as typical of any data driven model, be transferable to the data from other regions. In order to produce improved classification outcomes, future studies should be targeted at extending the algorithms either by parameter optimization or feature engineering.

6. REFERENCES

Abdul-Rahman, H., Takim, R., & Min, W. S. (2009). Financial-related causes contributing to project delays. *Journal of Retail and Leisure Property*, *8*(3), 225–238. https://doi.org/10.1057/rlp.2009.11

Abdul-Rahman, H., Wang, C., Takim, R., & Wong, S. (2011). Project schedule influenced by financial issues: Evidence in construction industry. *Scientific Research and Essays*, *6*(1), 205–212. https://doi.org/10.5897/SRE10.989

Ahmed, S. M., Azhar, S., Kappagantula, P., & Dharam, G. (2003). Delays in Construction: A Brief Study of the Florida Construction Industry. *ASC Proceedings of the 39th Annual Conference Clemson University*, 257–266. http://ascpro0.ascweb.org/archives/cd/2003/2003pro/2003/Ahmed03.htm

Akhund, M. A., Khoso, A. R., Memon, U., & Khahro, H. (2017). Time overrun in construction projects of developing countries. *Imperial Journal of Interdisciplinary Research (IJIR)*, *3*(5), 124–129.

Al-Hazim, N., Salem, Z. A., & Ahmad, H. (2017). Delay and Cost Overrun in Infrastructure Projects in Jordan. *Procedia Engineering*, *182*, 18–24. https://doi.org/10.1016/j.proeng.2017.03.105

Andrew, R. (2013). UK Construction: An economic analysis of the sector. In *Department for Business Information & Skills* (Issue July). https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/210060/bis-13-958-uk-construction-an-economic-analysis-of-sector.pdf

Asadi, A., Alsubaey, M., & Makatsoris, C. (2015). A machine learning approach for predicting delays in construction logistics. *International Journal of Advanced Logistics*, *4*(2), 115–130. https://doi.org/10.1080/2287108x.2015.1059920

Assaf, S. A., & Al-Hejji, S. (1995). Causes of delay in large construction projects. *International Journal of Project Management*, *24*(4), 349–357. https://doi.org/10.1016/j.ijproman.2005.11.010

Assaf, S. A., & Al-Hejji, S. (2006). Causes of delay in large construction projects. *International Journal of Project Management*, *24*(4), 349–357. https://doi.org/10.1016/j.ijproman.2005.11.010

Chen, G. X., Shan, M., Chan, A. P. C., Liu, X., & Zhao, Y. Q. (2019). Investigating the causes of delay in grain bin construction projects: the case of China. *International Journal of Construction Management*, *19*(1), 1–14. https://doi.org/10.1080/15623599.2017.1354514

Doloi, H., Sawhney, A., & Iyer, K. C. (2012). Structural equation model for investigating factors affecting delay in Indian construction projects. *Construction Management and Economics*, *30*(10), 869–884. https://doi.org/10.1080/01446193.2012.717705

Doloi, H., Sawhney, A., Iyer, K. C., & Rentala, S. (2012). Analysing factors affecting delays in Indian construction projects. *International Journal of Project Management*, *30*(4), 479–489. https://doi.org/10.1016/j.ijproman.2011.10.004

Egan, S. J. (2018). Rethinking the Report of the Construction Task Force. *Construction Task Force*. http://constructingexcellence.org.uk/wp-content/uploads/2014/10/rethinking_construction_report.pdf

Elawi, G. S. A., Algahtany, M., & Kashiwagi, D. (2016). Owners' Perspective of Factors Contributing to Project Delay: Case Studies of Road and Bridge Projects in Saudi Arabia. *Procedia Engineering*, *145*(480), 1402–1409. https://doi.org/10.1016/j.proeng.2016.04.176

Ellis, R. D., & Thomas, H. R. (2002). the Root Causes of Delays in Highway Construction. *Presentation at the 82nd Annual Meeting of the Transportation Research Board, Washington D.C.*, 1–14.

Flyvbjerg, B. (2014). What you should know about megaprojects and why: An overview. *Project Management Journal*, *45*(2), 6–19. https://doi.org/10.1002/pmj.21409

Gardezi, S. S. S., Manarvi, I. A., & Gardezi, S. J. S. (2014). Time extension factors in construction industry of Pakistan. *Procedia Engineering*, *77*, 196–204. https://doi.org/10.1016/j.proeng.2014.07.022

Głuszak, M., & Lešniak, A. (2015). Construction Delays in Clients Opinion - Multivariate Statistical Analysis. *Procedia Engineering*, *123*, 182–189. https://doi.org/10.1016/j.proeng.2015.10.075

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. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001736

Haq, S., Rashid, Y., Ph, D., & Aslam, M. S. (2017). Effects of Delay in construction Projects of Punjab-Pakistan : An Empirical Study Effects of Delay in construction Projects of Punjab-Pakistan : *Journal of Basic and Applied Scientific Research.*, *3*(January 2014), 87–96.

Ji, Y., Qi, L., Liu, Y., Liu, X., Li, H. X., & Li, Y. (2018). Assessing and prioritising delay factors of prefabricated concrete building projects in China. *Applied Sciences (Switzerland)*, *8*(11), 2324. https://doi.org/10.3390/app8112324

Mansfield, N., Ugwu, O., & Doran, T. (1994). Causes of delay and cost overruns in Nigerian construction projects. *International Journal of Project Management*, *12*(4), 254–260. https://doi.org/10.1016/0263-7863(94)90050-7

Mishmish, M., & El-Sayegh, S. M. (2018). Causes of claims in road construction projects in the UAE. *International Journal of Construction Management*, *18*(1), 26–33. https://doi.org/10.1080/15623599.2016.1230959

Odeh, A. M., & Battaineh, H. T. (2002). Causes of construction delay: Traditional contracts. *International Journal of Project Management*, *20*(1), 67–73. https://doi.org/10.1016/S0263-7863(00)00037-5

Odeyinka, H., & Adebayo, Y. (1997). "The Causes and Effects of Construction Delays on Completion Cost of Housing Projects in Nigeria." *Journal of Financial Management of Property and Construction*, *2*(3), 31-44. https://www.researchgate.net/publication/249643683

Ogunde, A. O., Dafe, O. E., Akinola, G. A., Ogundipe, K. E., Oloke, O. C., Ademola, S. A., Akuete, E., & Olaniran, H. F. (2017). Factors Militating Against Prompt Delivery of Construction Projects in Lagos Megacity, Nigeria: Contractors' Perspective. *Mediterranean Journal of Social Sciences*, *8*(3), 233–242. https://doi.org/10.5901/mjss.2017.v8n3p233

Orangi, A., Palaneeswaran, E., & Wilson, J. (2011). Exploring delays in Victoria-based Astralian pipeline projects. *Procedia Engineering*, *14*, 874–881. https://doi.org/10.1016/j.proeng.2011.07.111

Owalabi, J. D., Amusan, L. M., Oloke, C. O., Olusanya, O., Tunji-Olayeni, P., OwolabiDele, PeterJoy, & OmuhIganatious. (2014). Causes and effects of delay on project construction delivery time. *International Journal of Education and Research*, *2*(4), 197–208. www.ijern.com

Pedregosa, F., Grisel, O., Weiss, R., Passos, A., Brucher, M., Varoquax, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., & Brucher, M. (2011). Scikit-learn: Machine Learning in Python. *Journal of*

Machine Learning Research, *12*(85), 2825–2830. http://scikit-learn.sourceforge.net.

Ramanathan, C., Narayanan, S. P., & Idrus, A. B. (2012). Construction delays causing risks on time and cost - A critical review. In *Australasian Journal of Construction Economics and Building* (Vol. 12, Issue 1, pp. 37–57). Australian Institute of Quantity Surveyors. https://doi.org/10.5130/ajceb.v12i1.2330

Rhodes, C. (2019). Construction Industry: Statistics and policy. *House of Commons Library*, *01432*, 1–13.

Robinson T, G. (2015). Global construction market to grow $ 8 trillion by 2030 : driven by China , US and India. *Global Construction*, *44*(0).

Saunders, M., Lewis, P., Thornhill, A., Lewis, S., & Thornhill. (2009). *Research methods for business students fi fth edition* (Fifth edit). Pearson Education Limited. www.pearsoned.co.uk

Shah, R. K. (2016). An Exploration of Causes for Delay and Cost Overruns In Construction Projects: Case Study of Australia, Malaysia & Ghana. *Journal of Advanced College of Engineering and Management*, *2*, 41. https://doi.org/10.3126/jacem.v2i0.16097

Sullivan, A., & Harris, F. C. (1986). Delays on Large Construction Projects. *International Journal of Operations & Production Management*, *6*(1), 25–33. https://doi.org/10.1108/eb054752

Venkatesh, P. K., & Venkatesan, V. (2017). Delays in construction projects: A review of causes, need and scope for further research. *Malaysian Construction Research Journal*, *23*(3), 89–113.

Woetzel, J., Garemo, N., Mischke, J., Kamra, P., & Palter, R. (2017). Bridging Infrastructure Gaps Has the World Made Progress? *McKinsey Global Institute*, *October*, 10. https://www.mckinsey.com/~/media/McKinsey/Industries/Capital Projects and Infrastructure/Our Insights/Bridging infrastructure gaps has the world made progress/Bridging infrastructure gaps How has the world made progress v2/MGI-Bridging-infrastructure-gaps

Yaseen, Z. M., Ali, Z. H., Salih, S. Q., & Al-Ansari, N. (2020). Prediction of Risk Delay in Construction Projects Using a Hybrid Artificial Intelligence Model. *Sustainability*, *12*(4), 1514. https://doi.org/10.3390/su12041514