ORIGINAL ARTICLE

Using machine learning to analyze and predict construction task productivity

Laura Florez-Perez*[1] | Zhiyuan Song[2] | Jean C. Cortissoz[3]

[1] Bartlett School of Sustainable Construction, University CollegeLondon, London, WC1E 7HB, UK
[2] Department of Computer Science, University College London, London, WC1E 7HB, UK
[3] Department of Mathematics, Universidad de los Andes,Bogotá DC, 111711, Colombia

***Correspondence**
Laura Florez-Perez. University College London, London, WC1E 7HB, UK. Email: l.florez@ucl.ac.uk

Abstract

The factors that affect productivity is a major focus in construction.This article proposes a machine learning-based approach to predict task productivity by using a subjective measure (compatibility of personality), together with external and site conditions, and other workers characteristics. The approach integrates KNN, DNN, Logistic regression, SVM and ResNet18 to discover the mapping between input and output variables, alongside rigorous statistical analyses to interpret data. A database including 1,977 productivity measures is utilized to train, test, and validate the approach. Results test rules in the masonry industry, which do not seem to have been tested before: small crews are more productive than large crews; higher compatibility results in higher productivity in easy but not in difficult tasks; the relevance of experience to task productivity may depend on the difficulty of the

KEYWORDS
Construction crews, productivity prediction, machine learning, statistical

1. INTRODUCTION

The factors that affect productivity in construction sites is a major focus in the construction industry. In the field, superintendents are faced with multiple factors such as external conditions, site conditions, and workers characteristics that have an effect on the task productivity of construction crews. The factors' effects and interrelationships need to be considered by superintendents and project managers when planning work to better predict task productivity and identify which factors will have a negative impact, so that they can take the necessary actions such as reducing complexity of tasks, identifying which workers will be part of a crew, determining optimal crew size, and allocating workers and crews to the right tasks. Existing modeling approaches that aim to accurately predict task productivity of construction crews do not consider an essential worker characteristic (compatibility of personality) and the interrelationships between these characteristics, site and external conditions; thus they are not fit for the complexity of jobsites and the intrinsic subjectivity of the workforce. Therefore, a model to better understand the interactions between the factors and their combined effects to better predict task productivity needs to be developed.

Machine learning (ML) is a technology already impacting the global economy and has the potential to transform the construction industry with the use of data-based solutions to improve the way projects are delivered (Reich (1997), Adeli (2001); Teizer & Vela (2009); J. Wang & Ashuri (2016); Celebi-Aydin, 2016; Gao, Shi, Song, Zhang, & Zhang (2019); Fang et al. (2018)). Poor planning and unrealistic productivity predictions cause increased delivery costs and time. There is a potential to use existing data and experience to improve project planning with more accurate productivity predictions with the use of ML. However, challenges to apply ML techniques for predicting productivity remain unsolved. Firstly, the identification of factors has been mainly based on review of literature, surveys, and expert interviews (Ebrahimi, Fayek, & Sumati (2021)) with little use of quantitative robust methods to model the complexity due to non-linearity of underlying factors and determine the weights, relationships, and effects between the productivity factors (Dissanayake, Fayek, Russell, & Pedrycz. (2005)). Secondly, the data generation and quality does not reflect the reality and complexity of the site environment (Xu, Zhou, Sekula, & Ding (2021)). When dealing with productivity prediction, it is essential to not only build and train reliable ML models, but also consider how to integrate site realities and the knowledge of construction industry experts into the modelling process as a comprehensive framework (Bilal & Oyedele (2020)). This is because some essential subjective characteristics associated to the workforce affect productivity and have not been considered previously. Furthermore, industry experts have rules that have been used for years, often passed from one project to another, but these have yet to be tested with real life data. To test rules, ML techniques can be used, but these require high-quality training data to adaptively fit the data and form an accurate, robust model to make accurate predictions (Xu et al. (2021)). Although techniques such as Delphi (Alaloul, Liew, Zawawi, Mohammed, & Adamu (2018)) and Relief algorithm (Ebrahimi et al. (2021)) have been adopted to minimize the negative influence of subjective data collection, this does not completely remove the error caused by the subjective bias of experts. How to derive factors based on objective data and identify the relative importance and interrelationship of factors must be explored.

In this article, a combined ML approach that offers a solution based on experimental data, is proposed for predicting task productivity with increased confidence and certainty, and will be explored with data of masonry tasks with concrete masonry units (CMU). A common practice to build crews by superintendents based on a subjective measure of "how masons get along" (referred to in this paper as "compatibility") has also been tested, and although found not completely adequate by the ML methods employed, once a bit of statistical analysis is used, it reveals some interesting findings that might be helpful in applications. That is, ML methods were not only used to predict, but based on how the accuracy of the methods was affected by removing or fixing some inputs gave the motivation to perform some statistical analysis. All in all, with the results of such tests, strategies to better plan, schedule, and manage future projects are proposed. The rest of this article is organized as follows: Section 2 presents the literature review to introduce existing work on machine learning and productivity predictions, whereas the concepts of productivity and masonry construction are introduced in Section 3. Section 4 presents the methodology. Section 5 presents the ML techniques and details the implementation procedure of the proposed algorithms. Section 6 presents feature selection and Section 7 the statistical analyses. Finally conclusions and opportunities for future research are presented in Section 8.

2. LITERATURE REVIEW

The spread of ML techniques has raised various emerging opportunities in the construction industry. ML has been applied to construction research for more than two decades. Specifically, productivity studies have benefited from the application of ML because these techniques serve to determine the relationship between the influencing factors and productivity rates and the complexity of the combined effects between factors (Chao & Skibniewski (1994); Karim & Adeli (1999); Boussabaine & Kirkham (2008); Franco & Santurro (2020)). ML has been used in the design stage (As, Pal, & Basu (2018); Valikhani, Jahromi, Pouyanfar, Mantawy, & Aziz-inamini (2020); Rodrigues et al. (2017); Huang and Zheng, 2018; Luo & Paal (2020)), construction stage (Poh, Ubey-narayana, & Goh (2018); Bai et al. (2019); Xiong Huber, 2010; Yu et al. (2019)), and the operation and maintenance stage (Maeda, Kashiyama, Sekimoto, Seto, & Omata (2020); Okazaki, Okazaki, Asamoto, & jo Chun (2020); Yang & Su (2008); Farrar Worden, 2012; Zhang, Waller, & Jiang (2019)). ML has also been used in civil engineering applications such as damage assessment of buildings (Cheng, Behzadan, & Noshadravan (2021)), flood warning (Dong, Yu, Farahmand, & Mostafavi (2020)), risk assessment of infrastructure systems (Tomar & Burton (2021)). Supervised learning is a machine learning approach that learns a function that maps input to output based on example input-output pairs and infers a function from labelled training data consisting of a set of training examples (Caruana & Niculescu-Mizil (2006); Hastie, Tib- shirani, & Friedman (2008a)). Supervised learning, including logistic regression (Wilson, Sharpe, & Kenley (1987)), support vector machine (SVM) (Shu-quan et al. (2006)), random forest (Z. Liu, Sadiq, Rajani, & Najjaran (2010)), and K-nearest neighbor (KNN) (J. Wang & Ashuri (2016); Fang et al. (2018)) are the most widely used type of machine learning algorithms in the construction field. Unsupervised learning that focuses on data reduction and clustering problems is another approach of machine learning. No pre-labelled training examples are given, and the input data is automatically classified or grouped (Barlow (1989); Celebi & Aydin (2016); Hastie, Tibshirani, & Friedman (2008b)). Principal component analysis (Dobge- gah, Owusu-Manu, & Omoteso (2011)) and K-means (Lee & Chang (2005)) have also been used in construction. Because in the construction industry the unlabelled data has many limitations (relatively speaking) and the information that can be extracted is less than the labelled data, supervised learning is mostly used for data classification (Xu et al. (2021)). For processing unstructured data, such as images and videos, deep learning is a better choice than shallow learning (Fang et al. (2018); Gao et al. (2019)). Shallow learning is mainly used in architectural scenarios with structured data, such as safety monitoring early warning indicators. Deep learning applications can be divided into four categories: object detection, image segmentation, action recognition, and natural language processing (Xu et al. (2021)). The complex environment of the construction site causes various conditions such as poor lighting and object occlusion, which will affect the performance of deep learning methods. Therefore, it is imperative to embed vast amount of experience and knowledge into machine learning methods.

Despite the wide use of ML in construction, certain challenges remain unsolved. One challenge, which affects their performance is how to obtain large amounts of labelled data. This is because, when a data set is small, it may lead to a biased sample. This means that the sample points in a small data set do not represent the distribution of the target domain and the underlying patterns in the target data (Quionero-Candela, Sugiyama, Schwaighofer, & Lawrence (2019)), leading to bias in the trained ML model for prediction in the target domain. Another challenge is how to choose the most appropriate algorithm strategy in the face of different application scenarios. For instance, linear regression is not recommended for most practical applications. It oversimplifies real-world situations and is not a good fit for dealing with non-linear relationships. Logistic regression cannot deal with non-linear issues and cannot work out well unless all independent variables are identified. SVM cannot play its due role when there is noise in the data set, thereby reducing the accuracy of the results. At the same time, SVM cannot provide probability estimates, and it is challenging to understand the final model. As ML methods have their limitations and might not be suitable for construction, practitioners need to establish correspond- ing supervised learning methods and data collection based on specific architectural application scenarios. For instance, reinforcement learning is an algorithm based on trial and error, resulting in high development costs (Xu et al. (2021)). Reinforcement learning is currently very limited in the modeling of construction labor productivity compared to supervised learn- ing and unsupervised learning. Although methods such as ANN, random forest, and SVM are already applied for modelling construction labor productivity, the use of other advanced ML approaches is still very limited. Besides, ML approaches can be black boxes, as to how the model learns and how it generates predictions. The predictions generated by the model are useful, few complimentary methods for analyzing ML results are often used (Bilal & Oyedele (2020)). Thus, a framework to involve both applied ML, construction knowledge, and statistical analyses can expand practitioners' understanding of the model and actual engineering problems so that they can be correctly used in construction scenarios.

For the specific scenario to deal with the problem associated with productivity prediction, there are many ML approaches that have been proposed. Some approaches include artificial neural network (ANN) (Golnaraghi, Zangenehmadar, Moselhi, & Alkass (2019); Alaloul et al. (2018); El-Gohary, Aziz, & Abdel-Khalek (2017); Moselhi & Khan (2012); Nasirzadeh et al. (2020); Golnaraghi et al. (2019); Portas & AbouRizk (1997); Tsehayae & Fayek (2016); Badawy, Hussein, Elseufy, & Alnaas (2019), Heravi & Eslamdoost (2015)), computational intelligence (Dissanayake et al. (2005)), neurofuzzy (Boussabaine (2001); Mirahadi & Zayed (2016)), self organizing maps (Oral, Oral, & Andaç (2016)), random forest (Ebrahimi et al. (2021); Liu et al., 2018; Momade, Shahid, bin Hainin, Nashwan, & Umar (2020); Awada, Srour, & Srour (2021)), ML classifiers (Jassmi, Ahmed, Philip, Mughairbi, & Ahmad (2019), support vector machine (SVM) (Momade et al. (2020)). Determining the factors that affect the productivity of construction labor is often the first step in establishing ML models. The performance of these models greatly depends on the input factors. Factors include contract and delivery method (Alaloul et al. (2018)), management and supervision (leadership and competency, trust in foreman, fairness in review) (El-Gohary et al. (2017); Momade et al. (2020)), external conditions (temperature, humidity, wind speed, precipitation) (Golnaraghi et al. (2019); Dissanayake et al. (2005)), site conditions (equipment, floor level, work type, workload, complexity of task, congestion, interruptions) (Golnaraghi et al. (2019); Dissanayake et al. (2005); Ebrahimi et al. (2021); El-Gohary et al. (2017)), and workers characteristics (age, experience, skill, crew size, team spirit, happiness (Oral et al. (2016); Alaloul et al. (2018); El-Gohary et al. (2017); Jassmi et al. (2019)).

A noteworthy point is that many approaches that have been proposed to deal with the problem of predicting productivity ignore the correlation between different types of productivity factors and only consider these as independent and isolated factors. In addition, the extent to which the potential drivers of productivity are accurately identified determines the use- fulness of productivity measurement frameworks and tools for practitioners and policy makers (H. R. Thomas, Guevara, & Gustenhoven (1984)). This requires that the proposed approach is essentially capable of describing the production and construction process correctly and interpreting the construction output as accurately as possible based on the quantity and quality of the inputs used to generate it (Crawford & Vogl (2006)). Some studies have determined the relative importance by surveying experts (Momade et al. (2020); Alaloul et al. (2018) or a hybrid feature selection method (Ebrahimi et al. (2021)). Most studies simplify the correlation between the factors, or even ignore them (Nasirzadeh & Nojedehi (2013)). The more fundamental factors will play a more accurate role in predicting, which requires the model to first clarify the hierarchical relationship between the factors. How to derive factors based on objective data and identify the relative importance and inter- relationship of factors must be explored. These factors may vary from project to project and from one region to the other (El-Gohary & Aziz (2014); Enshassi, Mohamed, Mustafa, & Mayer (2007); Hafez (2014), Hamza, Shahid, Hainin, & Nash- wan (2019); A. V. Thomas & Sudhakumar (2014); Hiyassat, Hiyari, & Sweis (2016); Jarkas & Bitar (2012)), so a set of project and enterprise-based factor mining and analysis tools are needed. In contrast to models that depend on experts' opinions, future new models can obtain objective information from real data to export behavior or data-rules, identify critical factors and predict productivity performance. By neglecting the correlation between construction labor productivity factors, some existing studies have failed to identify the fundamental factors and develop sensible strategies to better predict labor productivity. Hybrid systems-based machine learning (ML), optimization algorithms, and simulation techniques have been applied in several construction problems because they are superior to sole artificial intelligence techniques (Ebrahimi et al. (2021)). Some studies have considered workers characteristics, but have not considered how the compatibility of personality between the workers in a crew can affect task productivity. In addition, some studies have presented technical methodologies, but have not tested rules that are widely used in the masonry industry to form crews.

3. PRODUCTIVITY AND MASONRY CONSTRUCTION

There are many definitions of productivity. Generally speak- ing, productivity is a measure of the full utilization of inputs to achieve an expected output. This measure can fit well with var- ious definitions of productivity in different contexts (Durdyev & Mbachu (2011)). Detailed methods to measure productivity can be categorized at the industry level, project level, and task level. In the field, productivity is measured at the task level for practical considerations. Therefore in this article, the task-level model will be used as single factor productivity, expressed as the the unit of work per labor hours (Shehata & El-Gohary (2011)). As masonry is one of the most labor-intensive trades in construction (Lowe (1987); Ng & Tang (2010)), it was used in the application of the ML approach. To detail the factors that impact task productivity, three sections namely external conditions, site conditions, and workers characteristics describe typical attributes of masonry jobsites.

3.1. External Conditions

These conditions refer to location and temperature. Location was stated regarding the building and the story the crews were working at a specific time the data was collected.The temperature, both low and high temperature, were recorded for the day at the time the data was collected.

3.2. Conditions in masonry sites

Extensive site observations and interviews with masonry practitioners (Florez (2015)) were used to inform typical site conditions related to walls.

FIGURE 1 Easy Wall (Wall Difficulty = 1). Picture by L. Florez-Perez.

3.2.1. Easy Walls (difficulty=1)

An easy wall is the most common type of wall in a masonry project (see Figure 1). It is a straight wall with no openings. Because it is a line and there are no openings, it is built using a string line. Since there is no difficulty in this wall, it is the fastest wall to build and the highest productivity rates are expected for this type of wall. To build an easy wall, a mason uses a string line and does not need to constantly level and mark cuts and details.

3.2.2. Normal Walls (difficulty=2)

A normal wall is the second most common type of wall in a masonry project (see Figure 2). It is a straight wall with a few openings such as doors, window frames. The spacing between the openings ranges between 15 ft. and 20 ft. Because it is a line with few openings, it is built using a string line. Since there is no difficulty in this wall, it is relatively fast to build and the second highest productivity rates are expected for this type of wall. To build a normal wall, a mason uses string line and for the openings may need to mark some cuts and details.

3.2.3. Difficult Walls (difficulty=3)

A difficult wall is a wall that has mostly detailed and technical work such as openings, intricate corners, leads, and penetrations (see 3). This type of wall may involve building curved walls, arches, piers, and columns. The spacing between the openings can be as small as 1 ft. Because of its shape and the amount of openings and details it has, it cannot be build using a fixed string line. Since there is difficulty in this wall, it requires a high level of technical work. It is the slowest wall to build and the lowest productivity rates are expected for this type of wall. To build a difficult wall, a mason uses a plumb rule and level and has to mark cuts and details.

FIGURE 2 Normal Wall (Wall Difficulty =2). Picture by L. Florez-Perez.

FIGURE 3 Difficult Wall (Wall Difficulty=3). Picture by L. Florez-Perez.

3.3. Workers' Characteristics

Masons have different levels of experience and age. In this article, experience is defined as a hierarchical skill, that is, workers with lower experience can do less compared to workers with higher experience. Workers with more experience are more educated or have spent more years working in the masonry trade (on the job training) and therefore can perform tasks faster. Age refers to the length of time a worker has lived. It is assumed that workers with less age can do more than workers with more age since they are physically more able to move and carry the weight of the blocks.

The size of crews were annotated as it happened onsite. Typically, the superintendent determines the size of crews based on the size of walls and the workload so that they can guarantee a constant flow and adequate control. During the data collection period, the superintendent used crews of one, two, and three masons.

3.3.1. Compatibility of Labor

When grouping masons in a crew, their personality characteristics are combined and how well they match can be stated by their compatibility. In (Florez, Armstrong, & Cortissoz (2020)), a way of measuring this construct was proposed and was based in the personality factors of the masons in a crew. In that study, it was found that if the personality differences of the masons in a crew, measured via tests based on the big five personality factors, are not excessively disimilar or incompatible, their interactions create an advantageous environment that facilitates their coordination and communication, result- ing in better performance and thus higher productivity. In this study, compatibility is defined as a measure of the capability of a group to interact and work well together to attain higher productivity, and in this case it was not measured using any particular test or psychological theory. Instead, during the extensive site visits and interviews with masonry practitioners in the United States (Florez (2015)), it was found that super-intendents form crews of masons that get along well, that is, masons that in their daily interactions seem to have a good relation and can establish a good working environment: the opinion of the superintendents is then used, via a Likert scale, to grade how compatible two masons are. By forming crews with compatible masons, superintendents are convinced that the productivity is increased: this belief is tested in this work.

3.4. Data Set

In this study, a data set from a combined masonry project located in Atlanta, GA in the United States, was used to deter-mine the relationships between factors and factors' effects on tasks productivity. The project consisted of two main buildings with an approximate area of 150,000 ft^2. Building A was mixed used space for upscale commercial stores and residential apartments. Building C had only upscale commercial stores. Up until the first storey, the floor use for both buildings was identical as well as the masonry units used. The second underground floor was used for parking, the first underground was used for storage for commercial clients, and the first floor were commercial stores. Building A had 12 more floors of residential apartments.

Data collection was taken between the Summer of 2013 and Spring 2014 for about 9 months. Construction started in July 2013 with Building A and in September 2013 Building C started. By April 2014 when data collection was finalized, both buildings were still under construction although the masonry parts in both buildings was almost complete. The collected data sets were documented on-site. The data set had 1,977 records, each of which includes the following information:

- ID of the construction site (name)
- Date productivity was measured
- Low and high temperature: the daily temperature (in degrees Celsius).
- Difficulty of the wall: the complexity of the wall [1= easy; 2= normal; 3=difficult]
- Size: number of masons in the crew
- Age: age (in years) of every mason
- Experience: experience (in years) working in the masonry trade
- Compatibility of the masons: pair-wise comparison between masons [0.1= not very well; 0.5= well; 0.9= excellent]
- Productivity: number of CMU blocks placed per crew every 5 min

The data was manually entered into note pads and then transcribed into a Microsoft Excel sheet, with the number of rows equal to the number of instances (1,977 rows), and a column for each of the aforementioned variables (14 columns). The data related to external conditions and site conditions was com- piled with site visits to the projects. The data regarding workers characteristics was complimented with the workers and the superintendent. The workers were asked to provide their age and experience. To determine compatibility, the superintendent was given a brief explanation of the compatibility scores and was provided witha table to fill in the scores, similarly to the ranking he gives to his masons. The compatibility score for each pair of masons was based on the knowledge the super- intendent had about the masons' performance from previous jobs and their characteristics. Further details can be found in Florez (2017). For simplicity, the superintendent was asked to state how well the masons get along: not very well, well, and excellent. The compatibility score between the masons was determined based on the superintendent's responses (0.1 = not very well, 0.5 = well, 0.9 = excellent). Note that the scores are similar to a Likert scale, which are scaled so that the distance on each item is equal. In this study, it is considered that the productivity of a mason that works in a compatible crew will increase and hence the productivity of the crew will increase as a result.

4. METHODOLOGY

Machine learning is an application of artificial intelligence (AI) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed. In this article, supervised ML algorithms serve as a powerful tool to classify the level of productivity of construction projects. In this work, deep neural network (W. Liu et al. (2017)), K-Nearest Neighbors (Cunningham & Delany (2020)), support vector machine (G. Wang (2008)), and logistic regression (Hosmer, Lemeshow, & Sturdivant (2013)) models were used in order to predict the level of productivity using the given information.

4.1. Data Processing

The dataset contains 1977 data samples with 22 dimensions, and to extract the important factors from the dataset for training and prediction, some irrelevant information is discarded from the dataset, such as date, construction site and name of masons. The empty data points are filled by 0 for padding purpose. The dataset is divided into training and testing datasets and each input data is labelled by its corresponding productivity. To measure the productivity of the construction task, the number of blocks built per minute per person is calculated. Since the productivity may vary due to human's internal factor even though all external factors are kept constant, e.g. a mason may build 5 blocks in the initial 5 minutes when he is energetic, but 4 blocks later when he is exhausted, the number of blocks built per minute per person is classified into different levels of productivity. As we have done several experiments to check how the ML techniques perform we have used two ways of classifying the productivity as follows. In the experiments with the whole data set the classification high (≥ 0.6), medium ($(0.2, 0.6]$) and low (≤ 0.2) was used, taking into account that the average productivity of the whole data set is 0.433 and the standard deviation is 0.182.

Since the input data varies in a large scale, therefore, pre-processing of the input data is required for this task. Both normalization and standardization have been implemented using Scikit-learn library, and by comparing the results, it is shown that standardization has a better performance on prediction of the productivity class using either deep neural network, k-nearest neighbor or logistic regression. Therefore, the input dataset is standardized using the formula below before feeding into the machine learning models using the usual standardization formula:

$$x_{standarisation} = \frac{X - \overline{X}}{\sigma X}$$

where $\bar{X}$ *and* σX are the empirical mean and standard deviation of the data.

Number	Feature
0	Low temperature
1	High temperature
2	Level of difficulty
3	Number of masons
4	Compatibility mason1
5	Compatibility mason 1&2
6	Compatibility mason 1&3
7	Compatibility mason 2&3
8	Age mason1
9	Age mason2
10	Age mason3
11	Experience mason1
12	Experience mason2
13	Experience mason3

TABLE 1 Features of the input data

4.2. Implementation

4.3. Deep Neural Network

In this case, a deep neural network model (DNN) is built in order to predict the level of productivity given the input dataset. Different network architectures have been explored.

In order to ensure a sufficient number of samples for training, validation and testing W. Liu et al. (2017), the dataset is shuffled and divided into training, validation and test. However, as the dataset is heavily imbalanced, with an 88% of the data collected just in the class of medium productivity, a sufficient amount of copies (duplication) of the data in the low and high classes where added so that in the new dataset, each class represented 1/3 of the total of the set, that is the data set is uniformly distributed.Then the dataset is divided into training, validation and test in the ratio 2400:700:711, keeping the training, validation and test part of the dataset still balanced.

A batch size of 64, which is a hyperparameter of gradient descent that controls the number of training samples to work through before the model's internal parameters are updated, is chosen. Rectified linear unit (ReLU) is chosen as it is a commonly used and well-performing activation function. In the output layer, log softmax is chosen to predict the class of the productivity level. A dropout rate of 0.2 added into the network to tackle the issue of overfitting. To train the neural network, cross entropy loss is selected as the loss function: it is a commonly used loss function in the type of classification problem considered in this paper. In this experiment, the DNN models used have a two-layer architecture (14-8-3), that is, there are 14 neurons in the input layer, 8 in the hidden layer and 3 in the output layer, and a three layer architecture (14-10-5-3). It was trained with an Adam optimizer and batch size of 64 and learning rate of 0.01. The deep learning model is trained for 150 epochs because by observing the plots of training and validation error, the training error is relatively constant after 150 epochs and the validation error starts to increase, which means 150 epochs of training is sufficient. The model with lowest validation loss is saved.

Then, the trained model is tested on the testing dataset. The best classification accuracy obtained, after probing with different architectures, is 92.7%. A confusion matrix is plotted in Figure 4 to provide an indication towards the classification results.

The confusion matrix is a performance measurement for machine learning classification problem to check the performance of a classification model on a set of test data for which the true values are known. The column represents for the ground truth of the classification and the row stands for the predicted classification results. For example, in Figure 4 , in the prediction of the 'low productivity' tasks, the DNN model correctly classified 210 of the samples, while misclassified 16 of the 'low productivity' as 'medium productivity'.

4.4. K-Nearest Neighbors

K-Nearest-Neighbour, KNN for short (Cunningham & Delany (2020)), is a simple, supervised machine learning algorithm that can be used to solve both classification and regression problems. A KNN classifier determines the class of a data point by a majority voting principle. For example, if K is set to 10,

Data Processing Operation	**Rationale**
Discard irrelevant information	Date, construction site and name of masons does no contribute to prediction
Padding with 0	Ensure input data has the same length
Label the productivity by classes	Create classification labels
Duplicate data from the classes with the lesser number of datapoints	To balance the dataset

Divide into Training, Validation and Testing	To train, validate and test
Standardisation	The input data varies in a large scale thus requires standardisation

TABLE 2 Data Preprocessing

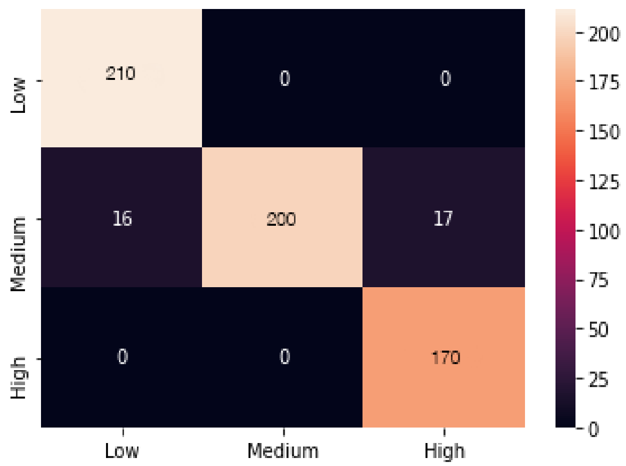

FIGURE 4 DNN - 2Layer Confusion Matrix

the classes of 10 closest points are checked and the is done according to the majority class. To determine how close the data points are from one another, Euclidean distance is one of most commonly used distance measurement.

In this case, a KNN model where K = 10 is built using the Scikit-learn library and achieved the classification accuracy of 97.5%. The value K = 100 was also tried with an accuracy of 81.4% Different values of K have been explored as shown in Table 3 , and when K = 10, the model achieves the highest accuracy. The confusion matrix is plotted in Figure 5 .

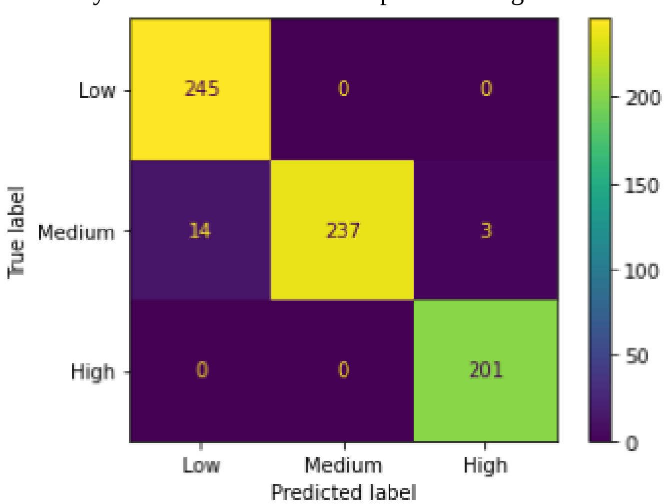

FIGURE 5 KNN Confusion Matrix, K = 10

4.5. Logistic Regression

Logistic regression (Hosmer et al. (2013)) is a statistical learning technique, categorized in supervised machine learning methods, and dedicated to classification tasks. Logistic regression uses the sigmoid function

$$f(x) = \frac{1}{1 + e^{-x}}$$

which takes any real value between zero and one.The logistic regression algorithm becomes a classification technique only when a decision threshold (default = 0.5) is brought into the picture. 'One VS Rest' strategy (Brownlee (2020)) is deployed for multi-class classification in this case.

In this case, a logistic regression model was built using the Scikit-learn library and Liblinear Fan,Chang, Hsieh, Wang, & Lin (2008) which is a good choice for small dataset, and achieved the classification accuracy of 87.2%. The confusion matrix for this classification method is plotted in Figure 6.

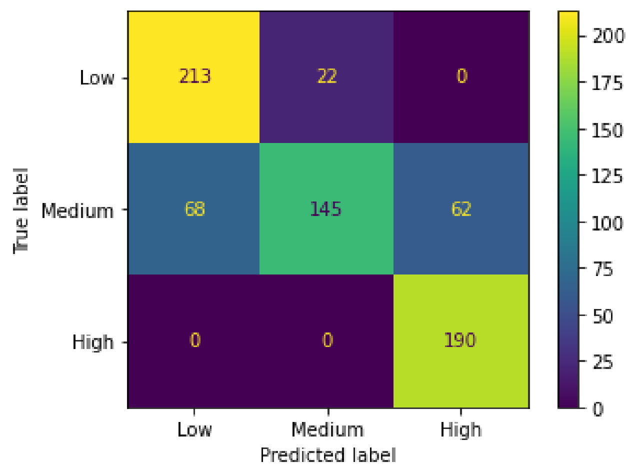

FIGURE 6 Logistic Regression Confusion Matrix

4.6. Support Vector Machine

A Support vector machine (SVM) (Hearst (1998)) is a machine learning technique to find a hyperplane in an N-dimensional space(N – the number of features) that distinctly classifies the data points. In this task, a SVM classifier with a sigmoid kernel was deployed to classify the level of productivity and the result and the accuracy obtained was 95.8%.

5. FEATURE SELECTION

To determine the factors that the greatest impact on predict- ing the productivity, feature selection has been deployed using permutation importance on both KNN and logistic regression models. For KNN model, the feature importance graph is shown in Figure 7 . From the graph above, it can be observed that the lowest and highest temperature of the day, followed by the level of difficulty of the task has the greatest impact on predicting the productivity (a fact that will be somewhat explained in Section 6). It also can be observed that other factors as the experience, number of masons in a crew, and their compatibility play a role in the predictive model.

For logistic regression model, the feature importance graph given in 8 . As we can see from the graph, the difficulty of the task still plays the most important role on predicting the class of productivity, however, the second most important factor changes to number of masons (again, an explanation might be extracted from the statistics shown in Section 6).

Using random forest classifier to reserve the five most important features and discard the rest, the accuracy of the logistic regression model decreases 78.2% and the accuracy of the KNN model increases a little bit from 97.5% to 97.7%. Nevertheless, the KNN after feature selection still remains the best performing method among all the models explored in this work, which proves that the feature selection is correctly implemented as the results almost do not change.

5.1. Performance Comparison

From Table 3 , the KNN model with K = 10 and feature extraction achieved the highest accuracy (97.7%) on predicting he level of productivity of the construction project, although he DNN model also worked quite well with a classification accuracy of 92.6%. By predicting the level of productivity of the masons, the project manager can hence make decisions on how to arrange the masons to maximum productivity.

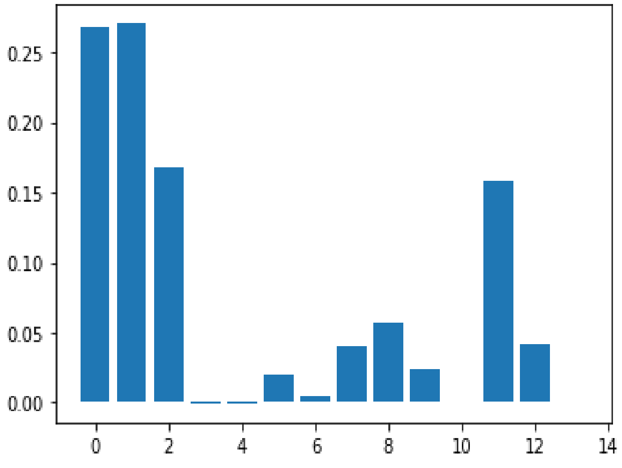

FIGURE 7 KNN Feature Importance, K = 10

changed. Indeed, by removing all compatibility features (compatibility of mason1, compatibility (mason1&mason2), compatibility (mason1&mason3), compatibility (mason2&mason3)), the necessity of the compatibility feature could be somewhat determined. The classification results on the dataset without compatibility features are shown in Table 5.

From Table 5, it can be observed that removing the

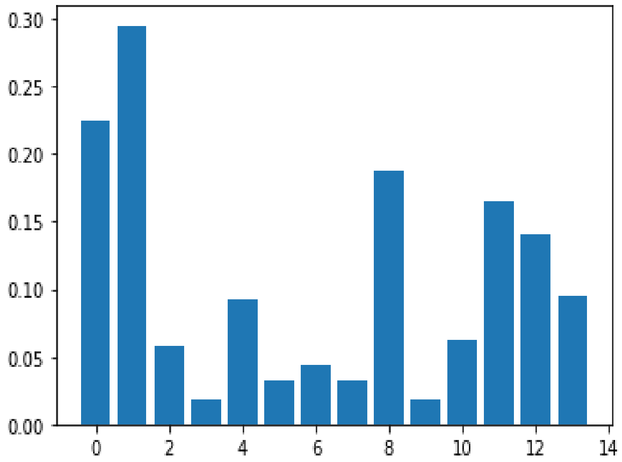

FIGURE 8 Logistic Regression Feature Importance

Task productivity was also classified into four different levels: low (strictly less than 0.2), medium low (larger or equal than 0.2 and less than 0.4), medium high (larger than or equal than 0.4 and less than 0.6), and high (larger or equal than 0.6). Here the results of experimenting with a residual neural network ResNet18 are included, as it gives better performance than the other DNN models: this hints that the use of more advanced ML models can contribute to better accuracy in the case of larger databases and finer productivity classifications (see Table 4). Residual Networks are mainly used to solve learning problems related to the vanishing gradient when using very deep neural networks with a traditional CNN structure; here it is shown that perform better than DNN when a finer classification of task productivity is introduced.

5.1.1. Removing Compatibility

Although the lowest and highest temperature, along with difficulty and crew size, seem to be the most important features when predicting the productivity of masons, the feature importance analysis shows that compatibility may play a role in predicting productivity. To try to measure the influence of compatibility as a predictive factor, it was removed to check how the accuracy of the algorithms

compatibility features from the input dataset gives mixed results regarding the accuracy of the classification. For instance, although the DNN model without compatibility has a lower accuracy, though not much, in the case of others as KNN with K = 100, there is some real improvement (about 5% better accuracy). As compatibility is a nonstandard feature used in this work, and its role in mason productivity is not well known, its relation with the average

productivity of masons will be analyzed below.

Machine Learning Model	Classification Accuracy
DNN with 2 layers	92.6%
DNN with 3 layers	88.2%
KNN (K=10)	97.5%
KNN with Feature Extraction (K=10)	97.7%
KNN (K=100)	81.4%
KNN with Feature Extraction (K=100)	92.8%
Logistic Regression	85.2%
Logistic Regression with Feature Extraction	78.2%
Sigmoid SVM	95.8%

TABLE 3 Classification Results

5.1.2. Fixing difficulty: Walls of difficulty 1

Since, as it will be shown below, the task productivity of masons when working on easy walls is higher that when work- ing on more difficult walls, and difficulty is one of the most prominent features used for prediction of task productivity, experiments were performed to test how ML algorithms work in the case when we considered task productivity for a walls of difficulty 1. Productivity classes were changed as follows: high (>0.6), medium ((0.4, 0.6]) and low (::; 0.4); this because for walls of difficulty 1 average productivity is 0.515, with standard deviation 0.146. In this case, the algorithms were trained on a balanced dataset, so that each class had about 1/3 of the data. The dataset is shuffled and the divided into training (1500 data samples), validation (400 data samples) and testing (290 data samples) separately. The architectures considered for the DNN model are (14-8-3) and (14-10-5-3).

The experiments reveal a poorer performance, although it is still moderately good, especially in the KNN model with $K = 10$ with feature extraction (accuracy: 75.1%) and the SVM with sigmoid kernel (77.7%). The results of the experiments are shown in Table 6 .

6. STATISTICAL ANALYSIS

The ML techniques used in this paper are not only useful in building predictive models, as they can be also be used to determine what factors exert the largest influence when try- ing to make predictions. Although ML techniques are often seen as a black box that produces some desired results, their performance in turn can be used to look more carefully into certain aspects of the data and analyze them to make some interpretations.

External factors such as temperature seem to play the largest role in task productivity, which is easy to explain for physically intensive labors such as masonry. As this is a difficult factor to control by a superintendent, the focus in the analysis that follows will be on those factors that in principle can be controlled when forming a crew. Some of the factors are shown to be more important by the feature selection analysis, and some of them, as compatibility, seem to be of minor importance, since when it is removed from the data, the accuracy of the ML methods is better, or not much worse than when it is included. However, as compatibility is used by superintendents to form crews, it might be important and interesting to analyze its relation with task productivity. In the course of this analysis, it was also found that crew size has a statistically significant relation with task productivity, which means that investigating which factors affect team work needs to be more carefully considered. Furthermore, this investigation must take into account what type of work is being performed by the teams (for instance labor intensive work such as masonry, versus more creative work such as designing a new product).

To perform this analysis, the wall difficulty variable has been fixed. This is done because difficulty is, along with the range of temperature of the day (and in some cases crew size -see below), the most important feature in predicting productivity, a fact that is revealed by the differences in average productivity for the different types of walls, being average productivity for walls of difficulty 1 clearly higher than for more difficult walls. Note that the average productivity for the three types of walls is statistically different (see Table 7), which suggests that the three level categorization of the difficulty of the walls seems to accurately represent masonry wall types

Machine Learning Model	Classification Accuracy
DNN with 2 layers	67.9%
DNN with 3 layers	63.6%
KNN (K=10)	70.7%
KNN (K=100)	61.5%
Logistic Regression	61.3%
Sigmoid SVM	73.7%
ResNet18	72.9%

TABLE 4 Using classification into low, medium-low, medium-high, and high.

Machine Learning Model	Classification Accuracy
DNN with 2 layers	91.9%
DNN with 3 layers	89.8%
KNN (K=10)	97.4%
KNN with Feature Extraction (K=10)	98.0%
KNN (K=100)	81.5%
KNN with Feature Extraction (K=100)	97.7%
Logistic Regression	85.4%
Logistic Regression with Feature Extraction	79.1%
Sigmoid SVM	96.0%

TABLE 5 Classification Results without Compatibility

Note that the feature selection analysis compatibility appears to have some moderate influence in the prediction of productivity (being this a novel subjective feature considered in this paper), it might be of interest to give a short statistical perspective on how compatibility affects productivity. In this case, the analysis shall focus on crews of two masons as the compatibility measure is easier to handle in this case. An interesting finding is that for the easier type of wall (difficulty = 1) crews with the higher compatibility score (0.9) have higher average productivity than those with lower compatibility score (0.5), and that this difference is statistically significant. Since productivity in the case of the data collected does not seem to follow a normal distribution (as checked using a Shapiro-Wilk test), a Mann-Whitney-Wilcoxon U test under the alternative hypothesis that the mean productivity of the group with compatibility 0.9 (average productivity = 0.540) is higher than the mean productivity of the group with compatibility 0.5 (average productivity= 0.526), shows that this difference is statistically significant (U = 62671, p-value=0.019).

However, when more difficult walls are considered, this tendency is reversed. That is, the average productivity for walls of difficulty 2 and 3 is higher for crews with lower compatibility. In the case of walls of difficulty 2, the average productivity of the crews with compatibility 0.5 is significantly higher than that of the crews with compatibility 0.9 (0.335 vs. 0.273), and this difference is statistically significant (U = 10939, p-value= 4.84×10^{-6}). This shows that the intuition behind a superintendents' choice of crews, that is, that a more compatible crew is more productive, is not right once the task at hand becomes difficult enough.

Machine Learning Model	Classification Accuracy
DNN with 2 layers	69.3%
DNN with 3 layers	65.8%
KNN (K=10)	74.2%
KNN with Feature Extraction (K=10)	75.1%
KNN (K=100)	66.1%
KNN with Feature Extraction (K=100)	76.9%
Logistic Regression	64.0%
Logistic Regression with Feature Extraction	45.9%
Sigmoid SVM	77.7%

TABLE 6 Classification Results, wall difficulty=1

Difficulty	Avg. productivity	Std deviation
1	0.515	0.1469128
2	0.332	0.1507142
3	0.212	0.1069510

TABLE 7 Average productivity by wall difficulty

Number of masons	Avg productivity	Std deviation
1	0.5449495	0.1686874
2	0.5086207	0.1261019
3	0.4388889	0.1744629

TABLE 8 Average productivity by number of masons in a crew, wall difficulty=1

The fact that compatibility, in the case of more difficult walls, is not related the right way with productivity might be explained by the fact that in these cases mason experience plays a more fundamental role, this dictated by the fact that the construction of these walls requires more technique and finesse. In fact, in the case of walls of difficulty 2, we examined the case of crews of two masons. We define the experience of the crew as the average of the experience of its members, and we found that the average experience of the crews with compatibility 0.5 is 24.913 years, whereas for the crews with compatibility 0.9 the average experiences is 22.750 years, and as revealed by a Mann-Whitney-Wilcoxon U test, the difference in average experience in the two different types of crews is statistically significant

(U = 12772, p-value = 8.72×10^{-15}).

Of course a question arises here: are just 2.16 more years of experience (in average) enough to improve performance (in average), for workers with more than 20 years of experience? Perhaps this has to do with the fact that masons have much less time of experience working with walls of higher difficulty than with walls of difficulty 1, and thus, in two years the cumulative experience of working in walls of high difficulty has more weight (using the data collected, almost 62 % of the walls have difficulty 1; however, this is just speculation, and there is not enough data to conclude).

There are other factors such as the age of the masons that might also affect task productivity. In this case a possible explanation is due to the fact that humans reach a physical peak at a certain age, and masonry is a physically intensive labor. In fact, it was found that for crews of two masons, the average age of the crew is slightly anticorrelated with productivity (= −0.0747, -value 0.043), and that the maximum age of a member of the crew anticorrelates with average productivity (-value 0.003); minimum age of the members of the crew seems to be independent of task productivity. So perhaps, as an application of this to enhance productivity, pairing a young mason with a more experienced one is better than always pairing experienced masons.

Feature selection analysis also shows that the number of masons in a crew have a role in the productivity. Again, we analyze the walls of difficulty 1, and we found the follow- ing. In the case of crews of 1 and 2 masons respectively, a Mann-Whitney-Wilcoxon test with alternative hypothesis that the average productivity is larger in the case of a crew with 1 mason than when having a team of 2, has a p-value of

1.43×10 . In the case of comparing crews of 2 and 3 masons the p-value of 0.002 obtained. This of course opens up the question of which factors are making larger teams less productive (at least in average), and what can be taken into account to make crews more satisfied with their work and to increase their productivity. Hence, having a better measure of how "compatible" a team is, or finding the right psychological construct that helps predicting better satisfaction and performance becomes an important task. Furthermore, although the data collected is for construction crews, this research might be relevant to other fields where team work is a requirement, but this of course requires collecting the relevant data.

7. CONCLUSIONS

ML techniques were used to predicting task productivity levels in construction crews, using as input some obvious and not so obvious choices of data that might influence it. The experiments show moderate (Logistic Regression and KNN) to very good (DNN and SVM) accuracy when using ML models to predict productivity with the data collected. This confirms the appreciation that ML methods can be use as decision making tools in managing crews in building construction, and its use deserves to be spreaded.

Also, feature selection methods were used to find the most important factors influencing the productivity of a crew from the ones selected in the collection of data. An interesting finding is that a subjective measure (compatibility) can be affecting the productivity of the masons in a crew. Indeed, the ML techniques used in this paper reveal that it plays a role in the predictive power of the different models employed. But beware that the fact that compatibility might play a role in the prediction of productivity levels is not related with either a positive or negative correlation with productivity when the whole data set is considered. That compatibility was found to be a moderately important feature, plus the fact that difficulty (excluding the range of temperature in a working day) was the most important feature lead us to find, using classical statistical analysis, that, whereas in the case of difficulty 1 (easy walls) productivity seems to correlate positively with average productivity in the case of more difficult walls is the opposite. This is, perhaps, a surprising discovery, taking into account that superintendents in the field use compatibility thinking that it is always the case that it has a positive correlation with productivity. Perhaps, this is due to the fact that in easy walls, masons have to constantly coordinate with other masons when to raise the line and this type of walls require more interactions. In the case of normal and difficult walls, masons spend more time measuring openings and making cuts and may have less interactions with other masons, so the technical ability may be the more relevant than compatibility. It might be appropriate to add that this finding is a bit in contrast with the work reported in Florez et al. (2020), where a different measure of compatibility, based on psycho- logical tests was defined, and was shown to correlate with task productivity. This immediately leads to the question of how these two different types of compatibility are related, and prompts to search other important psychological factors, such as motivation and commitment, that might affect the productivity of a team. This also subsumes an important critique that can be stated about this work: perhaps measuring compatibility requires more than the subjective opinion of a superintendent, even though this measurement is based on the longtime and probably careful observations of the masons she/he has man- aged. Also, crew size influences task productivity: in some cases, larger crews have lower average productivity; this topic needs a more in depth research.

However, the fact that the compatibility measure used in this paper was greatly subjective, yet seems to be useful, should prompt research studies where more objective ways of measuring "how well workers relate to each other" are investigated (see Florez et al. (2020)). Furthermore, future research should look into collecting more data, and in this case, given a dataset large enough, using more powerful and more data intensive ML classification algorithms is fully justified such as Enhanced Probabilistic Neural Networks (Ahmadlou & Adeli (2010)), Finite Element Machine for Fast Learning (Pereira, Piteri, Souza, Papa, & Adeli (2019)), and Dynamic Ensamble Learning Algorithm (Alam, Siddique, & Adeli (2019)).

Finally, if further research validates the findings of this work, as a practical application, the results of this study have the potential to benefit masonry construction companies in improving task productivity with strategies that can be controlled by the superintendent. For instance, in easy walls, assign crews with higher compatibility of personality; in difficult walls, assign crews with masons that are more experienced. Keep the crew size to a minimum as small size crews seem to be more productive than large crews. Also, the data collected for a workday, including environmental data and data related to the masons' experience, age and compatibility, can be used with the trained ML models used in this work to predict the most probable task productivity for a crew on that given workday.

8. ACKNOWLEDGEMENTS

The authors would like to thank University College London's Bartlett Research Grants Scheme 2020-2021 for funding this study. The third author wants to thank his home institution, Universidad de los Andes, for the support given during a semester of academic sabbatical (STAI).

References

Adeli, H. (2001). Neural networks in civil engineering: 1989-2000. *Computer-Aided Civil and Infrastructure Engineering*, *16*(2), 126–142.

Ahmadlou, M., & Adeli, H. (2010, July). Enhanced probabilistic neural network with local decision circles: A robustclassifier. *Integrated Computer-Aided Engineering*, *17*(3),197–210.

Alaloul, W. S., Liew, M. S., Zawawi, N. A. W., Mohammed, B. S., & Adamu, M. (2018). An artificial neural networks (ANN) model for evaluating construction project performance based on coordination factors. *Cogent Engineering*, *5*(1), 1507657.

Alam, K. M. R., Siddique, N., & Adeli, H. (2019). A dynamic ensemble learning algorithm for neural networks. *Neural Computing and Applications*, *32*(12), 8675–8690.

As, I., Pal, S., & Basu, P. (2018). Artificial intelligence in architecture: Generating conceptual design via deep learning. *International Journal of Architectural Computing*, *16*(4), 306–327.

Awada, M., Srour, F. J., & Srour, I. M. (2021). Data-driven machine learning approach to integrate field submittals in project scheduling. *Journal of Management in Engineering*, *37*(1), 04020104.

Badawy, M., Hussein, A., Elseufy, S. M., & Alnaas, K. (2019). How to predict the rebar labours' production rate by using ANN model? *International Journal of Construction Management*, *21*(4), 427–438.

Bai, S., Li, M., Kong, R., Han, S., Li, H., & Qin, L. (2019). Data mining approach to construction productivity prediction for cutter suction dredgers. *Automation in Construction*, *105*, 102833.

Barlow, H. (1989). Unsupervised learning. *Neural computation*, *1*(3), 295–311.

Bilal, M., & Oyedele, L. O. (2020). Guidelines for applied machine learning in construction industry—a case of profit margins estimation. *Advanced Engineering Informatics*, *43*, 101013.

Boussabaine, A. (2001). Neurofuzzy modelling of construction projects duration i: principles. *Engineering, Construction and Architectural Management*, *8*(2), 104–113.

Boussabaine, A., & Kirkham, R. (2008). Artificial neural network modelling techniques for applied civil and construction engineering research. In A. Knight & L. Ruddock (Eds.), *Advanced research methods in the built environ- ment* (p. 155-170). Hoboken, NJ: Wiley-Blackwel.

Brownlee, J. (2020). One-vs-rest and one-vs-one for multi-class classification.

Caruana, R., & Niculescu-Mizil, A. (2006). An empirical comparison of supervised learning algorithms. In *Proceedings of the 23rd international conference on machine learning* (pp. 161–168).

Celebi, M. E., & Aydin, K. (Eds.). (2016). *Unsupervised learning algorithms*. Springer International Publishing.

Chao, L.-C., & Skibniewski, M. J. (1994). Estimating construction productivity: Neural-network-based approach.*Journal of Computing in Civil Engineering*, *8*(2), 234– 251.

Cheng, C.-S., Behzadan, A. H., & Noshadravan, A. (2021). Deep learning for post-hurricane aerial damage assessment of buildings. *Computer-Aided Civil and Infrastructure Engineering*, *36*(6), 695–710.

Crawford, P., & Vogl, B. (2006). Measuring productivity in the construction industry. *Building Research & Information*, *34*(3), 208–219.

Cunningham, P., & Delany, S. J. (2020). *k-nearest neighbour classifiers: 2nd edition (with python examples).*

Dissanayake, M., Fayek, A. R., Russell, A. D., & Pedrycz, W. (2005). A hybrid neural network for predicting construction labour productivity. In *Computing in civil engineering (2005).* American Society of Civil Engineers.

Dobgegah, R., Owusu-Manu, D.-G., & Omoteso, K. (2011). A principal component analysis of project management construction industry competencies for the ghanaian. *Construction Economics and Building*, *11*(1), 26–40.

Dong, S., Yu, T., Farahmand, H., & Mostafavi, A. (2020).A hybrid deep learning model for predictive flood warn- ing and situation awareness using channel network sensors data. *Computer-Aided Civil and Infrastructure Engineering*, *36*(4), 402–420.

Durdyev, S., & Mbachu, J. (2011). On-site labour productivity of New Zealand construction industry: Key constraints and improvement measures. *Construction Economics and Building*, *11*(3), 18–33.

Ebrahimi, S., Fayek, A. R., & Sumati, V. (2021). Hybrid artificial intelligence HFS-RF-PSO model for construction labor productivity prediction and optimization. *Algorithms*, *14*(7), 214.

El-Gohary, K. M., & Aziz, R. F. (2014). Factors influencing construction labor productivity in Egypt. *Journal of Management in Engineering*, *30*(1), 1–9.

El-Gohary, K. M., Aziz, R. F., & Abdel-Khalek, H. A. (2017). Engineering approach using ANN to improve and predict construction labor productivity under different influences. *Journal of Construction Engineering and Management*, *143*(8), 04017045.

Enshassi, A., Mohamed, S., Mustafa, Z. A., & Mayer, P. E. (2007). Factors affecting labour productivity in building projects in the gaza strip. , *13*(4).

Fan, R.-E., Chang, K.-W., Hsieh, C.-J., Wang, X.-R., & Lin, C.-J. (2008). Liblinear: A library for large linear classification. *J. Mach. Learn. Res.*, *9*, 1871–1874.

Fang, Q., Li, H., Luo, X., Ding, L., Luo, H., Rose, T. M., & An, W. (2018). Detecting non-hardhat-use by a deep learning method from far-field surveillance videos. *Automation in Construction*, *85*, 1–9.

Florez, L. (2015). *Decision support system for masonry labor planning and allocation considering productivity and social sustainability* (Unpublished doctoral dissertation). Georgia Institute of Technology.

Florez, L., Armstrong, P., & Cortissoz, J. C. (2020). Does compatibility of personality affect productivity? exploratory study with construction crews. *Journal of Management in Engineering*, *36*(5), 04020049.

Franco, G. D., & Santurro, M. (2020). Machine learning, artificial neural networks and social research. *Quality & Quantity*, *55*(3), 1007–1025.

Gao, X., Shi, M., Song, X., Zhang, C., & Zhang, H. (2019). Recurrent neural networks for real-time prediction of TBM operating parameters. *Automation in Construction*, *98*, 225–235.

Golnaraghi, S., Zangenehmadar, Z., Moselhi, O., & Alkass, S. (2019). Application of artificial neural network(s) in predicting formwork labour productivity. *Advances in Civil Engineering*, *2019*, 1–11.

Hafez, S. M. (2014). Critical factors affecting construction labor productivity in Egypt. *American Journal of Civil Engineering*, *2*(2), 35.

Hamza, M., Shahid, S., Hainin, M. R. B., & Nashwan, M. S. (2019). Construction labour productivity: review of factors identified. *International Journal of Construction Management*, 1–13.

Hastie, T., Tibshirani, R., & Friedman, J. (2008a). Overview of supervised learning. In *The elements of statistical learning* (pp. 9–41). Springer New York.

Hastie, T., Tibshirani, R., & Friedman, J. (2008b). Unsupervised learning. In *The elements of statistical learning* (pp.485–585). Springer New York.

Hearst, M. A. (1998). Support vector machines. *IEEE Intelligent Systems*, *13*(4), 18–28.

Heravi, G., & Eslamdoost, E. (2015). Applying artificial neural networks for measuring and predicting construction-labor productivity. *Journal of Construction Engineering and Management*, *141*(10), 04015032.

Hiyassat, M. A., Hiyari, M. A., & Sweis, G. J. (2016). Factors affecting construction labour productivity: a case study of Jordan. *International Journal of Construction Management*, *16*(2), 138–149.

Hosmer, D., Lemeshow, S., & Sturdivant, R. (2013). *Applied logistic regression*. Wiley.

Jarkas, A. M., & Bitar, C. G. (2012). Factors affecting construction labor productivity in Kuwait. *Journal of Construction Engineering and Management*, *138*(7), 811–820. Jassmi, H. A., Ahmed, S., Philip, B., Mughairbi, F. A., &Ahmad, M. A. (2019). E-happiness physiological indicators of construction workers productivity: A machine learning approach. *Journal of Asian Architecture and Building Engineering*, *18*(6), 517–526.

Karim, A., & Adeli, H. (1999). OO information model for construction project management. *Journal of Construction Engineering and Management*, *125*(5), 361–367.

Lee, S., & Chang, L.-M. (2005). Digital image process- ing methods for assessing bridge painting rust defects and their limitations. In *Computing in civil engineering (2005).* American Society of Civil Engineers.

Liu, W., Wang, Z., Liu, X., Zeng, N., Liu, Y., & Alsaadi, F. E. (2017). A survey of deep neural network architectures and their applications. *Neurocomputing*, *234*(December 2016), 11–26. doi:

Liu, Z., Sadiq, R., Rajani, B., & Najjaran, H. (2010). Exploring the relationship between soil properties and deterioration of metallic pipes using predictive data mining methods. *Journal of Computing in Civil Engineering*, *24*(3), 289–301.

Lowe, J. G. (1987). The measurement of productivity in the construction industry. *Construction Management and Economics*, *5*(2), 101–113.

Luo, H., & Paal, S. G. (2020). Reducing the effect of sample bias for small data sets with double-weighted sup-port vector transfer regression. *Computer-Aided Civil and Infrastructure Engineering*, *36*(3), 248–263.

Maeda, H., Kashiyama, T., Sekimoto, Y., Seto, T., & Omata, H. (2020). Generative adversarial network for road dam-age detection. *Computer-Aided Civil and Infrastructure Engineering*, *36*(1), 47–60.

Mirahadi, F., & Zayed, T. (2016). Simulation-based construction productivity forecast using neural-network-driven fuzzy reasoning. *Automation in Construction*, *65*, 102–115.

Momade, M. H., Shahid, S., bin Hainin, M. R., Nashwan, M. S., & Umar, A. T. (2020). Modelling labour productivity using SVM and RF: a comparative study on classifiers performance. *International Journal of Construction Management*, 1–11.

Moselhi, O., & Khan, Z. (2012). Significance ranking of parameters impacting construction labour productivity. *Construction Innovation*, *12*(3), 272–296.

Nasirzadeh, F., Kabir, H. D., Akbari, M., Khosravi, A., Nahavandi, S., & Carmichael, D. G. (2020). ANN-based pre- diction intervals to forecast labour productivity. *Engineer- ing, Construction and Architectural Management*, *27*(9), 2335–2351.

Nasirzadeh, F., & Nojedehi, P. (2013). Dynamic modeling of labor productivity in construction projects. *International Journal of Project Management*, *31*(6), 903–911.

Ng, S. T., & Tang, Z. (2010). Labour-intensive construction sub-contractors: Their critical success factors. *International Journal of Project Management*, *28*(7), 732–740.

Okazaki, Y., Okazaki, S., Asamoto, S., & jo Chun, P. (2020). Applicability of machine learning to a crack model in concrete bridges. *Computer-Aided Civil and Infrastructure Engineering*, *35*(8), 775–792.

Oral, E. L., Oral, M., & Andaç, M. (2016). Construction crew productivity prediction: Application of two novel methods. *International Journal of Civil Engineering*, *14*(3), 181–186.

Pereira, D. R., Piteri, M. A., Souza, A. N., Papa, J. P., & Adeli, H. (2019). FEMa: a finite element machine for fast learning. *Neural Computing and Applications*, *32*(10), 6393–6404.

Poh, C. Q., Ubeynarayana, C. U., & Goh, Y. M. (2018). Safety leading indicators for construction sites: A machine learning approach. *Automation in Construction*, *93*, 375–386.

Portas, J., & AbouRizk, S. (1997). Neural network model for estimating construction productivity. *Journal of Construction Engineering and Management*, *123*(4), 399–410.

Quionero-Candela, J., Sugiyama, M., Schwaighofer, A., & Lawrence, N. (2019). *Dataset shift in machine learning*. The MIT Press.

Reich, Y. (1997). Machine learning techniques for civil engineering problems. *Computer-Aided Civil and Infrastructure Engineering*, *12*(4), 295–310.

Rodrigues, E., Sousa-Rodrigues, D., de Sampayo, M. T., Gaspar, A. R., Gomes, Á., & Antunes, C. H. (2017). Clustering of architectural floor plans: A comparison of shape representations. *Automation in Construction*, *80*, 48–65.

Shehata, M. E., & El-Gohary, K. M. (2011). Towards improving construction labor productivity and projects' performance. *Alexandria Engineering Journal*, *50*(4), 321–330.

Shu-quan, L., Xin-li, Z., Zhi-qiang, L., Li-xia, F., Lan, M., & Qiu-li, G. (2006). Dynamic monitoring on construction safety based on support vector machine. In *2006 international conference on management science and engineering.* IEEE.

Teizer, J., & Vela, P. (2009). Personnel tracking on construction sites using video cameras. *Advanced EngineeringInformatics*, *23*(4), 452–462.

Thomas, A. V., & Sudhakumar, J. (2014). Factors influencing construction labour productivity: an Indian case study. *Journal of Construction in Developing Countries*, *19*, 53–68.

Thomas, H. R., Guevara, J. M., & Gustenhoven, C. T. (1984). Improving productivity estimates by work sampling. *Journal of Construction Engineering and Management*, *110*(2),178–188.

Tomar, A., & Burton, H. V. (2021). Active learning method for risk assessment of distributed infrastructure systems. *Computer-Aided Civil and Infrastructure Engineering*, *36*(4), 438–452.

Tsehayae, A. A., & Fayek, A. R. (2016). System model for analysing construction labour productivity. *Construction Innovation*, *16*(2), 203–228.

Valikhani, A., Jahromi, A. J., Pouyanfar, S., Mantawy, I. M., & Azizinamini, A. (2020). Machine learning and image processing approaches for estimating concrete surface rough- ness using basic cameras. *Computer-Aided Civil and Infrastructure Engineering*, *36*(2), 213–226.

Wang, G. (2008). A survey on training algorithms for support vector machine classifiers. In *2008 fourth international conference on networked computing and advanced information management* (Vol. 1, p. 123-128). doi:

Wang, J., & Ashuri, B. (2016). Predicting ENR's construction cost index using the modified k nearest neighbors (KNN) algorithm. In *Construction research congress 2016.* American Society of Civil Engineers.

Wilson, O. D., Sharpe, K., & Kenley, R. (1987). Estimates given and tenders received: a comparison. *Construction Management and Economics*, *5*(3), 211–226.

Xu, Y., Zhou, Y., Sekula, P., & Ding, L. (2021). Machine learning in construction: From shallow to deep learning. *Developments in the Built Environment*, *6*, 100045.

Yang, M.-D., & Su, T.-C. (2008). Automated diagnosis of sewer pipe defects based on machine learning approaches. *Expert Systems with Applications*, *35*(3), 1327–1337.

Yu, Y., Li, H., Yang, X., Kong, L., Luo, X., & Wong, A. Y. (2019). An automatic and non-invasive physical fatigue assessment method for construction workers. *Automation in Construction*, *103*, 1–12.

Zhang, X., Waller, S. T., & Jiang, P. (2019). An ensemble machine learning-based modeling framework for analysis of traffic crash frequency. *Computer-Aided Civil and Infrastructure Engineering*, *35*(3), 258–276.