MDPI

Article

Forecasting Future Research Trends in the Construction Engineering and Management Domain Using Machine Learning and Social Network Analysis

Gasser G. Ali [1],*, Islam H. El-adaway [2], Muaz O. Ahmed [2], Radwa Eissa [2], Mohamad Abdul Nabi [2], Tamima Elbashbishy [2] and Ramy Khalef [2]

[1] Department of Civil Engineering, The University of Texas Rio Grande Valley, Edinburg, TX 78539, USA
[2] Department of Civil, Architectural, and Environmental Engineering, Missouri University of Science and Technology, Rolla, MO 65409, USA; eladaway@mst.edu (I.H.E.-a.); muaz.ahmed@mst.edu (M.O.A.); reissa@mst.edu (R.E.); mah59@mst.edu (M.A.N.); telbashbishy@mst.edu (T.E.); ramykhalef@mst.edu (R.K.)
* Correspondence: gasser.ali@utrgv.edu

Abstract: Construction Engineering and Management (CEM) is a broad domain with publications covering interrelated subdisciplines and considered a key source of knowledge sharing. Previous studies used scientometric methods to assess the current impact of CEM publications; however, there is a need to predict future citations of CEM publications to identify the expected high-impact trends in the future and guide new research efforts. To tackle this gap in the literature, the authors conducted a study using Machine Learning (ML) algorithms and Social Network Analysis (SNA) to predict CEM-related citation metrics. Using a dataset of 93,868 publications, the authors trained and tested two machine learning classification algorithms: Random Forest and XGBoost. Validation of the RF and XGBoost resulted in a balanced accuracy of 79.1% and 79.5%, respectively. Accordingly, XGBoost was selected. Testing of the XGBoost model revealed a balanced accuracy of 80.71%. Using SNA, it was found that while the top CEM subdisciplines in terms of the number of predicted impactful papers are "Project planning and design", "Organizational issues", and "Information technologies, robotics, and automation"; the lowest was "Legal and contractual issues". This paper contributes to the body of knowledge by studying the citation level, strength, and interconnectivity between CEM subdisciplines as well as identifying areas more likely to result in highly cited publications.

Keywords: construction management; construction engineering; citations; machine learning

1. Introduction

Construction Engineering and Management (CEM) is a broad domain within the extent of the Architecture, Engineering, and Construction (AEC) industry. In general, CEM spans multiple construction-related activities, issues, methods, and human factors such as cost and schedule estimations, quality control, constructability, sustainability, prefabrication, etc. Despite the broad scope of the CEM domain, combined with the practical nature of its facets, it is considered a relatively new domain within the civil engineering field [1]. Consequently, this has led to a significant amount of CEM-related research activities targeting knowledge expansions and development [2,3]. Peer-reviewed papers are considered the main source for knowledge sharing and diffusion among researchers in the academic community as well as practitioners in the AEC industry. In the academic community, decisions related to rewards, funding, hiring, and promotion have been extensively linked to the publications, their quality, as well as their impact [4,5]. With the evolving patterns and numbers of publications in the CEM domain, researchers found themselves more pressurized explicitly or implicitly to raise their metrics by being more productive and publishing more impactful and cited research work [6]. This phenomenon in the academic community is referred to as the "publish or perish" paradigm [7]. To this end, the increasing number of publications

(i.e., paper inflation) makes it hard for researchers to keep up with the literature and identify the research efforts that have made the most significant impact on the body of knowledge.

In that regard, various metrics have been utilized to assess the significance and impact of publications. The impact of a publication may be quantified based on the impact factor of its affiliated journal. However, this method is subject to numerous flaws, including that the journal impact factor does not provide insight into a specific publication, but it is a method for evaluation of the journal as a whole [4]. Another method used for the assessment of such impact is the cumulative citation count. The citation count was first proposed as a measure of how frequently subsequent publications tend to cite a specific publication [8]. The main logic behind the utilization of citation count is that an impactful publication will obtain a high citation count, which implies the outstanding reach and uptake rate of a publication and its influence on the advancement of knowledge [9]. Citation count is considered the most commonly accepted and used method for evaluating the impact of research articles [10,11]. However, it should also be noted that relying on citation counts solely is not a good representation of the quality of research and its contributions. In fact, award-winning papers may have citation counts that are lower than the average citation count [12].

Various previous studies have conducted scientometric research by utilizing the citation count in assessing the current impact and significance of publications in several topics and fields including computer science [13] and financial management [14], among others. Furthermore, researchers have used machine learning methods with scientometric research. For example, Weis and Jacobson [15] developed a machine learning framework, named DELPHI, which predicts whether a research work is likely to be impactful. The developed DELPHI framework is based on analyzing specific relationships between various features in a dataset related to the biotechnology literature represented by research papers published between 1980 and 2019 in 42 biotechnology-related journals.

Predicting the future impact of research work can help the scientific community in many directions such as the following: (1) researchers can better understand current trends and pursue promising directions, (2) funders can direct requests for proposal towards needed research areas, and (3) publishers, journal editors, and conference organizers can select trending research themes. Although there is research that has analyzed current research trends in CEM [2,12,16], there is no prior research that has attempted forecasting the future impact of publications in CEM, although this was attempted for other disciplines. Predicting the future impact of recent publications can reveal expected research trends and opportunities to guide new research efforts. This paper addresses this knowledge gap within the CEM domain.

2. Goal and Objectives

The goal of this paper is to explore the future impact of various topics in the CEM domain. Specifically, the research question of this paper is: what current CEM topics are expected to be highly impactful in the future as measured by their citation counts? This is achieved by (1) developing a predictive model using machine learning that can predict future citation counts of publications in civil engineering in general, (2) applying the model to research publications in the CEM domain, and (3) exploring the topics associated with the subset of impactful CEM papers, using Social Network Analysis (SNA), to answer the research question of this paper, as detailed in Section 5.3. Determining future impactful research trends can guide new research efforts and stakeholders towards future research needs and opportunities.

3. Background

Scientometric research is the field concerned with measuring and analyzing the scientific literature. It was introduced to assist in overcoming subjectivity issues in literature reviews [17]. Several previous studies have conducted scientometric research and analysis on the CEM domain by (1) conducting citation-metric-based studies (e.g., h-index

and citation counts) for specific sets of publications and/or authors over a defined time interval or (2) statistically analyzing publication datasets over a specific time span [12]. For example, concerning publication metrics and trends, Pietroforte and Aboulezz [18] studied construction-related publications and the trends of citations in the ASCE *Journal of Management in Engineering* for the period between 1985 and 2002. The authors concluded that the engineering management discipline has seen increased contributions related to organizational change, cultural issues, corporate strategies, and programs, as well as other project management topics such as quality planning and alternative project delivery systems. Jin et al. [2] reviewed published articles in the ASCE *Journal of Construction Engineering and Management* for the period between 2000 and 2018 to capture the latest research topics in the CEM domain. The findings highlighted trending topics such as project performance indicators; information and communication technologies including Building Information Modeling (BIM); and quantitative methods for CEM. Moreover, some researchers focused on analyzing publication metrics and trends in some specific CEM subdisciplines, such as construction labor productivity [19]; planning and scheduling [20]; building information modeling (BIM) [16]; and artificial intelligence adoption [21], among others. In addition, some studies focused on studying the quality of publications as well as impactful factors. For example, Bröchner and Björk [22] explored multiple journals in the construction management research area by conducting a survey for authors to analyze their choice of journals in relation with quality and service perception. El-adaway et al. [12] investigated the variables influencing citation metrics for publications in the extensive domain of civil engineering with a focus on the CEM. It was concluded that various factors, such as research topic trendiness, as well as other features associated with coauthors and their research output, collectively impact the citation metrics for authors as well as papers.

Despite the plethora of previous scientometric research on the CEM domain, there is no previous research that investigates forecasting the future impact of CEM publications and identifying impactful CEM research trends in the future. In that regard from other domains, Weihs and Etzioni [23] utilized machine learning regression techniques to forecast the future impact of authors, using the h-index, and of papers, using citation counts, within the computer science domain. Weis and Jacobson [15] developed a machine learning model, as previously highlighted, to forecast the future impact of biotechnology-related publications.

This paper covers this gap within the CEM body of knowledge by tackling the prediction of the future impact of CEM-related publications and identification of the expected high-impact research trends in the future. It is imperative to note that this paper differs from the previously highlighted papers related to impact prediction [15,23] by (1) focusing on the CEM domain specifically; (2) predicting the expected high-impact research trends and related topics in the future within the CEM domain rather than focusing only on the future impact of publications; (3) using different approaches in the utilization of machine learning techniques as well as different datasets; and (4) utilizing other techniques beside machine learning in achieving the research goal and associated objectives, as will be discussed under the Methodology Section.

4. Methodology

To achieve the aforementioned research goal and objectives, the authors conducted an exploratory and predictive analysis of citation metrics in the CEM domain using computational machine learning algorithms and SNA. Figure 1 provides an overview of the adopted methodology, which is further elaborated in the following subsections.
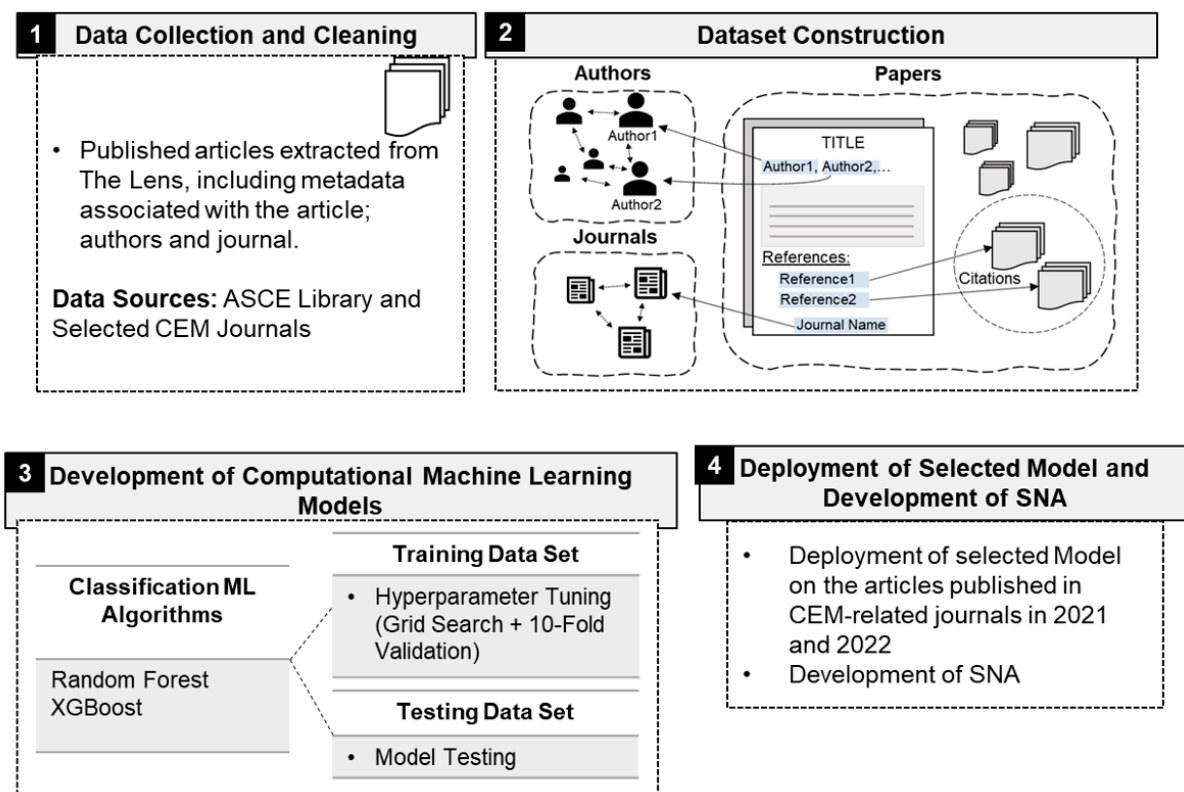
Figure 1. Research methodology.

4.1. Data Collection and Cleaning

The studied dataset was retrieved from The Lens (lens.org) [24], which is a platform that includes compiled scholarly metadata from Microsoft Academic, PubMed, CORE, and Crossref. The Lens is an extension of work started by Cambia and Queensland University of Technology in 1998 to combine scholarly and patent content sets to enable the discovery, analysis, and sharing of knowledge [25]. Retrieved data for this research included all peer reviewed journal articles published from 1985 up to 15 March 2022, in the ASCE library, in addition to the CEM journals, as shown in Table 1. The associated metadata for the articles includes the following: article ID, article title, author names, year published, citation count, and listed references. The dataset was cleaned by removing all records with missing information such as year published, authors, title, and journal name. Also, all papers that are editorials, book reviews, discussions, and conference papers were removed. After cleaning, the dataset included 93,868 articles. It should be noted that (1) all the data is used to train the model to enable generalization; (2) a subset of the data related to CEM for years 2021–2022 is used to create prediction for the CEM scope of this paper; and (3) the selected journals are not inclusive of all journals related to civil engineering or CEM, but rather represent a large representative sample of high quality publications to study the research trends.

Table 1. Journals included in data collection.

Domain	Journal Name
Structural	*Journal of Bridge Engineering* *Journal of Structural Engineering* *Journal of Cold Regions Engineering* *Journal of Performance of Constructed Facilities* *Practice Periodical on Structural Design and Construction*

Table 1. *Cont.*

Domain	Journal Name
Materials	*Journal of Composites for Construction* *Journal of Materials in Civil Engineering* *Journal of Nanomechanics and Micromechanics* *Journal of Engineering Mechanics*
Geotechnical	*International Journal of Geomechanics* *Journal of Geotechnical and Geoenvironmental Engineering* *GEOSTRATA Magazine*
Environmental and Water Resources	*Journal of Environmental Engineering* *Journal of Hydraulic Engineering* *Journal of Hydrologic Engineering* *Journal of Irrigation and Drainage Engineering* *Journal of Pipeline Systems Engineering and Practice* *Journal of Sustainable Water in the Built Environment* *Journal of Water Resources Planning and Management* *Journal of Waterway, Port, Coastal, and Ocean Engineering*
Transportation	*Journal of Highway and Transportation Research and Development (English Edition)* *Journal of Transportation Engineering, Part A: Systems* *Journal of Transportation Engineering, Part B: Pavements*
Cross-Disciplinary Civil Engineering	*Journal of Infrastructure Systems* *Journal of Hazardous, Toxic, and Radioactive Waste* *Natural Hazards Review* *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering* *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part B: Mechanical Engineering* *Journal of Architectural Engineering* *Journal of Urban Planning and Development* *Journal of Energy Engineering* *Journal of Computing in Civil Engineering* *Journal of Surveying Engineering*
Engineering Education and Practices	*Journal of Civil Engineering Education*
Aerospace Engineering	*Journal of Aerospace Engineering*
Others	*Civil Engineering Magazine* *Journal of Technical Topics in Civil Engineering* *Transactions of the American Society of Civil Engineers*
Construction Engineering and Management	*Journal of Construction Engineering and Management* *Leadership and Management in Engineering* *Journal of Legal Affairs and Dispute Resolution in Engineering and Construction* *Journal of Management in Engineering* *Automation in Construction* *International Journal of Project Management* *Engineering, Construction, and Architectural Management* *Construction Innovation* *Construction Management and Economics* *International Journal of Construction Management*

4.2. Dataset Construction

To build the machine learning model, the authors complemented the exported articles data and features exported from Lens with additional author-specific and journal-specific variables. This was done by designing a multidimensional network that links each article record to nodes for authors and nodes for journals as visualized in Figure 2. The created network functions as a large citation network that allows for calculating time-based bib-

liometric data for (1) authors, which included the number of published papers as well as the number of citations and (2) journals, which includes the number of published papers, the total citation count, and the mean value of citations for each paper. The network is domain-specific because it calculates the metrics using the papers in the network only. This allows for determining articles that have citations that are from high-quality peer-reviewed journals relevant to their fields. As such, the collected and cleaned data was compiled into a large multidimensional citation network, which was then used to create a unified dataset. It should be noted that both Random Forest (RF) and XGBoost machine learning algorithms used in this model are not influenced by the increased number of included features, neither are they are less susceptible to overfitting due to increased dimensionality; they are also more generalizable forms of modeling [26]. The final list of the resulting variables used in the model are summarized in Table 2. Finally, the data was converted to a classification problem based on a 95% percentile cut-off for citation counts within 5 years of the publication year to classify the papers into "High Impact" and "Non-High Impact" papers.
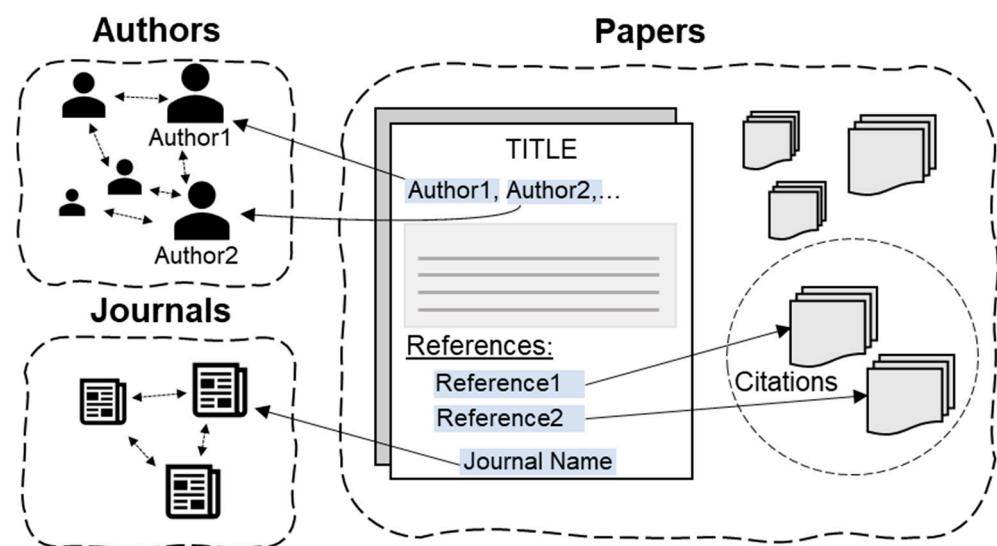

Figure 2. Dataset Extraction Process.

Table 2. Model Variables.

Category	Variable Name
Article	Publication year
	Number of authors
	Number of references
	Number of references in network [1]
	Total number of citations in network 5 years after publication [1]
Author	Total number of papers by authors
	Total number of citations for authors
	Number of papers per author
	Number of citations per author
Journal	Number of papers in the journal
	Number of citations per paper in journal
	Number of citations per paper in journal

[1] Number of citations and references in network includes citations and references from articles in the constructed network.

4.3. Machine Learning Models

Machine learning is a subdiscipline of artificial intelligence, which targets the utilization of computer-aided algorithms in building models based on sample data (training data) that enables making predictions or decisions. In general, machine learning is a data-based

technique, which incorporates various types of learning such as supervised, unsupervised, or reinforced learning and targets reducing human interference while making efficient and accurate predications or decisions [27]. In this paper, the following machine learning algorithms were utilized in developing a model that predicts the future impact of research studies in CEM domain: (1) RF Classifier and (2) XGBoost Classifier. The term "ensemble learning" refers to the fact that algorithms are built by combining predictions of multiple models using a given algorithm to enhance robustness over a single prediction of an individual model [28]. Ensemble learning methods are classified into: (1) bagging ensemble methods, where base learners are generated simultaneously (i.e., RF), and (2) boosting ensemble methods, in which the base models learn sequentially, using the knowledge of prior models' errors to boost performance (i.e., XGBoost) [29]. Both RF and XGBoost have proven satisfactory performance in previous research and are therefore considered in this paper [30–32].

4.3.1. RF Classifier

RF classifiers are fast and robust machine learning algorithms that fall under the category of ensemble learning algorithms [33]. RF is a machine learning algorithm that encompasses a combination of decision tree learning models to generate predictions [29]. For classification problems, the predicted class is yielded by a majority vote among the decision trees, i.e., the most frequent class [34]. Generally, to train an RF model, the following hyperparameters need to be selected and tuned: the number of decision trees, the splitting function, as well as the size of the random subsets of features [35]. RF models are more generalizable and less prone to overfitting [36]. As such, RF models are recognized for their increased classification performance and improved prediction accuracy [30].

4.3.2. XGBoost Classifier

XGBoost classifier is another ensemble learning algorithm which stands for Extreme Gradient Boosting and is a scalable implementation of gradient-boosting decision trees [37]. Similar to RF classifier, XGBoost uses multiple decision trees to build the algorithm. In XGBoost, the concept of "gradient boosting" stems from the general idea of "boosting" where a single weak model is enhanced by additively merging it with several other weak models in order to build a collectively superior model [37]. XGBoost algorithm has dominated many data science competitions in the past few years and is currently considered to have the leading combination of both prediction and computing performances [26].

4.4. Resampling for Imbalanced Data

The dataset used in this paper is imbalanced because the target is based on a 95% percentile split. More specifically, papers in the top 95 percentile by citation counts within 5 years of their publication year were classified as "High Impact" papers, while the remaining papers were classified as "Non-High Impact" papers; thus, this creates an imbalanced dataset. For machine learning models, imbalance in class distributions can generate biased classifiers that tend to be highly accurate over the majority class(es) but perform rather poorly over the minority class of interest [38]. To handle this concern, the Synthetic Minority Over-sampling Technique (SMOTE) has been applied in the training datasets to balance the ratio of the minority class of high-impact articles. SMOTE operates on the feature level of the dataset by generating synthetic instances of the minority class with respect to its nearest K-minority neighbors; hence, broadening the decision space for inductive learners such as decision tree-based or rule-based algorithms [38].

To account for the imbalanced nature of the model's target classification, the balanced accuracy score was adopted as a performance metric. Unlike raw accuracy scores, balanced accuracy prevents inflated estimations of performance on imbalanced datasets by averaging the recall scores per class. For binary classifiers, balanced accuracy is the mean of the specificity as well as the sensitivity metrics [39].

4.5. K-Fold Cross Validation, Hyperparameter Tuning, Model Performance Evaluation, and Selection

Figure 3 summarizes the adopted processes for cross-validation, model hyperparameter tuning, performance evaluation, and model selection. First, the constructed dataset was shuffled, stratified, and split into a training set with 80% of the records for training and validation of the models, and a testing set with the remaining 20% for robust performance evaluation. Second, a hyperparameter grid search was conducted to generate the optimal sets of hyperparameters in terms of the highest average of cross-validation balanced accuracy. Hyperparameter sets tuned for both RF and XGBoost classifiers are shown in Figure 3. As part of the hyperparameter grid search, 10 k-fold validation was performed. The average performance of all ten folds was used to evaluate the model. This technique ensured robustness and that overfitting was minimized. The model and hyperparameter combination with the highest mean 10-fold performance was selected and evaluated against the 20% testing set.
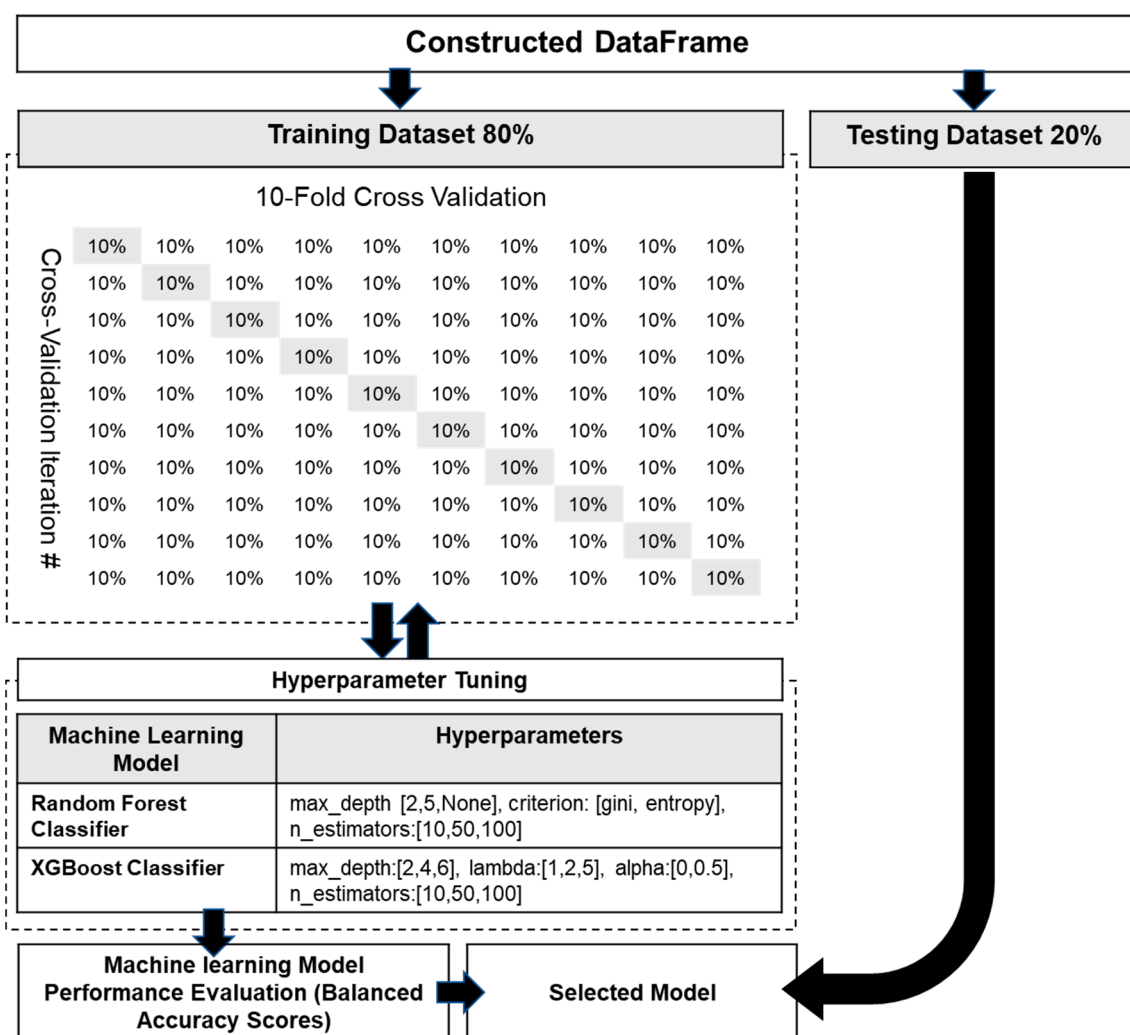

Figure 3. Machine learning model development summary.

4.6. Model Deployment

Upon selection and evaluation of the best-performing classification model, the authors utilized the selected model on the articles published in CEM-related journals in 2021 and 2022 to predict their probabilities of being impactful within 5 years after their publication year. The authors included articles published in CEM journals as shown in Table 1. The authors considered the articles with a probability of more than 90% to be highly impactful

within 5 years after their publication. A cut-off of 90% is selected by the authors such that it isolates approximately the top 10% of CEM papers published in 2021 and 2022. It is imperative to note that the authors considered a short-term span of 5 years because the future is changing, and the trendiness of research topics is continuously evolving and shifting. Previous studies have considered a time span up to 10 years as reasonable for identifying the research trends in various domains/applications [40,41]. That said, the short-term span of 5 years is considered reasonable to identify anticipated high-impact CEM research trends in the future considering the evolving nature of CEM research.

4.7. SNA Development

SNA is a graph theory-based mathematical method to analyze networks while taking into consideration the interconnectivity of its components [42]. In the context of scientometric studies, previous scientometric studies implemented SNA for analyzing the state of knowledge in relation to several topics and domains [43,44]. SNA enables researchers to unveil the knowledge structure of a specific field through the integration of co-occurrence analysis with network science [45]. The authors implemented SNA to identify promising CEM-related subdisciplines and trends in the future based on the results of the developed machine learning model. Nine CEM subdisciplines are considered in this paper and adapted from El-adaway et al. [12], which are as follows: D1: Legal and contractual issues; D2: Organizational issues; D3: Contracting; D4: Project planning and design; D5: Cost and schedule; D6: Labor and personnel issues; D7: Information technologies, robotics, and automation; D8: Strategy, decision making, risk, and finance; and D9: Contemporary issues.

The identified list of anticipated high-impact CEM articles was mapped with the CEM subdisciplines in the form of a matrix, referred to as reference matrix *Z*. In the developed reference matrix *Z*, the rows represent the CEM subdisciplines, and the columns represent the impactful CEM articles. If an article covers a topic related to a CEM subdiscipline, the value of its corresponding cell will be 1; otherwise, it will be 0. Figure 4 shows a descriptive example of the structure of a reference matrix. Let D_i denote a CEM subdiscipline where *I* is the total number of the CEM subdisciplines (9 subdisciplines). Further, let A_j denote an analyzed article where *J* is the total number of analyzed articles. For example, in Figure 4, the covered topic in the analyzed article A_{j+1} is related to the CEM subdisciplines D_i and D_I; thus, their corresponding cells have values of 1 whereas a value of 0 is entered for the remaining cells under the analyzed article A_{j+1}.

Figure 4. An illustrative example of the reference matrix.

Thereafter, the authors utilized SNA to quantitively analyze the developed reference matrix *Z*. In SNA, networks are visualized using nodes, representing the nine CEM subdisciplines, connected by links, representing their interconnectivity. Various mathematical methods can be used to analyze social networks to obtain valuable insights from their structures. Centrality is a main feature of SNA and its related metric, degree centrality (DC), is used to determine the number of links attached to each node [46]. In this paper, the authors applied DC as the SNA measure for evaluation of the CEM subdisciplines in terms of their consideration and co-occurrence with other subdisciplines in anticipated

impactful CEM research. The determination of DC for various nodes in the social network requires constructing an adjacency matrix A. The adjacency matrix A is determined, following Equation (1), by multiplying each reference matrix by its transpose and then replacing the diagonal values with zeros. A_{I*I} is an adjacency matrix of size ($I * I$), with I equal to the total number of the CEM subdisciplines (as previously highlighted, $I = 9$ in this paper); Z_{I*J} is a reference matrix, with J equal to the total number of the analyzed articles in the corresponding matrix; and i and j are the indices of the matrix rows and columns, respectively.

$$A_{N*N} = \begin{cases} Z_{I*J} * Z_{I*J}^T \ for \ i \neq j \\ 0 \ for \ i = j \end{cases} \quad (1)$$

Upon construction the adjacency matrix, DC is calculated for each CEM subdiscipline following Equation (2), where DC_i is the DC for the CEM subdiscipline i; and $V_{i,j}$ is the value of the cell in row i and column j of the adjacency matrix. It is worth noting that the value of DC does not represent the importance of the CEM subdiscipline but rather its frequency of consideration as well as interconnectivity with other CEM subdisciplines in the anticipated impactful CEM research in the future. This level of abstracting is considered acceptable as the main aim of utilizing SNA is to quantitatively identify impactful CEM research trends based on the predictions of the developed classification model in this research.

$$DC_i = \sum_{j:j \neq i} V_{i,j} \quad (2)$$

4.8. Tools and Software Used

The following tools and software were used: Python 3.9.10 Programming Language [47,48], NumPy 1.26.4 [49,50], Pandas 1.5.3 [51], Scikit-Learn 1.1.3 [52], Imblearn 0.9.1 [53], Matplotlib 3.8.3 and Seaborn 0.13.2 [54,55], VSCode 1.87.2 and Jupyter 5.7.2, Gephi 0.9.7 [56], and Microsoft 365 Excel.

5. Results and Analysis

5.1. Exploratory Analysis of the Constructed Dataset

Figure 5 shows the frequency distribution of the collected articles in terms of publication year, where an increasing trend can be observed from the year 1985 up to 2021. This increasing trend had been previously attributed by El-adaway et al. [12] to the collective impact of increasing publication pressure as well as the growing number of graduate students and faculty across the civil engineering disciplines. It should be noted that the data collection was conducted on 15 March 2022, and hence, articles published in 2022 were only up to this date.

In addition, Figure 6 shows the correlation heatmap that represents the correlation coefficients amongst each of the variables employed in the model.

It can be observed that the highest correlation exists between the "Number of Citations per Author" and the "Total Number of Citations for Authors" (0.89), followed by the correlation between the "Number of Citations per Paper in Journal" and the "Number of Citations for Journal" (0.84). This was expected as the correlated variables are related and cumulative. In fact, the correlation matrix is usually used to detect multicollinearity between the variables. In machine learning models, multicollinearity between the variables is a threat to their predictive ability and the reliability of the obtained results [57]. Nevertheless, both RF and XGBoost are decision tree-based computational algorithms and are hence immune to multicollinearity between variables [58]. As such, all variables, previously shown in Table 2, were considered in the development of the machine learning models in this paper.
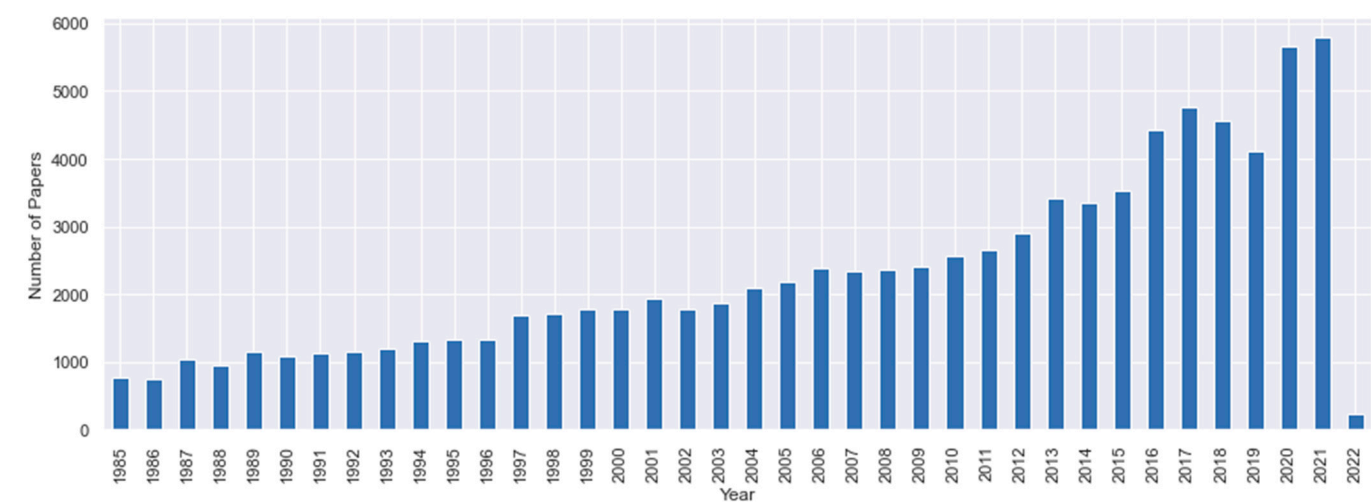
Figure 5. Number of papers by year.

Figure 6. Correlation heatmap.

5.2. Results of the Developed Machine Learning Models

5.2.1. Selection of the Best-Performing Prediction Model

Table 3 presents the results of both machine learning algorithms using selected sets of optimized hyperparameters. The most optimal performing model was the XGBoost classification model, with a mean cross-validation balanced accuracy of 79.5%. On the other hand, the RF model had an accuracy of 79.1%. Therefore, it was selected for the prediction of the impact of CEM publications in the future.

Table 3. Summary of machine learning models.

Algorithm	Best Set of Hyperparameters	Mean Cross-Validation Balanced Accuracy (%)
RF	Criterion = 'entropy', max_depth = 5, n_estimators = 50	79.1%
XGBoost	Alpha = 0.5, Lambda = 2, max_depth = 4, n_estimators = 10	79.5%

5.2.2. Evaluation of the Best-Performing Prediction Model

Using the testing dataset, the XGBoost classification model reached a balanced accuracy of 80.71%. Figure 7 presents the confusion matrices for the XGBoost classification model. The confusion matrix for the testing dataset illustrates that the model can correctly classify almost 82% of highly impactful articles, whereas 79.2% of the other articles were correctly classified. Figure 8 provides the feature importance of the model variables. The number of references in the network had the highest feature importance factor. As explained in the previous subsections, the number of references in the network is a count for references that are published in the civil engineering journals listed in Table 1. This significantly higher feature importance is in line with previous research which highlighted the correlation between the length of an article's references list and its citation count [59]. In a transitive manner, as the references in the network are considered as a selected civil engineering-specific subset of the overall references list, the correlation is extended and is further underlined to exceed the importance of the overall references count which ranked in 5th place. Previous studies attributed this correlation to multiple factors such as the following: (1) papers with a high number of references can be more comprehensive, such as literature reviews, which are naturally more often relied on by subsequent papers, and (2) the authors' knowledge of the field is more extensive and thus their paper is presenting research that is significant in the field. Other interpretations for this finding were furnished by Fox et al. [59], stating that (1) articles with more references usually cover a wide variety of arguments to support/counter the presented concepts and (2) a long list of references may increase an article's visibility on citation-based search engines (e.g., Web of Science and Google Scholar). The remaining features had relatively modest importance weights in comparison, as shown in Figure 8.
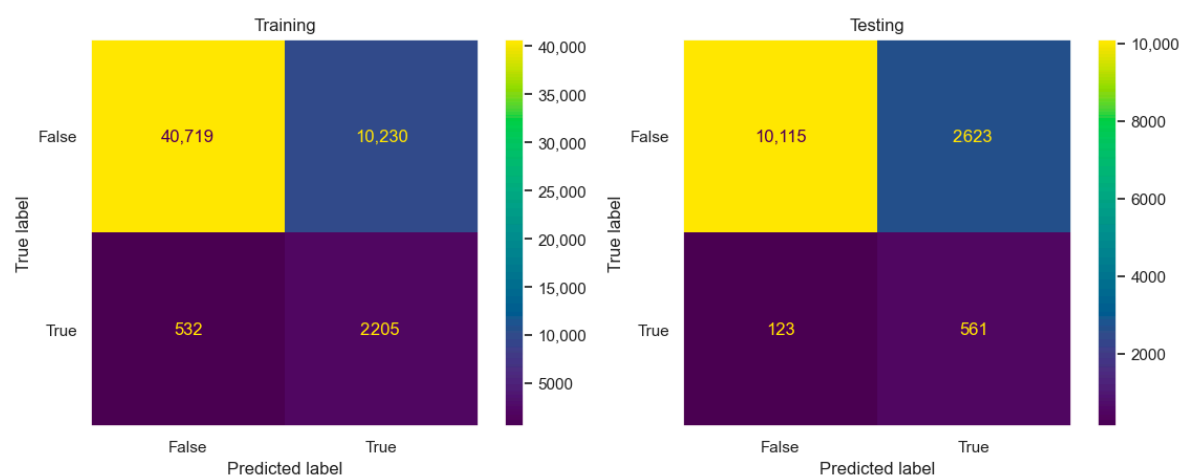
Figure 7. Confusion matrix for training (**left**) and testing (**right**).
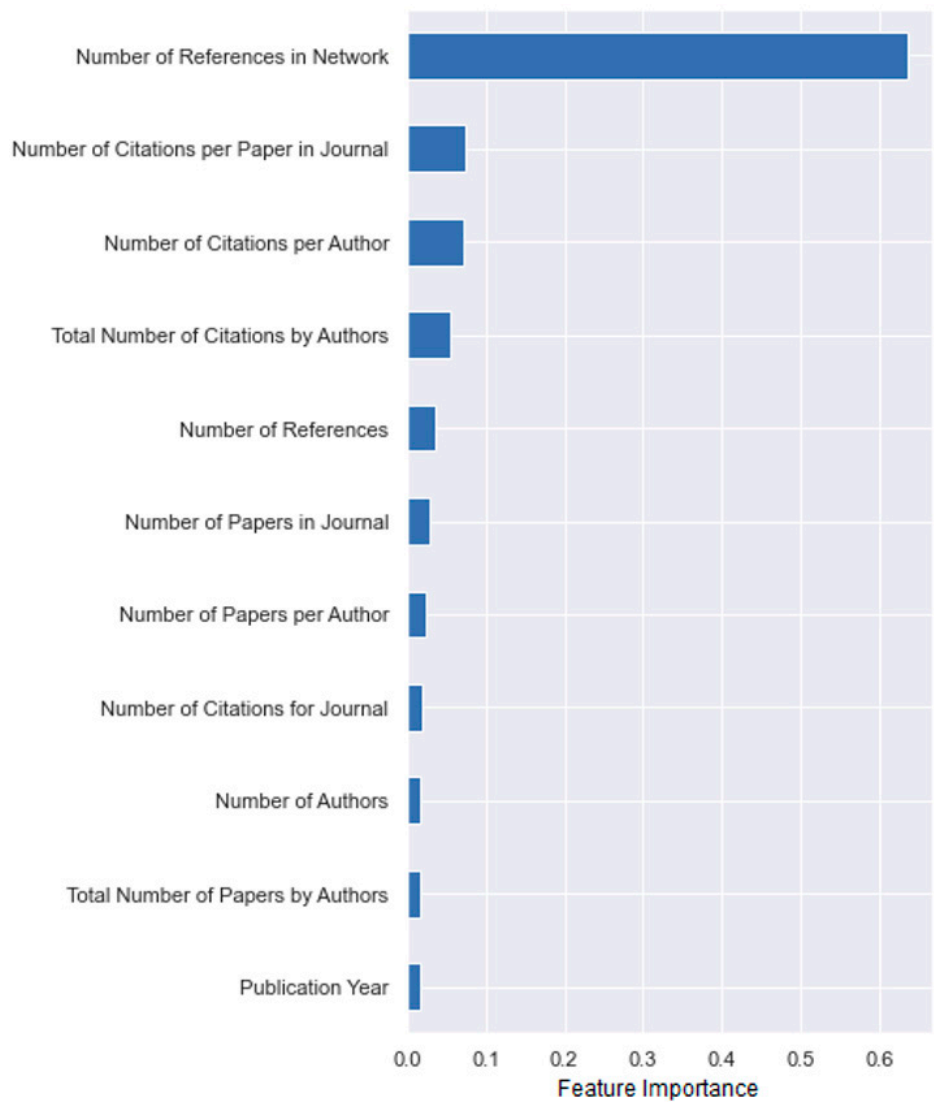

Figure 8. Feature importance.

5.3. Impactful CEM Research Trends

Figure 9 shows the number of predicted impactful CEM papers by subdiscipline. As previously highlighted, the authors considered the articles with a probability of more

than 90% to be highly impactful within 5 years after their publication, which resulted in a total of 197 articles. However, the cumulative number of articles as shown is more than 197 articles due to the fact that the topics of some articles are related to multiple CEM subdisciplines (i.e., double counting). That said, the top CEM subdisciplines based on the number of predicted impactful CEM papers are "*Project planning and design*", "*Organizational issues*", and "*Information technologies, robotics, and automation*". On the other hand, the last CEM subdiscipline with anticipated impactful CEM papers is "*Legal and contractual issues*". A more detailed interpretation regarding the growth of these trends is provided in the discussion section.
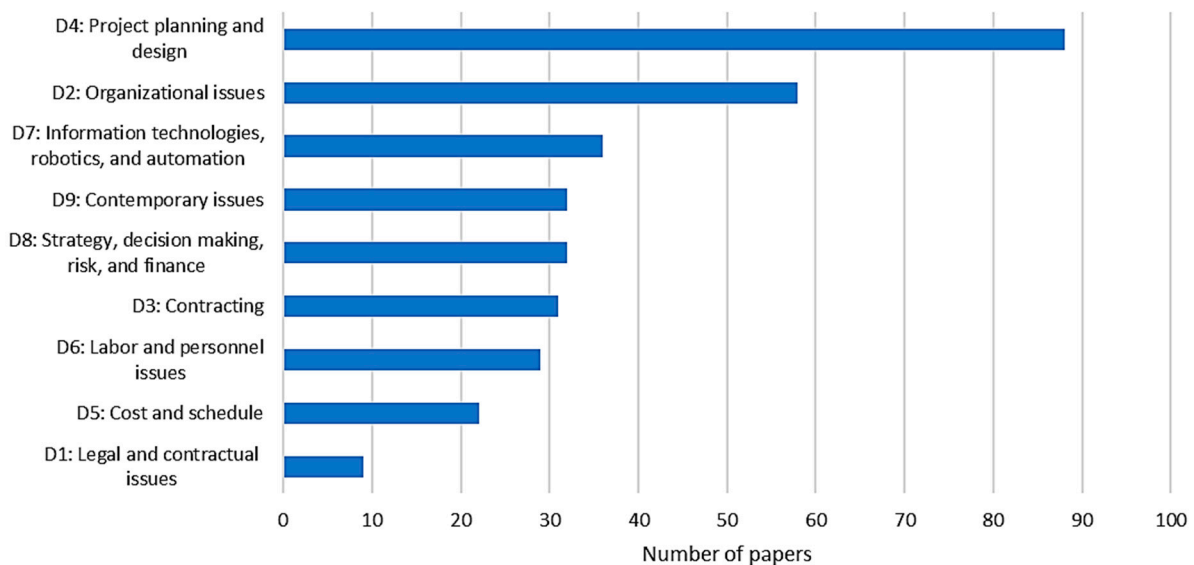

Figure 9. Number of predicted impactful CEM papers by CEM subdisciplines.

In addition, the authors utilized SNA to further understand and analyze the obtained results in terms of co-occurrence and interconnectivity among the CEM subdisciplines in anticipated impactful CEM research. It is imperative to note that interconnectivity between two subdisciplines is due to having one or more articles that cover both subdisciplines simultaneously. Figure 10 presents the diagram for the network between the CEM subdisciplines in the predicted impactful CEM papers, as well as the corresponding color-coded adjacency matrix *A*. The SNA results showed that "*Project planning and design*" is the subdiscipline with the most connectivity with other CEM subdisciplines, as shown in Figure 10. On the other hand, the SNA results showed that there is no connectivity between the "*Legal and contractual issues*" subdiscipline with the three subdisciplines "*Cost and schedule*", "*Labor and personnel issues*", and "*Contemporary issues*" in the predicated impactful CEM-related publications. Also, the SNA results indicated that there is no connectivity between "*Labor and personnel issues*" with the two subdisciplines "*Contracting*" and "*Cost and schedule*". By no connectivity, the authors mean that there is no predicted impactful CEM article that covers any pair of these subdisciplines simultaneously. Not being interconnected does not mean that the two subdisciplines are not related; instead, it may be that most of the predicted impactful CEM articles are more focused and concentrated on tackling one CEM topic/subdiscipline in a more detailed manner rather than covering multiple CEM topics/subdisciplines.
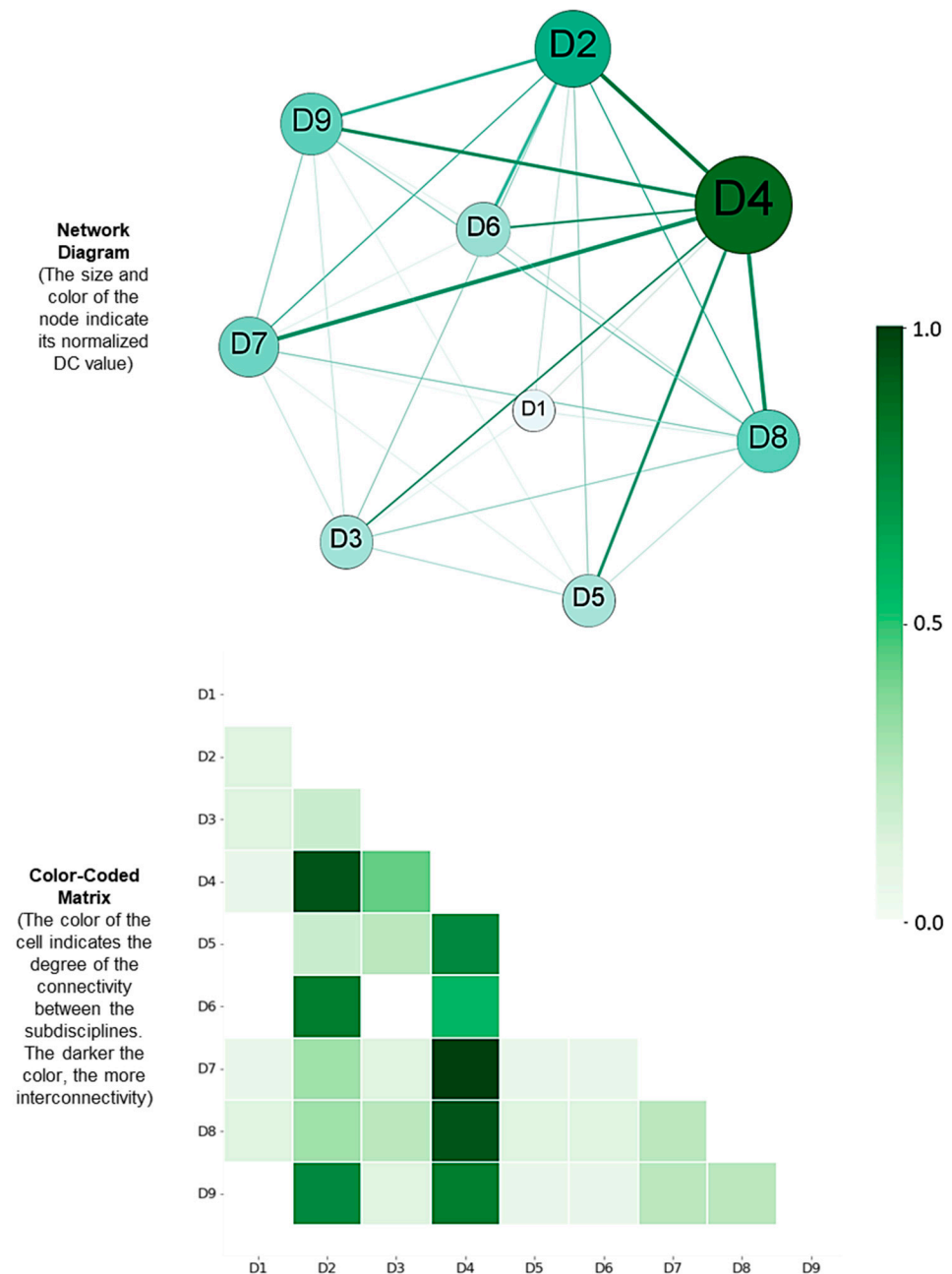
Figure 10. Results of the SNA of CEM subdisciplines.

6. Discussion

The results of the machine learning classification model (i.e., XGBoost) and SNA enabled the identification of the following impactful CEM research trends in the future:

- Results show that "Project planning and design" is considered a central CEM subdiscipline topic that is strongly connected to other subdisciplines, as shown by the links in the network diagram and the cells in the color-coded matrix in Figure 10. In the study by Jin et al. [2], it was found that topics related to the "Project planning and design" subdiscipline, such as scheduling and planning, were among the top studied topics in the period from 2000 to 2018 based on a quantitative analysis of keywords. The findings in this paper imply that the growth of the "Project planning and design" subdiscipline is expected to continue to grow. The "Project planning and design" is a primary area of CEM as it covers various vital topics within the CEM domain, including project management, scheduling, engineering design, and

construction methods, among others. As such, it may be considered central to the growth of CEM research.

- The "Organizational issues" subdiscipline tackles various trendy research topics in today's construction industry including equality and diversity, human resources management, relationships between project stakeholders, and project teams, among others. Topics related to equality and diversity in the construction industry have gained substantial attention since the publication of the well-known special issue by Dainty and Bagilhole [60]. Since then, various publications investigated the needed steps to address the lack of equality and diversity within the construction sector [61,62]. In addition, various publications emphasized the strong tie between the structure and culture of project teams, the relationship between stakeholders, and the success of construction projects [63,64]. Moreover, organizational issues, such as organizational work structures, virtual teams, and organizational resilience, were identified among the anticipated future research streams as a result of the COVID-19 pandemic [65].

- The "Information technologies, robotics, and automation" subdiscipline focuses on the adoption of new technologies and automation of construction processes using various techniques, including BIM, Geographic Information System (GIS), blockchain, Internet of Things (IoT), augmented reality, and virtual reality, among others. In relation to the CEM domain, El-adaway et al. [12] found that the number of publications on the "Information technologies, robotics, and automation" subdiscipline began to spike starting from the year 2010. Nowadays, the diffusion of the "Construction 4.0" concept reflects the trendy dynamic of the utilization of technologies to reshape the way projects are designed, constructed, and operated [66]. Ghaffar et al. [67] stated that "*the COVID-19 pandemic has forced many construction players to digitize to ensure safety and productivity, this dynamic will likely continue to be accelerated in the future years*". This emphasizes the anticipated significance and trendiness of this subdiscipline in the future as an assisting tool for much research subdisciplines and processes within the CEM domain.

- The "Legal and contractual issues" subdiscipline covers several topics including contractual provisions and guidelines, applied laws and regulations, jurisdiction, claims, and disputes, among others. As previously highlighted, the "Legal and contractual issues" subdiscipline possessed the least number of anticipated impactful CEM papers, as well as the least DC value in the conducted SNA. This result is in line, to some extent, with the findings of El-adaway et al. [12], which highlighted that "Legal and contractual issues" is the least cited CEM subdiscipline compared to others. This result can be ascribed to the impact of the research community size and their output on the citation metrics. A community of a smaller size is expected to have lower research output and fewer citations compared to other communities of a larger size. Moreover, according to de la Garza [68], the magnitude and quality of research output related to a specific topic depend on many factors, including funding availabilities and the interest of researchers. Overall, it is worth highlighting that possessing the least number of anticipated impactful CEM paper and/or DC value does not imply that the subdiscipline is less important compared to other CEM subdisciplines, because all subdisciplines collectively impact the CEM domain and the construction industry as a whole.

7. Recommendations

Based on the previous discussion, it can be seen that research trends are expected to keep growing in the subdiscipline of "Project planning and design" which includes topics such as project management, scheduling, engineering design, and construction methods, among others, followed by "Organizational issues" which includes topics such as equality and diversity, human resources management, relationships between project stakeholders, and project teams, among others. Accordingly, stakeholders in CEM research including publishers, journal editors, conference organizers, and funders can foster the

growth of research in these areas, considering that they are expected to have growing impact. However, there are limitations, as will be discussed in the following section, that should be considered in the decision process. From another perspective, the SNA results also highlight the difference in connections between the subdisciplines. In some instances, these connections are relatively weak compared to others, such as that between "Information technologies, robotics, and automation" and "Contracting". There is a need for more research that connects the topics and even cross-disciplinary research that creates advances in CEM.

8. Limitations

This study includes the following limitations: (1) The findings are based on a collective dataset analysis which does not necessarily uncover the relative impact of a paper within its subdiscipline. For example, the field of "Legal and contractual issues" generally has lower citation counts compared to other areas in CEM. As such, two highly "*impactful*" papers in different subdisciplines do not necessarily have comparable citation counts. Future research can work on introducing new metrics that can better describe the impact of a paper within its field. (2) The findings in this paper are based on a modeling approach with selected inputs about publications. However, expert opinion is ultimately needed to judge the quality and impact of a publication which may be exceptional and against the outcomes of the model. (3) Recommendations for research directions need to be supported by thorough literature reviews to determine possible research gaps and opportunities.

9. Conclusions

This research developed a computational model for forecasting the significance and impact of publications in the CEM domain. This was achieved by conducting an exploratory and predictive analysis of citation metrics using machine learning and SNA. A dataset of 93,868 publications related to the civil engineering field, with a focus on the CEM domain, was used. Two machine learning algorithms, RF and XGBoost, were tested to create a classification model. Validation of the RF and XGBoost resulted in a balanced accuracy of 79.1% and 79.5%, respectively. Accordingly, XGBoost was selected. Testing of the XGBoost model revealed a balanced accuracy of 80.71%. The findings showed that the number of references in a paper has a significant influence on its citation count. Also, results showed that the top three CEM subdisciplines in terms of the number of predicted impactful CEM papers are "Project planning and design", "*Organizational issues*", *and* "*Information technologies, robotics, and automation*". *On the other hand, the least number of impactful CEM publications belonged to the "Legal and contractual issues" subdiscipline*. Ultimately, this paper contributes to the CEM body of knowledge through studying the citation level, strength, and interconnectivity between CEM-related subdisciplines; identifying CEM research areas that are more likely to result in highly cited publications; providing early signs for recently published articles that are most likely to be of high impact in the future rather than just using the present day status; highlighting CEM subdisciplines with the highest abundance of potentially impactful publications; and capturing underlying changes in research interests and impactful research trends over any desired period of interest through incorporating adaptive learning methods.

References

1. Aboulezz, M.A. Mapping the Construction Engineering and Management Discipline. Ph.D. Thesis, Worcester Polytechnic Institute, Worcester, MA, USA, 2003.
2. Jin, R.; Zuo, J.; Hong, J. Scientometric Review of Articles Published in ASCE's *Journal of Construction Engineering and Management* from 2000 to 2018. *J. Constr. Eng. Manag.* **2019**, *145*, 06019001. [CrossRef]
3. Pietroforte, R.; Stefani, T.P. ASCE Journal of Construction Engineering and Management: Review of the Years 1983–2000. *J. Constr. Eng. Manag.* **2004**, *130*, 440–448. [CrossRef]
4. Carpenter, C.R.; Cone, D.C.; Sarli, C.C. Using Publication Metrics to Highlight Academic Productivity and Research Impact. *Acad. Emerg. Med.* **2014**, *21*, 1160–1172. [CrossRef]
5. National Science Foundation (NSF). *Proposal and Award Policies and Procedures Guide (PAPPG)*; NSF: Alexandria, VA, USA, 2022.
6. Aragón, A.M. A Measure for the Impact of Research. *Sci. Rep.* **2013**, *3*, 1649. [CrossRef]
7. Wilson, L. *The Academic Man: A Study in the Sociology of a Profession*; Routledge: New York, NY, USA, 2017; ISBN 978-1-315-13080-4.
8. Gross, P.L.K.; Gross, E.M. College Libraries and Chemical Education. *Science* **1927**, *66*, 385–389. [CrossRef]
9. Lawani, S.M. Citation Analysis and the Quality of Scientific Productivity. *BioScience* **1977**, *27*, 26–31. [CrossRef]
10. Fu, L.; Aliferis, C. Using Content-Based and Bibliometric Features for Machine Learning Models to Predict Citation Counts in the Biomedical Literature. *Scientometrics* **2010**, *85*, 257–270. [CrossRef]
11. Yang, J.; Liu, Z. The Effect of Citation Behaviour on Knowledge Diffusion and Intellectual Structure. *J. Informetr.* **2022**, *16*, 101225. [CrossRef]
12. El-adaway, I.H.; Ali, G.; Assaad, R.; Elsayegh, A.; Abotaleb, I.S. Analytic Overview of Citation Metrics in the Civil Engineering Domain with Focus on Construction Engineering and Management Specialty Area and Its Subdisciplines. *J. Constr. Eng. Manag.* **2019**, *145*, 04019060. [CrossRef]
13. Chakraborty, T.; Kumar, S.; Goyal, P.; Ganguly, N.; Mukherjee, A. Towards a Stratified Learning Approach to Predict Future Citation Counts. In Proceedings of the IEEE/ACM Joint Conference on Digital Libraries, London, UK, 8–12 September 2014; pp. 351–360.
14. Staszkiewicz, P. The Application of Citation Count Regression to Identify Important Papers in the Literature on Non-Audit Fees. *Manag. Audit. J.* **2018**, *34*, 96–115. [CrossRef]
15. Weis, J.W.; Jacobson, J.M. Learning on Knowledge Graph Dynamics Provides an Early Warning of Impactful Research. *Nat. Biotechnol.* **2021**, *39*, 1300–1307. [CrossRef]
16. Jin, R.; Zou, Y.; Gidado, K.; Ashton, P.; Painting, N. Scientometric Analysis of BIM-Based Research in Construction Engineering and Management. *Eng. Constr. Archit. Manag.* **2019**, *26*, 1750–1776. [CrossRef]
17. Hammersley, M. On 'Systematic' Reviews of Research Literatures: A 'Narrative' Response to Evans & Benefield. *Br. Educ. Res. J.* **2001**, *27*, 543–554. [CrossRef]
18. Pietroforte, R.; Aboulezz, M.A. ASCE *Journal of Management in Engineering*: Review of the Years 1985–2002. *J. Manag. Eng.* **2005**, *21*, 125–130. [CrossRef]
19. Yi, W.; Chan, A.P.C. Critical Review of Labor Productivity Research in Construction Journals. *J. Manag. Eng.* **2014**, *30*, 214–225. [CrossRef]
20. Larsen, J.K.; Ussing, L.F.; Brunø, T.D. Trend-Analysis and Research Direction in Construction Management Literature. In Proceedings of the ICCREM 2013: Construction and Operation in the Context of Sustainability, Karlsruhe, Germany, 10–11 October 2013; pp. 73–82. [CrossRef]
21. Pan, Y.; Zhang, L. Roles of Artificial Intelligence in Construction Engineering and Management: A Critical Review and Future Trends. *Autom. Constr.* **2021**, *122*, 103517. [CrossRef]
22. Bröchner, J.; Björk, B. Where to Submit? Journal Choice by Construction Management Authors. *Constr. Manag. Econ.* **2008**, *26*, 739–749. [CrossRef]
23. Weihs, L.; Etzioni, O. Learning to Predict Citation-Based Impact Measures. In Proceedings of the 2017 ACM/IEEE Joint Conference on Digital Libraries (JCDL), Toronto, ON, Canada, 19 June 2017; pp. 1–10.
24. The Lens. The Lens—Free & Open Patent and Scholarly Search. Available online: https://www.lens.org/lens (accessed on 15 March 2022).
25. Jefferson, O.A.; Koellhofer, D.; Warren, B.; Jefferson, R. The Lens MetaRecord and LensID: An Open Identifier System for Aggregated Metadata and Versioning of Knowledge Artefacts. 2019. Available online: https://osf.io/preprints/lissa/t56yh (accessed on 15 March 2022).
26. NVIDIA What Is XGBoost? Available online: https://www.nvidia.com/en-us/glossary/xgboost/ (accessed on 15 March 2022).
27. Khalef, R.; El-adaway, I.H. Automated Identification of Substantial Changes in Construction Projects of Airport Improvement Program: Machine Learning and Natural Language Processing Comparative Analysis. *J. Manag. Eng.* **2021**, *37*, 04021062. [CrossRef]
28. Piryonesi, S.M.; El-Diraby, T.E. Role of Data Analytics in Infrastructure Asset Management: Overcoming Data Size and Quality Problems. *J. Transp. Eng. Part B Pavements* **2020**, *146*, 04020022. [CrossRef]
29. Scikit-Learn Ensembles: Gradient Boosting, Random Forests, Bagging, Voting, Stacking. Available online: https://scikit-learn.org/stable/modules/ensemble.html (accessed on 15 March 2022).

30. Azeez, D.; Gan, K.B.; Ali, M.A.M.; Ismail, M.S. Secondary Triage Classification Using an Ensemble Random Forest Technique. *Technol. Health Care* **2015**, *23*, 419–428. [CrossRef]

31. Wu, J.; Ma, D.; Wang, W. Leakage Identification in Water Distribution Networks Based on XGBoost Algorithm. *J. Water Resour. Plan. Manag.* **2022**, *148*, 04021107. [CrossRef]

32. Wang, M.-X.; Huang, D.; Wang, G.; Li, D.-Q. SS-XGBoost: A Machine Learning Framework for Predicting Newmark Sliding Displacements of Slopes. *J. Geotech. Geoenvironmental Eng.* **2020**, *146*, 04020074. [CrossRef]

33. Xu, C.; Liu, X.; Wang, E.; Wang, S. Calibration of the Microparameters of Rock Specimens by Using Various Machine Learning Algorithms. *Int. J. Geomech.* **2021**, *21*, 04021060. [CrossRef]

34. IBM What Is Random Forest? Available online: https://www.ibm.com/topics/random-forest (accessed on 15 March 2022).

35. Ahmed, M.O.; Khalef, R.; Ali, G.G.; El-adaway, I.H. Evaluating Deterioration of Tunnels Using Computational Machine Learning Algorithms. *J. Constr. Eng. Manag.* **2021**, *147*, 04021125. [CrossRef]

36. Gupta, R.; Bruce-Konuah, A.; Howard, A. Achieving Energy Resilience through Smart Storage of Solar Electricity at Dwelling and Community Level. *Energy Build.* **2019**, *195*, 1–15. [CrossRef]

37. Chen, T.; Guestrin, C. XGBoost: A Scalable Tree Boosting System. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; Association for Computing Machinery: New York, NY, USA, 2016; pp. 785–794.

38. Chawla, N.V.; Lazarevic, A.; Hall, L.O.; Bowyer, K.W. SMOTEBoost: Improving Prediction of the Minority Class in Boosting. In Proceedings of the Knowledge Discovery in Databases: PKDD 2003, Cavtat-Dubrovnik, Croatia, 22–26 September 2003; Lavrač, N., Gamberger, D., Todorovski, L., Blockeel, H., Eds.; Springer: Berlin/Heidelberg, Germany, 2003; pp. 107–119.

39. Brodersen, K.H.; Ong, C.S.; Stephan, K.E.; Buhmann, J.M. The Balanced Accuracy and Its Posterior Distribution. In Proceedings of the 2010 20th International Conference on Pattern Recognition, Istanbul, Turkey, 23–26 August 2010; pp. 3121–3124.

40. Lemaire, C.; Rivest, L.; Boton, C.; Danjou, C.; Braesch, C.; Nyffenegger, F. Analyzing BIM Topics and Clusters through Ten Years of Scientific Publications. *J. Inf. Technol. Constr. (ITcon) Spec. Issue Archit. Inform.* **2019**, *24*, 273–298.

41. Onososen, A.O.; Musonda, I. Research Focus for Construction Robotics and Human-Robot Teams towards Resilience in Construction: Scientometric Review. *J. Eng. Des. Technol.* **2022**, *21*, 502–526. [CrossRef]

42. Otte, E.; Rousseau, R. Social Network Analysis: A Powerful Strategy, Also for the Information Sciences. *J. Inf. Sci.* **2002**, *28*, 441–453. [CrossRef]

43. Ali, G.G.; El-adaway, I.H. Distributed Solar Generation: Current Knowledge and Future Trends. *J. Infrastruct. Syst.* **2024**, *30*, 03123002. [CrossRef]

44. Elbashbishy, T.S.; Ali, G.G.; El-adaway, I.H. Blockchain Technology in the Construction Industry: Mapping Current Research Trends Using Social Network Analysis and Clustering. *Constr. Manag. Econ.* **2022**, *40*, 406–427. [CrossRef]

45. Choudhury, N.; Uddin, S. Time-Aware Link Prediction to Explore Network Effects on Temporal Knowledge Evolution. *Scientometrics* **2016**, *108*, 745–776. [CrossRef]

46. Freeman, L.C. Centrality in Social Networks Conceptual Clarification. *Soc. Netw.* **1978**, *1*, 215–239. [CrossRef]

47. Oliphant, T.E. Python for Scientific Computing. *Comput. Sci. Eng.* **2007**, *9*, 10–20. [CrossRef]

48. Millman, K.J.; Aivazis, M. Python for Scientists and Engineers. *Comput. Sci. Eng.* **2011**, *13*, 9–12. [CrossRef]

49. Oliphant, T.E. *A Guide to NumPy*; Trelgol Publishing: Provo, UT, USA, 2006; Volume 1.

50. Van der Walt, S.; Colbert, S.C.; Varoquaux, G. The NumPy Array: A Structure for Efficient Numerical Computation. *Comput. Sci. Eng.* **2011**, *13*, 22–30. [CrossRef]

51. McKinney, W. Pandas: A Foundational Python Library for Data Analysis and Statistics. *Python High Perform. Sci. Comput.* **2011**, *14*, 1–9.

52. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-Learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.

53. Lemaître, G.; Nogueira, F.; Aridas, C.K. Imbalanced-Learn: A Python Toolbox to Tackle the Curse of Imbalanced Datasets in Machine Learning. *J. Mach. Learn. Res.* **2017**, *18*, 559–563.

54. Hunter, J.D. Matplotlib: A 2D Graphics Environment. *Comput. Sci. Eng.* **2007**, *9*, 90–95. [CrossRef]

55. Waskom, M. Seaborn: Statistical Data Visualization. *JOSS* **2021**, *6*, 3021. [CrossRef]

56. Bastian, M.; Heymann, S.; Jacomy, M. Gephi: An Open Source Software for Exploring and Manipulating Networks. In Proceedings of the International AAAI Conference on Web and Social Media, San Jose, CA, USA, 19 March 2009; Volume 3, pp. 361–362.

57. Chakrabarti, S. Voltage Stability Monitoring by Artificial Neural Network Using a Regression-Based Feature Selection Method. *Expert Syst. Appl.* **2008**, *35*, 1802–1808. [CrossRef]

58. Wang, R.; Lu, S.; Li, Q. Multi-Criteria Comprehensive Study on Predictive Algorithm of Hourly Heating Energy Consumption for Residential Buildings. *Sustain. Cities Soc.* **2019**, *49*, 101623. [CrossRef]

59. Fox, C.W.; Paine, C.E.T.; Sauterey, B. Citations Increase with Manuscript Length, Author Number, and References Cited in Ecology Journals. *Ecol. Evol.* **2016**, *6*, 7717–7726. [CrossRef]

60. Dainty, A.R.J.; Bagilhole, B.M. Guest Editorial. *Constr. Manag. Econ.* **2005**, *23*, 995–1000. [CrossRef]

61. Baker, M.; French, E.; Ali, M. Insights into Ineffectiveness of Gender Equality and Diversity Initiatives in Project-Based Organizations. *J. Manag. Eng.* **2021**, *37*, 04021013. [CrossRef]

62. Al-Bayati, A.J.; Abudayyeh, O.; Fredericks, T.; Butt, S.E. Managing Cultural Diversity at U.S. Construction Sites: Hispanic Workers' Perspectives. *J. Constr. Eng. Manag.* **2017**, *143*, 04017064. [CrossRef]

63. Chinowsky, P.S.; Rojas, E.M. Virtual Teams: Guide to Successful Implementation. *J. Manag. Eng.* **2003**, *19*, 98–106. [CrossRef]

64. Albanese, R. Team-Building Process: Key to Better Project Results. *J. Manag. Eng.* **1994**, *10*, 36–44. [CrossRef]

65. Assaad, R.; El-adaway, I.H. Guidelines for Responding to COVID-19 Pandemic: Best Practices, Impacts, and Future Research Directions. *J. Manag. Eng.* **2021**, *37*, 06021001. [CrossRef]

66. Forcael, E.; Ferrari, I.; Opazo-Vega, A.; Pulido-Arcas, J.A. Construction 4.0: A Literature Review. *Sustainability* **2020**, *12*, 9755. [CrossRef]

67. Ghaffar, S.H.; Mullett, P.; Pei, E.; Roberts, J. (Eds.) *Innovation in Construction: A Practical Guide to Transforming the Construction Industry*; Springer International Publishing: Cham, Switzerland, 2022; ISBN 978-3-030-95797-1.

68. De La Garza, J.M. Sponsored Research and Its Impact on Universities, Faculty, and Journals. *J. Constr. Eng. Manag.* **2007**, *133*, 708–709. [CrossRef]