Overseas Construction Suppliers Assessment based on Clustering and Sentiment Analysis

Qing Mei Shi¹, Siva Shankar Ramasamy², Praveenkumar Somasundara³
linlin_xu@cmu.ac.th¹, sivashankar.r@cmu.ac.th², pravsoma@qti.qualcomm.com³

International College of Digital Innovation, Chiangmai University, Chiang Mai, Thailand¹,², Qualcomm Technologies Inc., USA³

Abstract. The Global construction projects requires a combination of contractors, suppliers, Engineers, Real estate people, Labors in foreign soil. It is better to know or assess the team before involving them with major projects or government-based projects. The article proposes an innovative approach to evaluate overseas construction suppliers using clustering and sentiment analysis techniques which analyze the textual data from various sources, such as customer reviews and supplier communications. The sentiment scores are utilized for supplier classification using machine learning algorithms. This work study and validates the effectiveness of our approach in assisting construction companies with supplier selection and risk management. Despite some limitations, our approach provides valuable insights into supplier performance and reputation. By leveraging sentiment analysis, construction companies can make informed decisions and enhance project outcomes through better supplier management. The method also involves K-Means clustering to set up the construction unit to increase the score of the construction supply to the projects. This research contributes to advancing sentiment analysis applications in the construction industry on a global scale.

Keywords: Sentiment analysis, Overseas construction suppliers, Classification techniques, Supplier evaluation, Risk management.

1 Introduction

After the world witnessed pandemic, people learned the significance of having house for their own. By this learning the construction suppliers, industry related to construction, has find a new curve and also significant rise in international collaborations as well as the engagement of overseas suppliers. While this globalization brings opportunities for increased efficiency and access to a wider range of resources, it also introduces challenges related to trust and reliability. Maintaining high standards of integrity and ensuring the credibility of suppliers are crucial factors for successful project execution and sustainable growth in the construction sector.

Traditionally, assessing the Assessment of suppliers has been a complex task, often relying on subjective evaluations or limited information. Supply requests and receiving messages are missing the traditional phone calls or human communication system. Old methods faced many problems and failures in Covid times. Supply communication, payment issues to suppliers and employees, evidence issues towards the salary and payment logbook is mis handled or missing.
sometimes. However, with the Digital platforms and mobile technologies are giving a promising avenue for revolutionizing the way supplier communication, employee maintenance, site operations shall be improved.

The company will add the suppliers and employees in App to the different groups. Employees are recommended to send the current location and image near the equipment to the contractor manager at the construction site. If order food in app is encouraged and deliver the food to the construction site. Within a few days, the company can understand and order the food for the entire team to the site itself. The supplier selection and construction workers shall be selected by the way they stay near the site. This paper aims to propose and analyze an assessment platform specifically designed for overseas construction suppliers. By leveraging the unique features of sentimental analysis for evaluating and rating the Assessment suppliers. Classification will be used to classify the impactable properties of the construction suppliers to be selected for analysis.

The study shares findings which exhibit significant implications for both the international collaborated construction industry. A robust as well as transparent integrity assessment platform can enhance trust among stakeholders, minimize risks associated with unreliable suppliers, and foster a more efficient and competitive construction market.

Employee Insurance will be handled by the supplier or employee contractor, which shall be submitted to the police station for the crime clearance and sent to the hospital for the medical records in digital way to check the Insurance, stay and bank account details) are all digitally transferred to the construction company and blockchain in the Chinese Company’s data. The blockchain process will be handled and maintained by the company itself.

Additionally, this study indirectly encourages to the utilization of supplier, construction resources data or knowledge on the application of blockchain technology. In realm it takes lot of time, but supply chain management and trust verification will be improving through years.

2 Literature Review

After agriculture, the construction sector is the second largest in the nation. It employs a sizable number of people and contributes significantly to the national economy. Mega-scale projects are now feasible because of the international collaborations towards the project management mechanisms and a variety of modern technology. Industry must overcome several obstacles in order to progress. Nonetheless, the sector continues to confront several significant obstacles, including as mass transit, housing, building that is resilient to natural disasters, and water management. A number of recent mega-projects have demonstrated that the industry is well-positioned for success. Out of all the professions in the nation, this is the second time that the field of civil engineering has returned home (Arghadeep Laskar and C. V. R. Murty, 2004).

The COVID-19 pandemic that struck China in the start of 2020 affected the country's economy, society, and people in a big way. It also hindered the growth of the construction sector. The impact on small and medium-sized construction companies that are not well-equipped to handle risk has been significant, making their survival dilemma worse. This study examines how COVID-19 has affected the expansion of small and medium-sized construction firms. This work develops a growth evaluation model based on factor analysis and generates a growth evaluation
index based on the features of small and medium-sized construction firms. The small and medium-sized enterprises board lists twenty-three construction companies, which are chosen as samples. The empirical analysis is conducted using the quarterly data from 2019 and 2020. The findings indicate that the epidemic has had a significant short-term impact on construction companies, as evidenced by the 16% decrease in the industry's total output value in the first quarter of 2020 compared to the same time the previous year. Over time, the epidemic has had little effect on the expansion of small and medium-sized construction companies. The growth score of businesses declined by just 1.95% in 2020's first quarter compared to the previous year, and it remained virtually unchanged in the next two quarters. The epidemic has had a significant short-term impact on profitability, capital expansion, and market expectations, but no effect on solvency, tangible resources, or intangible resources. Although it has a short-term effect, it helps to enhance an enterprise's capacity to operate. Lastly, specific recommendations are made at the policy and enterprise levels to both address the impact of the COVID-19 on small and medium-sized construction enterprises and to actualize their transformation and upgrading. The findings will promote decision-making for the healthy and orderly development of the follow-up construction sector and aid in understanding how the pandemic affected the expansion of construction companies (Wang W, Lin W, Bao Z, Dai X, Lin Q, 2022).

The Chinese construction market is expected to open up more and eventually join the global market following China's admission to the World Trade Organization (WTO). Contractors in different markets are likely to face distinct risks due to differences in historical and cultural contexts, social and economic systems, and other factors. This study uses an important assessment index and provides an importance evaluation of the numerous risks faced by Chinese contractors while contracting for projects in Chinese markets, based on questionnaires and case studies. Additionally, this report analyzes and compares the research findings with similar investigation data that are currently available. The current study uses the Cox–Stuart trend increase test method, and the findings show that as the risk event importance declines, the variance corresponding to the importance index value tends to grow. This trend indicates that individuals under investigation have a tendency to agree on risk events of greater significance. This study also uses Cronbach's Alpha Coefficient to assess the questionnaires' reliability. According to the research, irregular owner conduct and government department intervention in the construction industry are the biggest risks that Chinese contractors now face in domestic markets. International contractors now have more options to enter the Chinese construction business thanks to China's WTO entrance. Thus, the study findings presented in this work can offer useful information that will help foreign contractors better grasp the possible hazards present in the Chinese construction industry (Dongping Fang; Mingen Li; Patrick Sik-wah Fong; and Liyin Shen, 2004).

The Chinese construction industry is heading in the direction of digital transformation. This essay seeks to examine its current state, significant obstacles, and possible implications. A survey using questionnaires was conducted in order to accomplish this aim. The findings indicate that strategies for digital transformation have already been developed or made in 80% of the businesses where the industry experts are employed. Furthermore, the most popular digital technology was BIM software. Other significant obstacles included "data fragmentation," "lack of core technology," "weak allocation of digital infrastructure," "lack of technical personnel," and "lack of technical standards." Furthermore, it was believed that the majority of the impact of digital transformation would be felt at the project level in procurement management and the
business level in governance performance. These data can give academics and professionals a thorough grasp of how the Chinese construction industry is undergoing digital change. Additionally, they could assist legislators in creating suitable regulations to support the digital transformation (Zhang, N.; Ye, J.; Zhong, Y.; Chen, Z., 2023).

Sustainable development is negatively impacted by construction accidents, which pose a serious risk to the community. This essay provides an overview of China’s construction industry’s safety record for the previous ten years. A thorough examination of the fatal accidents that happened in 2018 reveals the patterns of spatiotemporal distribution and the characteristics of construction safety incidents. More often than not, managerial issues are to blame for building failures rather than technical issues. A case involving a significant mishap that occurred in 2018 when shield tunnel construction was underway in Foshan, Guangdong, is thoroughly examined. During the planning, geological study, design, and construction phase of the Foshan metro project, management concerns are examined using strategic environmental assessment (SEA). With a low overall SEA score, the SEA result demonstrates the extremely high safety risk. A safety construction management guideline for sustainability is suggested based on the analysis (Xin-Hui Zhou, Shui-Long Shen, Ye-Shuang Xu, and An-Nan Zhou, 2019).

Research on risk management techniques in the construction sector has increased. On the other hand, not much study has been done to comprehensively examine the general features of risk management from the viewpoints of different project participants. This study presents the results of an empirical industry survey conducted in China on the significance of project risks, the use of risk management strategies, the current state of the risk management system, and the perceived obstacles to risk management by the key project participants. Additionally examined were the risk management techniques used on the Three Gorges Project. According to the study, most project risks are typically concerning to project participants; the focus of the industry has shifted from risk transfer to risk reduction; existing risk management systems are insufficient to handle project risks; and the main obstacle to effective risk management is the absence of collaborative risk management mechanisms. Future research should be done to systematically enhance risk management in the construction industry using various strategies that enable players to share rewards fairly while managing risk effectively. In order to effectively utilize the corporate experience, personal knowledge, and judgment of all participants, such studies should also take into account the establishment of an open communication risk management process (Wenzhe Tang, Maoshan Qiang, Colin F. Duffield, David M. Young, and Youmei Lu, 2007).

Subcontractor selection, the use of blockchain technology in application design and deployment, the architecture and implementation procedures of blockchain-based Information Security Event Management Systems (ISEMS), the integration of next-generation platforms in the construction industry, the viability of public and private blockchain technologies in the sector, blockchain-based frameworks for automated term and condition enforcement, the adoption of environmentally sustainable blockchain-based digital twins (BCDT) in Construction Industry 4.0, and challenges unique to UK customs borders are all covered in the literature review.

The goal of Hartmann et al. (2010) is to investigate how important price and trust are in relation to one another when choosing a subcontractor. A unified blockchain as a service platform (uBaaS) that facilitates the creation and implementation of blockchain-based applications is presented by Lu et al. (2019). The architecture and implementation procedures of an Information
Security Event Management System (ISEMS) based on blockchain, which includes threat reporting, risk analysis, warning releases, and emergency responses, are described by Huang et al. (2020).

In their evaluation of the most recent research on building an integrated platform, Wang et al. (2020) offer an outlook on the integration of next-generation platforms in the construction industry. Through two industrial scenarios, Yang et al. (2020) investigates the viability of implementing both public and private blockchain technologies in the construction sector. Wei et al. (2020) uses a virtual machine agent model, which is working in distributed concept. This makes the process ease to investigate blockchain-based cloud data integrity protection techniques.

A distributed blockchain-based system is presented by Das et al. (2020) that permits the safe and transparent enforcement of terms and conditions pertaining to interim payments and the sharing of payment records at the project level. The goal of Teisserenc et al. (2021) is to create and use a theoretical framework. This framework will be deploying the blockchain-based digital twins (BCDT). The BCDT is digital platform environment friendly in the context of Construction Industry 4.0. In order to establish generalizability, Brookbanks et al. (2022) highlight issues unique to the UK customs boundaries and recommend replication in several scenarios.

So the work got clarity by the previous observations, studies by various scholars in the field, the study propose to make sure the supplier selection assessment shall be from the feedbacks of the construction projects, the words they are using at the time of work, the place the construction project store the supply workers, equipment’s and other resources as well. In addition, influential work by Piccininni et al. (2020) is mentioned, but the specific content is not provided. By synthesizing these literature studies, valuable insights and research limitations are obtained regarding subcontractor selection, the application of blockchain technology in the construction industry, digital platform integration, and sustainable development.

3 Methodology

The scope of the Trust in Overseas Construction Suppliers refers to the people deployed in the construction business, supply chain, communication methods, payment methods, conditions and feedbacks from the previous projects. geographic coverage, and timeframe within which the study will be conducted. The scope helps provide clarity and focus to the research objectives and ensures that the study remains manageable and feasible. Here is an example of the research scope:

Geographical Scope: The study will primarily focus on the Association of Southeast Asian Nations (ASEAN) region and select African countries. These regions have been chosen due to their significant involvement in international construction projects and the presence of Chinese construction companies.

Construction Industry Stakeholders: The research will involve various stakeholders in the construction industry, including construction companies, suppliers, project owners, regulatory bodies, and other relevant parties. Their perspectives and experiences will be considered to evaluate the trust challenges associated with overseas construction suppliers.
Trust Factors: The study will investigate key trust factors in the context of overseas construction suppliers, such as supplier reliability, quality control, delivery capabilities, compliance with regulations, and financial transparency. These factors will be examined to understand the underlying challenges and opportunities in building trust in these suppliers.

Figure 1. Internal and External Data Collection

A. K-Means Clustering

In this work, the proposed method, the K-means clustering algorithm is implemented to locate, identify the clusters of construction company and resources. Through data analysis, the K-means clustering method may determine an appropriate location for a supplier hub or distribution center, which can be used to determine the placement of a multilevel supplier distribution network and resources.

The K-means algorithm is popularly referred to as the Fast-Clustering Method. The fundamental aim of this clustering is to group each sample into the subclass that most closely resembles it. The distance between the sample data or resources and the centroid of the class, which is often the Euclidean distance, is typically used to measure similarity. The general description of the K-means algorithm’s principle is as follows:

N items are split into k classes, resulting in a high degree of similarity within classes and a low degree of similarity between classes.

The N data points are split into K classes, based on the distance information between them, and the specific flow is as follows:

- Randomly initialize k cluster centers.
- Determine the class that is closest to each data point by measuring the distance between each point and the class center, or distance($x_i, \mu_i$). This is known as the nearest-neighbor principle.
- Determine K new class centers, with each class center’s coordinates being the average coordinates of all the data points in the class.
- Re-assign each data point to the new centroid of each cluster by repeating the second step.
move to step 3 if there is a reassignment; else complete the process.

If you consider the Figure 2, the supplier team’s resources that are the data for the contractor, will be monitored by their Contractor’s regular report through Wechat or Line with location information and images of the current work status. Construction suppliers will be encouraged to have group to ensure the work in shared with rational under the agreement. So, every cluster will be doing their work and share to the Construction Company. The construction company will share that to the Organization which they committed to do the project. Widely used applications such as WeChat and Line can be used as communication channels. The companies can randomly call or ask any employee to send the current location or to attend the video call to make sure the workers are around the unit. Since the past records, the construction suppliers have undergone the selection strategy which is preliminary decided if a construction company wants to select employees who are available within 5 to 50 kms distance. In order to relate the distance and simple calculations, the employee’s geographic location position coordinates must be understood.

B. Sentiment Analysis

Language that is natural is the primary means by which people express their feelings and wishes. Concepts in this language are vague and undefined. In contrast, computers use numerical expressions to carry out their calculations and display their results. A communication gap exists between natural language and numerical language as a result. Utilizing NLP is one approach to resolving this communication issue. Our approach focuses specifically on identifying the tone of the experts’ contributions to the discussion. Natural language is more at ease when being used by experts to describe a subject.

Therefore, techniques that enable computers to operate utilizing natural language must be used. The majority of techniques described in the literature glean ideas and knowledge that the computer can utilize to manage the data.

Different techniques can be used to apply sentiment analysis. The bag-of-words strategy is one of the most well-liked. It is required to remove superfluous information, such as articles and prepositions, prior to producing the bags.
The unigrams must also be made simpler to prevent the placement of similar words with different prefixes or suffixes in the same bag. The following steps are involved in making the bags of words:

i. **Removal of additional information**: In this initial step, words, symbols, and numbers that are not necessary for learning from the text are removed. Removal of irrelevant content, such as articles or prepositions, results in comments that are immediately actionable. A bag of words is optimally generated as a result of this suppression. Additionally, it enhances the detection procedure.

ii. **Finding appropriate information**: Getting the pertinent information from the comments makes up this second stage. Utilizing the feedback from the previous phase and examining the word roots to identify unigrams are required for this stage. Similar to the previous stage, this one enables the bag of words to be clarified.

iii. **Separating the comments**: The ngram, where n ∈ N, is the portion of a comment. The comments are separated into words, or unigrams, if n = 1, and these words. It is possible to determine which language phrases correspond to each emotion by segmenting the comments into unigrams.

iv. **Building of the bags of words**: It is vital to define classes in accordance with the attitudes that we wish to track in order to generate the bags of words. We use four different classes in this document, and each unigram only appears once in the bag of words that corresponds to that class.

![Diagram](image)

Figure 3. Sequel of the Sentiment analysis path

Employee surveys, social media comments, performance reviews, analytics inside the domestic units, will be saved in different data type. They be unstructured data, so the extraction will be done in real time information processing. This data processing can assist businesses in analyzing employee sentiment. Using this knowledge will increase productivity, retention, and employee engagement.

In the Research the authors decided to Quantitative and Qualitative analysis as well.

- Quantitative analysis will take time and amount as major properties.
- Qualitative analysis will take communication and feedbacks as major properties.

The scope and methods outlined above provide a framework for conducting the research on building trust in overseas construction suppliers using a blockchain-based integrity assessment.
platform. However, it is important to note that the specific scope and methods may be further refined and adjusted during the research process based on practical considerations and emerging insights.

4 Results and Discussion

Identify and articulate the problem of evaluating the Assessment of overseas construction suppliers. The traditional methods of supplier evaluation, which rely on subjective assessments and limited information, often lead to inefficiencies, biases, and a lack of comprehensive insight into supplier integrity. This problem poses risks to construction projects and impedes the industry's growth and sustainability.

The Author has to come across the unavoidable areas of the construction company’s vertices. They are Limitations in suppliers, Bias on employee, Proof, Transparency, User friendly, Smart contracts, Accurate supplier assessment, Money transactions, Meetings, Demands and supply understanding and balancing, mitigate risks, Foster among the stakeholders, Potential challenges, technical barriers, Data privacy, Data abstraction, Selection process, Making policies, Reliable supplier and raw materials, decision making, global communication, addressing the language barriers, supporting employees and so on.

By addressing the problem statement and accomplishing these goals, the study hopes to improve supplier assessment procedures in the construction sector, offering a transparent and dependable framework for assessing foreign construction suppliers and encouraging integrity and trust in the sector.

The objectives of this research are

1. To assess the current challenges and limitations in geographical and evidence of payment system through available digital platforms. The author decided to use Wechat and like to do compare with the traditional manual entry system.

2. To improve communication, distribute payment methods to the supplier and evaluate the overall process in the construction industry.

3. To do sentimental analysis with the supplier reply and responses to the construction industry operators.

4. Try to use classification methods to identify the best supplier available by the terms of distance, budget, schedule as well.

The conceptual framework for the study on "Overseas Construction Suppliers Assessment based on Sentiment Analysis and Classification" provides a theoretical framework and structure for understanding and analyzing the factors influencing trust in overseas construction suppliers and the role of blockchain technology in enhancing trust. The conceptual framework helps guide the research process and provides a basis for organizing the study's findings. Here is an example of a conceptual framework for this study:
Trust and Assessment theory: Ground evidence is first parameter for the Trust. So, the author planned to use the messages sent and receive towards the supplier through mobile based application, such Wechat.

Factors Influencing Trust and Assessment: The conceptual framework will encompass various factors that influence trust in overseas construction suppliers. These factors may include supplier reputation, track record, past performance, communication effectiveness, relationship history, and transparency in business operations.

Assessment-building Mechanisms: The framework will identify and analyze the trust-building mechanisms that can be enabled by the integrity assessment platform. These may include supplier rating and feedback systems, verification of credentials and certifications, traceability of supply chain transactions, and real-time monitoring of supplier performance.

Impact and Outcomes: The conceptual framework will assess the impact and outcomes of data collected, terms of references available and how much they are finished are evaluated. The integrated assessment will be done by the terms of Time, Money, quality as well. Every property will be related to feedback system and the over feedback will give the data taken from the platform and given to the analysis part. This will determine the potential benefits, challenges, and implications for project outcomes, stakeholder relationships, and the overall construction industry ecosystem and Researchers. conclusion

The research advantages of the "Overseas Construction Suppliers Assessment based on Sentiment Analysis and Classification" are significant in revolutionizing the supplier Assessment or evaluation process and fostering trust within the construction industry. By leveraging the existing digital platforms, many construction companies can enhance their communication, saving the records, increasing the transparency, ensuring a clear and immutable payment record of supplier evaluations and global certifications. Stakeholders can access and verify evaluation data, fostering accountability and minimizing the risk of fraudulent activities.

Classifying the required data, properties can give more clarity in recommendation system of the suppliers and employees. The world is open and many constructions projects are taken by multinational companies and the companies are using the existing digital platforms to verify their contracts, supply, employees and payments systems by innovative integrated methods. Blockchain can support these methods for the future references, which can be handled by the companies itself.

5 Conclusions

The Global construction projects success rates depend on the combination and working hours of the contractors, suppliers, Engineers, Real estate people, Labors in the foreign soil. It is better assessed earlier and define the agreements before getting involved in there. Particularly the involvement in government-based projects requires more evidences to apply for the projects and auctions as well. The article proposed Sentimental analysis and clustering methods to assess the construction supplier to evaluate overseas construction suppliers from the textual data from various sources, such as customer reviews and supplier communications. The sentiment scores are used for supplier characteristics parameters and classification was used to monitor and maintenance. Despite some limitations, the proposed method provides valuable insights into
supplier performance and reputation. By leveraging sentiment analysis, construction companies can make informed decisions and enhance project outcomes through better supplier management. This research contributes to advancing sentiment analysis applications in the construction industry on a global scale.

Acknowledgement

This article is submitted to this conference based on the MOU signed between International College of Digital Innovation (ICDI), Chiang Mai University, Chiang Mai, Thailand and Christ (Deemed to be University) Pune, Lavasa Campus, India.

References

[30] ALOJAIL, A. A. M. Exploring the adoption of building information modelling technology using the TOE framework in the Saudi construction industry (Doctoral dissertation, RMIT University)