ENHANCING CYBER RISK MANAGEMENT IN THE CONSTRUCTION INDUSTRY

DISSERTATION

Submitted in Partial Fulfillment of
the Requirements for
the Degree of

DOCTOR OF PHILOSOPHY
(Civil Engineering)

at the

NEW YORK UNIVERSITY
TANDON SCHOOL OF ENGINEERING

by

Dongchi Yao

May 2024

ENHANCING CYBER RISK MANAGEMENT IN THE CONSTRUCTION INDUSTRY

DISSERTATION

Submitted in Partial Fulfillment of
the Requirements for
the Degree of

DOCTOR OF PHILOSOPHY
(Civil Engineering)

at the

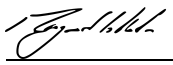
NEW YORK UNIVERSITY
TANDON SCHOOL OF ENGINEERING

by

Dongchi Yao

May 2024

Approved:



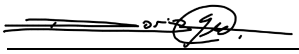
Department Chair Signature

05/06/2024

Date

Approved by the Guidance Committee:

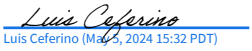
Major: Civil Engineering



Borja García de Soto
Assistant Professor
Tandon School of Engineering
05/06/2024
Date

Mike Wilkes
Adjunct Professor
Tandon School of Engineering
The Security Agency LLC
05/05/2024
Date

Semiha Ergan
Associate Professor
Tandon School of Engineering
05/06/2024
Date



Luis Ceferino
Assistant Professor
New York University
05/05/2024
Date

Vita

Dongchi Yao, born on April 13, 1994, in Anyue County, Ziyang City, Sichuan Province, China, completed his high school education at Sichuan Anyue High School in June 2012. He pursued both his bachelor's and master's degrees in Oil and Gas Storage and Transportation Engineering at Southwest Petroleum University, Chengdu City, Sichuan Province, earning his bachelor's degree in 2016 and his master's degree in 2019. During this time, Dongchi dedicated himself to the risk management of oil and gas pipelines and stations, which is essential for ensuring their integrity and operations. Dongchi's passion for scientific research was ignited by his master's degree supervisor, Prof. Xingyu Peng, an associate professor at the School of Petroleum Engineering at Southwest Petroleum University.

Dongchi began his Ph.D. journey in September 2020, focusing on cybersecurity risk management in the construction industry under the supervision of Prof. Borja García de Soto. Supported by the Global Ph.D. Student Fellowships from New York University Abu Dhabi, his research primarily employs machine learning, deep learning techniques, and probabilistic modeling for the quantitative analysis of cyber risks in construction projects. This research addresses a gap in the existing literature and facilitates a more objective and data-driven risk management process in practice. Dongchi is especially interested in developing and applying large language models (LLMs), among other sequence modeling techniques, to aid the risk management process. The long-term goal of his Ph.D. research is to create intelligent systems that can be embedded into construction companies' project management software. This intelligent system, deployed with advanced and automated risk management modules, would be capable of autonomously managing the entire cyber risk management process. Dongchi, who has actively published academic papers and presented at several international conferences, is dedicated to applying his passion for scientific research to solve real-world engineering problems. He remains committed to this path with the overarching aim of contributing to both the scientific community and the industry.

Acknowledgement

My Ph.D. journey has been a testament to the support, guidance, encouragement, and inspiration provided by many. At the forefront of this journey, I extend my deepest gratitude to Prof. Borja, my Ph.D. supervisor. In the vast sea of research, he stood as a guiding beacon, illuminating my path. His approach to problem-solving was not just instructive but also deeply supportive, teaching me to navigate research hurdles step by step. In moments of self-doubt, his comforting words and assurance were a balm, reminding me that research is a marathon, not a sprint. It demands a careful, pragmatic approach aimed at contributing to the scientific community. Under Prof. Borja's meticulous and patient mentorship, my research skills flourished, cementing my resolve to pursue a career in academia.

I am very thankful to Prof. Mike Wilkes for his guidance in cybersecurity, which ensured the precision and progress of my research on cyber risk management. My research also benefited greatly from the constructive feedback of my committee members, Prof. Semiha Ergan and Prof. Luis Ceferino, who gave me valuable suggestions and comments on improving this dissertation and propelled my work forward. My lab mates—Samuel A. Prieto Ayllón, Keyi Wu, Eyob Mengiste, Semih Sonkor, Xinghui Xu, and KC Lalropuia—have been instrumental in fostering a collaborative and enriching research environment. Our discussions and teamwork have been pivotal to my growth and the success of our collective projects.

On a personal note, my family's unwavering support and love have been my backbone. I carry my grandfather in my heart, hoping he finds peace and happiness in his next journey, after passing away on March 19, 2024. The music of the singer Kewei (Yisa) Yu has provided solace and companionship through the highs and lows of my Ph.D. life, becoming the soundtrack of my journey. As I step into my thirties, I reflect on the richness of experiences shared with teachers, friends, and family that have made this milestone truly memorable. With gratitude and anticipation, I embrace the path ahead, ready for new horizons and challenges.

ABSTRACT

ENHANCING CYBER RISK MANAGEMENT IN THE CONSTRUCTION INDUSTRY

by

Dongchi Yao

Advisor: Prof. Borja García de Soto, Ph.D.

Submitted in Partial Fulfillment of the Requirements for

the Degree of Doctor of Philosophy (Civil Engineering)

May 2024

The construction industry is in the midst of a digital transformation marked by Construction 4.0. While various digital technologies such as cloud computing, robotics, drones, and unmanned aerial vehicles have significantly boosted the efficiency of construction projects, the industry is lagging behind others in terms of cybersecurity. This is evidenced by a large number of cyber incidents that have been increasing over the years. However, studies on enhancing cybersecurity in this industry have been both scarce and fragmented, lacking a unified cybersecurity guideline and framework at the project level. To bridge this gap, this dissertation aims to systematically explore the integration of cybersecurity into construction projects. The outcomes from this dissertation will be beneficial in enhancing cyber risk management in the construction industry.

This dissertation contains four research projects, each having a unique contribution to enhancing cyber risk management in construction projects. (1) It identifies potential cybersecurity research topics by employing the Latent Dirichlet Allocation topic modeling

technique to analyze a large, meticulously collected and screened text corpus. This guides future research and industry practices, facilitating unified research progress in this area. One of the most important topics identified is cyber risk management, setting the stage for the remainder of the dissertation. (2) It identifies various cyber risks across construction phases by developing a construction cybersecurity-dedicated language model, which is enhanced by Supervised Fine-tuning and Reinforcement Learning from Human Feedback training techniques. The resulting prioritized cyber risk checklist serves as a new benchmark for the industry. Additionally, the developed language model shows great potential in serving as an intelligent cybersecurity consultant for industry-wide applications. (3) It identifies risk factors associated with construction projects that lead to common cyber risks, realized through a systematic methodological process of literature review, questionnaire surveys, and expert consultations. The identified risk factors, each divided into specific scales, lay the groundwork for quantitative risk assessments. (4) It develops a machine learning-centric approach for a more quantitative cyber risk assessment at the project level, providing a tool for project managers for reliable cyber risk decision-making and efficient risk mitigation efforts.

This dissertation represents one of the pioneering systematic studies on cybersecurity within the construction industry, aiming primarily to enhance cyber risk management for construction projects. It contributes to the existing body of knowledge by proposing and adapting methodologies and frameworks for cyber risk management, fostering interdisciplinary research among the fields of artificial intelligence, cybersecurity, and construction management. Additionally, it offers practical tools for construction practitioners to enhance cybersecurity, such as the identified research topics, the cyber risk checklist, and the developed machine learning model for quantitative risk assessment. The long-term goal is to create a cybersecurity digital platform tailored for construction projects, integrated into construction companies' software as mobile or web applications. This platform will be capable of performing various advanced functions and providing targeted cybersecurity answers, suggestions, and solutions, catering to both specific construction projects and general cybersecurity inquiries.

Table of Contents

List of Figures

List of Tables

Chapter 1
Introduction

The construction industry is undergoing a digital revolution, often referred to as Construction 4.0, where various digital technologies are integrated into all areas of the construction industry [1]. This integration aims to create a smarter work environment and foster efficiency and innovation in construction projects. The concept of Construction 4.0 is adapted from the broader umbrella of the Industry 4.0 revolution, wherein interconnectivity, automation, machine learning (ML), and real-time data are fully explored and incorporated to transform the manufacturing process. In Construction 4.0, various digital tools and technologies are utilized in different phases of construction projects to improve performance, reduce costs, and enhance safety [2], [3]. For example, in the planning phase of a construction project, Building Information Modeling (BIM) software enables architects, engineers, and specialists in the construction industry to coordinate building design and visualize the physical and functional characteristics of design objects. This benefits decision-making and coordination among stakeholders, also allowing for the identification of problems ahead of time, thus preventing errors to a large extent. During the construction phase, drones and Unmanned Aerial Vehicles (UAVs) are increasingly

used for site patrolling and monitoring, providing real-time data and images to track progress and inspect work to ensure compliance with safety protocols. Additionally, wearable technologies such as smart helmets and exoskeletons can monitor the health status of workers and provide physical support, improving their safety and productivity. In the operation and maintenance phase, Internet of Things (IoT) equipment and sensors can be embedded into buildings to monitor the structural health, energy consumption, and environmental conditions, contributing to smarter, sustainable operation of construction projects. Construction 4.0 represents a significant transformation of the construction industry, signifying higher digitalization, efficiency, and sustainability, where various digital tools are utilized in different phases of construction projects to achieve these objectives of higher digitalization, efficiency, and sustainability [1], [2].

The digitalization of the construction industry, while offering numerous benefits, also introduces cybersecurity challenges. As construction firms increasingly adopt digital tools, they become more vulnerable to cyber attacks. These technologies rely heavily on data and connectivity, making them prime targets for cyberattacks [4]. For example, one of the primary cybersecurity concerns is the risk of data breaches. Sensitive information, including architectural designs, structural plans, and personal data of workers, could be exposed or stolen, leading to financial losses, legal repercussions, and damage to reputation. Moreover, the interconnectedness of devices and systems increases the risk of malware or ransomware attacks, which could cause negative impact on operations, leading to significant delays and cost overruns. Another issue is the potential for sabotage through cyber-physical attacks, either from insider attacks or external attacks. Hackers could potentially take control of connected machinery and equipment, causing physical damage or endangering lives. Furthermore, the lack of standardized cybersecurity protocols across the industry exacerbates these risks, as does the insufficient cybersecurity awareness among construction professionals. As the industry moves further towards Construction 4.0, addressing these cybersecurity challenges becomes imperative to protect assets, maintain operations, and ensure the safety of workers and the public [4], [5], [6].

However, the construction industry is notably behind other sectors in cybersecurity awareness and preparedness, a fact made clear by looking into real-world incidents. Cyber incidents within the industry have increased dramatically, from nearly 10 in 2013 to almost 520 in 2022, marking an sharp increase of 5100% [7]. Table 1.1 showcases several cyber incidents from the past five years within the broader construction sector, highlighting the critical need for a comprehensive and effective cybersecurity strategy to protect construction projects.

Table 1.1. Cyber incident examples in the last five years

Year	Victim	Country	Attack Nature	Consequence	Reference
2018	Ingérop	France	Data breach/theft	65 GB of data related to nuclear power plants stolen, over 11,000 files from a dozen projects accessed, and personal details of more than 1,000 employees compromised.	[8]
2019	Bird Construction	Canada	Data breach	MAZE claims to have stolen 60 GB of data from the company, which landed 48 contracts worth $406 million with Canada's Department of National Defense between 2006 and 2015.	[9]
2019	Marous Brothers Construction	United States	Email fraud	When the major renovation project of the St. Ambrose Catholic Church was about to be completed, there were financial issues, and the $1.7 million payment that should have been made to the contractor was not received. This situation stems from the theft of email communication content,	[10]
2020	Bouygues Construction	United Kingdom	Ransomware attack	Forced the company to shut down its systems worldwide due to a ransomware attack at the end of January 2020.	[11]
2020	Bam Construct & Interserve	United Kingdom	Ransomware attack	Bam Construct faced ransomware encrypting files for ransom, and Interserve suffered a data breach potentially affecting 100,000 employees, with suspicions of targeting its anti-pandemic efforts.	[12]
2021	Oldsmar Water Treatment Plant	United States	Unauthorized access to control systems	Control system accessed, and the amount of sodium hydroxide in the water increased to dangerous levels; discovered before any harm to the public.	[13]

Table 1.1 (continued)

Year	Victim	Country	Attack Nature	Consequence	Reference
2021	Colonial Pipeline	United States	Cyberattack using a compromised password	Networks accessed using a compromised password, shutting off the largest fuel pipeline in the US until a $4.4 million ransom was paid, marking its first complete shutdown in 57 years.	[14]
2022	Interserve Group Ltd	United Kingdom	Data breach	Issued a £4.4 million fine for failing to secure the personal information of its staff, constituting a breach of data protection law.	[15]
2022	The Knauf Group	Germany	Ransomware attack	Resulted in emails and product-ordering software being taken offline, disrupting customer communications.	[16]
2023	Huntington Ingalls Industries	United States	Unauthorized access to data	Unauthorized access to sensitive consumer data, including personal, financial, and medical information, reported.	[17]
2023	Simpson Manufacturing Co., Inc.	United States	Malicious activity	IT infrastructure and applications disrupted, with steps taken to stop and remediate the activity.	[18]

However, as highlighted in Section 2.2, there is a significant gap in cybersecurity studies in the construction industry, highlighting the scarcity and fragmentation of existing literature. It spans from broad industry discussions to specific topics like project threat modeling and equipment cybersecurity assessment. The scarcity and fragmentation impede the development of cohesive cybersecurity strategies for construction projects. To bridge this gap, this dissertation first identifies a range of potential research topics aimed at informing future directions and enabling researchers to concentrate on areas that facilitate unified advancement. It then prioritizes and explores the cyber risk management process, tailored specifically for construction projects, which is the most important topic identified.

1.1 Research questions

This dissertation aims to answer four research questions, each of which will contribute to enhancing cyber risk management in the construction industry.

(1) What potential topics exist for cybersecurity studies in the construction industry? This question, studied in Chapter 3, is addressed by implementing the Latent Dirichlet Allocation (LDA) topic modeling technique. This study involves collecting a variety of text sources to formulate a text corpus, applying the LDA topic modeling technique, and discussing emergent topics that could guide future academic research and industry practice. Through the study, the current landscape of cybersecurity within the construction sector is comprehensively explored and preliminarily analyzed.

(2) Which cyber risks may be encountered during each phase of a construction project? This question, studied in Chapter 4, is addressed through a language modeling approach. This study utilizes the curated text corpus in Chapter 3 to identify cyber risks, which are organized into different phases of the construction project lifecycle. The formulated cyber risk list can serve as a new benchmark for the industry, which can help project managers formulate preventive measures.

(3) What specific risk factors could lead to common cyber risks in construction projects? This question, studied in Chapter 5, is addressed through a systematic process of literature review, questionnaire surveys and expert consultation. This study explores the root causes of cyber risks, seeking to identify relevant factors involving various aspects of construction projects such as project characteristics, project structure, information technology (IT) factors, operational technology (OT) factors, and human/management factors. The identified risk factors establish a foundation for a more quantitative cyber risk assessment of construction projects. They also inform construction practitioners on which project characteristics might be related to cybersecurity, thus warranting closer attention.

(4) Is it possible to establish a model for a more quantitative and efficient cyber risk assessment of construction projects? This question, studied in Chapter 6, is addressed by developing an ML-centric approach. This approach, being data-driven, can quantify the degrees of various cyber risks at the project level, and can provide project managers with insights on whether to take actions to mitigate risks and, if

so, which risk factors should be prioritized. Additionally, the study highlights the top risk factors of general importance across broad construction projects.

1.2 Research contributions

This dissertation, to the best of the author's knowledge, represents one of the pioneering systematic studies on cybersecurity within the construction industry, aiming primarily to enhance cyber risk management for construction projects. It contributes to the existing body of knowledge by proposing or adapting methodologies and frameworks for cyber risk management and by fostering interdisciplinary research among the fields of AI, cybersecurity, and construction. Additionally, it offers practical tools for construction practitioners to manage cybersecurity. The dissertation contains four main studies: topic modeling, project risk identification, risk factor identification, and project risk assessment, with each study making its own contributions.

(1) Chapter 3 utilizes the LDA topic modeling technique to analyze abundant text, aiming to identify emerging topics in cybersecurity research. The academic contributions of this study are dual: Firstly, it addresses research gaps and guides future directions by identifying emerging research topics within the construction industry and cybersecurity, directing researchers towards promising fields where their efforts could lead to significant and unified advancements. Secondly, at a time when LLMs are becoming increasingly prevalent, this study contributes a unique corpus specifically tailored to construction and cybersecurity. This paves the way for advanced language models capable of specialized tasks such as question-answering and automatic risk identification in these sectors. This work aligns with the digital transformation trends in construction, marking progress towards Construction 4.0. The practical contributions of this study include: Firstly, informing industry practices. By identifying emerging topics, practitioners in the construction industry can leverage the insights to adopt or enhance cybersecurity measures; they can develop proactive strategies for risk management and resilience. Secondly, guiding policy and standards development. The findings of this study can aid policymakers

and standards organizations in crafting regulations and guidelines tailored to the evolving cybersecurity needs within the construction sector. This is particularly relevant as digital technologies become increasingly embedded in construction processes.

(2) Chapter 4 develops a language model to utilize the established corpus constructed in Chapter 3 to identify 36 common cyber risks across different phases of construction projects. The academic contributions of this study include: Firstly, the novel development and application of techniques dedicated to construction cybersecurity, fostering interdisciplinary research in AI, cybersecurity, and construction management. Secondly, it bridges the gap in the comprehensive recognition of cyber risks across project phases, setting a new benchmark with a cyber risk checklist. Thirdly, it introduces the possibility of employing the developed framework for risk identification in other industries. Practically, the cyber risk checklist provides several advantages: Firstly, the curated cyber risk checklist can be used for a variety of construction projects, especially in aiding project managers to assess their cybersecurity status and develop proactive and preventive cybersecurity measures against the prioritized risks. Secondly, risk analysts can use this checklist for thorough risk analyses on specific construction projects, focusing their efforts on the most critical risks first. Lastly, with the checklist being regularly updated, various groups, including IT and cybersecurity teams, stakeholders, and general personnel of construction companies, can access the most current information on the cybersecurity landscape, ensuring they stay informed about prevailing trends.

(3) Chapter 5 utilizes literature review, questionnaire surveys and expert consultation to identify risk factors that could lead to the cyber risks outlined in Chapter 4. The academic contributions of this study are twofold: Firstly, it presents a set of risk factors that serve as a foundation for future quantitative cyber risk assessments, which is absent in existing literature; and secondly, it offers a systematic methodology framework for identifying risk factors associated with various risks. On the practical side, the study provides a list of risk factors that can raise awareness among project managers and construction practitioners about the types of project characteristics related to cybersecurity status, which should

receive more attention. Consequently, this enables a deeper understanding of the cybersecurity status of construction projects and their underlying causes.

(4) Chapter 6 develops an ML-centric approach designed to quantitatively assess the degree of cyber risk in construction projects. The academic contributions of this study are twofold: Firstly, it develops a quantitative cyber risk assessment approach, addressing a significant gap in existing literature by proposing a framework that enables more quantitative assessments of cyber risks at the project level. Secondly, the study pioneers interdisciplinary research by utilizing ML techniques at the intersection of cybersecurity, ML, and construction management. The practical contributions of this study are also twofold: Firstly, it offers a dynamic risk assessment tool. The developed approach, including the trained models and adapted feature analysis & greedy optimization methods, serves as a ready-to-use tool for dynamically assessing the cyber risk status throughout the progression of construction projects. This tool empowers project managers to make informed decisions about whether and to what extent actions should be taken to mitigate high-contribution risk factors, thus facilitating a more efficient risk reduction and prevention. Secondly, the study highlights the risk factors of general importance across a wide range of construction projects. This aids project managers in recognizing which risk factors should generally be prioritized and addressed, particularly in situations where project-specific risk assessments are not feasible.

1.3 Dissertation outline

The rest of the dissertation is organized as follows:

Chapter 2 presents the fundamentals and literature review. It begins by introducing the fundamentals of risk management, followed by their discussion within the construction industry context. Next, we review relevant cybersecurity studies, categorized into general discussions, review papers, and specific solutions. We then offer a detailed overview of ML, a crucial technique utilized in this dissertation, encompassing supervised learning, deep learning, loss functions, and optimization techniques.

Chapter 3 identifies emerging topics in the construction cybersecurity area. We first outline the steps taken to collect a large amount of text to formulate a corpus, including a series of text processing steps to ensure the corpus contains high-quality sentences. This developed corpus is then used to apply the LDA topic modeling technique to identify emerging cybersecurity research topics in the construction industry. These topics are discussed in-depth, and corresponding actions to take are suggested. The identified topic of risk management sets the stage for the rest of the dissertation.

Chapter 4 identifies cyber risks in construction projects. We develop a language model trained on the corpus developed in Chapter 3. The base model for the language model is chosen from three candidates: GPT-2, BERT model with a language modeling head, and T5 model with a language modeling head. The SFT and RLHF techniques are then performed to enhance the performance of the language model in understanding cybersecurity content and generating answers to our formulated cyber risk identification questions. The developed model is then utilized to identify cyber risks across project phases, with the likelihood of the identified risks provided. The language model and the identified risks are evaluated using a mix of quantitative and qualitative methods and sources.

Chapter 5 identifies risk factors that could contribute to the identified cyber risks within construction projects. The methodology involves iterative steps of literature review, questionnaire survey, and expert interview. The identified risk factors, closely relevant to construction project characteristics, are grouped into five categories: Overall Information of the Project, Project Structure, IT Factors, OT Factors, and Management and Human Factors. These factors capture both general IT vulnerabilities and specific vulnerabilities associated with construction projects. Each risk factor is assigned distinct scales. Subsequently, it discusses the advantages of these identified risk factors for conducting a more quantitative assessment of a wide range of cyber risks.

Chapter 6 proposes an ML-centric approach for the cyber risk assessment of construction projects, focusing on common cyber risks identified in Chapter 4, with the risk factors identified in Chapter 5 as ML inputs. The Monte Carlo simulation is utilized to generate

synthetic data for model training, which is labeled by an ensemble labeling method. Then, each cyber risk is tested with a variety of ML models to examine their capability to capture the nonlinearity in the dataset and the labels. The best model architecture for each risk, along with the optimal labeling weights in the ensemble labeling method, is chosen. Finally, a case study involving a real construction project is performed to demonstrate the effectiveness and applicability of the developed approach.

Chapter 7 summarizes the work conducted in this dissertation and outlines the research findings. Additionally, it points out directions for future research.

Chapter 2

Background and Fundamentals

In this chapter, we first introduce the fundamentals of risk management and then discuss them within the context of the construction industry. Subsequently, we review cybersecurity studies pertinent to construction, categorizing them into general discussions, review papers, and specific solutions. Following that, we provide a comprehensive exposition of machine learning, an important technique employed in this dissertation, covering supervised learning, deep learning, loss functions, and optimization techniques.

2.1 Risk management in construction projects

Risk management is a strategic and structured methodology employed to identify, assess, and manage potential risks occurring within projects or operational contexts. This process implements proactive measures to minimize detrimental consequences while concurrently recognizing and exploiting opportunities throughout the project lifecycle. Originating in the financial sector in the 20th century, risk management initially focused on mitigating financial risks using insurance instruments. However, over the decades, it has undergone substantial evolution and expanded into diverse areas of application [19].

Nowadays, risk management has evolved into a holistic framework that encompasses a

wide range of strategic and operational risk domains, playing a critical role in safeguarding the resilience and success of projects and organizations. This evolution reflects the increasing complexity of the contemporary business landscape, characterized by rapid technological advancements, expanding globalization, and stringent regulatory regimes. As a result, contemporary risk management practices have become more intricate and multidimensional, integrating advanced analytical tools, detailed risk assessment models, and comprehensive contingency plans. These are designed to tackle a broad spectrum of potential threats, including cyberattacks, supply chain disruptions, ecological disasters, and challenges in regulatory compliance, all of which underscore the importance of fostering resilience and adaptability [20]. Moreover, the evolution of risk management indicates a shift toward a more holistic understanding of risk, highlighting the necessity of adopting a panoramic view of organizational risks. This approach recognizes the unpredictable and multifaceted nature of risks, suggesting that they can emerge from various directions. Consequently, aligning the risk management process with the broader corporate strategy becomes vital, involving strategic synchronization to ensure that organizations remain not just responsive but also anticipatory in navigating the complex challenges of today's business environment. In essence, risk management has transformed from a mere protective measure into a strategic tool that empowers organizations to pursue and achieve their ambitious goals with increased confidence and security [19], [20], [21].

Risk management has progressed from straightforward risk detection to a sophisticated, all-encompassing approach permeating every phase of a project, encompassing risk recognition, evaluation, and reaction. Contemporary risk management revolves around persistent scrutiny and agile responses, underlining the fluidity of risks amid fluctuating project scopes, technological innovations, and market dynamics. Its pivotal components consist of [20], [22]:

- Risk identification: It is the preliminary stage that systematically pinpoints and records potential risks affecting the project. This phase encompasses a wide-ranging scrutiny of both intrinsic and extrinsic factors, leveraging methods such as

brainstorming sessions, expert evaluations, and risk catalogs to reveal prospective adversities and advantages alike.

- Risk assessment: Once risks have been identified, they undergo an assessment process to determine their probability and potential impact on the project. This dual-phase assessment incorporates both qualitative and quantitative dimensions. The qualitative assessment enhances risk management by ranking risks based on their severity and likelihood. Conversely, the quantitative assessment assigns numerical metrics to these risks, often leveraging statistical methodologies and models to gauge probable economic and temporal outcomes. This pivotal step determines the significance of risks and guides the efficient allocation of resources.
- Risk response: After assessing risks, launching suitable responses is important. This involves determining the most effective way to handle each key risk, which might involve one or more of the following strategies: avoidance, transfer, mitigation, or acceptance. Avoidance tactics could entail adjusting the project blueprint to evade the risk entirely; transfer strategies might include subcontracting tasks or securing insurance coverage; mitigation efforts focus on reducing either the probability or the impact of a risk, whereas acceptance implies acknowledging a risk without intervention, typically when the costs of other strategies exceed the projected impact. Each response strategy is meticulously calibrated to the unique attributes and requirements of the project, thus ensuring that risk management aligns seamlessly with the broader project goals and stakeholders' expectations.

The construction sector, characterized by its intricate web of stakeholders, logistical complexities, and environmental unpredictabilities, presents a multitude of scenarios ripe for the application of risk management. Construction projects are particularly vulnerable to an array of perils such as budget overruns, schedule delays, safety hazards, and regulatory compliance issues. In this context, robust risk management goes beyond mere adversity evasion; it facilitates seamless project execution through meticulous strategic

planning and continuous monitoring. Given the diversity of threats – financial, environmental, time-related, safety-centric, among others – construction projects necessitate a multidimensional risk management strategy. This requires the formulation of a comprehensive plan, the establishment of clear and effective communication channels, and the strategic adoption of technology. To execute this effectively, a profound understanding of the project ecosystem is essential, ensuring unimpeded stakeholder communication and the deployment of advanced analytical tools for risk evaluation and mitigation. Strategically, this involves establishing a thorough risk register, deeply integrating risk management into current project management practices, and leveraging simulation and modeling tools to anticipate and prepare for potential obstacles. This holistic approach not only identifies and mitigates emergent risks but also assists in capitalizing on arising opportunities, thereby enhancing the efficiency and effectiveness of project governance [20], [22], [23].

2.2 Cybersecurity studies in construction

Several general cybersecurity tools and frameworks have been developed for cybersecurity management. These include the National Institute of Standards and Technology (NIST) Cybersecurity Framework [24], General Data Protection Regulation (GDPR) [25], ISO/IEC 27000 series [26], Center for Internet Security (CIS) Controls [27], and the Common Vulnerability Scoring System (CVSS) [28]. Table 2.1 provides a comprehensive comparison of these tools and frameworks. While these general tools and frameworks provide general guidance for companies assessing their cybersecurity posture, they are not specifically tailored to the construction industry. This industry faces unique cybersecurity vulnerabilities and challenges, with the primary goal of ensuring the successful delivery of projects. However, these tools and frameworks do not provide an adaptable method to assess and manage project-specific risks across different project phases, such as assessing project-level risks using specific metrics. Additionally, these tools and frameworks typically require a manual approach involving domain experts, which can be time-

consuming and may not keep pace with the rapidly evolving cybersecurity landscape in construction. Therefore, there is a clear need for the development of tailored, easy-to-adopt, and time-efficient cyber risk management methods specific to the construction industry. However, only a few studies have been developed for the construction industry within the last few years. These studies can be classified into three categories: general discussions, review papers, and specific solutions [29].

2.2.1 General discussions

General discussions have emphasized the importance of research in bolstering the construction industry's defense against cybersecurity threats. Key areas include enhancing training to counteract social engineering, addressing 3D printing risks, understanding Construction 4.0's impact, mitigating digital vulnerabilities, and creating tailored cybersecurity frameworks for better management and standardization. Bello and Maurushat [30] emphasized the need for enhanced training to mitigate ransomware attacks initiated through social engineering. El-Sayegh et al. [31] discussed the benefits and cybersecurity risks associated with 3D printing in construction, highlighting material and regulatory challenges. García de Soto et al. [3] explored the workforce and organizational implications of Construction 4.0, predicting job evolution and the coexistence of conventional and robotic technologies. Mantha and García de Soto [32] focused on discussing the cybersecurity vulnerabilities introduced by digitalization, proposing a framework for risk identification. Turk et al. [33] called for a construction-specific cybersecurity solution, underscoring the industry’s unique vulnerabilities and the need for tailored solutions. Turk et al. [34] aimed to develop a cybersecurity framework for the construction industry to enhance the understanding, management, and standardization of cybersecurity issues in the industry, as well as to guide future research directions.

Table 2.1. Comparison among general tools and frameworks

Standards/tools	Purpose	Key Features	Process	Risk Assessment	Industries	Limitation to the construction industry	Expertise required
NIST Cybersecurity Framework [24]	Provides guidance for organizations to assess, manage, and improve cybersecurity posture, fostering resilience and risk management	Includes guidelines for identifying, protecting, detecting, responding, and recovering from cyber threats	Mostly Manual	Provides qualitative risk assessment guideline for companies	General	Provides a high-level approach. Needs tailoring to address construction-specific threats (e.g., drone misuse, BIM security)	Cybersecurity Expert
GDPR [25]	Protects personal data of EU citizens and regulates data handling and processing	Includes guidelines about personal data protection, individual rights, obligations for data handlers, and enforcement mechanisms for compliance	Mostly Manual	Does not provide a risk assessment framework	Sectors within Europe Union	Focuses on personal data protection, not project cybersecurity directly. May apply to construction companies handling EU citizen employee or client data	Data protection officer or a legal expert
ISO/IEC 27000 series [26]	Provides standards for implementing, maintaining, and improving information security management systems (ISMS) for organizations	Includes standards for implementing, maintaining, and improving information security management systems (ISMS) within organizations	Mostly Manual	Provides guidelines mainly for qualitative risk assessment, but allowing for quantitative one	General	Requires adaptation to construction project lifecycles and specific digital tools used (e.g., BIM)	Information Security Expert
CIS [27]	Provides a set of actions for cyber defense that provide specific and actionable ways to thwart the most pervasive attacks	Includes a set of critical security controls ranging from inventory and control of hardware and software assets to data protection and incident response	Mostly Manual	Provides a risk assessment framework combining qualitative and some quantitative elements	General	Offers best practices but may need industry-specific supplements for construction cybersecurity	Cybersecurity Expert
CVSS [28]	Provides a way to capture the principal characteristics of a vulnerability and produce a numerical score reflecting its severity	Provides standardized vulnerability scores. Enables IT managers to prioritize responses and resources according to threat severity	Semi-automatic	Provides a standardized measure of vulnerability severity regarding assets	General	Useful for assessing vulnerabilities in assets used in construction. Does not address broader project cybersecurity risks	Cybersecurity Expert

2.2.2 Review works

Review papers have collectively explored cybersecurity challenges in critical infrastructure, adapting cybersecurity frameworks for construction, ML in additive manufacturing, and gaps in construction cybersecurity research. They emphasized enhancing awareness, custom solutions, and the need for future research on technology-mediated countermeasures. Pärn and Edwards [35] presented a comprehensive review of cyber-threats, including industrial espionage, cyber-crime, and politically motivated cyber-interventions, that pose risks to the critical infrastructure vital for national wealth, health, safety, and welfare. Sonkor and García de Soto [36] reviewed existing cybersecurity frameworks and standards, then customized the most suitable one for the construction industry, specifically for sites using autonomous equipment. This review aimed to enhance cybersecurity awareness, advocate for tailored cyber assessments, and foster a security-minded culture within the construction sector. Goh et al. [37] comprehensively reviewed various machine-learning techniques used in additive manufacturing. The review described various ML techniques applied to additive manufacturing, including 3D design, material tuning, process optimization, monitoring, cloud services, cybersecurity, and potential applications in biomedicine, tissue engineering, and architecture. Pargoo and Ilbeigi [29] conducted a systematic scoping review of cybersecurity in the construction industry, identified gaps in current research, particularly in digital systems like robots and prefabrication platforms, and suggested future directions focusing on intrusion detection, response strategies, and exploring beyond Blockchain for technological countermeasures. Das et al. [38] reviewed information security requirements for collaborative BIM platforms and identified seven BIM security components. They proposed two conceptual frameworks: an encryption strategy for secure BIM storage and distribution, and a blockchain framework for recording changes.

2.2.3 Technical solutions

As for technical solutions, they can be generally categorized into Encryption,

Authentication and Access Control, Data Decentralization, Intrusion Detection and Prevention, Attack Detection, Physical Protection, and Risk Assessment [29].

As for Encryption, Authentication and Access Control, Ackerson et al. [39] explored the effectiveness of Recurrent Neural Networks (RNNs) in biometric authentication, expression recognition, anomaly detection, and aircraft applications, analyzing the advantages and disadvantages of various models. Patel and Patel [40] aimed to improve the privacy of architectural and engineering projects by employing Hybrid-K anonymization techniques to protect personal and sensitive data from cyber attacks in data mining. Das et al. [38] identified security components for collaborative BIM platforms, proposing encryption and access control as key strategies. It also explored frameworks leveraging encryption and blockchain to enhance data confidentiality and sharing security.

As for data decentralization, Shemov and García de Soto [41] explored the potential of blockchain to improve the efficiency and security of the construction supply chain. Das et al. [38] recommended blockchain to ensure a tamper-proof ledger for recording BIM changes, enhancing security and trust in the non-trusting environment of construction projects by providing a decentralized and transparent mechanism for managing and safeguarding data. Nawari and Ravindran [42] proposed integrating blockchain technology with BIM processes to enhance post-disaster recovery efficiency in the Architecture, Engineering, and Construction (AEC) industry. They suggested blockchain could automate building permit processes using Smart Contracts and Hyperledger Fabric, improving framework efficiency in post-disaster events. In [43], the authors further evaluated blockchain technology's applications in the built environment, exploring its integration with BIM processes, enhancing network security, data management, and automating construction design reviews through smart contracts and Hyperledger Fabric. Shi et al. [44] demonstrated the application and proof-of-concept of blockchain for Peer-to-Peer (P2P) collaboration in the construction industry. Tao et al. [45] presented a distributed BIM design-sharing framework using blockchain and InterPlanetary File System (IPFS), including performance testing.

As for Intrusion Detection & Prevention, Attack Detection, and Physical Protection, Sheikh et al. [46] proposed a Boolean Identification Strategy (BIS) to detect electrical faults and cyber attacks in Building Management Systems (BMS). It then employed ML classifiers to analyze and categorize detected attacks, enhancing intrusion detection and prevention capabilities for smart buildings. Patel and Patel [40] proposed the use of Hybrid-k anonymity technique to protect sensitive information in the AEC industry. By modifying and anonymizing data, it aims to enhance privacy and prevent cyber attacks from accessing sensitive project data. Pan et al. [47] proposed a context-aware intrusion detection system (IDS) for Building Automation and Control networks. It utilized anomaly-based behavior analysis to detect cyber-attacks and functional failures, classify attacks, and automatically initiate protective countermeasures. This system was demonstrated through evaluation in a Smart Building testbed. Skandhakumar et al. [48] proposed a novel approach using BIM graphs for access control applications in buildings. By representing building information models using graph theory, it enhances the understanding of building layouts and access points, which can contribute to physical protection measures.

As for Risk Assessment, Mantha and García de Soto [32] developed a framework to identify cybersecurity risks in Construction 4.0 and assessed vulnerability across project participants and entities using an agent-based model, aiding in comprehensive risk assessment. Mohamed Shibly and García de Soto [49] aimed to enhance cybersecurity awareness in the AEC industry by developing a preliminary threat modeling approach tailored for construction projects, focusing on cyber-physical systems. It also sought to demonstrate its practicality through the example of a robotic arm system used for 3D printing, identifying potential threats and promoting secure technology implementation. Some works specifically use the Common Vulnerability Scoring System (CVSS) for risk assessment; for example, Mantha et al. [50] evaluated the security vulnerabilities within the construction industry (Construction 4.0) using the CVSS. It aimed to quantify the vulnerability of key project participants like owners, contractors, and workers, contributing to a broader understanding of construction network security. Mantha and García de Soto

[51] assessed cybersecurity vulnerabilities in the AEC industry using the CVSS, implementing a novel methodology to evaluate and improve construction network security. It highlighted the practical and social implications of cybersecurity in digitized construction, aiming to enhance industry-wide security awareness.

2.2.4 Summary

Based on the comprehensive overview of cybersecurity studies within the construction industry, it is evident that while there has been some progress in addressing cybersecurity concerns, these efforts are relatively nascent compared to other sectors. The literature review reveals a multi-dimensional exploration of cybersecurity, ranging from general discussions to review papers and specific technical solutions. However, compared to cybersecurity studies in other industries or other risk studies in the construction industry, cybersecurity research specific to construction remains scarce. This indicates a significant gap in cybersecurity awareness and preparedness within the sector. This scarcity is further exacerbated by the fragmented nature of existing research. Although there are studies covering a broad range of topics—from the importance of training to counteract social engineering attacks and addressing risks associated with 3D printing, to the implications of Construction 4.0, and the development of industry-specific cybersecurity frameworks and technical solutions like encryption, blockchain applications, and intrusion detection systems—these efforts appear to be isolated. They lack a unified focus on the risk management process, which is crucial for developing comprehensive cybersecurity strategies for construction projects.

2.3 Machine Learning

This section introduces the fundamental concepts of ML that are primarily utilized in this dissertation. It begins with the basics of supervised ML algorithms, which form the foundation for further discussion. Next, we delve into the fundamentals of neural-network-based DL, highlighting its significance in contemporary ML applications. Among the

various architectures within DL, special attention is given to the state-of-the-art Transformer architecture, noted for its effectiveness in handling sequential data. Additionally, this section explores the various loss functions employed in this dissertation, outlining their roles in model training. Finally, we discuss optimization techniques, with a particular focus on the AdamW optimizer, chosen for its improvements in handling weight decay for DL models.

2.3.1 Supervised learning

ML can be broadly categorized into supervised learning, unsupervised learning and reinforcement learning, among which supervised learning is the most common type, where the model learns from labeled data. The training data includes input-output pairs, meaning that each example comes with the correct answer or label. The goal of the algorithm is to find patterns in this data that can be used to make predictions on new, unseen data [52]. This category is called 'supervised' because the learning process is supervised by the provided labels. Parametric supervised learning algorithms are characterized by a fixed number of parameters regardless of the dataset size. Examples of parametric algorithms include Linear Regression, Polynomial Regression, Neural Networks, and so on.

More formally, a parametric ML model learns from a training dataset D_{train} which consists of N labeled samples. Each sample is independent and identically distributed (i.i.d.), which is represented as a pair $(\boldsymbol{x}^{(i)}, y^{(i)})$ where $\boldsymbol{x}^{(i)}$ is a vector of d features corresponding to the i-th observation and $y^{(i)}$ is the associated label. The model learns patterns in this data and produces a mapping $f \in H: X \mapsto Y$ that relates $\boldsymbol{x} \in X$ to $y \in Y$. H is the hypothesis space and are often parametric functions. After the model finds the best hypothesis f^*, given a new observation $\boldsymbol{x}$, the learned model can map it to a prediction $\hat{y}$, expressed as Equation (2.1).

$$\hat{y} = f^*(\boldsymbol{x}; \boldsymbol{\theta}) \text{ where } \boldsymbol{x} = \left(x_1, x_2, \ldots, x_d\right) \tag{2.1}$$

The process of training an ML model to predict reliable results is typically achieved by minimizing a loss function. The loss for a given sample $\boldsymbol{x}^{(i)}$ measures the discrepancy

between the predicted value $\hat{y}^{(i)}$ and the true label $y^{(i)}$. This loss is usually calculated using a predefined function $l:(f(\boldsymbol{x}^{(i)}), y^{(i)}) \mapsto \mathbb{R}$. The average loss L across all samples in the training dataset is often minimized during training, which can be formulated as an optimization problem (referred to as Equation (2.2)).

$$L = \frac{1}{N}\sum_{i=1}^{N} l(f(\boldsymbol{x}^{(i)}; \boldsymbol{\theta}), y^{(i)}) \tag{2.2}$$

The training process aims to find the best hypothesis f^* from the hypothesis space H, by identifying its optimal parameters θ^*, as specified in Equation (2.3), that yield the lowest average loss L across all hypotheses in H.

$$\boldsymbol{\theta}^* = \underset{\boldsymbol{\theta}}{\operatorname{argmin}} \frac{1}{N}\sum_{i=1}^{N} l(f(\boldsymbol{x}^{(i)}; \boldsymbol{\theta}), y^{(i)}) \tag{2.3}$$

The ideal outcome is that the chosen hypothesis not only fits the training data D_{train} well but also generalizes effectively to new, unseen samples, collectively denoted by D which are not part of the training data. This capability to perform well on both the training data and unseen data is referred to as the model's generalization ability. To enhance model's generalization ability, regularization techniques such as L1 and L2 regularization [36] may be added to the loss function to prevent overfitting. Other modifications can include the use of sample weights to give different importance to certain samples in the loss calculation. Another important consideration is to ensure that the training dataset D_{train} closely mirrors the distribution of the full dataset D, making D_{train} representative of the true data distribution. This alignment can be achieved by carefully selecting real-world data samples for the training set or by generating synthetic data that adheres to probabilistic assumptions resembling those of the real world. The latter approach is one objective of the study in Chapter 6.

Among the various optimization techniques for minimizing the loss function — such as gradient descent, momentum, and Newton's method — gradient descent methods [53] are commonly preferred due to their simplicity, efficiency, and effectiveness in handling the

large datasets that are frequently encountered in ML applications. The basic idea of gradient descent is that the parameters are updated until the loss converges; at each time step, the parameters are moved towards the direction that can reduce the loss value, as shown in Equation (2.4), where α is the learning rate, a hyperparameter that controls the size of the steps taken towards the minimum of the loss function. It determines how much the parameters will be adjusted during each update. $\frac{\partial L}{\partial \boldsymbol{\theta}}$ is the partial derivative of the loss function L with respect to the parameters $\boldsymbol{\theta}$. It represents the gradient of the loss function, indicating the direction and rate of the steepest increase.

$$\boldsymbol{\theta} = \boldsymbol{\theta} - \alpha \cdot \frac{\partial L}{\partial \boldsymbol{\theta}} \tag{2.4}$$

2.3.2 Deep learning

Deep learning (DL), a subset of ML, employs artificial neural networks with multiple layers to extract intricate features from complex data. Unlike traditional ML algorithms like Naïve Bayes [54], Random Forests [55], etc., DL autonomously learns these features, making it adept at tasks like image recognition and NLP (NLP). Its hierarchical representation capability marks a transformative advancement in the broader field of ML.

2.3.2.1 Neural networks

Neural networks are a class of ML models inspired by the structure and function of the human brain [56]. They are composed of interconnected nodes or neurons, organized in layers that process and transmit information. The basic building blocks of neural networks are these artificial neurons which perform computations on input data and transfer the results to other neurons in subsequent layers. In a neural network, each neuron receives inputs, performs a weighted sum of those inputs (adjusted by biases), and then passes them through an activation function to produce an output. These outputs can be either final predictions or inputs for neurons in the next layer, creating a hierarchical representation of the data [57].

A classical topology is the Feedforward Neural Network (FNN). The FNN is one of the earliest and simplest artificial neural networks. It propagates information through a series of layers without forming cycles, unlike recurrent neural networks. The specific equations governing the propagation process are typically denoted by Equation (2.5) and (2.6), with the symbols defined in Table 2.2.

$$z^{(k)} = \boldsymbol{W}^{(k)} \boldsymbol{x}^{(k-1)} + \boldsymbol{b}^{(k)} \tag{2.5}$$

$$\boldsymbol{h}^{(k)} = a_k(z^{(k)}) \tag{2.6}$$

Table 2.2. Meaning of symbols in neural networks

Symbols	Meaning of symbols
$a_k(\cdot)$	The activation function of layer k neuron
$\boldsymbol{W}^{(k)} \in \mathbb{R}^{M_k \times M_{k-1}}$	The weight matrix from layer k - 1 to layer k
$\boldsymbol{b}^{(k)} \in \mathbb{R}^{M_k}$	The bias from layer k - 1 to layer k
$z^{(k)} \in \mathbb{R}^{M_k}$	Input from layer k neurons
$\boldsymbol{h}^{(k)} \in \mathbb{R}^{M_k}$	Output from layer k neurons

Some commonly used activation functions are shown in Table 2.3.

Table 2.3. Examples of commonly used activation functions

Activation Function	Formula
ReLU (Rectified Linear Unit) [58]	$a(x) = \max(0, x)$
LeakyReLU (Leaky Rectified Linear Unit) [59]	$a(x) = \begin{cases} x, & \text{if } x \geq 0 \\ \alpha \cdot x, & \text{otherwise} \end{cases}$ where α is a small positive constant
Tanh (Hyperbolic Tangent) [60]	$a(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$
Sigmoid (Logistic Function) [61]	$a(x) = \frac{1}{1 + e^{-x}}$

The significance of neural networks in ML is profound and is reflected in several key

aspects: (1) Non-linearity and Flexibility. Neural networks can model complex, non-linear relationships between inputs and outputs that traditional linear models may struggle with. This enables them to excel in tasks such as image and speech recognition, where patterns are highly intricate. (2) Feature Learning: Unlike many classical ML algorithms that require handcrafted features, neural networks can automatically learn relevant features from raw data. This ability is particularly useful in DL architectures where multiple layers allow for hierarchical feature extraction. (3) Scale and Performance: Large-scale neural networks have proven to be extremely powerful, achieving state-of-the-art performance in many domains, including computer vision, NLP, game playing, and autonomous systems. (4) Generalization: Despite being trained on specific datasets, neural networks often generalize well to unseen data, making them robust for real-world applications [57].

2.3.2.2 Briefing of deep learning

DL is a sophisticated area within the broader field of ML, distinguished by its reliance on artificial neural networks with multiple layers to learn and extract meaningful features from complex data. These networks, often referred to as "deep" because of their depth in layers, are designed to mimic the hierarchical organization and functioning of the human brain, where each layer learns to recognize different levels of abstraction. In traditional ML algorithms, feature engineering plays a critical role, requiring experts to manually design and extract features from raw data. However, in DL, the model itself learns these features automatically through training on large datasets. This ability to learn representations directly from raw inputs (like images, sound, or text) without manual intervention is one of the key strengths of DL.

The architecture of deep neural networks typically includes an input layer that receives raw data, several hidden layers for feature extraction and transformation, and an output layer that produces the final result. These layers can take various forms such as fully connected, convolutional, recurrent, or other specialized types depending on the task at hand. For instance, Convolutional Neural Networks (CNNs) excel at image and video recognition

tasks due to their ability to identify spatial patterns, while Recurrent Neural Networks (RNNs), including Long Short-Term Memory (LSTM) networks, are highly effective in processing sequential data like speech or text where context is crucial. Additionally, modern Transformer architectures have demonstrated adaptability to both spatial patterns and sequential data, making them versatile solutions in various ML tasks.

DL has revolutionized various fields such as computer vision, NLP, speech recognition, recommendation systems, and more, achieving state-of-the-art performance in many applications. Its power lies in its capacity to handle big data effectively, extract intricate patterns, and generalize them across diverse scenarios, making it a cornerstone technology in the modern AI landscape [57].

2.3.2.3 Transformer architecture

As a seminal model architecture in DL, introduced by Vaswani et al. in their 2017 paper "Attention is All You Need" [62], the Transformer architecture has revolutionized the NLP field, serving as the foundation for state-of-the-art models such as GPT-4, Gemini, and Baidu's ERNIE Bot. Distinct from traditional approaches, it avoids recurrent and convolutional layers in favor of self-attention mechanisms, enabling direct interaction between all positions in the input sequence. This innovation accelerates parallel processing and training efficiency. The architecture, detailed in Figure 2.1, consists of encoder and decoder components. Encoders generate context-aware representations through self-attention and feed-forward networks, while decoders, using similar mechanisms plus attention to encoder outputs, facilitate autoregressive predictions for tasks such as translation and text generation. Key features include input embeddings with positional information, multi-head self-attention for capturing complex dependencies, and layer normalization and residual connections in both encoder and decoder stacks to enhance training. Through these design choices, the Transformer significantly improves NLP performance by effectively processing and predicting sequence data.

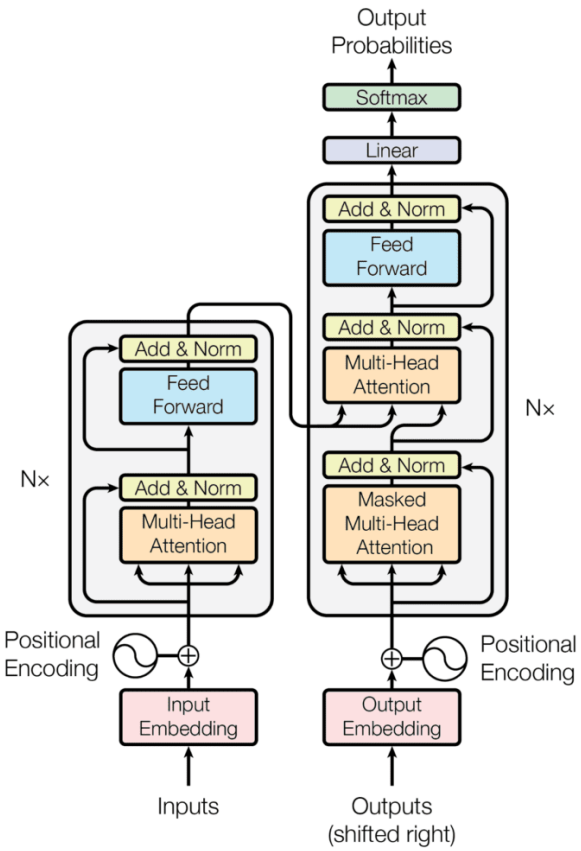


Figure 2.1. Transformer architecture [62]

(1) Input Embeddings

Input Embeddings consist of two parts: Token Embeddings and Positional Encodings. As for token embeddings, each word or token in the input sequence is transformed into a vector of a fixed size, denoted by d_{model}, which represents the semantic meaning of the token within the space of the model. As for positional encoding, since the Transformer architecture does not inherently capture sequence order (unlike RNNs or CNNs), it requires a mechanism to incorporate the order of the tokens. Positional Encodings are therefore added to the token embeddings to imbue them with information about the relative or absolute position of the tokens in the sequence. The original Transformer paper proposes using sine and cosine functions of different frequencies for these encodings, shown as

Equations (2.7) and (2.8).

$$PE_{(pos,2i)} = \sin\left(\frac{pos}{10000^{2i/d_{\text{model}}}}\right) \tag{2.7}$$

$$PE_{(pos,2i+1)} = \cos\left(\frac{pos}{10000^{2i/d_{\text{model}}}}\right) \tag{2.8}$$

Where *pos* is the position and *i* is the dimension. That is, each dimension of the positional encoding corresponds to a sinusoid. The wavelengths form a geometric progression from 2π to $10000 \cdot 2\pi$.

(2) The Multi-Head Self-Attention Mechanism

The Multi-Head Self-Attention Mechanism is a critical component of the Transformer architecture, enabling it to capture global dependencies within a sequence. The mechanism computes Query-Key-Value (QKV) attention, allowing each token to assess all other tokens in the sequence. Queries (***Q***), keys (***K***), and values (***V***) are derived from the input embeddings via linear transformations. The attention mechanism computes the compatibility of each query with all keys, yielding attention scores that are used to weight the values. These weighted sums focus the model's attention on the most relevant parts of the input when constructing the output representation for each token. The attention function is usually scaled by the inverse square root of the dimension of the key to avoid overly large dot product values leading to vanishing gradients during training. This scaling factor helps maintain a more manageable gradient flow. The process is summarized in the Equation (2.9).

$$\text{Attention}(\boldsymbol{Q}, \boldsymbol{K}, \boldsymbol{V}) = \text{softmax}\left(\frac{\boldsymbol{Q}\boldsymbol{K}^T}{\sqrt{d_k}}\right)\boldsymbol{V} \tag{2.9}$$

Here, d_k is the dimensionality of the keys (and queries), which ensures that the dot products are scaled appropriately. The softmax function is applied to convert the scores

into probabilities that sum to one. The attention scores are then used to create a weighted sum of the values, which forms the output of the attention mechanism.

Rather than calculating attention once, the Transformer model computes it multiple times in parallel through the Multi-Head Attention mechanism, where each "head" performs attention independently using distinct, learned linear projections. This process enables the model to capture information from different representational subspaces and consider various aspects of the input at different positions. The Multi-Head Attention can be described by Equation (2.10).

$$\begin{aligned} &\text{MultiHead}(\boldsymbol{Q}, \boldsymbol{K}, \boldsymbol{V}) = \text{Concat}(\text{head}_1, \ldots, \text{head}_h)\, \boldsymbol{W}^O \\ &\text{where head}_i = \text{Attention}(\boldsymbol{Q}\boldsymbol{W}_i^Q, \boldsymbol{K}\boldsymbol{W}_i^K, \boldsymbol{V}\boldsymbol{W}_i^V) \end{aligned} \tag{2.10}$$

In this extended equation, $\boldsymbol{Q}\boldsymbol{W}_i^Q, \boldsymbol{K}\boldsymbol{W}_i^K, \boldsymbol{V}\boldsymbol{W}_i^V$ are the projection matrices for queries, keys, and values respectively for the i-th head, and $\boldsymbol{W}^O$ is the output projection matrix. h represents the total number of heads in the multi-head attention mechanism. The concatenation of outputs from all heads is linearly transformed to produce the final values. This allows the model to capture information from different representation subspaces at different positions.

The Transformer model implements multi-head attention in three distinct ways: Encoder-decoder attention layers allow decoder positions to attend to encoder outputs, similar to traditional sequence-to-sequence models. Self-attention within the encoder lets each position consider all previous encoder outputs. In the decoder, self-attention is masked to maintain the auto-regressive property, permitting positions to attend only to previously decoded positions, preventing future information peeking.

(3) Position-wise Feed-Forward Networks

Besides the attention sub-layers, each layer in both the encoder and decoder includes an identical, position-wise fully connected feed-forward network. This network comprises two linear transformations with a ReLU activation function applied in between them, as

depicted in Equation (2.11).

$$FFN(x) = \max(0, xW_1 + b_1)\, W_2 + b_2 \tag{2.11}$$

Where $\max(0, xW_1 + b_1)$ is a ReLU activation function that outputs the input directly if it is positive; otherwise, it outputs zero. It's a piecewise function. $xW_1 + b_1$ represents the first layer transformation where x is the input vector. W_1 is the weight matrix associated with the first layer and b_1 is the bias vector for the first layer. W_2 is the weight matrix associated with the second layer and b_2 is the bias vector for the second layer.

(4) Layer Normalization and Residual Connections

Layer normalization [63] is applied after every multi-head attention and FFN block, but before the addition of the residual connection, to ensure stable gradients during training. Residual connections, which add the input directly to the output of each sub-layer, facilitate gradient propagation and improve the training of deep architectures.

(5) Encoder-Decoder Structure

The Transformer's encoder comprises six identical layers, each with a multi-head self-attention mechanism and a feed-forward network, leveraging residual connections and layer normalization to stabilize gradients and aid deep network training; all layers output a 512-dimension vector. The decoder mirrors the encoder's structure with an additional sub-layer for encoder output attention, modified self-attention to mask future tokens, ensuring autoregressive behavior where position *i* predictions rely solely on known outputs from positions before *i*. This design facilitates the encoding of input sequences into rich representations and the sequential decoding for tasks like machine translation, where each step's prediction influences the next.

2.3.3 Loss functions

Loss functions are mathematical functions utilized in ML models to quantify the error or

disparity between predicted outputs and actual (ground truth) values during the training process. The primary objective of training is to minimize this loss, leveraging optimization techniques such as gradient descent, as loss functions are typically chosen to be differentiable and convex to facilitate efficient optimization. There are a variety of basic losses, which can be broadly classified into two types: for regression and for classification. Losses for regression include Mean Squared Error (MSE), Mean Absolute Error (MAE), Huber loss, Smooth L1 loss, etc. These losses measure the discrepancy between predicted continuous values and ground truth labels, with each having its own characteristics in terms of sensitivity to outliers and robustness. Losses for classification, on the other hand, include Cross-Entropy Loss, Binary Cross-Entropy Loss (used for binary classification), and Categorical Cross-Entropy Loss (used for multiclass classification), Hinge Loss, etc. These losses are designed to measure the dissimilarity between predicted class probabilities and true class labels, facilitating the training of models to correctly classify input data into predefined categories [64]. In this dissertation, MSE, MAE, and RMSE losses are used for regression task, and Cross-Entropy Loss for language modeling (considered as a classification task), in addition to a combination and extension of these losses, to achieve various training objectives.

2.3.3.1 MSE loss

MSE loss is a widely adopted loss function in regression tasks, especially in ML and DL models designed to predict continuous values. MSE quantifies the average squared difference between the predicted values and the ground truth values. The mathematical formula for MSE loss is depicted in Equation (2.12).

$$L_{\text{MSE}} = \frac{1}{N} \sum_{i=1}^{N} \left(y^{(i)} - \hat{y}^{(i)} \right)^2 \tag{2.12}$$

By squaring the differences, MSE gives more penalty to larger errors, making it advantageous in scenarios where minimizing the overall deviation from the true values is crucial. However, it is sensitive to outliers because they contribute disproportionately due

to the squared term, potentially affecting the model's robustness, especially in datasets with extreme values. Minimizing the MSE during training enables the model to learn parameters that reduce the average squared error over all samples, consequently enhancing its accuracy in regression tasks.

2.3.3.2 RMSE loss

RMSE loss, shown in Equation (2.13), quantifies the average discrepancy between predicted and actual values. By squaring errors, it penalizes larger deviations more heavily. RMSE is sensitive to outliers due to its squared nature. Its square root renders it interpretable in the same units as the target variable, aiding in assessing model performance.

$$L_{\text{RMSE}} = \sqrt{L_{\text{MSE}}} \tag{2.13}$$

2.3.3.3 MAE loss

MAE loss, shown as Equation (2.14), computes the average absolute discrepancy between predicted and actual values, providing a robust measure of error in regression tasks. Unlike RMSE, it does not penalize larger deviations more heavily, making it less sensitive to outliers. MAE is straightforward to interpret and is suitable when the absolute magnitude of errors is more critical than their direction.

$$L_{\text{MAE}} = \frac{1}{N}\sum_{i=1}^{N} |y^{(i)} - \hat{y}^{(i)}| \tag{2.14}$$

Although it is typically not used as a loss function, the Coefficient of Determination (R^2), shown as Equation (2.15), is a useful metric for regression tasks. It measures the proportion of the variance in the dependent variable $\boldsymbol{Y}$ that is predictable from the independent variables $\boldsymbol{X}$. R^2 ranges from 0 to 1, where 1 indicates a perfect fit. It is commonly employed in linear regression models, yet its applicability can extend to other types of regression. This includes polynomial regression, ridge regression, and more. R^2 is a versatile metric that can be valuable for evaluating the predictive power of regression models.

$$R^2 = 1 - \frac{\sum_{i=1}^{N} (y^{(i)} - \hat{y}^{(i)})^2}{\sum_{i=1}^{N} (y^{(i)} - \bar{y})^2} \tag{2.15}$$

Where $\bar{y}$ is the mean value of the actual values.

2.3.3.4 Cross-Entropy loss

Cross-Entropy Loss, also known as Log Loss, is a commonly used loss function in both classification tasks and language modeling.

(1) Cross-Entropy Loss for Classification Tasks

In a classification problem, there are typically multiple classes to predict, such as in image classification with various categories. Cross-Entropy loss quantifies the disparity between the predicted probability distribution and the true label. The Cross-Entropy loss is shown in Equation (2.16). This equation computes the average Cross-Entropy loss over a dataset with *n* samples, each belonging to one of *C* possible classes. In ML, it is a common way to measure how well a classification model's predicted probabilities align with the actual labels.

$$L_{Cross-Entropy}(Y, P) = -\frac{1}{N} \sum_{i=1}^{N} \sum_{c=1}^{C} y_c^{(i)} \log(P(y = c | \boldsymbol{x}^{(i)})) \tag{2.16}$$

Where $\boldsymbol{Y}$ is a matrix of the one-hot encoded true labels for all *n* samples. $\boldsymbol{P}$ is a matrix of the predicted probabilities for all *n* samples. $y_c^{(i)}$ is the true label for class *c* of the *i*-th sample. $P(y = c | \boldsymbol{x}^{(i)})$ is the predicted probability that the *i*-th sample belongs to class *c*. *n* is the total number of samples. *C* is the total number of classes.

Because *y* is represented as one-hot encoding, Equation (2.16) can be simplified as Equation (2.17). For a sample *i*, we only consider the predicted probability for the correct class $c^{(i)}$, and take the logarithm of this probability. A higher probability for the correct class results in a lower loss. We sum these values across all classes and all samples, and then take the average by dividing by the number of samples *n*. This gives us a single value

that reflects the model's average performance across the entire dataset, with a lower value indicating better performance.

$$L_{Cross-Entropy}(\boldsymbol{Y},\boldsymbol{P}) = -\frac{1}{N}\sum_{i=1}^{N}\log(P(y=c^{(i)}|\boldsymbol{x}^{(i)})) \tag{2.17}$$

(2) Cross-Entropy Loss for Language Modeling

In NLP, language models predict the next word in a sequence of words. Each token prediction can be treated as a classification problem where the model predicts a probability distribution over the entire vocabulary. Thus, Cross-Entropy loss is commonly used to measure the performance of the model, as it quantifies the difference between the predicted probability distribution and the actual distribution of words. Consider a sequence of T words, where the target word at time step t is w_t, the loss of the sequence is shown in Equation (2.18) [65].

$$L_{sequence} = -\frac{1}{T}\sum_{t=1}^{T}\log(P(w_t|w_{1:t-1})) \tag{2.18}$$

The model outputs a probability distribution over the entire vocabulary given the context of the previous words $w_{1:t-1}$. The Cross-Entropy loss for this word prediction task is calculated by taking the negative log probability assigned to the true target word w_t and averaging this loss over all T words in the sequence.

2.3.3.5 Specially designed losses

It is important to recognize that a variety of fundamental loss functions can be combined or extended to design specialized loss functions tailored to specific objectives or challenges in ML tasks. By combining elements of different loss functions or incorporating additional components, researchers can create novel loss functions that address the unique requirements of their particular problem domains. For instance, the Elastic Net combines L1 and L2 regularization to mitigate overfitting while maintaining model sparsity, making

it ideal for regression tasks with many features [66]. In contrast, Focal Loss, by adjusting the contribution of each example based on its classification difficulty, effectively tackles the issue of class imbalance, enhancing the model's focus on hard-to-classify instances [67]. Similarly, Triplet Loss facilitates the learning of rich feature representations by encouraging relative distances between data points, proving invaluable in applications like facial recognition [68]. GAN Loss embodies the adversarial paradigm, propelling the concurrent training of generator and discriminator networks to produce remarkably realistic synthetic data [69]. The loss function designed for the InstructGPT language model study, which utilizes Supervised Fine-tuning and Reinforcement Learning from Human Feedback training techniques, integrates the loss representing rewards to generated sentences with pre-training losses and KL divergence [70]. The latter two components help prevent the policy model from over-optimizing for rewards, promoting a more robust and generalizable model by reducing the risk of overfitting to the reward signal.

2.3.4 Optimization

In the realm of ML, optimization plays a pivotal role in fine-tuning model parameters, aiming to minimize a loss function or maximize an objective function. This process is crucial for the effective learning and performance of various ML models. Among the numerous optimization methods, gradient descent [53] has been particularly noted for its simplicity and efficiency when dealing with differentiable objectives. This method's ability to iteratively adjust model parameters in the direction of the steepest decrease in the loss function makes it a fundamental technique for training a wide array of supervised learning models.

Within the broader family of gradient descent algorithms, Stochastic Gradient Descent (SGD) optimization is particularly notable [53], as illustrated in Algorithm 1. Unlike standard (batch) gradient descent, which computes the gradient of the loss function using the entire dataset at each iteration, SGD updates parameters incrementally, either by selecting a single data point $\left(\boldsymbol{x}_t, y_t \right)$ (in the case of true SGD) or a mini-batch of data (in

the case of Mini-batch SGD) at each step t to calculate the gradient $\frac{\partial L(\boldsymbol{\theta}, \boldsymbol{x}_t, y_t)}{\partial \boldsymbol{\theta}}$. This stochastic approach can significantly enhance computational efficiency, especially with large datasets, and it may also contribute to avoiding local minima. The versatility in handling individual samples or mini-batches makes SGD a versatile and preferred choice for training complex models in practical ML applications.

Algorithm 1: Stochastic Gradient Descent (SGD)

Input: Learning rate α, number of iterations T, dataset D
Output: Trained parameters $\boldsymbol{\theta}$
Initialize parameters $\boldsymbol{\theta}$
for $t = 1$ *to* T do
  Sample a data point $(\boldsymbol{x}_t, y_t)$ from the dataset D
  Compute the loss function value $L(\boldsymbol{\theta}, \boldsymbol{x}_t, y_t)$
  Compute the gradient $\frac{\partial L(\boldsymbol{\theta}, \boldsymbol{x}_t, y_t)}{\partial \boldsymbol{\theta}}$
  Update the parameters: $\boldsymbol{\theta} \leftarrow \boldsymbol{\theta} - \alpha \cdot \frac{\partial L(\boldsymbol{\theta}, \boldsymbol{x}_t, y_t)}{\partial \boldsymbol{\theta}}$
end
return $\boldsymbol{\theta}$

While SGD has been a cornerstone in training ML models due to its simplicity and effectiveness, it is not without its limitations, especially when applied to complex tasks with intricate loss function landscapes. One notable challenge is SGD's constant learning rate, which does not adjust to the specific features of the parameter space, potentially leading to suboptimal convergence rates. In response to these limitations, advanced optimizers like Adam [71] have been developed. Adam builds upon the strengths of SGD by introducing adaptive learning rates for each parameter, utilizing estimates of lower-order moments of the gradients. This sophistication allows Adam to navigate the complex landscapes of various loss functions more deftly, providing an edge in achieving faster and more stable convergence in a wide array of ML tasks.

Building upon the solid foundation laid by Adam, AdamW [72] is a variant that incorporates weight decay regularization directly into the optimization process. AdamW

extends Adam's adaptive learning rate methodology by decoupling weight decay from the loss function, thereby enabling a more straightforward and effective regularization method. This innovation not only leverages the strengths of Adam but also significantly enhances model generalization capabilities. The adoption of AdamW in this research underscores a commitment to harnessing the latest advancements in optimization techniques, with the aim of pushing the boundaries of model performance. The pseudocode for AdamW is shown as Algorithm 2.

Algorithm 2: Adaptive Moment Estimation with Weight Decay Correction [72]

Input: learning rate α, exponential decay rates for moment estimates β_1, β_2, weight decay parameter λ

Initialization: initialize parameters $\boldsymbol{\theta}_0$, first moment vector $\boldsymbol{m}_0 = 0$, second moment vector $\boldsymbol{v}_0 = 0$

for t = 1,2,... do

Compute gradient w.r.t. parameters: $\nabla_{\theta_t} L$

Update biased first moment estimate:

$$\boldsymbol{m}_t = \beta_1 \boldsymbol{m}_{t-1} + (1 - \beta_1) \nabla_{\theta_t} L$$

Update biased second moment estimate:

$$\boldsymbol{v}_t = \beta_2 \boldsymbol{v}_{t-1} + (1 - \beta_2) (\nabla_{\theta_t} L)^2$$

Correct the bias of the first moment estimate:

$$\widehat{\boldsymbol{m}}_t = \frac{\boldsymbol{m}_t}{1 - \beta_1^t}$$

Correct the bias of the second moment estimate:

$$\widehat{\mathbf{v}}_t = \frac{\boldsymbol{v}_t}{1 - \beta_2^t}$$

Apply weight decay correction separately

Perform Adam update without decoupling weight decay:

$$\boldsymbol{\theta}_{t+1} = \boldsymbol{\theta}_t - \frac{\alpha}{\sqrt{\widehat{\mathbf{v}}_t} + \varepsilon} \odot \widehat{\mathbf{m}}_t$$

Apply L2 regularization (weight decay)

Apply weight decay after the Adam update:

$$\boldsymbol{\theta}_{t+1} = \boldsymbol{\theta}_{t+1} - \alpha \cdot \lambda \cdot \boldsymbol{\theta}_{t+1}$$

end

return $\boldsymbol{\theta}$

2.4 Conclusions

In this chapter, we discussed risk management and placed it in the context of construction projects, underscoring its critical role in preventing and mitigating risks to ensure successful project delivery. We then examined cybersecurity studies within the construction industry, categorizing them into general discussions, review papers, and specific solutions. We concluded that the existing research on this topic is limited and fragmented, presenting a challenge to achieving a unified progression in the field of construction cybersecurity. Finally, we explored the fundamentals of ML, explaining the techniques central to this dissertation. In the forthcoming chapters, we will investigate how cyber risk management can be integrated into construction projects to enhance cybersecurity through the use of advanced technologies, including ML.

Chapter 3
Exploring Cybersecurity in the Construction Industry through Topic Modeling

In this chapter, potential research directions are identified by utilizing the topic modeling technique. A text corpus is established by collecting various text sources, which undergo a scrupulous text screening process, including keyword and semantic screening. The Latent Dirichlet Allocation (LDA) technique for topic modeling is then applied to the corpus to identify emerging research topics. Each identified topic is summarized along with high-level actions suggested, among which prioritized topics are recommended and discussed in detail. Cyber risk management, one of the most significant topics, sets the foundation for subsequent chapters of this dissertation.

3.1 Introduction

Identifying research topics in the construction industry is crucial for fostering innovation, improving efficiency, and addressing emerging challenges. It enables advancements in technology, sustainability practices, and safety protocols, driving progress and

competitiveness within the sector. Despite these broad benefits, a critical area within the construction industry that struggles with organized research efforts is cybersecurity. As noted in Section 2, current research on construction cybersecurity is scattered across general discussions, review papers, and targeted technical solutions. The topics explored are diverse, ranging from training to combat social engineering, managing 3D printing risks, understanding the impact of Construction 4.0, to developing sector-specific cybersecurity measures like encryption, blockchain, and intrusion detection systems. These efforts, while covering a wide range of topics, lack coordination and fail to create a comprehensive, accessible framework for practitioners not specialized in cybersecurity. This lack of organization prevents a unified understanding of the unique cybersecurity challenges in the construction industry and thus reveals a critical knowledge gap. Therefore, it is crucial to integrate these disparate studies into a unified framework, promoting a more systematic exploration of the field. This would not only address the current fragmentation but also enhance the accessibility and applicability of cybersecurity research within the construction industry, fostering a more robust defense against evolving digital threats.

The topic modeling technique can bridge the gap. Topic modeling [73], an ML technique, offers a powerful method for discovering hidden thematic structures in large text datasets. By algorithmically analyzing and grouping texts based on underlying content, this approach enables the systematic identification of prevailing themes, emerging trends, and areas requiring further investigation in cybersecurity within the construction industry. Through topic modeling, this study can not only consolidate fragmented topics but also discover underexplored ones, thus formulating a comprehensive reference for research directions. This facilitates a deeper understanding of the complexity and interconnectedness of cybersecurity challenges in the construction sector, contributing to unified progress in this field.

A variety of text sources available online can be utilized for topic modeling, each offering unique insights and perspectives on the subject at hand. For instance, news articles serve as a rich repository of timely information, capturing the latest trends, incidents, and

developments in cybersecurity within the construction industry. These articles can shed light on recent cyberattacks, emerging threats, and the evolving strategies employed by both attackers and defenders. Academic publications, on the other hand, provide in-depth analyses, theories, and empirical studies, contributing to a more nuanced understanding of cybersecurity challenges and solutions in construction. Blogs and industry reports offer practical insights and expert opinions, highlighting real-world applications, case studies, and best practices in cybersecurity management. Together, these diverse sources encompass a broad spectrum of knowledge, from theoretical frameworks to actionable intelligence, making them invaluable for topic modeling. By systematically extracting and analyzing topics from such a wide array of texts, we can construct a comprehensive landscape of cybersecurity in the construction industry, identifying key themes, gaps in the literature, and avenues for future research.

In this study, LDA is chosen for topic modeling due to its proven effectiveness in handling large datasets and its ability to uncover latent thematic structures within text corpora. Unlike other methods that may require predefined topics or depend heavily on the exact matching of terms, LDA operates on the assumption that documents are mixtures of topics, where each topic is characterized by a distribution of words. This allows for a more nuanced analysis of texts, revealing not just the dominant themes but also the intricate relationships between them. Furthermore, LDA's flexibility in adapting to the complexity and variability of language used across different sources makes it particularly suited for our comprehensive study on cybersecurity in the construction industry. By leveraging LDA, we can achieve a deep and systematic understanding of the topic landscape, which is essential for identifying emerging trends and guiding future research directions [74], [75].

3.1.1 Related works

As mentioned above, LDA excels in large-scale thematic analysis, revealing nuanced topic relationships and adapting to linguistic variability, making it ideal for comprehensive cybersecurity studies. It has become one of the most popular topic models and has been

widely used in various industries such as social media analysis [76], [77], [78], software engineering [79], [80], [81], finance [82], [83], [84], biology [85], [86], [87], medical sciences [88], [89], [90], and the construction field [91], [92], [93], [94].

In the field of social media analysis, Zhou et al. [76] utilized LDA to investigate the temporal potential themes on Twitter during Hurricane Laura in 2020. The results of this study could help quickly identify and extract situational awareness information for responders, stakeholders, and the public so that they can adopt timely responsive strategies and wisely allocate resources during hurricane events. Tsolmon and Lee [77] employed the LDA topic model to extract social events based on the weights of event items on the timeline and from reliable users. Since Twitter posts are short and noisy, this approach avoids the difficulty of extracting reliable events based solely on word frequency. de Groof et al. [78] applied LDA to analyze clustered text data in online social media to make further recommendations on topics related to the promotion of cultural products.

In software engineering field, Maskeri et al. [79] investigated using LDA to extract business domain topics from source code in large software systems, proposing a human-assisted approach and testing it on various software, with preliminary results showing promise for manual refinement. Linstead et al. [80] applied LDA technique to detect functional components in Eclipse and Argo UML source code, tracing their evolution and summarizing software dynamics for project management and comprehension, while also identifying refactoring events and general programming concepts. Asuncion et al. [81] proposed an automated traceability technique using topic modeling to semantically categorize artifacts and visualize software systems, implementing tools for artifact search, prospective traceability, and architecture navigation, applying it to datasets to enhance software traceability.

For the application of LDA in finance, Khatib [82] were the first to comprehensively construct a literature on the application of ML to finance. The authors analyzed 6,148 academic articles from 1990 to 2019 using LDA techniques, summarizing the progress of

ML in three major topics: investment analysis, asset modeling and forecasting, and risk management. Subsequently, in 2021, Aziz et al. [83] structured finance literature using LDA technique to identify 14 key research topics across 5,204 articles from 1990 to 2018, revealing topic evolution and offering a guide for integrating ML in finance research, emphasizing probabilistic topic modeling's comprehension benefits. Feuerriegel and Pröllochs [84] used LDA to pre-categorize the contents of more than 70,000 regulatory 8-K documents from U.S. companies. The authors then assessed the subsequent stock market reaction, benefiting managers, investors, and policymakers by indicating how regulatory filings should be structured and identifying the topics most likely to precede changes in stock valuations.

The LDA technique could also be advantageously applied to the large datasets of biomedical research. Yang et al. [85] developed an scRNA-seq analysis method using the LDA model to uncover cell function-gene patterns, outperforming classic methods in cell clustering accuracy. They demonstrated the method's ability to identify cell types and trace development, using benchmark and complex datasets. Shivashankar et al. [86] proposed using the LDA topic model for protein structure representation and a multi-viewpoint framework for retrieval, to identify close or remote homologs, demonstrating improved performance over state-of-the-art methods on a benchmark dataset by Kolodny and co-workers. Pinoli et al. [87] introduced two variants of the LDA algorithm for gene annotation prediction, utilizing collapsed Gibbs Sampling with distinct initialization approaches. They demonstrated their methods' efficiency by comparing them with tSVD predictions, validated against updated gene annotation datasets.

In the medical science field, Li et al. [88] utilized the LDA method to extract topic word sets from unstructured medical texts, enhancing the content for their KTI-RNN model, which aims to accurately recognize heart failure from large-scale electronic health records. Khalil et al. [89] applied the LDA technique to analyze 119,986 posts from Crohn's disease forums to uncover themes in patient experiences with perianal fistulae, revealing insights into the biopsychosocial impacts and treatment challenges, despite not identifying

apprehension about treatments as a separate theme. Frick et al. [90] used the LDA technique on Online Mendelian Inheritance in Man records to extract latent topics and measure disease similarity from textual descriptions, validating the approach against the Disease Ontology and exploring the use of topics for mapping concepts and tagging texts.

In recent years, LDA has also begun to be used in the construction field. Liu et al. [91] combined text mining and LDA with an information entropy TF-IDF scheme to analyze Chinese construction accident reports, identifying key safety accident factors and themes, enhancing accuracy and efficiency in understanding causal factors and improving safety management practices. Zhou et al. [92] employed the LDA technique and Correlation Explanation (CorEx) topic modeling to analyze 1,984 construction industry safety and health articles, clustering latent topics and keywords to identify research trends and inter-relationships, aiding in future research direction and topic selection. Zhong et al. [93] used LDA to visualize interdependencies between causal variables in construction accident narratives, complementing a CNN model for automated text feature extraction and classification, enhancing machine-assisted text interpretation and providing insights for on-site safety improvement. Zeng et al. [94] analyzed Instagram posts related to construction health and safety using LDA for topic extraction and Jaccard coefficient for co-occurrence cluster analysis, identifying key topics, collocations, active users, and sentiment trends to understand public perceptions and inform industry practices.

The extensive use of LDA across diverse fields, from social media analysis and software engineering to finance, biomedical research, and notably in the construction industry, underscores its versatility and effectiveness in extracting meaningful insights from vast textual data. This literature review highlights LDA's successful application in identifying key safety factors, thematic trends, and the inter-relationships of various topics within the construction domain. Given its proven efficacy, LDA presents a promising approach for identifying cybersecurity research topics in the construction industry, potentially enhancing safety management practices and informing future research directions by uncovering nuanced thematic relationships and evolving trends specific to cybersecurity challenges.

3.1.2 Objectives

The objectives of this study are twofold: (1) To create a text corpus comprising sentences related to the construction industry and cybersecurity. The sources of these texts will vary, encompassing news articles, databases, book chapters, academic publications, etc. This approach ensures a wide coverage of content and diversity in text content. The collected text is delicately filtered and processed into sentences suitable for topic modeling. (2) To perform topic modeling utilize the processed text corpus to identify emerging research directions of construction cybersecurity. Practitioners in the construction industry and researchers interested in cybersecurity can refer to the identified topics for insights into current trends and potential areas for future investigation. This could facilitate the development of innovative solutions and strategies to address the unique challenges at the intersection of construction and cybersecurity.

3.1.3 Contributions

The academic contributions of this study are twofold: (1) Addressing Research Gaps and Guiding Future Directions: Through topic modeling, this study identifies emerging research areas within the construction industry and cybersecurity, shedding light on significant gaps and under-explored topics. It directs researchers toward promising fields where their work could lead to substantial advancements, encouraging exploration at the intersection of these two areas. (2) Enriching Language Modeling Research with a Specialized Corpus: At a time when LLMs are increasingly prevalent, this study offers a unique corpus tailored to construction and cybersecurity. This contribution paves the way for advanced language models capable of specialized tasks such as question-answering and automatic risk identification in these sectors. This work aligns with the digital transformation trends in construction, marking a step forward into Construction 4.0 era.

The practical contributions of this study include: (1) Informing Industry Practices: Practitioners in the construction industry can utilize insights from this study to formulate or enhance cybersecurity measures. By identifying emerging topics and threats, they can

develop proactive strategies for risk management and resilience. (2) Guiding Policy and Standards Development: The findings of this study can inform policymakers and standards organizations in the creation of regulations and guidelines tailored to the evolving cybersecurity needs within construction. This is especially relevant as digital technologies become more integral to construction processes.

3.2 Methodology

Figure 3.1 illustrates the overall process of this study. It can be observed that there are 6 main steps leading to the final purpose of topic modeling, during which 4 intermediate corpora are formed and stored. The steps are explained in detail in Sections 3.2.1 to 3.2.6.

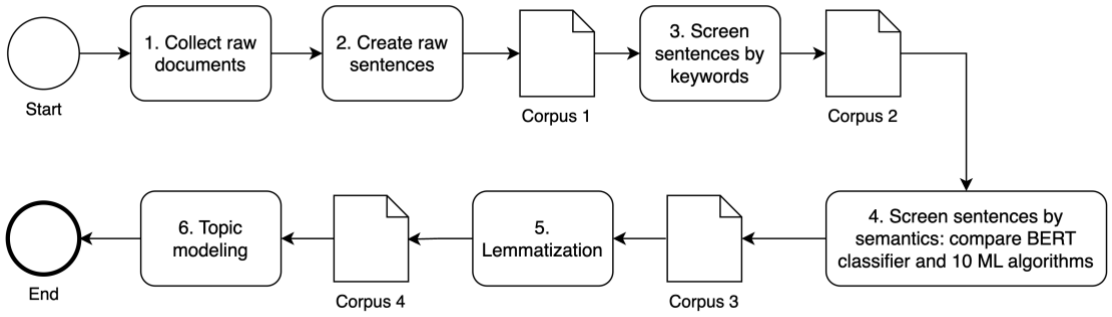


Figure 3.1. Overview of the main steps used in this study

3.2.1 Collect raw documents

Firstly, we collected as much relevant text as possible for topic modeling in the cybersecurity domain within the construction industry. The text was gathered from six sources to encompass most necessary data. Note that for most sources, we used the search criteria "cyber*" and "construction industry" to ensure relevance. In this context, "*" serves as a wildcard character to represent any sequence of characters. Therefore, "cyber" would match any word beginning with "cyber," such as "cybersecurity" or "cyberattack." Similarly, "construction industry" would match any occurrence of that exact phrase.

(1) **News articles and blogs.** They serve as valuable sources of domain-specific text, offering comprehensive insights into contemporary cyber incidents. They provide detailed information about the causes, procedures, and impacts of these incidents, facilitating a deeper understanding of cybersecurity within the construction industry. In this study, we utilized Octoparse software to crawl over 70 websites. We extracted the title, URL, and content of each piece of text from every website and saved them in CSV format.

(2) **LexisNexis database.** LexisNexis is a comprehensive digital database offering access to a vast repository of legal, journalistic, and business documents. It serves as an invaluable tool for professionals and researchers, providing detailed insights through legal case law, news archives, company information, and public records. Its advanced search capabilities enable users to navigate and extract precise information efficiently, supporting informed decision-making and research. The database includes documents dating back decades and adding about 54 million new documents each month. This study utilized the free API service provided by New York University (NYU) to request relevant documents using the Python query code. Through this process, we obtained nearly 4,000 pieces of text that were in JSON format. We then converted them into CSV format.

(3) **Academic papers.** Academic papers that are peer-reviewed serve as important data sources on cybersecurity in the construction industry due to the high quality and relevance of their professional knowledge. We conducted searches using the query words mentioned above in Google Scholar, Web of Science, and Scopus databases to identify the most relevant publications on this topic. From the articles, we removed duplicates and retained those that are highly relevant to our purpose, resulting in 78 publications saved as PDF files.

(4) **Books and book chapters.** Books and book chapters are also regarded as having high-quality content related to our study because they often provide comprehensive and in-depth analysis, theories, methodologies, and case studies developed by experts in the field. We applied the same searching criteria to search in Google

search engine. Although we could only find a small number of relevant books and book chapters, we were able to extract valuable information from those sources. Our search resulted in the inclusion of 13 books and chapters, which were saved as PDF files.

(5) **Specifications and standards.** Specifications and standards are important sources because they contain complete statements of legal or industry requirements related to cybersecurity. We conducted the initial search on Google search using the key words mentioned above. It should be noted that there are only a few standards directly related to cybersecurity in the construction industry. Therefore, we retained the standards, searched on Google, that are relevant to a broad range of industries and collected 37 documents for text extraction. including PAS 1192-5:2015, ISO/IEC 27001, etc., which were saved as PDF files.

(6) **Reports.** Reports from companies, including slides and manuals, can be useful sources of text because they provide insights into practical applications, company-specific cybersecurity practices, and real-world challenges and solutions. These documents often contain detailed information about the implementation of cybersecurity measures, case studies, and proprietary approaches to managing cyber risks. To gather this information, we used Google Search to locate 46 PDF files of company reports, manually copied their content, and saved it into TXT files.

The overview of the raw documents from various sources is presented in Table 3.1. The raw documents are open-sourced on the GitHub page [95].

Table 3.1. Overview of raw documents

Criteria	Text source	Tools	# of document	Data format
"cyber*" and "construction industry"	News articles and blogs	Octoparse crawling tool	75 websites	CSV
	LexisNexis databases	LexisNexis API through NYU	3,968 pieces of news	JSON, CSV
	Academic papers	Google Scholar, Web of Science, and Scopus	78 pieces	PDF
	Books (chapters)	Google search	13 files	PDF
	Specifications/Standards	Google search	37 files	PDF
	Company reports	Google search	46 files	TXT

3.2.2 Create raw sentences (Corpus 1)

This step aims to process the collected raw documents into a list of sentences, which forms our Corpus 1. Firstly, text was extracted from the raw documents. There are four types of raw documents: CSV, PDF, JSON, and TXT. We used open-source libraries to extract the text from these documents. The Python scripts and libraries used are summarized as follows: (1) The Slate3k library was used to extract text from one-column PDF files; (2) The Pytesseract library was used to extract text from PDF files that have two columns; (3) The Pandas library was used for processing CSV and JSON files and for text handling.

To segment the extracted text into a list of sentences, we used `mysentences.py` [96], a self-created Python script, to preliminarily clean and split the text from each source into sentences, which include the following processing steps.

- **Preparation**: Prepends and appends spaces to the text and replaces newline characters with spaces to ensure consistent spacing.
- **Handling Abbreviations and Websites**: Substitutes periods in recognized abbreviations, such as titles (Mr., Mrs., etc.), suffixes (Inc., Ltd., etc.), acronyms, and website domains (.com, .org, etc.), with **<prd>** to prevent them from being mistakenly identified as sentence boundaries.

- **Special Cases**: Adjusts for special cases like "Ph.D.", "et al.", "Fig.", and "Int." by replacing their periods with **<prd>** to maintain their integrity as single entities.
- **Acronym and Starter Adjustment**: Looks for sequences where an acronym is followed by a starter word (e.g., a capitalized word that typically begins a sentence) and introduces a **<stop>** tag to indicate potential sentence boundaries.
- **Suffix and Starter Adjustment**: Similar to the acronym and starter adjustment, this step deals with suffixes followed by starter words, using **<stop>** to mark potential sentence ends.
- **Consolidation of Periods in Acronyms**: Replaces sequences of periods in acronyms with **<prd>** to prevent them from being seen as multiple sentence ends.
- **Adjustment for Quotation Marks**: Ensures that periods, question marks, and exclamation points followed by quotation marks are treated correctly, adjusting the placement of punctuation to support accurate sentence splitting.
- **Cleaning and Standardizing Text**: Removes extra spaces and unnecessary hyphens, ensuring that the text is neat and standardized for further processing.
- **Sentence Splitting**: Finally, the text is split into sentences based on **<stop>** tags, which were introduced to mark the end of sentences. The **<prd>** placeholder is replaced back with a period, and each sentence is stripped of leading and trailing spaces.
- **Post-processing**: After the initial split, sentences shorter than a specified length (in terms of word count) are filtered out, leaving only those sentences considered substantial for analysis.

The statistics of Corpus 1 is shown in Table 3.2.

Table 3.2. Statistics of Corpus 1

Text source	Number of sentences	Total number of sentences
News articles and blogs	71 K	
LexisNexis databases	596 K	
Academic papers	26 K	802 K
Books (chapters)	73 K	
Specifications/Standards	22 K	
Company reports	14 K	

3.2.3 Screen sentences by keywords (Corpus 2)

Not all sentences were relevant to this study. Sentences containing noisy or irrelevant information were removed. To accomplish this, we created a list of 76 keywords to screen out related sentences. Some of the keywords are partial words, as there are different variations of the word, e.g., ‘'cyber' can be a part of ‘cybersecurity’, and some are abbreviations commonly used in the cybersecurity field, e.g., ‘dos ’ represents ‘denial of service’. According to the content of the keywords, they were grouped into five groups, as shown below. Some keywords might fit into multiple groups, but they were placed based on their primary association or most common context within cybersecurity discussions. This keyword filtering step resulted in our Corpus 2 (summarized in Section 3.3.2).

- **Cybersecurity Terms and Concepts:** {'cyber', 'threat', 'vulnerab', 'secur', 'safe', 'exploit', 'breach', 'malware', 'virus', 'ransom', 'trojan', 'worm', 'bot ', 'botnet', 'spyware', 'rootkit', 'ddos', 'phishing', 'spear', 'malicious', 'tamper', 'spoof', 'malfunction', 'weak', 'flaw', 'error', 'risk', 'attack', 'hack'}
- **Technology and Infrastructure:** {'internet', 'cloud', 'software', 'domain', 'virtual', 'vpn ', 'ip ', 'firewall', 'computing', 'infrastructure', 'network', 'iot ', 'mobile', 'application'}
- **Security Measures and Tools:** {'denial of service', 'emotet', 'clickjacking', 'deepfake', 'white hat', 'sql ', 'password', 'ware', 'intention', 'operat', 'disrupt', 'expos', 'inject', 'leak', 'insider', 'outsider', 'spear '}

- **Cybersecurity Attacks and Techniques:** {'denial of service', 'emotet', 'clickjacking', 'deepfake', 'white hat', 'sql', 'password', 'ware', 'intention', 'operat', 'disrupt', 'expos', 'inject', 'leak', 'insider', 'outsider', 'spear '}
- **Data and Information Security:** {'information', 'data', 'encrypt', 'decrypt', 'secret', 'expos', 'leak', 'insider', 'outsider', 'weak', 'flaw', 'error'}

3.2.4 Screen sentences by semantics (Corpus 3)

To ensure that all sentences in our study were meaningful and relevant, we employed a semantic screening process. This involved training a sentence classifier to filter out sentences that are not semantically relevant to construction cybersecurity.

Data: We randomly selected 2,000 sentences from Corpus 2, and manually labeled them sentences with the label "0" indicating "Exclusion" and the label "1" indicating "Inclusion." The criteria for including sentences in the training corpus are as follows: (1) the sentence should have a proper format and structure, and (2) the sentence should be directly relevant to the construction industry or cybersecurity. These criteria are clear and should effectively filter out texts suitable for training the model. Out of the 2,000 sentences, the number labeled 'Exclusion' is 1,381, while that labeled 'Inclusion' is 619.

Model: The classifier we employed is built upon a pre-trained BERT (Bidirectional Encoder Representations from Transformers) model [97], which is a groundbreaking method introduced by Devlin et al. for NLP tasks. BERT's architecture enables it to understand the context of words in a sentence by considering the words that come before and after, leveraging deep bidirectional training. To this foundation, we added a linear layer that serves as the classifier on top of the BERT embeddings. This setup allows the model to utilize the rich contextual representations generated by BERT, making it highly effective for a wide range of text classification tasks. By fine-tuning this composite model on our specific dataset, we aim to leverage the pre-trained BERT model's extensive knowledge of language, gained from training on vast corpora, to achieve superior performance in our

classification task.

Training: For training, we randomly divided the 2,000 sentences into a training set with 1,660 samples and a test set with 340 samples, achieving approximately an 80:20 data split. We experimented with different sets of hyperparameters and configurations for fine-tuning the model and ultimately selected the set of hyperparameters that achieved the highest performance on the test set. This model was trained using the Hugging Face's Transformers library [98].

Baseline comparison: For comparison to the BERT classifier, we also trained and tested 10 traditional statistics-based ML algorithms [99] as baselines: Naive Bayes (NB), Support Vector Machine (SVM), Logistic Regression (LR), Random Forest (RF), K-Nearest Neighbors (KNN), Gradient Boosting Decision Tree (GDBT), eXtreme Gradient Boosting (XGBoost), Stochastic Gradient Descent (SDG) Classifier, Decision Tree (DT), and Light Gradient Boosting Machine (LGBM) Classifier. To do this, we used the lemmatized version of Corpus 2 (see Section 3.2.5 for the explanation of lemmatization) to create Term Frequency-Inverse Document Frequency (TF-IDF) features and fed them into these ML algorithms. The training of these 10 algorithms was implemented using the scikit-learn library [100].

After training, the trained BERT classifier was used to screen all sentences in Corpus 2. The sentences labeled as "Inclusion" formed Corpus 3 (summarized in Section 3.3.2), which represents a corpus with high semantic relevance to construction cybersecurity.

3.2.5 Lemmatization (Corpus 4)

Now, Corpus 3 contains sentences that have high semantic relevance to construction cybersecurity. In order to perform topic modeling on this corpus, we need to create a lemmatized version of it. Lemmatization is the process of converting a word to its base form, or lemma, by considering its part of speech and meaning in the context of a sentence. This is more sophisticated than merely removing inflectional endings, as it requires

understanding the word's role and the specific grammar rules of the language. By converting words to their lemmas, lemmatization can significantly reduce the number of unique tokens in the corpus while maintaining the original level of information. This reduction in lexical complexity can improve the performance of NLP tasks, such as topic modeling, by focusing on the essential meanings of words and reducing noise. Generally, the lemmatization process involves the following steps: converting each sentence to lowercase, tokenizing each sentence into words using the NLTK library, removing stopwords and punctuation, and lemmatizing the words. The pseudo-code for this process is shown in Algorithm 3. A Python script named `myRaw2Lemmatized.py` [101] was developed to perform the entire lemmatization process in an integrated way, resulting in Corpus 4 (summarized in Section 3.3.2).

Algorithm 3: Lemmatization for sentences

```
For each text in the list of texts:
    Convert text to lowercase
    Tokenize the text into words using NLTK library
    Initialize a list for storing the lemmatized word
    For each word in the text:
        If the word is not a stopword and is alphabetic:
            Determine the part of speech (POS) tag for the word
            Lemmatize the word based on its POS tag
            Add the lemmatized word to the list
    Join the words in the list into a single string
```

3.2.6 LDA technique for topic modeling

Topic modeling is a technique for identifying the key themes or topics in a collection of texts. The LDA method [75] is a commonly used topic model whose main function is to discover hidden topic structures in text collections. It represents text data as distributions of topics and words and infers the relationships between documents (sentence in this study) and topics through statistical learning methods. Specifically, the LDA model assumes that each document (sentence in this study) can contain multiple topics and that each topic comprises multiple words. By inferring the statistical distribution of documents and words,

the model reveals the relationship between documents and themes, thereby uncovering hidden themes in text data. It essentially captures the core content of types of discussions or concepts.

To optimize the results of the LDA model, we must first determine the input variable that represents the number of topics to be generated. To compare different models with varying numbers of topics as the input variable, we can use the log perplexity metric and the coherence metric [102]. The log perplexity score, shown as in Equation (3.1), measures how well a probabilistic model predicts a sample, with a lower score indicating better model performance.

$$\text{Log Perplexity}(D) = -\frac{\sum_{d=1}^{M} \log(P(W_d))}{\sum_{d=1}^{M} N_d} \tag{3.1}$$

Where M is the total number of documents in the dataset (the sentences in Corpus 4 in this study). W_d represents the words in the d-th sentence. $P(W_d)$ is the probability of the word sequence in sentence d, as assigned by the model. N_d is the total number of words in the d-th document. The numerator $\sum_{d=1}^{M} \log(P(W_d))$ sums up the log probabilities of all sentences in the corpus. The denominator $\sum_{d=1}^{M} N_d$ is the total number of words across all sentences in the corpus, which normalizes the average log likelihood by the corpus size.

The coherence score assesses how related the top words in each topic are, with a higher score indicating that the topics are more meaningful and coherent. There are several coherence measures, but one of the most commonly used is C_V coherence [103]. The computation of C_V coherence involves a few steps and depends on multiple factors, including word co-occurrence and similarity scores. The general approach can be outlined as follows:

- **Segmentation**: For each topic, select the top *N* words based on their probability within the topic. These words are then segmented into pairs of words (sets of size 2).
- **Confirmation Measure**: Calculate a confirmation measure for each pair of words in the segment. One common measure is the pointwise mutual information (PMI), which for a pair of words w_i and w_j can be calculated as Equation (3.2):

$$\text{PMI}(w_i, w_j) = \log \frac{P(w_i, w_j)}{P(w_i)P(w_j)} \tag{3.2}$$

  Where $P(w_i, w_j)$ is the probability of observing w_i and w_j together, $P(w_i)$ and $P(w_j)$ are the probabilities of observing w_i and w_j independently.
- **Aggregation**: Aggregate the confirmation measures for all pairs in a topic to compute the topic coherence. One approach is to take the average.
- **Coherence Score**: Finally, the coherence score for the model is the average of the topic coherence scores across all topics.

Therefore, a model with a low perplexity score and a high coherence score is considered more accurate and produces more coherent topics. In essence, if the goal is to have topics that are meaningful and can be readily understood and used for further analysis, coherence should be the priority. We experimented with input variables ranging from 5 to 30 and evaluated the models based on their perplexity and coherence scores. The optimal input variable was chosen by considering a balance between minimizing perplexity and maximizing coherence. A Python script called `topic_modeling.py` [104] was developed to run the entire process. Section 3.3.3 describes the experimentation process and indicates how we ultimately chose 8 as the input variable (i.e., the number of topics). Section 3.4 discusses how we can summarize the topics from the keywords for each topic and the representative sentences.

3.3 Experiments and Results

In Section 3.3.1, we present the experiments and results of screening sentences by semantics and then present an overview of the statistical information for the four corpora established in Section 3.3.2. Section 3.3.3 describes the experiment process for deciding the input variable for topic modeling and presents the corresponding results.

3.3.1 Screen sentences by semantics

In Section 3.2.4, we described how to use BERT classifier to screen sentences semantically; we also planned to compare the BERT classifier with 10 ML algorithms. The training and experimenting of the BERT classifier were conducted on the Google Colab. A total of 10 experiments using the BERT classifier were conducted. The hyperparameter settings and training results are presented in Table 3.3. The main variables for experiments were the batch size for gradient accumulation, learning rate, class weights, and dropout rate. We set the total number of training epochs to 6, the optimizer with scheduler AdamW, and the loss function to Cross-Entropy. Note that in the 2,000 sentences, the number of label 'Exclusion' is 1,381 while that of 'Inclusion' is 619. To make the training dataset more balanced for training, we set different class weights when computing the loss, respectively 1:3, 1:2 and 1:1.4. Any ratio between 1:3 and 1:1 should be fine for experimentation.

Table 3.3. BERT classifier experiments

Setting number	Batch size	Learning rate	Class weights	Dropout rate	Epochs	Optimizer with scheduler	Loss
1	4	5e-5	1:3	0.2	6	AdamW	Cross-Entropy
2	8	5e-5	1:3	0.2			
3	4	5e-5	1:2	0.4			
4	8	5e-5	1:2	0.4			
5	4	5e-5	1:1.4	0.2			
6	8	5e-5	1:1.4	0.2			
7	4	4e-5	1:3	0.2			
8	8	4e-5	1:3	0.2			
9	4	4e-5	1:1.4	0.4			
10	8	4e-5	1:1.4	0.4			

We found that both Setting 7 achieved the highest testing accuracy of 0.903, demonstrating the BERT model's strong ability to fit training examples and its superiority in semantic understanding. This validates our labeling process on the 2,000 random samples and subsequent screening on the entire Corpus 2.

In comparison, we also applied 10 statistic-based ML algorithms to predict labels for the 340 test samples. The results, shown in Figure 3.2, indicate that the performance of these statistic-based ML algorithms is generally lower than that of the BERT model, regardless of the hyperparameter settings used for the BERT classifier. The highest accuracy among the 10 algorithms, recorded at 0.797, was achieved by both the XGBoost and GBDT models. The superiority of the BERT classifier is primarily due to its ability to adjust its pre-trained weights and better comprehend the semantics—and, consequently, the content—of the corpus during training.

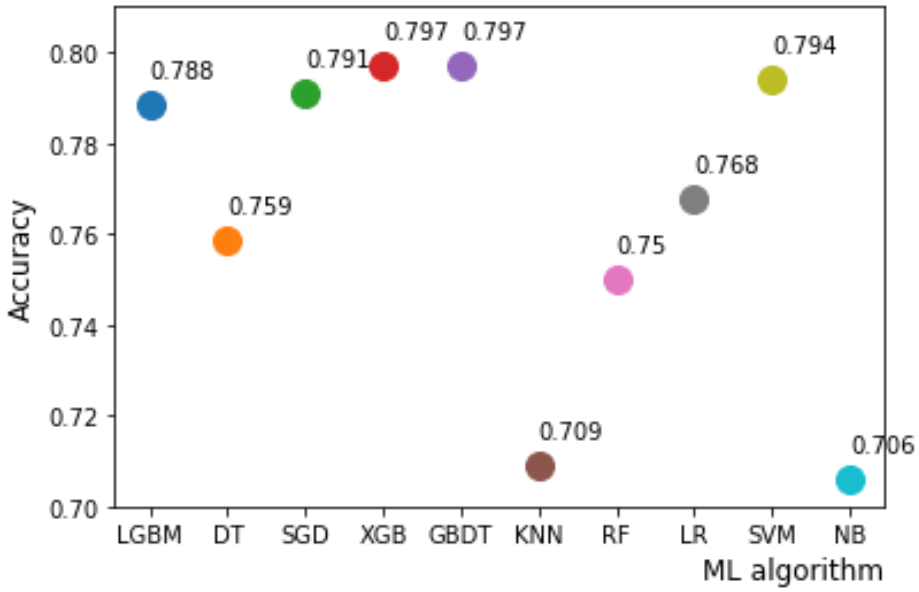


Figure 3.2. Testing accuracies of 10 ML algorithms

Therefore, we selected the BERT classifier as our final model, which was configured with

Setting 7 in Table 3.3. We then used this trained model to automatically label all sentences in Corpus 2, resulting in the creation of Corpus 3. For example, the 2nd sentence in Corpus 3 is "*For some construction companies, recent ransomware attacks have led to the loss of confidential data or a systems shutdown*", and the BERT classifier labels it as "Inclusion" with probability 0.998, meaning that this sentence should be included in our corpus with a high confidence score. The 189th sentence is "*Innovative building firms employ Building Information Modeling (BIM) as a central database for blueprints, designs, and other assets*", and the BERT classifier labels it as "Inclusion" with a probability of 0.687, meaning that this sentence can only be included in our corpus with a low confidence score. Sentences with probabilities lower than 0.5 should be excluded from our corpus.

3.3.2 The four corpora

Sections 3.2.1 through 3.2.3 involved collecting raw sentences and filtering them using keywords, resulting in the creation of Corpora 1 and 2. Section 3.2.4 involved the screening of sentences based on semantics, leading to the formation of Corpus 3. Section 3.2.5 discussed lemmatizing the corpus, yielding Corpus 4. This entire process culminated in a corpus that is richer in content and more semantically useful for our study, making it suitable for topic modeling. The statistical information for Corpora 1 through 4 is summarized in Table 3.4. At each step, there was a significant reduction in the number of sentences, words, vocabulary size, and storage size, with some reductions exceeding 50%. These reductions contribute to the creation of a corpus that is not only richer in content but also more semantically useful for our study. We used Corpus 4 for topic modeling.

Table 3.4. Statistics of the four corpora

Corpus No.	Description	No. of sentences (K)	No. of words (K)	Vocab size (K)	Size (MB)
1	Raw sentences	802	23,753	905	163
2	Screened by keywords	238	9,224	447	66
3	Screened by semantics	66	2,227	121	15
4	Lemmatization	66	1,305	34	11

3.3.3 Deciding the number of topics

In Section 3.2.6, we described how to use Corpus 4 for topic modeling through the LDA method. To determine the optimal input variable for the LDA model, we evaluated different numbers (5 to 30) using perplexity and coherence scores as metrics [102]. Figure 3.3 shows how the perplexity score decreases with an increase in the input variable while the coherence score fluctuates, generally decreasing. The input variable corresponding to 8 topics reaches the peak coherence score, indicating that the generated topics are highly coherent and meaningful. Also, its perplexity score is lower than for any number of topics fewer than 8. Although the perplexity score is slightly higher than for more topics, the high coherence score suggests that the topics are the most interpretable and semantically distinct at this number. Consequently, the final choice is 8 topics, which yield a perplexity score of -8.68 and a coherence score of 0.48.

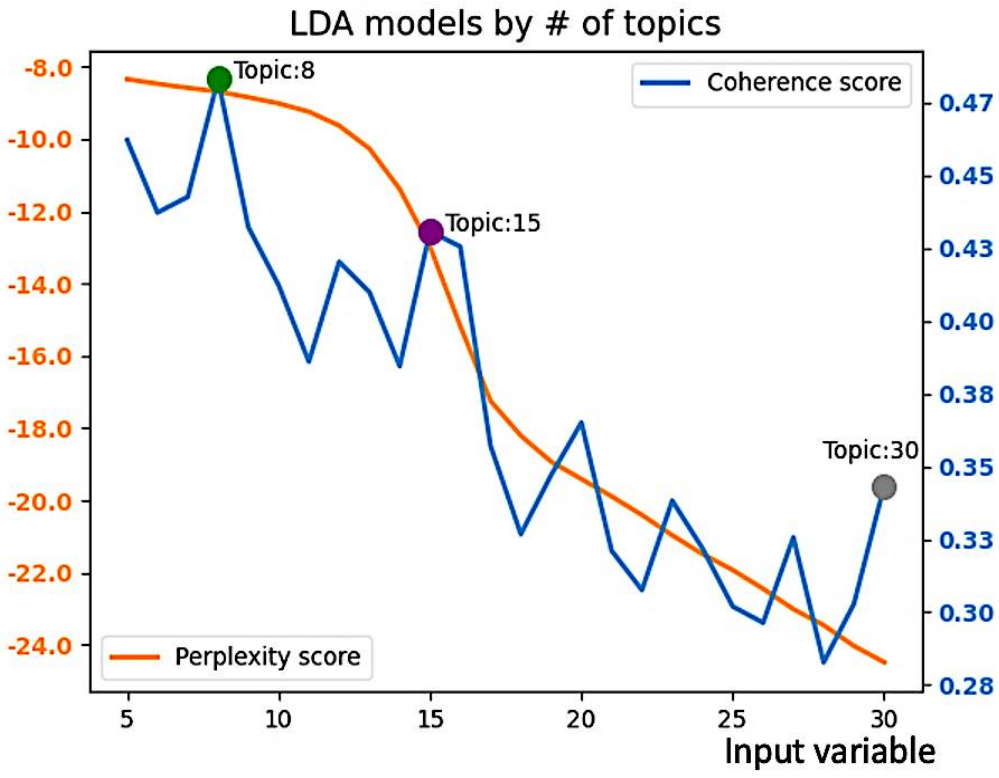


Figure 3.3. Perplexity and coherence scores

3.4 Topic summarization

Once the number of topics was selected, in this case, eight, they can be summarized based on two elements: (1) The top keywords associated with each topic. In LDA, each topic is characterized by a distribution over the vocabulary, and each word has a weight within a topic that signifies its importance or contribution to that topic. We extracted 10 keywords with the highest weight for each topic, which are displayed in Table 3.5; and (2) The representative original sentences for each topic. The representative sentences for each topic are those that have the highest percentage of contribution to that topic, as calculated using the `Gensim` library. We selected the 3 most meaningful sentences from a pool of the 10 sentences with the highest percentage of contribution for each topic, as shown in Table 3.6.

For the 1st topic, by combining the 10 keywords and the 3 representative sentences, we can infer and summarize the topic: *Increase awareness of construction firms by providing training and contact periodic risk management*. A similar process for Topic 1 was used to determine the remaining topics. The results are summarized in Table 3.7. In general, the 8 topics can be described as Perform Risk Management, Prevent the Increasing Cyber Incidents, Detect Ransomware, Strengthen Management Process, Protect Network Devices, Regulate Information Storage and Sharing, Protect Privacy, and Improve Authentication Process. The action to take to address these topics is also proposed in Table 3.7. These 8 identified topics and corresponding suggested actions can serve as a reference for researchers interested in studying future research directions in cybersecurity within the construction industry.

Table 3.5. Weights and keywords for each topic

No.	Weight	Keyword	Weight	Keyword	Weight	Keyword	Weight	Keyword	Weight	Keyword
1	0.046	security	0.038	cybersecurity	0.025	safety	0.023	risk	0.015	management
	0.015	training	0.014	area	0.013	provide	0.013	private	0.013	ensure
2	0.056	cyber	0.037	security	0.022	standard	0.020	sector	0.017	industry
	0.015	technology	0.014	development	0.014	dimension	0.014	report	0.013	impact
3	0.031	threat	0.030	infrastructure	0.029	national	0.029	critical	0.028	attack
	0.027	cyber	0.020	government	0.020	public	0.019	measure	0.015	communication
4	0.033	construction	0.023	risk	0.022	project	0.021	building	0.019	design
	0.018	system	0.018	management	0.018	build	0.013	personnel	0.012	may
5	0.040	system	0.029	http	0.029	use	0.029	control	0.028	security
	0.027	network	0.022	secure	0.021	access	0.019	internet	0.018	device
6	0.112	information	0.035	data	0.033	asset	0.028	personal	0.022	include
	0.022	use	0.020	level	0.017	access	0.013	specific	0.013	person
7	0.062	data	0.040	protection	0.035	activity	0.032	policy	0.029	breach
	0.023	require	0.018	state	0.017	privacy	0.015	key	0.015	manage
8	0.072	cybersecurity	0.043	develop	0.033	strategy	0.032	response	0.029	incident
	0.029	framework	0.028	security	0.025	number	0.023	practice	0.023	plan

Table 3.6. Representative sentences for each topic

Topic number	Percentage of contribution	Index in Corpus 4	Sentence
1	0.9562	41	Considerations for cyber security in construction industry. Given the risks, construction companies need to conduct comprehensive and frequent third-party cyber security management so they can identify and remediate vulnerabilities.
	0.9170	12799	Cyber security is not just an IT issue crucially, cyber risk should not be seen as an issue solely for your IT department or provider.
	0.9124	132	Many construction firms also lack awareness regarding cyber security.
2	0.8751	12623	Regardless of the industry, there is a worrying shift in the mindsets of security professionals.
	0.8750	3787	Rising cyber-attacks 75% of respondents reported an increase in cyber-attacks in the last 12 months.
	0.8541	24069	No doubt we are in a changing world, and cyber security is an ever-changing industry.

Table 3.6 (continued)

Topic number	Percentage of contribution	Index in Corpus 4	Sentence
3	0.9028	2079	You don't want to get ransomware, and the government also doesn't want you to get ransomware.
	0.8911	12676	Phishing emails also increased significantly during this period, by 20%, while malware soared by 423%.
	0.8908	12620	For example, supply-chain-based attacks, such as the SolarWinds SUNBURST attack, are not simple system vulnerabilities.
4	0.9126	9083	These risks can be physical and directly affect the insurance business, or they may be more transitional and affect insurers?
	0.8750	33717	Stakeholder-associated life cycle risks in construction supply chain.
	0.7368	31312	Due to known and unknown risks, the Company’s results may differ materially from its expectations and projections.
5	0.8488	27587	The Internet of Things (IoT) is a system where devices are connected wirelessly with the help of unique identifiers.
	0.8250	33593	Access to the software is not properly controlled.
	0.8125	20397	Centralized visibility allows organizations to control access for any device, connecting from any network.
6	0.8995	96	In addition to proprietary employee data, other potentially vulnerable information includes sensitive client data, tenant personally identifiable information (PII) and non-public material information.
	0.8542	52421	Are any assets (e.g., web server, web application) Internet-facing?
	0.8250	62329	What information is being shared, and what is the purpose of sharing it?
7	0.7912	48682	The attacked data was encrypted, and ransom payments were demanded in the Bitcoin cryptocurrency.
	0.7906	43072	After that, both processing and privacy should be determined before clarifying the requirement of granularity of trajectory data.
	0.7658	6360	They are triggered if: (1) there is a breach of contract, and (2) the penalty for breach is stated in the contract.
8	0.8257	15106	Here are some strategic approaches: Deploy Multi-Factor Authentication.
	0.8091	51	Implementing cyber security strategies to fortify endpoints and IT/OT integration while establishing a robust incident response plan will go a long way toward delivering peace of mind.
	0.7855	1426	Choose a good, strong password: Passwords are the weakest cybersecurity link.

Table 3.7. Summarized topics and actions to take

Topic No.	Keywords	Three representative sentences	Topic summarized	Some actions to take
1	security, cybersecurity, safety, risk, may, training, area, provide, private, ensure	• Construction companies need to conduct comprehensive and frequent third-party cyber security assessments so they can identify and remediate vulnerabilities. • Cyber security is not just an IT issue. Cyber risk should not be seen as an issue solely for your IT department or provider. • Many construction firms also lack awareness regarding cyber security.	Perform Risk Management	Researchers should take the lead to develop frameworks first.
2	cyber, security, standard, sector, industry, technology, development, dimension, report, impact	• Regardless of the industry, there is a worrying shift in the mindsets of security professionals. • Rising cyber-attacks 75% of respondents reported an increase in cyber-attacks in the last 12 months. • We are undoubtedly in a changing world, and cyber security is an ever-changing industry.	Prevent the Increasing Cyber Incidents	Industry practitioners should share the incident data.
3	threat, infrastructure, national, critical, attack, cyber, government, public, measure, communication	• You don't want to get ransomware, and the government also doesn't want you to accept ransomware. • Phishing emails have increased by 20%, while malware soared by 423%. • For example, supply-chain-based attacks, such as the SolarWinds SUNBURST attack, are not simple system vulnerabilities.	Detect Phishing and Malware	Software providers must constantly update their algorithms.
4	construction, risk, project, building, design, system, management, build, personnel, may	• These risks can be physical and directly affect the insurance business, or they may be more transitional and affect insurers. • Stakeholder-associated life cycle risks in the construction supply chain. • Due to known and unknown risks, the Company's results may differ materially from its expectations and projections.	Strengthen Management Process	Companies should adopt strict policies and train employees effectively.

Table 3.7 (continued)

Topic No.	Keywords	Three representative sentences	Topic summarized	Some actions to take
5	system, http, use, control, security, network, secure, access, internet, device	• The Internet of Things (IoT) is a system where devices are connected wirelessly with the help of unique identifiers. • Access to the software is not properly controlled. • Centralized visibility allows organizations to control access for any device, connecting from any network.	Protect Network Devices	Hardware manufacturers should integrate advanced security features.
6	information, data, asset, personal, include, use, level, access, specific, person	• Besides proprietary employee data, other potentially vulnerable information includes sensitive client data, and tenant personally identifiable information (PII). • Are any assets (e.g., web server, web application) Internet-facing? • What information is being shared, and what is the purpose of sharing it?	Regulate Information Storage and Sharing	Governments must establish strict data protection laws.
7	data, protection, activity, policy, breach, require, state, privacy, key, manage	• The attacked data was encrypted, and ransom payments were demanded in the Bitcoin cryptocurrency. • After that, both processing and privacy should be determined before clarifying the requirement of granularity of trajectory data. • They are triggered if: (1) there is a breach of contract, and (2) the penalty for breach is stated in the contract.	Protect Privacy	Users should be educated on best practices, and developers should prioritize privacy by design.
8	cybersecurity, develop, strategy, response, incident, framework, security, number, practice, plan	• Here are some strategic approaches: Deploy Multi-Factor Authentication. • Implementing cyber security strategies to fortify endpoints and IT/OT integration while establishing a robust incident response plan will go a long way toward delivering peace of mind. • Choose a good, strong password: passwords are the weakest cybersecurity link.	Improve Authentication Process	Companies should implement multi-factor authentication or biometrics.

3.5 Discussions

The LDA model used 8 as the input variable for topic modeling. The results are visualized in Figure 3.4 as a 2D mapping of the 8 topics grouped into clusters. The distances between the clusters indicate their correlation, with topics 1, 2, 3, and 8 being closely related, and topics 4, 5, 6, and 7 being more independent. The size of the bubble represents the prevalence of each topic, with topics 1, 2, and 3 being the most prevalent and, thus, prioritized.

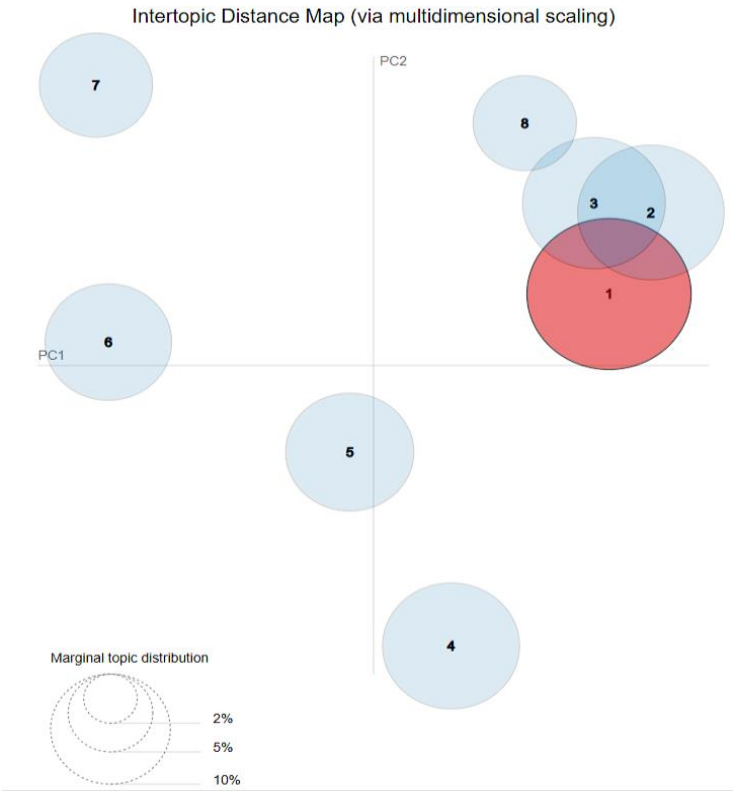


Figure 3.4. 2D mapping of eight topics

Based on the topic prevalence shown in Figure 3.4, we recommend prioritizing research on Topics 1, 2, and 3, which are respectively: Performing Risk Management, Preventing Cyber Incidents, and Detecting Phishing and Malware. These topics have significant implications for various stakeholders in the realm of cybersecurity and risk management.

They are discussed in detail in this section. Additionally, limitations and directions for future work will be addressed here.

Firstly, the topic of performing risk management encompasses the strategic overview and methodologies for identifying, assessing, and responding to cyber risks. It adopts a holistic approach that incorporates not just technical aspects, but also organizational, legal, and procedural frameworks to mitigate risks. Risk management is about establishing a comprehensive strategy that supports specific actions like incident prevention and response. Researchers play a pivotal role in this process. It is crucial for researchers to take the lead in proposing comprehensive frameworks or methodologies specifically designed to counter the evolving landscape of cyber threats. The development of these frameworks or methodologies must consider the unique vulnerabilities and characteristics inherent in construction projects. Moreover, it is vital for these efforts to be conducted in collaboration with industry practitioners. This collaborative approach ensures that the research is grounded in the practical realities of the industry. By pioneering robust frameworks and methodologies, researchers lay the groundwork for effective risk mitigation strategies that can be adapted and applied across the construction industry. This topic sets the stage for the subsequent exploration in the dissertation.

Secondly, the topic of preventing the increasing cyber incidents focuses on specific strategies and technologies designed to stop cyber threats before they inflict harm. Although it is a part of the broader risk management umbrella, prevention targets specific measures to avert incidents through enhanced security practices. These include securing network vulnerabilities, conducting employee training, and implementing advanced threat detection systems, all aimed at taking proactive steps to prevent threats from materializing. In this endeavor, researchers play a crucial role by developing sophisticated algorithms and models for predicting and preventing cyber incidents. Equally important are the contributions of industry practitioners. Their active participation in sharing incident data and collaborating with researchers and other stakeholders is vital. However, the sharing of cyber incident data currently faces several challenges, including concerns about company

reputation and privacy. Industry practitioners must realize that the open sharing of incident data fosters a collective enhancement in understanding emerging threats and strengthens defenses against cyberattacks.

Thirdly, the topic of detecting phishing and malware explores the technical detection mechanisms and solutions for identifying specific cyber threats. Detection serves as a crucial last line of reactive defense, initiating response actions when prevention measures fall short. This focus emphasizes the ability to quickly and accurately identify malicious activities to minimize damage. In combating phishing and ransomware, which have become increasingly prevalent in the construction industry in recent years, software providers hold a significant responsibility. Given the dynamic and constantly evolving landscape of cyber threats, email providers like Google and Outlook, as well as providers of construction-related software such as Autodesk Construction Cloud and Procore, must commit to the continuous and comprehensive updating of their detection algorithms. By prioritizing the development and deployment of advanced detection algorithms, software providers can empower organizations to swiftly identify and neutralize threats from ransomware, phishing, and other types of malware, thereby safeguarding vital assets and infrastructure.

In summary, the topic of performing risk management provides an overarching framework for identifying, assessing and responding to cyber risks through comprehensive strategies. The topic of preventing the increasing cyber incidents focuses on proactive measures to avert such incidents by securing vulnerabilities and enhancing security protocols. The topic of detecting phishing and malware specializes in identifying specific threats like phishing and malware, acting as a reactive measure of defense when preventive measures fail. Together, these three topics create a layered approach to cybersecurity, emphasizing the importance of holistic strategic planning, proactive prevention, and reactive detection in safeguarding against cyber threats.

One limitation of this study is that the corpus is not as clean as it could have been due to

the large workload of the sentence cleaning and time constraints; some samples contain irregular characters and symbols from the original websites crawled or the original files collected. Although it is believed that these irregular characters and symbols do not affect the quality of topic modeling because the LDA topic modeling technique is robust to noise and outliers, further cleaning of the corpus is recommended for future studies utilizing it. The refined corpus can be used for other purposes, such as training large language models to create question and answer systems (like the ChatGPT dialogue system released in November 2022 [105]), which can provide relevant individuals with preliminary consulting advice on cybersecurity management, such as risk identification and solution suggestion. This will be the focus of our future research.

3.6 Conclusions

To bridge the gap in the scattered landscape of cybersecurity research within the construction industry and foster cohesive understanding and unified progress in construction cybersecurity, this chapter employs topic modeling to uncover potential research directions in construction cybersecurity. We gathered 6 types of raw text sources, which were transformed into 802K raw sentences. Through a process of keyword and semantic screening, the data were reduced to 66K sentences, forming the corpus (Corpus 4) for our topic modeling analysis. We utilized the LDA method and tested input variables ranging from 5 to 30, evaluating and selecting the model based on perplexity and coherence scores. Our analysis revealed that 8 was the most suitable number of topics. The results of our analysis identified 8 potential topics for future research, and corresponding actions to take are also proposed.

This study recommends research on three main topics: performing risk management, preventing cyber incidents, and detecting phishing and malware. Together, these form a layered approach to cybersecurity tailored to the construction industry, combining holistic strategic planning, proactive prevention, and reactive detection. Risk management in cybersecurity involves a holistic approach, emphasizing strategic methodologies, technical,

organizational, legal, and procedural aspects, with crucial collaboration between researchers and industry practitioners. Preventing cyber incidents focuses on strategies and technologies to preemptively stop threats, involving network security, employee training, and advanced detection systems. Collaboration between researchers and practitioners is crucial, despite challenges in data sharing due to privacy and reputation concerns. Detecting phishing and malware is critical in cybersecurity, serving as the last defense line by detecting threats to minimize damage when prevention fails. Software providers, including email and construction-related software companies, must continuously update detection algorithms to combat evolving cyber threats effectively. The study's limitation lies in the corpus's cleanliness, with time constraints and workload leading to samples with irregular characters. Future research should aim to refine the corpus, making it suitable for applications such as training large language models that can serve as intelligent cybersecurity consultants.

Chapter 4
Enhancing Cyber Risk Identification in the Construction Industry: A Language Model Approach

In this chapter, we leverage the capability of a language model to analyze the text corpus created in Chapter 3, innovatively tailoring it to identify cyber risks relevant to diverse construction projects. A language model is trained with the corpus, enhanced by Supervised Fine-Tuning and Reinforcement Learning from Human Feedback techniques. Utilizing this model, a comprehensive prioritized checklist of cyber risks across various project phases is compiled. We present a two-layered evaluation approach: (1) evaluating the model's suitability in the task of identifying cyber risks and (2) validating the compiled list of cyber risks. The evaluation methods incorporate both quantitative and qualitative aspects. This study then discusses the applicability of the checklist, the high-level risk mitigation strategy, the prospect of the language model, and tips on enhancing the dataset for upgrading the model for industry-wide application.

4.1 Introduction

The construction industry, entering the Construction 4.0 era, significantly lags in cybersecurity awareness, making it a prime target for cyberattacks that could lead to project delays, financial losses, and other detrimental consequences [3]. This vulnerability has been highlighted by incidents such as Turner Construction falling victim to a spear-phishing scam [106], Marous Brothers Construction not receiving payment due to maliciously changed routing numbers [10], Bird Construction being breached by ransomware [107], and Hoffmann Construction reporting unauthorized access to employee information [108]. A key factor underlying these incidents, as noted in [109], is the industry stakeholders' lack of full recognition of the cyber risk, which arise from threats (malicious actions aimed at exploitation) and vulnerabilities (specific weak points susceptible to exploitation) inherent in construction projects. This lack of recognition results in the absence of a benchmark checklist and, consequently, a lack of comprehensive preventive measures, which are crucial for protecting sensitive data and maintaining operational integrity.

Therefore, it is imperative to clearly identify cyber risks, including threats and vulnerabilities. Doing so can raise the awareness of project managers in construction companies and facilitate the implementation of preventive and proactive measures against cyber attacks. To ensure broad relevance and applicability across construction projects, these cyber risks can be categorized according to the different project phases: initiation, design, construction & procurement, commissioning, operation & maintenance, renovation, and end of life. This categorization, as recognized in the literature [109], reflects the universal progression of construction projects, regardless of their size, location, or delivery method. Aligning risk identification with these phases enables the construction industry to recognize and tackle the widespread challenges of cybersecurity more effectively.

However, as discussed in Section 4.1.1, existing studies on cyber risk identification in the construction industry are limited and do not provide a comprehensive, industry-specific list

of cyber risks that could serve as a credible benchmark. Furthermore, the methods used in these studies exhibit several drawbacks: they rely heavily on manual implementation, which can introduce bias and oversights due to the varying levels of expertise among the implementers. These methods do not support dynamic risk identification in evolving scenarios, as they require starting from scratch for each new assessment, thus lacking adaptability and flexibility. Additionally, these approaches are time-inefficient, necessitating lengthy discussions and checks among implementers. Given the construction sector's growing dependency on computational tools and the rapidly evolving cybersecurity landscape, it becomes imperative to shift to an automated approach, which needs to offer greater comprehensiveness, dynamism and efficiency.

Recently, the surge in NLP techniques has led to an increased use of text analysis for risk identification and management across various industries, as shown by [110], [111], [112]. Considering the abundance of textual sources related to cybersecurity and construction, such as news articles, academic papers, and published reports [113]—both existing and forthcoming—these can be meticulously analyzed to identify emerging cyber risks in construction. Inspired by large language models like Baidu's Ernie Bot [114], OpenAI's GPT-4 [115], and Google's Gemini [116], LLMs trained with SFT (Supervised Fine-Tuning) and RLHF (Reinforcement Learning from Human Feedback) techniques can synthesize and analyze these abundant text sources, which allows to identify cyber risks comprehensively. In the future, the developed model can continue to receive periodic training with new text data. This will ensure it remains aligned with the evolving cybersecurity landscape in construction and facilitates regular updates to the identified cyber risks, enabling dynamic cyber risk identification. Continuous training with new data is more efficient than traditional methods, which require starting from scratch with human intervention, thereby enhancing efficiency [117].

4.1.1 Related works

Few studies have been found on threat identification in the construction industry. For

instance, Shemov et al. [41] presented a threat model for a specific case scenario at the initial stage of a construction supply chain to identify potential attacks and how to prevent them. This threat model took the form of a systematic process for cybersecurity risk management, ranging from data collection to the implementation of a security model, and included the analysis of system elements, weaknesses, threats, vulnerabilities, and response strategies. Mohamed Shibly and García de Soto [49] developed a construction-focused threat modeling approach based on the Quantitative Threat Modeling Method (QTMM). This method combines system definition, data flow diagrams, threat mapping with STRIDE, threat identification, quantitative risk assessment using attack trees and CVSS, and the planning and implementation of risk mitigation strategies. They applied this threat model to an offsite 3D concrete printing system to identify vulnerabilities and potential countermeasures. Mantha et al. [109] presented a preliminary cybersecurity threat model for the AEC (Architecture, Engineering, and Construction) industry, identifying cyber risks across different stages of construction projects. This model involves selecting a lifecycle phase, identifying critical assets, verifying cyber-enabled components, determining potential threats and attackers, identifying vulnerabilities, and developing countermeasures for effective cybersecurity management. They demonstrated its feasibility with an example from the commissioning phase of a building.

It can be seen that these studies predominantly focus on threat identification and there seems to be a lack of comprehensive vulnerability identification within these models. Additionally, the implementations of these methods are mostly manual, which is time-consuming and may incur bias. Furthermore, they are not updatable over time, failing to reflect the evolving nature of cybersecurity threats and vulnerabilities. This underscores the need for a more automated, objective, and dynamic solution for cyber risk identification, which involves both threat identification and vulnerability identification.

As stated in Section 4.1, the abundance of text-based resources relating to cybersecurity and construction — including news articles, academic papers, and reports — presents an opportunity for language models to analyze and identify emerging cyber risks (both threats

and vulnerabilities) within the construction sector. Language modeling is typically represented as the probability of a sequence of words ($w_1, w_2, \ldots, w_n$) to represent the joint probability of a sentence. Utilizing the chain rule of probability, this concept is broken down into a series of conditional probabilities, as demonstrated in Equation (4.1) [118].

$$P(w_1, w_2, \ldots, w_n) = P(w_1) \times P(w_2|w_1) \times P(w_3|w_1, w_2) \times \ldots \times P(w_n|w_1, w_2, \ldots, w_{n-1}) \quad (4.1)$$

Language models aim to maximize the probability of sentences that appear in the ground truth corpus dataset. This enables them to generate sentences that fall within the same distribution as the dataset, which has been exceptionally performed by modern LLMs. The growth of LLMs has significantly influenced the field of NLP, with applications extending across various industries. This development can be traced back to the introduction of the transformer architecture in 2017 by Vaswani et al. [62]. The transformer's self-attention mechanism and parallel computation enabled the training of more semantically rich word embeddings. In 2018, OpenAI introduced GPT [119], a transformer-based architecture trained via unsupervised pre-training and fine-tuning for specific NLP tasks. Google's BERT [97], also introduced in 2018, utilized transformer architecture and improved performance on many NLP tasks. In 2019, OpenAI released the larger GPT-2 [120] with remarkable language generation capabilities, and in 2020, they launched GPT-3 [121], notable for its zero-shot/few-shot learning ability.

Since the introduction of GPT-3, LLMs with billions of parameters have been developed, showing strong performance across many tasks, such as language generation, question answering, and machine translation. LLMs are significantly expanding AI's impact in various domains, including healthcare (e.g., medical image analysis, drug discovery), gaming (e.g., game development, chatbots), finance (e.g., fraud detection, risk management), robotics (e.g., natural language interaction, sensor data analysis), and enterprise software development (e.g., code completion, testing). In 2023, OpenAI introduced its latest model, GPT-4 [115], a large multimodal model capable of accepting text and image inputs and generating text outputs, exhibiting human-level performance on

various benchmarks, and offering improved reliability and creativity over its GPT-series predecessor. In addition to GPT-4, numerous other LLMs have been developed, including ChatGPT [122], PaLM [123], and LaMDA [124].

The SFT and RLHF techniques represent an important milestone for LLMs, as it significantly enhances their capabilities by incorporating human preferences into the fine-tuning process through the use of reinforcement learning. RLHF incorporates human evaluations to align models with nuanced human values. An initial language model is pre-trained on textual data and fine-tuned using a reward model generated from ranked human feedback. The optimization process employs algorithms like Proximal Policy Optimization (PPO) [125] or the REINFORCE algorithm [126] to maximize alignment with human preferences. SFT and RLHF techniques have been successfully applied in models like GPT-4 (OpenAI) [115], Gopher (DeepMind) [127], and Claude (Anthropic) [128], resulting in improved performance in natural language tasks. Through iterative updates, the techniques enable more human-like and reliable language models for diverse applications.

In summary, given the insufficient recognition of cyber risks in various construction projects, the limited literature where most methods require significant human involvement, the volume of construction cybersecurity-related text data both existent and expected, and the inspiration from large language model applications in other industries such as finance [129], this study aims to develop a language model for identifying cyber risks across project phases.

4.1.2 Objectives

The goal of this study is to comprehensively identify cyber risks across project phases, achieved by innovatively developing a language model dedicated to construction cybersecurity. This study is structured around three objectives:

(1) To pre-train a base model that encodes knowledge of both cybersecurity and construction. The candidate base models are the GPT-2 model, the BERT model

with a language modeling head, and the T-5 model with a language modeling head. These candidates will be trained with the text corpus created in Chapter 3. Due to limited computing resources, training can be accomplished with a single GPU, which serves as proof of the scalability and applicability of the model for larger models used for industry-wide applications.

(2) To enhance the model's ability to understand and generate answers about cybersecurity content and questions by implementing SFT and RLHF techniques. The SFT technique will be used to align it with our final downstream task, enabling it to understand the specific requirements of the task and produce more relevant and accurate outputs. The RLHF training technique will be used to make the model more generalizable to unseen questions.

(3) To compile a comprehensive cyber risk checklist categorized by project phases, which will serve as a new benchmark that can be referred to for the formulation of preventive measures. It should be reiterated that the identified cyber risks are applicable to diverse projects rather than specific ones, and the detailed risk mitigation strategies for each risk are beyond the scope of this study.

4.1.3 Contributions

The academic contributions of this study include: (1) The novel development and application of techniques dedicated to construction cybersecurity, promoting interdisciplinary research in AI, cybersecurity, and construction management. (2) Bridging the gap in the comprehensive recognition of cyber risks across project phases, providing a new benchmark in the form of a cyber risk checklist. (3) The possibility of using the developed framework for identifying risks in other industries.

Practically, the cyber risk checklist offers several benefits: (1) It is valuable for diverse construction projects, particularly valuable in assisting project managers to check their cybersecurity status and formulate proactive and preventive cybersecurity measures

against prioritized risk items. (2) Risk analysts can leverage this checklist for in-depth risk analyses on specific construction projects, prioritizing their efforts by starting with the most critical risks in the list. (3) IT and cybersecurity teams, stakeholders and general personnel of construction companies can benefit as well. As the checklist can be regularly updated, people will have access to the latest information on the cybersecurity landscape, ensuring they remain informed about current trends.

4.2 Methods

4.2.1 Language model development overview

Our language model development is based on the framework described in [70], where a language model named InstructGPT was trained by implementing SFT and RLHF techniques. By incorporating human feedback, these techniques enhance the model's ability to generate text that aligns more closely with human expectations. Building on this, our study adapts the SFT and RLHF techniques with the goal of training a language model that can understand cybersecurity content and generate answers to cyber risk identification queries.

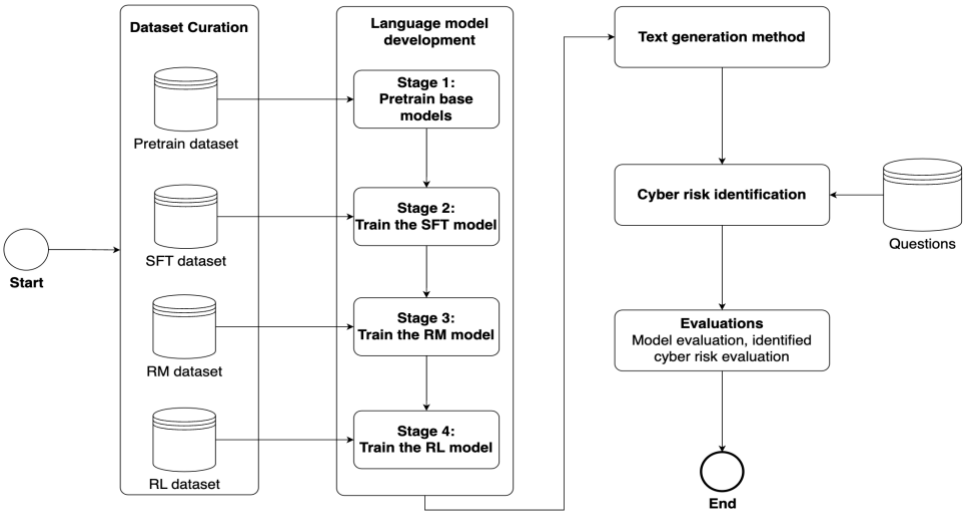


Figure 4.1. The methodological flowchart

As illustrated in Figure 4.1, the model development process encompasses four stages.

Stage 1: Collect a large corpus and pretrain a base model. We initially identified a substantial text corpus (pretrain dataset) on cybersecurity and construction. Then, we pretrained a language model on this corpus to encode the domain knowledge. This pretrained model serves as the base model.

Stage 2: Compile a question-answer dataset and train a supervised model (SFT model). We have developed a dataset of question-answer pairs related to construction cybersecurity (SFT dataset) to fine-tune the base model, which allows the model to capture diverse question formats and styles and results in our SFT model.

Stage 3: Compile a scoring dataset and train a reward model (RM model). We compiled a dataset of question-answer pairs, where each label represents a score indicating the degree of human preference for the answer to the question (RM dataset). A reward model, built on the SFT model, was trained to emulate the human scoring. This reward model played an important role in the RLHF technique implementation in the next stage.

Stage 4: Compile a question dataset and train a reinforcement learning model (RL model). We compiled a new dataset (RL dataset) and used the RM model to score answers from the RL model. This score was then used to fine-tune the RL model with a specially designed loss, enhancing its generalization to unseen prompts. This resulted in the final RL model.

After training the RL model, it was employed to generate text for answering questions related to cyber risk identification. From the answers, we compiled a list of cyber risks applicable to various phases of a construction project. The evaluation included both the assessment of the model's performance progress and the scrutiny of identified cyber risks. Specifically, the model evaluation concentrated on the model's advancement through different stages in understanding construction cybersecurity content and generating responses to cybersecurity questions. The evaluation of the identified cyber risks involved

diverse sources, such as comparisons with existing literature, expert evaluations, and evaluations by the state-of-the-art GPT-4 model.

4.2.2 The dataset for each stage

Four datasets were utilized, each related to either the construction industry or cybersecurity. The pretraining dataset is derived from our previous work [113], while the other three datasets were carefully curated by us over a four-month period. To ensure alignment with each model's objectives, all prompts underwent rigorous scrutiny.

4.2.2.1 Pretrain dataset for Stage 1

The pretrain data, from our previous work [113], comprises 66K sentences from a large text corpus spanning six cybersecurity-related sources in the construction industry. The corpus underwent semantic screening, using a BERT classifier to assess each sentence's relevance to construction cybersecurity by assigning a probability to its semantic suitability for the dataset. In this study, we excluded sentences assigned a weight not greater than 0.75 by the BERT classifier. This decision was based on our observation that probabilities below this threshold often approached or fell below 0.5, indicating the BERT classifier's low confidence in the semantic relevance of these sentences. This filtering left us with 61,841 sentences, constituting 93.7% of the original dataset, shown in Table 4.1. Recognizing the higher quality of academic writing, sentences from academic publications and book chapters were weighted at 1.5, making up 32.9% of our corpus. Therefore, each sentence thus received two weights: one for semantic quality from the BERT classifier and another for academic value. These weights were integrated into the loss function of the base model shown in Equations (4.2) and (4.3).

4.2.2.2 SFT dataset for Stage 2

For the SFT stage, we curated 326 questions related to the construction or cybersecurity domain, each with a complete answer. To enable the model to adapt to different linguistic

styles, each question was rephrased into four alternative sentences, capturing both interrogative and imperative forms. The interrogative form typically ends with a question mark and is used for seeking information, while the imperative form usually ends with a period and is used for giving commands or requests. This allowed the model to recognize the same query irrespective of its linguistic presentation and generate the same answer, enhancing its adaptability. Therefore, the dataset comprises 326 × 5 = 1630 question-answer pairs. 5 example question-answer pairs related to the cybersecurity domain are shown in Table 4.2, and the complete dataset can be found on our GitHub page [130].

Table 4.1. Details of the pretrain dataset

Text source	Descriptions	Sentence Count	Total count
News articles and blogs	Cover incidents, causes, processes, outcomes; offer timely insights and diverse perspectives	6 K	62 K
LexisNexis databases	Provides legal, business, news, and records; regularly updated for current information	46 K	
Academic papers	Deliver professional knowledge, emphasizing relevance and rigor across various research fields	2 K	
Books (chapters)	Comprehensive domain knowledge for in-depth understanding through detailed analyses	5 K	
Specifications/Standards	Detail cybersecurity legal/industry requirements, ensuring compliance and best practices	2 K	
Company reports	Reflect industry insights, strategies, and outcomes from a corporate perspective	1 K	

Table 4.2. SFT dataset examples

Question	Answer
How are cyber threats identified? How are cyber threats commonly detected? Can you explain how cyber threats are identified? Describe the process of identifying cyber threats. Explain the methods for identifying cyber threats.	Cyber threats may be identified through various means, including threat intelligence, security information and event management (SIEM), and user behavior analytics.

4.2.2.3 RM dataset for Stage 3

The RM dataset initially consisted of 150 questions. For each question, we used the SFT model to generate four answers, resulting in a total of 600 question-answer pairs. The

format of each pair is: "Question: {}; Answer: {}." Each pair was scored from 0 to 100 based on following criteria: Incomplete Answer (0-20) for answers missing essential information; Mention of Relevant Terms (20-40) for answers with related terms but lacking depth; Partially Related Answer (40-60) for answers somewhat addressing the question but missing key details; Logical Answer (60-80) for answers that are logical but not exhaustive; and Comprehensive and Logical Answer (80-100) for fully detailed and logically sound answers. To reduce bias, we held two scoring sessions spaced two days apart, presenting question-answer pairs in random order each time. The final score was the average of these two sessions. The RM dataset is available on our GitHub page [131].

4.2.2.4 RL dataset for Stage 4

We curated 353 questions (without answers) related to the construction industry or cybersecurity, which are different from those in the SFT dataset. These questions, created also in either interrogative or imperative form, were used for training the RL model and can be found on our GitHub page [132].

4.2.3 The model for each stage

This section presents the pretraining, SFT, and RLHF techniques for model training. Details of the model training and selections will be provided in Sections 4.3.1. Model progress evaluation methods and results are provided in Sections 4.2.6.1 and 4.3.2, respectively.

4.2.3.1 Base model for Stage 1

Pretraining a base model with a domain-specific corpus is crucial as it serves as a strong starting point for encoding cybersecurity knowledge and insights. In [70], the GPT-3 model, which has 175 billion parameters, was chosen as the base model. However, due to limited computing resources and the relatively smaller size of our dataset, we experimented with three types of smaller models.

- GPT-2: A generative language model by OpenAI, GPT-2 [120] is notable for human-like text generation, widely applied for task automation and insights generation. It is a precursor to GPT-4 [115], utilizing an autoregressive mechanism for word prediction based on prior context, and is pretrained on a comprehensive dataset.
- BERT-LM: Google's BERT [97] is a powerful transformer model acclaimed for its deep semantic understanding through bidirectional context analysis. We reset the configuration to make it only see the prior context and not what follows so that it can be used for autoregressive language modeling.
- T5-LM: Google's T5 [133], an encoder-decoder model, excels at translation and summarization via a uniform text-to-text approach. It features a BERT-like encoder and a GPT-2-like decoder. Our focus is on the decoder for language modeling, setting the encoder's input as "Language modeling task" without updating its weights.

The typical autoregressive training objective for language models uses Cross-Entropy loss [118], [57]. As stated in Section 4.2.2.1, considering the weight indicating each sentence's semantic relevance and academic value, we integrated the two weights into a combined weight (Equation (4.2)), which was then incorporated into the pretraining loss function in Equation (4.3) to scale the loss of the corresponding sentence.

$$w_i = \frac{w_{i,w_1} \cdot w_{i,w_2}}{\min\limits_{1 \le i \le N} \left(w_{i,w_1} \cdot w_{i,w_2} \right)} \tag{4.2}$$

$$L_{\text{weighted}} = -\frac{1}{\text{batch\_size}} \sum_{i=1}^{\text{batch\_size}} w_i \sum_{k=1}^{n_i} \sum_{c=1}^{|V|} y_{i,k,c} \log(P_{i,k,c}) \tag{4.3}$$

Where $|V|$ is the vocab size for each model; w_{i,w_1} is the original weight assigned to the *i*-sentence; w_{i,w_2} indicates the weight from academic sources, which is 1.5 if from academic

sources otherwise 1; w_i is the relative combined weight for the *i*-sentence; n_i is the number of words in the *i*-th sentence; $y_{i,k,c}$ is 1 if the ground-truth label for the *k*-th word in the *i*-th sentence is the *c*-th class of word, otherwise 0. The ground-truth label is the next word of the *k*-th word; $P_{i,k,c}$ is the predicted probability that the next word of the *k*-th word in the *i*-th sentence is the *c*-th class of word.

4.2.3.2 SFT model for Stage 2

Our ultimate goal is triggering the model to generate answers that highlight possible cyber risks, given a question or prompt. Although the base model has encoded construction cybersecurity knowledge, the model should be aligned with our final downstream task so that it can understand the specific requirements of the task and produce more relevant and accurate outputs [70]. To this end, we implemented SFT technique to further train the model, an approach proved effective by various LLM studies [70], [134], [135]. This process involves fine-tuning our base model using our customized question-answering pair dataset, which is the SFT dataset in our study. It includes question-answer pairs in various formats, styled to resemble diverse styles of questions asked by different humans, that the model is required to recognize. We concatenated each pair of questions and answers with a "\n" token in between to make it a coherent string available for the model's autoregressive training [70].

4.2.3.3 RM model for Stage 3

To increase the model's generalization ability of answering diverse questions beyond those in the SFT dataset, the RLHF technique [126] can be implemented. This involves providing feedback to the generated answers and optimizing the model against the feedback. In InstructGPT [70], the authors trained a reward model as the feedback provider, which is built on its SFT model and trained to predict a scalar value as the reward feedback. They designed a comparison mechanism between two different answers to the same question, which is incorporated into the loss function. Similarly, we added a regression head to our

SFT model to create the reward model. However, due to the smaller size of our model and dataset, we adopted a more straightforward approach for reward model training. The model is trained to predict a score on a continuous scale from 0 to 100, with the objective of minimizing the difference between the predicted and the ground truth score labeled by humans.

4.2.3.4 RL model for Stage 4

Following [70], we employed the RLHF technique to further fine-tune the SFT model. This step aims to enhance the model's ability of generalization to unseen questions. Figure 4.2 illustrates the RL fine-tuning process. Each question from the RL dataset is processed by the RL model to generate an answer, and the resulting question-answer pair is then evaluated by the RL model, the reward model, and the SFT model. Additionally, some data from the pretrain dataset are also processed by the RL model. These steps result in a loss function with three terms as shown in Equation (4.4), adapted from [70].

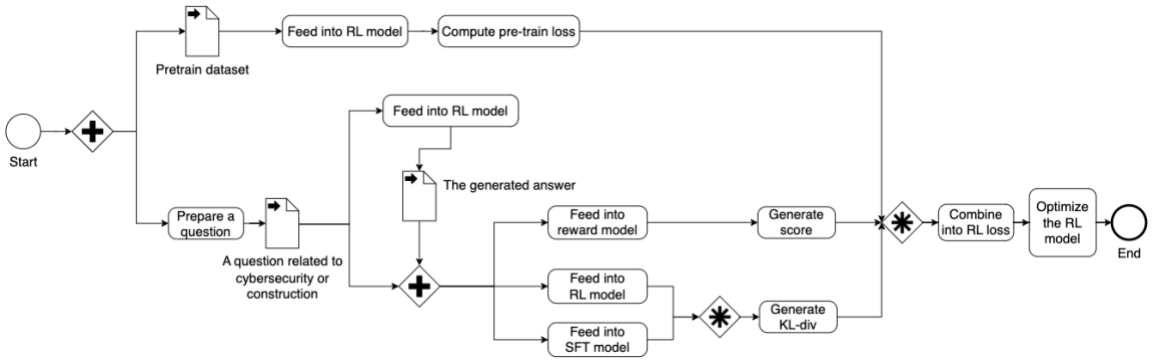


Figure 4.2. Overview of RL fine-tuning

$$l_{RL} = -w_1 \cdot \frac{1}{T}\sum_{i=1}^{T}\left(\log(a_i)\cdot\frac{\text{Score}(x,y)}{100}\right) + w_2 \cdot KL(\text{RL}(y|x),\ \text{SFT}(y|x)) + w_3 \cdot E_{x'\sim D_{pretrain}}\text{Pretrain}(x') \tag{4.4}$$

(1) Score term. This term incorporates the reward output by the RM model, evaluating the quality of the generated answer. This reward is normalized to fall between 0 and 1 and then

multiplied by the log probability (a_i) associated with each token (word) in the generated answer. The resulting product is averaged over the sentence length, T. This term is consistent with REINFORCE algorithm, a standard policy gradient optimization method [126].

(2) KL term. The per-token Kullback-Leibler (KL) divergence penalty, shown in Equation (4.5) [136], measures the discrepancy between the probabilities of sentences generated by the RL model and the SFT model. This penalty ensures that the answers generated by the RL model remain consistent with the knowledge acquired during the SFT stage, promoting more coherent answers. Additionally, it mitigates the risk of the RL model over-optimizing to the reward model, which could lead to the generation of nonsensical content [70].

$$KL(P_1, P_2) = \frac{1}{N}\sum_{i=1}^{N}\left(\sum_{j\in V} P_{1_i}(j)\, log\left(\frac{P_{1_i}(j)}{P_{2_i}(j)}\right)\right) \tag{4.5}$$

Where $P_{1_i}(j)$ is the probability of choosing the j-th token at the i-th position in the generated answer by RL model, $P_{2_i}(j)$ is the probability of choosing the j-th token at the i-th position in the generated answer by the SFT model, N is the number of tokens in the generated answer, and V is the set of all possible tokens in the vocabulary.

(3) Pretrain term. The pretrain term also helps constrain the RL model to avoid over-optimizing, ensuring the model retains its generative capability aligned with the distribution of the pretrained dataset. The inputs are minibatches from the pretrained dataset, and the mean Cross-Entropy loss within a minibatch is computed. Due to computing resource limitations, we only extract 20% of the sentences from the pretrained dataset with the highest combined weights (ensuring the model can still capture the distribution of the most weighted sentences) and divide them into a number of batches equal to the length of the dataset for the RL model.

4.2.4 Autoregressive text generation

Through this study, the standard algorithm for text generation is employed, which involves generating each word autoregressively [115], [118], [120], [124], [133]. The maximum length of tokens is to be 40 including the question part, which is deemed reasonable for a long sentence for the GPT-2 tokenizer [120]. Moreover, beam search [137] was adopted during the generation process to ensure the generated sentences have high joint probability. The pseudo code is shown as Algorithm 4.

Algorithm 4: Autoregressive Text Generation

```
BEGIN
    Tokenize the question using GPT-2 tokenizer [120]
    Convert each token into its numerical index in GPT-2 vocabulary
    ENCODE_SEQUENCE
        For each token in the question
            Encode the token
        End For
    END ENCODE_SEQUENCE
    GENERATE_TOKENS
        Initialize token_sequence with the encoded question
        While length of token_sequence
            Predict the next token using the model
            Append the predicted token to token_sequence
            Use beam search to generate token_sequence
        End While
    END GENERATE_TOKENS
    Decode token_sequence into a natural language sentence
    Return the decoded sentence as the answer
END
```

4.2.5 Cyber risk identification

The final RL model was tasked with identifying cyber risks across construction project phases. This was achieved by inputting crafted questions into the model, which then generated answers. To align the questions with the styles used in our SFT dataset, various phrasing structures were crafted, as shown in Table 4.3. To ensure the model could recognize different expressions of cyber risks and phases, we varied the keywords used

within these questions. According to [138], risk identification includes threat and vulnerability identification, so the first set of keywords, {key1}, encompasses “cyber risks”, “threats”, and “vulnerabilities”. The second set, {key2}, encompasses terms corresponding to different project phases, as shown in Table 4.4. This classification of phases aligns with the one in previous work on threat modeling [109]. The generated answers were subsequently reviewed and transformed into a checklist of identified cyber risks, ensuring the contextual accuracy and quality.

Table 4.3. Structuring of cyber risk identification questions

No.	Phrasing
1	What are the {key1} in the {key2} phase?
2	Can you identify {key1} in the {key2} phase?
3	{key1} in the {key2} phase include
4	Identify {key1} in the {key2} phase.
5	What types of {key1} should be considered during the {key2} phase?

Table 4.4. Different expressions for project phases (Key2)

Phase	Different terms
Initiation	pre-planning, concept, feasibility study, conception, early project definition, inception
Design	design development, schematic design, detailed design, planning
Construction &Procurement	construction, procurement, build, execution, implementation
Commissioning	handover, startup, completion, closeout
Operation & Maintenance	operation, maintenance, operations phase, facility management, service phase
Renovation & End of Life	renovation, end of life, demolition, decommissioning, retrofitting, rehabilitation, disposal

For each phase, the likelihood of each risk was derived from the RL model. We first formatted an identified risk, *r*, into a prompt: ‘{key1} in the {key2} phase includes {*r*}'. The prompts together make up the prompt set S_r for the phase. Then, the average probability over the number of tokens and the number of prompts in the set is computed, presented as Equation (4.6). The probability of all risks in the phase was normalized to a range from 1 to 5 using min-max normalization [57] for later comparisons with other assessors [139].

$$P(r) = \frac{1}{|S_r| \cdot T} \sum_{s \in S_r} \sum_{t=1}^{T} P(x_{s,t} | x_{s,1:t-1}) \quad (4.6)$$

Where T is the token count of prompt s, $x_{s,t}$ is the t-th token of s, $x_{s,1:t-1}$ are all tokens before the t-th token.

Detailed in Section 4.2.6.2, two industry experts and the GPT-4 [115] served as additional assessors to give likelihood assessment on a 5-point Likert scale [139], [140], [141], aiming to compare with our RL model for evaluation. These additional assessors were also requested to assess other attributes including relevance, impact, and consequence. Finally, for each risk, the likelihood and impact levels were averaged across all assessors. The risk value is calculated as the product of the averaged likelihood and impact levels [141], [142]. This value was then categorized into five categories according to the risk matrix adapted from the literature [141], [142], [143] as shown in Figure 4.3: Very Low (VL, 0-5), Low (L, 6-10), Medium (M, 11-15), High (H, 16-20), and Very High (VH, 21-25).

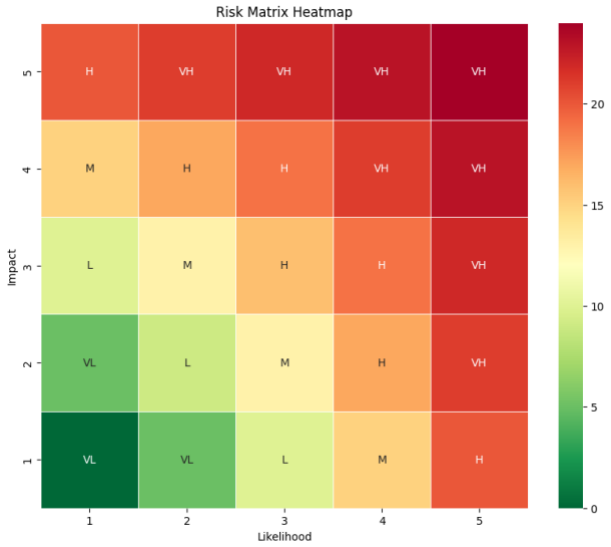


Figure 4.3. The 5 × 5 risk matrix

4.2.6 Evaluations

This study employs a two-layered approach for evaluation: (1) evaluating the model's

progress, demonstrating its suitability for the task of identifying cyber risks, and (2) evaluating the compiled list of cyber risks, validating its applicability. The methods are multifaceted and incorporate both quantitative and qualitative approaches.

4.2.6.1 Evaluating model progress

As language models are fundamentally focused on understanding and generating text [117], we evaluated our model's ability to understand cybersecurity content, facilitated by domain knowledge encoding, and to generate responses to cyber risk identification questions, enabled by SFT and RLHF training techniques. The original GPT-2 model, without training on pretrain data, served as the baseline.

Firstly, the ability of understanding cybersecurity content was evaluated [70], which can demonstrate the extent of cybersecurity knowledge the models have acquired gradually. Given phishing represents an emerging cyber risk in construction [36], [144], we evaluated our models' performance in classifying phishing emails using the public Phishing Email Detection [145]. This dataset consists of 11,929 email texts, with 43.5% labeled as phishing attempts. We compared the performance of each model using accuracy, precision, recall, and F1 score [141].

Secondly, the ability of generating answers regarding cybersecurity was evaluated [70]. The literature [109] was selected as the benchmark because, to the best of the authors' knowledge, it is the only related work that offers a relatively more comprehensive list of cyber risks for construction projects. The findings of work [109] are presented in Table 4.5, from which the questions were formulated as 'Cyber risks in the {key1} phase include'. Each reference answer was structured as 'Cyber risks in the {key1} phase include {key2} resulting from {key3}.' Two sets of metrics were employed: (1) BERTScore [146], encompassing Precision (semantic similarity between generated and reference answers), Recall (coverage of reference answer's words by the generated answer), and F1 Score (the harmonic mean of Precision and Recall) for a balanced assessment of answer similarity in both wording and context. (2) GPT-4 [115] is requested to score from 0 to 100 based on

Content (relevance, accuracy, and completeness), Clarity and Coherence (organization and readability), Specificity (detail and precision), and Risk Coverage (comprehensiveness of identified cyber risks) [147]. Priority is given to Content and Risk Coverage, with the assigned weights being 0.3, 0.2, 0.2, and 0.3, respectively.

Table 4.5. Cyber risk list from the selected benchmark [109]

Phase (Key1)	Potential Threats (Key2)	Potential Vulnerabilities (Key3)
Initiation	Data theft	Unsecured network transfer and cloud storage applications
Design	Proprietary information stolen	Unpatched software
Construction & Procurement	Performance degradation, physical damage	Excessive usage, fabricated chips
Commissioning	Data tampering, actuation tampering	Compromised dashboard and sensor
Operation & Maintenance	Spying, deliberate destruction	Chip insertion
Renovation & End of life	Data retrieval	Disposed sensors and equipment

4.2.6.2 Evaluating identified cyber risks

To achieve a comprehensive validation on the identified cyber risks, three sources were drawn: the selected benchmark [109], assessments by the two experts abovementioned, and assessments by GPT-4 [115]. To ensure a thorough and unbiased assessment of the experts and GPT-4, we designed their assessment process strategically:

- **Diverse Expertise Backgrounds:** Experts from different domains were invited to provide a broad spectrum of insights. The first expert, a cybersecurity specialist from the U.A.E., brought over ten years of industry experience. The second expert, a construction domain specialist from China, contributed six years of experience in the construction sector.
- **Multidimensional Judgement Criteria:** Assessors were tasked with evaluating each cyber risk from three attributes: its relevance to the construction phase, its likelihood of occurrence, and its potential impact on the project. Assessors evaluated each attribute using a 5-point Likert scale [139], [140], [141], where 1

signifies strong disagreement with the presence or significance of the attribute, and 5 indicates strong agreement.

- **Independent Assessment:** Assessments of the experts and GPT-4 were conducted independently. This approach ensured that each assessor evaluated the cyber risk list without prior knowledge of the other's opinions or ratings, thereby reducing the likelihood of conformity bias and encouraging impartial judgments.
- **Round Robin Evaluation:** We implemented a Round Robin evaluation method [148], dividing the assessment processes into six rounds, with each round corresponding to a project phase. Each assessor dedicated two days to the assessment, reviewing three project phases per day. Each phase was reviewed twice within the same day, and the average of the two assessments was taken to ensure an unbiased evaluation. This structured approach helped prevent assessment fatigue and ensured that each risk received adequate scrutiny.

Four dimensions of evaluation results are presented: (1) A detailed comparison with the benchmark literature [109]; (2) Qualitative feedback from two experts; (3) Analysis of the relevance of the identified cyber risks; and (4) Statistical analyses, aiming to verify the validity and consistency of likelihood assessments among all assessors. This involves comparing descriptive statistics to ensure likelihood assessments aligned with reality, utilizing the Friedman test [149] to ascertain the consistency of assessment criteria among all assessors holistically, employing the Wilcoxon Signed-Rank test [150] to determine if the assessment criteria between any two assessors are consistent, and applying the Spearman Rank Correlation test [151] to examine the consistency of risk prioritizations between any two assessors. These tests were chosen because they are non-parametric and do not assume a normal distribution, making them suitable for processing ordinal data, such as the risk likelihood levels. Additionally, they do not require a large sample size, rendering them appropriate for our analyses and providing reliable insights. This comprehensive set of results aims to further validate the effectiveness of our RL model and

the applicability of the identified cyber risks. The details of implementations and results are presented in Section 4.3.3.

4.3 Results

4.3.1 Model training and selection

The training sessions were conducted using the `PyTorch` computation framework with a fixed random seed to ensure replicability. Additionally, we employed the default seed value of 42 for dataset splitting via the `Scikit-learn` package. The configuration details for the four stages are outlined in Table 4.6. The results are as follows:

(1) **Base model for stage 1.** Table 4.7 shows the training outcomes for the base models. Model selection considers the balance of effectiveness and efficiency on the test set. All models showed similar test performance, but GPT-2 and BERT-LM were more time-efficient than T5-LM, resulting in T5-LM's exclusion. GPT-2 consistently surpassed BERT-LM, peaking at 3.603 after 4 epochs, and was chosen as the final base model for its sufficient encoding of construction cybersecurity knowledge.

(2) **SFT model for stage 2.** The training and test losses were computed solely on the answer part. The SFT model selection was based on its performance on the test set, as were the reward model and the RL model. The SFT model's peak performance on the test set was achieved at Epoch 12, with a loss of 0.015.

(3) **Reward model for stage 3.** The reward model achieved the lowest test loss at Epoch 69, exhibiting a prediction bias of ±6.5, which is considered acceptable and indicates reliable scoring predictions.

(4) **RL model for stage 4.** We experimented 32 combinations of weights for the three loss terms in Equation (4.4), ensuring each weight is non-zero, a single digit, and sums to 1. For each combination, the model underwent training illustrated in Table 4.6. The model achieved its best test loss of 3.22 in Epoch 8 with a weight combination of (0.5, 0.2, 0.3). The higher weight assigned to the first term (the

reward term) demonstrates the effectiveness of our reward model in improving the model's generalization ability of answering unseen questions.

Table 4.6. Training details of the four stages

Model Type	Dataset Split (train/test)	Epoch	Loss	Optimizer	Initial Learning rate	Batch size
Base	95% / 5%	20	Cross-Entropy loss using teacher forcing [118]	AdamW with scheduler	5e-5	32
SFT	80% / 20%	20	Cross-Entropy loss using teacher forcing [118]			
Reward	80% / 20%	100	MSE loss			
RL	80% / 20%	20	Designed loss in Equation (4.4)			

Table 4.7. Training results of base models (the first 10 epochs)

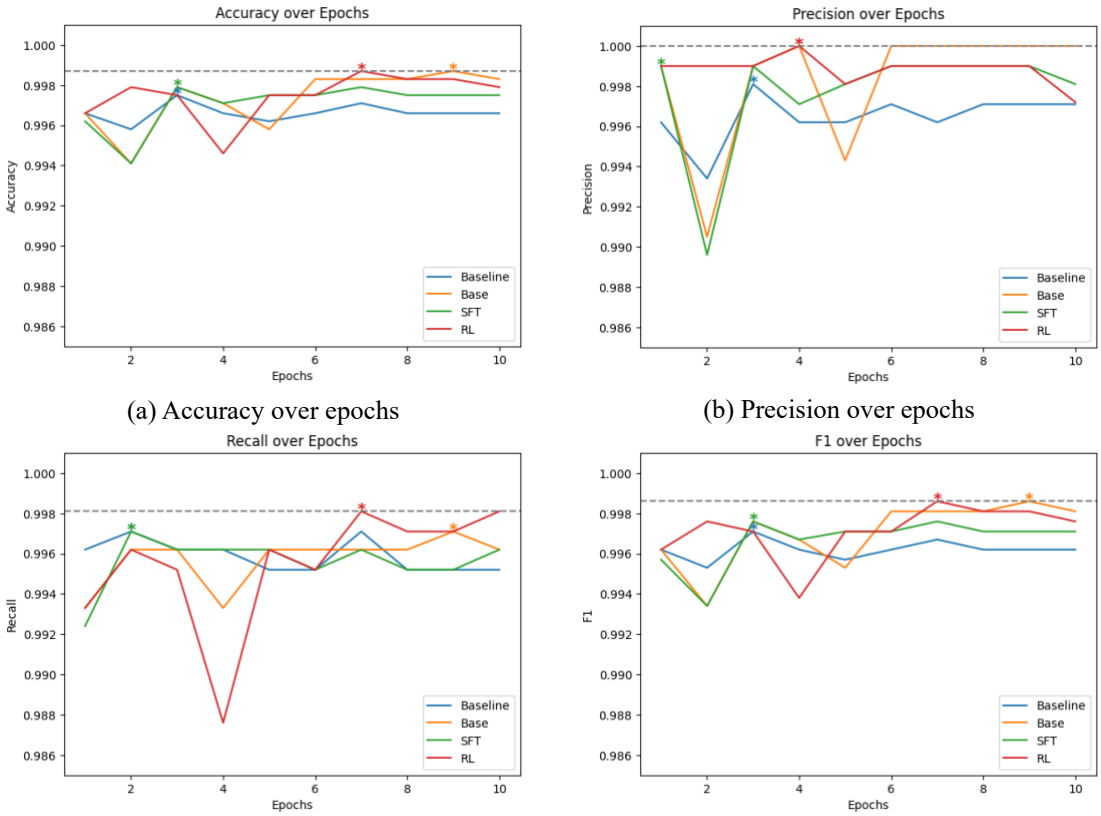
Model	Trainable parameters (million)	Training time (hours)	E_1	E_2	E_3	E_4	E_5	E_6	E_7	E_8	E_9	E_{10}
GPT-2	124	8.5	3.711	3.634	3.610	3.603	3.608	3.622	3.646	3.659	3.677	3.683
BERT-LM	109	9	3.861	3.728	3.671	3.662	3.676	3.698	3.728	3.754	3.767	3.779
T5-LM	113 (Decoder only)	19	3.713	3.615	3.556	3.517	3.494	3.478	3.467	3.460	3.456	3.454

4.3.2 Evaluating model progress

4.3.2.1 Progress in understanding cybersecurity content

Implementing the method in Section 4.2.6.1, the training was executed with a 5:1 ratio for splitting the training and testing data from the Phishing Email Detection dataset [145], spanning across 10 epochs. Figure 4.4 delineates the metric improvements for each model, where a "*" symbol denotes the peak performance of a model in a specific metric. Impressively, all models recorded scores exceeding 0.9 across a variety of metrics, each surpassing the baseline. This achievement underscores the enhanced cybersecurity comprehension of our models, a direct consequence of utilizing our specially curated pretraining dataset. The culmination of this progression is observed with the RL model,

which surpasses all others in every metric.



(a) Accuracy over epochs

(b) Precision over epochs

(c) Recall over epochs

(d) F1 over epochs

Figure 4.4. Results on phishing text classification

4.3.2.2 Progress in answering cybersecurity questions

Implementing the method in Section 4.2.6.1, Table 4.8 details the performance results, while Table 4.9 summarizes the average performance across all prompts. Generally, the models show improved performance on each metric, indicating a better ability to generate answers. Analyzing BERTScore metrics, the RL model surpasses others in all aspects. The RL model's higher precision and recall values indicate that its answers include relevant sections and cover a larger proportion of the reference, with the F1 score—the harmonic mean of both—highlighting its superior performance. The low baseline score (0.2139) by GPT-4 suggests its generated answer partially resembles the reference (from the moderate BERTScore) but lacks overall quality. Conversely, the RL model's high score (0.7041)

indicates its answers are not only more similar to the reference but also meet the high quality of our scoring criteria. The average score, considering all metrics, reveals the RL model as the top performer (0.6590), indicating it is the most effective in generating satisfying answers, outperforming the baseline by 33.64%.

Table 4.8. Detailed evaluation results of answer generation ability

Prompt & Reference No.	Model	Precision	Recall	F1	GPT-4 (Divided by 100)
1	Baseline	0.4677	0.599	0.5253	0.0927
	Base	0.5940	0.7015	0.6433	0.6403
	SFT	0.5726	0.7107	0.6342	0.6986
	RL	0.5734	0.7229	0.6354	0.7549
2	Baseline	0.5646	0.635	0.5977	0.3184
	Base	0.4588	0.722	0.5611	0.2447
	SFT	0.5186	0.6933	0.5933	0.3869
	RL	0.5308	0.7384	0.6004	0.5388
3	Baseline	0.5123	0.6594	0.5766	0.2162
	Base	0.5473	0.7059	0.6165	0.6027
	SFT	0.5647	0.7274	0.6358	0.6997
	RL	0.5855	0.7276	0.6454	0.7626
4	Baseline	0.5584	0.5781	0.5681	0.2344
	Base	0.5873	0.5966	0.5919	0.5634
	SFT	0.5746	0.6738	0.6203	0.8568
	RL	0.6205	0.6872	0.6411	0.7483
5	Baseline	0.5201	0.6964	0.5954	0.3093
	Base	0.5447	0.7506	0.6313	0.6628
	SFT	0.5778	0.7494	0.6525	0.7468
	RL	0.5691	0.7541	0.6487	0.7257
6	Baseline	0.5475	0.726	0.6242	0.1123
	Base	0.5585	0.7599	0.6438	0.2533
	SFT	0.5608	0.7505	0.6419	0.6897
	RL	0.5543	0.7271	0.6290	0.6940

Table 4.9. Answer generation evaluation results averaged across phases

Model	Precision	Recall	F1	GPT-4	Average
Baseline	0.5284	0.649	0.5812	0.2139	0.4931
Base	0.5484	0.7061	0.6146	0.4945	0.5909
SFT	0.5615	0.7175	0.6297	0.6798	0.6471
RL	**0.5723**	**0.7262**	**0.6333**	**0.7041**	**0.6590**

Table 4.10. Cyber risk identification checklist with risk assessment results

Phase	Cyber risks (threats and vulnerabilities)	Potential consequence	Expert 1			Expert 2			GPT-4			RL	Ave L	Ave I	Risk value	Risk level
			R	L	I	R	L	I	R	L	I	L				
Initiation	Weak identity and access management;	Lead to unauthorized access to systems and data, increasing the risk of data breaches, fraud, and compliance violations.	4	5	5	4	5	5	4	4	4	5	4.75	4.67	22.17	VH
	Unauthorized access to bidding documents and information;	Lead to the leakage of sensitive project details, giving competitors unfair advantages and potentially resulting in financial losses.	5	3	5	4	3	4	5	3	5	2	2.75	4.67	12.83	M
	Insecure communication channels;	Expose the transmission of critical and confidential data to interception by cyber attackers, risking the confidentiality and integrity of sensitive information.	5	2	4	5	2	5	4	2	5	2	2.00	4.67	9.33	L
	Weak permission controls for the use of assets and systems;	Allow unauthorized access and misuse of resources, leading to data leaks, system disruptions, and compromised security protocols.	4	2	4	4	1	5	4	2	4	2	1.75	4.33	7.58	L
	Insufficient data encryption for bidding documents and plans;	Can make it easier for unauthorized individuals to intercept and decipher confidential information, compromising the integrity and security of the project.	5	1	5	5	1	4	5	1	5	1	1.00	4.67	4.67	VL
	Phishing attacks targeting personnel of construction companies;	Result in unauthorized access to corporate systems and sensitive data, leading to data breaches and financial fraud.	3	1	3	4	1	4	4	1	5	1	1.00	4.00	4.00	VL
Design	Alteration of configuration data and/or information associated with digital facility operations;	Undermine the integrity of building systems, leading to operational inefficiencies and potential safety hazards.	5	5	5	4	5	5	5	5	5	5	5.00	5.00	25.00	VH
	Use of outdated versions of BIM and other design tools;	Leaving systems open to exploits and data breaches, and potentially leading to design flaws.	4	4	4	5	4	5	4	4	5	4	4.00	4.67	18.67	H
	Failure to update software with patches;	Serve as a gateway for cyberattacks, exposing systems to vulnerabilities and compromising the security of data and operations.	4	3	4	5	3	4	5	3	5	3	3.00	4.33	13.00	M
	Potential attempts to exploit vulnerabilities to access and erase data;	Disrupt operations, result in the loss of critical information, and significantly impact project timelines and costs.	5	3	5	4	2	4	5	2	4	2	2.25	4.33	9.75	L
	Lack of regulation in the data sharing process;	Lead to the inadvertent disclosure of sensitive information, violating privacy laws and undermining stakeholder trust.	5	2	4	5	3	3	5	3	3	3	2.75	3.33	9.17	L
	Unauthorized third-party access and utilization of confidential and/or proprietary information;	Lead to intellectual property theft, competitive disadvantage, and significant financial losses.	5	2	5	5	1	5	5	2	5	1	1.50	5.00	7.50	L
	Unauthorized access to design and construction information (e.g., access control card, door position, CCTV system, PINs, etc.);	Compromise the security of a facility, making it susceptible to physical breaches and endangering the safety of occupants.	5	2	5	5	1	5	5	1	4	1	1.25	4.67	5.83	L
	Fabrication of documents, images, and information pertaining to design and usage;	Result in non-compliance with regulations, jeopardize safety standards, and lead to financial and reputational damage.	5	1	5	5	1	5	4	1	5	1	1.00	5.00	5.00	VL

Table 4.10 (continued)

Phase	Cyber risks (threats and vulnerabilities)	Potential consequence	Expert 1			Expert 2			GPT-4			RL	Ave L	Ave I	Risk value	Risk level
			R	L	I	R	L	I	R	L	I	L				
Construction & Procurement	Unauthorized access to information on design and construction;	Lead to the exposure of proprietary designs and construction methodologies, jeopardizing competitive advantage and project integrity.	5	5	5	4	5	4	5	5	4	5	5.00	4.33	21.67	VH
	Difficulty accessing data or information about the contractor, operator, and maintenance owner;	Hinder project coordination, delay construction timelines, and impact overall project management efficiency.	4	2	4	5	2	5	4	2	5	2	2.00	4.67	9.33	L
	Interference with the operation or integrity of devices (also referred to as data theft) and manipulation of processes (IoT);	Compromise system functionality, endanger safety, and lead to unauthorized control or data breaches.	5	1	5	4	1	5	3	1	5	1	1.00	5.00	5.00	VL
	Manipulation of construction delivery services and other construction-related systems through cyber manipulation;	Disrupt project timelines, increase costs, and compromise safety protocols.	5	1	5	5	1	4	5	1	5	1	1.00	4.67	4.67	VL
Commissioning	Limited access controls and identity verification	Allow unauthorized individuals to access sensitive systems and data, increasing the risk of malicious activities and compromising the commissioning process.	5	5	5	5	5	4	5	5	4	5	5.00	4.33	21.67	VH
	Unauthorized access to commissioning data and systems	Compromise the integrity of the commissioning process, potentially leading to operational failures or safety hazards in the final infrastructure.	5	5	5	5	4	5	4	3	4	4	4.00	4.67	18.67	H
	Infiltration of malware into the commissioning systems	Disrupt critical operations, damage systems, and lead to costly delays and repairs, undermining the reliability of the commissioned asset.	5	2	5	5	2	5	5	1	5	2	1.75	5.00	8.75	L
	Social engineering attacks targeting commissioning personnel	Result in unauthorized access to systems and data, endangering the security of the commissioning operations and potentially leading to data breaches.	5	2	5	4	1	5	4	1	5	1	1.25	5.00	6.25	L
	Insufficient cybersecurity measures for IoT devices utilized in commissioning	Leave these devices vulnerable to attacks, compromising the functionality and security of critical systems involved in commissioning activities.	5	1	5	5	1	5	5	1	5	1	1.00	5.00	5.00	VL
	Vulnerable communication channels between commissioning teams and systems	Expose sensitive data to interception and manipulation, risking the confidentiality and integrity of the commissioning process.	5	1	4	3	1	3	4	1	2	1	1.00	3.00	3.00	VL

Table 4.10 (continued)

Phase	Cyber risks (threats and vulnerabilities)	Potential consequence	Expert 1			Expert 2			GPT-4			RL	Ave L	Ave I	Risk value	Risk level
			R	L	I	R	L	I	R	L	I	L				
Operation & Maintenance	Leakage of or interference with the critical asset information;	Undermine the security and integrity of essential services, leading to operational disruptions and potential safety hazards.	5	5	5	5	5	5	5	5	5	4	4.75	5.00	23.75	VH
	Unauthorized access to the design and construction data/information;	Lead to the exposure of proprietary or sensitive project details, potentially resulting in competitive harm and financial loss.	4	5	4	4	5	5	5	5	5	5	5.00	4.67	23.33	VH
	Leakage of or interference with the construction model, the operation and maintenance plan of the assets;	Disrupt operational efficiency, compromise safety protocols, and lead to increased maintenance costs.	5	2	5	4	3	5	5	3	5	3	2.75	5.00	13.75	M
	Malicious actors who obtain access to a building's systems and data, such as keystroke loggers, code-based logic bomb threats, file downloads, etc.	Introduce risks such as espionage, sabotage through logic bombs, and unauthorized access to sensitive information, compromising security and privacy.	5	1	5	4	1	4	4	2	5	2	1.50	4.67	7.00	L
	Attacks on the operational phase, which may lead to physical damage, theft of intellectual property, and damage to third-party assets;	Cause physical damage to infrastructure, result in the theft of intellectual property, and lead to financial liabilities due to damage to third-party assets.	5	1	5	5	1	4	5	1	4	1	1.00	4.33	4.33	VL
Renovation & End of life	Inadequate identity and access management for demolition-related systems.	Allow unauthorized access to critical controls and information, increasing the risk of sabotage, data theft, and manipulation of the demolition process.	5	5	5	5	5	5	5	4	4	5	4.75	4.67	22.17	VH
	Unauthorized access to demolition plans and schedules;	Lead to premature disclosures or alterations, potentially endangering workers and the public by compromising the planned safety measures.	5	4	5	5	3	5	5	4	5	3	3.50	5.00	17.50	H
	Disclosure or theft of sensitive demolition-related information;	Expose strategic or competitive details, leading to financial losses or unauthorized access to secured sites, compromising safety and security protocols.	5	3	5	5	1	5	4	2	4	2	2.00	4.67	9.33	L
	Insecure communication and data transmission among stakeholders;	Result in the interception or manipulation of sensitive data, undermining the integrity and confidentiality of the demolition operation.	5	2	4	5	2	5	5	2	5	2	2.00	4.67	9.33	L
	Social engineering attacks aimed at demolition personnel;	Trick individuals into divulging confidential information or granting access to restricted systems, jeopardizing the security of the demolition operation.	5	2	4	5	2	3	5	3	4	2	2.25	3.67	8.25	L
	Interference with demolition-related systems and safety controls;	Cause malfunctions or failures, leading to accidents, injuries, or uncontrolled collapses, significantly increasing risk to human life and surrounding properties.	5	1	5	5	2	4	5	2	3	2	1.75	4.00	7.00	L
	Unpatched or outdated software utilized in the demolition process;	Contain vulnerabilities that cyber attackers could exploit, potentially leading to system failures, data breaches, or unauthorized control of demolition activities.	4	1	4	3	1	4	4	2	4	1	1.25	4.00	5.00	VL

Note: R denotes Relevance; L denotes Likelihood; I denotes Impact; VH denotes Very High; H denotes High; M denotes Medium; L denotes Low; VL denotes Very Low.

In summary, the model is progressing in both its ability to understand cybersecurity and generate answers. This progression verifies the efficacy of our training strategy pathway, thus laying the foundation for our RL model's effectiveness in performing the task of cyber risk identification.

4.3.3 Evaluating identified cyber risks

Following the methods in Sections 4.2.5 and 4.2.6.2, the identified cyber risks and likelihoods given by RL model are shown in Table 4.10, and the evaluation results are presented in following sections.

4.3.3.1 Superiority over the benchmark

Compared to the results of the work [109] (shown in Table 4.5), our language model and results demonstrate superiority in comprehensiveness, adaptability and speed, as shown in Table 4.11.

4.3.3.2 Positive feedback from experts

The cyber risk checklist received strong endorsements from both experts. The cybersecurity expert praised its comprehensiveness and relevance, highlighting its utility in segmenting risks by project phase and its role in enhancing awareness and prioritization of cybersecurity efforts within the construction industry. He recommended the checklist as an essential tool for companies aiming to bolster their cybersecurity posture. Similarly, the construction domain expert recognized the checklist's value in raising awareness of cyber threats and vulnerabilities, endorsing its broad coverage and foundational role in advancing cybersecurity practices. He advocated for a proactive cybersecurity approach guided by the checklist, suggesting collaboration with cybersecurity professionals to refine and adapt strategies to the dynamic cyber threat environment.

Table 4.11. Comparison with the selected benchmark [109]

Aspect	Sub-aspect	Benchmark [109]	Our work
Comprehensi-veness	List of Cyber Risks	Not comprehensive, providing one or two identified risks for each phase	Comprehensive, providing detailed list across different project phases and ensuring broad applicability to construction projects
	Risk Prioritization	Not specified	Specified ranking of cyber risks (Table 4.10), aiding strategic resource allocation for risk prevention and mitigation
	Scenario Coverage	Not wide, may missing specific incident scenarios	Wide, covering a broad spectrum of incident scenarios and can be updated regularly
Adaptability	Framework Nature	Static, requiring complete restart for updating the list	Dynamically updateable, supporting model self-training for ongoing adaptation to new data
	Response to Cybersecurity Landscape Changes	Limited, necessitating expert intervention and framework restart for updates	Capable, efficiently aligning with changes, allowing periodic updates without training the model from scratch
	Updating Process	Complicated, requiring a group of experts for framework adjustments and restarting	Simple, allowing automatically updating with new data, minimizing the need for expert intervention
Speed	Time to Identify Risks	Time-consuming, requiring days to weeks per cycle, dependent on manual processes	Time-efficient, although initially taking around 87 hours for training and fine-tuning (excluding time for experimentations), but allowing for rapid adaptation to new data within hours
	Human Intervention	High, reliant on manual processes and expert discussions	Minimal, primarily for oversight, with the language model handling the bulk of processing
	Inference Speed	Not applicable, requiring restarting the whole process	Fast, requiring only seconds for answering a question, enabling swift risk re-identification and response to evolving threats

4.3.3.3 High relevance of identified risks

Table 4.12 displays the percentages of relevance levels by all assessors. It reveals that none of the risk items were rated with a relevance level below 3, with level 5 being the most common rating for both. This suggests that all the identified cyber risks on the list are considered relevant, demonstrating the effectiveness of our language models in identifying cyber risks.

Table 4.12. Percentage of relevance levels

Expert 1			Expert 2			GPT-4		
Level 3	Level 4	Level 5	Level 3	Level 4	Level 5	Level 3	Level 4	Level 5
2.78%	19.44%	75.00%	5.56%	25.00%	69.44%	2.78%	30.56%	66.67%

4.3.3.4 Valid and consistent assessments

Following Section 4.2.6.2, statistical analyses were performed to check the validity and consistency of all assessors.

(1) The assessments are demonstrated to be valid. Looking at the descriptive statistics in Table 4.13, the variation in mean values across different phases for each assessor reflects the distinct cybersecurity challenges inherent to each phase, underscoring the necessity of the difference for phase-specific assessment criteria. Furthermore, the fluctuation in likelihoods assigned by each assessor within a phase, evidenced by the standard deviation, confirms that assessors recognize the uniqueness of each cyber risk. These variations add validity to the assessment criteria of all assessors, including our RL model.

Table 4.13. Descriptive statistics and Friedman test results

Phase	RL model (Mean ± SD)	Expert 1 (Mean ± SD)	Expert 2 (Mean ± SD)	GPT-4 (Mean ± SD)	Test Statistic	P value
Initiation	2.167 ± 1.344	2.333 ± 1.374	2.167 ± 1.462	2.167 ± 1.067	1.000	0.801
Design	2.500 ± 1.414	2.750 ± 1.199	2.500 ± 1.414	2.625 ± 1.317	2.538	0.468
Construction & Procurement	2.250 ± 1.639	2.250 ± 1.639	2.250 ± 1.639	2.250 ± 1.639	-	-
Commissioning	2.333 ± 1.599	2.667 ± 1.700	2.333 ± 1.599	2.000 ± 1.528	7.000	0.072
Operation & Maintenance	3.000 ± 1.414	2.800 ± 1.833	3.000 ± 1.789	3.200 ± 1.600	2.400	0.494
Renovation & End of life	2.429 ± 1.178	2.571 ± 1.400	2.286 ± 1.278	2.714 ± 0.881	2.415	0.491

(2) The assessments are consistent among assessors overall. We conducted phase-specific Friedman test [149], which is suitable for ordinal datasets (in our case, levels) and does not require minimum sample size, to check for overall significant differences within each phase in assessors' criteria. Because the likelihoods for all assessors are identical for the Construction & Procurement phase, the test was omitted for this phase. The null hypothesis (H_0) for every other phase is set to "There is no statistically significant

difference in the assessments". The test results are presented in Table 4.13. It is evident that the p-value for these phases exceeds the alpha level of 0.05, so we fail to reject the null hypothesis H_0 for each phase. This suggests that the assessors are holistically adhering to the same criteria of phase-specific assessment, proving our RL model's effectiveness.

(3) The assessments are consistent between any two assessors. To make sure there is no significant differences between pairs of assessors that might have been overlooked by the Friedman test, we performed the pairwise Wilcoxon Signed-Rank test [150] as a post-hoc analysis. The null hypothesis (H_0) for each assessor pair in each phase is set to "There is no statistically significant difference in the assessments." The results are presented in Table 4.14. It shows that p-values across all phases are greater than the alpha level of 0.008, which has been adjusted using the Bonferroni correction [150] to mitigate the risk of errors from multiple pairwise comparisons, so we fail to reject H_0 for each assessor pair, indicating no significant difference in assessment criteria between any two assessors. This further confirms the consistency of the assessors, proving our RL model's effectiveness.

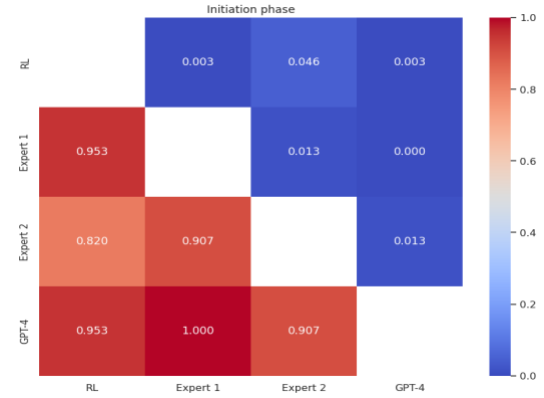
(4) The risk prioritizations are consistent among assessors. Spearman Rank Correlation test [151] was adopted to determine whether risk prioritization within each phase was consistent across assessors, as indicated by their correlations. The null hypothesis (H_0) for each pair of assessors in each phase was set as "There is no statistically significant correlation in the assessments." The test results are presented in Figure 4.5, where the right upper triangle of each heatmap contains the p-values, while the left lower triangle contains the correlation coefficients. All coefficients are greater than 0.5, indicating a strong correlation among all assessor pairs. An exception was Expert 1 - Expert 2 in the Renovation & End of Life phase, which had a p-value of 0.123. However, nearly all other p-values were significantly below the alpha level of 0.05, so we rejected the H_0 for these pairs, demonstrating statistically significant correlations. These high correlations indicate the consistency in risk prioritizations among assessors, further validating the accuracy across assessors and proving the effectiveness of our RL model.

Table 4.14. Wilcoxon signed-rank test results

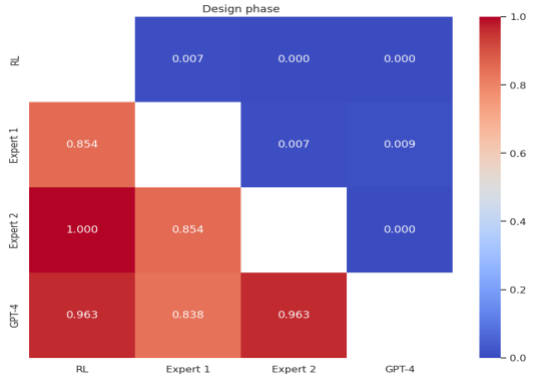
Initiation	RL model	Expert 1	Expert 2	GPT-4
RL model	-	1.000	1.000	1.000
Expert 1	0.000	-	1.000	1.000
Expert 2	1.500	0.000	-	1.000
GPT-4	1.500	0.000	1.500	-
Design	RL model	Expert 1	Expert 2	GPT-4
RL model	-	0.375	1.000	1.000
Expert 1	2.500	-	0.625	0.750
Expert 2	2.500	2.500	-	1.000
GPT-4	0.000	2.000	0.000	-
Construction & Procurement	RL model	Expert 1	Expert 2	GPT-4
RL model	-	1.000	1.000	1.000
Expert 1	0.000	-	1.000	1.000
Expert 2	0.000	0.000	-	1.000
GPT-4	0.000	0.000	0.000	-
Commissioning	RL model	Expert 1	Expert 2	GPT-4
RL model	-	0.500	1.000	0.500
Expert 1	0.000	-	0.500	0.250
Expert 2	0.000	0.000	-	0.500
GPT-4	0.000	0.000	0.000	-
Operation & Maintenance	RL model	Expert 1	Expert 2	GPT-4
RL model	-	0.750	1.000	1.000
Expert 1	2.000	-	1.000	0.500
Expert 2	1.500	0.000	-	1.000
GPT-4	0.000	0.000	0.000	-
Renovation & End of life	RL model	Expert 1	Expert 2	GPT-4
RL model	-	0.750	1.000	0.375
Expert 1	2.000	-	0.750	0.813
Expert 2	0.000	1.500	-	0.313
GPT-4	2.500	6.000	3.000	-

Notes: Bonferroni correction: $\alpha = 0.008$. Underlined values represent the Wilcoxon signed-rank test statistic; all other values are p-values.

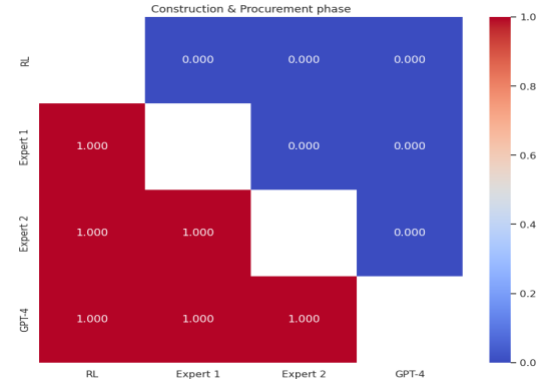
In summary, the results from the benchmark comparison, expert and GPT-4 assessments, and statistical analyses further validate the effectiveness of our RL model and the applicability of the identified risks. Table 4.10 presents the identified risks for each phase along with the consequences elicited from experts, ranked by their risk values.



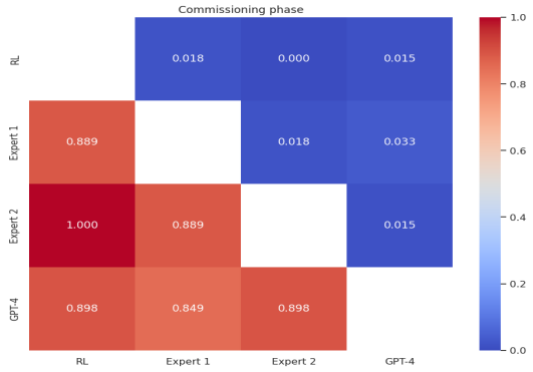
(a) Initiation phase



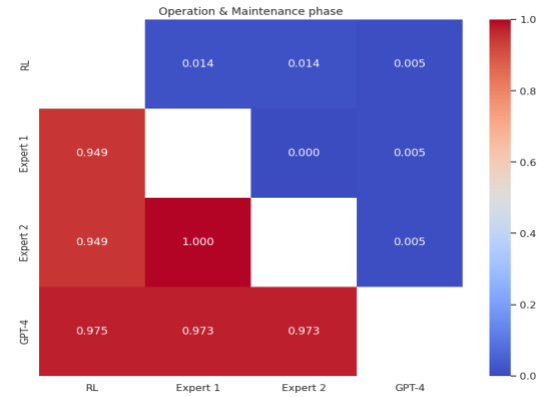
(b) Design phase



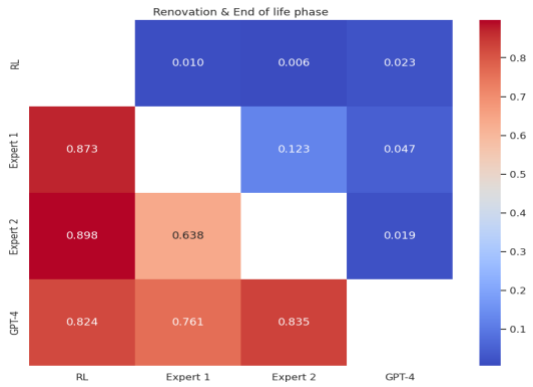
(c) Construction & Procurement phase



(d) Commissioning phase



(e) Operation & Maintenance phase



(f) Renovation & End of Life phase

- p-values , - Correlation coefficients, "0.000": $p < 0.001$

Figure 4.5. Spearman's Rank Correlation Coefficient test results

4.4 Discussions

4.4.1 The applicability of the checklist

Two main applications of the prioritized checklist are proposed: (1) It can serve as a new benchmark, which project managers can refer to for formulating proactive and preventive measures for their projects, focusing on the most significant risks. The process of developing these measures can involve collaboration between project managers, IT, and security teams. Such prioritization helps in effectively preventing high risks while ensuring more efficient resource allocation. (2) Risk analysts can use the checklist for in-depth risk analyses on specific projects, focusing first on the most significant risks. This process may involve identifying risk factors, performing quantitative risk assessments, and pinpointing critical risk factors, which should be addressed with priority.

In addition to the two primary applications, the checklist serves as a vital tool for IT and cybersecurity teams to evaluate and enhance existing security protocols. By identifying specific vulnerabilities, these teams can implement more robust security measures tailored to the unique needs of construction projects. Furthermore, stakeholders, including contractors and vendors, can use the checklist to verify that their systems and communication methods meet the necessary security standards. Moreover, the checklist can also play a crucial role in education and training, enabling personnel to understand the nature of common cyber risks, understand their potential impact, and recognize the importance of adhering to established security protocols. As the checklist can be regularly updated, people will have access to the latest information on the cybersecurity landscape, ensuring they remain informed about current trends.

4.4.2 Risk mitigation recommendations

Among the "High" and "Very High" risks, the primary concerns in construction projects can be summarized as weak identity and access management, unauthorized access across various project stages, falsification of critical data, reliance on outdated software tools, and

leakage of crucial asset information. These areas require attention. Additionally, it is evident that cyber risks can stem from both the IT sector and the construction sector, affecting the management and operational aspects of a project. This intersectionality of risk sources underscores the necessity for resolutions that involve collaboration between IT and construction professionals. By collaborating, these experts can ensure that cybersecurity measures are not only in the technical aspect but also encompass the optimization of management and operational guidelines, all of which should be practically applicable within the unique context of construction projects. The integration of IT security practices with construction management processes is integral in the project lifecycle.

We recognize that the cyber risks identified are high-level and broadly applicable to diverse projects. Considering the variations among individual projects, each has different exposures to these risks, attributed to the unique risk factors inherent to each project. Therefore, it is crucial to closely examine the specific factors contributing to each risk for individual projects, so as to formulate efficient and targeted risk mitigation strategies at the project level [152]. Key factors for investigation include various aspects of construction projects: general project information, project structure, IT-related factors, OT-related factors, and the human and management aspects [152]. Such a study will not only enable us to quantify the risk level specific to a construction project but also evaluate the effectiveness of risk mitigation strategies by directly addressing these critical factors.

4.4.3 The prospect of the language model

In Section 4.3.2.2, the quantitative evaluation of the models' progress has shown that the RL model excels in generating responses to cybersecurity queries. Providing an in-depth qualitative illustration, Table 4.15 compares the responses of our RL model and the baseline model (original GPT-2) to five varied prompts, all concerning the definition of phishing. The RL model consistently delivers relevant and informative responses, effectively explaining the phishing concept regardless of the phrasing differences. In contrast, the baseline model frequently yields answers that are irrelevant to the prompts. These

evaluations demonstrate the RL model's superior capability in handling cybersecurity queries, underscoring its potential as an effective cybersecurity consultant for construction personnel, particularly for those with limited knowledge.

Table 4.15. Illustration of answers by the baseline and RL model

Prompt	Baseline model	RL model
What is phishing?	We'll see what it looks like in the screenshots below, but before diving into the actual attack, please feel free to look at the phishing attacks in step 3.\n\nStep 2. Create a simple message or e-mail containing your password\n\nStep…	A phishing attack is a type of social engineering attack in which an attacker sends an unsolicited email or message, often disguised as a legitimate source, with the intent of tricking the recipient into divulging sensitive information.\nA phishing attack is a type of social…
How to define phishing?	The most common way for someone to use your e-mail address or your social media handle is to have them use it in the e-mail address or your social media handle and that address. This way, everyone uses your email address in order to spam you…	Phishing is a type of cyberattack in which attackers send fraudulent emails or messages to trick individuals into divulging sensitive information.\nPhishing is a type of cyber attack in which attackers send fraudulent emails or messages to trick individuals into divulging sensitive information.\nPh…
Can you explain phishing?	How do I get rid of it? (5:30)\n\nWhat are the main security features of phishing? (2:15)\n\nWhat are security flaws? (2:20)\n\nWhat does the system look like in real life…	Phishing is a type of cyber attack in which attackers send fraudulent emails or messages, often disguised as legitimate sources, with malicious attachments or links.\nA phishing attack can be conducted by attackers who view the emails or messages as legitimate sources. However, it can…
Discuss what phishing is.	Let's set up our system. The first thing you need to do is to do some common actions in your app:\n\nLaunch a new browser. In your app, go back to the menu bar (if using the built-in Chrome…	Phishing attacks, especially phishing attacks, can be conducted by attackers who view unsuspecting individuals or organizations as potential targets, using psychological manipulation techniques, or merely by sending fraudulent emails.\nPhishing attacks could be perpetrated…
Define phishing.	There are two types of phishers - those that make a mockery of your online service and those who make fun of you to ruin your online presence. Once you've discovered why your identity has been compromised, it's time to turn that around.\n\nA phishing attack…	Phishing is a type of cyber attack in which attackers send fraudulent emails or messages to trick individuals into divulging sensitive information or downloading malware.\nPhishing is a type of cyber attack in which attackers send fraudulent emails or messages to trick individuals into divulging sensitive information or…

Although the original training for the developed language model took up to 76 hours, for future updates, it is not necessary to train the models from scratch. We can employ continuous training techniques on the already established model, significantly reducing the time to just a few hours or even minutes. This makes it manageable for construction companies to use their own text dataset for training. Furthermore, the inference time—the time it takes to answer each query—is only a few seconds, which will be well within the operational limits of most construction companies. However, as we plan to expand data and model sizes in the future, it will be important for construction companies to balance their needs and budget, considering the extent to which they wish to expand the model. Training time will then be aligned with their specific requirements and affordability.

This upgraded model could be integrated into websites or apps, making it accessible to a wider audience within the construction industry, including contractors, project managers, safety officers, and stakeholders. With regular updates, the model can provide project-specific risk identification and analysis, dynamic risk identification and analysis, real-time risk monitoring, solution recommendations, and customized cybersecurity training for stakeholders. Updates should focus on incorporating new data and enlarging the model size, aiming to develop a construction cybersecurity-specific large language model that is 10^4 times larger, comparable to GPT-4 [115]. Its capacity for human-like interactions will improve the system's ability to comprehend and process a variety of inquiries from different users, thus enabling the model to grasp users' intentions more accurately. The model's responses will also be more comprehensible to users. This interactive feature can effectively assist individuals of diverse backgrounds and educational levels, increasing its utility for widespread industry application.

4.4.4 Enhancing dataset for model upgrade

Regarding a larger, more capable model, there are several considerations to keep in mind. Firstly, the dataset used to train our base model, compiled from online sources, needs update and expansion. Considering the vast amount of text generated daily, new data should

be collected periodically and organized, then integrated with existing datasets to enhance the model's understanding of construction cybersecurity. This approach will also keep the model current with cybersecurity trends. The frequency of updates will depend on available computing resources and the current demand. Secondly, increasing the diversity of the SFT dataset could be advantageous. Our current SFT question-answering dataset primarily includes simpler questions that focus on definitions, significance, or impacts of specific entities. Future iterations could introduce more complex and varied questions, incorporating logic and reasoning tasks to challenge the model with more sophisticated assignments. For example, questions might require the model to evaluate and discuss the evolving cybersecurity status based on detailed project information, thus aligning the model more closely with project-specific issues. Thirdly, the criteria for scoring answers in the RM dataset can be more granular and comprehensive. This aspect is crucial as it significantly influences the reward model's understanding of what constitutes a 'high-quality' answer, which, in turn, affects the RL model's response quality. Future evaluations can consider a broader array of scoring rubric, including format, content, typographical errors, repetitions, and logical coherence.

4.4.5 Limitations and future works

This study has limitations, including a dataset lacking project-specific information, limiting the model's ability to identify and assess unique project-specific cyber risks. Additionally, due to limited model size and computational resources, the risk identification process is only semi-automatic; while the model can generate content with risk identification results, it cannot fully automatically generate a checklist. Besides addressing the limitations, future efforts will focus on expanding our dataset and enhancing our model to match the capabilities of advanced models like Ernie Bot [114], GPT-4 [115] and Gemini [116]. We also plan to enable document analysis, including drawings, software codes, and schedules, to produce customized outputs such as figures, tables, and LaTeX code. Our final goal is to integrate this tool into web or mobile applications, offering an intelligent cybersecurity consultant accessible to various groups, especially those lacking cybersecurity expertise.

This integration can achieve advanced functions, including project-specific risk analysis, dynamic risk identification and analysis, real-time risk monitoring, solution recommendations, customized cybersecurity training for stakeholders, etc.

4.5 Conclusions

Given the insufficient recognition of cyber risks across project phases, this study developed a language model to thoroughly identify cyber risks across project phases, which are applicable to diverse construction projects. Our model, trained on 61,841 sentences of construction cybersecurity textual data and improved by the SFT and RLHF techniques, shows expected improvement in understanding cybersecurity content and in answering cybersecurity questions. This positions our language model as suitable for identifying cyber risks across project phases. The cyber risk checklist, ranked by risk values, surpasses the existing literature, has received positive feedback from industry experts, and proves to be highly relevant. The risk likelihood assessments by our model are consistent with those of two experts and GPT-4. These results collectively validate the effectiveness of our model and the applicability of the identified cyber risks.

This study provides in-depth discussions on the applicability of the identified checklist, emphasizing its usage for project managers to formulate risk preventive measures and for focused risk assessment. The study emphasizes the need for collaboration between IT and construction professionals and underscores the importance of examining specific risk factors in individual construction projects to formulate targeted mitigation strategies. Additionally, it emphasizes the importance of quantitative risk assessment at the project level. In terms of the language model, the study highlights the developed language model as suitable for serving as an intelligent cybersecurity consultant for construction personnel, capable of achieving various advanced functions. The study also discusses strategies for enhancing the dataset to improve the model's size and performance.

Compared to previous studies and practices that mostly propose frameworks or methods

reliant on manual effort and thus lack flexibility and efficiency, our developed language model offers a more comprehensive approach to cyber risk identification. It also enables dynamic identification in the future that ensures time efficiency. Moreover, two additional benefits are offered: (1) Customizability. The model can be fine-tuned with a company's own text corpus; fine-tuning requires only a few hours, and subsequent question-answering takes mere seconds per question. This capability facilitates efficient and tailored risk identification and analysis that considers the unique characteristics of each organization. (2) Intelligent cybersecurity consultant. The upgraded model can be used as an intelligent cybersecurity consultant deployed in construction companies' mobile or website applications, achieving various advanced functions and thus showing great potential for industry-wide utilization.

This study's limitations include the inability to identify project-specific risks and the semi-automatic process for compiling the general risks. Future efforts will focus on expanding the dataset and enhancing the model to match advanced capabilities. We aim to integrate the language model into construction companies' web or mobile applications for accessible cybersecurity consultancy, achieving advanced functions such as project-specific risk analysis, real-time monitoring, and customized cybersecurity training.

Chapter 5
Identifying Cyber Risk Factors Associated with Construction Projects

After identifying the potential cyber risks across project phases, in the chapter, we employ a systematic methodology to identify cyber risk factors associated with construction projects that have a causal relationship with the identified cyber risks. The identified 32 risk factors reflect general vulnerabilities prevalent in the IT sector and those arising from the complex and dynamic nature of construction projects specific to the construction industry. The risk factors are grouped into five categories: 7 about overall project information, 4 about project structure, 9 about IT, 5 about OT, and 7 about management and human aspects, which can be used to assess a diverse range of cyber risks. Each risk factor is also divided into specific categorical or ordinal scales. The set of identified risk factors is beneficial for a more quantitative risk assessment in the future, which is discussed in detail.

5.1 Introduction

In risk management, after identifying potential risks, it is crucial to dissect them to uncover their underlying factors. A comprehensive set of risk factors enhances understanding and

provides insights into which aspects to prioritize. This proactive approach allows for implementing preventative measures in a more granular manner [20]. In this dissertation, the identification of cyber risk factors associated with construction projects can help the industry better understand the causal relationship between project characteristics and cyber risks. This also enables the creation of a cyber risk assessment model for construction projects, facilitating a more targeted and effective approach to risk management. Based on the outputs of the risk assessment model, project managers can decide whether and how to address the specific risk.

Cyber risk factors are specific conditions, actions, or inactions that might expose construction projects to cyber risks. Cyber risk factors in construction projects broadly encompass those common across the IT sector, which might occur in any industry, as well as those unique to the construction industry. Common risk factors in the IT sector include weak passwords, outdated software, lack of encryption, insufficient network security, and human error, among others. These universal vulnerabilities can expose any organization to cyber risks. However, the construction industry faces unique cybersecurity challenges beyond those commonly encountered in the IT sector. These distinct vulnerabilities and challenges originate from the inherently multifaceted and dynamic nature of construction projects, characterized by the following five aspects.

(1) Frequent changes in teams. Construction projects have ever-changing teams based on phases and specialties. This dynamism, while essential, disrupts workflows, particularly in communication and cybersecurity. With each transition, vulnerabilities arise due to unfamiliarity with protocols or overlooked onboarding [32].
(2) Varied levels of cybersecurity knowledge among personnel. The construction sector spans roles from field workers to IT experts, leading to a broad spectrum of cybersecurity understanding. This diversity creates potential weak links: uninformed personnel might inadvertently jeopardize project security through mistakes or phishing vulnerabilities [153].

(3) Scattered and frequent communications. Communications in construction are widespread across different channels and locations. While necessary, this breadth increases potential security breach points. With diverse stakeholders communicating, risks like misinterpretation, data leaks, or unauthorized access grow [154].
(4) Frequent exchange of digital information. The digital era has amplified data exchanges in construction, especially in supply chains. Every digital interaction, from design blueprints to daily updates, poses a cybersecurity risk. Given the variety of data transferred, specialized security measures are crucial [35], [153].
(5) Overlap of personnel across multiple projects. It is common for personnel to handle multiple construction projects simultaneously. While resource-efficient, this can blur project boundaries, risking unintentional data leaks or access. For instance, an architect working on different projects using a single device might inadvertently mix or share data [113].

Given these distinct vulnerabilities in construction projects, a tailored set of risk factors is a must for effective industry-specific cyber risk management. However, such work is still notably absent, as detailed in Section 5.1.1. While ample literature exists on identifying risk factors for other types of risks in construction, such as delay and safety risks, there is a dearth of literature focused on industry-specific cyber risk factors.

5.1.1 Related works

The construction industry has witnessed extensive studies of identifying/discussing risk factors for a variety of other risks, including financial risks [155], [156], [157], [158], environmental and safety risks [159], [160], supply chain risks [161], procurement risks [162], technological application risks [163], [164], time performance risks [141], [165], [166], and a mix of these [167], [168], [169], [170], [171], [172], [173]. Each of these studies introduces a tailored set of risk factors, categorized based on specific criteria, aiming to either enhance the understanding of associated risks or facilitate their assessment.

In terms of financial risks, Sharma and Goyal [155] identified 55 risk factors associated

with cost overrun in construction projects through an intensive literature review and expert opinions. They then conducted a fuzzy assessment of the risk. Baloi and Price [156] identified 36 major global risk factors influencing cost-effectiveness through an extensive literature review and preliminary discussions with construction contractors, successfully developing a fuzzy decision-making framework for risk management. El-Karim et al. [157] researched and evaluated 70 contingency factors affecting costs based on the literature and opinions of practitioners/experts through fifty-nine questionnaires. Chileshe and Boadua Yirenkyi-Fianko [158] identified 25 risk factors that could impact cost of construction projects in Ghana, based on a literature review.

Environmental and safety risks are among the key concerns in projects. Hwang et al. [159] identified 42 risk factors through a literature review that are associated with environmental and safety risks in green residential construction, aiming to enhance understanding of the risks in this specific scenario. Aghaei et al. [160] identified 44 risk factors involved in 8 groups of accidents that could occur in construction projects. These risk factors were identified through a process in which 40 experts were asked to list the factors they believed were responsible for each risk based on their experience and observations during their work history.

As for supply chain risk and procurement risk, Rudolf and Spinler [161] argued that the absence of supply chain risk management (SCRM) has been a major cause of failure in large projects. They identified 114 risk factors, particularly relevant to SCRM in large-scale engineering and construction projects, through a systematic, multi-phased approach that includes literature reviews, expert opinions, and empirical data. Chan et al. [162] identified 34 risk factors specific to the procurement risk target cost contracts (TCC) or guaranteed maximum price (GMP) contracts in the construction industry. These were determined through a literature review, structured interviews with industry practitioners, and surveys among construction professionals, focusing on their experiences and perceptions of risks associated with these contract types in the Hong Kong construction industry.

As for technological application risks, Sun et al. [163] identified risk factors for BIM technology through a literature review focusing on BIM risks, obstacles, challenges, and success factors. From a preliminary collection, 27 highly cited and authoritative sources were selected, leading to the identification of 28 risk factors categorized into legal, industry, capital, technical, and management aspects. Chien et al. [164] identified 13 risk factors across general construction, IT projects, and Building Information Modeling (BIM) through literature reviews, surveys, and expert interviews. This multifaceted approach combined previous research findings, professional intuition, and practical experience to categorize risks into technical, management, environmental, financial, and legal dimensions.

Regarding time performance risks, Al Zubaidi and Al Otaibi [165] conducted a study to identify critical risk factors causing delays in Kuwait's building and infrastructure projects. Utilizing a combination of questionnaire development, a survey, and a case study analysis of 28 projects, it evaluated the responses to pinpoint the main causes of time overruns. Gondia et al. [141] identified 59 risk factors related to project delays, classified into 9 groups, with identification based on literature review and expert consultation. Assaf and Al-Hejji [166] identified 73 risk factors related to project delays, categorizing them into groups such as owner-related, consultant-related, and design-related risks, with identification based on literature review and discussions with various parties involved in the construction industry.

Besides the risk factors for a specific kind of risk, some studies also explored the risk factors that are for a mix of risks. Rezakhani [167] identified and classified the key risk factors that significantly impact the scope of construction projects, drawing on a comprehensive literature review and the expertise of professionals in the field of construction management. Zou et al. [168] identified 85 risk factors targeting construction projects in China, through a comprehensive process that included a literature review, consultation with industry experts, and pilot surveys, covering various types of risks, including cost, time, quality, and more. Jarkas and Haupt [169] identified the risk factors

considered by general contractors operating in the State of Qatar through a combination of previous research on construction risk management, consultations with local industry experts, practitioners, and professionals. This approach culminated in a structured questionnaire survey, which further refined the identification and categorization of 37 risk factors across four major groups.

Additionally, Renuka et al. [170] identified critical risk factors in international construction projects by reviewing 50 relevant articles published over the last 25 years. They suggest developing simple analytical tools for task-specific risk assessments to facilitate early risk identification and assessment during the bidding stage. Wuni et al. [171] identified risk factors for Modular Integrated Construction (MiC) through a systematic review and synthesis of 39 empirical studies examining the risks associated with MiC across different countries. This comprehensive analysis led to the identification of 73 risk factors, of which 30 were considered critical risk factors (CRFs) as they were reported in at least two studies. Zou and Zhang [172] identified risk factors in construction projects from life cycle and stakeholder perspectives using a conceptual methodological framework based on AS/NZS4360 and ISO31000 standards. This framework utilizes two-dimensional graphical presentations for analyzing key risks, allocating them to responsible stakeholders, and identifying their occurrence phases in the project life cycle. Mahmoud et al. [173] identified risk factors for highway construction projects in Egypt through an extensive literature review and consultation with experts, leading to a preliminary list. This list was refined via pilot surveys and expert consultations into 12 risk groups containing 73 risks. A questionnaire survey was then conducted among professionals from different sectors to evaluate these risks.

Despite numerous studies identifying risk factors across various domains, a distinct gap is evident in the literature regarding risk factors within the scope of cybersecurity, with minimal attention given to cybersecurity-related risks in the construction industry. The prevalent methodologies utilized in these works for risk factor identification include literature reviews, expert interviews, and surveys and questionnaires. As highlighted in

Chapter 2, although there are some discussions and efforts to raise awareness about cybersecurity within the construction sector, these are predominantly general in nature and a comprehensive set of risk factors related to cyber risks remains absent. This gap signifies a critical need for dedicated research to methodically identify these cyber risk factors. This study aims to bridge this gap by employing a systematic approach that integrates literature review, expert evaluations, and surveys and questionnaires to identify risk factors for cyber risks specifically related to construction projects.

5.1.2 Objectives

The study aims to bridge this significant research gap by developing a set of cyber risk factors tailored to the characteristics of construction projects. These risk factors can be applied to a wide range of cyber risks, including but not limited to ransomware, phishing, data breaches, insider attacks, and supply chain attacks. Two objectives are outlined to achieve the goal:

(1) To develop a comprehensive set of project-level cyber risk factors using a systematic methodology that combines literature review, expert evaluation, and questionnaire surveys. The identified risk factors will cover different aspects of construction projects to ensure comprehensiveness and coverage.
(2) To pinpoint the advantages of the identified risk factors for future cyber risk assessments, which includes incorporating the network structure of projects, integrating both macro and micro project aspects, allowing for a more quantitative risk assessment, and infusing information on construction-unique vulnerabilities into the risk assessment models.

5.1.3 Contributions

The academic contributions of this study include: (1) offering a set of risk factors that can serve as a reference for future cyber risk assessments by scholars, and (2) providing a

systematic methodology framework that can be applied to identify risk factors for various risks. The practical contribution involves presenting a list of risk factors that construction practitioners can use as a checklist for proactive risk management. This can raise awareness among project managers and practitioners about which project characteristics require more attention and focus, so that they can have a better understanding of the cybersecurity status of construction projects and the corresponding causes.

5.2 Project as network

The project structure is crucial for cybersecurity because it outlines how complex the communication is among stakeholders and shows the potential weak points where attacks could occur within a network. In simpler terms, the way a project is organized can highlight how secure or vulnerable it is to cyber threats [32]. We draw inspiration from [32], where project dynamics were portrayed through an agent-based model and involved stakeholders are simulated to assess the vulnerability propagation. Similarly, we conceptualize the project structure as a network where various teams are interlinked, exchanging information via communication channels. Figure 5.1 illustrates this. Every team in a construction project (owners, contractors, subcontractors, etc.) is depicted as nodes (represented by circles). The edges (indicated by arrows) exemplify the communications/data exchanges among them through mediums like emails, texts, video calls, or software. The graph is organized into three distinct layers (rings), each identified by a unique color and encompassing diverse teams or sub-teams. The network graph helps us understand the project's structure and how cyber risk is distributed across all teams involved. It provides additional data that enhances the risk assessment model, making it easier to identify and address potential cybersecurity issues in the project network.

The example graph (Figure 5.1) delineates three layers with varying numbers of teams (sub-teams) and communication channels. In this particular case, Layer 1 has 10 sub-teams (purple circles) and 90 channels (arrows); Layer 2 comprises 14 sub-teams (yellow circles) and 20 channels (arrows); Layer 3 consists of 6 sub-teams (red circles) and 10 channels

(arrows). Different project delivery methods yield diverse structures and network graphs. Figure 5.1 illustrates the Integrated Project Delivery (IPD) delivery method, which involves a complex network of connected teams and sub-teams, creating a dynamic and complicated digital ecosystem. In contrast, other delivery methods, such as Design-Bid-Build (DBB), can result in a more sequential and linear network graph due to their more straightforward and less complex communication patterns. When collecting data, it is beneficial to guide the project manager or interviewee in creating a rough network graph like this to extract useful information.

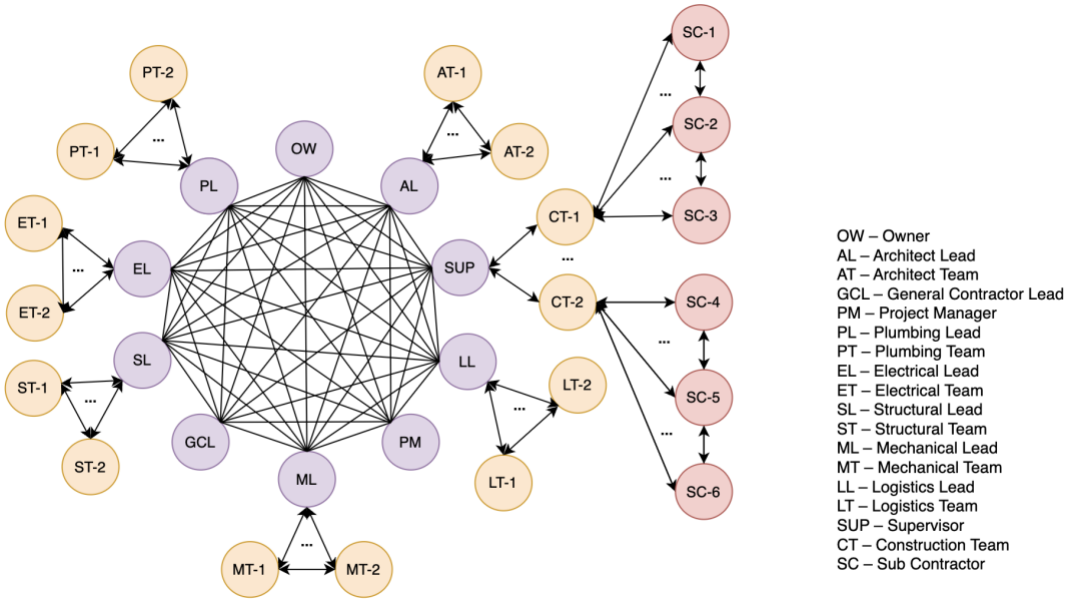


Figure 5.1. Network graph of a construction project (adapted from [32])

5.3 Steps and methods

This section presents the steps and methods used in our study to systematically identify a ready-to-use comprehensive set of risk factors that encompass various aspects of a construction project. Among the existing methods for identifying risk factors [20], we chose a combination of three: literature review, expert evaluation, and questionnaire survey for a well-rounded analysis and identification, borrowing insights from the literature

review in Section 5.1.1. The entire process of identifying risk factors can be visualized in Figure 5.2, which can be broadly divided into two stages: initial identification of risk factors based on literature review, followed by refining them through questionnaires and expert opinions.

5.3.1 Literature review

We began our study with a thorough review of the 20 publications on various types of risks in construction projects, as detailed in Section 5.1.1. These sources offer insights into risk factors related to project delays, supply chain issues, and beyond. Our objective was to extrapolate and adapt pertinent factors that might contribute to various risks, so that they can be included in the set of cyber risk factors. We also reviewed literature that discuss cybersecurity in the construction industry and thus might shed light on various other cyber-related factors, including [32], [50], [51], [174], [4], [6], [153]. Another source of literature review is the six textual sources concerning cybersecurity in the construction industry, published in our previous work as a database containing a large number of sentences [113]. These sources, collected from six types of construction cybersecurity literature, offer an in-depth exploration of cybersecurity challenges and risks in construction. Table 5.1 presents a statistical overview of these sources. We used the keyword “factor” to search the database and review the returned results.

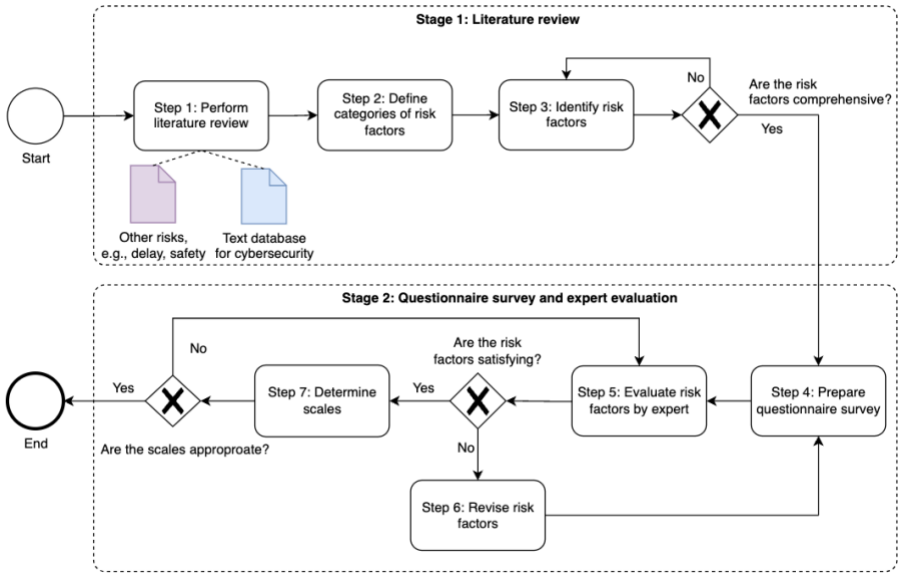


Figure 5.2. The process of risk factor identification

Table 5.1. Statistical information of the text database

Text source	Number of documents	Number of sentences	Total number of sentences
News articles and blogs	75 websites	71 K	
LexisNexis databases	3,968 pieces of news	596 K	
Academic papers	78 files	26 K	802 K
Books (chapters)	13 files	73 K	
Specifications/Standards	37 files	22 K	
Company reports	46 files	14 K	

5.3.2 Define risk factor categories

After reviewing the literature, we initially identified six aspects of a construction project that collectively offer a comprehensive view of the cybersecurity landscape, which are expounded as follows.

(1) Basic Information of the Project: This category includes general project information like company size, current project phase, and other foundational details.

Understanding these elements is vital as it sets the stage for evaluating the cybersecurity posture of the project.

(2) Project Structure: This refers to the organization and communication patterns of project teams and sub-teams. It's crucial to analyze this structure to comprehend the complexity and connectivity, which, in turn, influences the project's cybersecurity vulnerabilities and strengths.

(3) Cybersecurity Scores: Teams within the project are evaluated and scored based on their cybersecurity status. These scores, while offering insights into each team's cybersecurity posture, collectively provide a comprehensive view of the project's overall cybersecurity health, indicating areas of strength and vulnerability.

(4) Project Context: This category considers elements like governmental regulations and the construction company's financial health, which can externally influence the project. Assessing these factors is vital to understanding the impact of external pressures and resources on the project's cybersecurity resilience.

(5) Information Technology (IT) Factors: This category assesses the project's robustness and cybersecurity awareness from the perspective of information technology, focusing on factors such as the existence of a specialized IT department, the regularity of app updates, etc.

(6) Operational Technology (OT) Factors: This category evaluates the security measures in place for protecting physical and digital assets crucial to the project. Analyzing factors like access control mechanisms and the security of critical digital assets offers insights into the project's operational security and efficiency.

Categories 1: Basic Information of the Project and 4: Project Context, were inspired by the 20 publications mentioned in Section 5.3.1. After reviewing the literature, we found that basic information of the project (e.g., project type, project budget) and the context of the project (e.g., government regulation, financial health) can contribute to various risks, so we decided to include these two categories, assuming that they can contribute to cyber risk as well. Category 2, Project Structure, was inspired by the work on assessing the

cybersecurity vulnerability dispersion among the entities in a construction project [32], which reviews a project as a network and allows for the extraction of structural and communication patterns of the projects. Category 3: Cybersecurity Scores, was inspired by the CVSS scoring system for assessing the cyber risk status of entities in construction project networks, as included in works [50] and [51]. Category 5, Information Technology (IT), is inspired by the works [4], [6], [153] where the authors deeply discuss the vulnerabilities in the construction industry brought about by digital technologies, which can directly influence the cybersecurity status of construction projects. Category 6, Operational Technology (OT), was inspired by the work where the authors suggest that the cybersecurity status is also directly related to operational technologies in construction projects [174].

5.3.3 Internal identification and evaluation of risk factors

For each category, we recurrently reviewed the literature, referring back to Step 1, to identify, adapt, and extrapolate as many relevant risk factors as possible with the goal of achieving a comprehensive compilation. Simultaneously, we conducted regular internal discussions to assess these factors, totaling ten discussions over one month. In each discussion session, we assessed the previously identified risk factors considering four criteria: (1) their relevance to cybersecurity, (2) their significance in contributing to cyber risk, (3) the ease of understanding for industry professionals, and (4) the simplicity in collecting associated information. These criteria informed our decision on whether to include each factor. If there were disagreements on including certain risk factors, we kept them tentatively on the list. Later, we sought feedback on these disputed factors from external experts to gain additional insights. Ultimately, we compiled a list of 62 preliminary risk factors, as outlined in Table 5.3, with 9, 9, 6, 11, 16, and 11 factors allocated to each respective category. Throughout the study, each risk factor is formulated as a question. This approach enhances comprehension when presented to experts and industry practitioners, leading to higher response accuracy. The clarity aids experts in offering insightful feedback and promotes efficient data collection from the industry, particularly when selecting

specific scales of risk factors, a process detailed in Step 7.

5.3.4 Questionnaire survey

We developed a detailed questionnaire that was presented to experts to gather feedback on the 62 identified cybersecurity risk factors. This questionnaire is meticulously structured as follows:

- The basic information of the expert, including name, email, position/title.
- An executive summary of around 400 words provides background information, making it easy for company practitioners to understand our goals and objectives.
- The body is divided into six sections, each representing a specific category of risk factors, accompanied by an explanation of that category.
- Under each section are the corresponding risk factors initially identified (totaling 62), each accompanied by a definition, an in-depth definition, and an explanation of its potential impact on project cybersecurity. For clarity, we have included textual descriptions, illustrative graphics, or mathematical derivations where applicable.
- Each risk factor is followed by a multiple-choice scoring question, with options ranging from 1 to 5 (five levels). A level of 5 indicates the expert's perception that it is highly reasonable to include this risk factor in the cybersecurity evaluation.

5.3.5 Expert evaluation

The questionnaire was presented to three experts, two of whom specialize in cybersecurity and are affiliated with a cybersecurity company based in New York, United States. They were requested to provide a score ranging from 1 to 5, indicating their perception of the level of reasonability for including each risk factor from a cybersecurity perspective. We conducted online meetings and email communications to discuss the questionnaire, collect the scoring, and gather feedback. The third expert specializes in construction and is associated with one of the leading construction companies in Dubai, United Arab Emirates.

This expert provided feedback on the comprehensibility and appropriateness of the factors within the context of the construction industry at the same time. Communication with this expert was primarily conducted through email and phone calls. In later communications, after finalizing the risk factors, all three experts provided feedback on the scale design of these factors, detailed in Step 7.

Engaging experts from both cybersecurity and construction sectors enriches the validity and applicability of our identified risk factors. Their diverse insights ensure a well-rounded, cross-industry perspective, enhancing the robustness and relevance of the factors in both contexts. Table 5.2 provides a summarization of the experts and the collaborative process. Spanning approximately five months, this collaborative phase involved online meetings, emails, and phone calls. We have retained the recordings of these communications and copies of email correspondences, with permission, for future reference during the iterative risk factor revision process.

Table 5.2. Overview of expert information and collaborative process

No.	Expertise	Affiliation	Location	Years of Experience	Communication Method	Tasks and Feedback	Quantity
1	Cybersecurity	A cybersecurity scoring company	New York, U.S.	20+	Online ZOOM meetings & Email	Score and provide feedback on risk factors and their scales	12 meetings 60 emails 2 phone calls
2	Cybersecurity	A cybersecurity scoring company	New York, U.S.	8+	Online ZOOM meetings & Email		
3	Construction	A construction company	Dubai, UAE	10+	Email and phone calls	Provide feedback on risk factors and their scales	

5.3.6 Revise risk factors

The finalization of risk factors is an iterative process closely linked to step 5. After collecting scores and initial feedback, we adjusted the risk factors in the questionnaire, including their explanation, scope, and time span certain risk factors cover, etc. We then continued to consult with experts on these revisions to confirm if the risk factors were

appropriately refined at each stage; in the meantime, new risk factors, if any, were added as suggested. This approach allowed us to systematically incorporate feedback and align with expert insights. After concluding all discussions and communications, we eliminated risk factors with an average score below 3. One significant removal is the previous category, "Cybersecurity Scores." This change was prompted by all experts' advice, indicating that the risk factors within this category, primarily concerning the cybersecurity scoring of each team, are challenging to quantify and, therefore, not practical and should be eliminated from this study. We consolidate the average scores of the original risk factors and the suggested actions from the experts in Table 5.3. We originally planned to include a summary of the feedback in the table, but it has been omitted due to page constraints.

Table 5.3. The initial identification of risk factors and expert feedback

Category	NO.	Risk factor	Score	Action
1. Basic Information of the Project	1.1	Is your construction company global or local?	2	deleted
	1.2	What is the scale of your construction company?	4	revised
	1.3	What is the current phase of the construction project?	5	kept
	1.4	What is the weather of the current project phase?	1	deleted
	1.5	What is the total number of people involved in the project?	4	revised
	1.6	What is the percentage of people who have access to sensitive information?	5	kept
	1.7	What is the percentage of FTE (full-time employees) involved in the project?	1	deleted
	1.8	What is the percentage of people with over 10 years working in the organization?	1	deleted
	1.9	What is the region of the project?	5	revised
2. Project Structure	2.1	What is the project delivery method?	5	kept
	2.2	What is the number of sub-teams at different layers of the project?	3	kept
	2.3	What is the total number of teams in the project?	1	deleted
	2.4	What is the number of communication channels at different layers in the project?	4	kept
	2.5	What is the total number of communication channels among teams in the project?	1	deleted
	2.6	What is the average communication strength of channels at each layer?	2	deleted
	2.7	What is the overall communication strength of all channels?	2	deleted
	2.8	What is the average maturity of communication channels at different layers?	1	deleted
	2.9	What is the overall maturity of all communication channels?	1	deleted

Table 5.3 (continued)

Category	NO.	Risk factor	Score	Action
3. Cybersecurity Scores	3.1	What is the average risk score at each layer of the project?	2	deleted
	3.2	What is the average risk score over all the teams in the project?		
	3.3	To what extent are the risk scores spread out among the different teams in the project?		
	3.4	What is the percentage of teams that have high-risk scores (higher than 70)?		
	3.5	What is the highest value of risk scores over all the teams?		
	3.6	The IQR metric of the risk scores over all teams		
4. Project Context	4.1	What is the level of the cybersecurity impact and stakeholder engagement regarding cybersecurity?	2	deleted
	4.2	Whether there is a dedicated cybersecurity legal team?	5	kept
	4.3	What is the level of commitment to corporate governance and ethical practices regarding cybersecurity?	5	kept
	4.4	What is the percentage of the total project budget for cybersecurity management?	5	kept
	4.5	What is the level of financial risk?	2	deleted
	4.6	What is the frequency of daily information exchange?	2	deleted
	4.7	What is the average socioeconomic level of the involved people?	3	kept
	4.8	What is the degree of variation in the socioeconomic level of the involved people?	2	deleted
	4.9	What is the percentage of teams overlapping in different projects?	4	kept
	4.10	What is the average level of team member variability?	4	revised
	4.11	What is the average churn rate of all teams?	1	deleted
5. Information Technology (IT) Factors	5.1	Is the IT staff under-resourced for the size of the project?	2	deleted
	5.2	What is the number of user endpoints of digital devices?	5	kept
	5.3	What is the average computer/laptop security score?	1	deleted
	5.4	What is the ratio of Windows system vs non-Windows?	2	deleted
	5.5	What is the ratio of Android vs non-Android systems?	2	deleted
	5.6	Whether 90% of the computers/laptops have 90% of its applications are up to date?	2	deleted
	5.7	What is the construction-related APP/software maturity level?	2	deleted
	5.8	Is there a dedicated IT team for the project?	5	kept
	5.9	What is the level of stringency of cybersecurity policy?	2	deleted
	5.10	What is the level of commitment to cybersecurity policy?	2	deleted
	5.11	What is the average frequency of security training per year among all teams?	5	kept
	5.12	What is the percentage of people who fail phishing tests a second time after completing the required training?	4	revised
	5.13	What is the percentage of password reuse among employees in the project?	4	revised
	5.14	Is there any presence of exploitable critical findings in annual pen testing?	2	deleted

Table 5.3 (continued)

Category	NO.	Risk factor	Score	Action
	5.15	What is the estimated mean time to respond (MTTR) of the project in hours?	3	kept
	5.16	What is the number of production-impacting incident tickets per month in the project?	2	deleted
6. Operational Technology (OT) Factors	6.1	What is the total number of critical digital assets in the project?	5	kept
	6.2	What is the total number of important OT equipment and devices?	4	kept
	6.3	What is the average age of the important OT equipment?	3	kept
	6.4	Type of Network used in the project (Public Network or Private Network?)	5	kept
	6.5	What is the percentage of digital devices with firewalls or intrusion detection systems involved in the project?	5	kept
	6.6	What is the percentage of firewalls and endpoint detection systems with the latest security updates?	2	deleted
	6.7	What is the level of the physical access control mechanism?	3	revised
	6.8	Does access to the internet require Multi-Factor Authentication (MFA)?	4	revised
	6.9	What is the level of authentication mechanism to access the HMI (Human Machine Interface)?	4	kept
	6.10	What is the percentage of OT equipment in proximity to personnel during operation?	1	deleted
	6.11	What is the percentage of OT equipment isolated from the project's general network?	4	kept

As a result, we arrived at a final set of 32 risk factors, which have been re-categorized into 5 aspects: (1) Overall information of the project; (2) Project structure; (3) IT factors; (4) OT factors; (5) Management and human factors. This new categorization minimizes overlap among distinct categories while ensuring comprehensive coverage of construction project characteristics. The finalized set of risk factors is listed in Table 5.4, with counts of 7, 4, 9, 5, and 7, respectively. Importantly, Categories 3, 4, and 5 are tailored to evaluate a specific company and the project phase it is involved in, necessitating the consideration of its sub-teams. This led us to assign risk factors 3.1 and 3.2 to Category 3, even though they were initially intended for Category 1. This strategic reclassification aids companies engaged in future data collection to comprehend that these two factors, and the ones following, are aimed at assessing the distinct phase their company is involved in, ensuring targeted and relevant responses. Table 5.4 displays the finalized risk factors along with their corresponding numbering in the initial version. A detailed explanation of each risk

factor and its implications for project cybersecurity will be covered in Section 5.4.

Table 5.4. The finalized 32 risk factors

Category	NO.	Risk factor	Scales	Previous No.
1. Overall project information	1.1	What is the country of the project?	Asia, Europe, Africa, North America, South America, Antarctica, and Oceania. We initially asked for information about the country, and then derived the continent.	1.9
	1.2	What is the project budget?	<= $100,000, $100,000 - $500,000, $500,000 - $1 million, $1 million - $5 million, > $5 million	Newly added
	1.3	What is the percentage of the total project budget for cybersecurity management?	<= 1%, 1% - 2%, 2% - 3%, 3% - 4%, 4% - 5%, > 5%	4.4
	1.4	What is the project duration?	<= 3 months, 3 - 6 months, 6 - 12 months, 12 - 24 months, > 24 months	Newly added
	1.5	What is the total number of people involved in the project (labor excluded)?	<= 50, 51 - 100, 101 - 200, 201 - 300, 301 - 400, > 400	1.5
	1.6	What is the project type?	Transportation Infrastructure Projects, Government Projects, Healthcare Projects, Large-Scale Commercial Projects, Residential Projects, Other types	Newly added
	1.7	Whether there is a dedicated cybersecurity legal team for the project?	Yes, No, Unsure	4.2
2. Project structure	2.1	What is the project delivery method?	Design-Bid-Build (DBB), Design-Build (DB), Construction Manager at Risk (CMAR), Construction Management Multi-Prime (CMMP), Public-Private Partnership (PPP or P3), Integrated Project Delivery (IPD), Design/Build/Operate/Maintain (DBOM), Other types	2.1
	2.2	What is the number of sub-teams at different layers of the project?	Eight layers, each layer's choices are: <= 10, 11 - 20, 21 - 30, 31 - 40, > 40, N/A ("N/A" means this layer is not existent)	2.2

Table 5.4 (continued)

Category	NO.	Risk factor	Scales	Previous No.
	2.3	What is the number of communication channels at different layers in the project?	Eight layers, each layer's choices are: <= 50, <= 100, <= 150, <= 200, < 250, <= 300, > 300, N/A ("N/A" means this layer is not existent)	2.4
	2.4	What is the percentage of teams overlapping in different projects?	<= 20%, 21% - 40%, 41% - 60%, 61% - 80%, 81% - 100%	4.9
3. IT factors	3.1	What is the scale of your company?	Five choices: <= 30, 31 - 60, 61 - 100, 101 - 150, > 150	1.2
	3.2	What is the phase of the construction project when your company is involved?	Planning and Bidding phase, Design phase, Construction phase, Maintenance & Operation phase, Demolition phase	1.3
	3.3	Is there a dedicated IT team for the project?	Yes, No, Unsure	5.8
	3.4	What is the total number of critical digital assets?	<= 50, 51 - 200, 201 - 400, 401 - 600, > 600	6.1
	3.5	What is the total number of user endpoints of digital devices for the project?	<= 50, 51 - 200, 201 - 400, 401 - 600, > 600	5.2
	3.6	What is the percentage of digital devices with firewalls or intrusion detection systems involved in the project?	<= 20%, 21% - 40%, 41% - 60%, 61% - 80%, 81% - 100%	6.5
	3.7	What is the network type used for the project: Public or Private?	Public network, Private network, Both public and private network	6.4
	3.8	What is the percentage of individuals who fail phishing tests after completing mandatory training?	<= 20%, 21% - 40%, 41% - 60%, 61% - 80%, 81% - 100%	5.12
	3.9	What is the estimated Mean Time to Respond (MTTR) in hours?	Within 1 hour, 1 - 4 hours, 4 - 8 hours, 8 - 24 hours, Above 24 hours	5.15
4. OT factors	4.1	What is the total number of important OT equipment involved?	<= 30, 31 - 60, 61 - 90, 91 - 120, 121 - 150, > 150	6.2
	4.2	What is the level of physical access control mechanism to OT equipment?	Level 1, Level 2, Level 3, Level 4, Level 5	6.7
	4.3	What is the percentage of OT equipment isolated from the project's general network?	<= 20%, 21% - 40%, 41% - 60%, 61% - 80%, 81% - 100%	6.11

Table 5.4 (continued)

Category	NO.	Risk factor	Scales	Previous No.
	4.4	What is the average age of the important OT equipment, in years?	<= 1, 1 - 3, 4 - 7, 8 - 10, > 10	6.3
	4.5	What is the level of authentication mechanism to access the HMI (Human Machine Interface)?	Level 1, Level 2, Level 3, Level 4, Level 5	6.9
5. Management and human factors	5.1	What is the average level of commitment to corporate governance, ethical practices and cybersecurity policy?	Level 1, Level 2, Level 3, Level 4, Level 5	4.3
	5.2	What is the average frequency of security training per year?	<= 10, 11 - 20, 21 - 30, 31 - 40, 41 - 50, > 50	5.11
	5.3	Do you allow password reuse for any project-related software, systems, or accounts (e.g., project management tools, email, internal networks, file storage, etc.)?	Yes, No	5.13
	5.4	Does internet access within your construction project require Multi-Factor Authentication (MFA) or utilize other methods such as biometrics or face recognition?	Yes, No	6.8
	5.5	What is the percentage of people who have access to sensitive information in the project?	<= 10%, 11% - 30%, 31% - 50%, 51% - 70%, 71% - 90%, 91% - 100%	1.6
	5.6	What is the average team member variability over a 3-month period?	<= 20%, 20% - 40%, 40% - 60%, 60% - 80%, 80% - 100%	4.10
	5.7	What is the average socioeconomic level of the people involved in the project?	Level 1, Level 2, Level 3, Level 4, Level 5	4.7

5.3.7 Determine the scales of risk factors

In risk assessment, a quantitative approach is often preferred over a qualitative one, as it yields more numerically based and, thus, objective results. However, this approach requires quantified or fine-grained risk factor inputs, an element often missing in existing literature, including [32], [49], [51], [141]. As an illustration, in the study of predicting project delay risk [141], the authors subjectively classified risk factors as high or low risk without offering numerically-based or substantial evidence to back their classifications. Our study

aims to take a more numerical approach by incorporating risk factor scales, where each risk factor is categorized into distinct levels, categories, or numerical values, paving the way for more quantitative future risk assessments. For instance, risk factor 1.3 – the percentage of the total project budget for cybersecurity management – can be divided into six scales: <= 1%, 1% - 2%, 2% - 3%, 3% - 4%, 4% - 5%, > 5%. Similarly, risk factor 1.4 – the project duration – can be divided into five distinct time intervals: <= 3, 3 - 6, 6 - 12, 12 - 24, and> 24 months.

The initial determination of these scales was based on our expertise and fundamental understanding of construction projects. They were incorporated into the final version of the questionnaire with 32 risk factors presented to the experts, as outlined in Step 4. The determination of the final scales also underwent an iterative process. The feedback and insights garnered from them prompted refinements to the scales—some were expanded, others narrowed, and certain risk factors transitioned from level representation to numerical representation. The iterative refinements and expert validations have shaped the final scales that strike a balance between detailed granularity and practical feasibility. The finalized scales for each risk factor are presented in Table 5.4.

5.4 Results of the 32 risk factors

This section delves into the impact of each risk factor on the cybersecurity of construction projects, aiding in a deeper comprehension of their significance. Table 5.4 provides a detailed breakdown of the scales associated with each risk factor.

5.4.1 Category (1): Overall Information of the Project

This category encompasses seven factors that offer a comprehensive outlook on the project's foundational elements. They collectively give insights into the project's environmental, financial, temporal, and human aspects, along with legal considerations, setting the context for a tailored cyber risk assessment. It is recommended to involve a project manager familiar with the overall project to provide the necessary data.

- **Risk factor 1.1: What is the country of the project?** The project's country influences its cybersecurity due to varied regulations, technology infrastructure, and cyber threat levels in different nations. Each country's unique cybersecurity regulations and threat landscape necessitate tailored security measures. Therefore, the location can impact the complexity and nature of security strategies needed to mitigate potential cyber risks effectively.
- **Risk factor 1.2: What is the project budget?** The overall project budget is a relevant cybersecurity risk factor, as limited funding may constrain resources allocated to IT infrastructure, security controls, auditing, training, and expertise. Larger budgets allow more investment in robust cybersecurity measures and skilled personnel to implement best practices.
- **Risk factor 1.3: What is the percentage of the total project budget for cybersecurity management?** The allocation to cybersecurity affects a project's defense capability. Insufficient funds can lead to vulnerabilities, while adequate investment ensures robust security measures, skilled personnel, and effective responses to cyber threats, safeguarding project integrity and data security.
- **Risk factor 1.4: What is the project duration?** Project duration impacts cybersecurity due to evolving threats. Longer projects face changing, potentially escalating cyber threats requiring ongoing adaptations in security measures. Shorter timelines may minimize exposure but still necessitate comprehensive security plans to protect against breaches.
- **Risk factor 1.5: What is the total number of people involved in the project? (labor excluded)** The number of non-labor participants correlates with cybersecurity risk. More individuals increase potential points of vulnerability due to human error or insider threats. Managing and training a larger group in cybersecurity protocols becomes vital to mitigate risks of breaches and data leaks.
- **Risk factor 1.6: What is the project type?** Different project types present distinct cybersecurity challenges. Infrastructure, residential, commercial, or industrial

projects each have unique security needs and vulnerabilities. Identifying the project type helps tailor cybersecurity measures to specific risks and regulatory requirements, enhancing the effectiveness of security protocols.

- **Risk factor 1.7: Whether there is a dedicated cybersecurity legal team?** A cybersecurity legal team signifies enhanced preparedness and response to legal issues arising from cyber threats. Their absence can expose the project to regulatory non-compliance, liability risks, and inadequate legal response during cybersecurity incidents, impacting project security and integrity.

5.4.2 Category (2): Project Structure

Four factors in this category provide an overview of the project's organizational and communication architecture. They highlight the structural complexity and interconnectedness that could influence the project's vulnerability to cyber threats. A network graph, similar to Figure 5.1, can be drawn to visually depict these relationships, aiding in comprehending the risk factors and deriving needed statistical figures. It is recommended to involve a project manager familiar with the overall project to help with data collection.

- **Risk factor 2.1: What is the project delivery method?** Different delivery methods (e.g., Design-Bid-Build (DBB), Integrated Project Delivery (IPD)) have varying levels of collaboration, contractual relationships, and information sharing among stakeholders. The specific delivery method can impact the project's cybersecurity posture by affecting access to sensitive information, communication channels, and the implementation of security protocols.
- **Risk factor 2.2: What is the number of sub-teams at different layers of the project?** The number of sub-teams at various project layers can be an indicator of cyber risk. Typically, outer-layer sub-teams (shown in Figure 5.1), potentially smaller with limited cybersecurity resources, pose more risks. Inner-layer teams are

often larger, better equipped, and more aware of cybersecurity. Thus, tracking sub-teams across layers helps in targeted risk management.

- **Risk factor 2.3: What is the number of communication channels at different layers in the project?** This factor gauges the security implications of the number of communication channels within each layer of the project. A higher count can amplify vulnerabilities due to increased data exchange points, necessitating enhanced security protocols to protect sensitive information and maintain system integrity at each individual layer.
- **Risk factor 2.4: What is the percentage of teams overlapping in different projects?** The percentage of teams working on multiple projects simultaneously (overlap) increases cyber risks through shared resources, people, and potential security gaps. More overlap raises breach, gap, and dependency risks, requiring management to mitigate threats. Team overlaps matters because shared resources and personnel heighten cyber vulnerabilities.

Categories (1) and (2) emphasize the general information of the project, whereas Categories (3), (4), and (5) are tailored to specific phases of the project and are particularly relevant to the specific company engaged in those phases. This approach ensures that the risk assessment model to be established is comprehensive, incorporating both generic project data and phase-specific data. Furthermore, it's essential that Categories (3) through (5) take into account all sub-teams during the data-gathering process to ensure a thorough and accurate assessment.

5.4.3 Category (3): IT Factors

Nine elements in this category explore the company's IT infrastructure and behaviors, highlighting the integral role of IT in managing cybersecurity. These factors are crucial as they can reflect specific IT vulnerabilities of this company, enabling targeted defense strategies. By evaluating IT factors and forming dedicated mitigation strategies, companies

can enhance cyber resilience, ensuring project security against evolving cyber threats, making them essential for informed cybersecurity planning. It is recommended to involve both the project manager of this company and an IT professional to help with data collection.

- **Risk factor 3.1: What is the scale of your company?** Larger firms often have more resources for advanced security measures but can be targets for cyberattacks due to their visibility and data value. In contrast, smaller companies, though less visible, may be more vulnerable due to limited security resources and expertise despite potentially having fewer threats to contend with.
- **Risk factor 3.2: What is the phase of the construction project when your company is involved?** The project phase influences cybersecurity needs. For example, the design phase may focus on data protection and privacy, while the construction phase phases might emphasize system security and integrity. Each phase presents unique cyber challenges, requiring tailored strategies to mitigate risks effectively.
- **Risk factor 3.3: Is there a dedicated IT team for the project?** A dedicated IT team enhances real-time response and management of cybersecurity issues. Without it, a project may face delayed responses to threats, increased vulnerabilities, and potential breaches, impacting the security and integrity of project data and systems.
- **Risk factor 3.4: What is the total number of critical digital assets?** The quantity of critical digital assets indicates the potential risk exposure. These assets encompass various data types, including BIM files, project plans, management data, contracts, communication platforms, and databases. More assets mean increased vulnerability points, requiring enhanced security measures to ensure the integrity and confidentiality of sensitive data and systems amidst various cyber threats. We can estimate the total count by considering the number of stored digital documents across various project management systems, including cloud storage, PCs, and

mobile devices used for the project. Consider the number of stored digital documents such as BIM files, project plans, contracts, and other relevant files. While it may not provide an exact count, this estimation method gives a reasonable understanding of the overall volume of digital assets.

- **Risk factor 3.5: What is the total number of user endpoints of digital devices for the project?** As the number of connected devices, such as laptops and smartphones, increases in a network, so does the attack surface. Each additional device provides another potential entry point for cyber threats, making the network more vulnerable to security breaches. To acquire accurate data on user endpoints, we can utilize a centralized information or inventory system that catalogs all devices within a project. This resource allows for the identification and counting of every unique device, including laptops, desktops, and smartphones, integral to the digital infrastructure.
- **Risk factor 3.6: What is the percentage of digital devices with firewalls or intrusion detection systems involved in the project?** The percentage of devices with firewalls or intrusion detection systems reflects the project's defense depth against cyber threats. A higher percentage indicates stronger security, while lower figures suggest potential vulnerabilities and the need for enhanced protective measures.
- **Risk factor 3.7: What is the network type used for the project: Public or Private?** Choosing between a public or private network impacts a project's cybersecurity. Public networks can be vulnerable to attacks due to easier access, while private networks offer enhanced security controls but may be costly and complex to manage. Each type requires specific security approaches.
- **Risk factor 3.8: What is the percentage of individuals who fail phishing tests after completing mandatory training?** The percentage of individuals failing phishing tests post-training indicates the effectiveness of education programs and the remaining vulnerability to phishing attacks, highlighting areas for improvement in training and awareness to enhance cybersecurity defenses.

- **Risk factor 3.9: What is the estimated Mean Time to Respond (MTTR) in hours?** The MTTR estimates the average time required for the project team, including sub-teams, to respond to and resolve cybersecurity incidents. It is calculated by tracking the time from the moment an incident is detected to the time it is fully resolved. Add up the response and resolution times for all incidents, then divide by the total number of incidents to get the average MTTR. This process should be done for the team and all sub-teams, and the results should be averaged to get an overall MTTR for the project.

5.4.4 Category (4): OT Factors

This category, containing five factors, centers on the project's operational technology. It looks at the equipment and systems pivotal for managing physical processes, underscoring their vulnerability and the essentialness of strategic measures to enhance security and prevent unauthorized access. Important OT equipment in construction includes Industrial Control Systems (ICS); Programmable Logic Controllers (PLCs); Human-Machine Interfaces (HMIs); sensors and actuators; communication networks and specific protocols; Building Management Systems (BMS); access control, security systems such as surveillance cameras and intrusion detection systems; environmental monitoring systems; control panels and field devices; SCADA systems; and remote monitoring and control systems [174]. It is recommended to involve a manager well-acquainted with the company and deeply involved in the project during this phase to help with data collection.

- **Risk factor 4.1: What is the total number of important OT equipment involved?** The number of critical OT equipment pieces is directly proportional to the potential attack surface. A higher count indicates an increased risk, necessitating more complex and robust security protocols to mitigate the risks of operational disruptions, data breaches, and system failures, thereby ensuring the operational continuity and integrity of the construction project.

- **Risk factor 4.2: What is the level of physical access control mechanism to OT equipment?** Evaluating physical access control mechanisms, including secure entry points, surveillance, and ID badges, is vital to prevent unauthorized physical access to sensitive systems and data. More stringent controls like biometrics and visitor audits indicate a higher security level. Robust physical access control is crucial for maintaining cybersecurity.
- **Risk factor 4.3: What is the percentage of OT equipment isolated from the project's general network?** Measuring the percentage of operational technology assets separated from the general network assesses the extent of critical system isolation. Higher levels of OT equipment segregation into distinct networks enhance cyber resilience by preventing unauthorized access, containing threats, and minimizing impacts to maintain functionality.
- **Risk factor 4.4: What is the average age of the important OT equipment, in years?** The average age indicates potential vulnerabilities from aging infrastructure, like inadequate security, maintenance issues, and decreased performance. Tracking average age helps identify needs for upgrades, maintenance, and resources to ensure reliable, secure OT operations and mitigate risks from outdated equipment.
- **Risk factor 4.5: What is the level of authentication mechanism to access the HMI (Human Machine Interface)?** This factor is crucial given the pivotal role HMI plays as the interface connecting operators to machinery or plant control systems. It evaluates the strength of authentication protocols for HMI access. Higher levels indicate robust measures like multi-factor or biometric verification. Including this evaluates potential vulnerabilities, ensuring only authorized access and control through the HMI. Assessing the HMI access authentication mechanism enhances cybersecurity and safeguards the project's integrity and safety.

5.4.5 Category (5): Management and Human Factors

Seven factors in this category explore the company's governance, ethical standards, and

cybersecurity culture. It underscores the vital role of human elements and management practices in bolstering the project's overall cybersecurity posture, emphasizing a holistic approach that combines technology and human effort. It is recommended to involve a manager well-acquainted with the company and deeply involved in the project during this phase to help with data collection.

- **Risk factor 5.1: What is the average level of commitment to corporate governance, ethical practices and cybersecurity policy?** This reflects a project's commitment to incorporating principles of good governance and ethics into its cybersecurity efforts. It encompasses compliance with laws and regulations, transparency, accountability, and the making of ethical decisions. A strong commitment bolsters stakeholder trust and contributes to the project's success, whereas inadequate governance may result in reputational harm and legal complications.
- **Risk factor 5.2: What is the average frequency of security training per year?** The average yearly cybersecurity training frequency for all teams indicates how proactively knowledge and skills are enhanced. More frequent sessions promote awareness, inform on threats and best practices, and foster a culture of security. Regular training enables teams to effectively contribute to the project' s overall security posture.
- **Risk factor 5.3: Do you allow password reuse for any project-related software, systems, or accounts (e.g., project management tools, email, internal networks, file storage, etc.)?** Including this factor is crucial because password reuse can significantly heighten cybersecurity risk. It measures the project's vulnerability to unauthorized access and potential breaches. Assessing this aspect enables the implementation of stringent password policies to enhance overall security and data protection.
- **Risk factor 5.4: Does internet access within your construction project require Multi-Factor Authentication (MFA) or utilize other methods such as**

biometrics or face recognition? The use of MFA, biometrics, or face recognition enhances security by adding layers of authentication, reducing unauthorized access risks. The absence or inadequacy of these measures can increase vulnerabilities, emphasizing the need for robust authentication to secure internet access and protect sensitive project data and systems.

- **Risk factor 5.5: What is the percentage of people who have access to sensitive information in the project?** Here, sensitive information refers to any data that is protected against unwarranted disclosure, such as personal identification information, financial records, and proprietary project details. A larger percentage of people with access to sensitive data increases the risk of breaches. It underscores the need for stringent access controls and security protocols to protect sensitive information, minimize insider threats, and ensure that data confidentiality and integrity are maintained throughout the project's lifecycle.
- **Risk factor 5.6: What is the average team member variability over a 3-month period?** Team member variability refers to the extent of changes in team composition during a project. More frequent changes increase cyber risks, as new members may introduce vulnerabilities while departing members leave gaps. Frequent team changes impact security practices and knowledge retention. A higher level of variability indicates more frequent team changes, increasing project cybersecurity risks.
- **Risk factor 5.7: What is the average socioeconomic level of the people involved in the project?** This refers to the collective economic and social standing of the project's personnel, measured by income, education, and occupation. It is essential for assessing potential disparities in cybersecurity awareness and practices. Individuals with higher socioeconomic statuses often exhibit better cybersecurity attitudes and behaviors, influencing the overall project's cyber risk.

5.5 Discussions

The identified risk factors play a critical role in establishing the risk assessment model, which aims to achieve a more quantified assessment of risks, a topic to be studied in Chapter 6. Therefore, this section discusses the advantages of these risk factors for the model. Additionally, limitations and future work are discussed.

5.5.1 Capturing project structure dynamics

In our study, we innovate by viewing the project as a multi-layered network, providing a detailed understanding of its complex structure and associated cybersecurity vulnerabilities. Each layer includes a variety of teams, sub-teams, and communication channels. From these, we have identified specific risk factors in Category 2 that extract statistical features concerning the spread of sub-teams and communication channels across the project's layered network. This detailed, layer-specific information enhances future risk assessment models, enabling them to capture the dynamic and complex nature of modern construction projects. Including this structural information can improve the accuracy and reliability of the predictions of the risk assessment model, setting our approach apart from previous studies [41], [49], [109], which neglect the importance of a project's layered structure in assessing risks.

5.5.2 Enhancing specificity with contextual insight

Risk factors in Categories 1 and 2 offer a broad overview of the project, ensuring a comprehensive encapsulation of general information and establishing a foundational context for the risk assessment model. In contrast, risk factors in Categories 3, 4, and 5 derive from specific project phases, with data sourced directly from the companies involved, ensuring phase-specific specificity. This integration of broad and detailed data boosts the model's efficacy in making nuanced risk predictions that are specific to the phase the company is involved in without losing the contextual information of the project. As a result, the model makes risk predictions that are both comprehensive and applicable across

distinct project phases, marking an advancement in the field of cyber risk assessment.

5.5.3 Enabling a more quantitative risk assessment

Many works, including [32], [141], [175], mainly used qualitative analysis for risk assessment, where expert opinions and subjective judgments determined whether a risk factor, a stakeholder, or a system was considered risky without a concrete numerical standard for reference. Our study allows for more quantitative risk assessments by segmenting each risk factor into distinct scales and requiring data collection before determining the risk status. For example, risk factor 4.3 – the percentage of OT equipment isolated from the general network - is divided into five scales: <= 20%, 21% - 40%, 41% - 60%, 61% - 80%, and 81% - 100%. After data about OT equipment isolation is collected, it is compared with the predefined scales to determine the risk status of the risk factor, eliminating ambiguity and subjectivity as these scales have been vetted and validated. Although the interpretation of these scales is still influenced by the risk analyzer's criteria, establishing these criteria beforehand ensures that the numerical values-based assessments are more objective. This approach augments the consistency and comparability of risk evaluations across various contexts.

5.5.4 Addressing unique industry vulnerabilities

In addition to the risk factors addressing general IT vulnerabilities, some align with the distinct vulnerabilities of the construction industry, as delineated in Section 5.1. These correlations are explicitly mapped in Table 5.5, with an explanation of how these risk factors can be efficacious. By analyzing diverse risk factors, the model can balance between addressing general cybersecurity concerns while simultaneously addressing challenges specific to the construction sector.

5.5.5 Limitations and future works

This study has a limitation pertaining to the undetermined specific risk degree associated

with each scale of each identified risk factor. An example of this limitation is observed in the context of cybersecurity investments within a project. Here, the varying degrees of risk associated with different budget percentages allocated to cybersecurity remain undefined. Determining these risk degrees for each risk factor is important as it lays the foundation for developing a more quantitative risk assessment model. The resolution of this limitation is not addressed within the current scope of this study but is left for future research. This forthcoming research will explore a variety of qualitative and quantitative methodologies to ascertain the most effective approach for determining these risk degrees.

Table 5.5. Correlation mapping between industry vulnerabilities and risk factors

Construction Industry Vulnerability	Risk Factors	Explanation
Fluidity of Team Compositions	2.1 - 2.4	These factors address the project's structural and communication dynamics, revealing the challenges induced by changing team compositions and structures. This reflects the adaptability needed in various phases of construction projects.
Diverse Workforce	1.5, 3.8, 5.2	These factors provide insights into the diversity of the workforce's cybersecurity awareness. It highlights potential gaps and vulnerabilities, emphasizing the need for targeted training and awareness programs.
Widespread communications networks	2.3	This factor illuminates the expansive and multi-layered communication networks in construction projects, pinpointing potential vulnerabilities and areas for enhanced data protection and communication security.
Frequent Information/Data Exchange	3.4 - 3.9, 4.1 - 4.5	These IT and OT factors are pivotal in evaluating the risks and vulnerabilities emerging from the extensive digital information exchange, underscoring the need for robust, tailored security protocols.
Blurring of Project Boundaries	2.4, 5.5	These factors identify the potential for overlapping team roles and access to sensitive information across projects, signaling heightened risks of data leaks and the need for stringent access and information management protocols.

5.6 Conclusions

The identification of cyber risk factors associated with construction projects is beneficial for the industry to better understand the causal relationship between cyber risks and project characteristics; however, such studies are absent in the literature. To bridge this gap, this study adopted a systematic approach combining literature review, expert evaluation, and

questionnaire survey, which resulted in the identification of 32 risk factors covering five aspects of construction projects: Overall Information of the Project, Project Structure, IT Factors, OT Factors, and Management and Human Factors. These factors, capturing both general IT vulnerabilities and distinct vulnerabilities linked to construction projects, are capable of evaluating a broad spectrum of cyber risks.

The identified risk factors offer four advantages for cyber risk assessment in this domain: Treating the project as a network allows for the capture of complex project structures and, consequently, the associated cybersecurity vulnerabilities inherent in each layer's dynamics. The integration of both broad and phase-specific data enhances the model's predictive specificity to project phases without compromising the overall project's contextual insight, ensuring tailored and context-aware risk assessments. The incorporation of risk factor scales facilitates a more quantitative risk assessment, boosting objectivity and consistency in risk evaluations and mitigating the influence of subjective judgments and expert biases. The inclusion of industry-specific risk factors ensures that the model is not only comprehensive but also tailored to address unique vulnerabilities inherent to the construction sector. One of the limitations of the study is the lack of determination of specific risk degrees for identified factors. Future research aims to explore methods for determining these degrees, crucial for quantitative risk assessment model development.

Chapter 6
Assessing Cyber Risks in Construction Projects: A Machine Learning-Centric Approach

In this chapter, we develop an ML-centric approach to assess common cyber risks among the ones identified in Chapter 4 at the project level, which can serve as a tool for project managers for dynamic risk assessment throughout the progression of construction projects. The features of the ML models are the risk factors identified in Chapter 5. To this end, a simulated dataset is generated using Monte Carlo simulations and ensemble labeling methods, which are validated and utilized for model training. A two-phase model development strategy is proposed to select the optimal model and determine the optimal ensemble labeling weights. ML feature analysis methods are adapted to identify risk factors that are generally important to construction projects and those that have a high contribution to risks of specific projects. A greedy optimization algorithm is proposed to formulate the risk reduction strategy. To demonstrate the applicability of developed approach, a case study is conducted on a real construction project. This study then discusses the complex construction cybersecurity landscape, the generally important risk factors, and the

prospects of the developed approach for widespread industry application.

6.1 Introduction

The construction sector trails other industries in cybersecurity, experiencing a dramatic surge in cyber incidents from nearly 10 in 2013 to almost 520 in 2022—a 5100% increase [7]. This escalation is particularly notable in five distinct types of cyber risks:

(1) Ransomware is a prevalent cyber risk in the construction industry where attackers encrypt critical project data and demand a ransom for its decryption. This form of attack can cause severe disruptions to construction projects. Project managers may find essential documents such as blueprints and financial records encrypted, leading to project delays, significant financial losses due to ransom payments, operational shutdowns, and damage to client relationships. The repercussions of such attacks can be far-reaching, impeding the timely completion of projects and potentially souring professional relationships that are crucial to business success.

(2) Phishing risks in construction involve schemes designed to deceive individuals into disclosing sensitive information, often through misleading emails. For construction professionals, succumbing to phishing attempts can result in unauthorized access to financial accounts or confidential project data. The impact of phishing can be severe, leading to financial loss and jeopardizing the security of the entire project. As phishing becomes increasingly sophisticated, the need for vigilance and robust cybersecurity measures is more critical than ever to protect against such deceptive tactics.

(3) Insider threats in the construction sector are particularly insidious, occurring when individuals within an organization, often trusted employees, act with malicious intent. In the context of construction, insider threats can result in the theft or sabotage of key materials, or the leaking of proprietary information, which can lead to unforeseen expenses, delays in project timelines, and legal challenges concerning contract or intellectual property breaches. The threat posed by insiders underscores

the importance of comprehensive security protocols and employee monitoring to mitigate these risks.

(4) Data breaches represent another significant cyber threat, involving incidents where confidential, sensitive, or protected information is accessed, stolen, or disclosed without authorization. Construction projects with substantial digital data repositories are at risk of such breaches, which can lead to financial loss, legal complications, and long-lasting reputational damage that may adversely affect future business opportunities. Protecting this data is paramount, requiring robust cybersecurity frameworks to prevent unauthorized access.

(5) Supply chain attacks in the construction industry target vulnerable components within a supply network, compromising organizations by undermining software, hardware, or services through third-party relationships. These attacks can introduce severe disruptions in the complex construction supply chain, such as supplier insolvency, transportation issues, or substandard material quality. Such disruptions can cause project delays, escalate costs for alternative solutions, and risk the integration of inferior materials, potentially compromising the overall quality and integrity of construction projects.

The growing prevalence of cyber risks in the construction industry clearly indicates the necessity for strengthened cybersecurity management strategies tailored to construction projects [6]. Project managers equipped with predictive tools that can forecast potential cyber risks in the upcoming phases of a project will be able to preemptively address these risks. This foresight would enable them to align their risk mitigation actions with the specific risk tolerance levels of their projects—ranging from reducing risks to averting them entirely. In light of this, there is a preference for a risk assessment tool capable of dynamic risk assessment. Such a tool would ideally accept project information as inputs at any given time and use a predictive model to assess the level of cyber risk. It should also be able to pinpoint the key risk factors contributing to a heightened risk level and propose effective strategies for risk reduction tailored to the specific context of the project. The goal

is not merely to react to cyber risks but also to anticipate and neutralize them before they can impact the project's progress and success.

However, existing research on cybersecurity within the construction industry remains sparse, with a notable absence of comprehensive and dynamic cyber risk assessment tools. ML techniques, which have been widely explored to build risk assessment tools for other risks such as delay, financial, and environmental risks in construction projects, could potentially offer a robust solution for a comprehensive and dynamic cyber risk assessment.

6.1.1 Related works

As reviewed in Chapter 2, cybersecurity studies within the construction industry are limited. Among the three works on risk assessment, Mantha and García de Soto's study [32] highlighted the role of agent-based models in understanding and calculating the spread of vulnerability in the complex interactions among various stakeholders. However, their method only addressed stakeholders' vulnerabilities and overlooked the likelihood of corresponding risk occurrences, which should take into account project-specific factors. Additionally, it is an oversimplified model in its initial phase and is difficult to generalize across different construction projects. Their other work [51] utilized the Common Vulnerability Scoring System (CVSS), offering a systematic approach to assessing risk in construction networks. However, the scoring process was largely manual and lacked the flexibility for easy adaptation to new project scenarios. Mohamed Shibly and García de Soto [49] focused on quantitative risk assessment through attack tree-based threat modeling. They used the tampering threat associated with a 3D printer as a case study but acknowledged that the existing model's manual nature makes it cumbersome and inflexible for broader use in the construction industry. This is particularly because both tree creation and risk propagation are time-consuming manual processes. Furthermore, this work was conceptual and did not include concrete evidence from real-world projects for validation. In addition to the limitations of each study, the scope of the three studies is fragmented, ranging from project-level to single-threat-level assessments, and lacks a unified

framework for evaluation.

ML is increasingly applied as either a technique or a framework in the construction industry to assess a range of risks such as delays, safety, cost overruns, legislative compliance, disputes, defects, etc., which has the potential to address cyber risks in construction. A considerable portion of these works is dedicated to assessing delay, safety, and cost overrun risks. Within this scope, scholars utilize a spectrum of techniques from traditional ML algorithms to neural networks, as well as advanced DL models. This variety in algorithmic application illustrates the breadth and depth of ML in managing construction risks.

Delay risk is a critical concern due to its potential to disrupt timelines, inflate costs, and strain client relationships, necessitating effective management strategies. Regarding this, Sanni-Anibire et al. [176] developed an ML model for delay risk assessment in tall building projects by analyzing 36 identified delay risk factors through surveys from subject matter experts. They employed KNN, Artificial Neural Networks (ANN), SVM, and Ensemble methods, with ANN showing the highest accuracy at 93.75%. Gondia et al. [141] identified construction project delay risk sources, compiles a dataset, and applies decision tree and naïve Bayesian classification algorithms to develop predictive models. They evaluated these models using cross-validation and performance indices, finding the naïve Bayesian model superior in predictive performance. Fitzsimmons et al. [177] developed a hybrid ML model for construction schedule risk analysis, combining Gaussian Mixture Modelling, Empirical Bayesian Networks, Support Vector Machine, and Monte Carlo simulation. Trained on a dataset from 302 projects, it accurately predicted delays, outperforming traditional methods.

Safety risk in construction projects endangers workers, disrupts timelines, and can lead to legal consequences, necessitating rigorous safety measures to mitigate potential hazards. George [178] et al. developed an ensemble ML model for construction site risk assessment using a dataset from Occupational Safety and Health Administration (OSHA), focusing on accident-related attributes. They employed classifiers like Gradient Boosting Machine

(GBM) and Extreme Gradient Boosting (XGBoost), finding GBM with Critical Factors-2 (CR-2) attributes as the best predictor through resampling methods and ensemble modeling. Liu and Tian [179] developed a construction safety assessment model using extension cloud theory and distributed ML algorithms . They established a safety evaluation index system, apply the analytic hierarchy process for index weighting, and implement an early warning mechanism, demonstrating practicality and effectiveness through engineering examples. Poh et al. [180] used ML to develop safety leading indicators for construction sites, applying Boruta feature selection and decision tree to identify significant variables from a dataset, and training models with five algorithms. Random Forest showed the best performance in predicting accident occurrence and severity. Gondia et al. [181] developed an ML framework to predict construction site injury severity, utilizing glass-box and black-box models for interpretability and accuracy. They employed a ranking algorithm to identify key injury factors, validated model performance through robust metrics, and used the OSHA injury data set for demonstration, offering insights for targeted risk mitigation strategies.

Cost overruns disrupt budgets, strain finances, hinder project success, erode trust, delay completion, and lead to resource misallocation, jeopardizing stakeholder satisfaction and organizational credibility. Regarding this, Bilal et al. [182] proposed Applied Machine Learning (AML) guidelines for the construction industry, focusing on creating robust models for enterprise solutions, including profit margin estimation. The methods include interpretable ML, feature engineering, and a detailed AML process, developed through experience with the Construction Simulation Tool (CST) project. Arai et al. [183] investigated ML algorithms' efficacy in predicting cost overrun risks in construction projects, utilizing a case study on NYC school construction. They compared Decision Trees, Artificial Neural Networks, XGBoost, and Linear and Ridge Regressions, finding XGBoost superior for its predictive accuracy and model advantages. Aung et al. [184] explored ML algorithms—linear regression, support vector machines, and artificial neural networks—for predicting construction project cost overruns, comparing their effectiveness

against traditional estimation methods. The study demonstrated these algorithms' superior predictive accuracy and discusses their potential to enhance project management practices.

ML applications related to legislation compliance, dispute risk outcomes, and defect identification are also prevalent in the construction industry. Raliile et al. [185] investigated the applications of unsupervised ML algorithms for monitoring health and safety legislation and compliance on construction sites through a systematic literature review spanning from 2005 to 2020. The study highlighted the potential of ML to improve compliance and proposes future tool development for contractors. Anysz et al. [186] employed ML tools, specifically decision trees (DT) and artificial neural networks (ANN), to predict the outcome of disputes between general contractors and clients in construction projects. The analysis was based on a dataset from an undisclosed contractor, with a focus on accuracy and risk assessment. Fan [187] employed a combined ML method consisting of association rule mining (ARM) and a Bayesian network (BN) to identify relationships between defects and their occurrence probabilities. The Swiss cheese model (SCM) was used to analyze high-risk defects and establish risk factors and hierarchical relationships.

In summary, the literature on cybersecurity within the construction industry is limited, and the risk assessment methods currently applied are typically manual, inflexible, and fragmented at the assessment level. ML techniques are advantageous due to their ability to automate risk assessment processes and adapt flexibly to new data. They have been extensively used to assess various risks in the construction industry, which shed lights on their potential for assessing cyber risks. Therefore, this study proposes using ML as the main technique to assess cyber risks at the project level.

6.1.2 Objectives

The goal of this study is to develop an ML-centric approach for dynamic cyber risk assessment of construction projects. To support this goal, there are four objectives that need to be achieved.

(1) To develop ML models capable of processing project characteristics and predicting the risk degrees of cyber risks in the project, where risk degrees range from 0 to 1 in this study, representing the probability of risk occurrence. For each of the five risks, three categories of ML models will be explored: linear regression, non-linear regression, and neural networks, making a total of 13 model candidates. The exploration aims to select the optimal model that can appropriately capture the non-linearity in the dataset.

(2) To adapt ML feature analysis methods to conduct risk factor analysis, so as to identify risk factors of general importance to diverse projects and those that actually contribute to the risks of a given project.

(3) To propose a greedy optimization algorithm for formulating risk reduction strategies for a specific project. The algorithm will determine which and how many risk factors should be addressed to lower the predicted risk degree to a level that satisfies the project's risk tolerance.

(4) To demonstrate the applicability of the developed approach by conducting a case study of a real construction project. The project information will be collected and organized in a manner suitable for the inputs of the developed ML models; the risk will be predicted, high-contributing factors will be identified, and a risk reduction strategy will be suggested.

6.1.3 Contributions

The academic contribution of this study is twofold: (1) Quantitative Cyber Risk Assessment Development: This study fills a significant gap in existing literature by developing an approach that enables more quantitative cyber risk assessments at the project level. This approach represents a novel contribution to the domain, offering a structured framework to assess cyber risks quantitatively. (2) Interdisciplinary Research Pioneering: At the core of this study are ML techniques, which facilitate pioneering research at the intersection of cybersecurity, ML, and construction management. This interdisciplinary

effort broadens the scope of application for ML techniques, marking a significant stride in integrated research fields.

The practical contributions of this study are twofold: (1) Dynamic Risk Assessment Tool: The developed approach, including the trained models and adapted methods, provides a ready-to-use tool for dynamically assessing the cyber risk status throughout the progression of construction projects. This enables project managers to make informed decisions about whether and to what extent to take actions to address high-contribution risk factors, thus achieving preemptive risk reduction and prevention. (2) Identification of General Risk Factors: This study identifies risk factors of general importance to a broad range of construction projects. This assists project managers in understanding which risk factors should generally be prioritized and addressed, especially in scenarios where project-specific risk assessments are not feasible.

6.2 Methodology

Figure 6.1 is the flowchart for developing the ML-centric approach, involving primarily five steps. Step 1 explains the feature sources for the ML models, which are derived from the risk factors identified in Chapter 5. Step 2 involves data generation based on Monte Carlo simulation, alongside an ensemble labeling method that incorporates two techniques to ensure the labeling is as comprehensive and objective as possible. Step 3 introduces a two-phase model development strategy, which includes selecting the best model candidate for each risk and determining the optimal weight combination of different labeling methods. Step 4 adapts the ML feature analysis method for risk factor analysis to identify risk factors of general importance to diverse construction projects and pinpoint those significantly contributing to elevated risks in a specific project. Step 5 proposes a greedy optimization algorithm to assist in formulating risk reduction strategies more efficiently. The developed approach culminates in a dynamic cyber risk assessment tool, which is composed of three main components. These components are: (1) risk degree prediction by the trained ML models, resulting from the processes outlined in Steps 1 to 3; (2) risk factor analysis,

resulting from Step 4; and (3) risk reduction strategy formulation, resulting from Step 5.

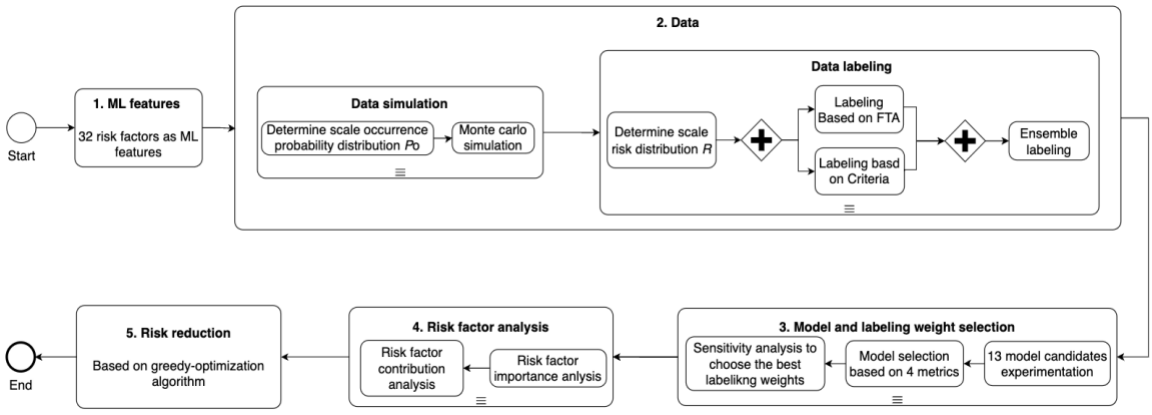


Figure 6.1. The development flowchart of the cyber risk assessment approach

6.2.1 Risk factors as ML features

The features of the ML models are derived from the risk factors associated with construction projects. We used the risk factors identified in Chapter 5 to assess the five cyber risks, which are shown in Table 5.4. Key characteristics of these factors are summarized as follows:

- The 32 risk factors encompass five facets of construction projects: (1) General project information; (2) Project structure; (3) IT factors; (4) OT factors; and (5) Management and human factors. They address both general vulnerabilities and those unique to the construction industry, ensuring a balanced risk assessment model.
- The projects are approached as a multi-layered network, with each layer consisting of diverse sub-teams and communication channels. From these, the authors identify Category 2 risk factors. This layer-specific approach enables the model to capture the dynamic nature of contemporary construction projects.
- Risk factors in Categories 1 and 2 provide an overview of the project, setting a context for the risk assessment model. Those in Categories 3, 4, and 5 are derived

from specific project phases and use data directly from the companies involved in those phases. This combination of broad and detailed data enables the model to make more contextually precise predictions.

- The risk factors allow for a more quantitative risk assessment by categorizing each into distinct scales and requiring data collection to determine each factor's risk status, thus enabling more detailed and objective decision-making in risk management.

6.2.2 Data simulation and labeling

The risk assessment is project-based, necessitating the training of an ML model with data from multiple projects. In the initial stages of this research, we faced a shortage of existing datasets. Inspired by the works [188], [189], [190], [191], [192] that utilize synthetic data, our study also employed synthetic project samples for ML model training, which were generated based on the Monte Carlo simulation [193]. Subsequently, an ensemble labeling approach was adopted, which involved the use of various methods to assign risk degrees to the synthetic project samples. Different combinations of weights for these labeling methods were explored in Section 6.2.3.2 for model training and development. The processes of data generation and labeling were rigorous, and the resulting labeled dataset underwent an evaluation process to ensure its alignment with the realities of the construction industry.

6.2.2.1 Data simulation

We opted for synthetic data generated through Monte Carlo simulations [193] due to their ability to model the uncertainty and variability inherent in real-world scenarios. This approach allows for extensive scenario analysis, robustness testing under diverse conditions, and the creation of a comprehensive dataset that supports the initial training of our models effectively.

The Monte Carlo simulation process is shown in Equation (6.1). To generate a project case *i*, probabilistic sampling is applied to all risk factors. Each risk factor has a unique probability distribution for its respective scales, denoted as $\boldsymbol{Po}_j$, indicating the likelihood of each scale's occurrence. We sampled from this probability distribution to select one scale for each risk factor, and the chosen scale was denoted by its index within that risk factor's range $K^{(j)}$.

$$S^{I}_{i=1}(j) = \sum_{k=1}^{K^{(j)}} (k-1) \times \text{Sample}(\boldsymbol{Po}_j) \quad \text{for all } j = 1, \ldots, J \tag{6.1}$$

Where,

$S^{I}_{i=1}(j)$ — The array of simulated scales for each risk factor *j* across all *I* project cases

$K^{(j)}$ — The number of scales of the *j*-th risk factor

$(k-1)$ — The index of the *k*-th scale, indexing from 0

$\boldsymbol{Po}_j$ — The probability distribution of occurrence across the scales of the *j*-th risk factor

$\text{Sample}(\boldsymbol{Po}_j)$ — This sampling term would return a 1 for the chosen scale and 0 for others, essentially picking *k*-th scale for the *j*-th risk factor based on $\boldsymbol{Po}_j$

Determining $Po = \{\boldsymbol{Po}_1, \boldsymbol{Po}_2, \ldots \boldsymbol{Po}_j \ldots, \boldsymbol{Po}_J\}$ is essential for the Monte Carlo simulation. For a risk factor *j*, two approaches are adopted: (1) Data sources. The resource we relied on is a large, published text database [113] that includes six textual sources related to cybersecurity in construction, which was the deliverable of Chapter 3. This approach resulted in the determination of 63% of the risk factors. (2) For the remaining risk factors for which the required data is not available, we determined $\boldsymbol{Po}_j$ based on our expertise. Finally, the distributions from both approaches were reviewed by experts with over 10 years of experience from a construction company based in the U.A.E. Their feedback aided in refining these distributions, ensuring they mirror real-world scenarios and meet industry expectations for reliable simulation results. The finalized Po is shown in the Appendix.

Because Risk Factors 2.2 and 2.3 pertaining to the layers of a construction project, which

include eight predefined layers, this adds seven additional factors for each of these two risk factors. Therefore, the total count of risk factors should be 32 plus 14 additional layer-specific ones, equaling 46 factors in total (259 scales in total). Following Equation (6.1), we generated 1,000 simulated project cases through Python, resulting in an array of dataset shown in Table 6.1. The array has 1,000 rows, with each row corresponding to an individual project case. The columns of this array represent the indices of the selected scales for each risk factor, derived from the probabilistic sampling.

Table 6.1. Data structure

No.	Risk Factor													
	1.1	1.2	1.3	1.4	1.5	1.6	1.7	...	5.2	5.3	5.4	5.5	5.6	5.7
1	2	2	3	2	2	3	1	...	2	0	0	1	0	2
2	3	3	3	3	3	4	1	...	3	1	1	1	1	3
3	2	2	3	2	2	3	1	...	2	0	1	1	1	2
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
998	5	4	5	4	4	4	2	...	4	1	1	3	2	4
999	0	0	1	1	1	1	0	...	1	0	0	0	0	1
1000	3	2	3	3	2	4	1	...	2	0	1	1	1	2

6.2.2.2 Data labeling

For model training, labels indicating the risk degrees for the five cyber risks are needed for each generated project case. This study defines the risk degree within the range [0,1], interpreted as the probability of occurrence. This definition is based on the concept of the probability of the top event (the risk) occurring in the Fault Tree Analysis method [194] discussed later. The labeling process can be illustrated by Equation (6.2).

$$p_i = f^l\left(r_{1,S_{i,1}}, r_{2,S_{i,2}}, \ldots, r_{j,S_{i,j}}, \ldots, r_{J,S_{i,J}}\right) \tag{6.2}$$

Where,

p_i — The labeled risk degree of a cyber risk for the *i*-th project, within [0,1]

f^l — The labeling process

$S_{i,j}$ — The selected scale of the *j*-th risk factor of the *i*-th project

$r_{j,S_{i,j}}$ — The risk degree of the *j*-th risk factor, which equals to the risk degree of the selected scale, representing the probability of the worst-case scenario occurring for this risk factor under this scale.

(1) Determine the Scale Risk Distribution (R)

For a risk factor, the risk degrees of its associated scales make up the risk distribution for that risk factor, denoted as $\boldsymbol{R}_j=\left(r_{j,1},\ r_{j,2},\ldots,r_{j,k},\ldots,r_{j,K^{(j)}}\right)$. The determination of this distribution can be approached in two ways, depending on whether the scales are ordinal or categorical.

(a) The risk distribution of ordinal scales. For certain risk factors, such as risk factor 1.4, where the scales are ordinal (a total of 38 risk factors), a specific scale can be chosen to represent the risk factor's worst scenario, with its risk degree set to 1. In such risk factors, the riskiest scale is typically either the first or the last based on the understanding of this risk factor. The risk degrees for the other scales are then considered to have a linear relationship with the worst scenario's risk degree, shown as Equations (6.3) and (6.4).

If the first scale represents the worst scenario:

$$r_{j,k}=\frac{K^{(j)}-(k-1)}{K^{(j)}} \tag{6.3}$$

If the last scale represents the worst scenario:

$$r_{j,k}=\frac{k}{K^{(j)}} \tag{6.4}$$

(b) The risk distribution of categorical scales. For certain risk factors, such as risk factor 1.7, where the scales are categorical (8 risk factors), it is necessary to determine the risk degree for each scale. To minimize uncertainty, fuzzy set theory [195], a branch of fuzzy mathematics, can be used. This method is designed to handle imprecise and ambiguous input data within complex systems, which has gained widespread acceptance in risk

analysis across various industries such as nuclear power, chemical, and oil and gas industries. When assigning risk likelihoods, a seven-level natural language set can be used, represented as {Very Low (VL), Low (L), Moderately Low (ML), Medium (M), Moderately High (MH), High (H), Very High (VH)} [196]. These assigned terms are then mapped to a set of fuzzy numbers, which are subsequently defuzzified into a point value to obtain the final risk probability. The two most common fuzzy number representations are triangular $\tilde{p}_A=\left(p_{a_1}, p_{a_2}, p_{a_3}\right)$ and trapezoidal $\tilde{p}_B=\left(p_{b_1}, p_{b_2}, p_{b_3}, p_{b_4}\right)$. The relationship between these terms and fuzzy numbers, derived from their membership functions [196], can be shown as Figure 6.2.

In this study, for the *k*-th scale of risk factor *j*, if the risk degree is assigned as "*L*", the probability fuzzy number for this scale should be $\tilde{r}_{j,k}=(0.1, 0.2, 0.3)$; if the risk degree is assigned as "*MH*", the probability fuzzy number should be $\tilde{r}_{j,k}=(0.5, 0.6, 0.7, 0.8)$. If the risk degree is assigned as "*VL*" or "*VH*", the fuzzy numbers are part of the trapezoidal probability fuzzy number, which should be reconstructed as $\tilde{r}_{j,k}=(0, 0, 0.1, 0.2)$ and $\tilde{r}_{j,k}=(0.8, 0.9, 1, 1)$, respectively.

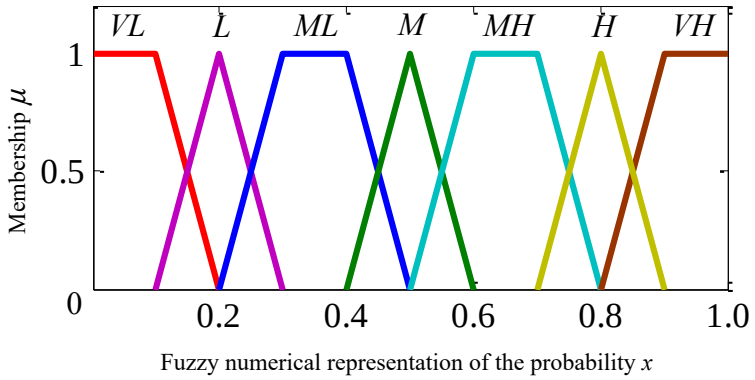


Figure 6.2. Fuzzy numerical representation of natural language terms [196]

For defuzzification, the Center of Area (CoA) method [195], known for its straightforwardness and practical application, is commonly used. The defuzzification

formula for the triangular and trapezoidal fuzzy number is given in Equations (6.5) and (6.6) respectively [197].

$$p_A^* = \frac{1}{3} \cdot \left(p_{a_1} + p_{a_2} + p_{a_3} \right) \tag{6.5}$$

$$p_B^* = \frac{1}{3} \cdot \frac{\left(p_{b_4} + p_{b_3} \right)^2 - p_{b_4} p_{b_3} - \left(p_{b_1} + p_{b_2} \right)^2 + p_{b_1} p_{b_2}}{p_{b_4} + p_{b_3} - p_{b_2} - p_{b_1}} \tag{6.6}$$

By implementing the methods above, we preliminarily determined the risk degrees for all 259 scales of the 46 factors. These were presented to the same experts from the construction company and underwent minor modifications based on their professional feedback. The finalized scale risk distribution R is displayed in Appendix, based on which the generated scales in Table 6.1 are converted into corresponding risk degrees, shown in Table 6.2.

Table 6.2. Data structure indicating risk degrees

No.	Risk Factor													
	1.1	1.2	1.3	1.4	1.5	1.6	1.7	...	5.2	5.3	5.4	5.5	5.6	5.7
1	0.90	0.6	0.50	0.6	0.50	0.8	1.0	...	0.67	1.0	0.2	0.33	0.2	0.6
2	0.65	0.8	0.50	0.8	0.67	1.0	1.0	...	0.50	0.2	1.0	0.33	0.4	0.4
3	0.90	0.6	0.50	0.6	0.50	0.8	1.0	...	0.67	1.0	1.0	0.33	0.4	0.6
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
998	0.60	1.0	0.17	1.0	0.83	1.0	0.7	...	0.33	0.2	1.0	0.67	0.6	0.2
999	0.80	0.2	0.83	0.4	0.33	0.4	0.4	...	0.83	1.0	0.2	0.17	0.2	0.8
1000	0.65	0.6	0.50	0.8	0.50	1.0	1.0	...	0.67	1.0	1.0	0.33	0.4	0.6

(2) Ensemble labeling methods

To mitigate the potential variance and bias inherent in a single labeling method f^l, an ensemble approach is utilized, integrating multiple methods. This approach combines the assessments of a particular cyber risk for each project from various methods, leveraging their individual strengths. Consequently, Equation (6.2) is revised to Equation (6.7) to reflect this integrated methodology.

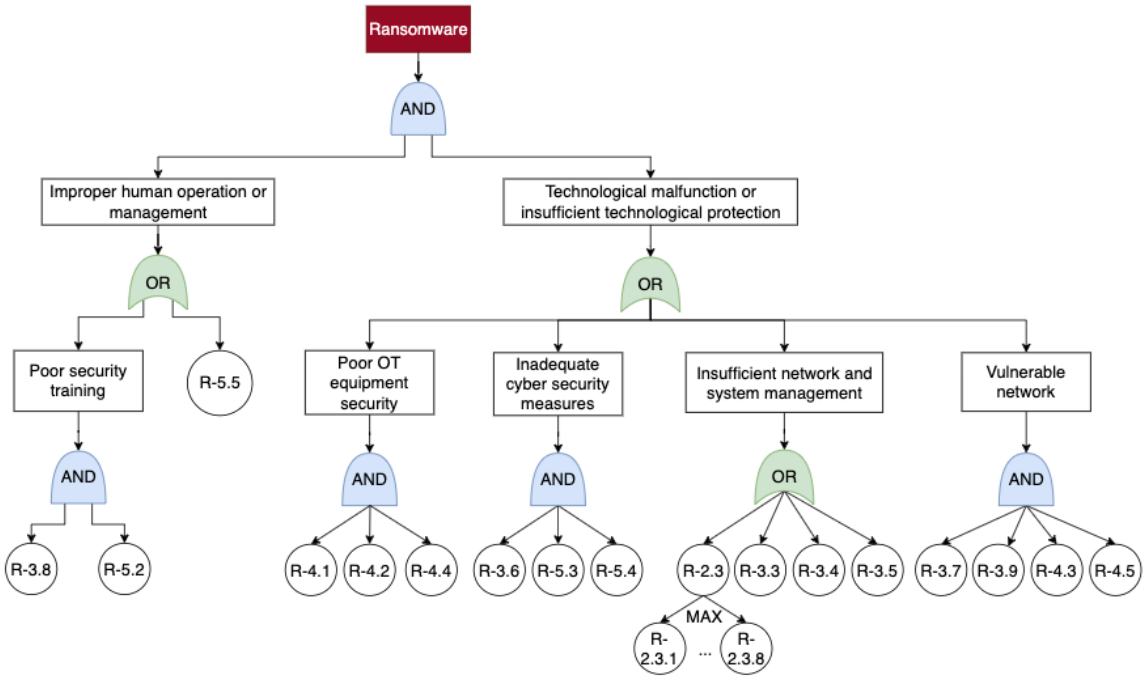
$$\bar{p}_i = \sum_{e=1}^{E} w_e^l \cdot f_e^l\left(r_{1,S_{i,1}}, r_{2,S_{i,2}}, \ldots, r_{j,S_{i,j}}, \ldots, r_{J,S_{i,J}} \right) \tag{6.7}$$

Where $\bar{p}_i$ is the ensembled risk degree of the *i*-th project; w_e^l is the weight of the *e*-th labeling method, $\sum_{e=1}^{E} w_e^l = 1$; f_e^l is the *e*-th labeling method.

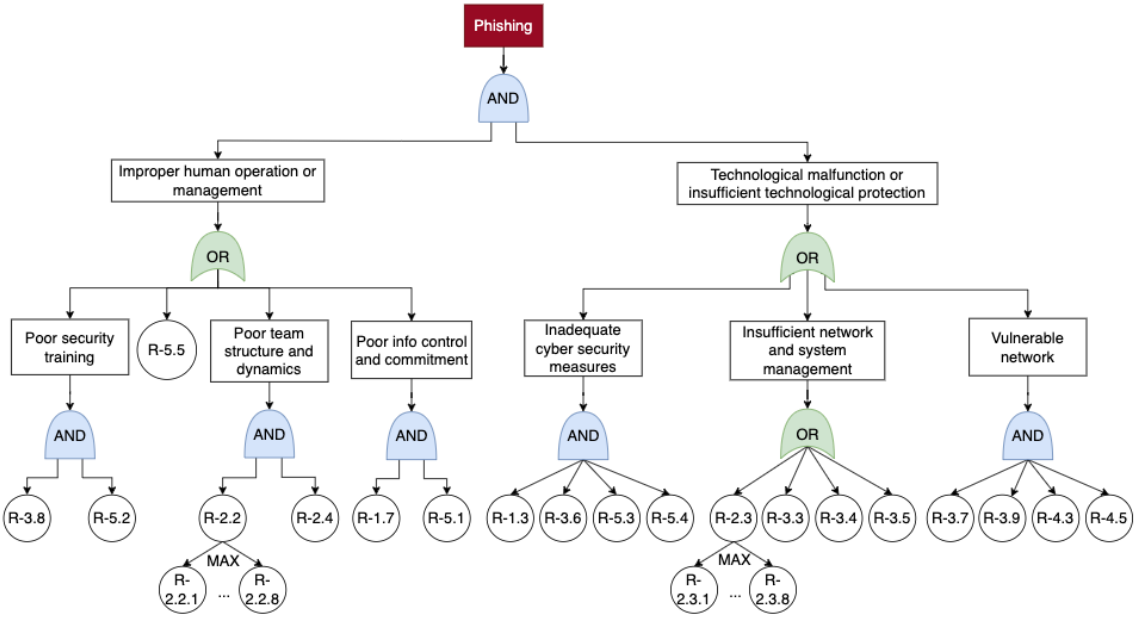
In this study, we adopted an ensemble approach that combines Fault Tree Analysis (FTA) [194] and the intuitive criteria-based method. FTA's systematic and detailed framework effectively models complex interdependencies among risk factors and aids in causal analysis that connects these factors to cyber risks. The criteria-based method offers simplicity, presenting an easily comprehensible way to associate risk factors with cyber risks.

(a) FTA labeling. FTA [194] is a versatile analytical method used in risk assessment across various industries, such as the oil and gas industry [196], chemical industry [197], and construction industry [160], to investigate the causes and probabilities of an undesired event. Following the methods in [194], a fault tree for each risk is established, shown in Figure 6.3 and verified by the same experts from the U.A.E. mentioned above. The undesired event (cyber risks in this study, enclosed in a rectangle) is positioned at the top of the tree and is referred to as the 'top event'. This event is logically decomposed into basic events at the bottom of the tree (risk factors in this study, enclosed in circles) using logic operators known as "gates". Not all risk factors are included in a fault tree; some may be omitted if they are not directly observed to be related.

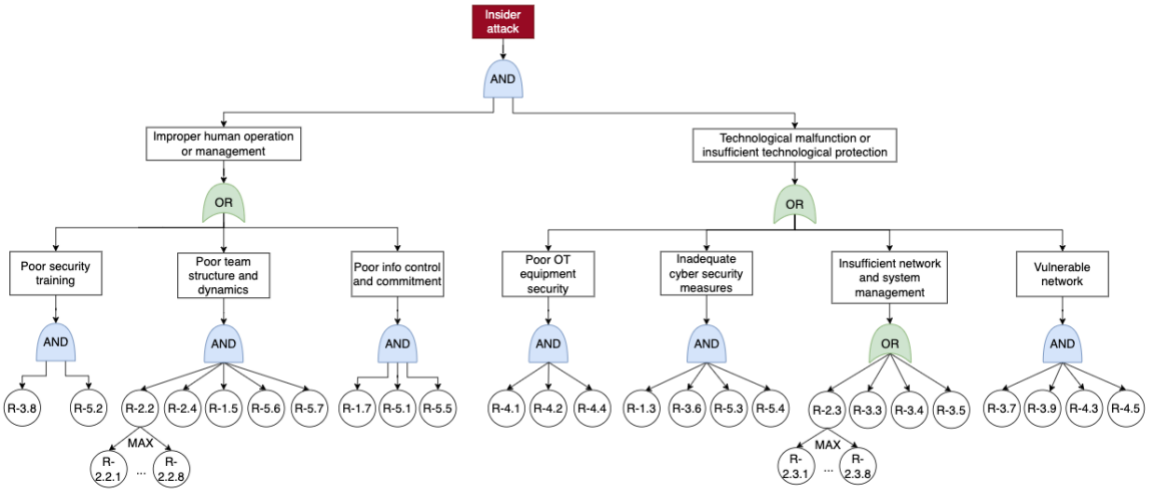
The probability of the top event can be derived from the basic events using Boolean logic operators on gates, which means the risk degrees associated with these factors can be propagated to determine the overall risk degree of cyber risk. Events enclosed in rectangles between the top events and the basic events are intermediate events. Events from which arrows originate are termed "parent events", while those the arrows point to are called "child events".



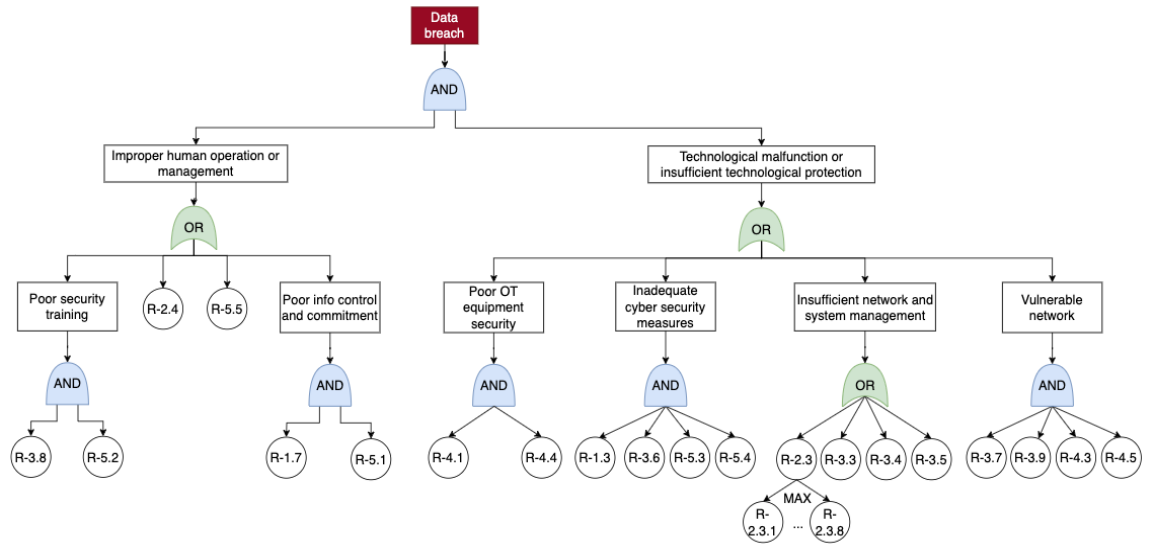
(a). Fault tree of ransomware



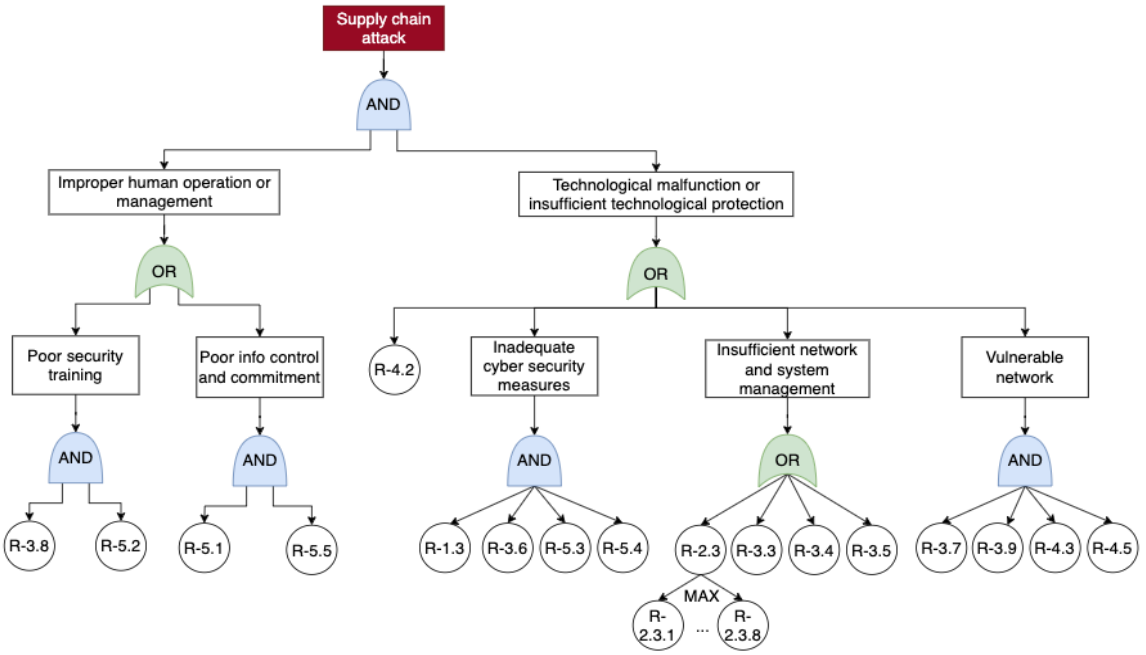
(b). Fault tree of phishing



(c). Fault tree of insider attack



(d). Fault tree of data breach



(e). Fault tree of supply chain attack

Figure 6.3. Fault trees developed for the five cyber risks

Where OR represents the OR gate; AND represents the AND gate; ▭ represents the intermediate events; ◯ represents the basic events.

For AND gate (all child events must occur for the parent event to occur) and for OR gate (any child event can trigger the parent event to occur), the probability of the parent event is calculated as Equations (6.8) and (6.9) respectively.

$$P(Parent) = P(A) \times P(B) \times \ldots \times P(N) \tag{6.8}$$

$$P(Parent) = 1 - (1 - P(A)) \times (1 - P(B)) \times \ldots \times (1 - P(N)) \tag{6.9}$$

Where $P(Parent)$ is the probability of the parent event; $P(A), P(B), \ldots, P(N)$ are probabilities of the child events.

(b) Criteria-based labeling. For the factors not covered in fault trees, the criteria-based

method can be adopted to derive the risk degree of cyber risks. This method is intuitive, defining a threshold T that a risk factor must satisfy to be considered significant in influencing the overall risk degree of the project, as shown in Equation (6.10).

$$p_{C_i} = f_C\left(r_{1,S_{i,1}}, r_{2,S_{i,2}}, \ldots, r_{j,S_{i,j}}, \ldots, r_{J,S_{i,J}}\right) = \frac{1}{J_C}\sum_{j=1}^{J_C} I(RF_j \in RF_C \cap r_{i,S_{i,j}} \geq T) \quad (6.10)$$

Where J_C is the number of risk factors not in the fault tree; RF_j is the j-th risk factor not in the fault tree; I is a function that returns 1 if the condition inside the parentheses is true and 0 otherwise.

For each of the 1,000 generated project case, both labeling methods are applied, resulting in two risk degrees (p_T and p_C) for each risk, the ensembled risk degree of this risk is shown as Equations (6.11) and (6.12). The weights of the labeling methods will be explored to train the ML models in Section 6.2.3.2.

$$\bar{p}_i = w_T \cdot p_{T_i} + w_C \cdot p_{C_i} \quad (6.11)$$

$$w_T + w_C = 1 \quad (6.12)$$

6.2.3 Model development for risk prediction

In this section, various model candidates are introduced and adapted for training on the dataset. The determination of the final model for each risk involves a two-phase selection strategy, which first identifies the best model candidate and then determines the most effective labeling weight combination.

6.2.3.1 Model candidate exploration

The ensemble labeling method inevitably introduces non-linearity to the relationship between the risk factors and the project risk degrees. To train a model that can capture this non-linearity neither too little nor too excessively, we explore different models. This dataset

contains both ordinal and categorical features. Given the inconsistency in the intervals between the rankings of the ordinal features and in the interest of maintaining consistent feature processing, we will treat these ordinal features as categorical. Consequently, both types of features will be represented using one-hot encoding. We used $Pr_{j,k}$ to denote the value of the *k*-th scale in the one-hot representation of the *j*-th risk factor, which is either 0 or 1, indicating whether this scale is selected or not. If the *k*-th scale is selected, $Pr_{j,k}=1$, and $\boldsymbol{Pr}_j=\boldsymbol{Pr}_j^{e_{j,k}}=\left(0,\ldots,1_{(k)},\ldots,0\right)_j$.

(1) Linear regression. We started from the linear regression models with the lower ability to capture non-linearity: basic linear regression [198], linear regression with an L1 penalty (Lasso regression) [199], and linear regression with an L2 penalty (Ridge regression) [200]. Both Lasso and Ridge regression incorporate regularization techniques to counteract overfitting. A basic linear regression model in context is shown in Equation (6.13).

$$\widehat{p}=\beta_0+\sum_{j=1}^{46}\sum_{k=1}^{K^{(j)}}\beta_{j,k}\cdot Pr_{j,k} \tag{6.13}$$

Where,
$\widehat{p}$ — The predicted risk degree of a certain risk
β_0 — The intercept of the linear regression model
$\beta_{j,k}$ — The coefficient of the *k*-th scale of the *j*-th risk factor, totally 259
$Pr_{j,k}$ — The value of the *k*-th scale in the one-hot representation of the *j*-th risk factor $\boldsymbol{Pr}_j$.

(2) Non-linear regression. Then, polynomial regression [198] was utilized to accommodate higher polynomial degrees of input features, enhancing the model's ability to detect non-linear relationships and capture the interactions between features. The polynomial regression of two features to the dependent variable *y* can be depicted as Equation (6.14).

$$y=\beta_0+\sum_{i=1}^{d}\beta_{1i}x_1^i+\sum_{i=1}^{d}\beta_{2i}x_2^i+\sum_{i=1}^{d-1}\sum_{j=1}^{d-i}\beta_{ij}x_1^i x_2^j \tag{6.14}$$

Where,

y — The dependent variable

x_1, x_2 — The features for exemplification

β_0 — The intercept

β_{1i} — The coefficients for the terms involving x_1 to the *i*-th power

β_{2i} — The coefficients for the terms involving x_2 to the *i*-th power

β_{ij} — The coefficients for the interaction terms, where x_1 is raised to the *i*-th power and x_2 is raised to the *j*-th power, with the constraint that $i+j$ does not exceed d

Given that the features in our study $Pr_{j,k}$ can only be 0 or 1 for one-hot encoding representation, all polynomial terms of a feature collapse to the original feature, so that Equation (6.14) can be simplified as Equation (6.15).

$$y=\beta_0+\sum_{i=1}^{d}\beta_{1i}x_1+\sum_{i=1}^{d}\beta_{2i}x_2+\sum_{i=1}^{d-1}\sum_{j=1}^{d-i}\beta_{ij}x_1x_2 \tag{6.15}$$

(3) Neural networks. Similarly, Neural Networks [201] can also capture the nonlinearity. They have the advantage of learning feature transformations automatically during training, eliminating the need for pre-processing required by polynomial regression. A feedforward NN, composed of neurons in multiple layers, utilizes hidden layers to perform this automatic transformation. The operation of a hidden layer with *m* neurons, following a layer with *n* neurons, is described by Equation (6.16).

$$h_j=a\left(\sum_{i=1}^{n}w_{ij}x_i+b_j\right) \tag{6.16}$$

Where,

h_j — Value of the *j*-th neuron in the hidden layer

x_i — Value of the *i*-th neuron in the previous layer

w_{ij} — Weight of the *i*-th neuron in the previous layer that connects to the *j*-th neuron in the hidden layer

b_j — Bias term for the *j*-th neuron in the hidden layer

a — Activation function. We experimented with three common activation functions: ReLU [58], LeakyReLU [59], and Tanh [60]

The model, with an input layer of 259 neurons, can be designed to predict the risk degree for single or multiple cyber risks. For the latter, its multi-task learning framework [202] offers two main benefits: it fosters information exchange across the five cyber risk tasks to provide complementary insights, potentially improving generalization and performance, and it streamlines training by requiring only one model for all tasks, enhancing efficiency. We tested nine NN architectures, aiming for a comprehensive experimentation, shown in Table 6.3.

Table 6.3. Model structure and training configuration of neural networks

Model Name	# of Hidden Layers	# of Neurons in Hidden Layers	Activation Function	# of Output Neurons
NN_1	1	100	ReLU	1
NN_2	2	150, 100	ReLU	1
NN_3	1	100	LeakyReLU	1
NN_4	3	200, 150, 100	ReLU	1
NN_5	1	100	Tanh	1
NN_6	2	200, 100	Tanh, ReLU	1
Combined_NN_1	1	100	ReLU	5
Combined_NN_2	2	150, 75	ReLU	5
Combined_NN_2	3	200, 100, 50	LeakyReLU	5

6.2.3.2 Training and selection

To develop the optimal model for each risk, the model candidate and labeling weight combination must be determined. The straightforward approach would involve training models for each labeling weight combination. This results in 13 model candidates across 11 labeling weights, with weights ranging from 0 to 1 in 0.1 intervals, totaling 143 training sessions. Although comprehensive, this method is time-intensive and inefficient. To

streamline the process, we adopted a two-phase selection inspired by the greedy algorithm's principles [203], which prioritize selecting the best option at each step for local optimization. This approach reduces the number of training iterations required while still aiming to identify the most effective model and weight combination.

(1) The model development strategy

Phase 1: For each risk, each of the 13 model candidates was trained with a weight of 0.5 for both labeling methods, treating them as equally important. The 1,000 generated samples are split into training, validation, and testing sets with a ratio of 7:2:1. Four metrics were adopted for performance evaluation: MSE, RMSE, MAE, and the Coefficient of Determination (R^2). The best model candidate is selected based on its performance on the test set. The four metrics were used because they can provide different interpretation perspectives: MSE highlights the average squared errors, emphasizing the penalty on larger errors; RMSE brings error metrics to the same scale as the target variable, making interpretations more intuitive; MAE offers a direct average of absolute errors, less sensitive to outliers and more representative of typical error magnitude; R^2 quantifies the proportion of variance explained by the model, indicating fit quality. This multifaceted approach can not only provide a comprehensive model evaluation but also can deliver different insights to stakeholders to understand model performance.

Phase 2: A sensitivity analysis of the labeling weights was conducted using the selected optimal model candidate. The best weights were chosen based on the model's performance on the two additional simulated projects, validated and labeled by three experts. This reduces the number of training iterations for each risk to 24 (13 for the initial selection plus 11 for the sensitivity analysis). It should be noted that the training configurations for the same model candidate remain consistent across the two phases.

Among the experts invited, two experts are U.S.-based cybersecurity specialists, and the other is a specialist in the Chinese construction domain. Table 6.4 illustrates the data structure of the two additional simulated projects used for sensitivity analysis, while Table

6.5 displays the labeling results provided by the experts. The labeling for each risk for each project is taken from the mode of votes from the three experts.

Table 6.4. The two simulated projects

Project	Risk Factors																						
	1.1	1.2	1.3	1.4	1.5	1.6	1.7	2.1	2.2.1	2.2.2	2.2.3	2.2.4	2.2.5	2.2.6	2.2.7	2.2.8	2.3.1	2.3.2	2.3.3	2.3.4	2.3.5	2.3.6	2.3.7
S_1	2	2	3	2	2	3	1	4	2	1	3	0	0	0	0	5	3	3	5	0	1	1	1
S_2	0	0	0	0	0	0	0	0	0	1	0	1	0	5	0	0	0	1	0	3	1	7	0
Project	Risk Factors																						
	2.3.8	2.4	3.1	3.2	3.3	3.4	3.5	3.6	3.7	3.8	3.9	4.1	4.2	4.3	4.4	4.5	5.1	5.2	5.3	5.4	5.5	5.6	5.7
S_1	6	2	1	2	1	2	2	3	1	1	2	2	2	3	1	2	3	2	0	0	1	1	2
S_2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table 6.5. Labeling of the two simulated projects

Project	Ransomware	Phishing	Insider Attack	Data Breach	Supply Chain Attack
S_1	Yes	Yes	Yes	Yes	No
S_2	No	Yes	No	No	No

(2) Training details

(a) For linear regression, the commonly used loss function is the MSE, which is presented in Equation (6.17). The loss functions for Lasso and Ridge regressions are represented in Equations (6.18) and (6.19), respectively.

$$L_{MSE} = \frac{1}{N}\sum_{i=1}^{N}(\hat{p}_i - p_i)^2 \tag{6.17}$$

$$L_{Lasso} = \frac{1}{N}\sum_{i=1}^{N}(\hat{p}_i - p_i)^2 + \lambda\sum_{j=1}^{46}\sum_{k=1}^{K^{(j)}}|\beta_{j,k}| \tag{6.18}$$

$$L_{Ridge} = \frac{1}{N}\sum_{i=1}^{N}(\hat{p}_i - p_i)^2 + \lambda\sum_{j=1}^{46}\sum_{k=1}^{K^{(j)}}\beta^2{}_{j,k} \quad (6.19)$$

Where N is the number of project cases for training. λ is the regularization strength. The training and testing sets were used for the three candidates of linear regression models. For both Lasso and Ridge regressions, we employed a grid search strategy coupled with 5-fold cross-validation to fine-tune the hyperparameter λ.

(b) For non-linear regression, to train the model, we used the loss function from Equation (6.17), applying 5-fold validation on the same dataset to find the optimal polynomial degree. We limited polynomial degrees to 2 and 3 to avoid overfitting, considering the dataset's complexity does not warrant higher degrees.

(c) For neural networks, two loss functions were used respectively for the single output and multioutput, respectively using the loss functions in Equations (6.17) and (6.20). All models used the same training setup, utilizing the Adam optimizer with a learning rate of 0.004, 200 epochs for training, and a batch size of 16. The best-performing model on the validation set was saved.

$$L_{multi-output} = \frac{1}{5N}\sum_{i=1}^{N}\sum_{r=1}^{5}(\hat{p}_{i,r} - p_{i,r})^2 \quad (6.20)$$

Where,
$\hat{p}_{i,r}$ — Predicted risk degree for the the r-th risk of project case i
$p_{i,r}$ — Actual risk degree for the r-th risk of project case i

(3) Model and labeling weights selection results

By implementing the Phase-1 strategy, Table 6.6 presents the performance results of the total 13 trained models on the test set across five cyber risks. The optimal model candidate for each risk, identified by its consistent superior performance on these metrics, is detailed in Table 6.7.

Table 6.6. Performance of the 13 models using the test set

Risk	Metrics	NN_1	NN_2	NN_3	NN_4	NN_5	NN_6	Linear	Ridge	Lasso	Polynomial	Combined_NN_1	Combined_NN_2	Combined_NN_3
Ransomware	MSE	2.06e-05	1.27e-05	1.48e-05	1.30e-05	1.91e-05	1.98e-05	2.02e-05	1.85e-05	4.22e-05	3.97e-05	2.01e-05	1.41e-05	1.99e-05
	RMSE	4.54e-03	3.57e-03	3.85e-03	3.60e-03	4.37e-03	4.45e-03	4.50e-03	4.31e-03	6.50e-03	6.30e-03	4.48e-03	3.75e-03	4.46e-03
	MAE	3.32e-03	2.63e-03	2.76e-03	2.30e-03	3.27e-03	3.28e-03	3.43e-03	3.22e-03	4.73e-03	4.78e-03	3.24e-03	2.66e-03	3.08e-03
	R^2	0.998	0.999	0.999	0.999	0.998	0.998	0.998	0.998	0.996	0.997	0.998	0.999	0.998
Phishing	MSE	6.12e-05	4.97e-05	7.82e-05	6.36e-05	7.58e-05	6.84e-05	1.27e-04	1.19e-04	1.23e-04	1.19e-04	6.42e-05	6.26e-05	5.99e-05
	RMSE	7.82e-03	7.05e-03	8.84e-03	7.97e-03	8.71e-03	8.27e-03	1.13e-02	1.09e-02	1.11e-02	1.09e-02	8.01e-03	7.91e-03	7.74e-03
	MAE	5.95e-03	5.24e-03	6.53e-03	5.40e-03	6.42e-03	5.87e-03	9.12e-03	8.76e-03	8.70e-03	8.58e-03	5.31e-03	5.89e-03	5.40e-03
	R^2	0.996	0.997	0.995	0.996	0.995	0.996	0.992	0.993	0.992	0.993	0.996	0.996	0.996
Insider attack	MSE	6.12e-06	2.76e-06	3.10e-06	3.36e-06	3.84e-06	2.96e-06	2.14e-06	1.99e-06	5.01e-06	4.73e-06	8.53e-06	9.44e-06	1.02e-05
	RMSE	2.47e-03	1.66e-03	1.76e-03	1.83e-03	1.96e-03	1.72e-03	1.46e-03	1.41e-03	2.24e-03	2.18e-03	2.92e-03	3.07e-03	3.20e-03
	MAE	1.83e-03	1.25e-03	1.35e-03	1.34e-03	1.54e-03	1.28e-03	1.09e-03	1.05e-03	1.83e-03	1.64e-03	2.32e-03	2.44e-03	2.43e-03
	R^2	1	1	1	1	1	1	1	1	1	1	0.999	0.999	0.999
Data breach	MSE	5.15e-05	5.62e-05	4.85e-05	5.24e-05	7.36e-05	5.15e-05	7.96e-05	7.38e-05	9.75e-05	1.07e-04	5.74e-05	4.36e-05	3.66e-05
	RMSE	7.18e-03	7.50e-03	6.96e-03	7.24e-03	8.58e-03	7.18e-03	8.92e-03	8.59e-03	9.87e-03	1.03e-02	7.58e-03	6.60e-03	6.05e-03
	MAE	4.33e-03	4.45e-03	4.44e-03	4.75e-03	6.18e-03	4.86e-03	6.89e-03	6.52e-03	7.24e-03	7.95e-03	4.46e-03	3.83e-03	4.11e-03
	R^2	0.995	0.994	0.995	0.995	0.992	0.995	0.992	0.992	0.99	0.989	0.994	0.995	0.996
Supply chain attack	MSE	1.24e-05	1.48e-05	1.15e-05	9.66e-06	7.14e-06	1.39e-05	9.92e-07	8.93e-07	1.98e-05	2.69e-05	9.59e-06	1.80e-05	1.42e-05
	RMSE	3.53e-03	3.84e-03	3.39e-03	3.11e-03	2.67e-03	3.72e-03	9.96e-04	9.45e-04	4.45e-03	5.19e-03	3.10e-03	4.24e-03	3.77e-03
	MAE	2.75e-03	2.90e-03	2.41e-03	2.24e-03	2.11e-03	2.55e-03	7.59e-04	7.09e-04	3.27e-03	3.76e-03	2.20e-03	2.61e-03	2.68e-03
	R^2	0.998	0.998	0.999	0.999	0.999	0.998	1	1	0.997	0.997	0.999	0.998	0.998

Table 6.7. Optimal base model selection results

Risk	Metric	Best Performance	Best Models	Final Model
Ransomware	MSE	1.27e-05	NN_2	NN_2
	RMSE	3.57e-03	NN_2	
	MAE	2.30e-03	NN_4	
	R^2	0.999	NN_2, NN_3, NN_4, Combined_Model_2	
Phishing	MSE	4.97e-05	NN_2	NN_2
	RMSE	7.05e-03	NN_2	
	MAE	5.24e-03	NN_2	
	R^2	0.997	NN_2	
Insider attack	MSE	1.99e-06	Ridge Regression	Ridge Regression
	RMSE	1.41e-03	Ridge Regression	
	MAE	1.05e-03	Ridge Regression	
	R^2	1	All except combined models	
Data breach	MSE	3.66e-05	Combined_NN_3	Combined_NN_3
	RMSE	6.05e-03	Combined_NN_3	
	MAE	3.83e-03	Combined_NN_2	
	R^2	0.996	Combined_NN_3	
Supply chain attack	MSE	8.93e-07	Ridge Regression	Ridge Regression
	RMSE	9.45e-04	Ridge Regression	
	MAE	7.09e-04	Ridge Regression	
	R^2	1	Linear Regression, Ridge Regression	

Implementing the Phase-2 strategy of sensitivity analysis, Table 6.8 illustrates the predicted risk degrees under different weight combinations, with a value of 0.5 or higher indicating the risk occurrence. It can be seen that most predictions aligned with actual labels, achieving over 80% accuracy, demonstrating our selected models' reliability. The weight combination of 0.6 for FTA and 0.4 for Criteria-based labeling consistently matched actual labels for all risks and both projects, leading to its selection as the final optimal weight combination for each risk in our study. We then retrained each selected base model with the optimal weight combination.

Table 6.8. Sensitivity analysis results of ensemble labeling

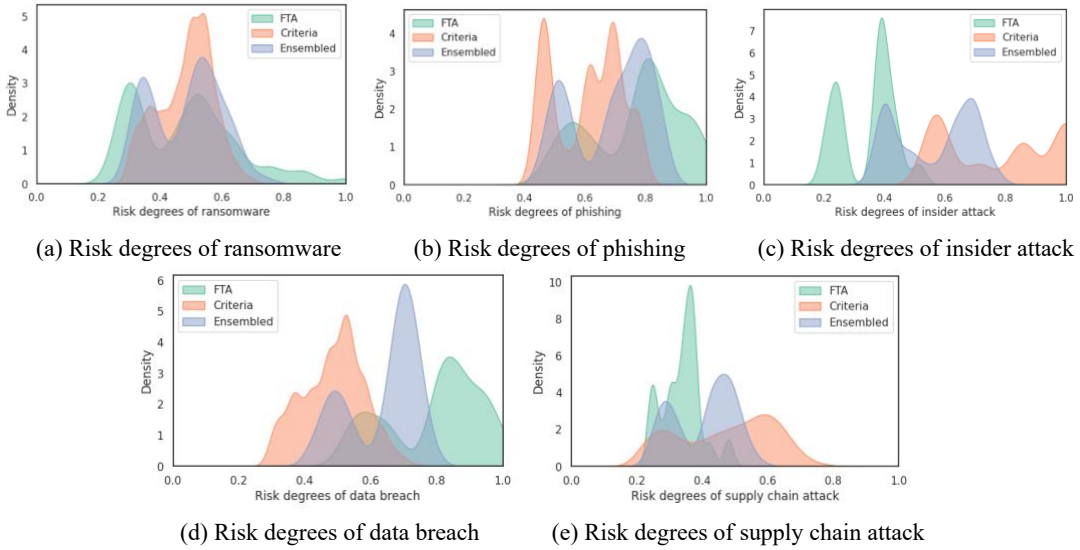
Weights		S_1					S_2				
FTA	Criteria	R_1	R_2	R_3	R_4	R_5	R_1	R_2	R_3	R_4	R_5
0.00	1.00	0.55	0.69	1	0.58	0.55	0.32	0.46	0.57	0.32	0.23
0.10	0.90	0.55	0.71	0.94	0.61	0.53	0.32	0.47	0.54	0.34	0.24
0.20	0.80	0.54	0.72	0.88	0.64	0.51	0.31	0.48	0.51	0.37	0.25
0.30	0.70	0.54	0.73	0.82	0.67	0.49	0.31	0.48	0.47	0.4	0.26
0.40	0.60	0.54	0.75	0.76	0.7	0.47	0.31	0.49	0.44	0.42	0.27
0.50	0.50	0.53	0.77	0.7	0.73	0.46	0.3	0.49	0.41	0.46	0.28
0.60	**0.40**	**0.53**	**0.78**	**0.64**	**0.76**	**0.44**	**0.3**	**0.51**	**0.38**	**0.48**	**0.29**
0.70	0.30	0.52	0.8	0.58	0.79	0.42	0.3	0.49	0.35	0.5	0.3
0.80	0.20	0.52	0.81	0.52	0.82	0.4	0.3	0.5	0.31	0.53	0.31
0.90	0.10	0.51	0.83	0.46	0.85	0.38	0.3	0.51	0.28	0.54	0.32
1.00	0.00	0.51	0.84	0.4	0.89	0.37	0.29	0.5	0.25	0.59	0.33

6.2.3.3 Analysis and Evaluations

A statistical analysis known as Kernel Density Estimation (KDE) [204] was conducted. This analysis focuses on the distributions of the integrated weighted risk degrees of the generated samples, enabling an examination of their alignment with reality. Figures 6.4 and 6.5 display the KDE plots of risk degrees for each risk category across 1,000 generated project cases, where risk degrees were derived from combining different weights.

Three observations are noteworthy: (1) The density distribution of the averaged risk degrees for all risks does not lean towards extremely high or low-risk degrees, confirming the unbiased nature and effectiveness of our ensemble labeling approach; (2) The density distribution of phishing and data breach tends slightly towards higher risk degrees, reflecting the prevalent occurrence of these cyber risks in the current construction cybersecurity landscape; (3) The density distribution for supply chain attack is skewed towards lower risk degrees, in line with their less frequent occurrence compared to phishing and data breach. These observations underscore the efficacy of

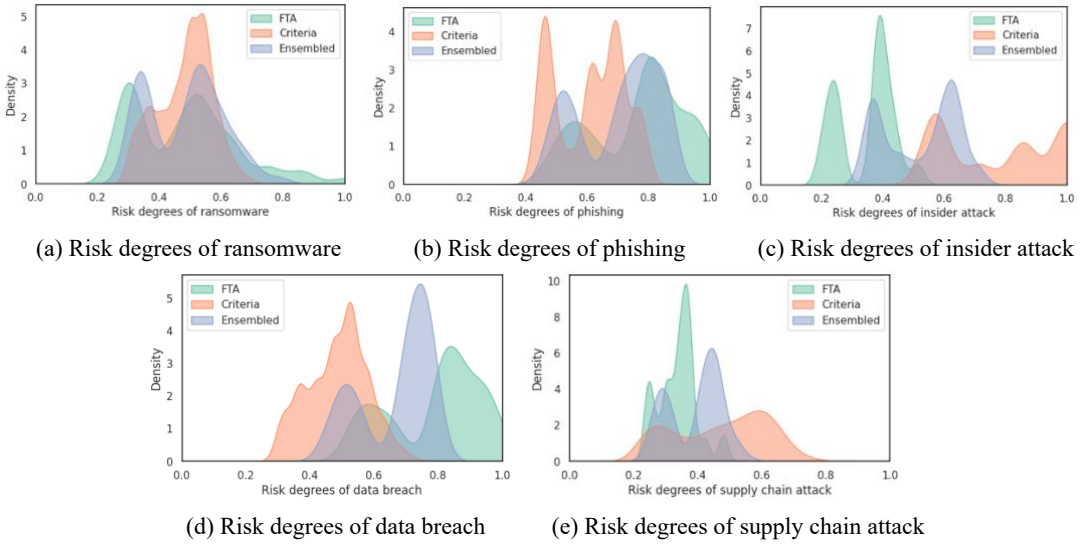
our ensemble labeling method as it appears unbiased and mirrors the industry realities.



(a) Risk degrees of ransomware (b) Risk degrees of phishing (c) Risk degrees of insider attack

(d) Risk degrees of data breach (e) Risk degrees of supply chain attack

Figure 6.4. KDE of the labeled risk degrees (FTA Weight: 0.5, Criteria Weight: 0.5)



(a) Risk degrees of ransomware (b) Risk degrees of phishing (c) Risk degrees of insider attack

(d) Risk degrees of data breach (e) Risk degrees of supply chain attack

Figure 6.5. KDE of the labeled risk degrees (FTA Weight: 0.6, Criteria Weight: 0.4)

Three insights derived from the KDE analysis and sensitivity analysis results, as displayed in Table 6.8, support the validity of our trained models: (1) The KDE analysis results for a labeling weight of (0.5, 0.5) confirm that the dataset used for Phase-1 model training aligns with reality, thereby making the candidate model selection process

reliable. (2) Similarly, the KDE analysis results for a labeling weight of (0.6, 0.4) demonstrate that the dataset derived from these labeling weights is also aligned with reality. This ensures that our final trained and selected models at Phase-2 is also reliable and has effectively encoded real-world information. (3) As shown in Table 6.8, with the final determined labeling weights of (0.6, 0.4), the model for each risk can accurately predict the occurrence or non-occurrence of the risk for both projects. This further validates the predictive accuracy and effectiveness of our developed models.

6.2.4 Risk factor analysis

The methods of Feature Importance Analysis (FIA) and Feature Contribution Analysis (FCA) [205] in ML were adapted to conduct analyses of risk factor importance and contribution. FIA was adapted to assess the general importance of risk factors relevant to broad construction projects. FCA was adapted to analyze the specific contribution of each risk factor to the project's risk at a particular time point in the progression of construction projects. Since this study involves various models, ranging from linear regression to neural networks, choosing an ML feature analysis method that is model-agnostic is advantageous, as it can be applied to a variety of different models.

Shapley Additive exPlanations (SHAP) [206] is recognized as a commonly used method in studies across various fields [207]. It boasts a model-agnostic nature, making it suitable for nearly all ML models. It adeptly manages data with high-dimensional features and can factor in interaction effects among features during analysis. For a specific sample, the contribution of each feature in the model can be quantified as the SHAP value demonstrated in Equation (6.21) [206], computed by iterating through all possible combinations of the features and determining how the introduction of the feature $Pr_{j,k}$ in question alters the prediction. By summing these variations over all combinations and averaging them, the SHAP value for a certain feature is obtained.

$$\mathrm{SHAP}_{j,k}=\sum_{Pr_S \subseteq Pr\backslash\{j,k\}} \frac{|Pr_S|!\,(|Pr|-|Pr_S|-1)\,!}{|Pr|!}\Big[f(Pr_S\cup\{Pr_{j,k}\})-f(Pr_S)\Big] \tag{6.21}$$

Where,

Pr — The set of all features

Pr_S — A subset of features not including $Pr_{j,k}$

$f(Pr_S \cup \{Pr_{j,k}\})$ — The prediction with both the Pr_S and $Pr_{j,k}$

$f(Pr_S)$ — The prediction with just the features in Pr_S

$\frac{|Pr_S|!(|Pr| - |Pr_S| - 1)!}{|Pr|!}$ — A combinatorial coefficient ensuring that each possible combination of features is weighted appropriately

In this study, the importance of risk factor *j* was determined as the mean absolute SHAP value for all its scales, averaged across all project cases in the testing dataset, described in Equation (6.22). This value indicates its average importance in influencing the risk status of general construction projects.

$$I_j = \frac{1}{N \cdot K^{(j)}} \sum_{i=1}^{N} \sum_{k=1}^{K^{(j)}} \left| \mathrm{SHAP}_{i,j,k} \right| \tag{6.22}$$

Where N is the number of project cases in the test dataset; $\mathrm{SHAP}_{i,j,k}$ is the SHAP value of the *k*-th scale of the *j*-th risk factor for the *i*-th project case

To compare risk factor importance uniformly across different models, we normalized the importance values over the 46 risk factors, providing a standardized index for assessment as shown in Equation (6.23).

$$\tilde{I}_j = \frac{I_j}{\sum_{j=1}^{46} I_j} \tag{6.23}$$

For a specific project *i*, the contribution of a risk factor to the risk was determined by the sum of contributions across all its scales, as shown in Equation (6.24).

$$C_{i,j} = \sum_{k=1}^{K^{(j)}} SHAP_{i,j,k} \tag{6.24}$$

6.2.5 Risk reduction strategy

The FIA analysis for each risk aids in identifying important risk factors relevant to a broad range of construction projects. This analysis assists project managers in understanding which risk factors should generally be prioritized and addressed if the project-specific risk assessment is available. When a project-specific risk assessment is performed, the FCA analysis can pinpoint the risk factors that actually contribute to the project's predicted risk degree. This insight informs the creation of targeted risk reduction strategies. To be efficient, the risk reduction strategies should be systematically developed, beginning with addressing the most critical ones. Equation (6.25) illustrates the brute force method of finding the global minimum risk degree by exploring all possible combinations of risk factors.

$$\boldsymbol{Pr}_1^{e_{1,k^*}}, \ldots, \boldsymbol{Pr}_j^{e_{j,k^*}}, \ldots, \boldsymbol{Pr}_J^{e_{J,k^*}} = \underset{1 \le e_{1,k} \le K^{(1)}, \ldots, 1 \le e_{J,k} \le K^{(J)}}{\arg\min} f\left(\boldsymbol{Pr}_1^{e_{1,k}}, \ldots, \boldsymbol{Pr}_j^{e_{j,k}}, \ldots, \boldsymbol{Pr}_J^{e_{J,k}}\right) \tag{6.25}$$

Where,

$e_{j,k}$ — It denotes that the k-th scale of the j-th risk factor is selected

$\boldsymbol{Pr}_j^{e_{j,k}}$ — The one-hot representation of risk factor j, of which the k-th scale is selected so that $Pr_{j,k}=1$, denoted as denoted as $\left(0, \ldots, 1_{(k)}, \ldots, 0\right)_j$

$\boldsymbol{Pr}_j^{e_{j,k^*}}$ — The one-hot representation of risk factor j, of which the k^*-th scale is the optimal selection so that $\boldsymbol{Pr}_{j,k^*}=1$, denoted as $\left(0, \ldots, 1_{(k^*)}, \ldots, 0\right)_j$.

The size of the search space is shown in Equation (6.26). This rapidly expanding search space is computationally demanding and time-consuming, making it impractical for the prompt formulation of risk reduction strategies.

$$N_{search\ space} = \prod_{j=1}^{J} K^{(j)} \tag{6.26}$$

Instead of brute force method, the greedy optimization approach based on the greedy principles [203] can be utilized. The pseudo-code of the optimization process is shown as Algorithm 5. The core idea is to make the best choice at each step. The highest contributing

risk factor is first selected for optimization. The scale that leads to the lowest predicted risk degree is optimal and selected. This process continues until the predicted risk drops below the threshold *T*. The final search space would be reduced to Equation (6.27) at most if *m* risk factors are optimized.

$$N_{greedy\ search\ space} = \sum_{j}^{m} K^{(j)} \tag{6.27}$$

Algorithm 5: Risk reduction strategy based on Greedy optimization

1: Initialize $j = 1$, the risk factor with the most risk contribution

2: While $f\left(\boldsymbol{Pr}_{j}^{e_{j,k}}, \ldots, \boldsymbol{Pr}_{J}^{e_{J,k}} \mid \boldsymbol{Pr}_{1}^{e_{1,k^*}}, \ldots, \boldsymbol{Pr}_{j-1}^{e_{j-1,k^*}} \right) > T$

3: $k^* = \underset{1 \le e_{j,k} \le K^{(j)}}{arg\ min} \left(f\left(\boldsymbol{Pr}_{j}^{e_{j,k}} = \left(0, \ldots, 1_{(k)}, \ldots, 0 \right)_j, \boldsymbol{Pr}_{j+1}^{e_{j+1,k}}, \ldots, \boldsymbol{Pr}_{J}^{e_{J,k}} \mid \boldsymbol{Pr}_{1}^{e_{1,k^*}}, \ldots, \boldsymbol{Pr}_{j-1}^{e_{j-1,k^*}} \right) \right)$

4: $\boldsymbol{Pr}_{j}^{e_{j,k^*}} = \left(0, \ldots, 1_{(k^*)} \ldots, 0 \right)_j$

5: $j \leftarrow j + 1$

6: *m* risk factors have been optimized; the risk reduction strategy is: $\left\{ \boldsymbol{Pr}_{1}^{e_{1,k^*}}, \ldots, \boldsymbol{Pr}_{j}^{e_{j,k^*}}, \ldots, \boldsymbol{Pr}_{m}^{e_{m,k^*}}, \boldsymbol{Pr}_{m+1}^{e_{m+1,k}}, \ldots, \boldsymbol{Pr}_{J}^{e_{J,k}} \right\}$

7: End

6.3 Case study

6.3.1 Data source

To prove the applicability of developed approach, we conducted a real-world application in collaboration with a leading engineering and contracting firm in the U.A.E. (Company A). This firm is a subsidiary of a major investment corporation renowned for its complex and iconic projects. Founded in 1999, this company has a workforce of over 12,000 and offers a range of services from design management to procurement. Our collaboration involved data collection from a commercial building project in the UAE, worth over 5 million dollars and lasting 24 months, where our partner handled the construction phase. Each risk factor was thoroughly explained in online meetings to highlight its relevance to construction projects. The manager and members of the IT and construction teams were

then asked to provide information on the 46 risk factors related to the project within a two-week period. The real project data is presented in Table 6.9. The "Scale" refers to the chosen scale's index, which can be referenced in the Appendix. Additionally, the teams were inquired about any cyber risks encountered, and they revealed that phishing and data breaches had been encountered.

Table 6.9. Project data from Construction Company A

Risk factor	1.1	1.2	1.3	1.4	1.5	1.6	1.7	2.1	2.2.1	2.2.2	2.2.3	2.2.4	2.2.5	2.2.6	2.2.7	2.2.8	2.3.1	2.3.2	2.3.3	2.3.4	2.3.5	2.3.6	2.3.7
Scale	0	4	2	4	5	3	1	1	2	4	5	5	5	5	5	5	1	3	7	7	7	7	7
Risk factor	2.3.8	2.4	3.1	3.2	3.3	3.4	3.5	3.6	3.7	3.8	3.9	4.1	4.2	4.3	4.4	4.5	5.1	5.2	5.3	5.4	5.5	5.6	5.7
Scale	7	0	4	2	0	4	4	1	2	2	1	5	1	0	2	1	1	0	0	1	1	0	1

The case study involves three components of implementation: (1) Risk Prediction: The developed model for each risk is applied to predict the risk degree. (2) Risk Factor Analysis: FIA and FCA analysis is performed to identify the risk factors that are generally important, as well as those that actually contribute to the predicted risk. (3) Risk Reduction Strategies: The greedy optimization algorithm is implemented for suggesting risk reduction strategies.

6.3.2 Risk prediction

The data from Table 6.9 were applied to the optimal model for each risk. Table 6.10 displays these predictions, indicating phishing and data breaches as the most probable risks. These predictions align with the actual events observed in the project, demonstrating a close match between predicted and actual occurrences. This consistency underscores the validity of our models.

Table 6.10. Comparison of predicted and actual risk occurrences

Risk	Optimal Model	Predicted Risk Degree	Predicted Occurrence	Actual Occurrence
Ransomware	NN_2	0.47	No	No
Phishing	NN_2	0.70	Yes	Yes
Insider attack	Ridge Regression	0.48	No	No
Data breach	Combined_NN_3	0.64	Yes	Yes
Supply chain	Ridge Regression	0.39	No	No

6.3.3 Risk factor analysis

6.3.3.1 General importance of risk factors

Based on the methods in Section 6.2.4, we conducted FIA for each optimal model to obtain the risk factor importance, with the basic linear regression as baseline. Table 6.11 presents the top 10 risk factors for each risk. It is obvious that there is minimal overlap (not greater than 50% overlap) between the top 10 risk factors of the optimal models and those of the basic linear regression models. The limitations of basic linear regression, such as its inability to capture non-linear relationships and feature interactions, diminish its effectiveness in ranking risk factor importance. This demonstrates that our selected models are better at reflecting the data's non-linearities.

Project managers seeking to address a specific type of risk can refer to the five ranked lists for the optimal models in Table 6.11 to formulate risk prevention strategies. For those aiming to address all risks holistically, they can refer to Table 6.12, which enumerates the ten risk factors that appear most frequently across the five ranked lists associated with the optimal models, highlighting their significance in the broader cybersecurity landscape of construction projects.

Table 6.11. Top 10 risk factors for each risk

	Ranking No. (from the most to the least important)		1	2	3	4	5	6	7	8	9	10
Ransomware	Optimal model	Risk factor	3.3	5.3	2.2.2	1.7	2.2.3	3.8	5.4	3.7	2.2.4	3.2
		Importance	0.0485	0.0402	0.0402	0.0379	0.0365	0.0338	0.0326	0.0304	0.0303	0.0281
	Linear regression	Risk factor	5.4	3.7	5.3	2.2.1	1.3	5.2	4.1	1.7	5.7	4.3
		Importance	0.1166	0.0676	0.0586	0.0427	0.035	0.0345	0.0337	0.0327	0.0316	0.0295
	Percentage of overlap		40%									
Phishing	Ranking No.		1	2	3	4	5	6	7	8	9	10
	Optimal model	Risk factor	3.7	3.3	5.3	5.4	2.4	1.4	1.7	2.2.1	4.5	3.6
		Importance	0.0449	0.044	0.0395	0.0388	0.0383	0.0365	0.034	0.0323	0.0311	0.0303
	Linear regression	Risk factor	1.7	5.3	3.7	5.4	2.2.1	1.3	4.1	3.1	5.7	5.5
		Importance	0.0725	0.067	0.0665	0.0608	0.0552	0.0404	0.0365	0.0327	0.032	0.0298
	Percentage of overlap		50%									
Insider attack	Ranking No.		1	2	3	4	5	6	7	8	9	10
	Optimal model	Risk factor	3.3	3.7	1.7	5.1	3.8	3.2	1.6	1.4	3.1	5.4
		Importance	0.1193	0.0626	0.0587	0.0576	0.0555	0.0533	0.0412	0.0386	0.0346	0.0335
	Linear regression	Risk factor	3.7	3.1	2.2.1	5.3	1.7	5.4	5.7	1.3	3.6	5.5
		Importance	0.06	0.0583	0.0562	0.0535	0.0516	0.0449	0.0363	0.0363	0.0361	0.0341
	Percentage of overlap		40%									
Data breach	Ranking No.		1	2	3	4	5	6	7	8	9	10
	Optimal model	Risk factor	2.2.2	2.2.3	3.7	3.3	2.2.4	2.4	3.1	3.6	1.7	2.2.5
		Importance	0.0565	0.0446	0.0434	0.0374	0.0356	0.0336	0.0313	0.028	0.0268	0.0268
	Linear regression	Risk factor	5.4	3.7	5.3	2.2.1	1.7	5.2	4.1	1.3	5.7	4.3
		Importance	0.1153	0.0666	0.0575	0.0447	0.0346	0.0345	0.0344	0.0336	0.0303	0.0301
	Percentage of overlap		20%									
Supply chain attack	Ranking No.		1	2	3	4	5	6	7	8	9	10
	Optimal model	Risk factor	5.4	2.2.2	2.2.3	2.4	2.2.5	3.3	3.8	2.2.4	5.1	2.2.6
		Importance	0.105	0.0411	0.0394	0.0368	0.0357	0.0349	0.0348	0.0332	0.032	0.0304
	Linear regression	Risk factor	5.4	2.2.1	3.7	5.3	5.2	4.3	3.4	3.5	5.7	1.3
		Importance	0.1042	0.0599	0.0593	0.046	0.0445	0.0429	0.0418	0.0411	0.0352	0.0333
	Percentage of overlap		10%									

Table 6.12. Top 10 risk factors based on appearance frequency

Risk factor	3.3	5.4	1.7	3.7	2.2.4	2.4	3.8	2.2.3	2.2.2	3.2
Appearance frequency	5	4	4	4	3	3	3	3	3	2

6.3.3.2 Risk factor contribution analysis

Based on the methods in Section 6.2.4, the FCA analysis was conducted for phishing and data breaches. The results, showcasing the top 10 risk factors contributing most significantly to the two risks, are plotted in Figure 6.6 and Figure 6.7, respectively.

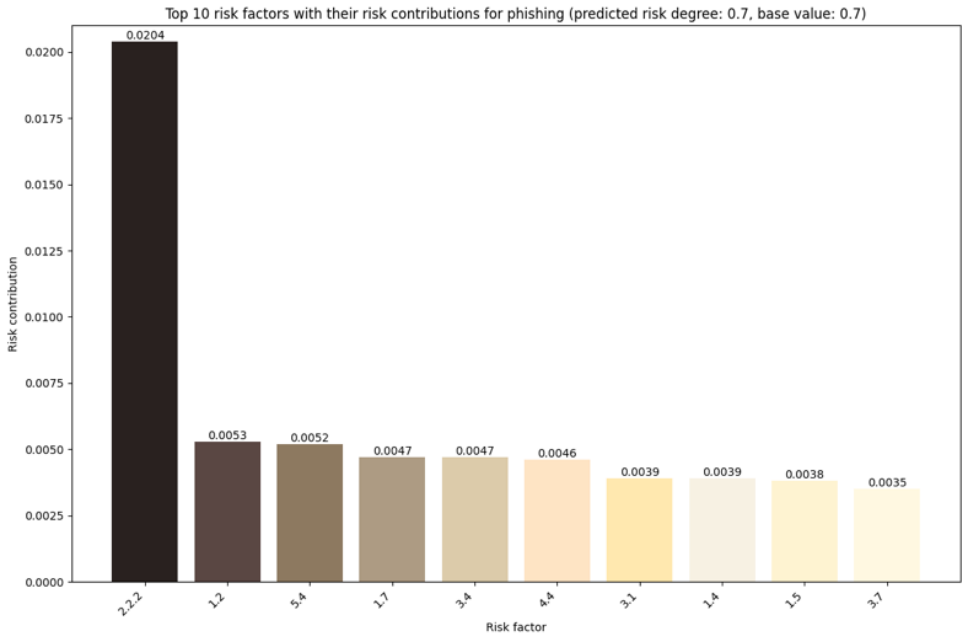


Figure 6.6. Top 10 risk factors with their risk contributions for phishing

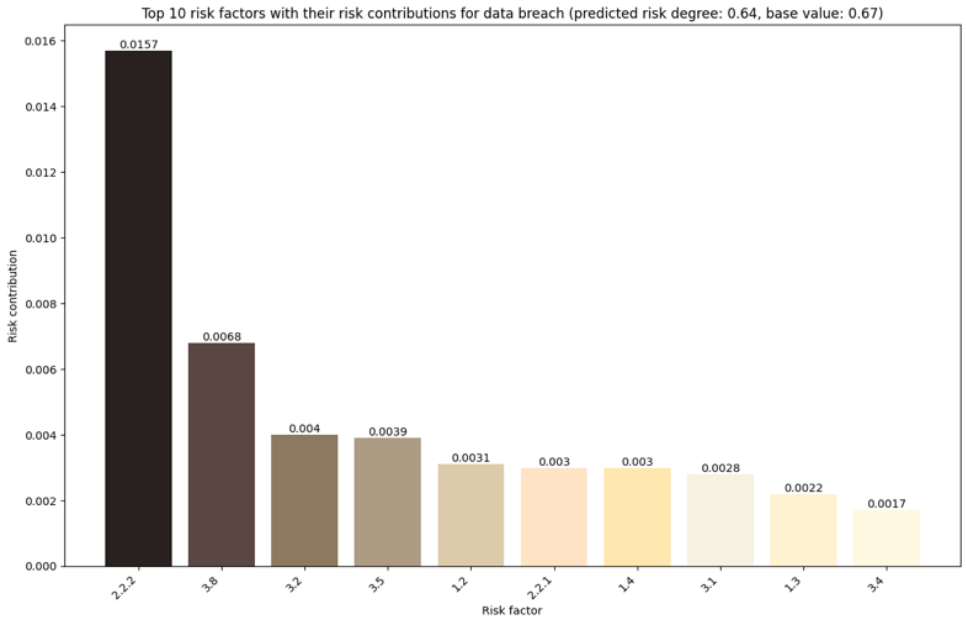


Figure 6.7. Top 10 risk factors with their risk contributions for data breach

The following observations can be made:

(1) It is evident that risk factor 2.2.2 is a major contributor to both types of risks. This factor indicates that the number of sub-teams at the second layer of the project is excessive, with over 40 sub-teams in the project data. This should receive significant attention because a higher number of teams increases exposure to cyberattacks. Furthermore, the sub-teams in the second layer of the project are less adept at implementing cybersecurity measures and defenses.

(2) Over 50% of the risk factors in the two top-10 lists overlap. This suggests that these risk factors have a concurrent effect: contributing to one risk might also contribute to another. This implies that by addressing these critical factors, there is a high likelihood that the risk degrees for all associated risks could be reduced.

(3) Risk factors in Category 1, related to general project information, rank first in their appearance on the top-10 contributing risk factors of phishing and second on that of data breach, with risk factors 1.2 (project budget) and 1.4 (project duration) both appearing. This highlights that the foundational elements, determined during the

planning phase, significantly influence the cybersecurity landscape of a construction project. As such, cybersecurity considerations should be integral to the early project planning stages.

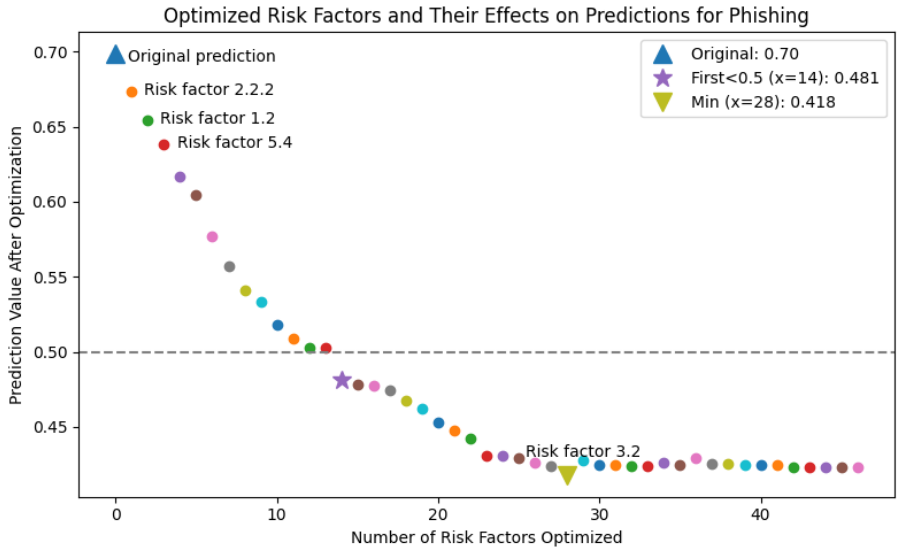
(4) Risk factors in Category 3, pertaining to the IT aspect, are also prominent in both top-10 lists. This highlights the critical role IT-related risk factors have in the cybersecurity of construction projects. As such, construction projects should ensure that their IT strategies are well-formulated and effectively implemented.

(5) Risk factors in Categories 4 and 5 are least prominent in the top-10 lists for both risks, indicating that Construction Company A is performing well in operational technology and human and management aspects.

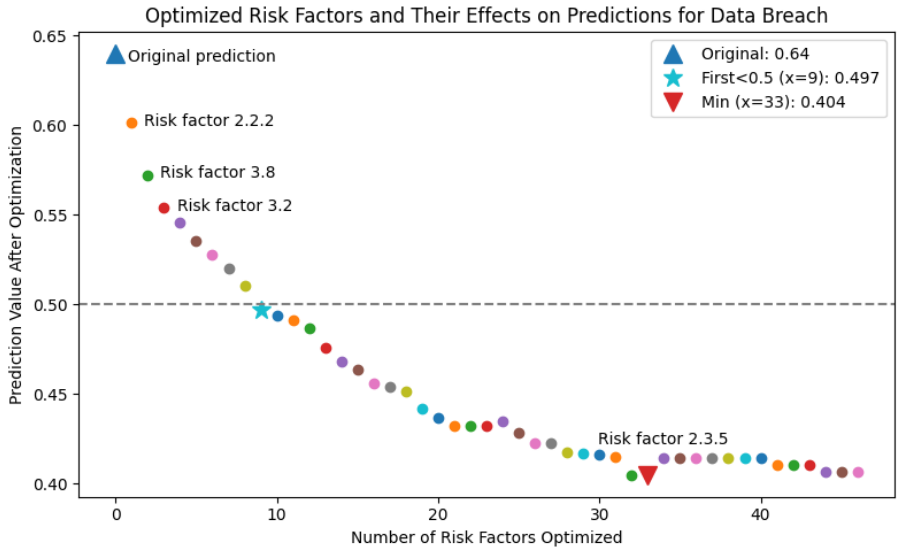
6.3.4 Risk reduction strategy

Utilizing the algorithm in Section 6.2.5, we implemented the greedy optimization process to mitigate these two cyber risks. Figure 6.8 shows the predicted risk degrees against the number of optimized risk factors. Table 6.13 displays more detailed information, where "Scale*" means the index of the optimal scale of the optimized risk factor. For instance, for phishing, the risk factor 1.5 – 'What is the total number of people involved in the project (labor excluded)?' – is recommended to choose the risk factor scale with index 1, which means 51 – 100 people as per the Appendix.

When the risk threshold is set at 0.5, we recommend prioritizing the top 14 risk factors significantly contributing to phishing risk, which is projected to reduce the risk degree to 0.481. Similarly, addressing the top 9 risk factors associated with data breach risk should lower its degree to 0.497. If a more conservative approach is preferred, the threshold can be adjusted as per the project manager's criteria, the project's risk tolerance, and the feasibility and cost of mitigating these risks. For risks predicted not to occur, continuous monitoring is advised throughout the project. This approach involves using the trained models for ongoing risk predictions as project data is updated in subsequent project phases.



(a) Phishing



(b) Data Breach

Figure 6.8. Predicted risk degrees against the number of optimized risk factors

Table 6.13. The risk predictions against the number of optimized risk factors

Phishing	No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
	Risk factor	2.2.2	1.2	5.4	1.7	3.4	4.4	1.4	3.1	1.5	3.7	3.8	4.1	3.5	2.2.5	1.6	2.2.1	1.3	4.3	2.2.6	2.2.7	2.3.2	2.2.8	2.2.4
	Scale*	0	0	0	0	0	0	1	0	1	0	0	0	4	0	1	1	1	1	0	1	2	0	0
	Predicted risk	0.673	0.654	0.638	0.617	0.604	0.577	0.557	0.541	0.534	0.518	0.509	0.503	0.503	0.481	0.478	0.478	0.475	0.468	0.462	0.453	0.448	0.442	0.431
	No.	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46
	Risk factor	2.3.8	3.9	5.5	2.3.3	3.2	2.3.6	2.2.3	5.3	2.3.4	4.5	2.3.7	5.6	2.3.5	5.1	5.2	2.1	3.3	3.6	1.1	2.4	4.2	5.7	2.3.1
	Scale*	1	0	2	2	1	1	0	0	1	1	1	4	0	2	0	7	0	1	4	0	1	0	1
	Predicted risk	0.431	0.43	0.426	0.424	0.418	0.428	0.425	0.425	0.424	0.424	0.426	0.424	0.429	0.425	0.425	0.424	0.424	0.424	0.423	0.423	0.423	0.423	0.423
Data breach	No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
	Risk factor	2.2.2	3.8	3.2	3.5	1.2	1.4	2.2.1	3.1	1.3	3.4	4.1	1.5	3.7	4.4	1.7	1.6	5.4	5.5	2.2.7	2.2.8	4.3	3.9	5.3
	Scale*	0	0	1	0	0	1	0	0	1	0	0	0	0	0	0	1	0	0	0	0	1	1	0
	Predicted risk	0.601	0.572	0.554	0.546	0.535	0.528	0.52	0.51	0.497	0.494	0.491	0.487	0.476	0.468	0.464	0.456	0.454	0.452	0.442	0.436	0.432	0.432	0.432
	No.	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46
	Risk factor	2.3.6	2.3.2	2.2.5	4.5	2.2.6	2.3.8	2.3.3	2.3.4	2.2.4	2.3.5	2.3.7	5.6	1.1	3.6	2.3.1	5.2	5.7	5.1	2.1	4.2	2.2.3	3.3	2.4
	Scale*	2	2	1	1	0	1	1	1	0	0	1	0	0	1	1	0	1	2	1	1	0	2	0
	Predicted risk	0.435	0.428	0.422	0.422	0.417	0.417	0.416	0.415	0.404	0.404	0.415	0.415	0.415	0.415	0.415	0.415	0.415	0.41	0.41	0.41	0.407	0.407	0.407

6.4 Discussions

6.4.1 The complexity of cybersecurity landscape

To explore the model behavior, R^2 values were extracted form Table 6.6 and plotted in Figure 6.9. R^2 assesses the variance in labels explained by the model, indicating the linearity between training data and labels. High R^2 values suggest effective variance capture by the model. For simpler models like regression, high R^2 indicates a strong linear relationship between the dataset and labels. Conversely, low R^2 values in simpler models, but high in complex models like neural networks, imply significant non-linearity in the dataset.

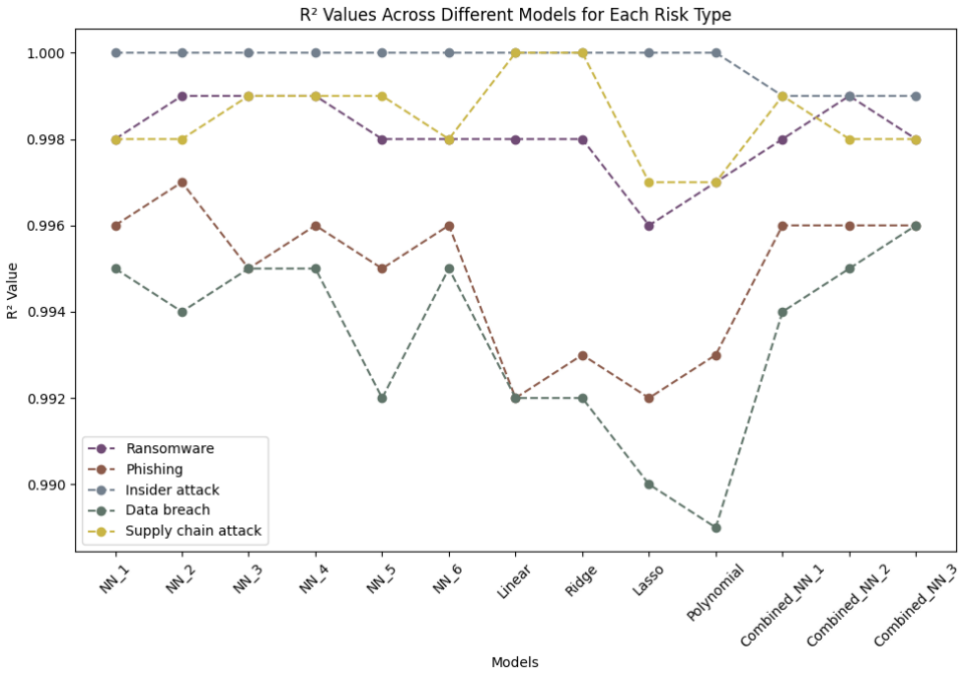


Figure 6.9. R^2 values across different models for five cyber risks

Figure 6.9 shows that the insider attack risk consistently achieves near-perfect R^2 values across all models, including simple ones like Linear, Ridge, and Lasso regression. This indicates a predominantly linear relationship between the dataset and labels for insider

attack, making basic models suitable for prediction. Similarly, supply chain attack risk also demonstrates high R^2 values with simpler models, confirming a linear relationship. In cases of such obvious linearity, complex models underperform on test data because of overfitting, affirming our choice of simpler regression models for these two risk types.

For ransomware attack, all models achieve high R^2 values, but complex models like neural networks surpass simpler regressions. This indicates less linearity between the dataset and labels for ransomware than for the two risks above, with non-linearity better captured by complex models. Similar patterns are seen for phishing and data breach risks. Notably, data breach risk has consistently lower R^2 values across all models, suggesting the highest non-linearity and the limited capacity of even sophisticated models to fully capture this. Such high non-linearity makes complex models, like neural networks, more suitable for predicting these risk degrees, justifying our selection of neural networks for the three risks.

The linearity analysis illuminates the complexity of mapping risk factors to cyber risks. Nearly all risks exhibit some non-linearities, signifying a complex relationship between risk factors and cyber risks. This necessitates that project managers in the construction industry have a comprehensive understanding of these risk factors to gain deeper insights into cyber risks, requiring a commitment to cybersecurity and a holistic approach. Additionally, different cyber risks demonstrate distinct non-linear relationships with their respective risk factors, suggesting that effectively addressing each type of cyber risk may require strategies specifically tailored to these unique relationships. Due to these non-linearities, effective reduction of cyber risk necessitates a coordinated approach that encompasses multiple risk factors to achieve significant risk reduction, as addressing individual factors in isolation may not result in a substantial decrease in risk.

6.4.2 Risk factor of general importance

Table 6.12 lists the top 10 critical factors for risk prevention, from which several key insights can be derived.

(1) IT-related risk factors are the most impactful on risk predictions in construction cybersecurity. Key among these is the presence of a dedicated IT team (risk factor 3.3), which is vital for the continuous monitoring and swift management of cybersecurity threats. The choice between public and private networks (risk factor 3.7) also emerges as critical; private networks markedly decrease vulnerability to external threats, enhancing security control. Additionally, the importance of anti-phishing training (risk factor 3.8) cannot be overstated, as it significantly elevates employee awareness and reduces the incidence of breaches due to human error. Project managers are advised to prioritize these areas in cybersecurity planning for effective resource allocation.

(2) Factors related to project structure, such as the number of stakeholders at different layers (risk factor 2.2) and percentage of teams overlapping in different projects (risk factor 2.4), significantly impact the cybersecurity status. The distribution of teams across project levels must be carefully managed, as stakeholders from various levels often have distinct perspectives and approaches to cybersecurity. These differences can result in inconsistencies in the project's cybersecurity defenses. Therefore, a careful balance must be struck between the distribution of team numbers across layers and the potential vulnerabilities that may arise. Moreover, minimizing team overlap across projects is critical because it complicates the management of information security protocols and can create gaps in the cybersecurity execution. When team members are allocated to multiple projects, it may dilute their focus and lead to inconsistencies in security practices, increasing the risk of errors, data breaches, or non-compliance with cybersecurity standards.

(3) Factors related to management side such as the implementation of MFA (risk factor 5.4) is crucial cybersecurity measures. MFA, including biometrics or face recognition, significantly enhances security by adding layers of verification, reducing unauthorized access risks.

(4) A dedicated cybersecurity legal team (Risk factor 1.7) is also important for cyber risk management because it ensures compliance with evolving legal requirements

and standards. This team can proactively identify potential legal issues related to cybersecurity, advise on risk mitigation strategies, and help navigate the complex landscape of international cyber laws, thus protecting the organization from legal 144and financial penalties.

(5) The phase of a construction project (risk factor 3.2) is also a determinant of its cybersecurity status. This observation is consistent with the understanding that different project phases may have varying types of threats and vulnerabilities and also varying levels of them, as demonstrated in the work [109]. Therefore, as a project progresses through its phases, a tailored set of cybersecurity plans and measures should be prepared and implemented to address the evolving risk landscape.

6.4.3 The practicality and prospect of the models

In this study, the models successfully predicted the occurrence of cyber risks in two projects labeled by experts and in a real project from Company A. This demonstrates the validity and effectiveness of our models. Additionally, risk factor contribution analysis through the models has accurately identified the factors contributing to the risks and the extent of their contributions. This analysis is invaluable for project managers, particularly during the project's lifecycle. The models enable the prediction of cyber risk status at any stage of a construction project, allowing project managers to implement immediate risk reduction strategies. These strategies are informed by the model’s greedy optimization algorithm, aimed at maximizing resource allocation efficiency. In the future, we plan to develop web and mobile applications for cyber risk assessment, specifically designed for practitioners and project managers. The backbone of these applications will be the trained models, enabling users to input project information and receive risk predictions and recommended response strategies directly.

6.4.4 Limitations and future works

The primary limitation of our study stems from the lack of an existing dataset, leading us to rely on a simulated dataset for training our models. Although we based our Monte Carlo simulation on rigorously determined probability distributions, there remains a chance that these distributions may not accurately reflect real-world projects. Additionally, due to data scarcity, our trained ML models have not been tested on a large external dataset, which could result in a degree of variance in the results. In the future, we plan to conduct sensitivity analyses on the probability distribution assumptions of the risk factor scales to enhance the robustness of our models, particularly when still using simulated datasets for training. Simultaneously, we are actively seeking collaborations with industry partners to gather as much real-world data as possible, aiming to eventually replace the simulated dataset and further enhance the robustness and reliability of our models.

6.5 Conclusions

The construction industry is becoming increasingly susceptible to cyber threats, potentially leading to severe consequences. However, there remains a scarcity of research on dynamic cyber risk assessment through the progression of construction projects. To bridge this gap, this study developed an ML-centric approach for assessing the five most common cyber risks in construction projects: ransomware, phishing, insider attacks, data breaches, and supply chain attacks. The developed approach enables to dynamically predict cyber risks throughout construction project progressions, identify the highest contributing risk factors, and suggest efficient strategies for risk reduction. A case study was conducted to demonstrate the applicability of the developed approach to a real construction project.

This study reveals the complex relationship between risk factors and cyber risks in construction, highlighting the need for project managers to deeply understand and strategically address these risks. Tailored strategies and a coordinated approach encompassing multiple factors are essential for effective cyber risk reduction. The study

highlights critical risk factors that are broadly relevant across construction projects, with a special emphasis on IT-related factors. These include the necessity for a dedicated IT team, the deployment of private networks, and comprehensive anti-phishing training. It stresses the imperative for implementing distinct security mechanisms tailored to each layer of the project, as well as the importance of minimizing personnel overlap across multiple projects to enhance security. Additionally, the study points out the significance of management practices, such as Multi-Factor Authentication (MFA), in establishing a robust multi-layer security defense. The inclusion of a dedicated cybersecurity legal team is underscored as crucial for ensuring adherence to the evolving legal standards and requirements pertaining to cybersecurity practices. Furthermore, the study identifies the project phase as a critical determinant of cybersecurity needs, suggesting that cybersecurity plans should be specifically tailored to each phase of the project. This approach is recommended in order to effectively address the varying risk landscapes encountered throughout the lifecycle of a project.

Future work will concentrate on replacing simulated data with real-world data to further improve the models' accuracy and applicability. Furthermore, we plan to develop web and mobile applications for cyber risk management tailored for practitioners and project managers. These applications will offer models that are regularly updated with new, real data, allowing users to input project details and instantly obtain risk predictions along with suggested strategies for reducing risks.

Chapter 7

Conclusions and Future Directions

This section summarizes the dissertation, reiterates the research findings of each study, and outlines future research directions.

7.1 Dissertation summary

This dissertation presents a comprehensive study on cybersecurity in the construction industry, structured into several chapters, each aimed at enhancing cyber risk management in modern construction projects. Chapter 3 utilizes the LDA topic modeling technique to identify construction cybersecurity research topics, aiming to unify fragmented research efforts and facilitate a cohesive understanding of cybersecurity challenges and opportunities in construction. Cyber risk management emerges as a key topic, guiding the focus of subsequent chapters. Chapter 4 introduces an innovative approach that employs a tailored language model to identify cyber risks relevant to various construction project phases, culminating in a comprehensive, prioritized checklist of cyber risks. This checklist serves as a new benchmark to improve industry-wide recognition of and preparedness for

cyber risks. Chapter 5 uses a systematic methodology, combining literature review, questionnaire surveys, and expert consultations, to identify risk factors associated with construction projects, which are beneficial for future quantitative risk assessments. Chapter 6 develops an ML-centric approach for assessing common cyber risks at the project level, utilizing the risk factors identified in Chapter 5 as ML features. Various models, each with a different ability to capture non-linearity, are tested. ML feature analysis methods are adapted for risk factor importance and contribution analysis, and a greedy optimization algorithm is formulated for a more efficient risk reduction strategy. A case study demonstrates the developed approach's applicability and effectiveness. In summary, this dissertation represents a pioneering systematic study to enhance cyber risk management for construction projects. It contributes to the existing body of knowledge by proposing methodologies and frameworks for cyber risk management, fostering interdisciplinary research among AI, cybersecurity, and construction management, and offering practical tools for industry practitioners.

7.2 Research findings

This dissertation explored cybersecurity in the construction industry. By implementing rigorous existing methodologies and integrating cutting-edge technology, this study aimed to significantly advance the construction industry's approach to cybersecurity, facilitating enhanced cyber risk management practices in the construction industry. The research findings of each chapter are as below:

(1) Chapter 3 identified eight topics within the domain of cybersecurity in the construction industry. These topics were delineated as follows: Performing Risk Management, Preventing the Increasing Cyber Incidents, Detecting Phishing and Malware, Strengthening Management Processes, Protecting Network Devices, Regulating Information Storage and Sharing, Protecting Privacy, and Improving Authentication Processes. These identified topics, along with their corresponding suggested actions, provided a framework for researchers interested in exploring future research directions.

Based on the topic prevalence analysis, we recommend prioritizing research on three main topics: Performing Risk Management, Preventing Cyber Incidents, and Detecting Phishing and Malware. Together, these form a layered approach to cybersecurity tailored to the construction industry, combining holistic strategic planning, proactive prevention, and reactive detection.

(2) Chapter 4 developed a language model to identify cyber risks across different phases in construction projects. The adapted SFT and RLHF training techniques enhanced the model's capacity to comprehend cybersecurity content and respond to related cybersecurity queries accurately. As a deliverable, this study compiled a cyber risk checklist categorized by project phases and ranked by risk values, which outperformed existing literature and obtained affirmations from industry experts for its relevance and usefulness. The checklist highlighted the inadequate identity and access management, unauthorized access, data falsification, dependency on obsolete software, and the leakage of essential information. These concerns underscored the necessity for collaborative efforts between IT and construction sectors to effectively integrate cybersecurity with project management practices.

Two primary applications of the prioritized checklist were suggested: Firstly, it can serve as a new benchmark for project managers to formulate proactive cybersecurity measures towards high-priority risks. Secondly, risk analysts can utilize the checklist for detailed project-specific risk assessments, prioritizing significant risks for developing targeted management strategies. Additionally, this study demonstrated the superiority of the developed language model in generating informative responses to cybersecurity inquiries, which evidenced the model's potential as a user-friendly, intelligent cybersecurity consultant for construction personnel with different expertise levels. The language model allows for customization for construction companies, enabling integration into their mobile or web applications. Being periodically updated and upscaled, it will become an integral part of the digitized and automated cybersecurity platform, offering various advanced cybersecurity management functions for industry-wide applications.

(3) Chapter 5 employed a comprehensive approach combining an iterative process of literature review, questionnaire survey and expert consultation for identifying risk factors that cause five common cyber risks. 32 risk factors covering five aspects of construction projects were determined: Overall Information of the Project, Project Structure, IT Factors, OT Factors, and Management and Human Factors. These identified risk factors have four advantages for future cyber risk assessment:

- Capturing project structure dynamics: The study innovated by analyzing the project as a multi-layered network, identifying risk factors related to the distribution of sub-teams and communication channels across layers. This approach enhanced the ability of risk assessment models to reflect the complex and dynamic nature of construction projects.
- Enhancing specificity with contextual insight: The identified risk factors provide a dual perspective, combining a broad overview of the project with detailed, phase-specific data sourced directly from companies. This dual perspective allows for nuanced phase-specific risk predictions enriched with project-wide information, making the predictions more aligned with reality.
- Enabling a more quantitative risk assessment: The risk factors, segmented into distinct numerical or ordinal scales, allow for a more quantitative assessment of cyber risks in construction projects. This segmentation can help reduce subjectivity during the risk assessment process and ensure consistency across different projects.
- Addressing unique industry vulnerabilities: The identified risk factors cover both general cybersecurity challenges and those unique to the construction industry, ensuring a balanced and relevant risk assessment tailored to the construction industry.

(4) Chapter 6 introduced an ML-centric approach for dynamic cyber risk assessment in construction projects, focusing on the five common cyber risks: ransomware, phishing, insider attacks, data breaches, and supply chain attacks. The developed models could not

only predict cyber risks as projects progressed but also identify the high-contributing risk factors and proposed effective risk reduction strategies. A detailed case study validated the approach's effectiveness for real-world construction projects. This study revealed the complex relationship between risk factors and cyber risks in construction, highlighting the need for project managers to deeply and holistically understand and strategically address these risks. Tailored strategies and a coordinated approach encompassing multiple factors are essential for effective cyber risk reduction.

The study highlighted critical risk factors relevant across construction projects, particularly focusing on IT-related factors. These include the necessity for a dedicated IT team, deployment of private networks, and comprehensive anti-phishing training. It emphasized the need for implementing distinct security mechanisms tailored to each project layer, and minimizing personnel overlap across multiple projects to enhance security. Additionally, the study underscored the significance of management practices, such as MFA, in establishing a robust multi-layer security defense. The inclusion of a dedicated cybersecurity legal team is also crucial for ensuring adherence to evolving legal standards and requirements. Furthermore, the study identified the project phase as a critical determinant of cybersecurity needs, suggesting that cybersecurity plans should be specifically tailored to each phase to address varying risk landscapes throughout the project lifecycle.

7.3 Future directions

7.3.1 Enhancing the language model

Our developed proof-of-concept language model has demonstrated a high ability to understand cybersecurity content and to generate cohesive and relevant answers to cybersecurity questions. To make it more accessible and to achieve more advanced functions, future updates will focus on expanding the model size and enhancing the coverage and quality of the training text data, with the goal of developing a construction

cybersecurity-specific large language model that is significantly more comprehensive and capable, similar to GPT-4.

Based on the findings presented in Table 6.12, which highlights the top critical factors of general importance for construction projects, the text data collection and preservation strategy should focus on areas with the highest impact. Specific types of documents that should be prioritized include:

- IT Team Operations: Operational records and logs from dedicated IT teams that detail daily security activities, incident response strategies, and regular maintenance schedules. These documents are invaluable for understanding the practical aspects of managing and mitigating cybersecurity threats.
- Network Configuration Records: Documentation on the architecture and security configurations of both public and private networks used within construction projects. This should include assessments of network vulnerabilities, network security policies, and changes in network setups over time.
- Anti-Phishing Training Materials: Curricula and training feedback reports related to anti-phishing measures, which are critical for raising employee awareness and reducing human-error-induced breaches. This would also include outcomes and effectiveness assessments of these training programs.
- Stakeholder and Project Team Structuring: Comprehensive records detailing the number and roles of stakeholders at various project levels, the structuring of project teams, and protocols for minimizing team overlaps across different projects. These documents help in understanding the complex interactions and potential inconsistencies in cybersecurity practices due to project team dynamics.
- Multi-Factor Authentication (MFA) Implementation: Documents related to the deployment and operation of MFA systems, including biometrics or face recognition technologies. This would encompass setup protocols, user feedback, and performance evaluations of these security systems.

- Cybersecurity Legal Compliance: Records from dedicated cybersecurity legal teams, including ongoing compliance checks, updates to cyber law adherence, and strategic recommendations to avoid legal and financial repercussions.
- Project Phase-Specific Security Plans: Documents that outline specific cybersecurity plans and measures tailored to different phases of construction projects. These should detail threat assessments and risk mitigation strategies that evolve as the project progresses.

Additionally, we also plan to enhance the model's ability to interpret and analyze more kinds of documents, including technical drawings, software codes, and schedules. This will enable the production of tailored outputs such as figures, tables, and LaTeX codes. The enhanced language model has the potential to serve as an intelligent cybersecurity consulting resource for personnel in the construction industry, regardless of their level of expertise. Integrating this model into web or mobile applications could significantly increase its accessibility to contractors, project managers, safety officers, and other stakeholders. This integration will offer not only real-time risk monitoring and solution recommendations but also project-specific and dynamic risk identification and analysis, as well as customized cybersecurity training and support across the construction industry.

7.3.2 Including resilience factors

In the 32 identified risk factors, we have not considered factors related to the resilience of the construction project for cybersecurity. Resilience factors are the ones related to qualities or mechanisms that enable systems, organizations, or projects to withstand, adapt to, and recover from adverse events, ensuring continuity and minimal impact. For example, this includes the ability of IT and OT systems to detect, resist, and recover from cyberattacks without compromising the integrity and functionality of the construction project. Other overlooked aspects include the adaptability of existing cybersecurity measures in the face of evolving threats, and the robustness of existing incident response plans to ensure

minimal disruption and swift recovery. Additionally, the redundancy of critical systems—namely, the presence of backup systems and data redundancy to ensure critical project operations can continue even if primary systems are compromised—was not fully explored. In the future, we will expand our risk assessment framework to include these resilience-related factors. This will involve developing a comprehensive evaluation of the project's cyber resilience capabilities, incorporating the latest cybersecurity technologies and strategies. By doing so, we aim to provide a more holistic approach to managing cyber risks in construction projects, ensuring they are better equipped to withstand and recover from cyber incidents.

7.3.3 Substituting the simulated dataset

In Chapter 6, although we based our dataset on rigorously determined probability distributions through Monte Carlo simulation, there remains a chance that these distributions may not accurately reflect real-world projects. In the future, we plan to conduct sensitivity analysis on the probability distribution assignment to make the trained models more robust. Simultaneously, we are actively seeking collaborations with industry partners to gather as much real-world data as possible. Incorporating this new dataset will gradually replace our simulated dataset, enhancing the robustness and reliability of our models. We aim to create a dataset that can be open-sourced and beneficial to researchers across the industry. To systematically and consistently collect and utilize data for model training or cybersecurity purposes, we will create a structured format that can serve as the basis for industry standardization. This standardization is necessary to guide what data should be collected, especially for OT data. The identified risk factors should be structured into a fixed format and presented to relevant companies for data collection. Implementing a standardized data collection protocol can guide the collection of pertinent data. This protocol should outline the types of data required for each risk factor, methods of collection, and the frequency of updates.

7.3.4 Creating cybersecurity digital platform

The culmination of this dissertation will be the creation of an all-encompassing cybersecurity digital platform tailored for construction projects. The main goal of this platform is to provide targeted cybersecurity answers catering both to specific construction projects and to general cybersecurity inquiries. The platform can output answers in a format that is both readable and understandable. Inputs can vary, including qualitative questions or structured risk factors that reflect the characteristics of the project. The digital platform will feature several modules, enabling a comprehensive range of functions including question answering, project-specific risk identification, quantitative cyber risk assessment, risk mitigation strategy recommendation, and personnel cybersecurity training, among others. These modules will undergo periodic updates, particularly with respect to the training dataset. Since the platform facilitates interaction between different models and users, the inputs, and interactions—subject to appropriate permissions—will serve to enrich the training dataset. This iterative improvement process ensures that the various modules advance over time. Integrating the digital platform into both web and mobile applications, which can be embedded into construction management software, presents significant benefits for advancing cybersecurity within the construction industry. It not only facilitates immediate access to cybersecurity support and information but also fosters a culture of security awareness and preparedness. Through continuous improvement and user engagement, the platform aims to establish a benchmark for cybersecurity practices within the industry, contributing to the safer execution of construction projects.

Bibliography

[1] R. Klinc and Ž. Turk, "Construction 4.0 – Digital Transformation of One of the Oldest Industries," *Economic and Business Review*, vol. 21, no. 3, pp. 393–496, Dec. 2019, doi: 10.15458/ebr.92.

[2] R. A. Vandegrift, *Construction 4.0: An Innovation Platform for the Built Environment*. New York: Routledge, 2020. doi: 10.4324/9781315697550-32.

[3] B. García de Soto, I. Agustí-Juan, S. Joss, and J. Hunhevicz, "Implications of Construction 4.0 to the Workforce and Organizational Structures," *International Journal of Construction Management*, vol. 22, no. 2, pp. 205–217, Jan. 2022, doi: 10.1080/15623599.2019.1616414.

[4] B. García de Soto, A. Georgescu, B. Mantha, Ž. Turk, and A. Maciel, "Construction Cybersecurity and Critical Infrastructure Protection: Significance, Overlaps, and Proposed Action Plan." May 12, 2020. doi: 10.20944/preprints202005.0213.v1.

[5] K. Alshammari, T. Beach, and Y. Rezgui, "Cybersecurity for Digital Twins in the Built Environment: Current Research and Future Directions," *Journal of Information Technology in Construction*, vol. 26, pp. 159–173, Apr. 2021, doi: 10.36680/j.itcon.2021.010.

[6] B. García de Soto, Ž. Turk, A. Maciel, B. Mantha, A. Georgescu, and M. S. Sonkor, "Understanding the Significance of Cybersecurity in the Construction Industry: Survey Findings," *Journal of Construction Engineering and Management*, vol. 148, no. 9, Sep. 2022, doi: 10.1061/(ASCE)CO.1943-7862.0002344.

[7] Deloitte, "Building Cybersecurity in the Construction Industry." Accessed: Sep. 30, 2023. [Online]. Available: https://www2.deloitte.com/ce/en/pages/real-estate/articles/ce-building-cybersecurity-in-the-construction-industry.html

[8] Cyware, "Hackers hit French firm Ingerop stealing 65 GB data relating to nuclear power plants," Cyware Hacker News. Accessed: Jul. 18, 2021. [Online]. Available: https://cyware.com/news/hackers-hit-french-firm-ingerop-stealing-65-gb-data-relating-to-nuclear-power-plants-f193b9ba/

[9] S. Coble, "Major Canadian Military Contractor Compromised in Ransomware Attack," Infosecurity Magazine. Accessed: Oct. 21, 2022. [Online]. Available: https://www.infosecurity-magazine.com/news/bird-construction-compromised-in/

[10] T. Sawyer and J. Rubenstone, "Construction Cybercrime Is On the Rise," Engineering News-Record (ENR). Accessed: Feb. 04, 2024. [Online]. Available:

https://www.enr.com/articles/46832-construction-cybercrime-is-on-the-rise

[11] R. Korman, "Bouygues Construction Unit Gradually Recovering After Ransomware Attack," Engineering News-Record (ENR). Accessed: Jul. 18, 2021. [Online]. Available: https://www.enr.com/articles/48637-bouygues-construction-unit-gradually-recovering-after-ransomware-attack

[12] D. Price, "Bam Construct and Interserve Hit by Cyber Attacks," Construction News. Accessed: Jul. 18, 2021. [Online]. Available: https://www.constructionnews.co.uk/contractors/bam-construct/bam-construct-hit-by-cyber-attack-13-05-2020/

[13] J. Margolin and I. Pereira, "Outdated Computer System Exploited in Florida Water Treatment Plant Hack," Abcnews. Accessed: Jul. 15, 2021. [Online]. Available: https://abcnews.go.com/US/outdated-computer-system-exploited-florida-water-treatment-plant/story?id=75805550

[14] W. Turton and K. Mehrotra, "Hackers Breached Colonial Pipeline Using Compromised Password," Bloomberg. Accessed: Aug. 10, 2021. [Online]. Available: https://www.bloomberg.com/news/articles/2021-06-04/hackers-breached-colonial-pipeline-using-compromised-password

[15] L. Steel, "Data Protection: Security Breach Results in £4.4m Fine for Interserve," Wright Hassall. Accessed: May 24, 2023. [Online]. Available: https://www.wrighthassall.co.uk/knowledge-base/data-protection-security-breach-results-in-4-4m-fine-for-interserve#:~:text=The%20Information%20Commissioner%27s%20Office%20%28ICO%29%20has%20issued%20a,affected%20due%20to%20a%20ransomware%20attack%20in%202020.

[16] The Stack, "Plasterboard Giant Knauf Group Pummelled by Ransomware," The Stack. Accessed: Jul. 21, 2023. [Online]. Available: https://www.thestack.technology/knauf-group-ransomware-attack-plasterboard-shortage/

[17] JDSUPRA, "Huntington Ingalls Industries Files Official Notice of Data Breach Affecting 43,643 Individuals," JDSUPRA. Accessed: Jan. 12, 2024. [Online]. Available: https://www.jdsupra.com/legalnews/huntington-ingalls-industries-files-3524071/

[18] P. Kunert, "US Construction Giant Unearths Concrete Evidence of Cyberattack," The Register. Accessed: Jan. 05, 2024. [Online]. Available: https://www.theregister.com/2023/10/12/simpson_manufacturing_security_incident/?td=readmore

[19] T. Aven, "Risk Assessment and Risk Management: Review of Recent Advances on Their Foundation," *European Journal of Operational Research*, vol. 253, no. 1, pp. 1–13, Aug. 2016, doi: 10.1016/j.ejor.2015.12.023.

[20] T. Meyer and G. Reniers, *Engineering Risk Management*. 2022. doi: 10.1515/9783110665338.

[21] G. Zhao, *Information Security Management and Risk Assessment*. Beijing: Tsinghua University Press, 2020.

[22] A. Ahmed, B. Kayis, and S. Amornsawadwatana, "A Review of Techniques for Risk Management in Projects," *Benchmarking: An International Journal*, vol. 14, no. 1, pp. 22–36, Mar. 2007, doi: 10.1108/14635770710730919.

[23] S. Iqbal, R. M. Choudhry, K. Holschemacher, A. Ali, and J. Tamošaitienė, "Risk Management in Construction Projects," *Technological and Economic Development of Economy*, vol. 21, no. 1, pp. 65–78, Jan. 2015, doi: 10.3846/20294913.2014.994582.

[24] National Institute of Standards and Technology, "The NIST Cybersecurity Framework (CSF) 2.0," National Institute of Standards and Technology, Gaithersburg, MD, NIST CSWP 29, Feb. 2024. doi: 10.6028/NIST.CSWP.29.

[25] European Union, "GDPR (General Data Protection Regulation)," European Union, 2018. Accessed: May 16, 2023. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:02016R0679-20160504

[26] ISO (International Organization for Standardization), "ISO/IEC 27000:2018 Information Technology — Security Techniques — Information Security Management Systems — Overview and Vocabulary." Accessed: Oct. 11, 2021. [Online]. Available: https://standards.iso.org/ittf/PubliclyAvailableStandards/c073906_ISO_IEC_27000_2018_E.zip

[27]CIS (Center for Internet Security), "Center for Internet Security Controls, Version 7.1." Accessed: Oct. 11, 2021. [Online]. Available: https://learn.cisecurity.org/20-controls-download?_gl=1*2ttlk*_ga*MjA0MDEzNDk4LjE2ODQyNTE4MDI.*_ga_N70Z2MKMD7*MTY4NDI1NDcwMS4yLjEuMTY4NDI1NDcxMy40OC4wLjA.*_ga_ZQVR7NM9HJ*MTY4NDI1NDcwMS4yLjEuMTY4NDI1NDcxMy4wLjAuMA..

[28] S. Radack, "The Common Vulnerability Scoring System (CVSS)," National Institute of Standards and Technology.

[29] N. Salami Pargoo and M. Ilbeigi, "A Scoping Review for Cybersecurity in the Construction Industry," *Journal of Management in Engineering*, vol. 39, no. 2, Mar. 2023, doi: 10.1061/JMENEA.MEENG-5034.

[30] A. Bello and A. Maurushat, "Technical and Behavioural Training and Awareness Solutions for Mitigating Ransomware Attacks," in *Advances in Intelligent Systems and Computing*, vol. 1226 AISC, 2020, pp. 164–176. doi: 10.1007/978-3-030-51974-2_14.

[31] S. El-Sayegh, L. Romdhane, and S. Manjikian, "A Critical Review of 3D Printing in Construction: Benefits, Challenges, and Risks," *Archives of Civil and Mechanical Engineering*, vol. 20, no. 2, p. 34, Jun. 2020, doi: 10.1007/s43452-020-00038-w.

[32] B. R. K. Mantha and B. García de Soto, "Cyber Security Challenges and Vulnerability Assessment in the Construction Industry," in *Proceedings of the Creative Construction Conference 2019*, Budapest University of Technology and Economics, 2019, pp. 29–37. doi: 10.3311/CCC2019-005.

[33] Ž. Turk, M. S. Sonkor, and R. Klinc, "Cybersecurity Assessment of BIM/CDE Design Environment Using Cyber Assessment Framework," *Journal of Civil Engineering and Management*, vol. 28, no. 5, pp. 349–364, May 2022, doi: 10.3846/jcem.2022.16682.

[34] Ž. Turk, B. García de Soto, B. R. K. Mantha, A. Maciel, and A. Georgescu, "A Systemic Framework for Addressing Cybersecurity in Construction," *Automation in Construction*, vol. 133, p. 103988, Jan. 2022, doi: 10.1016/j.autcon.2021.103988.

[35] E. A. Parn and D. Edwards, "Cyber Threats Confronting the Digital Built Environment: Common Data Environment Vulnerabilities and Blockchain Deterrence," *Engineering, Construction and Architectural Management*, vol. 26, no. 2, pp. 245–266, Mar. 2019, doi: 10.1108/ECAM-03-2018-0101.

[36] M. S. Sonkor and B. García de Soto, "Is Your Construction Site Secure? A View from the Cybersecurity Perspective," *Proceedings of the 38th International Symposium on Automation and Robotics in Construction (ISARC)*, no. ISARC, pp. 864–871, 2021, doi: 10.22260/isarc2021/0117.

[37] G. D. Goh, S. L. Sing, and W. Y. Yeong, "A Review on Machine Learning in 3D Printing: Applications, Potential, and Challenges," *Artificial Intelligence Review*, vol. 54, no. 1, pp. 63–94, Jan. 2021, doi: 10.1007/s10462-020-09876-9.

[38] M. Das, X. Tao, and J. C. P. Cheng, "BIM Security: A Critical Review and Recommendations Using Encryption Strategy and Blockchain," *Automation in Construction*, vol. 126, p. 103682, Jun. 2021, doi: 10.1016/j.autcon.2021.103682.

[39] J. Ackerson, R. Dave, and N. Seliya, "Applications of Recurrent Neural Networks for Biometric Authentication and Anomaly Detection," *Information*, vol. 12, no. 7, p. 272, Jul. 2021, doi: 10.3390/info12070272.

[40] T. Patel and V. Patel, "Data Privacy in Construction Industry by Privacy-Preserving Data Mining (ppdm) Approach," *Asian Journal of Civil Engineering*, vol. 21, no. 3, pp. 505–515, Apr. 2020, doi: 10.1007/s42107-020-00225-3.

[41] G. Shemov, B. García de Soto, and H. Alkhzaimi, "Blockchain Applied to the Construction Supply Chain: A Case Study with Threat Model," *Frontiers of Engineering Management*, vol. 7, no. 4, pp. 564–577, Dec. 2020, doi: 10.1007/s42524-020-0129-x.

[42] N. Nawari and S. Ravindran, "Blockchain and Building Information Modeling (BIM): Review and Applications in Post-Disaster Recovery," *Buildings*, vol. 9, no. 6, p. 149, Jun. 2019, doi: 10.3390/buildings9060149.

[43] N. O. Nawari and S. Ravindran, "Blockchain and the Built Environment: Potentials and Limitations," *Journal of Building Engineering*, vol. 25, p. 100832, Sep. 2019, doi: 10.1016/j.jobe.2019.100832.

[44] M. Shi, A. Hoffmann, A. Wagner, T. Huyeng, C. D. Thiele, and U. Rüppel, "Using Blockchain Technology to Implement Peer-to-Peer Network in Construction Industry," in *Lecture Notes in Civil Engineering*, vol. 98, 2021. doi: 10.1007/978-3-030-51295-8_58.

[45] X. Tao, M. Das, Y. Liu, and J. C. P. Cheng, "Distributed Common Data Environment Using Blockchain and Interplanetary File System for Secure BIM-Based Collaborative Design," *Automation in Construction*, vol. 130, p. 103851, Oct. 2021, doi: 10.1016/j.autcon.2021.103851.

[46] A. Sheikh, V. Kamuni, A. Patil, S. Wagh, and N. Singh, "Cyber Attack and Fault Identification of HVAC System in Building Management Systems," in *2019 9th International Conference on Power and Energy Systems (ICPES)*, Perth, Australia: IEEE, Dec. 2019, pp. 1–6. doi: 10.1109/ICPES47639.2019.9105438.

[47] Z. Pan, S. Hariri, and J. Pacheco, "Context Aware Intrusion Detection for Building Automation Systems," *Computers & Security*, vol. 85, pp. 181–201, Aug. 2019, doi: 10.1016/j.cose.2019.04.011.

[48] N. Skandhakumar, F. Salim, J. Reid, R. Drogemuller, and E. Dawson, "Graph Theory-Based Representation of Building Information Models for Access Control Applications," *Automation in Construction*, vol. 68, pp. 44–51, Aug. 2016, doi: 10.1016/j.autcon.2016.04.001.

[49] M. U. R. Mohamed Shibly and B. García de Soto, "Threat Modeling in Construction: An Example of a 3D Concrete Printing System," presented at the 37th International Symposium on Automation and Robotics in Construction, Kitakyushu, Japan, Oct.

2020. doi: 10.22260/ISARC2020/0087.

[50] B. R. K. Mantha, Y. Jung, and B. García de Soto, “Implementation of the Common Vulnerability Scoring System to Assess the Cyber Vulnerability in Construction Projects,” in *Creative Construction e-Conference 2020*, Budapest, Hungary: Budapest University of Technology and Economics, 2020, pp. 117–124. doi: 10.3311/ccc2020-030.

[51] B. R. K. Mantha and B. García de Soto, “Assessment of The Cybersecurity Vulnerability of Construction Networks,” *Engineering, Construction and Architectural Management*, vol. 28, no. 10, pp. 3078–3105, Nov. 2021, doi: 10.1108/ECAM-06-2020-0400.

[52] E. Alpaydın, *Introduction to Machine Learning*, Fourth edition. in Adaptive computation and machine learning. Cambridge, Massachusetts London: The MIT Press, 2020.

[53] S. Ruder, “An Overview of Gradient Descent Optimization Algorithms.” arXiv, Jun. 15, 2017. Accessed: Apr. 12, 2024. [Online]. Available: http://arxiv.org/abs/1609.04747

[54] D. Lowd and P. Domingos, “Naive Bayes Models for Probability Estimation,” in *Proceedings of the 22nd international conference on Machine learning - ICML ’05*, Bonn, Germany: ACM Press, 2005, pp. 529–536. doi: 10.1145/1102351.1102418.

[55] L. Breiman, “Random Forests,” *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001, doi: 10.1023/A:1010933404324.

[56] N. Lange, C. M. Bishop, and B. D. Ripley, “Neural Networks for Pattern Recognition.,” *Journal of the American Statistical Association*, vol. 92, no. 440, 1997, doi: 10.2307/2965437.

[57] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Erscheinungsort nicht ermittelbar: Alanna Maldonado, 2023.

[58] V. Nair and G. E. Hinton, “Rectified Linear Units Improve Restricted Boltzmann Machines,” in *ICML 2010 - Proceedings, 27th International Conference on Machine Learning*, 2010.

[59] A. L. Maas, A. Y. Hannun, and A. Y. Ng, “Rectifier Nonlinearities Improve Neural Network Acoustic Models,” in *in ICML Workshop on Deep Learning for Audio, Speech and Language Processing*, 2013.

[60] C. Nwankpa, W. Ijomah, A. Gachagan, and S. Marshall, “Activation Functions: Comparison of trends in Practice and Research for Deep Learning.” arXiv, 2018. doi:

10.48550/ARXIV.1811.03378.

[61] H. Pratiwi *et al.*, "Sigmoid Activation Function in Selecting the Best Model of Artificial Neural Networks," *Journal of Physics: Conference Series*, vol. 1471, no. 1, p. 012010, Feb. 2020, doi: 10.1088/1742-6596/1471/1/012010.

[62] A. Vaswani *et al.*, "Attention Is All You Need." arXiv, Aug. 01, 2023. Accessed: Mar. 02, 2024. [Online]. Available: http://arxiv.org/abs/1706.03762

[63] J. L. Ba, J. R. Kiros, and G. E. Hinton, "Layer Normalization." arXiv, Jul. 21, 2016. Accessed: Apr. 12, 2024. [Online]. Available: http://arxiv.org/abs/1607.06450

[64] J. Terven, D. M. Cordova-Esparza, A. Ramirez-Pedraza, and E. A. Chavez-Urbiola, "Loss Functions and Metrics in Deep Learning," 2023, doi: 10.48550/ARXIV.2307.02694.

[65] S. Aroca-Ouellette and F. Rudzicz, "On Losses for Modern Language Models." arXiv, 2020. doi: 10.48550/ARXIV.2010.01694.

[66] H. Zou and T. Hastie, "Regularization and Variable Selection Via the Elastic Net," *Journal of the Royal Statistical Society Series B: Statistical Methodology*, vol. 67, no. 2, pp. 301–320, Apr. 2005, doi: 10.1111/j.1467-9868.2005.00503.x.

[67] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal Loss for Dense Object Detection." arXiv, 2017. doi: 10.48550/ARXIV.1708.02002.

[68] F. Schroff, D. Kalenichenko, and J. Philbin, "FaceNet: A Unified Embedding for Face Recognition and Clustering," in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, MA, USA: IEEE, Jun. 2015, pp. 815–823. doi: 10.1109/CVPR.2015.7298682.

[69] I. J. Goodfellow *et al.*, "Generative Adversarial Networks." arXiv, 2014. doi: 10.48550/ARXIV.1406.2661.

[70] L. Ouyang *et al.*, "Training Language Models to Follow Instructions with Human Feedback." arXiv, Mar. 04, 2022. Accessed: Mar. 02, 2024. [Online]. Available: http://arxiv.org/abs/2203.02155

[71] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization." arXiv, 2014. doi: 10.48550/ARXIV.1412.6980.

[72] I. Loshchilov and F. Hutter, "Decoupled Weight Decay Regularization." arXiv, 2017. doi: 10.48550/ARXIV.1711.05101.

[73] B. V. Barde and A. M. Bainwad, "An Overview of Topic Modeling Methods and

Tools," in *2017 International Conference on Intelligent Computing and Control Systems (ICICCS)*, Madurai: IEEE, Jun. 2017, pp. 745–750. doi: 10.1109/ICCONS.2017.8250563.

[74] H. Jelodar *et al.*, "Latent Dirichlet Allocation (LDA) and Topic Modeling: Models, Applications, a Survey," *Multimedia Tools and Applications*, vol. 78, no. 11, pp. 15169–15211, Jun. 2019, doi: 10.1007/s11042-018-6894-4.

[75] D. M. Blei, A. Y. Ng, and J. B. Edu, "Latent Dirichlet Allocation," *Journal of Machine Learning Research*, vol. 3, pp. 993–1022, 2003.

[76] S. Zhou, P. Kan, Q. Huang, and J. Silbernagel, "A Guided Latent Dirichlet Allocation Approach to Investigate Real-Time Latent Topics of Twitter Data During Hurricane Laura," *Journal of Information Science*, vol. 49, no. 2, pp. 465–479, Apr. 2023, doi: 10.1177/01655515211007724.

[77] B. Tsolmon and K.-S. Lee, "An Event Extraction Model Based on Timeline and User Analysis in Latent Dirichlet Allocation," in *Proceedings of the 37th International ACM SIGIR Conference on Research & Development in Information Retrieval*, Gold Coast Queensland Australia: ACM, Jul. 2014, pp. 1187–1190. doi: 10.1145/2600428.2609541.

[78] R. de Groof, H. Xu, J. Zhang, and R. Liu, "Mining Significant Terminologies in Online Social Media Using Parallelized LDA for the Promotion of Cultural Products," in *Proceedings of the 14th International Conference on Data Science (ICDATA'18)*, Las Vegas, NV USA, Jul. 2018, pp. 3–9. Accessed: Apr. 06, 2024. [Online]. Available: http://www.cis.umassd.edu/~hxu/Papers/UMD/2009/CP/ICDATA-2018-XU.pdf

[79] G. Maskeri, S. Sarkar, and K. Heafield, "Mining Business Topics in Source Code Using Latent Dirichlet Allocation," in *Proceedings of the 1st India Software Engineering Conference*, Hyderabad India: ACM, Feb. 2008, pp. 113–120. doi: 10.1145/1342211.1342234.

[80] E. Linstead, C. Lopes, and P. Baldi, "An Application of Latent Dirichlet Allocation to Analyzing Software Evolution," in *2008 Seventh International Conference on Machine Learning and Applications*, San Diego, CA, USA: IEEE, 2008, pp. 813–818. doi: 10.1109/ICMLA.2008.47.

[81] H. U. Asuncion, A. U. Asuncion, and R. N. Taylor, "Software Traceability with Topic Modeling," in *Proceedings of the 32nd ACM/IEEE International Conference on Software Engineering - Volume 1*, Cape Town South Africa: ACM, May 2010, pp. 95–104. doi: 10.1145/1806799.1806817.

[82] A. Sameer El Khatib, "Machine Learning and Finance: A Review Using Latent

Dirichlet Allocation Technique (LDA)," *SSRN Journal*, 2020, doi: 10.2139/ssrn.3730256.

[83] S. Aziz, M. Dowling, H. Hammami, and A. Piepenbrink, "Machine Learning in Finance: A Topic Modeling Approach," *European Financial Management*, vol. 28, no. 3, pp. 744–770, Jun. 2022, doi: 10.1111/eufm.12326.

[84] S. Feuerriegel and N. Pröllochs, "Investor Reaction to Financial Disclosures Across Topics: An Application of Latent Dirichlet Allocation." arXiv, 2018. doi: 10.48550/ARXIV.1805.03308.

[85] Q. Yang, Z. Xu, W. Zhou, P. Wang, Q. Jiang, and L. Juan, "An Interpretable Single-Cell RNA Sequencing Data Clustering Method Based on Latent Dirichlet Allocation," *Briefings in Bioinformatics*, vol. 24, no. 4, p. bbad199, Jul. 2023, doi: 10.1093/bib/bbad199.

[86] S. Shivashankar, S. Srivathsan, B. Ravindran, and A. V. Tendulkar, "Multi-View Methods for Protein Structure Comparison Using Latent Dirichlet Allocation," *Bioinformatics*, vol. 27, no. 13, pp. i61–i68, Jul. 2011, doi: 10.1093/bioinformatics/btr249.

[87] P. Pinoli, D. Chicco, and M. Masseroli, "Latent Dirichlet Allocation Based on Gibbs Sampling for Gene Function Prediction," in *2014 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology*, Honolulu, HI, USA: IEEE, May 2014, pp. 1–8. doi: 10.1109/CIBCB.2014.6845514.

[88] D. Li *et al.*, "KTI-RNN: Recognition of Heart Failure from Clinical Notes," *Tsinghua Science and Technology*, vol. 28, no. 1, pp. 117–130, Feb. 2023, doi: 10.26599/TST.2021.9010093.

[89] C. Khalil *et al.*, "Patients' Perspectives, Experiences, and Concerns with Crohn's Perianal Fistulae: Insights from Social Media Platforms," *Inflammatory Bowel Diseases*, vol. 28, no. Supplement_1, pp. S92–S93, Jan. 2022, doi: 10.1093/ibd/izac015.149.

[90] J. M. Frick, R. Guha, T. Peryea, and N. T. Southall, "Evaluating Disease Similarity Using Latent Dirichlet Allocation." arXiv, Nov. 04, 2015. doi: 10.1101/030593.

[91] Y. Liu, J. Wang, S. Tang, J. Zhang, and J. Wan, "Integrating Information Entropy and Latent Dirichlet Allocation Models for Analysis of Safety Accidents in the Construction Industry," *Buildings*, vol. 13, no. 7, p. 1831, Jul. 2023, doi: 10.3390/buildings13071831.

[92] K. Zhou, J. Wang, B. Ashuri, and J. Chen, "Discovering the Research Topics on Construction Safety and Health Using Semi-Supervised Topic Modeling," *Buildings*,

vol. 13, no. 5, p. 1169, Apr. 2023, doi: 10.3390/buildings13051169.

[93] B. Zhong, X. Pan, P. E. D. Love, L. Ding, and W. Fang, "Deep Learning and Network Analysis: Classifying and Visualizing Accident Narratives in Construction," *Automation in Construction*, vol. 113, p. 103089, May 2020, doi: 10.1016/j.autcon.2020.103089.

[94] L. Zeng, R. Y. M. Li, T. Yigitcanlar, and H. Zeng, "Public Opinion Mining on Construction Health and Safety: Latent Dirichlet Allocation Approach," *Buildings*, vol. 13, no. 4, p. 927, Mar. 2023, doi: 10.3390/buildings13040927.

[95] D. Yao, "Data for Topic Modeling." Accessed: Mar. 16, 2023. [Online]. Available: https://github.com/Daniel-Yao-Chengdu/NLP-project/blob/master/Data%20for%20topic%20modeling

[96] D. Yao, "mysentences.py," GitHub Repository. Accessed: Jan. 12, 2024. [Online]. Available: https://github.com/Dongchi-Yao/Project/blob/master/mysentences.py

[97] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding." arXiv, May 24, 2019. Accessed: Mar. 02, 2024. [Online]. Available: http://arxiv.org/abs/1810.04805

[98] S. M. Jain, "Hugging Face," in *Introduction to Transformers for NLP*, Berkeley, CA: Apress, 2022, pp. 51–67. doi: 10.1007/978-1-4842-8844-3_4.

[99] B. Giuseppe, *Machine Learning Algorithms*. Packt Publishing Ltd, 2017.

[100] O. Kramer, "Scikit-Learn," Oldenburg: Springer, Cham, 2016, pp. 45–53. doi: 10.1007/978-3-319-33383-0_5.

[101] D. Yao, "myRaw2Lemmatized," GitHub repository. Accessed: Jan. 13, 2023. [Online]. Available: https://github.com/Daniel-Yao-Chengdu/NLP-project/blob/master/myRaw2Lemmatized.py

[102] A. Hosseiny Marani and E. P. S. Baumer, "A Review of Stability in Topic Modeling: Metrics for Assessing and Techniques for Improving Stability," *ACM Computing Surveys*, vol. 56, no. 5, pp. 1–32, May 2024, doi: 10.1145/3623269.

[103] M. Röder, A. Both, and A. Hinneburg, "Exploring the Space of Topic Coherence Measures," in *Proceedings of the Eighth ACM International Conference on Web Search and Data Mining*, Shanghai China: ACM, Feb. 2015, pp. 399–408. doi: 10.1145/2684822.2685324.

[104] D. Yao, "Topic Modeling," GitHub repository. Accessed: Jan. 08, 2023. [Online]. Available: https://github.com/Daniel-Yao-Chengdu/NLP-

project/blob/master/Topic%20modelling/topic_modeling.py

[105] OpenAI, "ChatGPT: Optimizing Language Models for Dialogue," OpenAI. Accessed: Jan. 08, 2023. [Online]. Available: https://openai.com/blog/chatgpt/

[106] A. Parks, "The Ongoing Risk of Phishing in the Construction Industry," Construction Management Association of America, 2021. Accessed: Mar. 15, 2022. [Online]. Available: https://resources.infosecinstitute.com/topic/phishing-attacks-construction-industry/

[107] C. Tunney, "Ransomware Attack on Construction Company Raises Questions About Federal Contracts," CBC News. Accessed: Mar. 15, 2021. [Online]. Available: https://www.cbc.ca/news/politics/ransomware-bird-construction-military-1.5434308

[108] R. Korman, "Hoffman Construction Reports Hack of Self-Insured Health Plan Data," Engineering News-Record. Accessed: Mar. 15, 2021. [Online]. Available: https://www.enr.com/articles/51232-hoffman-construction-reports-hack-of-self-insured-health-plan-data

[109] B. Mantha, B. García de Soto, and R. Karri, "Cyber Security Threat Modeling in the AEC Industry: An Example for the Commissioning of the Built Environment," *Sustainable Cities and Society*, vol. 66, p. 102682, Mar. 2021, doi: 10.1016/j.scs.2020.102682.

[110] B. Zhong, X. Pan, P. E. D. Love, J. Sun, and C. Tao, "Hazard Analysis: A Deep Learning and Text Mining Framework for Accident Prevention," *Advanced Engineering Informatics*, vol. 46, p. 101152, Oct. 2020, doi: 10.1016/j.aei.2020.101152.

[111] A. Bittar, S. Velupillai, A. Roberts, and R. Dutta, "Using General-purpose Sentiment Lexicons for Suicide Risk Assessment in Electronic Health Records: Corpus-Based Analysis," *JMIR Medical Informatics*, vol. 9, no. 4, p. e22397, Apr. 2021, doi: 10.2196/22397.

[112] M. Boholm and Å. Boholm, "Risk Identification: A Corpus-Assisted Study of Websites of Government Agencies," *Risk, Hazards & Crisis in Public Policy*, vol. 11, no. 3, pp. 242–269, Sep. 2020, doi: 10.1002/rhc3.12184.

[113] D. Yao and B. García de Soto, "A Corpus Database for Cybersecurity Topic Modeling in the Construction Industry," presented at the 40th International Symposium on Automation and Robotics in Construction, Chennai, India, Jul. 2023. doi: 10.22260/ISARC2023/0072.

[114] Baidu Inc., "Introducing ERNIE 3.5: Baidu's Knowledge-Enhanced Foundation Model Takes a Giant Leap Forward," Baidu Research. Accessed: Nov. 28, 2023.

[Online]. Available: http://research.baidu.com/Blog/index-view?id=185

[115] OpenAI *et al.*, "GPT-4 Technical Report." arXiv, Dec. 18, 2023. Accessed: Mar. 02, 2024. [Online]. Available: http://arxiv.org/abs/2303.08774

[116] Gemini Team *et al.*, "Gemini: A Family of Highly Capable Multimodal Models." arXiv, Dec. 18, 2023. Accessed: Mar. 02, 2024. [Online]. Available: http://arxiv.org/abs/2312.11805

[117] W. X. Zhao *et al.*, "A Survey of Large Language Models." arXiv, Nov. 24, 2023. Accessed: Mar. 02, 2024. [Online]. Available: http://arxiv.org/abs/2303.18223

[118] D. Jurafsky and J. H. Martin, *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*, 2. ed. [Nachdr.]. Upper Saddle River, NJ: Prentice Hall, 2009.

[119] A. Radford, K. Narasimhan, T. Salimans, and I. Sutskever, "Improving Language Understanding by Generative Pre-Training," OpenAI blog. Accessed: Mar. 02, 2024. [Online]. Available: https://openai.com/research/language-unsupervised

[120] Radford Alec, Wu Jeffrey, Child Rewon, Luan David, Amodei Dario, and Sutskever Ilya, "Language Models are Unsupervised Multitask Learners," OpenAI blog. Accessed: Mar. 02, 2024. [Online]. Available: https://cdn.openai.com/better-language-models/language_models_are_unsupervised_multitask_learners.pdf

[121] T. B. Brown *et al.*, "Language Models are Few-Shot Learners." arXiv, Jul. 22, 2020. Accessed: Mar. 02, 2024. [Online]. Available: http://arxiv.org/abs/2005.14165

[122] OpenAI, "Introducing ChatGPT," OpenAI. Accessed: Mar. 02, 2024. [Online]. Available: https://openai.com/blog/chatgpt

[123] A. Chowdhery *et al.*, "PaLM: Scaling Language Modeling with Pathways." arXiv, Oct. 05, 2022. Accessed: Mar. 02, 2024. [Online]. Available: http://arxiv.org/abs/2204.02311

[124] R. Thoppilan *et al.*, "LaMDA: Language Models for Dialog Applications." arXiv, Feb. 10, 2022. Accessed: Mar. 02, 2024. [Online]. Available: http://arxiv.org/abs/2201.08239

[125] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal Policy Optimization Algorithms." arXiv, Aug. 28, 2017. Accessed: Mar. 02, 2024. [Online]. Available: http://arxiv.org/abs/1707.06347

[126] R. S. Sutton, D. McAllester, S. Singh, and Y. Mansour, "Policy Gradient Methods for Reinforcement Learning with Function Approximation," in *Advances in Neural*

Information Processing Systems 12 (NIPS 1999), 1999, pp. 1057–1063.

[127] J. W. Rae *et al.*, "Scaling Language Models: Methods, Analysis & Insights from Training Gopher." arXiv, Jan. 21, 2022. Accessed: Mar. 02, 2024. [Online]. Available: http://arxiv.org/abs/2112.11446

[128] Y. Bai *et al.*, "Training a Helpful and Harmless Assistant with Reinforcement Learning from Human Feedback." arXiv, Apr. 12, 2022. Accessed: Mar. 02, 2024. [Online]. Available: http://arxiv.org/abs/2204.05862

[129] B. Workshop *et al.*, "BLOOM: A 176B-Parameter Open-Access Multilingual Language Model." arXiv, Jun. 27, 2023. Accessed: Mar. 02, 2024. [Online]. Available: http://arxiv.org/abs/2211.05100

[130] D. Yao, "SFT Training Dataset," GitHub Repository. Accessed: Apr. 28, 2023. [Online]. Available: https://github.com/Daniel-Yao-Chengdu/NLP-project/blob/master/paper-A%20Cybersecurity-focused%20Large%20Language%20Model%20for%20the%20Construction%20Industry:%20A%20Case%20Study%20on%20Risk%20Identification/SFT%20training%20dataset.pt

[131] D. Yao, "Reward Model Training Dataset," GitHub Repository. Accessed: Apr. 28, 2023. [Online]. Available: https://github.com/Daniel-Yao-Chengdu/NLP-project/blob/master/paper-A%20Cybersecurity-focused%20Large%20Language%20Model%20for%20the%20Construction%20Industry:%20A%20Case%20Study%20on%20Risk%20Identification/Reward%20model%20training%20dataset.pt

[132] D. Yao, "RL Fine-tuning Dataset," GitHub Repository. Accessed: Apr. 28, 2023. [Online]. Available: https://github.com/Daniel-Yao-Chengdu/NLP-project/blob/master/paper-A%20Cybersecurity-focused%20Large%20Language%20Model%20for%20the%20Construction%20Industry:%20A%20Case%20Study%20on%20Risk%20Identification/RL%20fine-tuning%20dataset.pt

[133] C. Raffel *et al.*, "Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer." arXiv, Sep. 19, 2023. Accessed: Mar. 02, 2024. [Online]. Available: http://arxiv.org/abs/1910.10683

[134] H. Zhang *et al.*, "Fine-tuning Pre-trained Language Models for Few-shot Intent Detection: Supervised Pre-training and Isotropization." arXiv, May 26, 2022. Accessed: Mar. 02, 2024. [Online]. Available: http://arxiv.org/abs/2205.07208

[135] Y. Yu, S. Zuo, H. Jiang, W. Ren, T. Zhao, and C. Zhang, "Fine-Tuning Pre-trained

Language Model with Weak Supervision: A Contrastive-Regularized Self-Training Approach." arXiv, Mar. 30, 2021. Accessed: Mar. 02, 2024. [Online]. Available: http://arxiv.org/abs/2010.07835

[136] J. Shlens, "Notes on Kullback-Leibler Divergence and Likelihood." [object Object], 2014. doi: 10.48550/ARXIV.1404.2000.

[137] B. Lowerre and R. Reddy, "The Harpy Speech Recognition System: Performance with Large Vocabularies," *The Journal of the Acoustical Society of America*, vol. 60, no. S1, pp. S10–S11, Nov. 1976, doi: 10.1121/1.2003089.

[138] C. Laorden, B. Sanz, G. Alvarez, and P. G. Bringas, "A Threat Model Approach to Threats and Vulnerabilities in On-line Social Networks," in *Computational Intelligence in Security for Information Systems 2010*, vol. 85, Á. Herrero, E. Corchado, C. Redondo, and Á. Alonso, Eds., in Advances in Intelligent and Soft Computing, vol. 85. , Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 135–142. doi: 10.1007/978-3-642-16626-6_15.

[139] R. Likert, "A Technique for the Measurement of Attitudes," *Archives of Psychology*, vol. 22, no. 140, p. 55.

[140] D. Banerjee Chattapadhyay, J. Putta, and R. M. Rao P, "Risk Identification, Assessments, and Prediction for Mega Construction Projects: A Risk Prediction Paradigm Based on Cross Analytical-Machine Learning Model," *Buildings*, vol. 11, no. 4, p. 172, Apr. 2021, doi: 10.3390/buildings11040172.

[141] A. Gondia, A. Siam, W. El-Dakhakhni, and A. H. Nassar, "Machine Learning Algorithms for Construction Projects Delay Risk Prediction," *Journal of Construction Engineering and Management*, vol. 146, no. 1, p. 04019085, Jan. 2020, doi: 10.1061/(ASCE)CO.1943-7862.0001736.

[142] N. Xia, R. Zhong, C. Wu, X. Wang, and S. Wang, "Assessment of Stakeholder-Related Risks in Construction Projects: Integrated Analyses of Risk Attributes and Stakeholder Influences," *Journal of Construction Engineering and Management*, vol. 143, no. 8, p. 04017030, Aug. 2017, doi: 10.1061/(ASCE)CO.1943-7862.0001322.

[143] University of Defense, Military Aacademy, Belgrade, N. Kovačević, A. Stojiljković, University of Novi Sad, Faculty of Economics, Subotica, M. Kovač, and "Educons" University, Faculty for Project and Innovation Management, Belgrade, "Application of the Matrix Approach in Risk Assessment," *Operational Research in Engineering Sciences: Theory and Applications*, vol. 2, no. 3, pp. 55–64, Dec. 2019, doi: 10.31181/oresta1903055k.

[144] Accenture Security, "2020 Cyber Threatscape Report I," Accenture, Feb. 2020.

Accessed: Mar. 01, 2024. [Online]. Available: https://www.accenture.com/content/dam/accenture/final/capabilities/technology/security/document/11177%20Cyber%20Threatscape%20Report_Digital_AW_SH.pdf

[145] C. Cop, "Phishing Email Detection." Accessed: Feb. 10, 2024. [Online]. Available: https://www.kaggle.com/datasets/subhajournal/phishingemails/data

[146] T. Zhang, V. Kishore, F. Wu, K. Q. Weinberger, and Y. Artzi, "BERTScore: Evaluating Text Generation with BERT." arXiv, Feb. 24, 2020. Accessed: Mar. 02, 2024. [Online]. Available: http://arxiv.org/abs/1904.09675

[147] D. Deutsch and D. Roth, "Understanding the Extent to which Content Quality Metrics Measure the Information Quality of Summaries," in *Proceedings of the 25th Conference on Computational Natural Language Learning*, Online: Association for Computational Linguistics, 2021, pp. 300–309. doi: 10.18653/v1/2021.conll-1.24.

[148] B. R. Lashley and C. F. Bond, "Significance Testing for Round Robin Data," *Psychological Methods*, vol. 2, no. 3, pp. 278–291, Sep. 1997, doi: 10.1037/1082-989X.2.3.278.

[149] M. Friedman, "The Use of Ranks to Avoid the Assumption of Normality Implicit in the Analysis of Variance," *Journal of the American Statistical Association*, vol. 32, no. 200, pp. 675–701, Dec. 1937, doi: 10.1080/01621459.1937.10503522.

[150] F. Wilcoxon, "Individual Comparisons by Ranking Methods," *Biometrics Bulletin*, vol. 1, no. 6, p. 80, Dec. 1945, doi: 10.2307/3001968.

[151] C. Spearman, "The Proof and Measurement of Association between Two Things," *The American Journal of Psychology*, vol. 100, no. 3/4, p. 441, 1987, doi: 10.2307/1422689.

[152] D. Yao, B. García De Soto, and M. Wilkes, "Identifying Cyber Risk Factors Associated with Construction Projects," *SSRN Journal*, 2023, doi: 10.2139/ssrn.4648243.

[153] S. Kurtz, "Cybersecurity Vulnerabilities in the Construction Industry," Total IT Information Technology. Accessed: Mar. 19, 2021. [Online]. Available: https://totalit.com/cybersecurity-vulnerabilities-in-the-construction-industry/

[154] K. Nyamuchiwa, Z. Lei, and C. Aranas, "Cybersecurity Vulnerabilities in Off-Site Construction," *Applied Sciences (Switzerland)*, vol. 12, no. 10, May 2022, doi: 10.3390/app12105037.

[155] S. Sharma and P. K. Goyal, "Fuzzy Assessment of the Risk Factors Causing Cost Overrun in the Construction Industry," *Evolutionary Intelligence*, vol. 15, no. 4, pp.

2269–2281, Dec. 2022, doi: 10.1007/s12065-019-00214-9.

[156] D. Baloi and A. D. F. Price, "Modelling Global Risk Factors Affecting Construction Cost Performance," *International Journal of Project Management*, vol. 21, no. 4, 2003, doi: 10.1016/S0263-7863(02)00017-0.

[157] M. S. B. A. Abd El-Karim, O. A. Mosa El Nawawy, and A. M. Abdel-Alim, "Identification and Assessment of Risk Factors Affecting Construction Projects," *HBRC Journal*, vol. 13, no. 2, pp. 202–216, Aug. 2017, doi: 10.1016/j.hbrcj.2015.05.001.

[158] N. Chileshe and A. Boadua Yirenkyi-Fianko, "An Evaluation of Risk Factors Impacting Construction Projects in Ghana," *Journal of Engineering, Design and Technology*, vol. 10, no. 3, pp. 306–329, Oct. 2012, doi: 10.1108/17260531211274693.

[159] B. G. Hwang, M. Shan, H. Phua, and S. Chi, "An Exploratory Analysis of Risks in Green Residential Building Construction Projects: The Case of Singapore," *Sustainability (Switzerland)*, vol. 9, no. 7, 2017, doi: 10.3390/su9071116.

[160] P. Aghaei, G. Asadollahfardi, and A. Katabi, "Safety Risk Assessment in Shopping Center Construction Projects Using Fuzzy Fault Tree Analysis Method," *Quality and Quantity*, vol. 56, no. 1, 2022, doi: 10.1007/s11135-021-01115-9.

[161] C. A. Rudolf and S. Spinler, "Key Risks in the Supply Chain of Large-Scale Engineering and Construction Projects," *Supply Chain Management*, vol. 23, no. 4, 2018, doi: 10.1108/SCM-09-2017-0292.

[162] D. W. M. Chan, A. P. C. Chan, P. T. I. Lam, J. F. Y. Yeung, and J. H. L. Chan, "Risk Ranking and Analysis in Target Cost Contracts: Empirical Evidence from the Construction Industry," *International Journal of Project Management*, vol. 29, no. 6, 2011, doi: 10.1016/j.ijproman.2010.08.003.

[163] C. Sun, H. Xu, and S. Jiang, "Understanding the Risk Factors of BIM Technology Implementation in the Construction Industry: An Interpretive Structural Modeling (ISM) Approach," *Engineering, Construction and Architectural Management*, vol. 27, no. 10, 2020, doi: 10.1108/ECAM-09-2019-0508.

[164] K.-F. Chien, Z.-H. Wu, and S.-C. Huang, "Identifying and assessing critical risk factors for BIM projects: Empirical study," *Automation in Construction*, vol. 45, pp. 1–15, Sep. 2014, doi: 10.1016/j.autcon.2014.04.012.

[165] H. Al Zubaidi and S. Al Otaibi, "An Empirical Approach for Identifying Critical Time-Ovrerun Risk Factors in Kuwait's Construction Projects," *Journal of Economic and Administrative Sciences*, vol. 24, no. 2, pp. 35–53, Dec. 2008, doi:

10.1108/10264116200800007.

[166] S. A. Assaf and S. Al-Hejji, “Causes of delay in large construction projects,” *International Journal of Project Management*, vol. 24, no. 4, pp. 349–357, May 2006, doi: 10.1016/j.ijproman.2005.11.010.

[167] P. Rezakhani, “Classifying Key Risk Factors in Construction Projects,” *The Bulletin of the Polytechnic Institute of Jassy, Construction and Architecture Section*, vol. 62, no. 2, 2012.

[168] P. X. W. Zou, G. Zhang, and J. Wang, “Understanding the key risks in construction projects in China,” *International Journal of Project Management*, vol. 25, no. 6, pp. 601–614, Aug. 2007, doi: 10.1016/j.ijproman.2007.03.001.

[169] A. M. Jarkas and T. C. Haupt, “Major construction risk factors considered by general contractors in qatar,” *Journal of Engineering, Design and Technology*, vol. 13, no. 1, 2015, doi: 10.1108/JEDT-03-2014-0012.

[170] S. M. Renuka, C. Umarani, and S. Kamal, “A Review on Critical Risk Factors in the Life Cycle of Construction Projects,” *Journal of Civil Engineering Research*, vol. 4, no. 2A, 2014.

[171] I. Y. Wuni, G. Q. P. Shen, and A. T. Mahmud, “Critical risk factors in the application of modular integrated construction: a systematic review,” *International Journal of Construction Management*, vol. 22, no. 2, 2022, doi: 10.1080/15623599.2019.1613212.

[172] P. X. W. Zou and G. Zhang, “Managing Risks in Construction Projects: Life Cycle and Stakeholder Perspectives,” *International Journal of Construction Management*, vol. 9, no. 1, 2009, doi: 10.1080/15623599.2009.10773122.

[173] Mahmoud Mohamed Mahmoud Sharaf and Hassan T. Abdelwahab, “Analysis of Risk Factors for Highway Construction Projects in Egypt,” *Journal of Civil Engineering and Architecture*, vol. 9, no. 5, 2015, doi: 10.17265/1934-7359/2015.05.004.

[174] M. S. Sonkor and B. García de Soto, “Operational Technology on Construction Sites: A Review from the Cybersecurity Perspective,” *Journal of Construction Engineering and Management*, vol. 147, no. 12, p. 04021172, Dec. 2021, doi: 10.1061/(ASCE)CO.1943-7862.0002193.

[175] M. Kalinin, V. Krundyshev, and P. Zegzhda, “Cybersecurity Risk Assessment in Smart City Infrastructures,” *Machines*, vol. 9, no. 4, p. 78, Apr. 2021, doi: 10.3390/machines9040078.

[176] M. O. Sanni-Anibire, R. M. Zin, and S. O. Olatunji, "Machine Learning Model for Delay Risk Assessment in Tall Building Projects," *International Journal of Construction Management*, vol. 22, no. 11, pp. 2134–2143, Aug. 2022, doi: 10.1080/15623599.2020.1768326.

[177] J. P. Fitzsimmons, R. Lu, Y. Hong, and I. Brilakis, "Construction Schedule Risk Analysis – A Hybrid Machine Learning Approach," *ITcon*, vol. 27, pp. 70–93, Jan. 2022, doi: 10.36680/j.itcon.2022.004.

[178] M. R. George, M. R. Nalluri, and K. B. Anand, "Application of Ensemble Machine Learning for Construction Safety Risk Assessment," *J. Inst. Eng. India Ser. A*, vol. 103, no. 4, pp. 989–1003, Dec. 2022, doi: 10.1007/s40030-022-00690-w.

[179] H. Liu and G. Tian, "Building Engineering Safety Risk Assessment and Early Warning Mechanism Construction Based on Distributed Machine Learning Algorithm," *Safety Science*, vol. 120, pp. 764–771, Dec. 2019, doi: 10.1016/j.ssci.2019.08.022.

[180] C. Q. X. Poh, C. U. Ubeynarayana, and Y. M. Goh, "Safety Leading Indicators for Construction Sites: A Machine Learning Approach," *Automation in Construction*, vol. 93, pp. 375–386, Sep. 2018, doi: 10.1016/j.autcon.2018.03.022.

[181] A. Gondia, M. Ezzeldin, and W. El-Dakhakhni, "Machine Learning–Based Decision Support Framework for Construction Injury Severity Prediction and Risk Mitigation," *ASCE-ASME J. Risk Uncertainty Eng. Syst., Part A: Civ. Eng.*, vol. 8, no. 3, p. 04022024, Sep. 2022, doi: 10.1061/AJRUA6.0001239.

[182] M. Bilal and L. O. Oyedele, "Guidelines for Applied Machine Learning in the Construction Industry—A Case of Profit Margins Estimation," *Advanced Engineering Informatics*, vol. 43, p. 101013, Jan. 2020, doi: 10.1016/j.aei.2019.101013.

[183] A. Khodabakhshian, U. Malsagov, and F. Re Cecconi, "Machine Learning Application in Construction Delay and Cost Overrun Risks Assessment," in *Advances in Information and Communication*, vol. 921, K. Arai, Ed., in Lecture Notes in Networks and Systems, vol. 921. , Cham: Springer Nature Switzerland, 2024, pp. 222–240. doi: 10.1007/978-3-031-54053-0_17.

[184] Theingi Aung, S. R. Liana, A. Htet, and Amiya Bhaumik, "Using Machine Learning to Predict Cost Overruns in Construction Projects," *JTIE*, vol. 2, no. 2, pp. 1–7, Jun. 2023, doi: 10.56556/jtie.v2i2.511.

[185] M. T. Raliile and T. C. Haupt, "Machine Learning Applications for Monitoring Construction Health and Safety Legislation and Compliance," *Proceedings of*

International Structural Engineering and Construction, vol. 7, no. 2, Nov. 2020, doi: 10.14455/ISEC.2020.7(2).CON-23.

[186] H. Anysz, M. Apollo, and B. Grzyl, "Quantitative Risk Assessment in Construction Disputes Based on Machine Learning Tools," *Symmetry*, vol. 13, no. 5, p. 744, Apr. 2021, doi: 10.3390/sym13050744.

[187] C.-L. Fan, "Defect Risk Assessment Using a Hybrid Machine Learning Method," *J. Constr. Eng. Manage.*, vol. 146, no. 9, p. 04020102, Sep. 2020, doi: 10.1061/(ASCE)CO.1943-7862.0001897.

[188] D. Rankin, M. Black, R. Bond, J. Wallace, M. Mulvenna, and G. Epelde, "Reliability of Supervised Machine Learning Using Synthetic Data in Health Care: Model to Preserve Privacy for Data Sharing," *JMIR Medical Informatics*, vol. 8, no. 7, p. e18910, Jul. 2020, doi: 10.2196/18910.

[189] B. N. Jacobsen, "Machine Learning and the Politics of Synthetic Data," *Big Data & Society*, vol. 10, no. 1, p. 205395172211453, Jan. 2023, doi: 10.1177/20539517221145372.

[190] A. Gupta, D. Bhatt, and A. Pandey, "Transitioning from Real to Synthetic data: Quantifying the bias in model." arXiv, May 10, 2021. doi: 10.48550/ARXIV.2105.04144.

[191] V. Wolf, A. Lugmayr, M. Danelljan, L. Van Gool, and R. Timofte, "DeFlow: Learning Complex Image Degradations from Unpaired Data with Conditional Flows," in *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Jun. 2021, pp. 94–103. doi: 10.1109/CVPR46437.2021.00016.

[192] V. Thambawita *et al.*, "SinGAN-Seg: Synthetic Training Data Generation for Medical Image Segmentation," *PLOS ONE*, vol. 17, no. 5, p. e0267976, May 2022, doi: 10.1371/journal.pone.0267976.

[193] RiskAmp, "What Is Monte Carlo Simulation?," *RiskAmp*, 2012.

[194] W. S. Lee, D. L. Grosh, F. A. Tillman, and C. H. Lie, "Fault Tree Analysis, Methods, and Applications ☐ A Review," *IEEE Trans. Rel.*, vol. R-34, no. 3, pp. 194–203, Aug. 1985, doi: 10.1109/TR.1985.5222114.

[195] L. A. Zadeh, "Fuzzy Sets," *Information and Control*, vol. 8, no. 3, 1965, doi: 10.1016/S0019-9958(65)90241-X.

[196] Y. Wang, C. Su, and M. Xie, "A Weakest T-Norm Based Fuzzy Fault Tree Approach for Probability Assessment of Natural Gas Pipeline," in *2022 6th International Conference on System Reliability and Safety, ICSRS 2022*, 2022. doi:

10.1109/ICSRS56243.2022.10067575.

[197] Y. E. Senol, Y. V. Aydogdu, B. Sahin, and I. Kilic, "Fault Tree Analysis of Chemical Cargo Contamination by Using Fuzzy Approach," *Expert Systems with Applications*, vol. 42, no. 12, pp. 5232–5244, Jul. 2015, doi: 10.1016/j.eswa.2015.02.027.

[198] S. Huang, "Linear Regression Analysis," in *International Encyclopedia of Education: Fourth Edition*, 2022. doi: 10.1016/B978-0-12-818630-5.10067-3.

[199] R. Tibshirani, "Regression Shrinkage and Selection Via the Lasso," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 58, no. 1, 1996, doi: 10.1111/j.2517-6161.1996.tb02080.x.

[200] A. E. Hoerl and R. W. Kennard, "Ridge Regression: Biased Estimation for Nonorthogonal Problems," *Technometrics*, vol. 12, no. 1, 1970, doi: 10.1080/00401706.1970.10488634.

[201] W. S. McCulloch and W. Pitts, "A Logical Calculus of the Ideas Immanent in Nervous Activity," *Bulletin of Mathematical Biophysics*, vol. 5, no. 4, pp. 115–133, Dec. 1943, doi: 10.1007/BF02478259.

[202] Y. Zhang and Q. Yang, "An Overview of Multi-Task Learning," *National Science Review*, vol. 5, no. 1, 2018, doi: 10.1093/nsr/nwx105.

[203] J. Edmonds, "Matroids and the Greedy Algorithm," *Mathematical Programming*, vol. 1, no. 1, 1971, doi: 10.1007/BF01584082.

[204] E. Parzen, "On Estimation of a Probability Density Function and Mode," *Ann. Math. Statist.*, vol. 33, no. 3, pp. 1065–1076, Sep. 1962, doi: 10.1214/aoms/1177704472.

[205] G. K. Rajbahadur, S. Wang, G. A. Oliva, Y. Kamei, and A. E. Hassan, "The Impact of Feature Importance Methods on the Interpretation of Defect Classifiers," *IEEE Transactions on Software Engineering*, vol. 48, no. 7, 2022, doi: 10.1109/TSE.2021.3056941.

[206] S. Lundberg and S.-I. Lee, "A Unified Approach to Interpreting Model Predictions." arXiv, Nov. 24, 2017. Accessed: Apr. 12, 2024. [Online]. Available: http://arxiv.org/abs/1705.07874

[207] C. Molnar, "Interpretable Machine Learning," GitHub Repository. Accessed: Jan. 22, 2024. [Online]. Available: https://christophm.github.io/interpretable-ml-book/

Appendix: Table

Table A1. Summary of the risk factors

Category	No.	Risk factor	Scales	Number of Scales	Scale Occurrence Probability Distribution (P_O)	Scale Risk Distribution (R)
1. Overall information of the project	1.1	What is the country of the project?	[Asia, Europe, Africa, North America, South America, Antarctica, and Oceania]. Although initially we planned to use countries as indexes, for simplicity, we finally used continents.	7	[0.3, 0.14, 0.2, 0.14, 0.14, 0.04, 0.04]*	[0.8, 0.7, 0.9, 0.65, 0.7, 0.6, 0.7]**
	1.2	What is the project budget?	[<= $100,000, $100,000 - $500,000, $500,000 - $1 million, $1 million - $5 million, > $5 million]	5	[0.25, 0.25, 0.2, 0.2, 0.1]*	[0.2, 0.4, 0.6, 0.8, 1.0]*
	1.3	What is the percentage of the total project budget for cybersecurity management?	[<= 1%, 1% - 2%, 2% - 3%, 3% - 4%, 4% - 5%, > 5%]	6	[0.1, 0.15, 0.25, 0.25, 0.15, 0.1]*	[1.0, 0.83, 0.67, 0.50, 0.33, 0.17]*
	1.4	What is the project duration?	[<= 3 months, 3 – 6 months, 6 – 12 months, 12 – 24 months, > 24 months]	5	[0.1, 0.2, 0.35, 0.25, 0.1]*	[0.2, 0.4, 0.6, 0.8, 1.0]*
	1.5	What is the total number of people involved in the project (labor excluded)?	[<= 50, 51 – 100, 101 – 200, 201 – 300, 301 – 400, > 400]	6	[0.2, 0.25, 0.25, 0.15, 0.1, 0.05]*	[0.17, 0.33, 0.50, 0.67, 0.83, 1.0]*
	1.6	What is the project type?	[Transportation Infrastructure Projects, Government Projects, Healthcare Projects, Large-Scale Commercial Projects, Residential Projects, Other types]	6	[0.15, 0.2, 0.1, 0.2, 0.3, 0.05]*	[0.17, 0.33, 0.50, 0.67, 0.83, 1.0]*
	1.7	Whether there is a dedicated cybersecurity legal team for the project?	[Yes, No, Unsure]	3	[0.3,0.6,0.1]*	[0.4, 1, 0.7] **
2. Project structure	2.1	What is the project delivery method?	[Design-Bid-Build (DBB), Design-Build (DB), Construction Manager at Risk (CMAR), Construction Management Multi-Prime (CMMP), Public-Private Partnership (PPP or P3), Integrated Project Delivery (IPD), Design/Build/Operate/Maintain (DBOM), Other types]	8	[0.2, 0.15, 0.1, 0.1, 0.1, 0.2, 0.1, 0.05]*	[0.7, 0.8, 0.7, 0.7, 0.7, 0.9, 0.7] **

Table A1 (continued)

Category	No.	Risk factor	Scales	Number of Scales	Scale Occurrence Probability Distribution ($\vec{Po}$)	Scale Risk Distribution ($\vec{R}$)
	2.2 (2.2.1 – 2.2.8)	What is the number of sub-teams at different layers of the project?	Eight layers, each layer's scales are: [<= 10, 11 – 20, 21 – 30, 31 – 40, > 40, N/A], "N/A" means this layer is not existent	47	[0.1, 0.2, 0.3, 0.25, 0.15, 0] [0.25, 0.3, 0.15, 0.15, 0.1, 0.05] [0.35, 0.25, 0.15, 0.1, 0.05, 0.1] [0.4, 0.2, 0.1, 0.1, 0.05, 0.15] [0.45, 0.15, 0.075, 0.05, 0.025, 0.25] [0.5, 0.15, 0.055, 0.02, 0.025, 0.25] [0.55, 0.1, 0.025, 0.015, 0.01, 0.3] [0.6, 0.025, 0.01, 0.01, 0.005, 0.35] **	[0.2, 0.4, 0.6, 0.8, 1.0, 0.0]* The same for other layers
	2.3 (2.3.1 – 2.3.8)	What is the number of communication channels at different layers in the model?	Eight layers, each layer's scales are: [<= 50, <= 100, <= 150, <= 200, < 250, <= 300, > 300, N/A], "N/A" means this layer is not existent	63	[0.1, 0.15, 0.25, 0.2, 0.15, 0.1, 0.05, 0] [0.13, 0.16, 0.18, 0.17, 0.12, 0.09, 0.045, 0.03] [0.16, 0.17, 0.16, 0.15, 0.1, 0.08, 0.046, 0.06] [0.19, 0.18, 0.14, 0.13, 0.08, 0.07, 0.044, 0.09] [0.22, 0.19, 0.12, 0.11, 0.06, 0.06, 0.042, 0.12] [0.25, 0.2, 0.1, 0.09, 0.04, 0.05, 0.042, 0.05] [0.28, 0.21, 0.08, 0.07, 0.02, 0.04, 0.038, 0.18] [0.31, 0.22, 0.06, 0.05, 0.01, 0.03, 0.036, 0.21] **	[0.14, 0.29, 0.43, 0.57, 0.71, 0.86, 1.0, 0.0]* The same for other layers
	2.4	What is the percentage of teams overlapping in different projects?	[<= 20%, 21% - 40%, 41% - 60%, 61% - 80%, 81% - 100%]	5	[0.25, 0.3, 0.25, 0.15, 0.05] **	[0.2, 0.4, 0.6, 0.8, 1.0]*

Table A1 (continued)

Category	No.	Risk factor	Scales	Number of Scales	Scale Occurrence Probability Distribution (Po)	Scale Risk Distribution (R)
3. IT factors	3.1	What is the scale of your company?	[<= 30, 31 – 60, 61 – 100, 101 – 150, > 150]	5	[0.3, 0.35, 0.2, 0.1, 0.05]*	[1.0, 0.8, 0.6, 0.4, 0.2]*
	3.2	What is the phase of the construction project when your company is involved?	[Planning and Bidding phase, Design phase, Construction phase, Maintenance & Operation phase, Demolition phase]	5	[0.15, 0.20, 0.25, 0.3, 0.1] **	[0.6, 1, 0.8, 0.6, 0.2] **
	3.3	Is there a dedicated IT team for the project?	[Yes, No, Unsure]	3	[0.35, 0.55, 0.1]*	[0.2, 1, 0.6] **
	3.4	What is the total number of critical digital assets?	[<= 50, 51 – 200, 201 – 400, 401 – 600, > 600]	5	[0.20, 0.30, 0.30, 0.15, 0.05]*	[0.2, 0.4, 0.6, 0.8, 1.0]*
	3.5	What is the total number of user endpoints of digital devices for the project?	[<= 50, 51 – 200, 201 – 400, 401 – 600, > 600]	5	[0.20, 0.30, 0.30, 0.15, 0.05]*	[0.2, 0.4, 0.6, 0.8, 1.0]*
	3.6	What is the percentage of digital devices with firewalls or intrusion detection systems involved in the project?	[<= 20%, 21% - 40%, 41% - 60%, 61% - 80%, 81% - 100%]	5	[0.05, 0.15, 0.25, 0.30, 0.25]*	[1.0, 0.8, 0.6, 0.4, 0.2]*
	3.7	What is the network type used for the project: Public or Private?	[Public network, Private network, Both public and private network]	3	[0.3, 0.4, 0.3] **	[1, 0.2, 0.6] **
	3.8	What is the percentage of individuals who fail phishing tests after completing mandatory training?	[<= 20%, 21% - 40%, 41% - 60%, 61% - 80%, 81% - 100%]	5	[0.35, 0.30, 0.20, 0.10, 0.05] **	[0.2, 0.4, 0.6, 0.8, 1.0]*
	3.9	What is the estimated Mean Time to Respond (MTTR) in hours?	[Within 1 hour, 1 – 4 hours, 4 – 8 hours, 8 – 24 hours, Above 24 hours]	5	[0.10, 0.30, 0.30, 0.20, 0.10] **	[0.2, 0.4, 0.6, 0.8, 1.0]*
4. OT factors	4.1	What is the total number of important OT equipment involved?	[<= 30, 31 – 60, 61 – 90, 91 – 120, 121 – 150, > 150]	6	[0.30, 0.25, 0.20, 0.15, 0.07, 0.03]*	[0.17, 0.33, 0.50, 0.67, 0.83, 1.0] *
	4.2	What is the level of physical access control mechanism to OT equipment?	[Level 1, Level 2, Level 3, Level 4, Level 5]	5	[0.10, 0.20, 0.30, 0.25, 0.15] **	[1.0, 0.8, 0.6, 0.4, 0.2] *
	4.3	What is the percentage of OT equipment isolated from the project's general network?	[<= 20%, 21% - 40%, 41% - 60%, 61% - 80%, 81% - 100%]	5	[0.05, 0.15, 0.30, 0.30, 0.20]*	[1.0, 0.8, 0.6, 0.4, 0.2] *
	4.4	What is the average age of the important OT equipment, in years?	[<= 1, 1 – 3, 4 – 7, 8 – 10, > 10]	5	[0.25, 0.35, 0.25, 0.10, 0.05]*	[0.2, 0.4, 0.6, 0.8, 1.0] *
	4.5	What is the level of authentication mechanism to access the HMI (Human Machine Interface)?	[Level 1, Level 2, Level 3, Level 4, Level 5]	5	[0.10, 0.40, 0.25, 0.15, 0.10] **	[1.0, 0.8, 0.6, 0.4, 0.2] *

Table A1 (continued)

Category	No.	Risk factor	Scales	Number of Scales	Scale Occurrence Probability Distribution (Po)	Scale Risk Distribution (R)
5. Management and human factors	5.1	What is the average level of commitment to corporate governance, ethical practices and cybersecurity policy?	[Level 1, Level 2, Level 3, Level 4, Level 5]	5	[0.05, 0.10, 0.20, 0.35, 0.30] **	[1.0, 0.8, 0.6, 0.4, 0.2] *
	5.2	What is the average frequency of security training per year?	[<= 10, 11 – 20, 21 – 30, 31 – 40, 41 – 50, > 50]	6	[0.20, 0.30, 0.20, 0.15, 0.10, 0.05]*	[1.0, 0.83, 0.67, 0.50, 0.33, 0.17] *
	5.3	Do you allow password reuse for any project-related software, systems, or accounts (e.g., project management tools, email, internal networks, file storage, etc.)?	[Yes, No]	2	[0.7, 0.3]*	[1, 0.2] **
	5.4	Does internet access within your construction project require Multi-Factor Authentication (MFA) or utilize other methods such as biometrics or face recognition?	[Yes, No]	2	[0.6, 0.4]*	[0.2, 1] **
	5.5	What is the percentage of people who have access to sensitive information in the project?	[<= 10%, 11% - 30%, 31% - 50%, 51% - 70%, 71% - 90%, 91% - 100%]	6	[0.45, 0.35, 0.10, 0.05, 0.03, 0.02]*	[0.17, 0.33, 0.50, 0.67, 0.83, 1.0] *
	5.6	What is the average team member variability over a 3-month period?	[<= 20%, 20% - 40%, 40% - 60%, 60% - 80%, 80% - 100%]	5	[0.55, 0.30, 0.10, 0.03, 0.02] **	[0.2, 0.4, 0.6, 0.8, 1.0] *
	5.7	What is the average socioeconomic level of the people involved in the project?	[Level 1, Level 2, Level 3, Level 4, Level 5]	5	[0.10, 0.20, 0.40, 0.20, 0.10] **	[1.0, 0.8, 0.6, 0.4, 0.2] *

Notes:
* in the Po column indicates the probability distribution is informed by the published database.
** in the Po column indicates the probability distribution is based on estimation and expert validation.
* in the R column indicates the risk degree is derived based on linear assumption.
** in the R column indicates the risk degree is derived based on fuzzy set theory.