

THESIS FOR THE DEGREE OF LICENTIATE OF ENGINEERING

Toward Accident Prevention Through Machine Learning Analysis of Accident Reports

MAY SHAYBOUN

Department of Architecture and Civil Engineering

CHALMERS UNIVERSITY OF TECHNOLOGY

Gothenburg, Sweden 2022

Toward Accident Prevention Through Machine Learning Analysis of Accident Reports
MAY SHAYBOUN

© MAY SHAYBOUN, 2022.

Technical report no 2022:3 Lic /Architecture and Civil Engineering / Chalmers University of Technology

Department of Architecture and Civil Engineering
Chalmers University of Technology
SE-412 96 Gothenburg
Sweden
Telephone + 46 (0)31-772 1000

Chalmers Reproservice
Gothenburg, Sweden 2022

Toward Accident Prevention Through Machine Learning Analysis of Accident Reports

MAY SHAYBOUN

Department of Architecture and Civil Engineering
Chalmers University of Technology

Abstract

Occupational safety remains of interest in the construction sector. The frequency of accidents has decreased in Sweden but only to a level that remains constant over the last ten years. Although Sweden shows to be performing better in comparison to other European countries, the construction industry continues to contribute to a fifth of fatal accidents in Europe. The latter situation pushes towards the need for reducing the frequency and fatalities of occupational accident occurrences in the construction sector. In the Swedish context, several initiatives have been established for prevention and accident frequency reduction. However, risk analysis models and causal links have been found to be rare in this context.

The continuous reporting of accidents and near-misses creates large datasets with potentially useful information about accidents and their causes. In addition to that, there has been an increased research interest in analysing this data through machine learning (ML). The state-of-art research efforts include applying ML to analyse the textual data within the accumulated accident reports, identifying contributing factors, and extracting accident information. However, solutions that are created by ML models can lead to changes for a company and the industry. ML modelling includes a prototype development that is accompanied by the industry's and domain experts' requirements. The aim of this thesis is to investigate how ML based methods and techniques could be used to develop a research-based prototype for occupational accident prevention in a contracting company. The thesis focus is on the exploration of a development processes that bridges ML data analysis technical part with the context of safety in a contracting company. The thesis builds on accident causation models (ACMs) and ML methods, utilising the Cross Industry Standard Process Development Method (CRISP-DM) as a method. These were employed to interpret and understand the empirical material of accident reports and interviews within the health and safety (H&S) unit.

The results of the thesis showed that analysing accident reports via ML can lead to the discovery of knowledge about accidents. However, there were several challenges that were found to hinder the extraction of knowledge and the application of ML. The identified challenges mainly related to the standardization of the development process and, the feasibility of implementation and evaluation. Moreover, the tendency of the ML-related literature to focus on predicting severity was found not compatible either with the function of ML analysis or the findings of accident causation literature which considers severity as a stochastic element. The analysis further concluded that ACMs seemed to have reached a mature stage, where a new approach is needed to understand the rules that govern the relationships between emergent new risks – rather than the systemization of risks themselves. The analysis of accident reports by ML needs further research in systemized methods for such analysis in the domain of construction and in the context of contracting companies – as only few research efforts have focused on this area regarding ML evaluation metrics and data pre-processing.

Key words: Accident report, accident causation models, construction, machine learning, prevention, health and safety.

List of Appended papers and authors' contributions

Paper I: A REVIEW OF MACHINE LEARNING FOR ANALYSING ACCIDENT REPORTS IN THE CONSTRUCTION INDUSTRY AND APPLICATION REQUIREMENTS (under review at The Journal of Information Technology in Construction (ITcon) submitted on 2022-02-10)

Shayboun, M, Kifokeris, D, and Koch, C (2022)

This paper is under review in The Journal of Information Technology in Construction (ITcon) submitted on 2022-02-10.

- May Shayboun contributed to formulating the research question, introduction, collecting the literature material, synthesising the literature and discussion and conclusion.
- Dimosthenis Kifokeris contributed to the development of the method and with feedback throughout the paper writing.
- Christian Koch contributed to the development of the method, development of the discussion and with feedback throughout the paper writing.

Paper II: A COMPARISON OF ACCIDENT CAUSATION MODELS (ACMS) AND MACHINE LEARNING (ML) FOR APPLIED ANALYSIS WITHIN ACCIDENT REPORTS

Shayboun, M, Koch, C, and Kifokeris, D (2021)

This paper was included in the proceedings of the 37th Annual ARCOM Conference, 6-7 September 2021, UK, Association of Researchers in Construction Management, 289-298

- May Shayboun contributed to writing the entire paper.
- Christian Koch contributed to the development of the method and with feedback and commentary throughout the paper writing.
- Dimosthenis Kifokeris contributed to the development of the method and with feedback and commentary throughout the paper writing.

Paper III: LEARNING FROM ACCIDENTS: MACHINE LEARNING PROTOTYPE DEVELOPMENT BASED ON THE CRISP-DM BUSINESS UNDERSTANDING

Shayboun, M, Koch, C, and Kifokeris, D (2021)

This paper appeared in the Proceedings of the Joint CIB W099 & W123 International Conference 2021: Changes and innovations for improved wellbeing in construction

- May Shayboun contributed to writing the introduction, conducting interviews, transcribing the interviews, writing of the method business understanding, empirical material, discussion and conclusion.
- Christian Koch contributed to writing the mapping of the context and with feedback and commentary throughout the paper writing.
- Dimosthenis Kifokeris contributed to writing the method and the status of ML development methods section and with feedback and commentary throughout the paper writing.

Acknowledgements

I want to thank everyone who had made this work possible and contributed with discussions, advice, and valuable insights. It had been a journey that involved much growth and development. I am very grateful that I had the opportunity to pursue my licentiate thesis. I want to start by recognizing the outstanding efforts of my main supervisor, Christian Koch and co-supervisor, Dimosthenis Kifokeris. Working with you has been a lifetime experience that I will always keep. I am incredibly thankful for the profound knowledge, the discussions, the guidance and the support that Christian Koch has provided throughout this work. I want to extend my gratitude to my co-supervisor, Dimosthenis Kifokeris, who has always contributed with learnings, continuous feedback, and close collaboration. I want to acknowledge the efforts of Christina- Claeson-jonsson for the engagement and for providing collaboration with the industry.

I would like to take the chance to thank all my colleagues within the division of Construction Management and the division of Building Design for contributing to the supportive academic dialogue. A special thanks go to my examiner Christine Räisänen for her valuable feedback, involvement, and support.

It was inspiring to meet the most impressive people in conferences, Ph.D. courses, and other academic activities during this journey. I am very thankful for that!

To all the respondents and reference group members and collaborators, thank you for sharing your experiences and taking the time to add insights and feedback.

Finally, I would like to acknowledge the support of my partner Ghaith and thank my mother, father and sisters. I would not have accomplished this journey without your continuous encouragement.

May Shayboun

Gothenburg, February 2022

Table of Contents

1. Introduction	1
1.1. Background	1
1.2. Aim and research questions	3
2. Theoretical framework	4
2.1. Accident causation models (ACMs)	4
2.2. Machine learning (ML)	5
3. Research design	12
3.1. Research process	13
3.2. Case description	14
3.3. Collection of empirical material	15
3.3.1. Accident reports	15
3.3.2. Interviews	16
3.4. Analysis of empirical material	17
3.4.1. Interview's analysis	17
3.4.2. Machine learning design	17
3.4.2.1. Understanding the data structure	17
3.4.2.2. Data pre-processing	19
3.4.2.3. Algorithm choice	19
3.5. Researcher own role	20
3.6. Ethical considerations	20
4. Summary of the papers	21
4.1. Paper I: A REVIEW OF MACHINE LEARNING FOR ANALYSING ACCIDENT REPORTS IN THE CONSTRUCTION INDUSTRY AND APPLICATION REQUIREMENTS (under review at The Journal of Information Technology in Construction (ITcon) submitted on 2022-02-10)	21
4.2. Paper II: A COMPARISON OF ACCIDENT CAUSATION MODELS (ACMS) AND MACHINE LEARNING (ML) FOR APPLIED ANALYSIS WITHIN ACCIDENT REPORTS	22
4.3. Paper III: LEARNING FROM ACCIDENTS: MACHINE LEARNING PROTOTYPE DEVELOPMENT BASED ON THE CRISP-DM BUSINESS UNDERSTANDING	23
5. Machine learning model design and analysis	25
5.1. Model design	25
5.2. Framework of understanding	27
5.3. ML analysis results	28
6. Results and discussion	30
7. Conclusion	35
8. Future work	38
9. References	39
List of abbreviations	44
Appendix	45

1. Introduction

This project focuses on identifying risks associated with occupational accidents within a Swedish Contracting Company, particularly those that result in injuries.

Contracting companies have improved their processes and registration of accident reports. These improvements have been parallel to developments in regulations and the available software for accident registration. There is, however, a desire and need to learn from failures by analysing the accidents that are now reported more consistently. Moreover, new opportunities to gain knowledge about the causes of accidents through the registration database have emerged, together with the increased interest in and the capabilities of machine learning (ML), both of which could potentially improve accident prevention.

1.1. Background

The health and safety (H&S) units within large contractor organisations continuously report their internal accidents and near-misses (incidents) that have occurred on construction sites. The accumulated reports form a database that details the different types of accidents. In addition, production personnel such as the site managers and safety engineers have practical knowledge of accidents and their causes. However, it is seldom that these accumulated accident reports are analysed. Nonetheless, there has been a recent increasing trend within publications in accident prevention to look at this data using ML (Xu et al. 2021). These include applying ML to analyse the textual data within the accumulated accident reports in contracting companies and national registries, identifying contributing factors, and extracting accident information (Hegde and Rokseth 2020, Sarkar and Maiti 2020, Khallaf and Khallaf 2021, Hou et al. 2021).

ML, designed to find underlying patterns in a dataset, can be used to predict situations that may pose the risk of accidents at construction sites (Baek et al. 2021, Vallmuur 2015). ML systems automatically improve their built-in functionality through experience (Jordan and Mitchell 2015). The use of ML – a subdivision within artificial intelligence (AI) - in a construction context is now being seen as a promising development (Vallmuur 2015, Kifokeris and Xenidis 2018, Pan and Zhang 2021). Another key tool in understanding the accumulated data is data mining, which is understood as "the process of discovering interesting patterns from large amounts of data" (Han et al. 2011).

Vallmuur (2015) reviewed eight ML system examples that analyse the database of registered occupational accidents. The ML systems use Bayesian networks (BN), decision trees (DT) and association rule mining. Examples of using ML in the prevention of occupational accidents are becoming more common within relevant published research studies (Hegde and Rokseth 2020). Algorithms such as DT, Random Forest (RF), Stochastic Gradient Tree Boosting (SGTB), artificial neural network (ANN), and natural language processing (NLP) for data pre-processing (Vallmuur 2015, Witten et al. 2016, Hegde and Rokseth 2020, Hou et al. 2021) can be used to analyse the data of injury cases. The purpose of the latter type of analysis includes the prediction of accident types, classification of causes, and information extraction. It is important to note that in this thesis, the author terms a "prototype" as the designed digital software that suggests a precise implementation for ML-based data analytics, i.e., that shows means of application and an interface that is ready for use.

Despite the attention being paid to the importance of safety in the workplace, the building industry has the highest frequency of fatalities (Arbetsmiljöverket 2021). Moreover, the rate of fatal injuries has not decreased since 2019 compared to other industries (such as transport and warehousing) (Arbetsmiljöverket 2021). The frequency of accidents (number of accidents / 1000 employees) has levelled out in the last decade, hanging roughly at around 11 (Byggforetagen 2021). The accident types show a rather complex and scattered pattern: body movement with physical overload (18%), injuries from tools and gear (16%), collapse, falls and rupture of material (12%), falls from a height (12%) and falls at the same level (tripping) (12%). Although Sweden's figures are better than other European countries (e.g., France, Portugal, and Germany), the construction industry accounts for one-fifth of all fatal accidents at work in the EU. Of these fatal accidents, 27 occurred within the construction sector (Eurostat 2018). There is a need to reduce the injury frequency and fatalities in the construction industry.

The construction industry is characterized by high complexity, uncertainty, and interdependence (Berglund et al. 2017). This situation creates difficulties in operational planning and creating a safe and disturbance-free workflow. An agenda for safer construction has been pursued by both practitioners and researchers alike. Multiple routines and approaches exist in Swedish projects and companies (Törner and Pousette 2009). Accident prevention research has developed several risk analyses and accident-related causal models (Behm and Schneller 2013, Berglund et al. 2017, Harms-Ringdahl 2013, Jørgensen 2002, Reason 2008). Some of these models systematically distribute different levels to different causes and systematize causes in a fault-tree analysis or a hierarchical analysis, assuming that multiple causes drive accidents. However, the use of such causal links is rare in the Swedish construction sector (Berglund et al. 2017). Safety in construction is affected and/or hindered by conditions such as the construction site's organization, management (Törner and Pousette 2009), equipment, and materials (Berglund et al. 2019). Those change from one workplace to another, making it more difficult to maintain sufficient, common safety routines (Albrechtsen and Hovden 2014, Lingard et al. 2012, Schwatka et al. 2016). In addition, safety and the safety culture are affected by several factors such as subcontractor and the main contractor cultures, organizational decision-making regarding safety considerations, and individual behaviour (Koch 2013, Zhou et al. 2015).

This thesis assumes that it is possible to prevent accidents by systematizing the learning and knowledge accumulated from registered accidents by investment in the latest digitization technology – in this case, ML (Berglund et al. 2017). However, IT research that implements ML data analysis can lead to changes not only for a company but for the entire industry (Bilal et al. 2016, Bilal and Oyedele 2020). Applying ML does not only include the development of a prototype, but also to address the industry requirements and collaborate with industry experts and ML analysts (Bilal and Oyedele 2020). Regardless, the current literature lacks concrete use cases and the required integration with domain and expert knowledge (Vallmuur 2015, Bilal et al. 2016). The aforementioned collaboration with domain experts needs an understanding of the context or the domain of the application and explaining the ML models to the humans involved (Gilpin et al. 2018). This thesis contributes to exploring development processes that bridge ML data analysis technical part with the context of safety in a contracting company. In understanding the context, the Cross Industry Standard Process Development Method (CRISP-DM) is of interest (Martínez-Plumed et al. 2019). CRISP-DM is a methodology consisting of six steps that catalogue and guide the process of data mining projects (Martínez-Plumed et al. 2019). Moreover, accident causation models (ACMs) are of interest in explaining ML models.

According to Kjellen and Albrechtsen (2017), ACMs are the "simplified representations of the processes in the real world that result in accidental loss" (p.25). ACMs are mature theoretical, conceptual models that have impacted the development of safety management methods and processes (Kjellen and Albrechtsen 2017).

1.2. Aim and research questions

This thesis aims to investigate how ML-based methods and techniques could be used to develop a research-based prototype for occupational accident prevention in a contracting company. The thesis focuses on exploring a development process that bridges ML data analysis technical part with the context of safety in a contracting company. The research context is accident prevention and H&S activities on-site, with the company being the case for the prototype development. The main (primary) data source is the registry of accident reports of the case company. Secondary data collection was undertaken through interviews with the company's H&S unit.

It is important for both business and academia to understand the use and the obstacles in introducing advanced technologies such as ML. Such understanding can benefit the construction industry through improved safety performance and better accident prevention strategies. As discussed earlier, this industry-wide interest in improving safety measures has been evident in the literature.

Therefore, this research project contributes to the development of ML for analysing accident reports and exploring methods for building a prototype to improve occupational accident prevention strategies while considering the context. The context of the contracting company and its safety processes are the targeted application domain for the digital system. The ML analysis is based on the data generated by different actors in the case company and is intended to be applied within its safety processes.

One overall research question and sub-questions were posed based on this research aim.

Overall research question: Does the application of ML on accident reports reveal new knowledge about accidents in the construction industry?

RQ1: What are the requirements for applied ML in the domain of accident prevention in a contracting company's occupational safety processes?

RQ2: What is the role of accident causation models (ACMs) as a theoretical framework for the ML results of analysed reported accidents in the construction industry – and what can be learned about ACMs through ML?

RQ3: What are the experiences and challenges of applying CRISP-DM's business understanding to assure a solid contextual embedding and an appreciation of local dynamics?

RQ4: What are the predictive attributes of accidents based on the application of ML to accident reports?

2. Theoretical framework

2.1. Accident causation models (ACMs)

ACMs may provide a foundation for accident investigation and feedback and, most importantly, highlight accidents' causal factors (Kjellen and Albrechtsen 2017). Moreover, ACMs were developed and adjusted over the last 100 years, resulting in different ACMs having their own characteristics (Pillay 2015, Fu et al. 2020). Thus, ACMs are different in causes representation and the logic behind the occurrence of accidents (Fu et al. 2020).

ACMs can be classified in different ways, such as linear and non-linear models according to the logical sequence of events that lead to accidents (Fu et al. 2020). Fu et al. (2020) further categorized the non-linear models into human-based, statistics-based, energy-based (e.g., the Bow-tie model, and the tripod beta model), and system-based (Systems Theoretic Accident Model and Processes (STAMP), AcciMap), while linear models included the Swiss cheese model (SCM), Heinrich domino theory, and the HFACS.

Kjellen and Albrechtsen (2017) distinguished between seven main ACMs categories. The categorization by Kjellen and Albrechtsen (2017) included causal-sequence models (the domino theory, the tripod model), process models (Occupational Accident Research Unit (OARU), Haddon's phase model), the energy model (the Swiss cheese model), logical tree models (fishbone diagram, Construction Accident Causation (ConAC)), system models (HFACS, MORT, AcciMap, STAMP).

The inclusion of models in Kjellen and Albrechtsen (2017) and Fu et al. (2020) demonstrates the complexity and diversity of ACMs, primarily evident in the difference in the typology of causes, levels of causes, the relationship between the levels, their application, and the mechanism within which events take place.

According to Woolley et al.'s (2019) categorization, accident causation models have three main categories based on their characteristics and their time of development:

- Simple linear models (1920s)
- Complex linear models (1950s–1990s)
- Complex non-linear models (1990s to present)

The simple linear models (e.g., the domino theory) represent the view on accidents as being predictable through a chain of events and that they could be prevented if one of the root causes was eliminated in the sequence of that chain of events (Woolley et al. 2019). This category usually concentrates on physical/mechanical and human error (Woolley et al. 2019). However, it is criticized for the lack of distinction of uncertain causal relationships at the personal, organizational, and management levels (Kjellen and Albrechtsen 2017).

Complex linear models (SCM, the Loughborough Construction Accident Causation Model, and the Causal Model of Construction Accident Causation) view the accident as being caused by the interaction between latent factors and unsafe human behaviour (Woolley et al. 2019). The SCM argues that accident causes can be traced back to the origins of organizational decision-making (Kjellen and Albrechtsen 2017). Although complex linear models introduce the organizational factors, they also

retain the sequencing of events and do not include factors outside the organization (Woolley et al. 2019).

Complex non-linear models took a broader view of system-related factors (Woolley et al. 2019, Kjellen and Albrechtsen 2017) as a response to the growing complexity and tighter couplings in industrial domains. System-based models (Fu et al. 2020, Kjellen and Albrechtsen 2017) are now surpassing the previous ACMs through their systematic and thorough concentration on managerial and organizational factors and their interaction with individuals, technology, and behaviour. System-based models assume the responsibility of everyone within the system (including politicians and regulators), and accidents are claimed to have been caused by the dynamic and non-linear interaction among multiple factors within the entire system (Woolley et al. 2019).

The development of ACMs initially focused on human behaviour within sequences of events. More recently, ACMs have tended to explore the more dynamic approach and consider higher levels of causation (Pillay 2015). This development seems to be based on the assumption that higher levels of causation can explain accidents. Moreover, the different types of ACMs assume stochasticity in accident severity since the accident impact level is not differentiated within any of the reviewed models. ACMs break down to multiple causation levels that are primarily not assigned weights for their importance but portray the interplay of causation factors as either single-rooted and linear, multiple-linear, and having multiple and dynamic causes.

In construction research, applied causation models range from technological and behavioural models (e.g., the domino theory) to the more advanced socio-technical and cultural models (e.g., the Loughborough construction accident causation model, the fault tree analysis, and the Swiss cheese Model) (Pillay 2015). System-based models were, in contrast, hardly, if ever, used in the published literature dealing with accident causation analysis within the construction sector (Woolley et al. 2019). The scarcity of system-based models points to the limited inclusion of governance and regulatory factors in accident analysis (Woolley et al. 2019). When system-based models are applied to analyse accidents, regulatory and governance factors are often overlooked (Pillay 2015). For example, physical processes, actor activities, equipment and environment, unsafe acts, and management decision-making are more prominent in system-based accident analysis in multiple industrial contexts rather than regulatory and other governmental factors (Hulme et al. 2019).

The limited inclusion of higher levels of causation hampers the understanding of whether the predictability of accidents increases from the advanced growth in the interactions of causes and the inclusion of factors outside the organization and limit benefits in prevention (Grant et al. 2018). They also act to hinder identifying the relationship between factors (Woolley et al. 2019). Since accidents persist in the construction industry, there is a need to revisit theories and models of accidents' causation and critically reflect on applied ACMs in construction – especially set against the quantitative data that is now being derived from many registered accidents.

2.2. Machine learning (ML)

Machine learning is defined as the “computational methods using experience to improve performance or to make accurate predictions” (Mohri et al. 2018). Experience refers to existing information mostly available in a digital form of data (Mohri et al. 2018). ML is also defined as a set of methods that automatically detect patterns and use those to predict future data (Murphy 2012), while according to

Carbonell et al. (1983) “the study and computer modelling of learning processes in their multiple manifestations constitutes the subject matter of machine learning.”

ML includes different types of learning: supervised and unsupervised learning are the main ones (shown in Figure. 1), but there is also semi-supervised learning, transductive inference, online learning, and reinforcement and active learning (Shalev-Shwartz and Ben-David 2014, Mohri et al. 2018, Murphy 2012).

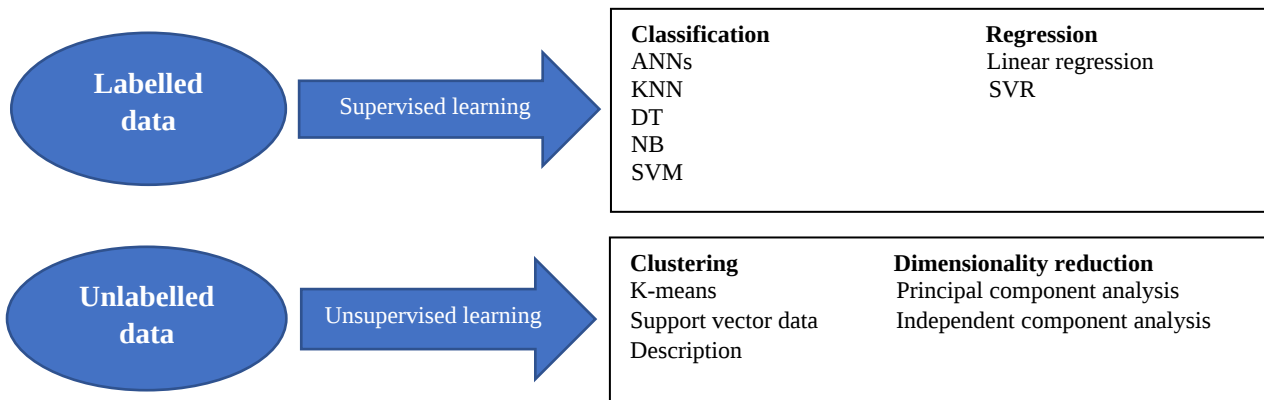


Figure 1. Machine learning summary.

- *Unsupervised learning* is a method of data exploration or description used for, among other things, clustering (Shalev-Shwartz and Ben-David 2014, Mohri et al. 2018, Murphy 2012). In unsupervised learning, there are no specific patterns to be followed or an error metric (Murphy 2012). Clustering can partition or group a set of objects into homogeneous subsets and is usually used in analysing large datasets (Mohri et al. 2018). The same sequence of objects can be clustered differently depending on the algorithm used; therefore, unsupervised learning does not always provide steady results (Shalev-Shwartz and Ben-David 2014). Another feature of unsupervised ML is that checking for accuracy and interpretation is subjective and requires expert knowledge for examining the results and inference (Shalev-Shwartz and Ben-David 2014).
- *Supervised learning* is an approach that learns a mapping from input to output and is used mainly for prediction (Murphy 2012). The input can be referred to as features, attributes, or covariates. At the same time, the output can be either categorical (a classification or a categorical problem) or numerical (a regression or ranking problem) (Murphy 2012). The data should be already labelled (input and output variables are known and identified through pre-assigned categorization). The purpose of using this type of learning is to predict or classify the labels of future examples as accurately as possible (Mohri et al. 2018). Moreover, supervised ML can be used in prediction or classification algorithms and assessed by calculating the potential loss in finding false instances (Shalev-Shwartz and Ben-David 2014).
- *Semi-supervised learning* is when the data is partially labelled and commonly when the unlabelled data is accessible but labelling the data unattainable (Mohri et al. 2018) or when the labelled part is used to infer the unlabelled part (El Naqa et al. 2019).

The ML model process

The ML process usually consists of multiple steps:

- Data exploration
- Data pre-processing
- Model training
- Model validation
- Model testing

It is important to note that the author of this thesis term an ML model as the specific mathematical or computational description that expresses the relationship between a set of input variables and one or more outcome variables studied or predicted.

Data exploration entails gaining knowledge into the attribute types (e.g., nominal or numerical), the entries contained in each attribute, and the distribution of the input and the output features (Han et al. 2011). Data pre-processing ensures data quality for a reliable ML analysis and consists of multiple tasks – including handling missing values, noise, and resolving inconsistencies and discrepancies (Han et al. 2011). Discrepancies might originate from the data entry form, human errors, system errors, and other reasons (Han et al. 2011). The data is then split into the training, validation, and testing datasets (Shalev-Shwartz and Ben-David 2014). The training data should not be used in testing the model to find out whether the ML model performs as well with data points that have not been used in its training (Han et al. 2011). The validation step is used to tune the model's parameters for the ML algorithms (Han et al. 2011).

The testing of ML performance depends on the type of ML problem and the employed algorithms. The Receiver operating characteristic (ROC) curves and the F1 measure are usually used in classification problems (Han et al. 2011). The F1 measure is based on the confusion matrix depicted in Table 1. TP, FP, FN, and TN stand for true positive, false positive, false negative, and true negative, respectively (Japkowicz and Shah 2015). In binary classification tasks, the class of interest is called the positive class while the other is the negative class (Gopal 2018). Accordingly, TP and TN are the accurate classifications that the algorithm achieves. FP and FN are referred to when the algorithm inaccurately classifies a positive when it is a negative in reality and a negative when it is positive, respectively (Gopal 2018). The latter can be variously combined to calculate specific performance metrics, as in the following (Japkowicz and Shah 2015, Han et al. 2011):

$$Accuracy = (TP + TN) / (P + N)$$

$$Precision = TP / (TP + FP)$$

$$Recall = TP / (TP + FN)$$

$$F1\ measure = (2 \cdot Recall \cdot Precision) / (Recall + Precision)$$

It is common that accuracy does not sufficiently evaluate a model's performance, such as in cases of data imbalance (Japkowicz and Shah 2015). Precision and recall also have shortcomings in showing how a classifier behaves in terms of showing the detailed negative and positive recognition (Japkowicz and Shah 2015). Alternatively, the ROC curve is another method for testing the performance of an ML algorithm when accuracy, precision or recall fall short (Han et al. 2011, Japkowicz and Shah 2015,

Gopal 2018) – e.g., when false-negative classifications are costly (such as in disease diagnostic applications). The ROC curve takes paired measurements of false-positive rates on the x-axis and the true-positive rates on the y-axis, with the highest value being 1 (Han et al. 2011).

Table 1. A generic confusion matrix

True class	Positive	Negative
True positive	TP	FP
True negative	FN	TN

Supervised and unsupervised learning have different algorithms characterized by different structures and application types (El Naqa *et al.* 2019). Supervised ML is more interpretable, testable and applicable when available data is labelled (Murphy 2012, Mohri et al. 2018). Furthermore, supervised ML algorithms can be organized into linear and non-linear models.

Linear Models

Linear regression (LR) and support vector machines (SVM) are linear algorithms that can be used for regression or classification (Shalev-Shwartz and Ben-David 2014). LR is a simple model without parameters to control model complexity (Ray 2019). However, if the data is not linearly separable, LR is not best to fit the data (Shalev-Shwartz and Ben-David 2014). Instead, a polynomial regression (which fits a non-linear function although being a statistical estimation problem) can be used (Shalev-Shwartz and Ben-David 2014), but the polynomial model is more complex than LR, and there is a risk for overfitting (Hawkins 2004). SVMs work by separating the dimension space into two classes in the case of a binary classification task (Shalev-Shwartz and Ben-David 2014). The margin of a hyperplane that separates the data is the smallest distance between a point in the training set and the hyperplane (Shalev-Shwartz and Ben-David 2014). This margin limits the performance of the linear SVM; if the margin is larger, the error decreases because the model becomes more tolerant to the disturbance in the data points. The SVM is regularized with using the parameter C – large values of C (smaller regularization) allow the model to fit the training data even in the case of a smaller margin, while larger regularization makes the model more tolerant to errors on individual data points (Bhavsar and Ganatra 2012, Singh et al. 2016).

Non-linear models

K-Nearest Neighbor (KNN) is one of the simplest ML algorithms used for regression and classification. It assumes that the close-by instances are likely to have the same labelling (Shalev-Shwartz and Ben-David 2014). The parameter K can take different values starting from 1, and then the algorithm looks at the single closest instance label to predict the label of another instance. The smaller K is, the more complex the model, and there is a risk of an overfitting decision boundary (Bhavsar and Ganatra 2012, Singh et al. 2016). The disadvantages of KNN models are the sensitivity to dimensionality (which can affect the algorithm’s performance) (Shalev-Shwartz and Ben-David 2014) and the compromise of accuracy because the algorithm assigns equal weights for the features and the sensitivity to the local structure of the data and the value of K (Bhavsar and Ganatra 2012).

Kernelized support vector machines (KSVM) are a variation of SVM that transform the data into a high dimensional space to allow for a linear classification for a feature space that is not linearly separable (Shalev-Shwartz and Ben-David 2014). KSVM are highly sophisticated models and one of

the most accurate models in binary classifications (Bhavsar and Ganatra 2012). On the other hand, DTs are characterized by the easy interpretation by simply visualizing the entire tree. However, feature importance rankings do not indicate which classes are predicted by a feature or the relationships between features (Singh et al. 2016). Moreover, the sample complexity of DTs might result in growing very large trees (deep trees) that are prone to overfitting (Shalev-Shwartz and Ben-David 2014). This situation, however, can be prevented by controlling the size of the tree by applying a reduced-error pruning method (Lee and El Naqa 2015). Random forest (RF) is a classifier consisting of a collection of DTs (Shalev-Shwartz and Ben-David 2014). The DTs within RF are built with random sample variations that are bootstrapped by a random feature split selection (Lee and El Naqa 2015). Although RFs are generally more accurate than simple DTs, they can be unstable, produce local optimal solutions instead of global ones, and have sampling errors (Ray 2019).

Artificial Neural Networks (ANNs) are computational models inspired by the structure of the brain's neurons and have recently reached high performance in different learning tasks (Shalev-Shwartz and Ben-David 2014). There are two main types of ANNs (feed-forward and back-propagation). Feed-forward networks – also called multi-layer perceptron (MLP) – take the idea of computing weighted sums of input features (like in logistic regression) but introduce a processing step that consists of several neurons as a hidden layer (hidden units) (Shalev-Shwartz and Ben-David 2014). The MLP complexity is affected by the number of units, layers, regularization and activation function (Lee and El Naqa 2015). Backpropagation has the same structure as an MLP but backwards learns the network's weights by employing a gradient descent to minimize the squared error between the network outputs and the target values of these outputs (Gopal 2018).

The characteristics of the previously presented ML algorithms are summarized in Table 2. The table characterizes the ML algorithms in strengths – represented in a plus sign – and weaknesses – represented in a minus sign. The table can be used to choose ML algorithms based on the task's requirements. This contributes to a systematic and informed choice of algorithms instead of the experimental approach.

Table 2. Summary of ML algorithms characteristics

Algorithm	NB	SVM	KSVM	DT	RF	KNN	LR	LogR
Interpretability	+	+	-	+	-	+	+	+
Parameters tuning	+	-	-	+	+	-	+	-
High dimensionality	+	+	+	+	-	-	-	-
Feature dependability	-	+	+	-	+	+	-	-
Generalization	-	+	+	-	+	+	-	-
Accuracy	-	+	+	-	+	+	-	+
Small data set	+	-	+	+	+	+	-	-
Large data set	+	+	+	-	+	-	+	+
Linearity	-	-	-	-	-	-	+	-
Low dimensionality	+	-	+	+	+	-	+	+

Dobbe et al. (2018) suggested that bias might originate from multiple sources when data is used in ML decision-making models. First, measurement bias can originate due to how the collected data is scaled,

and the people registering their entries are represented. The modelling bias is affected by the engineering of features and the selection of the model classes. These processes include reconstructing a complex phenomenon in a finite data sample. The optimisation bias is related to the model builder choices of designing and optimising the parameters of the ML algorithms, which affect the outcomes or decisions the model produces (Dobbe et al. 2018).

Dobbe et al. (2018) explained that the origins of bias acknowledge the need for understanding the epistemology of the specific context, and the role played by the creator of the ML model. Also, when applying ML to research in social sciences, Radford and Joseph (2020) suggest a “theory in” and “theory out” approach. Theory in means that known theories about a phenomenon should be considered in the pipeline for research that uses ML in analysing social data. A problem and task definition should rely on the knowledge gap in what is already known about the social world, starting from the conception of an ML model. Thus, theory help in identifying which problems are worth solving and frame why a problem is important. Moreover, theory help to define the ML outcome that validly captures the construct sought to be measured (Radford and Joseph 2020). Theory out refers to considering the model’s interpretability, explainability and theory building beyond the model’s technical parameters – in other words, using theory to understand why the model learned what it did and what can be learned about the world based on its results (Radford and Joseph 2020). Theory building here refers to the new knowledge about the social world that can be discovered from the results of our model (Radford and Joseph 2020). The implication of the latter described approach is that an ML model needs to be developed by supporting the relevant theories throughout the ML development process.

One of the most famous models in ML industrial applications is the Cross Industry Standard Process Development Method (CRISP-DM) (Martínez-Plumed et al. 2019). CRISP-DM consists of multiple steps (business understanding, data understanding, data preparation, modelling, evaluation, and deployment) (Martínez-Plumed et al. 2019, Figure 2)

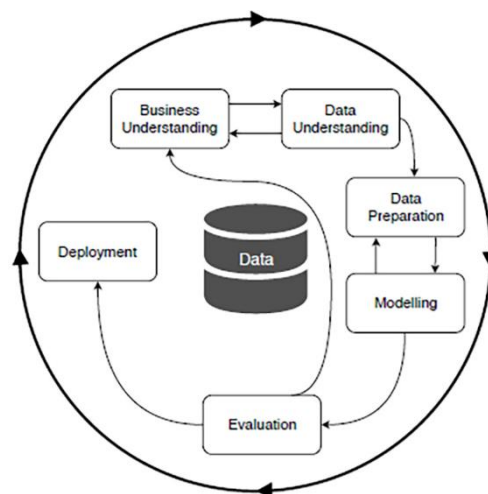


Figure 2. The CRISP-DM process model of data mining (Martínez-Plumed et al. 2019).

These steps can account for a contextualisation of the developmental process, starting with the initial step of business understanding. The business understanding plays a role in defining the business objective and offering a systemised process to mitigate the dependence only on data experimentation

(Chapman et al. 2000, Martínez-Plumed et al. 2019). The business understanding consists of four sub-tasks: determine business objectives, assess the situation, determine data mining goals, and produce a project plan (Chapman et al. 2000).

3. Research design

The approach chosen in this research is interpretive, where reality is deemed as the construction of the interaction between the researcher and the research (Alvesson and Sköldberg 2017). This approach assumes a reflexive research methodology that typically arises when different levels or elements of interpretation are played out against each other, and when none of the research components gains dominance throughout the entire research process (Alvesson and Sköldberg 2017). Reflexivity encourages creativity through the movement between different philosophical profundities and other empirical research elements (Alvesson and Sköldberg 2017). To follow a reflexive research methodology, multiple paradigms that are associated with this research were identified similar to mixed method approaches (Creswell and Clark 2017) – namely, the literature and texts, the empirical material, ACMs and the ML theory.

The research is mainly interested in understanding and interpreting ML methodologies and techniques as a process to develop ML models to be applied in existing practice. The research aim is motivated by the need to bridge the technical ML analysis and the context of the application that involves people. The focus here is on the H&S unit as both a generator of the accident reports dataset and the end-user of the ML intended prototype. The H&S unit consists of safety engineers, safety representatives, site supervisors, site managers, safety managers and safety strategists. The qualitative interpretive research is aligned with the research aim to cultivate interpretation and reflection as key elements of reflexive research (Alvesson and Sköldberg 2017). The premises of this research method is derived from the view that how researchers interpret phenomena is always perspectival and that facts are always theory-laden (Alvesson and Sköldberg 2017).

According to Alvesson and Sköldberg (2017) method, reflexive interpretation consists of four levels – namely, interaction with the empirical material, interpretation, critical interpretation, and reflection on text production and language use (p.331), also called the quadruple hermeneutics (p.122). ACMs, ML algorithms, and CRISP-DM provide the multiplicity needed in the interpretive approach - as illustrated in Figure 3.- each of these is used for interpreting the data analysis results. The formulated research questions are accordingly generated to support interaction across the aforementioned theoretical framework and the empirical material.

The primary data collection was done through the digital reporting system used by a contracting company: Synergi Life. Complementary data collection was done through twelve interviews with the H&S unit within the contracting company. Accident reports are highly dependent on the reporters, especially their interpretations of how accidents and their causes should be described (Dekker 2015). Thus, the challenges and opportunities of developing an ML-powered prototype and implementing it in the safety processes within the case company are ultimately dependent on the prevailing perceptual, theoretical and cultural assumptions within the case company. For the development and analysis of an applied ML model, the CRISP-DM (Cross Industry Standard Process Development Method) is chosen as a development process method. CRISP-DM is also used as a framework to understand the H&S objectives and identify ML utilisation propositions.

Mainly, primary data is analysed through the application of ML algorithms. ACMs is chosen as the theoretical framework for interpreting the ML model results. ACMs components that describe accident occurrences are used to contextualise and conceptualise the results of the ML analysis.

The critical interpretation level stems from the reflection on ACMs from the perspective of existing ML literature by a comparative analysis of the components and assumptions of ACMs against the components and assumptions of existing ML models. Moreover, the experience of conducting the CRISP-DM's business understanding analysis through the interviews was utilised to explore the fit of CRISP-DM to develop ML with the H&S unit.

The researcher finally reflects on their assumptions about the phenomenon and the limitation of the repertoire of interpretation.

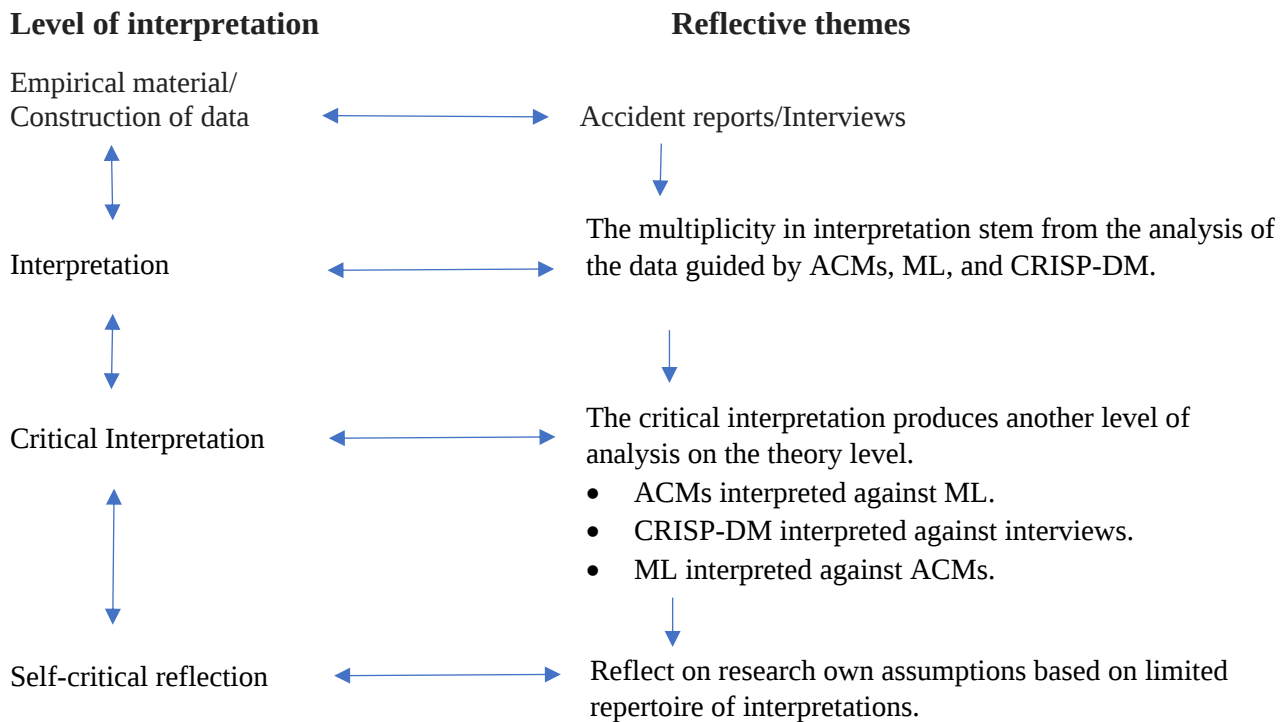


Figure 3. Illustration of reflexive methodology based on levels of interpretations (the quadruple hermeneutics, Alvesson & Sköldbberg 2017).

3.1. Research process

The research process consists of the four steps described in paper I, paper II, and paper III summarised in section 4 and ML results described in section 5. The papers and data analysis were conducted sequentially, represented in a process graph shown in Figure 4. Each paper provided a background for the contribution and the development of the following paper.

The first step was to review the existing body of literature to identify the requirements for applied ML in occupational accident prevention within a construction company. This first step in paper I investigates the possible development requirements for an ML model to be implemented in occupational construction safety. The reviewed literature provides an in-depth and detailed exposition on the uses of ML in analysing accident reports, and is synthesised in terms of used algorithms, data characteristics, data processing, and purpose and scope of the ML models.

Consequently, paper II was designed to investigate the role of ACMs as a theoretical framework in ML application for analysing accident reports in the construction industry. A framework of understanding the ML results should be established to place causes in meaningful categories – and

vice versa, to contribute to learnings about ACMs obtained through ML. The work is carried out in the form of a comparative desk study of the literature covering the application of ML to accident reports in the construction industry and ACMs. This contributed to providing a conceptualisation of ML models through the lens of ACMs' components.

After the means of understanding ML results through ACMs was established, the need for a systemised ML development method ensuring the contextual embedding led to the use of the CRISP-DM method for understanding the context of accident reporting in the case company, and analysing the experience and challenges in applying the “business understanding” stage (Martínez-Plumed et al. 2019). Paper III centres on five interviews with a safety strategist and four safety engineers at the case company in Sweden to answer this research question. Paper III contributed both to the identification of ML utilisation proposals that meet the needs of the H&S unit, and also adds value to accident prevention measures. Seven further interviews were conducted after paper III was published to gain further insights from different actors within the H&S unit of the case company.

The fourth step finally provides an ML analysis of accident reports from the case company based on an ML model design and interviews analysis. This work builds on the learnings and conclusions of papers I, II, and III. Furthermore, the continuation of the interviews contributed to decisions related to the purpose and design of the ML prototype. It is important to note that this thesis has not realised a prototype. However, paper I, paper II, paper III and the ML data analysis are all paving the path of the prototype development process.

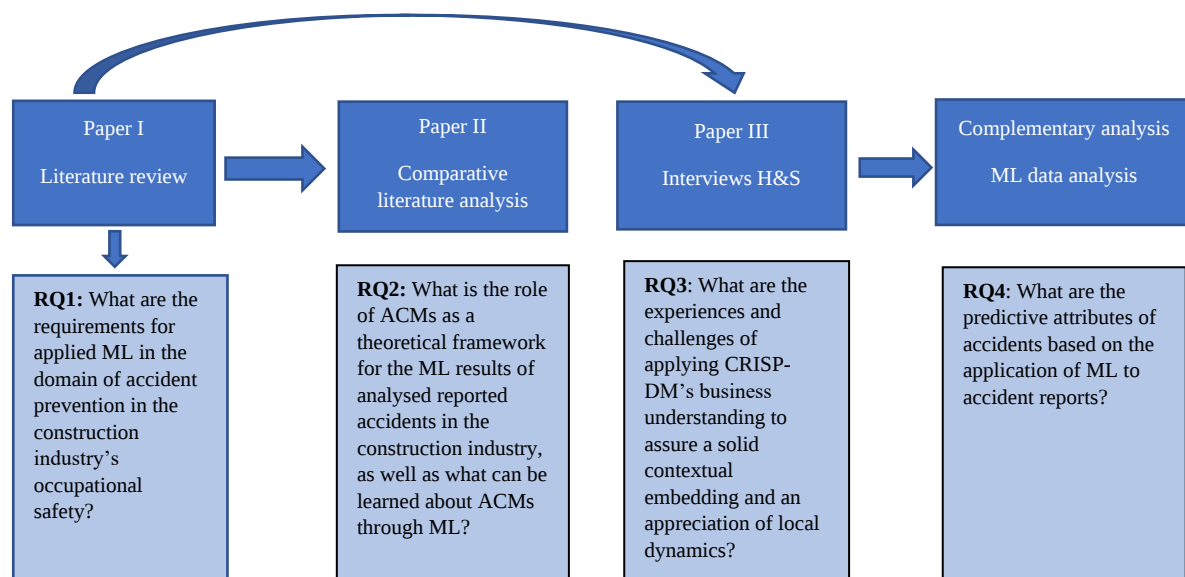


Figure 4. Research process.

3.2. Case description

This thesis's primary interest is in embedded accident prevention in the business setting. The contractor operates a project-based organisation. The building project is the most important value and turnover generator and cost transformer. The different building projects are produced in portfolios placed in divisions with slightly different business objectives, i.e., civil works, residential buildings, office buildings. The project commences with a contract with a client. The H&S work commences by documenting how H&S will be organised in the project in a bid for the customer. Typically, no risk

analysis is carried out by the safety engineers (SEs) this early; however, this is done once a contract is obtained. A particular job role, called BAS P (educated in design safety), is part of this process. From the beginning of work planning, the SEs inspect the plans with an H&S perspective. During production, the safety representatives (the so-called BAS U personnel – basic education for production) are responsible for a particular part of the building project and the building process. They collaborate with the on-site H&S, Quality, and Environment (HES) manager and the SEs. Together, they constitute a horizontal element of the H&S organisation and support the similarly horizontal building processes. H&S work is thus organised close to the single building project. Apart from this horizontal element, the company also encompasses a vertical hierarchy, where H&S is attached to several organisational levels. A central H&S unit is part of a corporate management HR unit. HES units are adjacent to several organisational levels. This cross-organisational H&S apparatus works with behaviour issues, analysis and reporting, digitalisation, and developing directives. In it, it is a common perception that accidents are mostly due to behaviours, so efforts are targeting this issue. Another workstream is related to analysing and reporting, digitalisation, driving projects, and the way the company benefits from machines and innovation. The third workstream is related to developing directive processes and procedures.

3.3. Collection of empirical material

Below I elaborate first of the more quantitatively oriented collection of data emanating from the accident reports, moving on to the rich knowledge and information in the interviews.

3.3.1. Accident reports

The accumulated accident data have a common method for registering and analysing single occurrences of accidents in the construction industry. The case company's data is mostly gathered by safety engineers, site managers, safety representatives, and workers. Accidents are registered through a digital software interface called Synergi Life, which is a complete quality, health, and safety risk management software package. Accident reports were extracted by the researcher into excel sheets and initially investigated in excel.

The software package offers the option of recording four types of reports:

- Accident: An event that led to personal injury.
- Incident: An unwanted, sudden event that could have led to a personal injury.
- Negative observation: An unwanted situation or risk that could have led to personal injury.
- Positive observation: A positive action or solution that has led to better health or safety.

The reporting process consists of four steps:

1. Registration, which is possible to be made by anyone working at the case company (either on the desktop or the mobile application software versions).
2. Appointment of a case handler.
3. Filling in the case with either investigations or a required action.
4. Closing of the case, which needs to be done by the health and safety unit.

The accident report consists of seven main sections (see Appendix).

- Where, what and who.
- General classification.

- Consequences.
- Potential loss.
- Causes.
- Prevention.
- Attached documents.

The data contains two forms of reporting: free text describing the accident, and pre-populated drop-down list options for causes, processes, consequences of severity, and personal injury-related information. The dataset contains 3,626 cases of accidents. The data status varies in terms of complete entries for every available case. Monetary loss information was only entered 109 times out of all cases, and prevention comments were only reported for 365 cases. Description of injury type was entered for only 139 cases. Moreover, there are usually two levels of entries: a general category and one that is more detailed- such as for the injured body part, the category of injury, Specific physical activity, and injury type.

The entered data shows that a number of entries did not belong to a known category, such as in injury type level 1 (310 cases) and specific physical activity level 1 (205 cases). Moreover, it is observed that the level of detail varies between the general level and the more detailed levels of the reported fields. The more detailed levels of “type of work in detail” contain 149 unique categories of entries, the “external factor that affected the incident” contains 159, and the “work process” contains 149.

Although these accidents report mainly describe the accident by pre-populated drop-down lists, the reporters select the causes and other information using their understandings. Dekker (2015) argues that the epistemology of accident descriptions implies that reporters can have different narratives for the same event, depending on multiple factors (such as the reporter’s perspective and experience).

3.3.2. Interviews

The interviews were considered a secondary source of empirical material that was complementary to the accident reports. The semi-structured interviews were conducted in a thematic format to explore and gather information and knowledge about accident reporting. Thematic semi-structured interviews are useful in exploring a particular organizational issue, and are characterized by connecting a series of questions within a particular theme (Cassell 2015). The intention behind the interviews was to gain an insight into the perspectives of the H&S unit on the meaning of safety (in general), the accident response process, the quality of collected reports, and the expectations from an ML-based prototype.

Mainly, the ML-related questions and discussions were formulated based on the business understanding framework of CRISP-DM (Chapman et al. 2000) and the recommended practice (RP) framework (DVN GL AS 2020). The intention for this formulation concerns developing an ML prototype situated within the needs and perspectives of the H&S unit with the explicit purpose of improving awareness of accident prevention measures within the case contracting company.

Interviews were chosen to provide the actor’s point of view on the needs of safety process and site accident prevention. The interviewees were selected based on the mapping of the H&S unit of the case company, as shown in case description section 3.2 and Table 3.

The interview guideline was organized into four thematic sections. The first focused on a background of position and experience, and the second on the meaning of safety and a description of daily safety

processes. The third part included questions about the reporting regarding assigning causes, levels of causation, credibility, quality, and overall value of reporting accidents. The fourth part investigated the potential for improvement in relation to accident prevention, based on learnings from or the utilization of accident reports. The questions then targeted the anticipated added value of a potential ML application, potentially benefitted users, advised propositions, work-process constraints, risks, and ethical considerations.

Table 3. Interview respondents and position

Positions	Respondents
Safety engineer	4
Safety representative	4
Site manager	1
Site supervisor	1
Safety manager	1
Safety strategist	1

3.4. Analysis of empirical material

3.4.1. Interview's analysis

The interviews were analysed using a qualitative method combining Kvale & Brinkmann (2009) approach to analysing interviews with Alvesson and Sköldbberg (2017) reflexive methodology. The themes of the interviews were organized based on the themes of the interview's questions.

- The meaning of safety at the contracting company
- A normal working day
- The response in the event of an accident
- The reporting of accidents
- Status of the data use and safety objectives
- The value of reporting of accidents and improvements
- Improvement in the safety process for accident prevention support
- Value proposition
- ML potential
- Proposals and ML risks
- Satisfaction with the reporting
- Success criteria for a prototype based on the reports' data

The analysis of the interviews was done in an iterative manner, where the interview questions were modified based on the gained insights from the previous interviews. The interviews continued until the responses began to be repeated and reached a state of saturation (Schutz 1972). Drawing on reflexive methodology in this context meant critically reflecting on the respondents' utterances, placing them in an organization and societal context and finally reflecting about the researcher own role and position - inspired by Alvesson and Sköldbberg (2017) concept quadruple hermeneutics.

3.4.2. Machine learning design

3.4.2.1. Understanding the data structure

The dataset can be characterised as relational, namely a collection of records in tabular format (sometimes called "relations") with columns that denote data features, and rows that indicate

individual observations of instances (Martyr and Rogers 2020). The dataset mainly consists of two types of features: structured and unstructured features. The data, in this case, might include a reference number that identifies the instances (Gopal 2018) – in this instance, the case unique number. This allows the features to be searched, filtered, and reorganised (Martyr and Rogers 2020). However, some features are labelled as accident title, description and health and safety category, cause description, comments, and prevention description, and are written by the reporter in free-text format (see Appendix). The free text data type is considered unstructured (Gopal 2018). Structured features can be handled differently than the free text, as the latter requires methods of data mining, NLP or unsupervised ML (Gopal 2018). In this thesis, the structured dataset acts as an investigative step for the predictability of accidents in building the information extraction on the first step of the prototype development and recommend prevention measures.

Another characteristic of the dataset is the definition of input and output features. In an application where an event happens at a specific point in time and in prediction models, data leakage must be prevented (Kaufman et al. 2012). Data leakage is defined as the introduction of information about the target of a data mining problem, from which it should not be legitimately available to mine (Kaufman et al. 2012). The input features chosen in the current case are listed in Table 4. The latter were chosen based on whether the features contained information that could be known before an accident occurred, since the data was generated as an occupational accident reporting. The downside is that most of the data described the event's outcome, which leaves only a few input attributes. Table 5 illustrates which existing feature could potentially be a target output for an ML analysis.

The data is generally nominal (which indicates that they are represented by symbols) – it can be represented numerically by coding the entries by a nominal encoding scheme (Han et al. 2011).

Table 4. Input features

	Input feature	Type of entries
1	Type of work in detail	Nominal
2	Involved substance / chemical	Nominal
3	Employment relationship	Nominal
4	Work environment	Nominal
5	Position	Nominal
6	Company name	Nominal
7	Specific physical activity level 1	Nominal
8	Specific physical activity level 2	Nominal
9	Shift or accident to / from work	Nominal
10	Experience in position (months)	Numerical
11	The last deviating event that preceded the injury	Nominal
12	Work Process	Nominal
13	External factor that affected the incident	Nominal

Table 5. Output features

	Output feature	Type of entries
1	Actual severity	Nominal
2	External factor that affected the incident	Nominal
3	Description of injury Type	Nominal
4	Description of damaged body part, Common	Nominal
5	The last deviating event that preceded the injury	Nominal
6	Injured body part	Nominal

7	Injury class	Nominal
8	Category of injury	Nominal
9	Injury type level 1	Nominal
10	Injury type level 2	Nominal
11	Cause category	Nominal
12	Circumstances of the accident	Nominal
13	Potential Severity - Most Severe	Nominal
14	Risk area	Nominal

3.4.2.2. Data pre-processing

Data pre-processing consists of several major tasks: data cleaning, data integration, data reduction, and data transformation (Han et al. 2011). Data cleaning is usually the first step in pre-processing the data and is done by handling missing values and noisy data (Han et al. 2011).

In the current case, the most interesting output feature was the actual severity. The distribution of the classes of this output was unbalanced and needed further data preprocessing. It is possible to combine the last three levels of injury in one category (called major injury) and the first two classes in another (called minor injury). This combination separates the severity into two categories, while the major injury class starts at the point where accidents result in the absence of a worker from the construction site.

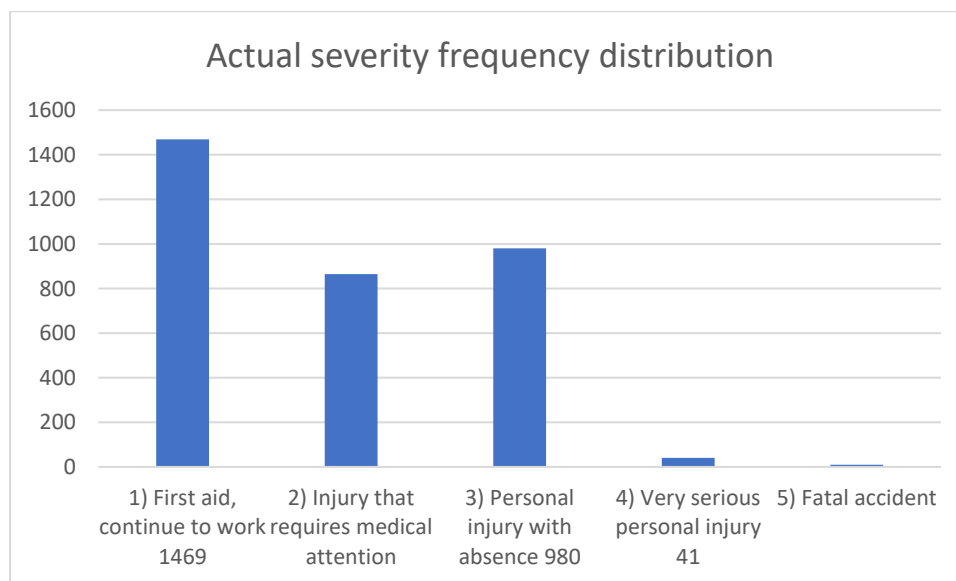


Figure 5. Actual severity frequency distribution

3.4.2.3. Algorithm choice

The goal of the first stage of the prototype is to predict the severity of construction processes. The following criteria influence the choice of algorithm for this first stage of the prototype (see Table 2), based on the purpose and the accident reports data structure:

- **Interpretability:** The algorithm must be interpretable, especially since there is a second step involving the prevention recommendation that is going to be connected to the prediction.
- **Parameter tuning:** A model that depends only on parameter tuning can be problematic because the model is then highly sensitive to the parameter's values – it is preferable that the chosen algorithm is less sensitive to parameter tuning, but it is deemed not as a strict requirement.

- High dimensionality: The high dimensionality criteria are not critical since our data structure is not high dimensional.
- Generalizability: Generalizability is one of the most essential features of ML algorithms, and there is a need for a model that generalizes well, especially for a relatively small dataset such as the one in this case.
- Accuracy: The model must produce high and accurate predictions, especially since safety is the application domain – and therefore, accuracy is crucial.
- Large dataset: The dataset is relatively small (fewer than 50000 instances), which is not a highly important criterion.
- Linearity: It is not known whether the data is linearly separable. Therefore, there is a need to experiment with linear and nonlinear algorithms to test which best classifies the severity level.
- Low dimensionality: The algorithm should perform well with low dimensional data since the dataset is relatively large compared to the number of input features.

Highly interpretable algorithms are NB, SVM, DT, KNN, LR and LogR (see Table 2). However, Only NB, DT, RF, LR and LogR perform well with low dimensional datasets. RF and KNN have very good generalization abilities. KNN, RF and LogR are usually good for this criterion in terms of accuracy. Only the LogR is suitable for building a linear model. Based on this breakdown of the algorithms and their characteristics, the chosen algorithms are KNN, DT, LogR and RF.

3.5. Researcher own role

In line with Alvesson and Sköldbberg (2017)'s suggestions, all activities in this licentiate thesis were exposed to a critical reflection of my own role and identity vis a vis not only collaboration partners, literature, interview respondents and supervisors, but also in reflecting and analysing literature, developing analytical insights and discussion, even when arriving at the main results. Being a middle-class woman with a middle east background involves advantages and disadvantages. Particular Swedish construction industry traits are more visible for externals and can be identified comparing with the researcher own background. On the other hand, social group differences between university employees and building sector professionals would constitute more of a disadvantage, given the mutual stereotyping of academics and site professionals. In sum, these conditions are at a time enabling and constraining the research in characteristic ways.

3.6. Ethical considerations

There were a number of ethical considerations that shaped the research design. The use of the data by the licentiate team was governed by a non-disclosure agreement. The data transfer was only performed through secure channels managed by the data owner. The case company and respondents remain anonymous. Moreover, due to the sensitive nature of the data and the researchers' use of ML for generating solutions that target specific individuals, the author chose not to consider any personal or enterprise information in the ML data analysis.

4. Summary of the papers

This part presents a summary of the collection of papers included in this licentiate thesis.

4.1. Paper I: A REVIEW OF MACHINE LEARNING FOR ANALYSING ACCIDENT REPORTS IN THE CONSTRUCTION INDUSTRY AND APPLICATION REQUIREMENTS (under review at The Journal of Information Technology in Construction (ITcon) submitted on 2022-02-10)

Artificial intelligence (AI) and ML have become more popular in solving construction management problems (Pan & Zhang 2021). Applied ML reviews have shown advancements in safety management and knowledge extraction by examining accident records (Pan and Zhang 2021, Hon et al. 2021). However, the literature still lacks a comprehensive analysis of what developing and applying ML entails in the domain of accident reports analysis within the construction industry. It is not clear what the implantation of ML-based analysis in a contracting company might require. This paper aims to answer the research question: what are the requirements of an ML model based on accident reports data to be implemented in occupational safety in a contracting company?

This paper contributes to the identification of prerequisites of ML development that arise from the specific conditions and the processes associated with managing H&S in a contracting company in the construction industry. The research question was answered by a literature review conducted using the concept-centric framework augmented by units of analysis (Webster and Watson 2002). It was based on searches related to the application of ML to the analysis of accident registries in the construction sector. The organization of the review was done to synthesize the literature into appropriate units of analysis, namely data characteristics, data pre-processing, algorithm type and training the ML model, testing algorithm performance, and implementation of ML analysis. Three citation indexes were selected: Web of Science, Elsevier, and Scopus. The review was conducted iteratively within the three databases and within Google Scholar by using the search terms “accident report,” “construction industry,” “machine learning”, and “construction occupational safety.” Nineteen articles were finally selected, four of which were found in all the searched databases.

The analysis of the literature showed that multiple requirements are necessary. One of the most important requirements is a careful implementation strategy that considers existing safety processes and their relation to other in-place processes such as design and project planning. Thus, the implementation of ML-based models requires feasibility and implementation analysis - in a prototype format, for instance - and the involvement of practitioners. Another crucial requirement is ML performance measurement and evaluation to assess the performance metrics and accuracy threshold.

Risk critical application such as safety and, more importantly, accident analysis imposes higher requirements of accuracy and trustworthiness in applied ML solutions. Accidents have been shown to generate imbalanced data in terms of accident types, causes and severity. The ROC was proposed as an ML performance metric because of the visualization benefits for comparing different combinations of errors. The ML classifiers that have lower error rates for a specific class can then be chosen (Gholizadeh et al. 2018). This proposed approach was shown to be especially beneficial in maximizing the prediction accuracy of minority classes in unbalanced data sets in the construction accident reports data (Gholizadeh et al. 2018).

Overall, implementing ML in the construction industry, needs a standardized development method, notably due to the difficulty in assessing the best approaches in data pre-processing and applied

algorithms. However, ML algorithms that are easily interpreted were found to fit the safety context because they allow for understanding ML results. The word embedding in the data pre-processing showed a pattern and potential improvement within the domain-specific corpus. Future research should experiment and conclude whether domain-specific dictionaries should be used in word embeddings in the pre-processing stage. Finally, there is a need for theoretical frameworks for guiding the contextualization of causal factors. This would assist when developing safety solutions and when they are being deployed further from centralized computing platforms for real-time decision-making support.

4.2. Paper II: A COMPARISON OF ACCIDENT CAUSATION MODELS (ACMS) AND MACHINE LEARNING (ML) FOR APPLIED ANALYSIS WITHIN ACCIDENT REPORTS

ACMs are theoretical frameworks and have had an impact on accident causation analysis. However, ACMs were not sufficiently addressed in the literature of applied ML in accident records analysis. ML-based analysis has been criticised for lacking interpretable recommendations, data quality issues, clear implementation cases, the integration with domain knowledge (Vallmuur 2015, Bilal et al. 2016), and generalizability (Xu et al. 2021, Sarkar and Maiti 2020). On the other side, ACMs can be categorised into many types characterised by different causation logic and focus of causation categories. The current literature on ML applications within the domain of accident analysis does not integrate ACMs as theoretical frameworks into the ML model development and analysis. The authors of this paper also assume that analysing accident reports using ML can contribute to learning about ACMs as well as occupational accidents. This research investigated the question of what ACMs can contribute to the ML results of analysed reported accidents in the construction industry, and what can be learned about ACMs from the application of ML in this domain. This paper contributes to conceptualising ML models through the lens of ACMs.

This paper is based on a desk study of the literature of applied ML in the analysis of construction accident reports and ACMs. The ML models are based on a literature review and the systemisation of the purpose of the ML, the included features, and the ranking of important factors. The themes are presented for an in-depth analysis. ACMs were selected based on crossing the models which were reviewed by Kjellen and Albrechtsen (2017), Fu et al. (2020) and Woolley et al. (2019). Three models were selected based on the types of ACMs and their common application in the construction industry.

ML analysis of accident reports usually results in components that are predictive of accident types or severity levels. The comparative study illustrated that the components extracted by ML could be compared to the typology of the BOW-Tie model and the SCM. However, one major difference was found in ML components in that they lack prevention measures which are a bottom-line building block in ACMs and, consequently, accident prevention. However, the levels of causations were found to be mostly those remaining close to the workplace and human behaviour factors. At the same time, ML results rarely included factors that are related to the higher levels of decision making within the organisation.

The lack of prevention measures or the inclusion of higher levels of causation factors is not necessarily a drawback of ML itself but the reporting that has repeatedly been missing the prevention measures suggestions. The more accident analysis considered factors further from the event, the harder it gets for further factors to become apparent in terms of their effect on the event. Furthermore, the mechanism

causing accidents seem to differ between the representation of ML and the SCM. Nevertheless, such a comparison remains ambiguous and need visualisation to make more conclusive comparisons. Finally, ML models in the reviewed literature tend to highlight severity as an outcome that needs to be predicted. In contrast, ACMs focus on accidents as events that occur regardless of their level of severity.

The paper concluded that the ML analysis of accident reports needs to be guided by ACMs to be useful in real-life implementation. This combination contributes to making sense of ML-based recommendations of accident prevention measures. As much as a prediction of the event of an accident or the magnitude of the consequence might seem to be preventive, in the future, ML analysis might be better utilised in the modelling of risk factors. The integration with ACMs such as the BOW-Tie model and the SCM provides the backbone for ML-based accident analysis models to be tuned towards accident risks and any corresponding prevention measures. From an ML point of view, explainable algorithms should be used. The conclusions of this paper for approaching the ML-based accident analysis provides a possible recipe for better understanding causation factors and the mechanisms through which accidents happen.

4.3. Paper III: LEARNING FROM ACCIDENTS: MACHINE LEARNING PROTOTYPE DEVELOPMENT BASED ON THE CRISP-DM BUSINESS UNDERSTANDING

The increased interest in ML evident in the literature and the growing use of ML in accident statistical analysis has been shown to be valuable in analysing large volumes of data. However, it is not beneficial to reinvent existing methods, so in truth, no new knowledge is provided by such solutions. Moreover, the analysis of the literature contained in paper I have shown that there is a need to contextualise understandings of ML-based analysis and define clear ML tasks. This paper explored the local and corporate context for ML-based analysis and the ML development method known as CRISP-DM for conducting such studies.

The aim of this paper is to analyse experiences and challenges in using the “business understanding” phase of CRISP-DM as the first step towards ML prototype development with respect to the context and local dynamics of a Swedish contracting company. The investigation adopted a bottom-up approach, where knowledge of accident registration procedures was the point of departure.

The overall method is an interpretive approach. A concept-centric literature review was conducted (Webster and Watson 2002) to review the status of ML-based solutions for accidents report analyses. For the empirical context, five interviews were carried out: four with safety engineers and one with a safety strategist at a high level in a Swedish contractor company. The ML related interview questions and discussions were focused on gathering the safety requirements for developing a data-driven prototype, inspired by the business understanding framework of CRISP-DM and the recommended practice (RP) framework (DVN GL AS 2020).

The business understanding phase begins with defining the client’s goal and deciding on a value proposition for the ML application. The interviews showed a difference in safety priorities between top management and operational level, especially the focus on behaviour and fatal accidents, the planning of safety tasks, and communication. This leads to conflict between single versus multiple-goal orientation compared to the CRISP-DM model, which suggests that a single goal should be identified. In response to this obstacle, this paper suggested that at the end of the first step of the

business understanding phase, there is a need for an intermediary step to agree on a common objective before proceeding to accident report analysis.

The second step of the business understanding phase requires a detailed analysis of the related resources, constraints, assumptions of the business objectives, risks of project failure, terminology, and cost-benefit analysis from a commercial perspective. The most important aspect of this application is to investigate the resources of the H&S unit and the characteristics of the data. The data incorporates valuable information, but the level of detail in reported accident causes is doubtful due to different experience levels among the personnel that do the reporting. Moreover, constraints can be found in the digital reporting system and the incorporation of safety planning between production's main objectives of meeting schedule and budget demands.

To sum up, following the recommendations of this business understanding phase reveals insights into possibilities and local constraints. However, it is not possible to cover all scenarios, especially if the first step of the business understanding phase was not concluded or aspects of the ethical consequences were very challenging to be identified through the interviews.

The following step would ideally be defining data-driven goals. These would include the ML prediction output and the model's acceptable accuracy. By the time the analysis arrived at this stage, this ideal had become more unattainable since the last two steps had not been completely closed. Moreover, the requirements of this step are highly dependent on the data condition. Thus, it is suggested that this step should be completed by adding an iteration as primary data analysis. This would then suggest realistic potentials and limitations to match the organisation's aspirations.

The previous analysis highlighted the application of CRISP-DM in the business understanding phase and the involvement of domain experts in a breakdown of daily processes and experiences. In project-based organisations such as the case contracting company, there is a need to investigate and analyse the business understanding phase on different organisational levels. The different organisational levels and their concentration on a very different set of priorities challenges the use of CRISP-DM. Moreover, the analysis showed that adding two intermediary steps was necessary to meet the challenges in defining ethical considerations, application design and data-driven goals.

5. Machine learning model design and analysis

This section describes a continuation of the thesis including the ML model design and analysis of accident reports. This section is organized after the summary of the papers (section 4) because the following analysis builds on the results of the previously described papers.

5.1. Model design

The following analysis encompasses a continuation of interviews based on the business understanding phase and a follow-up recapitulation of the proposed solutions. Seven further interviews were conducted by following the same main structure of the interview guide as in paper III. The conclusions of Paper III indicated the need to agree on a common organisational objective for the ML prototype design. The interviews which were conducted in paper III (section 4.3) were extended to include more actors from the H&S unit, including one site supervisor, four safety representatives, one safety manager and one site manager. The interviewees' collected ML application proposals are presented in Table 6.

Furthermore, a workshop was planned to discuss how the accumulated propositions identified by analysing the interviews could be integrated into the preliminary prototype development. The workshop included a presentation of the results of paper III and Table 6 as the intermediary step - suggested in paper III- of the business understanding phase.

The workshop included the following actors:

- Safety engineer from the contracting company (2)
- Trade union organiser
- Safety engineer contractor (1)
- Business Development Lead – Analytics contractor (3)
- Development leader Health and Safety contractor (3)
- Construction Workers' Union agent
- IT Solutions Manager contractor (3)
- Work environment manager contractor (2)

Table 6. Summary of interviews model propositions

Machine learning model propositions
To produce statistics on historical accident cases.
To pay attention to work steps where there are many accidents
To use Synergi more easily for safety work preparation and risk assessment
Tools for presenting information about safety risks to production people
Negative and positive observations to find the reasons behind workers not following the safety rules.
No safety improvement needs.

The workshop was organized in an online meeting and facilitated by the author of this thesis. The workshop resulted in a vote for the proposition that was considered the most value-adding for safety prevention, based on the ML analysis of the collected data from accident reports. Most of the workshop

participants thought the proposition for work steps and safety risk assessment was the most value-adding use of ML (see Figure 6).

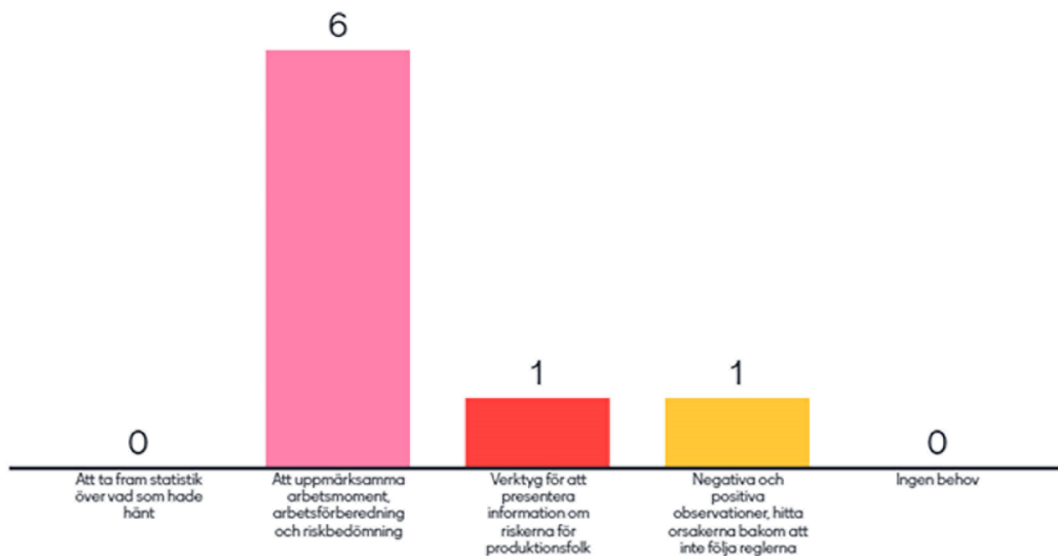


Figure 6. Workshop vote for ML model proposition.

The workshop then presented the preliminary data analysis of the case contractor's accident reports (section 3.4.2.1) and the corresponding prototype design (Figure 7). The prototype design consists of the input features, and the blue arrows represent a drop-list of predefined categories. The categories should be identical to those that constituted the accident reports for consistency. The prototype was designed to predict severity that is categorized as low risk and high risk. The high-risk category represents a prediction of a final outcome starting from the absence of workers towards outcomes of further severity.

Type of work in detail	▼
Involved (d) Substance / chemical	▼
Employment relationship	▼
Work Process	▼
Work environment	▼
Position	▼
Company name	▼
Specific physical activity level 1	▼
Specific physical activity level 2	▼
Shift or accident to / from work	▼
Experience in position (Months)	▼
External factor that affected the incident	▼
The last deviating event that preceded the injury	▼

Prediction output

Risk assessment ● Low risk ● High risk

Figure 7. prototype design illustration.

5.2. Framework of understanding

In response to the conclusions of paper II (see section 4.2), the Bow-Tie model was chosen as a framework of understanding for the ML model. In this section, I will categorise the selected input features for the ML model into the components of the Bow-Tie model (see Table 7).

The BOW-Tie model (see Figure. 8) consists of multiple components characterising accidents. The model analysis starts by identifying a hazard in the organisation or the surrounding environment. I interpret that the input factors “Type of work in detail”, “Shift or accident to/from work”, “Experience in position (Months)”, “Employment relationship”, “Work environment”, “Position” and “Company name” as representatives of the surrounding conditions of the work process. The hazard component is directly connected to the top event (see Figure 8), which I interpreted as the occurrence of an accident. The top event is at the centre of the BOW-Tie model and is caused by what the BOW-Tie model categorises as threats. The latter is represented by the input features “Specific physical activity level 1”, “Specific physical activity level 2”, “Involved (d) Substance / chemical”, and “The last deviating event that preceded the injury”. The “External factor that affected the incident” feature was interpreted as an escalation factor. Moreover, the consequences are interpreted as the level of severity of the accident. The input factors are summarised in Table 7.

Prevention barriers are very important components of the BOW-Tie model, and they are also reported in the accident reports. They are entered as free text, and the prototype design does take free-text data into consideration in this analysis (see section 5.1).

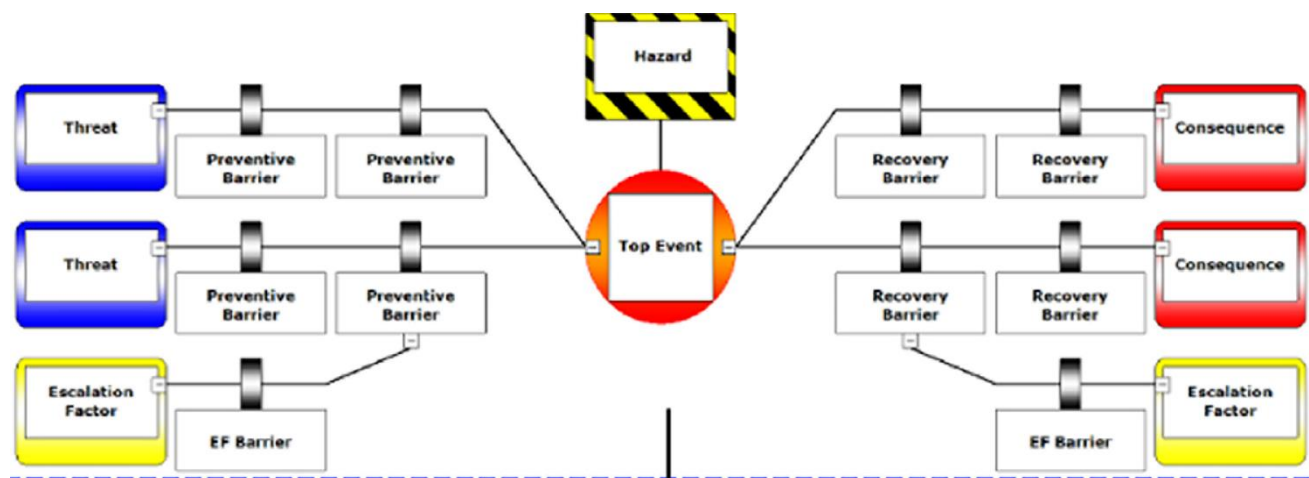


Figure 8. BOW-Tie, Fu et al. (2020)

Table 7. Input features categorization into the BOW-Tie framework.

	Input feature	
1	Type of work in detail	Hazards
2	Shift or accident to / from work	
3	Experience in position (Months)	
4	Employment relationship	
5	Work environment	
6	Position	

7	Company name	
8	Specific physical activity level 1	Threats
9	Specific physical activity level 2	
10	Involved (d) Substance / chemical	
11	The last deviating event that preceded the injury	
12	External factor that affected the incident	Escalating factor

5.3. ML analysis results

The input features represented in Table 8 were used to predict the level of accident severity in two different settings: the reported actual and potential severity. The actual and potential severity levels were reported and can be used as a prediction target. The potential severity represents the degree of severity of an accident that could have happened; for example, a minor accident requiring first aid medical attention could have resulted in a more severe injury and might have led to the worker having to take days off work. The ML model was performed by myself using the Pandas and Scikit-learn - Python 3.9.7 libraries version. Since almost all input features were nominal values, the data was encoded with the sklearn LabelEncoder function. The output features had five classes, which I then processed to make only two classes for binary classification. In particular, the first two initial severity levels “First aid, continue to work” and “Injury that requires medical attention” were merged into one category of low severity, and the highest three categories “Personal injury with absence”, “Very serious personal injury”, and “Fatal accident” into one high severity category. The results of the predictions are presented in Table 9.

The data analysis showed that the imbalanced state of the data had a considerable impact on the classification of severity. If we take the confusion matrix as a metric, the best prediction the DT algorithm can achieve is 383 true-positive cases of high severity, compared to 645 false-positive cases. However, if the accuracy is considered a metric, it was noted that it is not representative of how well the model classifies major and minor accidents. Since the interest here is to predict severity and the more critical one resulting in severe accident impact, accuracy alone is not enough as a metric. A good example here is the RF algorithm. The algorithm’s accuracy is 69.29%, while the confusion matrix shows that the classifier almost always assigns minor severity to the case. The results of classifying the potential consequences showed the models’ tendency to classify most cases as severe accidents – which is the most populated class in the potential severity case (see Table 9).

The unsuccessful prediction results can be attributed to the class imbalance, but they might also be attributed to the features. The features might be loosely correlated to the output, which explains the predictions. To test the features’ predictability, I performed a random under-sampling for the high populated class to classify actual and potential severity. The random under-sampling reduced the frequency of the high severity class to match the frequency of the low severity class. This resulted in a slight reduction in accuracy but improved the ROC metric due to the more balanced confusion matrix. According to the confusion matrix metric, the predictions of the undersampled data showed slight improvement. This result indicates that the balanced data is not the only problem for classifying the target values, but also that the features are not correlated with the output.

The results of severity prediction illustrate that the proposed prototype design (Figure 7) can not be realised based on the ML model design and analysis of this thesis.

Table 9. The results of ML algorithms classification of severity.

Under sampling				
Actual severity prediction	RF	LogR	KNN	DT
Accuracy %	60.02	56.57	55.01	53.94
ROC	0.6002	56.57	0.5501	0.5394
Confusion matrix	[622 406] [416 612]	[569 459] [434 594]	[579 449] [476 552]	[560 468] [479 549]
Original data form				
Accuracy %	69.29	69.05	65.33	60.12
ROC	0.5549	0.5102	0.5324	0.5373
Confusion matrix	[2123 209] [823 205]	[2273 59] [981 47]	[1968 364] [801 227]	[1637 695] [645 383]
Under sampling				
Potential severity prediction	RF	LogR	KNN	DT
Accuracy %	55.68	51.9	51.65	52.75
ROC	0.5568	0.519	0.5165	0.5275
Confusion matrix	[758 607] [603 762]	[1086 279] [1034 331]	[696 669] [651 714]	[713 652] [638 727]
Original data form				
Accuracy %	60.69	62.26	57.08	54.26
ROC	0.5443	0.5034	0.5187	0.5177
Confusion matrix	[397 968] [457 1803]	[28 1337] [31 2229]	[420 945] [611 1649]	[569 796] [862 1398]

6. Results and discussion

The overall aim of this thesis is to investigate how ML-based methods and techniques could be used to develop a research-based prototype for occupational accident prevention in a contracting company. The thesis focuses on exploring development processes that bridge ML data analysis technical part with the context of safety in a contracting company. The overall research question was formulated with the focus on accident prevention and H&S activities on-site, with the company being the case for the prototype development. The following four sub-questions were sequentially investigated and critically reflected upon to answer this research question. The following discussion around the theoretical framework and the data analysis is structured with these research questions in mind.

RQ1: What are the requirements for applied ML in the domain of accident prevention in a contracting company's occupational safety processes??

The review of current ML literature found that when analysing accident reports, a number of challenges originated from the characteristics of the data in terms of data format, availability, and content. Accident reports, which were discussed in the reviewed literature, existed in textual format and lacked labels. In the case of high volumes of data, the accident description content and causes were therefore not easily understood. On the other hand, accident reports in predefined reporting categories clearly illustrated the reported features. However, they often had shortcomings on the level of causation, as they mainly reported the factors close to the physical work environment. Only a few datasets included distal causal factors, such as the type of construction and project size (Choi et al. 2020) and monthly project-related attributes (Poh et al. 2018). The shallow description of causes in the literature studied is one of the most disrupting challenges because it indicates that accident reports do not possess sufficient detailed causation capacity to explain why accidents occur.

Ultimately, the data characteristics determine a considerable part of the data pre-processing step. Accident reports are characterised by their use of language and domain-specific terminology. Although the literature review did not reach a definite conclusion about the application of domain-specific NLP, the critical literature analysis suggested that word-embedding algorithms trained with domain-specific corpus achieved good results in pre-processing accident reports (Zhang 2019, Zhang et al., 2020, Baker et al. 2020).

Moreover, accidents happen at different frequencies – particularly severe and fatal accidents which are rare compared to the high frequency of reported minor injuries. The reviewed literature mostly used a range of data resampling methods to counteract the less frequent accidents in the reported accident dataset, such as Random Over Sampling (ROS), Synthetic Minority Oversampling Technique (SMOTE), Random Under Sampling (RUS), inversed proportional weights, and manual labelling. However, the reviewed literature lacked a justification for the choice of methods, and the consequences of using such methods were not considered. This leads to difficulty understanding ML models' results for them to be applied in real-life situations.

One of the essential recommendations arising from the ML literature review is the need for a systemised method for accident analysis with ML. This originated from the technical aspects of ML modelling and the need to integrate the theoretical and domain knowledge of safety practices. Considering the development context is an applied ML development general requirement and not exclusive for ML accident analysis. Nevertheless, it is suggested that the context around the data and

code systems, the data analysts, and the organisational expertise are the missing pieces of an ML modelling lifecycle (Garcia et al. 2018). Thus, in the construction safety domain, elements of digitalisation, change management, feasibility and implementation analysis depend heavily and necessarily on the involvement of domain experts.

The successful application of ML in the domain of safety requires reliable evaluation methods. More research is needed in the area of evaluating ML prediction models. The ROC was proposed as an efficient metric for maximising the prediction accuracy in construction accident reports, given the asymmetrical frequency of serious/non-serious accidents. Nevertheless, more research is needed in evaluating applied ML in construction safety, including external validation and implementation trials.

It is important to note that this literature analysis has been influenced by my thinking about the problem I intended to solve. I have been interested in finding the best way to conduct an ML model for analysing accident reports. It is not necessarily so that categorising the literature into main themes such as the data characteristics and the implementation of ML is original, but it has been shaped by my wish to find themes that could help my research. This probably explains my somewhat performative language use. Moreover, this observation also relies on being explicit about the fourth level in the reflective methodology, the researcher's role (Alvesson & Sköldbberg 2017).

RQ2: What is the role of ACMs as a theoretical framework for the ML results of analysed reported accidents in the construction industry, as well as what can be learned about ACMs from ML?

The accident causation models (ACMs) (i.e., the BOW-Tie, the SCM, and the STAMP model) were compared to the literature that applied ML to analyse accident reports in the construction industry. This comparison considered their level of causes, the relationship between causes, and the predictability of severity. This contributed to a re-conceptualising of ML-based models through the lens of ACMs. The study in paper II concluded that ML could benefit from integrating accident report components into the components of ACMs. The benefit is derived from conceptualising extracted features from free text and providing a foundation for prevention measures. Rule-based data mining and feature extraction methods were found to have shortcomings due to features and rules being prepared by a human and the weak generalisation of results (Pan and Zhang 2021). The SCM, the Bow-tie model, and several ACMs categorise accident causes into predefined categories. These provide somewhat well-defined causes levels – I say somewhat here because I have a reservation on how well accident causes are defined in ACMs. Nevertheless, ACMs provided a reference point that partly alleviates confusion of interpreting free text data used in ML-based accident analysis. I would say the same about accident reports in pre-populated format. These might have their implicit theory, and to use ACMs components to recategorise reported accident causes may clarify that.

There is also a missed opportunity to reflect on ACMs from the perspective of ML analysis. Methods such as data mining (Zhong et al. 2020) and semantic roles and rules analysis of accident components (Kim and Chi 2019) have visualised the relationships between causal variables. Generally, this contributes to creating a link between accident types and accident consequences. However, further research is needed to understand the nature of the relationships between causal variables and investigate if the levels of causation contribute to accidents equally.

A major difference was found by comparing the analysed ML literature and the SCM and the BOW-Tie. It was found that the approach taken by ML modelling to predict the severity of accidents is

contradictory with accident analysis and causation models that assume the outcome of accidents as involving an unpredictable stochastic element (Harms-Ringdahl 2013). Although the ML literature claimed success in severity predictions with internal validity (i.e., ML model accuracy and not in applied real-life situations), their results were not always consistent and did not show proof to counteract the assumptions of ACMs regarding the stochasticity of accident severity. I must say I would like that severity could be predicted. Maybe the academics who work in the same domain wish for that too. This might be a reflection of the desire to protect on-site personnel from a dangerous situation. One should be aware of such inclinations because they could lead to the opposite, such as increased exposure to the danger of minor accidents if predictions introduce overweight on instances leading to fatalities. This observation echoes Alvesson & Sköldbberg's (2017) observations on the role of the researcher and the structural, societal distance between ML developers and the building site.

RQ3: What are the experiences and challenges of applying CRISP-DM “business understanding to assure a solid contextual embedding and an appreciation of local dynamics?”

Based on the conducted literature review to answer RQ1, there was a recommendation to develop ML prototypes concerning the local context dynamics systematically. The CRISP-DM was investigated as a possible method. The application of this method by doing interviews, although posing relevant questions, was found too general to finalise a business understanding without adding multiple iterations. Nevertheless, the interviews revealed interesting insights into the safety processes and the perception of H&S personnel within the organisation. It is important to note that the description of the CRISP-DM method does not particularly advise doing interviews or a specific way for conducting the business understanding analysis. The decision to conduct interviews was my interpretation of using the CRISP-DM. It was somewhat challenging for me to start thinking about a prototype without understanding the existing safety processes and who would be using it.

While the respondents agreed on the meaning of safety (“everyone goes home injury-free”), it seems that ideas about achieving that goal were not as clearly a part of safety meaning. This indicated describing safety as the goal to be injury-free. There was much focus on the planning and preparing for safety measures on-site. Accordingly, processes were in place with a particular focus on fatal accidents and the behaviour of individuals. Individual risky behaviour was the shared major cause among top management and the safety engineers. However, another major cause mentioned by safety engineers and site managers was thought to be related to production time pressure and referring to contractual arrangements as an inevitable, unchangeable condition that produces safety risks.

There was also a prevailing assumption that effective accident prevention is achieved by systemically identifying the risks associated with accidents and taking measures to avoid their impact. Furthermore, the interviewees expressed conflicting views about risks associated with the prevailing safety assumptions about behaviour and the systemised risk analysis. These commonly held assumptions and associated risk evaluation techniques were often criticised in the literature for low inter-rater reliability, i.e., low degree of agreement among observers/raters/analysts in estimating frequencies and consequences (Harms-Ringdahl 2013). Analysts/raters tend to assume that a major consequence is automatically less probable (Harms-Ringdahl 2013).

The interviews showed a need and potential added value to accident prevention activities by improved safety planning and more accessible risk identification. It seems as if there is frustration with

anticipating what would cause the next accident. It might be because some safety professionals have a perception that all that can be done to prevent accidents is known and well-established, but accidents still occur. Some of the respondents pointed to the need to know why allegedly workers do not follow the safety work packages. They attributed that alleged behaviour to the workers' tendency to prefer to do the work before thinking about safety. So, considering workers' behaviour as a significant cause of accidents makes much sense for those safety professionals who deem the rules known and sufficient.

Ultimately, the CRISP-DM is applied to collect and understand the context requirement and identify the business's expected advantage of using ML data analysis. However, the interviewees' experiences and views indicated complexity in defining a goal that solves an existing problem -from their perspective. This presented a difficulty in deciding on an ML utilisation goal and more so to analyse the implications and more specific ML prototype design requirements. Paper III concluded that CRISP-DM might benefit from adding two iterative steps to the existing ones in its process.

RQ4: What are the predictive attributes of accidents based on ML application to accident reports?

Based on the ML data analysis in this thesis (see section 5), the pre-populated categories of accident description were the point of departure for the ML model design. However, in the accident reports data of the case contractor, only a few reported features could represent knowledge before the accident took place (such as "work environment"). Compared to those features registering the consequences such as the description of the damaged body part and description of injury type and cause category.

In this thesis's accident reports analysis (section 5.3), the same phenomenon of the low frequency of severe accidents was encountered. The ML model classified most accidents into the low severity category, which was the most populated class. Random undersampling (RUS) was employed to test the impact of the uneven severity frequency. However, it was found that the difference in frequency did not explain the ML model's results since the use of RUS for the more populated class did not result in a considerable improvement in the classification performance.

This result indicated that the same features' entries that characterised the work environment for a high impact accident were the same for the low impact ones, according to other accident research. It is important to note that the prototype design and the choice of severity as an output for the ML model were impacted by the available features within the case accident reports and influenced by my own inclination to predict severity. This finding implies that research on analysing accident reports by ML needs new ways of thinking by approaching the analysis differently. More research is needed about this development phase, and only a little research can be found in response to the methodological limitations for handling severe accident reports data. A systemised data pre-processing approach that implements clustering, chi-square test and principal component analysis (PCA) has been proposed in earlier published studies (Lee et al. 2020).

In this thesis, the BOW-Tie model was used to understand the consequence prediction ML analysis. It was found that the data organised as hazards and threats did not differentiate between which work conditions, physical activities and deviating events explained the level of severity in the event of accidents. This result is in line with basic assumptions in the BOW-tie model. This result raises two critical questions. One involves safety planning and what, in fact, the H&S unit knows about the

accidents that could be used in planning and prevention. This, in turn, relates to whether an overall prevention strategy is being implemented on-site.

The second important question involves the BOW-Tie model, or in general, whether the ACMs' assumption to prevent accidents by systematically analysing the risks is just an illusion. Building on the interviews with the staff of the H&S unit, this is probably true since general safety practices and accident analysis are more concerned with preventing the consequence/severity of an accident than the event itself. The ML analysis of accidents reports and the prediction model that was undertaken in this thesis supported this discussion. By using the reported potential accident consequence instead of the actual severity as a prediction target, there appears to be a stochastic element in accident outcomes. On a general note, there has been a change in the view of risk evaluation beyond probabilities and consequences and more towards a decision-making process that considers a broader view on the context of risk (Harms-Ringdahl 2013).

Intuitively, it could be assumed that the reported potential severity might alleviate some of the stochasticity in accident severity. However, the experimental ML analysis in this thesis (see Table 9) showed otherwise. This might be explained by the reported potential severity being overestimated. Most of the potential accident severities were estimated to be just in the next severity level above in the case accident reports. Although these results do not seem encouraging, accident reports and the application of ML could support other purposes instead of accident impact prediction. Such purposes could include causation modelling and extraction of unique accident cases that might present new knowledge. The case accident reports contained other data types such as causes and prevention measures and accident descriptions in free-text format (see Appendix). This type of data was not used in this thesis, but it can potentially be used in independent research that considers another ML model design.

Accident causation analysis of accident reports using ML provided an added value by giving means for efficient extraction of information. ML algorithms search for high frequency, often repeated patterns, and this approach is not compatible with finding new knowledge about all types of accident occurrences. The domain of accident prevention is mature, and much is known about accidents causation combined with developed ACMs that have evolved to include further levels of causes beyond the workplace and human behaviour. Therefore, expecting the next accident might not be associated with analysing frequent accidents cases but with discovering emergent risks. ACMs provide a stable foundation for putting emergent risks into perspective and support prevention strategies. However, ACMs seem to have reached a ceiling of causation levels and their representation. Thus, more development is needed in understanding the nature of the relationships between causes instead of adding further categorisations and levels of analysis.

Finally, according to my definition of a prototype" as the one suggesting a precise implementation for ML-based data analytics. One that shows means of application and a digital software interface that is ready for use". I can say that this thesis did not realise the aimed and designed prototype. Mainly, the results of the ML modelling of the case accident reports were not successful in being taken further to an implementation stage.

7. Conclusion

This thesis contributed to the exploration and understanding of ML development processes to bridge ML data analysis with the context of safety in a contracting company. The thesis has provided a method for choosing an ML algorithm based on the required criteria of the ML model. Moreover, the thesis discussion argued for the use of CRISP-DM as a method for understanding the context and gathering potential use of ML from the business perspective. ACMs were also essential in ML model interpretation, especially in identifying components of accident analysis and categories of causes.

ACMs provided a background for accident investigation and causation analysis for many years and developed over time. This has resulted in several classifications such as linear (e.g., SCM) and non-linear (e.g., Bow-tie model) models according to the assumed logical sequence of events that lead to accidents. Other classifications also exist, but ACMs have been divided into groups based on different stages and causations. The simple linear models attributed accidents to physical/mechanical and human errors. ACMs then became associated with complex linear models as they increasingly considered the interaction between latent organisational factors and unsafe behaviour. Complex non-linear models encouraged a broader view of system-related factors in response to the growing complexity and tighter couplings within industrial domains. They now explain accidents as being caused by the dynamic and non-linear interaction among multiple factors within the entire system, including political and regulatory factors.

ACMs evolved to include higher levels of causation. Moreover, ACMs assumed stochasticity in accident severity. Behaviour and advanced socio-technical and cultural models were used in the relevant domain literature in the construction research context, while the system-based models were hardly ever applied. Accidents continue to occur in the construction industry, and there is a need to investigate theories and models of accident causation against the quantitative data analysis that is now being derived from many registered accidents.

The ML-based approach to accident analysis includes supervised, unsupervised and semi-supervised learning. Unsupervised learning is a method of data exploration or description employed when there are no specific preassigned labels for the input or output features. In comparison, supervised machine-based learning depends on mapping an already labelled input to output. Semi-supervised machine-based learning consists of a combination of the latter two approaches. Supervised ML algorithms can be further categorised into linear and non-linear algorithms, each having different characteristics, strengths and weaknesses. The algorithms were organised based on their characteristics (e.g., interpretability, accuracy, generalizability).

This thesis has employed an overall qualitative-interpretive reflexive methodological approach. This approach combined different levels of interpretation. ACMs and ML accident analysis were chosen as the main theoretical frameworks for answering one overall research question and four related sub-questions. The associated empirical research included collecting accident reports and interviews conducted within the H&S unit in a contracting company. The CRISP-DM and ML algorithms were employed to develop and analyse an applied machine learning model. ACMs, ML algorithms, and CRISP-DM provided the desired multiplicity in interpreting the empirical material. The practical research method consisted of four sequential steps, including three papers and the ML-based analysis of accident data.

ML-based accident analysis can create knowledge about accidents in a contracting company or other organisations, such as objects and combinations of situations that cause accidents. The literature in this context showed various means by which the application of ML algorithms can enhance knowledge about accidents. ML models can be applied in severity estimation, accident type classification, information extraction and safety training scenario generation. However, the literature also showed that extracting new knowledge about accidents in a contracting company was hindered by an array of challenges. Mainly the systematisation of the ML process, the feasibility of implementation, and the focus on severity prediction.

The reviewed literature of applied ML in accident report analysis indicated the need for the standardisation of the development process regarding the feasibility of implementation and evaluation. However, implementing the CRISP-DM as a process did add essential components to understanding the context, such as context requirements, assumptions about safety processes and accident prevention. Although the CRISP-DM was found too general to provide specific guidelines for ML prototype development, it provided a backbone for the application domain. Further decisions on the ML system design can be established with a flexible-iterative model design process.

ACMs served as a theoretical framework for conceptualising reported accident features and understanding ML-based analysis and interpretation. It was concluded that the reported features that described the work environment do not explain severity. Although the domain produces less severe accidents that are not aligned with how the machine learning classification algorithms work, this was not the primary problem. The primary problem lies within the direction of the ML related literature to predict severity, which is stochastic.

Although ACMs have guided accident investigation and promoted successful prevention strategies, the promise of risk mitigation by systematically analysing accident risks has been undermined by the difficulties around the identification of unknown and emerging new types of risk. ACMs assume that the essence of prevention is by systemising risks and causes. However, in this mature field of study, what is needed is to understand better the rules that govern the relationships between emergent new risks. Accident report analysis using ML offers methods and means in this area. Data mining and unsupervised ML are proposed as a possible way forward to meet this ambition in that they are less explored in the ML models considered within the current literature.

This study suggests that there is a need for systemised machine learning modelling methods for analysing accident reports. Systemised methods should consider integrating an applied ML model within the context of domain experts responsible for implementing prevention measures and strategies. Moreover, there is a need for a development method that systemises the technical part of data pre-processing and the choice of algorithms along with the needed internal and external validation. Moreover, integrating a theoretical framework is essential for analysing accident reports, namely ACMs. The application of a theoretical framework proves to be particularly helpful in identifying components of accident prevention.

From a technical perspective, several methods in accident reports analysis in the construction industry were recommended. Specific NLP algorithms that consider the local domain language was recommended over ML algorithms trained with a general corpus. Moreover, data pre-processing and handling methods such as clustering and Chi-square were also recommended. These were explicitly

suggested to justify and explain the consequences of the chosen methods. The same applies to the evaluation metrics, such as the ROC metric. A definitive consensus about the best use of algorithms was not evident in the existing literature. However, this thesis suggested a method for selecting the ML algorithm based on its preferences and task definition and strengths and weaknesses of the relevant ML algorithm.

The ML model that is built on accident reports from the contracting company did not explain accident outcomes. It was found that the entries of the features that described the accidents did not differentiate between high severity and low severity accidents. This result indicated that ML models that mainly focus on accident severity prediction are less successful than they seem. Instead, this thesis advised that the focus should shift from accident severity level and use ML to identify emergent risks. The latter direction should involve close collaboration with domain experts and organisational change.

8. Future work

Future work might take in consideration unsupervised-ML based analysis and data mining methods on the unstructured reported accidents. Such an ML based approach will be useful in discovering and understanding the relationships between causes. Moreover, an in-depth analysis of the ML development process could also be a useful direction for future research, where more case studies could be taken in consideration in order to explore the development of a more generalized process. Such a development would benefit from investigating the prevention strategies implemented on site because much can be learned about successful safety process. Such an empirical investigation could be of benefit in identifying the methods and principles of prevention measures that allow for safe production rather than merely focusing on accident occurrences that drive current thinking.

9. References

[https://ec.europa.eu/eurostat/statistics](https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Accidents_at_work_statistics#Analysis_by_activity)

[explained/index.php?title=Accidents at work statistics#Analysis by activity](https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Accidents_at_work_statistics#Analysis_by_activity)

<https://byggforetagen.se/statistik/arbetsmiljo/>

- Albrechtsen, E., & Hovden, J. (2014). Management of emerging accident risk in the building and construction industry. *Working on safety*.
- Alvesson, M., & Sköldberg, K. (2017). *Reflexive methodology: New vistas for qualitative research*. sage.
- Baek, S., Jung, W., & Han, S. H. (2021). A critical review of text-based research in construction: Data source, analysis method, and implications. *Automation in Construction*, 132, 103915.
- Baker, H., Hallowell, M. R., & Tixier, A. J. P. (2020). Automatically learning construction injury precursors from text. *Automation in Construction*, 118, 103145.
- Behm, M., & Schneller, A. (2013). Application of the Loughborough construction accident causation model: a framework for organizational learning. *Construction Management and Economics*, 31(6), 580-595.
- Berglund, L., Johansson, J., Johansson, M., Nygren, M., Samuelsson, B., & Stenberg, M. (2017). Risker och säkerhetsarbete i byggbranschen: En kunskapssammanställning baserad på internationell forskning.
- Berglund, L., Johansson, M., Nygren, M., Samuelson, B., Stenberg, M., & Johansson, J. (2019). Occupational accidents in Swedish construction trades. *International Journal of Occupational Safety and Ergonomics*.
- Bhavsar, H., & Ganatra, A. (2012). A comparative study of training algorithms for supervised machine learning. *International Journal of Soft Computing and Engineering (IJSCE)*, 2(4), 2231-2307.
- Bilal, M., Oyedele, L. O., Qadir, J., Munir, K., Ajayi, S. O., Akinade, O. O., ... & Pasha, M. (2016). Big Data in the construction industry: A review of present status, opportunities, and future trends. *Advanced engineering informatics*, 30(3), 500-521.
- Bilal, M., & Oyedele, L. O. (2020). Guidelines for applied machine learning in construction industry—A case of profit margins estimation. *Advanced Engineering Informatics*, 43, 101013.
- Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., & Wirth, R. (2000). CRISP-DM 1.0 Stepby-step data mining guide, SPSS Inc. *CRISPMWP-1104*.
- Cheng, C. W., Lin, C. C., & Leu, S. S. (2010). Use of association rules to explore cause-effect relationships in occupational accidents in the Taiwan construction industry. *Safety science*, 48(4), 436-444.

- Carbonell, J. G., Michalski, R. S., & Mitchell, T. M. (1983). An overview of machine learning. *Machine learning*, 3-23.
- Cassell, C. (2015). *Conducting research interviews for business and management students*. Sage.
- Creswell, J. W., & Clark, V. L. P. (2017). *Designing and conducting mixed methods research*. Sage publications.
- Dekker, S. W. (2015). The psychology of accident investigation: epistemological, preventive, moral and existential meaning-making. *Theoretical Issues in Ergonomics Science*, 16(3), 202-213.
- DVN GL AS (2020) Recommended practice: framework for assurance of data-driven algorithms and models. Report No. DNVGL-RP-0510. Online: DVN GL AS.
- El Naqa, I., Li, R., & Murphy, M. (2019). Machine Learning in Radiation Oncology.
- Fu, G., Xie, X., Jia, Q., Li, Z., Chen, P., & Ge, Y. (2020). The development history of accident causation models in the past 100 years: 24Model, a more modern accident causation model. *Process safety and environmental protection*, 134, 47-82.
- Garcia, R., Sreekanti, V., Yadwadkar, N., Crankshaw, D., Gonzalez, J. E., & Hellerstein, J. M. (2018). Context: The missing piece in the machine learning lifecycle. In *KDD CMI Workshop* (Vol. 114).
- Gholizadeh, P., Esmaeili, B., & Memarian, B. (2018, April). Evaluating the performance of machine learning algorithms on construction accidents: An application of ROC curves. In *Construction Research Congress 2018* (pp. 8-18).
- Gopal, M. (2018). *Applied machine learning*. McGraw-Hill Education.
- Grant, E., Salmon, P. M., Stevens, N. J., Goode, N., & Read, G. J. (2018). Back to the future: What do accident causation models tell us about accident prediction?. *Safety Science*, 104, 99-109.
- Han, J., Pei, J., & Kamber, M. (2011). *Data mining: concepts and techniques*. Elsevier.
- Harms-Ringdahl, L. (2013). *Guide to safety analysis for accident prevention*. Stockholm: IRS Riskhantering.
- Hawkins, D. M. (2004). The problem of overfitting. *Journal of chemical information and computer sciences*, 44(1), 1-12.
- Hegde, J., & Rokseth, B. (2020). Applications of machine learning methods for engineering risk assessment—A review. *Safety science*, 122, 104492.
- Hon, C. K., Sun, C., Xia, B., Jimmieson, N. L., Way, K. A., & Wu, P. P. Y. (2021). Applications of Bayesian approaches in construction management research: a systematic review. *Engineering, Construction and Architectural Management*.
- Hou, L., Chen, H., Zhang, G. K., & Wang, X. (2021). Deep learning-based applications for safety management in the AEC industry: A review. *Applied Sciences*, 11(2), 821.

- Hulme, A., Stanton, N. A., Walker, G. H., Waterson, P., & Salmon, P. M. (2019). What do applications of systems thinking accident analysis methods tell us about accident causation? A systematic review of applications between 1990 and 2018. *Safety science*, 117, 164-183.
- Japkowicz, N., & Shah, M. (2015). Performance evaluation in machine learning. In *Machine Learning in Radiation Oncology* (pp. 41-56). Springer, Cham.
- Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science*, 349(6245), 255-260.
- Jørgensen, K. (2002). En taxonomi for arbejdsulykker: En systematisk beskrivelse af årsagssammenhænge.
- Khallaf, R., & Khallaf, M. (2021). Classification and analysis of deep learning applications in construction: A systematic literature review. *Automation in Construction*, 129, 103760.
- Kim, T., & Chi, S. (2019). Accident case retrieval and analyses: Using natural language processing in the construction industry. *Journal of Construction Engineering and Management*, 145(3), 04019004.
- Kjellén, U., & Albrechtsen, E. (2017). *Prevention of accidents and unwanted occurrences: Theory, methods, and tools in safety management*. CRC Press.
- Kifokeris, D., & Xenidis, Y. (2018). Application of linguistic clustering to define sources of risks in technical projects. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering*, 4(1), 04017031.
- Koch, C. (2013). From crew to country? Local and national construction safety cultures in Denmark. *Construction management and economics*, 31(6), 691-703.
- Kvale, S., & Brinkmann, S. (2009). *Interviews: Learning the craft of qualitative research interviewing*. sage.
- Lee, S., & El Naqa, I. (2015). Machine learning methodology. In *Machine Learning in Radiation Oncology* (pp. 21-39). Springer, Cham.
- Lingard, H. C., Cooke, T., & Blismas, N. (2012). Designing for construction workers' occupational health and safety: a case study of socio-material complexity. *Construction Management and Economics*, 30(5), 367-382.
- Martínez-Plumed, F., Contreras-Ochando, L., Ferri, C., Orallo, J. H., Kull, M., Lachiche, N., ... & Flach, P. A. (2019). CRISP-DM twenty years later: From data mining processes to data science trajectories. *IEEE Transactions on Knowledge and Data Engineering*.
- Martyr, A. J., & Rogers, D. R. (2020). *Engine Testing: Electrical, Hybrid, IC Engine and Power Storage Testing and Test Facilities*. Butterworth-Heinemann.
- Mohri, M., Rostamizadeh, A., & Talwalkar, A. (2018). *Foundations of machine learning*. MIT press.

- Murphy, K. P. (2012). *Machine learning: a probabilistic perspective*. MIT press.
- Pan, Y., & Zhang, L. (2021). Roles of artificial intelligence in construction engineering and management: A critical review and future trends. *Automation in Construction*, 122, 103517.
- Pillay, M. (2015). Accident causation, prevention and safety management: a review of the state-of-the-art. *Procedia manufacturing*, 3, 1838-1845.
- Ray, S. (2019, February). A quick review of machine learning algorithms. In *2019 International conference on machine learning, big data, cloud and parallel computing (COMITCon)* (pp. 35-39). IEEE.
- Reason, J. The Human Contribution: Unsafe Acts, Accidents and Heroic Recoveries.(2008). *Farnham: Ashgate Publishing*.
- Sarkar, S., & Maiti, J. (2020). Machine learning in occupational accident analysis: A review using science mapping approach with citation network analysis. *Safety science*, 131, 104900.
- Schutz, A. (1972). *The phenomenology of the social world*. Northwestern University Press.
- Schwatka, N. V., Hecker, S., & Goldenhar, L. M. (2016). Defining and measuring safety climate: a review of the construction industry literature. *Annals of occupational hygiene*, 60(5), 537-550.
- Singh, A., Thakur, N., & Sharma, A. (2016, March). A review of supervised machine learning algorithms. In *2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom)* (pp. 1310-1315). Ieee.
- Symon, G., & Cassell, C. (Eds.). (2012). *Qualitative organizational research: core methods and current challenges*. Sage.
- Törner, M., & Pousette, A. (2009). Safety in construction—a comprehensive description of the characteristics of high safety standards in construction work, from the combined perspective of supervisors and experienced workers. *Journal of safety research*, 40(6), 399-409.
- Xu, Y., Zhou, Y., Sekula, P., & Ding, L. (2021). Machine learning in construction: From shallow to deep learning. *Developments in the Built Environment*, 6, 100045.
- Vallmuur, K. (2015). Machine learning approaches to analysing textual injury surveillance data: a systematic review. *Accident Analysis & Prevention*, 79, 41-49.
- Webster, J., & Watson, R. T. (2002). Analyzing the past to prepare for the future: Writing a literature review. *MIS quarterly*, xiii-xxiii.
- Zhang, F. (2019). A hybrid structured deep neural network with Word2Vec for construction accident causes classification. *International Journal of Construction Management*, 1-21.
- Zhang, J., Zi, L., Hou, Y., Deng, D., Jiang, W., & Wang, M. (2020). A C-BiLSTM approach to classify construction accident reports. *Applied Sciences*, 10(17), 5754.

- Zhong, B., Pan, X., Love, P. E., Ding, L., & Fang, W. (2020). Deep learning and network analysis: Classifying and visualizing accident narratives in construction. *Automation in Construction*, 113, 103089.
- Zhou, Z., Goh, Y. M., & Li, Q. (2015). Overview and analysis of safety management studies in the construction industry. *Safety science*, 72, 337-350.

List of abbreviations

Abbreviations	Explanation
ACMs	Accident causation models
AcciMap	accident map model
AI	Artificial intelligence
ANN	Artificial neural network
ConAC	Construction Accident Causation
CRISP-DM	Cross Industry Standard Process Development Method
DT	Decision tree
FN	False negative
FP	False positive
HSACS	Human Factor Analysis and Classification System
KNN	K-nearest neighbour
KSVM	Kernelized support vector machine
LogR	Logistic regression
LR	Linear regression
ML	Machine learning
MLP	Multi-layer perceptron
MORT	The Management Oversight and Risk Tree
NB	Naïve Bayesian
NLP	Natural language processing
OARU	Occupational Accident Research Unit
RF	Random forest
ROC	Receiver operating characteristic
ROS	Random over sampling
RUS	Random under sampling
SCM	Swiss cheese model
SMOTE	Synthetic Minority Oversampling Technique
STAMP	Systems Theoretic Accident Model and Processes
SVM	Support vector machine
SVR	Support vector regression
TN	True negative
TP	True positive

Appendix

Accident report dataset description.

Where, what, who	
This section of the report consists of information about the date, time and notified authorities, followed by the health and safety category, and the company where the incident took place. A case title and description are asked for, as well as whether the accident involves in-house or subcontracted employees. Project information contains multiple levels of detail, indicating the project and divisions where the accident took place. The title of accident, its description, and the health and safety category are written by the reporter. The health and safety category report lists accident types such as machines and equipment, falling objects, walkways, access roads, lighting, etc.	
General classification	
The general classification section involves information about the detailed work process, material agent and the substance / chemical solution involved in the report and is listed in a pre-populated drop-down list. The work process concerns the general type of work, and the work process list concerns a more detailed process category. Both have multiple levels of detail, but despite being necessarily different from each other, they seem to be repeated in the data.	
Type of work in detail	<i>Reinforcement, Excavation, Concrete work, etc.</i>
Involved substance / chemical solution	<i>Gas, Cement, Bitumen, etc.</i>
Work Process	<i>Excavation, construction work, renovation, demolition New construction – house Etc.</i>
External factor that affected the incident	<i>Building and construction parts Facilities Etc.</i>
Consequences	
The consequences section indicates the severity of the accident, as well as details such as whether it resulted in personal injury, whether the worker was assigned alternative work, any financial losses, the units where the accident happened, and the work shift during which the accident occurred.	
Actual severity	<i>1) First aid, continue to work 2) Injury that requires medical attention 3) Personal injury with absence 4) Very serious personal injury 5) Fatal accident</i>
Monetary loss	<i>Monetary loss</i>
Employment relationship	<i>Part time employee, Own employee.</i>
Work environment	<i>Production site, factory, workshop Underground – mine Etc.</i>
Position	<i>Machine operator Supervisor Etc.</i>
Description of damaged body part, Common	<i>The leg / calf Torso Etc.</i>
Description of injury type	<i>Allergic reaction Electricity injury Etc.</i>
The last deviating event that preceded the injury	<i>Electrical problem due to defects in the installation - causes an indirect contact Fire, ignition Etc.</i>
Number of registrations Personal injuries	<i>The number of registered personal injuries.</i>

Experience in position (Months)	The number of months of experience.
Actual number of days with Alternative work	The actual number of days with alternative work.
Actual number of days of absence (Absence damage), calendar days	The actual number of days of absence
Company name	Main contractor Subcontractor
Injured body part	Finger (fingers) Teeth Etc.
Injury class	1) First aid, continue to work 2) Injury that requires medical attention 3) Personal injury with absence 4) Very serious personal injury 5) Fatal accident
Category of injury	Hit by moving objects, collision with – no Squeezing, crushing, getting stuck in, etc. Not specified. Etc.
Injury type level 1	Wounds and superficial injuries Dislocation, sprains, and strain Etc.
Injury type level 2	Superficial injury Dislocation and subluxations Etc.
Specific physical activity level 1	Working with hand-held tools - Not spec. Driving / staying on board transport equipment / handling equipment - Not spec. Etc.
Specific physical activity level 2	Working with hand-held tools – motorized Driving a means of transport or handling equipment - mobile and not motorized Etc.
Shift or accident to / from work	Day shift Evening shift Etc.
Loss potential	
Possible further consequence	Material damage Personal injury
Potential Severity - Most Severe	1) First aid, continue to work 2) Injury that requires medical attention 3) Personal injury with absence 4) Very serious personal injury 5) Fatal accident
Risk area	Less serious area (green traffic light) Serious area (yellow traffic light) Critical area (red traffic light)
Causes	
Comments	Potential comments offered for a case.
Circumstances of the accident	During travel between the home and the workplace At work: During work At the workplace but not in work tasks: Other premises than those arranged by the employer Etc.
Cause level 1	Inadequate risk assessment and / or risk assessment not carried out

	Removal of safety devices Etc.
Cause level 2	Unconcentrated / distracted Insufficient safety assessment Etc.
Causes - Cause description	Free text
Cause category	Prerequisites (Direct cause) Person-dependent factors (underlying cause) Etc.
Prevention	
Expiration Status of Action	Ended after due date Completed before due date No deadline
Comments	Free text
Prevention status	Prevented Rejected
Prevention type	Temporary Prevention
Action - Created Date, Time Period = Day	Date format
Action - Fixed, Time period = Day	Date format
Action description	Free text
Case handling	
Case Management Time	Number of days
Registration delay (Established date - Case date)	Number of days
Case management and status	All cases in the data set are closed