MACHINE LEARNING-DRIVEN ACCIDENT PREDICTION TO ENHANCE INHERENT SAFETY DESIGN THROUGHOUT PROCESS DESIGN LIFECYCLE

HASLINDA BINTI ABDUL SAHAK

MCS241004_MCST 1043

CHAPTER 1

UNIVERSITI TEKNOLOGI MALAYSIA

NOVEMBER 2024

CHAPTER 1

INTRODUCTION

1.1 Introduction

The chemical process industry (CPI) is one of the most significant engines among the various such engines that drive improvements in the economy of most countries. However, the activities and processes within CPI are sometimes associated with very high risks. The feature is true at least in the very long run. Despite several decades of safety practice improvement and technological advancement, accident rates remain a constant feature of this industry indicating critical flaws in safety management and prevention strategies. However, in practice, barriers result in problems like poor dissemination of accident knowledge, inadequacy in the quality of accident reports, and lack of historical evidence in predictive analysis, which is part of what hinders meaningful advancements in the reduction of accident frequencies (Lindberg et al., 2010).

The excessive reliance on procedural and organizational controls rather than on fundamental protections often causes safety management to fail when root causes lie hidden in plant design (Kletz, 2003). Consequently, an outer layer of protection rather than innate principles has caused recurring incidents maintained through disproportionate risk management processes and a poor connection between design and risk analyses (Kidam & Hurme, 2012). As technology evolves, more chemical plants face the inevitable aging of infrastructure and still higher operation thresholds. The complexity added to these chemical plants has produced entirely new accident scenarios and challenges for designers (Kidam et al., 2014).

In context, transforming the process design lifecycle by integrating machine learning technologies emerges as new and radical opportunities for improving inherent safety design (ISD). ML applications take from historical accident records to forecast risks, discover weaknesses in the design, and provide actionable suggestions for future preventive measures against hazards. In contrast to these traditional reactive approaches, accident prediction is externalized by ML, allowing decisions to be made based on evidence, which permits designers to take action early in a design phase when their strategies are still cost-effective and most efficient.

Machine learning would revolutionize accident prevention in the chemical, petroleum, and heavy industries by embedding inherent safety (IS) principles within the design lifecycle. Historical accident analyses would be systematically examined to create a framework for disseminating knowledge and a culture of learning from what has happened. The intention is to motivate design thinking that prioritizes safety and to show how critical life cycle assessments would be enabled through ML tools that pave the way for a more secure and sustainable future in the chemical industry.

1.2 Problem Background

Inherent safety (IS) issues have become important factors when designing chemical plants. Inherent or intrinsic means something has a permanent, essential, or characteristic attribute. To achieve the IS plant design, the critical step is to choose the "best" chemical process route due to its inherently less harmful to safety and health. According to Edwards and Lawrence (1993), a simple chemical process route refers to the raw material and reaction sequence transforming the raw material into the desired product. In response to this situation, Kletz proposed an inherently safe design (ISD) in the 1970s. The method proposes a method to eliminate the hazards in the chemical process. The method will be implemented throughout the process design, which usually begins during the research and development (R&D) phase and then enters the process development phase.

There are various phases of the process design lifecycle, such as conception and operation, during this period wherein safety should be a concern. Safety design today has mostly been reactive, with most accidents occurring before preventive measures against the same could be taken. More importantly, regardless of all the existing standards and regulations to govern accident prevention, many accidents still occur because certain risks are considered or underrated. The current safety measures still use a well-proven but static predictive approach. Machine learning asks about the possibility of making ISD better by forecasting accidents through data and insights into hidden risks for safety measures optimization throughout the design lifecycle. As yet, applying ML in safety design is still in its infancy, and several challenges continue to exist in integrating predictive models with real-world process design systems.

The chemical process industry (CPI) has struggled consistently to slash its accident rates despite considerable research and other findings regarding the causes of accidents. The existing safety framework makes minimal effort to identify, avoid, and control hazards early in the process design Kidam et al. (2016). Other than that, it has been shown that 95% of accidents in the CPI might be avoided with the available and current knowledge of accident research (Pasman, 2010). Unfortunately, there have been no proper channels to disseminate or even make use of accident data so the same accidents keep occurring with almost no learning from the previous incident (Jacobsson et al., 2010). Poor accident reports, improper analysis of data, and weak feedback mechanisms in the learning cycle such as Lindberg and Hasson (2006) are indeed considerable hindrances to effective accident prevention.

Unbalanced risk management in designing plants has also been identified as one of the singular causes of accidents in the CPI (Kidam & Hurme, 2012). Traditional practices typically depend on procedural safety measures and not on integrating inherently safer design principles, which target the elimination of hazards at source

(Amyotte et al., 2011). The safety measure is aimed at short-term cost savings but may attract a higher long-term cost in management and risk (CCPS, 1998). Furthermore, caused of rapid industrialization and technological advances, the complexity of chemical plant design greatly challenges the issue of maintaining safety and health standards (He et al., 2011).

The slow pace of adoption of ISD principles since the 1970s is evident by the serious lacunae in the application of ISD at the early design phases (Gupta & Edwards, 2002). It has resulted in the continuing high accident rates in the CPI, as seen in different areas of the world, including the United States, Asia, and Europe (Prem et al., 2012). Furthermore, the neglect of IS in design practices has hampered the industry from minimizing risks efficiently (Kidam et al., 2016).

In this case, the research provides a machine-learning-driven accident prediction model for the chemical process industry's design lifecycle to implement the principles of IS over time better. Thus, with early identification of design-related hazards, the systematic application of IS methodologies in the design process would improve safety performance with cost savings and reduced risks in the CPI.

1.3 Problem Statement

In the years after developing safety practices and safety technologies, the chemical process industry (CPI) accident rates remained high, mostly because of the failure to disseminate knowledge about accidents and poor learning from accidents. Although many studies and frameworks address IS in the CPI, the existing methods are not integrated with expertise and usability in early process design. These approaches, particularly HAZOP, are limited to applications since they require exhaustive process data which is hardly available during the preliminary design phase.

This gap ensures that hazards associated with designs are not identified and mitigated proactively, hence resulting in recurring accidents and meaningful lessons not learned.

In addition, the traditional methods of IS are rather complicated, protracted, and knowledge-rich, and demand thorough training and experience to use well, thus rendering them inaccessible for general application. The lack of a common, data-driven framework to integrate accident insights into process design increases this problem. The existing accident reporting and investigation systems are utilized minimally, and the fulfillment of their potential toward informing inevitably safe designs remains quite limited. This is a barrier for designers and engineers in systematically addressing hazards at the early design stages, where changes tend to be more cost-effective and have a deeper impact.

Designing safety into any process right across its life cycle from first conception through research, development, and ultimately deployment can only be accomplished through a machine-learning-guided approach using accident databases to improve predictive capability. It would be made possible through archiving taught accidents but, most importantly, phenomenon-based error modes, the end goals of which should be effective use and sharing of accident knowledge but also an end-user "friendly" tool facility for design decision-making and avoidance of accident recurrence FBI-based.

1.4 Research Questions

This research aims to visualize accident prediction using machine learning to enhance inherent safety design across the process design life cycle in the chemical industry. The research questions are:

1. What are the main causes of accidents in the chemical process industry across the different life-cycle phases of process design?

2. What prioritization should be given regarding process safety recommendations to prevent accidents at the different design phases?
3. In what way can accident data increase loss prevention at early process design decisions?

1.5 Research Aim

This research aims to develop a machine learning and prediction framework for data-driven analysis of accident-based information to enhance inherent safety and health throughout the process design life cycle in the chemical process industry (CPI). This framework will utilize historical accident data to identify, classify, and rank accident contributors and safety recommendations, providing actionable insights that support proactive, informed safety and health decision-making from early design phases through the process design life cycle. Ultimately, this research aims to reduce accident recurrence by embedding safety and health knowledge into CPI design practices.

1.6 Research Objective

The primary objective of this research is to enhance accident learning in the chemical process industry (CPI) by safety and health insights into process development and design. Specifically, the research aims:

1. To identify, classify, and rank key accident contributors and process safety recommendations throughout the process design lifecycle.

2. To analyze accident database information for inherent safety design, applying insights to:
    a. Research and Development (R&D)
    b. Preliminary Engineering Phase
    c. Basic Engineering Phase
    d. Detailed Engineering Phase
    e. Procurement, Fabrication, Commissioning and Start-up
    f. Operation/ Plant modification
3. To interconnect accident contributors against process safety hierarchy and determine the current preference of loss prevention in the industry.

1.7 Scope of Research

The scope of this research is to analyze a large volume of historical accident data and categorize insights relevant to safety and health integration across process design stages in the chemical process industry (CPI). The scope is defined as follows:

1. Data Collection: The research will retrieve at least 500,000 accident investigation reports from reputable sources, including the US Chemical Safety and Hazard Investigation Board (CSB), the US National Transportation Safety Board (NTSB), the US Environmental Protection Agency (EPA), Japan's Science and Technology Failure Knowledge Database (FKD), and the EU Major Accident Reporting System (e-MARS). These reports cover accidents from 1990 to 2024.

2. Accident Analysis: The study will focus on process safety-related accidents to identify critical contributors in the CPI. These contributors will be classified by design error categories, such as process condition, reactivity/incompatibility, unsuitable equipment/part, protection, construction

material, layout, utility setup, sizing, and automation/instrument issues, all linked to equipment failure types.

3. Classification of Accident Contributors and Root Causes: Identified contributors will be categorized into design, technical, human, organizational, and external errors. Root causes will be further classified by process design life cycle: research and development (R&D), preliminary engineering, basic engineering, detailed engineering, procurement, fabrication, commissioning start-up, and operation/ plant modification.

4. Recommendations Analysis: Recommendations within the reports will be classified based on preventive, protective, mitigative, and corrective actions. These safety actions are further categorized by inherent safety, passive-engineered, and active-engineered measures. Inherent safety and health recommendations will focus on ISD (Inherently Safer Design) principles: minimization, substitution, moderation, and simplification.

5. Statistical Ranking: Statistical analysis will classify and rank the accident contributors and safety recommendations based on frequency of occurrence, allowing the identification of critical design-related issues in CPI accidents.

6. Interconnection analysis of the contributors against process safety actions is conducted on single-cause accidents and overall accidents. For interconnection analysis of the overall accidents, a contributor-action mapping using brainstorming among researchers is used to identify the corresponding action(s) of each contributor as the reports generally list their recommendations.

1.8 Significance of Research

The present research has significant value since it advances the safety practices in the chemical industry. The machine learning-involved accident prediction integrates informing about potential hazards beforehand so that proactive measures can be taken to prevent accidents during the design life cycle. The strategy robustly supports ISD by nesting safety principles right into the design phase of the process, making processes inherently safer instead of depending on external protective measures.

Data-driven insights to engineers and safety professionals improve decision-making because they optimize process designs without excessive risk. Prevention of accidents not only increases safety but also reduces economic and environmental consequences, such as costly downtimes, material losses, and environmental damage; hence, safety and profit are ensured in chemical operations.

In addition, this research helps to develop standards in the industry by facilitating technology-enabled approaches to hazard prediction and mitigation. It aligns with Industry 4.0 objectives by integrating machine learning with safety management systems, supporting the chemical industry's transition to digitalization and smart manufacturing. Overall, this research offers a transformative approach to enhancing operational excellence and protecting lives, assets, and the environment.