Sci

MDPI

Systematic Review

A Systematic Review of Machine Learning Analytic Methods for Aviation Accident Research

Aziida Nanyonga [1], Ugur Turhan [2] and Graham Wild [2,*]

[1] School of Engineering and Technology, University of New South Wales, Canberra, ACT 2600, Australia; a.nanyonga@unsw.edu.au
[2] School of Science, University of New South Wales, Canberra, ACT 2612, Australia; u.turhan@unsw.edu.au
* Correspondence: g.wild@unsw.edu.au

Abstract

The aviation industry prioritizes safety and has embraced innovative approaches for both reactive and proactive safety measures. Machine learning (ML) has emerged as a useful tool for aviation safety. This systematic literature review explores ML applications for safety within the aviation industry over the past 25 years. Through a comprehensive search on Scopus and backward reference searches via Google Scholar, 87 of the most relevant papers were identified. The investigation focused on the application context, ML techniques employed, data sources, and the implications of contextual nuances for safety analysis outcomes. ML techniques have been effective for post-accident analysis, predictive, and real-time incident detection across diverse aviation scenarios. Supervised, unsupervised, and semi-supervised learning methods, including neural networks, decision trees, support vector machines, and deep learning models, have all been applied for analyzing accidents, identifying patterns, and forecasting potential incidents. Notably, data sources such as the Aviation Safety Reporting System (ASRS) and the National Transportation Safety Board (NTSB) datasets were the most used. Transparency, fairness, and bias mitigation emerge as critical factors that shape the credibility and acceptance of ML-based safety research in aviation. The review revealed seven recommended future research directions: (1) interpretable AI; (2) real-time prediction; (3) hybrid models; (4) handling of unbalanced datasets; (5) privacy and data security; (6) human–machine interface for safety professionals; (7) regulatory implications. These directions provide a blueprint for further ML-based aviation safety research. This review underscores the role of ML applications in shaping aviation safety practices, thereby enhancing safety for all stakeholders. It serves as a constructive and cautionary guide for researchers, practitioners, and decision-makers, emphasizing the value of ML when used appropriately to transform aviation safety to be more data-driven and proactive.

Keywords: machine learning; post-accident analysis; safety enhancement; aviation; review

check for updates

Academic Editor: Carson K. Leung

Received: 14 May 2025
Revised: 17 August 2025
Accepted: 2 September 2025
Published: 4 September 2025

Citation: Nanyonga, A.; Turhan, U.; Wild, G. A Systematic Review of Machine Learning Analytic Methods for Aviation Accident Research. *Sci* **2025**, *7*, 124. https://doi.org/10.3390/sci7030124

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Air transport stands at the centre of modern global connectivity, revolutionizing the way people traverse the world and facilitating rapid travel [1]. However, within this critical industry, the paramount concern is ensuring safety. The global aviation industry's complex procedures and high-stakes operations leave little margin for operational errors. Even the slightest mistake or system misconfiguration can lead to catastrophic accidents, with significant loss of life, potential property damage, substantial financial impacts, and

eroded customer trust [2]. The aviation industry has consistently demonstrated a comparatively remarkable commitment to safety, continually evolving and embracing technological advancements and comprehensive protocols. This dedication is particularly evident in the aftermath of accidents, where in-depth post-accident analyses play a pivotal role in understanding the root causes, mitigating future risks, and fortifying safety measures [3]. Government bodies like the International Civil Aviation Organization (ICAO), the European Union Aviation Safety Agency (EASA), and the U.S. Federal Aviation Administration (FAA) establish and promote aviation safety standards. Their guidelines, protocols, and procedures safeguard passengers, crew, and aircraft operations, forming the foundation of aviation safety.

The journey towards enhanced safety within the aviation sector is punctuated by significant milestones, each shaped by lessons derived from post-accident analyses of major (and sometimes minor) aviation accidents. These analyses entail multifaceted collaborations between regulatory bodies, aircraft manufacturers, airlines, aviation experts, and investigators [4]. While immediate safety occurrence causes are quickly examined, the analyses routinely extend beyond, aiming to uncover contributing factors and systemic issues. This approach has fostered advancements spanning aircraft design, maintenance engineering, air traffic management, crew training, and human factors analysis [5]. Addressing these challenges necessitates collaborative efforts. For example, NASA's Aviation Safety Reporting System (ASRS) collects safety occurrence reports voluntarily submitted by aviation professionals, offering insights into potential safety concerns and proactive accident prevention (ASRS program briefing, https://asrs.arc.nasa.gov/overview/summary.html accessed on 8th April 2023). Additionally, the FAA and NTSB play integral roles in investigating and enhancing aviation safety standards. While the FAA enforces regulations to maintain airspace safety, the NTSB investigates accidents, supplying invaluable data for safety improvement [6]. Collectively, these efforts have substantially reduced the rate of accidents [7]. The industry's commitment to continuous improvement reflects its determination to uphold a high safety record despite inherent risks.

The aviation sector's recent surge in growth fueled by escalating demand for air travel and infrastructure constraints has ushered in both opportunities and challenges [8]. The International Air Transport Association (IATA) forecasts global air travel demand doubling within the next two decades [9]. While this growth promises economic potential, it also increases pressure on existing infrastructure, underscoring safety as a paramount concern. IATA, a significant industry body, plays a pivotal role in advocating industry initiatives and standards. However, concerns have arisen regarding their representation of smaller airlines and potential conflicts of interest among their members. Increasing air traffic congestion and the strain on Air Traffic Management (ATM) systems further escalates safety risks where infrastructure developments often lag demand [10]. Considering aviation's recent growth surge, various approaches are being researched to bolster safety. These range from investigative methods that analyze data from accidents, incidents, and near misses to predictive strategies that proactively identify potential risk factors before accidents occur. Regulatory bodies and airlines maintain extensive repositories of data, and researchers diligently analyze these datasets to facilitate proactive measures against potential risks [11].

According to Annex 13 to the Convention on International Civil Aviation: Aircraft Accident and Incident Investigation (10th ed.) by the International Civil Aviation Organization (ICAO), an "accident" is defined as an occurrence associated with the operation of an aircraft which takes place between the time any person boards the aircraft with the intention of flight until such time as all such persons have disembarked, where a person is fatally or seriously injured, the aircraft sustains damage or structural failure, or the aircraft is missing or inaccessible. On the other hand, an "incident" is defined as an occurrence other than an

accident, associated with the operation of an aircraft, which affects or could affect the safety of operations. An "occurrence" is an accident, incident or any other operational issue that could impact aviation safety. These definitions provided by ICAO serve as fundamental frameworks for the investigation and analysis of aviation-related events, guiding safety protocols and procedures within the aviation sector [12].

Machine Learning (ML) has emerged as a transformative force in data analysis. A subset of Artificial Intelligence (AI), ML aims to equip systems to learn from data, make informed decisions, and detect patterns without explicit programming. With the patterns and process of vast datasets and employing advanced algorithms, ML offers substantial new potential to predict anomalies, extract patterns, and identify factors contributing to events [13]. As aviation continues to evolve, integrating ML into aviation safety occurrence investigation and analysis should help sustain or improve safety practices and operational procedures.

This systematic literature review explores ML's role in aviation safety, with a focus on post-accident analysis. Employing a systematic approach, the paper categorizes and synthesizes primary studies to address well-defined research questions. By analyzing 87 carefully selected studies from a pool of 3832 papers, this review aims to explore the state-of-the-art in the domain of ML for aviation accident analysis. Through the synthesis of these studies, this paper serves as a foundational resource for researchers, aviation stakeholders, and regulatory bodies seeking to harness the potential of machine learning to bolster safety practices. Formally, the primary objective of this systematic literature review (SLR) was to gather, evaluate, and synthesize existing research studies that use ML applications in post-accident analysis within the aviation sector. The associated research questions to be answered were:

1. How do variations in the demographics factors such as aviation application, electronic database, publication type, study year, trends, data sources, and authors influence the comprehensiveness, reliability, and applicability of findings in aviation accident research?
2. What types of machine learning approaches have been applied in post-accident analysis within the aviation industry?
    a. What specific machine learning types have been employed in this context?
    b. Which machine learning tasks have been targeted to enhance post-accident analysis?
    c. What are the predominant machine learning algorithms used in the aviation industry for post-accident analysis?
3. How have these machine learning techniques contributed to enhancing safety measures and providing insights into aviation accidents?

In the subsequent sections, the background and related work (Section 2) are covered, the methodology used in this systematic review (Section 3) is described, the results are presented and discussed (Sections 4 and 5) and conclusions based on the findings are presented (Section 6). This review aims to be a benchmark overview of machine learning applications in post-accident analysis within the aviation sector and identify associated opportunities for future research.

2. Background

In this section, the fundamental concepts that underpin the exploration of machine-learning techniques in the aviation sector are examined. It is essential to understand the difference between machine learning (ML) and artificial intelligence (AI) and to have an awareness of the various types of ML and the tasks they encompass.

2.1. Artificial Intelligence

Artificial Intelligence (AI) is the field dedicated to the study of intelligent agents, encompassing machines that mimic human-like behaviors in learning and problem-solving processes [14]. Utilizing computer-processing techniques, AI enables machines to learn, perceive, and process natural language, while also making decisions in ways reminiscent of human cognitive processes [15]. This transformative technology excels at processing substantial volumes of complex data and performing real-time decision-making, thus finding applications in a multitude of management domains.

Within the realm of AI, one prominent subset is machine learning (ML) because it plays a pivotal role in various industries, including aviation. The different types of ML and the diverse tasks they can accomplish will be explored in the following sections. With a clear understanding of these concepts, exploration into the integration of ML techniques to improve aviation safety and operational efficiency becomes feasible.

2.2. Machine Learning

Machine learning (ML) empowers computer programs to execute intricate tasks like prediction, diagnosis, planning, and recognition by assimilating insights from historical data. The synergy between data and algorithms is paramount for the proficiency of ML models. Enhanced accuracy often results from high-quality data and large datasets, so long as they are appropriately balanced to the key factors and the learning cases are appropriately sequenced [16,17]. Equally crucial is the selection of suitable algorithms tailored to the specific problems at hand, accommodating diverse types of datasets [14,18].

2.3. Machine Learning Types

ML encompasses several types, each with its unique characteristics and applications. These types enable the creation of models that can adapt and learn from data, making them versatile tools for addressing various challenges. These types are depicted in Figure 1. The following ML types are explored:

1. Supervised Learning: In supervised learning, models are trained on labeled datasets, where input data is associated with corresponding desired outputs. The model learns the relationships between inputs and outputs, enabling it to make predictions on new, unseen data. Supervised learning is widely employed in tasks such as classification, where the goal is to assign input data to predefined categories, and regression, where continuous values are predicted based on inputs [19,20].
2. Unsupervised Learning: Unsupervised learning involves training models on unlabeled data, with the aim of discovering patterns and structures within the data. Clustering is a common task in unsupervised learning, where similar data points are grouped together. Dimensionality reduction is another application, simplifying the data while preserving its key characteristics [21,22].
3. Semi-Supervised Learning: This type of learning combines labeled and unlabeled data to enhance model performance. Often, labeled data is scarce, but unlabeled data is more abundant. By leveraging both types of data, models can generalize better and make more accurate predictions [23].
4. Reinforcement Learning: Reinforcement learning is rooted in the concept of learning through interactions, often simulation of varying fidelity and often in environments with significant variability and uncertainty [24]. An algorithm interacts with a system and optimizes variables to maximize cumulative rewards based on an objective function and through trial and error, the system learns optimal strategies to achieve its goals [25].

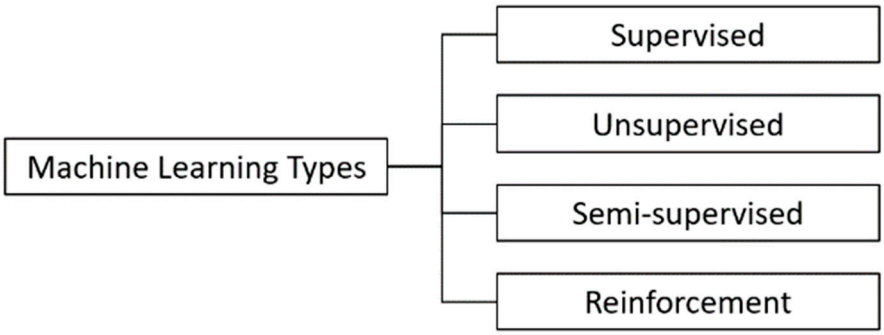


Figure 1. The four most common board machine learning types.

2.4. Machine Learning Tasks

ML encompasses a diverse array of tasks, each serving a specific purpose and leveraging data-driven techniques to extract valuable insights. These tasks enable the development of models that can autonomously make decisions, predictions, and classifications as shown in Figure 2. The following six common ML tasks are explicated:

1. Classification: Classification involves assigning input data to predefined categories or classes. The model learns patterns from labeled training data, enabling it to classify new, unseen data accurately. Applications range from email spam detection to medical diagnosis [19].
2. Regression: Regression is concerned with predicting continuous values based on input features. The model learns the relationships between variables from training data and can then make predictions on new data points. Examples include predicting housing prices or stock market trends [26].
3. Clustering: Clustering involves grouping similar data points together based on inherent patterns in the data. The model identifies clusters or segments within the data without requiring predefined categories. This task finds applications in customer segmentation and anomaly detection [27].
4. Anomaly Detection: Anomaly detection focuses on identifying rare or unusual instances within a dataset. The model learns the normal patterns and detects deviations from these patterns, making it valuable for fraud detection and network security [6,28].
5. Data Reduction: Data reduction involves techniques to reduce the complexity of large datasets while retaining crucial information. These techniques help minimize computational overhead, improve model efficiency, and avoid overfitting. Methods like Principal Component Analysis (PCA) and feature selection are commonly used for data reduction [21].
6. Natural Language Processing (NLP): NLP is a specialized task involving the interaction between computers and human language. It encompasses various sub-tasks such as sentiment analysis, text generation, and language translation. NLP enables machines to understand, interpret, and generate human language, with applications spanning from chatbots to language translation services [29,30]. Because of its role in communicating, NLP is often a necessary sequential addition to enable the other listed ML tasks. Due to the diverse ontology at play in each domain, it is often necessary to systematically tailor NLP to new domains like extending a dictionary or lexicon [31,32].

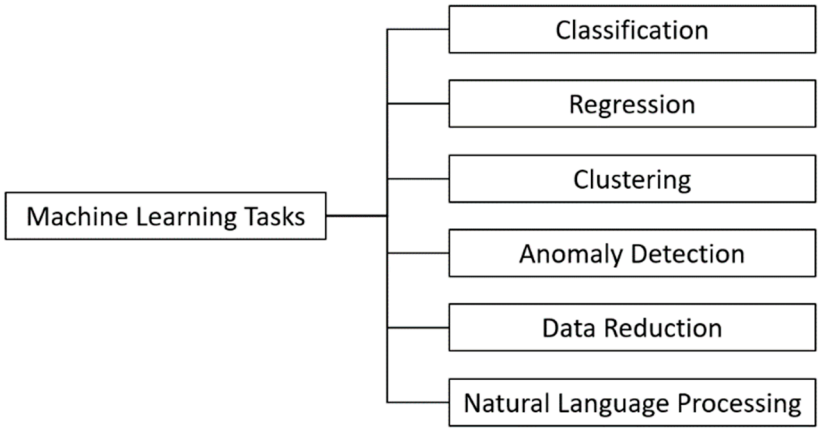


Figure 2. The six most common machine learning tasks.

Machine Learning Techniques in Aviation Safety

ML techniques enable the analysis of vast aviation datasets to uncover patterns, make predictions, and extract valuable information. Below are some explanations of key techniques along with notable studies that have employed specific ML techniques in aviation safety research:

1. Support Vector Machine (SVM). SVM is a powerful ML technique commonly used for classification and regression tasks. It works by finding the optimal hyperplane that maximally separates data points belonging to different classes. This hyperplane is selected in such a way that the margin between the classes is maximized, allowing for the effective classification of new, unseen clusters. In the context of aviation safety analysis, SVM has proven to be a valuable tool for categorizing aviation accidents based on their characteristics, such as severity, causal factors, and contributing variables [33]. In that research, SVM was used alongside other techniques such as Random Forest and Naive Bayes to classify aviation accidents over 20 years using the ASRS dataset.
2. Decision Tree: This is a powerful tool used for both classification and regression tasks in ML. It takes the form of a tree-like structure where each internal node represents a decision based on a feature, each branch represents an outcome of that decision, and each leaf node represents a classification label or a predicted value. Decision Trees have found applications in aviation safety and analysis due to their ability to handle complex decision-making processes and their transparency. For instance, refs. [34–36] utilized Decision Trees, along with other ML algorithms, to classify and reduce data from airline databases spanning 42 years. This research enabled the classification of aircraft accident data into different categories, contributing to insights into warning levels in accidents. Also, another study extended the use of Decision Tree in the aviation safety domain by applying them in conjunction with NLP techniques. Their study involved extracting word-level meaning from safety report narratives and using these meanings to analyze a large dataset of 186,000 reports over 39 years. The combination of NLP and Decision Trees allowed for the extraction of valuable insights from textual data, contributing to improved safety analysis and risk assessment in the aviation industry [37].
3. K-Nearest Neighbors (KNN): KNN algorithm is a fundamental ML technique used for classification and regression tasks. It is a non-parametric, instance-based learning method that makes predictions based on the similarity between input data points. KNN is intuitive and easy to understand, making it a popular choice for beginners in ML, noting it leverages the ML to distinguish nearest neighbor in often very multi-

factor dimensional space (10 or more). The algorithm operates on the principle that similar data points tend to share common characteristics and attributes. The KNN algorithm's performance depends on the choice of the parameter "k" and the distance metric used to measure similarity. A smaller "k" value can lead to a noisy prediction, while a larger "k" value can result in a smoother but potentially biased prediction. Additionally, selecting an appropriate distance metric is crucial, as it affects how the algorithm measures similarity between data points. Koteeswaran et al., [38] employed the KNN technique as part of their study on predicting the topmost causes of aviation accidents using data mining algorithms. In this research, along with other ML algorithms like Naive Bayes (NB) and Decision Trees (DT), KNN was used to classify aviation accidents spanning 95 years using data from the FAA. By applying KNN to the FAA dataset, the study identified commonalities and patterns in accidents that could contribute to better accident prevention strategies and safety measures.

4. Neural Networks: This is a class of ML algorithms inspired by the structure and functioning of the human brain's NNs (Neural Networks). They are designed to process and recognize patterns in data, making them highly suitable for tasks involving complex relationships and non-linear interactions. The determinations from neural networks can be difficult to trace, even when they satisfy objective functions and outperform baseline performance of humans or other more explainable AI [39]. NNs consist of interconnected nodes, or artificial neurons, organized in layers, each responsible for specific computations. In the context of aviation safety research, ref. [40] harnessed the power of NNs to address a critical task in aviation safety management systems. Their study involved classifying risk factors within the aviation domain using data obtained from the ASRS over 24 years. Also, the study conducted by [41] involved the application of NN to predict Human Factors Analysis and Classification System (HFACS) unsafe acts based on the pre-conditions of those unsafe acts.
5. Naive Bayes (NB): This technique is a probabilistic ML algorithm rooted in Bayes' theorem. It is particularly well-suited for classification tasks, where the goal is to assign predefined categories or labels to input data based on observed features. Naive Bayes assumes feature independence, which simplifies calculations and makes it a relatively efficient algorithm for text classification and other categorical data. In their study, Koteeswaran et al., [38] used NB to predict the topmost cause of accidents in aviation as a data mining technique. This technique was compared to other techniques such as SVM and KNN explained above and in this instance was the most effective when for their classification tasks.
6. Random Forest (RF): RF "combines several randomized decision trees and aggregates their predictions by averaging" [42]. It is particularly effective for classification and regression tasks. RF is based on the concept of creating multiple decision trees during the training phase and combining their predictions to improve accuracy. Each decision tree is constructed using a random subset of the training data and a random subset of features, which helps to introduce diversity and robustness to the model [43]. Zhang & Mahadevan, [9] utilized the Random Forest technique along with Deep Neural Networks (DNN) and SVM for NLP and classification of aviation incident reports from the ASRS dataset. The study focused on analyzing and classifying aviation incident reports spanning an 11-year period. By incorporating RF in their analysis, the researchers leveraged its ensemble capabilities to improve the accuracy and reliability of their classification model, ultimately enhancing the understanding of safety incidents in aviation.
7. Latent Dirichlet Allocation (LDA): LDA is a widely used ML technique for topic modeling and document clustering. It's particularly applicable to textual data analysis,

such as the analysis of aviation accident reports [44–49]. LDA assumes that each document is a mixture of a small number of topics or themes, and each topic is characterized by a distribution of words. The goal of LDA is to uncover these hidden topics and their associated word distributions from a collection of documents. In [50] study, "Text Mining Classification and Prediction of Aviation Accidents Based on TF-IDF-SVR Method," various machine learning techniques, including LDA, SVM, NB, RF, and LR, are employed to analyze 20,000 aviation accidents spanning 59 years from NTSB data. LDA, a natural language processing technique, aids in uncovering underlying topics within accident reports, contributing to NLP-driven classification efforts for improved accident prediction and prevention.

8. Long Short-Term Memory (LSTM): LSTM is a type of Recurrent Neural Network (RNN) architecture that is well-suited for processing data sequences, such as time series or sequences of text. It is particularly effective in capturing long-range dependencies and patterns within sequential data due to its ability to maintain and update information over extended sequences [51–53]. This makes LSTM suitable for tasks that involve sequential data, where past information can significantly impact future predictions and potentially where variability or uncertainty tends to obfuscate or confuse classical algorithms [54]. The research conducted by Zeng et al., focuses on the application of Long Short-Term Memory (LSTM) techniques for aviation safety prediction [55]. Specifically, the study employs LSTM with variable selection methods, including LASSO (Least Absolute Shrinkage and Selection Operator), to predict aviation safety-related outcomes. The data set used in this research is sourced from the ASRS, which collects and analyzes incident and accident reports from aviation professionals. In the context of aviation safety prediction by Zeng et al., LSTM is used to model and analyze the temporal patterns of safety-related incidents [55]. The LSTM network is designed to learn from historical data and capture complex relationships between variables, allowing it to make predictions about potential safety outcomes based on past incident reports.
9. Principal Component Analysis (PCA): PCA is a widely used dimensionality reduction technique in ML and data analysis. It aims to transform high-dimensional data into a lower-dimensional space while preserving as much of the original data's variance as possible. This reduction in dimensionality helps in simplifying the dataset and removing redundant or less informative features, making it easier to work with and potentially improving the performance of machine learning algorithms [56]. In the study conducted by İnan & Gökmen İnan, [57] PCA was applied in conjunction with NNs and DTs (Decision Trees) to classify survivor and non-survivor passengers in fatal aviation accidents based on data reduction. The researchers aimed to identify significant patterns and features that could distinguish between passengers who survived accidents and those who did not.

3. Materials and Methods

This research employed a rigorous Systematic Literature Review (SLR) approach, adhering to the comprehensive guidelines established by Kitchenham et al., [58], which are still widely used today [59]. This section provides details of the methodology, encompassing the formulation of research questions, the development of the search strategy, the establishment of study selection criteria, the application of a quality assessment methodology, the data collection procedures, and effective data synthesis techniques. Throughout the research process, the SLR guidelines and protocols outlined by Kitchenham et al., [58] were adhered to, ensuring a systematic and methodical process. Additionally, in alignment with PRISMA 2020 guidelines, the study documented and reported the review process

using a PRISMA flow diagram. Although this review was not registered in databases such as PROSPERO, due to its focus on aviation safety and machine learning, which falls outside the registry's scope, all relevant PRISMA checklist items were addressed where applicable, and justifications are provided for any deviations.

The SLR was conducted manually, consistent with the position of Van Dinter et al. [59], who noted that while AI/ML tools are available for automating reviews, manual analysis remains essential for transparency and control. A well-defined search strategy was designed to retrieve pertinent articles, and strict inclusion/exclusion criteria were applied to screen and assess the studies.

Upon defining the research questions, a robust search strategy was designed to retrieve pertinent articles from the Scopus database. Subsequently, specific study selection criteria were applied to the retrieved articles, refining the initial pool into a focused subset for more rigorous evaluation during the quality assessment phase. Details of the search, selection, and analysis (as illustrated in Figure 3) are presented in the following sections, and the complete selection process is illustrated in Figure 4 using the PRISMA flow diagram.

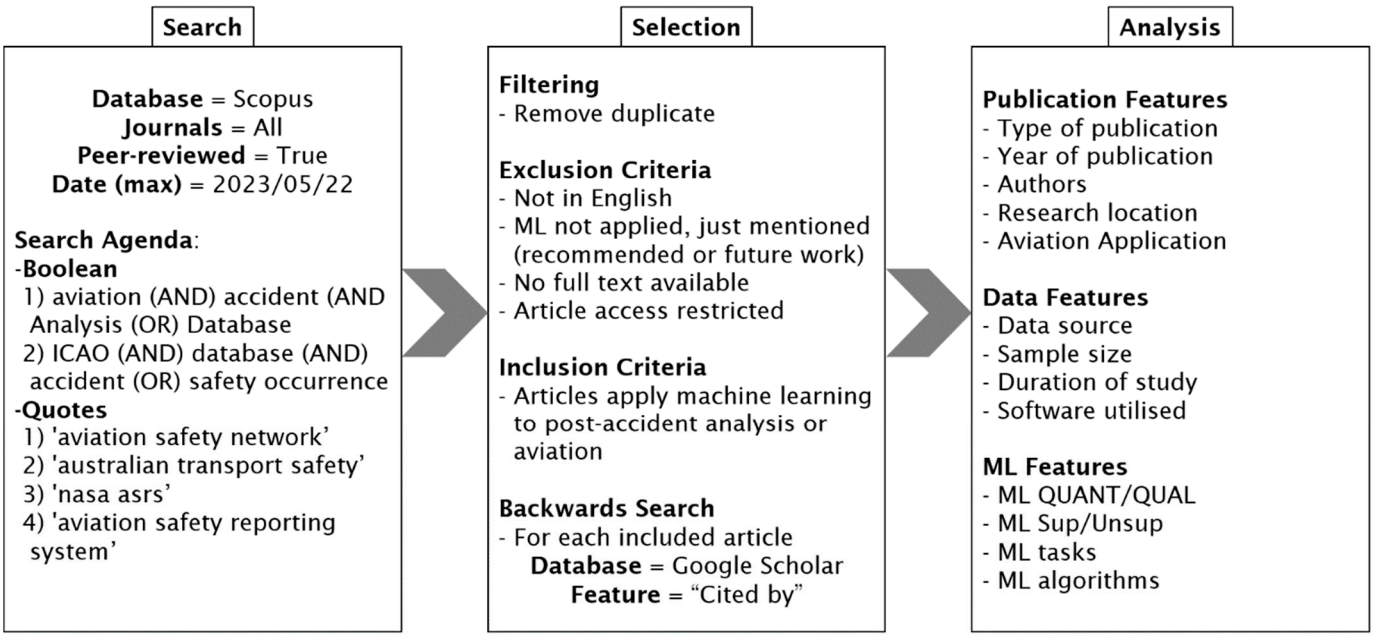


Figure 3. General framework of the process, from search, through selection, to analysis.

3.1. Search Strategy

The data source selection and search strategy were designed to ensure the thorough inclusion of pertinent studies. The search process involved several key steps, including database searches, backward search from references, and a structured data extraction process. The initial search commenced with 3832 articles identified. After the removal of duplicates (33 articles) and irrelevant records (3731 articles), a refined dataset was obtained for further analysis. The methodology employed for this systematic literature review (SLR) was designed to ensure thorough exploration of ML applications in post-accident analysis within the aviation industry. The review's purpose was to gather a diverse range of relevant studies that could enhance the understanding of how ML contributes to aviation safety and insights. The search process encompassed a series of well-defined and structured steps, which are detailed as follows.

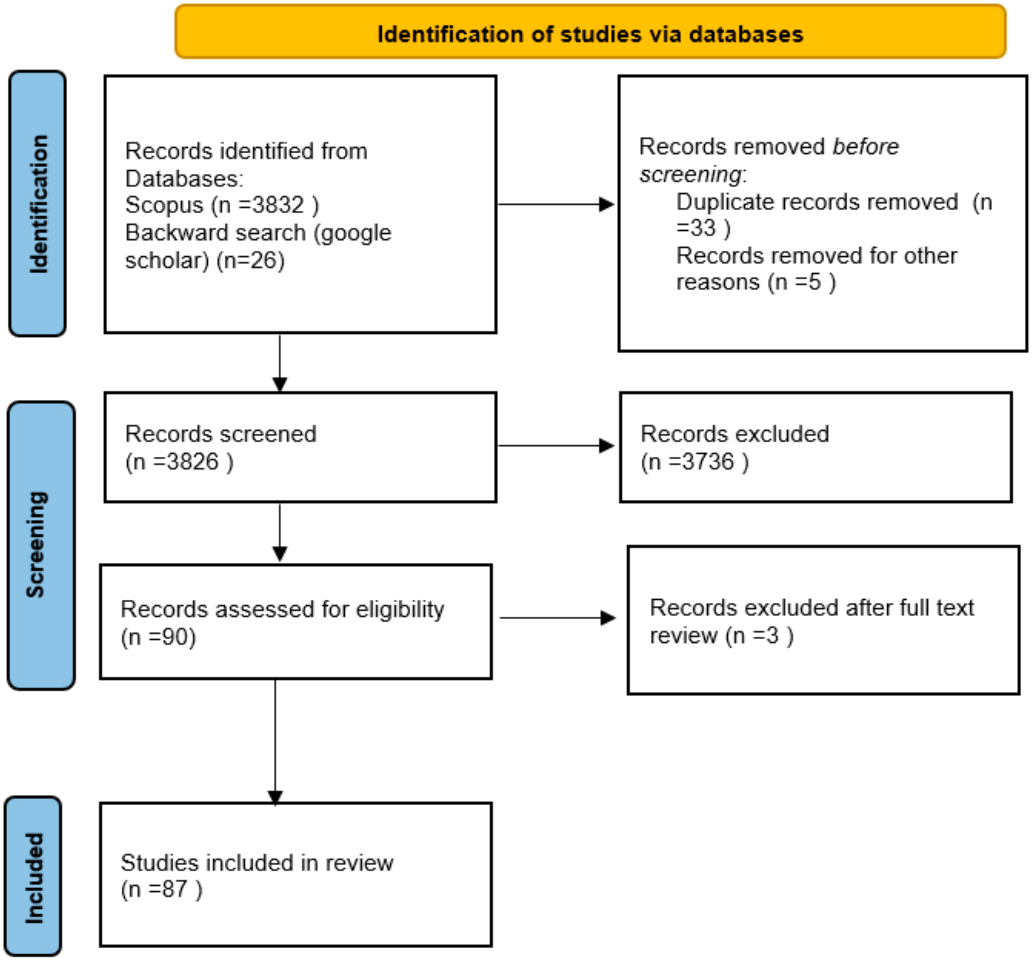


Figure 4. PRISMA flow diagram showing the selection process used in the systematic review.

3.2. Database Searches

The search was primarily conducted in the Scopus database due to its coverage of academic literature across multiple disciplines, making it a valuable resource for conducting SLRs [60,61]. This choice is aimed at ensuring the comprehensiveness of the review. The initial search using carefully chosen keywords yielded many articles, which were then refined through a stepwise process. After the removal of irrelevant articles (3731) and duplicates (33), the number was reduced to 68 articles that met the preliminary criteria for further evaluation.

The database searched included key sources such as "NASA (National Aeronautics and Space Administration) Aviation Safety Reporting System," "ICAO Accident Database," "Accident Safety Network," "NTSB AND Aviation AND Accident AND Database," "ICAO AND Safety AND Occurrence AND Database," and "Australian AND Transport AND Safety". The chosen keywords were designed to capture articles that may reference or discuss information from these sources within broader academic literature. The search criteria were expertly devised using Boolean operators (AND) and quotation marks ("") to precisely refine the search parameters, ensuring a high level of accuracy, and these final 87 papers are illustrated in Table A1.

3.3. Backward Search

Beyond the database searches, an additional approach was used to enhance the study's breadth. A backward search strategy was executed by reviewing the references cited within the identified studies. This approach, recommended for SLRs, helps identify relevant

articles that may not have been captured by the initial search [58,62,63]. By including this strategy, an additional 26 articles were retrieved and added to the review, thereby broadening the scope and depth of analysis, and ensuring a comprehensive coverage of the research landscape as depicted in Table 1.

Table 1. Number of papers initially identified by each search.

Search Criteria	Initial No.	Duplicates	Irrelevant	Final
Nasa asrs	9	-	6	3
ICAO accident database	15	-	13	2
Accident safety network	8	-	7	1
NTSB Aviation accident database	86	-	78	8
ICAO safety occurrence database	4	-	3	1
Australian transport safety	29	-	29	0
Aviation Accident Analysis AI	12	-	12	0
Aviation safety reporting system	159	2	132	25
Aviation Accident Analysis ML	61	5	44	10
Aviation accident database	457	10	440	7
Aviation accident analysis	2992	16	2972	4
Scopus article Total	**3832**	**33**	**3736**	**61**
Backward search (Google Scholar)	-	-	-	26
Total				**87**

3.4. Selection Criteria

The criteria for inclusion and exclusion were established to ensure the relevance and quality of the selected articles [58]. The selection process involved multiple steps, including the initial screening of abstracts and titles, followed by a more detailed assessment of the full-text articles. The criteria outlined below guided the selection process.

- Inclusion Criteria
  - Relevance to Aviation Accidents: Articles must directly address the application of machine learning techniques within the aviation industry.
  - Publication Language: Only articles written in the English language were considered to ensure accessibility for the research team.
  - Study Focus: The primary focus of the article should be on accident analysis within the aviation/air transport sector.
  - Study Type: Peer-reviewed journal articles, book chapters, theses, and conference proceedings were included only if they presented full-length papers with rigorous methodology and were published in reputable venues such as IEEE or AIAA.
- Exclusion Criteria
  - Non-English Language: Articles published in languages other than English were excluded due to language limitations.
  - Irrelevant Focus: Studies unrelated to aviation accident analysis were excluded.
  - Non-Aviation Applications: Articles discussing machine learning applications in contexts other than aviation accident analysis were excluded.
  - Publication Type: Abstract-only or lightly detailed conference proceedings and other non-peer-reviewed sources were excluded during full-text screening to ensure scholarly rigor.
  - Duplication: Duplicated records were removed to maintain the uniqueness of the dataset.

- Restricted Access: Articles for which full-text content was not accessible due to restrictions were excluded.

3.5. Data Extraction, Coding, and Quality

The data extraction process was conducted employing a structured tool developed through Microsoft Forms. This form encompassed a comprehensive set of 22 questions, addressing the selection criteria, statements assessing the quality of the article, and other important dimensions. These aspects included study methodologies, utilized datasets, primary research goals, participant details, publication year, author information, techniques, and relevant keywords. To ensure consistency and coherence, the extracted data were systematically organized using Microsoft Excel 365 (office 2021 build) [59]. This approach yielded a well-structured overview of the research landscape, facilitating efficient analysis and synthesis of the acquired information.

To evaluate the quality of the selected papers, a quality assessment process was included in the data extraction tool. This process aimed to establish the robustness and credibility of the studies, thereby bolstering the overall reliability of the review's findings. The quality assessment was systematically conducted by closely scrutinizing each selected paper, guided by a set of predefined quality assessment questions as presented in Table 2. These questions were designed to assess diverse dimensions of study quality, encompassing research methodology, data collection procedures, analysis techniques, and presentation of results.

Table 2. Quality assessment questions and the corresponding scope of papers they were applied to.

Quality Control Questions	Scope
1. Is the research objective clearly defined?	All
2. Is the context of the research clearly defined?	All
3. Does the study bring value to academia or industry?	All
4. Are the findings clearly stated and supported by the results?	All
5. Are limitations explicitly mentioned and analyzed?	All
6. Is the methodology clearly defined and justified?	All
7. Is the experiment clearly defined and justified?	All
8. Has the utilization of ML techniques been comprehensively described and justified?	ML
9. Are the chosen ML types and tasks appropriate for addressing the research questions?	ML
10. Is there clarity on the ML algorithms employed, including their rationale and suitability?	ML
11. Have biases and ethical considerations in the application of ML techniques been addressed?	ML
12. Are the implications of utilizing ML in post-accident analysis discussed?	ML
13. Is the integration of ML insights with safety measures thoroughly explored and elucidated?	ML

The quality assessment questions were formulated by drawing from well-established criteria in the realm of systematic reviews and were tailored to the specific context of machine learning applications in aviation post-accident analysis. Each paper underwent a meticulous assessment, with the evaluations' outcomes recorded to ensure consistency and accuracy throughout the evaluation process. It is important to note that the quality assessment was not solely reliant on quantitative metrics; rather, it encompassed a holistic evaluation of factors such as the alignment of the study with research objectives, the clarity of research questions, the validity of methodologies, and the appropriateness of interpretations.

The quality assessment procedure led to the exclusion of an additional 7 papers. This stringent process ultimately resulted in the final selection of 87 papers that met the rigorous quality standards established by the review [58].

3.6. Data Synthesis

The process of data synthesis involved systematic aggregation and analysis of the extracted data from the curated subset of 87 papers. This step aimed to distill and concisely summarize the key findings, trends, and insights across the studies to provide an overview of ML applications in aviation analysis. The extracted data were systematically organized and categorized based on common themes and research aspects. Patterns and recurring trends were identified by comparing the methodologies, approaches, algorithms, and outcomes of the studies. This facilitated the development of a coherent narrative illuminating diverse ways ML techniques enhance aviation safety through the post-accident analysis and identified gaps, challenges, and future research directions. Meaningful conclusions were drawn, providing a holistic perspective and contributing to an understanding of the implications of ML in aviation post-accident analysis. The next section presents outcomes of the systematic literature review, offering insights into a wide array of learning approaches, techniques, and their implications for aviation safety in post-accident analysis.

4. Results

The SLR on ML applications in post-accident analysis within the aviation industry identified 87 relevant articles. These studies collectively illustrate the breadth of ML techniques employed, their capacity to enhance aviation safety and the valuable insights they provide into the causes and patterns of aviation accidents. This section also addresses Research Question 1 by outlining the types of ML techniques and tasks used across the literature. The remaining research questions, 2a (ML techniques used), 2b (ML tasks performed), and 2c (data sources utilized), are each discussed in detail in Section 4.9, specifically in Section 4.9.1, Section 4.9.2, and Section 4.9.3, respectively.

4.1. Aviation Application

The systematic review categorized the post-accident analysis studies based on their specific aviation applications. The included studies were thoroughly examined to identify and analyze the aviation applications employed. Most of the studies focused on ALL aviation applications (44%), Regular Passenger Transport (S/RPT) (31%), and general aviation (GA) (7%). Other studies addressed Military aviation (3%) and Cargo (1%). This distribution is visually represented in Figure 5 as a Pareto plot, which illustrates that ALL received a more comprehensive focus in the studies compared to other aviation applications, at around 55% of all the studies included.

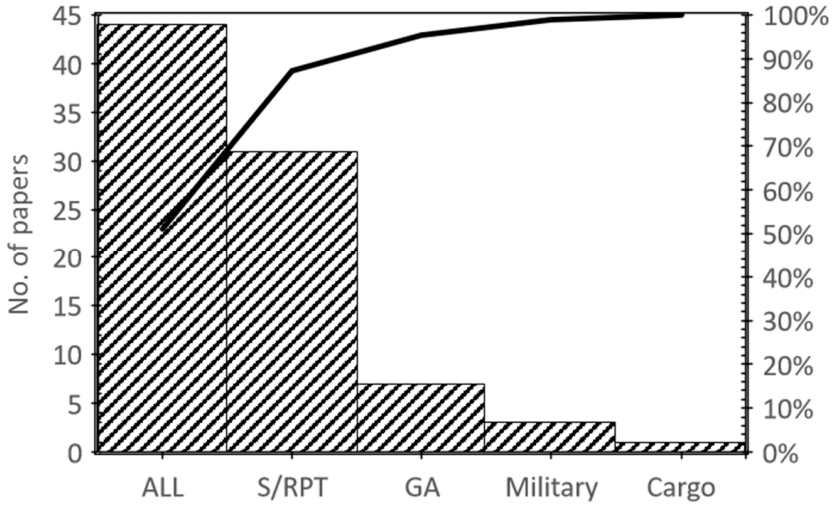


Figure 5. Pareto plot of the publication count for different aviation types identified in this study where GA = General Aviation, and S/RPT = Regular Passenger Transport.

4.2. Sources and Types of Publications

Although the Scopus database was used exclusively to search for relevant studies, the inclusion of the backward search offered a slightly wider pool of readily identifiable studies, using a reproducible approach. As such, Scopus was the primary source for included studies of suitable quality, with 69% of the articles found directly though Scopus. However, for those 61 articles from Scopus, the backward search in Google Scholar, using the "cited by" feature, a further 26 unique articles that directly referenced one or more of those 61 Scopus articles were also relevant and of suitable quality. Figure 6 below shows the breakdown of both the source and type of publication. As expected, a larger proportion of Google Scholar results are conferences, due to the limited number of such events and proceedings which are indexed by Scopus. The same is true of research theses.

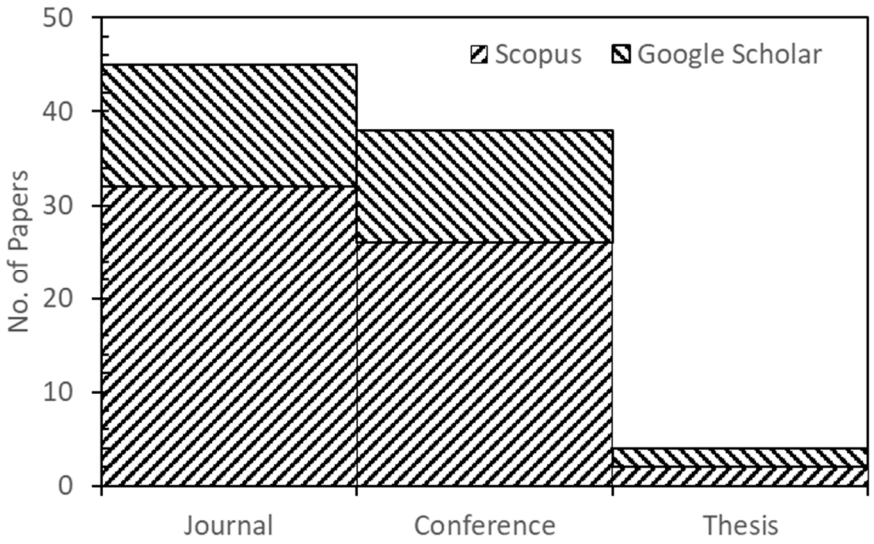


Figure 6. Stacked bar chart shows the distribution of source and type of publication.

The categorization of AI/ML post-accident analysis studies based on their types of publication revealed distinct patterns. The included studies were examined to identify and analyze their distribution across various publication types. Most of the studies were centered on journal articles, constituting 52% of the total. Conference papers accounted for 44% of the publications, signifying their significant presence in the field. A smaller portion of the studies were Theses (4%). Figure 6 visually represents the distribution of these publication types, clearly indicating that journal articles were the primary publication type.

4.3. Duration of Studies

The dataset periods ranged widely across the studies, encapsulating various periods. Some studies utilized extensive datasets that spanned many years or even multiple decades, such as those covering 62 years [64] and 101 years [65]. Other studies focused on more recent periods, with dataset durations of 1 year [66] and 2 years [67,68]. Also of note is the most extreme example being that by Koteeswaran et al. spanning 96 years. Figure 7 visually represents the distribution of the number of publications/studies by the number of years of accidents or incidents studied, i.e., the duration of the study in years. The distribution is skewed to shorter durations, with a peak in the 10-to-19-year category, when it is broken down into decades. The total of the 67 studies that included clear information about their duration was 24.7 years, with a median of 15 years. Duration is estimated as the top year minus the start year plus 1, so a study is at least one year long, although this makes the error for each year at ± 0.5 years.

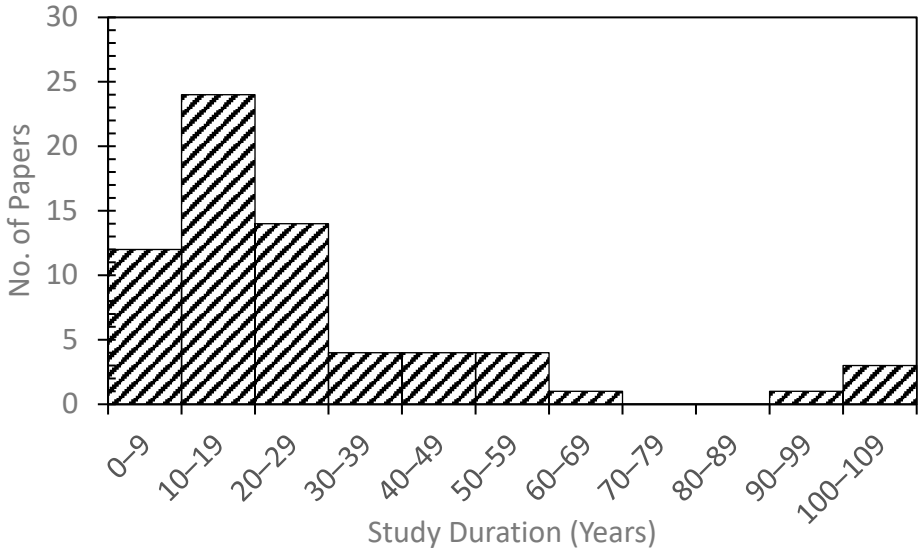


Figure 7. Bar chart shows the distribution of the publication count relative to the study duration in years.

4.4. Publication Trends

The analysis also revealed diverse variations in the years of publication. Studies encompassing earlier years, such as [69], provided insights into the field's historical progression. More recent studies, like those of Choi & Gibson and Nogueira et al., [37,70], added contemporary perspectives to the evolving landscape of ML in aviation safety analysis.

This study focused on data from 1998 to mid-2023 to analyze the trends in ML applications in post-accident analysis within the aviation industry. The analysis revealed interesting patterns in the publication of relevant studies over the years. Notably, the year 2020 exhibited a peak in publication activity, followed by 2021 and 2019, as illustrated in Figure 8. The graph depicting the publication trends highlights a distinct evolution. From 1998 to 2008, there was a limited number of studies applying ML techniques to post-accident analysis in aviation, with only one paper being published per year. Subsequently, a gradual increase was observed, with the number of published papers reaching six by 2018. The most significant surge in publications occurred from 2019 onwards, when the rate reached ten papers per year. It is important to note that the range from 1998 to the present is due to the availability of qualifying papers within the Scopus database, signifying the emergence and subsequent growth of this field after 1998.

The stratification of data in Figure 8 suggests further analysis is warranted. That is, three distinct regions in terms of publication counts can be identified looking at the plot; these are three generations of research, each marked by notable shifts in focus and intensity. The first generation, spanning from 1998 to 2008, reflects a foundational period characterized by an average of 0.4 papers per year. This era witnessed the emergence of key methodologies starting with decision trees in 1998 and logistic regression in 2004. In the second generation, extending from 2009 to 2018, research activity increased, with an average of 3.4 papers per year. During this period, methodologies such as support vector machines in 2010 and deep learning in 2017 gained prominence. Finally, the current third generation, covering 2019 to mid-2023 (2022), exhibits a significant surge in research output, averaging 12.5 papers per year. This phase signifies a rapid increase in research efforts, underscoring the dynamic evolution of AI/ML methodologies applied to post-accident analysis in aviation.

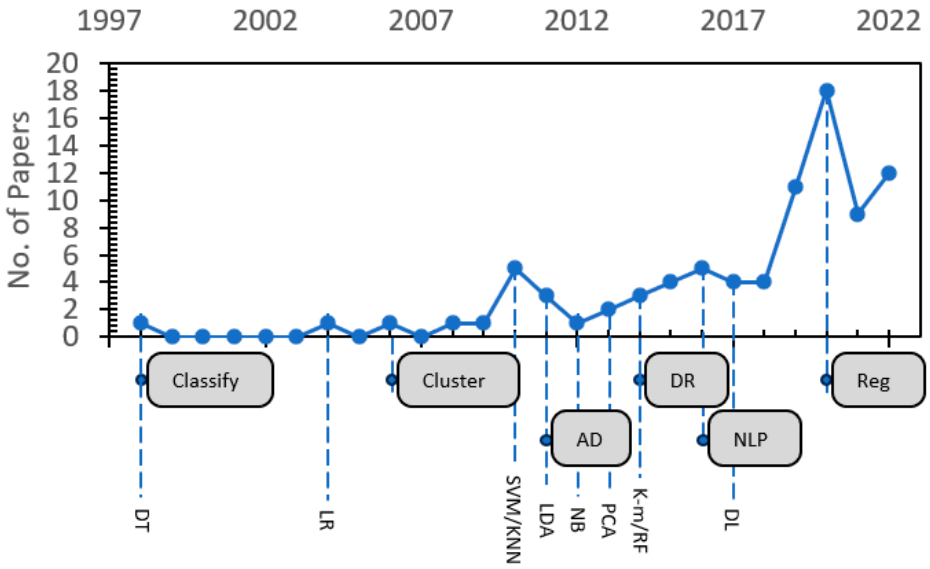


Figure 8. Line chart illustrating the publication trend from 1998 to 2022. Also shown are the first implementation of each ML task (boxes) and the first implementation of each specific ML algorithm (vertical text).

This trajectory indicates the growing recognition and adoption of ML in post-accident analysis within the aviation industry. The steady increase in publications from 2009 to 2018 could be attributed to the gradual integration of data-driven approaches in accident investigation. The substantial rise in publications from 2019 to 2020 signifies a significant shift towards more sophisticated ML techniques and their applications, reflecting the industry's acknowledgment of the potential of these methods in enhancing safety measures and accident insights. The trend is also likely driven by the fidelity of the NLP techniques to infer the ontology of aviation safety.

In general, the evolution of research generations, with increasing efforts, indicates the growing recognition and adoption of ML in post-accident analysis within the aviation industry. The steady increase in publications from 2009 to 2018 could be attributed to the gradual integration of data-driven approaches in accident investigation. The substantial rise in publications from 2019 to 2020 signifies a significant shift towards more sophisticated ML techniques and their applications, reflecting the industry's acknowledgment of the potential of these methods in enhancing safety measures and accident insights. The trend is also likely driven by the fidelity of the NLP techniques to infer the ontology of aviation safety.

4.5. Software Packages Utilized for Analysis

The analysis of software packages employed in the reviewed studies highlights the diverse tools utilized for conducting machine learning research within the context of aviation post-accident analysis. Among the various options available, Python programming language emerged as the most prominently favored choice, accounting for 40% of the studies. Python's popularity since it was released in 1995 can be attributed to its versatility, extensive ML libraries, batch scripting capabilities, and the broader data science community it supports [71].

In addition to Python, several other software packages were prevalent in the studies. The Waikato Environment for Knowledge Analysis (WEKA) software package was identified as a key tool, utilized in 20% of the studies [72]. WEKA's user-friendly interface and comprehensive suite of ML algorithms made it a preferred option for researchers. MATLAB, known for its robust numerical computing capabilities, was another significant choice, adopted by 16% of the studies.

While most studies leaned towards Python, a notable subset opted for R programming packages, constituting 11% of the total. R's statistical capabilities and data visualization functionalities were reasons behind its selection.

Beyond these primary software packages, a smaller proportion of studies incorporated other tools. Structured Query Language (SQL), often used for managing and querying databases, was employed in 4.5% of the studies. Similarly, the Statistical Package for the Social Sciences (SPSS) and the Statistical Analysis System (SAS), renowned statistical software packages, each contributed to 4.5% and 2% of the studies, respectively. RM Studio, a platform for statistical analysis and graphical representation, was utilized in 2% of the studies.

Visualized in Figure 9, the distribution of software packages underscores Python's commonality, followed by WEKA and MATLAB. The fact that 80% of the studies correspond to almost 40% of the different software packages indicates that there is currently no consensus on the software that is best to utilize for aviation safety analysis; likely due to a variety of selection criteria related to goals and the specific requirements. However, Python, as a free open-source package, appears to be becoming dominant.

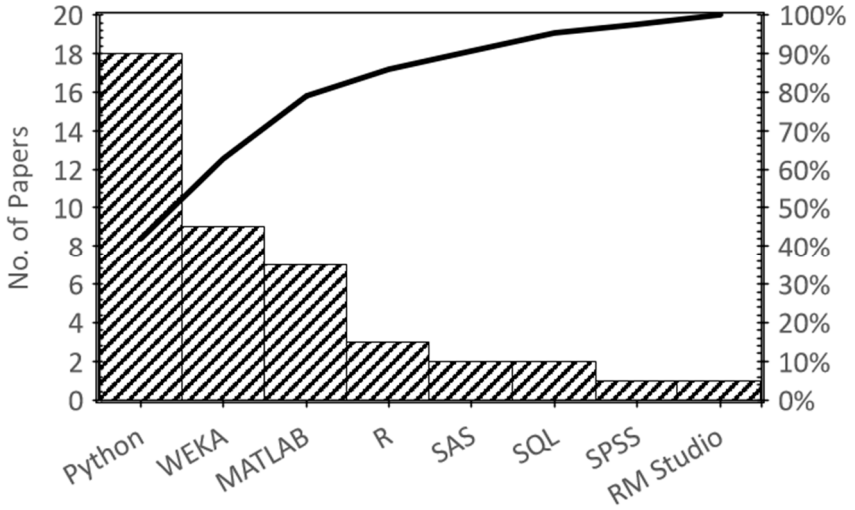


Figure 9. Pareto plot of publication count as categorized by the software package utilized in the study, to implement the relevant ML tools and techniques.

4.6. Data Source Utilized in the Studies

The Aviation Safety Reporting System (ASRS) dataset emerged as the most prominently utilized data source, featuring in 41% of the studies. This was followed by the National Transportation Safety Board (NTSB) dataset, which constituted another significant data source, contributing to 25% of the studies. Other data sources collectively constituted 18% of the total. This category encompassed a diverse array of sources that researchers drew upon to address their specific research questions.

A smaller fraction of studies incorporated additional data sources to enrich their analyses. The FAA dataset, comprising aviation-related data, was employed in 8% of the studies. Some researchers conducted surveys to gather primary data, representing 3% of the total. Accessing data from the Aviation Safety Network (ASN) and the Australian Transport Safety Bureau (ATSB) contributed to 4% and 1% of the studies, respectively. Figure 10 visually encapsulates the distribution of data sources, clearly illustrating the prevalence of the ASRS and NTSB datasets across the studies and reflecting the 41 and 25 per cent of overall air miles flown in these air domains.

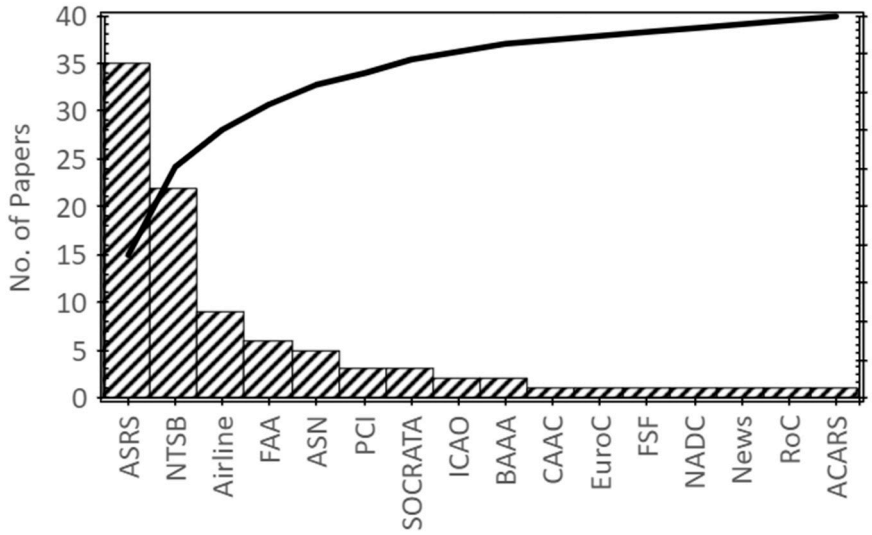
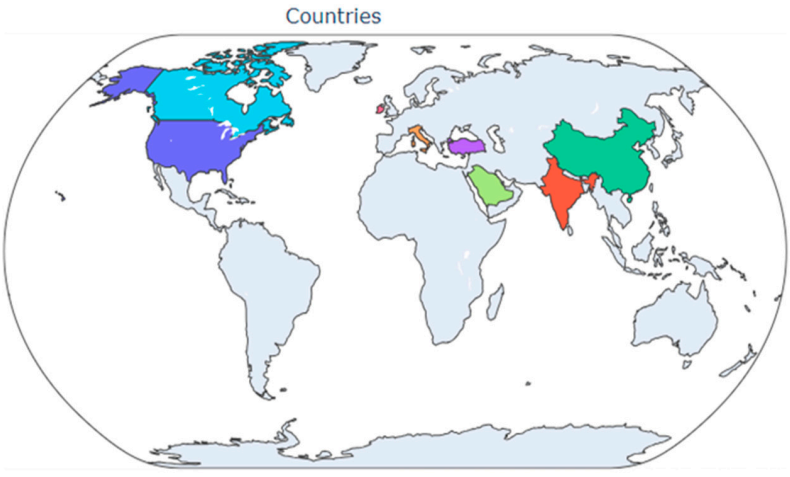


Figure 10. Pareto plot showing the publication count as categorized by the data sources utilized.

4.7. Country of Publication Analysis

The geographic distribution of research output offers contextual insights into global engagement with ML applications in aviation post-accident analysis. While this factor may not directly influence the analytical outcomes of ML models, it helps situate the research within broader patterns of technological advancement and regional emphasis. For instance, air travel remains dominant in developed regions such as North America and Europe, which is reflected in research activity. Notably, the United States (USA) emerges as a leading contributor with 45 publications, reflecting sustained academic and institutional investment in advancing ML for aviation safety. India follows with 12 publications, indicating strong regional engagement. China, with 11 publications, demonstrates increasing attention to ML integration in aviation. Canada, Türkiye, Italy, and Saudi Arabia each contribute 3 to 4 publications, showing diverse regional participation. Furthermore, countries such as Australia, the UK, and Germany each offer one publication, enriching the global discourse on ML in aviation safety. Figure 11 visually represents the distribution of research output across these countries. Although this analysis does not impact the methodological rigor of the review, it helps illustrate the global nature of scholarly interest in using ML to enhance aviation safety.



Region	Count
North America	45
Asia	21
Europe	9
Middle East	2

Figure 11. Geographic distribution of publications by country (**Left**), and count by ICAO Region/Continent (**Right**).

4.8. Notable Contributors in the Literature

In this study, authors who have made noteworthy contributions to the realm of ML applications in aviation post-accident analysis have been identified. Table 3 presents an overview of authors who have exceeded two publications in this domain, accompanied by their respective affiliations. The recurring presence of certain authors across multiple papers accentuates their active involvement and proficiency in driving the advancement of research within this field. Notably, Khan, L and Mavris, DN stand out as prolific contributors, each boasting five papers in the study period of 1998–2023. Likewise, Christopher, AA, and Puranik, TG have each authored four papers, while Alias Balamurugan, SA, Mahadevan, S, Rao, AH, Robinson, SD, and Zhang, X have made significant contributions with three papers each.

Table 3. Authors with notable contributions, showing their relevant number of publications, along with the years of the first and last (most recent) publication.

Author	Count	First	Last	Affiliation
Khan, L	5	2008	2020	Department of Computer Science, The University of Texas at Dallas
Mavris, DN	5	2020	2022	Georgia Institute of Technology, Atlanta, GA 30332, United States
Christopher, AA	4	2013	2022	Research scholar Anna University Tamil Nadu, India
Puranik, TG	4	2020	2022	Universities Space Research Association, NASA Ames Research Center, Moffett Field, CA, USA
alias Balamurugan, SA	3	2013	2022	Research Scholar Anna University, Tamilnadu, India
Mahadevan, S	3	2015	2021	Department of Civil and Environmental Engineering, Vanderbilt University, Nashville, TN, USA
Rao, AH	3	2018	2020	Collins Aerospace, 400 Collins RD, MS 124–319 Cedar Rapids, USA
Robinson, SD	3	2015	2019	Parks College of Engineering, Aviation and Technology, Saint Louis University, Saint Louis, MO 63103, USA
Zhang, X	3	2015	2018	Department of Civil and Environmental Engineering, School of Engineering, Vanderbilt University, Nashville, TN, 37235, USA

4.9. Machine Learning Dimensions

This section addresses Research Questions 2a, 2b, and 2c by providing insights into various dimensions of machine learning, specifically focused on data analysis in accident research within the aviation industry. Through systematic analysis of collected data and literature, each dimension is explored to understand the intricacies of machine learning applications in post-accident analysis. This includes investigating machine learning types, tasks, and algorithms utilized in the field.

4.9.1. Machine Learning Approaches

In response to Research Question 2, this section shows the research approaches adopted in the selected studies. It was found that most of these studies (62%) employed a quantitative approach, emphasizing the collection and analysis of numerical data to draw conclusions and make inferences. Also, a considerable proportion of studies (30%) utilized a mixed-methods approach, integrating both quantitative and qualitative methods while a smaller portion of studies (8%) employed a qualitative-only approach, focusing on understanding underlying meanings, motivations, and nuances of the phenomena under investigation.

The prevalence of quantitative and mixed methods approaches signifies a commitment to robust empirical data collection and analysis, enhancing the credibility and validity of the research findings.

4.9.2. Machine Learning Types

In response to Research Question 2a, the researchers explored various ML types used across the selected studies. The analysis revealed that a large portion of the studies (65%) applied supervised learning techniques. Supervised learning involves training models using labelled data, allowing them to make predictions or classifications based on patterns identified in the training set [19].

Furthermore, 30% of the studies embraced unsupervised learning, a technique aimed at discovering patterns or structures in data without the need for predefined labels. Unsupervised learning algorithms enable researchers to uncover hidden insights and relationships within datasets [21,22]. While a smaller portion of studies (5%) incorporated semi-supervised learning, which combines elements of both supervised and unsupervised learning. Semi-supervised learning leverages a limited amount of labelled data with a larger pool of unlabeled data to enhance the efficiency and performance of the model [23].

To provide a visual representation of these findings, Figure 12 depicts the distribution of ML types employed in the context of aviation post-accident analysis.

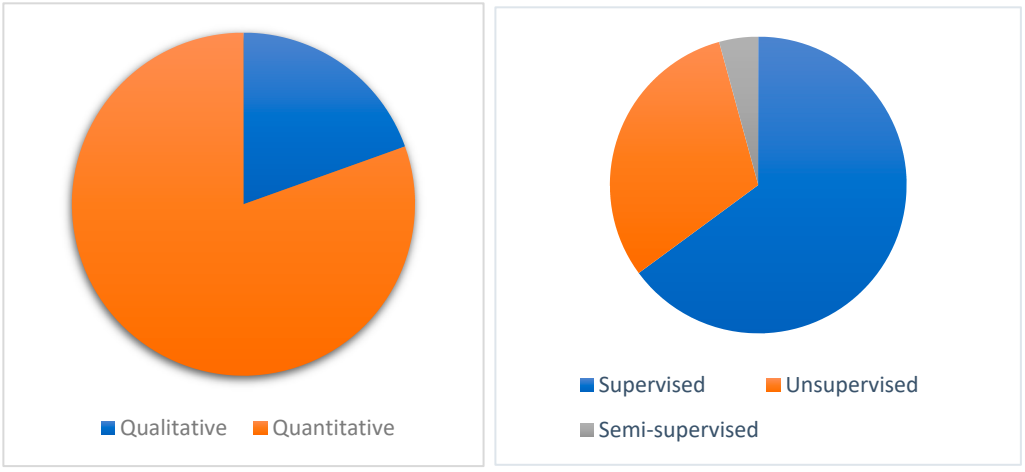


Figure 12. (**Left**) the breakdown of data types used in the ML studies, and (**Right**) the broad type of ML used.

4.9.3. Machine Learning Tasks

Addressing Research Question 2b, researchers investigated machine learning tasks employed to enhance post-accident analysis within the aviation industry. The investigation revealed a range of tasks that researchers engaged with to extract valuable insights from accident data.

The most prominent ML task undertaken was classification, accounting for 59% of the studies. Classification involves categorizing data into predefined classes or categories, aiding in the identification of patterns that differentiate various accident scenarios [73,74].

Furthermore, 18% of the studies focused on NLP, a task that involves the extraction of meaningful information from textual data. NLP techniques are instrumental in unraveling insights embedded in textual descriptions of accidents and incidents [75]. Anomaly detection, with 8% representation, emerged as another significant task. This involves identifying unusual or exceptional patterns in the data that could indicate potential safety hazards or exceptional circumstances [66]. Clustering, contributing to 9% of the studies, entails

grouping similar data points together, helping to unveil inherent structures within accident datasets [75].

A smaller fraction of studies (5%) targeted data reduction, aiming to distill relevant information from large datasets, and a mere 1% of studies involved regression, a task focused on predicting numerical values based on input data. For a visual representation of the distribution of these ML tasks, please refer to Figure 13. Classification accounts for 60%, indicating that while not dominant, the task is the primary consideration for aviation accident analysis with ML. This exploration of machine learning tasks elucidates the multifaceted strategies employed to glean insights from aviation accident data, contributing to the advancement of safety measures within the industry.

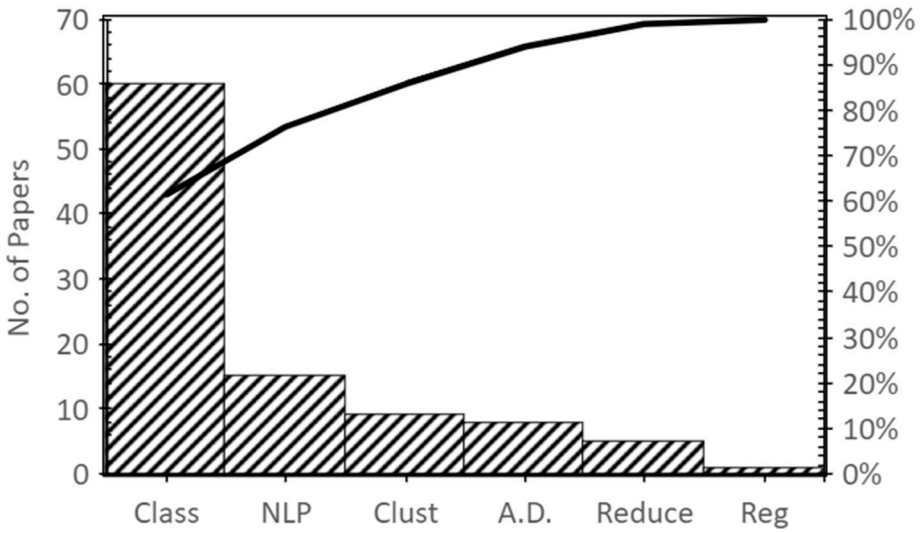


Figure 13. Pareto plot showing the publication count for each of the ML tasks utilized.

4.9.4. Machine Learning Algorithms

In response to Research Question 2c, researchers explored various machine learning algorithms utilized for post-accident analysis within the aviation industry. The analysis revealed diverse algorithms deployed to uncover meaningful insights from complex aviation accident-related data. The top 10 machine learning techniques used in this context are detailed in this section.

The most widespread algorithm employed was DL, constituting 21% of the studies. Deep Learning's capacity to capture intricate patterns and representations in complex data has made it a favored choice for unraveling hidden insights from aviation accident data [67,68].

Decision Trees (DT) were utilized in 16% of studies, providing a visual representation of decision-making processes and enabling the identification of key factors in accidents. SVM followed closely, being adopted in 14% of studies. SVM is a robust classification and regression technique that effectively handles both linear and non-linear data. PCA was applied in 4% of studies, facilitating dimensionality reduction and feature extraction from high-dimensional accident data. Linear Discriminant Analysis (LDA) played a role in 6% of studies, offering a powerful technique for dimensionality reduction and classification. LR was employed in 6% of studies, providing insights into the relationships between accident features and outcomes.

KNN emerged in 10% of studies, enabling the classification of accidents based on the similarity to neighboring data points. NB was utilized in 12% of studies, offering a probabilistic approach to classification based on Bayes' theorem. RF algorithms were applied in 9% of studies, highlighting their utility in ensemble learning for classification

and regression tasks. K-Means, a clustering algorithm, was used in 2% of studies, aiding in grouping similar accidents for pattern recognition [76]. It is important to note that many studies have tried multiple approaches and so are represented often.

For a visual representation of the prevalence of these ML algorithms, please refer to Figure 14. This analysis highlights the diverse toolkit of algorithms that researchers have harnessed to extract meaningful insights from aviation accident data, fostering advancements in safety measures and accident prevention. However, the fact that the 80% mark of the counts corresponds to half of the algorithm indicates that there is currently no primary focus in the ML algorithms utilized in aviation accident analysis. DL is an early leader, but it will be interesting to see how this aspect continues to evolve.

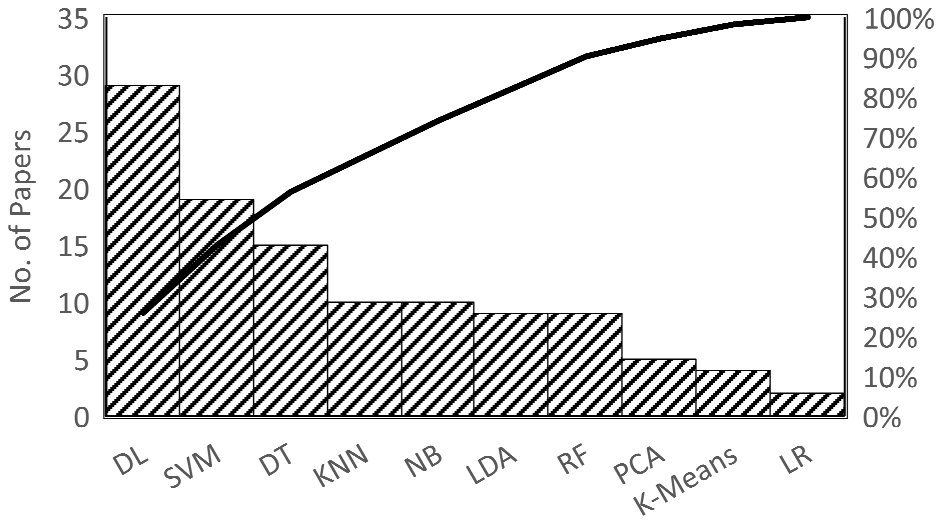


Figure 14. Pareto plot of the publication counts based on the ML algorithms utilized.

4.9.5. Machine Learning Summary

The systematic review unveiled ML techniques harnessed in post-accident analysis within the aviation sector. Ranging from conventional algorithms like DT and SVM to advanced methods such as DL and NN, these techniques were found to be crucial tools. They facilitated tasks like clustering, classification, and regression, contributing across various stages of accident analysis, including identifying factors that led to accidents and predicting potential outcomes. Examples include the works of Paul; Pimm et al., [77,78] who utilized ASRS and ECCAIRS datasets within the domain of civil aviation. Zhou et al., [68] employed ML techniques on ACARS (Aircraft Communication Addressing & Reporting Systems) data, thus offering insights into civil aviation. In parallel, Ackley et al., delved into the application of ML, using FOQA (Flight Operational Quality Assurance) data within the commercial aviation context [79].

4.10. Enhancements to Safety

The application of ML algorithms within the aviation industry has markedly advanced safety enhancement by facilitating novel insight extraction from aviation accident data. Unlike traditional techniques such as logistic regression and proportion analysis, ML approaches have enabled the identification of previously unknown risk factors and patterns.

For instance, unsupervised learning techniques, such as clustering and latent topic modeling [80] have been used to reveal underlying factors such as unusual flight control deviations and latent fatigue-related behaviors. Basora et al. demonstrated that ML models applied to ACARS data detected warning signs of anomalies not previously captured through traditional statistical methods [81]. Similarly, Ackley et al. [79] employed FOQA

data in deep learning models to highlight temporal signal irregularities that would be difficult to isolate using conventional methods. These findings emphasize the unique contributions of ML to aviation safety research. Specific algorithms such as SVM, KNN, LDA, and DL was responsible for advancing risk factor detection beyond the capabilities of traditional methods.

4.10.1. Risk Assessment and Prediction

Machine learning algorithms like SVM, RF, and NB have applied to analyze historical accident datasets to forecast potential risks and hazards in aviation systems [33,67,82]. While these studies indicate encouraging predictive capabilities, further validation in operational environments is needed to establish ML's consistent effectiveness in risk forecasting across diverse flight contexts.

4.10.2. Anomaly Detection

Algorithms such as DT and DL models have been utilized to identify anomalies deviations from expected operational behavior in aviation data [66,83]. These studies demonstrate ML's potential to serve as an early warning mechanism for operational disruptions. However, broader generalization and domain expert involvement in validating flagged anomalies remain limited.

4.10.3. Classification of Accidents and Factors

A wide range of classification algorithms including SVM, KNN, and GB have shown promise in categorizing aviation accidents and identifying underlying causal factors [74,84,85]. While several studies report strong classification accuracy, the effectiveness of these models in operational decision-making has yet to be systematically benchmarked against traditional analytical techniques.

4.10.4. Natural Language Processing (NLP) Applications

NLP-based algorithms like LDA and RNN have been explored for analyzing textual narratives in aviation safety reports [46,75,86]. These approaches have shown potential in extracting hidden trends and summarizing incident insights. Nonetheless, the integration of NLP-derived outputs into regulatory or investigative workflows remains nascent.

4.10.5. Clustering and Pattern Identification

Clustering algorithms such as K-Means, DBN and Self-Organizing Maps (SOM) have been applied to uncover latent structures in aviation accident data [87,88]. These unsupervised approaches offer useful exploratory insights, although they typically require supplementary validation from domain experts to ensure practical utility and relevance.

4.10.6. Dimensionality Reduction

Techniques like PCA and NMF have aided in reducing the dimensionality of complex data, thereby enabling a more manageable representation for analysis [89,90]. This simplification facilitates the identification of crucial features and relationships within the data.

4.10.7. Human Factors Analysis

Some ML algorithms have been applied to explore human performance and decision-making patterns in aviation accidents, offering new perspectives on behavioral and cognitive contributors to incidents [70]. These exploratory studies suggest that ML may support human factors research; however, robust interdisciplinary validation with aviation psychologists and domain experts is essential to ensure reliability and operational value.

In essence, the combination of various ML algorithms with aviation accident data has led to a comprehensive enhancement of safety measures and the generation of insights that contribute to informed decision-making within the aviation industry.

5. Discussion

By examining the contextual nuances that influence the application of ML and delving into the varied techniques employed, an understanding of the dynamic landscape of ML in aviation safety has been provided. While the study encompasses a timeframe of 25 years (1998–2023), the study has summarized all the ML used in various aviation applications (Table 1, Table 2) and discussed different software packages (Figure 9), data sources (Figure 10), country of publication (Figure 11), ML types (Figure 12), ML Tasks (Figure 13) and ML algorithms (Figure 14). This study has reviewed the ML algorithms used in the literature and summarized them in Table A1 for each chosen article (see the Appendix A). The review aimed to provide a comprehensive analysis of past research and offer a forward-looking perspective on the potential impact of ML in further enhancing aviation safety standards.

5.1. Diverse Machine Learning Techniques and Applications

The unique challenges posed by the aviation industry necessitate a focused approach to enhancing safety through ML applications. This review underscores the importance of contextualizing ML methods in alignment with aviation-specific safety needs. ML models have been widely applied not only to post-accident investigations but also to predictive modeling, ATM, and safety enhancement in general aviation, passenger transport, and military aviation [91]. Given the diversity of tasks, this study categorized ML techniques according to their learning paradigms and target applications. These include supervised models such as (SVM, DT) unsupervised methods such as (clustering, LDA), and deep learning architectures such as (CNNs, LSTMs). Each of these models offers specific benefits in dealing with the high-dimensional and often unstructured data typical of aviation accident records [92]. While many studies reported strong performance metrics such as this review did not conduct a comparative evaluation of ML models against traditional statistical techniques because such comparison was beyond this research's scope and not among the formulated research questions. Instead, the objective was to map out ML trends, use-cases, and data types within the aviation safety literature.

5.2. Data Sources and Quality Assessment

The foundation of effective ML-driven post-accident analysis lies in the quality and relevance of the data sources employed. Notably, the extensive utilization of datasets such as ASRS and NTSB highlights the industry's reliance on real-world aviation data to unveil intricate patterns and contribute to safety enhancement. The rigorous quality assessment protocols followed across the selected studies uphold the integrity, validity, and reliability of research findings, elevating the credibility of outcomes [14,18].

However, it is important to acknowledge the inherent differences in these data sources and how they may affect generalizability. The ASRS, being a voluntary reporting system, is susceptible to underreporting and reporting bias, as incidents are self-reported by aviation personnel. This may lead to a dataset that overrepresents certain event types while underrepresenting others, particularly more serious or sensitive occurrences. In contrast, the NTSB database is based on official investigations and generally provides more comprehensive and structured accounts, though it may be limited to more severe events. These differences in data scope, completeness, and reporting criteria can influence the training

and performance of ML models, potentially affecting their applicability across different operational or regional contexts.

It is pivotal to recognize the decisive role that high-quality datasets play in generating actionable insights for aviation safety. The review not only categorizes the datasets in use but also delves into their implications for research outcomes. This in-depth understanding empowers stakeholders to choose appropriate data sources and underpins data-driven decision-making, amplifying the precision and pertinence of aviation safety analyses.

5.3. Safety Enhancement and Insights Generation

Machine learning (ML) techniques hold promises for improving safety standards within the aviation industry. These methodologies excel at identifying concealed patterns and underlying factors that might have contributed to accidents, thereby empowering stakeholders to proactively address safety concerns. Additionally, the predictive ability of ML algorithms equips practitioners with the means to anticipate potential occurrences, facilitating the development of precise preventive strategies. Through furnishing valuable insights into aviation accidents, these techniques play a pivotal role in informed decision-making and fostering an all-encompassing enhancement of safety measures [40,67,82,93].

While the reviewed studies cover a wide range of aviation contexts, their practical implementation varies considerably. Some ML applications have demonstrated meaningful contributions to operational settings, such as automating narrative report classification, flagging safety-critical anomalies, and supporting predictive maintenance. Models trained on real-world data from the NTSB and ASRS datasets have been used to uncover hidden patterns and causal factors behind incidents, providing investigators with faster and more comprehensive analyses. Despite these advances, widespread adoption of ML tools in aviation safety management systems remains limited. Most reviewed studies present proof-of-concept models or offline experiments rather than deployed systems in live environments. Key barriers include regulatory uncertainty, interpretability concerns, data access limitations, and integration challenges with existing safety workflows.

Therefore, while ML's potential is clear, realizing its full practical value requires cross-disciplinary collaboration between data scientists, domain experts, software engineers, and regulators. Future studies should also incorporate case studies or implementation roadmaps to bridge the gap between theory and deployment.

5.4. Variability in Study Types and Purposes

This review underscores the imperative of considering diverse study types and objectives to yield meaningful advancements in aviation safety. By showcasing the array of methodologies and algorithms (Figure 14) adopted, insights into the spectrum of research endeavors are highlighted.

5.5. Critical Reflections on Cited Studies

This review acknowledges that many studies included in this synthesis possess limitations that warrant discussion. One notable limitation is the lack of prospective validation across most studies. While retrospective analysis is prevalent, very few works validate their models in real-world operational settings, which limits the generalizability of their findings. Additionally, there is a scarcity of collaboration with aviation domain experts, such as accident investigators, air safety analysts, or regulatory personnel, in the model development and evaluation process. This disconnect can result in ML models that are technically sound but misaligned with operational needs, regulatory standards, or safety-critical decision-making workflows. Future studies should prioritize interdisciplinary co-design, involving aviation safety professionals from early stages of data labeling, model evaluation, and interpretation, to ensure that ML insights are both actionable and trust-

worthy. Furthermore, several studies depend on datasets like the NTSB database, which in earlier decades (especially between 1980–1995) lacked detailed and consistent reporting. This incompleteness challenges the integrity of machine learning outcomes due to the "garbage-in-garbage-out" principle [94]. Without rigorous data preprocessing and domain-informed feature engineering, ML models trained on sparse or noisy data may yield unreliable or misleading results. These limitations necessitate caution in interpreting findings and reinforce the importance of comprehensive data validation, domain-aligned model development, and method transparency.

5.6. Future Research Directions

The literature underscores the significance of ethical considerations and bias mitigation in ML applications. The effectiveness of aviation safety analysis in the future relies on the implementation of transparent, fair, and unbiased ML models. As the ML field evolves, seven exciting avenues for future research emerge, charting the course for further advancements in ML-driven aviation safety analysis.

5.6.1. Interpretable and Explainable AI

The aviation industry's emphasis on transparency and accountability necessitates the development of ML models that provide understandable explanations for their decisions, particularly when influencing safety-critical choices. Future research should explore interpretable models such as decision trees, SHAP values, LIME, and attention-based deep learning frameworks. These approaches can bridge the gap between model complexity and human understanding, making them essential for stakeholder trust. Moreover, studies should assess how interpretability impacts decision-making in real-time operational settings and whether it can improve collaboration between human operators and AI systems [39,95,96].

5.6.2. Real-Time Accident Prediction

Real-time accident prediction represents one of the most transformative applications of ML in aviation. Research should focus on building systems that leverage data from sensors, flight data recorders, and ACARS in real time to predict potential incidents. This entails using online learning algorithms, adaptive models, and anomaly detection techniques capable of updating predictions as new data arrives. The integration of predictive models with cockpit alert systems or air traffic control workflows could revolutionize preemptive safety strategies. Case studies and pilot implementations in commercial or general aviation could provide the empirical foundation needed to validate such systems [2].

5.6.3. Hybrid Models

Hybrid models, which combine ML techniques with traditional expert systems or physics-based simulations, offer promising avenues for improving predictive performance and interpretability. These models benefit from the data-driven strengths of ML and the domain expertise embedded in rule-based systems. Future research can examine hybrid ensemble approaches that incorporate both quantitative and qualitative safety indicators. Applications could include turbulence forecasting, component failure prediction, and root cause analysis using a blend of probabilistic reasoning and ML inference.

5.6.4. Handling Unbalanced Data Sets

Given the inherent imbalance in aviation accident datasets, where the number of accidents is relatively small compared to non-accident instances, innovative research should focus on methods that effectively address this challenge. Developing techniques that can navigate imbalanced data will enhance the models' capability to identify rare but pivotal

events. One such technique that has shown promise is the use of cost-sensitive learning [97]. This approach involves assigning different costs to misclassification errors based on the class distribution, ensuring that the model prioritizes correctly classifying the minority class (accidents) even at the expense of higher error rates in the majority class (non-accidents). By incorporating the relative importance of each class into the learning process, cost-sensitive learning can significantly improve the model's ability to detect rare events while maintaining overall classification accuracy.

5.6.5. Privacy and Data Security

As the handling of sensitive aviation accident data requires a delicate balance between research needs and privacy protection, future studies should delve into privacy-preserving ML techniques and secure data-sharing mechanisms. Striking this balance will ensure ethical data use and enhance collaboration while safeguarding individual privacy. An example of a privacy-preserving ML technique is differential privacy. Differential privacy aims to provide guarantees that the output of a computation does not reveal sensitive information about any individual data point in the dataset [98]. By adding noise to the data or query responses in a carefully controlled manner, differential privacy enables the analysis of sensitive datasets while protecting the privacy of individuals. Additionally, secure multi-party computation (SMPC) protocols allow multiple parties to jointly compute a function over their inputs while keeping those inputs private [99]. These techniques, among others, offer promising avenues for ensuring privacy and data security in aviation accident data analysis.

5.6.6. Human–Machine Interfaces for Safety Professionals

The effectiveness of ML tools in aviation safety depends on their usability and interpretability by human operators. Future work should focus on designing intuitive human–machine interfaces (HMIs) that support decision-making under pressure. Research could investigate how pilots, air traffic controllers, and accident investigators interact with ML outputs. Incorporating insights from cognitive psychology, ergonomics, and visual analytics can result in systems that enhance human performance rather than overwhelm users with complex model outputs. Usability testing and feedback loops from practitioners should inform them about iterative design processes.

5.6.7. Regulatory Implications

The growing use of ML in aviation safety necessitates clear regulatory frameworks to ensure ethical, legal, and safe deployment. Future studies should explore how regulators such as the FAA, EASA, and ICAO can incorporate AI/ML standards into their certification and compliance procedures. This includes defining accountability for ML-driven decisions, ensuring auditability of algorithms, and managing algorithmic drift. Research should also propose international standards for AI deployment in safety-critical aviation environments, drawing from successful regulatory models in healthcare and autonomous vehicles.

6. Conclusions

This SLR has explored the emerging potential of ML applications in supporting aviation safety, particularly in post-accident analysis. While definitive impacts remain subject to further empirical validation, the integration of ML techniques signals a promising shift toward more proactive and data-informed safety management practices. By synthesizing findings from studies published over the past 25 years, this review has illuminated the growing role of ML across diverse aviation contexts.

The analysis of ML algorithms, data sources, and applications demonstrates their adaptability and growing utility in addressing complex challenges in aviation safety. From

classification and anomaly detection to predictive modeling and incident recognition, ML tools are enabling new capabilities for accident understanding and risk mitigation. These techniques have shown particular relevance across seven major safety domains, as discussed in Section 5.

Importantly, this review highlights the role of ethical considerations, data transparency, and stakeholder trust as foundational pillars in the adoption of ML for safety-critical environments. As these technologies advance, emphasis must remain on fairness, interpretability, and responsible deployment.

Looking forward, the review identifies seven specific avenues for future research that could accelerate innovation and improve the robustness of ML-driven safety analysis (see Section 5.6). Progress in these areas can strengthen the aviation industry's safety infrastructure while offering transferable lessons to similarly sensitive domains such as healthcare, autonomous systems, and industrial operations.

In an era of increasingly complex flight operations and data abundance, ML offers aviation stakeholders a valuable toolkit for enhancing safety, not as a replacement for human expertise but as a complement that unlocks insights previously hidden within aviation data. It is hoped that this review provides a foundation for ongoing collaboration, ethical innovation, and continued improvement in aviation safety practices.

Author Contributions: A.N.: conceptualization, methodology, software, data curation, validation, writing—original draft preparation, formal analysis; U.T.: writing—review and editing; and G.W.: data collection, supervision, final draft. All authors have read and agreed to the published version of the manuscript.

Funding: This research received funding from the Tuition Fee Scholarship at UNSW.

Data Availability Statement: This study is a systematic review based on data extracted from previously published literature. All sources used in the review are publicly available and cited appropriately within the manuscript.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

Abbreviation	Full Form
ACARS	Aircraft Communications Addressing and Reporting System
ADTs	Alternating Decision Trees
AI	Artificial Intelligence
ASN	Aviation Safety Network
ASR	Automatic Speech Recognition
ASRS	Aviation Safety Reporting System
BMR	Bayesian Multi-label Regression
BN	Bayesian Network
CNN	Convolutional Neural Network
DBN	Deep Belief Network
DL	Deep Learning
DNN	Deep Neural Network
DT	Decision Tree
FAA	Federal Aviation Administration
GB	Gradient Boosting
ICAO	International Civil Aviation Organization
KNN	K-Nearest Neighbors
LDA	Latent Dirichlet Allocation
LR	Logistic Regression

LSA	Latent Semantic Analysis
LSTM	Long Short-Term Memory
ML	Machine Learning
NB	Naive Bayes
NN	Neural Network
NLP	Natural Language Processing
NTSB	National Transportation Safety Board
PCA	Principal Component Analysis
RF	Random Forest
RNN	Recurrent Neural Network
RST	Rough Set Theory
SBA	State-Based Approach
SVDA	Support Vector Discriminant Analysis
SVM	Support Vector Machine

Appendix A

Table A1. General description of the 87 systematized articles.

ID	Reference	Title	Algorithms Used	Number of Accidents (Period)	ML Tasks	Data Source
1.	[33]	Setting up new standards in the aviation industry with the help of artificial intelligence—machine learning application	SVM, RF, NB	3000 (20 years)	Classification, Anomaly Detection	ASRS
2.	[34]	Prediction of warning level in aircraft accidents using data mining techniques	DT, KNN, SVM, NN and NB, PCA	- (42 years)	Classification, Data reduction	Airline databases
3.	[64]	Prediction of Warning Level in Aircraft Accidents Using Classification Techniques: An Empirical Study	DT, KNN, SVM, NN, NB	500 (62 years)	Classification	FAA
4.	[35]	Large-scale data analysis on aviation accident database using different data mining techniques	DT, NB, SVM, KNN, NN		Clustering	
5.	[100]	Feature selection techniques for prediction of warning level in aircraft accidents	PCA, DT	500 (62 years)	Classification	FAA
6.	[101]	Enabling the Discovery of Recurring Anomalies in Aerospace Problem Reports using High-Dimensional Clustering Techniques	SVM, NB, LDA, LR, ADT	62 (-)	Clustering	ASRS
7.	[46]	Text mining of accident reports using semi-supervised keyword extraction and topic modeling	LDA	37,678 (9 years)	Data Reduction	ASRS, PHMSA

Table A1. *Cont.*

ID	Reference	Title	Algorithms Used	Number of Accidents (Period)	ML Tasks	Data Source
8.	[102]	An Ensemble Machine and Deep Learning Model for Risk Prediction in Aviation Systems	DL, SVM, NB	6 (12 years)	Classification	ASRS
9.	[103]	Analyzing Aviation Safety Reports: From Topic Modeling to Scalable Multi-Label Classification.	LDA, BMR	66,309 (-)	Classification	ASRS
10.	[104]	Knowledge Graph–Deep Learning: A Case Study in Question Answering in the Aviation Safety Domain	DT	4000 (53 years)	NLP	NTSB
11.	[105]	Deep learning for extracting word-level meaning from safety report narratives	DT	186,000 (39 years)	NLP	ASRS
12.	[106]	A state-based approach to modeling general aviation accidents	-	6180 (33 years)	Classification	NTSB
13.	[107]	Natural Language Processing of Aviation Safety Reports to Identify Inefficient Operational Patterns	DT	4195 (23 years)	Anomaly Detection	ASRS
14.	[108]	Predicting General Aviation Accidents Using Machine Learning Algorithms	DT, NN, RF, LR, GB	27,786 (20 years)	Classification	NTSB
15.	[86]	Augmenting topic findings in the NASA aviation safety reporting system using topic modeling	LDA	100 (3 years)	NLP	ASRS
16.	[109]	Incorporation of Pilot Factors into Risk Analysis of Civil Aviation Accidents from 2008 to 2020: A Data-Driven Bayesian Network Approach.	BN	163 (12 years)	Classification	NTSB
17.	[110]	Multi-concept document classification using a perceptron-like algorithm.	Perceptron-like algorithm	-	Classification	ASRS
18.	[66]	Computational Solution to Prevent Aeronautic Accidents Cause by Wake Turbulence Using Machine Learning	KNN, NN, DT, NB	(1 year)	Anomaly Detection	EUROCONTROL
19.	[68]	Deep learning-based approach for civil aircraft hazard identification and prediction.	RNN, KNN, LSTM, NN	1244 (2 years)	Anomaly Detection	ACARS

Table A1. *Cont.*

ID	Reference	Title	Algorithms Used	Number of Accidents (Period)	ML Tasks	Data Source
20.	[111]	Machine learning for helicopter accident analysis using supervised classification: Inference, prediction, and implications	KNN, DT, ADT, RF, NB, DNN	13,055 (11 years)	Classification	NTSB
21.	[83]	Advanced text mining algorithms for aerospace anomaly identification.	NN, NB, ADT, LDA, LR, SVM	9910 (24 years)	Anomaly Detection	ASRS
22.	[37]	The effect of COVID-19 on self-reported safety incidents in aviation: An examination of the heterogeneous effects using causal machine learning	RF	7246 (2 years)	Anomaly Detection	ASRS
23.	[112]	Hybrid safety analysis method based on SVM and RST: An application to carrier landing of aircraft	SVM, RST	635	Classification	NADC
24.	[113]	Textual indicator extraction from aviation accident reports	SVM, DNN	61,687 (35 years)	NLP	NTSB
25.	[90]	Civil aviation safety evaluation based on deep belief network and principal component analysis.	DBN, PCA	0 (5 years)	Classification, Data Reduction	
26.	[114]	A hybrid data-driven approach to analyzing aviation incident reports	DNN, SVM	64,573 (11 years)	NLP, Classification	ASRS
27.	[9]	Ensemble machine learning models for aviation incident risk prediction	DNN, SVM	64,573 (11 years)	Classification	ASRS
28.	[115]	Bayesian network modeling of accident investigation reports for aviation safety assessment.	BN	2243 (24 years)		NTSB
29.	[116]	Classification of aviation safety reports using machine learning.	RF, SVM, KNN, NN, NB	73,000 (5 years)	Classification	ICAO
30.	[117]	The analysis of fatal aviation accidents more than 100 dead passengers: an application of machine learning	SVM, NN, PCA, LR, DT	220	Classification, Data Reduction	ICAO
31.	[118]	Semi-supervised learning with semantic knowledge extraction for improved speech recognition in air traffic control	DNN	6004	NLP	ASR

Table A1. *Cont.*

ID	Reference	Title	Algorithms Used	Number of Accidents (Period)	ML Tasks	Data Source
32.	[119]	Predicting airline crash due to bird strike using machine learning	DT, KNN, NB	-	Classification	NTSB
33.	[120]	Deep learning-based Time Series Forecasting of Go-around Incidents in the National Airspace System	LSTM	3835 (24 years)	Classification	ASRS
34.	[67]	An Innovative Approach to Modeling Aviation Safety Incidents	CNN, LSTM, SVM, RF, NB, LR	158,070	Classification	ASRS
35.	[76]	Flight crash investigation using data mining techniques	K-Mean	5268 (101 years)	Clustering	-
36.	[121]	Applying Distilled BERT for Question Answering on ASRS Reports	BERT	1,625,738 (43 years)	NLP	ASRS
37.	[122]	Application of Machine Learning to mapping Primary Causal Factors in self-reported safety narratives	LSA	7484 (4 years)	NLP	ASRS
38.	[123]	Visual representation of safety narratives	LSA	4497 (2 years)	NLP	ASRS
39.	[124]	Temporal topic modeling applied to aviation safety reports: A subject matter expert review	LDA	64,776 (14 years)	Clustering	ASRS
40.	[38]	Data mining application on aviation accident data for predicting topmost causes of accidents	NN, SVM, KNN, DT, NB	1610 (95 years)	Classification	FAA
41.	[82]	Application of machine learning techniques for incident-accident classification problem in aviation safety management	SVM, NB, DT, RF	84,262 (57 years)	Classification	NTSB
42.	[70]	Learning Methods and Predictive Modeling to Identify Failure by Human Factors in the Aviation Industry.	NN, RF	1105 (10 years)	Classification	ASN
43.	[75]	Application of structural topic modeling to aviation safety data.	LDA	386 (8 years)	NLP	ASRS, NTSB
44.	[87]	Natural language processing-based method for clustering and analysis of aviation safety narratives.	PCA, K-Means	13,336 (10 years)	Clustering	ASRS

Table A1. *Cont.*

ID	Reference	Title	Algorithms Used	Number of Accidents (Period)	ML Tasks	Data Source
45.	[125]	A textual analysis of dangerous goods incidents on aircraft.	SVDA	383 (10 years)	NLP	ASRS
46.	[65]	Prediction of injuries and fatalities in aviation accidents through machine learning	DT, KNN, SVM, NN	31,974 (27 years)	Classification	FAA
47.	[126]	Prediction of aviation accidents using logistic regression model.	LR	7415	Classification	ASN
48.	[127]	Airline Safety Data: How Predictable Are Accidents and Fatalities?	NN	10 (61 years)	Classification	FAA
49.	[88]	Aircraft safety analysis using clustering algorithms	K-Means	1500 (25 years)	Clustering	-
50.	[128]	Analysis of General Aviation fixed-wing aircraft accidents involving inflight loss of control using a state-based approach	SBA	5726 (18 years)	Clustering	NTSB
51.	[129]	Apriori algorithm for association rules mining in aircraft runway excursions.	AR	434 (10 years)	Classification	ASN
52.	[73]	Using correlation-based subspace clustering for multi-label text data classification	KNN	15, 000	Classification	ASRS, Reuters, 20 News-groups
53.	[130]	A model fusion strategy for identifying aircraft risk using CNN and Att-BiLSTM	CNN, BLSTM, DNN	32 (10 years)	Classification	ASRS
54.	[41]	Using Neural Networks to predict HFACS unsafe acts from the pre-conditions of unsafe acts	NN	523 (24 years)	Classification	ROC
55.	[40]	A data-mining approach to identification of risk factors in safety management systems	NB	168,227 (24 years)	Classification	ASRS
56.	[131]	Causes and risk factors for fatal accidents in non-commercial twin engine piston general aviation aircraft.	LR	376 (10 years)	Classification	NTSB
57.	[132]	Examination of Aircraft Accidents That Occurred in the Last 20 Years in the World.	KNN, NB, DT, LR, GBM	588 (20 years)	Classification	-

Table A1. *Cont.*

ID	Reference	Title	Algorithms Used	Number of Accidents (Period)	ML Tasks	Data Source
58.	[85]	Classification of aviation accidents using data mining algorithms	DT, NB, SMO	588 (20 years)	Classification	-
59.	[93]	Predictive safety analytics: Inferring aviation accident shaping factors and causation	BN	315 (23 years)	Classification	NTSB
60.	[133]	Analysis of Helicopter Accidents and Certification Categories Using Machine Learning.	RF, DT	1576 (10 years)	Classification	NTSB
61.	[84]	Application of machine learning for aviation safety risk metric	GBM, RNN, SVM	10,634 (20 years)	Classification	NTSB, MOR, ASIAS
62.	[134]	Descriptive and predictive analyses of data representing aviation accidents.	DT, KNN, RF	25,000 (4 years)	Classification	FAA
63.	[135]	Flight Accident Modeling and Predicting Based on Least Squares Support Vector Machine	SVM	40 years	Classification	NTSB
64.	[55]	A novel method of aviation safety prediction based on Lstm-Rbf model	LSTM, RBF	1 year	Classification	
65.	[136]	On the chaos analysis and prediction of aircraft accidents based on multi-timescales	SVM	59,511 (55 years)	Classification	NTSB
66.	[137]	Critical parameter identification for safety events in commercial aviation using machine learning.	NB, RF, DT, KNN	70 (6 years)	Classification	FOQA
67.	[138]	PIA Accidents Analysis Using Naïve Bayes Classifier	NB	22 (6 years)	Classification	-
68.	[139]	Failing &! Falling (F&! F): Learning to Classify Accidents and Incidents in Aircraft Data	DT, NN	137,236	Classification	FAA
69.	[140]	Subjectivity classification and analysis of the ASRS corpus	SVM, ADT	140,599 (2 years)	Classification	ASRS
70.	[141]	Using random forests to diagnose aviation turbulence	RF, KNN, LR	778 (2 years)	Classification	NTSB
71.	[142]	Understanding general aviation accidents in terms of safety systems.	-	2303 (10 years)	Classification	NTSB

Table A1. *Cont.*

ID	Reference	Title	Algorithms Used	Number of Accidents (Period)	ML Tasks	Data Source
72.	[143]	Using Machine Learning Models to Study Human Error-Related Factors in Aviation Accidents and Incidents	NB, RF, LR, SVM, NN	90,000 (47 years)	Classification	NTSB
73.	[144]	Using structural topic modeling to identify latent topics and trends in aviation incident reports.	LDA	25,706 (5 years)	NLP	ASRS
74.	[145]	Automated aviation occurrences categorization	NN	12,500 (6 years)	Classification	ASRS
75.	[89]	Understanding large text corpora via sparse machine learning.	PCA, LDA, LASSO	20,000 (4 years)	NLP	ASRS
76.	[146]	Sparse machine learning methods for understanding large text corpora.	PCA, LDA, LASSO	20,000 (4 years)	NLP	ASRS
77.	[50]	Text Mining Classification and Prediction of Aviation Accidents Based on TF-IDF-SVR Method.	LDA, SVM, NB, RF, LR	20,000 (59 years)	NLP, Classification	NTSB
78.	[147]	Cause identification from aviation safety incident reports via weakly supervised semantic lexicon construction	SVM	140,599 (9 years)	Classification	ASRS
79.	[148]	Analysis of Aviation Accidents Data.	RF, NB, KNN, DT, GBT	19,455 (14 years)	Classification	NTSB
80.	[149]	Document classification using nonnegative matrix factorization and underapproximation.	NMF	21,519 (1 year)	Clustering	ASRS
81.	[150]	Towards online prediction of safety-critical landing metrics in aviation using supervised machine learning	LSTM, NN, RF	623	Regression	FOQA
82.	[151]	Identifying Incident Causal Factors to Improve Aviation Transportation Safety: Proposing a Deep Learning Approach	LSTM	200,000 (32 years)	Classification	ASRS
83.	[69]	Recent Experiences with Data Mining in Aviation Safety	DT	1256 (9 years)	Classification	ASRS
84.	[152]	Safer Approaches and Landings: A Multivariate Analysis of Critical Factors	DT, LR	287 (16 years)	Classification	NTSB, ASRS

Table A1. *Cont.*

ID	Reference	Title	Algorithms Used	Number of Accidents (Period)	ML Tasks	Data Source
85.	[74]	Multi-label asrs dataset classification using semi-supervised subspace clustering	KNN	10,000	Clustering	ASRS, Reuters, 20 News-groups
86.	[153]	Sequential Classification of Aviation Safety Occurrences with Natural Language Processing.	LSTM, BLSTM, GRU, RNN	27,000 (15 year)	Classification	NTSB
87.	[57]	Classification of Survivor/Non-Survivor Passengers in Fatal Aviation Accidents: A Machine Learning Approach	NN, DT, PCA	100 (1 year)	Classification, Data Reduction	BAAA

Note: DT means decision trees; BN means Bayesian networks; NN means neural networks; SVM means support vectors machine; DL means deep learning; KNN means k-nearest neighbor; NB means naïve Bayes; PCA means principle component analysis; STS means Semantic Text Similarity; LR means logistic regression; LSTM means Long Short-Term Memory; Bidirectional Long Short-Term Memory (BLSTM); Gated Recurrent Unit (GRU); Recurrent Neural Network (RNN); Non-Negative Matrix Factorization (NMF); Latent Dirichlet Allocation (LDA); Latent Semantic Analysis (LSA); Convolutional Neural Networks (CNNs); Deep Neural Networks (DNNs); Random Forest (RF); Rough Set Theory (RST); Adaboost Decision Tree (ADT).

References

1. Dileep, M.R.; Kurien, A. *Air Transport and Tourism: Interrelationship, Operations and Strategies*; Routledge: London, UK, 2021.
2. Netjasov, F.; Janic, M. A review of research on risk and safety modelling in civil aviation. *J. Air Transp. Manag.* **2008**, *14*, 213–220. [CrossRef]
3. Aderibigbe, A. Root cause analysis of a jet fuel tanker accident. *Int. J. Appl. Eng. Res.* **2017**, *12*, 14974–14983.
4. Tanguy, L.; Tulechki, N.; Urieli, A.; Hermann, E.; Raynal, C. Natural language processing for aviation safety reports: From classification to interactive analysis. *Comput. Ind.* **2016**, *78*, 80–95. [CrossRef]
5. Wiener, E.L.; Nagel, D.C. *Human Factors in Aviation*; Gulf Professional Publishing: Woburn, MA, USA, 1988.
6. Janakiraman, V.M.; Nielsen, D. Anomaly detection in aviation data using extreme learning machines. In Proceedings of the IEEE 2016 International Joint Conference on Neural Networks (IJCNN), Vancouver, BC, Canada, 24–29 July 2016; pp. 1993–2000.
7. Vasigh, B.; Fleming, K.; Tacker, T. *Introduction to Air Transport Economics: From Theory to Applications*; Routledge: London, UK, 2018.
8. Wild, G.; Baxter, G.; Srisaeng, P.; Richardson, S. Machine learning for air transport planning and management. In Proceedings of the AIAA Aviation 2022 Forum, Chicago, IL, USA, 27 June–1 July 2022; p. 3706.
9. Zhang, X.; Mahadevan, S. Ensemble machine learning models for aviation incident risk prediction. *Decis. Support Syst.* **2019**, *116*, 48–63. [CrossRef]
10. Ladenbauer, S. *European Union Policymaking in the Field of Air Traffic Management: The Endeavor to Implement Functional Airspace Blocks in Light of Fragmented National Interests: A Case Study on the Functional Airspace Block Europe Central (FABEC)*; University of Zurich: Zurich, Switzerland, 2012.
11. Verma, S.; Kumar, P. A Comparative Overview of Accident Forecasting Approaches for Aviation Safety. *J. Phys. Conf. Ser.* **2021**, *1767*, 012015.
12. Weber, L. *International Civil Aviation Organization*; ICAO: Montreal, QC, Canada, 2023.
13. Kasula, B.Y. Machine Learning Unleashed: Innovations, Applications, and Impact Across Industries. *Int. Trans. Artif. Intell.* **2017**, *1*, 1–7.
14. Ongsulee, P. Artificial intelligence, machine learning and deep learning. In Proceedings of the IEEE 2017 15th International Conference on ICT and Knowledge Engineering (ICT&KE), Bangkok, Thailand, 22–24 November 2017; pp. 1–6.
15. Mohammadpour, A.; Karan, E.; Asadi, S. Artificial intelligence techniques to support design and construction. In Proceedings of the International Symposium on Automation and Robotics in Construction (ISARC), Banff, AB, Canada, 21–24 May 2019; IAARC Publications: Montreal, QC, Canada, 2019; Volume 36, pp. 1282–1289.

16. Cody, T.; Lanus, E.; Doyle, D.D.; Freeman, L. Systematic training and testing for machine learning using combinatorial interaction testing. In Proceedings of the 2022 IEEE International Conference on Software Testing, Verification and Validation Workshops (ICSTW), Valencia, Spain, 4–13 April 2022; pp. 102–109.
17. Lanus, E.; Freeman, L.J.; Kuhn, D.R.; Kacker, R.N. Combinatorial testing metrics for machine learning. In Proceedings of the 2021 IEEE International Conference on Software Testing, Verification and Validation Workshops (ICSTW), Porto de Galinhas, Brazil, 12–16 April 2021; pp. 81–84.
18. Kang, Z.; Catal, C.; Tekinerdogan, B. Machine learning applications in production lines: A systematic literature review. *Comput. Ind. Eng.* **2020**, *149*, 106773. [CrossRef]
19. Hastie, T.; Tibshirani, R.; Friedman, J. Overview of supervised learning. In *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*; Springer: New York, NY, USA, 2009; pp. 9–41.
20. Ibrahim, K.; Sorayya, M.; Aziida, N.; Sazzli, S.K. Preliminary study on application of machine learning method in predicting survival versus non-survival after myocardial infarction in Malaysian population. *Int. J. Cardiol.* **2018**, *273*, 8. [CrossRef]
21. Hastie, T.; Tibshirani, R.; Friedman, J. Unsupervised learning. In *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*; Springer: New York, NY, USA, 2009; pp. 485–585.
22. Bishop, C.M.; Nasrabadi, N.M. *Pattern Recognition and Machine Learning*; Springer: Berlin/Heidelberg, Germany, 2006; Volume 4.
23. Chapelle, O.; Scholkopf, B.; Zien, A. Semi-supervised learning (chapelle, o. et al., eds.; 2006) [book reviews]. *IEEE Trans. Neural Netw.* **2009**, *20*, 542. [CrossRef]
24. Wiering, M.A.; Van Otterlo, M. Reinforcement learning. *Adapt. Learn. Optim.* **2012**, *12*, 729.
25. Sutton, R.S.; Barto, A.G. *Reinforcement Learning: An Introduction*; MIT Press: Cambridge, UK, 2018.
26. James, G.; Witten, D.; Hastie, T.; Tibshirani, R. *An Introduction to Statistical Learning*; Springer: Berlin/Heidelberg, Germany, 2013; Volume 112.
27. Alpaydin, E. *Introduction to Machine Learning*; MIT Press: Cambridge, UK, 2020.
28. Chandola, V.; Banerjee, A.; Kumar, V. Anomaly detection: A survey. *ACM Comput. Surv.* **2009**, *41*, 1–58. [CrossRef]
29. Collobert, R.; Weston, J.; Bottou, L.; Karlen, M.; Kavukcuoglu, K.; Kuksa, P. Natural language processing (almost) from scratch. *J. Mach. Learn. Res.* **2011**, *12*, 2493–2537.
30. Meurers, D. Natural language processing and language learning. *Encycl. Appl. Linguist.* **2012**, *10*, 4193–4205. [CrossRef]
31. Bayat, B.; Bermejo-Alonso, J.; Carbonera, J.; Facchinetti, T.; Fiorini, S.; Goncalves, P.; Jorge, V.A.; Habib, M.; Khamis, A.; Melo, K.; et al. Requirements for building an ontology for autonomous robots. *Ind. Robot. Int. J.* **2016**, *43*, 469–480. [CrossRef]
32. Mahmud, F. *Human-Intelligence and Machine-Intelligence Decision Governance Formal Ontology*; Old Dominion University: Norfolk, VA, USA, 2018.
33. Andrei, A.; Balasa, R.; Semenescu, A. Setting up new standards in aviation industry with the help of artificial intelligent–machine learning application. In *Journal of Physics: Conference Series*; IOP Publishing: Bristol, UK, 2022; p. 012014.
34. Christopher, A.A.; alias Balamurugan, S.A. Prediction of warning level in aircraft accidents using data mining techniques. *Aeronaut. J.* **2014**, *118*, 935–952. [CrossRef]
35. Christopher, A.A.; Vivekanandam, V.S.; Anderson, A.A.; Markkandeyan, S.; Sivakumar, V. Large-scale data analysis on aviation accident database using different data mining techniques. *Aeronaut. J.* **2016**, *120*, 1849–1866. [CrossRef]
36. Malek, S.; Hui, C.; Aziida, N.; Cheen, S.; Toh, S.; Milow, P. Ecosystem monitoring through predictive modeling. *Encycl. Bioinform. Comput. Biol.* **2019**, *3*, 1–8. [CrossRef]
37. Choi, Y.; Gibson, J.R. The effect of COVID-19 on self-reported safety incidents in aviation: An examination of the heterogeneous effects using causal machine learning. *J. Saf. Res.* **2023**, *84*, 393–403. [CrossRef] [PubMed]
38. Koteeswaran, S.; Malarvizhi, N.; Kannan, E.; Sasikala, S.; Geetha, S. Data mining application on aviation accident data for predicting topmost causes for accidents. *Clust. Comput.* **2019**, *22*, 11379–11399. [CrossRef]
39. Minh, D.; Wang, H.X.; Li, Y.F.; Nguyen, T.N. Explainable artificial intelligence: A comprehensive review. *Artif. Intell. Rev.* **2022**, *55*, 3503–3568. [CrossRef]
40. Shi, D.; Guan, J.; Zurada, J.; Manikas, A. A data-mining approach to identification of risk factors in safety management systems. *J. Manag. Inf. Syst.* **2017**, *34*, 1054–1081. [CrossRef]
41. Harris, D.; Li, W.-C. Using Neural Networks to predict HFACS unsafe acts from the pre-conditions of unsafe acts. *Ergonomics* **2019**, *62*, 181–191. [CrossRef]
42. Biau, G.; Scornet, E. A random forest guided tour. *Test* **2016**, *25*, 197–227. [CrossRef]
43. Aziida, N.; Malek, S.; Aziz, F.; Ibrahim, K.S.; Kasim, S. Predicting 30-day mortality after an acute coronary syndrome (ACS) using machine learning methods for feature selection, classification and visualisation. *Sains Malays.* **2021**, *50*, 753–768. [CrossRef]
44. Nanyonga, A.; Joiner, K.; Turhan, U.; Wild, G. Does the Choice of Topic Modeling Technique Impact the Interpretation of Aviation Incident Reports? A Methodological Assessment. *Technologies* **2025**, *13*, 209. [CrossRef]
45. Nanyonga, A.; Wild, G. Analyzing Aviation Safety Narratives with LDA, NMF and PLSA: A Case Study Using Socrata Datasets. *arXiv* **2025**, arXiv:2501.01690. [CrossRef]

46. Ahadh, A.; Binish, G.V.; Srinivasan, R. Text mining of accident reports using semi-supervised keyword extraction and topic modeling. *Process. Saf. Environ. Prot.* **2021**, *155*, 455–465. [CrossRef]
47. Nanyonga, A.; Wasswa, H.; Wild, G. Topic Modeling Analysis of Aviation Accident Reports: A Comparative Study between LDA and NMF Models. In Proceedings of the IEEE 2023 3rd International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON), Bangalore, India, 29–31 December 2023; pp. 1–2.
48. Nanyonga, A.; Wasswa, H.; Turhan, U.; Joiner, K.; Wild, G. Comparative Analysis of Topic Modeling Techniques on ATSB Text Narratives Using Natural Language Processing. In Proceedings of the IEEE 2024 3rd International Conference for Innovation in Technology (INOCON), Bangalore, India, 1–3 March 2024; pp. 1–7.
49. Nanyonga, A.; Joiner, K.; Turhan, U.; Wild, G. Applications of natural language processing in aviation safety: A review and qualitative analysis. In Proceedings of the AIAA SCITECH 2025 Forum, Orlando, FL, USA, 6–10 January 2025; p. 2153.
50. Zhao, L.; Zhang, L.; Wang, J. Text Mining Classification and Prediction of Aviation Accidents Based on TF-IDF-SVR Method. In Proceedings of the IEEE 2022 4th International Conference on Frontiers Technology of Information and Computer (ICFTIC), Qingdao, China, 2–4 December 2022; pp. 322–327.
51. Nanyonga, A.; Wasswa, H.; Wild, G. Comparative Study of Deep Learning Architectures for Textual Damage Level Classification. In Proceedings of the IEEE 2024 11th International Conference on Signal Processing and Integrated Networks (SPIN), Noida, India, 21–22 March 2024; pp. 421–426.
52. Nanyonga, A.; Wild, G. Classification of Operational Records in Aviation Using Deep Learning Approaches. In Proceedings of the 2025 International Conference on Pervasive Computational Technologies (ICPCT), Greater Noida, India, 8–9 February 2025.
53. Nanyonga, A.; Wasswa, H.; Wild, G. Aviation Safety Enhancement via NLP & Deep Learning: Classifying Flight Phases in ATSB Safety Reports. In Proceedings of the IEEE 2023 Global Conference on Information Technologies and Communications (GCITC), Bangalore, India, 1–3 December 2023; pp. 1–5.
54. Van Houdt, G.; Mosquera, C.; Nápoles, G. A review on the long short-term memory model. *Artif. Intell. Rev.* **2020**, *53*, 5929–5955. [CrossRef]
55. Zeng, H.; Ren, B.; Zhang, H.; Wu, J.; Liu, C.; Ren, H. A novel method of aviation safety prediction based on Lstm-Rbf model. In Proceedings of the 12th International Conference on Quality, Reliability, Risk, Maintenance, and Safety Engineering (QR2MSE 2022), Emeishan, China, 27–30 July 2022; pp. 1592–1598.
56. Greenacre, M.; Groenen, P.J.; Hastie, T.; d'Enza, A.I.; Markos, A.; Tuzhilina, E. Principal component analysis. *Nat. Rev. Methods Primers* **2022**, *2*, 100. [CrossRef]
57. İnan, D.; Tolga, T. Classification of Survivor/Non-Survivor Passengers in Fatal Aviation Accidents: A Machine Learning Approach. *Int. J. Aviat. Aeronaut. Aerosp.* **2022**, *9*, 8. [CrossRef]
58. Kitchenham, B.; Brereton, O.P.; Budgen, D.; Turner, M.; Bailey, J.; Linkman, S. Systematic literature reviews in software engineering–a systematic literature review. *Inf. Softw. Technol.* **2009**, *51*, 7–15. [CrossRef]
59. Van Dinter, R.; Tekinerdogan, B.; Catal, C. Automation of systematic literature reviews: A systematic literature review. *Inf. Softw. Technol.* **2021**, *136*, 106589. [CrossRef]
60. Chadegani, A.A.; Salehi, H.; Yunus, M.M.; Farhadi, H.; Fooladi, M.; Farhadi, M.; Ebrahim, N.A. A comparison between two main academic literature collections: Web of Science and Scopus databases. *arXiv* **2013**, arXiv:1305.0377. [CrossRef]
61. Burnham, J.F. Scopus database: A review. *Biomed. Digit. Libr.* **2006**, *3*, 1–8. [CrossRef]
62. Khallaf, R.; Khallaf, M. Classification and analysis of deep learning applications in construction: A systematic literature review. *Autom. Constr.* **2021**, *129*, 103760. [CrossRef]
63. Slikboer, R.; Muir, S.D.; Silva, S.S.M.; Meyer, D. A systematic review of statistical models and outcomes of predicting fatal and serious injury crashes from driver crash and offense history data. *Syst. Rev.* **2020**, *9*, 1–15. [CrossRef] [PubMed]
64. Arockia Christopher, A.; Appavu alias Balamurugan, S. Prediction of warning level in aircraft accidents using classification techniques: An empirical study. In *Intelligent Computing, Networking, and Informatics, Proceedings of the International Conference on Advanced Computing, Networking, and Informatics, Chhattisgarh, India, 12–14 June 2013*; Springer: New York, NY, USA, 2014; pp. 1217–1223.
65. Burnett, R.A.; Si, D. Prediction of injuries and fatalities in aviation accidents through machine learning. In Proceedings of the International Conference on Compute and Data Analysis, Lakeland, FL, USA, 19–23 May 2017; pp. 60–68.
66. Leite, D.V.; Weigang, L.; Barreto, A.B.; Crespo, A.M. Computational Solution to Prevent Aeronautics Accidents Cause by Wake Turbulence Using Machine Learning. In Proceedings of the IECON 2020 The 46th Annual Conference of the IEEE Industrial Electronics Society, Singapore, 18–21 October 2020; pp. 124–129.
67. Shi, D.; Cao, S.; Zurada, J.; Guan, J. An Innovative Approach to Modeling Aviation Safety Incidents. In Proceedings of the 55th Hawaii International Conference on System Sciences, Maui, HI, USA, 4–7 January 2022.
68. Zhou, D.; Zhuang, X.; Zuo, H.; Wang, H.; Yan, H. Deep learning-based approach for civil aircraft hazard identification and prediction. *IEEE Access* **2020**, *8*, 103665–103683. [CrossRef]

69. Harris, E.; Bloedorn, E.; Rothleder, N.; Chaudhuri, S.; Dayal, U. Recent experiences with data mining in aviation safety. In Proceedings of the Special Interest Group on Management of Data, Data Mining and Knowledge Discovery (SIGMOD-DMKD) Workshop, Seattle, WA, USA, 2–4 June1998.
70. Nogueira, R.P.; Melicio, R.; Valério, D.; Santos, L.F. Learning methods and predictive modeling to identify failure by human factors in the aviation industry. *Appl. Sci.* **2023**, *13*, 4069. [CrossRef]
71. Dhruv, A.J.; Patel, R.; Doshi, N. Python: The most advanced programming language for computer science applications. In Proceedings of the International Conference on Culture Heritage, Education, Sustainable Tourism, and Innovation Technologies (CESIT 2020), Online, 17 September 2020; pp. 292–299.
72. Hall, M.; Frank, E.; Holmes, G.; Pfahringer, B.; Reutemann, P.; Witten, I.H. The WEKA data mining software: An update. *ACM SIGKDD Explor. Newsl.* **2009**, *11*, 10–18. [CrossRef]
73. Ahmed, M.S.; Khan, L.; Rajeswari, M. Using correlation based subspace clustering for multi-label text data classification. In Proceedings of the 2010 22nd IEEE International Conference on Tools with Artificial Intelligence, Arras, France, 27–29 October 2010; pp. 296–303.
74. Ahmed, M.S.; Khan, L.; Oza, N.C.; Rajeswari, M. Multi-label ASRS Dataset Classification Using Semi Supervised Subspace Clustering. In Proceedings of the CIDU, Mountain View, CA, USA, 5–6 October 2010; pp. 285–299.
75. Rose, R.L.; Puranik, T.G.; Mavris, D.N.; Rao, A.H. Application of structural topic modeling to aviation safety data. *Reliab. Eng. Syst. Saf.* **2022**, *224*, 108522. [CrossRef]
76. Sharma, S.; Sabitha, A.S. Flight crash investigation using data mining techniques. In Proceedings of the IEEE 2016 1st India International Conference on Information Processing (IICIP), Delhi, India, 12–14 August 2016; pp. 1–7.
77. Paul, S.; Purkaystha, B.S.; Das, P. Nlp Tools Used in Civil Aviation: A Survey. *Int. J. Adv. Res. Comput. Sci.* **2018**, *9*, 109–114. [CrossRef]
78. Pimm, C.; Raynal, C.; Tulechki, N.; Hermann, E.; Caudy, G.; Tanguy, L. Natural Language Processing (NLP) tools for the analysis of incident and accident reports. In Proceedings of the International Conference on Human-Computer Interaction in Aerospace (HCI-Aero), Brussels, Belgium, 13 September 2012.
79. Ackley, J.L.; Puranik, T.G.; Mavris, D. A supervised learning approach for safety event precursor identification in commercial aviation. In Proceedings of the AIAA Aviation 2020 Forum, Virtual, 15–19 June 2020; p. 2880.
80. Nanyonga, A.; Joiner, K.; Turhan, U.; Wild, G. Semantic Topic Modeling of Aviation Safety Reports: A Comparative Analysis Using BERTopic and PLSA. *Aerospace* **2025**, *12*, 551. [CrossRef]
81. Basora, L.; Olive, X.; Dubot, T. Recent advances in anomaly detection methods applied to aviation. *Aerospace* **2019**, *6*, 117. [CrossRef]
82. Rukabu, O. Application of Machine Learning Techniques for Incident-Accident Classification Problem in Aviation Safety Management. Master's Thesis, University of Rwanda, Kigali, Rwanda, 2021.
83. Bluvband, Z.; Porotsky, S. *Advanced Text Mining Algorithms for Aerospace Anomaly Identification*; Taylor & Francis Group: Boca Raton, FL, USA, 2012.
84. Bati, F.; Withington, L. Application of machine learning for aviation safety risk metric. In Proceedings of the 2019 IEEE/AIAA 38th Digital Avionics Systems Conference (DASC), Mission Bay, the Hilton San Diego Resort and Spa, San Diego, CA, USA, 8–12 September 2019; pp. 1–9.
85. Kuşkapan, E.; Sahraei, M.A.; Çodur, M.Y. Classification of aviation accidents using data mining algorithms. *Balk. J. Electr. Comput. Eng.* **2021**, *10*, 10–15. [CrossRef]
86. Paradis, C.; Kazman, R.; Davies, M.; Hooey, B. Augmenting topic finding in the NASA Aviation Safety Reporting System using topic modeling. In Proceedings of the AIAA Scitech 2021 Forum, Virtual, 11–15 and 19–21 January 2021; p. 1981.
87. Rose, R.L.; Puranik, T.G.; Mavris, D.N. Natural language processing based method for clustering and analysis of aviation safety narratives. *Aerospace* **2020**, *7*, 143. [CrossRef]
88. Čokorilo, O.; De Luca, M.; Dell'Acqua, G. Aircraft safety analysis using clustering algorithms. *J. Risk Res.* **2014**, *17*, 1325–1340. [CrossRef]
89. El Ghaoui, L.; Pham, V.; Li, G.C.; Duong, V.A.; Srivastava, A.; Bhaduri, K. Understanding large text corpora via sparse machine learning. *Stat. Anal. Data Mining ASA Data Sci. J.* **2013**, *6*, 221–242. [CrossRef]
90. Ni, X.; Wang, H.; Che, C.; Hong, J.; Sun, Z. Civil aviation safety evaluation based on deep belief network and principal component analysis. *Saf. Sci.* **2019**, *112*, 90–95. [CrossRef]
91. Nanyonga, A.; Wild, G. Impact of Dataset Size & Data Source on Aviation Safety Incident Prediction Models with Natural Language Processing. In Proceedings of the IEEE 2023 Global Conference on Information Technologies and Communications (GCITC), Bangalore, India, 1–3 December 2023; pp. 1–7.
92. Nanyonga, A.; Joiner, K.; Turhan, U.; Wild, G. Natural Language Processing for Aviation Safety: Predicting Injury Levels from Incident Reports in Australia. *Modelling* **2025**, *6*, 40. [CrossRef]

93. Ancel, E.; Shih, A.T.; Jones, S.M.; Reveley, M.S.; Luxhøj, J.T.; Evans, J.K. Predictive safety analytics: Inferring aviation accident shaping factors and causation. *J. Risk Res.* **2015**, *18*, 428–451. [CrossRef]
94. Kilkenny, M.F.; Robinson, K.M. Data quality:"Garbage in–garbage out". *Health Inf. Manag. J.* **2018**, *47*, 103–105. [CrossRef]
95. Nanyonga, A.; Wasswa, H.; Joiner, K.; Turhan, U.; Wild, G. A Multi-Head Attention-Based Transformer Model for Predicting Causes in Aviation Incident. *Modelling* **2025**, *6*, 27. [CrossRef]
96. Nanyonga, A.; Wasswa, H.; Joiner, K.; Turhan, U.; Wild, G. Explainable Supervised Learning Models for Aviation Predictions in Australia. *Aerospace* **2025**, *12*, 223. [CrossRef]
97. Thai-Nghe, N.; Nghi, D.; Schmidt-Thieme, L. Learning optimal threshold on resampling data to deal with class imbalance. In Proceedings of the IEEE RIVF International Conference on Computing and Telecommunication Technologies, Hanoi, Vietnam, 1–4 November 2010; pp. 71–76.
98. Dwork, C. Differential privacy. In Proceedings of the International Colloquium on Automata, Languages, and Programming, Venice, Italy, 10–14 July 2006; pp. 1–12.
99. Yao, A.C. Protocols for secure computations. In Proceedings of the IEEE 23rd Annual Symposium on Foundations of Computer Science (SFCS 1982), Chicago, IL, USA, 3–5 November 1982; pp. 160–164.
100. Christopher, A.A.; alias Balamurugan, S.A. Feature selection techniques for prediction of warning level in aircraft accidents. In Proceedings of the IEEE 2013 International Conference on Advanced Computing and Communication Systems, Coimbatore, India, 19–21 December 2013; pp. 1–6.
101. Srivastava, A.N. Enabling the discovery of recurring anomalies in aerospace problem reports using high-dimensional clustering techniques. In Proceedings of the 2006 IEEE Aerospace Conference, Big Sky, MT, USA, 4–11 March 2006; p. 17.
102. Alkhamisi, A.O.; Mehmood, R. An ensemble machine and deep learning model for risk prediction in aviation systems. In Proceedings of the IEEE 2020 6th Conference on Data Science and Machine Learning Applications (CDMA), Riyadh, Saudi Arabia, 4–5 March 2020; pp. 54–59.
103. Agovic, A.; Shan, H.; Banerjee, A. Analyzing Aviation Safety Reports: From Topic Modeling to Scalable Multi-Label Classification. In Proceedings of the CIDU, Mountain View, CA, USA, 5–6 October 2012; pp. 83–97.
104. Agarwal, A.; Gite, R.; Laddha, S.; Bhattacharyya, P.; Kar, S.; Ekbal, A.; Thind, P.; Zele, R.; Shankar, R. Knowledge graph-deep learning: A case study in question answering in aviation safety domain. *arXiv* **2022**, arXiv:2205.15952.
105. Chanen, A. Deep learning for extracting word-level meaning from safety report narratives. In Proceedings of the IEEE 2016 Integrated Communications Navigation and Surveillance (ICNS), Herndon, VA, USA, 19–21 April 2016; pp. 5D2-1–5D2-15.
106. Rao, A.H.; Marais, K. A state-based approach to modeling general aviation accidents. *Reliab. Eng. Syst. Saf.* **2020**, *193*, 106670. [CrossRef]
107. Miyamoto, A.; Bendarkar, M.V.; Mavris, D.N. Natural language processing of aviation safety reports to identify inefficient operational patterns. *Aerospace* **2022**, *9*, 450. [CrossRef]
108. Baugh, B.S. *Predicting General Aviation Accidents Using Machine Learning Algorithms*; Embry-Riddle Aeronautical University: Daytona Beach, FL, USA, 2020.
109. Zhang, C.; Liu, C.; Liu, H.; Jiang, C.; Fu, L.; Wen, C.; Cao, W. Incorporation of Pilot Factors into Risk Analysis of Civil Aviation Accidents from 2008 to 2020: A Data-Driven Bayesian Network Approach. *Aerospace* **2022**, *10*, 9. [CrossRef]
110. Woolam, C.; Khan, L. Multi-concept document classification using a perceptron-like algorithm. In Proceedings of the 2008 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology, Sydney, NSW, Australia, 9–12 December 2008; pp. 570–574.
111. Xu, Z.; Saleh, J.H.; Subagia, R. Machine learning for helicopter accident analysis using supervised classification: Inference, prediction, and implications. *Reliab. Eng. Syst. Saf.* **2020**, *204*, 107210. [CrossRef]
112. Dai, Y.; Tian, J.; Rong, H.; Zhao, T. Hybrid safety analysis method based on SVM and RST: An application to carrier landing of aircraft. *Saf. Sci.* **2015**, *80*, 56–65. [CrossRef]
113. Hu, X.; Wu, J.; He, J. Textual indicator extraction from aviation accident reports. In Proceedings of the AIAA Aviation 2019 Forum, Dallas, TX, USA, 17–21 June 2019; p. 2939.
114. Zhang, X.; Mahadevan, S. A hybrid data-driven approach to analyze aviation incident reports. In Proceedings of the 2018 Aviation Technology, Integration, and Operations Conference, Atlanta, GA, USA, 25–29 June 2018; p. 3982.
115. Zhang, X.; Mahadevan, S. Bayesian network modeling of accident investigation reports for aviation safety assessment. *Eng. Syst. Saf.* **2021**, *209*, 107371. [CrossRef]
116. de Vries, V. Classification of aviation safety reports using machine learning. In Proceedings of the IEEE 2020 International Conference on Artificial Intelligence and Data Analytics for Air Transportation (AIDA-AT), Singapore, 3–4 February 2020; pp. 1–6.
117. İnan, T.T.; Gökmen İnan, N. The analysis of fatal aviation accidents more than 100 dead passengers: An application of machine learning. *Opsearch* **2022**, *59*, 1377–1395. [CrossRef]

118. Srinivasamurthy, A.; Motlicek, P.; Himawan, I.; Szaszak, G.; Oualil, Y.; Helmke, H. Semi-supervised learning with semantic knowledge extraction for improved speech recognition in air traffic control. In Proceedings of the Interspeech, Stockholm, Sweden, 20–24 August 2017; pp. 2406–2410.
119. Lakshman, N.; Raj, R.; Mukkamala, Y. Bird strike analysis of jet engine fan blade. In Proceedings of the 2014 IEEE Aerospace Conference, Big Sky, MT, USA, 1–8 March 2014; pp. 1–7.
120. Subramanian, S.V.; Rao, A.H. Deep-learning based time series forecasting of go-around incidents in the national airspace system. In Proceedings of the 2018 AIAA Modeling and Simulation Technologies Conference, Kissimmee, FL, USA, 8–12 January 2018; p. 0424.
121. Kierszbaum, S.; Lapasset, L. Applying distilled BERT for question answering on ASRS reports. In Proceedings of the IEEE 2020 New Trends in Civil Aviation (NTCA), Prague, Czech Republic, 23–24 November 2020; pp. 33–38.
122. Robinson, S.D.; Irwin, W.J.; Kelly, T.K.; Wu, X.O. Application of machine learning to mapping primary causal factors in self reported safety narratives. *Saf. Sci.* **2015**, *75*, 118–129. [CrossRef]
123. Robinson, S. Visual representation of safety narratives. *Saf. Sci.* **2016**, *88*, 123–128. [CrossRef]
124. Robinson, S.D. Temporal topic modeling applied to aviation safety reports: A subject matter expert review. *Saf. Sci.* **2019**, *116*, 275–286. [CrossRef]
125. Walton, R.O.; Marion, J.W. A textual analysis of dangerous goods incidents on aircraft. *Transp. Res. Procedia* **2020**, *51*, 152–159. [CrossRef]
126. Mathur, P.; Khatri, S.K.; Sharma, M. Prediction of aviation accidents using logistic regression model. In Proceedings of the IEEE 2017 International Conference on Infocom Technologies and Unmanned Systems (Trends and Future Directions)(ICTUS), Dubai, United Arab Emirates, 18–20 December 2017; pp. 725–728.
127. Rogers, P.; Pavur, R. Airline Safety Data: How Predictable Are Accidents and Fatalities? *Fed. Bus. Discipl. J.* **2019**, *8*, 19–29.
128. Majumdar, N.; Marais, K.; Rao, A. Analysis of General Aviation fixed-wing aircraft accidents involving inflight loss of control using a state-based approach. *Aviation* **2021**, *25*, 283–294. [CrossRef]
129. Distefano, N.; Leonardi, S. Apriori algorithm for association rules mining in aircraft runway excursions. *Civ. Eng. Archit.* **2020**, *8*, 206–217. [CrossRef]
130. Zhou, D.; Zhuang, X.; Zuo, H.; Cai, J.; Zhao, X.; Xiang, J. A model fusion strategy for identifying aircraft risk using CNN and Att-BiLSTM. *Reliab. Eng. Syst. Saf.* **2022**, *228*, 108750. [CrossRef]
131. Boyd, D.D. Causes and risk factors for fatal accidents in non-commercial twin engine piston general aviation aircraft. *Accid. Anal. Prev.* **2015**, *77*, 113–119. [CrossRef]
132. Kuşkapan, E.; Çodur, M.Y. Examination of Aircraft Accidents That Occurred in the Last 20 Years in the World. *Düzce Üniversitesi Bilim Ve Teknol. Derg.* **2021**, *9*, 174–188. [CrossRef]
133. Mangortey, E.; Speirs, A.; Bendarkar, M.V.; Bui, V. Analysis of Helicopter Accidents and Certification Categories Using Machine Learning. In Proceedings of the AIAA SCITECH 2022 Forum, San Diego, CA, USA, 3–7 January 2022; p. 0249.
134. Babič, F.; Lukáčová, A.; Paralič, J. Descriptive and predictive analyses of data representing aviation accidents. In Proceedings of the New Research in Multimedia and Internet Systems, Wroclaw, Poland, 17–19 September 2014; pp. 181–190.
135. Xusheng, G.; Jingshun, D.; Wei, C. Flight accident modeling and predicting based on least squares support vector machine. In Proceedings of the IEEE 2010 International Conference on Educational and Information Technology, Chongqing, China, 17–19 September 2010; pp. V3-256–V253-259.
136. Yu, H.; Li, X. On the chaos analysis and prediction of aircraft accidents based on multi-timescales. *Phys. A Stat. Mech. Its Appl.* **2019**, *534*, 120828. [CrossRef]
137. Lee, H.; Madar, S.; Sairam, S.; Puranik, T.G.; Payan, A.P.; Kirby, M.; Pinon, O.J.; Mavris, D.N. Critical parameter identification for safety events in commercial aviation using machine learning. *Aerospace* **2020**, *7*, 73. [CrossRef]
138. Bhanbhro, J.; Yousuf, F.; Narejo, S.; Furqan, M. PIA Accidents Analysis Using Naïve Bayes Classifier. In Proceedings of the International Conference on Computational Sciences and Technologies (INCCST'20), Virtual, 1–4 July 2020; pp. 17–19.
139. Carson, J.; Hollingsworth, K.; Datta, R.; Segev, A. Failing &! falling (f&! f): Learning to classify accidents and incidents in aircraft data. In Proceedings of the 2019 IEEE International Conference on Big Data (Big Data), Los Angeles, CA, USA, 9–12 December 2019; pp. 4357–4365.
140. Switzer, J.; Khan, L.; Muhaya, F.B. Subjectivity classification and analysis of the ASRS corpus. In Proceedings of the 2011 IEEE International Conference on Information Reuse & Integration, Las Vegas, NV, USA, 3–5 August 2011; pp. 160–165.
141. Williams, J.K. Using random forests to diagnose aviation turbulence. *Mach. Learn.* **2014**, *95*, 51–70. [CrossRef]
142. Fuller, J.G.; Hook, L.R. Understanding general aviation accidents in terms of safety systems. In Proceedings of the 2020 AIAA/IEEE 39th Digital Avionics Systems Conference (DASC), San Antonio, TX, USA, 11–15 October 2020; pp. 1–9.
143. Kazi, N.M.S. Using Machine Learning Models to Study Human Error Related Factors in Aviation Accidents and Incidents. Doctoral's Thesis, National College of Ireland, Dublin, Ireland, 2020.

144. Kuhn, K.D. Using structural topic modeling to identify latent topics and trends in aviation incident reports. *Res. Part C Emerg. Technol.* **2018**, *87*, 105–122. [CrossRef]
145. Marev, K.; Georgiev, K. Automated aviation occurrences categorization. In Proceedings of the IEEE 2019 International Conference on Military Technologies (ICMT), Brno, Czech Republic, 30–31 May 2019; pp. 1–5.
146. El Ghaoui, L.; Li, G.-C.; Duong, V.-A.; Pham, V.; Srivastava, A.N.; Bhaduri, K. Sparse machine learning methods for understanding large text corpora. In Proceedings of the CIDU, Mountain View, CA, USA, 19–21 October 2011; pp. 159–173.
147. Abedin, M.; Ng, V.; Khan, L.J. Cause identification from aviation safety incident reports via weakly supervised semantic lexicon construction. *Artif. Intell. Res.* **2010**, *38*, 569–631. [CrossRef]
148. Kerfoot, D.; Hofmann, M. Analysis of Aviation Accidents Data. In *CERI 2018 Proceedings*; Civil Engineering Research Association of Ireland: Dublin, Ireland, 2018; pp. 350–355.
149. Berry, M.W.; Gillis, N.; Glineur, F. Document classification using nonnegative matrix factorization and underapproximation. In Proceedings of the 2009 IEEE International Symposium on Circuits and Systems, Iasi, Romania, 9–10 July 2009; pp. 2782–2785.
150. Puranik, T.G.; Rodriguez, N.; Mavris, D.N. Towards online prediction of safety-critical landing metrics in aviation using supervised machine learning. *Transp. Res. Part C Emerg. Technol.* **2020**, *120*, 102819. [CrossRef]
151. Dong, T.; Yang, Q.; Ebadi, N.; Luo, X.R.; Rad, P.J. Identifying incident causal factors to improve aviation transportation safety: Proposing a deep learning approach. *J. Adv. Transp.* **2021**, *2021*, 5540046. [CrossRef]
152. Heinrich, D.J. *Safer* Approaches and Landings: A Multivariate Analysis of Critical Factors*; Capella University: Minneapolis, MN, USA, 2004.
153. Nanyonga, A.; Wasswa, H.; Turhan, U.; Molloy, O.; Wild, G. Sequential classification of aviation safety occurrences with natural language processing. In Proceedings of the AIAA AVIATION 2023 Forum, San Diego, CA, USA, 12–16 June 2023; p. 4325.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.