# AI based Human Factor Approach in Predicting Loss of Control Inflight Factors during Initial Climb in General Aviation


By

Fatima Khalil

(Registration No: 00000363903)

Department of Transportation Engineering

School of Civil and Environmental Engineering

National University of Sciences & Technology (NUST)

Islamabad, Pakistan


(2024)

# AI based Human Factor Approach in Predicting Loss of Control Inflight Factors during Initial Climb in General Aviation


By

Fatima Khalil

(Registration No: 00000363903)


A thesis submitted to the National University of Sciences and Technology, Islamabad,

in partial fulfillment of the requirements for the degree of

Master of Science in
Transportation Engineering

Supervisor: Dr. Sameer-ud-Din (P.E.)


School of Civil and Environmental Engineering

National University of Sciences & Technology (NUST)

Islamabad, Pakistan


(2024)

# THESIS ACCEPTANCE CERTIFICATE

It is certified that final copy of MS thesis written by **Ms. Fatima Khalil**, Registration No. **00000363903**, of MS Transportation Engineering SCEE, (NICE), has been vetted by undersigned, found completed in all respects as per NUST Statutes/Regulations, is free of plagiarism, errors, and mistakes and is accepted as partial fulfilment for award of MS degree.

Signature: ____________________

Supervisor: Dr. Sameer-ud-Din

Date: 26/02/2024

Signature (HoD) HoD Transportation Engineering
NUST Institute of Civil Engineering
School of Civil & Environmental Engineering
National University of Sciences and Technology

Date: 27-02-2024

Dr. S. Muhammad Jamil
Associate Dean
NICE, SCEE, NUST

Signature (Associate Dean, NICE) ____________________

Date: 27/2/2024

Signature (Principal & Dean): --------------------
PROF DR MUHAMMAD IRFAN
Principal & Dean
SCEE, NUST

Date: 28 FEB 2024

## Certificate of Approval

This is to certify that the research work presented in this thesis, entitled "AI based Human Factor Approach in Predicting Loss of Control Inflight Factors during Initial Climb in General Aviation" was conducted by Ms. Fatima Khalil under the supervision of Dr. Sameer-ud-Din.

No part of this thesis has been submitted anywhere else for any other degree. This thesis is submitted to the National University of Sciences and Technology (NUST) in partial fulfillment of the requirements for the degree of Master of Science in field of Transportation Engineering from NUST institute of Civil Engineering (NICE), School of Civil and Environmental Engineering (SCEE), NUST

Student Name: Fatima Khalil Signature: Fatimay

Examination Committee:

a) GEC Member 1: Dr. Muhammad Usman Hassan
Assistant Professor (SCEE, NUST) Signature

b) GEC Member 2: Dr. Muhammad Asif Khan
Assistant Professor (SCEE, NUST) Signature

Supervisor Name: Dr. Sameer-ud-Din Signature:

Name of HOD: Dr. Arshad Hussain Signature:

HOD Transportation Engineering
NUST Institute of Civil Engineering
School of Civil & Environmental Engineering
National University of Sciences and Technology

Name of Associate Dean: Dr. S. Muhammad Jamil Signature:

Dr. S. Muhammad Jamil
Associate Dean
NICE, SCEE, NUST

Name of Principal & Dean: Prof. Dr. Muhammad Irfan Signature:

PROF DR MUHAMMAD IRFAN
Principal & Dean
SCEE, NUST

# AUTHOR'S DECLARATION

I Fatima Khalil hereby state that my MS thesis titled "AI based Human Factor Approach in Predicting Loss of Control Inflight Factors during Initial Climb in General Aviation" is my own work and has not been submitted previously by me for taking any degree from National University of Sciences and Technology, Islamabad or anywhere else in the country/ world.

At any time if my statement is found to be incorrect even after I graduate, the university has the right to withdraw my MS degree.

Name of Student: Fatima Khalil

Date: 26th February 2024

# PLAGIARISM UNDERTAKING

I solemnly declare that research work presented in the thesis titled "AI based Human Factor Approach in Predicting Loss of Control Inflight Factors during Initial Climb in General Aviation" is solely my research work with no significant contribution from any other person. Small contribution/ help wherever taken has been duly acknowledged and that complete thesis has been written by me.

I understand the zero-tolerance policy of the HEC and National University of Sciences and Technology (NUST), Islamabad towards plagiarism. Therefore, I as an author of the above titled thesis declare that no portion of my thesis has been plagiarized and any material used as reference is properly referred/cited.

I undertake that if I am found guilty of any formal plagiarism in the above titled thesis even after award of MS degree, the University reserves the rights to withdraw/revoke my MS degree and that HEC and NUST, Islamabad has the right to publish my name on the HEC/University website on which names of students are placed who submitted plagiarized thesis.

Student Signature: Fatimah

Name: Fatima Khalil

## DEDICATION

"To My Beloved and Supporting Parents, My Siblings, and My Dedicated Teachers, Without Whom This Milestone Would Not be Achieved."

# ACKNOWLEDGEMENTS

I would like to express my extreme gratitude to my research supervisor, **Dr. Sameer-ud-Din**, for his unwavering support, accurate guidance, and absolute motivation and encouragement. I would also like to thank my GEC members, **Dr. Muhammad Usman,** and **Dr. Muhammad Asif**, for their guidance throughout my research. Lastly, I would like to thank my MS fellows at National University of Sciences and Technology, for their support and motivation.

# TABLE OF CONTENTS

# LIST OF TABLES

**Page No.**

# LIST OF FIGURES

**Page No.**

# LIST OF SYMBOLS, ABBREVIATIONS AND ACRONYMS

| | |
|---|---|
| FAA | Federal Aviation Authority |
| FTA | Fault Tree Analysis |
| GA | General Aviation |
| GAJSC | General Aviation Joint Steering Committee |
| HFACS | Human Factor Analysis and Classification System |
| IATA | International Air Transport Association |
| ICAO | International Civil Aviation Organization |
| ICL | Initial Climb |
| LOC-I | Loss of Control In-Flight |
| NTSB | National Transportation Safety Board |
| ML | Machine Learning |
| RF | Random Forest |
| SHAP | SHapley Additive exPlanations |

# ABSTRACT

Human factors are increasingly leading the General Aviation's (GA) accident causation although the total number of accidents have significantly improved over past few decades. The actions majorly taken so far correspond to reactive safety approaches rather than proactive ones. GA has been neglected a lot in terms of safety and risk mitigation as the fatality rate has been almost constant for many years now. In this research study, the probable causes of GA Loss of Control-In Flight (LOC-I) accidents under Initial Climb (ICL) phase of flight are obtained from National Transportation Safety Board (NTSB). Each accident is classified into one of the 9 LOC-I accident categories defined by International Air Transport Association (IATA). The preprocessed and feature engineered dataset is fed to a Random Forest (RF) model to be trained. The prediction model gives an accuracy and F-1 score of 88% on the test set. Feature importance and SHapley Additive exPlanation (SHAP) analysis of RF model is performed to get the most influencing features on prediction. The most influential features of the RF model vulnerable are connected to the Human Factor Analysis and Classification System (HFACS) to get insights into the most vulnerable HFACS levels.

**Keywords:** Aviation Safety, HFACS, Machine Learning, Random Forest, Loss of Control-In Flight.

# CHAPTER 1: INTRODUCTION

## 1.1. Background

Aviation accidents are among the deadliest in the transportation world, resulting in a significant loss of lives, numerous injuries, and the destruction of the aircraft body destroyed, gaining widespread public attention. Technological advancements since the 1920's have significantly contributed to mitigating the occurrences. The establishment of mandatory reporting and thorough investigation for every accident has further improved protocols. Among the top aviation accident, the shooting down of a Ukrainian Aircraft in Iranian airspace in January 2020 resulted in 176 casualties. The second deadliest accident was the crash of Pakistan Airline Flight with 98 causalities in May 2020 [1].

Focusing on Pakistan's aviation accidents, Dawn, a Pakistani print media, has reported 18 serious commercial airplane crashes in the country's history. These range from the first crash on 20th May 1965, which resulted in 124 casualties, to the most recent one on 22nd May 2020, with 98 casualties [2]. Investigation by the Pakistan Civil Aviation Authority revealed that the PIA Airbus A320 crash, on 22nd May 2020 near Karachi airport, was due to human error [3]. The cockpit crew was flying the plane above the recommended speed and altitude, and the control tower granted permission for it. This negligence towards critical parameters led to the plane crash [3]. One of the most disastrous crashes was the Airblue Airbus 321 crash in the Margalla mountains of Islamabad, causing fatal injuries to all 152 passengers on board [2]. The pilot's mishandling due to bad weather conditions came out to be the reason behind the accident in the preliminary investigation reports.

## 1.2. Factors Leading to General Aviation Accidents

The aviation accidents can be attributed to three main factors: human, weather and aircraft. In the early days of aviation, mechanical faults were a common cause of accidents due to continuous changes in machine design. However, even with advancements in aircraft design, accidents persisted, leading to the realization that human error might be more threatening than aircraft issues (Mason 1993,[4] cited in Murray 1997) [5]. Trollip and Jensen (1991) stated that over 80% of the General Aviation (GA) accidents were a result of pilot error, often streaming from wrong judgements of the situation [6]. However, solely blaming pilots was a simplistic approach, as accidents are rarely due to single absolute reason or the fault of a single person (Helmreich, Peterson, and Roos, 1980) [7]. In aviation, accidents typically involve multiple faults at various organization levels, with the cockpit crew's unsafe acts being the final cause [8][9] [10].

Jensen and Benel (1977) reported that 52% of fatal GA accidents, between 1987 to 1989 in the United States (US), were due to poor decision-making by pilots. This finding was further supported by another research study in 1991, which analyzed US civil aviation accidents during the same period and concluded that 56% of accidents resulted from poor pilot judgement [11]. Despite the disastrous consequences of human error, it is important to question why pilots continue to adopt such behaviors.

## 1.3. Understanding General Aviation Accident Categories

To study aviation accidents and enhance safety parameters, one must understand the classification and taxonomy of aviation occurrence categories. These categories classify accident data into sub-sets based on different phases of an aircraft accident. The

International Civil Aviation Organization (ICAO) has classified the aircraft accident phases into thirty-eight (38) occurrence categories [12]. Research study by ICAO, analyzing accident data over a10 year period (2011-2021), showed that Loss of Control In-Flight (LOC-I) caused the highest number of fatal accidents (49% of the total considered accidents), even though the total number of LOC-I accidents was less than other categories (8% of the total considered accidents). This emphasized the high fatality rate of LOC-I compared to other accident categories [13].

The Federal Aviation Authority (FAA) established the General Aviation Joint Steering Committee (GAJSC) in 1997 to decrease the number of fatal GA accidents [14]. The GAJSC Accident Data Set of GA fatal accidents from 2001 to 2010 revealed that LOC-I was the most common occurring accident category, accounting for 40% of the 3136 total fatal GA accidents [15]. This data set was utilized in a 2012 report by the GAJSC, which aimed to determine the causes of fatal GA accidents and preventing future occurrences [16].

The FAA also introduced the Joint Safety Analysis Team (JSAT) to enhance aviation safety and identified the three most frequent occurrence categories as Controlled Flight into Terrain (CFIT), Loss of Control in Flight (LOC-I), and Approach and Landing [17]. An analysis report by the International Air Transport Association (IATA) for the period of 2009-2018 confirmed that LOC-I accidents were the most fatal among all categories [18]. The report revealed that 94% of analyzed LOC-I accidents resulted in fatalities among passengers or flight crew, emphasizing the severity of LOC-I accidents [18]. JSAT defined LOC-I as “a significant, unintended departure of the aircraft from controlled flight, the operational flight envelope, or usual flight attitudes, including ground events. "Significant" implies an event that results in an accident or incident. This definition excluded catastrophic explosions, CFIT, runway collisions, complete loss of thrust that did not involve loss of control, and any other accident scenarios in which the crew retained control" [17].

Aviation occurrence categories are linked to the phases of flight of an aircraft, which ICAO classifies into 13 distinguishing groups, including Takeoff (TOF), Initial Climb (ICL), Cruise (CRZ) etc. [19]. LOC-I accidents are particularly catastrophic when they occur during the Initial Climb (ICL) phase of flight, as reported by IATA in their accident analysis report (2009-2018) [3]. Approximately 26% of all LOC-I accidents happened during this flight phase, with a fatal to non-fatal accident ratio of 94% [18]. The FAA defined the Initial Climb (ICL) phase as "The initial climb begins when the airplane leaves the surface, and a climb pitch attitude has been established. Normally, it is considered complete when the airplane has reached a safe maneuvering altitude, or an en-route climb has been established" [20].

### 1.4. Accident Mitigation Techniques Through Decades

Failures in organizations are often viewed as a negative outcome that must be avoided at all costs. However, the concept of learning from failures is crucial for organization growth, even though such organizations are still rare. Organizational failures can also have a positive impact, as they provide opportunities to develop countermeasures and build resilience within the organization, preventing the repetition of the same mistakes [21][22][23][24]. Although small failures may be disregarded due to their limited impact, they can accumulate over time, leading to more significant disasters [22]. Taleb N., in his book, refers to this phenomenon as ‘Black Swan’, where small errors combine in a chain reaction to result in a larger catastrophe [25].

The analysis of failure/accident in any system has always been a topic of concern. However, the focus was often on blaming human operator rather than considering the overall system structure and environment. An example of this is the ‘theory of dominos’ presented by a safety engineer H. W. Heinrich in 1931[26]. It described accident as ‘a linear sequence of event or stones (dominoes)’ standing edge-to-edge, where if one falls, the others follow, leading to a system collapse. While considered one of the best accident defining theories, it faced criticism for putting excessive blame on individuals and neglecting the effects of organization and management [27].

Over time, with research advancements, several statistical and analytical techniques were introduced and used for accident causation analysis. Widely used techniques include Why-because analysis, Root Cause analysis, Sequentially Timed and Events Plotting, Management Oversight and Risk Tree (MORT), Fault Tree Analysis (FTA), and Event Tree Analysis (ETA) [28][29]. These techniques offer a more comprehensive approach to understanding the underlying causes of accidents and allow for a more systemic view rather than solely focusing on human operators as the source of failure.

Fault Tree Analysis (FTA) is a deductive analytical technique based on Boolean Algebra, Probability theory, and Reliability theory. It utilizes a sequential diagram to analyze accidents, starting from the basic events at the bottom and leading to the undesired event at the top. The main aim of FTA is to identify the causes of the event. Represented as an inverted tree, the branches (causing events) connect to form the trunk (accident) through logic gates like AND, OR, and NOT. These gates' function like those in electric circuits or piping networks, allowing few, all, or none to pass based on the gate type. FTA does not require the inclusion of all possible types and causes of system failures, but it must cover all the causes of a specific failure (top event).

### 1.5. Reason’s Swiss Cheese Model

James Reason introduced the "Swiss Cheese" model in 1990 to describe human error causation, initially focusing on nuclear power plant operators [8]. However, this model can be adapted for use in any organization with minor adjustments. The model represents four stages of human failure represented as slices of cheese, influencing each other, and culminating in a failure event. The holes in the slices symbolize vulnerability to human failure, categorized as latent failures or active failures. Latent failures remain dormant until they contribute to mishaps, while active failures refer to unsafe acts by humans. Mishaps occur when latent and active failures align, as shown in **Figure 1.1**. The “Swiss cheese” underscores how small, and neglected errors can combine to cause catastrophic accidents.

In the context of aviation accidents, the "Swiss Cheese" model highlights “Unsafe Acts” as the nearest failure leading to accidents, often attributed to aircrew members or pilots. Investigations commonly concentrate on this level, as it reveals the cause of most accidents, whether directly or indirectly.

The following three levels are based on latent failures. The first level, “Preconditions to Unsafe Acts”, addresses mental and physical conditions of operators. Crew Resource Management (CRM) aims to mitigate these conditions, as they can lead to poor decision-making and accidents. The third slice, “Unsafe Supervision”, emphasizes the lack of controlled supervision over aircrew and their practices, leading to failures. Last but not the least, Reason links the failure back to “Organizational Influences” addressing problems at higher organization levels affecting supervision (CRM) and contributing to unsafe acts by aircrew members.

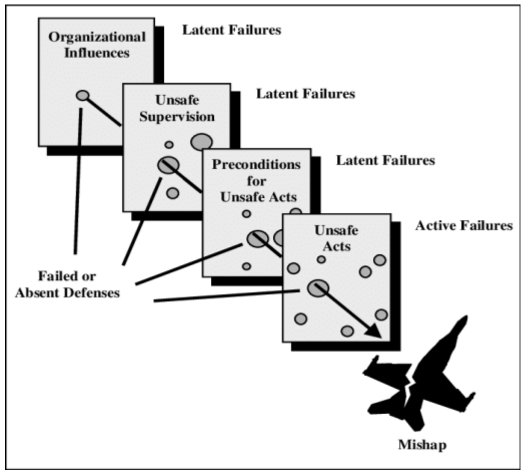


**Figure 1.1:** Reason’s "Swiss cheese" model of human error [8]

## 1.6. Human Factor Analysis and Classification System (HFACS)

Following Reason’s theory, Dr. Scott Shappell and Dr. Douglas Weigmann developed the Human Factor Analysis and Classification System (HFACS) in 1997 to address high-rate aviation accidents in the United States Navy [9]. The cause of the accidents was found to be violations committed by experienced Navy/Marine officers and new enlistees [30]. Later, HFACS further classifies latent and active failures into sub-categories to analyze aviation accidents. This research will focus solely on discussing ‘Unsafe Acts’ and ‘Precondition for Unsafe Acts’ in detail, given its scope.

Unsafe Acts- Operator’s unsafe acts can be broadly classified as errors and violations. Errors are an undeniable part of human nature, where one fails to accomplish the required task. Whilst violations are deliberate actions taken to disregard the set rules and regulations of an organization. Among these two, the enactment of violations by aircrew rules over the causation of aviation accidents. Although pilots are highly literate and experienced in their field, they are still humans, and making mistakes is a normal trait. Therefore, it is crucial to minimize the occurrence and consequences of these traits [31]. HFACS further divides errors into three categories and violations into two sub-categories.

The first division of error, decision errors, includes situations where the proper intended plan to carry out an operation becomes improper and inadequate under those circumstances. These are often referred to as ‘honest mistakes’, meaning that the doer has good intentions but made a poor decision. Decision errors include poor decision-making, inadequate procedures, failure to perceive emergencies, and incorrect reactions to them. In the context of aviation, pilots often make poor decisions in critical situations, leading to accidents. For example, on 20th April 2012, M/s Bhoja Air Flight BHO-213 was flying from Karachi to Islamabad with 127 passengers, including 6 crew members. The plane crashed, resulting in the loss of all onboard lives and the complete destruction of the aircraft. The investigation report by Pakistan Civil Aviation Authority (PCAA) clearly stated that the reason for the accident was the incorrect decision-making of the cockpit crew under adverse weather conditions (AWC), and non-adherence to prescribed rules and regulations [32].

The second division of error is skilled-based errors, which involves instinctive behavior by pilots without conscious thinking. These errors can be attributed to timely visual breakdown, memory failures due to lack of proper attention, skipping a step in some procedure, and use of substandard techniques, among others.

The third and final division of error is perceptual errors, which occur when a person's perception of the world around them differs from reality. In the case of aviation, such errors occur when the input data from the senses becomes delusional, leading to misjudge the altitude or speed of the aircraft.  Perceptual errors can also be attributed to visual illusions and space-disorientation.

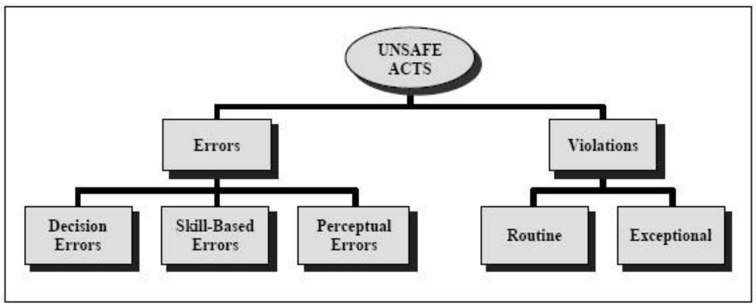


**Figure 1.2:** Unsafe Acts; HFACS [9]

Violations, as discussed earlier, refer to voluntary actions or behaviors of pilots that lead them to disregard rules and regulations governing safe flight operations, often resulting in unfortunate incidents involving fatalities. It's worth noting that pilots with a history of violations are more likely to repeat such behavior in the future [33]. Additionally, many aviation accidents occur due to indecisive transitions from VFR (Visual Flight Rules) to IMC (Instrument Meteorological Conditions) under unexpected weather conditions [34]. Most disastrous aviation accidents, by percentage, involve continued VFR to IMC situations [35]. This behavior can also be attributed not only to poor decision-making but also to presumptuous behavior in handling adverse situations [36].

Based on frequency and possibility of committing violations, there are two types. The first type is routine violations, which are repeated periodically and become a habit for the doer. Unfortunately, such violations are often overlooked by supervisory personnel [9]. For example, flying the plane slightly faster than the prescribed speed for a specific zone may be accepted without proper consequences, exacerbating the problem. Implementing proper warnings or fines for routine violations can significantly reduce their occurrences.

The second type of violation is exceptional violations. The term ‘exceptional’ here does not imply that the act is extraordinary, but rather that such violations are not typical behavior for the pilot and are not accepted by the authorities [9]. The most challenging aspect of such violations is that they do not align with the pilot’s usual behavior, leaving them with few exceptions when asked about the reasons behind their actions.

# CHAPTER 2: LITERATURE REVIEW

## 2.1. Background

Aviation has always been associated with accidents and incidents. From the first crash happening on 17th September 1908, Wright brothers improved the aircraft design and thus officially started the aviation safety system as we call it today [45]. James Reason described the objective of safety as the avoidance of damage as much as possible, including injuries, fatalities, aircraft, and surroundings [46]. Human factors are the cause of more than 90% of the incidents that happen in the aviation industry and others such as maritime, highway etc. [47]. Aviation safety is not only the task of a single person i.e., pilot. It involves all levels of management, supervision and even organization. Textron explained that safety starts from the top levels of management, and so all the individuals including pilots, supervisors, mechanics etc. should be properly trained to prevent all levels of injuries [48]. To further strengthen the concept, Reason showed a broader understanding of humans causing aviation accidents. He argued that the ones in the proximity of the system might be more probable to cause an accident (active failures), however it is a dire need to understand the unseen factors that contributed to such outcome (latent failures) [49]. In today's world, we call this approach 'Safety Management Systems'.

## 2.2. Human Factors Contribution to GA Accidents

A significant number of studies have been carried out in the past that declared that human error majorly contributes to GA accidents either directly or indirectly. Annual report issued by U.S. General Aviation Accident Data 2006 [50] showed that 91% of the 1494

GA accidents happened in the year 2006 were personnel-related, with 95% personnels being pilots. Other personnels not aboard the aircraft also caused 8% of the accidents, including flight instructors, maintenance personnels and airport staff [50]. It also showed that student pilot's aircraft control and decision-making errors were the main cause of the flight instruction GA accidents [50]. The 33rd AOPA Air Safety Institute Accident Report [51] encompassed non-commercial accidents from 2012 to 2021. It showed that 647 out of 938 accidents (69%) studied in the report were pilot-related [51]. Out of these accidents, 62% were fatal accidents [51]. Jensen and Benel (1977) reported that 52% of fatal GA accidents, between 1987 to 1989 in the United States (US), were due to poor decision-making by pilots [52]. This finding was further supported by another research study in 1991, which analyzed US civil aviation accidents during the same period and concluded that 56% of accidents resulted from poor pilot judgement [53].

Over the past few decades, aviation accidents have reduced largely in numbers, with the advancements in technology and safety systems. However, the contribution of human factors as the leading cause has increased drastically [54][55]. Human-factors not only correspond to pilots but also the air traffic staff, air crew and maintenance personnels. The related factors can be psychological factors such as stress and anxiety, health issues, errors, and violations, experience-related etc. Kara A. Latorella studied the role of aviation maintenance in the causation of the accident [56]. It was found that fatigue, miscommunication among the aviation workers and less knowledge were the leading causes of the aviation accidents [56].

## 2.3. Machine Learning in Aviation Accidents Study

In recent years, ML has proved to be the most powerful tool to play with the past datasets to derive future predictions. In contrast to the traditional regression analysis, it has shown better and more accurate prediction results [57]. This led to its larger use in aviation accidents studies [58] and the related human factors analysis on it [59]. Olja Čokorilo conducted a cluster analysis study over 1500 aircraft accidents occurring between 1985 and 2010 to compare these accidents based on aircraft, environment, and traffic characteristics [60]. Bradley S. Baugh in 2020 analyzed GA accidents from 1998 to 2018 using ML and data mining techniques [37]. To find the best fitted model, he used both the unstructured (text format) and tabular data from NTSB, and applied ML methods of Decision Tree, Gradient Boosting, Logistic Regression, Neural Network, and Random Forest [37].

Tomas Madeira et al. (2021) identified human factors from the aviation incident reports from 2000 to 2020. The HFACS framework was modified by making it compatible with the investigation reports of Aviation Safety Network (ASN). The engineered dataset was then fed into the machine learning algorithms; semi-supervised Label Spreading (LS) algorithm, supervised Support Vector Machine (SVM). Micro F1 scores of 0.9, 0.779 and 0.875 were obtained for predictions for unsafe supervision, preconditions to unsafe acts and unsafe acts respectively [61]. A. Khattak (2023) studied the missed approaches (MAPs) at Hong Kong International Airport (HKIA) from 2017 to 2021 under low-level wind shear conditions [62]. Machine learning techniques of random forest (RF), light gradient boosting machine (LGBM) and extreme gradient boosting machine (XGBoost) were used in the analysis. For the hypermeters tuning, Bayesian optimization and 10-fold cross validation is used. The LGBM got the highest precision of 75.23%, while RF got

69.11% precision [62]. Rui P. R. Nogueira (2023) studied human factors' involvement in the failure of aviation operations [63]. The study was mainly focused on predicting the occurrences of fatal accidents because of human-related input features. HFACS is used to find the relationships among the input features and human factors. For the model training, two supervised algorithms (Random Forest & Artificial Neural Networks) and one semi-supervised (Active Learning) were used. The RF model gave the prediction accuracy of 90% and outperformed ANN models [63].

Some of the researchers took detailed aviation accidents written reports and applied NLP techniques to find out the main topics of the accidents in the form of words [64]. The drawback with such a study is that those words are not necessarily used as the cause of accident but can be used anywhere in the report. To handle such a wide range of multilabel classifications, S. Robinson took 4000+ accidents from ASRS dataset and trained in a Latent Semantic Analysis (LSA) model [65]. The results of the study were not very good with an F1 score of 0.409 only, the reason for which was the little sample size used to classify all the contributing factors [65]. D. Tianxi devised a deep learning model to make the extraction of causal factors from aviation accidents easy [66]. Around 200,000 incident reports from Aviation Safety Reporting System (ASRS) are used to train, validate, and test the model. To obtain causal factors, first NLP and then long short-term memory (LSTM) model is used [66].

### 2.4. Research Gap

1. Government and semi-government institutions conduct limited research to reduce general aviation (GA) accidents [37]. Their focus primary lies in reactive safety

approaches, which deals with addressing incidents that have already occurred. Safety mitigations are generally categorized into reactive safety, proactive safety, and predictive safety. Reactive safety deals with past incidents, proactive safety identifies and corrects preconditions of potential incidents, while predictive safety involves predicting incidents using models build from past data. Proactive safety is rarely applied for GA accidents, while it is more common in the case of commercial aviation (CA) [37]. Similarly, predictive safety is also utilized in GA research, despite its potential for addressing safety deficiencies. Utilizing machine learning algorithms can significantly overcome GA safety and risk mitigation by actively employing the predictive safety approach.

2. Another compelling reason to focus on GA accidents in research is that their total number of occurrences exceeding commercial aviation accidents. Bradley S. Baugh research [37] showed a significant disparity between Part 91 (GA) and Part 121 (CA) accident rates. Of particular concern is the fatality rate, which has decreased for CA accidents due to improved safety measures and risk mitigations adopted since 2002. In contrast, GA fatality rates have remained relatively constant over the same 15-year period.

3. Autonomous vehicles, such as Unmanned Aerial Vehicles (UAVs) and drone operations, will be controlled under Unmanned Aircraft System Traffic Management (UTM) in cooperation with FAA to ensure smooth aerial operations without conflicting with other aviation operations [38]. Taiwan has already implemented UTM services under the Civil Aeronautical Administration (CAA) since March 31, 2020 [39]. Small UAVs flying below 400 ft are controlled by local

governments, while those flying above 400 ft fall under the jurisdiction of CAA [39]. Federal Aviation Regulation (FAR) part 91.119 states that all aviation operations must maintain a minimum flying altitude of 500 ft above rural areas and 1000 ft above urban areas [40]. As GA includes low-line operations, the integration of UAVs with GA is expected to change the landscape of GA operations, which is the focus of this research.

4. The initial climb (ICL) is a critical phase of flight, as flight crews face heavy workload coordinating with Air Traffic Control (ATC) and following standard operating procedures (SOPs). While the importance of the ICL flight phase in Loss of Control In-flight (LOC-I) accidents has been highlighted in CA accident reports [41][42][43], it has received limited attention in GA accident analyses. This research aims to address this gap, found in literature, by solely focusing on LOC-I accidents occurring during the ICL flight phase.
5. Although Human Factor Analysis and Classification System (HFACS) is valuable in evaluate GA accidents causation, it has some limitations. Accident investigation reports issued by governing bodies often fail to attribute organizational factors as contributing factor [44], making the application of HFACS challenging. Since organizational influence plays a significant role in aviation accidents, it cannot be ignored. Additionally, the specific categories in the HFACS framework make it difficult to classify the diverse range of aviation accidents accurately. Hence, this research combines HFACS with machine learning (ML) algorithms to get more detailed results, keeping in view model limitations of statistical tools for analysis.

6. In understanding the psychological factors contributing to aviation accidents, many studies involve interviews with relevant personnel or experts' opinions. However, such classifications can be influenced by personal thinking or bias. To avoid ambiguity, this research relies solely on unbiased methods using the original accident data.

# CHAPTER 3: METHODOLOGY

To engineer human factors contributing to the GA LOC-I accidents in the ICL phase of flight, two major sections in methodology are adopted. The first section focuses on the acquiring of desired dataset and implementing the machine learning algorithm to better understand the dataset and its trends. Another major outcome of the study is to find the important input variables influencing the predictions made by the trained model.

The second section of the study focuses on the modification of the HFACS framework to adapt to the obtained dataset. The modified framework is then used to identify the HFACS's level corresponding to the important features and hence suggesting required reforms at those levels respectively.

## 3.1. Data Collection

The general aviation accidents dataset was obtained from the National Transportation Safety Board (NTSB) official website, identified as a 'publicly available dataset'. Access to the dataset was facilitated through the 'custom search' feature within the CAROL Query section of the website [a]. The applied filters during the custom research focused on the aviation operation labelled as General Aviation and investigation fields specified as 'Probable Cause' and 'Finding Text'. The dataset contained GA accidents recorded from January 2008 to January 2023, with a total count of 78,905 accidents. Once the dataset was accessible, it underwent preprocessing procedure.

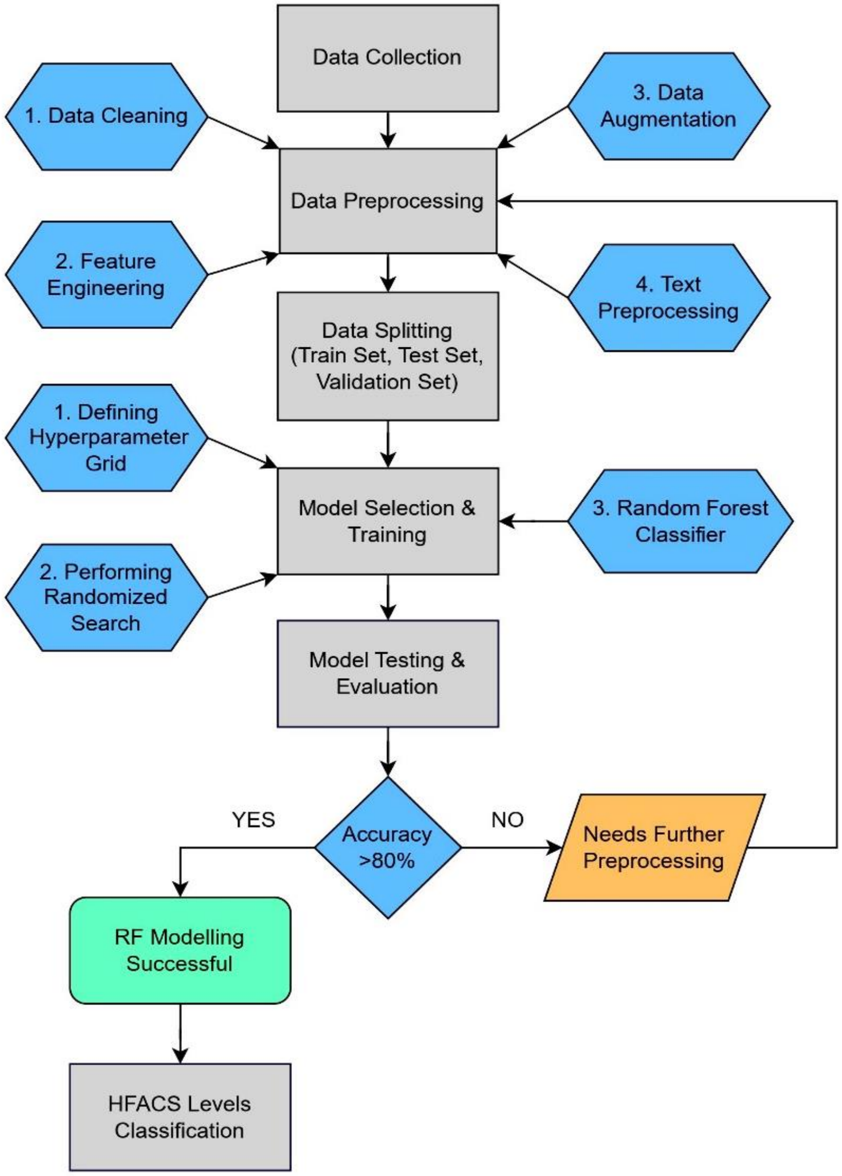


**Figure 3.1:** Methodology Framework

## 3.2. Data Preprocessing

Data preprocessing is a critical step that involves identifying and rectifying inaccuracies, addressing discrepancies, and eliminating extraneous data. Datasets acquired from open sources are typically raw and necessitate extensive preprocessing to enhance data quality and prepare them for subsequent analysis. This process ranges from straightforward tasks like removing incomplete entries to more intricate procedures such as encoding the entire dataset into a format compatible with specific algorithms.

### *3.2.1. Data Cleaning and Filtering*

Data filtering involves selecting specific data points from a dataset based on predefined criteria and conditions. In this study, ‘categorical filtering’ was employed, applying specific filters such as setting the phase of flight as ‘initial climb’, accident occurrence category as ‘loss of control in flight’ and selecting only the FAR 91 and FAR 137 GA accidents. These filters narrowed down the accidents count to 719.

Subsequently, data cleaning was conducted, eliminating all empty or incomplete accident entries. Entries containing ‘not determined’ or ‘unknown’ in the findings were also removed to further refine the dataset. This additional cleaning reduced the dataset to 703 accident entries.

### *3.2.2. Feature Engineering*

Feature engineering includes modifying, deleting, or creating new features to prepare the data for machine learning (ML) algorithms. Commonly used techniques involve selection, transformation, and amalgamation.

The ‘findings’ of the GA accidents, extracted from the dataset, were categorized into three main types: aircraft, personal issues, and environmental issues. These three categories were further subdivided based on the causes of the accidents. Each accident could have multiple findings or causes, totalling more than 550, which is overwhelming for input feature ‘x’ in a machine learning algorithm. The main concern lies not in the high number of input features but in the relatively lower number of available accidents. To address this issue, features were grouped together, and some were removed based on importance and frequency. The clustering of features falls under the domain of 'data transformation' as 'aggregation of features'. The table below shows a few feature aggregations used in this study.

**Table 3.1:** Aggregation of few Inputs Features

| Lack of Inspection | Decision Making/ Judgement | Improper Aircraft Control |
|---|---|---|
| Aircraft - Aircraft handling/service - Maintenance/inspections - Time limits - Not inspected | Personnel issues - Action/decision - Info processing/decision - Decision making/judgment – Pilot | Personnel issues - Task performance - Use of equip/info - Aircraft control - Pilot |
| Aircraft - Aircraft handling/service - Maintenance/inspections - Scheduled maintenance checks - Not inspected | Personnel issues - Action/decision - Info processing/decision - Decision making/judgment - Instructor/check pilot | Personnel issues - Task performance - Use of equip/info - Aircraft control - Student/instructed pilot |
| Aircraft - Aircraft structures - Doors - Service doors - Not inspected | Personnel issues - Action/decision - Info processing/decision - Decision making/judgment - Student/instructed pilot | Personnel issues - Task performance - Use of equip/info - Aircraft control - Flight crew |
| Aircraft - Aircraft handling/service - Maintenance/inspections - (general) - Not inspected | Personnel issues - Action/decision - Info processing/decision - Decision making/judgment - Owner/builder | Personnel issues - Task performance - Use of equip/info - Aircraft control - Instructor/check pilot |
| Aircraft - Aircraft systems - Flight control system - Aileron control system - Not inspected. | Personnel issues - Action/decision - Info processing/decision - Decision making/judgment – Copilot. | Personnel issues - Task performance - Use of equip/info - Aircraft control - Not specified |
| Personnel issues - Miscellaneous - Intentional act - Stolen/unauthorized – Pilot | Personnel issues - Action/decision - Info processing/decision - Decision making/judgment - Flight crew. | Personnel issues - Task performance - Use of equip/info - Aircraft control - Pilot of other aircraft. |

### 3.2.3. *Data Augmentation*

Data augmentation is applied in case of class imbalance in classification tasks to enhance the dataset and the robustness of the machine learning. Typically, the dataset exhibits a significant bias towards a few features that occur most frequently. Such class imbalance affects the accuracy of the ML model due to insufficient entries for less common classes. Addressing this issue involves generating one-to-one relationships between inputs and outputs, automatically improving the accuracy of the ML model.

Given that the available dataset was insufficient for achieving a higher accuracy ML model, a data augmentation technique called 'oversampling' was applied. This technique equalizes the weight differences among the various classes of the output 'y', effectively increasing the dataset size by 3.2 times the original. The augmented dataset is now used for further analysis.

### 3.2.4. *Text Preprocessing*

Natural Language Processing (NLP) plays a vital role while dealing with text databases. The dataset in this research involves input features 'x' and output classes 'y' in the form of sentences/ words, it is essential to transform the text into a readable format for the ML algorithm. To achieve this, the oversampled dataset obtained in the previous step underwent 'tokenization' and then 'one-hot encoding'. Tokenization is a text preprocessing technique that involves converting words into individual tokens.

One-hot encoding, falling under the domain of 'NLP feature extraction', takes these tokens and transforms them into binary vectors. Each feature is converted into a column with the number of rows equal to the sample size of the dataset. In this process, entries in the dataset containing a specific feature are assigned 1, indicating presence, while others are assigned 0, indicating absence.

### *3.2.5. Finalized Dataset*

Consider a multiclass classification problem with a Boolean feature space. This space refers to a dataset where features take values of 0 or 1, representing the presence (1) or absence (0) of that feature. In other words, it is a one-hot encoded dataset as discussed in 3.2.4. The input space can be defined as follows:

x = 0, 1B, where b are the input features having the value 0 or 1.

**Table 3.2:** Input features and their % presence in the dataset

| Abb. | Input Features | % | Abb. | Input Features | % |
|---|---|---|---|---|---|
| **B1** | Action/ Decision | 18.21 | **B31** | Improper Use/ Operation of Aircraft Propeller System | 0.71 |
| **B2** | Aircraft Airspeed Performance/ Capability | 36.13 | **B32** | Improper Use/ Operation of Flight Control System | 2.42 |
| **B3** | Aircraft Altitude Performance/ Capability | 2.42 | **B33** | Inadequate Inspection | 5.12 |
| **B4** | Aircraft Angle of Attack Performance/ Capability | 24.89 | **B34** | Inadequate Service/ Maintenance | 4.41 |
| **B5** | Aircraft CG/ Weight Distribution | 1.99 | **B35** | Inadequate Training | 1.00 |
| **B6** | Aircraft Climb Performance/ Capability | 1.85 | **B36** | Incorrect Service/ Maintenance | 3.13 |
| **B7** | Aircraft Climb Rate Performance/ Capability | 2.56 | **B37** | Incorrect/ Use Operation of Landing Gear System | 0.57 |
| **B8** | Aircraft Engine & Systems Failure | 4.13 | **B38** | Lack of Communication | 0.28 |
| **B9** | Aircraft Engine & Systems Malfunction | 4.27 | **B39** | Lack of Inspection | 1.28 |

| | | | | | |
|---|---|---|---|---|---|
| **B10** | Aircraft Fluids (Fuel)/ Misc Hardware | 4.69 | **B40** | Lack of Service/ Maintenance | 1.28 |
| **B11** | Aircraft General Performance/ Capability | 0.43 | **B41** | Lateral/ Bank Control Parameters | 5.26 |
| **B12** | Aircraft Maximum Weight Capability | 3.56 | **B42** | Light Condition | 2.99 |
| **B13** | Aircraft Performance/ Control Parameters | 17.78 | **B43** | Operation Through High Denstiy Altitude | 4.55 |
| **B14** | Aircraft Structures Failure/ Improper Use | 3.84 | **B44** | Operation Through Icing Condition | 1.42 |
| **B15** | Aircraft Systems & Equipment | 0.57 | **B45** | Operation Through Turbulence | 1.42 |
| **B16** | Aircraft Takeoff Distance Performance/ Capability | 0.43 | **B46** | Operation Through Winds | 11.95 |
| **B17** | Aircraft Yaw Control Performance/ Capability | 1.85 | **B47** | Operation Under High Temperature | 0.43 |
| **B18** | Airport Facilities/ Design | 0.57 | **B48** | Organizational Issues | 0.85 |
| **B19** | Ceiling/ Visibility/ Precipitation | 4.84 | **B49** | Perception/ Orientation/ Illusion | 5.83 |
| **B20** | Damaged/ Degraded Aircraft Engine & Systems | 1.42 | **B50** | Physical Environment Obstruction | 5.83 |
| **B21** | Decision Making/ Judgement | 15.65 | **B51** | Pitch Control Parameters | 5.83 |
| **B22** | Directional Control Parameters | 4.55 | **B52** | Planning/ Preparation | 6.12 |
| **B23** | Experience/ Knowledge | 7.25 | **B53** | Psychological Factors | 4.84 |
| **B24** | Fatigue/ Wear/ Corrosion of Aircraft Engine & Systems | 1.14 | **B54** | Qualification Certification | 3.70 |
| **B25** | General Health/ Fitness | 2.70 | **B55** | Runway/ Land/ Takeoff/ Taxi surface | 1.14 |
| **B26** | Identification/ Recognition | 1.28 | **B56** | Terrain Physical Environment | 1.85 |
| **B27** | Impairment/ Incapacitation due to Health Issues | 1.71 | **B57** | Use of Alcohol/ Illicit Drugs | 1.56 |
| **B28** | Improper Aircraft Control | 68.14 | **B58** | Use of Equipment/ Information | 8.53 |
| **B29** | Improper Use Operation of Aircraft Fuel System | 1.14 | **B59** | Use of Medication Drugs | 3.56 |
| **B30** | Improper Use/ Operation of Aircraft Power Plant (Engine) | 3.56 | | | |

And the output space can be defined as follows:

y = {1, 2, …, C}, consisting of C possible output classes.

The input features, obtained through the data preprocessing steps, are shown in **Table 3.2** below. The % represents the proportion of entries being TRUE (1) in the respective feature column in the one-hot encoded dataset.

The output classes are derived from the LOC-I accident categories defined by IATA in the Loss of Control In-Flight Accident Analysis Report (Edition 2019) [18]. According to the report, LOC-I accidents in the ICL flight phase confenciated in these nine (9) broader categories as shown in **Table 3.3.**

**Table 3.3:** Output Classes for LOC-I Accidents

| Abb. | LOC-I (ICL) Accident Categories | Size |
|---|---|---|
| **C1** | Aircraft System Malfunction | 94 |
| **C2** | Operating Outside Aircraft Limitations | 34 |
| **C3** | Poor Manual Handling | 249 |
| **C4** | Poor, or Lack of Decision-Making | 94 |
| **C5** | Inadequate Crew Monitoring and Cross Checking | 18 |
| **C6** | Operating in Adverse Meteorological Conditions (AMC) | 26 |
| **C7** | Non-Adherence to Standard Operating Procedures (SOPs) | 42 |
| **C8** | Inadequate Implementation of Safety Management Systems (SMS) | 79 |
| **C9** | Incorrect Response to the Scenario Faced | 67 |

All the accidents were carefully analysed based on their findings and probable causes. After careful analyses, each accident was assigned the most suitable output class and put into one of the 9 above mentioned output classes. These individual accidents' interpretation was done between the researcher and the research supervisor.

## 3.3. Data Splitting

Machine learning algorithms partition the dataset into three smaller datasets for training, testing, and validation of the model. This segregation guarantees that the model being assessed is distinct from the one it is trained on. The training set is the most critical, constituting 80% of the overall dataset, while testing and validation sets typically each account for 10% [b]. The testing set evaluates the performance of the trained ML model on unseen data, providing insights into the accuracy of predictions. Although not mandatory, the validation set proves beneficial for hyperparameters tuning and model optimization.

## 3.4. Model Selection and Training

Random Forest (RF) is an ensemble learning method employing trees, utilized for both classification and regression tasks. It involves multiple decision trees trained on slightly varied datasets through techniques like feature bagging, bootstrap aggregating, and a voting mechanism.

### *3.4.1. Bagging and Bootstrapping*

Each decision tree involves conditions at each node, culminating in a final prediction at the leaf node, expressed mathematically as:

$$Xi \leq d \text{ or } Xi > e \tag{3.1}$$

where, Xi is a feature and d is a threshold.

In feature bagging, every decision tree randomly selects a unique set of features from the dataset, offering varied perspective of studying distinctive features and their correlations. Bootstrap aggregating involves drawing and replacing dataset entries, ensuring each decision tree has a unique dataset, thus adding larger scope and variety to the model. The voting mechanism aggregates predictions from all decision trees through a 'majority vote', reducing overfitting concerns associated when using a single decision tree.

### *3.4.2. Splitting Criterion*

The splitting criterion, a metric in RF, guides data division at each node for every decision tree in the ensemble. Three metrics, entropy, information gain, and Gini impurity, are employed for classification problem. Entropy measures dataset disorder mathematically expressed as:

$$\text{H(S)} = -\sum_{i=1}^{C} p_i \, log_2(p_i) \tag{3.2}$$

where,

S Set of data points with C classes

Pi Proportion of data points in class i

Information gain represents the reduction in entropy resulting from dataset splitting based on a particular feature.

$$\text{IG (S,V)} = \text{H(S)} - \sum_{i=1}^{k} \frac{|S_i|}{|S|} H(S_i) \tag{3.3}$$

Gini impurity, like entropy, quantifies disorder, with lower values indicating less disorder:

$$\text{Gini (S)} = 1 - \sum_{i=1}^{C} (p_i)^2 \tag{3.4}$$

The objective of these criteria is to select a node splitting that minimizes entropy and Gini impurity while maximizing information gain.

### *3.4.3. Hyperparameter Optimization*

Random Forest (RF) encompasses numerous hyperparameters crucial for tuning based on the dataset and analysis type. Tuning can be performed through Grid Search, where values are individually selected, or Randomized Search, involving providing a range for each hyperparameter and letting the algorithm choose the optimal set after a specified number of iterations.

The hyperparameters in this research comprise of n-estimators, max depth, min sample split, min samples leaf, max features, bootstrap, and random state. Their explanations are detailed in the **Table 3.4** below:

**Table 3.4:** Random Forest Hyperparameters

| **Hyperparameters** | **Explanation** |
|---|---|
| **n-estimators** | Total number of decision trees in the forest |
| **max depth** | The maximum depth of the tree |
| **min sample split** | Min sample split at an internal node of decision tree |
| **min sample leaf** | Min samples at the leaf node |
| **max features** | Maximum number of features at the best split |
| **bootstrap** | Randomizing the samples of each decision tree |
| **random state** | Number controlling the model's randomness |

### *3.4.4. Random Forest Classifier*

The RF classifier is executed in 'Jupyter Notebook' using Python as the kernel. This utilizes the optimal set of hyperparameters chosen through randomized search. Once the model is trained, it undergoes evaluation based on accuracy with the test set and validation set.

### 3.5. Model Testing and Validation

The final step in RF modelling is to access the accuracy by applying the test set and validation tests. Once a satisfactory value is achieved, it can be concluded that the model is prepared for future data usage. If the accuracy falls outside the acceptable range, additional steps must be taken in data preprocessing and feature engineering to enhance the data's suitability for RF modelling.

### 3.6. HFACS Modified Framework

The aviation accidents are mostly due to combination of environmental impact, aircraft malfunction and personnel's errors rather than a single reason. This makes it difficult to assign accidents directly to one of the HFACS levels/ sub-levels. Hence, rather than linking the whole accident to HFACS, the input features of the ML model are linked to HFACS levels to design a modified framework. These frameworks cover the levels of organizational influences, unsafe supervision, preconditions to unsafe acts, unsafe acts, and external factors. The modified frameworks are shown in **Figure 3.2, 3.3 & 3.4.**

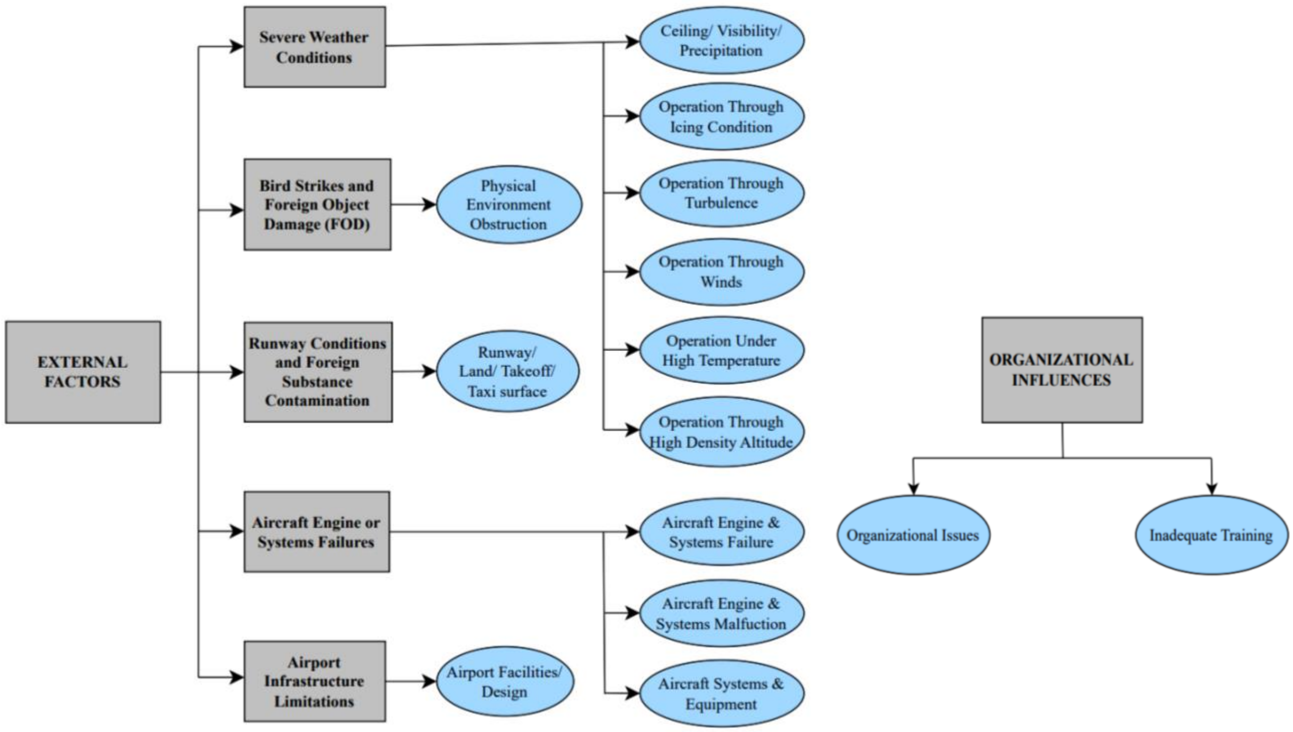


**Figure 3.2:** Input Features corresponding to External Factors & Organizational Influences

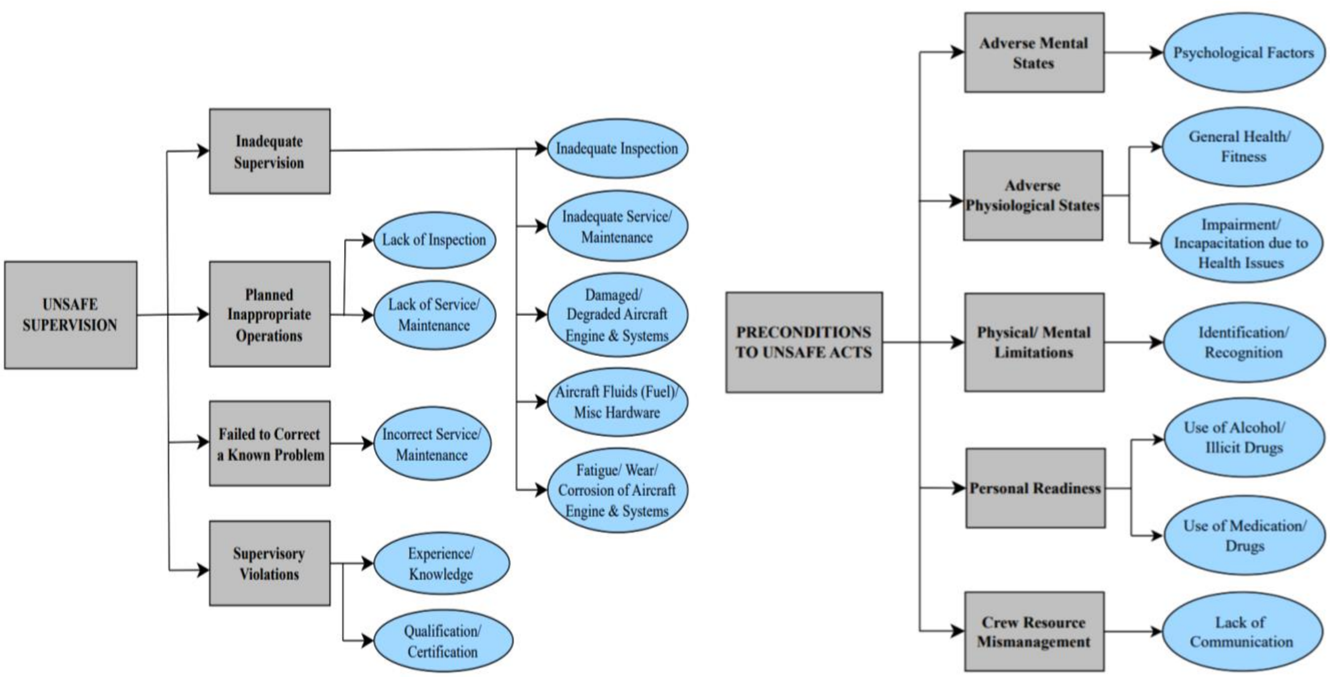


**Figure 3.3:** Input Features corresponding to Unsafe Supervision & Preconditions to Unsafe Acts

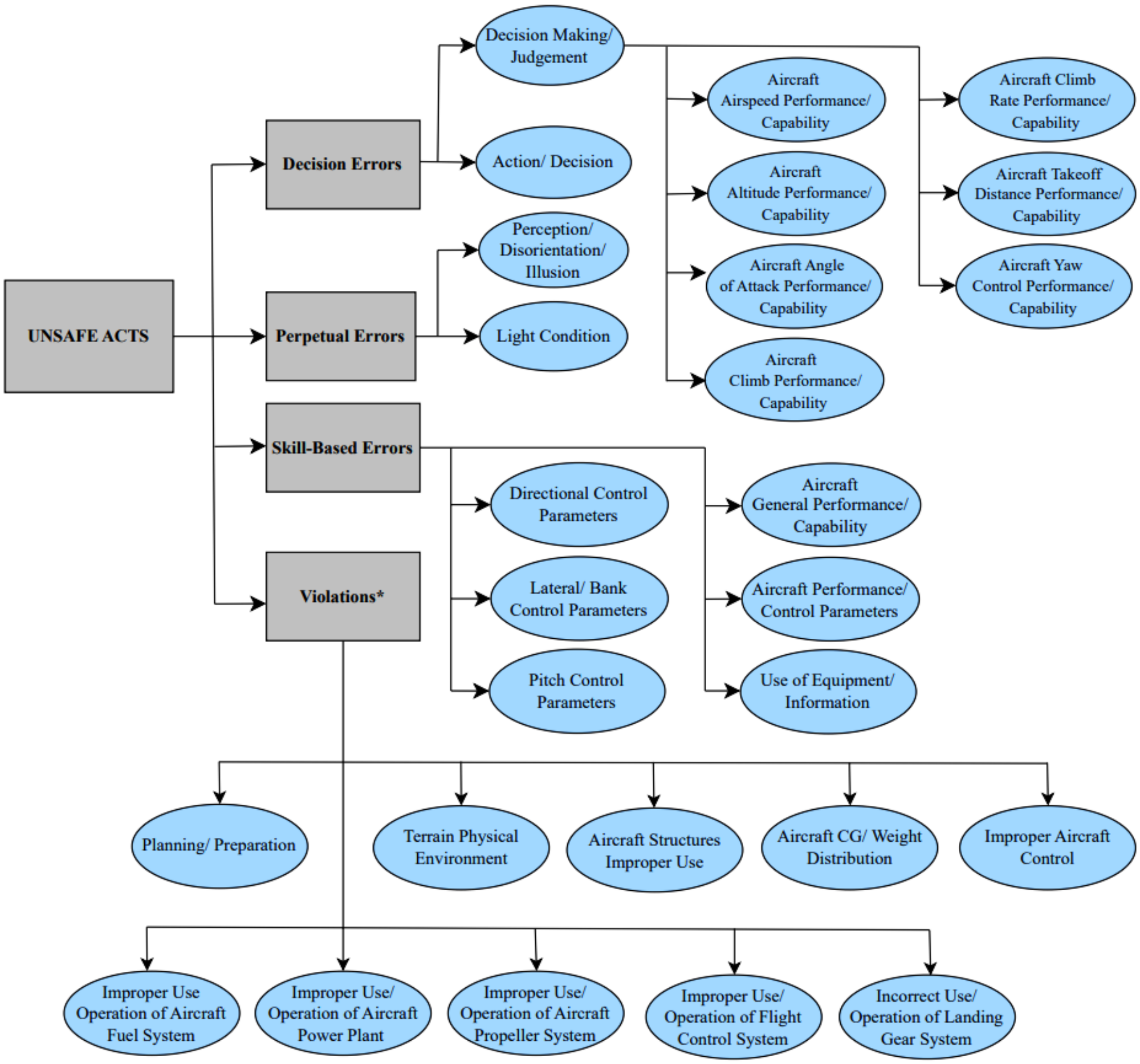


**Figure 3.4:** Input Features corresponding to Unsafe Acts - HFACS

# CHAPTER 4: RESULTS & DISCUSSION

Random Forest Classifier was used on a dataset of 2241 accidents, having 59 input features and 9 output classes. The best-fitted RF model hyperparameters obtained through randomized search are shown in **Table 4.1**. A range of values were given to each of these hyperparameters and run on the model to get the best values for the highest accuracy. The relationship between these hyperparameters and their effect on accuracy of predictions can be depicted into a surface 3D plot as shown in **Figure 4.1**. X-axis has n-estimators (number of Decision Trees), y-axis has the max-depth (maximum depth of tree) and z-axis has accuracy of the model.

**Table 4.1:** Best Hyperparameters from Randomized Search

| Hyperparameters | Best Randomized Search Values |
|---|---|
| **n-estimators** | 100 |
| **max depth** | 50 |
| **min sample split** | 2 |
| **min sample leaf** | 1 |
| **max features** | log2 |

The overall RF model accuracy was 85% and 88% on the validation set and test set respectively. This overall classification accuracy is a general performance metric which is a ratio of correct predictions to the total number of predictions. This method is usually dominant class biased and can be misleading in a lot of ways. So, to normalize the results and to give equal importance to all the classes including the rare-occurring ones, different performance metrics are used to evaluate the RF model.

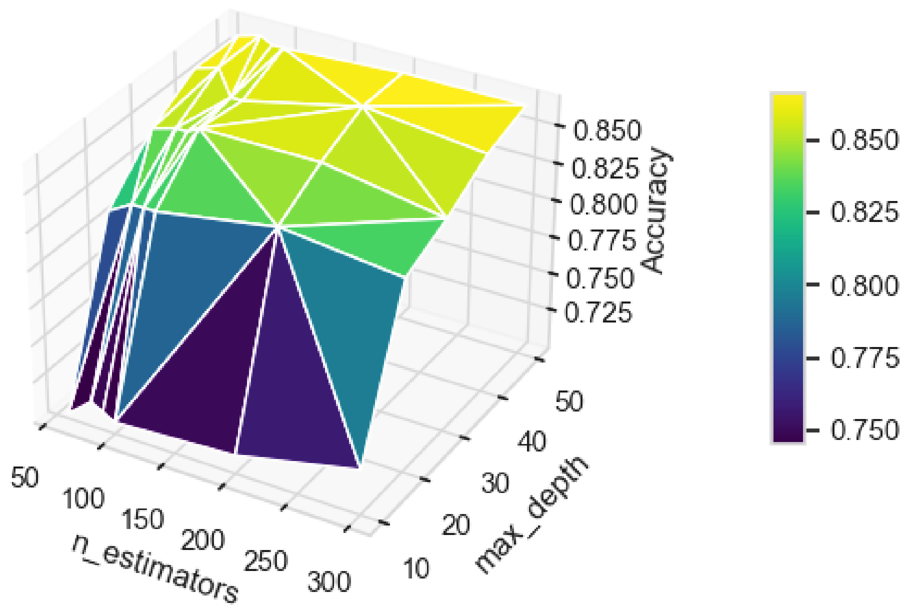


**Figure 4.1:** Surface 3D Plot of RF Hyperparameters

## 4.1. Performance Metrics

A classification report is generated from the Random Forest model that shows the overall performance of each of the output classes in the prediction model. The performance metrics evaluated in this report are the precision, recall, and F1-score.

$$\text{Precision} = \frac{\text{TP}}{(\text{TP+TN})} \tag{4.1}$$

$$\text{Recall} = \frac{\text{TP}}{(\text{TP+FN})} \tag{4.2}$$

$$\text{Accuracy} = \frac{\text{TP+TN}}{(\text{TP+FN+FP +TN})} \tag{4.3}$$

$$\text{F1-Score} = 2 \text{ x } \frac{\text{Precision x Recall}}{(\text{Precison+Recall})} \tag{4.4}$$

where,

TP True Positives

TN True Negatives

FP False Positives

FN False Negatives

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| 0 | 0.79 | 1.00 | 0.89 | 27 |
| 1 | 0.96 | 1.00 | 0.98 | 22 |
| 2 | 0.90 | 0.60 | 0.72 | 30 |
| 3 | 0.87 | 0.90 | 0.88 | 29 |
| 4 | 1.00 | 0.91 | 0.95 | 22 |
| 5 | 0.84 | 1.00 | 0.92 | 27 |
| 6 | 0.93 | 1.00 | 0.96 | 25 |
| 7 | 0.91 | 0.84 | 0.87 | 25 |
| 8 | 0.81 | 0.72 | 0.76 | 18 |
| accuracy | | | 0.88 | 225 |
| macro avg | 0.89 | 0.89 | 0.88 | 225 |
| weighted avg | 0.89 | 0.88 | 0.88 | 225 |

**Figure 4.2:** Classification Report Generated from the RF Model

It can be seen from **Figure 4.2** that class 4 has the highest precision, classes 0, 1, 5, 6 has the highest recall and class 1 has the highest F1-score out of the 225 cases in the test set. The highest weighted average accuracy is 89% which shows the overall good performance of the classification model.

## 4.2. RF Model Performance

As the data was split into three datasets of training, validation and testing, the performance of the model was checked from the validation set in the form of confusion matrix (**Figure 4.3**). It breaks down each class's predictions against all output classes in the form of number of accidents of the validation set predicted for that class. Each row in the matrix corresponds to the actual output class while the columns represent the predicted classes. Here the main diagonal of the matrix represents the true positives of each class. For instance, Actual 0 refers to the actual class 1 'Aircraft System Malfunction' and all the 27 actual output cases were predicted correctly, and the other columns have 0 instances.

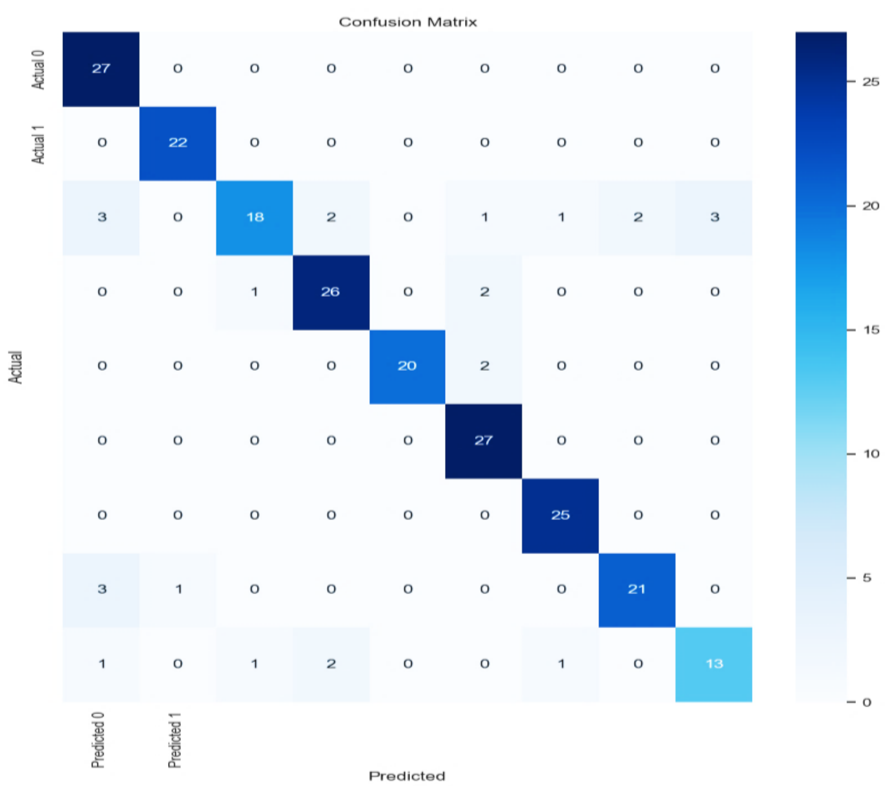


**Figure 4.3:** Confusion Matrix for Output Classes

**Table 4.2:** Confusion Matrix Percentages Table

| | **0** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** |
|---|---|---|---|---|---|---|---|---|---|
| **0** | 12 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **1** | 0 | 9.78 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **2** | 1.33 | 0 | 8 | 0.89 | 0 | 0.44 | 0.44 | 0.89 | 1.33 |
| **3** | 0 | 0 | 0.44 | 11.56 | 0 | 0.89 | 0 | 0 | 0 |
| **4** | 0 | 0 | 0 | 0 | 8.89 | 0.89 | 0 | 0 | 0 |
| **5** | 0 | 0 | 0 | 0 | 0 | 12 | 0 | 0 | 0 |
| **6** | 0 | 0 | 0 | 0 | 0 | 0 | 11.11 | 0 | 0 |
| **7** | 1.33 | 0.44 | 0 | 0 | 0 | 0 | 0 | 9.33 | 0 |
| **8** | 0.44 | 0 | 0.44 | 0.89 | 0 | 0 | 0.44 | 0 | 5.78 |

When the values in the confusion matrix are obtained as a percentage of the total instances, a confusion matrix table is obtained. The non-diagonal values show the percentage of incorrectly predicted classes and help in understanding and improving the dataset better. An extremely low percentage of incorrectly predicted classes can be seen scattered in **Table 4.2.**

## 4.3. Human Factor Analysis, HFACS

All the input features are linked to the HFACS levels in the modified framework as shown in the 'methodology' section of this study. Few of the features do not fall directly under the human factors definition and hence are assigned to external factors. After finding the model's performance in different perspectives, the next step is to find the promising human factors prevailing in the dataset. For that purpose, feature importance is performed.

### *4.3.1. Feature Importance*

Feature importance is performed on the RF model to obtain the top 10 features that are the most influential in making predictions. It is calculated based on average impurity reduction of a specific feature during the formation of decision trees in the forests. The more the impurity reduction, the more the feature importance. It also considers the number of times a feature is used in splitting the data, which also increases its importance. This ultimately helps to align our concentration to the most impactful features while feature selection. If there are features with very less feature importance value, these can be removed from the dataset to improve the overall accuracy of the model. **Figure 4.4** shows the top ten features in making predictions in the descending order with the topmost holding the highest importance.

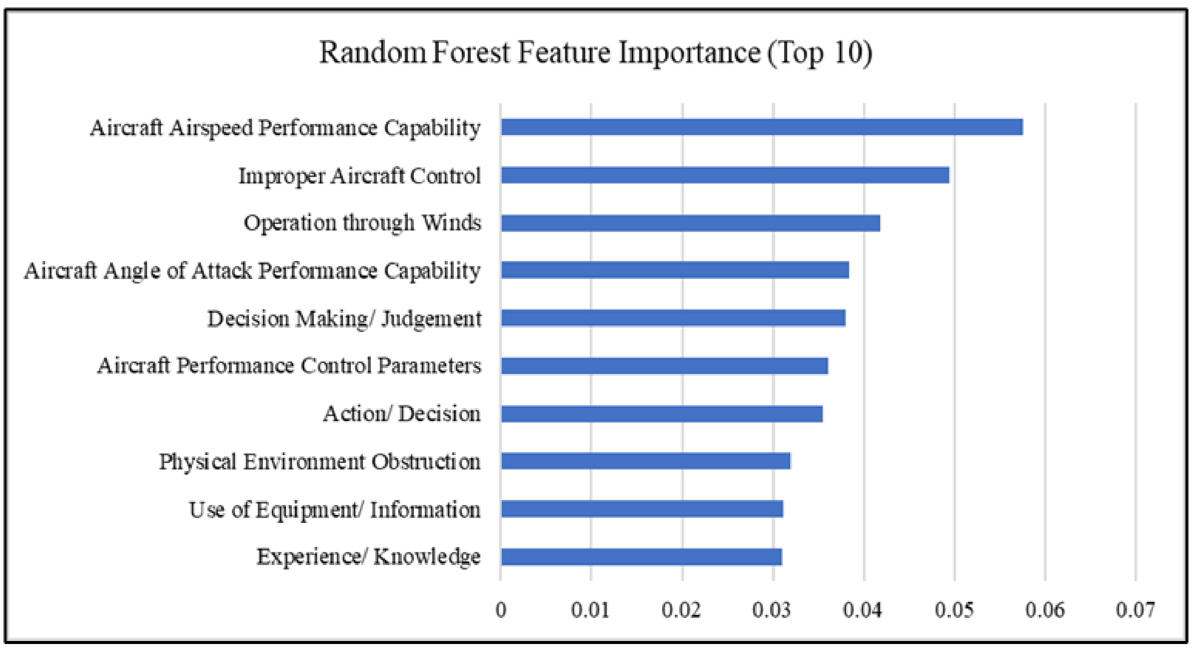


**Figure 4.4:** Feature Importance (Top 10 Features)

The important features listed above are linked back to the HFACS framework shown in 'methodology' section. **Table 4.3** lists the top 10 important features and their corresponding HFACS levels and sub-levels.

**Table 4.3:** HFACS Levels and Sub-Levels of Top 10 Most Important Features

| S. No. | Top 10 Important Features | HFACS Level | HFACS Sub-Level |
|---|---|---|---|
| **1** | Aircraft Airspeed Performance Capability | Unsafe Acts | Decision Errors |
| **2** | Improper Aircraft Control | Unsafe Acts | Violations |
| **3** | Operation through Winds | External Factors | Severe Weather Conditions |
| **4** | Aircraft Angle of Attack Performance Capability | Unsafe Acts | Decision Errors |
| **5** | Decision Making/ Judgement | Unsafe Acts | Decision Errors |
| **6** | Aircraft Performance Control Parameters | Unsafe Acts | Skill-Based Errors |
| **7** | Action/ Decision | Unsafe Acts | Decision Errors |
| **8** | Physical Environment Obstruction | External Factors | Bird Strike and Foreign Object Damage |
| **9** | Use of Equipment/ Information | Unsafe Acts | Skill-Based Errors |
| **10** | Experience/ Knowledge | Unsafe Supervision | Supervisory Violations |

The blue highlighted rows are the features that come under HFACS levels while the grey ones are the external factors. Many of these features correspond to 'Unsafe Acts' level and one comes under 'Unsafe Supervision'. It is a crucial step as it highlights the most important HFACS levels and sublevels that contribute most to the predictions of GA LOC-I accidents in ICL flight phase. This is the crux of our study, and hence it is explored well for better understanding.

### *4.3.2. Pearson's Correlation Matrix*

Now that we have the most important human factors contributing to accidents of our scope, relationships among these features are found through Pearson's Correlation Matrix. It is a method of establishing a linear relationship of one feature with another. It has values ranging from -1 to 1, while -1 means a complete negative relationship and vice versa. The main diagonal has value of 1, which is the highest for each class against itself. **Figure 4.5** shows a correlation matrix of 11 features. These features are all HF-based and are obtained from feature importance list (top 20) of the model, excluding the features relating to external factors.

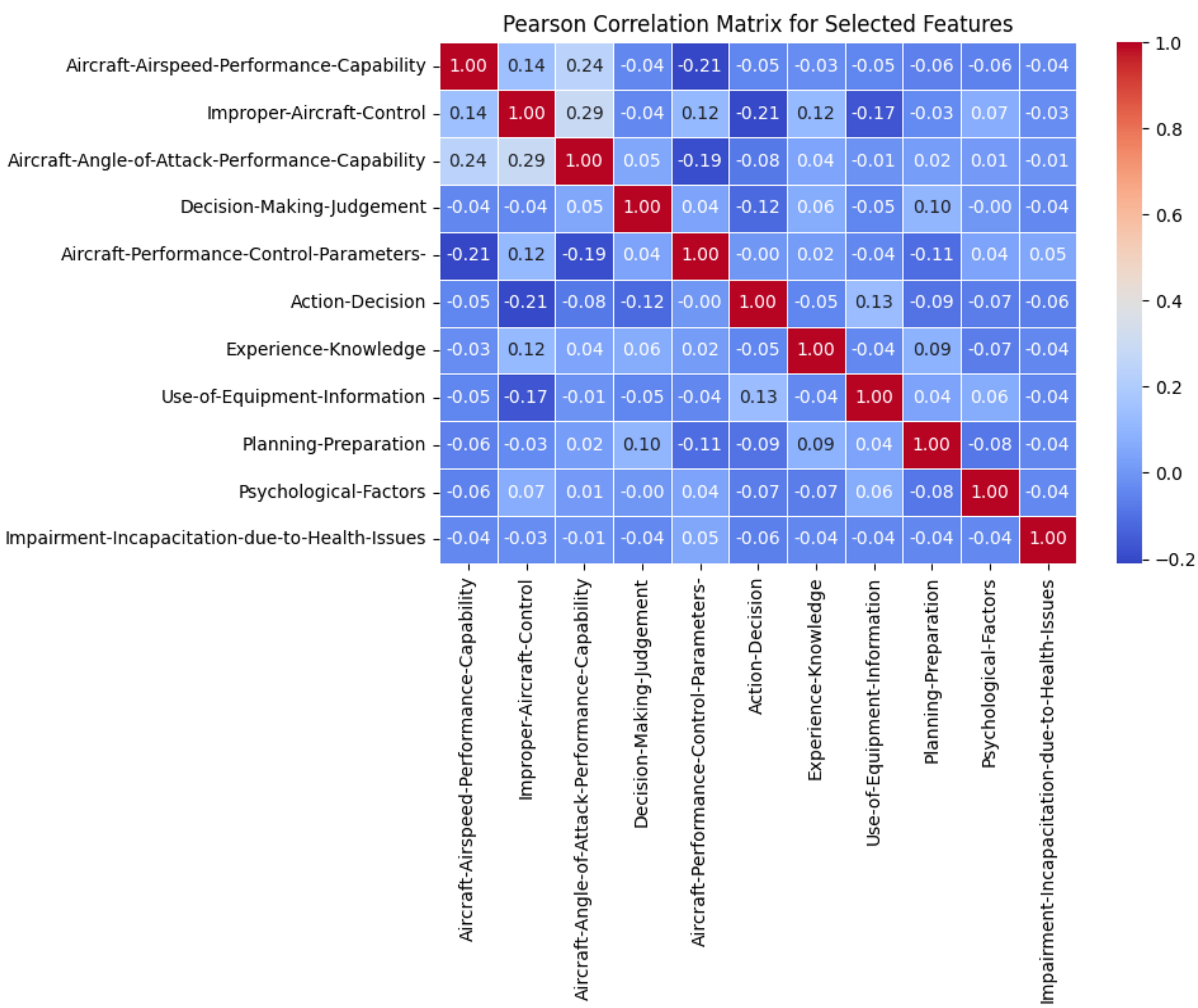


**Figure 4.5:** Pearson's Correlation Matrix for Top 11 HFACS Features of the Trained RF-Model

The non-diagonal values are on the lower side and even negative. This indicates a weak correlation among HF-based features. It might seem discouraging at first, but there are positive insights to this analysis.

- The absence of strong relationships among HF-based features shows how aviation accidents are caused by numerous numbers of triggering factors. This calls for a holistic safety approach which is focused on multiple aspects rather than one.

- Aviation accidents are not caused by a single HF-based factor, but a combination of factors. Optimistically, we can say that different levels of hierarchy in the aviation industry need to work together for safety improvements. There is a dire need to continuously improve the safety strategies and to adapt to the future unpredictability.

- This calls for better and detailed training programs for the aviation personnels so that a wide range of skills is obtained to mitigate the risks of human factors.

- Instead of implementing a solution based solely on a specific correlation between two human factors (strong correlation), it is better to tailor a customized intervention can be introduced based on the unique combination of many human factors (weak correlation) of the accidents.

- Such a weak correlation among the human factors calls for better communication among the personnels of aviation sectors. There should be effective collaboration so that they can work on identifying and addressing the multiple human-related errors and violations together.

### 4.3.3. *SHAP (Shapley Additive exPlanations) Analysis*

SHAP analysis is performed on the validation set of RF model to get the average impact of each feature on the overall output of the model. It works by considering marginal contribution of each feature against all the possible feature combinations. **Figure 4.6** shows the impact of each feature on the overall as well as each output class's prediction. The features are listed from top to bottom according to the absolute average SHAP values while the length of the bar shows the impact magnitude of that feature to the model's prediction accuracy.

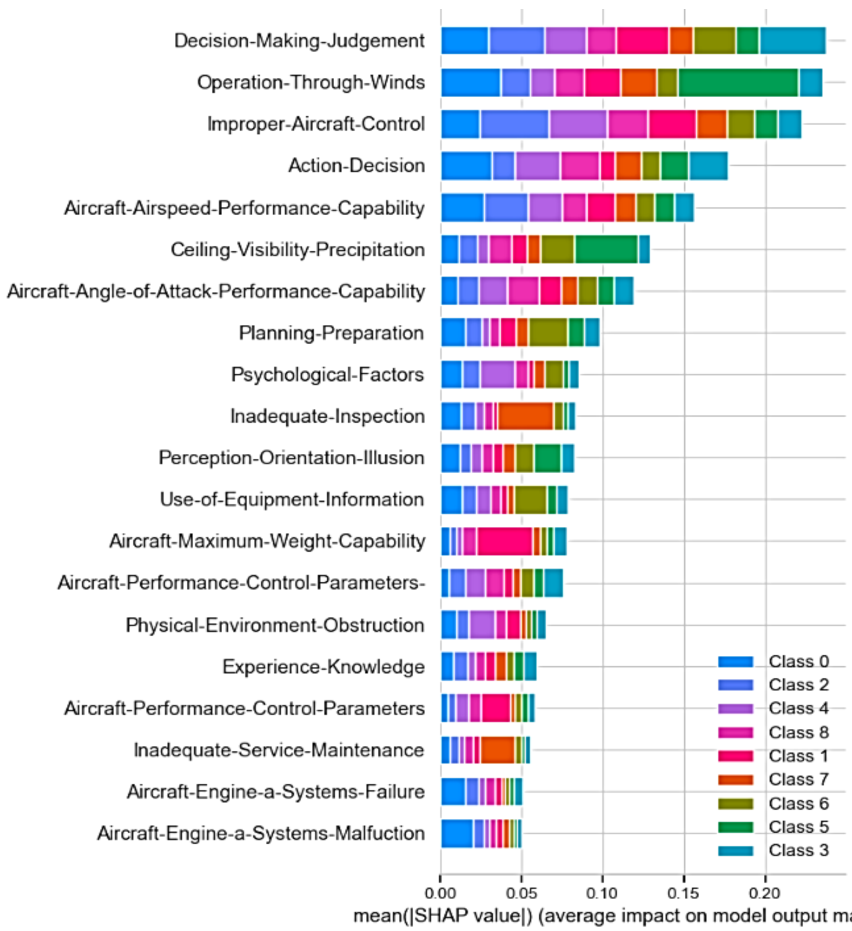


**Figure 4.6:** Mean SHAP Value Graph

The output classes are distributed among the important SHAP features. This helps in identifying the impact of a specific feature on each of the output classes. To further dive into the feature-class relationship specific to HFACS, another SHAP analysis of interaction plots is performed. For this analysis, 7 out of the 9 output classes were used. It is because 'Aircraft System Malfunction' (C1) and 'Operating in Adverse Meteorological Conditions (AMC)' (C6) are non-HF accident categories. Furthermore, HF-related features from **Figure 4.6** are separated into a list and were utilized into getting interaction plots as shown in **Figure 4.7**, against each of the 7 HF-related classes.

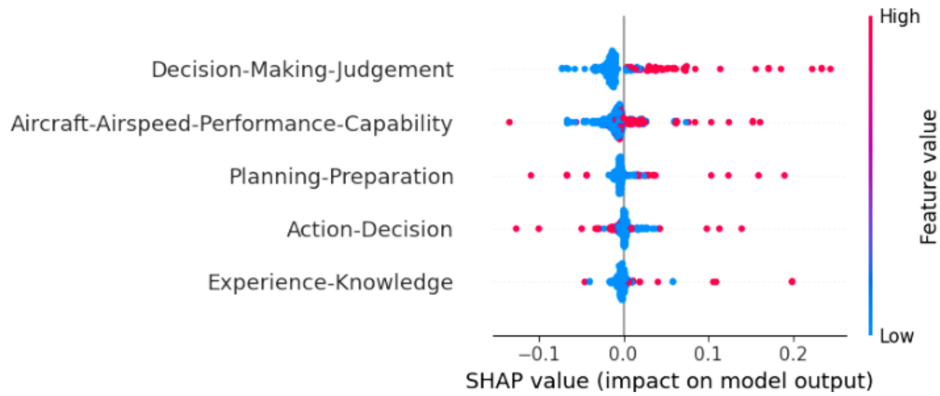


**Figure 4.7:** Output Class-C2, Operating Outside Aircraft Limitations

The blue and red dots against the input features represent the number of accidents falling under that feature for that specific output class, while red color shows higher positive impact on the Quadrant-I of the graph while blue color shows lower negative impact in the Quadrant-IV and vice versa. From **Figure 4.7**, decision making, and judgement of the pilot highly influences the capability of him flying the aircraft inside its set limitations. Similarly, improper planning of flight and inexperience of the pilot also contribute a lot to the accidents under the output class C2.

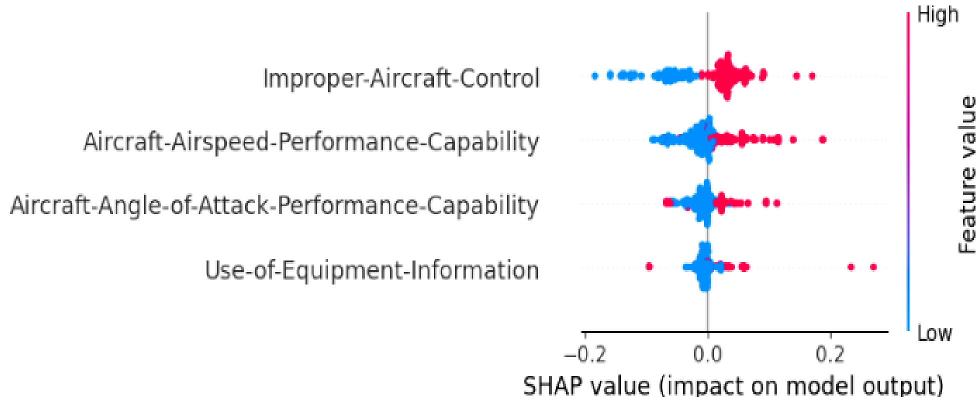


**Figure 4.8:** Output Class-C3: Poor Manual Handling

Poor manual handling of the aircraft flight resulted majorly from the improper aircraft control by the pilot. Improper use of flight controls, errors in aircraft handling and failure to get stability at unusual altitudes leads to accidents falling under output class C3.

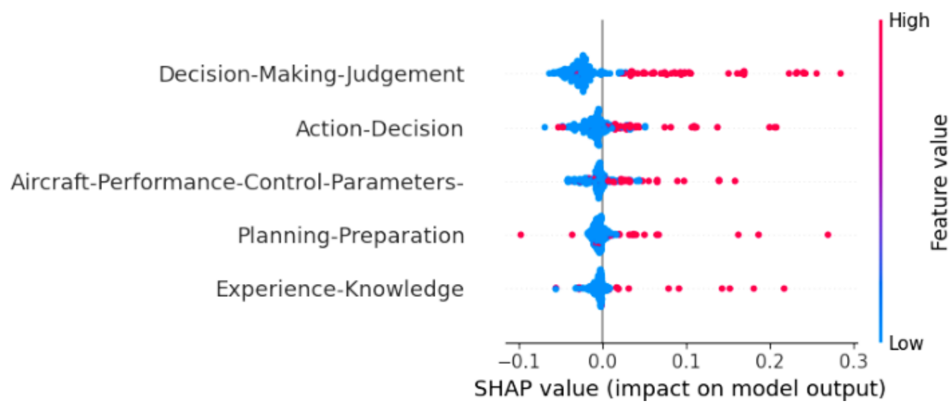


**Figure 4.9:** Output Class-C4: Poor, or Lack of Decision-Making

A high accident occurring class of poor decision making has a lot of contributing factors, while the actions and decisions of the pilots tops the list. Poor decision making and incorrect action performance during the initial climb phase of flight results in the loss of control in flight of the aircraft. Not having a proper preparation for the unexpected scenarios during the flight and being inexperienced to make best decisions could lead to loss of control in flight.

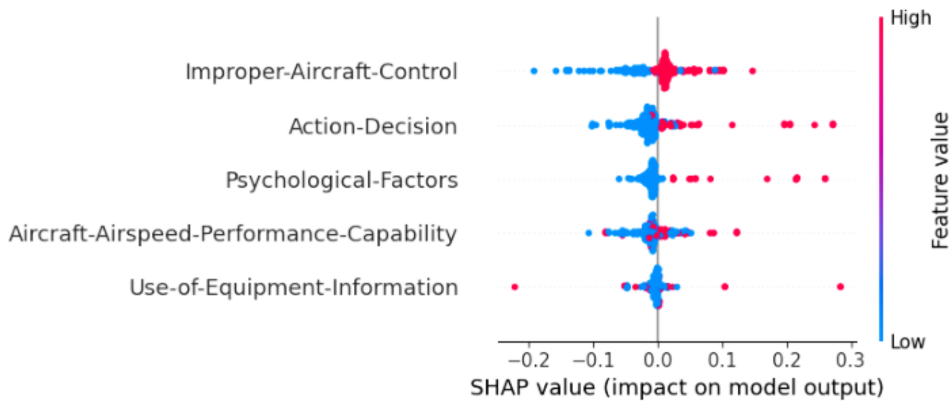


**Figure 4.10:** Output Class C-5, Inadequate Crew Monitoring & Cross Checking

Improper aircraft control affects adversely the crew monitoring and cross checking as it can easily divert attention of the crew members on the instability of the aircraft. Decision making among the crew members during difficult scenarios could be problematic as different members could have different opinions and psychological states of mind.

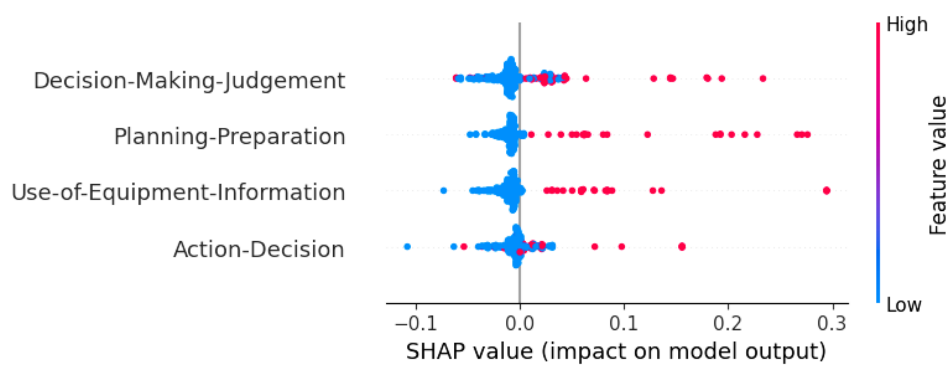


**Figure 4.11:** Output Class-C7: Non-Adherence to SOPs

If the crew members of the aircraft are making decisions that does not follow the established protocols of the flight such as SOPS regarding specific speed, altitude, or angle of attack during a specific encounter, this could lead to loss of control inflight. In addition to that, crew members should be well prepared and should correctly use equipment according to their respective use. For example, not utilizing the navigational aids according to SOPs could result in disastrous accidents.

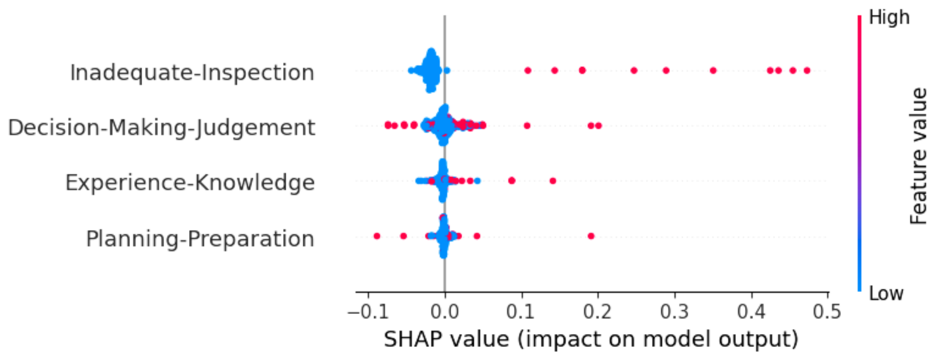


**Figure 4.12:** Output Class-C8: Inadequate Implementation of SMS

Safety management during the flight is the foremost priority. Failure in identifying any underlying issues related to the aircraft engine or flight procedures during inspection could lead to unsafe flight operations often resulting in accidents. Pilot in command is also expected to put safety first in all his flight decisions and pre-flight preparations.

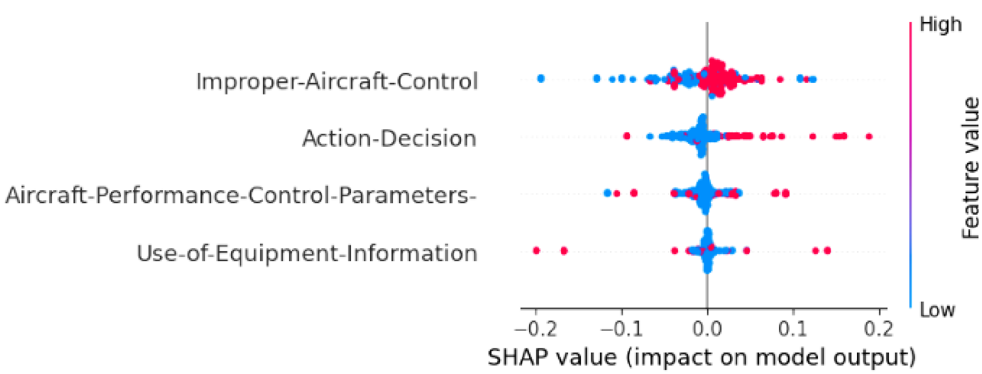


**Figure 4.13:** Output Class-C9: Incorrect Response to the Scenario Faced

Pilots often encounter situations where a very abrupt action is needed to keep the flight operation safe and smooth. Miscalculations in the actions and improper handling of the aircraft while being exposed to adverse conditions lead to loss of control in flight and ultimately cause accidents.

# CHAPTER 5: CONCLUSIONS

General Aviation accidents are caused by a wide set of features covering all the human, environmental and aircraft mechanical factors. No accident can be contributed solely to one contributing factor as different factors, either active or latent, contribute to the unsafe act. Instead of running raw or a little preprocessed dataset on the machine learning algorithm, it is extremely time saving to study the dataset in detail and apply different techniques (e.g. feature engineering, NLP, etc.) to make it suitable for the ML algorithm before training. This will not only increase the prediction accuracy drastically, but also will help in getting purified and error-free outcomes i.e., contributing factors.

Random forest model shows an accuracy of 85% on the test set, and 88% accuracy on the validation set. This concludes that the preprocessed dataset and the model training were compatible, and the model can be used for further predictions. 80% of the top 10 most important contributing factors obtained from the RF model are human-related. Out of which 50% are Decision Errors, 25% are Skill-Based Errors, and 25% are Violations. All of these human factors correspond to the HFACS level of Unsafe Acts.

The Person's correlation matrix shows a weak relationship among the most important human factors. This calls for a holistic safety approach which focuses on multiple aspects. Different hierarchy levels need to work together to have effective collaboration and communication amongst themselves.

SHAP interaction plots of the human-related output classes showed that the most concerning human factors are 'Action/ Decision', 'Decision Making/ Judgement', Planning/ Preparation', 'Experience/ Knowledge', and 'Improper Aircraft Control'. 'Action/ Decision' factor contributes majorly to the accident category of 'Poor or Lack of Decision Making', and 'Non-Adherence to SOPs'. Poor execution of decisions or taking inappropriate actions during the initial climb phase can contribute to the loss of control. Similarly, if pilots make decisions that deviate from established procedures or take actions contrary to SOPs, it can lead to non-compliance and increase the risk of accidents.

'Decision making/ judgement' majorly contributes to the accident category of 'Operation Outside Aircraft Limitations'. Pilots may make decisions that lead to operating the aircraft beyond its designated limitations, possibly due to misinterpretation of information or inadequate decision-making skills. 'Planning/ preparation' contributes to the accident categories of 'Non-Adherence to SOPs' and 'Operation Outside Aircraft Limitations'. If there's a deficiency in planning or preparation, pilots may not be well-equipped to follow the established procedures during the initial climb, potentially leading to deviations from SOPs.

'Improper Aircraft Control' contributes majorly to the accident categories of 'Poor Manual Handling' and 'Incorrect Response to the Scenario Faced'. It could imply instances where pilots are not effectively managing the aircraft, leading to a loss of control. This could be due to factors such as inadequate pilot training, lack of situational awareness, or failure to respond appropriately to changing conditions. 'Experience/ Knowledge' highly contributes to the accident category of 'Inadequate Implementation of SMS'. Inexperienced individuals or those lacking specific knowledge related to safety procedures might inadvertently neglect aspects of the Safety Management System, leading to inadequate implementation.

## FUTURE RECOMMENDATIONS

In future works, the focus should be shifted from the Reactive Safety Approach towards the Proactive safety approach. All hierarchy levels of aviation industry must undergo an evaluation and improvement training programs to lessen the human- related features causing accidents. From airport facilities to maintenance and pilots/ staff's psychological and skill-based training, everything should be improved based on the results. The organizations should be critical about choosing the individuals based on experience and knowledge and keep in check their periodic performance.

# REFERENCES


[1] "Aviation deaths rise worldwide in 2020 even as fatal incidents, flights fall", International News Organization "Reuters."

Aviation deaths rise worldwide in 2020 even as fatal incidents, flights fall | Reuters

[2] "A timeline of major air crashes in Pakistan", Dawn News.

A timeline of major air crashes in Pakistan - Pakistan - DAWN.COM

[3] "Accident of PIA Flight Pk8303 Airbus A320-214 Reg No AP-BLD Crashed Near Karachi Airport On 22-05-2020", Pakistan Civil Aviation Authority.

AAIB-431.pdf (caapakistan.com.pk)

[4] Mason, C.D. (1993). *Personality profile of pilots compared to nonflying civilians. A summary of findings*. Unpublished manuscript, supplied by the author.

[5] Murray, S.R. (1997). *Deliberate decision making by aircraft pilots: A simple reminder to avoid decision making under panic*. The International Journal of Aviation Psychology, 7, 83-100

[6] Trollip, S.R., & Jensen, R.S. (1991). *Human factors for general aviation*. Englewood, CO: Jeppesen Sanderson.

[7] Heinrich, H.W., Petersen, D., & Roos, N. (1980). *Industrial accident prevention: A safety management approach* (5th ed.). New York: McGraw-Hill

[8] Reason, J. (1990). *Human error*. New York: Cambridge University Press

[9] Shappell, S.A. and Wiegmann D.A. (1997). A human error approach to accident investigation: The taxonomy of unsafe operations. *The International Journal of Aviation Psychology, 7*, 269-91.

[10] Bird, F. (1974). *Management guide to loss control*. Atlanta, GA: Institute Press.

[11] Diehl. A. (1991). The effectiveness of training programs for preventing aircrew "error." In R.S. Jensen, *Proceedings of the 6th Symposium on Aviation Psychology (Vol. 2, pp. 640-655).* Columbus: Ohio State University.

[12] ICAO Aviation Occurrence Categories (Oct, 2013)

OccurrenceCategoryDefinitions.pdf (ntsb.gov)

[13] ICAO *"Analysis Loss of Control Inflight (LOC-I) Accidents"*

PowerPoint Presentation (icao.int)

[14] General Aviation Joint Safety Committee

General Aviation Joint Safety Committee – Partners in Aviation Safety (gajsc.org)

[15] Federal Aviation Administration. (2011). GAJSC accident data set. Retrieved from http://www.aabi.aero/News&Calendar/2011ppts/GA%20Add%20on.pdf

[16] General Aviation Joint Steering Committee. (2012). Loss of control work group: Approach and landing. Retrieved from *130327safety-committee.pdf (aopa.org)

[17] Commercial Aviation Safety Team (CAST), *"JSAT Loss of Control," CAST Approved Final Report, 2000*

[18] International Air Transport Association (IATA), *"Loss of Control In-Flight Accident Analysis Report Edition 2019"* *Loss of Control In-Flight Accident Analysis Report 2019 Edition (iata.org)

[19] ICAO "*Phase of Flight (April 2013)"*

Microsoft Word - DLV Phase of Flight Definitions English v1.3.doc (ntsb.gov)

[20] FAA "*Airplane Flying Handbook"*

Airplane Flying Handbook (FAA-H-8083-3B) Chapter 5

[21] Labib, A., & Read, M. (2013). Not just rearranging the deckchairs on the Titanic: Learning from failures through Risk and Reliability Analysis. *Safety Science*, *51*(1). https://doi.org/10.1016/j.ssci.2012.08.014

[22] Sitkin, S. B. (1992). LEARNING THROUGH FAILURE: THE STRATEGY OF SMALL LOSSES. In *Research in organizational behavior* (Vol. 14).

[23] Edmondson, A. C. (2011). Strategies of learning from failure. *Harvard Business Review*, *89*(4).

[24] Cannon, M. D., & Edmondson, A. C. (2005). Failing to learn and learning to fail (intelligently): How great organizations put failure to work to innovate and improve. *Long Range Planning*, *38*(3 SPEC. ISS.). https://doi.org/10.1016/j.lrp.2005.04.005

[25] Taleb, N. N. (2007). The Black Swan: The Impact of the Highly Improbable (Random House, 2007). In *The Review of Austrian Economics* (Vol. 21, Issue 4).

[26] Heinrich, Herbert William. 1931. *Industrial accident prevention: A scientific approach*. New York, McGraw-Hill.

[27] Poor Sabet, P. G. (2013). Application of Domino Theory to Justify and Prevent Accident Occurance in Construction Sites. *IOSR Journal of Mechanical and Civil Engineering*, *6*(2). https://doi.org/10.9790/1684-0627276

[28] Blackett, Claire. (2006). Combining Accident Analysis Techniques for Organisational Safety.

[29] Johnson, C., 2003. Failure in Safety Critical Systems: A Handbook of Incident and Accident Reporting Glasgow University Press

[30] The HFACS Framework; the Official HFACS website.

HFACS, Inc | History

[31] Allnutt, M. (1982). Human factors: Basic principles. In R. Hurst & L.R. Hurst (Eds.). *Pilot errors: The human Factors* (pp. 1-21). St. Albans, England: Granada.

[32] Aircraft Accident Investigation into M/s Bhoja Air Flight BHO-213, Boeing 737-236A, REG # AP-BKC crashed on 20th April, 2012 near BBIAP, Islamabad.

Microsoft Word - Bhoja Air Final Report CAA Website 200115 (caapakistan.com.pk)

[33] Li, G., & Baker, S. P. (2007). *Crash risk in general aviation*. JAMA, 297, 1596–1598. https://doi.org/10.1001/jama.297.14.1596

[34] Detwiler, C., Holcomb, L., Hackworth, C., & Shappell, S. (2008). *Understanding the human factors associated with visual flight rules flight into instrumental meteorological conditions* (DOT/FAA/AM, 8/12). Federal Aviation Administration

[35] Aircraft Owners and Pilots Association (AOPA). (2019). *28th Joseph T. Nall report.* https://www.aopa.org/training-and-safety/air-safety-institute/accidentanalysis/joseph-t-nall-report

[36] Goh, J., & Wiegmann, D. (2001). An investigation of the factors that contribute to pilots' decisions to continue visual flight rules into adverse weather. Proceedings of the Human Factors and Ergonomics Society, Annual Meeting, 1, 26-29. https://doi.org/10.1177/154193120104500205

[37] Bradley S. Baugh (2020), Predicting General Aviation Accidents Using Machine Learning Algorithms, PhD Thesis, Embry-Riddle Aeronautical University commons.erau.edu/cgi/viewcontent.cgi?article=1548&context=edt

[38] Unmanned Aircraft System Traffic Management (UTM), Federal Aviation Administration. Unmanned Aircraft System Traffic Management (UTM) | Federal Aviation Administration (faa.gov)

[39] Lin, C. E., Shao, P. C., & Lin, Y. Y. (2020). System operation of regional UTM in Taiwan. *Aerospace*, *7*(5). https://doi.org/10.3390/AEROSPACE7050065

[40] FAA Guide to Low-Flying Aircraft, FAA Guide to Low Flying Aircraft

[41] Jacobson, S.R. 'Airplane Loss of Control Causal Factors and Mitigation Challenges,' AIAA Guidance, navigation and control conference, 2010. https://doi.org/10.2514/6.2010-8007

[42] Belcastro, C., Groff, L., Newman, R., Foster, J., Crider, D. and Klyde, D., Preliminary Analysis of Aircraft Loss of Control Accidents: Worst Case Precursor Combinations and Temporal Sequencing. (AIAA 2014-0612), 2014 Guidance, Navigaion and Control Conference, National Harbor, Maryland, USA, 13-17 January 2014. https://doi.org/10.2514/6.2014-0612

[43] IATA, Loss of Control In-Flight Accident Analysis Report 2010-2014. 1st ed., International Air Transport Association, Montreal, Canada, 2015

[44] Madeira, T., Melício, R., Valério, D., & Santos, L. (2021). Machine learning and natural language processing for prediction of human factors in aviation incident reports. *Aerospace*, *8*(2). https://doi.org/10.3390/aerospace8020047

[45] Martin, J. D. (1999). The first United States Army aircraft accident report. Washington, DC: Department of the Army. Retrieved on March 11, 2020, from https://apps.dtic.mil/docs/citations/ADA382312

[46] Reason, J. (2000b). Safety paradoxes and the safety culture. Injury Control and Safety Promotion, 7(1), 3-14. https://doi.org/10.1076/1566-0974(200003)7:1;1-V;FT003

[47] Azadeh, A., & Zarrin, M. (2016). An intelligent framework for productivity assessment and analysis of human resource from resilience engineering, motivational factors, HSE and ergonomics perspectives. *Safety Science*, *89*. https://doi.org/10.1016/j.ssci.2016.06.001

[48] Textron. (2018). 2018 corporate responsibility report: Workplace safety. Retrieved March 23, 2020, from https://www.textron.com/assets/CR/2018/CRR_OurPeople_1_WorkplaceSafety.html

[49] Reason, J. (2016). Managing the risks of organizational accidents. Taylor & Francis

[50] Annual Review of U.S. General Aviation Accident Data 2006

Microsoft Word - 2006 GA Review Cover (erau.edu)

[51] The Richard G. McSpadden Report, 33rd AOPA Air Safety Institute Accident Report (2021).

Richard G. McSpadden Report Figure View - AOPA

[52] Jensen,R. S. ; Benel,R. A., Judgment Evaluation and Instruction in Civil Pilot Training (1977). Judgment Evaluation and Instruction in Civil Pilot Training. (dtic.mil)

[53] Diehl. A. (1991). The effectiveness of training programs for preventing aircrew "error." In R.S. Jensen, *Proceedings of the 6th Symposium on Aviation Psychology (Vol. 2, pp. 640-655).* Columbus: Ohio State University.

[54] Santos, L.F.F.M.; Melicio, R. Stress, pressure and fatigue on aircraft maintenance personal. Int. Rev. Aerosp. Eng. (IREASE) 2019, 12, 35. Stress, Pressure and Fatigue on Aircraft Maintenance Personal | Santos | International Review of Aerospace Engineering (IREASE) (praiseworthyprize.org)

[55] Satow, .R.L. The Role of Human Factors in Improving Aviation Safety. AERO. 1999, 8,.

[56] Latorella, K.A.; Prabhu, P.V. A review of human error in aviation maintenance and inspection. Int. J. Ind. Ergon. 2000, 26, 133–161. A review of human error in aviation maintenance and inspection - ScienceDirect

[57] Stolzer, A., & Halford, C. (2007). Data mining methods applied to flight operations quality assurance data: A comparison to standard statistical methods. Journal of Air Transportation, 12(1), 6-24.

[58] Maheshwari, A., Davendralingam, N., & DeLaurentis, D. A. (2018). A comparative study of machine learning techniques for aviation applications. In 2018 Aviation Technology, Integration, and Operations Conference (p. 3980). https://doi.org/10.2514/6.2018-3980

[59] Carnahan, B., Meyer, G., & Kuntz, L. A. (2003). Comparing statistical and machine learning classifiers: alternatives for predictive modeling in human factors research. Human Factors, 45(3), 408-423. https://doi.org/10.1518/hfes.45.3.408.27248

[60] Čokorilo, O., De Luca, M., & Dell'Acqua, G. (2014). Aircraft safety analysis using clustering algorithms. Journal of Risk Research, 17(10), 1325-1340. https://doi.org/10.1080/13669877.2013.879493

[61] Madeira, T., Melício, R., Valério, D., & Santos, L. (2021). Machine learning and natural language processing for prediction of human factors in aviation incident reports. Aerospace, 8(2). https://doi.org/10.3390/aerospace8020047

[62] Khattak, A., Chan, P. W., Chen, F., Peng, H., & Mongina Matara, C. (2023). Missed Approach, a Safety-Critical Go-Around Procedure in Aviation: Prediction Based on Machine Learning-Ensemble Imbalance Learning. *Advances in Meteorology*, *2023*. https://doi.org/10.1155/2023/9119521

[63] Nogueira, R. P. R., Melicio, R., Valério, D., & Santos, L. F. F. M. (2023). Learning Methods and Predictive Modeling to Identify Failure by Human Factors in the Aviation Industry. *Applied Sciences (Switzerland)*, *13*(6). https://doi.org/10.3390/app13064069

[64] L. Tanguy, N. Tulechki, A. Urieli, E. Hermann, and C. Raynal, "Natural language processing for aviation safety reports: from classification to interactive analysis," Computers in Industry, vol. 78, pp. 80–95, 2016. (PDF) Natural Language Processing for aviation safety reports: from classification to interactive analysis (researchgate.net)

[65] S. Robinson, "Multi-label classification of contributing causal factors in self-reported safety narratives," Safety, vol. 4, no. 3, p. 30, 2018. Safety | Free Full-Text | Multi-Label Classification of Contributing Causal Factors in Self-Reported Safety Narratives (mdpi.com)

[66] Dong, T., Yang, Q., Ebadi, N., Luo, X. R., & Rad, P. (2021). Identifying Incident Causal Factors to Improve Aviation Transportation Safety: Proposing a Deep Learning Approach. Journal of Advanced Transportation, 2021. https://doi.org/10.1155/2021/5540046