

## Original Article

<https://doi.org/10.12985/ksaa.2026.34.2.001>  
ISSN 1225-9705(print) ISSN 2466-1791(online)

## 머신러닝 기반 항공기 사고 인명 상해 등급에 영향을 미치는 요인 분석

이정열\*, 전정환\*\*

## Analysis of Factors affecting Injury Severity in Aircraft Accidents based on Machine Learning

Jeoung yeoul Lee\*, Jeong hwan Jeon\*\*

## ABSTRACT

This study utilized accident data from the National Transportation Safety Board (NTSB) to develop a predictive model for factors affecting human injury severity in aircraft accidents using the XGBoost algorithm. After addressing class imbalance, the model achieved accuracy of 0.850, recall of 0.842, and F1 score of 0.815 at a learning rate of 0.7. SHAP analysis identified distance to the airport, elapsed time since inspection, cumulative flight hours, seasonal factors, and wind speed as the most influential factors. Unlike previous studies that primarily focused on mechanical causes, this study demonstrates the importance of environmental and operational conditions, highlighting the need for data-driven safety management and practical implications for real-time risk detection and improving global ATC operations.

**Key Words :** Aircraft Accidents(항공기 사고), Machine Learning(머신러닝), XGBoost(엑스지부스트), SHAP(샤프 분석), Aviation Safety(항공 안전), NTSB(국가교통안전위원회)

## 1. 서 론

세계화와 생활 수준의 향상, 그리고 이동 시간 단축에 대한 욕구, 저비용항공사(LCC)의 성장, 항공운임 인하, 항공기 성능 향상 등으로 인해 세계 각국의 항공 운송 시장은 꾸준히 성장하고 있다(Yoon and Son, 2023). 국내 항공시장 역시 지난 20년간 연평균 6%대의 높은 성장률을 나타내고 있다(Kim and Hwang, 2025). 이에 따라 항공 운송량은 증가하고 있으며, 공항의 관제시스템 운영 효율화와 항공 안전 확보는 점

차 더 중요해지고 있다. 항공운송에 있어서 안전은 가장 중요한 요소이며, 항공 안전을 향상하기 위하여 항공기 사고의 분석은 필수적이다(Kim, 2020; Yoo et al., 2024).

국내의 인천국제공항, 김포국제공항, 김해국제공항, 제주국제공항 등 주요 공항에서도 항공 운항 편수는 꾸준히 증가하는 추세에 있으며, 특히 출발과 도착이 집중되는 시간대에는 활주로 포화 현상이 나타나고 있다(Shin et al., 2024). 이러한 포화 현상은 항공 안전 관리와 관제시스템에도 영향을 주게 되는데, 관제사와 조종사 간의 통신 오류, 활주로 진입 착오, 이륙 혹은 착륙의 지연으로 인한 스트레스 등의 사고 혹은 준사고 사례가 보고되고 있다(Shin and Jo, 2021). 그러나 현재의 항공 관제시스템은 여전히 규칙 기반 통신과 감시 기능에 크게 의존하고 있으며, 사고 발생 이후 대

Received: 24. Dec. 2025, Revised: 15. Jan. 2026,

Accepted: 2. Feb. 2026

\* 대한민국 공군 시험평가단

\*\* 경상국립대학교 산업시스템공학과

연락처 E-mail: jhjeon@gnu.ac.kr

연락처 주소: 경상남도 진주시 진주대로 501 (52828)

응하는 후행적 방식에서 크게 벗어나지 못하고 있다. 이로 인해 실제 사고 가능성을 조기에 감지하거나, 선제적으로 위험을 차단하는 데에는 뚜렷한 한계가 존재한다.

특히 현행 시스템은 실시간 데이터 기반의 예측 기능이 미흡하며, 복합적 위험 요인 간 상호작용을 반영한 정량적 분석도 부족하다. 기상 변화, 항공기 속도, 활주로 점유 시간, 관제사 응답 지연 등 다양한 요소가 동시에 작용하는 이착륙 단계의 특성을 기존 관제 체계는 충분히 포착하지 못하고 있으며, 관제사의 경험과 직관에 크게 의존하는 운영 방식은 예기치 못한 상황에서 오류 가능성을 높이고 있다. 따라서 보다 정교한 예측 시스템의 도입과 관제시스템 구조의 개선이 시급한 과제로 대두되고 있다.

본 연구에서는 미국 국가교통안전위원회(NTSB)<sup>1)</sup>에서 제공하는 항공사고 데이터를 가지고 머신러닝(machine learning) 기법의 하나인 XGBoost(eXtreme gradient boosting)을 활용하여 항공기 사고 데이터에서의 인명 피해 등급에 영향을 미치는 가장 중요한 요인을 예측하는 모형을 구축하였다. XGBoost는 비선형적 관계와 변수 간 상호작용을 효과적으로 포착할 수 있으며, 불균형 데이터 처리, 결측치 자동 보정, 정규화 기반 과적합 제어 등 고급 기능을 제공한다. 또한 <sup>2)</sup>SHAP(shapley additive explanations)과 같은 해석 기법을 통해 예측 결과에 영향을 미친 주요 요인을 정량적으로 확인할 수 있어, 단순한 예측을 넘어 정책적 시사점을 도출하는 데 강점을 지닌다. 따라서 본 연구에서는 XGBoost를 활용하여 항공기 사고의 발생 가능성을 예측하고, 이를 통해 항공안전관리와 관제시스템 개선을 위한 실증적 근거를 제시하고자 하였다.

## II. 이론적 배경

### 2.1 항공기 사고

국제민간항공기구(ICAO) 부속서 13에 따르면, 항공

기 사고(aircraft accident)란 항공기가 사람을 태우기 위해 탑승을 시작한 시점부터 모든 사람이 항공기에서 내릴 때까지의 동안 발생한 사건 중, 사람이 사망하거나 중상을 입었거나 항공기가 중대한 손상을 입었거나 항공기가 실종되었거나 접근할 수 없는 상태가 된 경우를 말한다.

항공기 사고의 주요 유형은 발생단계에 따라 지상 사고(ground accident), 이착륙 사고(takeoff & landing accident), 순항 중 사고(en-route accident)로 구분할 수 있으며, 이 중 이착륙 사고의 발생 빈도가 가장 높다.

항공기 사고는 단일 원인보다는 여러 요인이 복합적으로 작용한 결과로 발생하는 경우가 많다. 일반적으로 항공기 사고 원인은 인적 요인, 환경 요인, 기계·기술적 결함, 조직·시스템 요인으로 분류된다(Hong, 2002). 이 가운데 조종사 판단 오류, 절차 미준수, 관제사 실수, 승무원의 판단 착오, 의사소통 오류 등의 인적 요인에 기인한 사고가 전체 항공기 사고 중 가장 큰 비중을 차지한다(Hong, 2005). 환경 요인은 기상 악화, 난기류, 활주로 상태, 지형적 특성과 같이 외부 조건에서 기인한다. 기계·기술적 결함은 기체 노후화, 부품 결함, 센서 혹은 장비 이상 등 기술적 요인으로 설명된다. 조직·시스템 요인은 관제 인력 부족, 안전 절차의 미비, 규제 체계의 한계 등 제도적·구조적 측면에서 발생한다(Hong, 2005).

이들 요인은 개별적으로 사고를 유발하기도 하지만, 실제로는 다중 요인이 상호작용하면서 사고로 이어지는 경우가 대부분이다. 예컨대, 불리한 기상 조건이 존재할 때 조종사의 피로 누적과 기체의 성능 저하가 동시에 발생하면 사고 위험은 기하급수적으로 증가한다(Paek et al., 2022). 따라서 항공기 사고를 이해하고 예방하기 위해서는 특정 요인을 단일하게 분석하기보다는, 인적·환경적·기계적·시스템적 요인이 상호작용하는 복합적 맥락을 정량적으로 분석하는 접근이 필요하다(Oster et al., 2013).

### 2.2 항공 안전 연구의 데이터 기반 분석

항공 안전 연구는 전통적으로 사고 발생 이후 축적된 데이터를 바탕으로 한 통계적·확률적 분석에 의존해 왔다. 항공기는 고도의 안전성을 목표로 설계·운영되는 시스템이므로 항공기 사고나 중대 준사고는 발생 빈도가 매우 낮은 희소한 사건이다. 이로 인해 수백만

1) 국가교통안전위원회(NTSB, National Transportation Safety Board): 미국 연방 정부 산하 독립기관으로, 항공을 포함한 교통사고를 조사하고 원인 규명 및 안전 개선 권고를 수행한다.

2) SHAP(Shapley Additive exPlanations): 머신러닝 모델의 예측 결과를 설명하기 위한 기법(Explainable AI, XAI) 중 하나로, 각 특성(feature)이 예측에 얼마나 기여했는지를 수치로 제시.

회의 정상 운항 데이터와 비교해 사고 사례는 극히 일부에 불과하며, 항공 안전 데이터는 정상 운항 상태 데이터가 압도적으로 많은 구조를 갖는다. 이러한 특성은 안전 시스템의 안정성을 입증하는 긍정적 지표이지만, 통계적 분석이나 위험 예측 모델링 관점에서는 심각한 데이터 불균형을 초래한다. 데이터 불균형 환경에서는 사고 사례가 분석 과정에서 충분히 반영되지 못해, 사고율 산출, 회귀분석, 빈도 기반 비교와 같은 기존 통계적 접근은 사고 원인의 사후적 규명이나 제도적 개선에는 유용할 수 있으나, 사고 이전 단계에서의 실시간 위험 징후를 민감하게 탐지하는 데에는 구조적 한계를 지닌다.

더 나아가 항공기 사고는 개별 변수의 변화가 결과에 미치는 영향이 일정하지 않고, 다른 변수의 상태나 조합에 따라 그 영향의 크기와 방향이 달라지는 비선형적 변수 상호작용의 특성을 갖는다. 즉, 항공기 사고는 단일 요인에 의해 발생하기보다는 조종사 상태, 기상 조건, 항공기 시스템 상태, 운항 환경 등 여러 위험요인이 복합적으로 작용하면서 발생하며, 이 과정에서 위험은 선형적으로 누적되기보다 특정 조건이나 조합에서 급격히 증폭되거나 완화되는 양상을 보인다. 특히 일부 변수는 일정 수준까지는 사고 발생에 미미한 영향을 미치다가 임계치를 초과하는 순간 위험이 급격히 증가하는 특성을 나타내는데, 이러한 비선형적·임계적 구조는 전통적인 선형 통계 모형으로는 충분히 포착하기 어렵다. 결과적으로 데이터 불균형 문제와 비선형적 변수 상호작용이라는 구조적 특성으로 인해, 기존의 통계적 분석 방법은 예방적 안전성 확보와 실시간 대응력 강화를 위한 도구로 활용되기에는 근본적인 제약을 지닌다.

최근에는 빅데이터와 머신러닝 기술의 도입으로 항공 안전 분석의 패러다임이 사후 대응에서 사전 예방과 실시간 예측 중심으로 전환되고 있다(Yoon, 2025). 항공기 운항 과정에서 생성되는 방대한 센서 데이터, ATC 로그, 정비 이력, 기상 자료 등을 통합 분석함으로써 이상 징후를 조기에 포착하고, 사고 가능성을 사전에 예측하는 연구들이 확산하고 있다. XGBoost<sup>3)</sup>, LSTM<sup>4)</sup>, Random Forest<sup>5)</sup> 등 다양한

머신러닝 알고리즘이 활용되고 있으며, 적용 범위 또한 시계열·텍스트·네트워크 데이터까지 확대되고 있다. 이러한 연구들은 해외 주요 항공안전기관과 학계에서 실증적으로 검증되었으며, 사고 예방 가능성과 실시간 대응 역량에서 긍정적인 평가를 받고 있다.

Shin & Jo(2021)는 항공 보안 · 관계 시스템 성능 향상을 위해 머신러닝 기반 능동형 모델 개선 기법을 적용하였다. 그 결과, 불확실 데이터 선별과 이상 탐지 · 분류 모델을 통해 탐지율 향상과 오탐률 감소에 따른 대응 시간 단축 효과를 확인하였다.

Yoon(2025)의 연구에서는 GBC(gradient boosting classifier)와 RF(random forest) 등의 분류 기반 머신러닝 모델을 적용해 인명 피해 등급을 예측하고, 예측 정확도, 재현율(recall), F1-score 등의 성능 지표를 중심으로 검증하였다. 그러나 GBC(gradient boosting classifier)<sup>6)</sup>는 과적합 방지, 결측값 처리, 데이터 불균형 대응 측면에서 한계를 보였다.

국내 항공 안전 분야에서도 머신러닝 기반 실시간 예측 모델의 도입은 이루어지고 있으나, 여전히 초기 단계에 머물러 제한적으로 적용되고 있고 빅데이터와 머신러닝 기반의 연구가 더 활발히 이루어질 필요가 있다. 이에 본 연구는 고도화된 부스팅 기법인 XGBoost를 적용하여, 학습률 조정, 정규화, 트리 가지치기 등을 통한 과적합을 제어하고, 데이터 불균형을 고려한 학습 전략과 결측값 자동 처리와 SHAP 기반의 설명 가능한 예측 결과 해석 등 기술적·해석적 확장성을 확보하였다. 또한 기존 연구보다 정밀한 인명 피해 등급 예측 성능을 검증하고, 정책적 적용 가능성까지 제시하고자 한다(Table 1).

## 2.3 머신러닝

머신러닝은 대규모 데이터로부터 패턴을 학습하여 목표 특성에 영향을 미치는 요인의 중요도를 도출하거나, 주요 요인과 목표 특성 간의 관계를 예측하는 인공지능(AI)의 하위 분야다. 이러한 분석을 통해 의사결정

(RNN: recurrent neural network)의 한 종류로, 시계열 데이터나 순차적 데이터(문장, 음성, 센서데이터 등) 학습에 특화된 딥러닝 모델.

3) XGBoost(eXtreme Gradient Boosting) : Gradient Boosting을 극대화한 알고리즘으로, 빠른 속도·높은 성능·과적합 억제 기능 덕분에 머신러닝 대회와 산업 현장에서 가장 많이 사용되는 모델 중 하나.

4) LSTM (Long Short-Term Memory): 순환신경망

5) Random Forest : 여러 의사결정을 무작위로 조합한 앙상블 기법.

6) GBC (gradient boosting classifier): Gradient Boosting 알고리즘을 기반으로 한 분류(classification) 모델.

Table 1. Summary of previous studies and their limitations

| 선행연구                 | 연구 내용                                                                 | 연구 방법                                         |
|----------------------|-----------------------------------------------------------------------|-----------------------------------------------|
| Yoon (2025)          | Airplane accident prediction model                                    | Machine learning<br>DT, RF, GBC               |
| Shin & Jo (2021)     | SIEM system performance enhancement mechanism                         | Supervised learning &<br>Incremental learning |
| Liu et al. (2025)    | Dynamic determinants of aviation accidents                            | Random parameter bivariate probit approach    |
| Silagyi & Liu (2023) | Prediction of severity of aviation landing accidents using SVM models | Support vector Machine models                 |
| Rosadi et al. (2024) | Predict classification of aircraft damage                             | Machine learning                              |

을 자동화할 수 있으며, 항공 안전과 관련하여 다양한 연구와 적용이 시도되고 있다. 첫째, 이착륙 사고 예측은 대표적인 사례로, 고도·속도·기상·조종 패턴 등의 데이터를 활용해 활주로 이탈(runway excursion), 고강도 착륙(hard landing) 등을 사전 예측하고 경고가 가능한 모델이 개발되고 있다. LSTM 기반 연구에서는 착륙 30초 전 충격 강도를 예측하는 성과가 보고되었다. 둘째, 정비 예측(predictive maintenance) 분야에서는 Random Forest나 XGBoost가 부품 결함 가능성을 추정하여 예방 정비를 가능케 하고, 불필요한 비용을 절감할 수 있다. 셋째, 공역 혼잡 및 충돌 위험 분석에서는 시계열 모델과 그래프 기반 학습(graph neural network 등)이 적용되어 지연, 홀딩, 충돌 가능성을 정밀하게 분석한다. 넷째, 관제 커뮤니케이션 분석은 음성 인식과 자연어 처리(NLP)를 결합해 관제사-조종사 교신의 오류, 비표준 표현 등을 자동 탐지하는 방식으로 발전하고 있다.

머신러닝 적용의 실효성은 단순한 정확도 향상을 넘어 설명 가능성 확보와 실제 운영 시스템과의 통합 여부에 달려 있다. SHAP, LIME<sup>7)</sup>과 같은 기법은 모델

예측 결과에서 변수별 영향력을 정량적으로 제시하여 관제사·조종사·정책결정자가 결과를 신뢰할 수 있도록 지원한다. 예컨대 활주로 이탈 예측 시 시정 불량, 감속률 저하, 횡풍 강도 등이 어떻게 위험을 증폭시키는지를 시각화할 수 있다. 그러나 복잡한 딥러닝 모델은 여전히 블랙박스 성격을 지니며, 데이터 편향이나 불균형은 경고 과잉(false alarm) 또는 경고 누락(missed detection)을 초래할 수 있어 신중한 검증이 요구된다. 또한 예측 결과를 실제 경고 시스템과 어떻게 연계할지, 임계값 설정과 책임 소재를 어떻게 설계할지도 남은 과제다. 이처럼 머신러닝의 항공 분야 적용은 단순한 기술적 도입을 넘어 예측-설명-통합이라는 다차원적 과제를 수반한다. 이를 통해 사고 예방 능력과 운영 효율성을 동시에 제고하는 것이 향후 항공 안전 연구의 핵심 방향이라 할 수 있다.

본 연구에서는 예측 결과에 대한 설명 가능성을 확보하기 위하여 SHAP(Shapley additive explanations) 기법을 적용하였다. SHAP은 게임이론(game theory)의 Shapley Value 개념을 기반으로, 각 피처(feature)가 모델의 예측값에 기여하는 강도를 정량적으로 계산한다.

이를 수식적으로 표현하면, 모델  $f$ 와 피처 집합  $N=\{1, 2, \dots, n\}$ 이 주어질 때, 특정 피처  $i$ 의 기여도  $\phi_i$ 는 다음과 같이 정의된다.

$$\phi_i = \sum_{S \subseteq N \setminus i} \frac{|S| \text{vert}! (|N \text{vert}| - |S \text{vert}| - 1)!}{|N \text{vert}|!} \left[ \int (S \cup i) - \int (S) \right] \quad (1)$$

### III. 연구 설계

본 연구의 전체 분석 절차는 Fig. 1에 제시되어 있다. 연구는 자료수집과 전처리 과정을 거쳐 머신러닝 기반 예측모형을 설계하고, 모형의 성능을 평가한 후 SHAP 기반 분석을 통해 변수의 중요도를 해석하는 단계로 구성된다.

#### 3.1 자료수집

항공기 사고의 원인을 체계적으로 분석할 수 있는 데이터는 제한적이다. 이러한 제약에도 불구하고, 본 연구에서는 신뢰성과 활용도가 높은 NTSB의 항공기 사고 데이터를 모형 구성 및 항공기 사고 요인 분석에 사용하기로 한다. 이 데이터는 2008년 1월부터 2025

7) LIME (local interpretable model-agnostic explanations) : 복잡한 머신러닝 모델의 예측 결과를 국소적(local)으로 해석할 수 있게 돕는 기법.

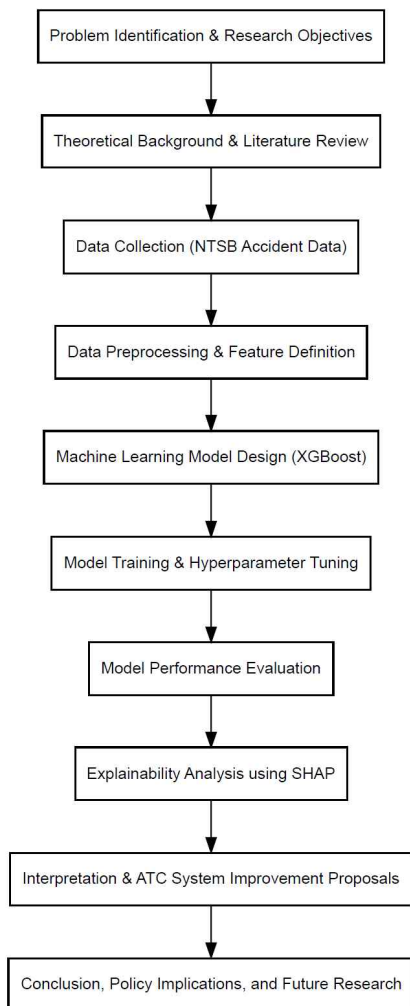


Fig. 1. Research process flowchart

년 5월까지 항공기 사고(accident, ACC) 및 준사고(incident, INC)와 관련하여 사고 요인이 될 수 있는 사고 요인이 될 수 있는 다양한 변수와 사고별 특성을 제공한다. NTSB에서 제공하는 avall.mdb 파일은 20개의 테이블로 구성되며, 이 가운데 항공기 사고와 관련이 깊은 것으로 평가되는 3개의 테이블인 events, Flight\_Crew, aircraft을 병합하여 하나의 통합 데이터로 정리하여 분석하였다.

이 통합 데이터에는 28,893건의 사고와 관련된 기록 데이터가 포함되어 있다. 인명 피해 등급과 관련된 ev\_highest\_injury 변수는 결측값이 1,313건이 있어 이를 제외하면 27,580건의 데이터가 남는다. 이 데이터에는 NONE(상해 없음)이 14,589건(52.9%)으로 가

장 높게 나타났으며, 다음으로 FATL(사망) 5,976건(21.7%), MINR(경미) 3,846건(13.9%), SERS(중상) 3,169건(11.5%) 순으로 나타났다. 그리고 설명 변수 가운데 중요한 요인인 afm\_hrs, cert\_max\_gr\_wt, total\_seats, num\_eng의 경우 결측값으로 나타나는 경우가 11,083건, 10,004건, 7,001건, 5,027건으로 많다. 설명 변수에 결측치가 포함된 행을 제거하여 완전한 데이터셋을 구성한 결과 총 2,350건이 남았다.

본 연구에서 결측치 제거 이후 남은 표본은 무작위 추출된 하위 집합이 아니라, 상대적으로 조사 수준과 기록 품질이 높은 사고 사례로 구성된 선택적 표본이다. 따라서 본 연구의 분석 결과는 전체 항공기 사고에 대한 일반적 결론이라기보다는, 조사 정보가 충분히 확보된 사고 사례 집단에서의 요인 중요도를 설명하는 결과로 해석되어야 한다.

이러한 데이터 축소는 단순한 데이터 손실이라기보다, NTSB 데이터의 수집·구성 방식과 분석 변수의 특성에서 기인한 구조적 결과로 이해할 필요가 있다.

첫째, NTSB 사고 데이터는 장기간에 걸쳐 다양한 사고 유형과 조사 수준을 포괄적으로 수집한 자료로, 모든 사고 사례에 대해 같은 수준의 정보가 체계적으로 기록되지 않는다. 특히 사고의 심각도가 낮거나 조사 범위가 제한적인 사례의 경우, 기상 조건, 인적 요인, 항공기 시스템 상태와 같은 세부 변수들을 누락시킨 경우가 빈번하다.

둘째, 항공기 사고 데이터는 사고 발생 빈도가 낮고, 고위험 사고일수록 조사 강도가 높아 상대적으로 정보가 충실한 특성을 가진다. 따라서 결측치 제거 이후 남은 2,350건의 사례는 무작위로 선택된 표본이 아니라, 비교적 조사 정보가 충분히 확보된 사고 사례들로 구성된 하위 집합이라고 볼 수 있다. 이는 분석의 신뢰성을 저해하기보다는, 오히려 변수 간 관계를 명확히 탐색할 수 있는 데이터 구조를 형성한다.

셋째, 표본 수의 감소에도 불구하고 최종 분석에 사용된 2,350건의 사고 사례는 항공 안전 연구 분야에서 여전히 충분한 규모이며, 머신러닝 모델링에 필요한 최소 표본 요건을 충족한다. 따라서 본 연구에서의 데이터 축소는 분석의 한계라기보다, 데이터 품질과 해석 가능성을 우선한 선택적 정제 과정의 결과로 해석될 수 있다.

한편, 다수의 결측치를 가진 행을 선별적으로 제거한다거나 결측치 대체(imputation)를 수행하는 경우, 여러 분류 알고리즘에서 정확도와 재현율이 각각 0.6

과 0.2 수준을 거의 넘지 못하는 것으로 보고되었다(Yoon, 2025). 결측치가 제거된 2,350건의 데이터를 데이터 불균형을 조정한 경우 0.8을 상회하는 예측 성능을 보이므로, 본 연구에서는 결측치 대체(imputation)를 수행하지 않고, 하나라도 결측값이 존재하는 경우 해당 관측치는 분석에서 제외하였다. 이러한 완전 사례 분석(listwise deletion) 방식은 표본 수를 감소시키는 단점이 있으나, 변수 간 관계를 왜곡 없이 해석하고 모델 안정성을 확보하기 위해 일반적으로 사용되는 보수적인 접근이다.

### 3.2 변수 정의 및 전처리

해당 사고의 구분을 위한 식별자 `ev_id` 및 `Aircraft_Key`는 항공기 사고 예측과 상관이 없으므로 제외하였다. 또한 분석모형의 목표 특성이 될 수 있는 항공기 사고 당시 인명 피해 등급을 나타내는 `ev_highest_injury`의 빈도는 결측치 제거 및 전처리 이후 분석에 사용된 2,350건의 데이터 기준으로 NONE(상해 없음)이 1,353건(57.57%)으로 가장 높게 나타났으며, 다음으로 MINR(경미) 428건(18.21%), FATL(사망) 293건(12.47%), SERS(중상) 276건(11.74%) 순으로 나타났다. 클래스 간 불균형은 존재하나 그 수준이 극단적이지는 않다. 이에 본 연구에서는 인명 피해 등급을 나타내는 `ev_highest_injury`를 예측모형의 목표 특성으로 설정하였다. `ev_highest_injury`의 심각성 정도의 분포는 Fig. 2에 제시하였다.

모델의 예측대상은 사고 당시 부상 중증도(`ev_highest_injury`)이므로 사후 결과에 해당하는 기체 손상 심각도

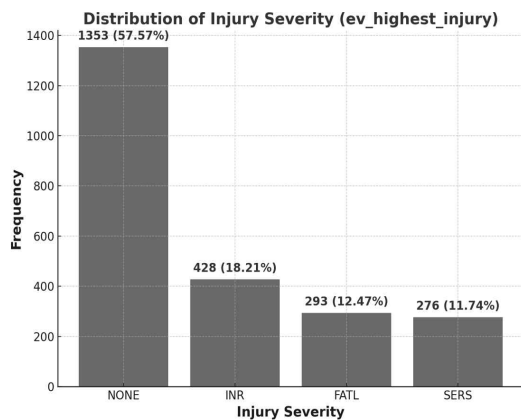


Fig. 2. Distribution of aircraft accident injury severity levels

(damage) 변수는 제외하였다. 다음 항목은 분산이 거의 없어 과적합과 잡음을 늘리므로 제거하였다. `ev_country`는 USA가 99.96%, `crew_no`는 1명이 99.7%, `ev_type`은 ACC(accident)가 98.99%로 대부분을 차지하며, `wx_cond_basic`은 VMC가 98.8%로 다른 기상 조건의 분포가 매우 제한적이므로 제거하였다. `ev_city`는 좌표(`dec_latitude/dec_longitude`)와 `ev_state`가 있어 대체 가능하므로 제거하였다. 다음의 33개의 변수를 최종 분석 변수로 확정하였다. 기체/성능은 `acft_category`, `cert_max_gr_wt`, `total_seats`, `num_eng`, `homebuilt`, `afm_hrs`, `afm_hrs_last_insp`의 7개의 변수를 사용하였다. 위치/환경은 `dec_latitude`, `dec_longitude`, `apt_elev`, `apt_dist`, `vis_sm`, `wind_dir_deg`, `wind_vel_kts`, `gust_ind`, `gust_kts`, `light_cond`, `sky_*`, `wx_temp`, `wx_dew_pt`의 16개의 변수를 사용하였다. 조종사 관련은 `crew_age`, `crew_sex`, `crew_category`, `med_certf`, `med_crtf_vldty`의 5개의 변수를, 그리고 기타 `ev_year`, `ev_month`, `ev_dow`, `ev_state`, `ev_nr_apt_loc`의 5개 변수를 사용하였다.

### 3.3 머신러닝 모델 설계

NTSB 데이터를 활용하여 결측 행과 일부 특성을 제거한 뒤, `ev_highest_injury`를 목표 특성으로 하고 나머지 33개 변수를 입력 특성으로 하여 항공기 사고의 인명 피해 등급을 분류·예측하는 모형을 XGBoost로 구축하였다.

본 연구에서는 XGBoost 모형의 주요 하이퍼파라미터를 교차검증을 통해 탐색하였다. 트리 수는 최대 1,000개까지 학습을 시도하였으며, 조기 종료(early stopping)를 적용하여 최적 반복 수를 자동으로 선택하였다. 트리 깊이(`max_depth`)는 3~10 범위에서 탐색한 결과 6으로 결정하였다. 리프 노드의 최소 가중치 합(`min_child_weight`)은 기본값인 1로 설정하였다. 또한 학습률은 0.1~0.9 범위에서 변화시키며 가장 재현율이 높게 나타나는 학습률을 탐색하여 적용하였으며, 전체 2,350건의 항공기 사고 데이터를 훈련 데이터(70%)와 테스트 데이터(30%)로 나누었다.

학습률 변화에 따른 정확도(accuracy), 정밀도(precision), 재현율(recall), F1-score의 검증 성능은 Fig. 3에 제시하였다. 인명사고 등급은<sup>8)</sup> 'MINR',

8) 인명사고 등급 : 국제민간항공기구(ICAO)와 항공사

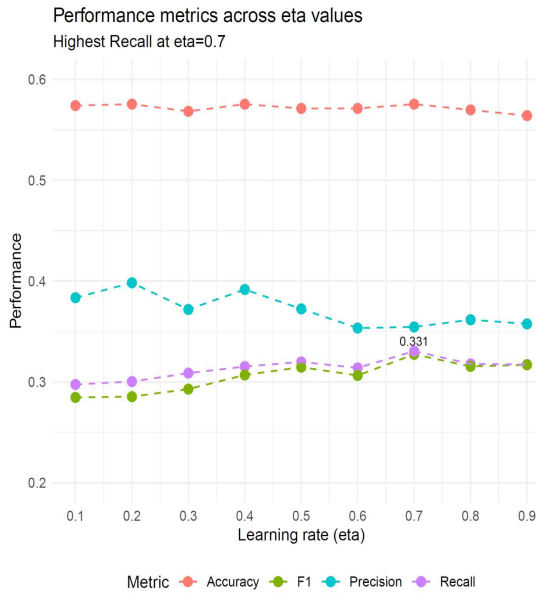


Fig. 3. XGBoost performance metrics by learning rate (accuracy, F1, precision, recall)

‘SERS’, ‘FATL’에 대해서는 클래스 빈도를 고려한 가중 평균을 통해 정밀도, 재현율, F1-score를 산출하였다. 분석 결과, 정확도는 데이터 불균형의 영향으로 큰 변화는 관찰되지 않았으며, 이는 ‘NONE’ 등급이 다수를 차지하는 분포 특성에 기인한 것으로 해석된다. 학습률 변화에 따른 비교 결과, 본 연구에서 설정한 하이퍼파라미터 조건에서는 학습률 0.7에서 재현율이 가장 높게 관찰되었다. 다만 이는 특정 실험 조건에 기반한 결과로, 다른 하이퍼파라미터 조합이나 데이터 구성에 따라 상이한 결과가 도출될 수 있다.<sup>9)</sup>

## IV. 분석 결과

### 4.1 혼동행렬

혼동행렬(confusion matrix)<sup>10)</sup>은 XGBoost 모형

FOQA·SMS(안전관리시스템)에서 사고/사건의 인명피해 심각도(Fatal : 사망, Serious Injury : 골절, 장기 손상, 화상 등 중상, Minor Injury : 단기 치료로 가능한 경상, No Injury : 인명피해 없음).

9) 후속 연구에서는 0.01-0.3 범위를 중심으로, 학습률을 포함한 여러 하이퍼파라미터의 결합 탐색을 통해 정밀한 최적화가 필요하다.

10) 혼동행렬(confusion matrix): 분류 모델의 성능을 평가하기 위해, 실제 값(ground truth)과 모델의 예측값(prediction)을 비교하여 정리한 표.

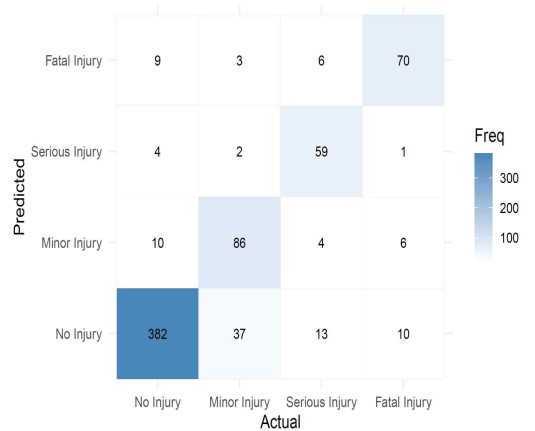


Fig. 4. Confusion matrix of the XGBoost model

의 예측 결과를 실제 값과 비교하여 분류 성능을 직관적으로 파악할 수 있도록 시각화된 분석 도구이다. 혼동행렬을 히트맵(heat map) 형태로 제시함으로써, 정확도(accuracy), 정밀도(precision), 재현율(recall), F1-score와 같은 주요 성능 지표를 종합적으로 평가할 수 있으며, 범주별로 올바르게 분류된 사례와 오분류된 사례를 구체적으로 확인할 수 있다. 이러한 특성으로 인해 혼동행렬은 머신러닝 분류 모형의 성능을 진단하는 데 효과적인 분석 기법으로 활용된다.

분석 결과, 불균형 보정 이후 XGBoost 모형은 전체 정확도(accuracy) 85.0%를 나타냈다. Fig. 4에 나타난 것과 같이 재현율(recall)은 No Injury 86.4% (382/442), Minor Injury 81.1%(86/106), Serious Injury 89.4%(59/66), Fatal Injury 79.5%(70/88)을 나타내었고, 정밀도(Precision)는 No Injury 94.3% (382/405), Minor Injury 67.2%(86/128), Serious Injury 72.0%(59/82), Fatal Injury 80.5%(70/87)로 확인되었다.

### 4.2 변수 중요도 분석

Table 2와 Fig. 5에 나타난 것처럼, XGBoost 모형의 SHAP 분석 결과에 따르면, 항공기 사고를 설명하는 데 가장 큰 SHAP 값을 나타낸 요인은 사고 발생 지점과 가장 가까운 착륙공항까지의 거리(apt\_dist)로 나타났다. 이는 사고 위험이 기체 내부 요인보다도 공항 접근 과정에서의 공간적 맥락에 크게 영향을 받을 수 있음을 시사한다. 특히 이착륙은 비행 중 가장 위험도가 높은 구간으로 알려져 있으며, 활주로의 거리,

Table 2. SHAP feature importance ranking (Top 20)

| Ranking | Variable           | Mean (SHAP) |
|---------|--------------------|-------------|
| 1       | apt_dist           | 0.133       |
| 2       | afm_hrs_last_insp  | 0.101       |
| 3       | afm_hrs            | 0.078       |
| 4       | ev_month           | 0.072       |
| 5       | wind_vel_kts       | 0.052       |
| 6       | dec_latitude       | 0.044       |
| 7       | crew_age           | 0.042       |
| 8       | altimeter          | 0.041       |
| 9       | cert_max_gr_wt     | 0.040       |
| 10      | wx_temp            | 0.035       |
| 11      | ev_year            | 0.032       |
| 12      | apt_elev           | 0.031       |
| 13      | dec_longitude      | 0.030       |
| 14      | wind_dir_deg       | 0.028       |
| 15      | wx_dew_pt          | 0.026       |
| 16      | total_seats        | 0.022       |
| 17      | acft_category.HEL1 | 0.000       |
| 18      | crew_category.FLTI | 0.000       |
| 19      | crew_category.PLT  | 0.000       |
| 20      | ev_dow.Mo          | 0.000       |

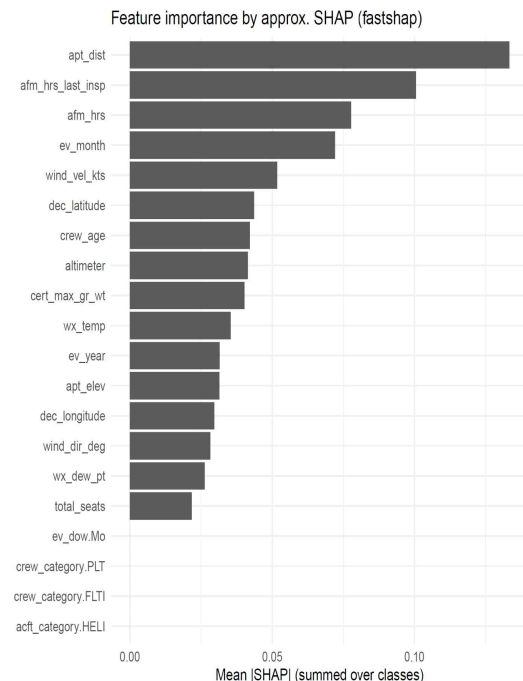


Fig. 5. Feature importance by SHAP

접근 경로의 복잡성, 지형적 특성과 같은 요인이 사고 발생 가능성에 직접적인 영향을 미칠 수 있다.

본 연구에서 apt\_dist가 가장 큰 SHAP 값으로 나타난 결과를 사고 발생 지점과 가장 가까운 공항까지의 거리(apt\_dist) 자체의 직접적 인과 효과로 해석하는 것은 과대 해석의 위험이 있다. 이에 본 연구에서는 apt\_dist를 사고 발생 위치와 공항 접근·이착륙 단계의 공간적 맥락을 대표하는 대리변수(proxy)로 해석하는 것이 타당하다고 본다. 즉, apt\_dist는 단순한 거리 효과라기보다, 이착륙 단계에서 나타나는 운항 복잡성, 공항 인근의 교통 밀도, 지형적 제약, 접근 경로의 난이도, 관제 부담 및 환경 요인의 복합적 영향을 간접적으로 반영하는 변수로 이해될 수 있다.

두 번째로 큰 SHAP 값을 나타난 변수는 항공기의 마지막 점검 이후 경과시간(afm\_hrs\_last\_insp)으로 나타났는데, 이는 정비 주기와 안전성 간의 밀접한 연관성을 보여준다. 세 번째로 변수는 항공기의 누적 운항시간(afm\_hrs)으로 나타났으며, 이는 장기간 운항에 따른 기체 피로도가 인명 상해 등급과 관련될 수 있음을 시사한다. 네 번째 변수는 사고 발생 월(ev\_month)로 나타났는데, 이는 계절적·기상적 요인이 사고 발생 위험을 증폭시킬 수 있으며, 이러한 결과는 계절적 패턴에 기반한 사전적 안전 대책 수립의 필요성을 뒷받침한다.

이러한 분석 결과는 기존 연구와의 차별성을 분명히 한다. 선행연구에서는 주로 항공기의 누적 운항시간(afm\_hrs)을 핵심 요인으로 제시하며, 기체의 노후화, 부품 마모, 정비 이력과 같은 기계적 특성을 사고의 주요 원인으로 설명해 왔다. 반면 본 연구에서는 기체 이력보다 운항 환경 요인(apt\_dist)이 더 큰 설명력을 갖는 것으로 나타났다. 즉, 선행연구가 “기체 관리와 정비”에 초점을 두었다면, 본 연구는 “환경적 요인의 작용”을 실증적으로 부각했다는 점에서 의의가 있다. 이는 이착륙 사고 예방을 위한 안전 정책이 단순히 정비 체계 강화에 국한되지 않고, 공항 접근 환경 관리 및 계절별 위험 대응 전략을 포함하는 통합적 안전관리 체계로 확장될 필요가 있음을 시사한다.

추가적으로, 풍속(wind\_vel\_kts)은 다섯 번째 요인으로 나타났으며, 위도(dec\_latitude)와 같은 지리적 요인은 여섯 번째, 승무원 연령(crew\_age)과 같은 인적 요인은 일곱 번째로 나타났다. 이는 항공기 사고가 단일 요인에 의해 설명되기보다, 기상·지리·인적 요인 간의

복합적 상호 작용으로 발생할 수 있음을 보여준다.

이상의 분석 결과는 항공기 이착륙 사고 예방 전략이 기존의 기체 중심 정비 관리 패러다임을 넘어, 공항 접근 환경과 운항 맥락을 종합적으로 관리하는 방향으로 전환될 필요성을 시사한다. 특히 사고 발생 지점과 가장 가까운 공항까지의 거리( $apt\_dist$ )가 가장 SHAP 값이 큰 설명 변수로 도출되었다는 점은, 사고 위험이 특정 공항 반경 및 접근 구간에서 구조적으로 집중될 가능성을 의미한다. 이는 향후 관제시스템이 단순한 이착륙 허가 기능을 넘어, 공항 접근 단계에서의 위험도를 사전에 정량화하고 경고하는 기능을 포함하도록 발전해야 함을 시사한다.

## V. 결 론

본 연구에서는 NTSB 항공기 사고 데이터를 기반으로 XGBoost 알고리즘을 적용하여 인명 상해 등급을 분류·예측하고, 관제시스템 개선방안을 모색하였다. 분석 과정에서는 결측치 및 불필요한 변수를 제거한 후,  $ev\_highest\_injury$ 를 목표 특성으로 설정하고 나머지 33개 변수를 입력 특성으로 활용하였다. 분석 결과, 데이터 불균형으로 인해 전체 정확도에는 큰 변화가 없었으나, 인명 상해가 발생한 사고를 탐지하는 재현율(recall) 측면에서 학습률( $\eta$ )의 변화가 큰 영향을 미치는 것으로 나타났다. 특히 학습률이 0.7일 때 재현율이 가장 높게 도출되어, 항공 안전관리에서는 중대 사고와 같은 소수 클래스 탐지 성능을 중시하는 접근이 유효할 수 있음을 시사한다. 불균형 보정 이후 XGBoost 모형은 Accuracy(0.8500), Recall (0.8416), F1-score(0.8146) 등에서 우수한 성능을 보였으며, SHAP 분석 결과, 사고 지점에서 가장 가까운 공항까지의 거리, 정비 주기, 계절 요인, 풍속 등이 주요 설명 변수로 도출되었다. 이는 기존 연구가 기체 노후화와 같은 기계적 요인에 주로 초점을 두었던 것과 달리, 운항 환경 및 계절적 특성의 중요성을 실증적으로 확인했다는 점에서 의의가 있다.

본 연구의 분석 결과는 향후 실시간 위험 탐지 모델 개발을 위한 기초자료로 활용될 수 있으며, 이를 실제 관제시스템에 적용하기 위해서는 비사고(정상운항) 데이터를 포함한 대규모 운항 데이터 확보와 시간축 기반 모델링이 선행되어야 한다.

항공 안전 관제시스템의 정책적 개선 방향 측면에서는 머신러닝 기반 위험 탐지 모듈의 단계적 도입이 필

요하다. 이를 통해 기존의 규칙 기반 대응이 가능할 것으로 기대한다. 또한 국내 항공사고 데이터의 제한성을 고려할 때, 항공사·공항·관제기관 간 비행운항품질보증(FOQA) 자료와 정비 이력 데이터를 연계하는 통합 관리 체계 구축이 요구된다. 나아가 머신러닝 기반 예측 모델의 설명 가능성을 토대로 연구자, 실무자, 정책 결정자 간 협력 체계를 강화함으로써 정책적 수용성과 신뢰성을 높일 필요가 있다. 이는 학문적·기술적 성과가 실제 정책으로 연계되는 중요한 기반이 될 것이다.

본 연구의 한계로는 분석에 사용한 데이터가 미국 NTSB 단일 데이터에 한정되어 있어, 연구 결과를 국내 항공 안전 환경에 적용하는 데에는 제약이 있다는 점을 들 수 있다. 또한 결측치 제거 과정에서 표본 수가 감소하였고, 목표 특성의 데이터 불균형과 XGBoost 단일 알고리즘을 사용했다는 점 역시 한계로 지적될 수 있다. 향후 연구에서는 국내 항공사고 데이터를 포함한 분석과 함께 다양한 머신러닝 알고리즘과의 비교를 통해 연구 결과의 일반화 가능성을 높일 필요가 있다.

본 연구는 머신러닝 기법을 활용하여 항공기 사고 시 인명 상해 등급에 영향을 미치는 주요 요인을 실증적으로 규명하였다는 점에서, 국내에서는 아직 초기 단계에 해당하는 연구로서 학술적 기여를 지닌다. 더불어 본 연구의 분석 결과는 향후 국내 항공 안전 정책 수립과 관제시스템 개선에 참고 자료로 활용될 수 있는 실무적 함의를 제공한다. 이러한 점에서 본 연구는 항공 안전 분야의 학문적·실천적 논의 확산에 기여할 것으로 기대된다.

## References

1. Yoon, M. G., and Son, W. M., "A study on the long-term forecasting of the international passengers based on the changes of the air travel demand structure in Korea", Journal of the Korean Society for Aviation Management, 21(4), 2023, pp.57-80.
2. Kim, H., and Hwang, E., "Comparative analysis of arima and sarima for forecasting air passenger demand by time", Journal of the Korean Data Analysis Society, 27(4), 2025, pp.1123-1140.

3. Kim, D. H., "A study on the accident model from the system safety perspective: Focused on aircraft accident", *Journal of the Korean Society for Aviation and Aeronautics*, 28(2), 2020, pp.63-70.
4. Yoo, J. H., Lee, J. H., and Kim, H. D., "Study on aviation safety management analysis and improvement strategies for aviation services using text mining", *Journal of the Korean Society for Aviation and Aeronautics*, 32(4), 2024, pp.8-15.
5. Shin, J., Kim, S., Baik, H., and Lee, J., "Study on airport operation characteristics analysis and capacity enhancement strategies using the capacity coverage chart (CCC)", *Journal of the Korean Society for Aeronautical Science and Flight Operations*, 32(3), 2024, pp.35-43.
6. Shin, Y. S., and Jo, I. J., "Siem system performance enhancement mechanism using active model improvement feedback technology", *Journal of the Korea Contents Association*, 21(12), 2021, pp.896-905.
7. Hong, S. J., "A study on human factors and organizational culture from the perspective of aviation safety", *The Journal of Aerospace Industry*, 2002, pp.88-100.
8. Hong, S. J., "Historical review on human factors of aviation safety", *Journal of the Korean Air Transport Association*, (1), 2005, pp.115-128.
9. Paek, H., Kim, J. H., Lim, J. J., Jeon, S., and Choi, Y. J., "Quantitative safety risk assessment using aviation safety data", *Journal of the Korean Society for Aviation and Aeronautics*, 30(4), 2022, pp.145-158.
10. Oster, C. V., Jr., Strong, J. S., and Zorn, C. K., "Analyzing aviation safety: Problems, challenges, opportunities", *Research in Transportation Economics*, 43(1), 2013, pp.148-164.
11. Yoon, H. S., "Machine learning based prediction model and factor appraisal of aircraft accident", *Journal of Digital Industry and Information*, 21(1), 2025, pp.15-26.
12. Liu, Q., Song, P., and Li, F., "Exploring the dynamic determinants of general aviation accidents across flight phases and time: a random parameter bivariate probit approach with heterogeneity in means", *Analytic Methods in Accident Research*, 2025, pp.100386.
13. Silagyi, D. V., II, and Liu, D., "Prediction of severity of aviation landing accidents using support vector machine models", *Accident Analysis & Prevention*, 187, 2023, pp.107043.
14. Rosadi, I., Franciscus, F., and Widanto, M. H., "A machine learning-based method for predicting the classification of aircraft damage", *EAI Endorsed Transactions on Internet of Things*, 10, 2024.