

## 자연어처리 알고리즘을 이용한 위험기반 항공안전데이터 자동분류 방안 연구

# A Study on Auto-Classification of Aviation Safety Data using NLP Algorithm

양성훈\* · 최 영 · 정소영 · 안주현\*\*  
항공안전기술원 데이터분석연구센터

Sung-Hoon Yang\* · Young Choi · So-young Jung · Joo-hyun Ahn\*\*

Data Analysis & research Center, Korea Institute of Aviation Safety Technology, Incheon 22581, Korea

### [요 약]

항공기 제작 및 운송 기술 발달로 국내 항공산업은 비약적인 발전을 이루었으나, 항공안전 사고는 지속해서 발생하고 있다. 관리 감독기관에서는 위험기반 항공안전데이터를 기반으로 위해 요인과 위험도를 분류하고, 운송사업자별 안전 경향성 파악과 취약분야를 도출하여 사전점검을 수행함으로써 사건·사고를 사전 예방중에 있다. 그러나 자연어 형식으로 기술된 항공안전데이터의 휴먼 분류는 지식과 경험, 성향에 따라서도 다른 분류 결과를 초래하고, 이벤트 내용의 의미 파악 및 분류를 위한 작업에 상당한 시간을 소요케 한다. 이에, 본 논문에서는 KoBERT 모델을 fine-tuning하고 5천 건 이상의 항공안전데이터를 기계학습 시켜 신규 데이터의 분류 값을 예측한 결과 79.2%의 정확성을 보였다. 그리고 유사 이벤트에 대해 동일한 결과 예측과 fail 된 데이터 중 일부는 휴먼 에러에 의한 오류임을 확인할 수 있었다.

### [Abstract]

Although the domestic aviation industry has made rapid progress with the development of aircraft manufacturing and transportation technologies, aviation safety accidents continue to occur. The supervisory agency classifies hazards and risks based on risk-based aviation safety data, identifies safety trends for each air transportation operator, and conducts pre-inspections to prevent event and accidents. However, the human classification of data described in natural language format results in different results depending on knowledge, experience, and propensity, and it takes a considerable amount of time to understand and classify the meaning of the content. Therefore, in this journal, the fine-tuned KoBERT model was machine-learned over 5,000 data to predict the classification value of new data, showing 79.2% accuracy. In addition, some of the same result prediction and failed data for similar events were errors caused by human.

**Key word** : Auto-Classification, Aviation Safety Data, Aviation Safety Inspector, KoBERT, NLP.

<http://dx.doi.org/10.12673/jant.2022.26.6.528>



This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Received 17 November 2022; Revised 1 December 2022  
Accepted (Publication) 16 December 2022 (30 December 2022)

\*Corresponding Author; Joo-Hyun Ahn

Tel: +82-02-2666-8692

E-mail: sh\_yang@kiast.or.kr

## I. 서론

항공기 제작 및 운송 기술 발달로 국내 항공산업은 비약적인 발전을 이루었으나, 운항 중 발생한 사건·사고 건수가 2017년부터 2021년까지 1백만 운항 횟수 당 연평균 1.6건으로 지속해서 발생하고 있다. 이러한 항공사고는 항공기의 치명적인 손상이나 승객의 사망사고로 이어지는 경우가 높기 때문에 이를 예방하기 위한 법·규정·지침 등을 마련하여 운영 중이다 [1].

국제민간항공기구(ICAO, International Civil Aviation Organization)는 2019년 Annex 19 Safety Management를 제정하여 항공안전데이터와 안전정보를 수집, 저장, 공유 등의 방법을 통한 사전예방형 안전관리시스템 구축 및 운영에 관한 사항을 규정하고 있다 [2].

국내에서는 항공안전법을 통해 항공운송산업의 안전감독을 시행하고 있으며 운송사업자별 담당 항공안전감독관을 배정하여 직접적인 안전관리를 감독하고 있다. 항공안전감독관은 안전관리 업무를 수행하며 발생한 각종 항공안전데이터를 기반으로 안전 경향성 파악과 운송사업자별 취약분야를 도출하고 항공기 운항 전 점검을 시행하여 막대한 손실로 이어질 수 있는 사건·사고를 사전 예방 중이다 [3].

이때 활용되는 항공안전데이터는 항공운송사업자로부터 수집 받는 항공안전이벤트 내용으로써 항공안전감독관이 직접 표준분류체계를 기반으로 안전데이터의 위해요인과 위험도를 분류하고 그 결과를 안전성과 모니터링 및 감독계획 등에 이용하고 있다 [4].

그러나, 수집된 안전이벤트 내용이 자연어 형식으로 기술되어 있어 용어, 표현방식, 첨부 자료 등의 형식이 비표준화되고, 이벤트 등록 담당자별로 두괄식, 개괄식, 개조식 등의 내용 전개로 감독관들이 내용 파악에 있어 어려움이 있다. 이는 내용이 같거나 유사한 이벤트에 대해 안전감독관별 지식과 경험, 성향에 따라 다른 분류 결과를 초래하고, 기술된 이벤트 내용의 의미 파악 및 분류를 위한 작업에 상당한 시간이 소요된다.

이러한 문제점을 해결하고자 본 논문에서는 자연어처리 알고리즘을 기반으로 일관성 있는 분류와 시간 소모를 절감할 수 있는 항공안전데이터의 자동분류 방안을 제시한다. 본 논문의 2장에서는 관련 연구 및 표준분류체계와 자연어처리 알고리즘에 대한 내용을 기술하고, 3장에서는 실험모델 및 결과와 fail된 예측 데이터의 원인분석에 관해 기술하고, 마지막 4장에서는 결론과 향후 연구 방향에 대해 제시한다.

## II. 연구배경

### 2-1 관련연구

자연어처리 기술이 발달하면서 항공분야에서의 활용을 위한 다양한 연구가 진행 중이다.

논문 [5].에서는 제주국제공항의 최근 5년(2015~2019)간의

표 1. kg2vec+CNN 분류모델의 평균 정확도 비교

Table 1. Comparison of Ave. accuracy of Suggested Model

Algorithms	Average Accuracy
one-hot+CNN	56.8%
word2vec+CNN	64.9%
<b>kg2vec+CNN</b>	<b>68.6%</b>
Naive Bayes	55.6%
Support Vector Machine	65.2%

운항 및 기상 데이터 분석을 통해 기상요인과 운항 지연간의 상관관계를 분석하고 그 결과에 머신러닝을 적용한 지연 예측모형을 제안하였다. 제안한 예측모형은 PYTHON Scikit-learn 라이브러리의 Random Forest 알고리즘을 기반으로 구축하고, 337,774개 학습 데이터와 84,444개 테스트 데이터를 이용해 모델을 테스트하였다. 그 결과 76%의 정확성을 보였으며, 해당 모델을 통해 제주공항 출발 항공편의 지연 여부를 예측하였다.

논문 [6].에서는 Text와 Knowledge Graph를 결합한 항공안전 보고서의 자동분류 알고리즘을 제안하였다. 제안한 알고리즘은 항공안전에 대한 정보를 포함하고 있는 지식 그래프를 논문에서 제시한 KG2vec 워드벡터 모델에 적용하여 단어 간의 연관정보를 추출하고, 합성곱 신경망(CNN, Convolutional Neural Network) 알고리즘을 이용해 차원을 축소시키며 정보를 학습하는 모델이다. 이후, 8년(2010~2018)간의 항공안전보고서 96,019개의 실험데이터를 제안 모델에 적용하여 8개(Human Factors, Aircraft, Procedure, Ambiguous, Company Policy, Weather, Environment, ATC Equipment) 도메인으로 분류하고, 각 도메인별 정확도를 다른 4개의 다른 알고리즘별로 비교하여 표 1.과 같이 평균 68.6%의 가장 좋은 성능을 보였다.

논문 [7].에서는 방대한 양의 항공안전정보로부터 안전사고 발생 시 유사 이벤트의 추출·제공을 위해 AKG(Aviation Knowledge Graph) 및 딥러닝 기반의 QA(Question Answering) 시스템 모델을 제안하였다. 제안 모델은 항공안전정보로부터 도출된 KG를 구성하고 데이터베이스 쿼리를 통해 Q를 제출하면 BERT 및 GPT3와 같은 NLP 알고리즘을 이용해 KG로부터 A를 추출하는 방식이다.

표 2. KGQA+GPT3-QA 모델의 정확도 비교

Table 2. Comparison of accuracy of Suggested Model

Algorithms	Accuracy
KGQA	56.0%
BERT-QA	22.7%
GPT3_QA	76.0%
KGQA+BERT-QA	63.0%
<b>KGQA+GPT3-QA</b>	<b>85.3%</b>

이후, NTSB(National Transportation Safety Board)로부터 4,000개 보고서를 수집하고 제안한 KGQA와 더불어 BERT 등과 혼합한 알고리즘에 적용한 결과 KGQA+GPT3-QA 모델의 정확도가 85.3%로 가장 높았음을 표 2와 같이 확인하였다.

**2-2 항공안전데이터 표준분류체계**

국내 항공운송사업자를 통해 받는 항공안전데이터는 감독관에 의해 발생분야, 발생원인, 위해요인을 구분하고 해당 항공사의 안전 증진에 활용되고 있다. 이를 구분하는 표준분류체계는 3 Level로 구분하여 1차 발생분야(Occurrence), 2차 발생원인(Cause), 3차 위해요인(Hazard)으로 정의하고 있으며 기본 구성 요소는 Table 3과 같다.

표 3. 항공안전데이터 표준분류체계

Table 3. Standard classification system of Aircraft Safety Data

Occurrence (14)	Cause (51)	Hazard (260)
Flight Operation	Pilot Error and 4 others	Pilot Injury/illness and 49 others
Maintenance	Mechanic Error and 7 others	Lack of, or poor maintenance release and 68 others
Cabin	Cabin crew Error and 3 others	Time / Business pressure and 10 others
Cargo/Dangreous Goods	Lack of, or poor Dangerous Goods control and 2 others	Improper cargo loading and distribution and 11 others
A i r Operation/Mainten ance Operation	Error and 4 others	Lack of, or poor scheduling for Cockpit crew/Cabin crew and 15 others
Design & Manufacturing	Aircraft System	Changes in safety requirements after design and 3 others
Weather/Natural Disasters	T u r b u l e n c e encounter and 2 others	Clear Air Turbulence and 10 others
Air Traffic Control	ATC Error and 2 others	U n s t a b i l i z e d Approach due to ATC and 15 others
Air Navigation	Faulty Navigational Aids	Lost communication
Other Aircrafts	Aircraft Proximity and 1 other	ACAS RA due to military aircraft and 4 others
Airfield Apron Operation	Runway Operation and 7 others	Lack of coordination with Air Traffic Control(ATC) and 17 others
Ground Operation	Ground Operation Error and 3 others	Time / Business pressure and 14 others
Other	Passenger and 2 others	Passenger failure to follow guidance and 9 others
Investigating	Investigating	-

1차 발생분야는 조종, 정비 등과 같이 분야를 식별할 수 있는 체계이며 13개로 분류되고, 2차 발생원인은 1차 발생분야 별 발생원인을 분석하기 위한 정의로 전체 51개로 분류되어있다. 3차 위해요인은 2차 발생원인을 발현케한 요인을 정의한 것으로 전체 260개로 구성되어있다. 항공안전감독관은 안전 이벤트 데이터를 기반으로 발생내용을 파악하고, 표준분류체계를 기준으로 발생분야, 발생원인, 위해요인을 직접 분류하는데 이는 서론에서 기술한 잠재적인 문제를 내포하고 있다.

**2-3 자연어처리 모델**

자연어 처리(NLP; Natural Language Processing)란 인공지능의 한 분야로 컴퓨터가 사람의 언어를 이해하고 처리하는 기술을 말하며, 항공, 항공우주, 방위산업 등 다양한 분야에서 활용되고 있다. 그러나, 한국어의 모호성, 불연속적인 데이터, 교착어, 의문문 등의 특성을 갖는 자연어처리를 위해서는 적합한 언어모델 선정이 필요하다 [8].

LSTM(Long Short Term Memory)은 단방향 순차 데이터를 분석하는 알고리즘으로써 문장의 길이가 길어질수록 앞서 처리한 정보가 열어지는 장치의존성(Long Term Dependency)<sup>1)</sup> 문제를 보완한 모델이다. 즉, “The sun is up in the SKY”와 같은 짧은 문장에서 SKY를 예측한다고 했을 때, 앞에서 기술된 단어들을 통해 예측할 수 있다. 그러나, “I grew up in Korea. ~ I can speak Korean fluently.”에서 Korean을 예측하기 위해서는 앞쪽의 문장에서 관련 정보를 확인해야 하지만, 관련 정보가 목표로 하는 단어와 거리(Gap)가 멀어질수록 연계 학습이 어려워지는 것이다. LSTM은 그림 1.과 같이 hidden state를 통해 불필요한 정보는 삭제하고 필요 정보만 전달하는 Cell state 구조를 채택하여 LTD 문제를 해결하였다 [9],[10].

ELMo(Embeddings from Language Models)는 기존의 단어 임베딩 방식의 알고리즘이 갖는 동음이의어 및 시제 구분 불가 문제를 보완한 알고리즘이다.

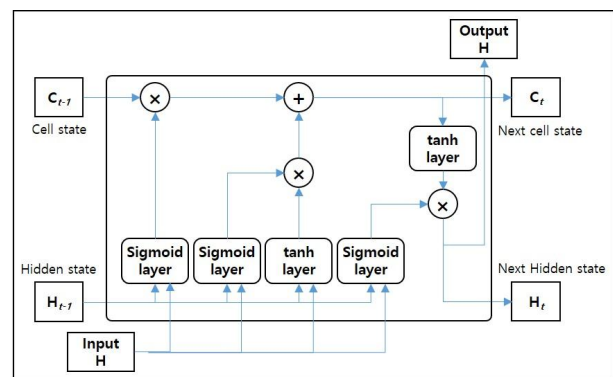


그림 1. LSTM 아키텍처

Fig. 1. LSTM Architecture

1) 시퀀스 데이터의 길이가 길어질수록, 과거의 중요한 정보에 대한 학습이 어려워지는 문제

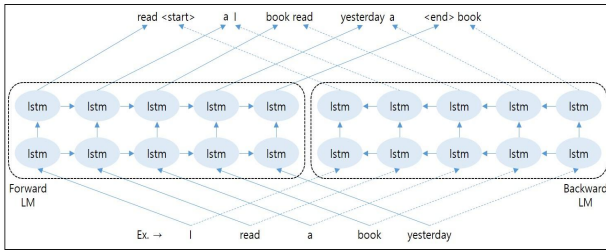


그림 2. ELMo 아키텍처  
Fig. 2. ELMo Architecture

기존 임베딩 방식은 “① I’ll book the Hotel. ② I’ll read a book”에서 모든 “book”을 동일한 벡터값(ex; ①-book : 0.65, ②-book : 0.65)으로 인코딩하므로 의미 구분이 어렵지만, ELMo는 다중 레이어를 이용해 각각 다른 벡터값(ex; ①-book : 0.65, ②-book : 0.75)으로 처리하므로 각각의 의미 구분이 가능하다. 그리고 “① I read a book. ② I read a book yesterday”의 경우 “yesterday”로부터 역방향 분석을 통해 ②-read가 과거 시제임을 구분 한다 [11],[12].

BERT(Bidirectional Encoder Representations Transformers)는 자연어의 양방향 분석을 지원하며 코퍼스라는 대량의 텍스트 데이터를 기반으로 다음 문장 예측(NSP : Next Sentence Prediction)과 마스킹 된 단어 예측(MLM : Masked Language Model)에 대한 학습을 사전훈련(Pre-training)시킨 언어모델이다. BERT의 입력 데이터는 3가지(최소 단어 단위로 구분하는 토큰, 문장을 구분하는 세그먼트, 각 토큰의 위치를 나타내는 포지션) 임베딩 벡터값으로 변환하고, 각 벡터값에 양방향 분석 처리하도록 Transformer 인코딩 모델을 적용하여 최종 데이터를 추출한다. 이때, MLM은 입력 데이터의 15%를 MASK 토큰으로 변경하여 마스킹 된 단어를 예측하고, NSP는 문장의 시작과 끝을 나타내는 토큰([CLS], [SEP])을 이용하여 문장 간의 연속성 및 연관 관계를 예측한다. 이후, 사전훈련된 학습모델에 신규 데이터 처리를 위한 파라미터 조정(Fine-tuning)을 통해 목표 데이터의 분류, 분석, QnA 등의 자연어처리 성능을 향상할 수 있다 [13],[14].

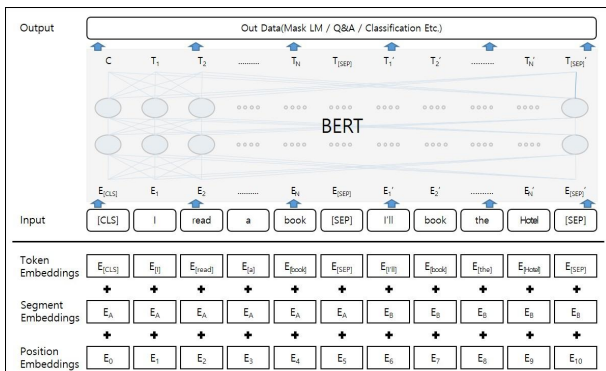


그림 3. BERT 아키텍처  
Fig. 3. BERT Architecture

KoBERT(Korean Bidirectional Embedding Representation Transformer)는 SKTBrain에서 공개하였으며 BERT 모델에 500만 개 이상의 한국어 위키 데이터와 2천만 개 이상의 문장으로 구성된 한국어 뉴스데이터를 추가로 학습시켜 한국어 처리 능력을 향상시켰다. 국내에서는 KoBERT를 이용한 댓글 분석, 여론 분석, 감성 분석, 분류 및 Q&A 등을 위한 많은 연구가 진행되고 있으며, 한국어 처리에 성능이 타 모델에 비해 뛰어난 것을 증명하였다 [15],[16]. 이에, 본 논문에서는 교착어, 비표준화의 특성을 갖는 목표 데이터의 자동분류를 위해 KoBERT 모델을 기반으로 다중 분류가 가능토록 Fine-tuning 하였다.

### III. 항공안전데이터 자동분류

#### 3-1 Data 및 Dataset

자동분류를 위한 Raw 데이터는 위험기반 항공안전감독 시스템을 통해 최근 4년(‘19.01~’22.06)간 발생한 국제항공운송사업자 대상 항공기 사고·준사고 및 항공안전장애 등의 안전이벤트데이터 5,409건을 수집한 것이며, Raw 데이터에 대한 예시는 표 4.와 같다. Data Field는 이벤트의 내용과 원인, 발생시기, 항공기 종류, 기번 등의 내용이 기술되어 있으며, Category Field는 4개(Field, Occurrence, Cause, Hazard) 카테고리로 구분되어 각 이벤트마다 표 3.의 표준분류체계를 기준으로 카테고리 명이 레이블링 된다. 본 논문에서 수집된 Raw 데이터의 레이블링 개수는 1차 Field - 2개, 2차 Occurrence - 14개, 3차 Causes-42개, 4차 Hazard-189개 카테고리로 구분된다. 이후, Null 데이터 삭제 및 불용어 처리를 통해 최종 5,041개의 학습 데이터를 추출하였다.

표 4. 로우데이터 (예시)

Table 4. Sample of RawData

Category	Data Field	Category Field			
		Field	Occurrence	Cause	Hazard
Sample Value	20XX-XX-XX XX:XX (UTC), HLXXXX (B000-XF/XX000) A non-scheduled cargo flight from YEG to KIX exceeded the No.3 Engine EGT Limit at an altitude of approximately 15,000 ft during takeoff climb and the "ENG 3 FAIL" message was displayed. So, pilots shut down the engine and returned aircraft to YEG airport after fuel dumping.	Air operation	Maintenance	Faulty parts(Engine)	Faulty parts (Normal operation after replacement of the faulty part)
Count		2	14	42	189

표 5. 실험 데이터 셋

Table 5. Dataset for Experiment

Data set	Data	Data Cnt.
D1	EventData + (Field, Occurrence, Causes, Hazard)	5,041
D2	EventData + Field	20,164
	EventData + Occurrence	
	EventData + Causes	
	EventData + Hazard	
D3	EventData + Field	20,164
	(EventData, Field) + Occurrence	
	(EventData, Field, Occurrence) + Causes	
	(EventData, Field, Occurrence, Causes) + Hazard	
D4	EventData + (Occurrence, Causes, Hazard)	20,164
	EventData + (Field, Causes, Hazard)	
	EventData + (Field, Occurrence, Hazard)	
	EventData + (Field, Occurrence, Causes)	

본 연구에서는 분류의 성능 향상을 위해 표 5.와 같이 4가지 데이터 셋으로 설정하여 연구를 진행하였다. 데이터 셋의 구성 방식은 분류 대상인 이벤트데이터와 각 카테고리의 결합형태로 구분한다. 이는 논문에서 제안한 분류모델이 학습할 학습데이터의 구성형태에 따른 성능의 차이를 확인하고자 한 것이다. 예를 들어, D1의 학습데이터는 한 이벤트에 대해 4개의 카테고리가 한꺼번에 결합되고 1회의 입력 데이터로 처리되어 총 5,041개의 데이터를 학습하는 것이고, D2는 한 이벤트에 각 카테고리별로 결합되고 4회의 입력데이터로 처리되어 총 20,164개의 데이터를 학습하는 것이다. 이러한 방식으로 D3는 이벤트 데이터에 각 카테고리가 1개씩 추가되며, D4는 이벤트데이터와 1가지 카테고리가 제외된 셋이다.

3-2 성능 실험 환경

본 논문에서 Fine-tuning된 모델의 최적화 알고리즘은 AdamW Optimizer를 사용하고, 설계모델이 예측한 값과 원래 값의 오차를 판단하는 손실함수(loss function)로는 cross entropy를 이용하였다. 모델의 임베딩 층(embedding layer)은 768개, 초기 학습률(learning rate)은 0.00002로 설정하고, 1회당 훈련 가능한 데이터 개수(batch size)는 700, 전체 데이터 셋의 학습 횟수(epoch)는 35로 설정하였다. 그리고, 데이터의 비율은 훈련(8) : 테스트(1) : 검증(1)로 분할하여 실험을 진행하였다.

3-3 성능 실험 결과

표 6.은 본 논문에서 제시한 4가지 데이터 셋의 성능을 나타낸 것으로, 각 지표는 제안한 모델의 평가를 수치화 한 것이다.

표 6. 데이터 셋 별 평가결과

Table 6. Evaluation Result of each Dataset

Data Set	Accuracy	F1-score	Precision	Recall
D1	0.259	0.130	0.106	0.169
D2	0.618	0.465	0.482	0.449
D3	0.684	0.518	0.537	0.500
D4	<b>0.792</b>	<b>0.691</b>	0.687	0.696

정확도(Accuracy)는 전체 예측 데이터에서 실제 올바른 데이터 대비 참으로 예측한 데이터의 비율을 의미하는 것으로, 4가지 데이터 셋의 순위가 D4(79.2%) > D3(68.4%) > D2(61.8%) > D1(25.9%) 순서로 확인되었다. 그리고, F1-score는 Precision(정밀도)와 Recall(재현율)의 조화평균으로, 정밀도는 분류모델이 올바른 값이라고 예측한 결과 중 실제 참인 값의 비율을, 재현율은 실제 참인 데이터로부터 분류모델에 의해 올바른 값이라고 예측한 비율을 의미한다. 결국, F1-score 순위 또한 D4가 69.1%로 가장 높았는데 실제 참인 데이터 중 올바른 데이터를 예측한 비율이 69.6%, 올바른 데이터로 예측한 값 중 실제 참인 비율이 68.7%인 것으로 측정되었다. 이때, D1 데이터 셋은 4가지 측정 지표 모두 가장 낮게 나타났는데 이는 데이터의 구성방식이 이진 분류에 적합한 형태인 것으로 판단된다.

한편, D4의 정밀도(68.7%)와 재현율(69.6%)의 비슷한 측정값은 항공안전데이터 분류에 있어 정답 데이터를 바르게 예측한 것과 정답이라고 예측한 것에서 올바른 데이터와의 차이가 매우 낮다는 것을 의미한다. 실제 항공안전데이터의 분류 목적이 이미 발생한 사건·사고를 바탕으로 위험도가 높은 분야를 찾아 사전 점검함으로써 발생 가능한 막대한 자원손실을 예방함에

표 7. 제안모델과 기존 연구모델과의 정확도 비교

Table 7. Comparison the accuracy of suggested model and each Model

Algorithms	Accuracy	Remark
KoBERT	<b>79.2%</b>	Suggested
Random Forest	76.0%	Ref. [5]
one-hot+CNN	56.8%	
word2vec+CNN	64.9%	Ref. [6]
kg2vec+CNN	68.6%	
Naive Bayes	55.6%	
Support Vector Machine	65.2%	Ref. [7]
KGQA	56.0%	
BERT-QA	22.7%	
GPT3_QA	76.0%	
KGQA+BERT-QA	63.0%	
<b>KGQA+GPT3-QA</b>	<b>85.3%</b>	

있기 때문에 제안 모델을 통해 도출된 올바른 분류데이터의 균형 잡힌 밸런스가 본 논문에서 제시한 자동분류 방안으로 적합함을 시사한다. 이와 함께, 표 7.은 기존 연구모델과의 정확도를 비교한 것으로, 제안 모델의 정확도가 KGQA+GPT3-QA 모델에 비해 6.1% 차이가 있음을 보인다. 그러나, 영어 자연어처리에 최적화된 KGQA+GPT3-QA 모델과 다르게, 제안 모델은 교착어의 특징을 갖는 한글 자연어처리를 위해 KoBERT를 fine-tuning한 모델이며, 표 7.의 알고리즘 중 BERT에 기반한 BERT-QA, KGQA+BERT-QA 모델과 한글 자연어를 처리하는 Random Forest 모델보다 성능이 우수함을 알 수 있다.

최종적으로, 최근 4년간 수집된 5천 건 이상의 데이터를 학습하고 결과를 예측하는데 소요된 시간은 3시간이었으며, 유사하거나 동일한 이벤트에 대해서도 일관된 결과값을 추출하였음을 표 8.을 통해 확인할 수 있다. 이러한 결과는 제안 모델만으로도 항공안전 이벤트에 대한 정확하고 신속한 분류가 가능함을 시사하며, 휴먼 분류가 갖는 시간과 비용, 분류 결과의 통일성 문제를 효과적으로 개선할 수 있을 것으로 기대한다.

3-4 오류분석

표 9.는 예측된 데이터 중 fail 된 데이터 일부를 발췌한 내용

표 8. 유사 이벤트 데이터에 대한 예측 테이블(샘플)

Table 8. Prediction Result Table about Event Data of Similar

NO	Overview	Answered Data				Predicted Data			
		Field	Occurrence	Causes	Hazard	Field	Occurrence	Causes	Hazard
1	The results of checking the status of major manual revision according to the standardization of maintenance manual are mostly appropriate, but in the case of training program for maintenance education, some contents are not reflected in the the standardization of maintenance manual.	A i r operation	Maintenance	Regulation , Rule and Manual	Improper Regulation , Rule and Manual	A i r operation	Maintenance	Regulation , Rule and Manual	Improper Regulation , Rule and Manual
2	Maintenance manual has been completely revised and many has been changed because of the standardization work, but the table of contents was not changed in accordance with operation criteria.	A i r operation	Maintenance	N o n compliance with applicable regulations	Time/Business Pressure	A i r operation	Maintenance	N o n compliance with applicable regulations	Improper Regulation , Rule and Manual
3	As a result of checking the safety glasses stored in the tool room of the maintenance contractor, the No.2 lens was broken, so it is corrected on site to be replaced.	A i r operation	Maintenance	N o n compliance with applicable regulations	Inappropriate of incorrect use of tools for the task	A i r operation	Maintenance	N o n compliance with applicable regulations	Lack of, or poor maintenance release
4	The status of right forward Escape Slide Bottle Pressure Viewer was not worked properly, so improvement is recommended.	A i r operation	Maintenance	Faulty defect	Faulty parts Normal operation after replacement of the faulty part	A i r operation	Maintenance	Faulty defect	Lack of, or poor maintenance release

으로, 모델이 바르게 예측했음에도 불구하고 휴먼에 의한 분류 값과 서로 상이하어 fail 처리가 된 데이터 예시이다. 예를 들어, 표 8. NO 1의 Overview 예시를 보면, 위탁 정비사에게 초기 교육만 시키고 정기 교육과정을 제공하지 않고 있다는 내용에 대해 분류모델은 “교육 훈련 프로그램 미흡”이라는 올바른 결과를 예측하였다. 그러나 휴먼에 의해 이미 분류된 값은 “작업 카드 작성 시 부정확한 정비 데이터 또는 표기 오류”였기 때문에 분류모델의 예측 결과가 올바른 값을 알 수 있다. 그리고 NO 3은 정상적인 비행 과정에서 타 항공기의 근접 비행 또는 관제기관에서의 관제제어 미흡인 내용이었고 Overview에 기술된 내용 자체가 “관제기관 또는 상대 항공기”로 기술되어 원인 분류의 결과는 2가지 경우 모두 올바른 결과에 해당하며, 분류 모델 또한 정확한 데이터를 예측한 것을 알 수 있다.

IV. 결 론

4-1 결론

본 연구에서는 항공기 사고·준사고 및 안전장애 내용을 담고 있는 위험기반 항공안전데이터의 Field, Occurrence, Causes, Hazard 분류를 예측하기 위한 자연어처리 알고리즘 기반의 자동분류 모델을 제시하였다. 이를 위해 5천 건 이상의 데이터를

표 9. 오류 데이터(샘플)

Table 9. Failed data sample

NO	Overview	Answered Data				Predicted Data			
		Field	Occurrence	Causes	Hazard	Field	Occurrence	Causes	Hazard
1	Maintenance Contractor VAECO mechanics are not provided with regular training courses except initial ones. For example China Eastern Airlines is providing Recurrent Training Outstation courses to VAECO Mechanics and has verified completion records.	A i r operation	Maintenan ce	Non-compliance with procedures	Imprecise maintenance data or transcription errors when creating job-card	A i r operation	Maintenan ce	Non compliance with applicable regulations	Lack of or poor training program
2	Defects in the Green Servo Control driving the Rudder Surface of the aircraft Primary Flight Control	A i r operation	Maintenan ce	Faulty parts	Faulty parts Normal operation after replacement of the faulty part	A i r operation	Maintenan ce	Faulty defect	Flight Control ATA 27 defect
3	Our aircraft was in the normal flight process, but it is presumed to have been caused by the corresponding ATC or the other aircraft.	Non-Air operation	O t h e r Aircrafts	Aircraft Proximity	A C A S T A / R A Warning rate	Non-Air operation	Airport Operation	Aircraft Proximity	A C A S T A / R A Warning rate
4	Procedural mistakes due to lack of skills during training flight	A i r operation	F l i g h t Operation	Error	Improper execution of procedures in all flight phases	A i r operation	F l i g h t Operation	Error	Lack of or poor training program

수집하여 4가지 타입의 데이터 셋을 설계하고 Fine-tuning 된 모델에 기계학습을 시킨 결과 D4 데이터 셋의 예측 결과가 79.2%로 가장 높은 정확성을 보였다. 또한, 유사 이벤트 내용에 대해 같은 결과를 예측하였고, fail 된 데이터 중 일부는 휴먼 에러에 의한 오류임을 예측된 결과를 통해 확인할 수 있었다. 결국, 본 논문에서 제시한 자동분류 모델이 항공안전데이터를 분류에 있어 휴먼 분류가 갖는 시간과 비용을 절감하고 분류 결과의 통일성 문제를 개선함과 동시에 대량의 항공안전 이벤트에 대해서도 신속하고 정확한 자동분류가 가능함을 시사한다.

4-2 한계점 및 향후 과제

본 논문에서는 자연어처리 알고리즘에 기반한 항공안전데이터의 자동분류 모델을 제시하여 그 가능성을 확인하였다. 그러나, 예측된 결과 중 올바른 데이터 임에도 불구하고 휴먼 에러로 인해 fail 된 데이터의 정확한 비율과 원인 파악에 시간적 한계가 있다. 이는, 논문에서 제시한 분류모델의 효율성 향상을 위한 요소 중 하나로써 심도 있는 연구 분석이 필요하고, 실제 fail 된 데이터의 문맥과 단어 패턴 등을 분석하여 그에 적합한 모델을 제시할 필요가 있는 것으로 판단한다.

이에 향후 연구계획에는 fail 된 데이터의 정확한 분석과 여러 알고리즘을 조합한 분류모델을 설계 및 제시하여 성능을 최적화시키고자 한다.

References

[1] MOLIT(Ministry of Land, Infrastructure and Transport), *The Aviation Safety White Book*, pp. 29-35, 2022.

[2] E. J. Kim, “Study of the Introduction on the Aviation Safety Data Protection System,” *The Korean Journal of Air & Space Law and Policy*, Vol. 33, No. 1, pp. 81-120, 2018.

[3] Ministry of Land, Transport & Maritime Affairs Republick of Korea. Aviation Safety Inspector Manual. Korean Law Information Center [Internet]. Available: <https://law.go.kr/LSW/admRulLsInfoP.do?chrClsCd=&admRulSeq=2100000212305>.

[4] KIAST(Korea Institute of Aviation Safety Technology) : Final Report on the Development of System-based Aviaion Safety Oversight Support Technology, MOLIT, Technical Report OTKCRK210001, 2020.

[5] C. S. Lee, Z. M. Paing, H. M. Yeo, D.S. Kim, and H.J. Baik, “Development of a Prediction Model and Correlation Analysis of Weather-induced Flight Delay at Jeju International Airport Using Machine Learning Technique,” *Journal of the Korean Society for Aeronautical and Flight Operation*, Vol. 29, No. 4, 2021, pp. 1-20.

- [6] A. Agarwal, R. Gite, S. Laddha, P. Bhattacharyya, S. Kar, A. Ekbal, P. Thind, R. Zele, and R. Shankar, "Knowledge Graph - Deep Learning: A Case Study in Question Answering in Aviation Safety Domain," in *Proceeding of the LREC 2022 : 14th Conference on Language Resources and Evaluation*, Marseille, arXiv:2205.15952 [cs.CL], 2022.
- [7] W. Zhang, H. Shi, Y. Yang, and Y. Luo, "Research on the Classification of Aviation Safety Reports Based on Text and Knowledge Graph," *Journal of Physics: Conference Series*, Vol. 1646, No. 1, 2020. pp. 1-6.
- [8] K. H. Kim, *Natural Language Processing with PyTorch*, Hanbitmedia, p.520, 2019.
- [9] H. Sak, A. Senior, and F. Beaufays, "Long short-term memory recurrent neural network architectures for large scale acoustic modeling," in *Proceeding of the 15th Annual Conference of the International Speech Communication Association on Computer Science*, Singapore, pp. 338-342, 2014.
- [10] Colah's blog. Understanding LSTM Networks [Internet]. Available: <http://colah.github.io/posts/2015-08-Understanding-LSTMs/>
- [11] J. H. Lee, The review of Deep learning, Master's Thesis, Ewha Womans University, Korea, Dec. 2018.
- [12] M. E. Peters, M. Neumann, M. Iyyer, M. Gardner, C. Clark, K. Lee, and L. Zettlemoyer, "Deep Contextualized Word Representations," in *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, Louisiana, pp. 2227-2237, Mar. 2018.
- [13] J. Devlin, M. W. Chang, K. Lee, and K. Toutanova, "BERT:Pre-training of Deep Bidirectional Transformers for Language Understanding," in *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, Minnesota, pp. 4171-4186, May. 2019.
- [14] S. H. Hwang, D. H. Kim, "BERT-based Classification Model for Korean Documents," *The Journal of Society for e-Business Studies* Vol. 25, No. 1, 2020, pp. 203-214.
- [15] J. Y. Choi, H. S. Lim, "E-commerce data based Sentiment Analysis Model Implementation using Natural Language Processing Model," *Journal of the Korea Convergence Society*, Vol. 11., No. 11, 2020, pp. 33-39.
- [16] H. R. Cho, H. Y. Im, J. W. Cha, and Y. M. Yi, "Automatic Score Range Classification of Korean Essays Using Deep Learning-based Korean Language Models - The Case of KoBERT & KoGPT2," *Journal of the International Network for Korean Language and Culture*, Vol. 18, No. 1, 2021, pp. 217-241.



**양 성 훈 (Sung-Hoon Yang)**

2007년 2월 : 목포대학교 정보보호학과 (공학석사)  
 2007년 3월 ~ 2016년 2월 : 한국건설기술연구원 ICT융합연구실 석사후연구원  
 2016년 11월 ~ 2020년 10월 : (주)Innerinfo 기술연구소 소장/부장  
 2020년 11월 ~ 현재 : 항공안전기술원 데이터분석연구센터 선임연구원  
 ※관심분야 : 센서네트워크, 클라우드 컴퓨팅, 시스템 정보분야



**최 영 (Young Choi)**

2019년 2월 : 숭실대학교 정보통신전자공학부 (공학사)  
 2016년 11월 ~ 2017년 12월 : 한국지역정보개발원 지식정보부 선임  
 2020년 2월 ~ 2020년 8월 : 한국항공협회 항공산업정보실 대리  
 2020년 9월 ~ 현재 : 항공안전기술원 데이터분석연구센터 연구원  
 ※관심분야 : 항공안전, 데이터 분석, 자연어분석



**정 소 영 (So-Young Jung)**

2019년 5월 : Embry-Riddle Aeronautical University Aeronautics(이학사)  
 2021년 1월 ~ 현재 : 항공안전기술원 데이터분석연구센터 연구원  
 ※관심분야 : 항공안전, 데이터 분석, 빅데이터



**안 주 현 (Joohyun Ahn)**

한국항공대학교 항공운항관리학과 박사수료  
 항공안전기술원 데이터분석연구센터 선임연구원  
 ※관심분야 : 항공안전, 위험기반안전감독, 빅데이터 분석