

# 인공지능을 활용한 교통사고 발생 예측에 대한 연구

김가을, 김정현, 손혜지, 김도현\*

부산가톨릭대학교

## A Study on the Prediction of Traffic Accidents Using Artificial Intelligence

Ga-eul Kim · Jeong-hyeon Kim · Hye-ji Son · Dohyun Kim\*

Catholic University of Pusan

E-mail : autumn2663@naver.com / kimjeonghyeon1368@gmail.com / kuy06046@naver.com / dohyun@cup.ac.kr

### 요 약

국민의 안전을 위해 교통사고를 방지하고자 교통 규제는 계속 확대되고 있지만, 교통사고는 여전히 줄어들지 않고 있다. 본 연구에서는 기상청의 날씨 예측 데이터, 도로교통공단의 요일, 시간대, 장소 별 교통사고 발생 데이터, 특정 위치 정보 등 다양한 요인들의 연관관계를 인공지능을 활용하여 분석함으로써 특정 시간, 장소에 대한 교통사고 발생 확률을 예측하고자 한다. 본 연구는 이전의 수많은 교통사고 발생에 대한 객관적인 데이터와 기존의 다른 연구들에서 활용되지 않은 다양한 추가 요소들을 접목시켜 더욱 향상된 교통사고 발생 확률 예측 모델을 도출한다. 본 연구 결과는 국민의 안전한 삶을 위한 다양한 교통 관련 서비스에 유용하게 활용될 수 있을 것이다.

### ABSTRACT

Traffic regulations are expanding to prevent traffic accidents for people's safety, but traffic accidents are not decreasing. In this study, the probability of traffic accidents occurring at a specific time and place is estimated by analyzing various factors such as weather forecast data from the Meteorological Agency, day of the week, time of day, location data, and location information. This study combines objective data on the occurrence of numerous previous traffic accidents with various additional elements not considered in previous studies to derive a more improved traffic accident probability prediction model. The results of this study can be effectively used for various transportation-related services for the safety of people.

### 키워드

Traffic accident prediction, Artificial intelligence, Deep learning, Machine learning

### 1. 서 론

국민의 안전을 위해 교통사고를 방지하고자 교통

규제는 계속 확대되고 있지만, 도로교통공단에  
서 제공하는 자료에 따르면 교통사고는 여전히 줄  
어들지 않는 상황이다. 다음의 <표 1>은 2010년도  
부터 2019년까지의 교통사고 발생 건수에 관한 통  
계자료다[1].

---

\* corresponding author

<표1> 2010~2019년도 교통사고 건수

분석지표	사건건수
2010	226,878
2011	221,711
2012	223,656
2013	215,354
2014	223,552
2015	232,035
2016	220,917
2017	216,335
2018	217,148
2019	229,600

또한 <표2>에 따르면, 기상상태라는 변수에 따라 교통사고 발생의 빈도수가 다르다는 것을 알 수 있다[2].

<표2> 기상상태별 교통사고 통계자료

구분	발생건수			사망자수		부상자수	
	(건)	구성비	상대사망사과율	(명)	구성비	(명)	구성비
계	5,165	100	1.00	5,392	100	3,273	100
맑음	4,007	78.7	0.94	4,233	78.5	2,545	77.8
흐림	392	7.6	1.56	407	7.5	253	7.7
비	587	11.4	1.18	625	11.6	420	12.8
안개	31	0.6	4.25	35	0.6	19	0.6
눈	61	1.2	0.96	63	1.2	33	1.0
기타/불명	27	0.5	0.83	29	0.5	3	0.1

따라서 본 연구에서는 날씨라는 하나의 변수뿐만 아니라 요일, 시간대, 장소별 교통사고 발생 데이터 등 다양한 요인들의 연관관계를 인공지능을 활용하여 분석함으로써 특정 시간, 장소에 대한 교통사고 발생 확률 예측 모델을 도출하고자 한다.

또한 본 연구의 결과를 교통 관련 서비스로 활용함으로써, 국민에게 사고 위험도를 알리고 안전한 삶을 제공하고자 한다.

## II. 관련 연구

사례 기반 추론은 과거에 적용되었던 사례와 그 결과들을 참조하여 새로운 사례에 대한 결과값을 예측하는 것으로 새로운 사례와 가장 비슷한 과거의 사례를 일부 추출하여 추출한 사례의 특징 지식을 통해 문제를 해결하는 방식이다. k-NN 알고리즘으로 교통사고 발생 예측이 가능하다.

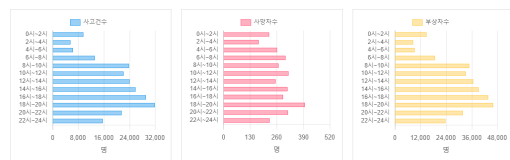
리즘으로 교통사고 발생 예측이 가능하다.

K-NN 알고리즘은 비모수모형으로 종속변수와 설명변수 간의 관계를 정의하지 않고 예측하고자 하는 관측치가 있을 때 이와 가장 가까운 거리에 있는 k 개의 데이터를 결정한 후 이 특성들을 이용해 관심 예측치를 예측 하는 과정을 거치는 알고리즘이며 단순하고 효율적이며 훈련 데이터의 양이 많을수록 효과적이기 때문에 대용량의 데이터를 활용하는 연구에 적합하다. 이를 이용한 사고 위험 예측의 설계는 전체적으로 변수의 선정과 불균형 데이터 처리, 데이터 정규화, 분석 데이터 설정, k-NN 알고리즘에 필요한 하이퍼파라미터 선정의 과정으로 이루어졌다. 거리 기반 알고리즘에서 중요한 데이터의 정규화 과정을 거쳐 분석 데이터 설정하며 파라미터 값은 k값 설정, 거리 산정 방법, 거리 가중치 산정 방법, 변수 가중치 산정 방법으로 정의하였고 k값 설정과 거리 가중치 산정 방법은 경험적인 방법을 통해 적합한 하이퍼파라미터를 선정하였으며 사고위험 예측은 실제 사고 발생은 70%이상 가능함을 나타냈다[3].

UNIST는 LSTM 알고리즘을 활용하여 울산시 기준으로 실시간으로 15분 단위로 도로교통 상황을 예측할 수 있는, 명확한 도로 상황 예측 정보를 제공하는 기술이 개발했다. 인공지능 기술을 이용해 교통정체의 원인을 파악하고, 특정 도로의 가까운 미래 상황을 예측해 시각화하는 시스템을 개발하였다. 기존 확률통계 분석에 딥러닝 기술을 도입하여 특정 도로구간에서 15분 후에 벌어질 교통상황을 평균 4km/h내외의 오차로 예측이 가능하다[4]. 여기서 설명하는 기술은 미래 상황을 예측하는 것이기 때문에 본 논문에서 이 서비스를 이용할 수 있을 것이라 판단하였다. 따라서 이 기술에 사용되었던 시스템을 활용하여 본 논문의 서비스에 적용 시키고자 한다.

## III. 향상된 교통사고 발생 예측을 위한 연구

도로교통공단은 TASS (Traffic Accident Analysis System)라는 교통사고분석시스템을 통해 다양한 교통사고에 대한 통계 정보를 제공한다. 이것은 요일별, 시간대별, 연령층별, 도로종류별, 사고 유형별 정보 등 다양한 정보를 포함한다.



<그림 1> TASS의 시간대별 교통사고 정보

우리는 이러한 TASS로부터 수집한 정보들과 기

상정보로부터 수집할 수 있는 실시간 기상 정보, 그리고 실제 운전자의 연령대와 시간대별 운행 예상 지역 정보 등을 결합하여 전처리한다.

그 후 기존에 연구된 70% 정도의 정확도를 확인한 K-NN 알고리즘과 그 후 더 개량된 성능을 갖는 LSTM 알고리즘의 결과를 결합함으로써 K-NN과 GRU 알고리즘을 사용할 것이다.

우리의 연구는 기존 연구들에서는 고려하지 않았던 교통사고 예측에 필요한 추가적인 많은 정보들을 포함하고 있기 때문에 운전자의 상황에 맞는 더욱 정교한 교통사고 발생 예측을 할 수 있을 것으로 예상된다.

#### IV. 결 론

본 연구는 교통 규제가 확대되고 있음에도 불구하고, 교통사고가 줄어들지 않고 있다는 문제를 인식했다. 교통사고 발생 수를 줄이기 위해서는 날씨, 요일, 장소 등 여러 변수에 따른 교통사고 발생 확률을 예측하는 것이 필요하다. 따라서 본 연구에서는 다양한 요인들의 연관관계를 k-NN, GUR 알고리즘을 통해 분석함으로써 교통사고 발생률을 예측하고자 한다. 또한 본 연구에서 얻은 결과를 내비게이션 서비스와 결합하여 길안내와 교통사고 위험 예측 서비스를 함께 제공함으로써 교통사고 예방에 도움이 되고자 한다.

#### References

- [1] TASS Traffic Accident Analysis System - Road Traffic Authority <https://www.koroad.or.kr/>
- [2] Young-Suk Chung, Rack-Koo Park, Jin-Mook Kim, "Study on predictive modeling of incidence of traffic accidents caused by weather conditions" Journal of the Korea Convergence Society, vol.5, no.1, Mar.31.2014, pp.9 - 15
- [3] Min-Ji Kang, "A Study on Traffic Accident Risk Prediction Methodbased on k-NN Algorithm Using DSRC Data" Department of Urban Planning and Transportation Engineering Graduate School Keimyung University, Feb.2018
- [4] Chug-Gi Lee, Yeon-Jun Kim, Seung-Min Jin, Dong-Min Kim, Ross Maciejewski, David Ebert, Sung-Ahn Ko, "A Visual Analytics System for Exploring, Monitoring, and Forecasting Road Traffic Congestion", IEEE Transactions on Visualization and Computer Graphics, vol.26, no.11, p.3133 - 3146, Nov.12 020