https://doi.org/10.15813/kmr.2024.25.1.011

클러스터링 기법을 이용한 이륜차 사고의 특징 분류

Classification of Characteristics in Two-Wheeler Accidents Using Clustering Techniques

허원진 (Heo, Won-Jin) 경희대학교[1)]
강진호 (Kang, Jin-ho) 경기대학교[2)]
이소현 (Lee, So-hyun) 경기대학교[3)]

〈 국문초록 〉

최근 배달문화의 확산으로 이륜차 수요가 증가하면서 이륜차 운행도 함께 증가하고 있다. 이륜차 운행은 혼잡한 교통 상황이나 경제적으로 효율적이지만 이륜차 난폭 운전과 명확하게 정립되지 않은 이륜차에 대한 교통 법규로 이륜차 사고는 새로운 사회문제로 나타나고 있다. 이륜차는 차체 특성 상 치사율이 높기 때문에 이륜차 사고가 발생하면 그 심각성 및 위험이 크다. 그러므로, 이륜차 사고에 대한 특성을 분석함으로써 이륜차 사고의 특성을 제대로 파악하는 것이 필요하다. 그리하여, 본 연구에서는 이륜차 사고 데이터를 기반으로 K-prototypes 알고리즘을 이용하여 이륜차 사고의 특성을 분류하였다. 그 결과, 이륜차 사고 특성에 따라 4개의 군집으로 분류되었다. 각 군집마다 사고발생 도로, 주요 위반 법규, 사고 유형, 사고 발생 시간 등에서 다른 특성을 나타내었다. 이를 기반으로 이륜차 사고 예방을 위한 구체적인 방안을 제안한다. 각 사고 특성에 따른 단속 방법 및 규율을 개정함으로써 수도권 지역의 이륜차 사고 발생을 최소화하고 궁극적으로는 도로 안전성 향상에 기여한다. 더불어, 머신러닝 기법을 도시교통 및 안전 분야에 적용함으로써 관련 문헌 확장에도 기여한다.

주제어: 이륜차 사고, K-prototypes, 머신러닝, 클러스터링

1) 제1저자, 9341kkk@khu.ac.kr
2) 제2저자, jinho5080@gmail.com
3) 교신저자, dlthgus1010@gmail.com

논문접수일 2024년 03월 01일 | 1차 수정 2023년 03월 20일 | 게재확정일 2024년 03월 22일

1. 서론

전 세계적으로 배달 서비스가 확장되면서 이륜차 수요는 빠르게 증가하고 있다. 바쁜 현대 사회에서 물품이나 배달음식 등의 빠른 배송을 원하는 수요가 커지면서 혼잡한 교통 상황에서 신속한 배송을 위한 이륜차 활용이 증가하고 있다. 특히 COVID-19 발생 이후 혼밥문화가 확산되면서(김진솔 등, 2021) 배달앱에 월 평균 15만원 이상 사용하는 인구가 2018년 32.4%에서 2020년 62.8%로 30.4% 증가하면서 배달 산업이 성장을 이루면서 배달 이륜차 라이더 역시 증가하였다(KOSIS, 2021).

젊은 세대들은 효율적이고 비교적 경제적인 이동 수단이나 취미의 목적으로 이륜차를 선택하는 경향이 두드러지고 있다. 이륜차란 도로교통법 제2조 제19항에 1인 또는 2인의 사람을 운송하기에 적합하게 제작되어진 이륜의 자동차이다. 특히, 도로교통법 상 배기량이 125cc 이상이면 이륜차, 125cc이하이면 원동기장치 자전거로 구별된다(도로교통법, 2024). 국토교통부의 이륜차 신고 현황을 보면 2024년 1월 기준 약 220만대가 등록되어 있다(국토교통부, 2024). 이륜차 운전자의 대부분은 안전수칙 불이행으로 인한 교통사고가 많으며, 특히 신호 위반, 중앙선 침범, 안전모 미착용 등으로 그 사고가 심각해지는 경우가 많았다(KBS뉴스, 2024).[1] 또한, 도로 교통 공단이 배달 이륜차의 위험운전 행동 유형을 분석한 결과, 배달 이륜차 10대중 4대가 위험 운전을 하고 있는 것으로 나타났다(M Today, 2023).

특히, 이륜자동차 배달업 시장 특성상 시간이 수입과 직결되어 있기 때문에 주문을 배당 받아야 하는 배달 기사 입장에서 강제 배당과 배송 시간의 압박으로 인해 위험 운전을 감수하며 무리한 주행을 하게 된다(이지선, 이희원, 2021). 최근 10년 동안의 이륜차 사고와 관련하여 연평균 사망자 수는 감소 추세이지만 연평균 사고 건수, 부상자 수는 증가하는 추세이다. 2022년, 이륜차 교통사고 건수는 총 18,295건이고, 사망자 수는 484명으로 전년대비 5.4% 증가하였다. 단속 현황으로서는 2019년 약 27만 5천건에서 2022년 47만 9천건까지 계속 증가하고 있다(공공데이터 포털, 2023). 전체 교통사고의 발생 건수나 부상자 수는 큰 변화가 없는 것에 비해, 이륜차 사고는 그 증가와 감소를 반복하는 경향을 보인다. 이는 이륜차 사고 예방에 관한 방안이 다른 교통 수단에 비해 명확히 규정되어 있지 않다는 것으로 해석할 수 있다.

2022년 기준 이륜차 사고는 서울, 경기 등 수도권 지역이 전체 사고 건수의 약 50% 정도를 차지하고 있다. 이륜차의 신호위반, 안전 운전 불이행 등 사고의 주요 원인은 이륜차의 주행자를 위협할 뿐만 아니라, 보행자 및 승용차주에게도 큰 위협을 끼친다. 이륜차 수요의 증가와 이를 통한 교통사고 발생의 지속적인 증가로 이륜차 사고 관련 연구들도 다양하게 수행되고 있다. 이륜차 사고에 관한 선행 연구들은 이륜차 특성, 이륜차 사고의 영향 요인, 이륜차 사고 치사율 분석 및 대안에 관한 연구들이 많았다(최정우, 금기정, 2014; 배현식, 김준용, 김주영, 김진태, 2015; 채범석, 2005). 하지만, 이륜차 사고의 특성에 따라 이륜차 사고를 분류한 연구는 드물었다. 특히, 이륜차는 주행할 수 있는 도로의 범위가 더 넓으므로 도로 유형별 특징과 그 밖의 이륜차 사고의 특징을 나타내는 데이터를 기반으로 한 분류가 필요하다.

그리하여, 본 연구에서는 도로형태 등과 같은 이륜차 사고 관련 데이터를 기반으로 이륜차 사고 특성을 분류하고, 이를 기반으로 이륜차 사고 예방을 위한 방안을 제시한다. 특히, 사고 지역(위치) 및 도로별 특성에 따른 실질적인 단속 방법과 사고 예방 방안을 제시한다. 본 연구에서는 현대 사회문제로 떠오르고 있는 이륜차 사고 예방을 위해 머신러닝 기법의 K-prototype 알고리즘을 활용하여 이륜차 사고의 특성을 분류한 것에 학술적 의의가 있다. 또한, 도시교통 분야에 새로운 기술을 적용한 융합

1) https://news.kbs.co.kr/news/pc/view/view.do?ncd=7893137&ref=A

연구로써 도시교통, 정보시스템 분야의 문헌 확장에 기여한다. 더불어, 본 연구결과는 수도권 지역의 이륜차 사고 예방을 위한 이륜차 단속 방법 및 규정을 개정하는데 활용할 수 있으며, 이는 사회적으로 안전한 교통 환경 조성에 기여한다. 나아가 본 연구를 통해 도로 교통 안전에 대한 사회적 관심을 보다 향상시킬 수 있을 것을 기대한다.

2. 개념적 배경

2.1. 이륜차 사고 특징

이륜차 특성상 다른 차종들에 비해 사고 발생 시 죽거나 다칠 확률이 월등히 높다. 도로교통공단에서 발표한 2023년판 교통사고 통계분석 보고서에 따르면 사망사고 차종별 현황에서 사망율을 보면 승합차나 화물차는 각각 67.5%, 65.9%로 나타나는데 비해 이륜차는 82% 사망률을 보인다. 이렇듯 이륜차 사고는 다른 사고에 비해 높은 치사율을 보이는 특징이 있다. 이륜차 사고의 시간대별 사고건수를 보면 18시~20시에 많이 발생하는 현황을 보였다.

이륜차 사고 사고유형별 현황(TASS 2022)으로는 전체 사고건수 중 차대차 사고가 72.7%로 가장 많이 차지하였고, 차대사람 사고가 17.7%, 차량단독 사고가 9.6%를 차지하였다. 차대차 사고 유형에서는 측면 충돌이 6,023건으로 가장 많았고, 차대사람 사고 유형에서는 횡단 중이 1,543건으로 가장 많았다.

이륜차 사고의 법규 위반별 현황으로는 총 18,295건 중 안전운전의무불이행이 9,655건으로 가장 많이 차지하고 있고 그 뒤로 신호위반(3,771건), 안전거리미확보(1,227건), 교차로운행방법위반(839건)순으로 나타나고 있다. 이륜차 사고의 도로, 환경적 특성으로는 교차로 형태에서의 사고 건수가 9750건으로 단일로에서 발생한 사고건수(7,504건)보다 많았다. 도로 선형별 현황으로 보면 직선 도로에서 전체 사고 건수의 92.3%를 차지하고 커브 및 곡각도로 사고는 6.2%였지만 이륜차 사고 사망자 수의 23.3%를 차지한 것으로 보아 커브길에서 일어나는 사고가 치사율이 높은 것을 확인할 수 있다. 또한 이륜차 사고가 발생한 도로 종류별 통계량을 보면 특별광역시도와 시도에서 발생 비율이 75%로 지방이나 외곽 도로보다는 수도권이나 도시지역에서 많이 발생하고 있다.

이륜차 운전자 수는 늘어나는 추세이고 거기에 촉박한 배달시간 등 여러가지 이유 때문에 이륜차 운전자들의 위협운전이 늘어나면서 이에 따른 교통사고도 증가하는 추세이다. 하지만 도로에 있는 단속 카메라들은 사륜차 운전자들에게 안전 운전을 강제할 수 있도록 하는 것이 주 목적이기 때문에 이륜차를 단속하기에는 한계가 있다. 이륜차 단속은 주로 현장 단속이나 '경찰청 SMART 국민제보'를 통한 시민들의 제보에 의존하고 있다. 현장 단속의 경우 인력 문제와 2차 교통사고 유발 우려 등의 어려움이 있다(조용빈 등, 2022). 그러므로, 이륜차 사고 예방을 위한 단속 카메라 방법이나 교통 법규에 대한 개정이 필요하다. 특히 인력이 지속적으로 배치해야 하는 현장 단속의 경우 최소의 인력으로 단속 효과를 극대화하기 위해서는 지역 및 시간에 따른 차별적 단속 인력 배치와 단속 방법에 대한 구체적인 방안이 필요하다.

2.2. 선행연구

배달 문화의 확산과 혼잡한 도시 교통 상황에서 빠르고 효율적인 이동이 가능한 이륜차는 계속해서 증가하고 있는 추세이다. 복잡한 도시 교통에서 효율적으로 운행할 수 있는 이륜차의 이점과 대비하여 이륜차의 위협운전이 교통사고의 문제로도 대두되고 있다. 이륜차 증가를 통한 이륜차 사고 문제는 사회문제 중 하나로 언급되고 있으며 이를 분석 및 해결하고자 하는 연구들도 수행

되고 있다. 다음 <표 1>은 이륜차 사고 관련 선행 연구를 나타내었다.

채범석(2005)은 이륜차 운전자를 대상으로 한 설문조사를 통해 이륜차 사고는 청소년의 사고빈도가 높고, 사고 시간대는 18시-20시에 가장 많이 발생한 것으로 발견하였다. 이러한 이륜차 사고 특성 분석을 통해 이륜차 안전운행을 위한 등록제도, 검사제도, 면허제도, 교육제도, 보험제도 등 제도개선 대책방안을 제시하였다. 배현식 등(2015)은 이륜차 사고의 발생 원인과 사고 특성을 파악하기 위하여 도로교통공단에서 제공하는 교통사고 통계자료를 이용하여 분석을 진행하였다. 분석 결과, 이륜차 사고는 교차로와 단일로에서 모두 차량과의 측면 직각충돌에 의해 발생한다는 특성을 발견하였으며, 전방주시 태만, 핸들 과다조작과 같은 안전운전 불이행 등이 사고의 원인이라고 분석하였다.

유진화 등(2020)은 이륜차사고 데이터 분석을 통해 교차로보다는 단일로 사고에서의 사고가 치사율이 더 높았으며, OECD 회원국과 비교한 결과 고령자의 이륜차 사망사고가 최하위국보다 3.7배 많아 그에 따른 정책이 필요한 상황임을 나타내었다. 연구결과를 기반으로 이륜차 사고를 감소시키기 위해 배달업 전반의 의식개선, 안전모 착용에 대한 홍보와 안전모 미착용에 따른 상향된 범칙금을 제안하였다. 이지선, 이희원(2021)은 이륜자동차 현황 및 법제도, 사업자제도, 보험제도 등을 분석하여 정비업, 배달업, 관리정책 등의 개선방안을 제시함과 동시에 개선방안 우선순위를 평가하여 개선방안 추진전략을 제안하였다. 최정우, 금기정(2014)은 순서형 프로빗 모형을 활용하여 이륜차 사고 심각도에 영향을 주는 요인들을 분석하였다. 유의한 변수로 사고발생시간, 사고유형, 도로선형, 이륜차 배기량이며, 좌측선형, 내리막선형 등의 요인이 심각도를 증가시키고, 이륜차의 구조적 특성상 단독사고의 경우여도 전복, 낙하, 시설물 충격으로부터 보호해줄 수 있는 차체가 없으므로 심각도에 긍정적인 영향을 주는 요인이라는 것을 알 수 있었다. 이를 통해 어떠한 요인들이 이륜차사고 심각도에 유의한 영향을 주는지 확인할 수 있었다.

〈표 1〉 선행연구

구분	연구	연구내용	연구방법
이륜차사고 영향요인 분석	최정우, 금기정(2014)	이륜차사고 심각도에 영향을 주는 요인을 파악 및 분석	순서형 프로빗모형
	Rifaat, S et al(2011)	이륜차사고 심각도에 영향 미치는 요인 분석	로짓모형, 혼합선택모형, 순서형 로짓모형
이륜차사고 현황 및 데이터분석	배현식 등 (2015)	이륜차사고 발생 원인과 사고특성 분석	통계 데이터 분석
	채범석 (2005)	이륜차사고 특성 분석 및 안전 운행 정책 개선방안 제시	데이터 분석, 설문조사
	유진화 등 (2020)	이륜차사고분석을 통해 교통사고 예방대책 도출	데이터 분석
	이지선, 이희원(2021)	이륜자동차 산업측면과 관리정책 측면에서 안전관리 문제 분석 및 개선방안 제시	자료조사
이륜차 단속 인식알고리즘	조용빈 등(2022)	이륜차 단속방안 제안 및 단속 가능여부 확인	딥러닝(광학 문자인식(OCR)) 기법
	인병철 등(2023)	이륜차 단속 기술 개발 후 현장실험 및 성능평가	CNN, LPD, LPR 기법
무인교통단속장비 적정 설치대수	소형준 등(2020)	무인교통 단속장비가 필요한 곳 도출 및 적정대수 산출	군집분석, 분산분석

Rifaat 등(2011)은 도로에서의 오토바이 충돌 심각도에 영향을 미치는 요인을 다양한 모델링 방법론을 활용하여 분석하였다. 분석 결과, 우측커브의 도로선형, 교차로에서의 좌회전, 트럭과의 충돌, 음주운전, 과속 충돌이 심각도를 높인다는 것을 발견하였다. 이를 기반으로 교차로의 안정성을 향상시키기 위해 투자해야 하며, 이륜차 운전자들의 위협운전방지를 위한 캠페인을 진행해야 한다고 결론을 내렸다. 조용빈 등(2022)은 이륜차 자격증 ID 체계를 활용한 단속 실험을 수행해 이 단속의 가능여부를 D-MESO 프로그램을 통해 확인하였다. 이륜차에 번호판처럼 자격증 ID를 부착시킨 후, 단속 시나리오대로 실험을 수행하며 OCR 번호 인식 처리 기법을 활용하여 번호판 문자열을 추출 및 인식하였다. OCR 분석 결과, 주관 시간대는 66.7%, 캠코더의 경우 100% 성공하였지만 야간 시간대의 경우에는 단속이 어려운 것으로 나타내었다.

인병철 등(2023)은 기존 이륜차 단속의 문제점을 검토한 후 CNN을 활용해 이미지 공간 정보를 유지하면서 학습이 가능하고 객체 검지 및 분류가 가능한 현장 실험을 수행하였다. 여기에서는 LPD 기법을 활용하여 번호판을 추출하고, LPR 기법을 통해 문자열을 구분하여 정보를 인식하는 기술을 활용하였다. 성능평가 결과, 주간 94.4%, 야간 100.0%로 나타나 AI기반의 영상분석 방식을 적용한 무인단속 기술의 도입이 가능하다는 것을 확인하였다. 소형준(2020)은 교통사고데이터와 [사망자 수 + 중상자 수]값인 KSI를 활용해 교통사고 심각도를 나타내는 ARI지표를 개발하고 일원배치 분산분석, 군집분석을 활용해 4가지 도로유형으로 분류하였다. 설치판단기준을 사고건수, 사고 심각도를 포함할 수 있게 유클리드 거리를 활용한 ARI 기준에 따라 전국의 사고 잦은 곳에 5,244대의 무인교통 단속장비 설치가 필요한 것으로 분석하였다.

이륜차 사고 관련 선행 연구들은 대부분 공공데이터와 통계자료를 이용하여 이륜차 사고의 주요 특성을 분석하였다. 즉, 이륜차 사고가 발생과 관련한 주요 요인을 도출하고, 이를 기반으로 이륜차 사고 예방에 대한 논의를 진행하였다. 하지만, 이륜차 사고가 발생하는 원인 요인들을 통합하여 분석하고, 이를 통해 사고별 특성을 파악한 연구는 드물었다. 이륜차 사고의 주요 요인을 파악하는 것도 중요하지만, 이륜차 사고의 특징이나 원인이 될 수 있는 요인들을 통합하고 이를 통해 사고의 특징을 분류한다면 각 사고 유형을 명확하게 파악할 수 있다. 이를 기반으로 이륜차 사고 특성에 따른 이륜차 단속 및 교통법규를 새롭게 개정할 수 있을 것이다. 따라서, 본 연구에서는 이륜차 사고의 특성을 뚜렷하게 파악하기 위해 군집분석을 수행해 각 사고유형을 명확하게 분류하고, 이를 기반으로 이륜차 사고 예방을 위한 구체적인 방안을 제시한다.

3. 연구방법론

3.1. 클러스터링 기법

클러스터링(Clustering)이란 탐색적 데이터 분석에서 널리 사용되는 기법 중 하나이다. 클러스터링은 레이블이 지정되지 않은 데이터 집합내에서 객체를 유사성을 가진 그룹들로 조직화하는 비지도 학습에 해당한다. 하나의 집합(Cluster)을 이루고 있는 객체들끼리는 유사하지만 다른 클러스터의 객체와는 유사하지 않다. 이 알고리즘은 다양한 알고리즘들이 있으며, 그 중 가장 일반적인 거리기반 알고리즘은 몇 가지 거리 측정법을 사용하여 유사도를 계산함으로서 이루어진다(A Tutorial on Clustering Algorithms, 2023).

클러스터링 기법은 클러스터의 구조적 종류에 따라 크게 계층적 클러스터링 기법과 비계층적 클러스터링 기법

으로 나뉜다. 계층적(hierarchical) 클러스터링 기법은 상향식(병합형) 또는 하향식(분할형) 방법이 있다. 상향식 방법은 처음 각 데이터들을 개별 클러스터로 간주한 후 가장 가까운 클러스터 쌍을 결합해 하나의 클러스터를 만들고 나머지 (n-2)개의 데이터들을 각각 하나의 클러스터를 이루도록 하는 과정이다. 결론적으로 모든 데이터들을 묶어 하나의 클러스터로 만드는 단계까지 나무모양의 계층적 구조를 형성하는 방법이다. 반면, 하향식 방법은 처음 모든 데이터들을 하나의 클러스터로 간주 후 모든 데이터가 개별적인 클러스터가 될 때까지 분할을 한다(A Survey of Advances in Hierarchical Clustering Algorithms and Applications, Amr Munshi, 2022). 비계층적 클러스터링 기법은 클러스터링 이전에 클러스터의 크기와 클러스터의 수 등 여러 기준을 설정한 후, 기준에 맞춰 데이터들을 여러 개의 클러스터로 나누어 적합한 클러스터로 재배치하는 기법이다(정영미, 2012).

본 연구에서는 비계층적 클러스터링의 K-prototypes를 사용했다. K-prototypes 군집분석 알고리즘은 연속형 데이터와 범주형 데이터가 동시에 존재하는 혼합 데이터를 클러스터링하는 방법이다(Huang, 1998). 연속형 변수에는 K-means, 범주형 변수에는 k-modes 방법을 사용한다. K-prototypes 알고리즘은 데이터를 k개의 클러스터로 군집화 하기 위해 cost function을 최소화하는 것을 목표로 한다. 이 알고리즘은 초기 단계에서 k를 지정하면 다음과 같은 네 단계를 순차적으로 수행한다. 첫째, k개의 부분집합으로 데이터를 분할한다. 둘째, 각 클러스터의 중심값으로 연속형 데이터는 평균값, 범주형 데이터는 최빈값으로 결정한다. 셋째, 각 데이터를 적절한 클러스터에 재할당한다. 이 때 연속형 데이터는 유클리드 거리를, 범주형 데이터는 적합한 비유사성 측정을 사용한다. 마지막으로, 전체 데이터 객체에 대한 유사성을 다시 테스트한다. 만약 데이터 객체가 현재 클러스터가 아닌 다른 클러스터와 유사성이 더 높게 나온다면 해당 클러스터로 다시 할당하고, 할당된 클러스터의 변화가 없다면 종료한다(김정화, 진서훈 2018; Ji et al., 2012). 본 연구에 활용된 데이터는 연속형 데이터와 범주형 데이터가 혼합된 데이터를 사용하므로 K-prototypes기법을 적용하며, 이를 통해 이륜차 사고 데이터를 분류한다.

3.2. 데이터 수집

본 연구의 데이터 분석에 사고가 발생한 지점의 명확한 위치, 시간, 도로형태 등의 사고 분석을 위한 데이터를 수집하여 사용하였다. 또한, 제공된 데이터의 신뢰성 및 추후 연구를 위한 데이터의 지속적인 최신화도 배제할 수 없었다. 따라서, 본 연구에서 가장 활용도가 높은 데이터를 수집하고자 다방면으로 데이터를 수집하였고, TAAS(교통사고 분석시스템)의 공공데이터가 본 연구와 가장 적합하다고 판단하여 활용하였다.

한국정보화진흥원의 연구(송석현, 2014)에 따르면 공공데이터는 두가지 특성을 가지고 있다. 첫 번째는 신뢰성이다. 데이터의 목적이 공공인 경우가 대부분으로, 왜곡이나 조작이 발생할 가능성이 적다. 두번째 특성은 주기성이다. 공공기관에서는 기간 단위로 국가 및 국민과 관련된 공공업무를 지속적으로 수행한다(권성아 등, 2015)는 본 연구에 반드시 필요한 특성을 가지고 있는 데이터였다. 이러한 특성들뿐만 아니라 교차로 및 단일로에 관한 정보가 포함되어 있어 사고발생지의 특성을 분류하기에 본 연구 목표에 가장 적합한 데이터라고 판단되어 수집하여 연구를 진행하였다.

한국정보화진흥원의 연구에 따르면 공공데이터는 두가지 특성을 가지고 있다. 첫 번째는 신뢰성이다. 데이터의 목적이 공공인 경우가 대부분으로, 왜곡이나 조작이 발생할 가능성이 적다. 두번째 특성은 주기성이다. 공공기관에서는 기간 단위로 국가 및 국민과 관련된 공공업무를 지속적으로 수행한다.

〈표 2〉 데이터의 종류 및 형태

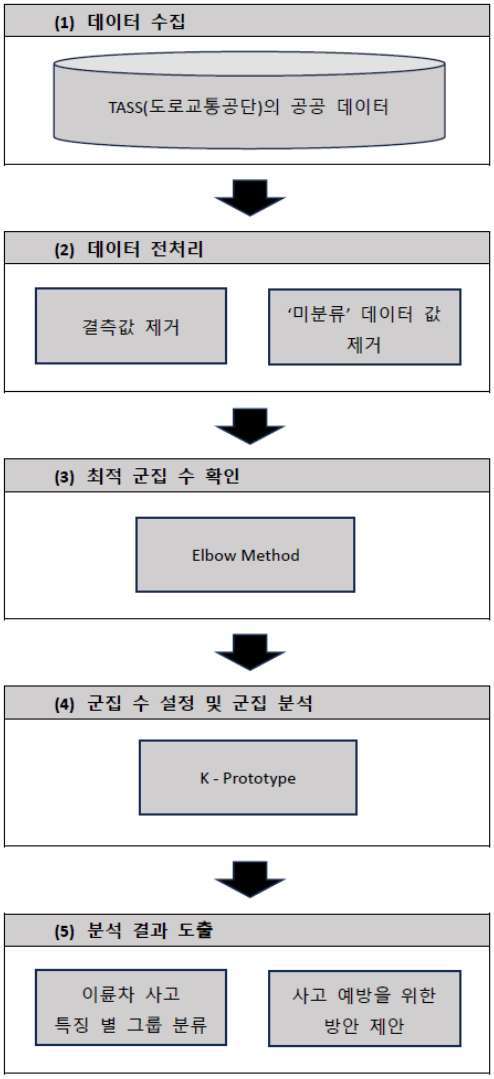
구분	데이터 이름	데이터 형태	데이터 예시
1	노면 상태	범주형 데이터	젖음, 건조 등
2	법규 위반	범주형 데이터	신호 위반, 안전거리 미확보 등
3	도로 형태	범주형 데이터	교차로, 단일로 등
4	날씨	범주형 데이터	맑음, 비, 눈 등
5	요일	수치형 데이터	월요일 ~ 금요일
6	시간	수치형 데이터	00시 ~ 23시
7	사고 유형	범주형 데이터	차대차, 차대 사람
8	지역(시)	범주형 데이터	성남시, 안성시 등
9	지역(구)	범주형 데이터	분당구, 단원구, 강남구 등

본 연구에 활용된 데이터는 2022년 수집된 경기도 이륜차의 사고 특성이다. 이륜차가 피해차종 혹은 가해차종에 포함된 데이터만을 활용했으며, 2022년 1월1일부터 12월 31일까지의 사고 현황을 활용하였다. 연구에 활용된 총 데이터의 개수는 2,036개이다. 이와 같이 수집한 데이터는 범주형 데이터와 수치형 데이터가 섞여 있는 데이터를 수집, 활용하였다. 2022년 경기도의 이륜차 사고를 각 요소별로 구별하였고, 비교적 뚜렷하게 분류된 요소들은 법규위반, 도로형태, 시간, 사고유형, 구(장소) 값으로 구별할 수 있었다.

3.3. 연구 설계 및 분석 절차

본 연구의 분석 절차는 다음과 같은 순서대로 진행되었다. TAAS에서 제공하는 데이터를 기반으로 데이터 수집, 데이터 전처리, 최적 군집 수 확인, 군집 수 설정 및 군집분석, 군집 별 지역 특성의 순서대로 진행하였다 (<그림 1> 참고). 첫번째 단계인 데이터 수집 단계에서는, TAAS에서 제공하는 데이터를 수집하였다. 본 연구는 이륜차 사고 특성 분석에 관한 연구이기 때문에 가해차량 혹은 피해 차량에 이륜차가 포함되어있는 데이터만을 수집하였다.

두번째 단계인 데이터 전처리 과정에서는, 명확하게 분류할 수 없는 부분들이 포함 되어있는 값을 제거하였

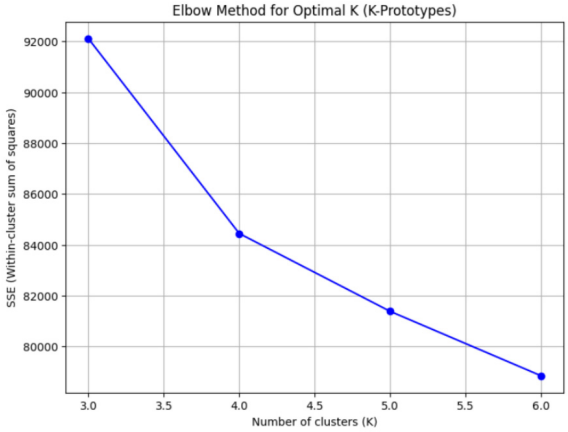


〈그림 1〉 분석 절차

다. 데이터 값이 '미분류'이거나, 결측값이 있는 데이터 행을 삭제할 수 있도록 하였다. 그 후에, '피해 운전자 차종' 및 '가해 운전자 차종' 열의 데이터 값이 '이륜차' 혹은 '원동기' 값의 데이터만 남기고, 나머지 사고 데이터는 삭제하는 과정을 실행했다.

선택한 독립 변수는 '가해 운전자 차종', '피해 운전자 차종', '사고 유형', '요일', '도로형태', '법규위반', '시군구', '사고일시', '노면상태', '기상상태' 변수를 활용했다. 선택한 변수 중'시군구' 값은 한 데이터에 시 혹은 군, 구 혹은 읍이 합쳐져 있어 각각의 시, 군, 구의 값을 활용하기 위해 공백을 기준으로 분류하여 '시', '구' 변수값으로 새로운 열로 추가하여 데이터에 활용하였다.

세번째 단계에서는, 최적 군집 수 확인을 위해 엘보우 메소드(Elbow method)를 활용하였다. 엘보우 메소드는 클러스터의 개수 k를 입력 받는 클러스터링 알고리즘에서 적절한 클러스터의 수를 결정하는 방법 중 하나이다. 클러스터 비용은 각 클러스터에서 중심점과 클러스터 안의 다른 데이터 간 거리합의 평균이다. 이때, k값에 따라 비용이 급격하게 감소하는 구간과 비용이 완만하게 감소하는 구간이 존재할 수 있다. 비용이 급격하게 감소하는 구간과 완만하게 감소하는 구간 사이의 k값이 엘보우이다(김동혁 외, 2015). <그림 2>와 같이, 엘보우 메소드를



〈그림 2〉 엘보우 메소드를 활용한 결과 그래프

실행해 본 결과 최적의 분석 군집 수를 4개로 설정하는 것이 가장 적합하다는 결과를 얻을 수 있었다.

네번째 단계에서는, 실질적으로 군집 분석을 실행하였다. 엘보우 메소드를 통해 얻은 k값을 4로 설정하고, scikit-learn 패키지의 K-prototypes를 활용하였다. 군집별로 어떤 변수에서 뚜렷하게 차이가 나는지 확인해보고, 이후 마지막 단계에서 활용할 수 있도록 하였다. 마지막 단계에서는, 네번째 단계에서 얻은 결과값을 활용하여 군집에서 나타나는 지역별 특성 정리 및 새로운 단속 구역 확장을 제안할 수 있도록 하였다. 각각의 군집들은 법규위반, 도로형태, 시간 값 개수, 지역별로 뚜렷한 특성을 보이고 있었다. 이에 따라, 본 연구에서는 뚜렷한 독립변수들을 중심으로 군집 별 특성을 정리하여 이륜차 사고 방안을 제안하고자 한다.

4. 분석결과

본 연구에서는 K-prototype을 통해 네 개의 군집으로 분류되었다(<표 3> 참조). 클러스터링을 통해 이륜차 사고의 군집 별 특성을 파악할 수 있다. 특히, 이륜차 사고의 주요 사고 발생 도로, 주요 위반 법규와 사고 유형, 그리고 사고 발생 시간적 특성에 따라 각 특징을 도출할 수 있었다. 분석 결과, 네 개의 군집 모두 공통적으로 안전운전불이행이 주요 위반 법규에 포함되었다. 차별적으로는 군집 1은 교차로에서 신호위반을 통한 사고가 많이 발생하는 것으로 차대차(이륜-사륜차) 유형이 가장 많았다. 군집 2는 교차로에서 신호등이 없거나 교차로 통행 방법을 무시하여 사고가 많이 발생하는 것으로 차대차(이륜-사륜차) 유형이 가장 많았다. 군집 3은 학교나 주거지 인근 보도 근처 단일로에서 사고가 많이 발생하는 것으로 차대사람(이륜-사람) 유형이 가장 많았다. 군집 4는 아파트, 마트 근처 단일로에서 사고가 많이 발생하는 것

〈표 3〉 분석 결과 특성 정리

구분	군집1 (신호위반)	군집2 (교차로 운행방법 위반)	군집3 (차대사람)	군집4 (아파트 마트 주변)
주요 특성	교차로에서 신호위반을 통한 사고가 많이 발생	교차로에서 신호등이 없거나 교차로 통행 방법을 모르거나 무시하여 사고가 많이 발생	학교나 주거지 인근 보도 근처 단일로에서 사고가 많이 발생	아파트, 마트 근처 단일로에서 사고가 많이 발생
주요 사고발생 도로 및 특성	교차로	교차로	단일로	단일로
주요 위반 법규	신호위반, 안전운전불이행	교차로운행방법위반, 안전운전불이행	보행자보호의무위반, 안전운전불이행	안전거리미확보, 안전운전불이행
주요 사고 유형	차대차(이륜-사륜차)	차대차(이륜-사륜차)	차대사람(이륜-사람)	차대차(이륜-사륜차)
사고 발생 시간적 특성	비교적 차량이 적은 시간대	교통량 혼잡 시간대	보행자 이동량이 많은 시간대	교통량 혼잡 시간대

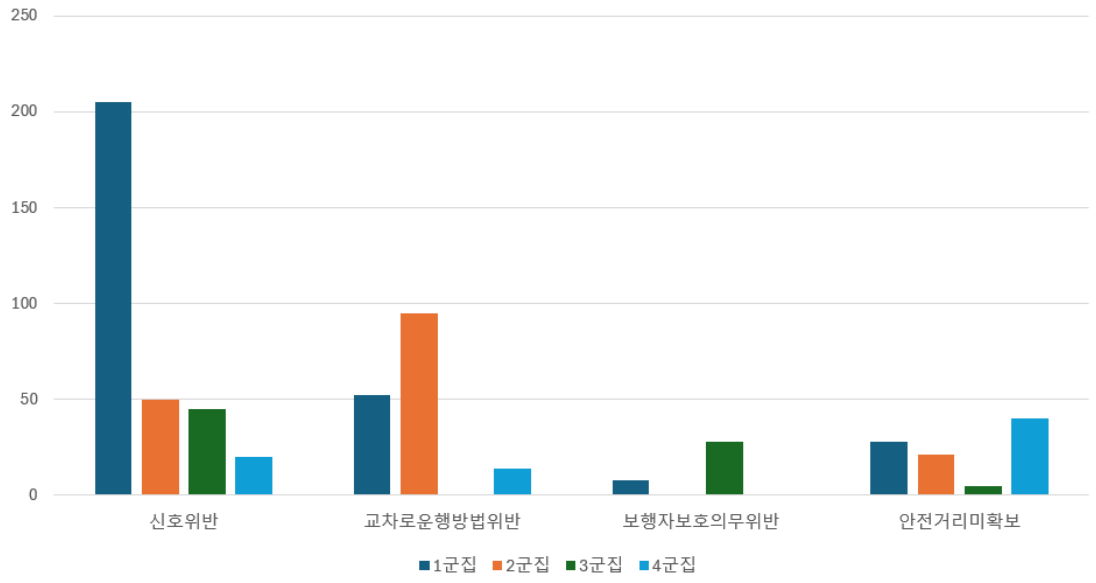
으로 차대차(이륜-사륜차) 유형이 가장 많았다. 특히, <그림 3>은 각 군집의 주요 위반 법규 빈도 수를 나타낸 것으로 신호위반은 군집 1, 교차로 운행 방법 위반은 군집 2, 보행자 보호 의무 위반은 군집 3, 안전거리미확보는 군집 4가 가장 높은 빈도를 차지하였다.

군집별 특성을 살펴보면 군집 1에서는 비교적 차량 유동량이 많지 않은 시간대에 교차로에서 신호위반을 통한 사고 발생이 많았다. 군집 2에서는 교통량 혼잡 시간대(출, 퇴근 등) 신호등이 없거나, 회전교차로 등과 같이 통행 방법을 명확하게 알고 있지 못하거나, 무시하는 경우 사고 발생이 많았다. 또한, 우회전하는 차량과 이륜차 충돌의 발생이 잦다는 결과가 나타났다. 군집 3에서는 차대사람, 즉 보행자 사고가 많이 나타나는 것으로 나타났는데 학교와 주거지 인근 보도 근처 단일로에서 많이 발생하였다. (보행자 겸용도로 등). 사고유형과 마찬가지로, 유동인구가 많은 시간대에 많이 발생하였다. 군집4에서는 아파트, 마트 근처 단일로에서 많이 발생하는 결과값이 나타났다. 이어 <그림 3>과 <표 3>을 토대로 각 군집별 교통 법규 위반과 특성을 연관지어 분석해 보았다.

군집 1의 경우에서는 법률 위반 중 신호위반이 전체의 약65%를 차지하고 있고, 비교적 차량 유동량이 적은 시간대에 사고가 발생하고 있다. 비교적 차량이 적은 시간대이다 보니 신호위반을 하는 차량이 많아지고, 이에 따른 사고가 발생하고 있는 모습으로 보여 진다. 예를 들어, 본 연구에서 사용한 데이터의 특성에서 안산시 버스터미널 사거리 같은 경우, 차량이 적은 시간대임에도 불구하고 버스가 주기적으로 이동하며 비교적 차량이 많은 특성을 가진 장소이다. 현재 이륜차 교통사고 다발지역으로 분류된 장소이기도 하다(도로교통공단,2023). 클러스터링 결과를 통해 이러한 특성이 도로 안전에 부정적인 영향을 미칠 수 있음을 확인할 수 있었다.

군집 2의 경우는 법률 위반 중 교차로 운행 방법 위반이 전체의 60%로 나타나고 있다. 교차로 운행 방법이란, 도로교통법 제19357호 제25조상의 6가지 항으로 정의되어 있다(경찰청, 2023). 대부분의 경우 신호등이 없거나, 회전교차로 등과 같이 통행방법이 어려운 장소이다. 뿐만 아니라, 교통량이 혼잡한 시간대의 고속도로 입구 및 출구 등 끼어들기 사고가 많이 일어나고 있다. 신호등이 없는 교차로는 평소에는 차량 이동량이 많지 않아 점멸등을 사용하는 도로이지만 출, 퇴근시간대에 차량의 이동량이 많아지면 사고발생이나 교통체증을 일으키고 있는 도로이다.

군집 3의 경우는 법률 위반 중 보행자보호의무 위반이 전체의 89%를 차지하고 있다. 보행자보호의무 위반이라



〈그림 3〉 사고 유형별 군집의 데이터 수 분석

함은, 도로교통법 제14839호 제27조상의 5가지 항으로 정의 되어있다(경찰청, 2017). 학교와 주거지 인근 보도 근처 단일로에서 자주 발생하는 모습을 보이는데, 횡단보도나 보행자 겸용도로 등에서 많이 나타나고 있다. 전체적으로 유동인구가 많은 시간대에 발생하는데, 특히 출, 퇴근 시간과 같이 보행자가 많은 시간대에 발생한다. 차대 사람의 형태로 일어나는 사고이기에 인명 피해에 가장 큰 영향을 끼치는 사고 유형이 될 수 있다.

군집 4의 경우는 법률 위반 중 안전거리 미확보가 전체의 41%를 차지하고 있다. 한국은 현재 자동차 등록 대수가 2500만대를 돌파했다(통합 데이터 지도,2022). 운행 중에 차간 거리를 좁은 상태로 운행하게 되면 추돌 사고의 위험이 높아진다. 앞차의 급정거로 인한 사고를 대비하기 위해서는 차간 거리를 반드시 지켜야 한다(Autopostkorea, 2022). 하지만 현재 한국에서의 자동차 등록 대수의 증가 추이와 수도권에 몰려 있는 인구의 수 때문에 안전거리를 확보하게 되면 앞차가 끼어들고, 다시금 거리를 확보하면 다시 앞차가 끼어드는 악순환을 겪고 있다. 아파트 혹은 마트 입구에서 끼어들기 위해 좁은 도로위에서 얌체 운전자들이 있고, 이러한 얌체 운전자들을 대비하기 위해 운전자들을 스스로 안전거리 확보를 포기하고, 안전거리 미확보로 인한 사고로 이어진다.

도로 교통 안전은 현대 사회에서 가장 중요시 되는 이슈 중 하나이다. 본 연구에서 도출된 군집 분석 결과는 도로 교통 안전에 대한 효율적인 관점을 제시한다. 본 연구의 데이터 분석을 통해 법률적 위반과 사고 발생 장소의 특성을 명확하게 파악할 수 있다. 각 군집별로 눈에 띄는 특성들을 가지고 있으며, 이를 토대로 교통 안전 대책 및 캠페인 활용 등 이륜차 사고 예방 방안을 제안할 수 있다.

5. 결과 토의 및 시사점

5.1. 연구결과 토의 및 활용 방안

본 연구에서는 K-prototype 분석을 활용하여 이륜차 사고 발생의 주요 특성을 나타내는 4개의 군집을 도출하

〈표 4〉 이륜차 사고 예방을 위한 방안 제안

사고 그룹의 분류	해결 방안
(군집1) 교차로 – 교통량이 적은 시간대	신호위반 단속 강화 – 무인 단속 카메라, 이동식 단속 등의 강화 – 신호위반 벌금 및 벌점 강화
(군집2) 교차로 – 교통량 혼잡 시간대	교통 표지판 및 도로 상 안내 활용 – 회전 교차로 진입 전 표지판 안내, 도로 상 연결선 안내 등 교차로 통행 방법에 대한 교육 업데이트 – 면허 획득 시나 후에도 교차로 통행 등 복잡한 도로교통 통행 방법에 대한 추가 교육 제공
(군집3) 단일로 – 보행자 이동이 많은 시간대	보행자를 우선으로 하는 운행 제도 – 보행자 우선도로로 지정 및 속도 제한 시행 – 안전시설 설치로 보행자와 차 간 물리적 경계 설정
(군집4) 단일로 – 교통량 혼잡 시간대	도로 환경 개선을 위한 안전거리 유지 – 차대차 안전거리 유지 규율 및 법률 강화 – 갓길 주정차 금지 시행 및 단속 강화

였다(<표 3> 참조). 각 군집은 사고유형, 주요 발생 지역, 위반 법규, 사고 시간적 등 특성을 나타내었다. 이를 토대로 이륜차 사고 특성별로 사고 예방을 위한 해결방안을 제안한다(<표 4>참고).

본 연구결과에서 군집 1은 교차로에서 교통량이 비교적 적은 시간대로 이때 신호위반을 하는 차량이 증가하면서 사고가 많이 발생하는 특성이 있다. 군집 2는 교차로에서 교통량이 많은 시간대로 교차로(회전 교차로) 통행 방법을 명확히 알지 못하거나, 무시해서 사고가 많이 발생하는 특성이 있었다. 특히, 우회전하는 차량과 이륜차 충돌의 발생이 잦은 결과를 보아 이륜차 운전자들이 안전운전을 하지 않고 우회전할때 무리하게 추월하려다 충돌이 발생하는것으로 볼 수 있다. 군집 3은 학교나 주거지 인근 보도의 단일로에서 보행자 이동이 많은 시간대에 사고가 많이 발생한 것이 확인하였다. 군집 4는 마트 근처의 단일로에서 교통향이 혼잡한 시간대에 안전거리미확보로 사고가 많이 발생하였다. 군집 1, 군집 2, 군집 4는 주로 이륜차대 사륜차 사이에 일어나는 사고 유형이고, 군집 3은 이륜차대 사람의 형태로 사고가 일어나는 사고 유형이 많은 것으로 확인되었다. 본 연구에서는 4개의 군집별로 이륜차 사고의 특성을 분석하였고, 이를 토대로 각 군집 별로 맞춤형 교통 안전 대책을 제안하고자 한다.

교차로에서 비교적 차량이 한적한 시간대에 신호 위반으로 인한 사고가 많이 발생하는 지역인 경우 신호위반 행동을 유발하는 특정 상황(한적한 교차로, 교통경찰관의 부재)에 해당(조성호, 2013)하므로 무인단속 카메라를 추가하거나, 경찰의 이동식 단속을 강화하여 운전자들에게 신호위반행동을 조절하도록하는 환경단서를 활용하는 방안을 고려할 수 있다. 또한, 신호위반 시 벌금이나 벌점을 강화할 수 있다.

교차로에서 교통량이 혼잡한 시간대에는 통행방법이 익숙하지 않은 이륜차 운전자들이 많고 무신호 교차로가 많기 때문에 교차로(회전교차로)에 진입하기 전 구체적인 표지판 안내를 통해 통행 방법을 숙지시켜줄 수 있다. 또한, 사고예방 메시지를 추가하여 따르지 않을 경우 사고가 촉진된다는 것을 강조하는 메세지 프레이밍 효과를 활용하여 안전한 도로 환경을 조성하는 방안을 고려할 수 있다(유병호, 황조혜, 2018). 그 밖에, 도로 위에 연결선으로 표시하여 헷갈리지 않고 주행할 수 있는 방안을 고려할 수도 있다. 추가적으로 이륜차나 사륜차 면허 획득 시나 획득 후에도 교차로 통행 방법 등과 같은 복잡하

거나 변경되는 도로교통 통행 방법에 대한 교육을 추가할 수 있다.

학교나 주거지 인근 단일로에서 보행자 이동이 많은 시간대에는 보행자대 이륜차 사고 비율이 많아 치사율이 높으므로 보도 근처 단일로 경우 보행자 운선도로로 지정하여 일방통행 및 속도제한을 시행 및 강화할 수 있다. 이와 함께 보도 근처에 펜스 등의 안전시설 설치하여 보행자와 차 간의 물리적 경계를 설정하여 안전한 보행 환경을 조성하는 등의 방안을 고려할 수 있다.

단일로에서 교통 혼잡 시간대에는 안전거리 미확보로 인한 사고가 많이 발생하였으므로 도로 환경을 개선하기 위해 차량 간의 안전거리를 유지할 수 있는 규율 및 법률을 강화할 수 있다. 특히, 교통 혼잡 시간대에 교통단속을 강화하여 안전거리 미확보나 교통 규칙을 위반하는 운전자들을 단속하는 방안을 고려할 수 있다. 경찰대 치안정책연구소에서 시행한 보행자대 차 사고 발생원인 조사결과, 갓길 주정차의 원인이 가장 높기 때문에 자동차 정차나 주차 금지 등의 조치를 시행할 수도 있다(이세원 등, 2021). 마지막으로 본 연구의 모든 군집에서 공통적인 이륜차 사고 원인으로 꼽히는 이륜차 운전자의 안전운전 불이행을 줄이기 위해 안전운전 불이행 단속에 따른 횟수와 방법에 따라 추가적으로 벌금을 징수하고 안전운전 관련 교육을 이수하도록하는 방법을 고안할 수 있다.

5.2. 연구의 한계 및 향후 연구방향

본 연구는 다음과 같은 한계점과 향후 연구 방향을 가진다. 첫째, 본 연구에서는 2022년 1월부터 12월까지의 경기도 지역의 이륜차 사고 데이터만을 사용하였다. 즉, 수도권 중 경기도 지역 데이터만을 사용한 것뿐만 아니라 한정된 기간의 데이터를 사용하여 연구를 진행한 것에 한계점이 있다. 향후 연구에서는 경기도뿐만 아니라 서울, 부산 등 이륜차 사고가 많이 발생하는 주요 광역시들의 데이터를 통합하여 분석함으로써 그 특성을 분석할 필요가 있다. 더불어 더욱 광범위한 기간의 데이터를 수집하여 분석 결과의 일관성과 신뢰성을 높일 수 있다.

둘째, 사용한 데이터의 속성을 더욱 다양화할 수 있다. 본 연구에서는 이륜차 사고 특성을 분류하기 위하여 주요 사고 발생 도로, 주요 위반 법규와 사고 유형, 사고 발생 시간적 특성 등의 데이터를 수집하여 그 특성을 분류하였다. 현재 데이터뿐만 아니라 좀 더 다양한 데이터들을 수집하여 분석한다면 더 구체적인 정보를 기반으로 한 분류가 가능할 것이다. 예를 들어, 이륜차 이용자의 연령, 성별 등의 인구통계학적 정보와 이륜차 이용의 주요 목적, 이용 시간 등의 이륜차 사용에 대한 정보를 추가한다면 좀 더 의미 있는 결과를 도출할 수 있을 것이다. 이는 연령 및 성별에 따라 주로 사고가 발생하는 도로 형태나 사고 유형을 파악할 수 있을 뿐만 아니라 이륜차 사용 목적이나 사용 시간에 따라 사고의 특성이 다르게 도출될 수 있다. 특히, 본 연구결과에서 모든 군집의 사고 특성과 관련하여 안전운전불이행의 원인이 나타났다. 하지만, 안전운전불이행에도 다양한 세부 요인들이 존재하므로 향후 연구에서는 안전운전불이행의 세부 데이터를 추가로 수집하여 분석에 활용한다면 보다 의미 있는 결과를 도출할 수 있을 것이다.

셋째, 본 연구의 분석결과에 대한 검증을 위한 추가적인 분석 및 확인이 필요하다. 본 연구에서는 이륜차 사고 특성을 분류하고 분류된 군집들을 기반으로 이륜차 사고 예방을 위한 방안을 제시하였다. 하지만, 각 군집의 특성을 제대로 해석하고, 해석한 내용과 관련된 현황과 문제에 대한 전문적인 파악이 필요하다. 즉, 도시교통 법규 및 단속, 그리고 이륜차 사고와 관련하여 제대로 이해하고 있는 전문가 대상 인터뷰를 통해 분석한 결과에 대한 해석 및 확인이 추가된다면 연구결과의 신뢰가 높아질 것이다. 뿐만 아니라, 분석 결과를 해석하는데 더 풍부한

정보를 제공할 수 있다. 향후 연구에서는 분류한 결과를 해석 및 강화하기 위한 전문가 인터뷰 또는 검증을 위한 추가 분석을 고려할 수 있다.

5.3. 학술적 시사점

본 연구는 수도권 지역의 이륜차 사고 데이터를 K-prototype을 이용하여 그 특징을 분류하고, 이를 기반으로 사고 예방을 위한 구체적인 방안을 제시한 것으로 학술적 시사점은 다음과 같다.

첫째, K-prototype을 이용한 이륜차 사고 데이터의 특징을 분류한 연구로 클러스터링 기술을 도시교통 분야에 적용함으로써 정보시스템 뿐만 아니라 도시교통 분야의 문헌 확장에 기여한다. K-prototype 알고리즘은 범주형 데이터와 연속형 데이터를 혼합한 분석이 가능한 것으로 다양한 분야에서 활용되고 있다. 본 연구에서는 도시교통 분야의 이륜차 사고에 대한 위반 법규 유형, 도로 형태, 사고 유형, 위치 등의 범주형과 수치형 데이터를 이용하여 사고의 특징을 분류하고, 이를 기반으로 사고 예방을 위한 방안을 제시함으로써 문헌 확장에 기여한다. 이는 이륜차 사고에 대한 특성과 패턴의 분석뿐만 아니라, 교통 안전 및 예방에 관한 타 연구 분야에도 영향을 줄 수 있다. 더불어, 도시 교통 및 안전 분야뿐만 아니라 클러스터링 및 머신러닝 기술을 활용한 연구 영역에서도 이 연구의 결과는 새로운 연구 방향을 제시할 수 있을 것으로 기대된다.

둘째, 본 연구에서는 이륜차 사고와 관련한 공공데이터를 이용하여 분석을 진행하였다. 선행 연구에서도 교통사고 관련 공공데이터를 이용한 분석이 이루어졌지만 대부분이 사륜차를 중심으로 수행되었다. 이륜차 관련 데이터를 활용한 연구들은 이륜차 사고 발생 원인과 사고 특성 자체를 분석한 연구는 있었지만, 이륜차 사고의 원인 및 특성 데이터를 기반으로 이륜차 사고의 특성을 분류한 연구는 드물었다. 본 연구에서는 이륜차 사고 발생이 많은 수도권 지역의 데이터에서 도로형태, 위반 법규 유형, 사고 유형, 사고 시간, 위치 등을 분석에 이용함으로써 사고의 특징이 명확히 구분되었다. 즉, 실제 이륜차 사고 데이터를 기반으로 도로형태를 포함한 주요 요인들의 데이터를 분석함으로써 이륜차 사고의 유형별 특징을 파악하였다.

더불어, 본 연구의 결과는 지식경영 분야에도 중요한 영향을 미친다. 데이터 분석과 지식의 공유를 통해 조직적인 학습에 기여할 수 있다. 이륜차 사고 데이터 분석을 통해 얻은 지식을 조직 내 다른 프로젝트 혹은 국책 수립 과정에 적극 활용하여 조직 내 지속적인 학습과 발전을 촉진시켜 교통 안전 측면의 연구의 다양성을 제공한다.

5.4. 실무적 시사점

본 연구는 위와 같은 학술점 시사점뿐만 아니라, 다양한 분야에서의 실무적인 시사점을 포함하고 있다. 구체적이고 지역특성에 맞춰진 방안들은 다음과 같은 실무적 시사점을 나타낸다.

첫째, 본 연구는 최근 배달문화가 전 세계적으로 확산되고 도시교통 문제가 사회적 문제로 이슈가 되고 있는 상황에서 새로운 사회문제로 언급되고 있는 이륜차 사고 문제를 다루고 있다. 이륜차 운전자들의 위협 운전으로 인한 사고 발생이 증가하고 있으며, 이러한 이륜차 사고는 그 위험성이 높아 사륜차에 비해 사망으로 이르게 하는 비율이 높다. 본 연구에서는 이러한 이륜차 사고에 대한 심각성을 인지하고 이를 예방하기 위한 방안을 제안함으로써 사회문제 해결 분야에 기여한다.

둘째, 본 연구는 이륜차 사고 특성을 분류하고 이를 기반으로 이륜차 사고 예방을 위한 구체적인 방안을 제시하였다 (<표 4> 참조). 본 연구에서 분류한 이륜차 사고 특성을 기반으로 사고 예방을 위한 사회적, 실무적, 정책

및 법률적 측면의 구체적인 방안을 제시하였다. 다시 말해서, 본 연구에서는 이륜차 사고 특성을 분류한 것뿐만 아니라 이를 기반으로 이륜차 단속 및 규정 개정을 위한 방안을 제시한 것에 의의가 있다. 특히, 사고 다발지역을 중심으로 교통 인프라의 개선 및 안전 시설을 설치하여 사고 발생 가능성을 줄이고 교통 안전을 향상시킬 수 있다. 이는 연구 결과를 현장에서 실질적으로 활용하여 교통안전성을 향상시키고 사고 발생을 줄이는 데에 직접적인 도움을 줄 수 있다.

마지막으로, 본 연구는 이륜차 사고 예방을 위한 구체적인 방안을 제시함으로써, 지식을 실무적인 문제 해결에 적용하는 방법을 보여준다. 이는 지식경영 분야에서 강조하는 지식의 적용 및 실행 측면과 연결되어, 조직이 보유한 지식을 현실의 문제 해결에 활용할 수 있는지에 대한 사례를 제공한다. 따라서, 본 연구는 지식경영의 실무적인 실행 가능성과 효과를 보여주는 중요한 사례로 활용될 수 있다.

〈참고문헌〉

[국내 문헌]

1. 권성아, 백승환, 이지연 (2015). 국내 공공데이터 개방현황 및 활용사례 분석. **한국정보관리학회 학술대회논문집 22**, 37-45
2. 김주영, 김준용, 배현식, 김진태 (2015). 국내 자전거 및 이륜차 교통사고 특성분석 연구. **2015년 한국ITS학회 춘계학술대회**, 70-70.
3. 김정화, 진서훈 (2018). K-prototypes 군집분석에서 가중치 선택에 관한 연구. **대한설비관리학회지, 24**(2), 51-58.
4. 김진솔, 신동훈, 김희웅. (2021). 비정형 빅데이터를 이용한 COVID-19 주요 이슈 분석. **지식경영연구, 22**(2), 145-165.
5. 소형준, 김용만, 김남선, 황재성, 이철기 (2020). 무인교통단속장비 설치 판단 기준 및 설치대수 산정 연구. **한국ITS학회논문지, 19**(6), 49-60.
6. 송석현 (2014). **공공데이터 민간개방의 경제적 파급효과 분석연구**. 한국정보화진흥원.
7. 이지선, 이희원 (2021). **이륜자동차 사고원인 분석 및 교통안전 개선방안**. 한국교통연구원.
8. 이세원, 김원중, 심태일 (2021). 이면도로 보행안전 법 · 제도 개선 연구. **치안정책연구, 35**(2), 1-35.
9. 인병철, 유성준, 한음, 이경진, 박성호 (2023). 이륜차 무인교통단속장비 개발 및 표준규격 연구. **한국ITS학회논문지, 22**(1), 126-142.
10. 유진화, 최인범, 송서하 (2020). 이륜차 교통사고 분석을 통한 사고예방 연구. **대한교통학회지, 17**(1), 29-36.
11. 유병호, 황조혜 (2018). 메시지 프레이밍이 소셜미디어 신뢰도 및 행동의도에 미치는 영향. **지식경영연구, 19**(2), 69-90.
12. 조용빈, 김진태, 임준범, 오상태 (2022). 배달 이륜차 라이더 교통 법규 위반 단속 연구. **한국ITS학회논문지, 21**(1), 182-192.
13. 조성호 (2013). 운전자들의 신호위반 심리요인 분석. **한국범죄심리학회지, 9**(3), 139-165.
14. 최정우, 금기정 (2014). 순서형 프로빗모형에 의한 이륜차 사고심각도의 영향요인 분석. **한국도로학회논문집, 16**(5), 143-154.
15. 채범석 (2005). 이륜차 사고 특성 및 대책에 관한 연구. **대한교통학회 학술대회지, 48**, 685-694.

[국외 문헌]

16. Munshi, A. (2022). A survey of advances in hierarchical clustering algorithms and applications. ***IJCSNS International Journal of Computer Science and Network Security, 22***(5), 17-24.
17. Ji, J., Pang, W., Zhou, C., Han, X., & Wang, Z. (2012). A fuzzy k-prototype clustering algorithm for mixed numeric and categorical data. ***Knowledge-Based Systems, 30***, 129-135.
18. Rifaat, S. M., Tay, R., & De Barros, A. (2012). Severity of motorcycle crashes in Calgary. ***Accident Analysis & Prevention, 49***, 44-49.
19. Huang, Z. (1988). Extensions to the k-means algorithm for clustering large data sets with categorical values. ***Data Mining and Knowledge Discovery, 2***(3), 283-304.

[URL]

20. A Tutorial on Clustering Algoritms. http://www.aistudy.com/pattern/clustering.htm
21. Korean Statistic Information Service (2021). http://www.kosis.kr
22. 국토교통부 (2023). https://www.molit.go.kr/portal.do
23. 도로교통법 (2024). https://www.law.go.kr/lsInfoP.do?lsiSeq=250023&efYd=20240101#0000
24. 엠투데이 (2023). https://www.autodaily.co.kr/news/articleView.html?idxno=506394

저 자 소 개

허 원 진 (Won-jin Heo)

현재 경희대학교 빅데이터 응용학과 경영학 석사과정에 재학중이다. 주요 관심 분야는 딥러닝, 데이터 애널리틱스, NLP 등이다.

강 진 호 (Jinho Kang)

현재 경기대학교 산업경영공학과 학사과정에 재학 중이다. 주요 관심 분야는 데이터 애널리틱스, 딥러닝, 이상 탐지 등이다.

이 소 현 (So-Hyun Lee)

Xi'an Jiaotong University 경영대학에서 근무한 후, 현재 경기대학교 산업경영정보공학과 조교수로 재직 중이다. 주요 관심 분야는 디지털 비즈니스, 데이터 애널리틱스, 생성형 AI 등이다. 관련 연구들은 Information Systems Research, Communications of ACM, Information and Management, International Journal of Information Management, Internet Research 등에 논문이 게재되었다.

〈 Abstract 〉

Classification of Characteristics in Two-Wheeler Accidents Using Clustering Techniques

Heo, Won-Jin*, Kang, Jin-ho, Lee**, So-hyun***

The demand for two-wheelers has increased in recent years, driven by the growing delivery culture, which has also led to a rise in the number of two-wheelers. Although two-wheelers are economically efficient in congested traffic conditions, reckless driving and ambiguous traffic laws for two-wheelers have turned two-wheeler accidents into a significant social issue. Given the high fatality rate associated with two-wheelers, the severity and risk of two-wheeler accidents are considerable. It is, therefore, crucial to thoroughly understand the characteristics of two-wheeler accidents by analyzing their attributes. In this study, the characteristics of two-wheeled vehicle accidents were categorized using the K-prototypes algorithm, based on data from two-wheeled vehicle accidents. As a result, the accidents were divided into four clusters according to their characteristics. Each cluster showed distinct traits in terms of the roads where accidents occurred, the major laws violated, the types of accidents, and the times of accident occurrences. By tailoring enforcement methods and regulations to the specific characteristics of each type of accident, we can reduce the incidence of accidents involving two-wheelers in metropolitan areas, thereby enhancing road safety. Furthermore, by applying machine learning techniques to urban transportation and safety, this study adds to the body of related literature.

Key words: Two-wheeled vehicle Accidents, K-prototypes, Machine learning, Clustering

* 9341kkk@khu.ac.kr
** jinho5080@gmail.com
*** dlthgus1010@gmail.com