

치안현장 의사결정 지원을 위한 상황 대응 분류 모델

권은정, 변성원, 박현호, 장동만, 정의석
한국전자통신연구원

ejkwon@etri.re.kr, swbyon@etri.re.kr, hyunhopark@etri.re.kr, dmjang@etri.re.kr,
esjung@etri.re.kr

A Situational Response Classification Model to Support Decision-Making in Police Field

Eun-Jung Kwon, Sungwon Byon, Hyunho Park, Dongman Jang, Euisuk Jung
Electronic and Telecommunications Research Institute (ETRI)

요 약

본 논문은 수시로 변화하는 치안현장에서 발생하는 데이터를 분석하여 현장경찰관들이 접하는 각종 사건 유형에 따라 대응정보를 제공하는 분류모델을 제안한다. 이를 위하여 112 신고를 받고 출동하는 현장경찰관에게 제공되는 출동지령 데이터, 사건현장에서 현장경찰이 발화하는 질의 데이터의 특징을 분석하여 치안업무를 지원하기 위한 상황대응 모델을 제시한다. 현장 상황 유형 기준에 따라 특징이 되는 주요 변수 및 데이터의 주요 키워드를 가중치 단어로 선정하여 모델 학습 수행 및 서비스 가능성을 제공한다.

I. 서 론

범죄현장에 출동하는 현장 경찰관은 지령정보를 통해 사건발생현장의 상황을 인지하고 현장으로 출동하게 된다. 이 과정에서 현장 상황에 따라 경찰은 현장 상황에 필요한 법령 및 경찰 업무준수 지침 등을 숙지할 필요가 있다. 또한, 현장의 위험 상황에 따라 신속하게 대응하기 위한 방안을 강구해야 한다. 이러한 판단은 현장 경찰관 개인의 능력 및 경험에 따라 대응 조치가 달라질 수 있으므로 효과적인 현장 대응이 가능한 지원 시스템 필요하다. 통계에 따르면 2019 년 112 신고에 대한 출동지령은 연 총 12 백 54 만 5 천건으로, 중요범죄 40 만건의 출동지령이 포함된다. 이러한 이유로 본 논문에서는 현장경찰이 각종 사건에 대응 시 업무지원을 효과적으로 제공받을 수 있도록 치안현장 데이터기반 상황 대응 분류 모델을 제안한다.

II. 관련연구

112 신고가 접수되면 출동지령을 통해 사건 발생 위치로 신고 종류와 긴급성에 따라 출동 경찰에게 신고 내용과 출동 단계가 전달된다. 현장 경찰은 전달받은 신고 내용을 이해하고 순찰차로 이동하는 과정에서 현장의 상황을 추측, 도착 후 상황 업무를 수행하게 된다. 현장 상황에 대한 대응 방안은 업무범위 지침내에서 현장경찰의 판단에 의해 이루어진다. 현장경찰에게 경찰 순찰차 내부 탑재 단말기기 또는 경찰 보유 개인 휴대기기를 통해서 현장상황에 대응하기 위한 적합한 안내정보를 적시에 제공받게 된다면 치안현장의 안전은 더욱 강화될 것이다.

범죄는 강력범죄, 절도, 폭력범죄, 지능범죄, 풍속범죄, 교통범죄 등 약 15 개의 죄종으로 분류되고, 각 죄종은 다시 세부 항목으로 정의되어 범죄 발생 시간, 발생장소

기준으로 세부적인 통계분석 결과 확인이 가능하다[1]. 정형적인 데이터 분석 결과 외에 112 신고내용으로부터 텍스트 마이닝 분석을 통해 행위 주체와 대상, 행위 유형, 장소에 관한 키워드 분류를 통해 112 신고 분석시스템 및 범죄발생 예측시스템의 구축에 대한 시사점을 연구한 바 있다[3]. 또한, 최근 언어모델이 단순히 순차적으로 특정 단어 기준으로 이후 단어를 예측하지 않기 때문에 신경망의 모든 레이어에서 문장 내 단어 순서, 단어 간의 관계, 문맥 정보를 확인할 수 있는 딥러닝 기반 언어모델을 활용하여 기존 전통적인 방법의 비정형 데이터 분석의 한계를 극복한 사례들이 제시되고 있다. 이러한 연구성과는 기존 데이터 분석 방법 대비 대용량의 학습데이터가 필요하고 이로 인해 학습 시간도 증가하지만, 벡터의 저장공간에 대한 희소성 문제점, 동음 이의어 구별 등 기존 언어모델의 한계점을 극복할 수 있기때문에, 사전 학습된 언어모델을 통해 도메인 말뭉치를 이용하여 다운스트림 태스크 성능의 개선으로 활발히 이용되고 있다[3].

III. 치안 현장 상황대응 분류 모델링

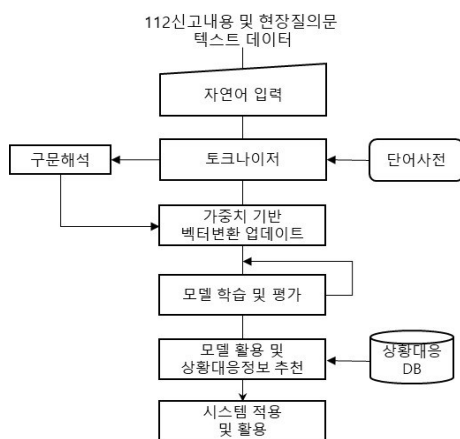
본 장에서는 112 상황실에 접수되어 현장경찰의 이동 기기로 현장상황 대응을 위한 정보를 자동 제공하기 위한 모델 구조와 학습용 데이터로 사용되는 데이터 구조 및 속성에 대해 설명한다. 모델 학습을 위한 데이터 중 개인정보가 포함된 부분은 정규 표현식, 고유 개체명 인식기를 이용하여 데이터 전처리 과정을 수행한다. 또한, 상황 별 문맥상 의미가 통하지 않는 데이터도 학습 데이터에서 제외한다.

그럼 1 은 본 논문에서 제안하는 모델학습을 위한 전체 워크플로우 처리에 대한 개념을 표현한 것이다.

그림 1 은 경찰 출동지령시스템을 통해 신고접수 중별 분류코드 기준에 따라 텍스트 형태로 현장 경찰의 이동 단말로 대응정보를 전달하기 위한 모델 구조이다. 상황 대응에 대한 세부기준은 현장상황의 위험도를 고려해야 하기 때문에 출동지령 데이터를 구성하는 주요 단어 및 문맥의 의미에 가중치를 부여하는 구조를 갖는다. 이에 따라 본 논문에서는 치안 도메인에서 상황 별 단어가 가지는 빈도와 위험 유형에 따라 해당 단어에 대한 가중치를 설정, 출동지령 텍스트 문장 구분을 위한 1 단계 작업을 수행한다. 예를 들어, ‘폭력’으로 분류된 출동지령 텍스트는 ‘때리고’, ‘맞음’, ‘과격’과 같은 단어에 대한 가중치 값이 ‘방화’로 분류된 출동지령 텍스트의 주요 키워드 값보다 높은 값을 갖도록 한다. 가중치 기법을 기준으로는 출동지령 텍스트 문장의 전체 문맥에 대한 이해가 어렵기 때문에 문맥에 대한 의미까지 파악하기 위해 학습 알고리즘이 필요하다. 본 논문에서는 상황 별 가중치가 높은 단어는 치안 도메인 용어사전을 정의하여 선정한다. 출동지령 텍스트 데이터는 토큰나이저를 이용하여 문장을 공백 단위로 분절, 분절된 단어를 다시 서브단어로 구분하여 벡터 값으로 변환된 값을 통해 모델 학습이 이루어진다. 이 과정에서 사전 학습된 언어모델을 통해 다운스트림 태스크 성능이 개선될 수 있도록 모델을 학습하고 그 결과에 대한 성능을 검증한다.

모델 학습에 사용되는 데이터세트의 일부 항목에 대한 예는 표 1 과 같다. 112 신고 접수 시 기록되는 사건중별과 현장경찰이 사건 현장에서 종결처리과정에서 분류하는 값이 다를 수 있다. 다시 말해, 신고를 받고 출동하는 현장경찰은 출동지령 시 제공되는 일부 정보만으로는 현장의 다양한 상황을 대응하기 어렵다고 이해할 수 있다. 실제 실험에 적용된 출동지령 데이터가 딥러닝 알고리즘을 수행하기에 충분하지 않다면 특정 범죄 최종으로 편향(overfitting)될 수 있기 때문에 성능 개선을 위해 데이터 증강 및 랜덤 오버 샘플링 기법을 통한 학습방식이 요구된다.

본 논문에서 데이터를 증강하는 방식은 치안현장 도메인 용어사전을 정의하고 활용한다. 학습용 데이터로 문장을 구성하는 주요 단어 중심으로 자주 등장하는 단어와 관련된 단어를 관계어로 정의, 추가 생성 하여 학습용 데이터로 보강한다. 이방법은 표 1 의 ‘신고내용’ 텍스트 문장에 대해 예측되는 ‘사건중별’ 값을 획득하여 상황대응 DB로부터 관련된 대응정보 추천 받을 수 있도록 시스템에 적용한다.



(그림 1) 현장 상황 대응 분류 모델 적용을 위한 워크플로우 흐름도

(표 1) 상황대응분류를 위한 가정폭력 데이터 샘플 예시

코드	사건중별	중결분류	중결시사건코드	신고내용
C1	폭력	현장중결	가정폭력	위증에서 양정싸우는소리가 들린다/ 신고자201호에 살고있고/
C1	소음	현장중결	가정폭력	가등 옆집들에서 조등 딸을 심하게 때려서 우는소리가 크게 들린다/
C1	가정폭력	잠거	가정폭력	남편이 베트남인 채로 아이들을 위협 다급
C1	가정폭력	현장중결	가정폭력	가족이 싸움다/
C1	가정폭력	현장중결	가정폭력	아버지가 치매이신데, 어머니를 때리고 있다고, 어머니한테 연락 왔더라, 어머니,
C1	가정폭력	현장중결	가정폭력	부부싸움하고 있다/ 불리오라/
C1	가정폭력	잠거	가정폭력	등상이 술마시고 정신없이 날뛰고있음/
C1	가정폭력	허위오인	가정폭력	술취한 아들이 행패를 부린다/그동안 신고를 많이 했는데 조치가 잘 되지 않아서 불안하다/조치가 필요하다고/
C1	가정폭력	현장중결	가정폭력	열집에서 싸우고 있다고가정폭력 없다고 502호에서
C0	가정폭력	현장중결	가정폭력	[CODE] 인지행)생명의 위협을 느낀다/아버지가 용동이 들고 뛰어온다/신고자 핸드폰 놓고와서 연락적 없다고/
C1	가정폭력	잠거	가정폭력	아들이 행패를 하고 있다고
C2	가정폭력	현장중결	가정폭력	아침에 아들과 좋지 않은 일이 있어서 경찰서 갔다/지구대 불렀다 집 근처에 오면 112에 신고하고라고 한다/
C1	가정폭력	비중등 중결	가정폭력	정신이 이상한것 같다는/둘째 연인을 때린다/불리와달라고
C1	가정폭력	현장중결	가정폭력	행이 아버지하고 싸움이 있다/행이 불렀다/누나도 다쳤다/

현장에 출동한 경찰이 실제 출동 당시 전달받은 내용과 현장 상황이 다른 양상으로 전개가 된 경우, 경찰 업무 지원 휴대기기에 설치된 현장지원 어플리케이션을 통해 질의문을 받고, 변경된 현장 상황에 따른 적합한 대응정보를 수신 받을 수 있도록 한다. 모델의 성능 및 정성적인 평가를 위해서 사건 중별 레이블링 된 값이 정확히 매칭되었는지 확인하기 위해서는 고차원 단위의 ‘신고내용’ 텍스트 문장을 구성하는 단어들의 벡터값을 2 차원으로 축소 변환 후 ‘중결시사건코드’와 관계를 t-SNE 를 통해 시각적으로 표현하여 확인한다. 또한 분류모델 평가지표인 F1 Score 에 따라 합성 데이터를 추가 생성하거나 딥러닝 모델의 하이퍼파라미터를 조정하여 모델 성능 개선을 위한 파인튜닝 작업을 수행한다.

IV. 결론

빅데이터 기반 지식 구축 시장은 꾸준히 성장하고 있고 공공분야에서도 새로운 서비스를 통한 경쟁력 제고 필요성이 대두되고 있다. 본 논문은 현장경찰의 업무를 지원하기 위해 치안현장의 각종 상황들에 대응하기 위해 112 출동지령 및 현장 경찰의 주요 질의내용을 통해 적합한 대응정보 제공을 위한 전체 워크플로우를 제안하였다. 성능이 우수한 모델은 향후 테스트 케이스를 설계하여 실제 시스템에 적용 가능성을 확인할 계획이다.

ACKNOWLEDGMENT

이 논문은 2021 년도 정부(경찰청)의 재원으로 지원받아 수행된 연구결과임 [내역사업명: 112 긴급출동 의사결정 지원 시스템/연구개발과제번호: PR08-03-000-21]

참 고 문 헌

- [1] 경찰범죄통계-범죄 발생 상황 관련 특성, 경찰청, https://stat.kosis.kr/nsieu/view/tree.do?task=branchView&id=132_13204*MT_OTITLE&hOrg=132
- [1] 박진수, 박현호, 조준택, " 112 신고내용에 대한 텍스트 마이닝 분석과 시사점, "한국치안행정논집, 제 3 회(2018), pp. 159-186
- [3] J. Devlin, M.W. Chang, K. Lee, K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," arXiv:1810.04805