스팸 문자 데이터의 인공지능 분석을 통한 교육적 대처 방안 연구

임수빈[1], 민무홍[2]
[1]성균관대학교 인공지능융합교육과 석사과정
[2]성균관대학교 컴퓨터교육과 조교수
cndqnsgo22@skku.edu, iceo@skku.edu

The Study on Educational Countermeasures Through AI Analysis of Spam Message Data

Subeen Lim[1], Moohong Min[2]
[1]Dept. of Education, Applied Artificial Intelligence, Sungkyunkwan University
[2]Dept. of Computer Education, Sungkyunkwan University

요 약

최근 스팸 문자 메시지의 급증은 사용자들에게 심각한 불편을 초래할 뿐만 아니라, 불법 도박 및 금융 사기와 같은 중대한 사회적 문제로 이어지고 있다. 본 연구는 이러한 스팸 문자 메시지의 패턴과 텍스트를 심층적으로 분석하여, 스팸 메시지의 경향을 파악하고 청소년 도박 문제에 대한 교육적 대처 방안을 도출하는 것을 목표로 한다. 2022년 8월부터 2024년 3월까지 약 1년 7개월 동안 신고된 스팸 문자 데이터를 바탕으로, 스팸 메시지의 신고량 추이와 키워드 군집 분포를 분석하였다. 분석 결과, 스팸 메시지는 주로 주중 방과 후 시간대인 12시에서 18시 사이에 집중되었으며, 도박 관련 스팸 메시지가 전체의 62.7%로 가장 큰 비중을 차지하였다. 텍스트 분석을 통해 도출된 주요 키워드는 신규 가입 유도 및 금전적 유인과 관련된 것으로, 청소년들이 도박에 유인될 가능성이 높은 요소들이 포함된 것으로 분석되었다. 본 연구는 스팸 문자 데이터의 인공지능 분석을 통해 청소년 도박 문제에 대한 중요한 시사점을 제공한다.

1. 서론

최근 스팸 문자 메시지의 급증은 사용자들의 일상에 심각한 불편을 초래할 뿐만 아니라, 불법 도박, 금융 사기 등 중대한 사회적 문제로 이어지고 있다[1,2]. 특히, 불법 도박, 사이버 도박과 같은 부적절한 내용의 스팸 메시지는 사용자들에게 심리적 스트레스와 경제적 피해를 유발할 가능성이 크다[3].

이러한 스팸 메시지들은 특정 카테고리에 집중되며, 문법 오류, 비정상적인 링크 등 독특한 언어적 특징을 통해 기존 스팸 필터링 시스템을 우회하려는 경향이 있다[4]. 이에 따라, 스팸 필터링의 한계를 극복하기 위한 다양한 기술들이 연구되어왔다. 예를 들어, 딥러닝을 이용한 비정상 문자 조합 탐지 기법에서는 변형된 스팸 키워드를 복원하여 스팸 여부를 판단하는 방법이 제안되었으며[5], 허니팟 시스템에서는 SMS 스팸에서 URL 정보를 자동으로 추출하고 불법 온라인 도박 사이트를 감지하기 위한 방법이 제시된 바 있다[6].

또한, 스마트폰 이용량이 증가함에 따라 청소년 도박 경험률도 꾸준히 높아지고 있다[7]. 주로 온라인 도박으로 행해지며 주된 도박 동기는 유흥과 금전적 요인이다[8]. 청소년 도박은 단순한 놀이 문화로 치부되는 경향이 있으며, 참여자는 도박 자금 마련을 위해 2차 범죄까지 저지르는 경우도 있다[9].

본 연구는 이러한 배경을 기반으로, 스팸 문자 메시지의 패턴과 텍스트를 심층적으로 분석하여 스팸 메시지의 경향을 파악하고자 한다. 특히, 스팸 메시지의 신고량 추이와 메시지 내 키워드의 군집 분포를 다각적으로 분석함으로써, 청소년 도박에 관한 교육적 대처 방안을 마련하는 것을 목표로 한다.

2. 본론

2.1. 분석 방법

본 연구는 2022년 8월 마지막 주부터 2024년 3월 셋째 주까지 약 1년 7개월간 사용자들로부터 신고된 스팸 문자 데이터를 기반으로 분석을 진행하였다. 연구는 크게 두 가지로 나누어졌으며, 첫 번째로는 스팸 문자 신고 현황을 분석하였고, 두 번째로는 스팸 문자 텍스트 분석을 수행하였다.

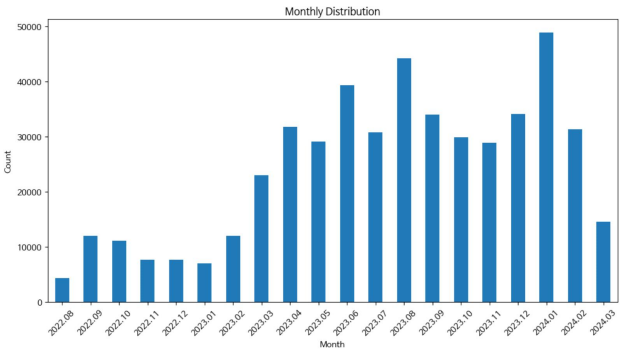
스팸 문자 신고 현황 파악을 위해 월별, 시간대별, 요일별, 카테고리별 분석을 하였다. 데이터를 날짜별로 분류하여 월별 스팸 문자 신고량을 집계하였고,

하루 24시간을 기준으로 시간대별 스팸 문자 신고량을 분석하였다. 또한 일주일을 기준으로 요일별 스팸 문자 신고 패턴을 파악하였고, 특정 유형의 스팸 문자가 차지하는 비율을 분석하였다.

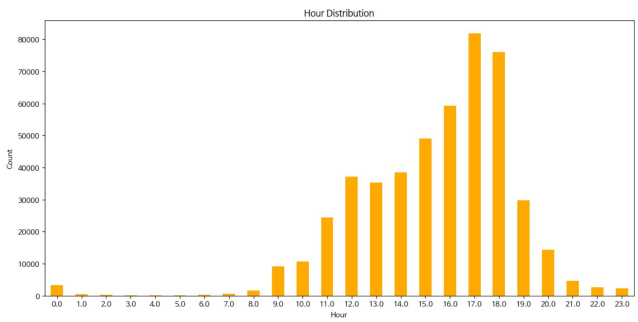
스팸 문자 텍스트 분석을 위해 키워드 분석, 워드 클라우드 시각화 및 클러스터링 분석을 하였다. 키워드 분석을 위해 최빈값 분석과 함께 TF-IDF 분석을 실시하였다. 또한 키워드 분석 결과를 시각화하고 클러스터링 기법을 적용하여 주요 키워드들이 어떻게 군집을 형성하는지 분석하였다. 이때 최적의 군집 개수를 파악할 수 있는 엘보우 방법과 K-Means 기법을 활용하여 총 3개의 클러스터로 구분하였다.

2.2. 분석 결과

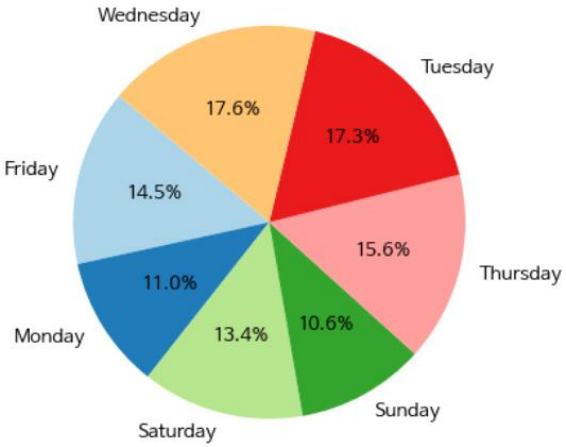
스팸 문자 신고 현황을 분석한 결과 스팸 문자 신고량이 2024년 1월, 2023년 8월, 2023년 6월 순으로 많았고, 2023년 중반 이후 급증하면서 전반적인 증가 추세를 보이고 있다. 또한 스팸 문자는 12시에서 18시 사이에 집중적으로 신고되었고, 수요일, 화요일, 목요일 비중이 50.5%로 상대적으로 한 주의 중반에 많이 신고되었다. 스팸 문자의 카테고리 중 '도박(불법 도박)'이 62.7%를 차지하여 '기타'로 분류된 카테고리를 제외하고는 스팸 문자의 대부분이 도박 관련 스팸 문자임을 확인할 수 있었다.



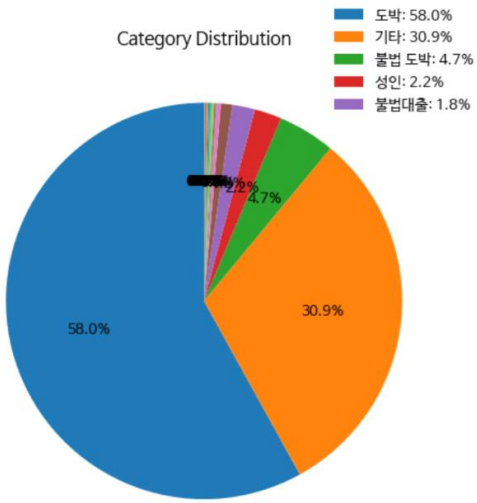
(그림 1) 월별 스팸 문자 신고량



(그림 2) 시간대별 스팸 문자 신고량



(그림 3) 요일별 스팸 문자 비율



(그림 4) 카테고리별 스팸 문자 비율

스팸 문자 텍스트를 분석한 결과 최빈값 분석에서는 '이용', '이벤트', '만', '원', '휴면', '애용' 등 일반적인 단어들이 많이 포함된 반면, TF-IDF 분석에서는 '신규', '코드', '이용', '가입', '이벤트' 등 새로운 가입을 유도하는 문구가 다수 출현하였다. 키워드 분석 결과를 직관적으로 파악하기 위해 워드 클라우드를 활용해 시각화하였다. 중복 출현 단어로는 '이용', '이벤트', '휴면', '애용', '보유', '가입' 등이 파악되었다.

(그림 5) 최빈값 워드 클라우드

(그림 6) TF-IDF 워드 클라우드

클러스터링 결과로 세 가지 주요 군집이 도출되었으며, 첫 번째 군집은 금전 거래 유도와 관련된 내용을, 두 번째 군집은 기술적 문제와 관련된 내용을, 세 번째 군집은 스팸 필터링이나 불필요한 마케팅 메시지와 관련된 내용을 다루고 있었다.

<표 1> 군집별 주요 키워드

	Cluster 1	Cluster 2	Cluster 3
1	입금	휴먼	앞전
2	무한	처리	스팸
3	보증	게이트	신고
4	환전	접속	보스
5	바다이야기	문의	잔금
6	아이디	해주시	단장
7	주소	보유	전액
8	혜택	디지털	제목
9	매일	놀이터	전체
10	스포츠	양해	아이디

구체적으로 보자면 첫 번째 군집의 주요 키워드는 '입금', '무한', '보증', '환전', '바다이야기', '아이디', '주소', '혜택', '매일', '스포츠'로, 금전적 거래와 도박 관련 스팸 메시지가 주를 이룬다. '입금', '환전'은 금융 거래를 유도하며, '보증', '무한'은 사용자가 계속 자금을 투입하게 만드는 도박 메시지에서 자주 사용된다. '바다이야기', '스포츠'는 도박 게임 종류를 구체적으로 언급하며, '혜택', '매일'은 거래를 촉진하기 위한 보상을 제시하는 문구에 해당한다. 따라서 이 군집은 금전 거래 및 도박을 유도하는 스팸 메시지의 핵심을 담고 있다.

두 번째 군집의 주요 키워드는 '휴먼', '처리', '게이트', '접속', '문의', '해주시', '보유', '디지털', '놀이터', '양해'로, 시스템 문제 해결이나 고객 지원을 가장한 스팸 메시지를 포함한다. '처리', '게이트', '접속'은 기술적 문제를 다루며, '문의', '해주시'는 고객 지원과 관련된 내용을 포함한다. '휴먼' 키워드는 주로 '휴먼 유저', '휴먼 고객' 등으로 사용되고, 스팸 문자의 경우 자동 필터링을 우회하기 위해 대체어를 사용하는 경향을 반영한다면 '휴먼'은 '휴면'으로 이해할 수 있다. 이는 휴면 계정과 관련된 메시지로, 계정 활성화를 유도하는 키워드로 사용된다. 이 군집은 주로 기술 지원 및 고객 서비스 관련 스팸 메시지를 설명하는데 활용될 수 있다.

세 번째 군집의 주요 키워드는 '앞전', '스팸', '신고', '보스', '잔금', '단장', '전액', '제목', '전체', '아이디'로, 스팸 메시지 자체나 금전적 문제와 관련된 메시지가 주를 이룬다. '스팸', '신고', '제목'은 스팸 메시지 자체와 이를 관리하는 내용으로, 사용자가 원치 않는 메시지의 공통적 특징을 반영한다. '앞전', '잔금', '전액'은 이전 거래나 잔금을 언급하며 추가 거래를 유도한다. 또한 '전체', '아이디'는 대량 발송된 마케팅 메시지나 스팸 메일에 활용되는 키워드로, 특정 아이디나 개인 정보를 활용한 메시지일 가능성이 있다. 이 군집은 스팸 메시지 필터링이나 불필요한 마케팅 메시지의 특성을 분석하는 데 유용하다.

실루엣 점수는 군집 내 데이터 포인트의 일관성과 군집 간 구별력을 평가하는 지표로, 1에 가까울수록 군집이 잘 형성된 것을 의미한다. 클러스터링 기법을 통해 도출된 세 개의 군집에 대해 실루엣 점수를 계산한 결과, 0.849이라는 높은 점수가 도출되었다. 이는 각 군집이 명확히 분리되고 데이터가 적절히 배치되었음을 나타낸다.

2.3. 연구 결과

스팸 문자 데이터 분석을 활용하여 교육적 대처 방안을 도출할 수 있다. 첫째, 스팸 문자 수치 분석 결과, 신고량이 증가하며 주로 주중 12-18시 사이에 집중되는 경향이 나타난다. 이 시간대는 학생들이 방과 후 시간을 보내는 동시에 보호자들이 주로 근무 중인 시간으로, 학생들이 자율적으로 시간을 활용하기 쉬운 시기다. 따라서 교사와 보호자들은 이 시간대에 학생들이 스팸 문자를 통해 도박에 노출되지 않도록 주의할 필요가 있다. 둘째, 키워드 분석에 따르면 '놀이터', '이벤트', '신규', '가입' 등 놀이처럼

보이는 단어들이 사용되어 신규 가입을 유도하고 있다. 그러나 이는 표면적으로는 놀이처럼 보이지만, 실제로는 도박과 관련된 내용임을 인식해야 한다. 마지막으로, 첫 번째 군집에서는 금전적 유인과 관련된 키워드가 많이 나타나며, 이는 청소년들이 도박에 빠지는 주요 동기와 밀접한 관련이 있다. 청소년들이 금전적 유인에 휩쓸려 도박에 빠지지 않도록 주의를 기울여야 하며, 자금 마련을 위해 2차 범죄로 이어질 수 있는 위험성을 경계해야 한다.

3. 결론

본 연구에서는 1년 7개월 동안 사용자들로부터 신고된 스팸 메시지 데이터를 분석하여 스팸 메시지의 패턴과 텍스트 내용을 조사하였다. 연구 결과, 스팸 메시지의 신고량이 시간이 지남에 따라 증가하는 추세를 보였으며, 특히 12시에서 18시 사이, 그리고 주중반에 상대적으로 많이 신고되는 경향이 나타났다. 다양한 스팸 카테고리 중에서 도박 관련 메시지가 전체의 62.7%를 차지하며, 스팸 생태계에서 도박 관련 스팸이 큰 비중을 차지하고 있음을 보여준다. 이는 도박 스팸이 사용자들로부터 많이 신고될 뿐만 아니라, 도박 스팸 메시지의 발송량 자체가 다른 유형의 스팸에 비해 상대적으로 많기 때문에, 그 결과 신고량이 더 많아지는 주요 원인 중 하나로 작용했을 가능성을 시사한다.

텍스트 분석에서는 최빈값 분석과 TF-IDF 분석을 통해 도박이나 금전적 유인 관련 용어들이 자주 사용된다는 점이 드러났다. 클러스터링 분석을 통해 스팸 메시지가 크게 세 가지 그룹으로 분류되었으며, 각각 금전 거래, 기술적 문제 해결, 일반적인 스팸 관리 및 불필요한 마케팅과 관련된 메시지로 구분되었다. 이러한 연구 결과는 스팸 문자에 대한 심층적인 이해를 바탕으로 청소년을 대상으로 한 교육적 대처 방안 마련에 중요한 시사점을 제공한다. 스팸 문자가 주로 방과 후 시간대에 집중되며, 이는 학생들이 자율적으로 시간을 활용하는 시기이므로 교사와 보호자의 주의가 필요하다. 또한, '놀이터', '이벤트', '신규', '가입' 등의 키워드로 도박을 놀이처럼 가장해 유도하는 점을 인식하고 경각심을 키워야 한다. 마지막으로, 청소년이 금전적 유인에 의해 도박에 빠지지 않도록 교육을 강화해야 한다.

그러나 본 연구에는 몇 가지 제한점이 있다. 첫째, 신고된 스팸 문자 데이터에만 의존하여 실제 모든 스팸 메시지를 포괄하지 못했다는 점에서 분석 결과가 일부 제한될 수 있다. 둘째, 텍스트 분석 과정에서 불용어나 특수문자, URL을 제거하여 비정상적인 문자 조합이 분석에서 제외되었을 가능성이 있다. 이러한 제한점들로 인해 교육적 대처 방안이 모든 유형의 스팸 메시지나 상황에 적절히 적용되지 않을 가능성이 있다. 그럼에도 불구하고, 본 연구는 스팸 문자 데이터의 인공지능 분석을 통해 청소년 도박 문제에 대한 중요한 시사점을 제시했으며, 향후 연구에서 이 한계를 보완한다면 더욱 효과적인 대응 방안을 마련할 수 있을 것이다.

* 본 연구는 한국인터넷진흥원 KISA(Korea Internet & Security Agency)로부터 스팸 문자 데이터를 제공받아 수행되었음

참고문헌

[1] 한국인터넷진흥원(KISA), "2023년 하반기 스팸 유통현황 보고서", 한국인터넷진흥원, 2024.

[2] 강서구, "당신을 괴롭히는 스팸문자 줄지 않는 이유 : 스팸문자에 숨은 늑장 대응과 탐욕", 더스쿠프,(527), 34-37. 2023.

[3] 이연호, 박영화, 배영목, 한광석, "불법도박 참여 실태와 결정요인", 사회과학연구, 57(1), 101-131. 2018.

[4] 서수연, 허민석, 황재호, 장지원, 신지호, 민무홍. "스팸 메시지 발생 빈도 및 패턴 분석에 관한 연구". 2024 한국컴퓨터종합학술대회(KCC 2024), 제주. 2024.6. 447-449.

[5] 김가현, 유헌창, "딥러닝을 이용한 비정상 문자 조합으로 구성된 스팸 문자 탐지 기법", ACK 2023 학술발표대회 논문집, 서울, 2023, pp. 583-586.

[6] Moohong Min, Jemin Justin Lee, Hyunbeom Park, Kyungho Lee. "Honeypot System for Automatic Reporting of Illegal Online Gambling Sites Utilizing SMS Spam". 2021 World Automation Congress (WAC), Taipei, Taiwan, August 1-5, 2021.

[7] 이승현, 서민수, 조윤오. "청소년 사이버도박 실태 및 대응방안 연구". 한국형사법무정책연구원. 2019.12.

[8] 한국도박문제예방치유원, "청소년 도박문제 선별 안내서", 한국도박문제예방치유원, 2022.

[9] 정현정. "청소년 사이버도박의 실태 분석 및 예방대책에 관한 연구". 한국자치경찰논총, 1(2), 117-136. 2022.