

사이버비행 요인 파악 및 예측모델 개발: 혼합방법론 접근*

손새아** · 신우식*** · 김희웅****

〈 목 차 〉

I. 서론	5.2 예측모델 및 평가
II. 개념적 배경	5.3 예측모델 결과
2.1 사이버비행 배경이론	VI. 토의 및 시사점
2.2 사이버비행 관련 선행연구	6.1 모형 분석 결과 토의
III. 데이터 및 연구방법	6.2 예측모델 변수 중요도
3.1 데이터와 변수	6.3 한계점 및 향후 연구 방안
3.2 분석 절차	6.4 시사점
IV. 사이버비행 요인 파악	참고문헌
4.1 분석 모형	Appendix A. 변수 및 측정항목
4.2 분석 결과	Appendix B. 변수 간 상관관계
V. 사이버비행 예측 모델	<Abstract>
5.1 분석 방법	

I. 서론

최근 스마트폰 보급률이 높아짐에 따라 국내 외에서 사이버범죄 문제는 점차 심각해지고 있고 있다. 경찰청 조사에 따르면 2020년 약 18만 건의 사이버범죄가 발생하였고 이는 2017년 대비 약 37% (131,734 건) 상승한 수치이다 (경찰청, 2020). 특히, 청소년이 연관된 사이버범죄

는 2016년 2,000건을 넘으며 2012년 약 900건이었던 것에 반해 4년 만에 2배의 수치를 기록하였으며(교육부, 2018), 2018년 기준으로 청소년이 사이버범죄에 가담하거나 피해를 입은 경우는 전체 학생의 29.5%로 2017년보다 4.7% 상승하였다 (방송통신위원회, 2019). 또한, 미 여론조사기관 퓨 리서치 센터(Pew Research Center)에 따르면, 미국 청소년의 59%가 사이버

* 이 논문은 2018년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임 (NRF-2018S1A3A2075114)

** LG CNS, saerowah@gmail.com(주저자)

*** 연세대학교 정보대학원, woosik.shin@yonsei.ac.kr

**** 연세대학교 정보대학원, kimhw@yonsei.ac.kr(교신저자)

불링(cyberbullying)을 당한 경험이 있다고 응답하였고, 2016년에는 미국 텍사스주의 한 18세 여학생이 사이버불링에 시달리다 가족들이 보는 앞에서 권총으로 자살하는 사건이 발생하였다 (동아일보, 2019). 전 세계적으로 청소년의 사이버비행 문제는 양적 측면에서 높은 상승 추이를 보이는 것뿐 아니라 질적 측면에서도 피해가 심각해지고 있는 실정이다 (이수봉, 한상철, 2018). 따라서, 청소년 비행이 발생하는 원인요인을 파악하여 예방책 수립에 기여하고 비행 고위험군 청소년을 판별하는 접근이 필요하다.

2000년대 인터넷 사용률 증가와 사이버범죄 피해의 심각성이 증대됨에 따라 사이버비행 요인을 파악하는 다양한 선행 연구들이 진행되어 왔다. 그러나 기존 연구들은 청소년 사이버비행을 설명하는 배경이론을 선택적으로 차용하여 대부분 단일 관점에서 연구가 진행되었으며 (Paez, 2018; Choi et al., 2017; Udris, 2016), 여러 배경이론을 종합적으로 다룬 연구는 드물다. 단일 배경이론을 활용한 연구들은 청소년들의 사이버비행을 야기하는 여러 요인을 포괄적으로 파악하는 것에 한계가 있으므로 다양한 이론을 통합 적용한 종합적인 관점에서의 연구가 필요하다. 게다가 지금까지 진행된 청소년 사이버비행 관련 연구는 사이버비행 가해 요인 탐색에 그치고 있다 (Lee et al., 2017; Paez, 2018; 이창문 등, 2019; 김수진, 강현아, 2016; 성경주, 2019). 요인들을 파악하는 것만으로는 실제 사이버비행 가해자를 탐지 및 예측하여 사이버비행 예방에 기여하는 것에 한계가 있다. 따라서, 다양한 요인 변수들을 활용하여 잠재적 사이버비행 가해자를 예측하는 것이 실질적인

사이버비행 예방하는 것이 필요하다. 또한 연구 방법 측면에서도 일회성 설문조사를 통해 연구를 진행한 경우가 대부분이며 이로 인해 장기간 동안 개인의 특성을 반영한 사이버범죄 요인 탐색 및 가해자 탐지에 대한 연구는 매우 부족하다. 따라서, 개인 특성을 반영한 데이터를 통해 특정 이론이 아닌 종합적 차원에서 사이버범죄 요인을 탐색하고 잠재적 사이버비행 청소년을 예측하는 연구가 필요하다.

본 연구의 목적은 청소년 비행을 설명하는 네 가지 배경이론을 통합적으로 활용하여 사이버비행 요인을 탐색 및 테스트하고 나아가 사이버비행 가담 확률을 예측하여 청소년 사이버비행 현상 예방에 기여를 하는 것이다. 이를 위해 혼합방법론의 연구 방법을 채택하여 연구를 진행한다. 배경이론들을 통해 추출한 잠재 원인 요인들을 계량모형을 통해 테스트하고 비선형 관계성을 밝히며 예측력을 높일 수 있는 딥러닝 기반 예측 모델을 이용하여 사이버비행 가담자를 예측한다. 요인 탐색 및 예측 모델 개발이라는 혼합방법론의 분석 방법은 설명 가능한 인공지능을 구현할 수 있다는 시너지 효과가 존재하고 최종적으로는 청소년 사이버비행 현상을 예방하고 방지할 수 있도록 분석 결과를 제시한다. 이를 위해 한국아동·청소년패널조사 2010(KCYPS 2010)로부터 5개년의 패널 데이터를 사용하여 분석을 수행한다.

II. 개념적 배경

2.1 사이버비행

사이버비행(cyber delinquency) 또는 사이버 일탈(cyber deviance)은 사이버 공간에서 일어나는 법 및 규범 위반 행위를 의미하는 말로 가해자가 청소년일 때 주로 사용된다 (배상률 등, 2017). 비행이라는 용어가 전통적인 범죄라는 용어를 피해 청소년에게만 사용된 것이나, 일반적으로 청소년이 법을 위반하는 행위를 포함하여 청소년의 지위에 맞지 않는 문제 행위 및 기대되는 모범적 행위를 벗어난 일탈 행위를 포괄한다 (송광섭, 2003). 따라서, 청소년 사이버 비행은 좁게는 사이버 저작권을 침해하는 불법 콘텐츠 다운로드와 같은 법에 저촉되는 사이버 범죄부터 넓게는 단순한 욕설 혹은 사이버폭력을 포함한 청소년의 지위에 맞지 않는 온라인 상의 모든 행위를 포괄한다고 할 수 있다 (배상률 등, 2017).

2.2 사이버비행 배경이론

본 연구는 통합적으로 청소년 사이버비행을 설명하는 잠재요인 파악을 위해 네 가지의 기존 사이버범죄 관련 이론을 활용한다. 일반긴장이론 (general strain theory), 사회학습이론 (social learning theory), 사회유대이론 (social bonding theory), 일상행위이론(routine activity theory)을 기반으로 다양한 관점에서 사이버비행 행동을 야기하는 동기들을 살펴보고자 한다 (Agnew, 1992; Akers, 1998; Hirschi, 1969; Cohen & Felson, 1979). 첫 번째로 Agnew (1992)의 일반긴장이론에서는 부모와의 갈등이나 학교 생활 부적응 및 친구 관계에서의 소외 등과 같은 일상생활에서의 긴장이 비행의 원인이 될 수 있다고 주장한다 (이성식, 전신현,

2013). Agnew는 비행의 원인이 되는 긴장의 원천에 대해 목표달성의 실패, 긍정적 자극의 소멸, 부정적 자극의 발생이라는 세 가지 범주를 제시하였으며, 부정적 자극의 발생은 부모의 학대, 친구들의 괴롭힘, 교사의 폭언과 체벌 등 청소년에게 고통을 주는 사건들에 기인한다 (황성현, 이강훈, 2013). 이러한 긴장 상황들이 청소년들에게 화, 우울 등의 부정적 감정을 일으키고, 이를 해소하려는 대처 방식으로 비행이 발생한다고 보는 것이다 (이성식, 2007). 두 번째로 Akers(1998)의 사회학습이론은 사회구성원들 간의 상호작용을 통해 비행이나 일탈이 학습된다고 주장하며, 주요 개념은 차별적 접촉, 정의, 차별적 강화, 모방이 있다 (조은미 등, 2016). 이중 차별적 접촉은 비행에 대해 우호적인 태도를 보이는 친구들과의 직접적인 접촉 및 상호작용을 의미하며, 정의는 주어진 행위에 대해 개인이 부여하는 의미와 태도를 말한다 (황성현, 이강훈, 2013). 즉, 비행에 대해 우호적이거나 합리화를 잘 하는 경향이 강할수록 혹은 그런 친구들과 가까이 지낼수록 자신도 그런 비행을 할 가능성이 커진다 (황성현, 이강훈, 2013).

세 번째로 Hirschi(1969)의 사회유대이론은 사회와의 유대가 약화되었을 때 비행이 발생한다는 전제에서 출발한다. 즉, 청소년들이 모두 잠재적으로 비행 동기를 가지고 있으나 비행 행동으로 원초집단과의 관계가 깨질 것을 두려워하여 비행을 저지르지 않는다고 본다 (황성현, 이강훈, 2013). 네 가지 하위 개념 중 애착 (attachment)은 타인에게 가지는 애정과 존경심, 기대에 부응하기 위한 노력의 정도를 의미하며, 관여(commitment)는 미래를 위한 교육

및 저축 등의 관습적인 활동에, 참여 (involvement)는 학교, 가족, 여가 생활 등의 인 습적인 활동에 열중하는 정도를 말하고, 마지막으로 신념(belief)은 사회적 가치와 규범을 지키 려는 믿음의 정도를 의미한다 (황성현, 이강훈, 2013). 네 번째로 Cohen and Felson(1979)의 일상행위이론은 범죄자의 특성보다 범죄가 발 생하는 상황에 더 집중하여, 가해자와 적합한 대상이 있고 감시가 없는 시간 및 공간에서 대 부분의 범죄가 발생한다고 본다. 이에 따라 비 대면성 및 익명성이 보장되는 사이버공간에서 보내는 시간이 많은 청소년들이 사이버비행을

저지를 상황이나 기회여건에 더 많이 노출되어 실제 사이버비행을 경험할 가능성이 높아질 수 있다 (정혜원, 2010; 고은희, 김은정, 2015). 본 연구에서는 네 가지 이론을 통해 포괄적으로 사이버비행 잠재 요인들을 탐색하고 이를 바탕으로 원인 요인 테스트 및 예측 모델 개발을 수 행한다.

2.3 사이버비행 관련 선행연구

기존의 청소년 사이버비행 관련 연구를 살펴 보면, 사이버비행 현황에 관해서만 연구를 진행

<표 1> 청소년 사이버비행 관련 선행연구

선행연구	배경이론			
	일반긴장이론	사회학습이론	사회유대이론	일상행위이론
Patching and Hinduja (2011)	○ (학업성적, 가족관계, 친구관계)			
Navarro and Jasinski (2012)				○ (인터넷 사용 빈도, 사용 유형)
Jang et al. (2014)	○ (학업긴장, 부모긴장, 비행피해)	○ (비행 친구)		
Udris (2016)			○ (학교 애착, 이웃 애착, 가족식사)	
Choi et al. (2017)		○ (비행 친구, 사이버비행 인식)		
Lee (2018)		○ (비행 친구)		
Paez (2018)	○ (학교 생활, 친구관계, 학업긴장)			
정혜원 (2010)	○ (학업긴장)			○ (인터넷 사용시간)
배성만 (2016)	○ (부모긴장, 학업긴장, 친구갈등, 비행피해)		○ (부모애착, 부모지도 감독, 교사관계)	○ (인터넷 사용, 오프라인 비행)

하거나 (Aricak et al., 2008) 청소년이 사이버비행에 가담하는 원인을 찾아내는 연구가 대부분이다 (Dehue et al., 2008; Patchin & Hinduja, 2011; Navarro & Jasinski, 2012; Jang et al., 2014; Ang, 2015; Udris, 2016; Choi et al., 2017; Lee, 2018; Lee et al., 2018; Paez, 2018; Baldry et al., 2019). 네 가지 배경이론 중 하나 또는 그 이상을 적용한 선행연구를 <표 1>에 정리하였다.

우선, 일반긴장이론을 적용하여 부모, 학업, 친구관계 및 비행 피해경험 등에서 오는 일상 긴장요인을 청소년 사이버비행의 주원인으로 본 연구 (Patchin & Hinduja, 2011; Paez, 2018)가 있다. 이들은 학업에서 오는 스트레스나 친구와의 갈등 등이 청소년의 일상에 긴장을 유발하고 이를 완화하기 위해 사이버비행을 저지른다고 보았다. 나아가 Jang et al. (2014)은 일상 긴장뿐만 아니라 사회학습이론을 적용하여 비행 친구와의 차별적 접촉이 청소년 사이버비행을 일으킨다는 것을 보였다. 일반 비행 친구 외에도 사회학습이론을 적용하여 사이버공간의 비행 친구 수가 청소년 사이버비행을 유의하게 설명한다는 연구 (Lee, 2018)와 사이버비행 친구 수 및 사이버비행에 대한 인식을 청소년 사이버비행의 주요인으로 파악한 연구가 있다 (Choi et al., 2017). 사회학습이론을 통해 주변에 일반 비행이나 사이버비행을 저지르는 친구가 있고 이에 대해 심각하게 여기지 않거나 합리화를 통해 따라 하게 될 때 사이버비행을 경험하기 쉽다고 봤다.

일상행위이론은 해킹, 사이버불링과 같은 사이버비행의 원인 요인을 설명하기 위해 다양하게 활용되어왔다 (e.g., 김민수, 김희웅, 2020;

Navarro & Jasinski, 2012). 구체적으로 기존 연구에서 일상행위이론을 적용하여 인터넷 사용 빈도 및 사용 유형과 이에 대한 부모의 감독 행태를 청소년 사이버비행의 주요인으로 파악하였다 (Navarro & Jasinski, 2012). 인터넷 사용이 많고 인스턴트 메시지(IM), 소셜 네트워크 서비스(SNS), 온라인 채팅룸 등을 주로 사용할 때 사이버비행을 저지를 가능성이 높고, 부모님이 웹사이트 필터 소프트웨어를 사용할 때 청소년 사이버비행을 예방하는 데 영향을 미치는 것으로 나타났다. 정혜원 (2010)은 비행 기회 측면에서 인터넷 사용 시간과 함께 학업 성적으로 인한 일상긴장을 청소년 사이버비행의 주요인으로 보았다. 마지막으로, 사회유대이론을 적용하여 가족과의 식사 시간 등 참여 요인을 청소년 사이버비행의 주요인으로 본 연구가 있다 (Udris, 2016). 즉, 사회에 대한 애착이나 참여가 높을수록 사이버비행을 저지릴 가능성이 낮다고 봤으며, 이에 더해 일상 생활에서의 긴장, 인터넷 사용 및 오프라인 비행을 청소년 사이버비행의 주요인으로 파악한 연구도 있다 (배성만, 2016). 이처럼 대부분의 청소년 사이버비행 관련 선행연구는 단일 이론을 기반으로 사이버비행 원인 요인을 탐색하고 있고 여러 배경이론을 종합적으로 다룬 연구는 많지 않다. 또한, 청소년 사이버비행 예측 모델을 수행하는 연구는 전혀 없는 실정이다. 이에 본 연구는 네 가지 배경이론을 이용하여 다양한 관점에서의 청소년 사이버비행 원인 요인을 탐색 및 검증하고 딥러닝 기반 예측 모델링을 진행해 청소년 사이버비행 예방을 위한 통합적 분석을 진행하고자 한다.

Ⅲ. 데이터 및 연구 방법

3.1 데이터와 변수

본 연구는 한국청소년정책연구원의 KCYPS 2010 데이터를 이용한다. KCYPS 2010은 아동 및 청소년의 성장과 발달의 다양한 양상을 종합적으로 파악하는 것을 목적으로 하고 있으며 (한국청소년정책연구원, 2017), 본 연구는 중1 코호트 중 사이버비행 항목이 조사된 중학교 2학년부터 고등학교 3학년까지 5개년의 데이터를 분석 대상으로 하였다.

본 연구 데이터의 종속변수는 사이버비행 연간 경험여부로서, 5개의 세부 응답 항목 중 하나라도 경험이 있다고 응답한 경우 사이버비행 경험이 있는 것으로 코딩(있다: 1, 없다: 0)하여 사용하였다. 구체적으로 종속변수의 세부 항목은 (1) 다른 사람 컴퓨터나 웹사이트 해킹하기, (2) 다른 사람의 아이디나 주민번호 허락 없이 사용하기, (3) 불법 소프트웨어 다운받아 사용하기, (4) 채팅/게시판 등에서 욕이나 폭력적인 언어 사용하기, (5) 채팅/게시판 등에 일부러 거짓 정보 올리기가 있다. 본 연구에서 사용한 설명변수는 <표 2>와 같으며, 네 가지 배경이론과 각 이론 내의 주요 개념에 따라 변수를 추출하였다. 일상긴장이론에서는 관계가 좋지 않거나 갈등이 있을 때 일상적인 긴장을 유발할 수 있는 교우관계, 교사관계, 비행 피해 경험을 고려하고, 이로 인한 부정적 감정 측면의 우울을 포함시켰다. 사회학습이론의 차별적 접촉 측면에서는 비행 행동경험 친구 수를 통해 비행 친구와의 교류 정도를 고려하였다 (Jang et al., 2014).

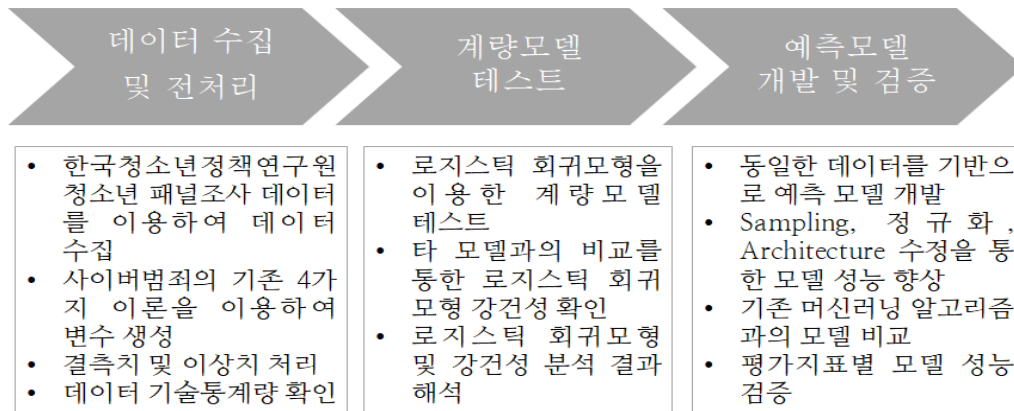
사회유대이론의 네 가지 주요 개념에 해당되는 변수를 모두 고려하였다. 우선, 사회유대이론을 적용하여 지역 및 이웃 간 애착이 사이버비행에 유의한 영향을 미친다(Udris, 2016)는 연구를 기반으로 지역사회에 대한 인식이 좋을 경우 사회에 대한 애정이 강할 것으로 보고 애착 측면에서 지역사회 인식을 선택하였다. 학습활동을 열심히 하거나 체육시간 중 운동시간이 많을 경우 미래를 위한 관습적 활동에 열중하는 것으로 보고 관여 측면에서 학습활동과 체육시간 중 운동시간을 선택하였다. 또한, 체험활동, 교내외 동아리활동, 문화활동이 많을 때 사회에 대한 참여가 높을 것으로 보고 참여 측면에서 위 세 가지 변수를 선택하였다. 마지막으로 학교 규칙을 잘 지키거나 공동체 의식이 높을 경우 사회적 가치와 규범을 지키려는 믿음이 클 것으로 보고 신념 측면에서 학교규칙과 공동체의식을 선택하였다. 일상행위이론의 비행 기회 측면에서는 사이버공간에서 보내는 시간이 많은 경우를 고려하여 온라인 팬덤활동 빈도를 새로운 변수로 포함하였고, 댓글달기 등 세부적인 목적의 컴퓨터 이용빈도 및 휴대전화 이용빈도를 고려하였다. 이는 기존 연구에서 인터넷 사용 빈도를 비행 기회 변수로 활용한 기존 문헌과 일치한다 (Navarro & Jasinski, 2012). 설명변수에 대한 자세한 내용은 appendix에서 확인할 수 있다.

3.2 분석 절차

본 연구는 크게 데이터 전처리 및 탐색, 계량 모델 테스트, 예측 모델 개발 및 검증의 3단계로 분석을 진행한다. 계량경제모델 검증과 예측

<표 2> 배경이론을 이용한 독립변수 선정 내용

이론	주요 개념	변수 설명	변수 척도
일반긴장이론 (Agnew, 1992)	부정적 감정	우울	4
	일상 긴장	교우관계 교사관계 연간 비행 피해경험 유무	4 4 0 or 1
사회학습이론 (Akers, 1998)	비행 친구와의 접촉	연간 비행 행동경험 친구 수	0 이상 정수
사회유대이론 (Hirschi, 1969)	애착	지역사회 인식	4
	관여	학습활동 체육시간 중 운동시간	4 5
	참여	연간 체험활동 참여경험 유무 연간 동아리활동 참여경험 유무 연간 문화활동 경험횟수	0 or 1 0 or 1 0 이상 정수
	신념	학교규칙 공동체 의식	4 4
일상행위이론 (Cohen & Felson, 1979)	비행 기회	연간 비행 행동경험 유무 컴퓨터 용도별 이용 빈도 휴대전화 용도별 이용 빈도 온라인 팬덤활동 빈도	0 or 1 5 5 5



<그림 1> 전체 분석 절차

모델 생성을 거쳐 혼합방법론의 구성을 완성하여 청소년 사이버비행 예방을 위한 통합적 분석 결과를 제시한다. 이를 통해 특정 가해자 탐지, 가해자의 가해 요인 탐색을 결합한 설명 가

능한 예측 모델을 개발하며 최종적인 메타인퍼런스를 도출한다. 이후 한국 사회에서 청소년의 사이버비행 방지 및 예방을 위한 방안을 제안한다. 전체 데이터 분석 절차는 다음 <그림 1>

과 같다.

먼저 데이터 전처리 및 탐색 단계에서는 KCYPS 2010 데이터를 통해 데이터를 수집한다. 이후 네 가지의 배경이론을 종합하여 사이버비행 이론에 부합하는 변수를 추출해낸다. 이를 통해 기초 데이터 세트를 완성하며 전처리 과정을 진행한다. 전처리가 완료된 데이터에 대

해서 기초 통계 분석을 시행해 사이버비행 현상을 파악한다. 다음으로 계량 모델 테스트 과정을 진행한다. 계량 모델 테스트의 경우 로지스틱 회귀 모형을 이용하여 분석을 진행한다. 이후 모형 분석 결과를 기반으로 청소년 사이버비행에 유의미한 영향을 끼치는 변수가 무엇인지 파악한다. 마지막으로 예측 모델 개발 및

<표 3> 각 변수에 대한 기술통계량

변수	최소값	최대값	평균	표준편차	
성별	0	1	0.53	0.499	
체육시간 중 운동시간	1	5	2.59	1.292	
우울	1	4	1.90	0.564	
연간 비행 행동경험 유무	0	1	0.22	0.414	
연간 비행 피해경험 유무	0	1	0.05	0.211	
연간 비행 행동경험 친구 수	0	330	5.90	15.497	
학습활동	1	4	2.81	0.555	
학교규칙	1	4	2.93	0.5110	
교우관계	1	4	3.12	0.433	
교사관계	1	4	2.89	0.605	
지역사회 인식	1	4	2.65	0.588	
공동체의식	1	4	2.89	0.550	
연간 체험활동 참여경험 유무	0	1	0.64	0.480	
연간 교내외 동아리활동 참여경험 유무	0	1	0.53	0.499	
연간 문화활동 경험횟수	0	100	4.53	5.808	
온라인 팬덤활동 빈도	0	4	0.87	0.986	
컴퓨터 이용 빈도	공부/학습 관련 정보 검색 및 사이트 이용	0	4	2.49	1.191
	학습 이외의 정보 검색 및 자료 이용	0	4	2.83	1.237
	게임 및 오락	0	4	2.65	1.370
	채팅이나 메신저 및 블로그, SNS 활동	0	4	1.97	1.102
	동호회, 카페, 커뮤니티 활동	0	4	1.82	1.135
	댓글 달기	0	4	1.88	1.162
	19세 이상 사이트 이용	0	4	1.30	0.847
휴대폰이용 빈도	가족과의 통화	0	4	3.31	0.889
	가족과의 문자메시지 (카카오톡, 라인 등 포함)	0	4	3.07	0.956
	친구와의 통화	0	4	3.35	0.918
	친구와의 문자 메시지 (카카오톡, 라인 등 포함)	0	4	3.56	0.881
	게임 및 오락	0	4	2.75	1.179
사이버비행 연간 경험여부	0	1	0.27	0.443	

검증 단계를 진행한다. 계량 모델 테스트에 이용된 동일한 데이터를 기반으로 예측 모델 개발을 진행한다. 예측 모델은 딥러닝 알고리즘인 long short-term memory(LSTM)을 통해 진행하며 샘플링 기법, 단위 정규화, 네트워크의 구조 수정 과정을 통해 예측 성능을 높인다. 이후 재현율(recall), 정밀도(precision) 등 다양한 평가지표를 기준으로 개발된 모델을 평가한다.

데이터 수집 및 전처리 과정을 살펴보면, 우선 청소년 사이버비행을 설명하는 네 가지 이론에 근거하여 원인 변수들을 추출하였고 응답 항목이 많은 변수의 경우 평균값을 계산하여 사용하였다. 다음으로 사이버비행 항목이 무응답인 개체는 제거하였고 이상치 및 결측치로 표기된 항목을 가진 개체는 적절한 값으로 대체하는 전처리 과정을 진행하였다. 데이터 전처리 후 1,740명의 청소년을 대상으로 5개년(중 2~고3)에 걸친 균형 패널이 구성되어 8,700개의 샘플을 확보하였다. 변수에 대한 기술통계량은 <표 3>과 같으며, 평균적으로 연간 약 27%의 청소년들이 적어도 한 종류의 사이버비행을 경험한 것으로 나타났다. 변수 간 상관관계는 appendix에서 확인할 수 있으며, 휴대폰 이용빈도 내의 변수 간 상관관계수가 높게 나타나 다중공선성 (multicollinearity) 우려가 제기되었으나 분산 팽창 인수(variance inflation factor; VIF)를 구하여 본 결과 기준치인 5를 넘는 변수가 없었고, 최대치가 3.23이고 평균이 1.77로서 다중공선성에 문제가 없는 것으로 판단하였다. 따라서 앞서 제시한 설명변수를 모두 포함하여 분석을 진행하였다.

IV. 사이버비행 요인 파악

4.1 분석 모형

청소년 사이버비행 경험여부를 설명하는 주요 요인을 파악하기 위해 고정효과 로지스틱 회귀모형(fixed effects logistic regression model)을 활용한다. 종속변수가 사이버비행 경험이 있을 경우 1, 없을 경우 0인 이항변수(binary variable)이므로 이에 적합한 로지스틱 회귀모형을 선택하였으며, 관측되지 않은 개체들의 고유 특성들과 시간 추세를 통제하기 위해 고정효과모형을 적용하였다. 모형식은 다음과 같다.

$$y_{it} = \begin{cases} 1 & \text{if } y_{it}^* > 0 \\ 0 & \text{if } y_{it}^* \leq 0 \end{cases}$$

$$y_{it}^* = x_{it}\beta + u_i + \tau_t + \varepsilon_{it}$$

y_{it} 는 청소년 i 의 t 시점에서의 사이버비행 경험여부 관측 값이고, y_{it}^* 는 청소년 i 의 t 시점에서의 실제 사이버비행 경험여부를 나타내는 latent variable이다. x_{it} 는 네 가지 이론을 기반으로 추출한 설명변수 벡터이고, u_i 는 시간에 따라 변하지 않는 개별 청소년의 특성을 반영하는 고정효과이며, τ_t 는 시간(year-level)의 특성을 반영하는 시간 고정효과이다. ε_{it} 는 오차항이다. 본 연구에서는 STATA 13.0에서 제공하는 로지스틱 회귀에 고정효과 모형을 적용하여 분석을 수행하였다.

4.2 분석 결과

고정효과 로지스틱 회귀모형을 추정된 결과는 <표 4>와 같다. 일반긴장이론의 일상 긴장

<표 4> 로지스틱 회귀 모형 추정 결과

이론	변수 카테고리	설명변수	추정치 (표준오차)		
일반긴장이론 (Agnew, 1992)	부정적 감정	우울	0.112 (0.098)		
		일상 긴장	교사관계	-0.134 (0.083)	
	교우관계		0.116 (0.123)		
	비행 피해경험		0.613*** (0.087)		
사회학습이론 (Akers, 1998)	비행 친구와의 접촉	비행 행동경험 친구 수	0.006* (0.001)		
사회유대이론 (Hirschi, 1969)	애착	지역사회 인식	-0.192* (0.094)		
	관여	학습활동	-0.054 (0.097)		
		체육시간 중 운동시간	0.081* (0.037)		
	참여	체험활동	0.228* (0.092)		
		교내외 동아리활동	교내외 동아리활동	0.125 (0.089)	
			문화활동	0.013 (0.007)	
	신념	학교규칙	-0.255* (0.106)		
		공동체의식	-0.222** (0.095)		
일상행위이론 (Cohen & Felson, 1979)	비행 기회	비행 행동경험	0.313** (0.105)		
		온라인 팬덤활동 빈도	0.322*** (0.056)		
		컴퓨터	공부/학습 관련 정보 검색 및 사이트 이용	-0.049 (0.050)	
			학습 이외의 정보 검색 및 자료(음악·영화 파일 등) 이용	0.095 (0.052)	
			게임 및 오락	0.090 (0.051)	
			이용 빈도	채팅이나 메신저 및 블로그, SNS 활동	-0.083 (0.058)
				동호회, 카페, 커뮤니티 활동	-0.060 (0.052)
		댓글 달기		0.130* (0.052)	
		19세 이상 사이트 이용	0.313*** (0.060)		
		휴대폰	가족과의 통화	-0.057 (0.079)	
			가족과의 문자 메시지 (카카오톡, 라인 등 포함)	-0.124 (0.066)	
			이용 빈도	친구와의 통화	0.026 (0.078)
				친구와의 문자 메시지 (카카오톡, 라인 등 포함)	0.173* (0.079)
				게임 및 오락	0.039 (0.043)
		Log Likelihood		-5591.3887	

*p<0.05, **p<0.01, ***p<0.001

측면에서 비행 피해경험과 사이버비행을 저지를 가능성이 유의한 양의 상관관계를 보였고, 사회학습이론의 비행 친구 접촉 측면에서 비행 행동경험 친구 수와 사이버비행 경험 가능성 역시 유의한 양의 상관관계가 있는 것으로 나타났다. 사회유대이론의 애착 측면에서 지역사회 인식과 신념 측면에서 학교 규칙 및 공동체 의식은 사이버비행을 저지를 가능성과 유의한 양의 상관관계를 보인 반면, 관여 측면에서 체육시간 중 운동시간과 참여 측면에서 체험활동은 사이버비행 경험 가능성과 유의한 양의 상관관계를 갖는 것으로 나타났다. 일상행위이론의 비행 기회 측면에서는 비행 행동경험 및 온라인 팬덤활동 빈도가 사이버비행 경험 가능성과 유의한 양의 상관관계를 보였다. 또한, 댓글 달기 및 19세 이상 사이트 이용을 목적으로 컴퓨터를 많이 사용할 때와 친구와 문자 메시지를 주고받기 위해 휴대폰을 많이 사용할 때 사이버비행을 저지를 가능성과 유의한 양의 상관관계가 있는 것으로 나타났다.

V. 사이버비행 예측 모델

5.1 분석 방법

청소년 사이버비행을 예측하기 위해서 본 연구는 딥러닝 알고리즘인 LSTM을 주된 분석 방법으로 이용한다. LSTM은 recurrent neural network (RNN) 계열의 딥러닝 알고리즘으로 데이터의 시간적 관계를 반영하는 것이 가능하다 (De Bruyn et al., 2020). 또한 RNN의 경우 long-term dependency에서 경사 소실 (gradient

vanishing) 경사 폭주 (gradient exploding) 문제에 취약한 반면에 LSTM은 단순 RNN과는 달리 긴 시간 차이에 따른 변수의 영향력 변화를 반영할 수 있다는 장점이 있다 (Kim & Won, 2018). LSTM은 memory cell과 더불어 forget gate, input gate, output gate 등 세 개의 게이트로 구성되어 있으며 세 게이트를 통해 memory cell에 다음 timestamp로 전달할 정보 비율을 결정하는 방법으로 학습이 이루어진다. 사이버비행 행동이 지난해의 사이버비행 행동과 요인 변화에 영향을 받는다는 점에서 LSTM을 이용하여 사이버비행을 예측하는 모델을 본 연구의 주 방법으로 채택하였다.

5.2 예측모델 및 평가

본 연구에서는 예측 모델 개발에서도 네 가지 이론을 종합하여 데이터를 정제된 계량경제 모형 검증 단계의 데이터를 이용하나 변수 간 단위 규모가 다르기 때문에 MinMax 정규화 방식을 통해 예측 모델에서는 모든 변수의 단위를 정제하였다. 정제된 데이터는 무작위로 70%는 훈련 데이터(training data)로 나머지 30%는 테스트 데이터(test data)로 나누었다. 다음으로 최적의 LSTM 네트워크를 구축하기 위해 하이퍼 파라미터 조절 과정을 거쳤다. LSTM 레이어의 개수를 최소 1개에서 최대 3개까지 조절하였으며 각 레이어 이후 배치 정규화(batch normalization) 적용 여부, 각 레이어마다 dropout 적용 여부 및 비율 최적화 작업을 진행하였다. 배치 크기(Bath size)와 에포크(epoch)는 32에서 512 그리고 10에서 50의 범위에서 실험을 진행하였다. 이때, 1년을 하나의 시간

<표 5> 각 모델별 성능 지표 평가

	정밀도	재현율	F1-score
LSTM	0.82	0.89	0.86
로지스틱 회귀	0.70	0.50	0.59
의사결정나무	0.74	0.80	0.77
SVM	0.50	0.64	0.56
K-NN	0.73	0.61	0.65
Random forest	0.95	0.70	0.81
XGBoost	0.94	0.71	0.81

단위로 설정하여 LSTM이 시계열 정보를 학습할 수 있도록 네트워크를 구축하였다. 최적화 작업 이후 최종적으로 총 2개의 LSTM 레이어를 가졌으며 각 LSTM 레이어 사이에 배치 정규화를 적용한 최적의 네트워크를 도출하였다. 첫 레이어에는 0.1의 dropout이 적용되었으며 활성화함수로는 시그모이드(sigmoid)함수를 이용하였다. 옵티마이저(optimizer)로는 root means square propagation(RMSProp)을 이용하였고 32의 배치 크기로 총 40 에포크 학습을 진행했다. 성능평가 방법으로 정밀도, 재현율, F1-score를 종합적으로 비교하였다. 정밀도의 경우, 사이버비행 행동으로 예측된 청소년 중 실제 사이버비행에 가담한 비율을 의미한다. 반면, 재현율은 실제 사이버비행에 가담한 청소년을 모델이 사이버비행 가담자로 예측하는 비율을 의미한다. 즉, 사이버비행에 가담하였지만 이를 탐지해내지 못하는 경우를 방지하기 위해서는 높은 재현율 값을 확보해야 한다. F1-score는 재현율과 정밀도의 조화평균으로 두 성능 지표를 종합적으로 고려하는 것이 가능하다. 본 연구의 종속변수가 불균형하다는 점을 고려하

여 재현율과 F1-score를 주요 성능 지표로 모델을 평가한다.

5.3 예측 모델 결과

본 연구는 LSTM 모델 성능을 평가하기 위해 기존 기계학습 알고리즘인 로지스틱 회귀, 의사결정나무, support vector machine, KNN, random forest, XGBoost 모델들과 비교하였다. 기존 기계학습 알고리즘은 시계열 특성을 반영하지 않고 각각의 샘플에 대해서 사이버비행을 예측하는 모델로 구성하였다. 따라서 성능 비교를 통해 패널 데이터를 바탕으로 예측 모델을 개발하는 데 있어 LSTM 네트워크가 시계열 정보를 반영할 수 있다는 점에서 다른 모델보다 우수함을 확인할 수 있다. 각 모델별 여러 성과 지표에 따른 평가 결과는 다음 <표 5>와 같다. 정밀도 측면에서는 LSTM, XGBoost, random forest 모형이 모두 비슷한 수준을 보여주고 있으나 재현율 및 F1-score에서 본 연구의 LSTM 모델이 가장 우수한 성능을 보여주고 있다. 특히 재현율의 경우 LSTM 모델을 제외하고 모든

모델이 0.8 미만의 수치를 보여주었는데 본 연구의 LSTM 모델은 0.89의 재현율을 보여주고 있어 실제 사이버비행 가해자를 예측하는데 가장 우수한 성능을 보이고 있음을 확인하였다.

VI. 토의 및 시사점

6.1 모형 분석 결과 토의

본 연구는 네 가지의 배경이론을 통해 청소년의 사이버비행 가해 요인을 통합적으로 탐색한 뒤 사이버비행 가해자를 예측하는 모델을 개발하였다. 먼저 계량 모델 테스트 단계에서 분석 모형 추정 결과에 따라 청소년 사이버비행을 설명하는 유의한 변수를 확인할 수 있었다. 일반긴장이론의 일상긴장 측면에서 비행 피해경험은 사이버비행 경험유무 및 경험횟수와 양의 상관관계를 보였다. 심한 놀림이나 조롱, 집단 따돌림, 폭행, 욕설 등의 비행 피해를 경험하는 것이 청소년들에게 일상적인 긴장을 유발하는 부정적 자극이 된다. 이를 해소하기 위한 행동으로 청소년의 사이버비행이 증가하는 것으로 추정된다 (Patchin & Hinduja, 2011; Jang et al., 2014; 배성만, 2016). 사회학습이론의 차별적 접촉 측면에서 비행 행동경험 친구 수는 사이버비행을 저지를 가능성과 양의 상관관계를 갖는 것으로 나타났다. 비행을 저지르는 친구가 많은 청소년이 이들과의 교류를 통해 비행을 학습하고 따라할 수 있다. 이 때 일반 비행 행동 뿐만 아니라 사이버비행을 경험할 가능성도 높아지는 것으로 볼 수 있다 (Jang et al., 2014; Choi et al., 2017).

사회유대이론의 관여 측면에서 체육시간 중 운동시간과 참여 측면에서 체험활동은 기존 문헌들과의 결과와 달리 사이버비행을 저지를 가능성과 양의 상관관계를 보였다. 체육시간 중 운동시간과 체험활동의 증가는 관습적인 활동 및 참여를 높이는 차원에서 사회와의 유대가 강화되어 사이버비행이 감소할 것으로 기대되었는데, 이러한 청소년들이 오히려 사이버비행을 경험할 가능성이 높을 것으로 나타났다. 청소년들은 운동시간과 체험활동 시간을 통해 다른 청소년들과 유대관계를 형성하기도 하고 다양한 환경에 노출된다. 청소년의 성향에 따라 친구와의 유대 관계 유지와 과도한 외부 환경에 대한 노출은 오히려 청소년들에게 스트레스 등의 부정적인 자극으로 이어지기도 한다 (위휘, 2005). 이러한 부정적 자극을 해소하기 위해 청소년들이 사이버비행을 경험할 가능성이 높아지는 것으로 판단된다. 반면, 사회유대이론의 애착 측면에서 지역사회 인식과 신념 측면에서 공동체의식은 사이버비행 경험유무 및 경험횟수와 음의 상관관계를 보였다. 지역사회에 대한 인식이 좋으면 사회에 대한 애착이 많고, 공동체의식이 높으면 사회적 가치나 규범을 지키려는 믿음, 즉 신념이 높다고 할 수 있다. 이러한 애착 및 신념 측면에서 사회와의 유대가 강할 때 사이버비행을 저지를 가능성이나 횟수가 낮아진다고 볼 수 있다 (배성만, 2016; 황성현, 이강훈, 2017).

일상행위이론의 비행 기회 측면에서는 온라인 팬덤활동 빈도와 댓글달기 및 19세 이상 사이트 이용을 목적으로 한 컴퓨터 이용빈도가 사이버비행을 저지를 가능성과 양의 상관관계를 보였다. 이는 비대면성과 익명성을 특징으로

하는 온라인 공간에 많이 머물수록 사이버비행 기회에 더 많이 노출됨에 따라 사이버비행 행동 가능성도 높아지는 것으로 볼 수 있다 (정혜원, 2010; Navarro & Jasinski, 2012). 또한 비행 행동경험이 사이버비행 경험유무 및 경험횟수와 유의한 양의 상관관계를 보였다. 일반 비행 행동경험이 많은 청소년이 환경적으로 사이버비행을 저지를 기회에도 더 많이 노출될 것이다. 따라서 실제 사이버비행을 저지를 가능성과 횟수 모두 높을 것으로 추정된다 (배성만, 2016). 친구와의 문자 메시지를 위한 휴대폰 이용빈도는 사이버비행을 저지를 가능성 및 횟수와 양의 상관관계를 보인 반면, 채팅이나 메신저 및 블로그, SNS 활동을 위한 컴퓨터 이용빈도는 사이버비행 경험횟수와 음의 상관관계를 갖는 것으로 나타났다. 친구와의 문자 메시지, 채팅이나 메신저 사용, 블로그 및 SNS 활동은 모두 커뮤니케이션이 목적인 행동이라 할 수 있다. 이를 위해 컴퓨터와 휴대폰을 이용할 때 사이버비행 경험과의 관계가 각각 반대로 나타난 것은 기기의 특성에 기인하는 것으로 보인다. 컴퓨터보다 휴대폰이 더 사적인 기기(private device)로서 항상 지니고 다니며 일상 생활의 많은 부분을 함께한다. 즉, 감시나 비밀 노출에서 보다 자유롭게 휴대폰을 이용한다는 점에서 커뮤니케이션을 목적으로 한 컴퓨터 이용빈도는 적고 휴대폰 이용빈도는 많을 때 사이버비행 경험횟수가 높아질 것으로 판단된다.

이후 예측 모델 개발 단계에서는 LSTM기반 사이버비행 예측모델을 개발하였다. LSTM 모델이 재현율과 F1-score에서 가장 우수한 성능을 기록하였으며 LSTM 모델을 구축할 때, 1년

을 기준으로 시계열 정보를 반영하도록 학습한 것을 고려하면 청소년 사이버비행 가해 여부를 예측할 때 시간적 특성이 매우 중요하게 작용하고 있음을 알 수 있다. 특히 재현율의 경우 구축한 LSTM 모델이 기존 로지스틱 회귀모델보다 월등히 성능이 좋았던 것을 통해 사이버비행에 가담한 청소년을 탐지해내는 것은 시간적 특성을 포함해 복잡한 비선형적 모습을 띠고 있음을 유추할 수 있다.

6.2 예측모델 변수 중요도

이후 본 연구에서는 딥러닝 모델의 black box 문제를 보완하기 위한 과정을 진행하였다 (Lundberg & Lee, 2017; Ribeiro et al., 2016). 최근 정보시스템 분야의 연구에서 딥러닝 모델의 black box 문제를 해결하고 설명 가능한 모델을 만들기 위해 주요 변수들을 규명하고 영향력을 측정하는 shapley value와 LIME(local interpretable model-agnostic explanations)과 같은 기법들이 활용되고 있다 (조보근 등, 2020). 본 연구에서는 게임 이론에서의 shapley value 개념을 이용하였다 (Shapley, 1953; Roth, 1988). SHAP value는 각 변수가 최종 정확도에 영향을 미치는 정도를 측정하여 딥러닝 및 머신러닝 모델의 설명력을 보충하는 데 이용되고 있다(Antwarg et al., 2019; Mokhtari et al., 2019). 본 연구에서는 LSTM 모델의 변수별 SHAP value를 통해 각 변수가 예측 성능 향상에 미친 영향도를 측정하였으며 각 변수별 SHAP value와 계량모델 테스트 단계에서의 유의성을 정리한 결과는 다음 <표 6>과 같다.

<표 6> 계량모델 테스트의 변수 유의성과 LSTM의 SHAP Value 비교

배경이론	변수 명		추정치 (표준오차)	정규화된 SHAP value (Ranking)
일반긴장이론 (Agnew, 1992)	우울		0.112 (0.098)	0.079 (17)
	교사관계		-0.134 (0.083)	0.050 (19)
	교우관계		0.116 (0.123)	0.022 (24)
	비행피해경험		0.613*** (0.087)	0.003 (26)
사회학습이론 (Akers, 1998)	비행 행동경험 친구 수		0.006* (0.001)	0.030 (23)
사회유대이론 (Hirschi, 1969)	지역사회 인식		-0.192* (0.094)	0.328 (4)
	학습활동		-0.054 (0.097)	0.033 (21)
	체육시간 중 운동시간		0.081* (0.037)	0.190 (10)
	체험활동		0.228* (0.092)	0.405 (2)
	교내외 동아리 활동		0.125 (0.089)	0.067 (18)
	문화활동		0.013 (0.007)	0.100 (13)
	학교 규칙		-0.255* (0.106)	0.009 (25)
	공동체 의식		-0.222** (0.095)	0.034 (20)
일상행위이론 (Cohen & Felson, 1979)	비행 행동경험		0.313** (0.105)	0.383 (3)
	온라인 팬덤 활동 빈도		0.322*** (0.056)	0.172 (11)
	컴퓨터 이용빈도	공부/학습	-0.049 (0.050)	0.226 (6)
	컴퓨터 이용빈도	정보 검색	0.095 (0.052)	0.101 (12)
	컴퓨터 이용빈도	게임 및 오락	0.090 (0.051)	0.096 (16)
	컴퓨터 이용빈도	채팅, SNS	-0.083 (0.058)	0.552 (1)
	컴퓨터 이용빈도	카페, 커뮤니티	-0.060 (0.052)	0.200 (7)
	컴퓨터 이용빈도	댓글 달기	0.130* (0.052)	0.100 (14)
	컴퓨터 이용빈도	19세 이상 사이트	0.313*** (0.060)	0.232 (5)
	휴대전화 이용빈도	가족과의 통화	-0.057 (0.079)	0.100 (15)
	휴대전화 이용빈도	가족과의 문자	-0.124 (0.066)	0.032 (22)
	휴대전화 이용빈도	친구와의 통화	0.026 (0.078)	0.002 (27)
	휴대전화 이용빈도	친구와의 문자	0.173* (0.079)	0.191 (9)
	휴대전화 이용빈도	게임 및 오락	0.039 (0.043)	0.199 (8)

*p<0.05, **p<0.01, ***p<0.001

패널 데이터를 활용한 계량경제 모형인 로지스틱 회귀 모형에서 변수 유의미성과 예측 모델인 LSTM 알고리즘에서의 SHAP value를 파악한 결과 로지스틱 회귀모형에서 유의미한 변수가 SHAP value 역시 높은 경향이 있음을 알 수 있다. 대부분의 유의한 변수는 정규화된 SHAP value가 0.1 이상을 기록하고 있다. 하지만 유의미한 변수이나 SHAP value가 낮게 도출되거나 유의미하지 않지만 SHAP value가 높게 도출된 경우도 존재한다. 비행피해경험, 학교 규칙은 로지스틱 회귀 모형 결과 유의미한 변수이지만 SHAP value는 매우 낮은 수치를 기록하고 있다. 반면, 컴퓨터를 통한 채팅과 커뮤니티 활동, 휴대폰을 이용한 게임 및 오락은 유의미하지 않은 변수이지만 매우 높은 정규화된 SHAP value 수치를 기록하고 있다.

이러한 차이의 원인으로는 변수 간 관계의 선형성과 비선형성을 들 수 있다. 본 연구의 인과관계 모형의 경우, 독립변수와 종속변수의 선형 관계를 가정으로 통계적으로 유의미한 요인들을 탐색하기 때문에 변수 간 비선형적인 관계를 설명하는 데 한계가 있다 (Friston et al., 1994). 반면에 LSTM모형은 활성화 함수를 통해 독립변수와 종속변수의 비선형적 관계를 반영하는 데 계량경제모형보다 장점이 있다 (Lee et al., 2016; Wong et al., 2020). 이를 바탕으로 계량테스트에서 통계적으로 유의하지는 않으나 높은 SHAP value 값을 지닌 변수는 실무적으로 예측 모델을 개발할 때 매우 중요한 역할을 할 수 있으며, 반대의 경우 예측 모델 개발에는 큰 효용이 없지만, 원인 탐색에는 의미미해 사이버비행 예방 방안 마련에는 유효한 변수라고 해석할 수 있다.

본 연구는 계량경제모형 테스트와 예측 모델 개발의 두 단계로 이루어진 혼합방법론을 적용하였다. 서로 다른 두 가지 방법론을 적용함으로써 각각의 단일 분석에서는 파악할 수 없었던 새로운 시사점을 도출할 수 있다. 또한, 두 분석을 결합함으로써 통합적 청소년 사이버비행 분석 프레임워크 구축 등 다양한 효과를 지닌다 (Lundberg & Lee, 2017; Ribeiro et al., 2016). 먼저, 계량모형 테스트를 통해 청소년 사이버비행 가해 요인을 파악할 수 있으며 이는 다시 사이버비행 예방 방안 제안과 연결된다. 이후의 예측 모델 개발은 사이버비행 청소년을 높은 정확도로 예측할 수 있도록 만들어 주며 각 청소년별로 사이버비행 가해 확률을 도출해 잠재적 청소년 사이버범죄 가해자를 탐색할 수 있도록 한다. 이는 사이버비행 예방을 위한 대상, 시점, 원인, 방안과 연결이 되는 분석 결과이다. 또한 이 두 분석을 결합함으로써 딥러닝 모델을 이용하였던 예측 모델의 설명 원리를 파악할 수 있고 black box 문제를 해결할 방안이 될 수 있다 (Lundberg & Lee, 2017; Ribeiro et al., 2016). 이는 딥러닝 기반 예측 모델과 계량경제 모형의 장점을 결합한 것이라고 볼 수 있다.

6.3 한계점 및 향후 연구 방안

본 연구는 청소년 사이버비행 방지를 위해 혼합방법론을 적용해 종합적 차원에서 청소년 사이버비행을 분석하였지만 다음과 같은 한계점을 지닌다. 먼저, 본 연구는 KCYPS 2010에서 5년간 지속적으로 응답을 한 청소년들 위주로 패널 데이터를 구성하여 선택 편향(selection

bias)이 있을 수 있고, 청소년의 내적 성향이나 개인적 환경에 대한 응답 항목들은 시간 갭이 존재하는 경우가 많아 통제변수로 다양하게 반영하지 못하였다는 한계를 지닌다. 또한, 본 연구에서는 종속변수를 다섯 가지 사이버비행에 대한 연간 경험 여부로 측정하고 있어 사이버비행 간 성격이나 심각성에 대해 구분하지 못한다는 한계가 있다. 따라서 향후 연구에서 다양한 청소년 사이버비행을 특성에 따라 구분하면 더욱 심층적으로 청소년 사이버비행에 대한 원인 요인을 파악할 수 있는 연구가 될 것이다. 본 연구에서 사용한 KCYPS 2010 데이터에는 일상 비행에 대해서는 조사가 되었으나 사이버비행 피해 경험에 대한 항목이 조사되지 않아 사이버비행 피해에 대한 변수를 고려하지 못하였다. 일상 비행 피해 경험이 사이버비행을 야기한다는 점을 고려하면 사이버비행에 대한 피해도 청소년에게 일상 긴장의 요인으로 작용하여 사이버비행 가담에 유의한 영향을 미칠 수 있으므로 향후 연구에서는 사이버비행 피해에 대한 변수를 반영하여 사이버비행의 원인 요인을 파악하는 연구를 진행할 필요가 있다. 본 연구에서는 청소년 비행을 설명하는 이론 중 일반이론(general theory)을 고려하지 못하였다. Gottfredson and Hirschi(1990)의 일반이론은 대부분의 비행이나 범죄가 충동적으로 발생한다는 점에 주목하여 이러한 충동성이나 순간적인 만족감을 통제 및 조절할 수 있는 내적성향, 즉 자기통제력을 비행의 설명 요인으로 보았다(이성식, 2010). 본 연구에서 사용한 KCYPS 2010 데이터에는 청소년들의 자기통제력 항목이 조사되지 않아 일반이론을 제외한 네 가지 이론을 통합적으로 활용하여 연구를 진행하였

다. 따라서 본 연구의 네 가지 배경이론은 물론 일반이론에서 이야기하고 있는 자기통제력과 관련한 변수 및 문항을 추가한다면 더욱 종합적인 차원에서의 청소년 사이버비행 원인 요인 파악이 가능할 수 있고 더욱 성능이 우수한 예측 모델을 개발하는 연구가 진행될 수 있다.

6.4 시사점

본 연구의 학술적 시사점으로는 첫째, 청소년 사이버비행을 설명하는 네 가지 이론을 바탕으로 기존에 고려되지 않은 변수들을 추가하여 통합적인 분석을 진행하였다는 점을 들 수 있다. 기존의 연구들(Patchin & Hinduja, 2011; Navarro & Jasinski, 2012; Paez, 2018) 대부분이 단일 이론 및 관점에서 진행되었으며, 여러 배경이론을 종합적으로 다룬 연구는 많지 않다. 사회유대이론을 적용하여 청소년 사이버비행의 주원인을 파악하고자 했던 기존 연구들(배성만, 2016; Udris, 2016)은 주로 애착 개념의 변수들을 선택하여 관여나 신념 측면에서 고려되지 못한 변수가 존재한다. 각 이론 내에서 사용되지 않았던 개념적 변수를 추가하고, 네 가지 이론을 통합 적용하여 종합적인 관점에서 청소년 사이버비행에 대한 분석을 진행하였다는 점에서 학술적 기여가 있다. 특히, 본 연구는 비행 및 범죄 행동 관련 기존 이론을 딥러닝 기반의 사이버비행 예측모델에 적용한 연구로 기존 사이버비행 관련 AI 예측 모델들과 달리 모델 개발 측면에서 이론적 타당성이 있는 변수 생성 과정을 거쳤다. 이는 기존의 예측 모델 개발 연구에서 이론적 배경 없이 변수를 추출하는 것과 차별화되는 시사점이다.

둘째, 본 연구는 청소년 사이버비행 가해 요인을 찾아내는 것뿐 아니라 혼합방법론을 적용하여 딥러닝 및 데이터 사이언스 기법을 기반으로 각 청소년이 사이버비행에 가담하였을 확률을 계산하는 예측 모델을 개발하였다. 이는 청소년 연구, 사이버비행 연구 등의 분야에서 새로운 접근이며, 계량경제모형과 예측 모델링을 혼합하여 통합적으로 청소년 사이버비행 문제 해결에 대한 접근을 이루었다는 시사점이 존재한다. 기존 연구들이 모두 단순한 원인 파악에만 머물러 미미한 검거율이나 우선적인 교육 대상 청소년 선별 등 시급한 문제를 해결하는데 한계점이 있던 것과는 달리 새로운 방법론의 적용을 통해 더욱더 높은 시사점을 기대할 수 있다 (Navarro & Jasinski, 2012; Jang et al., 2014; Ang, 2015; Udris, 2016; Choi et al., 2017). 그뿐만 아니라 예측 모델 개발 이전에 원인 요인을 동일 데이터로 탐색함으로써 기존 예측 모델의 블랙박스 (black box) 문제를 보완해 줄 수 있다.

본 연구의 실무적 기여로는 첫째, 청소년 개인이 사이버비행을 가해하였는가에 대한 확률을 기반으로 예측 모델을 만듦으로써 사이버비행 관리를 위한 시스템 구축에 기여한다. 본 연구에서 개발한 청소년 사이버비행 예측 모델은 입력받은 데이터를 기반으로 자동적으로 각 청소년이 사이버비행에 가담하였을 확률을 추출해줄 수 있다. 게다가 계량경제모형 테스트와 예측 모델 개발을 연동함으로써 하나의 프로세스 생성에 기여를 할 수 있다. 이 프로세스는 예측 모델을 활용하여 사이버비행 고위험군 학생을 선정하고 계량경제모형 테스트로부터 검증된 주요 사이버비행 원인 요인을 바탕으로

사이버비행 가담 학생들에게 맞춤형 상담을 제공하는 것에 활용이 될 수 있으므로 높은 실무적 활용성을 지닌다.

둘째, 본 연구의 결과는 사이버비행 예방을 위한 교육을 목적으로 활용될 수 있다. 계량모형 테스트를 통해 밝혀낸 유의한 설명변수를 토대로 사이버비행의 원인이 되는 주요 행동 특징을 파악하여 청소년들을 대상으로 사이버비행 예방 교육을 실시할 수 있다. 예를 들면, 사이버비행 예방을 위한 교육 콘텐츠 제작에 본 연구결과가 실무적으로 활용이 될 수 있다. 특히 계량경제모형 분석 결과 커뮤니케이션 목적 외의 스마트폰 이용 빈도가 높을 수록 사이버비행 가해 확률이 높아졌으므로 정부적 차원에서 스마트폰 중독 예방을 위한 정책을 마련하여 청소년의 사이버비행을 낮출 수 있다. 또한 학교 차원에서도 규칙 및 공동체 생활에 대해 교육하고 체험하는 프로그램을 만들어 학생들로 하여금 사회적 유대감을 높여 청소년 사이버비행을 예방하는 것이 필요하다. 나아가 예측 모델을 활용하여 잠재적 사이버비행 가해 청소년을 파악하고 해당 교육의 대상으로 선별하는 것은 효과적인 예방프로그램을 운영하는 것에 기여할 수 있다.

마지막으로, 사이버비행이 이루어지는 주 매체인 컴퓨터와 휴대폰의 이용빈도를 세부적으로 나누어 분석함으로써 특히 어떤 목적으로 컴퓨터나 휴대폰을 이용할 때 사이버비행 경험과 관련이 높은지 파악하여 대비할 수 있다. 19세 이상 사이트와 같은 성인 매체로부터 청소년들을 보호하기 위해 유해 정보 필터링 및 모니터링을 강화할 수 있으며 (배상률 등, 2017; Navarro & Jasinski, 2012), 게임을 하거나 댓글

을 달면서 사이버공간 상의 상대를 비방하거나 폭력적인 언어를 사용하는 것을 방지하기 위한 교육이 필요하다. 이는 휴대폰으로 친구들과 문자메시지를 주고받을 때나 온라인상에서의 팬덤활동 즉, 팬카페 가입 혹은 포털 사이트 기사에 댓글쓰기와 같은 활동 시에 경험할 수 있는 사이버비행을 예방하기 위해서도 필요하다. 이를 위해 청소년들을 대상으로 인터넷 예절(네티켓) 교육과 홍보 활동이 이루어져 하며 (배상률 등, 2017), 시청각자료를 통한 형식적인 교육보다 학생들이 실제 경험하며 느낄 수 있는 실습 위주의 교육이 효과적일 수 있다 (곽대훈, 임숙희, 2017). 사이버비행으로 인한 피해를 입는 당사자가 된다고 가정하여 어떤 일이 일어나고 어떤 감정을 겪게 될지 생각해보는 교육 등이 한 예가 될 수 있을 것이다. 또한, 학습 이외의 정보를 검색하고 음악이나 영화 파일 등의 자료를 찾을 때 불법 다운로드를 할 가능성이 높아질 수 있으므로, 이를 방지하기 위한 사이버 윤리의식 함양 교육이 이루어져야 한다 (곽대훈, 임숙희, 2017).

참고문헌

경찰청, 「사이버범죄통계」, 2020, Retrieved from <https://www.police.go.kr/www/open/publice/publice0204.jsp>

곽대훈, 임숙희, “비행청소년의 사이버 일탈행동 실태 및 대응방안 연구,” 충남대학교 산학협력단, 2017.

고은희, 김은정, “우울, 주의집중의 어려움, 인터넷 사용빈도 및 교사/친구관계가 청

소년 사이버비행에 미치는 영향: 조절된 매개효과 검증,” 청소년학연구, 제22권, 제3호, 2015, pp. 1-26.

교육부, 2018년 1차 학교폭력 실태조사, 2018.

김민수, 김희용, “온라인 해킹 불법 시장 분석: 데이터 마이닝과 소셜 네트워크 분석 활용,” 정보시스템연구, 제29권, 제2호, 2020, pp. 221-242.

김수진, 강현아, “사회적 지지가 청소년의 사이버불링 경험에 미치는 영향: 자기통제력의 매개효과를 중심으로,” 청소년복지연구, 제18권, 제3호, 2016, pp. 171-93.

동아일보, “美청소년들도 ‘온라인 괴롭힘’ 피해 호소,” 2019.07.17., Retrieved from <http://www.donga.com/news/article/all/20190717/96511036/1>

방송통신위원회, 2018년 사이버폭력 실태조사 결과, 2019.

배성만, “긴장요인, 사회유대, 자기통제, 인터넷 사용 및 오프라인 비행이 청소년의 사이버 비행에 미치는 영향-위계적 선형모형을 통한 4년 종단자료 분석,” 상담학연구, 제17권, 제6호, 2016, pp. 381-394.

성경주, “청소년의 학교폭력 피해경험, 차별경험이사이버불링에 미치는 영향,” 한국융합과학회지, 2019, p. 200.

송광섭, “청소년비행 관련법규의 현재와 미래,” 형사정책, 제15권, 제2호, 2003, pp. 193-223.

이수봉, 한상철, “청소년의 감각추구성향과 또래동조성이 사이버불링에 미치는 영향,”

- 미래청소년학회지, 제15권, 제2호, 2018, pp. 79-105.
- 위휘, “청소년 스트레스 영향요인에 관한 연구,” 정신간호학회지, 제14권, 제2호, 2005, pp. 129-138.
- 이성식, “청소년 사이버비행의 동기, 기회와 통제요인을 통한 모형의 구성과 검증,” 정보화정책, 제14권, 제3호, 2007, pp. 3-15.
- 이성식, “낮은 자기통제력과 비행기회요인들의 오프라인 및 온라인 비행에의 영향 비교: 일반이론의 검증,” 형사정책연구, 제21권, 제1호, 2010, pp. 203-223.
- 이성식, 전신현, “청소년비행 설명에 있어 역할 정체성 관련 긴장과 통제: 수정된 두 이론의 검증,” 형사정책연구, 제24권, 제4호, 2013, pp. 233-255.
- 이창문, 문진영, 박주원, “매체환경과 청소년 사이버비행과의 관계에 대한 연구: 성인 매체몰입을 중심으로,” 디지털융복합연구, 제17권, 제4호, 2019, pp. 365-379.
- 조보근, 박경배, 하성호, “기계학습 알고리즘을 활용한 지역 별 아파트 실거래가격지수 예측모델 비교: LIME 해석력 검증,” 정보시스템연구, 제29권, 제3호, 2020, pp. 119-144.
- 조은미, 진영선, 임정아, 조아미, “Akers 사회학습이론의 주요 개념이 청소년 비행에 미치는 영향,” 청소년학연구, 제23권, 제5호, 2016, pp. 357-378.
- 한국청소년정책연구원, 한국아동 청소년패널 조사 2010 (KCYPs 2010) 유저 가이드, 2017.
- 황성현, 이강훈, “청소년비행의 원인에 관한 사회학습, 사회유대, 일반긴장이론적 접근,” 한국청소년연구, 제24권, 제3호, 2013, pp. 127-145.
- 황성현, 이강훈, “사이버비행에 영향을 미치는 청소년들의 심리요인 분석,” 사회과학연구, 제24권, 제1호, 2017, pp. 157-175.
- Agnew, R., “Foundation for a General Strain Theory of Crime and Delinquency,” *Criminology*, Vol. 30, No. 1, 1992, pp. 47-88.
- Akers, R. L., *Social Learning and Social Structure: A General Theory of Crime and Deviance*. Transaction Publishers, 2011.
- Ang, R. P., “Adolescent Cyberbullying: A Review of Characteristics, Prevention and Intervention Strategies,” *Aggression and Violent Behavior*, Vol. 25, No. 1, 2015, pp. 35-42.
- Antwarg, L., Miller, R. M., Shapira, B., and Rokach, L., “Explaining Anomalies Detected by Autoencoders Using SHAP,” arXiv: 1903.02407, 2019.
- Aricak, T., Siyahhan S., Uzunhasanoglu, A., Saribeyoglu, S., Ciplak, S., Yilmaz, N., and Memmedov, C., “Cyberbullying among Turkish Adolescents,” *Cyberpsychology & Behavior*, Vol. 11, No. 3, 2008, pp. 253-261.
- Baldry, A. C., Sorrentino, A., and Farrington,

- D. P., "Cyberbullying and Cybervictimization versus Parental Supervision, Monitoring and Control of Adolescents' Online Activities," *Children and Youth Services Review*, Vol. 96, No. 3, 2019, pp. 302-307.
- Choi, K. S., Lee, S. S., and Lee, J. R., "Mobile Phone Technology and Online Sexual Harassment among Juveniles in South Korea: Effects of Self-control and Social Learning," *International Journal of Cyber Criminology*, Vol. 11, No. 1, 2017, pp. 110-127.
- Cohen, L. E., and Felson, M., "Social Change and Crime Rate Trends: A Routine Activity Approach," *American Sociological Review*, Vol. 44, No. 4, 1979, pp. 588-608.
- De Bruyn, A., Viswanathan, B., Beh Y. S., Brock, J. K. U., and von Wangenheim, F., "Artificial Intelligence and Marketing: Pitfalls and Opportunities," *Journal of Interactive Marketing*, Vol. 51, 2020, pp. 91-105.
- Dehue, F., Bolman, C., and Völlink, T., "Cyberbullying: Youngsters' Experiences and Parental Perception," *CyberPsychology & Behavior*, Vol. 11, No. 2, 2008, pp. 217-223.
- Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J. P., Frith, C. D., and Frackowiak, R. S., "Statistical Parametric Maps in Functional Imaging: a General Linear Approach," *Human Brain Mapping*, Vol. 2, No. 4, 1994, pp. 189-210.
- Hirschi, T., *Causes of Delinquency*, Berkeley, CA: Free Press, 1969.
- Gottfredson, M. R., and Hirschi, T., *A General Theory of Crime*, Stanford University Press, 1990.
- Jang, H., Song, J., and Kim, R., "Does the Offline Bully-Victimization Influence Cyberbullying Behavior among Youths? Application of General Strain Theory," *Computers in Human Behavior*, Vol. 31, 2014, pp. 85-93.
- Kim, H. Y., and Won, C. H., "Forecasting the Volatility of Stock Price Index: A Hybrid Model Integrating LSTM with Multiple GARCH-Type Models," *Expert Systems with Applications*, Vol. 103, 2018, pp. 25-37.
- Lee, B. H., "Explaining Cyber Deviance among School-Aged Youth," *Child Indicators Research*, Vol. 11, No. 2, 2018, pp. 563-584.
- Lee, J. M., Hong, J. S., Yoon, J., Peguero, A. A., and Seok, H. J., "Correlates of Adolescent Cyberbullying in South Korea in Multiple Contexts: A Review of the Literature and Implications for Research and School Practice," *Deviant Behavior*, Vol. 39, No. 3, 2018, pp. 293-308.
- Lee, V. H., Foo, A. T. L., Leong, L. Y., and

- Ooi, K. B., “Can Competitive Advantage be Achieved through Knowledge Management? A Case Study on SMEs,” *Expert Systems with Applications*, Vol. 65, 2016, pp. 136-151.
- Lundberg, S. M., and Lee, S. I., “A Unified Approach to Interpreting Model Predictions,” *Proceedings of the 31st International Conference on Neural Information Processing Systems*, 2017, pp. 4768-4777.
- Mokhtari, K. E., Higdon, B. P., and Başar, A., “Interpreting Financial Time Series with SHAP Values,” *Proceedings of the 29th Annual International Conference on Computer Science and Software Engineering*, 2019, pp. 166-172.
- Navarro, J. N., and Jasinski, J. L., “Going Cyber: Using Routine Activities Theory to Predict Cyberbullying Experiences,” *Sociological Spectrum*, Vol. 32, No. 1, 2012, pp. 81-94.
- Paez, G. R., “Cyberbullying among Adolescents: A General Strain Theory Perspective,” *Journal of School Violence*, Vol. 17, No. 1, 2018, pp. 74-85.
- Patchin, J. W., and Hinduja, S., “Traditional and Nontraditional Bullying among Youth: A Test of General Strain Theory,” *Youth & Society*, Vol. 43, No. 2, 2011, pp. 727-751.
- Ribeiro, M. T., Singh, S., and Guestrin, C., “‘Why should I Trust You?’ Explaining the Predictions of Any Classifier,” *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016, pp. 1135-1144.
- Roth, A. E., *Introduction to the Shapley Value*, Cambridge: Cambridge University Press, 1988.
- Udris, R., “Cyber Deviance among Adolescents and the Role of Family, School, and Neighborhood: A Cross-National Study,” *International Journal of Cyber Criminology*, Vol. 10, No. 2, 2016, pp. 127-146.
- Wong, L. W., Leong, L. Y., Hew, J. J., Tan, G. W. H., and Ooi, K. B., “Time to Seize the Digital Evolution: Adoption of Blockchain in Operations and Supply Chain Management among Malaysian SMEs,” *International Journal of Information Management*, Vol. 52, 2020, p. 101997.

손 새 아 (Shon, Sae Ah)



연세대학교 정보대학원에
서 석사학위를 취득하였으며,
주요 관심분야는 Business
Analytics, Machine Learning,
Digital Marketing 등이다.

신 우 식 (Shin, Woo Sik)



연세대학교 정보대학원에
서 비즈니스 빅데이터 분석 트
랙 박사과정이다. 현재 디지털
서비스 연구실에서 빅데이터
분석 관련 연구를 진행 중이
다. 주요 연구분야는 Digital
Marketing, Cyberbullying,
Business Analytics 등이다.

김 희 웅 (Kim, Hee Woong)



National University of
Singapore 정보시스템학과에
서 근무한 후, 현재 연세대학
교 정보대학원 교수로 재직 중
이다. 주요 연구분야는 디지털
비즈니스, 정보시스템 관리 및
활용 등이다. 관련 연구들은
MIS Quarterly, Information
Systems Research, Journal of
Management Information
Systems 등에 60여 편의 논문
이 게재되었다. JAIS, IEEE
TEM의 편집위원으로 활동했
고, KrAIS 회장을 역임했다.

Appendix A. 변수 및 측정항목

대영역	중/소영역	변수명	변수 유형
배경변인	일반 신상	성별 (GENDER)	범주형 (1: 남자, 0: 여자)
개인발달	신체발달 (건강)	체육시간 중 운동시간 (PHY3C)	5 4시간 이상 4 3시간 3 2시간 2 1시간 1 없다
	사회정서발달 (우울)	우울 (PSY1E)	4 매우 그렇다 3 그런 편이다 2 그렇지 않은 편이다 1 전혀 그렇지 않다
	비행 (비행 피해)	연간 행동경험 유무 (DLQ1A) 담배 피우기 술 마시기 무단결석 가솔 다른 사람 심하게 놀리거나 조롱하기 다른 사람 집단따돌림(왕따)시키기 패싸움 다른 사람 심하게 때리기 다른 사람 협박하기 다른 사람 돈이나 물건 뺏기(뺏기기) 다른 사람 돈이나 물건 훔치기 성관계 성폭행이나 성희롱 (위 변인 중 하나라도 해당되면 피해경험 있는 것으로 코딩)	범주형 (1: 있다, 0: 없다)
	연간 피해경험 유무 (DLQ2A) 심한 놀림이나 조롱당하기 집단따돌림(왕따)당하기 심하게 맞기(폭행) 협박당하기 돈이나 물건 뺏기(뺏기기) 성폭행이나 성희롱 심한 욕설 (위 변인 중 하나라도 해당되면 행동경험 있는 것으로 코딩)	범주형 (1: 있다, 0: 없다)	
	연간 행동경험 친구 수 (DLQ3A) 담배 피우기 (#명) 술 마시기 (#명) 무단결석 (#명) 가솔 (#명) 다른 사람 심하게 놀리거나 조롱하기 (#명) 다른 사람 집단따돌림(왕따)시키기 (#명) 패싸움 (#명) 다른 사람 심하게 때리기 (#명) 다른 사람 협박하기 (#명) 다른 사람 돈이나 물건 뺏기(뺏기기) (#명) 다른 사람 돈이나 물건 훔치기 (#명) 성관계 (#명) 성폭행이나 성희롱 (#명) (위 변인 모두 합하여 코딩)	연속형 (#명)	

발달환경	교육환경 (학교적응)	학습활동 (EDU2A) 학교규칙 (EDU2B) 교우관계 (EDU2C) 교사관계 (EDU2D)	4 매우 그렇다 3 그런 편이다 2 그렇지 않은 편이다 1 전혀 그렇지 않다								
	지역사회환경 (지역사회관)	지역사회 인식 (COM1A) 공동체 의식 (COM2A)									
	매체환경 (컴퓨터)	용도별 이용빈도 (MDA1D01~03,47,06,09,10) 1) 공부 및 학습 관련 정보 검색 및 사이트 이용 2) 학습 이외의 정보 검색 및 자료(음악·영화 파일 등) 이용 3) 게임 및 오락 4) 채팅이나 메신저 및 블로그, SNS 활동 5) 동호회, 카페, 커뮤니티 활동 6) 댓글달기 7) 19세 이상 사이트 이용	4 자주 사용한다 3 가끔 사용한다 2 거의 사용하지 않는다 1 전혀 사용하지 않는다 0 컴퓨터 없음								
	매체환경 (휴대전화)	용도별 이용빈도 (MDA2B01~05, MDA2B07) 1) 가족과의 통화 2) 가족과의 문자 메시지 (카카오톡, 라인 등 포함) 3) 친구와의 통화 4) 친구와의 문자 메시지 (카카오톡, 라인 등 포함) 5) 게임 및 오락	4 자주 사용한다 3 가끔 사용한다 2 거의 사용하지 않는다 1 전혀 사용하지 않는다 0 휴대폰 없음								
	활동/문화환경	체험활동 연간 참여경험 유무 (ACT1A) <table border="1" style="width: 100%; border-collapse: collapse;"> <tr><td>건강/보건 활동</td></tr> <tr><td>과학/정보 활동</td></tr> <tr><td>교류 활동</td></tr> <tr><td>모험/개척 활동</td></tr> <tr><td>문화/예술 활동</td></tr> <tr><td>봉사 활동</td></tr> <tr><td>직업체험 활동</td></tr> <tr><td>환경보존 활동</td></tr> <tr><td>자기(인성) 개발 활동</td></tr> </table> (아래 활동 중 하나라도 경험이 있으면 1, 아니면 0으로 코딩) 교내외 동아리활동 연간 참여경험 유무 (ACT2A) (교내외 하나라도 경험이 있으면 1, 아니면 0으로 코딩) 문화활동 연간 경험횟수 (ACT3B)	건강/보건 활동	과학/정보 활동	교류 활동	모험/개척 활동	문화/예술 활동	봉사 활동	직업체험 활동	환경보존 활동	자기(인성) 개발 활동
건강/보건 활동											
과학/정보 활동											
교류 활동											
모험/개척 활동											
문화/예술 활동											
봉사 활동											
직업체험 활동											
환경보존 활동											
자기(인성) 개발 활동											
	온라인 팬덤활동 빈도 (ACT6B) (아래 활동에 대한 응답 평균) 1) 팬클럽이나 팬카페 가입 2) 홈페이지 방문 3) 이메일 또는 팬레터 쓰기 4) 포털 사이트나 기사에 댓글 달기 5) 인터넷 사이트에 동영상 올리기	연속형 (#회) 4 자주 했다 3 가끔 했다 2 거의 하지 않았다 1 전혀 하지 않았다 0 선호 연예인/운동선수 없음									

<Abstract>

Juvenile Cyber Deviance Factors and Predictive Model Development Using a Mixed Method Approach

Shon, Sae Ah · Shin, Woo Sik · Kim, Hee Woong

Purpose

Cyber deviance of adolescents has become a serious social problem. With a widespread use of smartphones, incidents of cyber deviance have increased in Korea and both quantitative and qualitative damages such as suicide and depression are increasing. Research has been conducted to understand diverse factors that explain adolescents' delinquency in cyber space. However, most previous studies have focused on a single theory or perspective. Therefore, this study aims to comprehensively analyze motivations of juvenile cyber deviance and to develop a predictive model for delinquent adolescents by integrating four different theories on cyber deviance.

Design/methodology/approach

By using data from Korean Children & Youth Panel Survey 2010, this study extracts 27 potential factors for cyber deviance based on four background theories including general strain, social learning, social bonding, and routine activity theories. Then this study employs econometric analysis to empirically assess the impact of potential factors and utilizes a machine learning approach to predict the likelihood of cyber deviance by adolescents.

Findings

This study found that general strain factors as well as social learning factors have positive effects on cyber deviance. Routine activity-related factors such as real-life delinquent behaviors and online activities also positively influence the likelihood of cyber deviance. On the other hand, social bonding factors such as community commitment and attachment to community lessen the likelihood of cyber deviance while social factors related to school activities are found to have positive impacts on cyber deviance. This study also found a predictive model using a deep learning algorithm indicates the highest prediction performance. This study contributes to the prevention of cyber

deviance of teenagers in practice by understanding motivations for adolescents' delinquency and predicting potential cyber deviants.

Keyword: Juvenile cyber deviance, Econometric analysis, Deep learning

* 이 논문은 2021년 4월 12일 접수, 2021년 5월 11일 1차 심사, 2021년 5월 27일 2차 심사, 2021년 5월 27일 게재 확정되었습니다.