

A Predictive Model for Digital Sexual Crime of Adolescents Using Random Forests

Choonbum Cho* & Hyun Lee**

Abstract: This study is to examine the predictive model for digital sexual crime (distribution and consumption of sexually explicit material) of adolescents using random forests. A total of 802 youths in Gyeonggi province were used in the empirical analysis. The distribution and consumption of sexually explicit material were classified into sale, purchase, and upload behaviors. Each of the behavior-models was established and estimated. As for the importance of predictors, status delinquency such as runaway, online gambling, drinking, and smoking was high in the sale and purchase models. On the other hand, sexual delinquency such as prostitution and watching adult broadcasting showed high importance in the upload model. Based on these results, an individualized approach is required for each of digital sexual crimes, and prevention and intervention are necessary to target on antecedent problems.

Key Words: Random Forests, Adolescent, Digital Sexual Crime, Distribution, Consumption

랜덤 포레스트 기법을 활용한 청소년의 디지털 성범죄 관련 예측요인 탐색 조 춘 범 · 이 현

요약: 본 연구는 랜덤 포레스트 기법을 활용하여 청소년의 디지털 성범죄 중 유포 및 소비형 가해행위와 관련된 예측요인을 탐색하는 것을 목적으로 한다. 실증분석을 위해 수도권 서·남부 지역의 청소년을 대상으로 진행된 총 802명의 설문조사 결과를 활용하여 분석하였다. 유포 및 소비형 가해행위를 판매, 구입, 업로드로 구분하여 각 행위를 종속변수로 하는 모형을 설정하고 추정하였다. 설명변수의 중요도에서는 디지털 성범죄 콘텐츠 판매와 구입 행위에서 청소년들의 지위 비행과 관련된 가출, 온라인 도박, 음주, 흡연이 높게 나타났으며, 업로드 행위에서는 성매매 경험, 성인방송 시청과 같은 성비행 범주의 변수들이 높은 중요도를 보였다. 이러한 결과를 바탕으로 볼 때, 세부 디지털 성범죄 가해행위마다 개별화된 접근이 필요하며 지위 비행이나 성비행 등의 선행문제를 중심으로 표적화된 예방과 개입이 요구된다.

핵심어: 랜덤포레스트, 청소년, 디지털 성범죄, 유포, 소비

□ 접수일: 2020년 12월 9일, 수정일: 2020년 12월 26일, 게재확정일: 2020년 12월 28일

* 주저자, 성결대학교 사회복지학과 교수

(First Author, Professor, Sungkyul Univ., Email: flytiger7@hanmail.net)

** 교신저자, 연세대학교 사회복지대학원 BK 사업단 박사후 연구원

(Corresponding Author, PhD, Yonsei Univ., Email: hyunlee1237@gmail.com)

I. 서론

최근 디지털 성범죄가 연일 미디어를 장식하면서 우리나라의 어두운 단면이 여실히 드러났다. 그중에서도 충격적인 사실은 디지털 성범죄 사건으로 알려진 ‘N번방’의 주동자와 소비자로 가담한 가해자의 64%가 미성년자였다는 점이다(장아름, 2020). 표면으로 드러난 소수의 사례로 암수범죄를 추정하기는 어렵지만, 정보통신기술에 친숙한 청소년들이 디지털 성범죄에 용이한 점은 이전부터 지적됐다. 디지털 성범죄의 용어가 공식적으로 사용된 시점은 최근 발생한 텔레그램 n번방 사건으로 지능화, 조직화한 진화 형태로 밝혀진 것이 계기이지만 새롭게 나타난 범죄의 유형은 아니다(윤덕경, 2020).

디지털 성범죄는 현재 공식적인 법적 용어는 아니며 다양한 관련 법률로부터 규정된 범죄의 유형으로 이해할 수 있다. 디지털 성범죄는 이전의 불법촬영, 비동의 유포 및 재유포, 유통과 소비, 유포 협박, 성적 괴롭힘의 범죄행위 유형을 포함한다.¹⁾ 불법 촬영이나 성적 괴롭힘의 경우 직접 범죄 콘텐츠를 생산하고 피해자에게 해를 가한다는 공감대가 형성돼 있으나, 성적 촬영물이 범죄 영상이며 소비와 유포만으로도 성폭력이 될 수 있다는 인식은 부족하다(서승희, 2017).

이러한 맥락에서 디지털 성범죄를 ‘동의하지 않는 이에게 성적 폭력을 가하거나 신체 등을 촬영’하는 1차 가해행위와 1차 행위로부터 생성된 ‘불법 콘텐츠를 피해자의 동의 없이 유포하거나 소비함’으로써 발생하는 2차 가해행위로 구분할 수 있다. 1차 가해행위는 자신의 행위가 불법적인 기준에 저촉한다는 명백한 인지가 있다. 하지만 온라인에서 콘텐츠의 소비는 비대면으로 벌어지고 즉각적으로 포착되기도 힘들어 개인 내면의 통제 시스템이 허술해진다. 온라인의 익명 기반 관계 속에서 불법 성적 촬영물의 비평과 변형의 사회화된 성폭력으로 진화한 것이다(김소라, 2018). 피해 촬영물을 누가 언제든지 복제와 재생산을 통해 피해를 확산시킬 수 있다는 디지털 성폭력의 특징(김숙희 외, 2018)으로 인해 단순히 1차 가해행위의 근절뿐만 아니라 2차 가해행위인 유포 및 소비형 가해행위에 주목할 필요가 있다. 따라서 본 연구에서는 2차 디지털 성범죄 가해행위(유포 및 소비형 가해)에 집중하여 드러나지 않는 곳에서 불법 촬영물을 소비하고 재유포하여 직·간접적으로 범죄에 가담하고 있는 가해자의 구체적인 판매, 구입, 업로드 행위에 초점을 맞추고자 한다.

현재까지 정부는 디지털 성범죄 피해방지 종합대책 수립(2017. 9. 26.)을 시작으로 「성폭력

1) 정부는 디지털 성범죄의 정의와 이와 관련된 현재까지의 주요 대책, 관련 사업 등의 정보를 『대한민국정책브리핑』에서 제공하고 있다(2020. 11. 15. 검색).

방지 및 피해자보호 등에 관한 법률 일부개정법률안」의 국회 통과(2020. 1. 9.)까지 디지털 성범죄 피해자를 보호·지원하는 목적의 정책을 우선순위로 진행해왔다. 최근에 와서야 디지털 성범죄 근절대책(2020. 4. 23.)을 통해 가해자 처벌에 실효성을 확보하고 엄중성을 강화한다는 방향을 내세우고 있다. 하지만 현재 디지털 성범죄의 법적 체계 형성과 가해자의 처벌에 초점이 집중된 사이 어두운 곳에서부터 양산되고 있는 디지털 성범죄 문화에 대한 대처는 늦어질 수 있다(김경희·김수아·김은경, 2020). 사건 발생 이후의 처벌보다 다수를 차지한 청소년 가해자가 나타나지 않도록 예방의 측면을 강화하는 것이 디지털 성범죄 대처에 효과적일 수 있다.

예방적 대책 수립을 위해서는 가해자를 대상으로 특성을 파악하고 가해행위를 예측하는 실증적 연구결과의 축적이 뒷받침되어야 한다. 그러나 현재 디지털 성범죄 관련 선행연구들은 디지털 성범죄를 기존의 사이버 성범죄 유형을 참고하여 실태를 파악하고(김한균, 2017) 차별된 새로운 범죄로 개념화하거나(김소라, 2018), 현재까지 추진해온 디지털 성범죄 피해자 지원책(권미경, 2018; 서승희, 2017)과 가해자 처벌 정책(김한균, 2020; 김현아, 2017)을 분석하여 문제점을 파악하고 개선방안을 제시하는 내용이 주를 이뤘다. 디지털 성범죄자의 특성을 살펴본 윤정숙(2020)의 연구가 있으나 아동·여성 성착취 영상물 범죄자나 아동 그루밍 성범죄자로 연구대상을 제한하였고 문헌고찰 중심으로 탐색한 방법론적 제한이 있다. 이나현·조운오(2020)는 법무부 성범죄자 신상정보등록 시스템에 등록된 성범죄자들을 대상으로 전통적 성범죄(강간, 추행) 집단과 디지털 성범죄 집단을 실증적으로 비교하였으나 인구 사회학적 요인의 차이를 발견했을 뿐 디지털 성범죄 집단의 가해행위에 대한 구체적인 영향요인은 탐색하지 못하였다. 따라서 디지털 성범죄 가해자의 가해행위에 영향을 주는 요인을 파악하여 실질적인 예방 대책 수립에 도움이 되고자 가해자의 상당한 비중을 차지하고 있는 청소년을 대상으로 하여 실증적 연구를 진행하고자 한다.

이에 본 연구에서는 청소년의 디지털 성범죄에 영향을 미치는 다양한 요인들을 보다 효과적으로 탐색하기 위하여 대표적인 머신러닝(machine learning) 기법 중 하나인 랜덤 포레스트 알고리즘을 사용하여 예측하고자 한다. 본 연구의 구체적인 연구문제는 다음과 같다.

- 연구문제 1. 청소년의 유포형 디지털 성범죄 행위인 판매행동을 예측하는 주요 변수는 무엇이며 상대적 중요도는 어떠한가?
- 연구문제 2. 청소년의 소비형 디지털 성범죄 행위인 구입행동을 예측하는 주요 변수는 무엇이며 상대적 중요도는 어떠한가?
- 연구문제 3. 청소년의 디지털 성범죄 콘텐츠 재생산 행위인 업로드행동을 예측하는 주요 변수는 무엇이며 상대적 중요도는 어떠한가?

II. 연구방법

1. 자료 및 연구대상

본 연구의 실증분석을 위해 수도권 서·남부 지역의 청소년을 대상으로 진행된 설문조사 결과를 활용하였다. 조사는 비확률표집인 할당표집을 활용하여 6개의 시(고양시, 광명시, 군포시, 부천시, 안양시, 의왕시)를 선정한 뒤 지역별 중학교와 고등학교에 협조를 요청하여 설문이 가능한 12개 학교의 학생들을 대상으로 설문을 시행하였다. 조사는 2014년 6월부터 동년 7월에 진행되어 총 1,000부의 설문이 수거되었으나 응답이 불확실하거나 불성실한 설문지를 제외하고 실제 분석에는 총 802명의 유효응답이 활용되었다.

2. 측정도구

본 연구에서는 청소년의 디지털 성범죄 가해 유형을 측정하기 위해 심진숙(2008)이 사용한 사이버 비행 측정도구를 수정·보완하여 활용하였다. 연구의 취지에 맞게 연구자는 불법 음란물 콘텐츠를 판매, 구입, 업로드한 경험이 있는지 묻고 있는 3개의 문항을 활용하였다. 각 변수는 행위의 경험을 여부로 구분한 이항변수로 설정하였다. 전체 디지털 성범죄- 유포 및 소비형 가해는 3가지 유형의 한번이라도 경험한 적이 있는 경우를 기준으로 유무를 구분하였다.

본 연구는 종속변수를 예측하고자 청소년의 주요 체계를 중심으로 다양한 위험요인과 보호요인을 설명변수로 설정하였다. 랜덤 포레스트 분석은 설명변수가 다량으로 투입되더라도 자유도 감소가 발생하지 않는 장점이 있어 참여자가 응답한 다양한 영역의 정보들을 변수로 활용할 수 있다. <표 1>은 분석에 활용된 최종 5개의 범주의 19개의 설명변수를 보여준다.

<표 1> 설명변수 및 범주

범주	설명변수 ²⁾
인구통계학적 특성(4개)	성별, 연령, 성적, 생활수준
개인 수준(2개)	우울(4점), 자기통제력(5점)
주요 체계 수준(4개)	부모자녀관계(5점), 또래관계(4점), 멘토 유무, 사회활동 참여시간(6점)
지위 비행(5개)	비행친구 유무, 가출(5점), 음주(5점), 흡연(5점), 온라인도박(5점)
성 비행(4개)	성경험(5점), 성매매(5점), 포르노 사이트 이용(5점), 성인방송시청(5점)

3. 자료분석방법

청소년의 디지털 성범죄(유포 및 소비형 가해)에 영향을 미치는 주요 변수를 예측하고자 랜덤 포레스트(random forest) 기법을 사용하였다. 기존에 많이 활용된 데이터 마이닝 기법인 의사결정나무(decision tree)은 하나의 표본에서 얻어진 결과가 다른 자료에서는 낮은 예측력을 보여 모형의 안정성과 결과의 신뢰도 면에서 단점을 보인다(김영식·이은정·주현준, 2019). 하지만 랜덤 포레스트 기법은 자료 간 성능평가와 변수 선택에 있어 무작위성을 높이고 대수의 법칙에 기반한 반복을 통해 도출된 의사결정나무 결과를 종합(ensemble)함으로써 높은 예측력과 모형 안정성을 확보(유진은, 2015)하는 이점이 있다.

본 연구에서는 예측모형 평가를 위해 일반적으로 다수의 연구에서 적용하는 7:3의 비율로 전체 표본 데이터를 훈련(training)데이터와 시험(test data)데이터로 무작위 분류하였다(유진은, 2015; 최필선·민인식, 2018; 홍기혜, 2020). 여기서 훈련데이터는 예측모형 추정, 그리고 시험데이터는 예측모형의 일반화 검증에 활용된다.

실제 분석에는 R(ver. 4. 0. 2)의 randomForest package(Breiman, 2015)를 사용하였다. 설명변수의 결측치는 rfImpute 함수를 통해 얻은 유사값으로 대체하였고 범주형인 본 연구의 종속변수를 반영하여 분류(classification) 방식으로 500회의 부트스트랩(bootstrap) 표본을 사용하여 분석을 수행하였다. 청소년의 디지털 성범죄 행위 예측에 중요한 설명변수를 파악하고자 각 모형별로 중요도 지수(importance index)를 확인하였다. 모형별로 중요도 지수가 높은 상위 5개 변수들은 반응변수에 미치는 한계효과를 부분 의존성 도표(PDP: partial dependence plot)로 시각적으로 나타내었다.

III. 연구결과

1. 예측 모형 성과분석

<표 2>는 청소년들의 디지털 성범죄 유형 별 모형에 대한 예측 성과이다.

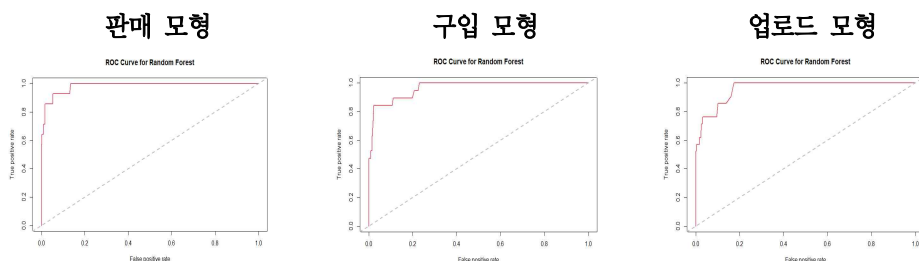
- 2) 우울(9문항, 4점, 평균화)과 부모자녀관계(10문항, 5점, 평균화)는 한국아동·청소년 패널조사의 척도를 수정·보완하여 사용하였다. 자기통제력(20문항, 5점, 평균화)는 류진아(2003)의 척도, 또래관계(27문항, 4점, 평균화)는 황옥경(1997)의 척도를 활용하였다.

〈표 2〉 디지털 성범죄 유형 별 랜덤포레스트 예측 성과³⁾

	판매 모형(%)	구입 모형(%)	업로드 모형(%)
Accuracy (정분류율)	99.13	98.4	98.26
Sensitivity (민감도)	100	100	100
Specifcity (특이도)	57.14	42.11	42.86

실제 범주를 올바르게 추정한 전체 비율인 정분류율은 판매 모형이 99.13%, 구입 모형이 98.4%, 업로드 모형이 98.26%로 높게 나타났다. 민감도에서도 세 모형에서 모두 100%로 나타나 실제 디지털 성범죄 경험이 있는 청소년을 양성으로 예측하는 성능이 우수하게 여겨진다. 상대적으로 특이도에서는 낮은 비율을 보여 실제 디지털 성범죄 경험이 없는 청소년을 음성으로 예측하는 성능 면에서는 미흡하다고 볼 수 있다.

다음으로 청소년들의 디지털 성범죄 유형 별 모형의 ROC 곡선을 〈그림 1〉과 같이 제시하였다. ⁴⁾ 세 모형의 ROC 곡선이 모두 북서쪽의 모서리에 가까워지면서 면적이 1에 가까워지는 이상적인 모습으로 우수한 예측성가를 보여준다.

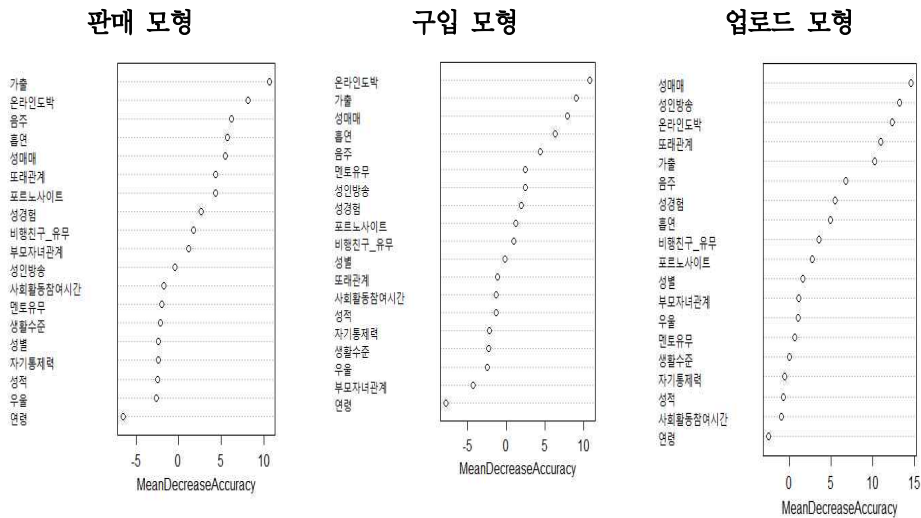


〈그림 1〉 디지털 성범죄 유형별 ROC 곡선

- 3) 본 연구와 같이 범주를 예측하는 모형의 평가지표로는 정오분류표(confusion matrix)를 통해 얻어진 정분류율(accuracy), 민감도(sensitivity), 특이도(specificity)가 활용된다. 정분류율은 실제범주를 제대로 추정한 전체 비율로서 본 연구에서는 디지털 성범죄 경험이 있는 청소년과 디지털 성범죄 경험이 없는 청소년을 올바르게 예측한 비율을 의미한다. 민감도는 실제 디지털 성범죄 경험이 있는 청소년을 양성으로 판정한 비율이다. 특이도는 실제 디지털 성범죄 경험이 없는 청소년을 음성으로 판정한 비율이다.
- 4) 민감도와 특이도를 활용하여 ROC(receiver operating characteristic) 곡선을 표현하면 시각적으로 기법의 우수함을 확인할 수 있다. x축에 1-특이도, y축에 민감도를 위치하였을 때 곡선의 면적이 1에 가까울수록(북서쪽에 위치할수록) 높은 예측력을 의미한다(전지혁, 2012).

2. 설명변수의 중요도

머신러닝 기법은 통계적 유의성 검정방식이 아닌 예측에 기여한 상대적 중요도를 통해 설명변수를 평가한다(김영식·이은정·주현준, 2019). 랜덤 포레스트는 의사결정나무 분석을 반복하여 얻은 모형정확도 개선 지수(MDA: mean decrease accuracy)와 지니 개선 지수(MDG: mean decrease Gini)로 설명변수의 상대적 중요도를 파악할 수 있다. 본 연구에서는 오분류율(1-정분류율)을 기준으로 계산된 모형정확도 개선 지수(MDA: mean decrease accuracy)를 활용하였다.



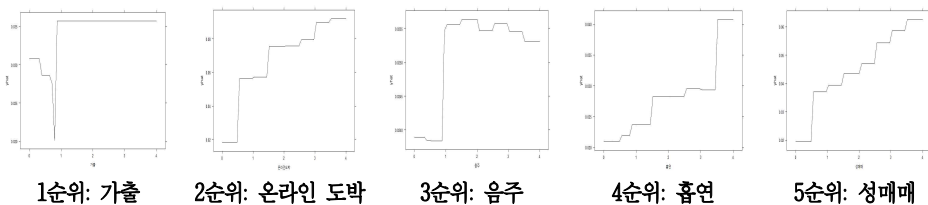
〈그림 2〉 디지털 성범죄 유형별 설명변수의 중요도 지수

〈그림 2〉는 디지털 성범죄 행위별 모형에서 상대적 중요도가 높은 설명변수를 보여준다. 디지털 성범죄 콘텐츠 판매를 예측변수로 하는 판매모형과 디지털 성범죄 콘텐츠 구입을 예측변수로 하는 구입모형에서는 상대적 중요도가 높은 상위권의 설명변수들이 유사하게 나타났다. 가장 높은 중요도는 가출과 온라인 도박, 음주, 흡연 등의 지위비행 범주의 설명변수들이었다. 성매매 경험도 두 모형에서 상위의 중요도를 지닌 변수로 나타났다. 하지만 상대적으로 개인의 성별, 성적, 연령, 생활수준의 인구사회학적 특성이나 개인적 수준의 우울과 자기통제력의 변수들은 판매와 구입 행위 예측에 낮은 중요도를 보였다. 디지털 성범죄 콘텐츠 업로드를 예측변수로 하는 업로드 모형에서는 성 비행 범주의 변수들이 높은 중요도를

나타냈다. 성매매 경험, 성인방송 시청이 대표적인 변수로서 청소년의 성 인지에 영향을 줄 수 있는 요인들이다. 또한 또래관계와 비행친구 유무와 같은 설명요인을 통해 청소년의 디지털 성범죄 업로드를 통한 유포행위에 또래와의 관계적 요인이 중요하게 영향을 줄 수 있음을 확인하였다. 한편 판매 및 구입 모형에서 높은 중요도를 보인 온라인 도박과 가출은 업로드 모형에서도 높은 상위권에 위치하였다.

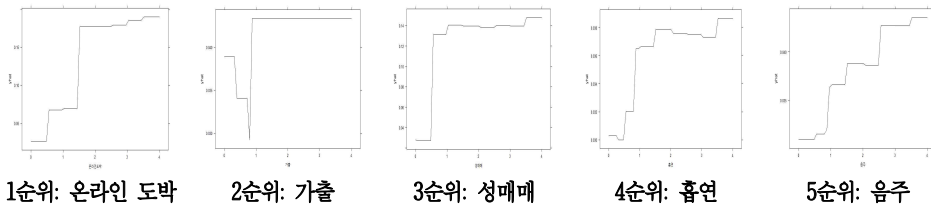
3. 부분 의존성 도표

부분 의존성(partial dependence)은 관심 설명변수가 종속변수에 미치는 한계효과(marginal effects)를 측정han 것으로 타 설명변수의 값이 일정할 때 관심 설명변수의 특정한 값에서 해당 사건이 발생할 확률이 높아진다고 해석할 수 있다. 다음은 디지털 성범죄 행위별 모형에서 상대적 중요도가 높은 상위 5개의 설명변수의 부분 의존성을 도표로 나타낸 결과이다.



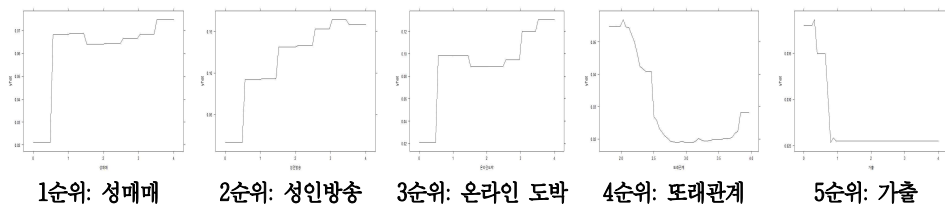
〈그림 3〉 부분 의존성 도표: 판매 모형

〈그림 3〉은 청소년의 디지털 성범죄 콘텐츠 판매에 영향을 미치는 변수의 부분 의존성을 순위대로 시각화하였다. 1순위의 가출은 가출 경험 1회 이상에서 반복횟수와 관계없이 판매 행위의 확률이 높게 유지되었다. 즉, 가출의 빈도보다는 경험 유무가 더 큰 의미를 지녔다. 2순위의 온라인 도박, 4순위 흡연은, 5순위의 성매매에서는 횟수가 늘어남에 따라 판매행위 확률이 지속해서 증가하는 것으로 보아 해당 경험이 축적되지 않도록 중단하는 것을 개입방안으로 생각해볼 수 있다. 3순위 음주에서는 경험 없음과 1회의 차이가 크게 나타났으나 반복횟수에서는 별다른 확률의 변화를 보이지 않아 음주의 횟수보다는 경험 여부가 판매 행위에 기여하는 것을 알 수 있었다.



〈그림 4〉 부분 의존성 도표: 구입 모형

〈그림 4〉는 청소년의 디지털 성범죄 콘텐츠 구입에 영향을 미치는 상위 5개의 변수에 대한 부분 의존성 도표의 결과이다. 1순위 온라인 도박, 2순위 가출과 3순위 성매매는 경험이 없는 경우보다 1회의 경험에서 구입행위 확률에 상승 변화를 보였지만 그 이후로는 유지되는 양상이 있었다. 즉, 각 비행행위의 반복횟수보다 경험 유무가 구입행위에 의미가 있었다. 대표적인 지위비행인 흡연과 음주는 횟수의 증가와 더불어 단계적인 확률의 증가폭이 눈에 띄었다.



〈그림 5〉 부분 의존성 도표: 업로드 모형

〈그림 5〉는 청소년의 디지털 성범죄 콘텐츠 업로드에 영향을 미치는 상위 5개의 변수에 대한 부분 의존성 도표를 나타낸다. 1순위 성매매는 경험 유무에 있어 업로드행위 확률에 커다란 차이를 보였다가 횟수에서는 안정적인 양상을 보였다. 다음의 순위를 차지한 성인방송과 온라인 도박은 횟수에 따라 단계적인 확률의 증가가 나타났다. 4순위의 포레관계에서는 타 포레들과 원만한 관계를 형성할 경우 업로드행위 발생확률이 크게 감소하는 것이 확인되었다.

IV. 논의 및 결론

본 연구는 데이터 마이닝 기법 가운데 높은 예측력과 모형 안정성을 보이는 랜덤 포레스트 분석을 적용하여 청소년의 디지털 성범죄- 유포 및 소비형 가해행위 예측요인을 탐색하였다. 이로부터 밝혀진 주요 결과와 함의는 다음과 같다.

첫째로 청소년들의 디지털 성범죄 유형별 모형에 대한 예측 성과가 우수하게 나타났다. 세부 지표로 살펴보면 실제 범주를 올바르게 추정한 전체 비율인 정분류율은 판매 모형이 99.13%, 구입 모형이 98.4%, 업로드 모형이 98.26%로 높게 나타났다. 일반적으로 민감도와 특이도가 음의 관계를 보이는 것(전치혁, 2012)과 암수(暗數)율이 높아 예측하기 어려운 범죄를 예측하는 선행연구(Alves et al., 2018; Nguyen et al., 2017)들의 성과 지표와 본 연구의 결과가 비견될 수준임을 감안할 때 본 예측모형의 성과는 신뢰할 만한 수준으로 여겨진다. 아직 디지털 성범죄 관련 실증분석을 실행한 선행연구가 매우 부족한 현실점에 우수한 예측력으로 해당 문제의 이해도를 높이고 정책 입안과 실천적 개입방안 마련의 근거로 기여할 수 있겠다.

둘째로 청소년의 디지털 성범죄 행위별 예측하는 설명변수들의 유사점과 차이점이 나타났다. 디지털 성범죄 콘텐츠 판매와 구입 행위에서는 청소년들의 지위 비행과 관련된 가출, 온라인 도박, 음주, 흡연이 높은 중요도를 보였으나 상대적으로 인구사회학적 특성(성별, 성적, 연령, 생활수준)과 개인 수준의 변수(우울과 자기통제력)의 낮은 예측력을 나타냈다. 즉, 청소년이 디지털 성범죄 콘텐츠를 판매하거나 구입하는 행위는 개인의 특성이나 내적 상태보다는 일련의 비행 행동과 관련지어 생각해볼 수 있다. 디지털 성범죄 콘텐츠 판매와 구입을 다양한 비행유형을 경험한 청소년의 비행행위 가운데 하나로 파악하거나 불법 콘텐츠를 사고팔아 경제적인 이익을 얻어 향락적인 ‘온라인 도박, 음주, 흡연’에 소비하는 과정과 연관 지어 예상해볼 수 있다(이현철·김경식·최성보, 2016). 업로드 행위에서는 성매매 경험, 성인방송 시청과 같은 성 비행 범주의 변수들이 높은 중요도를 보였다. 청소년 시기의 성 비행의 경험은 왜곡된 성 인지를 형성하는데 기여(김유정, 2010; 이수정, 2015; 최명일, 2009)하기 때문에 디지털 성범죄 업로드 행위에 더욱 관대하게 참여할 수 있었을 것이다. 또한 높은 중요도를 보인 또래관계와 비행친구 유무는 유포행위에 있어 또래와의 관계적 요인이 중요하게 작용할 수 있음을 보여준다(김정만, 2000; 이지선·이성식, 2019). 판매와 구입으로 물질적인 이득을 취하지는 않지만 오히려 또래 간의 불법 콘텐츠를 주고받으며 비공식적이고 역기능적인 성 교육이 광범위하게 이뤄질 수 있다(김경희·김수아·김은

경, 2020; 이효빈·최윤경, 2016; 정익중·이은주, 2010). 이러한 가해 행위별 설명요인 중요도의 공통점과 차이점은 청소년의 디지털 성범죄 가해행위에 접근하는 데 있어 개별화된 접근이 필요하며 지위 비행이나 성비행 등의 선행문제를 중심으로 표적화된 예방과 개입이 요구된다는 것으로 이해할 수 있다.

셋째로 부분 의존성 도표 분석을 통해 중요도가 높게 나타난 설명요인의 경험 유무가 중요한지 혹은 빈도(횟수)에 따라 디지털 성범죄 행위에 변화가 있는지 파악하였다. 디지털 성범죄 콘텐츠 판매와 구입에 높은 중요도를 보인 가출의 경우, 가출의 횟수보다는 경험 유무가 가해행위에 중요하였다. 가출이 청소년 비행의 하나의 유형이기도 하나 다른 비행으로 연계되는 관문으로서의 위험성(백혜정·방은령, 2009)에 주목할 필요가 있다. 따라서 이혼, 사별 등으로 인한 구조의 변화를 겪거나 부모가 적절한 보호 및 양육기능을 담당하지 못할 수 있는 위험가정의 청소년들이 가정 밖으로 방출되지 않도록 학교나 지역사회의 안전망을 견고히 할 필요가 있다(이유신, 2012). 현재 여성가족부와 교육부가 추진하는 디지털 성범죄 예방교육이 교육현장을 중심으로 이뤄지고 있으나(조린, 2020), 청소년안전망 사업의 청소년 동반자·청소년지원단에게도 교육을 확대할 필요가 있다. 반면에 온라인 도박, 흡연, 음주 등의 설명요인들은 횟수에 따라 판매 및 구입 행위의 확률을 높이고 있어 지위비행 행위들이 만성화되기 이전인 초기단계에 개입하는 것이 유용할 수 있다. 즉, 호기심이나 스트레스 해소의 수단으로 시작한 경비행이 반복되지 않도록 그 위험성과 폐해를 알리고 지역사회의 유해환경을 근절하는 접근이 필요하다. 업로드 행위의 위험요인으로서 청소년의 성매매 경험 유무와 성인방송의 시청 횟수가 중요하게 나타났다. 이미 개인방송 운영과 SNS의 사용이 용이한 온라인 환경이 구축되어 있고 이를 활용하는데 익숙한 청소년에게 성인방송에 접근하는 것이 어렵지 않다. 2018년 청소년 매체이용 및 유해환경 실태조사(여성가족부, 2018)에서 나타난 ‘성인용 콘텐츠를 보지 않으려고 해도 일상생활에서 쉽게 노출된다’는 청소년의 높은 응답률은 유해환경의 조성을 방지하고 청소년의 유해매체 접촉 방식을 개선해야 함을 시사한다. 또한 가장 큰 중요도를 보인 성매매 경험은 횟수에 상관없이 업로드 행위에 영향을 주는 것으로 나타나 김은실·김귀정·김봉한(2011)의 연구결과와 일치하였다. 성매매는 성을 소비의 수단으로 판단하고 행하는 중대한 성비행이다. 선정적으로 피해자를 소비하는 것만으로도 성범죄의 가해자 혹은 공범일 수 있다는 적극적인 인식변화가 필요하다(김경희·이조은, 2020).

본 연구는 개념적 논의와 피해자 대응을 중심으로 진행된 디지털 성범죄 주제를 높은 예측력과 안정된 결과가 검증된 랜덤 포레스트 기법을 활용하여 실증적 연구로서 접근했다

는 점에서 의의를 지닌다. 또한 유포 및 소비형 디지털 성범죄 가해행위를 종속변수로 각 모델 간의 설명변수의 중요도를 비교함으로써 구체적인 시사점을 도출하고 정책수립과 개입 방안의 근거를 마련하였다. 하지만 몇 가지 연구의 한계점이 존재한다. 본 연구는 경기도 일부 지역의 청소년을 대상으로 진행되어 모집단을 대표하기에 미흡하며 조사 시점이 오래 되어 최근의 실태를 반영하지 못할 가능성이 있다. 반면 매스컴을 통해 불법 콘텐츠의 유포와 소비가 성범죄라는 인식이 성립되기 이전의 조사로 역으로 현재의 암수적 수치가 응답에 반영될 가능성도 존재한다. 그럼에도 여전히 분석 결과의 해석에 주의를 기울여야 하며 후속연구에서는 전국의 청소년을 대상으로 확률표집을 적용하여 얻어진 자료와의 비교가 요구된다. 또한, 단일 시점의 횡단 조사로 종속변수와 설명변수와의 인과관계를 엄격히 담보하지 못하기 때문에 시간의 선행성 조건을 충족하기 위한 종단자료의 구축이 필요하겠다.

【참고문헌】

- 권미경(2018), “디지털성범죄 피해자 지원 프로세스에 대한 연구”, 『이화젠더법학』, 10(3): 39-76, <DOI: 10.22791/ewhagl.2018.10.3.002>.
- 김경희·김수아·김은경(2020), “중·고등학생의 디지털 성폭력에 대한 다면적 이해와 정책 시사점”, 『이화젠더법학』, 12(2): 257-289, <DOI:10.22791/ewhagl.2020.12.2.007>.
- 김경희·이조은(2020), “피해자를 선정적으로 소비하는 당신, N번방의 공범이다”, 『월간 복지동향』, (259): 50-55.
- 김소라(2018), “디지털 성폭력의 변화 양상과 음란성(obscurity)을 근거로 한 규제의 한계”, 『아시아여성연구』, 57(1): 163-199.
- 김숙희·김영미·김현아·서승희·장윤정(2018), “디지털 성폭력 피해 영상물 삭제와 처리 개선방안에 관한 연구”, 『이화젠더법학』, 10(1): 41-66, <DOI:10.22791/ewhagl.2018.10.1.002>.
- 김영식·이은정·주현준(2019), “랜덤 포레스트를 활용한 대학 수시 전형 선택 관련 예측 요인 탐색”, 『교육과학연구』, 50(4): 233-255, <DOI:10.15854/jes.2019.12.50.4.233>.
- 김유정(2010), “청소년들의 개인적 요인이 인터넷 음란물 이용과 이용 후의 성태도 형성에 미치는 영향”, 『한국방송학보』, 24(6): 45-84.
- 김은실·김귀정·김봉한(2011), “고등학생들의 사이버 음란물 접촉과 성범죄와의 관계성 분석”, 『한국콘텐츠학회논문지』, 11(6): 8-17.
- 김정만(2000), “보호관찰 청소년의 성비행에 영향을 미치는 위험요인과 보호요인 연구”, 이화여자대학교 사회복지대학원 석사학위논문.
- 김한균(2017), “사이버성범죄·디지털성범죄 실태와 형사정책”, 『이화젠더법학』, 9(3): 27-57, <DOI:10.22791/ewhagl.2017.9.3.002>.
- 김한균(2020), “디지털성범죄 차단과 처단 -기술매개 젠더기반 폭력의 형사정책-”, 『저스투스』, (178): 369-392, <DOI:10.29305/tj.2020.06.178.369>.
- 김현아(2017), “카메라 등 이용촬영죄 처벌의 문제점과 개선방안”, 『이화젠더법학』, 9(2): 1-32, <DOI:10.22791/ewhagl.2017.9.2.001>.
- 대한민국정책브리핑(2020), “디지털 성범죄”, <<https://www.korea.kr/special/policyCurationView.do?newsId=148853543>(2020.11.15. 검색)>.

- 류진아(2003), “청소년의 인터넷 중독에 영향을 미치는 생태체계 변인”, 숙명여자대학교 대학원 박사학위논문.
- 백혜정·방은령(2009), “청소년 가출 현황과 문제점 및 대책 연구”, 『한국청소년정책연구원 연구보고서』, 3-9.
- 서승희(2017), “사이버성폭력 피해의 특성과 근절을 위한 대응방안”, 『이화젠더법학』, 9(3): 59-95, <DOI:10.22791/ewhagl.2017.9.3.003>.
- 심진숙(2008), “청소년의 사이버 비행행동 연구”, 이화여자대학교 대학원 박사학위논문.
- 여성가족부(2018), 『2018년 청소년 매체이용 및 유해환경 실태조사』, 세종: 여성가족부.
- 유진은(2015), “랜덤 포레스트: 의사결정나무의 대안으로서의 데이터 마이닝 기법”, 『교육평가연구』, 28(2): 427-448.
- 윤덕경(2020), “디지털 성범죄 관련 법·제도 현황과 개선과제”, 『KWDI Brief』, (55): 1-10.
- 윤정숙(2020), “디지털 성범죄의 특성과 가해자 심리”, 『한국심리학회 학술대회 자료집』, 95-95.
- 이나현·조운오(2020), “디지털 성범죄자에 대한 프로파일링: 신상정보 등록자를 중심으로”, 『사회과학연구』, 27(2): 135-156.
- 이수정(2015), “국내 청소년 성범죄자의 재범요인으로서의 성일탈적 경향성”, 『한국심리학회지: 사회 및 성격』, 29(4): 1-16, <DOI:10.21193/kjspp.2015.29.4.001>.
- 이유신(2012), “청소년 가출비행의 상습화와 심화”, 『한국사회학회 사회학대회 논문집』, 927-945.
- 이지선·이성식(2019), “대학생들의 스마트폰을 통한 음란물유통 행위의 설명: 위험요인과 보호요인의 상호작용효과와 성별 차이를 중심으로”, 『한국범죄학』, 13(3): 5-24.
- 이현철·김경식·최성보(2016), “한국 청소년의 지위비행·범죄행위에 대한 중단분석”, 『Korea Journal』, 24(2): 125-147, <DOI:10.35151/kyci.2016.24.2.007>.
- 이효빈·최윤경(2016), “자극추구 성향과 또래동조성이 청소년 성비행에 미치는 영향”, 『청소년학연구』, 23(1): 257-278.
- 장아름(2020), “아동·청소년 대상 성범죄 가해자 64%가 미성년자”, 『연합뉴스』(6월 29일).
- 전치혁(2012), 『데이터마이닝 기법과 응용』, 서울: 한나래.
- 정익중·이은주(2010), “비행또래집단과 청소년비행 간의 종단적인 역동적 관계”, 『사회복지연구』, 41(1): 119-144, <DOI:10.16999/kasws.2010.41.1.119>.

- 조린(2020), “디지털 성범죄 예방을 위한 7가지 안전수칙 제안”, 『여성가족부 정책뉴스』, <http://www.mogef.go.kr/nw/enw/nw_enw_s001d.do?mid=mda700&bbtSn=708205>(2020. 12. 20. 검색)>.
- 최명일(2009), “인터넷 음란물 노출에 영향을 미치는 요인과 음란물 노출이 실제 성 태도에 미치는 영향에 관한 연구”, 『한국언론학보』, 53(2): 49-69.
- 최필선·민인식(2018), “머신러닝 기법을 이용한 대졸자 취업예측 모형”, 『직업능력개발연구』, 21(1): 31-54, <DOI:10.36907/krivet.2018.21.1.31>.
- 홍기혜(2020), “랜덤포레스트 머신러닝 알고리즘 기반 남·여 청소년의 자살생각 예측 및 분석”, 『한국사회복지학』, 72(3): 157-180, <DOI:10.20970/kasw.2020.72.3.007>.
- 황옥경(1997), “아동의 대인지각과 문제해결 행동 및 또래지위와의 관계”, 숙명여자대학교 대학원 박사학위논문.
- Alves, L. G., Ribeiro, H. V., & F. A. Rodrigues(2018), “Crime prediction through urban metrics and statistical learning”, 『Physica A: Statistical Mechanics and its Applications』, (505): 435-443.
- Breiman, L. (2015), “randomForest: Breiman and Cutler’s Random Forests for Classification and Regression”, 『R package version』, 4: 6-12.
- Nguyen, T. T., Hatua, A., & A. H. Sung(2017), “Building a learning machine classifier with inadequate data for crime prediction”, 『Journal of Advances in Information Technology』, 8(2): 141-147.