

랜덤 포레스트를 활용한 고등학생의 학교도서관 이용 예측 요인 탐색*

Exploring Predictive Factors of High School Students' School Library Use Using Random Forest

임 정 훈 (Jeonghoon Lim)**

< 목 차 >

I. 서론	IV. 연구 결과
II. 이론적 배경	V. 논의 및 제언
III. 연구 방법	

요약: 본 연구는 랜덤 포레스트를 활용하여 고등학생의 학교도서관 이용을 예측하는 설명변수를 탐색하는 데 목적이 있다. 이를 위해 대전 지역 고등학생을 대상으로 설문조사를 실시하여 총 613명의 응답을 분석에 활용하였다. 가정 환경, 학교 환경, 학습 태도 및 습관, IT 환경, 개인 특성의 5개 영역 109개 설명변수를 설정하고, 학교도서관 이용 여부에 대한 변수의 상대적 중요도를 분석하였다. 분석 결과, 학교도서관 담당 교사 만족도, 수업 시간 도서 활용 활동, 동아리 활동 참여, 숏폼 시청 시간, 부모님 도서 추천, SNS 사용 시간, 학교도서관 관심 주제 도서 보유 등이 고등학생의 학교도서관 이용 여부를 예측하는 주요 변수로 확인되었다. 이러한 분석 결과를 근거로 학교도서관 담당 교사의 전문성 강화 및 역할 확대, 교과 연계 도서관 활용 수업 확대 및 동아리 활동과의 연계, 학생의 자율적 독서 환경 조성과 디지털 미디어의 도서관 서비스 연계, 가족 독서 문화 조성을 위한 지원 등을 제언하였다.

주제어: 학교도서관, 도서관 이용, 랜덤 포레스트, 고등학생

ABSTRACT: This study aims to explore predictor variables for high school students' use of school libraries by employing random forest. A survey was conducted with high school students in the Daejeon region, and a total of 613 responses were used for the final analysis. A total of 109 predictor variables across five domains – home environment, school environment, learning attitudes and habits, IT environment, and personal characteristics – were established, and the relative importance of each variable in predicting school library use was analyzed. The analysis revealed that student satisfaction with school library teachers, library book utilization activities during class time, participation in club activities, short-form video viewing time, parental book recommendations, social media usage time, and availability of books on topics of interest in the school library were confirmed as key predictors of high school students' school library use. Based on these findings, this study suggests strengthening the expertise and expanding the roles of school library teachers, expanding curriculum-linked library utilization activities and linking club activities with library use, fostering student-driven reading environments with integration of digital media into library services, and supporting the development of a family reading culture.

KEYWORDS: School Library, Library Use, Random Forest, High School Students

* 2024년 한국도서관·정보학회 동계 학술대회에서 발표한 내용을 수정·보완한 것임.

** 계명대학교 문헌정보학과 조교수(mictoxic@kmu.ac.kr / ISNI 0000 0004 8339 2694)

• 논문접수: 2026년 2월 19일 • 최초심사: 2026년 3월 7일 • 게재확정: 2026년 3월 10일

• 한국도서관·정보학회지, 57(1), 317-336, 2026. <http://dx.doi.org/10.16981/kliss.57.1.202603.317>

© Copyright © 2026 Korean Library and Information Science Society

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 (<https://creativecommons.org/licenses/by-nc-nd/4.0/>) which permits use, distribution and reproduction in any medium, provided that the article is properly cited, the use is non-commercial and no modifications or adaptations are made.

I. 서론

학교도서관은 학생들의 지적 성장과 학업 성취를 뒷받침하는 핵심 교육 인프라이다. 오늘날 학교도서관은 도서의 수집과 보관이라는 전통적 기능에 더하여, 학생들의 정보 리터러시를 함양하고, 자기주도적 학습을 지원하며, 창의적 사고력을 증진하는 교육의 중심지로 변모하고 있다. 특히 디지털 기술의 발달로 방대한 정보 속에서 필요한 정보를 선별하고 활용하는 능력이 더욱 중요해짐에 따라, 학습공간으로서 학교도서관의 역할이 그 어느 때보다 강조되고 있다. 실제로 Ralph와 Hegeman(2025)는 학교도서관 이용 참여도가 높은 학생일수록 학업 부진이 유의미하게 감소함을 확인하고, 학교도서관 이용이 학업 성취와 밀접하게 관련될 수 있음을 보고하였다.

그러나 이러한 중요성에도 불구하고 최근 디지털 미디어의 발달과 학생들의 독서 습관 변화로 인해 학교도서관 이용률이 지속적으로 감소하는 추세를 보이고 있다(문화체육관광부, 2021). 이에 학생들의 도서관 이용 행태를 정확히 파악하고 이를 바탕으로 효과적인 활성화 방안을 모색할 필요가 있다. 이세나(2024)는 대학생의 학교도서관 이용 경험에 대한 에세이 분석을 통해 휴식 공간 제공, 만화·장르 소설 장서 확대, 동아리 활동 강화 등 구체적인 이용 활성화 방안을 제시하였으며, 이승연과 김기영(2023)은 잠재프로파일 분석을 활용하여 초등학교 교사들의 학교도서관에 대한 인식 유형을 도출하고, 인식 유형에 따른 맞춤형 운영 및 활성화 방안을 제안하였다. 또한 임정훈(2020)은 학교도서관의 SNS 마케팅 활용 사례를 분석하여, 방문자 증가 및 이용자와의 소통 측면에서 SNS 활용의 효과를 확인하고 학교도서관 SNS 마케팅 활성화 방안을 제안하였다. 한편, Loh와 Binny(2024)는 싱가포르 중등학교(secondary school)의 도서관 재설계 사례를 분석하여, 학생들이 편안하고 다양한 활동이 가능한 도서관 환경을 선호하며, 이러한 공간 구성이 학생들의 자발적 독서와 도서관 이용을 지원하는 환경으로 인식되고 있음을 보고하였다.

앞서 살펴본 바와 같이 학교도서관 활성화를 위한 연구들이 다양하게 이루어져 왔으나, 대부분 질적 분석이나 소수의 변수를 대상으로 한 선행적 통계 분석에 의존하고 있어 다수의 변수를 동시에 고려하거나 변수 간의 비선형적 관계와 상호작용을 파악하는 데 한계가 있다. 학교도서관 이용은 이용자의 행동적 이용 패턴이나 명시적 요구사항뿐 아니라, 도서관 이용에서 비롯되는 감정, 심리 상태, 동기와 같은 내면적 요소에도 영향을 받을 수 있다(이병기 외, 2015). 이처럼 다양한 요인들이 복합적으로 작용한다는 점에서, 이를 체계적으로 분석할 수 있는 새로운 분석 방법이 요구된다.

이에 본 연구에서는 관련 선행연구를 분석하여 학교도서관 이용에 영향을 미치는 예측 변수를 도출하고, 머신러닝 기법 중 하나인 랜덤 포레스트를 활용하여 이들 변수의 상대적 중요도와 비선형적 관계를 탐색함으로써, 학생들의 학교도서관 이용 패턴을 다차원적으로 이해하고, 나아가 학교도서관 활성화를 위한 실천적 시사점을 제시하고자 한다.

II. 이론적 배경

1. 학교도서관의 역할 및 이용 요인

IFLA의 학교도서관 가이드라인에 따르면, 학교도서관은 “독서, 탐구, 연구, 사고, 상상, 창의성이 중심이 되는 물리적·디지털 학습공간으로서 학생들의 정보·지식 전환 과정과 개인적·사회적·문화적 성장을 지원하는 공간”으로 정의된다(IFLA, 2015). 우리나라 도서관법 제40조 제2항에서도 학교도서관의 역할을 자료의 수집·정리·보존에 한정하지 않고, 독서교육, 정보활용교육, 협동수업 지원, 멀티미디어 자료 제공 등 폭넓은 교육적 기능을 수행하는 공간으로 규정하고 있다.

학교도서관은 특히 학생들의 독서교육과 리터러시 향상에 핵심적인 역할을 담당한다. 학생들의 독서 능력은 모든 교과 학습의 기초가 되며, 지식정보 사회에서 필수적으로 요구되는 역량이다(변우열, 2015). 박주현과 사공복희(2014)는 도서관 방문 빈도가 독서 방법, 독서 태도, 읽기 성취도에 유의미한 영향을 미치며, 도서관을 자주 방문할수록 독서 태도와 읽기 능력이 향상될 수 있음을 보고한 바 있다.

교육부는 학교 독서교육의 중요성을 인식하여 지속적인 정책을 추진하고 있다. 특히 ‘제4차 학교도서관 진흥 기본계획(2024-2028)’을 통해 학교도서관의 교수·학습 지원 강화와 학생 독서교육 활성화를 중점 과제로 제시하고 있으며, 2022 개정 교육과정에서는 ‘읽기의 기초’ 단원을 신설하는 등 읽기 영역을 확대하고, 자기주도적 독서 활동 및 디지털 문해력 역량을 강조함으로써 공교육 내 독서교육을 체계적으로 강화하고 있다. 아울러 ‘2025년 학교도서관 진흥 시행 계획’에서는 하이브리드 독서 프로그램과 학교도서관 전담 인력 확대를 통해 디지털 대전환 시대에 부합하는 독서 친화적 환경 조성을 도모하고 있다.

그동안 학교도서관 이용에 영향을 미치는 요인에 대한 연구도 활발히 이루어져 왔다. 박유리(2010)는 중학생의 학교도서관 이용에 최신 자료 및 흥미로운 분야의 도서관 자료, 사서교사의 친절함, 학교도서관 행사, 이용 교육, 방학 중 도서관 개방 등이 영향을 미친다고 보고하였다. 유가현(2018)은 중학생과 고등학생의 도서관 이용 만족도와 개선점이 상이하게 나타남을 분석하고, 학교급에 따라 도서관 이용 목적과 요구사항이 달라질 수 있음을 확인하였다. 이연옥(2007)은 고등학생의 정보요구 유형과 정보원 선호를 조사하여, 진로·직업, 수업 및 학습, 친구 관계, 건강, 여가, 사회적 이슈 등 다양한 영역에서 정보요구가 나타남을 분석하였다. 학생들은 인터넷과 또래 집단을 주요 정보원으로 활용하는 반면, 도서관과 교사에 대한 의존도는 상대적으로 낮은 것으로 나타났다. 이병기(2023)는 페퍼토리 그리드 기법을 적용하여 학교도서관 서비스의 특성을 파악하고, 인적자원 간의 인식 차이를 분석하였다. 연구 결과, 자료관리, 도서관협력수업, 독서·정보 상담 서비스가

영향력이 큰 요소로 확인되었으며, 자기 역할의 분명한 정도, 실제 수행 정도, 자기조절능력의 구성 개념이 중요하게 작용하는 것으로 나타났다.

해외에서도 이와 관련된 논의가 활발히 전개되어 왔다. Williams et al.(2002)은 학교도서관 관련 선행연구를 비판적으로 검토하여, 도서관 자원(예산 규모, 장서의 규모와 다양성), 인적 자원(도서관 인력 규모, 사서의 교육적 역할 및 수업 참여), 서비스 및 운영 방식(공공도서관과의 협력, 온라인 정보 접근성, 장서개발 정책), 교육적 실천(사서교사와 일반교사의 협력적 수업, 정보활용교육)을 학교도서관의 효과적인 운영과 관련된 핵심 요인으로 제시하였다. Lai(2007)는 타이페이의 한 초등학교 학교도서관을 대상으로 한 질적 사례연구를 통해, 장서의 다양성과 흥미도, 공간·시설의 쾌적성과 접근성, 사서의 친절한 상호작용 등 도서관 요인과 학업 부담, 또래, 교사, 부모의 영향, 정서적 인식처 등 비도서관 요인을 학생들의 도서관 이용에 영향을 미치는 요인으로 제시하였다.

한편 학교도서관 이용의 장애 요인과 이용자 요구에 주목한 연구들도 다수 존재한다. 송인정(2009)은 시간 부족, 보유 자료의 부족, 독서에 대한 흥미 부족 등을 장애 요인으로 지적하였으며, 오수연(2010)은 전문계 고등학생의 경우 진로 탐색 및 진학 정보, 수업 관련 자료, 여가 생활을 위한 서비스 등에 대한 요구가 높음을 확인하였다. 또한 정대근과 홍현진(2013)은 도서관 이용이 고등학생의 독서 만족, 몰입, 독서 지속 의지에 영향을 미치며, 특히 대출이나 도서관 프로그램 참여를 목적으로 방문하거나 도서관에 대해 편안함을 느끼는 학생일수록 독서 유효성이 높아짐을 밝혔다.

이상의 선행연구를 종합하면, 학교도서관 이용에는 자료, 인적 자원, 시설, 서비스, 이용자의 심리적 요인 등 다양한 변수가 복합적으로 작용하고 있음을 알 수 있다. 따라서 효과적인 학교도서관 활성화 방안을 마련하기 위해서는 이러한 변수를 포괄적으로 탐색할 필요가 있다.

2. 랜덤 포레스트

랜덤 포레스트는 다수의 의사결정나무를 결합하여 예측 정확성을 높이는 앙상블 기반 머신러닝 기법이다. Breiman(2001)에 의해 제안된 이후, 통계, 의학, 공학, 자연과학 분야뿐만 아니라 사회과학 영역에서도 그 활용도가 크게 증가하고 있다(윤홍주, 2021).

랜덤 포레스트의 기본 원리는 약한 학습자인 개별 의사결정나무를 다수 결합하여 강한 학습자를 구성하는 것이다. 이 방식은 다수의 후보 모델을 동시에 훈련시켜 결합함으로써 하나의 최적 모델을 탐색하는 데 소요되는 시간을 줄이고 모델의 다양성을 확보한다(Breiman, 2001).

랜덤 포레스트는 붓스트랩(bootstrap) 샘플링과 변수 선택 과정에서 무작위성을 도입한다. 이 과정에서 각 의사결정나무는 전체 데이터의 일부만을 사용하여 구축되며, 각 노드에서 분할 시

무작위로 선택된 일부 변수만을 고려한다. 이러한 방식은 개별 나무의 다양성을 증가시키고, 다양한 나무들을 앙상블로 결합함으로써 결과적으로 모델의 예측 성능이 향상된다(김미림, 박민호, 2019; Hastie et al., 2009).

랜덤 포레스트의 주요 특성 중 하나는 다중 공선성 문제에서 자유롭다는 점이다. 이로 인해 수백 개 이상의 변수를 동시에 모형에 투입할 수 있으며, 상관관계가 높은 예측변수를 모형에 포함하더라도, 각 나무가 무작위로 선택된 변수의 부분집합만을 활용하기 때문에 개별 나무 간의 상관성이 낮아지고, 예측의 분산이 감소한다(최필선, 민인식, 2018).

랜덤 포레스트 분석은 종속변수의 유형에 따라 분류(classification) 모델과 회귀(regression) 모델로 구분된다. 이항 변수와 같은 범주형 변수의 경우 다수결을 적용하여 분류 모델을 구축하며, 연속형 변수의 경우 평균값을 산출하는 회귀 모델을 구축한다. 아울러 랜덤 포레스트는 다양한 예측 변수의 영향력을 안정적으로 파악할 수 있고, 중요도 지수와 부분 의존성 도표 등을 활용하여 결과를 해석할 수 있으며, 하이퍼파라미터 튜닝을 통해 모형 성능을 개선할 수 있다는 특성이 있다(Breiman, 2001; Hastie et al., 2009).

랜덤 포레스트의 분석 절차는 총 4단계로 구분할 수 있다(Hastie et al., 2009). 첫째, 붓스트랩 샘플링을 통해 여러 개의 서로 다른 훈련 데이터 세트를 생성한다. 둘째, 각 훈련 데이터 세트에 대해 의사결정나무를 학습한다. 셋째, 학습된 각 의사결정나무를 사용하여 테스트 데이터에 대한 예측을 수행한다. 넷째, 각 의사결정나무의 예측을 결합하여 최종 예측을 생성한다. 이러한 과정을 통해 랜덤 포레스트는 다양한 변수의 상호작용과 비선형 관계를 모델링할 수 있으며, 이상치의 영향을 상대적으로 덜 받는 장점이 있다. 따라서 본 연구와 같이 학생들의 학교도서관 이용과 관련된 다양한 변수를 포괄적으로 탐색하는 데 적합한 방법론이라 할 수 있다.

III. 연구 방법

1. 연구 대상

본 연구의 분석을 위해 대전 지역 고등학생을 대상으로 설문조사를 실시하였다. 조사 기간은 2024년 11월 1일부터 11월 10일까지로, 총 632부의 설문지가 회수되었다. 이 가운데 불성실하게 응답된 사례를 제외하고 최종적으로 613부를 분석에 활용하였다. 랜덤 포레스트 기법은 표본 수가 적은 경우에도 강건한 성능을 보이는 것으로 보고되어 있어(Qi, 2012), 본 연구의 표본 크기는 분석에 적합한 것으로 판단하였다. 연구 대상자의 인구통계 특성은 <표 1>과 같다.

〈표 1〉 연구 참여자 정보

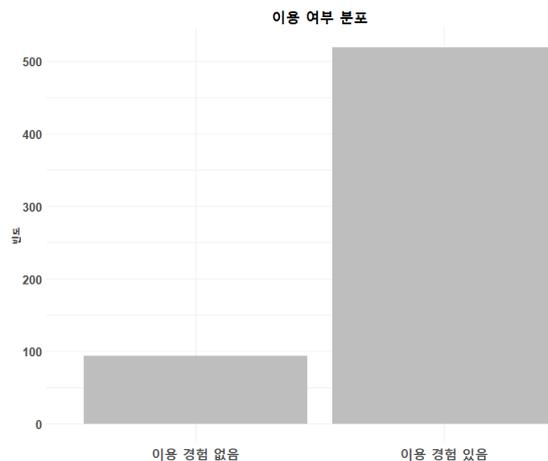
변인		인원(%)
성별	남	307(50.1)
	여	306(49.9)
학년	1학년	310(50.6)
	2학년	146(23.8)
	3학년	157(25.6)

2. 변수 설정

본 연구에서는 고등학생의 학교도서관 이용 요인을 탐색하기 위해 학교도서관 방문 경험을 종속변수로 설정하고, 가정 환경, 학교 환경, 학습 태도 및 습관, IT 환경, 개인 특성을 설명변수로 설정하였다. 종속변수는 고등학교 재학 기간 중 학교도서관 방문 여부를 묻는 문항으로, 방문 경험이 없으면 0, 있으면 1로 구성하여 이항 변수로 정의하였다.

학교도서관 이용 경험에 대한 응답 결과, 재학 중 이용 경험이 있는 학생이 519명(84.7%), 이용 경험이 없는 학생이 94명(15.3%)으로 나타나 데이터 비율의 불균형이 확인되었다(〈그림 1〉 참조). 이러한 불균형 문제를 해결하기 위해 본 연구에서는 김한용과 이우주(2017)의 연구를 참고하여 SMOTE(Synthetic Minority Over-sampling Technique) 기법으로 데이터를 오버샘플링하였다.

랜덤 포레스트 기법은 다중공선성 문제에서 자유롭기 때문에 다수의 설명변수를 동시에 투입할 수 있다(최필선, 민인식, 2018). 이에 본 연구에서는 관련 선행연구(박유리, 2010; 박주현, 사공복희, 2014; 이병기 외, 2015)를 참고하여 설명변수를 가정 환경, 학교 환경, 학습 태도 및 습관, IT 환경, 개인 특성의 5개 영역으로 구분하고, 총 109개 문항으로 설정하였다. 구체적인 변수 구성은 〈표 2〉와 같다.



〈그림 1〉 종속변수 분포

〈표 2〉 설명변수의 범주

범주		설명변수
가정 환경	독서 경험	가정에서 독서 여부, 부모님과 도서관 방문 경험, 부모님이 책 읽어준 경험
	부모 대화	부모님과 일상적인 대화, 책을 주제로 부모님과 대화, 주 평균 부모님과 대화시간
	부모 기대	부모님의 독서 관심, 책 읽기 권장, 도서 추천
	독서 여건	가정 책 보유 정도, 독서 분위기
	가정 정서	가족 관계, 가정의 엄격한 규칙
학교 환경	교육(수업) 환경	학급문고 보유, 독서 시간 제공, 수업 시간 도서 활용 활동, 독서지도 경험, 내적·외적 보상
	학교도서관 환경	학교도서관 담당교사 만족도, 학교도서관 독서프로그램, 학교도서관 관심 주제 도서 보유, 과제 해결에 필요한 책 보유, 학교도서관의 다양한 책, 학교도서관 방문 즐거움, 학교도서관 독서 적합 공간, 학교도서관 쉼 공간
	교사 관계	선생님과의 관계, 선생님의 독서 권장, 선생님의 도서 추천
	동료 관계	친구의 도서관 방문, 동아리 활동 참여
학습 태도 및 습관	학습 시간	주 평균 개인학습 시간, 사교육 시간
	학습 태도	학업에 대한 열의, 흥미, 태도
IT 환경	IT 기기 사용	IT 기기 처음 사용 시기, 디지털 기기 사용 시간 등
	미디어 사용	유튜브, 숏폼, SNS, OTT, 게임 시간 등
개인 특성	자기이해	자신의 장점 인식, 즐거움 감정, 삶의 만족도
	자아존중감	자기 존중, 독서 자신감, 끈기
	사회성	협력적 문제해결, 교우 관계, 대화
	학습동기	배움의 즐거움, 새로운 사실 발견, 독서 가치 인식
	진로	미래 희망, 진로 계획, 자기 결정력
	취미활동	주 평균 운동시간, 여가 시간, 좋아하는 연예인 등
독서 태도	종이책 선호, 전자책 선호, 독서 활동의 즐거움, 독서에 대한 좋은 기억 등	

3. 분석 방법

본 연구에서는 고등학생의 학교도서관 이용과 관련된 주요 변수를 탐색하기 위해 랜덤 포레스트 기법을 활용하였다. 랜덤 포레스트는 다수의 의사결정나무 결과를 통합함으로써 높은 예측력과 모형 안정성을 확보할 수 있으며, 다수의 설명변수를 투입하더라도 자유도 감소 문제를 야기하지 않는 장점이 있다. 이에 본 연구에서는 R 4.4.1 버전을 사용하여 613개의 설문조사 자료를 바탕으로 선정된 설명변수를 통해 고등학생의 학교도서관 이용 여부를 분석하였다.

분석 절차는 다음과 같다. 첫째, 데이터 전처리 과정을 통해 데이터를 분석에 적합한 형태로 변환하였다. 둘째, 전체 데이터를 학습용 데이터(75%)와 테스트용 데이터(25%)로 분할하였다. 셋째, SMOTE 기법을 적용하여 클래스 불균형 문제를 해결하였다. 넷째, 10회 교차 검증을 통해 랜덤 포레스트 모델의 주요 하이퍼파라미터를 튜닝하였다. 다섯째, 최적의 하이퍼파라미터($mtry = 58$, $trees = 988$, $min_n = 22$)를 사용하여 최종 랜덤 포레스트 모델을 학습하였다.

모델의 예측 성능을 다각도로 평가하기 위해 정확도, 정밀도, 민감도, 특이도, F1 점수, AUC 등의 지표를 활용하였다. 특히 AUC는 모델의 전반적인 분류 성능을 나타내는 지표로, Lantz(2019)의 기준에 따라 해석하였다.

한편, 변수 중요도 분석을 위해 순열 중요도(Permutation Importance) 기법을 사용하였다. 순열 중요도는 각 변수의 값을 무작위로 섞은 후 모델의 성능 변화를 관찰함으로써 해당 변수의 예측 기여도를 평가하는 방법으로, 기존의 중요도 지수보다 안정적이고 편향이 낮은 결과를 제공한다(Breiman, 2001). 아울러 부분 의존성 도표(Partial Dependence Plot)를 사용하여 각 설명 변수의 값 변화에 따른 예측 확률의 변화를 시각적으로 파악하였다.

이러한 분석 방법을 통해 고등학생의 학교도서관 이용과 관련된 주요 변수의 상대적 중요도와 변수별 관계 패턴을 구체적으로 파악하고자 하였다.

IV. 연구 결과

1. 예측 모형 성과 분석

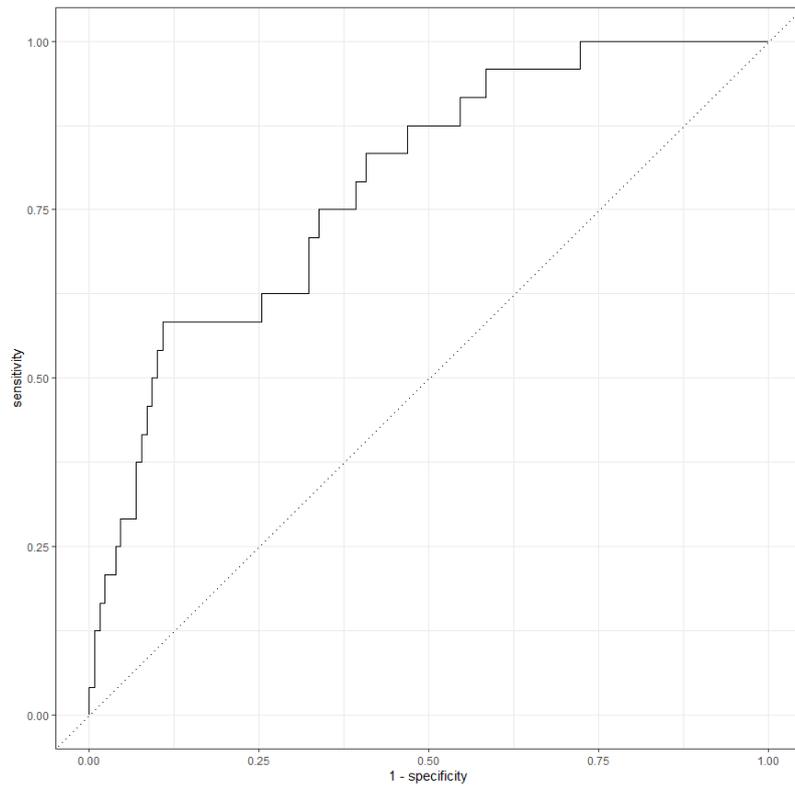
랜덤 포레스트 모델의 예측 성능을 평가한 결과는 <표 3>과 같다.

<표 3> 모델 성능

	정확도	정밀도	재현율	F1
시험자료	0.844	0.868	0.962	0.912

본 연구에서 구축한 모델은 고등학생의 학교도서관 이용 여부를 높은 정확도로 예측하는 것으로 나타났다. 시험자료 분석 결과, 모델의 전반적인 분류 정확도는 84.4%로 확인되었으며, 실제 도서관 이용자를 식별하는 재현율은 96.2%로 나타났다. 도서관 이용으로 예측한 사례 중 86.8%가 실제 이용자와 일치하였으며, 이 두 지표의 조화평균인 F1 점수는 0.912로 모델이 균형 잡힌 성능을 보이는 것으로 확인되었다.

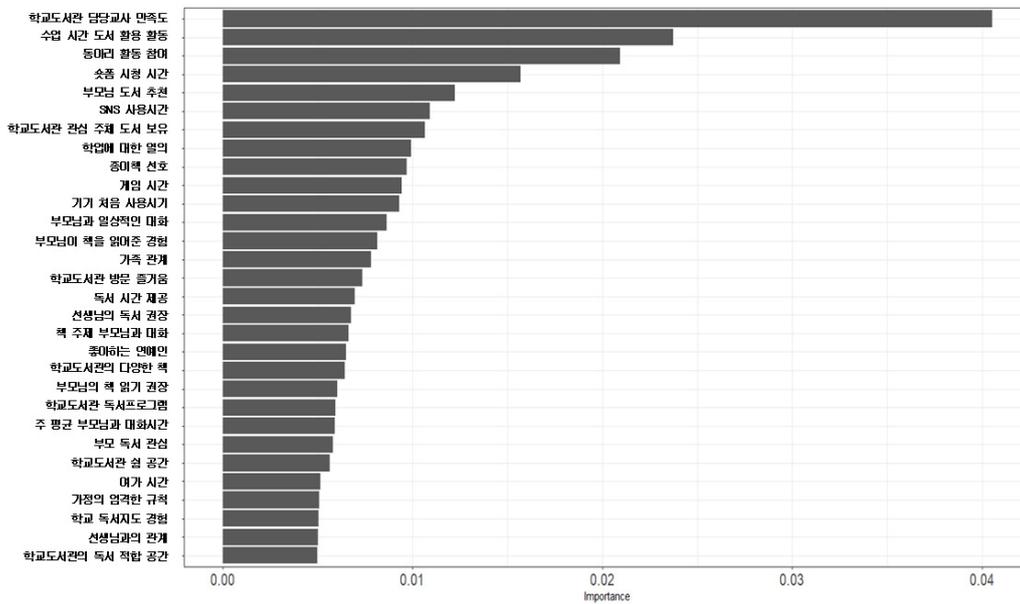
한편 모델의 분류 성능을 평가하기 위해 ROC 분석을 실시하였다. ROC 곡선은 다양한 임계값에서의 모델 성능을 시각화한 것으로, AUC는 이 곡선의 아래 면적을 의미하며 0에서 1 사이의 값을 가지고 1에 가까울수록 우수한 분류 성능을 나타낸다. 분석 결과, AUC 값은 0.788로 산출되었으며(<그림 2> 참조), 이는 Lantz(2019)의 기준에 따르면 '양호'한 수준의 예측 성능에 해당한다.



〈그림 2〉 ROC-곡선과 AUC-면적

2. 설명변수의 중요도

머신러닝은 통계적 유의성을 검증하는 대신, 예측에 기여하는 설명변수의 상대적 중요성을 평가한다(김영식 외, 2019). 본 연구에서는 순열 중요도 방법을 적용하여 각 설명변수가 고등학생의 학교도서관 이용 여부 예측에 미치는 상대적 기여도를 산출하였다. 순열 중요도는 특정 변수의 값을 데이터셋 내에서 무작위로 섞어 모델의 성능 변화를 관찰함으로써 해당 변수의 중요도를 평가하는 방식으로, 중요도 점수가 높을수록 해당 변수가 모델의 예측에 더 크게 기여하는 것으로 해석할 수 있다(Breiman, 2001; Fisher et al., 2019). 상위 30개에 해당하는 설명변수의 중요도 결과는 〈그림 3〉과 같다.



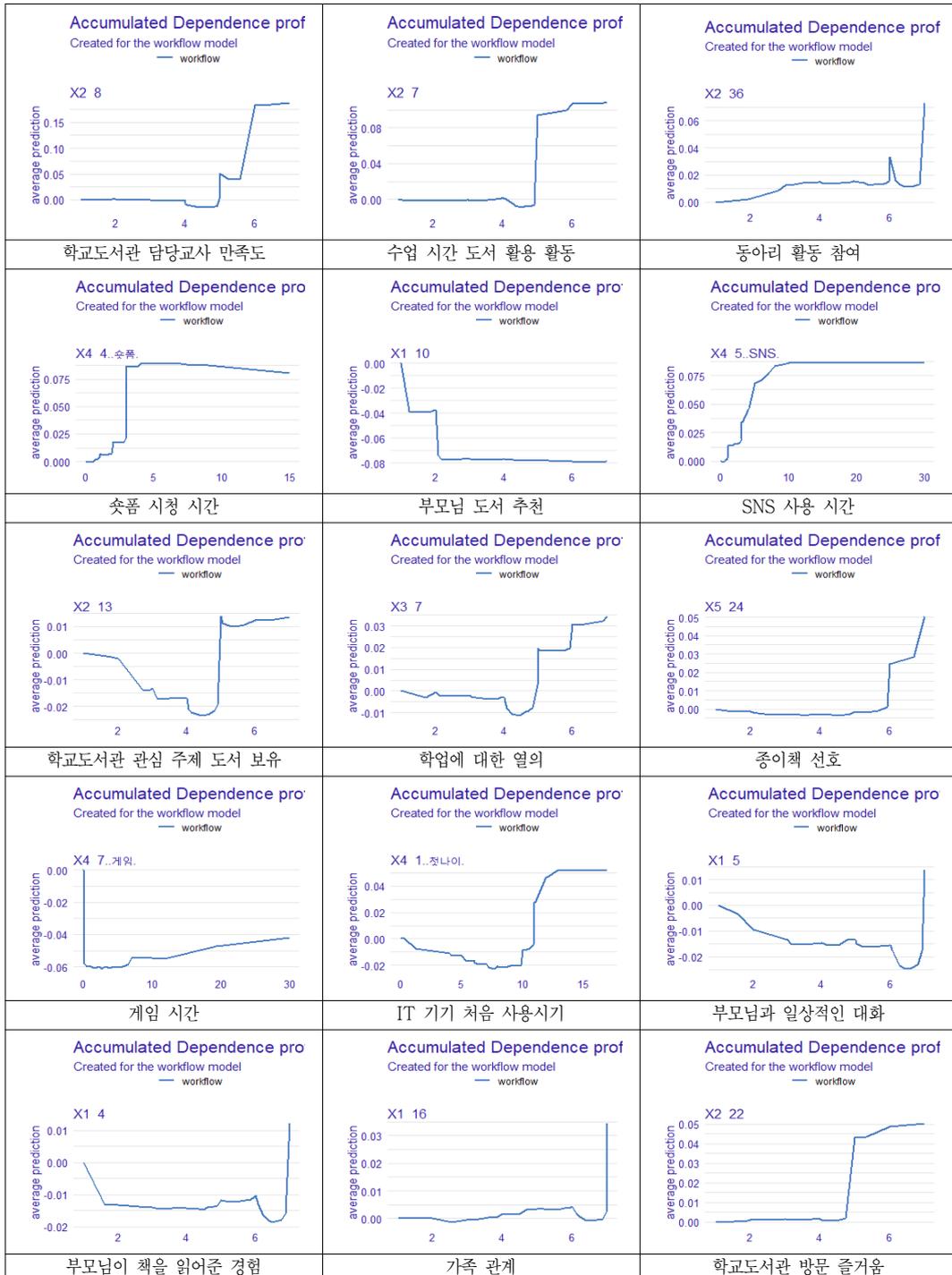
〈그림 3〉 설명변수의 중요도(Feature Importance) 결과

분석 결과, 고등학생의 학교도서관 이용 여부를 예측하는 데 가장 중요한 변수는 ‘학교도서관 담당교사 만족도’로 나타났으며, ‘수업시간 도서 활용 활동’과 ‘동아리 활동 참여’가 그 뒤를 이었다. 상대적으로 높은 중요도를 나타낸 이 세 가지 변수는 모두 학교 환경 요인에 해당한다는 공통점이 있다. 이어서 ‘숏폼 시청 시간’, ‘부모님 도서 추천’, ‘SNS 사용 시간’, ‘학교도서관 관심 주제 도서 보유’, ‘학업에 대한 열의’, ‘종이책 선호’, ‘게임 시간’, ‘기기 처음 사용시기’, ‘부모님과 일상적인 대화’, ‘부모님이 책을 읽어준 경험’ 순으로 중요도가 높게 확인되었으며, 이는 학교도서관 이용 예측에 학교 환경뿐만 아니라 가정 환경, 학습 태도, IT 환경, 개인 특성 등 다양한 요인이 복합적으로 관여하고 있음을 나타낸다.

이어서 학교도서관 이용과 관련된 주요 변수를 부분 의존성 도표를 통해 분석하였다. 부분 의존성 도표는 특정 설명변수와 모형의 예측 결과 간 함수 관계를 시각화하는 도구로(Friedman, 2001), 개별 설명변수의 값 변화에 따른 예측 확률의 변화 방향과 패턴을 분석할 수 있다(Greenwell, 2017). 부분 의존성 도표는 특정 설명변수의 값에 따른 사건 발생 확률의 변화를 한계 효과로 시각화한다는 점에서(이현, 이경원, 2021), 고등학생의 학교도서관 이용과 관련된 변수의 관계 방향을 직관적으로 파악하는 데 유용하다.

〈그림 4〉는 학교도서관 이용 여부 예측과 관련된 상위 15개 변수에 대한 부분 의존성 도표를 제시한 것이다. 그래프의 X축은 각 변수의 값, Y축은 부분 의존성 값을 나타낸다. 부분 의존성 도표에서 설명변수와 종속변수 간 관계는 1) 양의 관계, 2) 음의 관계에서 양의 관계로 전환,

랜덤 포레스트를 활용한 고등학생의 학교도서관 이용 예측 요인 탐색



〈그림 4〉 주요 변수별 부분 의존성 도표(Partial Dependence Plot)

3) 양의 관계에서 음의 관계로 전환, 4) 특정 지점에서 예측값이 급격히 증가 혹은 감소하는 계단 형태, 5) 음의 관계로 분류할 수 있다(박소영, 2024; 이주연 외, 2022). 그러나 본 연구에서는 2가지 이상의 패턴이 복합적으로 나타나는 변수가 다수 발견되었다. 대부분의 변수는 비선형적 관계를 나타내고 있어 변수 값의 변화에 따른 학교도서관 이용 예측 확률의 변동이 일정하지 않으며, 특히 일부 변수에서는 뚜렷한 임계점이 관찰되어 해당 변수가 특정 수준에 도달할 때 예측 패턴이 급격히 변화하는 양상을 보였다. 이러한 특성을 고려하여 유사한 패턴을 보이는 변수를 크게 세 가지 유형으로 분류하면 다음과 같다.

첫째, 특정 지점에서 예측값이 급격히 증가하거나 감소하는 계단 형태를 보이는 변수이다. '학교도서관 담당교사 만족도', '수업시간 도서 활용 활동', '동아리 활동 참여', '종이책 선호', '가족 관계', '학교도서관 방문 즐거움'은 특정 지점에서 예측값이 급격히 증가하는 양상을 보였으며, '부모님 도서 추천'은 특정 지점에서 예측값이 급격히 감소하는 양상을 보였다. 둘째, 특정 구간에서 음의 관계가 양의 관계로 전환되는 양상을 보이는 변수로는 '학교도서관 관심 주제 도서 보유', '학업에 대한 열의', '게임 시간', 'IT 기기 처음 사용 시기', '부모님과 일상적인 대화', '부모님이 책을 읽어준 경험'이 확인되었다. 셋째, 특정 구간에서 양의 관계가 음의 관계 혹은 일정 수준으로 유지되는 양상을 보이는 변수로는 '숏폼 시청 시간', 'SNS 사용 시간'이 확인되었다.

V. 논의 및 제언

본 연구에서는 고등학생의 학교도서관 이용에 관련된 변수를 탐색하기 위해 랜덤 포레스트 기법을 적용하였다. 연구 결과에 나타난 주요 특징은 다음과 같다.

순열 중요도 분석 결과, '학교도서관 담당교사 만족도', '수업 시간 도서 활용 활동', '동아리 활동 참여', '숏폼 시청 시간', '부모님 도서 추천', 'SNS 사용 시간', '학교도서관 관심 주제 도서 보유', '학업에 대한 열의', '종이책 선호' 등이 고등학생의 학교도서관 이용을 예측하는 데 주요한 변수로 확인되었다. 중요도 상위 30개 변수를 영역별로 살펴보면, 학교 환경 13개, 가정 환경 9개, IT 환경 4개, 개인 특성 3개, 학습 태도 1개 순으로 나타났다. 영역별 투입 변수 수가 동일하지 않아 단순 비교에는 한계가 있으나, 상위 중요도 변수에 학교 환경과 가정 환경 영역의 변수가 다수 포함된 것으로 확인되었다.

다음으로 부분 의존성 도표 분석에서 확인된 주요 특징을 논의하면 다음과 같다.

첫째, 특정 지점에서 예측값이 급격히 증가하거나 감소하는 계단 형태를 보이는 변수가 다수 확인되었다. 이 유형에 해당하는 변수는 '학교도서관 담당교사 만족도', '수업시간 도서 활용 활동', '동아리 활동 참여', '종이책 선호', '가족 관계', '학교도서관 방문 즐거움', '부모님 도서 추천'이며,

이들 변수는 공통적으로 특정 임계점 이하에서는 학교도서관 이용 예측 확률이 뚜렷한 변화가 없다가 해당 임계점을 넘어서면 이용 확률이 급격히 변화하는 비선형적 특성을 보였다. 세부 변수 별로 살펴보면 학교도서관 담당교사에 대한 만족도가 중간 수준을 넘어설 때 도서관 이용 예측 확률이 급증하였는데, 이는 담당교사와의 관계 형성과 효과적인 서비스 제공이 학교도서관 이용과 밀접하게 관련될 수 있음을 시사한다. 수업시간 도서 활용이나 동아리 활동 참여 역시 일정 수준을 넘어서면 도서관 이용 예측 확률이 크게 증가하는데, 이는 수업 및 동아리 활동을 통한 독서 참여가 학생들의 도서관 이용과 긍정적으로 관련될 수 있음을 보여준다. 종이책 선호와 학교도서관 방문 즐거움 역시 그 정도가 높아질수록 도서관 이용 예측 확률이 증가하였다. 이는 독서에 대한 긍정적 태도와 도서관에 대한 정서적 친밀감이 도서관 이용과 관련될 수 있음을 보여준다. 가족 관계의 경우 매우 긍정적인 관계가 유지될 때 학교도서관 이용 확률이 높아지는 경향을 보였는데, 이는 가족 간의 긍정적인 관계가 학업이나 독서에 대한 지지와 관련될 가능성이 있다. 반면, 부모님 도서 추천의 경우, 직접적인 도서 추천은 청소년의 자발적인 도서관 이용과 부정적으로 관련되는 것으로 나타났다. 이는 부모의 과도한 개입이 청소년의 자율적인 도서관 이용 동기와 상충될 수 있음을 보여주는 결과로 해석된다.

둘째, 특정 구간에서 음의 관계가 양의 관계로 전환되는 양상을 보이는 변수가 확인되었다. 이 유형에 해당하는 변수는 '학교도서관 관심 주제 도서 보유', '학업에 대한 열의', '게임 시간', 'IT 기기 처음 사용 시기', '부모님과 일상적인 대화', '부모님이 책을 읽어준 경험'이다. 세부 변수 별로 살펴보면, 학교도서관에서 학생들이 선호하는 작가나 주제와 관련된 도서 보유량이 일정 수준을 넘어서면 도서관 이용이 급격히 증가하는 것으로 나타났는데, 이는 학생들의 관심사를 반영한 장서 개발 정책의 중요성을 뒷받침하는 결과로 볼 수 있다. 학업에 대한 열의 역시 학교도서관 이용과 강한 양의 관계를 보였는데, 학업에 대한 열의가 높은 학생일수록 학습 자원으로서 도서관을 적극적으로 활용할 가능성이 높은 것으로 해석된다. 게임 시간의 경우 1~2시간의 비교적 짧은 시간에도 도서관 이용 확률이 크게 감소하는 것으로 확인되었다. 그러나 게임 시간이 일정 수준을 넘어서면 도서관 이용 확률이 다시 증가하는 양상을 보였는데, 이는 게임 시간과 도서관 이용 간의 관계가 단순한 부적 관계가 아닌 복합적인 양상일 수 있음을 의미한다. IT 기기 처음 사용 시기의 경우, IT 기기를 비교적 늦게(10세 이후) 처음 사용한 학생들의 도서관 이용 예측 확률이 가장 높게 나타난 반면, 비교적 일찍(10세 이전) 사용하기 시작한 학생들은 도서관 이용 확률이 점진적으로 감소하는 양상을 보였다. 이는 IT 기기의 조기 노출이 도서관 이용과 부정적으로 관련될 수 있음을 시사한다. 한편, 부모님과 일상적인 대화와 부모님이 책을 읽어준 경험이 낮은 수준일 때는 도서관 이용 예측 확률에 뚜렷한 변화가 나타나지 않으나, 일정 수준 이상이 되면 양의 관계로 전환되는 것으로 나타났다. 이는 부모와의 충분한 의사소통과 어린 시절의 풍부한 독서 경험이 도서관 이용 습관 형성과 관련될 수 있으며, 부모의 적극적인 독서 지원이

자녀의 도서관 이용과 밀접하게 연관될 수 있음을 시사하는 결과로 이해할 수 있다.

셋째, 특정 구간에서 양의 관계가 음의 관계로 전환되거나 일정 수준으로 유지되는 변수가 확인되었다. 이 유형에 해당하는 변수는 '숏폼 시청 시간', 'SNS 사용 시간'이다. 숏폼 콘텐츠나 SNS를 전혀 사용하지 않거나 매우 적게 사용할 때는 도서관 이용 확률이 낮지만, 사용 시간이 증가하면서 도서관 이용 확률이 급격히 증가하였다. 그러나 사용 시간이 숏폼의 경우 주 평균 5시간, SNS의 경우 주 평균 10시간을 초과하면 도서관 이용 확률이 점진적으로 감소하거나 일정 수준으로 유지되는 것으로 나타났다. 이는 적정 수준의 숏폼이나 SNS 사용이 호기심을 자극하고 정보 탐색 욕구를 촉진하여 도서관 이용과 긍정적으로 관련될 가능성이 있으나, 과도한 사용은 오히려 도서관 이용과의 긍정적 관련성을 약화시킬 가능성이 있는 것으로 해석된다.

이상의 연구 결과를 바탕으로 연구 제언을 제시하면 다음과 같다.

첫째, 학교도서관 담당 교사의 전문성 강화 및 역할 확대가 필요하다. '학교도서관 담당교사 만족도'가 도서관 이용에 가장 중요한 변수로 나타났으며, 특히 만족도가 중간 수준을 넘어설 때 도서관 이용 확률이 급격히 증가하는 것으로 확인되었다. 이를 고려할 때, 담당 교사의 정보서비스 능력, 독서지도 능력, 학생 상담 능력 등을 향상시키는 전문적인 연수 프로그램을 지원할 필요가 있다. 또한 담당 교사가 학생들과 긍정적인 관계를 형성하고 효과적인 도서관서비스를 제공할 수 있도록, 단순한 자료 관리자의 역할에서 나아가 독서 멘토, 협력 수업 지원 등 교육자로서의 역할까지 업무 범위를 확대하고 이에 대한 제도적 지원을 마련할 필요가 있다.

둘째, 교과 연계 도서관 활용 수업의 확대 및 동아리 활동과의 연계가 필요하다. '수업시간 도서관 활용 활동'과 '동아리 활동 참여'가 도서관 이용 예측에 중요한 변수로 확인되었으며, 특히 이들 활동이 일정 수준 이상 이루어질 때 도서관 이용 확률이 크게 증가하는 것으로 나타났다. 따라서 수업 시간에 독서를 활용하는 활동의 빈도를 확대하고, 독서 관련 동아리 활동을 적극 지원하여 도서관 이용과 연계하는 방안을 모색할 필요가 있다. 예를 들어, 동아리별 추천 도서 코너를 운영하거나 동아리 활동 결과물을 도서관에서 전시하는 등의 방안을 고려할 수 있을 것이다.

셋째, 학생 주도적 독서 문화 조성 및 디지털 미디어의 균형적 활용이 필요하다. '부모님 도서 추천'이 학교도서관 이용과 부정적으로 관련될 가능성이 있는 것으로 나타난 점을 고려할 때, 학생들의 자율성을 존중하는 독서 환경을 조성할 필요가 있다. 학생들이 스스로 책을 선택하고 독서 활동을 주도할 수 있도록 또래 추천 시스템이나 개인화된 도서 추천 서비스 등을 활용하는 방안을 고려할 수 있을 것이다. 한편, '숏폼 시청 시간'과 'SNS 사용 시간'이 적정 수준에서는 도서관 이용 확률과 긍정적으로 관련되는 것으로 나타난 점을 고려하여, 디지털 미디어를 도서관 서비스와 연계하는 방안을 모색할 필요가 있다. 예를 들어, 숏폼 콘텐츠를 활용한 북트레일러 제작, SNS 기반 독서 챌린지, 학생 서평단 운영 등을 통해 학생들의 자발적인 참여를 유도할 수 있을 것이다. 다만, 과도한 숏폼이나 SNS 사용은 도서관 이용 확률의 감소 또는 정체와 관련될 가능성이 있으

므로, 균형 잡힌 미디어 소비에 대한 교육도 함께 이루어질 필요가 있다.

넷째, 가족 독서 문화 조성을 위한 지원이 필요하다. ‘가족 관계’, ‘부모님과 일상적인 대화’, ‘부모님이 책을 읽어준 경험’ 등이 도서관 이용 예측에 중요한 변수로 확인되었으며, 특히 이들 변수가 일정 수준 이상일 때 도서관 이용 확률이 높아지는 것으로 나타났다. 다만, ‘부모님 도서 추천’이 도서관 이용과 부정적으로 관련될 가능성이 확인된 점을 고려할 때, 직접적인 부모 개입보다는 가족이 함께 참여할 수 있는 독서 관련 활동을 제공하는 것이 효과적일 수 있다. 예를 들어, 가족 독서 캠프 등을 통해 부모와 자녀가 함께 독서의 즐거움을 경험할 수 있는 기회를 제공하거나, 부모를 대상으로 청소년기 자녀의 독서지도 방법에 대한 교육을 실시하여 강압적이지 않은 방식으로 가정에서의 독서 문화를 조성할 수 있도록 지원할 필요가 있다.

본 연구는 랜덤 포레스트 기법을 적용하여 학교도서관 이용과 관련된 변수를 포괄적으로 탐색하였다는 점에서 의의가 있다. 기존의 선형적 분석 방법으로는 파악하기 어려운 복잡한 상호작용과 비선형적 관계를 확인할 수 있었으며, 다수의 설명변수를 동시에 고려함으로써 각 변수의 상대적 중요도를 산출할 수 있었다. 또한 부분 의존성 도표를 통해 각 변수와 학교도서관 이용 간의 관계 패턴을 시각적으로 제시함으로써 보다 구체적인 해석이 가능하였다.

그러나 본 연구는 랜덤 포레스트 기법의 특성상 변수 간 인과관계를 명시적으로 설명하기 어려우며, 횡단적 데이터를 사용하였기 때문에 시간에 따른 변화를 파악하는 데도 한계가 있다. 또한 제시된 변수들이 학교도서관 이용 행동을 설명하는 모든 요인을 포괄하지 못했을 가능성도 존재한다. 따라서 후속 연구에서는 본 연구에서 도출된 주요 예측 변수들 간의 인과관계를 심층적으로 분석하고, 종단 연구를 통해 이러한 요인들의 시간에 따른 변화와 상호작용을 탐구할 필요가 있다. 더불어 질적 연구 방법을 병행하여 분석 결과를 보완하고 더 깊이 있는 해석을 제공하는 것도 의미 있는 연구 방향이 될 것이다.

참 고 문 헌

- 김미림, 박민호 (2019). 랜덤 포레스트를 활용한 대학생의 최초 취업 사교육 참여 시점별 특성 분석. *교육연구논총*, 40(1), 1-33. <http://doi.org/10.18612/cnujes.2019.40.1.1>
- 김영식, 김민석, 이길재 (2019). 랜덤 포레스트를 활용한 고교생의 독서활동 예측 요인 분석. *교육재정경제연구*, 28(4), 137-156. <http://doi.org/10.46967/jefe.2019.28.4.137>
- 김한용, 이우주 (2017). 불균형적인 이항 자료 분석을 위한 샘플링 알고리즘들: 성능비교 및 주의점. *응용통계연구*, 30(5), 681-690. <https://doi.org/10.5351/KJAS.2017.30.5.681>
- 문화체육관광부 (2021). 2021년 국민 독서실태 조사. 출처:

- http://www.mcst.go.kr/site/s_policy/dept/deptView.jsp?pSeq=1574&pDataCD=0417000000
- 박소영 (2024). 랜덤 포레스트를 활용한 고등학교 교사의 수업 개선 활동에 영향을 미치는 요인 분석. 박사학위논문, 이화여자대학교.
- 박유리 (2010). 중학생들의 학교도서관 이용에 영향을 미치는 요인: 학교도서관 이용 빈도를 중심으로. 석사학위논문, 대전대학교.
- 박주현, 사공복희 (2014). PISA 2009에서 도서관방문이 독서태도 및 독서방법과 읽기성취도에 미치는 영향. 한국문헌정보학회지, 48(3), 169-191.
<http://doi.org/10.4275/KSLIS.2014.48.3.169>
- 변우열 (2015). 독서교육의 이해. 서울: 한국도서관협회.
- 송인정 (2009). 학교도서관 이용자 만족도에 관한 연구: 목포지역 중학교를 중심으로. 석사학위논문, 전남대학교.
- 오수연 (2010). 전문계 고등학교 도서관의 이용자 요구 분석: 공업고등학교를 중심으로. 석사학위논문, 중앙대학교.
- 유가현 (2018). 학교도서관 이용자 만족도 분석: 인천의 중등학교를 중심으로. 석사학위논문, 중앙대학교.
- 윤홍주 (2021). 랜덤 포레스트를 활용한 초등학생의 진로개발역량 예측요인 탐색. 학습자중심교과교육연구, 21(23), 819-833. <http://doi.org/10.22251/jlcci.2021.21.23.819>
- 이병기 (2023). 레퍼토리 그리드 기법을 통한 학교도서관 서비스 특성 분석에 관한 연구. 한국도서관·정보학회지, 54(3), 249-270. <http://doi.org/10.16981/kliss.54.3.202309.249>
- 이병기, 송기호, 김성준 (2015). 학교도서관 이용에 대한 플로우 경험과 만족도의 인과관계에 관한 연구. 한국도서관·정보학회지, 46(3), 119-140.
<http://doi.org/10.16981/kliss.46.3.201509.119>
- 이세나 (2024). 대학생의 경험 분석을 통한 학교도서관 이용 활성화 방안. 한국비블리아학회지, 35(2), 47-64. <http://doi.org/10.14699/kbiblia.2024.35.2.047>
- 이승연, 김기영 (2023). 초등학교 교사의 학교도서관 역할 인식유형이 이용행태에 미치는 영향에 관한 연구. 정보관리학회지, 40(4), 375-401. <http://doi.org/10.3743/KOSIM.2023.40.4.375>
- 이연옥 (2007). 고등학생의 정보요구 해결을 위한 도서관서비스 방안. 한국도서관·정보학회지, 38(1), 107-134. <http://doi.org/10.16981/kliss.38.1.200703.107>
- 이주연, 정채윤, 윤혜주, 손민지, 홍아정 (2022). 랜덤 포레스트 회귀를 활용한 일반대학원생의 학습몰입에 미치는 영향요인 탐색. 교육학연구, 60(2), 145-175.
<http://dx.doi.org/10.30916/KERA.60.2.145>

- 이현, 이경원 (2021). 다문화 가정 어머니의 문화적응 유형과 분류 예측요인 탐색연구: 잠재프로파일과 랜덤포레스트를 활용하여. *문화교류와 다문화교육*, 10(2), 171-193.
<http://doi.org/10.30974/kaice.2021.10.2.7>
- 임정훈 (2020). SNS를 활용한 학교도서관 마케팅 활성화 방안 연구. *한국도서관·정보학회지*, 51(1), 21-41. <http://doi.org/10.16981/kliss.51.1.202003.21>
- 정대근, 홍현진 (2013). 청소년의 독서유효성에 영향을 미치는 요인에 관한 연구: G지역 특성화고를 중심으로. *한국도서관·정보학회지*, 44(3), 235-259.
<http://doi.org/10.16981/kliss.44.3.201309.235>
- 최필선, 민인식 (2018). 머신러닝 기법을 이용한 대졸자 취업예측 모형. *직업능력개발연구*, 21(1), 31-54. <http://doi.org/10.36907/krivet.2018.21.1.31>
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5-32.
- Fisher, A., Rudin, C., & Dominici, F. (2019). All models are wrong, but many are useful: Learning a variable's importance by studying an entire class of prediction models simultaneously. *Journal of Machine Learning Research*, 20(177), 1-81.
- Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. *Annals of Statistics*, 29(5), 1189-1232. <http://doi.org/10.1214/aos/1013203451>
- Greenwell, B. M. (2017). pdp: An R Package for Constructing Partial Dependence Plots. *The R Journal*, 9(1), 421-436.
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (2nd ed.). New York: Springer Science & Business Media.
- IFLA (2015). *IFLA School Library Guidelines* (2nd ed.). Available: <https://repository.ifla.org/items/e61cb0b6-8d19-45e8-9be1-24b45309464e>
- Lai, K. Y. (2007). Motives that drive students to use the school library in Taipei Municipal Mingchuan Elementary School: A field study. 2007 IASL Conference Proceedings: Cyberspace, D-World, E-Learning: Giving Libraries and Schools the Cutting Edge. <https://doi.org/10.29173/iasl7630>
- Lantz, B. (2019). *Machine learning with R* (3rd ed.). 윤성진, (주)크라슬랩 옮김 (2020). *R을 활용한 머신 러닝*. 서울: 에이콘출판.
- Loh, C. E. & Binny, A. L. (2024). I use the library sometimes just to lose myself: school libraries supporting students' wellbeing and flourishing. *IFLA Journal*, 51(2), 399-408. <http://dx.doi.org/10.1177/03400352241276840>

- Qi, Y. (2012). Random forest for bioinformatics. In Zhang, C. & Ma, Y. eds. Ensemble machine learning. New York: Springer, 307-323.
http://dx.doi.org/10.1007/978-1-4419-9326-7_11
- Ralph, M. C. & Hegeman, M. (2025). Remodeled high school library increases use and improves course performance. *IFLA Journal*, 51(4), 1045-1057.
<http://dx.doi.org/10.1177/03400352251363297>
- Williams, D., Coles, L., & Wavell, C. (2002). Impact of school library services on achievement and learning in primary schools: Critical literature review of the impact of school library provision on achievement and learning in primary level students. OpenAIR@RGU. Available: <http://openair.rgu.ac.uk>

• 국한문 참고문헌의 영문 표기

(English translation / Romanization of references originally written in Korean)

- Byun, Woo-Yeoul (2015). *Understanding Reading Education*. Seoul: Korea Library Association.
- Choi, Pilsun & Min, Insik (2018). A predictive model for the employment of college graduates using a machine learning approach. *Journal of Vocational Education & Training*, 21(1), 31-54. <http://doi.org/10.36907/krivet.2018.21.1.31>
- Jeong, Dae-keun & Hong, Hyun-jin (2013). A study on the factors that affect adolescents' reading effectiveness: represented mainly from special high school in G city. *Journal of Korean Library and Information Science Society*, 44(3), 235-259.
<http://doi.org/10.16981/kliss.44.3.201309.235>
- Kim, Hanyong & Lee, Woojoo (2017). On sampling algorithms for imbalanced binary data: performance comparison and some caveats. *The Korean Journal of Applied Statistics*, 30(5), 681-690. <https://doi.org/10.5351/KJAS.2017.30.5.681>
- Kim, Meereem & Park, Minho (2019). Analysis of the characteristics of college students according to the first-time participation of private tutoring using random forest. *CNU Journal of Educational Studies*, 40(1), 1-33. <http://doi.org/10.18612/cnujes.2019.40.1.1>
- Kim, Young-sik, Kim, Minseok, & Lee, Giljae (2019). Analysis on the factors predicting reading activities of high school students. *The Journal of Economics and Finance of Education*, 28(4), 137-156. <http://doi.org/10.46967/jefe.2019.28.4.137>
- Lee, Byeong-Kee (2023). A study on the analysis of characteristics of school library services

- using the repertory grid technique. *Journal of Korean Library and Information Science Society*, 54(3), 249-270. <http://doi.org/10.16981/kliss.54.3.202309.249>
- Lee, Byeong-Kee, Song, Gi-Ho, & Kim, Sung-Jun (2015). A study on the structural relationships between flow experience and satisfaction about school library use. *Journal of Korean Library and Information Science Society*, 46(3), 119-140. <http://doi.org/10.16981/kliss.46.3.201509.119>
- Lee, Hyun & Lee, Kyeong-Won (2021). A study on acculturation profiles and predictors of mothers in multicultural families: using latent profile analysis and random forest. *Cultural Exchange and Multicultural Education*, 10(2), 171-193. <http://doi.org/10.30974/kaice.2021.10.2.7>
- Lee, Juyeon, Jeong, Chaeyun, Yun, Hyejoo, Son, Minjie, & Hong, Ah-Jeong (2022). An application of random forest regression for exploring factors influencing graduate learning engagement. *Korean Journal of Educational Research*, 60(2), 145-175. <http://dx.doi.org/10.30916/KERA.60.2.145>
- Lee, Sena (2024). A study on revitalizing the use of school libraries through analysis of university students' experiences. *Journal of the Korean Biblia Society for Library and Information Science*, 35(2), 47-64. <http://doi.org/10.14699/kbiblia.2024.35.2.047>
- Lee, Seungyeon & Kim, Giyeong (2023). A study on the impact of elementary school teachers' perceptions of the school library roles on their library use behaviors. *Journal of the Korean Society for Information Management*, 40(4), 375-401. <http://doi.org/10.3743/KOSIM.2023.40.4.375>
- Lee, Yeon-Ok (2007). School library information services for satisfying academic high school students' information needs. *Journal of Korean Library and Information Science Society*, 38(1), 107-134. <http://doi.org/10.16981/kliss.38.1.200703.107>
- Lim, Jeonghoon (2020). A study of the use of SNS for the activation of school library marketing. *Journal of Korean Library and Information Science Society*, 51(1), 21-41. <http://doi.org/10.16981/kliss.51.1.202003.21>
- Ministry of Culture, Sports and Tourism (2021). 2021 National reading survey. Available: http://www.mcst.go.kr/site/s_policy/dept/deptView.jsp?pSeq=1574&pDataCD=0417000000
- OH, Soo-yeon (2010). A Study on User Needs for Vocational High School Library: the case of technical high schools. Master's thesis, Chung-ang University, Korea.

- Park, Juhyeon & Sakong, Bok-Hee (2014). Effects of the library use factor on reading achievement, reading attitude and behavioral factors of PISA 2009. *Journal of the Korean Society for Library and Information Science*, 48(3), 169-191.
<http://doi.org/10.4275/KSLIS.2014.48.3.169>
- Park, So-Young (2024). Analysis of Factors Affecting the Class Improvement Activities of High School Teachers Using Random Forest. Doctoral dissertation, Ewha Womans University. Korea.
- Park, Yu-ri (2010). The Factors that Affection School Library Use of Middle School Students: By Focusing on School Library Using Rate. Master's thesis, Daejin University, Korea.
- Song, In-jeong (2009). A Study on User Satisfaction of School Library: Based on Middle Schools in Mokpo. Master's thesis, Chonnam National University, Korea.
- Yu, Ga-hyun (2018). An Analysis of the User Satisfaction in School Libraries: Focused on the Middle and High School in Incheon. Master's thesis, Chung-ang University, Korea.
- Yun, Hong-Ju (2021). Exploring predictive factors for career development competencies of elementary students with using random forest. *The Journal of Learner-Centered Curriculum and Instruction*, 21(23), 819-833.
<http://doi.org/10.22251/jlcci.2021.21.23.819>