

중학교 3학년 청소년의 행복감 관련 요인: 머신러닝 기반 비교 분석

이 상 미¹

¹ 동양대학교

초 록

본 연구는 중학교 3학년 청소년의 행복감 관련 요인을 머신러닝을 활용하여 파악하기 위해 육아정책연구소의 2023년 제16차 한국아동패널 II 자료를 이용하였으며, 분석대상은 중학교 3학년 학생 1,154명이었다. 행복감을 결과변수로 하고, 청소년의 인구사회학적·심리적·학업적·관계적 및 생활습관 특성을 예측변수로 구성하였다. 분석에서는 평균예측모형을 기준으로 다중선형회귀, 엘라스틱넷, 랜덤포레스트 및 XGBoost의 예측 성능을 비교하였다. 또한 변수 중요도와 SHAP 분석을 통해 주요 변수의 상대적 중요도와 예측 기여 방향을 확인하고, 랜덤포레스트의 부분의존도 분석을 통해 비선형적 예측 양상을 탐색하였다. 분석 결과, 교차검증과 최종 검증에서 모형 간 성능 순위는 일관되지 않았으며 전반적으로 유사한 수준을 보였다. 트리 기반 머신러닝 모형(랜덤포레스트 및 XGBoost)에서 자아존중감이 가장 중요한 변수로 나타났다. 우울, 또래애착, 자아탄력성, 학업스트레스, 부모애착, 자기조절학습, 스마트폰 과의존 및 주관적 사회경제적 지위도 주요 관련 요인으로 확인되었다. SHAP 분석에서 주요 변수의 기여 방향은 두 모형 간 대체로 일관되었으며, 부분의존도 분석에서는 일부 주요 변수와 행복감 간의 비선형적 관계 양상이 나타났다. 이러한 결과는 중학교 3학년 청소년의 행복감이 다양한 심리적·관계적·학업적 요인과 관련되어 있음을 보여주며, 청소년 행복감 이해를 위한 기초자료를 제공한다.

주요 용어: 청소년, 행복감, 머신러닝, 랜덤포레스트, XGBoost

알기 쉬운 요약

이 연구는 왜 했을까? 청소년기는 학업 부담, 부모·또래 관계, 정서 변화 등 다양한 요인이 행복감에 영향을 미치는 시기이다. 본 연구는 중학교 3학년 청소년의 행복감 관련 요인을 살펴보고, 머신러닝을 활용해 주요 요인을 파악하고자 수행하였다.

새롭게 밝혀진 내용은? 분석 결과, 행복감과 가장 중요하게 관련된 요인은 자아존중감이었으며, 우울, 또래애착, 자아탄력성, 학업스트레스, 부모애착, 자기조절학습, 스마트폰 과의존 등도 주요 요인으로 나타났다. 일부 요인은 행복감과 비선형적인 관계를 보였다. 이는 청소년의 행복감이 개인의 심리적 특성, 가족·또래 관계, 학업 및 생활습관 등 다양한 요인과 복합적으로 관련됨을 보여준다.

앞으로 무엇을 해야 하나? 청소년의 행복감을 높이기 위해 자아존중감과 자아탄력성을 강화하고, 우울과 학업스트레스를 줄이며, 부모·또래와의 긍정적 관계를 지원할 필요가 있다. 앞으로는 종단연구를 통해 요인과 행복감의 변화 관계를 확인하고, 학교와 가정에서 활용할 행복 증진 프로그램을 개발할 필요가 있다.

이 논문은 2025년도 동양대학교 학술연구비의 지원으로 수행되었음.

■ 투 고 일: 2026. 04. 11.
■ 수 정 일: 2026. 08. 18.
■ 게재확정일: 2026. 09. 15.

I. 서론

1. 연구의 필요성

행복감은 개인이 자신의 삶을 긍정적으로 평가하고 일상에서 긍정적 정서를 경험하는 주관적 웰빙의 핵심 요소로, 삶에 대한 인지적 만족과 정서적 안녕을 포괄한다(Diener et al., 2018; Steinmayr et al., 2019). 특히 청소년기의 행복감은 현재의 정서적 적응과 삶의 질을 보여주는 지표일 뿐 아니라 이후 성인기의 행복과 적응에도 관련되는 중요한 발달적 자원이다(김동심, 변종석, 2023). 그러나 우리나라 청소년의 정신건강과 주관적 안녕 관련 지표는 지속적인 관심이 필요함을 보여준다. 아동·청소년 삶의 질 2025 보고서에 따르면, 청소년의 스트레스 인지율은 2014년 37.0%에서 2024년 42.3%로 증가하였고, 일상생활에서 불안을 느끼는 비율도 2020년 11.2%에서 2024년 14.1%로 증가하였다. 또한 우리나라 15세 청소년의 삶의 만족도는 OECD 국가 중 하위권으로 보고되었다(심수진, 김은아, 2025). 이러한 결과는 우리나라 청소년의 행복감과 관련된 요인을 구체적으로 파악할 필요성을 보여준다.

중학교 3학년은 고등학교 진학을 앞두고 학업 부담과 진로 탐색이 구체화되며, 또래관계와 자아정체감의 변화가 함께 나타나는 발달적 전환기이다(박지선 외, 2019). 이 시기의 청소년은 학교급 전환을 준비하는 과정에서 학업과 진로에 대한 심리적 부담을 경험할 수 있으며, 학교환경과 학업 경험은 청소년의 정서적·사회적 안녕과 밀접하게 관련된다(OECD, 2017). 따라서 15세 전후에 해당하는 중학교 3학년 청소년의 행복감 관련 요인을 파악하는 것은 이들의 학교적응과 정신건강을 지원하기 위한 기초자료를 마련한다는 점에서 중요하다.

청소년의 행복감은 단일 요인보다 인구사회학적 배경, 심리적 자원, 학업 경험, 부모 및 또래관계, 생활습관 등 다양한 요인과 관련된다. 선행연구에서는 자아존중감, 자아탄력성 및 자율성이 행복감과 정적인 관련을 보인 반면, 우울과 스트레스는 부적인 관련을 보였다(서경현, 2016; 송주빈, 신정택, 2020; Yang & Moon, 2024; Wu et al., 2025). 학업스트레스와 자기조절학습 등 학업 관련 경험도 행복감과 관련되며(김선아, 2020; 전해경, 2024), 부모 및 또래와의 긍정적인 관계는 삶의 만족과 행복감에 관련되는 중요한 관계적 자원으로 보고되었다(강인경, 김춘경, 2016; 이현주, 2018). 또한 스마트폰 과의존, 수면 및 신체활동과 같은 생활습관과 사회경제적 특성도 청소년의 행복감 및 주관적 웰빙과 관련되는 것으로 나타났다(신승배, 2018; 김무현, 김준호, 2024; Kim et al., 2022; Nie et al., 2025). 이처럼 청소년의 행복감은 여러 영역의 요인과 관련된 다차원적 특성을 지니므로, 다양한 관련 요인을 하나의 분석틀에서 통합적으로 검토할 필요가 있다.

기존 청소년 행복감 연구의 상당수는 선형 회귀모형이나 구조방정식모형을 활용하여 변수 간 직접적인 관련성이나 매개경로를 검증해 왔다(김선아, 2020; 송주빈, 신정택, 2020; Yang & Moon, 2024). 이러한 접근은 이론에 근거하여 변수 간 관계를 검증하는 데 유용하지만, 예측변수와 행복감 간의 함수 형태를 미리 설정하기 때문에 복잡한 비선형적 예측 양상을 탐색하는 데에는 한계가 있다. 이에 최근에는 머신러닝을 활용하여 청소년의 행복감과 주관적 웰빙을 예측하고 주요 관련 변수를 탐색하려는 연구가 이루어지고 있다. 국외에서는 트리 기반 알고리즘을 적용하여 청소년의 삶의 만족과 주관적 웰빙을 예측하고 주요 심리·사회적 변수를 확인하였다(Zhang et al., 2019; Pan & Cutumisu, 2024). 국내에서도 청소년건강행태조사 자료를 활용하여 정신건강, 건강행태 및 사회경제적 요인의 행복감 예측 중요도를 분석한 연구가 보고되었으며(Kim et al., 2025), 한국아동패널 종단자료를 이용하여 행복감 변화 궤적을 유형화하고 여러 머신러닝 모형의 분류 성능과 SHapley Additive exPlanations(SHAP) 값을 분석한 연구도 이루어졌다(송소연 외, 2026).

이러한 연구들은 청소년 행복감 연구에서 머신러닝의 활용 범위를 확장하였다는 의의가 있다. 그러나 건강행태 중심의 자료를 사용한 연구에서는 자아존중감과 자아탄력성 등의 심리적 자원, 부모 및 또래관계, 학업 경험을 함께 고려하는 데 한계가 있었다(Kim et al., 2025). 또한 행복감 변화 궤적을 예측한 연구는 종단적 변화 유형을

분류하는 데 초점을 두었으므로(송소연 외, 2026), 특정 발달시점의 행복감을 연속형 결과로 설정하여 선형모형, 정규화 회귀모형 및 트리 기반 모형의 예측 성능을 비교한 연구와는 예측 과제가 다르다. 특히 중학교 3학년 청소년의 심리적·학업적·관계적 요인과 생활습관 및 사회경제적 특성을 포괄하여 행복감을 예측하고, 주요 변수의 상대적 중요도와 예측에 대한 기여 방향 및 비선형적 예측 양상을 함께 분석한 연구는 제한적이다.

머신러닝 모형은 변수와 결과 간 함수 형태를 사전에 구체적으로 설정하지 않고 비선형적 예측 양상을 탐색할 수 있다는 장점이 있다. 그러나 모형의 복잡성이 증가한다고 해서 새로운 자료에 대한 예측 성능이 반드시 향상되는 것은 아니므로, 자료의 특성을 고려하여 선형모형과 트리 기반 모형의 성능을 비교할 필요가 있다(James et al., 2021). 이에 본 연구에서는 평균예측모형, 다중선형회귀, 엘라스틱넷, 랜덤포레스트 및 XGBoost의 예측 성능을 비교하고자 한다. 또한 랜덤포레스트와 XGBoost의 변수 중요도 및 SHAP 분석을 통해 주요 변수의 상대적 중요도와 예측에 대한 기여 방향을 확인하고, 랜덤포레스트에서 중요도가 높게 나타난 변수를 대상으로 부분의존도(Partial Dependence Plot, PDP) 분석을 실시하여 행복감과의 비선형적 예측 양상을 탐색하고자 한다. 이러한 접근은 특정 모형의 우수성을 전제하기보다 서로 다른 모형의 예측 성능과 해석 결과를 비교함으로써 중학교 3학년 청소년의 행복감 관련 요인을 다각적으로 이해하고, 행복감 지원 영역을 설정하기 위한 기초자료를 제공할 수 있을 것이다.

2. 연구의 목적

본 연구의 목적은 중학교 3학년 청소년의 행복감 관련 요인을 머신러닝을 활용하여 파악하는 데 있다. 구체적인 목적은 다음과 같다. 첫째, 중학교 3학년 청소년의 행복감 예측을 위한 평균예측모형, 다중선형회귀, 엘라스틱넷, 랜덤포레스트 및 XGBoost 모형의 예측 성능을 비교한다. 둘째, SHAP 분석을 통해 변수의 상대적 중요도와 예측 기여 방향을 확인하고, PDP를 활용하여 주요 변수와 행복감 예측값 간의 비선형적 예측 양상을 탐색한다.

II. 문헌고찰

1. 청소년의 행복감

청소년 행복감은 청소년이 자신의 삶을 전반적으로 얼마나 긍정적이고 만족스럽게 인식하는지에 대한 주관적 평가를 의미하며, 기쁨과 만족과 같은 긍정정서뿐 아니라 삶 전반에 대한 인지적 평가를 포함한다(González-Carrasco et al., 2017; Steinmayr et al., 2019). 행복, 삶의 만족, 주관적 웰빙, 삶의 질은 개념적으로는 구분되지만 서로 밀접하게 관련된 개념으로 이해되며, 청소년 연구에서는 상호 보완적인 안녕(well-being) 지표로 활용되어 왔다(Diener et al., 2018; Steinmayr et al., 2019; Ross et al., 2020). 이들 연구에서 행복은 자신의 삶에 대한 주관적 판단에 기초한 주관적 웰빙으로 이해될 수 있으며, 단순한 쾌락적 즐거움을 넘어 자신의 가치와 신념에 맞게 살아가고 있다고 느끼는 의미와 성장의 차원까지 포괄하는 개념으로 확장되고 있다.

청소년기의 행복감은 학업, 대인관계, 정서조절, 건강행동 등 다양한 삶의 영역과 긴밀하게 연결되어 있다. OECD(2017)는 학생의 주관적 웰빙이 학업적 기능뿐 아니라 정서적·사회적 적응을 가능하게 하는 핵심 요소라고 보고하였다. 또한 청소년기 행복감은 연령이 증가할수록 감소하는 경향이 있으며, 이러한 감소는 여학생에게서 더욱 두드러진 것으로 나타났다(González-Carrasco et al., 2017). 이는 청소년 행복감이 단지 현재의 기분 상태가 아니라 발달과업 수행 및 정신건강과 밀접히 연결된 발달 지표임을 의미한다. Ross et al. (2020)은 청소년 웰빙을

건강, 연결성, 학습과 역량, 안전, 자아정체감과 목적의식, 주체성과 회복력 등이 조화를 이루는 상태로 설명하였다. 이러한 관점은 청소년의 행복감이 단일한 정서 반응이 아니라 건강, 관계, 교육, 환경, 자원 접근성 등이 통합된 결과임을 보여준다. 따라서 청소년 행복감을 이해하기 위해서는 개인 내 심리 요인뿐 아니라 가족, 학교, 사회문화적 맥락을 함께 고려하는 다차원적 접근이 필요하다.

2. 청소년 행복감의 관련 요인

초기 청소년 행복감 연구는 성별, 연령, 사회경제적 수준 등 인구사회학적 특성을 기본 분석 틀로 삼아 행복감의 차이를 확인하는 연구가 주를 이루었다(황환, 김영환, 2017; 박영신, 김의철, 2009; Crede et al., 2015). 이후 연구가 축적되면서 자아존중감, 자아탄력성, 우울과 같은 심리적 요인이 행복감과 밀접한 관련을 가지는 핵심 변인으로 제시되었으며, 행복감을 개인의 심리적 자원과 정신건강 측면에서 이해하려는 연구가 활발히 이루어졌다(서경현, 2016; 김선아, 2020; 송주빈, 신정택, 2020). 최근에는 청소년 행복감이 개인의 심리적 특성만으로 설명되기 어렵다는 관점이 확산되면서, 학업 경험, 또래 및 부모와의 관계, 학교환경, 수면과 신체활동, 디지털 미디어 사용 등 다양한 생활습관과 환경적 요인을 함께 고려하는 생태학적·다차원적 접근으로 연구가 확대되고 있다(라종민, 김소영, 2022; 김무현, 김준호, 2024; 이민영 외, 2025; Kleszczewska et al., 2018; Ross et al., 2020; Nie et al., 2025). 선행연구에서도 개인 요인뿐 아니라 가족환경 및 학교환경 요인이 청소년의 행복감과 관련되는 것으로 보고하였다(황환, 김영환, 2017; 이민영 외, 2025; Jang et al., 2025). 이처럼 청소년 행복감은 다양한 요인이 복합적으로 작용하는 결과로 이해되고 있어, 본 연구에서는 선행연구를 바탕으로 행복감 관련 변인을 인구사회학적 요인, 심리 요인, 학업 요인, 관계 요인 및 생활습관 요인으로 구분하여 고찰하였다.

청소년 행복감의 기초적 배경요인으로 성별, 연령, 사회경제적 수준 및 부모의 교육수준과 같은 인구사회학적 특성이 지속적으로 연구되어 왔다. 선행연구에 따르면 남학생이 여학생보다 행복감이 높고, 학년이 증가할수록 행복감은 감소하는 경향을 보였으며(신승배, 2018; 송주빈, 신정택, 2020; 유민상, 2025), 사회경제적 수준이 높을수록 삶의 질과 주관적 웰빙 수준이 높은 것으로 보고되었다(신승배, 2018; 유민상, 2025; Steinmayr et al., 2019). 부모의 교육수준과 행복감 또는 삶의 만족 간의 관계는 연구마다 다소 상반된 결과를 보였으나(라종민, 김소영, 2022; 김동심, 변종석, 2023; Crede et al., 2015), 부모의 사회경제적 수준은 부모의 주관적 웰빙과 양육태도 등을 통해 자녀의 웰빙과 관련되는 것으로 보고되었다(김명숙, 2016).

선행연구에서 청소년 행복감과 가장 밀접하게 관련된 영역은 심리 요인이었다. 자아존중감은 국내외 연구에서 행복감 및 삶의 만족의 가장 중요한 관련 요인으로 보고되었으며(송주빈, 신정택, 2020; 전해경, 2024; 최정현, 2025; Kleszczewska et al., 2018; Yang & Moon, 2024), 자아탄력성 역시 행복감과 주관적 웰빙을 증진하는 핵심 심리자원으로 확인되었다(서경현, 2016; Wu et al., 2025). 반면 우울은 행복감을 저해하는 대표적인 위험요인으로 보고되었으며(Yang & Moon, 2024), 자율성은 기본심리욕구 충족을 통해 행복감과 정적 관련성을 보이는 중요한 심리적 자원으로 제시되었다(윤영아, 최진아, 2025; Ryan et al., 2019).

학업은 청소년기의 핵심 발달과업으로, 학업 경험은 행복감과 심리적 안녕에 밀접하게 관련된다. 특히 학업스트레스가 높을수록 회복탄력성, 주관적 웰빙 및 행복감이 낮았으며(서경현, 2016), 학업스트레스는 친구나 외모스트레스보다 행복감과 더 밀접하게 관련되는 것으로 나타났다(김선아, 2020). 또한 학업스트레스는 학업 번아웃과 인터넷 중독을 통해 주관적 안녕과 부적으로 관련되는 것으로 보고되었다(Cheng & Lin, 2023). 반면 자기조절 학습, 학업성취 및 부모의 학업적·교육적 지원은 행복감을 증진하는 요인으로 제시되었다. 자기조절학습은 행복감 및 삶의 만족과 정적 관련을 보였으며(전해경, 2024), 학업성취와 성취에 대한 만족은 자아존중감과 삶의 만족을 향상시키고(Crede et al., 2015; Jang et al., 2025), 부모의 학업적 참여와 지원은 학업만족, 심리적 행복감 및

주관적 웰빙과 관련되는 것으로 보고되었다(강현권, 2015; Liu et al., 2024).

청소년 행복감은 또래 및 부모와의 관계를 비롯한 관계적 자원과 밀접하게 관련되는 것으로 보고되어 왔다. 또래관계의 질은 학교생활 적응과 관련되며(윤정희, 황혜정, 2022), 삶의 만족과 행복감을 향상시키는 핵심 요인으로 확인되었으며(강인경, 김춘경, 2016; 최정현, 2025), 자아존중감을 매개로 행복감에 영향을 미치는 것으로 나타났다(최정현, 2025). 부모애착 역시 삶의 만족과 심리적 안녕에 긍정적인 영향을 미치고(Tan et al., 2023), 정서조절을 매개로 삶의 만족에 간접적으로 기여하는 것으로 보고되었다(이현주, 2018). 또한 부모감독은 자율성과 미디어기기 사용 등을 통해 청소년의 정서적 적응 및 행복감과 관련되며(관다영, 박정운, 2024; 최유석, 임지영, 2023), 부모의 관심과 감독 수준이 높을수록 행복감이 증가하는 것으로 보고되었다(이주연, 한세영, 2021). 이러한 결과는 또래 및 부모와의 긍정적인 관계가 청소년 행복감의 중요한 관계적 자원임을 시사한다.

생활습관 역시 청소년 행복감의 중요한 관련 요인으로 보고되어 왔다. 스마트폰 과의존은 청소년의 건강, 수면 및 우울에 부정적인 영향을 미치며(김무현, 김준호, 2024), 과의존 수준이 높을수록 삶의 만족과 행복 수준이 낮아지는 것으로 나타났다(유민상, 2025). 또한 중국 대학생을 대상으로 한 연구에서는 스마트폰 의존과 주관적 웰빙의 관계에서 외로움의 매개역할이 보고되었다(Su & He, 2024). 수면 역시 중요한 생활습관 요인으로, 주말과 주중의 수면시간 차이가 클수록 삶의 만족은 감소하는 것으로 나타났다(Kim et al., 2022). 반면 신체활동은 주관적 행복감과 정적 관계를 보이며(목형균 외, 2016), 자아존중감, 행복감 및 삶의 만족과도 정적으로 관련되는 것으로 나타났다(신승아, 강민구, 2022), 메타분석에서도 삶의 만족과 유의한 정적 관련성이 보고되었다(Nie et al., 2025). 이와 같이 선행연구에서는 청소년의 인구사회학적, 심리, 학업, 관계, 생활습관 등 다양한 측면의 요인들이 청소년의 행복감과 심리적 안녕에 영향을 미치는 것으로 보고되어 왔으며, 이를 통합적으로 고려하여 행복감의 관련 요인에 대해 탐색할 필요가 있다.

3. 머신러닝 기반 청소년 행복감 연구 동향

최근 청소년의 행복감을 설명하기 위한 연구는 전통적인 회귀분석을 넘어 머신러닝 기반 예측모형을 활용하는 방향으로 확대되고 있다. 머신러닝은 변수 간 비선형 관계와 복잡한 상호작용을 효과적으로 반영할 수 있어, 다수의 심리·사회·환경 요인이 동시에 작용하는 청소년 행복감 연구에 적합한 방법으로 제시되고 있다(Kern et al., 2019; James et al., 2021). 실제로 랜덤포레스트를 활용하여 스트레스, 불안, 외로움 및 학업성적 등이 청소년 행복감 예측에서 높은 중요도를 보이는 것으로 확인되었다(Kim et al., 2025). 또한 한국아동패널 종단자료를 이용한 연구에서는 잠재계층성장분석을 통해 초기 청소년의 행복감 변화 궤적을 유형화하고, 다항 로지스틱 회귀, 랜덤포레스트, XGBoost 및 다층 퍼셉트론의 분류 성능을 비교하였다. 이 연구에서는 랜덤포레스트가 비교적 안정적인 성능을 보였으며, SHAP 분석을 통해 개인, 또래, 가족, 사회경제적 자원 및 미래인식 관련 변수가 행복감 변화 궤적 분류에 기여하는 것으로 확인되었다(송소연 외, 2026). 국외 연구에서도 자아존중감, 삶의 의미, 긍정적 감정, 학교 소속감 등 다양한 심리·사회적 요인이 행복감과 삶의 만족을 예측하는 주요 변수로 확인되었다(Zhang et al., 2019; Pan & Cutumisu, 2024).

이처럼 최근에는 여러 머신러닝 모형의 성능을 비교하고 SHAP을 활용하여 행복감 변화 궤적의 주요 예측변수를 해석하는 연구까지 확대되고 있다(송소연 외, 2026). 다만 행복감 변화 궤적을 범주형 결과로 예측한 분류연구와 달리, 특정 발달시점의 행복감을 연속형 결과변수로 설정하여 선형모형과 트리 기반 모형의 예측 성능을 비교하고, SHAP과 PDP를 함께 활용해 주요 변수의 예측 기여 방향과 비선형적 예측 양상을 검토한 연구는 제한적이다. 따라서 청소년 행복감을 보다 종합적으로 이해하기 위해서는 다양한 예측모형의 성능을 비교하고, 주요 변수의 상대적 중요도와 예측 패턴을 함께 해석하는 접근이 필요하다.

III. 연구 방법

1. 연구설계

본 연구는 2023년도 제16차 한국아동패널 II 자료를 이용하여 중학교 3학년 청소년의 행복감 관련 요인을 알아보기 위한 이차자료 분석 연구이다.

2. 연구대상 및 자료

본 연구에서 이용한 육아정책연구소의 한국아동패널 II 자료의 모집단은 2008년에 출생한 신생아 및 해당 신생아의 가구이며, 조사 모집단은 2006년 기준 연간 분만 건수 500건 이상인 전국 의료기관에서 2008년 4~7월 사이 출생한 신생아로 정의하였다(육아정책연구소, 2025). 표본은 층화다단계추출법을 적용하여 전국을 6개 권역으로 층화한 뒤, 건강보험심사평가원 의료기관 명부를 표본추출틀로 계통추출을 통해 30개 의료기관(1차 추출단위)을 선정하고, 해당 기관에서 조사 참여에 동의한 산모를 대상으로 신생아 가구 명부를 구축하여 신생아 가구(2차 추출단위)를 표집하였다. 이를 통해 예비표본 2,562가구, 원표본 2,078가구, 최종 전체 표본 2,150가구를 구축하였다.

본 연구에서는 출생 시부터 추적한 패널이 만 15세에 이른 2023년의 제16차 자료를 이용하였다. 16차 조사는 2023년 7월부터 12월까지 조사원 방문에 의한 아동과 보호자 대상의 태블릿을 이용한 대면면접조사, 어머니와 아버지 대상의 지필조사, 담임교사 대상의 온라인 조사를 통해 실시되었으며 총 1,258명의 아동이 참여하였다. 본 연구에서는 예측변수와 결과변수에 결측이 있는 사례를 제외하여 최종 1,154명을 분석 대상으로 포함하였다. 최종 분석 대상 아동의 성별은 남자 589명(51.0%), 여자 565명(49.0%)이었다. 어머니의 최종학력은 대학교(2~4년제) 졸업이 797명(69.1%)으로 가장 높은 비율을 차지하였으며, 고등학교 졸업 277명(24.0%), 대학원 졸업 77명(6.7%), 중학교 이하 3명(0.3%) 순이었다. 아버지의 최종학력은 대학교(2~4년제) 졸업이 717명(62.1%)으로 가장 많았으며, 고등학교 졸업 289명(25.0%), 대학원 졸업 145명(12.6%), 중학교 이하 3명(0.3%) 순이었다(표 1).

표 1. 연구대상자의 인구사회학적 특성

(N=1,154)

특성	구분	n	%
성별	남자	589	51.0
	여자	565	49.0
어머니 최종학력	중학교 이하	3	0.3
	고등학교	277	24.0
	대학교(2~4년제)	797	69.1
	대학원	77	6.7
아버지 최종학력	중학교 이하	3	0.3
	고등학교	289	25.0
	대학교(2~4년제)	717	62.1
	대학원	145	12.6

3. 연구도구

본 연구의 측정도구는 한국아동패널 II의 2023년도 제16차 조사에서 사용된 측정문항과 척도로 구성하였다. 각 도구의 출처와 문항 구성은 해당 조사 보고서에 제시된 내용을 바탕으로 기술하였다(조미라 외, 2023).

가. 결과변수: 전반적 행복감

전반적 행복감은 Millennium Cohort Study(MCS, 2008)의 아동 자기보고식 척도를 기반으로 측정하였다. 본 척도는 공부, 외모, 가족, 친구, 학교생활 및 삶 전반에 대한 만족을 측정하는 6개 문항으로 구성되어 있으며, 청소년의 삶의 다양한 영역에 대한 주관적 만족을 통합하여 전반적 행복감을 평가하도록 개발되었다. 본 연구에서는 문항들의 평균값을 전반적 행복감 점수로 사용하였다. 각 문항은 4점 Likert 척도(1점=전혀 행복하지 않다, 4점=매우 행복하다)로 측정되며, 점수가 높을수록 전반적 행복감 수준이 높은 것을 의미한다. 본 연구에서의 신뢰도는 Cronbach's $\alpha=.748$ 이었다.

나. 예측변수

1) 인구사회학적 요인

인구사회학적 요인은 성별, 주관적 사회경제적 지위, 가구의 경제적 걱정 빈도와 부모의 최종학력을 포함하였다. 성별은 남아(1점), 여아(2점)로 응답된 값을 분석 과정에서 남아 0점, 여아 1점으로 재코딩하여 사용하였다. 주관적 사회경제적 지위는 아동이 지각하는 가구의 사회경제적 수준을 평가하기 위해 사다리 모형 문항을 사용하여 측정하였다. 응답자는 대한민국 사회 구조를 나타내는 사다리에서 자신의 가구가 위치한다고 생각하는 수준을 표시하도록 하였으며, 응답은 1점(매우 못 산다)부터 10점(매우 잘 산다)까지의 10점 척도로 구성되었다. 점수가 높을수록 아동이 자신의 가구를 더 높은 사회경제적 지위에 있다고 인식함을 의미한다. 가구의 경제적 걱정 빈도는 가족이 돈 문제로 걱정하는 정도를 측정하는 단일 문항으로 조사하였다. 응답은 1점(전혀 걱정하지 않음)에서 4점(항상 걱정함)까지의 4점 척도로 구성되었으며, 점수가 높을수록 경제적 걱정 수준이 높음을 의미한다. 부모의 최종 학력은 무학 1점부터 대학원 졸업 7점까지의 범주형 응답을 사용하였다.

2) 심리 요인

심리 요인은 자아존중감, 자아탄력성, 자율성과 우울을 포함하였다. 자아존중감은 Rosenberg(1965)가 개발한 자아존중감 척도에서 일부 문항을 발췌하여 MCS(2012)에서 재구성한 5문항 척도를 사용하여 측정하였다. 각 문항은 4점 Likert 척도(1점=전혀 그렇지 않다, 4점=매우 그렇다)로 구성되며 점수가 높을수록 자아존중감 수준이 높은 것을 의미한다. 본 연구에서 도구의 신뢰도는 Cronbach's $\alpha=.889$ 이었다. 자아탄력성은 Block과 Kremen(1996)이 개발한 척도를 유성경과 심혜원(2002)이 번역·수정한 도구를 사용하였다. 총 14문항으로 구성되며 전혀 그렇지 않다(1점)부터 매우 그렇다(4점)까지의 4점 Likert 척도로 측정하였다. 점수가 높을수록 자아탄력성 수준이 높은 것을 의미하며 본 연구에서 도구의 신뢰도는 Cronbach's $\alpha=.838$ 이었다.

자율성은 The Children's Society(2012)에서 개발한 척도를 육아정책연구소 연구진이 번안한 5문항을 사용하여 측정하였다. '그렇지 않다(1점)부터 '그렇다(3점)까지의 3점 Likert 척도로 구성되며, '나는 내 생활 속에서 압박감에 시달린다' 문항은 역코딩하여 분석하였다. 점수가 높을수록 자율성 수준이 높은 것을 의미하며 본 연구에서

도구의 신뢰도는 Cronbach's $\alpha=.704$ 였다. 우울은 Weissman et al.(1980)의 CES-DC 척도를 기반으로 허만세 외(2017)가 표준화한 한국어판 CES-DC 11문항 척도를 사용하여 측정하였다. 각 문항은 지난 1주 동안의 정서 상태를 기준으로 '극히 드물다(일주일 동안 1일 이하)' 1점부터 '거의 대부분(5~7일)' 4점까지의 4점 척도로 응답하도록 구성되어 있으며 점수가 높을수록 우울 수준이 높은 것을 의미한다. '행복했다'와 '즐겁게 지냈다' 문항은 역코딩하여 분석하였다. 본 연구에서 도구의 신뢰도는 Cronbach's $\alpha=.843$ 이었다.

3) 학업 요인

학업 요인은 학업스트레스, 자기조절학습, 주관적 학업성취와 부모의 학업적·교육적 지원을 포함하였다. 학업스트레스는 한국청소년정책연구원의 한국청소년패널조사에서 사용된 척도를 활용하였다. 학교 성적, 숙제 및 시험, 공부, 대학입시 또는 취업 부담과 관련된 총 4문항으로 구성되며 1점(그런 적 없다)부터 5점(항상 그렇다)까지의 5점 Likert 척도로 측정하였다. 점수가 높을수록 학업스트레스 수준이 높음을 의미하며 본 연구에서 도구의 신뢰도는 Cronbach's $\alpha=.824$ 였다. 자기조절학습은 신현숙과 염시창(2017)의 중학생용 학교생활참여척도 단축형 중 자기조절학습 하위척도 5문항을 사용하였다. 각 문항은 1점(전혀 그렇지 않다)부터 4점(매우 그렇다)까지의 4점 Likert 척도로 구성되며 점수가 높을수록 자기조절학습 수준이 높은 것을 의미한다. 본 연구에서 도구의 신뢰도는 Cronbach's $\alpha=.813$ 이었다.

주관적 학업성취는 국어, 영어, 수학, 사회, 과학의 5개 교과목에 대한 아동의 자기보고 성취도를 활용하여 측정하였다. 각 교과목의 성취도는 A(1점)~E(5점) 등급으로 응답하도록 구성되어 있으며, 분석에서는 E등급을 1점, A등급을 5점으로 역코딩한 후 다섯 교과목 점수의 평균값을 산출하여 사용하였다. 점수가 높을수록 아동이 지각한 학업성취 수준이 높은 것을 의미하며, 본 연구에서 도구의 신뢰도는 Cronbach's $\alpha=.913$ 이었다. 부모의 학업적·교육적 지원은 한국교육종단연구2013 3차년도 조사 학생 설문지를 수정·보완한 도구를 사용하였다. 공부 분위기 조성, 숙제 확인, 공부방법 조언, 성적 관리 관심, 진로·진학 정보 수집, 일정 관리, 학부모 모임 및 정보 탐색 등을 보호자가 응답하도록 하는 총 7문항으로 구성되며 1점(전혀 그렇지 않다)부터 5점(매우 그렇다)까지의 5점 Likert 척도로 측정하였다. 점수가 높을수록 부모의 학업적·교육적 지원 수준이 높은 것을 의미하며 본 연구에서 도구의 신뢰도는 Cronbach's $\alpha=.878$ 이었다.

4) 관계 요인

관계 요인은 또래애착, 부애착, 모애착과 부모감독을 포함하였다. 또래애착은 Armsden과 Greenberg(1987)의 부모·또래 애착 척도(IPPA)에서 발췌한 9문항을 사용하여 측정하였다. 본 척도는 의사소통(3문항), 신뢰(3문항), 소외(3문항)의 세 하위영역으로 구성되며 4점 Likert 척도(1점=전혀 그렇지 않다, 4점=매우 그렇다)로 응답하도록 되어 있다. 애착의 부정적 측면을 나타내는 소외 3문항은 역코딩한 후 전체 문항의 평균값을 산출하였다. 점수가 높을수록 또래애착 수준이 높은 것을 의미하며 본 연구에서 도구의 신뢰도는 Cronbach's $\alpha=.827$ 이었다.

부모애착은 Armsden과 Greenberg(1987)의 부모·또래 애착 척도(IPPA)를 이정립 외(2017)가 변안하여 타당화한 도구를 사용하였다. 본 척도는 의사소통(6문항), 신뢰(3문항), 소외(3문항)의 세 하위영역으로 구성되어 있으며 아버지 애착과 어머니 애착을 각각 12문항으로 측정하였다. 각 문항은 1점(전혀 아니다)부터 5점(항상 그렇다)까지의 5점 Likert 척도로 응답하도록 되어 있으며 소외 3문항과 신뢰 1문항(나는 아버지/어머니에게 화가 난다.)은 역코딩하여 평균값을 산출하였다. 점수가 높을수록 부모애착 수준이 높은 것을 의미하며 본 연구에서 도구의 신뢰도는 부애착 Cronbach's $\alpha=.875$, 모애착 Cronbach's $\alpha=.862$ 로 나타났다. 부모감독은 허묘연(2004)의 부모 양육행동 척도 중 감독 하위요인 4문항을 사용하였다. 각 문항은 1점(전혀 아니다)부터 4점(매우 그렇다)까지의

4점 Likert 척도로 측정되며 점수가 높을수록 부모의 감독 수준이 높은 것을 의미한다. 본 연구에서 도구의 신뢰도는 Cronbach's $\alpha=.863$ 이었다.

5) 생활습관 요인

생활습관 요인은 스마트폰 과의존, 주말-주중 수면시간 차이와 신체활동을 포함하였다. 스마트폰 과의존은 신광우 외(2011)의 스마트폰 중독 자가진단 척도를 사용하였다. 총 15문항으로 구성되며 일상생활장애(5문항), 가상세계지향성(2문항), 금단(4문항), 내성(4문항)의 네 하위요인을 포함한다. 각 문항은 1점(전혀 그렇지 않다)부터 4점(매우 그렇다)까지의 4점 Likert 척도로 측정되며 점수가 높을수록 스마트폰 과의존 수준이 높은 것을 의미한다. 본 연구에서 도구의 신뢰도는 Cronbach's $\alpha=.726$ 이었다. 주말-주중 수면시간 차이는 주말 수면시간에서 주중 수면시간을 차감한 값으로 산출하였다. 값이 클수록 수면 패턴의 변동성이 큰 것을 의미한다. 신체활동은 최근 7일 동안 한 번에 30분 이상 격렬하거나 땀이 날 정도의 신체활동을 수행한 날의 수를 측정하는 단일 문항을 사용하였다. 응답은 '없음'부터 '5일 이상'까지의 6개 범주로 구성되며, 분석에서는 이를 0점에서 5점까지로 재코딩하였다. 점수가 높을수록 신체활동 빈도가 높은 것을 의미한다.

4. 자료 분석

본 연구는 중학교 3학년 청소년의 행복감 관련 요인을 파악하기 위하여 기초통계 분석과 머신러닝 기반 분석을 수행하였다. 기초통계 분석에서는 연구대상자의 일반적 특성과 연구변수의 분포를 확인하기 위해 빈도분석과 기술통계 분석을 실시하였으며, 왜도와 첨도를 통해 변수의 분포를 확인하였다. 각 척도의 내적 일관성은 Cronbach's α 를 산출하여 검토하였고, 주요 심리변수 간의 관련성을 확인하기 위해 Pearson 상관분석을 실시하였다. 기초통계 분석에는 SPSS Statistics 28.0을 이용하였다. 머신러닝 분석은 Python 3.13 기반의 Jupyter Notebook 환경에서 수행하였으며, pandas(2.3.3), NumPy(2.3.5), SciPy(1.16.3), pyreadstat(1.3.3), scikit-learn(1.7.2), XGBoost(2.1.3) 및 SHAP(0.50.0) 패키지를 이용하였다. 원자료는 pyreadstat을 이용하여 데이터프레임으로 변환한 후 변수 코딩과 역코딩을 실시하였다.

예측모형으로 평균예측모형, 다중선형회귀, 엘라스틱넷, 랜덤포레스트 및 XGBoost를 구축하였다. 평균예측모형은 결과변수의 평균을 예측값으로 반환하는 기준모형으로 활용하였다(James et al., 2021). 엘라스틱넷은 L1과 L2 정규화를 결합하여 모형의 복잡도를 조절하는 정규화 회귀모형이며(Zou & Hastie, 2005), 다중선형회귀와 엘라스틱넷에는 StandardScaler를 포함한 파이프라인을 적용하였다. 랜덤포레스트는 다수의 의사결정나무의 예측을 평균화하는 배깅 기반 앙상블 모형이고(Breiman, 2001), XGBoost는 이전 단계의 예측오차를 순차적으로 보완하는 부스팅 기반 앙상블 모형이다(Chen & Guestrin, 2016). 본 연구에서는 선형모형과 트리 기반 모형의 예측 성능을 비교하고, 두 트리 기반 모형의 변수 중요도 결과를 함께 검토하였다.

자료는 학습자료 80%(n=923)와 최종 검증자료 20%(n=231)로 분할하였다. 엘라스틱넷, 랜덤포레스트 및 XGBoost의 하이퍼파라미터는 학습자료에서 무작위 5-fold 교차검증과 평균 RMSE 최소화를 기준으로 RandomizedSearchCV를 이용하여 각각 30회, 40회 및 50회 탐색하였다. 최종 선정값은 엘라스틱넷의 $\alpha=.05995$ 와 $l1_ratio=.15$, 랜덤포레스트의 $n_estimators=200$, $max_depth=15$, $min_samples_split=6$, $min_samples_leaf=1$ 및 $max_features=.5$ 였다. XGBoost는 $n_estimators=200$, $max_depth=2$, $learning_rate=.08$, $subsample=.7$, $colsample_bytree=.8$, $min_child_weight=7$, $gamma=.01$, $reg_alpha=.5$ 및 $reg_lambda=10$ 으로 선정되었다. 자료분할, 하이퍼파라미터 탐색, 반복 교차검증 및 모형 학습에는 모두 random seed 42를 적용하였

다.

최종 선정된 하이퍼파라미터를 적용하여 학습자료에서 각 모형을 적합한 후, 별도로 분리해 둔 최종 검증자료에서 예측 성능을 평가하였다. 모형의 예측 성능은 결정계수(R^2), 평균제곱근오차(RMSE) 및 평균절대오차(MAE)를 이용하여 비교하였으며, 랜덤포레스트는 OOB(Out-of-Bag) R^2 를 추가로 산출하였다. 또한 최종 선정된 하이퍼파라미터를 고정한 상태에서 전체 분석 자료를 대상으로 5-fold 교차검증을 3회 반복하여 모형 성능의 변동성을 보조적으로 확인하였다. 변수 중요도는 랜덤포레스트의 불순도 기반 중요도와 XGBoost의 Gain 기반 중요도를 산출하였다. 또한 최종 검증자료에서 각 변수를 30회씩 무작위로 치환하는 순열 중요도(Permutation importance)를 R^2 기준으로 산출하여 두 모형의 중요도 결과가 일관되는지 확인하였다. SHAP 분석 역시 최종 검증자료를 대상으로 실시하여 각 변수의 예측에 대한 기여 방향과 상대적 중요도를 평가하였다. 마지막으로 랜덤포레스트의 평균 절대 SHAP 값이 높은 상위 10개 변수를 선정하고, 최종 검증자료에서 PDP를 산출하여 각 변수와 행복감 예측값 간의 관계 양상을 확인하였다.

IV. 연구 결과

1. 연구변수의 기술통계 및 주요 심리변수 간 상관관계

본 연구변수의 기술통계 분석 결과는 <표 2>와 같다. 종속변수인 행복감의 평균은 2.83 ± 0.45 점으로 중간 이상의 수준을 보였다. 심리 요인에서는 자아존중감(3.17 ± 0.55), 자아탄력성(2.82 ± 0.43), 자율성(2.60 ± 0.39)은 비교적 높은 수준이었으며, 우울(1.52 ± 0.42)은 비교적 낮은 수준으로 나타났다. 학업 요인에서는 학업스트레스(2.71 ± 0.90)와 자기조절학습(2.69 ± 0.56)은 중간 수준이었고, 주관적 학업성취(3.80 ± 1.05)와 부모의 학업적·교육적 지원(3.36 ± 0.68)은 비교적 높은 수준을 보였다. 또한 또래애착, 부모애착 및 부모감독은 전반적으로 긍정적인 수준이었으며, 스마트폰 과의존은 낮은 수준으로 나타났다. 대부분의 변수는 심각한 비정규성을 보이지 않았으나, 주말-주중 수면시간 차이의 첨도는 10.08로 다소 높게 나타났다. 그러나 본 연구에서 사용한 랜덤포레스트와 XGBoost는 비모수적 특성을 가지는 트리 기반 머신러닝 기법으로 자료의 정규성을 가정하지 않으며, 변수의 비선형성과 이상치에 비교적 강건한 특성을 갖는다(Kern et al., 2019). 따라서 해당 변수는 분석에서 제외하지 않고 포함하였다. 주요 심리변수 간 상관분석 결과<표 3>, 행복감은 자아존중감($r=.618$, $p<.001$), 자아탄력성($r=.476$, $p<.001$), 자율성($r=.389$, $p<.001$)과 유의한 정적 상관을 보였으며, 우울($r=-.493$, $p<.001$)과는 유의한 부적 상관을 나타냈다.

표 2. 연구 변수의 기술통계

						(N=1,154)
영역	변수명	평균	표준편차	왜도	첨도	
예측변수	인구사회학적 요인	주관적 사회경제적 지위	6.13	1.38	-0.01	0.67
		가구의 경제적 걱정 빈도	1.91	0.66	0.60	1.10
	심리 요인	자아존중감	3.17	0.55	-0.46	0.86
		자아탄력성	2.82	0.43	0.37	0.40
		자율성	2.60	0.39	-0.95	0.30
		우울	1.52	0.42	1.42	2.65
		학업스트레스	2.71	0.90	0.00	-0.31
	학업 요인	학업스트레스	2.71	0.90	0.00	-0.31

영역	변수명	평균	표준편차	왜도	첨도
관계 요인	자기조절학습	2.69	0.56	-0.42	0.91
	주관적 학업성취	3.80	1.05	-0.76	-0.19
	부모의 학업적·교육적 지원	3.36	0.68	-0.18	0.12
	또래애착	3.13	0.47	-0.34	0.74
	부애착	3.49	0.61	-0.20	0.35
	모애착	3.70	0.57	-0.01	-0.34
	부모감독	3.40	0.54	-0.66	0.39
생활습관 요인	스마트폰 과의존	1.99	0.35	0.17	-0.18
	주말·주중 수면시간 차이	1.47	1.46	1.00	10.08
	신체활동	2.55	1.73	-0.05	-1.19
결과변수	행복감	2.83	0.45	-0.17	0.44

표 3. 주요 심리변수 간 상관관계

(N=1,154)

변수명	자아존중감	자아탄력성	자율성	우울	행복감
자아존중감	1				
자아탄력성	.519***	1			
자율성	.457***	.277***	1		
우울	-.432***	-.212***	-.423***	1	
행복감	.618***	.476***	.389***	-.493***	1

*** $p<.001$.

2. 예측모형의 성능 비교

행복감 예측모형의 성능을 비교하기 위하여 평균예측모형, 다중선형회귀, 엘라스틱넷, 랜덤포레스트 및 XGBoost를 구축하였다<표 4>. 반복 5-fold 교차검증 결과, 엘라스틱넷의 평균 R^2 이 .556으로 가장 높았으며, 다중선형회귀(.553), XGBoost(.545), 랜덤포레스트(.537) 순으로 나타났다. 평균 RMSE는 엘라스틱넷(.296)이 가장 낮았고, 다중선형회귀(.297), XGBoost(.300), 랜덤포레스트(.303) 순이었다. 반면 최종 검증자료에서는 랜덤포레스트($R^2=.546$)와 XGBoost($R^2=.543$)가 엘라스틱넷($R^2=.527$)과 다중선형회귀($R^2=.529$)보다 다소 높은 예측 성능을 보였다. 랜덤포레스트는 최종 검증자료에서 가장 높은 R^2 (.546)과 가장 낮은 RMSE(.286) 및 MAE(.217)를 보였으며, OOB 결정계수는 .535로 나타났다. 즉, 교차검증과 최종 검증 간 모형 성능 순위는 일부 차이를 보였으나 전반적인 성능 차이는 크지 않았다.

표 4. 예측모형의 반복 교차검증 및 최종 검증 성능 비교

모형	반복 교차검증 R^2 (평균±SD)	반복 교차검증 RMSE (평균±SD)	최종 검증 R^2	최종 검증 RMSE	최종 검증 MAE	OOB R^2
평균예측모형	-.005±.007	.446±.021	-.013	.427	.325	-
다중선형회귀	.553±.033	.297±.019	.529	.291	.222	-
엘라스틱넷	.556±.031	.296±.019	.527	.292	.223	-
랜덤포레스트	.537±.024	.303±.018	.546	.286	.217	.535

모형	반복 교차검증 R^2 (평균±SD)	반복 교차검증 RMSE (평균±SD)	최종 검증 R^2	최종 검증 RMSE	최종 검증 MAE	OOB R^2
XGBoost	.545±.034	.300±.018	.543	.287	.217	-

3. 변수 중요도 분석

랜덤포레스트와 XGBoost의 변수 중요도를 비교한 결과는 <표 5>와 같다. 랜덤포레스트는 불순도 기반 변수 중요도를, XGBoost는 Gain 기반 변수 중요도를 산출하였으며, 변수 중요도의 안정성을 확인하기 위하여 최종 검증자료를 이용한 순열 중요도를 함께 제시하였다. 분석 결과, 자아존중감은 두 모형과 두 가지 중요도 산출 방법에서 모두 가장 높은 중요도를 보여 행복감을 예측하는 가장 핵심적인 변수로 확인되었다. 또한 우울, 또래애착 및 자아탄력성 역시 공통적으로 상위 변수로 나타났으며, 학업스트레스, 부애착, 모애착, 자기조절학습 및 스마트폰 과의존도 주요 예측변수로 확인되었다. 특히 순열 중요도에서도 상위 변수의 순위가 대체로 유사하게 나타나 변수 중요도의 안정성을 확인할 수 있었다.

표 5. 랜덤포레스트와 XGBoost의 변수 중요도 비교

변수	랜덤포레스트		XGBoost	
	불순도 기반 중요도	순열 중요도	Gain 기반 중요도	순열 중요도
자아존중감	.277	.215	.327	.184
우울	.125	.070	.123	.048
또래애착	.112	.045	.094	.049
자아탄력성	.094	.037	.092	.039
학업스트레스	.044	.019	.041	.026
부애착	.055	.015	.046	.013
모애착	.047	.022	.034	.020
자기조절학습	.034	.012	.030	.015
스마트폰 과의존	.047	.012	.027	.007
주관적 사회경제적 지위	.021	.002	.028	.002

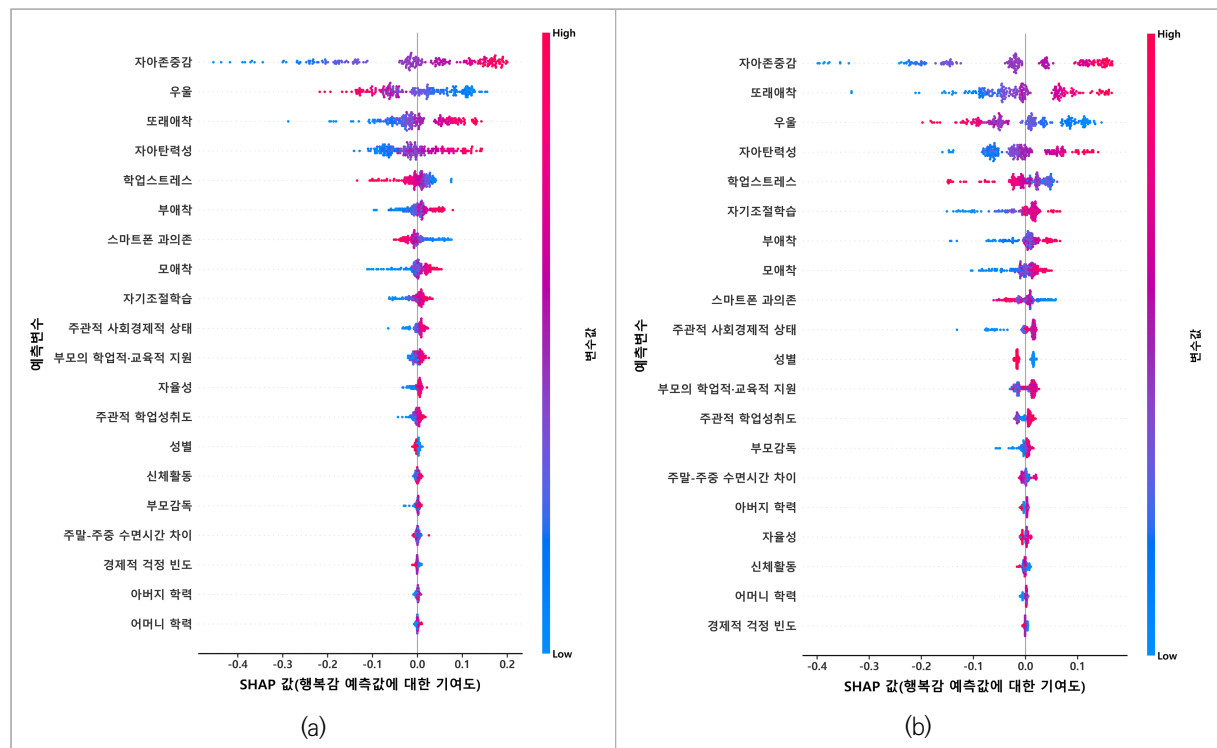
주: 상위 10개 변수는 두 모형의 변수 중요도, 순열 중요도 및 SHAP 중요도 등 총 6개 지표의 평균순위를 기준으로 선정하였다. 각 지표는 방법 내 상대적 순위를 중심으로 해석하였다.

4. SHAP 분석 결과

SHAP 분석 결과는 [그림 1]에 제시하였다. 랜덤포레스트와 XGBoost 모두에서 자아존중감이 행복감 예측모형에서 가장 높은 중요도를 보이는 변수로 나타났으며, 자아존중감이 높을수록 행복감 예측값을 증가시키는 방향으로 기여하였다. 우울과 또래애착은 두 모형에서 모두 상위 예측변수로 확인되었으나, 랜덤포레스트에서는 우울의 상대적 중요도가 다소 높았고, XGBoost에서는 또래애착의 상대적 중요도가 다소 높게 나타났다. 또한 자아탄력성은 두 모형 모두에서 높은 기여도를 보였으며, 값이 높을수록 행복감 예측값을 증가시키는 방향으로 기여하였다. 반면 학업스트레스와 스마트폰 과의존은 높은 수준일수록 행복감 예측값을 감소시키는 방향으로 나타났으며, 부애착, 모애착 및 자기조절학습은 행복감 예측값을 증가시키는 방향으로 기여하였다. 주관적 사회경제적 지위의 상대적 중요도는 낮았으나 두 모형에서 모두 행복감 예측값을 높이는 방향으로 기여하였다. 전반적으로 두 모형은

주요 변수의 예측 기여 방향과 상대적 중요도에서 유사한 결과를 나타냈으며, 변수 중요도 및 순열 중요도 분석에서 확인된 상위 변수와도 일관된 경향을 보여 행복감 예측모형에서 높은 중요도를 보이는 변수가 일관되게 확인되었다.

그림 1. 랜덤포레스트와 XGBoost 기반 SHAP 분석 결과

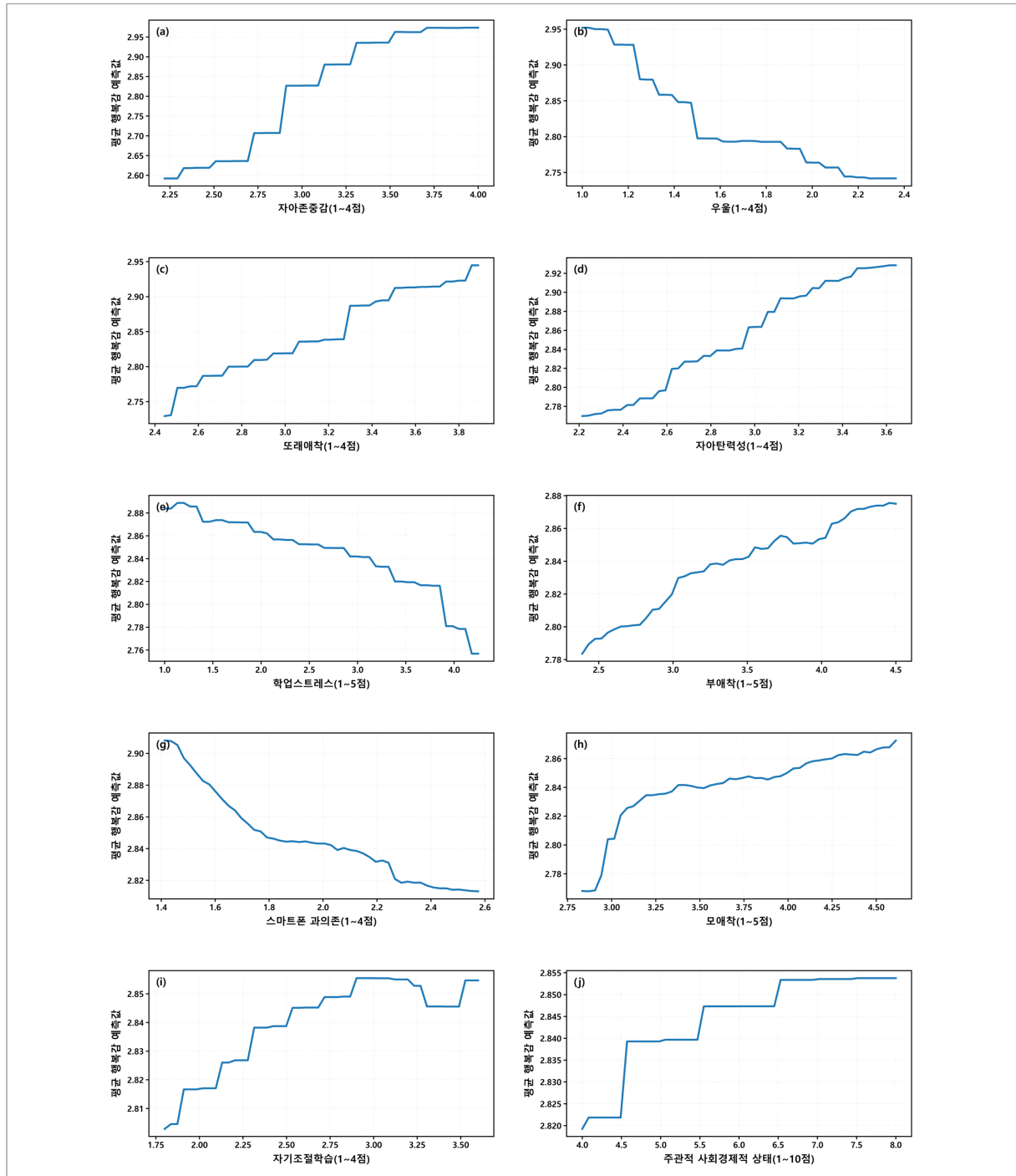


주: (a)는 랜덤포레스트, (b)는 XGBoost 결과이다. x축은 SHAP 값, y축은 평균 절대 SHAP 값이 큰 순으로 정렬된 예측변수를 나타낸다. 점의 색은 변수값의 크기를 의미한다(파란색=낮음, 붉은색=높음).

5. 랜덤포레스트 모형의 부분의존도 분석

랜덤포레스트 모형의 PDP 분석 결과는 [그림 2]에 제시하였다. 다른 변수들의 관측 분포에 대해 평균화했을 때 자아존중감과 자아탄력성이 높아질수록 행복감의 평균 예측값이 증가하였으며, 일부 구간에서 증가 기울기가 달라지는 비선형적 양상이 나타났다. 또래애착과 부모애착도 값이 증가할수록 행복감의 평균 예측값이 높아졌으며, 특히 또래애착의 변화폭이 상대적으로 크게 나타났다. 반면 우울, 학업스트레스 및 스마트폰 과의존은 값이 증가할수록 행복감의 평균 예측값이 낮아지는 양상을 보였다.

그림 2. 랜덤포레스트 모형의 행복감 주요 예측 변수별 부분의존도(PDP) 도표



주: 검증자료(n=231)를 이용하여 평균 절대 SHAP 값이 높은 상위 10개 변수의 부분의존도(PDP)를 산출하였다. x축은 각 변수의 값이며, y축은 다른 변수들의 관측 분포에 대해 평균화하여 산출한 행복감의 평균 예측값이다.

V. 결론 및 논의

본 연구는 2023년도 제16차 한국아동패널 II 자료를 활용하여 중학교 3학년 청소년의 행복감 예측모형을 비교하고, 변수 중요도, SHAP 및 PDP 분석을 통해 주요 관련 요인의 중요도와 예측 양상을 탐색하였다. 분석 결과, 평가방법에 따라 모형의 성능 순위는 달랐으나 전반적인 예측 성능은 유사하였으며, 트리 기반 모형에서는 자아존중감, 우울, 또래애착 및 자아탄력성이 공통적으로 높은 중요도를 나타냈다. 이를 바탕으로 모형의 예측 성능과 행복감의 주요 관련 요인을 중심으로 논의하고자 한다.

첫째, 행복감 예측모형의 성능을 비교한 결과, 반복 교차검증에서는 엘라스틱넷과 다중선형회귀가, 최종 검증자료에서는 랜덤포레스트와 XGBoost가 상대적으로 높은 성능을 보였다. 그러나 전반적으로 모형 간 R^2 와 RMSE의 차이는 크지 않았으며, 평가방법에 따라 성능 순위도 달랐다. 초기 청소년의 행복감 변화 궤적을 분류한 선행연구에서는 랜덤포레스트가 다른 모형보다 상대적으로 안정적인 분류 성능을 보였다(송소연 외, 2026). 본 연구에서도 랜덤포레스트가 최종 검증자료에서는 가장 높은 성능을 보였으나, 이러한 우위가 반복 교차검증에서는 일관되게 나타나지 않았다. 다만 두 연구는 결과변수의 형태, 예측 과제, 변수 구성 및 평가 지표가 서로 달라 직접 비교하는데 한계가 있으며, 모형의 상대적 성능이 분석 맥락에 따라 달라질 수 있음을 시사한다. 또한 본 연구 결과는 모형의 복잡성이 증가한다고 해서 예측 성능이 반드시 향상되는 것은 아니며, 모형의 성능이 자료의 특성과 편향-분산 간 균형에 따라 달라질 수 있다는 통계학적 논의와 부합한다(James et al., 2021).

둘째, 본 연구에서 자아존중감은 랜덤포레스트와 XGBoost 모두에서 가장 높은 변수 중요도를 나타냈으며, 순열 중요도와 SHAP 분석에서도 일관되게 가장 높은 중요도를 보여 행복감 예측모형에서 가장 안정적으로 확인된 변수였다. 이러한 결과는 자아존중감이 청소년 행복감과 밀접하게 관련된다는 선행연구(송주빈, 신정택, 2020; 전해경, 2024; Yang & Moon, 2024)와 일치한다. 특히 기존 회귀분석 연구에서도 자아존중감은 행복감을 설명하는 가장 중요한 변수 중 하나로 보고되어 왔으며, 본 연구에서도 분석 방법이 달라졌음에도 유사한 결과가 확인되었다는 점에서 의의가 있다. 또한 자아탄력성과 우울 역시 두 모형에서 공통적으로 높은 중요도를 나타냈다. 이는 자아탄력성이 행복감과 정적인 관련을 보이고(서경현, 2016; Wu et al., 2025), 우울이 행복감과 부적 상관관계를 보인다는 선행연구(전해경, 2024; Yang & Moon, 2024)와 일치한다. 이는 청소년 행복감이 긍정적 심리자원과 부정적 정서가 함께 작용하는 복합적인 특성을 가짐을 시사하며, 청소년 행복감을 이해할 때 자아존중감과 자아탄력성 같은 긍정적 심리자원뿐 아니라 우울과 같은 부정적 정서를 함께 고려할 필요가 있음을 보여준다.

셋째, 학업요인에서는 학업스트레스가 높은 중요도를 나타냈으며, 자기조절학습은 행복감과 긍정적인 관련성을 보였다. 이러한 결과는 학업스트레스가 청소년의 행복감을 저해하고(김선아, 2020), 자기조절학습이 행복감과 긍정적인 상관성을 보인다는 선행연구(전해경, 2024)를 지지한다. 이는 청소년 행복감을 이해할 때 단순한 학업성취보다 경쟁적 학업환경에서 경험하는 심리적 부담과 자기조절 역량을 함께 고려할 필요가 있음을 시사한다.

넷째, 관계요인에서는 또래애착이 부모 관련 변수보다 상대적으로 높은 중요도를 나타냈다. 이는 또래애착이 삶의 만족과 행복감에 밀접하게 관련된다는 선행연구(강인경, 김춘경, 2016; 최정현, 2025)를 지지하는 결과이다. 특히 중학교 3학년은 또래가 중요한 정서적 참조집단으로 기능하는 시기라는 점에서, 또래애착이 높은 중요도를 보인 결과는 발달적 특성과도 부합한다. 반면 부애착과 모애착은 또래애착보다 상대적 중요도가 낮았으나, 두 트리 기반 모형에서 모두 행복감의 주요 관계요인으로 확인되었다. 이는 부모와의 애착이 청소년 삶의 만족 및 정신건강과 관련된다는 선행연구(이현주, 2018; Tan et al., 2023)와 맥을 같이하며, 중학교 3학년 시기에도 부모 애착이 행복감과 관련된 관계적 자원임을 보여준다.

다섯째, 생활습관 요인에서는 스마트폰 과의존이 두 트리 기반 모형에서 모두 비교적 높은 중요도를 나타냈다. SHAP 분석에서도 스마트폰 과의존 수준이 높을수록 행복감 예측값이 낮아지는 양상이 확인되었다. 이는 스마트폰 과의존이 행복감 및 주관적 웰빙과 부적으로 관련된다는 선행연구(유민상, 2025; Su & He, 2024)와 맥을

같이한다. 이러한 결과는 스마트폰 과의존이 청소년 행복감 예측에서 부정적인 기여 방향을 보인 주요 생활습관 변수임을 보여준다. 따라서 청소년 행복감 지원방안을 마련할 때 단순한 사용시간뿐 아니라 스마트폰의 사용 목적과 양상 및 정서적 안녕과의 관련성을 함께 고려할 필요가 있음을 시사한다.

여섯째, 인구사회학적 요인 중 주관적 사회경제적 지위는 6개 중요도 지표를 통합한 평균순위에서 상위 10개 변수에 포함되었다. 특히 XGBoost의 Gain 중요도와 두 모형의 SHAP 중요도에서는 상위 10개에 포함되었으나, 순열 중요도는 두 모형에서 모두 낮게 나타났다. 이는 청소년이 인식하는 사회경제적 상태가 행복감 예측과 관련 될 가능성을 보여주지만, 다른 예측변수들을 함께 고려했을 때 독립적인 예측 기여도는 제한적일 수 있음을 시사한다. 반면 성별, 부모학력 및 경제적 걱정 빈도의 변수 중요도는 전반적으로 낮았다. 이는 이러한 특성이 행복감과 관련되지 않는다는 의미라기보다, 본 연구에 포함된 심리적·관계적·학업적 요인 및 생활습관 요인을 함께 고려했을 때 상대적 예측 기여도가 낮았기 때문으로 해석할 수 있다.

두 트리 기반 모형의 SHAP 분석에서는 주요 변수의 예측 기여 방향이 대체로 유사하게 나타났다. 한편 랜덤포레스트를 대상으로 실시한 PDP 분석에서는 자아존중감, 또래애착, 학업스트레스 및 스마트폰 과의존 등 일부 주요 변수와 행복감 예측값 간에 비선형적 양상이 확인되었다. 다만 PDP에서 나타난 예측값의 변화폭은 인과적 효과나 실제 중재 효과의 크기를 의미하지 않는다. 이러한 결과는 설명가능한 머신러닝 기법을 통해 변수의 상대적 중요도뿐 아니라 예측에 대한 기여 방향과 비선형적 예측 양상도 함께 탐색할 수 있음을 보여준다. 종합하면, 중학교 3학년 청소년의 행복감은 심리적·관계적·학업적 요인과 생활습관 등 다양한 영역의 요인과 관련되어 있었다. 따라서 청소년 행복감을 이해하고 지원 방안을 마련할 때 자아존중감과 자아탄력성, 우울, 또래 및 부모와의 관계, 학업스트레스와 자기조절학습, 스마트폰 사용 습관을 통합적으로 고려할 필요가 있다.

한편 최종 검증자료에서 랜덤포레스트와 XGBoost의 R^2 는 약 .54로 나타났다. 이는 본 연구에 포함된 요인들이 행복감 예측에 상당한 정보를 제공하였으나, 이들만으로 행복감의 변동을 모두 예측하기는 어렵다는 점을 보여준다. 선행연구에서도 주관적 안녕감은 개인의 심리적 특성뿐 아니라 사회적 관계, 경제적 조건 및 발달적·환경적 맥락과 관련된 다차원적 개념으로 설명되고 있다(Diener et al., 2018; Steinmayr et al., 2019; Ross et al., 2020). 따라서 모형에서 충분히 예측되지 않은 변동에는 본 연구에 포함되지 않은 개인적·환경적 요인이 관련되어 있을 가능성이 있다. 향후 연구에서는 학교풍토, 교사지원, 지역사회 환경 및 성격 특성 등 다양한 변인을 추가하여 모형의 예측 성능과 해석 가능성을 검증할 필요가 있다.

본 연구는 다음과 같은 제한점이 있다. 첫째, 본 연구는 특정 시점의 자료를 활용한 단면 분석이므로 변수 간 인과관계를 확인할 수 없다. 따라서 향후 연구에서는 종단자료를 활용하여 변수 간 시간적 선후관계와 변화 양상을 검토하고, 필요한 경우 인과추론 설계를 병행할 필요가 있다. 둘째, 본 연구에서 활용한 SHAP과 PDP는 머신러닝 예측모형의 해석을 위한 기법으로, 변수의 예측 기여도와 예측값 변화 양상을 설명할 뿐 인과적 효과를 의미하지 않는다. 따라서 향후 연구에서는 구조방정식모형이나 인과추론 기법을 병행하여 변수 간 관계를 보다 종합적으로 검증할 필요가 있다. 셋째, 본 연구에서는 랜덤포레스트의 불순도 기반 변수 중요도와 XGBoost의 Gain 중요도뿐 아니라 순열 중요도 및 SHAP 분석을 함께 활용하여 변수 중요도의 안정성을 확인하였다. 그러나 트리 기반 변수 중요도는 변수의 측정 특성에 영향을 받을 가능성을 완전히 배제할 수 없으므로, 향후 연구에서는 다양한 변수 중요도 지표를 활용한 추가 검증이 필요하다. 넷째, 본 연구의 PDP 분석에서는 주요 변수의 비선형적 예측 패턴을 확인하였으나, 예측값의 변화폭이 실제 중재 효과나 정책적 효과의 크기를 의미하는 것은 아니다. 따라서 향후 연구에서는 효과크기와 실제 적용 가능성을 함께 고려한 후속 연구가 필요하다. 다섯째, 본 연구는 머신러닝 기반 예측모형의 성능 비교를 목적으로 수행되었으므로 복합표본설계에 따른 표본가중치를 적용하지 않았다. 따라서 본 연구 결과는 한국아동패널 II 제16차 조사에 참여한 중학교 3학년 청소년을 대상으로 한 예측 결과로 해석할 필요가 있으며, 전체 중학교 3학년 청소년으로 일반화할 때에는 신중한 해석이 요구된다.

이상미는 동양대학교 간호대학 교수이자 학생건강관리센터 센터장으로 재직하고 있다. 주요 관심 분야는 소아청소년의 건강증진과 정신건강으로, 과체중 및 비만 소아청소년의 대사건강, 청소년의 심리 정서 및 행동발달, 다문화 청소년의 적응과 우울, 조산아의 성장발달 등 취약집단 소아청소년의 건강 관련 요인을 중심으로 연구를 수행하고 있다.

(E-mail: smlee7839@gmail.com)

참고문헌

- 강인경, 김춘경. (2016). 청소년의 또래애착이 삶의 만족도를 매개로 자기조절학습능력에 미치는 영향에 대한 종단연구. *교육심리연구*, 30(2), 411-428. <https://doi.org/10.17286/KJEP.2016.30.2.06>
- 강현권. (2015). 청소년의 부모지지와 자아실현욕구가 학업만족 및 심리적 행복감에 미치는 영향. *청소년문화포럼*, 43, 7-30. <https://doi.org/10.17854/ffyc.2015.07.43.7>
- 관다영, 박정윤. (2024). 청소년이 지각한 부모감독이 자녀의 행복감에 미치는 영향 -자율성의 매개효과 검증-. *아동과 권리*, 28(4), 547-564. <https://doi.org/10.21459/kccr.2024.28.4.547>
- 김동심, 변종석. (2023). 청소년기는 초기 청년기의 행복에 영향을 미치는가? *조사연구*, 24(4), 1-25. <https://doi.org/10.20997/SR.24.4.1>
- 김명숙. (2016). 부모의 사회경제적 수준이 부모의 주관적 웰빙과 자율성 지지 양육의 매개효과를 통해 자녀의 웰빙에 미치는 영향. *교육과정평가연구*, 19(1), 275-301.
- 김무현, 김준호. (2024). 청소년의 성별에 따른 스마트폰 의존이 건강, 수면, 우울에 미치는 영향에 대한 종단연구. *한국아동·청소년패널조사2018(KCYPs 2018)을 중심으로 한국응급구조학회지*, 28(1), 21-34. <https://doi.org/10.14408/KJEMS.2024.28.1.021>
- 김선아. (2020). 남녀청소년의 다양한 스트레스(부모, 학업, 친구, 외모, 경제), 학교생활적응과 행복감의 관계성 연구: 매개관계를 중심으로. *Journal of the Korea Academia-Industrial Cooperation Society*, 21(9), 161-169. <https://doi.org/10.5762/KAIS.2020.21.9.161>
- 라종민, 김소영. (2022). 개인요인, 가정요인, 학교요인이 청소년의 행복에 미치는 영향. *교육과 학습: Global Creative Leader*, 12(4), 57-76. <https://doi.org/10.34226/gcl.2022.12.4.57>
- 목형균, 조규희, 이승철, 이준협. (2016). 청소년의 신체활동과 주관적 행복감 간의 관련성. *한국학교·지역보건교육학회지*, 17(1), 101-112.
- 박영신, 김의철. (2009). 한국 청소년의 행복: 심리적, 관계적, 경제적 자원과 학업성취의 영향. *한국심리학회지: 문화 및 사회문제*, 15(3), 399-429.
- 박지선, 김재홍, 김현정, 정은주. (2019). 중·고등학교 전환기 학생 지원 방안 연구(Ⅰ): 교수·학습 특성 및 지원 실태 분석 (RRI 2019-3). 한국교육과정평가원.
- 서경현. (2016). 청소년의 학업스트레스와 주관적 웰빙 및 행복감 간의 관계: 회복탄력성의 역할을 중심으로. *청소년학연구*, 23(11), 137-157. <https://doi.org/10.21509/KJYS.2016.11.23.11.137>
- 송소연, 방윤석, 최지영. (2026). 초기 청소년의 행복감 변화 궤적에 대한 예측요인 탐색: 잠재계층성장분석과 머신러닝을 활용하여. *보건사회연구*, 46(2), 329-349. <https://doi.org/10.15709/hswr.2026.46.2.329>
- 송주빈, 신정택. (2020). 중학생의 스트레스, 자아존중감 및 행복감 간의 관계. *교육발전*, 40(2), 429-447. <https://doi.org/10.34245/jed.40.2.429>
- 신광우, 김동일, 정여주. (2011). 스마트폰 중독 진단척도 개발 연구. 서울: 한국정보화진흥원.
- 신승배. (2018). 한국 청소년 삶의 질 결정요인. *사회과학연구*, 29(1), 195-217. <https://doi.org/10.16881/jss.2018.01.29.1.195>
- 신승아, 강민구. (2022). 중학생의 신체활동 수준에 따른 자아존중감, 행복감 및 삶의 만족도의 차이 검증. *한국스포츠학회지*, 20(2), 159-167. <https://doi.org/10.46669/kss.2022.20.2.014>
- 신현숙, 염시창. (2017). 중학생용 학교생활참여척도(K-SES-M) 단축형의 개발 및 타당화. *한국심리학회지: 학교*, 14(2), 149-180. <https://doi.org/10.16983/kjsp.2017.14.2.149>
- 심수진, 김은아. (2025). 아동·청소년 삶의 질 2025. 국가데이터연구원.
- 유민상. (2025). 한국 아동의 행복 수준과 격차: 2023 아동종합실태조사의 주관적 웰빙 결과를 중심으로. *보건복지포럼*, (343), 66-84.

- 유성경, 심혜원. (2002). 적응 유연한 청소년들의 심리적 보호요소 탐색. *교육심리연구*, 16(4), 189-206.
- 육아정책연구소. (2025). *한국아동패널 제8차-16차 조사 데이터 사용자 지침서*. 육아정책연구소.
- 윤영아, 최진아. (2025). 청소년이 지각한 모 애착이 행복감에 미치는 영향: 자율성 충족과 신체상의 순차적 매개효과. *놀이치료연구*, 29(3), 23-38. <https://doi.org/10.32821/JPT.29.3.23>
- 윤정희, 황혜정. (2022). 중학생의 정서조절능력과 학교생활적응 간의 관계에서 또래관계 질의 매개효과. *학습자중심교과교육연구*, 22(18), 711-725. <https://doi.org/10.22251/jlcci.2022.22.18.711>
- 이민영, 임혜숙, 유진나. (2025). 한국 청소년의 주관적 행복감 국내 연구 동향분석: 주제범위 문헌고찰. *학습자중심교과교육연구*, 25(1), 1-12. <https://doi.org/10.22251/jlcci.2025.25.1.1>
- 이정림, 김지현, 이규림, 조혜주, 임종한, Carlo, G. (2017). 한국 어린이 성장발달 코호트 모형 구축 및 타당화 연구 (연구보고 2017-21). 육아정책연구소
- 이주연, 한세영. (2021). 부와 모의 자녀에 대한 관심감독 및 청소년의 성에 따른 청소년 자녀의 행복감. *한국생활과학회지*, 30(1), 63-77. <https://doi.org/10.5934/kjhe.2021.30.1.63>
- 이현주. (2018). 청소년의 부모애착, 정서조절과 삶 만족도 간의 종단적 관계: 청소년 초·중기 패널의 비교와 중학교 2학년에 대한 코호트 효과. *교육심리연구*, 32(4), 649-670. <https://doi.org/10.17286/KJEP.2018.32.4.03>
- 전혜경. (2024). 자기조절학습, 자아존중감, 자아탄력성, 전반적 행복감 및 우울이 중학교 1학년의 삶의 만족도에 미치는 영향. *아동부모학회지*, 10(3), 39-62. <https://doi.org/10.23079/2024.10.03-02>
- 조미라, 최효미, 김동훈, 김혜진, 강민권. (2023). *한국아동 성장발달 종단연구 2023(한국아동패널 II)*. 육아정책연구소
- 최유석, 임지영. (2023). 부모감독, 초등학교의 미디어기기 과다사용과 행복감 간의 종단적 관계. *가정과삶의질연구*, 41(3), 31-44. <https://doi.org/10.7466/JFBL.2023.41.3.31>
- 최정현. (2025). 청소년이 지각한 또래애착, 부모의 성취압력, 자아존중감, 행복감 간의 관계. *디지털정책학회지*, 4(1), 85-95. <https://doi.org/10.23149/JDP.2025.4.1.085>
- 허만세, 이순희, 김영숙. (2017). 한국어판 CES-DC 11의 개발: 신뢰도 및 타당도 분석을 중심으로. *정신건강과 사회복지*, 45(1), 255-285. <https://doi.org/10.24301/MHSW.2017.03.45.1.255>
- 허묘연. (2004). 청소년이 지각한 부모 양육행동 척도의 개발 및 타당화 연구. *청소년상담연구*, 12(2), 170-189.
- 황환, 김영환. (2017). 국내 청소년 행복의 영향요인 연구에 대한 체계적 문헌고찰. *청소년학연구*, 24(7), 1-28. <https://doi.org/10.21509/KJYS.2017.06.24.7.1>
- Armsden, G. C., & Greenberg, M. T. (1987). The inventory of parent and peer attachment: Individual differences and their relationship to psychological well-being in adolescence. *Journal of Youth and Adolescence*, 16(5), 427-454. <https://doi.org/10.1007/BF02202939>
- Block, J., & Kremen, A. M. (1996). IQ and ego-resiliency: Conceptual and empirical connections and separateness. *Journal of Personality and Social Psychology*, 70(2), 349-361. <https://doi.org/10.1037/0022-3514.70.2.349>
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5-32. <https://doi.org/10.1023/A:1010933404324>
- Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785-794). Association for Computing Machinery. <https://doi.org/10.1145/2939672.2939785>
- Cheng, X., & Lin, H. (2023). Mechanisms from academic stress to subjective well-being of Chinese adolescents: The roles of academic burnout and internet addiction. *Psychology Research and Behavior Management*, 16, 4183-4196. <https://doi.org/10.2147/PRBM.S423336>
- Crede, J., Wirthwein, L., McElvany, N., & Steinmayr, R. (2015). Adolescents' academic achievement and life satisfaction: The role of parents' education. *Frontiers in Psychology*, 6, 52. <https://doi.org/10.3389/fpsyg.2015.00052>
- Diener, E., Oishi, S., & Tay, L. (2018). Advances in subjective well-being research. *Nature Human Behaviour*, 2(4), 253-

260. <https://doi.org/10.1038/s41562-018-0307-6>
- González-Carrasco, M., Casas, F., Malo, S., Viñas, F., & Dinisman, T. (2017). Changes with age in subjective well-being through the adolescent years: Differences by gender. *Journal of Happiness Studies*, 18(1), 63–88. <https://doi.org/10.1007/s10902-016-9717-1>
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021). *An introduction to statistical learning: With applications in R* (2nd ed.). Springer. <https://doi.org/10.1007/978-1-0716-1418-1>
- Jang, H., Park, G. R., & Kim, J. (2025). Adolescents' satisfaction with academic achievement and the development of psychological resources: The moderating role of parental satisfaction. *School Psychology International*, 46(4), 339–360. <https://doi.org/10.1177/01430343251320282>
- Kern, C., Klausch, T., & Kreuter, F. (2019). Tree-based machine learning methods for survey research. *Survey Research Methods*, 13(1), 73–93. <https://doi.org/10.18148/srm/2019.v1i1.7395>
- Kim, E. J., Kim, S. K., Jung, S. H., & Ryu, Y. S. (2025). Predictive factors of adolescents' happiness: A random forest analysis of the 2023 Korea Youth Risk Behavior Survey. *Child Health Nursing Research*, 31(2), 85–95. <https://doi.org/10.4094/chnr.2024.049>
- Kim, J., Noh, J. W., Kim, A., & Kwon, Y. D. (2022). The impact of weekday-to-weekend sleep differences on health outcomes among adolescent students. *Children*, 9(1), 52. <https://doi.org/10.3390/children9010052>
- Kleszczewska, D., Dzielska, A., Salonna, F., & Mazur, J. (2018). The association between physical activity and general life satisfaction in lower secondary school students: The role of individual and family factors. *Community Mental Health Journal*, 54(8), 1245–1252. <https://doi.org/10.1007/s10597-018-0309-x>
- Liu, X., Fan, H., Shang, X., Li, W., He, X., Cao, P., & Ding, X. (2024). Parental involvement and children's subjective well-being: Mediating roles of the sense of security and autonomous motivation in Chinese primary school students. *Behavioral Sciences*, 14(7), 603. <https://doi.org/10.3390/bs14070603>
- MCS. (2008). *MCS Age 7 Child Self-Completion Form*. UCL Social Research Institute. https://cls.ucl.ac.uk/data_documentation/mcs-age-7-child-self-completion-form/
- MCS. (2012). *MCS5 child self-completion questionnaire*. https://cls.ucl.ac.uk/wp-content/uploads/2017/07/MCS5_MS_Child-Self-Completion-Questionnaire_CORE_ESNI_FINAL_PRINT.pdf
- Nie, Y., Wang, W., Guo, M., Tang, W., Qu, G., Wang, X., Zhou, F., Ma, W., Qiu, W., Gao, J., & Liu, C. (2025). Analysis of the association between adolescent physical activity and life satisfaction: a systematic review and meta-analysis. *BMC Psychology*, 13(1), 738. <https://doi.org/10.1186/s40359-025-02847-1>
- OECD. (2017). *PISA 2015 Results (Volume III): Students' well-being*. PISA, OECD Publishing. <https://doi.org/10.1787/9789264273856-en>
- Pan, Z., & Cutumisu, M. (2024). Using machine learning to predict UK and Japanese secondary students' life satisfaction in PISA 2018. *The British Journal of Educational Psychology*, 94(2), 474–498. <https://doi.org/10.1111/bjep.12657>
- Rosenberg, M. (1965). *Society and the Adolescent Self-Image*. Princeton University Press.
- Ross, D. A., Hinton, R., Melles-Brewer, M., Engel, D., Zeck, W., Fagan, L., Herat, J., Phaladi, G., Imbago-Jácome, D., Anyona, P., Sanchez, A., Damji, N., Terki, F., Baltag, V., Patton, G., Silverman, A., Fogstad, H., Banerjee, A., & Mohan, A. (2020). Adolescent well-being: A definition and conceptual framework. *Journal of Adolescent Health*, 67(4), 472–476. <https://doi.org/10.1016/j.jadohealth.2020.06.042>
- Ryan, R. M., Ryan, W. S., Di Domenico, S. I., & Deci, E. L. (2019). The nature and the conditions of human autonomy and flourishing: Self-determination theory and basic psychological needs. In R. M. Ryan (Ed.), *The Oxford handbook of human motivation* (2nd ed., pp. 89–110). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780190666>

453.013.6

- Steinmayr, R., Wirthwein, L., Modler, L., & Barry, M. M. (2019). Development of subjective well-being in adolescence. *International Journal of Environmental Research and Public Health*, 16(19), 3690. <https://doi.org/10.3390/ijerph16193690>
- Su, P., & He, M. (2024). The mediating role of loneliness in the relationship between smartphone addiction and subjective well-being. *Scientific Reports*, 14(1), 4460. <https://doi.org/10.1038/s41598-024-54546-3>
- Tan, R., Yang, Y., Huang, T., Lin, X., & Gao, H. (2023). Parent-child attachment and mental health in young adolescents: a moderated mediation analysis. *Frontiers in Psychology*, 14, 1298485. <https://doi.org/10.3389/fpsyg.2023.1298485>
- The Children's Society. (2012). *The good childhood report 2012: A review of our children's well-being*. The Children's Society. <https://www.childrenssociety.org.uk/sites/default/files/2023-08/GCR%202012.pdf>
- Weissman, M. M., Orvaschel, H., & Padian, N. (1980). Children's symptom and social functioning self-report scales. Comparison of mothers' and children's reports. *The Journal of Nervous and Mental Disease*, 168(12), 736-740. <https://doi.org/10.1097/00005053-198012000-00005>
- Wu, J., Zhang, Z., Chai, T., Shen, Y., & Yu, X. (2025). Adolescent psychological resilience and subjective well-being: A meta-analysis. *Journal of Psychology in Africa*, 35(6), 771-790. <https://doi.org/10.32604/jpa.2025.067273>
- Yang, J.-E., & Moon, S. (2024). Predictors of adolescents' subjective happiness based on Roy's adaptation model: A longitudinal study. *The Open Public Health Journal*, 17, e18749445352096. <https://doi.org/10.2174/0118749445352096241122113615>
- Zhang, N., Liu, C., Chen, Z., An, L., Ren, D., Yuan, F., Yuan, R., Ji, L., Bi, Y., Guo, Z., Ma, G., Xu, F., Yang, F., Zhu, L., Robert, G., Xu, Y., He, L., Bai, B., Yu, T., & He, G. (2019). Prediction of adolescent subjective well-being: A machine learning approach. *General Psychiatry*, 32(5), e100096. <https://doi.org/10.1136/gpsych-2019-100096>
- Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 67(2), 301-320. <https://doi.org/10.1111/j.1467-9868.2005.00503.x>

Factors Associated with Happiness among Korean Ninth-Grade Students: A Comparative Machine Learning Analysis

Lee, Sangmi¹

¹ Dongyang University

Abstract

This study aimed to identify factors related to happiness among Korean ninth-grade students using machine learning. Data were drawn from the 16th wave (2023) of the Panel Study on Korean Children (PSKC), conducted by the Korea Institute of Child Care and Education, and the final sample included 1,154 ninth-grade students. Happiness was used as the outcome variable, while sociodemographic, psychological, academic, relational, and lifestyle characteristics were included as predictors. Multiple linear regression, Elastic Net, Random Forest, and XGBoost models were compared. Feature importance, SHapley Additive exPlanations (SHAP), and Random Forest partial dependence plots (PDPs) were used to examine predictor importance, direction of predictive contribution, and nonlinear prediction patterns. Model performance rankings differed between cross-validation and final validation, with the models demonstrating broadly comparable predictive performance. In the tree-based machine learning models—Random Forest and XGBoost—self-esteem was the most important predictor, followed by depression, peer attachment, ego-resilience, academic stress, paternal and maternal attachment, self-regulated learning, smartphone overdependence, and subjective socioeconomic status. SHAP analyses showed broadly consistent directions of predictive contribution across the two models, while the Random Forest PDPs indicated nonlinear prediction patterns for several major variables. These findings demonstrate that happiness among Korean ninth-grade students is associated with diverse psychological, relational, and academic factors and provide foundational evidence for understanding adolescent happiness.

Keywords: Adolescents, Happiness, Machine Learning, Random Forest, XGBoost