

부모소진을 예측하기 위한 머신러닝 알고리즘의 적용

엄연용¹⁾ · 송원영²⁾

요약

본 연구의 목적은 아동종합실태조사 자료를 활용하여 부모소진을 예측하는 모델을 개발하고 성능평가 및 예측변수의 중요도를 파악하는 것이다. 이를 위해 2018년 아동종합실태조사 자료 중 만 17세 이하 자녀를 둔 어머니 2,659명의 자료를 분석하였다. 부모요인, 자녀양육요인 및 가족환경요인 차원에서 13개의 예측요인과 부모소진 간의 관계를 설정하고 랜덤포레스트, 에이다부스트, 인공신경망, 스택킹 모델을 적용하였다. 연구 결과, 네 가지 모델 모두 F1 .70 이상으로 나타났다. 부모소진을 예측하는 변수 중에서 양육효능감, 부모의 우울, 자녀 훈육, 가족관계가 중요한 영향을 미쳤다. 본 연구를 통해 아동종합실태조사 자료를 이용한 머신러닝 모델이 부모소진을 예측하는 데 유용하다는 것을 검증하였다. 부모소진은 부모요인, 자녀양육요인, 가족환경요인에 의해 예측되는 다결정요인임을 확인하였고 부모소진을 예방하기 위한 실증적인 자료와 토대를 마련하였다는 점에서 의의가 있다.

주요용어 : 부모소진, 랜덤포레스트, 에이다부스트, 인공신경망, 스택킹

1. 서론

초저출산이 지속됨에 따라 사회적으로 아동의 가치가 우선시 되고 자녀 중심의 양육방식이 만연해졌다. 자녀 양육은 부모에게 분명 중요하고 가치 있는 일이지만 자녀가 태어나면서부터 성인이 될 때까지의 양육은 지속되는 길고도 힘든 일이기도 하다. 부모가 된다는 사실만으로도 수많은 역할과 책임이 요구되며 부모들은 양육스트레스로 인한 어려움을 경험한다(Abidin, 1995). 부모 소진은 만성적인 양육스트레스로 인하여 대처 자원이 고갈되었을 때 발생하는 어려움을 말한다(Roskam, Brianda, & Mikolajczak, 2018). 부모소진은 심해지면 가정불화, 이혼, 아동방임 및 학대와 같은 사회적 문제를 야기한다(Mikolajczak, et al, 2018a).

최근의 연구자들은 부모소진의 위험성을 주목하였고 관련 연구가 급증하고 있다. 초기 연구에서는 특정 질병을 가진 아동 부모가 경험하는 부모소진을 측정하였지만 최근에는 일반부모를 대상으로 한 전반적인 연구가 진행되고 있다. 특히 부모소진이 부모, 자녀 및 양육, 가족기능과 환경에 의해 유발되는 다결정요인으로 부모뿐만 아니라

1) 주저자. 천안시 동남구 충절로 107, 건양대학교 심리상담학과, 박사수료, E-mail: eey83@hanmail.net

2) 교신저자. 충청남도 논산시 대학로 121, 건양대학교 심리상담치료학과, 교수, E-mail: song@konyang.ac.kr

자녀와 가족체계까지 영향을 미친다는 관점이 우세하다(Mikolajczak, et al, 2018b). 양육자 중에는 주 양육자 역할을 감당하는 어머니가 부모소진에 취약하였으며, ADHD를 비롯한 자녀의 문제도 부모소진에 직접적인 영향을 미친다(Hatice, & Sibel, 2019; Lindahl, et al, 2014; Slavka, Martina, & Alena, 2020). 전통적인 성 역할, 완벽주의, 사회적으로 요구되는 부모역할도 부모소진을 유발하는 중요한 요인이다(Bastiaansen, Verspeek, & Bakel, 2021; Lin, et al, 2021). 2020년 이후로는 전 세계적 감염병인 코로나19로 인하여 경제적 어려움, 이동통제로 인한 자녀의 돌봄 및 학업 공백에 영향을 끼침에 따라 부모들이 더 높은 소진을 경험하고 있다(Syazwani, et al, 2020). 최근 42개국에서 부모소진을 조사한 결과, 부모소진의 발병률은 5%로 나타났으며 일부 국가(벨기에, 폴란드)에서는 8% 이상으로 나타나 문화적 차이에 따라서도 다소 다른 현상이 관찰되고 있다(Roskam et al, 2021). 하지만 우리나라에서는 부모소진 연구가 매우 부족하며 한국 문화에 따른 부모소진을 예측하기 위한 척도 및 연구도 거의 전무한 실정이다. 부모소진이 가족의 위험 신호로 밝혀짐에 따라 국내에서도 부모소진에 대한 연구가 필요한 상황이다.

2. 이론적 배경

2.1 부모소진의 정의

부모소진은 양육스트레스로 인하여 신체적, 심리적으로 탈진(burnout)된 상태를 말한다(Roskam, 2017; Hubert & Aujoulat, 2018). 이는 부모 특성, 자녀 특성, 양육 스트레스, 가족과 환경의 특성에 의해 유발되는 다결정 증후군으로 특히, 양육과 관련된 불안, 지나친 책임, 완벽주의, 부모역할에서의 과부하가 부모소진에 직접적인 영향을 미친다(Mikolajczak, et al, 2018b; Mishina, 2012).

Roskam, Raes과 Mikolajczak(2017)은 직무 소진 개념을 토대로 부모소진(Parental Burnout)을 3차원으로 개념화하였다. 첫 번째 차원은 정서적 고갈이다. 자녀양육 과정에서 부모역할로 인한 과부하를 경험함에 따라 신체적, 정신적으로 탈진된 상태를 말한다. 두 번째 차원은 자녀와의 정서적 거리감이다. 만성적 스트레스로 인하여 지친 부모는 자녀에게 관심이 줄어들고 기능적/ 도구적 측면에서 부모역할이 제한된다. 이러한 현상은 정서적 고갈에 대한 자기 보호적 반응이지만 자신의 행동에 대한 죄책감을 유발하여 악순환을 야기할 수 있다(Mikolajczak, Gross, & Roskam, 2019). 세 번째 차원은 부모역할에 대한 무능감이다. 소진으로 인한 역기능적 양육행동은 부모역할에 대한 성취감을 저하시키고 이는 부모로서 부족하다고 생각하게 되어 자신을 부정적으로 평가하게 된다. 결국 부모역할에 대한 무능감은 부모로서 좌절감과 불행을 유발하여 부모소진을 더욱 심화시킨다.

이후 같은 연구자들(Roskam, Brianda와 Mikolajczak, 2018)이 Hubert와 Aujoulat(2018)의 연구를 토대로 부모 소진을 4차원으로 재 개념화 하였다. 첫째, 부모역할에 대한 탈진은 부모역할에 대한 책임감이 압도되어 신체적, 정서적으로 기진맥

진된 상태이다. 둘째, 부모로서 자기 대조는 소진 전의 모습과 소진 후 모습이 대조되는 상태를 말한다. 셋째, 포화는 부모역할에 대한 싫증으로 부모역할에서 더 이상 즐거움을 경험하지 못하는 상태이다. 마지막 자녀와 정서적 거리두기는 부모역할에 대한 불안, 두려움, 자책으로 인하여 자녀와 정서적 유대감이 소원해진 상태를 말한다.

2.2 부모소진의 배경

양육스트레스는 자녀를 양육하는 과정에서 경험하는 스트레스로 양육 맥락 안에서 부모와 자녀에게 부정적 영향을 미친다(김기현, 강희경, 1997; Abidin, 1995). 양육스트레스는 일상생활 속에서 지속적으로 경험됨에 따라 역기능적 양육행동을 유발하고 만성적으로 누적되어 부모소진을 유발할 수 있다(Hubert & Aujoulat, 2018; Roskam, Brianda & Mikolajczak, 2018).

양육결정모델에 따르면 양육과정에서 경험하는 스트레스는 부모, 자녀, 환경 등 부모의 삶 전반적인 측면에서 영향을 받는다(Chase-Lansdale & Pittman, 2002). 첫째, 부모요인은 부모의 정신건강을 포함한 심리적 요인, 성격 등으로 직접적인 영향을 미친다. 둘째, 자녀요인은 자녀의 기질, 성별, 성격뿐만 아니라 자녀양육과 관련된 요인들이 포함된다. 마지막 환경요인으로 가족기능, 경제수준, 부부갈등, 부모 및 자녀와 연결된 사회적 네트워크까지 직, 간접적인 영향을 미친다. 결국 양육스트레스는 부모, 자녀, 환경요인 안에서 부정적 요인들이 서로 상호 관계적 영향을 미쳐 발생한다. 가족스트레스도 부모 개인과 가족 및 환경으로 인해 유발된다. 가족스트레스 모델을 살펴보면 가족스트레스는 개별적으로 경험되지 않고 개인의 특성, 가족 차원, 환경적 차원 등 여러 가지 스트레스가 상호 연관적으로 발생하며 누적되는 특성이 있다(Lavee, McCubbin & Patterson, 1985). 그리고 자녀의 발달단계에 따라 변화하는 가족생활주기는 부모 역할에서 불균형을 초래해 스트레스를 유발한다(Grant, McMahon, & Austin, 2008; Pass, 2017). 부모역할에 대한 기술과 양육효능감 부족은 걱정, 불안, 두려움, 죄책감 등 심리적 스트레스를 유발하여 부모소진으로 이어질 수 있다(Mousavi, 2020).

2.3 부모소진에 영향을 미치는 요인

부모소진에 영향을 미치는 부모 요인으로서는 신경과민, 높은 완벽주의 수준, 연령, 교육 수준, 직업 등이 있다(김종두, 2018; Gerain & Zech, 2018; Lin et al, 2021; Mikolajczak et al, 2018b). 문화적 양육가치, 개인주의 문화, 전통적인 성 역할이 부모소진을 유발하였으며 특히 어머니가 부모 소진에 더욱 취약하였다(한수희, 2019; Mousavi, 2020; Sorkkila & Aunola, 2020). 자녀양육 요인으로서는 자녀의 질병이 부모역할에 부담을 주어 부모소진에 영향을 미치고 양육불안과 양육효능감 부족도 부모소진과 부모 삶의 질에도 부정적 영향을 미친다(김종두, 2018; Enkeu, & Enung, 2019; Mishina et al, 2012). 또한 자녀가 ADHD나 만성적 질환을 가진 경우 가족의 경제적 문제와도 관련되어 더 높은 수준의 부모소진을 야기한다(Hatice & Sibel, 2019;

Lindahl et al, 2014; Slavka, Martina & Alena, 2020). 최근의 환경적 요인으로서는 COVID-19 감염병 확산방지로 인해 시행된 이동 제한 명령(Movement Control Order)이 자녀 교육 환경 변화와 돌봄 문제를 야기하여 부모소진에 직접적인 영향을 미치고 있다(Syazwani et al, 2020).

중요한 점은 부모소진이 시급히 개입해야 할 사회적 문제라는 것이다. 부모소진은 양육스트레스보다 훨씬 위험하며 우울, 가족갈등 및 이혼에 직접적인 영향을 미치고 역기능적인 양육과 폭력적 양육행동을 유발하여 아동방임 및 학대를 초래할 수 있다(Mikolajczak 2018a; Mikolajczak et al, 2019a). 그리고 부모소진은 아동의 삶에 대한 불만족, 정서적 외로움, 반사회적 행동을 유발하며, 청소년에게도 적응에 부정적 영향을 미쳐 우울과 불안을 야기한다(Beiming et al, 2021).

그러나 국내에서는 현재까지 양적연구 1편, 질적연구 1편, 부모소진 척도(PBA) 타당화 연구 1편 정도의 연구만이 이루어져 관련 연구가 현저히 부족한 실정이다. 기존 연구에서 밝혀졌듯이 부모소진은 문화적 영향에 따라 연구되어야 할 분야로서 이미 많은 나라에서 연구가 진행되고 있다. 그리고 부모소진은 부모와 자녀 그리고 가족붕괴를 위협하는 요인이며 아동학대 예방과 개입을 위해서도 부모소진을 야기하는 요인을 시급히 식별하고 중요도를 확인하여야 할 필요가 있다(엄문설, 이양희, 2020).

2.4 머신러닝 알고리즘

부모소진을 예측하는 기존연구들은 대부분 회귀모델 또는 구조방정식 모델 검증을 통해 검증되어 왔다. 전통적 통계검증의 경우 많은 예측변수를 투입하면 영향력이 분산되는 한계점이 있다. 또한 예측변수간의 다중공선성 문제로 기본 가정을 위배하여 높은 상관관계를 보이는 변수들을 적용하기에는 제한적이다. 반면 머신러닝 모델은 빅데이터를 통해 효율적으로 모델을 학습시켜 분류, 예측할 수 있으며 새로운 데이터를 적용 및 예측하여 일반화(generalization) 할 수 있다. 이러한 점에서 기존연구는 부모요인, 자녀요인, 환경요인을 선택적으로 검증하여왔으나 머신러닝 모델의 경우 부모요인, 자녀 및 양육요인, 환경요인을 동시에 검증할 수 있다는 장점이 있다. 또한 알려진 요인간의 예측뿐만 아니라 모델을 구축할 수 있으며 중요도에 따라 어떤 요인이 관심변인에 영향을 미치는지 살펴 볼 수 있다. 다만 알고리즘에 따라 편향이 있어 어느 모델이 더 우월하다고 평가할 수 없으므로 가장 좋은 모델을 선택하기 위해 최소 몇 가지 모델을 비교하는 것이 권장된다(Sebastian, & Vahid, 2017). 머신러닝에서의 지도학습(Supervised Learning)이란 예측변수와 결과변수 간의 관계를 학습시켜 데이터로부터 모델을 구축하는 것이다. 본 연구에서는 지도학습을 통해 랜덤포레스트(Random forest), 에이다부스트(Adaboost), 인공신경망(Neural Network), 스택킹(Stacking) 모델들을 적용하여 부모소진 관련 변인들을 분석하고자 한다.

랜덤포레스트(Random forest)는 결정 트리의 앙상블(ensemble)을 사용하여 예측하는 모델로 결정 트리 집합을 통해 모델을 구축한다(Breiman, 2001). 훈련데이터에서 개별 트리를 개발할 때 임의의 특성 하위 집합을 통해 최종 모델을 결정한다. 일반화 성능이 높고 과대적합 위험이 적으며 앙상블 모델이 개별 결정 트리가 만든 잡음으로부터 매우 안정되어 있어 하이퍼파라미터 튜닝에 많은 노력을 기울이지 않아도 견고한 모델을 만들 수 있다는 장점이 있다. 에이다부스트(Adaboost)는 앙상블 메타 알고리즘으로

성능을 향상시키기 위하여 다른 형태의 학습 알고리즘과 결합하여 사용한다(Yoav, & Robert, 1997). 이전 분류기에 의해 잘못 분류된 것들을 수정한다는 점에서 다양한 상황에 적용할 수 있다. 약한 학습기를 훈련할 때 훈련 세트 전체를 사용하며 훈련 샘플은 반복마다 가중치가 다시 부여되며 이전 학습기의 실수를 학습하는 강력한 분류기를 만들어 편향을 감소시킬 수 있다. 최고의 발군 분류기로 다른 학습 알고리즘보다 과적합(overfitting)에 덜 취약하여 본 연구에 적용하였다. 인공신경망(Neural Network)은 생물학 신경망을 적용한 통계학적 학습 알고리즘으로 패턴을 분석하여 예측, 분류 모델을 구축할 수 있다(Hopfield, 1982). 지도 학습 알고리즘으로 외부 입력의 입력층(input layer), 결과 출력의 출력층(output layer), 입력층과 출력층 사이에 위치한 은닉층(hidden layer)으로 구성된다. 데이터 세트에 대한 학습을 통해 각 층은 여러 개 노드로 구성되며 다른 층의 노드와 연결 가중치로 연결된다. 즉, 입력층은 활성화 함수를 거쳐 은닉층을 통해 출력층으로 전달되는 역전파 알고리즘으로 데이터를 학습하며 입력층(input layer)과 출력층(output layer) 사이에 은닉층(hidden layer)이라고 하는 하나 이상의 비선형 레이어가 있다는 점에서 로지스틱 회귀와 다르다. 본 연구에서는 역전파 기법과 경사 하강법을 기반으로 학습이 이루어지는 Sklearn의 다층퍼셉트론 MLP(Multi-layer Perceptron)을 적용하였다. 스택킹(Stack multiple models)은 학습 알고리즘을 사용하여 메타 모델을 계산하는 효율적인 앙상블 기법이다(Syarif et al, 2012). 각 레이어가 모델 조합으로 구성되어 계층의 수를 증가시키며 각 계층의 오분류율(error-rate)을 최소화하여 성능을 개선하는 딥러닝 구조와 유사한 특징이 있다. 스택킹 모델은 계층 알고리즘을 사용하며 기존 학습된 모델로부터 도출된 예측 값을 학습 데이터로 삼아 새로운 예측 집합을 형성하기 때문에 복잡한 데이터 집합에서 최적의 성능으로 모델을 예측할 수 있어 본 연구에 적용하고자 하였다.

이상으로 본 연구의 목적은 영아부터 청소년까지 부모요인, 자녀양육요인, 가족환경요인을 복합적으로 조사한 아동종합실태조사 자료를 통해 부모소진을 예측하는 머신러닝 모델을 구축하고 성능 평가 및 다결정요인임을 밝히고자 한다. 더불어 부모소진은 자녀 연령에 따라 예측변수가 다르게 영향을 미칠 수 있음을 고려하여 자녀 연령에 따라 세 집단으로 자료를 분리, 모델에 적용하여 집단에 따라 영향을 미치는 예측변수의 중요도를 살펴보고자 한다. 부모소진을 예측하기 위한 머신러닝 모델 구축은 부모와 아동 복지의 근거 자료로서 활용할 수 있으며 부모소진을 예방할 수 있는 실증적 자료가 될 것이다.

3. 연구방법

3.1 연구대상

본 연구는 통계청 마이크로데이터 서비스를 통해 제공받은 2018년 아동종합실태조사 자료 중 어머니의 응답 자료를 사용하였다. 2018년 아동종합실태조사는 전국 만 18세 미만 아동 가구를 대상으로 2019년 1월 7일부터 2019년 2월 28일까지 진행된 횡단조사이다. 본 연구에서는 주 양육자인 어머니를 대상으로 2,659명의 어머니 자료만을 사

용하였다. 그리고 자녀의 연령에 따라 만 5세 이하 자녀를 둔 어머니 639명, 만 6세-11세 자녀를 둔 어머니 964명, 만 12세-17세 자녀를 둔 어머니 1,056명, 세 집단으로 구분하여 각각 분석하였다. 수집된 연구대상의 일반적 특성은 <표 3.1>과 같다.

<표 3.1> 연구대상의 일반적 특성(N=2,659)

구분	빈도(퍼센트)
어머니학력	고졸
	907(34.1%)
	전문대졸
	768(28.9%)
	대졸이상
	984(37.0%)
취업여부	미취업
	1,424(53.6%)
	취업
	1,235(46.4%)
가구구분	일반가구
	2,124(79.9%)
	저소득층가구
	535(20.1%)
자녀성별	남자
	1,380(51.9%)
	여자
	1,279(48.1%)
자녀연령	만 5세 이하
	639(24%)
	만 6세-11세
	964(36.3%)
	만 12세-17세
	1,056(39.7%)

3.2 측정도구

본 연구의 목적은 부모소진을 예측하고, 예측에 중요한 역할을 하는 주요 요인들을 파악하는 것이다. 이에 부모소진을 결과변수로 설정하고 부모요인, 자녀양육요인, 환경요인의 세 가지 차원에서 부모의 우울, 부모 건강만족, 부모 삶 만족, 부모의 걱정, 자녀 성별, 양육효능감, 자녀 훈육, 자녀 체벌, 가구특성, 소득수준, 생활수준만족, 가족관계, 부부싸움 빈도 등을 예측변수로 설정하여 다음과 같이 구성하였다.

3.2.1 결과변수 : 부모소진

본 연구에서는 부모소진을 결과변수(outcome)로 설정하였다. 부모소진은 Abidin(1995)이 개발하고 김기현과 강희경(1997)이 번안한 Parental Stress Index Short Form의 양육스트레스 척도의 문항 중 정서적 고갈(‘아이가 태어난 후 예전만큼 나의 생활이 즐겁지 않다’, ‘모임에 가면서도 아이 때문에 그리 즐겁지 않을 것이라고 생각한다’), 정서적 거리감(‘아이로부터 도망치고 싶을 때가 있다’, ‘피곤할 때 아이가 놀아달라고 보채면 귀찮은 생각이 든다’), 제한된 기능적/도구적 역할(‘아이를 잘 키울 수 있을지 자신이 없다’, ‘아이가 정서적으로 불안정한 모습을 보이면 내 탓인 것 같아 괴롭다’)를 반영한 10개 문항을 선별하였다. ‘아동종합실태조사에서는 5점 Likert 척도로 측정되었으나 머신러닝 알고리즘에서 최적의 성능 향상을 위해 특성(0,1) 범위 값으로 변환하여 사용하는 경우가 일반적이다(Sebastian, & Vahid, 2017). 부모소진을 긍정 문항(4점 “대체로 그렇다”, 5점 “매우 그렇다”)의 경우 1 값을 부여하여 부모소진 유로 설정하고 중립(3점 “보통이다”)과 부정 문항(1점 “전혀 그렇지 않다”, 2점 “그렇지 않다”)의 경우는 부모소진 무로 가변환하여 이항변수를 분석에 사용하였다. 어머니의 부모소진 빈도는 <표 3.2>와 같다.

<표 3.2> 어머니의 부모소진 빈도표

	만 5세 이하 (n=639)	만 6세-11세 (n=964)	만 12세-17세 (n=1,056)	전체 (n=2,659)
부모소진 유	150 (23.5)	236 (24.5)	281 (26.6)	667 (25.1)
부모소진 무	489 (76.5)	728 (75.5)	775 (73.4)	1,992 (74.9)

* 괄호 안은 %

<표 3.2>에 제시한 바와 같이 만 5세 이하 유아를 양육하는 어머니 중에서 부모소진을 경험한 어머니는 150명으로 23.5%로 나타났다. 만 6세에서 11세 아동을 양육하는 어머니 중에서 부모소진을 경험한 어머니는 236명으로 24.5%로 나타났다. 만 12세에서 17세 청소년을 양육하는 어머니 중에서 부모소진을 경험한 어머니는 281명으로 26.6%로 나타났다. 전체 2,659명의 어머니 중에서 부모소진을 경험한 어머니는 모두 667명으로 25.1%에 해당한다.

3.2.2 예측변수

본 연구에서는 양육결정모델과 가족스트레스 모델을 기반으로 부모요인, 자녀양육요인, 가족환경요인의 3차원에서 예측변수를 설정하였다. 부모요인으로는 부모의 우울, 부모 건강만족, 부모의 삶 만족, 부모의 걱정을 포함하였다. 자녀양육요인으로 자녀의 성별, 자녀 훈육, 자녀 체별, 양육효능감을 포함하였다. 가족환경요인으로 가구특성(일반가정, 저소득층), 소득수준, 생활수준만족, 가족관계, 부부싸움 빈도를 포함하였다. 예측변수는 부모요인 4개, 자녀양육요인 4개, 가족환경요인 5개로 총 13개의 변수이다.

3.2.2.1 부모요인

부모의 우울 CES-D(Center for Epidemiological Studies of Depression Scales)은 Radloff(1977)이 개발한 척도이다. 11문항으로 이루어져 있으며 ‘먹고 싶지 않고 식욕이 없다’, ‘상당히 우울했다’, ‘모든 일들이 힘들게 느껴졌다’, ‘마음이 슬펐다’, ‘도무지 뭘 해 나갈 엄두가 나지 않았다’ 등의 문항으로 Likert식 4점 척도 구성되어 있으며 점수가 높을수록 우울이 높은 것으로 해석된다. 부모건강만족 PWI(personal Wellbeing Index)은 Cummins(2003)이 개발한 척도로 건강에 대하여 만족하는지를 0점(전혀 만족하지 않는다)에서 10점(매우 만족한다)로 측정한다. 부모의 삶 만족(The Cantril Self-Anchoring Striving Scale)은 Cantril(1965)이 개발한 척도로 현재 삶에 대해 얼마나 만족하는지를 0점(최악의 삶)에서 10점(최상의 삶)로 측정한다. 부모의 걱정은 OECD에서 Measuring Well-Being에 포함된 척도로 현재 얼마나 걱정하는지를 0점(전혀 걱정하지 않았다)에서 10점(매우 걱정했다)로 측정한다.

3.2.2.2 자녀양육요인

자녀의 성별은 남, 여로 구분하여 사용하였다. 자녀 훈육은 2008년 아동종합실태조

사에서 사용한 문항으로 자녀를 훈육하는 5가지 방법으로 구성되어 있다. ‘신체적 체벌’, ‘벌세우기’, ‘말로 야단치기’, ‘벌로써 장난감이나 게임기, 스마트폰 사용 제한하기’, ‘칭찬과 보상 사용하기’, 문항으로 1점(거의 사용하지 않는다)에서 4점(매우 자주 사용 한다)로 측정한다. 자녀 체벌은 2013년 아동종합실태조사에서 사용한 문항으로 자녀를 양육할 때 신체적 체벌에 대한 인식에 대하여 1점(전혀 필요하지 않다)에서 4점(꼭 필요하다)로 측정한다. 양육효능감은 신숙재, 정문자(1998)가 개발한 척도로 9문항으로 이루어져 있으며 자녀 양육효능감에 대하여 1점(전혀 그렇지 않다)에서 5점(매우 그렇다)로 측정한다. 점수가 높을수록 양육효능감이 높은 것으로 해석된다.

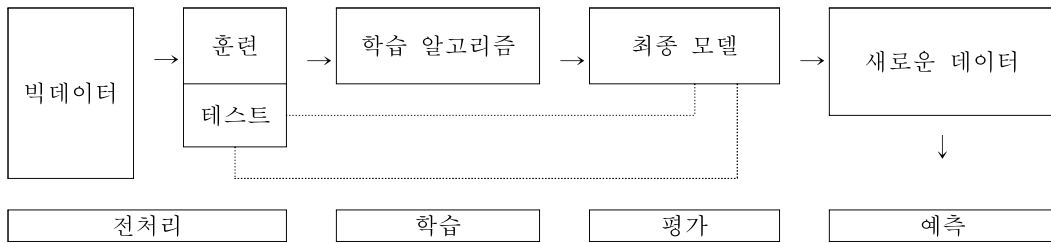
3.2.2.3 가족환경요인

가구특성은 일반가구와 저소득층가구로 구분하여 사용하였다. 일반가구는 우리나라에 거주하는 만 0세-18세 미만 아동가구, 저소득층가구는 행정구역에 거주하는 만 0세-18세 미만 아동가구 중 수급권자가 있는 가구이다. 소득수준은 중위소득 100% 미만과 중위소득 100% 이상으로 구분하여 사용하였다. 생활수준만족 PWI(personal Wellbeing Index)은 Cummins(2003)이 개발한 척도로 생활수준에 대하여 만족하는지를 0점(전혀 만족하지 않는다)에서 10점(매우 만족한다)로 측정한다. 가족관계는 양옥경, 김연수(2007)가 개발한 척도로 15문항으로 이루어져 있으며 ‘우리 가족은 가족끼리 대화가 잘 된다’, ‘우리 가족은 서로 지지해주고 격려해준다’, ‘우리 가족은 도움이 필요할 때 언제든지 서로 도움을 청할 수 있다’, ‘우리 가족은 가족들 간의 서로 다른 의견을 존중 한다’, ‘우리 가족은 서로의 감정을 존중 한다’, 등의 문항으로 1점(전혀 그렇지 않다)에서 5점(항상 그렇다)로 측정한다. 점수가 높을수록 가족관계가 좋은 것으로 해석된다. 부부싸움 빈도는 2008년 아동종합실태조사에서 사용한 문항으로 배우자와 싸우는 빈도에 대하여 1점(거의 싸우지 않는다)에서 6점(일주일에 1-2회 이상 싸운다)로 측정한다.

3.3 분석방법

본 연구에서 부모소진을 예측하기 위한 머신러닝 알고리즘 적용 및 성능을 검증하기 위하여 파이썬(python) 패키지 시스템 아나콘다(anaconda)와 아나콘다에서 제공되는 라이브러리 Orange 3.6.1를 활용하였다. Orange는 Slovenia의 Ljubljana 대학교에서 개발한 소프트웨어 프로그램이다(Demsar et al, 2013). 일반적인 머신러닝 모델 구축 과정은 [그림 1]과 같다.

데이터 전 처리는 일반적으로 데이터를 정제하고 검증하는 과정이다. 본 연구에서의 데이터 전 처리 과정은 다음과 같다. 첫째, 설명변수의 측정 단위가 다양하여 데이터를 표준/정규화(min-max scaling) 하였고 결측치와 극단치 처리, 변수의 값의 범위 설정, 관측값(feature)과 목표값(target)을 설정하였다. 둘째, 집단 간 차이를 살펴보기 위해 전체 데이터를 자녀 연령에 따라 만 5세 이하, 만 6세에서 11세, 만 12세에서 17세, 세 집단으로 구분하였다. 셋째, 학습용 데이터(training)와 평가용 데이터(test, evaluation)를 일반적으로 권장되는 7:3 비율, 랜덤 방식으로 설정하였다.

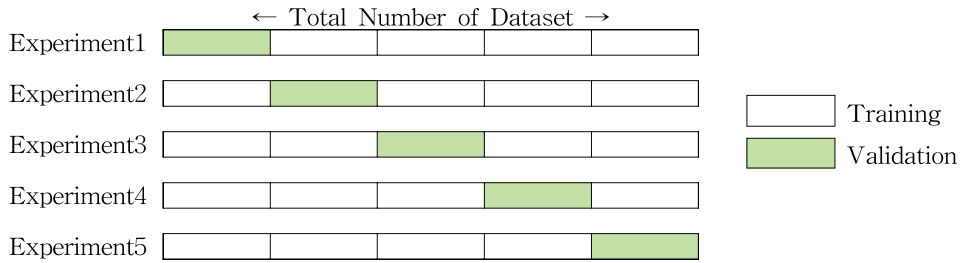


[그림 3.1] 머신러닝 모델 구축 과정

분석과정은 랜덤포레스트(Random forest), 에이다부스트(Adaboost), 인공신경망(Neural Network) 모델을 사용하여 지도학습(Supervised Learning)을 통해 예측변수와 결과변수를 학습시켰다. 그 다음 학습된 데이터로부터 모델을 구축하고 평가용 데이터에 적용하여 성능을 검증하였다. 그리고 이미 학습시킨 모델들을 합쳐서 사용하는 앙상블 기법 스택킹 모델(Stacking)을 적용하여 최적의 성능을 산출하였다. 하이퍼파라미터(hyperparameter)는 머신러닝 알고리즘에 따라 상이하고 기본 하이퍼파라미터 설정은 최적의 성능을 기대할 수 없기 때문에 연구자가 튜닝과정을 통해 과대적합(overfitting) 또는 과소적합(underfitting) 문제를 조율하여 성능이 가장 우수한 모형을 선택(model selection) 하여 일반화(generalization) 할 수 있다(Sebastian, & Vahid, 2017). 본 연구에서 적용한 하이퍼파라미터는 <표 3.3>과 같다.

<표 3.3> 하이퍼파라미터

Model		Hyperparameter	
Random Forest	n_estimators: 10		
	max_features: auto		
	min_samples_split: 5		
	Others are set to defaults		
Adaboost	Base estimator: Tree		
	Number of estimators: 50		
	Learning rate: 0.001, 0.005, 0.01		
	Classification algorithm: SAMME.R		
Neural Network	Neurons per hidden layer: 100		Cross-Validate number of folds=5 (usually 5 or 10)
	Activation function for the hidden layer: ReLu		
	Solver for weight optimization: Adam		
	Alpha: L2 penalty parameter		
	Max iterations: 200		
	Other parameters are set to sklearn's defaults		
Stacking	Random Forest		
	Adaboost		
	Neural Network		



[그림 3.2] Cross-Validate

성능평가는 ‘Orange 3.6.1’ 툴을 사용하여 혼동행렬을 산출하고 정확도(CA: accuracy), 가중평균(F1), 정밀도(Precision), 재현율(Recall), AUC(area under the ROC curve)를 계산하였다. AUC는 모델에 따라 ROC 결정곡선으로 시각화 하였다. 부모소진을 예측하는 변수들의 중요도는 Rank의 Information Gain으로 산출하였다. Rank 위젯은 구축한 모델을 기반으로 내부 득점에 기초하여, 결과변수와의 관계에 따라 n개의 최고 순위를 산출하고 자동으로 변경 사항이 반영된다. Information Gain은 각 노드의 엔트로피를 계산하고 노드의 속한 레코드의 개수를 가중치 하여 엔트로피를 평균한 값으로 속성에 대한 정보이득 양을 계산한다. 즉, Information gain 값이 클수록 변별력이 좋을 것을 의미한다. <표 3.4>는 모델 성능을 평가하기 위해 사용한 혼동행렬(confusion matrix)이다.

<표 3.4> Confusion Matrix(혼동 행렬)

		Predicted(예측값)	
		Positive(양성)	Negative(음성)
Actual(실제 값)	Positive(양성)	True Positive(TP)	False Negative(FN)
	Negative(음성)	False Positive(FP)	True Negative(TN)

정확도(CA: Classification Accuracy) = $TP + TN / TP + TN + FP + FN$

정밀도(Precision) = $TP / TP + FP$

재현율(Recall) = $TP / TP + FN$ (실제 양성인 모든 데이터 수)

가중평균(F1) = $2 * (Precision * Recall / Precision + Recall)$

ROC 곡선 = y축(sensitivity; probability that target=1 when true value=1)

x축(1-specificity; probability that target=1 when true value=0)

4. 연구결과

4.1 예측 모델 성능

부모소진을 예측하기 위하여 랜덤포레스트, 에이다부스트, 인공신경망, 스택킹 모

텔을 적용하여 데이터를 학습시키고 최적의 하이퍼파라미터를 적용하여 성능을 평가하였다. 예측변수로 부모요인(우울, 건강만족, 삶 만족, 걱정), 자녀양육요인(성별, 훈육, 체벌, 양육효능감), 가족환경요인(가구특성, 소득수준, 생활수준, 가족관계, 부부싸움빈도), 총 13개의 예측변수를 투입하였고 결과변수로 부모소진을 투입하였다. 머신러닝 모델 성능은 <표 4.1>과 같다.

<표 4.1>에 제시한 바와 같이 정확도(CA: Classification Accuracy)는 부모소진을 정확하게 예측한 비율이다. 전체 데이터에서는 스택킹 모델이 .77, 랜덤포레스트 모델이 .76, 인공신경망 모델이 .76, 에이다부트스 모델이 .74로 나타났다. 집단 별로 살펴보면 만 5세 이하 집단에서는 스택킹 모델이 .77, 랜덤포레스트 모델이 .77, 인공신경망 모델이 .75, 에이다부트스 모델이 .72로 나타났다. 만 6세에서 11세 집단에서는 스택킹 모델이 .78, 랜덤포레스트 모델이 .78, 인공신경망 모델이 .78, 에이다부트스 모델이 .78, 모두 동일하였다. 만 12세에서 17세 집단에서는 스택킹 모델이 .73, 랜덤포레스트 모델이 .73, 인공신경망 모델이 .73, 에이다부트스 모델이 .70으로 나타났다. 정확도는 모든 집단에서 스택킹 모델이 가장 높게 나타났다.

<표 4.1> 부모소진 예측 성능

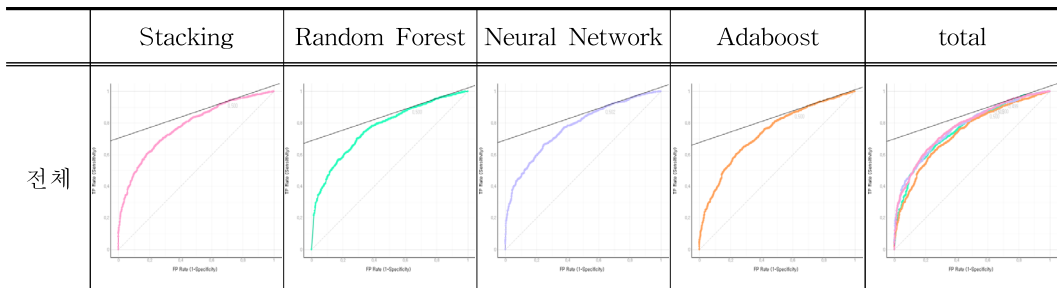
	AUC	CA	F1	Precision	Recall
전체집단					
Stacking	.77	.77	.74	.75	.77
Random Forest	.76	.76	.74	.73	.76
Neural Network	.77	.76	.74	.74	.76
AdaBoost	.73	.74	.73	.73	.74
만 5세 이하					
Stacking	.75	.77	.73	.74	.77
Random Forest	.75	.77	.75	.75	.77
Neural Network	.74	.75	.74	.73	.75
AdaBoost	.64	.72	.71	.71	.72
만 6세-11세					
Stacking	.79	.78	.76	.76	.78
Random Forest	.76	.78	.76	.75	.78
Neural Network	.78	.78	.77	.76	.78
AdaBoost	.75	.78	.76	.76	.78
만 12세-17세					
Stacking	.74	.73	.70	.70	.73
Random Forest	.73	.73	.71	.70	.73
Neural Network	.74	.73	.71	.70	.73
AdaBoost	.68	.70	.70	.69	.70

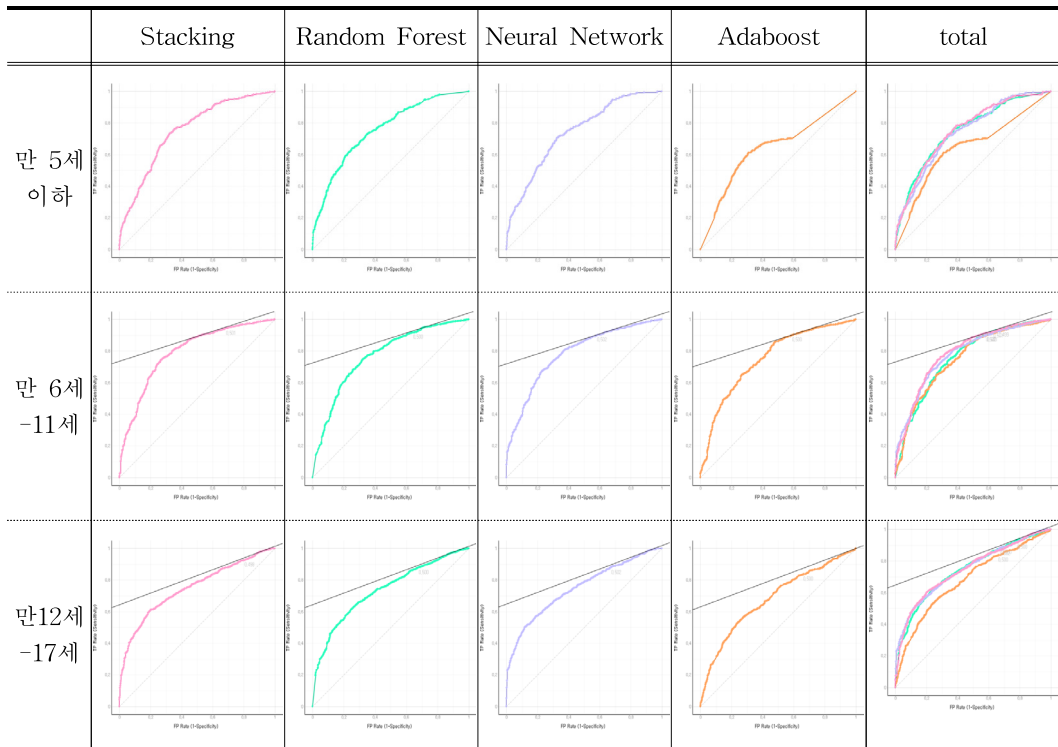
정밀도(Precision)는 양성진단 정확성으로 양성으로 예측한 데이터 중 실제 양성인 값을 계산한 값이다. 전체 데이터에서는 스택킹 모델이 .75, 랜덤포레스트 모델이 .73, 인공신경망 모델이 .74, 에이다부트스 모델이 .73으로 나타났다. 집단 별로 살펴보면

만 5세 이하 집단에서는 스택킹 모델이 .74, 랜덤포레스트 모델이 .75, 인공신경망 모델이 .73, 에이다부트스 모델이 .71으로 나타났다. 만 6세에서 11세 집단에서는 스택킹 모델이 .76, 랜덤포레스트 모델이 .75, 인공신경망 모델이 .76, 에이다부트스 모델이 .76으로 나타났다. 만 12세에서 17세 집단에서는 스택킹 모델이 .70, 랜덤포레스트 모델이 .70, 인공신경망 모델이 .70, 에이다부트스 모델이 .69으로 나타났다. 정밀도는 전체 집단에서는 스택킹 모델이 가장 높게 나타났으나 만 5세 이하 집단에서는 랜덤포레스트 모델이 가장 높게 나타났고 만 6세에서 11세 집단과 만 12세에서 17세 집단에서는 모델 성능이 유사하였다.

재현율(Recall)은 양성 탐색 효과성으로 전체 양성 중에서 예측으로 양성을 찾아낸 비율이다. 전체 데이터에서는 스택킹 모델이 .77, 랜덤포레스트 모델이 .76, 인공신경망 모델이 .76, 에이다부트스 모델이 .74로 나타났다. 집단 별로 살펴보면 만 5세 이하 집단에서는 스택킹 모델이 .77, 랜덤포레스트 모델이 .77, 인공신경망 모델이 .75, 에이다부트스 모델이 .72으로 나타났다. 만 6세에서 11세 집단에서는 스택킹 모델이 .78, 랜덤포레스트 모델이 .78, 인공신경망 모델이 .78, 에이다부트스 모델이 .78, 모두 동일하였다. 만 12세에서 17세 집단에서는 스택킹 모델이 .73, 랜덤포레스트 모델이 .73, 인공신경망 모델이 .73, 에이다부트스 모델이 .70으로 나타났다. 재현율은 모든 집단에서 스택킹 모델이 가장 높게 나타났다.

마지막으로 가중평균(F1)은 정밀도(Precision)와 재현율(Recall)을 가중평균으로 계산한 비율로 모델 성능 평가에서 매우 유용한 지표로 사용된다. 전체 데이터에서는 스택킹 모델이 .74, 랜덤포레스트 모델이 .74, 인공신경망 모델이 .74, 에이다부트스 모델이 .73으로 나타났다. 집단 별로 살펴보면 만 5세 이하 집단에서는 스택킹 모델이 .73, 랜덤포레스트 모델이 .75, 인공신경망 모델이 .74, 에이다부트스 모델이 .71으로 나타났다. 만 6세에서 11세 집단에서는 스택킹 모델이 .76, 랜덤포레스트 모델이 .76, 인공신경망 모델이 .77, 에이다부트스 모델이 .76으로 나타났다. 만 12세에서 17세 집단에서는 스택킹 모델이 .70, 랜덤포레스트 모델이 .71, 인공신경망 모델이 .71, 에이다부트스 모델이 .70으로 나타났다. 가중평균은 전체 집단에서는 에이다부트스 모델을 제외하고 동일하게 나타났으나 만 5세 이하 집단에서는 랜덤포레스트 모델이 가장 높게 나타났고 만 6세에서 11세 집단에서는 인공신경망 모델, 만 12세에서 17세 집단에서는 랜덤포레스트 모델과 인공신경망모델 성능이 높게 나타났다.





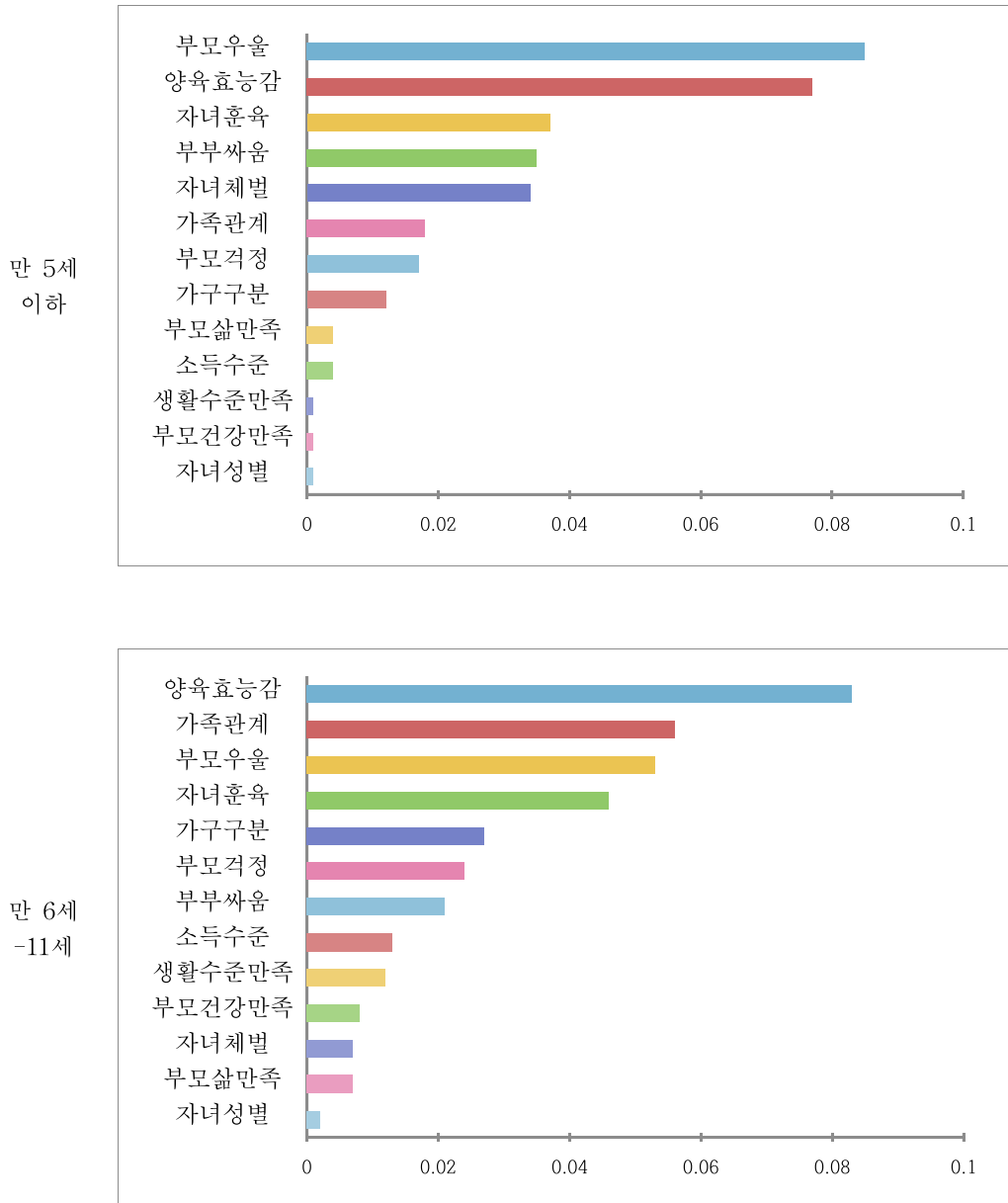
[그림 4.1] ROC 결정곡선

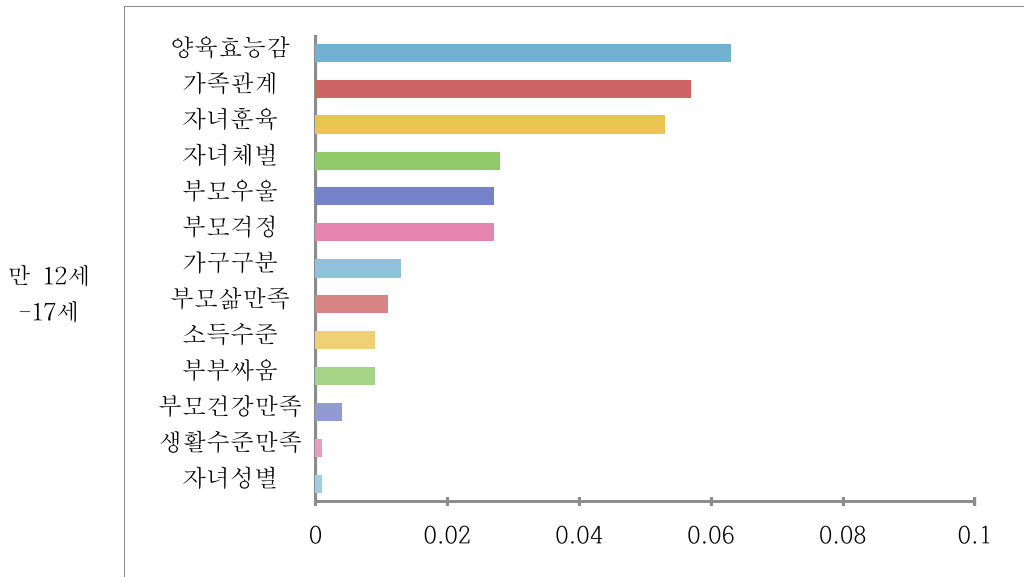
부모소진을 예측하는 머신러닝 모델 성능 평가를 살펴보면 랜덤포레스트 모델, 에이다부스트 모델, 인공신경망 모델, 스택킹 모델, 네 가지 모델의 예측력이 유사하였고 부모소진을 분별하는 데 유용함을 확인하였다. 집단에 따른 최적의 성능을 보인 모델을 살펴보면 전체집단에서는 스택킹 모델이 정확도 .77, 정밀도 .75, 재현율 .77, 가중평균 (F1) .74, 가장 높은 성능을 보였다. 만 5세 이하 집단에서는 랜덤포레스트 모델이 정확도 .77, 정밀도 .75, 재현율 .77, 가중평균(F1) .75, 가장 높은 성능을 보였으며 만 6세에서 11세 집단에서는 인공신경망 모델이 정확도 .78, 정밀도 .76, 재현율 .78, 가중평균(F1) .77, 가장 높은 성능을 보였다. 마지막으로 만 12세에서 17세 집단에서는 랜덤포레스트 모델과 인공신경망모델 성능이 정확도 .73, 정밀도 .70, 재현율 .73, 가중평균 (F1) .71, 동일하게 나타났다.

ROC 곡선은 모델 성능을 시각적으로 보여주는 그래프로 모델들의 성능을 비교할 수 있다. 'Orange 3.6.1' 툴의 ROC 위젯을 사용하였으며 기본 값(500)으로 설정하였다. [그림 3]에 제시한 바와 같이 모델에 따른 ROC 결정곡선을 살펴보면 전체 데이터는 .64~.75 범위로 나타났다. 집단별로 살펴보면 만 5세 이하 집단에서는 .64~.75범위로 나타났다. 만 6세에서 11세 집단에서는 .75~.79이었고 만 12세에서 17세 집단에서는 .68~.74범위로 나타났다.

4.2 예측 변수의 중요도

집단에 따라 결과 변수에 영향을 미치는 예측 변수 중요도를 살펴보기 위해 'Orange 3.6.1' 툴의 Rank 위젯 Information Gain을 통해 중요도를 산출하였다. 예측변수의 중요도는 [그림 4.2]와 같다.





[그림 4.2] 예측변수 중요도

만 5세 이하 집단에서 상위 5개 중요도를 살펴보면 부모의 우울 .085, 양육효능감 .077, 자녀 훈육 .037, 부부싸움빈도 .035, 자녀 체벌 .034로 나타났다. 만 6세에서 11세 집단과 만 12세에서 17세 집단보다 부모의 우울 중요도가 높았다. 이러한 결과는 자녀의 연령에 따른 집단 차이가 있으며 영유아 자녀를 양육하는 부모집단에서 부모의 우울이 부모소진에 더 많은 영향을 미친다는 것을 의미한다. 만 6세에서 11세 집단의 상위 5개 중요도를 살펴보면 양육효능감 .063, 가족관계 .057, 자녀 훈육 .053, 자녀 체벌 .028, 부모의 우울 .027로 나타났다. 만 6세에서 11세 집단의 경우 자녀의 학령기로 인하여 부모역할의 차이가 있음을 보여주는 결과이며 양육효능감이 부모소진에 미치는 영향력을 확인한 결과이다. 부모소진은 자녀 발달 단계에 따른 부모역할 맥락에서 영향을 받는 변인임을 알 수 있다. 마지막으로 만 12세에서 17세 집단의 상위 5개 중요도를 살펴보면 양육효능감 .083, 가족관계 .056, 부모의 우울 .053, 자녀 훈육 .046, 가구구분 .027으로 나타났다. 양육효능감과 가족관계의 중요도는 만 6세에서 11세 집단과 동일하게 나타났다. 그러나 부모의 우울과 가구구분의 중요도는 자녀 연령에 따라 부모가 경험하는 어려움이 상이함을 확인한 결과이다.

5. 논의 및 제언

본 연구는 아동종합실태조사 자료를 활용하여 랜덤포레스트, 에이다부스트, 인공신경망, 스택킹 모델을 통해 부모소진 예측 모델을 개발하였다. 자녀 연령에 따라 세 집단으로 구분하여 모델 예측 성능을 비교하였으며 부모소진에 미치는 주요 변수의 중요도를 확인하였다. 그 결과 첫째, 부모소진을 예측하는 랜덤포레스트 모델, 에이다

부스트 모델, 인공신경망 모델, 스택킹 모델은 F1 값을 기준으로 모두 .70 이상의 예측력을 보였다. 집단에 따라 전체집단에서는 스택킹 모델, 만 5세 이하 집단에서는 랜덤포레스트 모델, 만 6세에서 11세 집단에서는 인공신경망 모델, 만 12세에서 17세 집단에서는 랜덤포레스트 모델과 인공신경망모델이 최적의 성능을 보였다.

둘째, 부모소진을 예측하는 중요도는 집단에 따라 상이하였다. 우선 만 5세 이하 유아를 양육하는 어머니 집단에서 부모소진을 예측하는 가장 중요한 변수는 부모의 우울이었다. 영유아기 자녀를 양육하는 어머니는 출산 전과 다른 일상생활 변화과정에서 24시간 자녀를 도맡아 돌봐야 하는 역할로 인하여 육체적, 심리적 어려움을 경험한다. 아동의 발달 특성상 어린 자녀를 돌보는 어머니는 수면이 부족하고, 때론 식사도 편하게 할 수 없으며 자녀와 잠시 떨어져 자신만의 시간을 갖는 것도 어렵다. 그래서 어머니의 우울은 출산 후 가장 흔하게 경험하는 부정정서로 여성의 50~80%에서 나타난다(O'Hara & McCabe, 2013). 임상적 연구와 조사에 따르면 어머니의 우울은 영유아기 자녀를 키우는 동안 지속되기도 하였다(김민정 외, 2018). 어머니의 우울은 영유아기 자녀를 양육하는 과정에서 부정적 사고와 무력감으로 인하여 양육 어려움을 가중시킨다. 그리고 어머니의 우울이 어머니와 자녀에게 부정적인 영향을 미친다는 선행연구와도 동일한 결과이다(이연실, 2017; Mikolajczak et al, 2019b).

양육효능감은 자녀양육에서 발생하는 어려움이나 문제를 대처하는 인지적 특성으로 만 6세에서 11세 아동을 양육하는 어머니 집단에서 가장 중요한 변수였다. 그리고 만 12세에서 17세 청소년을 양육하는 어머니 집단에서도 양육효능감이 가장 중요한 변수로 나타났다. 돌봄이 강조되는 유아기와 달리 자녀의 학령기는 어머니에게 일상생활지도, 인성지도, 행동지도, 학업지도, 또래관계지도까지 많은 역할이 요구된다. 그리고 자녀의 학업성취를 부모의 중요한 과업으로 여기는 한국의 사회적 풍토는 자녀를 통제하고 다루는 능력뿐만 아니라 자녀의 학업성취까지 부모의 책임으로 부과함에 따라 양육의 위기를 초래 할 수 있다. 이는 자녀의 학업 성취와 사회성이 어머니의 양육효능감에 직접적 영향을 미쳤으며 양육효능감이 낮을 경우 부모-자녀 관계 및 아동 정서에 부정적 영향을 미친다는 연구들과 맥을 같이하는 결과이다(최형성, 2005; Mousavi, 2020). 즉, 어머니의 정서와 인지적 요인이 부모소진을 예측하는 데 중요한 요인임을 확인하였다.

또한, 자녀훈육도 모든 집단에서 부모소진을 예측하는 변수로 나타났다. 훈육은 자녀 행동을 수정하고 바람직한 습관을 형성시키는 교육과정으로 자녀양육에서 불가결한 역할임에 분명하다. 하지만 어머니는 훈육과정에서 자녀의 저항, 우발적인 처벌, 부정적 감정으로 인하여 실망, 후회, 죄책감 등을 경험하고 이는 심리적 스트레스를 야기하는 선행연구와도 일치된 결과이다(김인지, 이숙현, 2012; Hubert & Aujoulat, 2018). 마지막으로 가족관계는 만 6세에서 11세 아동을 양육하는 어머니 집단과 만 12세에서 17세 청소년을 양육하는 어머니 집단에서 부모소진을 예측하는 주요 변수로 나타났다. 가족의 긍정적 상호작용은 역할분담을 수월하게 하고 부모역할에 대한 부담을 낮추어 양육 어려움을 감소시킨 연구들과 동일한 결과이다(최윤희, 문혁준, 2017; Mikolajczak et al, 2018b).

본 연구는 머신러닝 알고리즘을 적용하여 부모소진을 예측하는 변수들을 확인하였고

부모요인, 자녀양육요인, 가족환경요인에 의해 예측되는 다결정요인임을 검증하였다는 점에서 의의가 있다. 이는 자녀 발달과 부모역할을 고려한 양육 환경적 맥락에서 부모소진을 예방하기 위한 방안이 모색되어야 함을 의미한다. 부모 소진을 예방하기 위해서 일관적이고 형식적 지원이 아닌 자녀 연령에 따라 부모가 경험하는 어려움이 상이하다는 것을 인식하여 지지체계 및 지원에 초점을 맞추어야 할 필요가 있다. 그리고 자녀 연령에 따라 부모소진에 미치는 부모의 우울, 양육효능감, 자녀 훈육, 가족관계의 중요도를 확인하였고 예측변수의 차이와 유사점을 밝혔다라는 점에서도 기존 연구와의 차별된 점이다. 이러한 결과를 통해 건강한 가정 복지 지원차원에서 부모소진으로 어려움을 호소하는 부모를 지원하기 위해 자녀 연령에 따라 개입 방향성 및 효과적인 프로그램을 계획할 수 있을 것이다. 마지막으로 본 연구는 부모소진을 예측하는 모델을 구축하고 최적의 성능을 검증하였다는 점에서 의의가 있다. 구축된 모델을 통해 부모소진을 예방적 차원에서 선별하고 지원할 수 있는 시스템 도입 및 부모 복지와 건강한 가족 지원에 실증적 자료로 활용할 수 있을 것이다.

본 연구는 다음과 같은 몇 가지 제한점을 가지고 있으며 후속연구를 위한 제안은 다음과 같다. 첫째, 아버지의 응답수가 작아 어머니만을 대상으로 하였다는 점에서 한계가 있다. 추후 연구를 통해 아버지를 포함한 부, 모 모두를 살펴 볼 필요가 있다. 둘째, 랜덤포레스트 모델, 에이다부스트 모델, 인공신경망 모델, 스택킹 모델을 통해 부모소진을 예측하는 변수와 중요도를 살펴볼 수 있었으나 예측변수들 간의 상호관계성을 확인할 수는 없었다. 따라서 후속연구에서는 구조방정식 모델 검증 분석을 통해 예측변수들의 인과성을 확인 할 필요가 있다. 셋째, 본 연구에서 사용한 네 가지 머신러닝 알고리즘 이외에 그라디언트 부스팅(Gradient Boosting)이나 SVM 등 다양한 모델 탐색을 제안한다. 마지막으로 부모다움은 한국 가족제도 속에서 구성되어 온 실체이며 부모들에게 희생과 헌신의 당위성을 통해 지나친 책임이 강요되고 있다. 이러한 맥락에서 부모소진은 부모 개인만의 문제가 아니며 사회적 위험신호임을 주목하여야 할 것이다. 그리고 선별적이고 예방적 차원에서 부모소진을 예측할 수 있는 시스템 도입과 건강한 가족 복지 지원을 위해 현실적으로 필요한 양육지원개입과 정책방안 논의 및 실증적인 연구가 필요하다.

“본 논문은 통계청이 주관하는 「2021년 제19회 통계청 논문 공모전」에서 수상작으로 선정된 논문(을 수정 보완한 논문)입니다.”

(2021년 9월 3일 접수, 2021년 9월 24일 수정, 2021년 10월 15일 채택)

참고문헌

- 김기현, 강희경 (1997). 양육스트레스 척도의 개발, <대한가정학회>, 35(5), 141-150.
- 김미숙, 김기현, 김진석 (2015). <아동종합실태조사 심층 분석 연구 2013>, 보건복지부, 서울.
- 김민정, 박진아, 성웅형, 홍수중, 이경숙 (2018). 산후우울 중단 코호트 연구: 모의 정신건강 및 모-자녀관계, <영유아아동정신건강연구>, 11(2), 53-73.
- 김인지, 이숙현 (2012). 어머니의 자녀 훈육 경험, <가족과 문화>, 24(2), 1-138.
- 김종두, 오가영 (2018). 양육스트레스가 부모 소진에 미치는 영향: 가정어린이집 원아 어머니를 대상으로, <인문사회21>, 9(3), 911-922. <DOI: 10.2143/HS21.9.3.67>
- 신숙재, 정문자 (1998). 어머니의 양육스트레스, 사회적 지원과 부모효능감이 양육행동에 미치는 영향, <한국아동학회>, 19(1), 27-42.
- 양옥경, 김연수 (2007). 축약형 가족관계척도 구성과 타당도 연구, <한국가족관계학회지>, 12(2), 103-129.
- 엄문설, 이양희 (2020). 한국판 부모 소진 척도 타당화 연구, <위기관리 이론과 실천>, 16(10), 119-133. <DOI: 10.14251/crisisonomy.2020.16.10.119>
- 이봉주, 광금주, 구인회, 김민화, 김선숙, 김지현, 김혜란, 박혜준, 신혜은, 윤명희, 이강이, 이광호, 정익중, 진미정 (2008). <한국 아동청소년 종합실태조사>, 보건복지부, 서울.
- 이연실 (2017). 유아기 자녀를 둔 어머니가 지각한 우울과 불안 성향이 자녀 양육유형에 미치는 영향, <한국자치행정학보>, 31(4), 309-334.
- 최윤희, 문혁준 (2017). 유아기 자녀를 둔 어머니의 양육스트레스와 온정적 양육행동의 관계에서 가족상호작용의 매개효과, <유아교육학논집>, 21(6), 417-435.
- 최형성 (2005). 남녀아 어머니 양육효능감의 매개적 역할: 양육행동의 인과관계 모형, <인간발달연구>, 12(4), 121-137.
- 한수희 (2019). 영유아기 자녀를 둔 어머니의 양육소진에 관한 합의적 질적 연구, <한국가족복지학>, 24(1), 69-98. <DOI: /10.13049/kfwa.2019.24.1.5>
- Berg, L. and Lyhagen, J. (1998). Short and Long-Run Dependency in Swedish Stock Returns, *Applied Financial Economics*, 8, 435-443.
- Conley, T.G., Hansen, L.P., Luttmer, E.G.J. and Scheinkman, J.A. (1997). Short-term interest rates as subordinated diffusions. *The Review of Financial Studies*, 10, 525-577.
- http://nso.go.kr/newcms/webjin/journal/journal_info.htm
- Rao, J.N.K. (2003). *Small Area Estimation*, John Wiley & Sons, New York.
- Abidin, R.R. (1995). *Parenting Stress Index: Professional Manual(3 ed)*, Odessa, FL: Psychological Assessment Resources.
- Bastiaansen, C., Verspeek, E. and Bakel, H.V. (2021). Gender Differences in the

- Mitigating Effect of Co-Parenting on Parental Burnout: The Gender Dimension Applied to COVID-19 Restrictions and Parental Burnout Levels, *Social sciences*, 10(4), 127-137. <DOI: 10.3390/socsci10040127>
- Beiming, Y., Bin-Bin, C., Yang, Q. and Yuanfei, Z. (2021). Impacts of Parental Burnout on Chinese Youth's Mental Health: The Role of Parents' Autonomy Support and Emotion Regulation, *Journal of Youth and Adolescence*, <DOI: 10.1007/s10964-021 -01450-y>
- Breiman, L. (2001). Random Forests, *In Machine Learning*, 45(1), 5-32.
- Cantril, H. (1965). *The pattern of human concerns*, New Brunswick NJ: Rutgers University Press.
- Chase-Lansdale, P.L. and Pittman, L. (2002). Welfare reform and parenting: Reasonable expectations, *Children and Welfare Reform*, 12(1), 166-185.
- Cummins, R.A., Eckersley, R., Pallant, J., VanVaguet, J. and Misajon, R. (2003). Developing a National Index of Subjective Wellbeing: The Australian Unity Wellbeing Index, *Social Indicators Research*, 64(3), 159-190.
- Demsar, J., Curk, T., Erjavec, A., Gorup, C., Hocevar, T., Milutinovic, M., Mozina, M., Polajnar, M., Toplak, M., Staric, A., Stajdohar, M., Umek, L., Zagar, L., Zbontar, J., Zitnik, M. and Zupan, B. (2013). Orange: Data Mining Toolbox in Python, *Journal of Machine Learning Research*, 14, 2349-2353.
- Enkeu, A. and Enung, H. (2019). Burnout of Parent Who Have Children with Cerebral Palsy Disability in Bandung City, *Global Journals*, 19(2), 24-31.
- Gerain, P. and Zech, E. (2018). Does informal caregiving lead to parental burnout? Comparing parents having(or not) children with mental and physical issues, *Front in Psychol*, 884(9), 1-10. <DOI: 10.3389/fpsyg.2018.00884>
- Grant, K.A., McMahon, C. and Austin, M.P. (2008). Maternal anxiety during the transition to parenthood: A prospective study, *Journal of Affective Disorders*, 108, 101-111.
- Hatice, P. and Sibel, A.K. (2019). An Examination of the Perceived Social Support, Burnout Levels and Child-Raising Attitudes of Mothers of Children Diagnosed with Attention Deficit Hyperactivity Disorder in Turkey, *International Journal of Caring Sciences*, 12(2), 1165-1174.
- Hopfield, J.J. (1982). Neural networks and physical systems with emergent collective computational abilities, *PNAS*, 79(8), 2554-2558. <DOI: 10.1073/ pnas.79.8.2554>
- Hubert, S. and Aujoulat, I. (2018). Parental Burnout: When Exhausted Mothers Open Up, *Frontiers in Psychology*, 1021(9), 1-9. <DOI: 10.3389/fpsyg.2018.01021>
- Lavee, Y. McCubbin, H.I. and Patterson, J.M. (1985) The Double ABCX Model of

- Family Stress and Adaptation: An Empirical Test by Analysis of Structural Equations with Latent Variables, *Journal of Marriage and Family*, 47(4), 811-825.
- Lin, G-X., Szczygiel, D., Hansotte, L., Roskam, I. and Mikolajczak, M. (2021). Aiming to be perfect parents increases the risk of parental burnout, but emotional competence mitigates it, *Article de recherche*, 294(3), 1-33. <DOI: 10.1007/s12144-021-01509-w>
- Lindahl, N.A., Mellgren, K., Winiarski, J. and Forinder, U. (2014). Relationship between problems related to child late effects and parent burnout after pediatric hematopoietic stem cell transplantation, *Pediatric Transplantation*, 18, 302-309. <DOI: 10.1111/ petr.12228>
- Mishina, H., Takayama, J.I., Aizawa, S., Tsuchida N. and Sugama, S. (2012). Maternal childrearing anxiety reflects childrearing burden and quality of life, *Pediatrics International*, 54, 504-509.
- Mikolajczak, M., Brianda, M.E., Avalosse, H. and Roskam, I. (2018a). Consequences of parental burnout: Its specific effect on child neglect and violence, *Child Abuse & Neglect*, 80, 134-145. <DOI: 10.1016/j.chiabu.2018.03.025>
- Mikolajczak, M., Gross, J.J., Stinglhamber, F., Norberg, A.L. and Roskam, I. (2019a). Is parental burnout distinct from job burnout and depressive symptomatology?, *Clinical Psychological Science*, 8(4), 1-50. <DOI: 10.1177/2167702620917447>
- Mikolajczak, M., Gross, J.J. and Roskam, I. (2019b). Parental burnout: What is it, and why does it matter?, *Clinical Psychological Science*, 7(6), 1319-1329. <DOI: 10.1177/2167702619858430>
- Mikolajczak, M., Raes, M.E., Avalosse, H. and Roskam, I. (2018b). Exhausted parents: Sociodemographic, child-related, parent-related, parenting and family-functioning correlates of parental burnout, *Journal of Child and Family Studies*, 27(2), 602-614. <DOI: 10.1007/s10826-017-0892-4>
- Mousavi, S.F. (2020). The Burnout and Exhaustion Due to Parental Responsibilities The Role of Parent-Child Demographic Variables, *MEJDS*, 9, 123-133.
- OECD (2015). *How's Life? 2015: Measuring Well-Being*, OECD.
- O'Hara, M.W. and McCabe, J.E. (2013). Postpartum Depression: Current Status and Future Directions, *Clinical Psychology*, 9, 379-407. <DOI:10.1146/annurev-clinpsy-050212-185612>
- Pass, L., Mastroyannopoulou, K., Coker, S., Murray, L. and Dodd, H. (2017). Verbal Information Transfer in Real-Life- When Mothers Worry About Their Child Starting School, *Journal of child and family studies*, 26(8), 2324-2334.
- Radloff, L.S. (1977). The CES-D scale: A self report depression scale for research

- in the general population, *Applied Psychological Measurement*, 1, 385-401.
- Roskam, I., Aguiar, J., Akgun, E., Arikan, G., Artavia, M., Avalosse, H. and Mikolajczak, M. (2021). Parental burnout around the globe: a 42-country study, *Affective Science*, 2, 58-79. <DOI:10.1007/s42761-020-00028-4>
- Roskam, I., Brianda, M.E. and Mikolajczak, M. (2018). A step forward in the conceptualization and Measurement of Parental Burnout: the parental burnout assessment(PBA), *Frontiers in Psychology*, 9(758), 1664-1078. <DOI: 10.3389/ fpsyg.2018.00758>
- Roskam, I., Raes, M.E. and Mikolajczak, M. (2017). Exhausted Parents: Development and Preliminary Validation of the Parental Burnout Inventory, *Frontiers in Psychology*, 163(8), 1~2. <DOI: 10.3389/fpsyg. 2017.00163>
- Sebastian, R and Vahid, M. (2017) Python Machine Learning - Second Edition. 박해선(역) (2019). 머신러닝 교과서 with 파이썬, 사이킷런, 텐서플로. 서울: 길벗.
- Slavka, M., Martina, R. and Alena, S. (2020). Burnout in parents of sick children and its risk factors: a literature review, *Cent Eur J Nurs Midw*, 11(4), 196-206. <DOI: 10.15452/CEJNM.2020.11.0015>
- Sorkkila, M. and Aunola, K. (2020). Risk factors for parental burnout among Finnish parents: the role of socially prescribed perfectionism, *Journal of Child and Family Studies*, 29(3), 648-659. <DOI: 10.1007/s10826-019-01607-1>
- Syazwani, A.M., Iylia, M., Hazhari, I. and Nurul, I.Y. (2020). Working Parents And Emotionally Parental Burnout During Malaysia Movement Control Order(Mco), *European Journal of Molecular & Clinical Medicine*, 7(2), 4930-4953.
- Syarif, I., Zaluska, E., Prugel-Bennett, A. and Wills, G. (2012). Application of bagging, boosting and stacking to intrusion detection, *International Workshop on Machine Learning and Data Mining in Pattern Recognition*, 7376(8), 593-602. <DOI: 10.1007/ 978-3-642-31537-4_46>
- Yoav, F. and Robert, E.S. (1997) A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting, *Journal of Computer and System Sciences*, 55(1), 119-139. <DOI: 10.1006/jcss.1997.1504>

Application of Machine Learning Algorithm to Predict Parental Burnout

Yun-Yong Eom³⁾ . Wonyoung Song⁴⁾

Abstract

The purpose of this study is to predict parental burnout using machine learning algorithms. To this end, 2,659 mothers with children under the age of 17 were analyzed from the comprehensive survey data of children in 2018. The relationship between 13 predictors and parental burnout was established and Random forest, Adaboost, Neural Network, and Stacking models were applied. Studies have shown that all four models are F1.70 or higher. Among the variables predicting parental burnout, parenting efficacy, parental depression, child discipline, and family relationships had important effects. Parental burnout is a multi-determined syndrome by three sets of factors: Parent factors, child rearing factors and family environmental factors. Through this study, it is meaningful in that it has prepared empirical data and foundations to prevent parental burnout.

Key words : Parental Burnout, Random forest, Adaboost, Neural Network, Stacking

3) First Author, Completion of Doctoral Course, Konyang Univ., 107 Chung-ro, Dongnam-gu, Cheonan-si, Korea. Email: eyy83@hanmail.net

4) Corresponding Author, Professor, Konyang Univ., 121 Daehak-ro, Nonsan-si, Chungcheongnam-do, Korea. E-mail: song@konyang.ac.kr