

「기계학습을 통한 양육 스트레스 관련 변수 탐색」 토론문

김은정(한국보건사회연구원 인구모니터링평가센터 부연구위원)

본 연구는 별점회귀모형을 사용하여 어머니의 양육 스트레스와 관련된 요인을 탐색하였다. 본 연구는 회귀모형의 단점을 보완하여 800개가 넘는 변수를 동시에 투입·분석하였고, 기존 연구에서 발견하지 못한 새로운 요인들을 발견했다는 점에서 우수하다. 다만 몇가지 점에서 부문에서 보완됐으면 하는 점들이 있다.

첫째, 본 연구는 한국 영유아 교육·보육 패널 1차년도(생후 4개월 이후 만 0세 아동) 데이터를 활용하여 변수들을 살펴보았는데, 육아정책연구소에서는 1차년도 데이터와 함께 임신 시기 기초조사 및 생후 1개월 조사 데이터도 같이 제공하고 있는 것으로 알고 있다. 해당 데이터들도 함께 활용하면 더 풍부한 결과를 얻을 수 있을 것으로 판단된다. 예를 들어, 기초조사에서는 해당 자녀가 계획 임신이었는지, 그리고 임신기간 동안 자녀교육을 수강한 경험이 있는지 등을 조사하고 있는데, 해당 변수들도 같이 살펴보면 좋겠다.

둘째, 패널에서 제공하고 있는 원 변수 이외에도 가공 변수를 만들어 분석하는 방법도 고민하면 좋겠다. 예를 들어, 조절변수 등을 만들어 변수 간의 상호작용(interaction) 효과도 함께 분석하는 방법도 고민하면 좋겠다. 예를 들면, 자녀의 성별과 어머니 연령을 교차하여, 여아를 키우는 30대 이하 어머니와 남아를 키우는 30대 이하 어머니의 양육 스트레스를 비교하거나, 여아를 키우는 30대 이하 어머니와 40대 이상 어머니의 양육 스트레스 등을 비교해 보면 좋겠다.

셋째, 본 연구는 1차년도 자료만을 살펴본 단일 시점 연구인데, 이는 현재 육아정책 연구소에 1차년도 데이터만 제공하고 있기 때문으로 알고 있다. 이에, 보완책으로 아동 월령으로 나누어 살펴보는 방법도 고민해 보길 바란다. 예를 들어 5개월 이하, 6개~7개월, 8개월 이상으로 나누어, 월령에 따라 어머니의 양육 스트레스 요인이 어떻게 다른지를 비교할 수 있을 것이다. 아동발달학적 관점에서 생후 6~7개월부터는 아동이 기기 시작하는 시기로 어머니의 스트레스 요인이 유의미하게 달라질 가능성이 높다.

넷째, 본 연구에서는 양육행동을 양육 스트레스의 예측 변수로 설명하고 있는데, 오히려 반대일 가능성이 더 높다. 즉, 역인과 관계일 가능성이 더 높다. 자녀를 따뜻하게

대하기 때문에 양육 스트레스가 낮은 것이 아니라, 어머니의 양육 스트레스가 낮기 때문에 자녀를 따뜻하게 대하는 것일 수 있다.

마지막으로, 저자는 본 연구의 한계로 Elastic net 대신 Group lasso 또는 Group Mnet을 사용하여 분석을 진행할 수도 있다고 썼는데, Elastic net이 Group lasso 및 Group Mnet보다 더 좋은점과 나쁜점이 무엇이며, 그럼에도 불구하고 왜 본 연구에서는 Elastic net을 사용했는지 설명이 보완될 필요성이 있다.