

영아 미디어 노출 유형과 분류 요인: 잠재프로파일 분석과 랜덤포레스트를 활용한 탐색*

구채원¹⁾ 정정희²⁾

요약

본 연구는 1세 영아의 미디어 노출 유형을 분류하고 각 유형을 분류하는 주요 요인을 탐색하는 것을 목적으로 한다. 이를 위해 「한국 영유아 교육·보육 패널」 2차년도 자료의 1,486가구를 대상으로 영아의 개인 특성, 부모 요인, 가구 환경 등 총 106개의 분류 변수를 활용하였다. 이후 잠재프로파일 분석을 통해 미디어 노출 유형을 도출하고 One-vs-One(OvO) 방식의 랜덤포레스트 모형을 적용하였다. 연구 결과, 미디어 노출 유형은 '미이용', '후기·저이용', '조기·고이용'의 세 가지로 나타났다. 랜덤포레스트 모형의 AUC는 .621~.759 범위였으며, '미이용 vs 조기·고이용' 모형이 가장 양호한 분류 성능을 보였다. 주요 변인으로는 영아의 수면·식습관, 아버지의 흡연 여부, 지역사회 양육환경 만족도 등이 확인되었다. 본 연구는 영아기 미디어 노출이 다층적 요인과 관련되어 있음을 보여주며, 아버지의 양육 참여 확대와 지역 기반의 양육 환경 개선 등 구체적 실천 방안을 제시한다.

주제어: 영아 미디어 노출, 잠재프로파일 분석, 랜덤포레스트, 머신러닝

I. 서론

최근 미디어 이용 연령이 급격히 하향화되고 있다. 2013년 국내 영유아의 스마트기기 최초 이용 평균 연령은 2.27세였으나(이정림·도남희·오유정, 2013), 10년이 지난 2023년에는 생후 12개월 무렵부터 디지털 미디어를 접하기 시작하는 것으로 나타났다(배운진·박

* 본 논문은 제16회 한국아동패널 & 제3회 한국 영유아 교육·보육 패널 공동학술대회(2025. 8. 29)에서 발표한 논문을 수정·보완한 것임.

1) 경북대학교 아동학부 박사과정

2) 경북대학교 아동학부 교수

은정·김자연 외, 2023). 조기 노출뿐 아니라 장시간 노출 역시 주요한 우려 요소로 지적된다. 특히 미디어 이용 시간 기준 위험군의 비율이 영아기에서 15%에 달해 유아기보다도 과다 노출 가능성이 높은 시기로 확인되었다(오주현·박용완, 2019). 이처럼 미디어 이용 시기가 돌 전후까지 앞당겨지고 고이용 집단도 증가하면서, 미디어는 ‘공갈 젖꼭지’나 ‘베이비시터’로 불릴 만큼 일상적인 육아 도구로 자리 잡았다(김연수·정윤경, 2018).

영아기의 미디어 조기 및 과다 노출은 언어, 정서, 인지, 운동, 사회성 등 거의 모든 주요 발달 영역에 부정적인 영향을 미치는 것으로 보고되고 있다. 예컨대 1세 영아를 대상으로 한 연구에서 미디어 노출 시간이 길수록 수용언어와 표현언어 점수가 낮았으며(강민아·이영신, 2020), 스마트기기의 최초 이용 시점이 빠를수록 운동 발달과 자조 행동에 부정적인 영향을 미친다는 결과도 제시되었다(성지현·변혜원·남지혜, 2015). 이는 미디어의 시각 중심 자극이 고차원적 사고 및 자기조절을 담당하는 전두엽의 활성화를 충분히 유도하지 못하고, 미디어 이용 시간이 증가함에 따라 놀이, 부모-자녀 상호작용, 또래 관계를 통한 발달 경험이 제한되기 때문으로 해석된다(이경숙·정석진·김명식, 2015). 즉, 영아기 미디어 노출은 단순한 이용 행태에 그치지 않고 이후 전반적 발달 경로를 재구성하는 요인으로 작용할 수 있다는 점에서 그 심각성이 더욱 크다.

이처럼 영아기 미디어 노출의 발달적 위험성이 분명함에도 불구하고, 국내 선행연구는 여전히 청소년기와 유아기에 집중되어 영아기 미디어 이용 실태와 그 하향화 추세를 충분히 반영하지 못하고 있다. 영아기는 ‘후속 연구’로만 언급될 뿐, 실증적 탐색은 부족하여 여전히 연구의 사각지대로 남아 있다(조준오·김은정, 2019). 특히 영아기 미디어 노출이 왜, 어떻게 발생하는지에 대한 탐색은 여전히 제한적이며, 조기·과다 노출의 예측 요인을 규명하려는 시도 역시 드물다. 따라서 미디어 노출 양상과 관련 요인을 실증적으로 규명하는 연구가 요구된다.

선행연구를 살펴보면 영아 미디어 노출에 영향을 주는 요인은 크게 영아 개인 특성 요인, 부모 요인, 가정 환경 요인으로 구분된다. 이 중 영아 개인 특성 요인으로는 생리적 규칙성, 정서 반응성과 같은 기질의 하위영역이 주요하게 지목된다. 가령 9개월 시점에서 잠 들기 어려워하거나 자주 깨는 등의 수면 관련 생리적 규칙성에 결함을 보인 영아는 이후 더 많은 미디어 시청 시간을 보유했다(Radesky, Silverstein, Zuckerman et al., 2014). 또한 정서 반응성과 관련해서 부정적 정서가 높은 영아의 경우, 미디어 이용 시간이 유의하게 길다는 결과도 보고되었다(Coyne, Shawcroft, Gale et al., 2021). 이러한 결과는 부모가 기질적으로 까다로운 영아를 진정시키거나 주의를 전환시키기 위한 수단으로 미디어를 활용했을 가능성을 시사한다.

하지만 영아의 미디어 이용에 가장 큰 영향을 미치는 요인은 부모이다. 청소년과 달리

영아는 미디어 기기를 직접 소유하지 않기에, 대부분 부모의 기기를 통해 간접적으로 접근한다. 이에 따라 부모는 실질적인 ‘게이트키퍼’로 기능하며 영아기 자녀의 미디어 접근 여부에 중요한 역할을 수행한다(이병호, 2012). 실제로 영유아 미디어 사용 목적을 분석한 연구에서 ‘공공장소에서 조용히 시키기 위해’, ‘일을 방해 없이 하기 위해’ 같은 도구적 목적이 높은 비율로 나타났다(이정원·박원순·엄지원, 2021). 이는 영아의 미디어 첫 노출과 이후 이용 양상에는 부모의 선택과 필요가 관련될 가능성을 시사한다(오주현·박용완, 2019).

이때 부모 요인은 주로 어머니를 중심으로 연구되어 왔으며, 특히 양육 스트레스, 양육 효능감, 미디어 중재 전략이 다루어졌다. 예를 들어 어머니의 양육 스트레스가 높을수록 자녀의 미디어 사용을 방임적으로 중재하여 결과적으로 자녀의 미디어 과의존 가능성이 높아지는 경향이 관찰되었다(장지우·김선희, 2022). 반면 양육 스트레스가 낮고 양육 효능감이 높은 어머니일수록 영아 자녀가 스마트폰을 전혀 사용하지 않는 경우가 많았다(류미향, 2014). 이는 양육 스트레스가 높고 효능감이 낮은 어머니일수록 자녀와의 직접적 상호작용에 어려움을 겪고 미디어를 양육 대체 수단이나 정서적 완충 장치로 활용하는 경향을 보여준다.

한편, 아버지 관련 요인은 비교적 적게 연구되었지만, 어머니의 양육을 보조하는 맥락에서 다루어진 바 있다. 아버지가 자녀의 양육과 놀이에 적극적으로 참여할수록 자녀의 미디어 사용 시간을 보다 감소시킬 수 있다는 결과가 보고되었다(이소현·김윤희, 2020; 이지인, 2020). 요컨대, 영아의 미디어 노출은 부모-특히 어머니-의 양육 행동과 관련하여 연구되어 왔으며, 이는 미디어가 오락 수단을 넘어 일상적 양육 맥락 속에서 전략적으로 활용되고 있음을 시사한다.

이와 더불어, 가구의 전반적인 환경 역시 영아의 미디어 이용에 영향을 미치는 중요한 맥락으로 작용한다. 대표적으로는 소득 수준에 따라 미디어 활용 양상이 다르며, 고소득 가구는 교육적 콘텐츠를, 저소득 가구는 돌봄 수단으로 활용하는 경향이 있다(Nikken & Oprea, 2018). 또한, 가정의 사회적 지지망의 유무도 영향을 미친다. 육아에 대한 지원 인력이 없는 가정에서는 미디어가 일종의 ‘가상 보모’ 역할을 수행하며 영아의 미디어 이용 빈도가 잦았다(Cingel & Krcmar, 2013). 아울러, 문화시설이나 여가시설의 접근성 또한 영아의 미디어 사용에 영향을 주는 요인으로 밝혀졌다. 관련 연구에 따르면 공공문화 및 여가 인프라 이용이 용이한 지역일수록 미디어 이용 시간이 줄어드는 경향이 있었다(김대웅·김창숙·박윤현, 2019). 이처럼 가정의 물리적·지역적 조건은 자녀의 미디어 이용 양상과 관련될 수 있으며, 이러한 관계에는 부모의 양육 여건이나 양육 전략이 함께 고려될 필요가 있다. 이는 양육 자원과 환경적 여건이 충분하지 않은 상황에서 미디어가 부모의 양육 부담을 일시적으로 덜어주는 이른바 ‘공갈 젖꼭지’와 유사한 역할을 할 가능성을 시

사한다.

이처럼 영아기 미디어 노출은 단일한 원인으로 환원할 수 없는 복합적 양상으로, 영아의 개인 특성, 부모, 가정 환경 등 다양한 수준의 요인이 작용하는 다층적 현상이다. 그러나 기존 연구는 미디어 이용의 부정적 결과를 단편적으로 경고하거나, 제한된 요인에 집중해 조기 및 과다 노출이 발생하는 구체적 경로와 구조적 조건을 총체적으로 조망하는 데에는 한계가 있다. 특히 어머니 요인에 편중된 기존 접근은 미디어 과의존을 개인화된 양육 문제로 축소할 위험이 있다.

이에 본 연구는 영아기 미디어 노출 현상을 입체적으로 이해하기 위해, 잠재프로파일 분석으로 미디어 노출 유형을 도출하고, 랜덤포레스트를 활용해 분류 요인을 탐색하였다. 잠재프로파일 분석은 이질적 하위 집단을 식별할 수 있는 사람 중심 통계 기법으로, 노출 시간이라는 단일 지표에만 의존했던 기존 분석의 한계를 보완할 수 있다(이혜림·박재완, 2018). 랜덤포레스트는 의사결정나무 기반 머신러닝 기법으로 변수 간의 비선형 관계와 상호작용을 반영하고 다중공선성에 강인하다는 점에서 복합적 구조 분석에 유리하다(Breiman, 2001). 본 연구에서는 영아의 개인 특성, 부모 요인, 가정 환경 등 총 106개의 분류 요인을 투입하였고, 주요 변인의 분류 확률 변화 양상을 파악하기 위해 부분의존도 분석을 병행하였다. 본 연구는 영아 미디어 노출 유형을 분류하고 분류 요인을 종합적으로 탐색함으로써 고위험군에 대한 선별적 대응의 기초자료를 제공하는 것을 목적으로 한다. 이러한 배경하에 본 연구는 다음과 같은 연구문제를 설정하였다.

- 연구문제 1. 영아의 미디어 노출 여부, 최초 노출 시기, 주중·주말 노출 시간을 기준으로 미디어 노출 유형은 몇 개로 분류되며, 각 집단의 특성은 어떠한가?
- 연구문제 2. 영아 미디어 노출 유형을 구분하는 주요 분류 요인은 무엇이며, 상대적 중요도는 어떠한가?
- 연구문제 3. 도출된 주요 분류 요인은 영아 미디어 노출 유형을 어떻게 설명하는가?

II. 연구방법

1. 연구대상

본 연구는 2023년 육아정책연구소에서 실시한 「한국 영유아 교육·보육 패널(K-ECEC-P: Korean Early Childhood Education & Care Panel Study)」 2차년도 자료를 활용하였

다. 2차년도에는 1세 영아의 주양육자(2,555부), 어머니(2,042부), 아버지(1,473부), 담임 교사(309부)를 대상으로 조사되었으며, 본 연구에서는 영아의 미디어 노출 양상과 관련된 요인을 다각적으로 탐색하고자 주양육자, 어머니, 아버지가 모두 응답한 1,486부의 자료를 최종 분석에 사용하였다. 연구대상의 일반적 특성은 <표 1>에 제시하였다.

<표 1> 연구대상의 일반적 특성

(N=1,486)

구분		빈도(%)	구분		빈도(%)		
영아	월령	16~17개월	354(23.8%)	가구소득	200만원 이하	28(1.9%)	
		18~19개월	574(38.6%)		201~400만원	627(42.2%)	
		20~21개월	362(24.4%)		401~600만원	567(38.2%)	
		22~23개월	196(13.2%)		601~800만원	155(10.4%)	
	성별	남	729(49.1%)		801~1000만원	87(5.9%)	
		여	757(50.9%)		1001만원 이상	22(1.5%)	
아버지	연령	20대	75(5.0%)	어머니	연령	20대 미만	1(0.1%)
		30대	1048(70.5%)			20대	153(10.3%)
		40대	341(22.9%)			30대	1172(78.9%)
		50대 이상	15(1.0%)			40대	159(10.7%)
		무응답	7(0.5%)			50대 이상	1(0.1%)
	최종 학력	중졸 이하	8(0.5%)	최종 학력	중졸 이하	15(1.0%)	
		고졸	261(17.6%)		고졸	198(13.3%)	
		전문대졸	252(17.0%)		전문대졸	274(18.4%)	
		대졸	835(56.2%)		대졸	859(57.8%)	
		대학원졸	130(8.7%)		대학원졸	140(9.4%)	

2. 측정도구

가. 종속변인

본 연구의 종속변인은 ‘영아 미디어 노출 유형’으로, 이는 유튜브, 애니메이션 등 동영상 시청을 목적으로 한 미디어 이용 여부 및 이용 특성을 바탕으로 구성되었다. 우선, ‘동영상 시청(애니메이션, 유튜브) 용도로 미디어를 이용’이라는 문항에 ‘이용하지 않음’이라고 응

답한 481명은 ‘미이용’ 집단으로 분류하였다.

반면 ‘이용’이라고 응답한 1,005명의 영아는 이용 집단으로 분류되었으며, 이들을 대상으로 ‘이용 시작 월령’, ‘주중 1일 이용 시간’, ‘주말 1일 이용 시간’의 세 변수를 활용하여 잠재프로파일 분석을 실시하였다. 따라서 ‘미이용’ 집단에 이용 집단의 잠재집단을 추가하여 영아 미디어 노출 유형을 구성하였으며, 이렇게 구성된 세 집단을 랜덤포레스트 분류 모형의 종속변인으로 설정하였다.

나. 분류변인

영아의 미디어 노출 유형을 분류하는 데 기여하는 변인을 탐색하기 위해 K-ECEC-P 조사에서 수집된 106개의 변인을 랜덤포레스트 모형에 투입하였다. 전체 변수 중 종속변수와 관련성이 낮다고 판단되거나 결측률이 20%를 초과하는 문항은 분석에서 제외하였다(Breiman, 2001). 분석에 투입된 변인의 결측률은 최대 5.2%로 비교적 낮은 수준이었다. 이에 본 연구에서는 R의 mice 패키지(Ver. 3.18.0)를 활용하여 연속형 변수는 평균값으로, 범주형 변수는 최빈값으로 대체하였다.

최종적으로 분석에 투입된 분류변인의 목록은 <표 2>에 제시하였다³⁾. 심리적 척도 중 스마트폰 과의존, 양육효능감과 같이 하위영역이 존재하는 경우에는 각 하위영역의 평균값을 산출하였고, 그릿, 자아존중감과 같이 단일 척도로 구성된 경우에는 전체 문항의 평균값을 산출하여 분석에 활용하였다. 반면 수면 특성, 식습관과 같이 각 문항의 의미가 독립적으로 해석되는 경우에는 개별 문항을 그대로 분석에 포함하였다. 또한, 도시규모, 맞벌이 여부 등 명목형 변수는 원-핫 인코딩(one-hot encoding) 방식으로 0과 1의 이진형 벡터로 변환하여 투입하였다. 결과적으로 영아 개인 변인 29개, 아버지 관련 변인 27개, 어머니 관련 변인 27개, 가구 관련 변인 23개로 구성된 총 106개의 분류변수가 랜덤포레스트 모형에 최종 투입되었다.

3) 기울임체는 복수 문항으로 구성된 척도 변인으로, 하위영역 또는 전체 문항의 평균값을 산출하여 분석에 활용한 변인을 의미함.

〈표 2〉 랜덤포레스트 모형에 투입한 분류 변인

구분		변인명
영아	인구학적 특성	월령, 성별, 출생순위
	일상생활 특성	수면 특성('밤새 10~12시간 잔다' 외 3개), 밤잠 시간, 밤잠 중 깨는 횟수, 낮잠 횟수, 낮잠 시간, 식생활 특성('컵을 잘 사용한다' 외 4개), 식습관('식사 시간은 규칙적입니까?' 외 4개), 기관 이용 여부
	건강·발달 특성	평소 건강, 장애 판정 여부, 대근육운동발달, 소근육운동발달, 인지발달, 언어발달, 사회성발달
아버지	인구학적·근로 특성	연령, 최종학력, 취업/학업상태
	건강 특성	장애 여부 및 정도, 흡연 여부, 음주습관, 스마트폰과의존('조절실패' 외 2개)
	심리·정서적 특성	평소 스트레스, 그릿, 자아존중감, 우울, 결혼만족도, 부부갈등
	양육 특성	양육부담('신체적 부담' 외 4개), 양육스트레스, 양육행동, 양육효능감('유능감' 외 2개), 공동양육 참여도, 공동가사 참여도
어머니	인구학적·근로 특성	연령, 최종학력, 취업/학업상태
	건강 특성	장애 여부 및 정도, 흡연 여부, 음주 습관, 스마트폰과의존('조절 실패' 외 2개)
	심리·정서적 특성	평소 스트레스, 그릿, 자아존중감, 우울, 결혼만족도, 부부갈등
	양육 특성	양육부담('신체적 부담' 외 4개), 양육스트레스, 양육행동, 양육효능감('유능감' 외 2개), 공동양육 참여도, 공동가사 참여도
가구	인구학적·일반적 특성	쌍둥이 가구 여부, 다문화 가구 여부, 맞벌이 여부, 가구원 수, 가구 형태
	경제적 특성	저소득 가구 여부, 주관적 사회경제적 지위, 가구소득
	물리적·지역사회환경 특성	가구평수, 도시규모, 지역사회 양육 환경 만족도, 사회적 지원 인원, 사회적 지원 정도('정서적 지원' 외 3개), 지역사회 환경에 대한 인식('지금 살고 있는 지역은 깨끗하다' 외 6개)

3. 분석방법

본 연구에서는 미디어 이용 집단 1,005명을 대상으로 Mplus 8.3을 활용하여 잠재프로파일 분석을 실시하였다. 분석지표는 이용 시작 월령, 주중 1일 이용 시간, 주말 1일 이용 시간이었으며, 1개부터 6개까지의 잠재집단 모형을 비교하였다. 최적 모형은 AIC, BIC, SABIC, Entropy, BLRT와 Elbow plot, 집단별 분류율을 종합적으로 고려하여 결정하였다. 이후 도출된 두 잠재집단에 '미이용' 집단을 추가하여 최종 세 가지 미디어 노출 유형

을 구성하고, 집단 간 이용 시작 시점과 이용 시간의 차이를 확인하였다.

이후 미디어 노출 유형을 분류하는 모형을 구축하기 위해 R 4.5.0 프로그램을 사용하여 랜덤포레스트 분석을 수행하였다. 본 연구에서 랜덤포레스트를 활용한 것은 106개의 분류 변인이 투입된 상황에서 영아 개인 특성, 부모 특성, 가구 및 지역사회 요인 간의 복합적인 관계를 탐색하기 위함이다. 랜덤포레스트는 전통적인 로지스틱 회귀모형과 달리 변수 간 관계를 선형적으로 가정하지 않으며, 다수의 분류 변인 사이에서 나타날 수 있는 비선형적 관계와 상호작용을 사전에 특정하지 않고도 반영할 수 있다. 또한 변수 중요도를 통해 잠재 집단 유형 분류에 상대적으로 크게 기여하는 요인을 탐색할 수 있다는 장점이 있다(Breiman, 2001). 다만, 이러한 접근은 전통적인 혼합모형의 3단계 접근법과 달리 잠재집단 분류 과정에서 발생할 수 있는 확률적 오차를 모형 내에 직접 반영하는 데에는 한계가 있다. 따라서 연구 결과는 개별 변인의 인과적 효과가 아니라 미디어 노출 유형을 구분하는 주요 분류 요인의 관점에서 해석하였다.

본 연구에서는 클래스 불균형 문제를 보완하기 위해 One-vs-One(OvO) 방식으로 이진 분류 모형을 구성한 후 smotefamily 패키지를 활용하여 소수의 클래스 데이터를 안정적으로 보완하였다(김길재·박강윤·문명현, 2023). 또한 과적합을 방지하기 위해 전체 자료는 학습용 70%와 검증용 30%로 분할하였으며, 그리드 서치와 10-fold 교차검증을 통해 하이퍼파라미터를 설정하였다. ntree는 100, 300, 500, 700, 1,000, 1,500을 비교한 결과, 세 개의 OvO 모형에서 모두 500으로 결정되었다. mtry는 분류 모형에서 일반적으로 활용되는 전체 분류 변인 수의 제곱근 기준을 적용하였으며, 이에 따라 최종 분류 변인 106개의 제곱근인 10으로 설정하였다(Breiman, 2001). 모형의 성능은 CA(Classification Accuracy), Precision, Recall, F1, ROC(Receiver Operating Characteristic) 곡선 및 AUC(Area Under the Curve)를 통해 평가하였다. 또한 MDG(Mean Decrease Gini)를 기준으로 상위 10개 변수의 상대적 중요도를 비교하였다.

마지막으로, 분류 성능이 가장 높고 위험군 분류에 효과적인 '미이용 vs 조기·고이용' 모형을 중심으로 부분의존성 도표(PDP: Partial Dependence Plot)를 활용해 주요 변수에 따른 분류 확률 변화 양상을 시각화하였다. PDP는 X축에 분류 변인 값, Y축에 해당 값에 따른 종속변인의 분류 확률을 나타내어 블랙박스 모델인 랜덤포레스트의 해석 한계를 보완하는 데 효과적이다(최필선·민인식, 2018).

III. 연구결과

1. 영아 미디어 노출에 대한 잠재프로파일 분석

가. 영아 미디어 노출의 잠재프로파일 수 결정

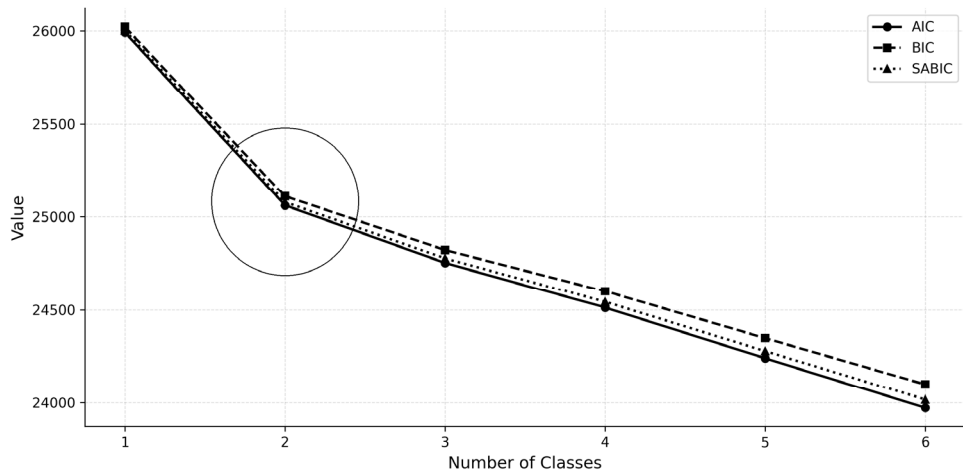
전체 1,486명의 영아 중 미디어를 이용한다고 응답한 1,005명을 대상으로 잠재프로파일 분석을 실시하였다. 잠재집단의 적절한 수를 결정하기 위해서 집단의 수가 1개부터 6개인 모형을 추정하고 각 모형의 적합도를 비교하였다. 적합도는 AIC, BIC, SABIC와 Entropy, BLRT의 검정값을 통해 확인하였으며, 그 결과는 <표 3>과 같다.

<표 3> 영아 미디어 노출의 잠재프로파일 적합도 지수

잠재 집단	AIC	BIC	SABIC	Entropy	BLRT	분류율(%)					
						1	2	3	4	5	6
1	25991.69	26021.16	26002.11			100					
2	25060.83	25109.96	25078.2	0.97	938.86***	88	12				
3	24753.84	24822.62	24778.15	0.98	314.99***	88	12	0.5			
4	24513.02	24601.45	24544.28	0.92	248.82***	62	27	10	0.5		
5	24239.96	24348.04	24278.16	0.97	281.06***	61	27	2	1	0.2	
6	23970.64	24098.37	24015.8	0.98	197.33***	61	27	9	2	1	0.4

*** $p < .001$

정보지수인 AIC, BIC, SABIC은 수치가 낮을수록 모형 적합도가 우수함을 의미하며, Entropy는 분류의 명확도를 나타내는 지표로서 1에 가까울수록 높은 분류 품질을 나타낸다. BLRT는 n 개의 집단으로 구성된 모형이 $n-1$ 개의 집단보다 유의하게 우수한지를 확인하는 검정으로 유의한 p 값은 분류력이 향상되었음을 뜻한다. 분석 결과, 본 연구에서는 집단 수가 증가함에 따라 AIC, BIC, SABIC 값이 꾸준히 감소하였으며 Entropy는 0.92~0.98로 모든 집단에서 우수한 값을 보였다. 또한 모든 단계에서 BLRT의 검정값이 유의하게 나타나, 집단 수를 증가시킨 모형의 적합도가 이전 모형보다 유의하게 개선되었다.



[그림 1] 잠재집단 수에 따른 정보지수의 Elbow plot

그러나 표본 수가 큰 데이터에서 집단 간 분산이 클 경우, 집단 수가 많아질수록 극단적 특성을 갖는 사례들이 새로운 집단으로 계속 분리되기도 한다. 따라서 모형 선택 시에는 [그림 1]과 같은 Elbow plot을 통해 정보지수 감소가 완만해지는 지점을 확인하고(김영서·홍세희, 2021), 각 집단의 분류율과 해석 가능성을 함께 고려할 필요가 있다. 본 연구에서는 최소 분류율이 1% 미만인 집단은 비교와 해석이 어렵다는 기준에 따라(Jung & Wickrama, 2008), 해당 집단이 포함된 3개 이상 모형을 제외하였다. 최종적으로는 정보지수가 완만해지는 경계에 위치하면서 최소 분류율이 12%로 나타난 잠재집단이 2개인 모형을 선정하였다.

나. 영아 미디어 노출 유형에 따른 집단별 특성

미디어를 이용하는 영아 1,005명을 대상으로 잠재프로파일 분석을 실시한 결과, 두 개의 잠재집단이 도출되었다. 잠재집단 1은 평균 13개월경 미디어 노출을 시작하였으며 1일 시청 시간이 약 30분으로 나타나 ‘후기·저이용’ 집단으로 명명하였다. 잠재집단 2는 평균 11개월경에 미디어 노출을 시작하여 1일 시청 시간이 약 130분으로 나타나 ‘조기·고이용’ 집단으로 명명하였다. 미디어를 전혀 이용하지 않는 481명의 영아는 ‘미이용’ 집단으로 분류하였다. 각 집단의 이용 시작 월령, 주중 1일 이용 시간과 주말 1일 이용 시간은 <표 4>와 같다.

〈표 4〉 영아 미디어 노출 유형과 특성

분류	N(%)	이용 시작 월령(개월)	주중 1일 이용 시간(분)	주말 1일 이용 시간(분)	
이용집단	후기·저이용 (잠재집단 1)	889(59.8%)	13.13 (3.19)	34.17 (20.65)	36.95 (26.11)
	조기·고이용 (잠재집단 2)	116(7.8%)	11.29 (3.50)	133.53 (50.05)	132.05 (56.24)
미이용 집단		481(32.4%)	-	-	-

2. 영아 미디어 노출 유형 집단에 대한 머신러닝

가. 영아 미디어 노출 유형 분류 모형의 성능 평가

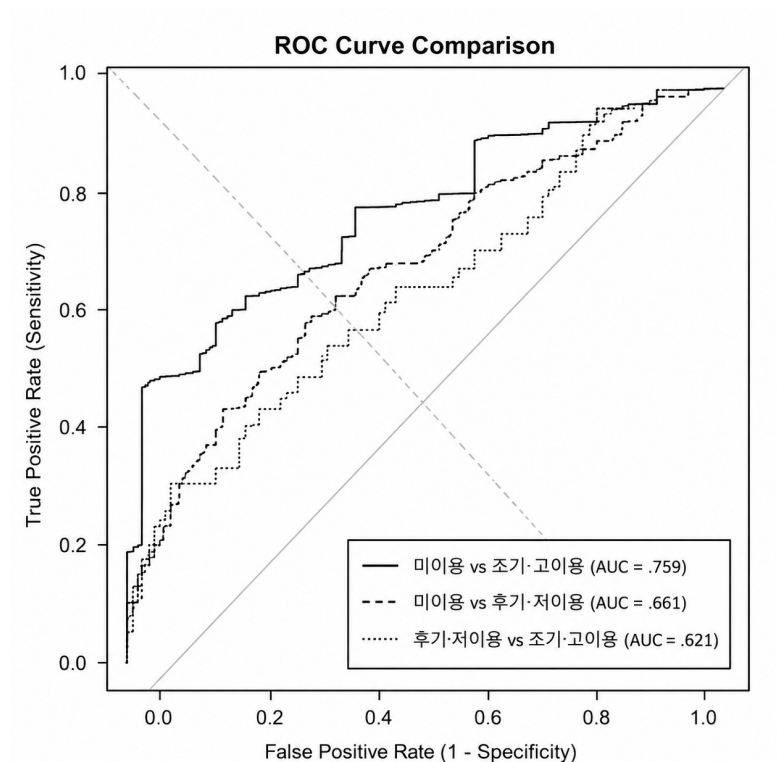
본 연구에서는 영아 미디어 노출 유형이 세 가지 집단으로 구분됨에 따라, 다중 분류 알고리즘을 적용할 수 있으나, ‘조기·고이용’ 집단이 전체의 7.8%로 비교적 소수임을 고려하여 OvO 방식으로 총 세 개의 이진 분류 모형을 구성하였다. OvO 방식은 클래스 수가 k 일 때, $k(k-1)/2$ 개의 이진 분류기를 구축하여 전체 문제를 단순화하며, 특히 소수 클래스가 포함된 불균형 데이터에 효과적인 분류 전략으로 알려져 있다(Hsu & Lin, 2002).

랜덤포레스트 알고리즘을 기반으로, 총 106개의 분류 변인을 투입하여 세 개의 이진 분류 모형을 구성하고 그 성능을 평가하였다. 성능 평가지표로는 CA, Precision, Recall, F1, AUC를 활용하였으며, 그 결과는 〈표 5〉와 같다. ‘미이용 vs 조기·고이용’ 모형은 전체 데이터의 78.9%를 정확히 분류하였으며(CA = .789), 분류된 ‘조기·고이용’ 집단 중 82.9%가 실제 ‘조기·고이용’ 집단으로 나타나 Precision이 높았다. 또한 실제 ‘조기·고이용’ 집단 중 92.3%를 올바르게 분류하여 Recall 역시 높았으며, F1 점수(.873)와 AUC(.759)도 우수하여 모든 지표에서 양호한 성능을 나타냈다.

〈표 5〉 미디어 노출 유형 분류 성능

모형	CA	Precision	Recall	F1	AUC
미이용 vs 조기·고이용	.789	.829	.923	.873	.759
미이용 vs 후기·저이용	.575	.715	.419	.528	.661
후기·저이용 vs 조기·고이용	.596	.910	.602	.725	.621

[그림 2]는 세 개의 모형에 대한 ROC 곡선을 시각화한 것이다. ROC 곡선의 아래 면적(AUC)이 1에 가까울수록 해당 분류기가 탁월한 분류 성능을 보유하고 있음을 의미한다. ROC 곡선 분석 결과, '미이용 vs 조기·고이용' 모형의 AUC가 .759로 가장 넓은 면적을 차지하고 있어, 다른 모형에 비해 양호한 분류 능력을 지닌 것으로 확인되었다.

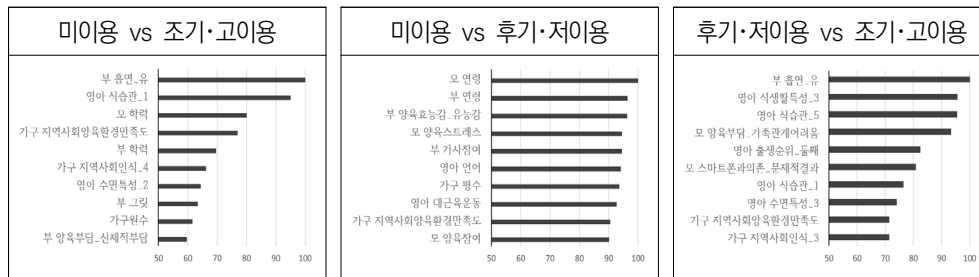


[그림 2] ROC curve

나. 영아 미디어 노출 유형을 분류하는 설명변수의 상대적 중요도

영아의 미디어 노출 유형을 분류하는 데 기여한 변수의 상대적 중요도를 확인하기 위해, 각 모형에 대해 MDG 기준 상위 10개의 변수를 산출하였으며, 결과는 [그림 3]과 같다. '미이용 vs 조기·고이용'과 '미이용 vs 후기·저이용' 모형에서는 아버지 관련 변수가, '후기·저이용 vs 조기·고이용' 모형에서는 영아 개인 특성 변수가 상대적으로 중요하게 나타

났다. 세 모형에서 공통적으로 중요도가 높았던 변수는 ‘가구 지역사회 양육 환경 만족도’였다.

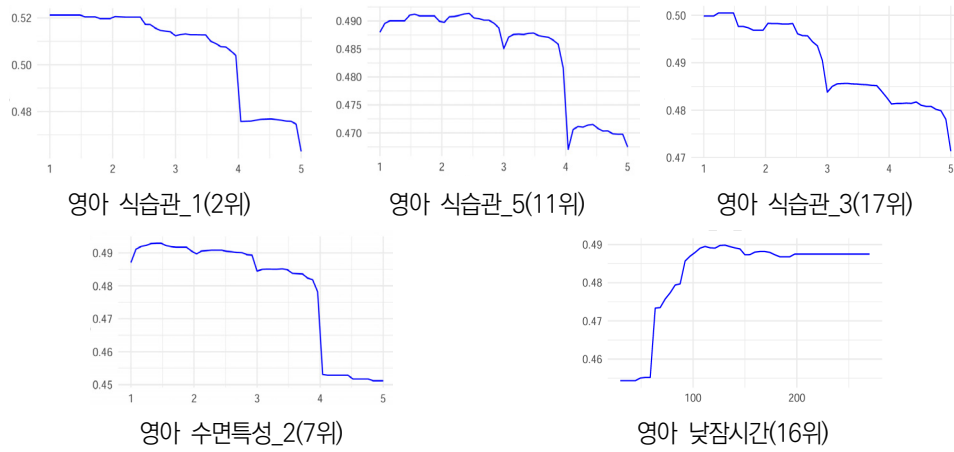


[그림 3] 미디어 노출 유형 분류 모형의 변인 중요도

3. 부분의존성 도표

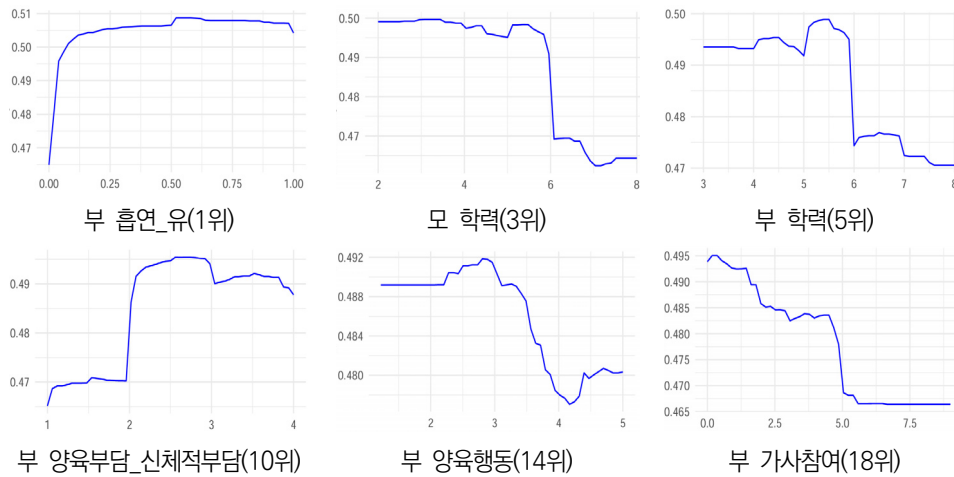
높은 분류 정확도를 보이며 위험군 분류에 효과적인 ‘미이용 vs 조기·고이용’ 모형을 중심으로 부분의존성 도표(PDP)를 분석하였다. 본 연구에서는 중요도 상위 20위 이내 변수 중 영아 개인 요인, 부모 요인, 가구 요인을 구분하여 대표 변수에 따른 분류 확률의 변화 양상을 해석하였다.

[그림 4]는 영아 개인 요인의 PDP 결과를 나타낸다. 분석 결과, 영아의 식사 습관과 수면 패턴은 ‘조기·고이용’ 집단 분류에 중요한 변인임이 나타났다. ‘영아 식습관_1’은 ‘아이의 식사 시간은 규칙적입니까?’를, ‘영아 식습관_5’는 ‘성인의 도움 없이 아이가 스스로 식사를 잘합니까?’를, ‘영아 식습관_3’은 ‘아이가 음식을 골고루 먹습니까?’를 묻는다. 따라서 식습관이 규칙적일수록, 스스로 먹을수록, 골고루 먹을수록 ‘조기·고이용’ 집단으로 분류될 가능성이 낮음을 알 수 있다. ‘영아 수면특성_2’는 ‘낮잠을 자지 않은 날은 저녁을 먹다가 잠이 들기도 한다.’를, ‘영아 낮잠시간’은 ‘하루 동안의 낮잠 시간은 총 몇 분입니까?’를 묻는다. 따라서 불규칙한 수면 리듬 및 장시간의 낮잠을 보이는 영아는 ‘조기·고이용’ 집단으로 분류될 가능성이 높다.



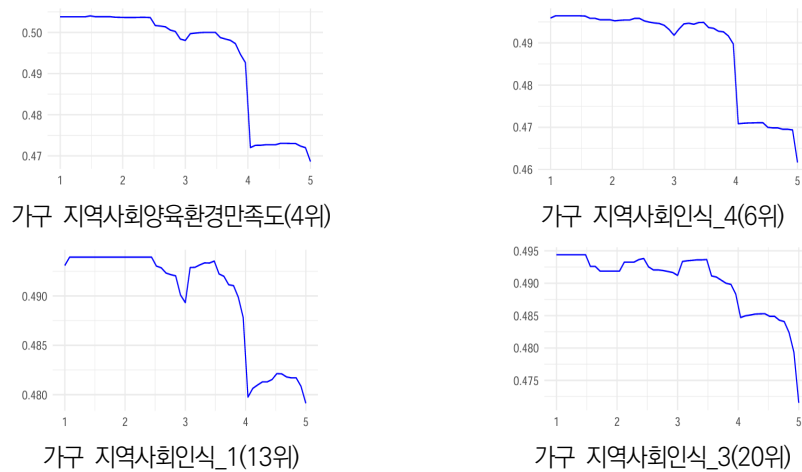
[그림 4] 영아 개인 요인 PDP

[그림 5]는 부모 요인의 PDP 결과를 나타낸다. 부모 요인 중에서는 특히 아버지 관련 변수의 중요성이 두드러졌다. 아버지가 흡연을 하거나 신체적 피로함과 관련된 양육부담이 클 때 ‘조기·고이용’ 집단으로 분류될 가능성이 높아졌다. 반면 어머니와 아버지의 학력 수준이 6 이상(4년제 대학 졸업 이상)일 때, 아버지의 양육행동과 가사참여도가 높을 때 ‘조기·고이용’ 집단으로 분류될 가능성은 낮아졌다.



[그림 5] 부모 요인 PDP

[그림 6]은 가구 요인의 PDP 결과를 나타낸다. 분석 결과, 지역사회 양육 환경 만족도가 높을수록 ‘조기·고이용’ 집단으로 분류될 가능성이 낮아졌다. ‘가구 지역사회인식_4’는 ‘사는 지역의 인도와 주변 도로를 안전하다고 느끼십니까?’를, ‘가구 지역사회인식_1’은 ‘사는 지역이 치안 측면에서 자녀를 키우기에 안전하다고 느끼십니까?’를, ‘가구 지역사회인식_3’은 ‘사는 지역에 놀이터, 공원, 산책로, 도서관 등 공공 여가 공간과 시설의 이용을 편리하다고 느끼십니까?’를 묻는다. 따라서 가구가 속한 지역의 도시 기반 시설과 공공 여가시설이 편리하고 안전하다고 느껴질수록 ‘조기·고이용’ 집단으로 분류될 가능성이 낮아졌다.



[그림 6] 가구 요인 PDP

종합적으로 PDP 분석 결과, 영아의 식습관과 수면습관, 아버지의 양육 자원 및 참여도, 가구의 지역사회 인식 및 만족도가 ‘조기·고이용’ 집단 분류에 상대적으로 크게 기여한 핵심 변수로 나타났다. 특히, 생활 습관의 규칙성과 가정 및 지역 환경의 양육 친화성이 확보될수록 ‘조기·고이용’ 집단으로의 분류 가능성은 감소하는 경향을 보였다.

IV. 논의 및 결론

본 연구는 「한국 영유아 교육·보육 패널」 2차년도 자료를 활용해 잠재프로파일 분석으로 1세 영아의 미디어 노출 유형을 분류하고, 랜덤포레스트로 분류 요인을 탐색하였다. 연구 결과를 바탕으로 한 주요 논의는 다음과 같다.

첫째, 미디어 이용 집단을 대상으로 잠재프로파일 분석을 실시한 결과, 두 개의 잠재집단이 도출되었다. ‘후기·저이용’ 집단(59.8%)은 13개월 무렵에 미디어를 처음 이용하기 시작하여 하루 30분가량 시청하는 양상을 보였고, ‘조기·고이용’ 집단(7.8%)은 11개월 무렵에 미디어를 접하기 시작하여 하루 130분가량 시청하는 특성이 있었다. 미디어를 전혀 사용하지 않는 ‘미이용’ 집단은 32.4%로 나타났다.

먼저, ‘조기·고이용’ 집단은 ‘후기·저이용’ 집단보다 미디어 이용 시작 시점이 평균 2개월 빠름에도 불구하고 시청 시간은 약 4배에 달했다. 이는 미디어 초기 노출 시점이 빠를수록 과다 이용으로 이어진다는 기존 연구(이수기·이현경·홍혜경, 2014)와 일치한다. 이는 영아기 미디어 노출에 대한 개입이 노출 시작 시점을 늦추는 예방적 전략까지 포함할 필요가 있음을 시사한다.

또한 잠재집단 내에서 주중과 주말의 미디어 노출 시간에 큰 차이가 없었다는 점은 주목할 만하다. 기존 연구에서는 보통 주말에 시청 시간이 더 길다고 보고된 바 있으나(최윤희, 2022), 본 연구에서는 뚜렷한 요일 차이가 관찰되지 않았다. 이는 영아기의 미디어 이용이 여가 시간보다 식사, 외출 준비 등 반복적인 일상 루틴 속에서 활용되기 때문일 수 있다(김광래, 2022).

이와 함께, 동일한 잠재집단 내부에서도 미디어 이용 양상이 균질하지 않다는 점을 고려할 필요가 있다. 두 이용 집단 모두 이용 시간의 표준편차가 작지 않았으며, 특히 ‘조기·고이용’ 집단에서 그 폭이 상대적으로 크게 나타났다. 이는 미디어를 이른 시기부터 많이 이용한 영아의 경우, 미디어 이용이 특정 시간대에 제한되지 않고 놀이, 달래기, 식사 보조, 휴식, 부모의 돌봄 공백 완충 등 다양한 양육 장면으로 확장되었을 가능성을 시사한다. 따라서 ‘조기·고이용’ 집단은 평균 이용 시간이 높은 집단일 뿐 아니라, 미디어가 일상에 편입되는 방식과 강도 역시 다양하게 나타나는 집단으로 이해할 필요가 있다.

둘째, 영아의 미디어 노출 유형을 분류하기 위해 적용한 랜덤포레스트 모델은 AUC 기준으로 .621~.759의 성능을 보였다. ‘미이용 vs 조기·고이용’ 모형은 AUC가 .759로 분류 성능이 양호한 편이었는데(오미애·최현수·김수현 외, 2017), 이 두 집단은 이용 여부나 시청 시간 등에서 명확한 차이를 보여 분류 성능이 높게 나타난 것으로 판단된다.

각 모형에서 주요 분류 변인의 구성이 상이하게 나타났다는 점도 주목할 필요가 있다. ‘미이용 vs 후기·저이용’ 및 ‘미이용 vs 조기·고이용’ 모형에서는 아버지 관련 변수의 비중이 컸으나, ‘후기·저이용 vs 조기·고이용’ 모형에서는 영아의 식습관과 같은 개인 특성이 상대적으로 높은 중요도를 보였다. 이는 영아의 미디어 이용 여부가 부모, 특히 아버지의 양육 태도나 행동 양식과 밀접하게 관련되는 반면, 미디어 사용이 시작된 이후의 이용 강도는 영아 개인 특성에 따라 달라질 수 있음을 시사한다. 즉, ‘노출의 시작’은 부모의 결정

에 의해 유도되지만, '사용의 지속'은 영아의 특성과의 상호작용 속에서 점차 고착화되는 경향을 보였을 가능성이 있다.

한편 모든 모형에서 공통적으로 중요도가 높게 나타난 변수는 '가구 지역사회 양육 환경 만족도'였다. 이는 영아의 미디어 노출이 단지 부모나 영아 개인 차원에서 발생하는 문제가 아니라, 지역사회 수준의 환경적 요인과 긴밀하게 연결되어 있음을 의미한다. 즉, 지역의 문화·여가 인프라, 공공 보육자원, 안전한 놀이 공간에 대한 접근성 등은 양육자들이 미디어 의존 대신 다양한 양육 자원을 활용할 수 있는 조건을 제공한다(김대웅·김창숙·박윤현, 2019).

셋째, 분류 성능이 가장 우수했던 '미디어 vs 조기·고이용' 모형을 중심으로 부분의존성 도표를 통해 주요 분류 변인을 분석한 결과, 아버지 흡연 유무, 영아 식습관, 어머니 학력, 가구 지역사회 양육 환경 만족도, 아버지 학력이 주요 분류 변인으로 나타났다. 이는 미디어 노출이 영아의 개인 특성은 물론, 부모와 가구 환경 등 다양한 수준의 요인이 함께 관련되어 있음을 시사한다.

영아 개인 특성 중에서는 수면 패턴이 불규칙하거나 식사 시 양육자에 대한 의존성이 높은 경우 '조기·고이용' 집단으로 분류될 가능성이 높았다. 이는 까다로운 기질을 가진 영아에 대해 부모가 미디어를 행동 조절 또는 진정 수단으로 활용할 가능성을 제시한 선행연구와 맥락을 같이한다(Coyne, Shawcroft, Gale et al., 2021).

부모 요인에서는 특히 아버지 관련 변수가 주요 분류 변인으로 나타났다. 아버지가 흡연을 하고, 학력이 낮으며, 신체적 양육 부담이 크고, 양육행동 및 가사참여 수준이 낮을수록 자녀가 '조기·고이용' 집단으로 분류될 가능성이 높았다. 아버지의 양육시간이 증가할수록 흡연 시간이 감소했다는 선행연구(최지은·김현경, 2019)를 고려할 때, 흡연 여부는 아버지의 양육 참여와 관련될 가능성이 있다. 또한 흡연은 건강행동을 나타내는 지표일 뿐 아니라, 학력, 가구소득, 정신건강, 근무형태 등을 복합적으로 반영하는 생활맥락 변수이기도 하다(문성미, 2016). 따라서 본 연구에서 아버지의 흡연 여부가 주요 분류 변인으로 나타난 결과는 흡연 자체의 직접적 효과라기보다, 아버지의 근무환경, 스트레스, 돌봄 참여 가능성 등 보다 넓은 가족 맥락이 함께 반영된 결과로 이해할 필요가 있다.

또한 아버지가 신체적 양육에 부담을 느낄수록, 퇴근 후의 피로감 등으로 인해 자녀와의 직접적인 상호작용보다는 미디어를 활용한 간접적 돌봄 방식에 의존했을 가능성이 있다(이소현·김윤희, 2020). 반면, 아버지가 적극적으로 양육과 가사에 참여하면 어머니 부담이 줄고, 자녀에게는 '놀이 친구'로 작용했을 수 있다(오윤경·이미현·최혜순, 2016; 이지인, 2020).

다만 이러한 결과를 어머니 변인이 영아의 미디어 노출과 관련이 없다는 의미로 해석하기는 어렵다. 랜덤포레스트의 변수 중요도는 특정 변인의 절대적 영향력을 의미하기보다, 주어진 자료 안에서 집단을 구분하는 데 기여한 상대적 정보를 나타낸다. 따라서 어머니

관련 변인이 집단 간 비교적 유사한 분포를 보인 반면, 아버지 관련 변인이 집단 구분에 더 큰 변별력을 제공했을 가능성이 있다. 이는 영아기 미디어 노출을 설명할 때 기존 연구에서 주로 강조되어 온 어머니 요인뿐 아니라, 아버지의 생활양식과 실제 양육 참여 맥락을 함께 고려할 필요가 있음을 시사한다.

가구 환경 요인 중에서는 지역의 치안이 불안정하거나 도로의 안전성이 낮고, 놀이터·공원 등 여가 공간을 불편하게 인식할수록 자녀가 ‘조기·고이용’ 집단으로 분류될 가능성이 높았다. 이는 부모 입장에서 위험한 외부 환경에 아이를 노출시키는 것보다, 가정 내에서 미디어를 활용하는 것이 안전한 양육 선택지로 여겨질 수 있음을 시사한다(Wang, Ma, Harada, et al., 2021). 아울러, 공원·놀이터·도서관 등 자녀와 함께 시간을 보낼 수 있는 지역사회 양육 자원이 충분하지 않은 환경에서는 미디어 이외의 대안적 활동 선택지가 상대적으로 제한될 가능성이 있으며, 이러한 환경적 조건이 영아의 미디어 이용 양상과 관련될 수 있다(김대웅·김창숙·박윤희, 2019). 이러한 맥락에서, 지역사회의 물리적 환경은 단순한 여가 활동의 배경이 아니라, 영아기 양육 전략과 미디어 이용 양상을 이해하는 데 함께 고려해야 할 맥락적 요인으로 볼 수 있다.

V. 정책적 제언 및 시사점

본 연구는 잠재프로파일 분석과 랜덤포레스트 알고리즘을 활용해 1세 영아의 미디어 노출 유형과 분류 변인을 탐색하였다. 세계보건기구는 2세 미만의 미디어 사용을 제한하고, 2세 이후 유아도 하루 1시간 이내 이용을 권고한다(WHO, 2019). 그러나 본 연구에서 확인된 ‘후기·저이용’과 ‘조기·고이용’ 집단은 모두 권고 기준을 초과하는 조기 노출에 해당했으며, 특히 ‘조기·고이용’ 집단은 영아기 이후에도 권장되지 않는 수준의 과도한 이용을 보여 영아기 미디어 이용에 대한 정책적 관심과 지원의 필요성을 시사한다.

첫째, 부모교육의 실효성을 높이는 동시에 모든 영아 가정이 자연스럽게 정보를 접할 수 있는 보편적 전달체계를 마련할 필요가 있다. 본 연구에서 주요 분류 변인으로 나타난 부모의 학력이나 생활양식은 정보 접근성, 근무 여건, 스트레스, 돌봄 자원 등 다양한 생활 조건을 함께 반영할 수 있으며, 교육 참여 가능성 역시 가정의 자원에 따라 달라질 수 있다. 따라서 영유아 건강검진, 예방접종, 보건소 방문, 어린이집 입소 등 대부분의 부모가 이용하는 기존 서비스 안에서 미디어 이용 지침과 대안적 놀이 정보를 반복적으로 제공할 필요가 있다. 또한 시간과 돌봄 자원이 부족한 가정을 위해 모바일 기반의 짧은 자료, 방문·전화 상담 등 참여 부담이 낮은 지원 방식도 함께 마련할 필요가 있다.

둘째, 영아기 미디어 이용과 관련한 부모 지원에서 아버지의 역할과 참여 여건을 함께 고려할 필요가 있다. 본 연구에서는 아버지의 돌봄 참여 여건과 관련된 변수가 영아의 미디어 노출 유형을 구분하는 주요 변인으로 나타났다. 따라서 부모교육과 미디어 이용 가이드라인에서 아버지의 역할을 명시하고, 아버지가 자녀와의 놀이와 미디어 이용 조절에 참여할 수 있도록 구체적인 실천 지침을 제공할 필요가 있다. 예를 들어 독일의 ‘Väterkurse(아버지 교육)’와 같은 성별 특화 부모교육 프로그램을 참고하여(Pietsch & Heckt, 2016), 아버지를 위한 미디어 이용 지도 교육, 온라인 콘텐츠, 캠페인, 상담 서비스를 마련할 수 있다. 다만 이러한 문제를 아버지 개인의 인식이나 의지만을 변화시키는 교육으로 해결하기는 어렵다. 따라서 아버지가 실제 양육과 자녀와의 놀이에 참여할 수 있도록, 부모교육과 가족지원 정책에서도 아버지의 참여 여건을 함께 고려할 필요가 있다.

셋째, 지역사회 양육환경을 점검하고 미디어를 대체할 수 있는 활동 자원에 대한 접근성을 높일 필요가 있다. 본 연구에서는 모든 분류 모형에서 ‘가구 지역사회 양육 환경 만족도’가 공통적으로 상위 중요 변수로 나타났으며, 치안, 도로 안전성, 여가시설 접근성 등 지역 환경 요인이 ‘조기·고이용’ 집단 분류와 관련되었다. 이는 영아의 미디어 이용이 가정 내 양육 방식뿐 아니라, 가정 밖에서 이용할 수 있는 놀이·문화 자원의 접근성과도 관련될 가능성을 시사한다. 따라서 단기적으로는 지자체와 육아종합지원센터가 영아 동반 가족이 이용할 수 있는 도서관, 공동육아나눔터, 실내의 놀이공간 등의 정보를 제공하고, 기존 지역 자원과의 연계를 강화할 필요가 있다. 중장기적으로는 부산광역시의 ‘들락날락’과 같이 생활권 내에서 책, 놀이, 문화·체험 활동을 제공하는 어린이복합문화공간의 확대를 검토할 수 있다. 이러한 지역사회 기반의 지원은 가정이 선택할 수 있는 대안적 활동의 범위를 넓힌다는 점에서 의미가 있다.

본 연구는 기존 연구에서 상대적으로 주목받지 못했던 영아기를 대상으로 미디어 노출 유형을 실증적으로 분류하고, 머신러닝 기반 분류 모형을 적용했다는 점에서 의의를 지닌다. 특히 영아의 미디어 노출을 단순히 이용 시간의 많고 적음으로 파악하지 않고, 최초 노출 시기와 주중·주말 이용 시간을 함께 고려하여 ‘미이용’, ‘후기·저이용’, ‘조기·고이용’ 유형으로 구분했다는 점에서 기존 연구의 한계를 보완하였다. 또한 어머니 요인에 집중되어 온 기존 논의와 달리, 아버지의 생활양식과 양육 참여, 가구의 지역사회 환경이 영아 미디어 노출 유형을 구분하는 데 중요한 변인으로 나타났다는 점은 영아기 미디어 이용을 개별 부모의 양육 문제로만 볼 것이 아니라, 가정 내 돌봄 맥락과 지역사회 환경이 결합된 다층적 현상으로 이해할 필요가 있음을 보여준다.

그럼에도 본 연구에는 몇 가지 한계가 있다. 첫째, 본 연구는 단일 시점의 횡단자료에

기반하므로, 분류 변인과 미디어 노출 유형 간 시간적 선후관계와 인과적 방향성을 확정하기는 어렵다. 향후에는 종단자료를 활용하여 미디어 노출 유형의 변화와 이후 발달 및 이용 양상과의 관계를 검증할 필요가 있다.

둘째, 본 연구는 미디어 이용 여부와 최초 노출 시기, 이용 시간을 중심으로 유형을 분류하였으나, 미디어의 콘텐츠 유형, 사용 목적, 동반 시청 여부, 부모-자녀 상호작용 수준 등 구체적인 이용 맥락을 충분히 반영하지 못하였다. 향후 연구에서는 노출량을 넘어 이용 맥락을 함께 고려할 필요가 있다.

셋째, 주양육자, 어머니, 아버지가 모두 응답한 가구만을 포함하였기 때문에 아버지 응답이 누락된 가구, 한부모 가구, 양육 참여가 낮은 가구의 특성이 충분히 반영되지 못했을 가능성이 있다. 따라서 후속 연구에서는 다양한 가족 구조와 양육 조건을 포함할 필요가 있다.

넷째, 본 연구는 랜덤포레스트의 변수 중요도를 활용하여 주요 분류 변인을 탐색하였으나, 변수 중요도는 특정 변인의 인과적 효과를 의미하지 않는다. 또한 본 연구에서 활용한 MDG 기준 중요도는 변수의 측정 수준이나 범주 수, 변수 간 상관 구조에 따라 달라질 수 있으므로, 본 연구의 변수 중요도 순위는 미디어 노출 유형 분류에 기여한 변수의 상대적 탐색 결과로 이해할 필요가 있다. 후속 연구에서는 permutation importance, SHAP 등 다양한 모형 해석 기법을 적용하여 분류 변인의 기여도를 보다 안정적으로 검토하고, 로지스틱 회귀 및 구조방정식 모형 등 이론 기반 분석을 병행하여 주요 변인의 영향력을 검증할 필요가 있다.

마지막으로 본 연구의 분류 모형은 동일 자료 내 학습용·검증용 자료 분할을 통해 성능을 평가하였으므로, 향후 독립된 외부 자료를 활용하여 일반화 가능성을 검증할 필요가 있다.

참고문헌

- 강민아·이영신(2020). 디지털 영상매체 이용시간과 가정문화환경 수준에 따른 영아의 언어 발달 차이. *학습자중심교과교육연구*, 20(19), 973-999.
- 김광래(2022). 유아의 디지털미디어 활용 형태가 보호자의 미디어 영향력에 대한 보호자의 인식과 제한의 차이 연구. *유아교육·보육복지연구*, 26(3), 133-160.
- 김길재·박강윤·문명현(2023). 랜덤포레스트를 활용한 청소년의 비행 예측 요인 탐색. *청소년학연구*, 30(5), 267-290.
- 김대웅·김창숙·박윤희(2019). 공공 문화시설 및 여가시설 이용 편리성, 양육 스트레스, 유아 미디어 이용시간의 관계. *열린유아교육연구*, 24(5), 1-18.
- 김연수·정윤경(2018). 어머니의 스마트기기 중독경향성과 어머니-유아 상호작용 간 관련성에 대한 단기종단연구. *인간발달연구*, 25(1), 25-42.
- 김영서·홍세희(2021). 고등학생의 청소년활동 참여에 따른 잠재프로파일 유형과 공동체의식 및 주관적 행복감의 차이. *조사연구*, 22(1), 91-125.
- 류미향(2014). 영아의 스마트폰 사용실태 및 어머니 인식: 어머니의 양육효능감, 양육스트레스와의 관련성 탐구. *한국영유아보육학*, 86, 307-329.
- 문성미(2016). 국내 임금근로자의 흡연 유형과 관련요인 연구. *한국보건간호학회지*, 30(3), 495-511.
- 배윤진·박은정·김자연·조경진·이혜민·강민권(2023). 한국 영유아 교육·보육 패널 연구 2023. 서울: 육아정책연구소.
- 성지현·변혜원·남지혜(2015). 유아의 스마트기기 이용과 발달 수준 및 공감 능력과의 관계 탐색. *유아교육연구*, 35(2), 369-394.
- 오미애·최현수·김수현·장준혁·진재현·천미경(2017). 기계학습(Machine Learning) 기반 사회보장 빅데이터 분석 및 예측 모형 연구. 세종: 한국보건사회연구원.
- 오윤경·이미현·최혜순(2016). 유아기 자녀를 둔 아버지의 양육참여와 스마트폰 과몰입 간의 관계. *어린이미디어연구*, 15(4), 169-191.
- 오주현·박용완(2019). 영유아의 스마트 미디어 사용 실태 및 부모 인식 분석. *육아정책연구*, 13(3), 3-26.
- 이경숙·정석진·김명식(2015). 1-3세 영유아 영상물 과몰입이 일반적, 정서사회적 발달에 미치는 영향. *한국심리학회지*, 28(1), 117-132.
- 이병호(2012). 가정에서 유아의 스마트 기기 사용에 관한 탐색 연구. *유아교육학논집*, 16(6), 511-532.
- 이소현·김윤희(2020). 어머니의 양육 참여가 유아의 미디어 이용에 미치는 영향에서 아버지

- 양육 참여의 조절 효과. *어린이미디어연구*, 19(2), 1-16.
- 이수기·이현경·홍혜경(2014). 유아의 스마트기기 사용현황과 게임 과몰입 경향성. *유아교육연구*, 34(3), 239-262.
- 이정림·도남희·오유정(2013). 영유아의 미디어 매체 노출실태 및 보호대책. 서울: 육아정책연구소.
- 이정원·박원순·엄지원(2021). 영유아의 미디어 이용 적정화를 위한 정책 방안 연구. 서울: 육아정책연구소.
- 이지인(2020). 어머니 양육참여와 아버지 놀이참여가 자녀 미디어중독에 미치는 영향. *학습자중심교과교육연구*, 20(19), 1431-1453.
- 이혜림·박재완(2018). 결혼이주여성의 문화 적응: 유형과 특성의 변화 분석. *한국정책학회보*, 27(4), 329-358.
- 장지우·김선희(2022). 어머니의 양육스트레스, 스마트미디어 방임적 중재 및 유아의 스마트미디어 과다 사용의 구조적 관계. *한국보육지원학회지*, 18(6), 1-19.
- 조준오·김은정(2019). 유아 어머니의 양육스트레스와 유아 미디어중독 간의 관계에서 양육효능감의 매개 및 조절효과. *학습자중심교과교육연구*, 19(9), 1189-1208.
- 최윤희(2022). 어머니 우울 증상이 유아의 스마트미디어 과의존에 미치는 영향: 어머니 스마트폰 중독경향성과 스마트미디어 중재의 순차매개효과와 유아 발달 인식의 조절된 매개효과. 박사학위논문. 숙명여자대학교 대학원.
- 최지은·김현경(2019). 아버지의 양육참여와 아버지의 신체적·정신적 적응: 생활시간조사와 한국아동패널을 이용한 시대, 연령, 코호트 간 차이 분석. *한국사회학*, 53(3), 1-41.
- 최필선·민인식(2018). 머신러닝 기법을 이용한 대졸자 취업예측 모형. *직업능력개발연구*, 21(1), 31-54.
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5-32.
- Cingel, D. P., & Krcmar, M. (2013). Predicting media use in very young children: The role of demographics and parent attitudes. *Communication Studies*, 64(4), 374-394.
- Coyne, S. M., Shawcroft, J., Gale, M., Gentile, D. A., Etherington, J. T., Holmgren, H., & Stockdale, L. (2021). Tantrums, toddlers and technology: Temperament, media emotion regulation, and problematic media use in early childhood. *Computers in Human Behavior*, 120, 106762.
- Hsu, C. W., & Lin, C. J. (2002). A comparison of methods for multiclass support vector machines. *IEEE Transactions on Neural Networks*, 13(2), 415-425.
- Jung, T., & Wickrama, K. A. S. (2008). An introduction to latent class growth

- analysis and growth mixture modeling. *Social and Personality Psychology Compass*, 2(1), 302-317.
- Nikken, P., & Oprea, S. J. (2018). Guiding young children's digital media use: SES-differences in mediation concerns and competence. *Journal of Child and Family Studies*, 27, 1844-1857.
- Pietsch, M., & Heckt, M. (2016). Family Literacy in Hamburg. Hält das Programm, was es verspricht. *Die Deutsche Schule, Beiheft*, 13.
- Radesky, J. S., Silverstein, M., Zuckerman, B., & Christakis, D. A. (2014). Infant self-regulation and early childhood media exposure. *Pediatrics*, 133(5), 1172-1178.
- Wang, Q., Ma, J., Harada, K., Kobayashi, S., Sano, H., & Kim, H. (2021). Associations among outdoor playtime, screen time, and environmental factors in Japanese preschoolers: The 'eat, be active, and sleep well' study. *Sustainability*, 13(22), 12499.
- World Health Organization (WHO). (2019). Guidelines on physical activity, sedentary behaviour and sleep for children under 5 years of age. <https://www.who.int/publications/i/item/9789241550536>

- 논문접수: 6월 1일 / 수정본접수: 7월 14일 / 게재 승인: 8월 13일
- 교신저자: 정정희, 경북대학교 아동학부 교수, chchung1@hanmail.net

Abstract

Infant Media Exposure Profiles and Key Variables for Classification: A Latent Profile Analysis and Random Forest Approach

Cheawon Goo and Chunghee Chung

This study aimed to classify patterns of media exposure among one-year-old infants and identify key variables for classification. Using data from the second wave of the Korean Early Childhood Education and Care Panel, this study analyzed 1,486 households and included 106 input variables, encompassing infant, parental, and household characteristics. Latent profile analysis was first conducted to identify patterns of infant media exposure, followed by random forest models using a One-vs-One (OvO) classification strategy. Three distinct patterns of media exposure were identified: “non-use,” “later/low use,” and “early/high use.” The AUC values of the random forest models ranged from .621 to .759, with the “non-use vs. early/high use” model showing the best classification performance. Key variables contributing to classification included infants’ sleep and eating habits, paternal smoking status, and satisfaction with the community caregiving environment. These findings suggest that infant media exposure is associated with multilevel factors and provide practical implications for expanding paternal involvement in caregiving and improving community-based caregiving environments.

Keywords: Infant Media Exposure, Latent Profile Analysis, Random Forest, Machine Learning