

## TV홈쇼핑 이용자의 상품구매 입문 요인 분석: 랜덤포레스트 분석을 활용한 개인의 미디어 레퍼토리 데이터를 중심으로\*

이창준\*\* · 우형진\*\*\* · 박성복\*\*\*\*

### 초 록

본 논문은 TV홈쇼핑 구매 경험이 없는 소비자들 중 1년 후 TV홈쇼핑 구매에 입문한 소비자의 미디어 매체, 연결, 행위로 구성되는 미디어 레퍼토리 패턴을 분석하고, 해당 소비자가 어떤 요인으로 인해 TV홈쇼핑에서 상품을 처음 구매하게 되는지에 대한 요인을 랜덤포레스트와 의사결정나무 알고리즘을 활용하여 분석하였다. 사용 가능한 171개의 잠재 변수를 분석에 활용하여 랜덤포레스트 알고리즘이 데이터를 학습하여 기존 문헌에서 발견되지 않았던 새로운 요인들을 발견 할 수 있었다. 예를 들어, 개인정보에 대한 우려 정도, 영화관에서 의 지출 정도, 적극적 정보 탐색자 여부와 같은 요인들도 TV홈쇼핑 구매로 이어지는 중요한 요인이라는 것을 보였다. 분석에서는 연령별, 성별로 총 12개의 하위 샘플을 나누고, 각각의 샘플에 대하여 미디어 레퍼토리 관점에서 TV홈쇼핑 입문 요인을 살펴보았다. 이후, 샘플별로 선정된 상위 30개 요인들만을 가지고 최종 의사결정나무를 구성하여 TV홈쇼핑 구매 입문 과정을 최종 분류해 보았다. 본 연구는 기존의 이론주도 요인 분석에 데이터주도 방식을 혼합하여 향후 미디어 상품과 서비스에 대한 연구가 나아갈 방향을 제시한다.

**주제어;** TV홈쇼핑, 요인 분석, 미디어 레퍼토리, 랜덤포레스트, 의사결정나무

논문접수일 2020. 3. 4, 논문수정일 2020. 3. 31, 논문게재확정일 2020. 4. 7

\* 본 논문은 ‘한국방송학회-롯데홈쇼핑 2019년도 최초연구 지원’에 의해 수행되었음.

\*\* 제1저자, 한양대학교 ERICA 정보사회미디어학과 조교수, changjunlee@hanyang.ac.kr

\*\*\* 교신저자, 한양대학교 ERICA 정보사회미디어학과 교수, hyungjinw@hanyang.ac.kr

\*\*\*\* 한양대학교 ERICA 정보사회미디어학과 부교수, octobermine@hanyang.ac.kr

## 1. 서론

TV홈쇼핑은 현대인들이 사용하는 전자 매체 중 가장 오랜 역사를 갖고 있는 텔레비전을 통해 다양한 상품을 소개하고, 주거 공간에서 소비자에게 편리한 의사결정을 도와주게 하는 플랫폼이다. 1995년 국내에서 24개 케이블 방송이 송출을 시작하고 그 중에 2개 채널이 홈쇼핑 전용 채널로 배정된 이후, TV홈쇼핑은 2010년까지 15년 간 연 평균 10.2%라는 엄청난 성장을 이루었다. 2018년 현재 TV홈쇼핑은 7개 사업자로 TV홈쇼핑 시장은 과점상태이고, 전체 매출액은 약 5조원에 이른다.<sup>1)</sup> TV홈쇼핑은 케이블TV산업, 택배산업, 중소기업과 같은 연관 산업들을 성장시키고, 시장에서 높은 고용창출을 만들어내는 산업이라고 할 수 있다(곽승준 · 이주석 · 최은철, 2018). 그러나 최근 TV홈쇼핑은 온라인 및 모바일 유통산업들과 경쟁하면서 그 성장세가 점점 정체되고 있는 실정이다. 미디어 환경의 변화와 TV시청자의 미디어 소비 행태 변화가 TV홈쇼핑 성장세에 부정적 영향을 미치고 있다. 조사회사 엠브레인의 조사결과에서 나타났듯이, 여러 세대가 TV홈쇼핑에 대해 관심을 가지고 있지만 주로 이용하는 세대는 장년층이고, 젊은 층은 온라인 및 모바일 홈쇼핑을 이용한다(EMBRAIN, 2017). TV홈쇼핑에 대한 소비자 신뢰도가 낮다는 것은 TV홈쇼핑이 언제라도 외면 받을 수 있다는 의미이고, 주요 이용자가 중장년층이라는 것은 지속적 성장이 불안정할 수 있다는 점을 말한다. 이런 점에서 국내 TV홈쇼핑의 성장이 정체 국면으로 들어선 것은 어쩌면 당연한 구조적 문제일 수 있다. TV홈쇼핑이 다양한 상품을 저렴하게 제공하고, 결제, 배송, 환불, 반품 서비스가 뛰어난 것은 사실이지만, 새로운 소비자 집단이나 틈새 소비자를 찾지 못한다면, 향후 TV홈쇼핑의 성장은 요원할 수 있다.

소비자가 새로운 방법으로 상품을 구매할 때, 첫 구매로 연결되기까지 여러 가지 진입 장벽이 존재한다(Lian and Yen, 2013). TV홈쇼핑 구매도 마찬가지로 상품구매자의 인구통계학적 특성, 미디어 이용 특성, 개인적 특성, 미디어 레퍼토리 등 다양한 변수들이 소비자의 상품 구매 과정에 개입할 것이다. 우리나라 소비자의 주요 이용 유통 통로가 대형할인마트이고, TV홈쇼핑이 전체 이용자의 13.9% 정도만 이용한다는 뜻은 역으로 해석하면 그만큼 TV홈쇼핑을 경험하지

1) 한국TV홈쇼핑협회 홈페이지, <http://kota.re.kr> 참조

못한 소비자들이 많다는 것을 의미한다. 2018년 기준, 소비자의 49.5%는 여전히 TV홈쇼핑의 구매경험이 전무하다고 밝히고 있다<sup>2)</sup>. TV홈쇼핑을 알고 있는 것과 실제로 TV홈쇼핑을 통해 상품을 구매한 것은 완전히 다른 이야기이다. TV홈쇼핑 시청자에서 TV홈쇼핑 구매자로 변화하는 과정에 어떤 변인들이 어떤 방식으로 조합을 이루는가는 TV홈쇼핑 운영자의 큰 관심사일 것이다. TV를 외면하는 젊은 세대의 경우, 적극적 정보 탐색자로서 한 시간에 한 카테고리의 상품을 편성하는 TV홈쇼핑의 긴 호흡을 참지 못하고, 대부분 본인의 스마트폰으로 온라인 쇼핑을 한다. 이런 변화의 흐름들은 TV홈쇼핑의 경쟁자가 대형할인마트나 백화점과 같은 오프라인 유통이 아니라 소비자의 관심과 시간을 TV로부터 완전히 뺏은 넷플릭스나 웨이브와 같은 온라인 콘텐츠 플랫폼 사업자들이라는 것이다. 소비자들도 지금처럼 케이블TV나 IPTV의 TV홈쇼핑 채널을 통해 상품을 구매할지, 아니면 이 중 하나의 매체로만 집중할지, 더 나아가 넷플릭스와 같은 콘텐츠 플랫폼을 여러 개 모아서 구독하는 시스템을 통해 홈쇼핑을 할지 고민할 것이다. 이에 변화하는 미디어 환경, 심화되는 시장 경쟁, 미디어 이용자의 소비패턴 변화 등을 동시에 잘 이해하고 이들을 관통하는 공통된 패턴을 파악할 필요가 있다.

본 연구는 심화된 경쟁 미디어 환경에서 TV홈쇼핑 구매 경험이 전혀 없다가 입문하게 된 상품 구매자의 인구통계학적 특성, 개인적 특성, 미디어 이용 특성, 미디어 레퍼토리를 분석한다. 기존의 요인 분석 연구와 달리, TV홈쇼핑에 처음으로 입문한 소비자가 어떤 미디어를 어떤 장소에서 얼마나 소비했는지 등의 패턴을 분석하고, TV홈쇼핑을 통한 상품 구매 입문의 분기점을 예측하여 연령대별 및 성별 분류를 통해 변화된 미디어 환경 속에서 TV홈쇼핑이 어떻게 대응해야 할지 전략적 시사점을 제시 하고자 한다.

## II. 이론적 배경

TV홈쇼핑 시청자는 프로그램 시청 후, 전화로 제품을 주문하고, 신용카드나 온라인 입금방식을 통해 대금을 지불할 수 있는 편리성 때문에 쇼핑에 드는 노력과 시간을 줄이는 경험을 한다(유수현 · 차경옥, 2004). 선행연구에 따르면, TV

2) 2018 KISDI 미디어 패널을 이용한 저자의 통계 분석

홈쇼핑 경험 만족도는 구매과정에서 제품 및 서비스에 대한 호의로 규정되고, 구매과정에서 저렴한 가격, 다양한 정보, 편리성이 만족도를 높이는 주요 원인이라고 밝혀졌다(임채이 · 전현진, 2016). TV홈쇼핑 시청자들은 상품 구매 경험에 따라 신뢰추구형, 합리적 선택추구형, 호의적 선호형, 유용성 추구형으로 구분될 수 있다고 한다(박웅기 · 한예진, 2014). 기존 연구들에서 TV홈쇼핑을 하는 주요 동기는 편리성이었다(한성희 · 양세정, 2009; 이상봉, 2011). 즉, TV홈쇼핑 경험의 가장 큰 특징은 구매자가 직접 매장에 가지 않아도 자신이 원하는 상품을 문제없이 획득할 수 있다는 것이고, 이런 점에서 TV홈쇼핑은 상품구매의 편리한 도구로 인식된다는 것이다.

TV홈쇼핑이 일상생활에서 어떤 의미를 갖는지 살펴본 기존 연구에 의하면, TV홈쇼핑 시청자들은 홈쇼핑 방송이 여러 텔레비전 장르의 특성이 혼합되어 있는 것으로 인지했고, 생방송임을 알고 있지만 그것이 진정성이나 진실성을 내포하는 것으로 보지 않았으며, TV홈쇼핑 채널을 통해 다양한 즐거움을 획득하고, 비대면성이라는 장점을 이용하여 친구, 가족, 이웃처럼 편리하게 교류하게 해준다고 인식했다(윤태진 · 옥미애, 2007). 즉, TV홈쇼핑은 상품 구매자들에게 일상, 소비, 오락, 효용이라는 ‘혼종성’(윤태진 · 옥미애, 2007, p.69)을 경험하게 해주고, 올드미디어인 TV와 인터넷, 모바일, T커머스와 같은 다양한 네트워크 매체 이용을 통해 상품구매의 전방위적 수단 제공의 복합적 미디어 이용 경험을 하게 해준다.

TV홈쇼핑 주요 구매자들은 여성, 젊은층, 높은 사회적 고립감, 유색인종보다 백인, 전통적 결혼관의 특성을 가지고 있다고 밝혔다(James and Cunningham, 1987). 구매자 다수가 여성이며, 60%이상이 대학 교육 이상의 교육수준을 가진 고소득자였다. TV홈쇼핑 이용률이 높은 집단은 주부, 취학 전 아동을 양육하는 파트타임 취업여성, 40세 미만의 독신여성, 40대 여성가장 가구와 중간소득 계층의 가정 홈쇼핑, 대도시 거주자인 것으로 나타났다(Darian, 1987). 국내연구를 살펴보면, 대도시보다 지방도시 거주자의 TV홈쇼핑 선호도가 높고, 교육수준이 높을수록 TV홈쇼핑에 대한 선호도가 낮은 것으로 나타났다(김주영, 1996, 정순희 · 이초인 · 차경옥, 2004에서 재인용). TV홈쇼핑 구매자들은 비구매자들과 비교했을 때, 낮은 연령, 높은 교육수준, 미혼에 여가시간이 많은 유형이라고 밝히고 있다(박광희, 1995). 다수의 TV홈쇼핑 구매자들은 소득수준, 사회계층, 교

육수준, 직업 등에서 평균 이상의 사회경제적 수준을 보이고, 장년층의 기혼여성이 많은 것으로 나타났다(심미영, 2002). TV홈쇼핑의 구매방법인 무이자 할부구매나 적립금 지급에 있어 남성보다 여성이 긍정적이라고 밝혔다(이지은, 2005). 2017년 국내 노인인구가 전체인구의 14%인 고령사회로 진입하면서 실버층에 대한 TV홈쇼핑 이용이 증가하고 있다. 이들에 대한 TV홈쇼핑 구매선택 요인에 있어서도 연령, 결혼여부, 소득, 종교, 주거형태는 유의미한 요인으로 나타났다(김하나 · 김숙웅, 2019). 2018년 기준, 한국홈쇼핑협회에서 밝힌 구매고객 구성비를 보면, 남성(18.1%)보다 여성(81.9%)이 훨씬 많으며, 주로 40대(22.3%)와 50대(27.6%) 여성이 전체 구매자의 절반에 해당한다고 밝혔다(한국홈쇼핑협회 홈페이지 참조). TV홈쇼핑 관련 선행연구에서 인구통계학적 변인들은 TV홈쇼핑의 중요한 구매요인으로 밝혀지고 있으나 연구자별, 국가별, 시대별로 채택된 인구통계학적 변인들이 상품구매와 관련하여 항상 같은 영향력을 일관성 있게 나타내는 것은 아니다. 이에 새로운 미디어 환경 속에서 TV홈쇼핑 이용자의 인구통계학적 특성의 영향력을 재 고찰할 필요가 있다.

이용과 충족이론은 텔레비전 시청행태를 관습적 시청(ritualized viewing)과 도구적 시청(instrumental viewing)으로 나눈다(Rubin, 1994). 전자는 비의도적이고 습관적으로 미디어를 이용하거나 시간을 보내기 위한 시청이라면, 후자는 목적 지향적으로 미디어를 선택해서 유용성 고려에 따른 의도적 미디어 이용행위이다(이상봉, 2011). TV홈쇼핑 구매자는 제품을 직접 보거나 기능을 시도해보고 구매하는 것이 아니고, 전적으로 텔레비전 화면에 나타나는 제품의 모습과 쇼호스트의 정보 전달에 의지하여 구매결정을 한다. 이런 미디어 이용행위는 적극성을 수반한 자기주도적이고 목적 지향적 의도를 가지고 있다. 그러므로 TV를 통한 상품 구매는 구매자의 성격이 반영된다고 볼 수 있다. TV홈쇼핑 구매자의 신상품에 대한 태도, 구매과정의 신중함, 직접보거나 경험하지 않고 구매하기 때문에 발생하는 지각된 위험에 대한 극복 또는 도전정신 등은 TV홈쇼핑 구매 과정에서 중요한 영향을 미칠 것이다. 개혁신산이론이나 기술수용모델은 새로운 기술이나 상품의 채택에 있어서 혁신적 태도를 가진 개인이 해당 대상을 수용, 구매할 가능성이 높다고 밝혔다(손승혜 외, 2011). TV홈쇼핑 구매자들은 이용 편리성과 함께 채널 사업자에 대한 신뢰성을 매우 높이 평가하지만 가격이나 신속성, 다양성은 낮게 평가한 것으로 나타났다(이광훈, 2007). 다시 말해서, TV홈

쇼핑을 통한 상품 구매가 직접 구매보다 손해를 보거나 예상과 달리 마음에 들지 않는 물건을 배송 받을 수 있기 때문에 보다 구매 결정에 있어서 세심하다는 것이다. 또한 TV홈쇼핑 구매자들은 본인이 원래 구매하고자 했던 본 제품 이외에 사은품 제공, 부가혜택, 무이자 할부와 같은 요인에 대해서도 관심이 많아 제품을 선택 및 구매할 때, 신중하게 접근한다(한성희 · 양세정, 2009). TV홈쇼핑은 전통적인 상품 구매 행위와 다르고, TV시청이라는 미디어 이용행위를 통해서 구매행위를 하는 것이다. 따라서 TV홈쇼핑은 TV시청행위와 동시에 구매행위를 하는 이중적 속성을 가지고 있어서 시간보내기, 오락성뿐만 아니라 가격요인, 편의성 요인 등도 개입한다(Cortese and Rubin, 2010). 더불어 빠르고 안전한 배송, 편리한 결제도 중요한 TV홈쇼핑 이용 요인이 될 수 있다(박정우 · 이영주, 2014). 특히, TV홈쇼핑을 이용하는 과정에서 지각된 위험들(정보유출위험, 교환반품위험, 사회심리적 위험)은 개별 홈쇼핑 기업에 대한 불만족에 유의미한 영향을 미치는 것으로 나타났다(배지양, 2018). TV홈쇼핑 방송의 시청동기는 미디어의 특성과도 관련이 있는데, 김봉철 외(2014)에 따르면, 제품구매, 지식습득, 시간보내기, 쇼핑호스트에 대한 호감이 통계적으로 유의한 시청 동기임을 밝혔다.

‘미디어 레퍼토리’라는 개념은 미디어의 수용자에서 사용자(user) 개념으로 바뀌면서, 사용자가 다양한 미디어 중 일부를 선택해서 능동적으로 이용한다는 개념이다. 연구 흐름은 Ferguson (1992)의 연구에서 VCR과 케이블 TV를 소비자가 리모컨이라는 미디어를 활용해 어떻게 채널을 조합하고 있는지에 대한 것으로 시작되었다. 이후 미디어의 종류가 바뀌고 다변화 되면서 웹, 장르, 프로그램 등에 대한 레퍼토리 분석으로 연구가 진행되었다. 특히나 모든 것이 연결된 트랜스 미디어 시대에 들어 표면적으로 사용자가 선택할 수 있는 미디어 조합의 수가 거의 무한대로 증가하였지만, 두드러지게 사용 되는 미디어 레퍼토리는 한정되어 있다는 것이 밝혀지면서 이에 대한 연구 필요성이 커지고 있다. 본 연구에서의 미디어 레퍼토리는 이용자에게 주어진 콘텐츠를 소비하는 매체, 행위, 연결 등의 모든 사용 가능한 조합 중 자신의 선호를 토대로 주로 사용하게 되는 특정 미디어의 묶음이라고 할 수 있다. 사용자들 고유의 미디어 레퍼토리 중에 특별히 TV홈쇼핑 구매에 입문 확률이 높은 조합이 있는지를 찾는 것이 본 연구의 목적이다.

### III. 연구 방법

#### 3.1 연구 문제

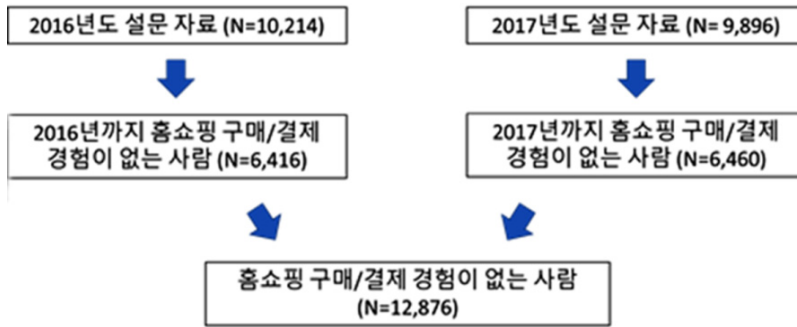
본 연구에서는 다음의 두 가지 연구 문제를 바탕으로 연구를 진행하였다.

**연구 문제 1:** TV홈쇼핑 구매 경험 유무에 따라 TV홈쇼핑 이용자의 미디어 레퍼토리가 매체, 행위, 연결 차원에서 어떤 차이를 나타내는가?

**연구 문제 2:** TV홈쇼핑 이용자의 예측변수들(인구통계학적 특성, 미디어 사용 특성, 성격과 라이프 스타일, 미디어의 매체, 연결, 행위에 대한 레퍼토리) 중, TV홈쇼핑 구매 의사 결정으로 이어지는데 결정적인 분기점 역할을 하는 요인들은 무엇인가?

#### 3.2 자료 수집 절차

방송과 통신 산업이 융합되면서 미디어 이용자들이 단말기, 플랫폼, 콘텐츠 패키징 방식에 따라 다양한 미디어 콘텐츠 이용이 가능해졌고 이에 다매체 환경에서 미디어 이용 행태를 측정할 수 있는 사회 조사가 필요해졌다. 이를 위해 본 연구에서는 정보통신정책연구원에서 제공하는 한국 미디어 패널 데이터를 사용하여 다매체 환경에서 미디어 사용자의 이용 행태를 추출한다. 미디어 패널은 전국 17개 시도의 인구 분포를 고려한 층화추출 방식으로 총 5,000여 가구 및 가구 내 개인에 대한 미디어 이용행태를 가구방문 면접조사를 통해 2010년부터 매년 조사하는 정부 주도의 사회조사이다. 본 연구에서는 TV홈쇼핑에 대한 설문이 시작된 2016에서 2018년까지 3개년의 자료를 활용하여 TV홈쇼핑 구매 결정 요인 분석을 위한 샘플을 추출하였다. 홈쇼핑 구매 경험이 없었다가 어떤 요인에 의해서 생기게 되는지를 보기 위해 2016년도와 2017년의 설문 대상 각각 10,214명과 9,896명의 자료에서 홈쇼핑 구매 경험이 없었던 6,416명과 6,460명을 추출하여 하나의 샘플로 구성하였다. 앞의 샘플을 A라고 하고 뒤의 샘플을 B라고 한다면 A 샘플에서는 2017년에 TV홈쇼핑 구매 경험이 생기는지 여부를, B 샘플에서는 2018년에 구매 경험이 생기는지 여부를 추적하였다. 연구 샘플의 구성은 <그림 1>과 같다.



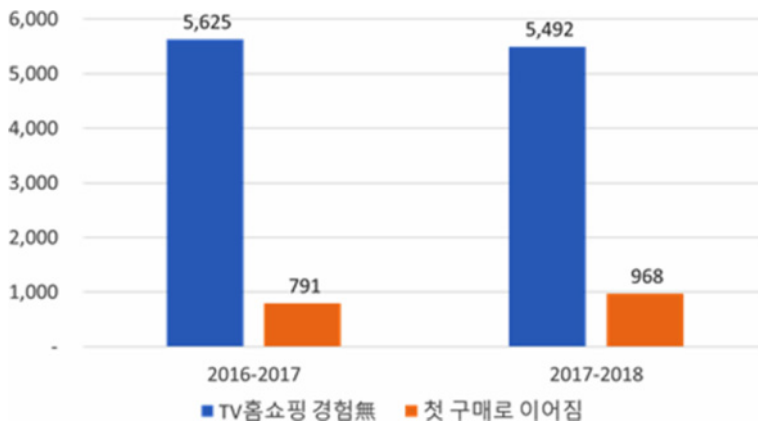
〈그림1〉 연구 샘플의 구성

### 3.3 연구 방법

#### 3.3.1 변인의 측정

##### 3.3.1.1 결정 변수

본 연구에서 사용한 결정 변수는 TV홈쇼핑 구매 경험이 없다가 1년 경과 후 생기는 것으로 정의한다. 사용하고 있는 데이터가 패널 조사이므로 동일한 사람의 TV홈쇼핑 구매 경험을 추적 조사할 수 있다. 이러한 장점을 살려 두 가지 샘플; A(2016-2017)와 B(2017-2018)을 구성하였고, 12.3%(791명)과 15.0%(968명)에 해당하는 미디어 사용자들이 1년 후 TV홈쇼핑 첫 구매가 이루어진 것을 알 수 있다. 〈그림 2〉는 연구 샘플 내 결정 변수에 대한 기술 통계를 나타낸다.



〈그림 2〉 결정변수에 대한 기술통계



### 3.3.1.2 예측 변수(1) 인구통계학적 변인

본 연구에서 기계 학습을 위해 사용된 총 변수들의 수는 171개이며, 부록 <표 A1>에 자세히 나와 있다. 변수들은 크게 다섯 가지로 구분할 수 있으며, 설문자의 인구통계학적 특성, 성격과 라이프 스타일, 미디어 사용 특성, 그리고 매체, 행위, 연결에 대한 미디어 레퍼토리가 그것이다. 인구통계학 변수로는 사회 과학에서 다루는 기본 요소인 성별, 나이, 최종학력, 결혼 여부, 개인 월평균 소득을 다루고 있으며, 그 외에도 거주지, 가구원 수, 가족 구성, 직업 유무, 종교 유무, 반려동물 유무 등으로 이루어져있다.

### 3.3.1.3 예측 변수(2) 개인적 특성

개인의 성격과 라이프 스타일은 새로운 방식으로 구매를 시도하는 데 있어 많은 영향을 끼치며, 이를 반영하기 위해 본 연구에서는 제품 구매 행태, 문제 해결 방식, 삶의 만족도 측면에서의 개인적 특성들을 기계학습 알고리즘에 추가하였다. 먼저, 최신 상품에 대한 선호도를 측정하기 위해 아래 세 가지 설문을 활용하였다.

- \* 제품 구매 시 최신 상품인지가 중요하다.
- \* 제품 구매 시 남들이 잘 쓰지 않는 상품이 더 매력적이다.
- \* 제품 구매 시 유행에 민감하다.

두 번째로 제품 구매 시 얼마나 신중한 편인지를 측정하기 위해 아래 설문 항목들을 활용하였다.

- \* 상품을 직접 눈으로 확인하고 가격을 흥정하기 위하여 관련 매장을 돌아다니며 구매한다.
- \* 내가 사용할 상품은 내가 직접 구입한다.
- \* 상품에 대해서 알아보고 저렴하게 구입하기 위해서 인터넷을 활용한다.

또한, 제품 구매시 주위 사람이나 인터넷에서의 평판이 중요한지, 기능보다는 가격이 중요한지, 유명 브랜드 여부가 중요한지, 여가활동으로서의 쇼핑 활동 빈도 등도 포함하였다. 사용자가 얼마나 호기심이 크며, 문제 해결을 위해 공을 들이는지에 대한 측정을 위해서는 아래 설문들을 활용하였다.

- \* 나는 어떤 문제에 대해 새로운 해결방법을 배울 때 즐겁다.
- \* 나는 복잡한 문제를 보면 왠지 그것을 분석해 보고 싶은 생각이 든다.
- \* 나는 잘 모르는 것에 대해서 더 알고 싶어 한다.
- \* 나는 어렵고 시간이 많이 드는 문제도 가능한 한 끝까지 풀기위해 노력한다.

삶의 만족도에 대한 변수는 개인적 측면 만족도, 관계적 측면 만족도, 집단적 측면 만족도에 대한 평균으로 측정하였고, 주관적으로 본인 성격이 미래에 대해서 낙관적인지 아니면 불안하고 우울한지에 대한 변수도 포함하였다.

#### 3.3.1.4 예측 변수(3) 미디어 특성

개인의 미디어 특성은 개인이 어떤 통신사와 어떤 유형의 핸드폰을 사용하는지에 대한 내용과 디지털 리터러시, 그리고 온라인 활동에 대한 내용을 포함한다. 통신사는 KT, SKT, LG U+, MVNO 중 어떤 통신사를 사용하고 있는지, 핸드폰은 몇 개를 보유하고 있으며 해당 핸드폰이 어떤 유형의 네트워크를 사용하는지(i.e., 3G, LTE, LTE-A)에 대한 설문 항목을 이용하여 변수를 만들었다. 미디어 지출 항목으로는 월평균 휴대폰 이용 총 금액, 월평균 기기 할부금, 그리고 디지털 콘텐츠에 대한 지출 금액, 유료 애플리케이션 구입 경험여부, 유/무료 애플리케이션 내 아이템 및 확장기능 구입 경험여부, 애플리케이션 관련 지출 금액, 극장에서의 영화 관람 지출 금액, 공연 관람 지출 금액 등을 포함하였다. 디지털 리터러시에 대한 부분은 지난 3개월 동안 인터넷 동호회/카페/클럽 글 읽기 활동 빈도, 댓글 달기 활동 빈도, 게시글 스크랩 활동 빈도, 글 쓰기 활동 빈도, 글 쓰기 활동 빈도를 활용하였고, SNS 활용 정도에 대한 부분은 지난 3개월 동안 SNS 답변, 글 쓰기 활동 빈도, SNS 정보 공유 활동 빈도, SNS 추천, 평점 주기 기능 활동 빈도 등으로 구성되었다.

#### 3.3.1.5 예측 변수(4) 미디어 레퍼토리

미디어 레퍼토리 변수는 이용자에게 주어진 콘텐츠를 소비하는 모든 사용 가능한 조합 중에서 특정 미디어 소비에 주로 사용하는 매체, 해당 매체를 통한 미디어 행위, 그리고 미디어 소비로 이어지는 연결 등의 묶음이다. 먼저, 매체 이용은 미디어 패널 데이터 내 총 41개의 매체 이용 변수를 사용하였다. 전통 미디어

에 속하는 신문/책/잡지, 가정용TV부터 디지털 미디어에 속하는 데스크탑, 노트북, 스마트 미디어에 속하는 스마트폰과 태블릿PC, 마지막으로 공간 미디어에 속하는 영화관, 공연장 등도 포함된다. 미디어 연결과 행위를 측정하기 위해서는 미디어 애용에 사용한 매체를 어떤 연결(i.e., 와이파이)로 어떤 행위(i.e., 동영상 시청)를 했는지에 대한 변수들을 사용하였다.

### 3.3.2. 결정요인 탐색방법

#### 3.3.2.1 의사 결정 나무

의사결정나무는 실제 의사 결정 또는 일어날 법한 결과나 이벤트의 확률을 나무 모양의 가지치기를 모형화 하여 최적화하는 방법론이다. 어떤 조건들의 분기점이 최적의 의사 결정 또는 결정 변수가 되는 결과로 이어지는지를 나타내는 기계학습의 한 방법으로 운영 과학, 의사 결정, 전략 분석 등에 자주 사용된다. 분기점을 결정하는 알고리즘은 지니 불순도를 최소화 하는 것이며 분기점에 의해 분리된 집합에 결정 변수의 결과가 얼마나 섞여 있는지를 측정하는 지표이다. 예를 들어, 본 연구에서 지니 불순도는 다음과 같이 정의된다. TV홈쇼핑 구매로 이어지는지 여부를  $J$  클래스라고 한다면, 구매로 이어지거나 이어지지 않는 두 가지의 결과로 이루어진 클래스는  $J=2$ , 로 나타낼 수 있다. 한 분리된 집합에서 클래스  $J$ 를 지닌 요소를 선택할 확률 또는 집합 내의 클래스  $J$ 를 가지는 요소의 비율을  $P_j$ 라고 한다면 지니 불순도  $G_j$ 는 아래식과 같이 정의된다.

$$G_j = \sum_{j=1}^2 P_j(1 - P_j)$$

이렇게 정의된 지니 불순도는 매 분기점마다 목적함수 Min()에 따라 최적화를 하면서 의사결정나무를 만들어 나가게 된다. 이 때 너무 많은 가지치기는 모델로 하여금 과적합 문제를 가지게 하며, 하나의 케이스에 대한 예측력은 높아지지만 일반화가 어려워지는 문제에 봉착한다. 하지만 수많은 요인들을 기반으로 결정 변수를 예측하게 되면 어느 정도의 과적합은 피할 수 없다.

### 3.3.2.2 랜덤포레스트

본 연구에서는 성별(남/여), 연령별(10대/20대/30대/40대/50대/60대 이상)으로 12개의 하위 샘플을 만들고, 다섯가지 예측 변수 대분류에 각각 랜덤 포레스트를 적용하여 중요 분기점을 만드는 5가지 변수를 각각 뽑아낸다. 이 후, 대분류별 대표 5가지 변수들을 합쳐 20개의 변수들로 성별, 연령별 TV홈쇼핑 구매에 대한 의사결정나무를 각각 구축해보았다.

수많은 요소를 기반으로 결정 변수를 예측할 때 하나의 의사결정나무에 의한 과적합 문제를 해결하기 위해 본 연구에서는 랜덤 포레스트를 사용하였다. 예측 변수로 사용되는 변수의 수가 170여개인 상황에서 하나의 의사 결정 나무로 최적화 한다면 나무의 가지가 너무 많아질 것이고, 이는 과적합을 야기할 수밖에 없다. 더욱이 최적화 문제를 푸는 것만으로 접근한다면 170여개의 변수 중 아주 성능이 좋은 분류 노드는 아니지만 다른 몇 개의 조건이 맞을 때 아주 좋은 분기점이 될 수도 있는 잠재 요인들은 배제될 수밖에 없다. 하지만 170개의 요인 중에서 무작위로 5개의 요인을 선택해서 하나의 결정 나무를 만들고, 또 170개 중 무작위로 5개의 요인을 선택하여 또 다른 결정 나무를 만든 식으로 300개 이상의 숲을 구성한 다음 나무들이 내는 예측 값들을 다수결로 결정한다면, 앞서 이야기 했던 과적합의 문제와 소수의 좋은 분류기가 소외되는 두 가지 문제를 모두 해결할 수 있다. 이렇게 의사 결정의 분기점을 샘플링 하는 과정에 무작위성을 더하여 수많은 의사결정 나무를 만들고 여러 나무들의 의견을 종합하는 방법을 랜덤 포레스트라고 한다.

## IV. 연구결과

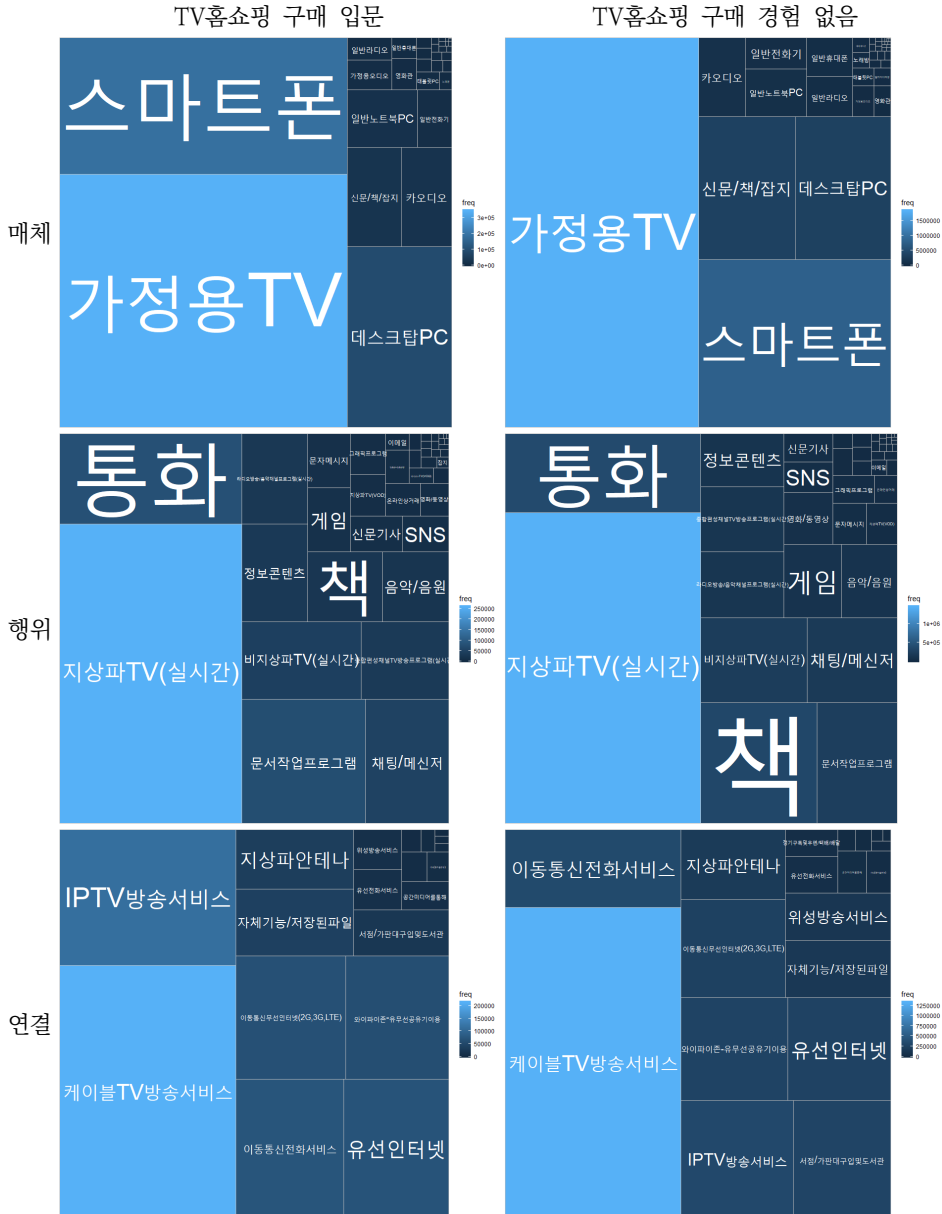
### 4.1 기술 통계 결과

TV홈쇼핑 구매 경험이 전무한 최종 샘플(N=10,354) 중에서 15.5%인 1,631명이 TV홈쇼핑 구매에 입문하였다. <표 1>은 TV홈쇼핑 구매에 입문한 샘플과 그렇지 않은 샘플에 대한 인구 통계학 변수의 기술 통계표이다. TV홈쇼핑 구매 입문 샘플의 특징을 살펴보면, 그렇지 않은 샘플에 비해 여성(66%)과 기혼(73.7%)의 비율이 크고, 3인 이상으로 구성된 가족 구성(79.5%)에 서울과 수도

권 거주 비율이 38.2%로 높은 것을 알 수 있다. 소득 수준으로는 월 평균 100에서 200만원 구간에서 차이가 많았고(TV홈쇼핑 구매 경험 없는 샘플이 8.8%, TV홈쇼핑 구매 입문 샘플에서 1.7%). 400만원 이상의 소득 구간에서도 차이를 보였다(TV홈쇼핑 구매 경험 없는 샘플이 26.6%, TV홈쇼핑 구매 입문 샘플에서 35%).

〈표-1〉 결정 요인 별 인구 통계학적 기술 통계 비교

		TV홈쇼핑 구매 경험 없음		TV홈쇼핑 구매 입문	
		평균	(표준편차)	평균	(표준편차)
나이(만)		51.85	(20.82)	47.17	(13.20)
		N(8,722)	(%)	N(1,631)	(%)
성별	남성	4792	54.9%	555	34.0%
	여성	3930	45.1%	1076	66.0%
혼인상태	미혼	2415	27.7%	305	18.7%
	기혼	5065	58.1%	1202	73.7%
가족구성	1인가구	846	9.7%	91	5.6%
	2인가구	2327	26.7%	244	15.0%
	3인이상	5549	63.6%	1296	79.5%
	3대동거	723	8.3%	131	8.0%
교육수준	초등학교 졸업	1939	22.2%	81	5.0%
	중학교 졸업	1042	11.9%	112	6.9%
	고등학교 졸업	2877	33.0%	726	44.5%
	대학교 졸업	2743	31.4%	683	41.9%
	대학원 이상	121	1.4%	29	1.8%
거주지	서울/수도권	2209	25.3%	623	38.2%
	광역시	2744	31.5%	467	28.6%
소득수준 (월평균)	100이하	3005	34.5%	565	34.6%
	100-200	770	8.8%	28	1.7%
	200-300	772	8.9%	80	4.9%
	300-400	1843	21.1%	387	23.7%
	400-500	1275	14.6%	299	18.3%
	500-600	641	7.3%	180	11.0%
	600-700	230	2.6%	52	3.2%
	700이상	186	2.1%	40	2.5%



〈그림 3〉 결정 요인 별 미디어 레퍼토리의 차이 비교: 나무지도 시각화

같은 방법으로 두 샘플의 미디어 레퍼토리를 비교해보았다. 〈그림 3〉은 각각 두 샘플에 대한 미디어 레퍼토리의 차이를 미디어의 매체, 행위, 연결에 대한 사

용 시간의 면적을 나무지도(Treemap)로 시각화 한 것이다. 먼저 미디어 매체에 대한 레퍼토리를 보면 두 샘플이 가정용TV, 스마트폰, 데스크탑, 신문/책/잡지, 카오디오 순으로 시간을 쓰는 것에서는 동일하나, 그 비율에 있어 TV홈쇼핑 구매에 입문한 샘플이 그렇지 않은 샘플에 비해 스마트 미디어(스마트폰과 테블릿 PC)와 디지털 미디어(데스크탑, 노트북) 사용 비율이 높고, 공간 미디어로서 영화관에서 보내는 시간이 더 많은 특징을 보였다.

미디어 행위에 대한 레퍼토리를 보면 지상파TV를 보거나 통화를 하는 행위가 두 샘플에서 가장 높은 비율의 레퍼토리를 차지한 것까지는 동일했으나 세번째부터는 차이를 나타내었다. TV홈쇼핑 구매 입문 샘플은 문서 작업을 하고 채팅이나 메신저 앱을 사용하는 행위가 레퍼토리의 많은 비율을 차지했지만, TV홈쇼핑으로 구매가 이어지지 않은 샘플에서는 책을 보거나 비지상파TV를 시청하는 행위가 많은 비율을 차지하는 특징을 보였다.

마지막으로 미디어 연결에 대한 레퍼토리를 보면 매체와 행위 레퍼토리에서 보이지 않았던 큰 차이를 볼 수 있다. TV홈쇼핑 구매로 이어진 샘플에서는 케이블TV방송에 연결된 미디어에 대한 시간 소비 비율이 그렇지 않은 샘플에 비해 확연히 적고 IPTV방송에 연결된 미디어에 시간을 보내는 비율이 두 번째로 큰 특징을 가진다. 그 외에도 3G나 LTE 등의 통신망에 연결된 인터넷 사용 비율도 TV홈쇼핑 구매 입문 샘플에서 확연하게 큰 것을 볼 수 있었다.

## 4.2 랜덤 포레스트 결과

〈표 2〉는 성별, 연령별 12개의 하위 샘플을 만들고, 예측 변수의 여섯가지 대분류에 각각 랜덤 포레스트를 적용하여 중요 분기점을 만드는 5가지 변수를 뽑아낸 것에 대한 10대와 20대의 샘플 결과이다. 여기서 선정된 상위 5가지 변수들은 여섯가지 예측 변수의 대분류 영역에서 TV홈쇼핑 구매로 이어지는 의사 결정의 중요한 분기점이라는 의미이다.

전체적으로 공통된 트렌드를 분석해 보면, 인구 통계학적 요인에서는 소득수준, 직업유무, 종교유무가, 미디어 특성에서는 개인정보에 대한 우려 정도, 월평균 모바일 지출과 영화관 지출이, 성격과 라이프 스타일 영역에서는 탐구적 성격, 적극적 정보 탐색자, 신상 선호 여부, 삶의 만족도, 여가활동으로서 쇼핑 빈도가 순서는 다르지만 모든 샘플에서 중요했다. 미디어 매체 레퍼토리에서는 스

마트폰, 데스크탑, 가정용TV, 신문이 미디어 연결 레퍼토리에서는 모바일 핸드폰, 모바일 인터넷, 케이블TV가, 마지막으로 미디어 행위 레퍼토리에서는 메신저, 통화, 지상파TV가 모든 샘플에서 공통적으로 상위 5개의 주요 요인으로 뽑혔다.

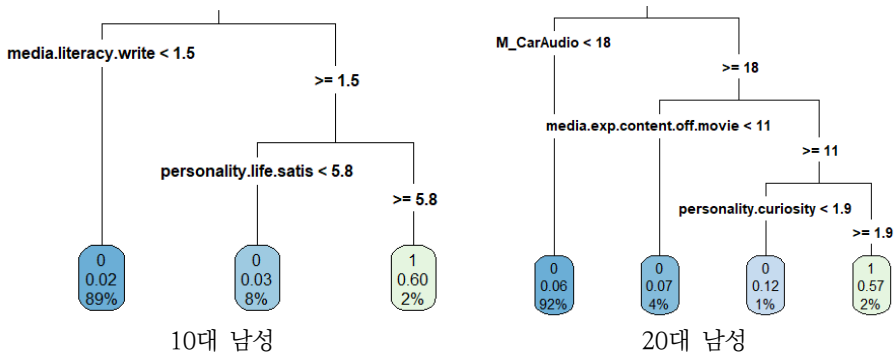
〈표-2〉 랜덤 포레스트로 학습한 TV홈쇼핑 구매에 대한 연령별, 성별 주요 요인들(10, 20대)

대분류	10대		20대	
	남성	여성	남성	여성
인구통계 학적 요인	소득수준	소득수준	소득수준	소득수준
	종교여부	종교여부	광역시거주여부	직업유무
	수도권거주여부	교육수준	종교여부	교육수준
	교육수준	직업유무	반려동물유무	반려동물유무
	직업유무	반려동물유무	직업유무	종교여부
미디어 특성	개인정보에 대한 우려정도	개인정보에대한 우려정도	개인정보에대한 우려정도	개인정보에대한 우려정도
	월평균모바일기기지출	월평균모바일지출	SNS활용정도	월평균모바일지출
	월평균모바일지출	SNS활용정도	월평균모바일지출	SNS활용정도
	월평균영화관지출	월평균영화관지출	월평균영화관지출	월평균모바일기기지출
	인터넷활동 (글쓰기빈도)	온라인활동빈도	월평균모바일기기지출	월평균영화관지출
성격과 라이프 스타일	탐구적성격	탐구적성격	삶의만족도	탐구적성격
	적극적정보탐색자	삶의만족도	탐구적성격	삶의만족도
	신상선호자	신상선호자	신상선호자	신상선호자
	삶의만족도	쇼핑빈도	적극적정보탐색자	적극적정보탐색자
	쇼핑빈도	적극적정보탐색자	쇼핑빈도	쇼핑빈도
미디어 레퍼토리 (매체)	스마트폰	가정용TV	스마트폰	가정용TV
	데스크탑	신문	가정용TV	스마트폰
	가정용TV	스마트폰	카오디오	카오디오
	신문	노트북	데스크탑	데스크탑
	가정용오디오	데스크탑	신문	신문
미디어 레퍼토리 (연결)	모바일인터넷	도서관	모바일인터넷	모바일인터넷
	저장된파일	모바일인터넷	케이블TV	모바일핸드폰
	모바일핸드폰	IPTV	와이파이	와이파이
	케이블TV	케이블TV	모바일핸드폰	IPTV
	와이파이	모바일핸드폰	유선인터넷	케이블TV
미디어 레퍼토리 (행위)	메신저채팅	지상파TV	메신저채팅	지상파TV
	그래픽프로그램	정보검색	통화	통화
	지상파TV	독서	지상파TV	메신저채팅
	음악스트리밍	통화	음악스트리밍	문서작업
	통화	메신저채팅	문서작업	정보검색

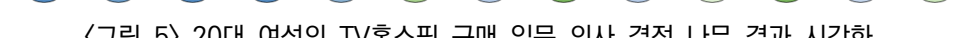


이렇게 선정된 30개(5개의 중요 요인 × 6개의 대분류)의 요인들로 다시 최종 의사결정나무를 구성해 본 결과를 <그림 4, 5>에 나타내었다. 10대 남성의 경우, 다른 샘플에서 발견되지 않았던 분류기인 인터넷 활동 중 글쓰기 빈도가 주요 분기점 역할을 한다. 일주일에 1.5회 이상 인터넷 동호회, 카페, 클럽 등에서 글을 쓰고 삶의 만족도가 7점 만점에서 5.8점 이상일 때 샘플의 2%에서 60%의 확률로 TV홈쇼핑 구매에 입문하는 사용자를 분류할 수 있다. 20대 남성의 경우, 미디어 매체 레퍼토리에서 카오디오 사용 시간, 공간 미디어인 영화관에서의 지출, 개인의 호기심이 분류기로 사용 되었다. 카오디오 사용 시간은 20대 남성의 경우 구매력과 연결될 수 있고 영화관에서의 지출은 문화생활 빈도를 대표하며 호기심은 새로운 구매 방법으로 입문 하는 것에 유리하다는 점에서 해석해 볼 수 있다. 주어진 샘플에서는 하루에 18분 이상 카오디오를 사용하고 영화관에서의 지출이 5만원 이상이며 호기심(평균 2.3)이 1.9이상일 때 샘플의 2%에서 57% 확률로 TV홈쇼핑 구매로 입문하였다.

10대 여성의 경우는 의사결정나무에서 불순도를 최소화 하는 과정에서 2개 이상의 분기점이 나오지 않았다. 20대 여성의 경우, <그림 5>에 나타나 있듯이 12개의 분기점이 존재한다. 가장 큰 확률로 TV홈쇼핑 입문을 예측할 수 있는 분류기는 카오디오를 하루에 13분 이상 사용하고, 6만 5천원 요금제 이상을 사용하는 것으로 70% 확률로 입문 한다는 것을 알 수 있다. 두 번째로 카오디오는 13분 미만으로 사용하지만 케이블TV에 연결된 시간이 68분 이상, 신문을 보는 시간이 2.5분이 채 안되며 6만 5천원 이상의 요금제를 사용하고 개인정보에 대한 우려가 크다면 86%의 확률로 TV홈쇼핑에 입문한다는 것을 보여주고 있다.



<그림 4> 10대, 20대 남성의 TV홈쇼핑 구매 입문 의사 결정 나무 결과 시각화



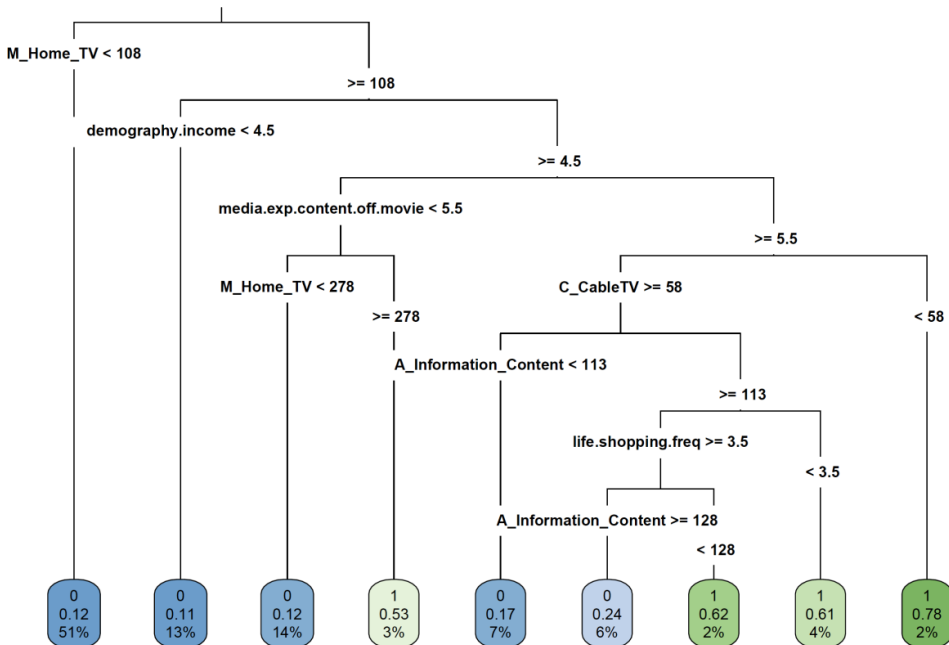
〈표-3〉 랜덤 포레스트로 학습한 TV홈쇼핑 구매에 대한 연령별, 성별 주요 요인들(30, 40대)

대분류	30대		40대	
	남성	여성	남성	여성
인구통계학적 요인	소득수준	소득수준	소득수준	소득수준
	교육수준	교육수준	교육수준	광역시거주여부
	광역시거주여부	미혼	3대거주여부	교육수준
	종교여부	기혼	반려동물유무	반려동물유무
	수도권거주여부	종교여부	수도권거주여부	종교여부
미디어 특성	개인정보에 대한 우려정도	개인정보에 대한 우려정도	개인정보에 대한 우려정도	개인정보에 대한 우려정도
	월평균모바일지출	월평균영화관지출	월평균영화관지출	월평균영화관지출
	월평균영화관지출	SNS활용정도	월평균모바일지출	월평균모바일지출
	SNS활용정도	월평균모바일지출	SNS활용정도	월평균모바일기기지출
	월평균모바일기기지출	월평균모바일기기지출	월평균모바일기기지출	SNS활용정도
성격과 라이프 스타일	적극적정보탐색자	삶의만족도	삶의만족도	삶의만족도
	삶의만족도	탐구적성격	탐구적성격	탐구적성격
	탐구적성격	적극적정보탐색자	적극적정보탐색자	신상선호자
	신상선호자	신상선호자	신상선호자	적극적정보탐색자
	쇼핑빈도	쇼핑빈도	쇼핑빈도	쇼핑빈도
미디어 레퍼토리 (매체)	가정용TV	스마트폰	가정용TV	가정용TV
	스마트폰	가정용TV	스마트폰	스마트폰
	데스크탑	데스크탑	데스크탑	데스크탑
	카오디오	신문	카오디오	신문
	핸드폰	핸드폰	노트북	카오디오
미디어 레퍼토리 (연결)	모바일인터넷	케이블TV	모바일인터넷	모바일인터넷
	모바일핸드폰	모바일핸드폰	유선인터넷	모바일핸드폰
	케이블TV	모바일인터넷	모바일핸드폰	케이블TV
	유선인터넷	와이파이	케이블TV	와이파이
	IPTV	IPTV	와이파이	IPTV
미디어 레퍼토리 (행위)	문서작업	지상파TV	지상파TV	지상파TV
	지상파TV	메신저채팅	문서작업	통화
	통화	통화	통화	메신저채팅
	메신저채팅	정보검색	메신저채팅	정보검색
	정보검색	음악스트리밍	라디오	비지상파TV

이렇게 선정된 30개의 요인들로 다시 최종 의사결정나무를 구성해 본 결과를 〈그림 6-8〉에 나타내었다. 30대 남성의 경우, 가장 먼저 가정용TV 시청 시간이 하루 평균 108분이 넘어야 한다. 그러면서 월 평균 소득 수준이 350만원 보다

크고 연 평균 영화관 지출이 3만원 이상이면 TV홈쇼핑 입문에 있어 이미 좋은 분류기 역할을 할 수 있다. 조금 더 자세하게 보면, 케이블TV에 연결된 시간이 58분 미만이면 가장 큰 확률인 78%로 예측 가능하고, 케이블TV 연결 시간이 58분 이상이라면 정보 검색 시간이 113분 이상이고 여가로서 쇼핑 빈도가 낮을 때 61% 확률로 TV홈쇼핑 구매로 입문할 가능성이 있다.

30대 여성의 경우, 의사결정나무의 가지 수도 많고 상위 노드의 성능이 뛰어나지 않아 복잡한 분류로 보이지만, 가장 우측 최종 분류에서 샘플의 12% 중 64%의 확률로 결정 요인을 분류하고 있기 때문에 예측의 양적인 면에 있어서는 우수하다고 볼 수 있다. 소득이 250~300만원 미만이지만 연 평균 영화 관람에는 3만원 이상을 지출하고 삶의 만족도가 5.5미만일 경우(평균 4.8, 최대값 7)가 이에 해당한다. 그 외에도 샘플의 3%에서 82% 확률로 예측하는 분류 가지 등도 존재하나 12%의 64% 확률 예측기가 다른 예측기에 비해 비중이 크기 때문에 나머지 분류기에 대한 설명은 생략하기로 한다.



〈그림 6〉 30대 남성의 TV홈쇼핑 구매 입문 의사 결정 나무 결과 시각화



40대 남성의 경우는 의사결정나무에서 불순도를 최소화 하는 과정에서 2개 이상의 분기점이 나오지 않았다. 40대 여성의 경우 역시 총 20개의 가지(노드)를 이루는 복잡한 의사결정나무로 보이지만 최우측 분류기를 보면 단 두개의 노드로 샘플의 23% 에서 68%를 예측하는 우수한 분류기가 존재한다. 영화관 지출이 연 3만원 이상이며 가정용TV 미디어에 하루 평균 178분 이상의 시간을 보내고 있다면 이에 속한다. 총 170여개의 잠재 분류기(예측 변수)들 중에서 각 카테고리 별로 300개의 의사결정나무를 통해 5개의 중요 분류기를 추려내고 다시 최종 의사결정나무로 최적화 했을 때 단 두개의 분류기로 이런 예측력을 보인다는 것은 40대 여성에게 있어 영화관에서의 지출과 가정용TV와 보내는 시간은 경제력, 성격, 가구 구성 등 많은 것들을 내포하고 있다는 의미이다.

〈표 4〉는 50대와 60대 이상의 샘플에 대한 랜덤 포레스트 학습 결과이다. 전체적으로 공통된 트렌드를 분석해 보면, 인구 통계학적 요인에서는 소득수준, 교육수준, 수도권 거주 여부, 종교유무가, 미디어 특성에서는 개인정보에 대한 우려 정도, 월평균 모바일 지출, 모바일 기기에 대한 지출과 영화관 지출이, 성격과 라이프 스타일 영역에서는 탐구적 성격, 적극적 정보 탐색자, 신상 선호 여부, 삶의 만족도, 여가활동으로서 쇼핑 빈도가 중요했다. 미디어 매체 레퍼토리에서는 스마트폰, 데스크탑, 가정용TV가 미디어 연결 레퍼토리에서는 모바일 핸드폰, 모바일 인터넷, IPTV가, 마지막으로 미디어 행위 레퍼토리에서는 메신저, 통화, 지상파TV가 모든 샘플에서 공통적으로 상위 5개의 중요 요인으로 뽑혔다.

〈표-4〉 랜덤 포레스트로 학습한 TV홈쇼핑 구매에 대한 연령별, 성별 주요 요인들(50, 60대)

대분류	50대		60대	
	남성	여성	남성	여성
인구통계학적 요인	소득수준	소득수준	소득수준	교육수준
	교육수준	교육수준	교육수준	소득수준
	수도권거주여부	종교여부	수도권거주여부	기혼
	종교여부	광역시거주여부	종교여부	수도권거주여부
	광역시거주여부	수도권거주여부	반려동물유무	종교여부

▶▶ 표계속

대분류	50대		60대	
	남성	여성	남성	여성
미디어 특성	개인정보에 대한 우려정도	개인정보에 대한 우려정도	개인정보에 대한 우려정도	개인정보에 대한 우려정도
	월평균영화관지출	월평균영화관지출	월평균모바일지출	월평균영화관지출
	월평균모바일지출	월평균모바일지출	월평균영화관지출	월평균모바일지출
	SNS활용정도	월평균모바일기기지출	월평균모바일기기지출	월평균모바일기기지출
	월평균모바일기기지출	SNS활용정도	인터넷활동(글읽기빈도)	SNS활용정도
성격과 라이프 스타일	삶의만족도	삶의만족도	탐구적성격	탐구적성격
	탐구적성격	탐구적성격	적극적정보탐색자	삶의만족도
	적극적정보탐색자	신상선호자	삶의만족도	신상선호자
	신상선호자	적극적정보탐색자	신상선호자	적극적정보탐색자
	쇼핑빈도	쇼핑빈도	쇼핑빈도	쇼핑빈도
미디어 레퍼토리 (매체)	스마트폰	스마트폰	가정용TV	스마트폰
	가정용TV	가정용TV	스마트폰	가정용TV
	데스크탑	데스크탑	데스크탑	데스크탑
	카오디오	카오디오	카오디오	신문
	핸드폰	핸드폰	라디오	핸드폰
미디어 레퍼토리 (연결)	모바일핸드폰	케이블TV	케이블TV	모바일핸드폰
	케이블TV	모바일핸드폰	모바일핸드폰	케이블TV
	모바일인터넷	모바일인터넷	모바일인터넷	모바일인터넷
	유선인터넷	IPTV	IPTV	IPTV
	IPTV	와이파이	안테나	와이파이
미디어 레퍼토리 (행위)	지상파TV	지상파TV	지상파TV	지상파TV
	통화	메신저채팅	통화	통화
	메신저채팅	통화	문자메세지	메신저채팅
	문서작업	정보검색	라디오	문자메세지
	라디오	비지상파TV	메신저채팅	정보검색

〈표 4〉는 50대와 60대 이상의 샘플에 대한 랜덤 포레스트 학습 결과이다. 전체적으로 공통된 트렌드를 분석해 보면, 인구 통계학적 요인에서는 소득수준, 교육수준, 수도권 거주 여부, 종교유무가, 미디어 특성에서는 개인정보에 대한 우려 정도, 월평균 모바일 지출, 모바일 기기에 대한 지출과 영화관 지출이, 성격과 라이프 스타일 영역에서는 탐구적 성격, 적극적 정보 탐색자, 신상 선호 여부, 삶의 만족도, 여가활동으로서 쇼핑 빈도가 중요했다. 미디어 매체 레퍼토

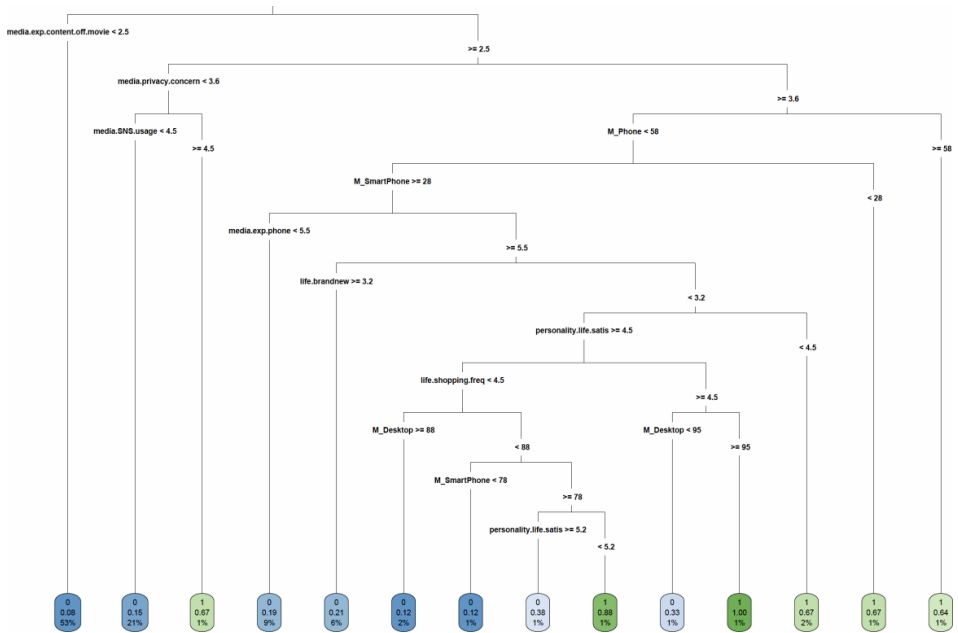
리에서는 스마트폰, 데스크탑, 가정용TV가 미디어 연결 레퍼토리에서는 모바일 핸드폰, 모바일 인터넷, IPTV가, 마지막으로 미디어 행위 레퍼토리에서는 메신저, 통화, 지상파TV가 모든 샘플에서 공통적으로 상위 5개의 중요 요인으로 뽑혔다.

50대 남성은 TV홈쇼핑 구매 입문으로의 비율 자체가 낮는데 소수의 구매 입문자들이 특이한 미디어 레퍼토리를 지나다보니 의사결정나무 자체가 조금 복잡하게 구성되었다. <그림 9>를 보면, 샘플 1%의 100% 확률로 예측을 하고 있는 분류 가지들을 따라가 볼 수 있다. 극장에서의 소비가 연 1만원 이상이며, 개인 정보에 대한 우려가 평균 이상, 핸드폰(스마트폰 제외) 매체를 사용하는 시간은 하루 평균 58분 미만이지만 스마트폰 매체에 대한 시간은 28분 이상, 5만 5천원 이상의 요금제 사용, 상품이 신상인지 여부는 그렇게 중요하게 생각하지 않지만 삶의 만족도가 높고 여가로서의 쇼핑 빈도도 1년에 4.5회 이상, 마지막으로 데스크탑PC 매체 이용 시간이 하루 평균 95분 이상이라면 이에 속한다.

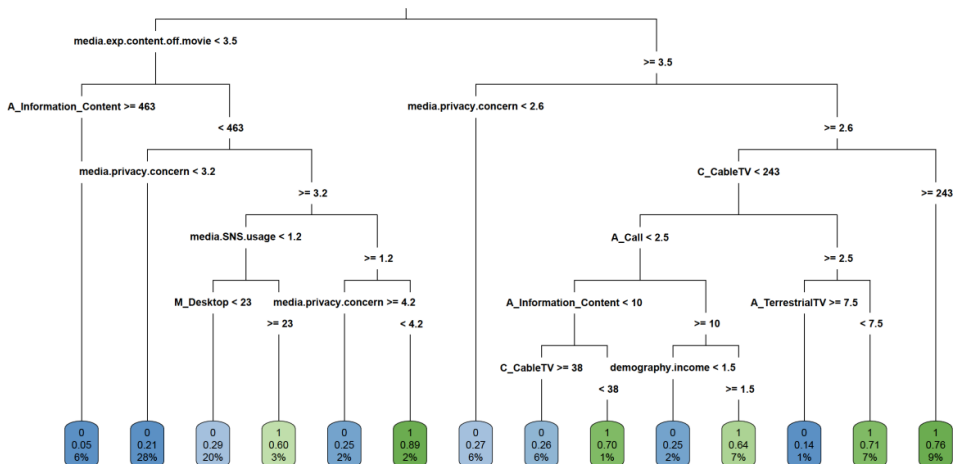
50대 여성의 의사결정나무는 보다 간결하다. 극장에서의 소비가 연 2만원 이상이며, 개인정보 침해에 대한 우려가 아주 낮지 않고(5점 만점에서 2.6이상), 미디어 연결 레퍼토리에서 케이블TV에 연결된 시간이 하루 평균 243분이 넘어 가면, 즉 4시간 이상 시청하는 사람이라면 샘플의 9%에서 76%확률로 TV홈쇼핑 구매 입문을 예측할 수 있고, 그렇게 많은 시간 케이블 TV에 연결되어 있지 않더라도 통화 2.5분 이상, 지상파TV 7.5분 미만을 시청한다면 샘플의 7%에서 71%의 확률로 TV홈쇼핑 입문을 예측할 수 있다.

60대 이상 남성의 경우, 분류기로는 극장 관람 지출, 호기심, 그리고 서울과 수도권 거주 여부가 사용되었다. 극장 관람 지출이 연 3만원 이상이고 개인의 호기심 지수가 2.9이상(평균 2.3), 그리고 서울 및 수도권에 거주한다면 64%의 확률로 TV홈쇼핑에 입문한다(<그림 11> 참조). 60대 이상 여성의 경우는 극장 관람 지출이 2만원 이상이며 SNS 활용지수가 높으면 79% 확률로 TV홈쇼핑에 입문하며 SNS 활용지수가 높지 않을 때는 통화 행위에 10분 이상, 케이블TV 연결에 148분 이상을 쓰고 있을 때 76% 확률로 입문한다는 것을 알 수 있다.

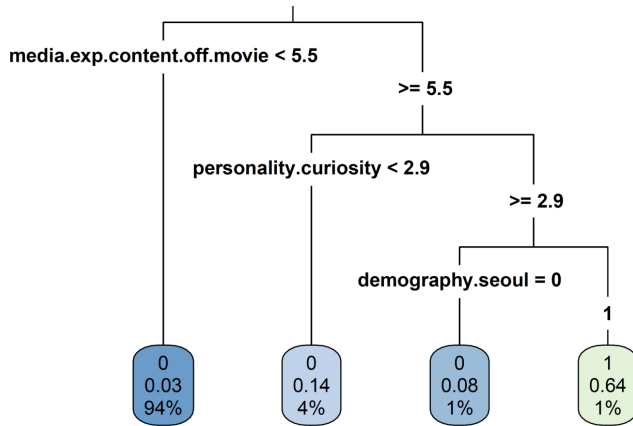




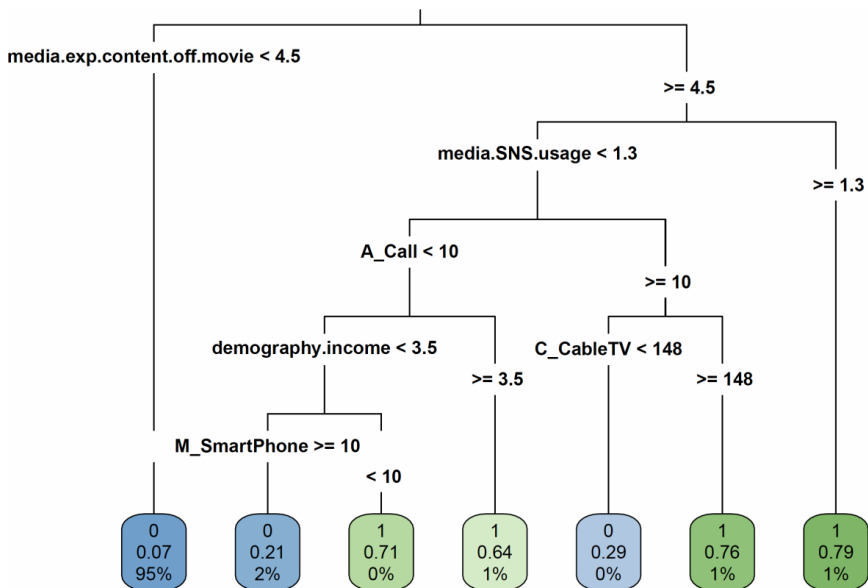
〈그림 9〉 50대 남성의 TV홈쇼핑 구매 입문 의사 결정 나무 결과 시각화



〈그림 10〉 50대 여성의 TV홈쇼핑 구매 입문 의사 결정 나무 결과 시각화



〈그림 11〉 60대 이상 남성의 TV홈쇼핑 구매 입문 의사 결정 나무 결과 시각화



〈그림 12〉 60대 이상 여성의 TV홈쇼핑 구매 입문 의사 결정 나무 결과 시각화

## V. 토의 및 결론

본 논문은 TV홈쇼핑 구매 경험이 없다가 1년 후 구매를 경험하게 되는 사용자는 그렇지 않은 사용자와 비교해 미디어 레퍼토리가 연결, 행위, 매체 차원에서

어떻게 다른지 보기 위해 미디어 패널 조사를 활용하여 2016년과 2017년에 TV 홈쇼핑 구매이력이 없는 사용자들만 추출하여 1년 후 구매로 이어지는지 여부를 추적 조사 하였다. 그 결과, TV홈쇼핑 구매에 입문한 샘플이 그렇지 않은 샘플에 비해 스마트 미디어와 디지털 미디어 사용 비율이 높고, 공간 미디어로서 영화관에서 보내는 시간이 더 많은 특징을 보였다. 또한 문서 작업과 채팅, 메신저 앱을 사용하는 행위가 미디어 레퍼토리의 많은 비율을 차지하고 케이블TV방송에 연결된 미디어보다 IPTV방송에 연결된 미디어에 시간을 보내는 비율이 큰 특징을 보였다. 그 외에도 모바일 통신망에 연결된 인터넷 사용 비율도 TV홈쇼핑 구매 입문 샘플에서 확연하게 크게 나타났다.

다음으로 미디어의 연결, 행위, 매체에 대한 사용자의 레퍼토리 중 어떤 요인이 TV홈쇼핑 구매 의사 결정으로 이어지는데 결정적인 분기점 역할을 하는지에 대한 연구 질문을 풀기 위해 기계학습 알고리즘 의 하나인 랜덤 포레스트를 도입 하였다. 성별과 연령별로 12개의 하위 샘플에서 인구 통계학적 요인, 개인의 성격과 라이프 스타일, 미디어 사용, 미디어 레퍼토리(매체, 행위, 연결)의 여섯 가지 대분류 별로 주요 분기점으로 쓰인 5가지 요인들을 추출하였다. 마지막으로 추출된 30개의 주요 요인들을 가지고 의사결정나무를 구성하여 샘플별로 구체적으로 어떻게 TV홈쇼핑 구매 입문의 의사 결정으로 이어지는지에 대해 알아보았다. 소득수준, 직업유무, 여가활동으로서 쇼핑 빈도 등은 어느 정도 기존 문헌을 통해 예측할 수 있었으나 개인의 미디어 특성에서 나타나는 개인정보에 대한 우려 정도, 영화관에서의 지출 정도, 적극적 정보 탐색자 여부 등이 전 연령, 모든 성별에서 주요 분류기로 뽑힌 것이 흥미로운 결과였다. 이를 통해 기존 문헌을 통해 선정된 통념적 변수들 외에도 결정 요인을 분류하는데 있어 생각지 못한 변수들이 발견되어 분류 효율을 올릴 수 있었다. TV홈쇼핑은 도구적 성격이 강한 목적 지향적 시청에 속하므로 적극적 정보 탐색자와 개인적 호기심 등의 분류 기들이 주요한 역할을 하는 것으로 보인다. 개인정보에 대한 우려가 너무 높으면 인터넷 쇼핑과 TV홈쇼핑 등 모든 원격 쇼핑에 대한 우려가 있기 때문에 TV홈쇼핑 입문 확률도 낮은 것으로 추측되고, 적당히 높으면 인터넷 쇼핑 보다는 그나마 TV홈쇼핑으로의 입문이 수월한 것으로 보인다. 마케팅 관점에서 보면 상품의 속성이나 개인의 속성 중에 관찰 가능한 변수들을 최대한 고려하려고 하지만 개인적 호기심과 같은 관찰되지 않는 속성들이 상품 판매에 있어 훨씬 더 중요한

요소일 수 있기 때문에 이 부분에 대한 추가적인 연구가 필요하다. 또한 이렇게 통상적으로 잘 알려지지 않은 변수들이 주요한 분류기로 발견되는 과정은 데이터를 통해 학습한 랜덤 포레스트 알고리즘을 사용했기 때문에 가능한 결과였다.

샘플별로 선정된 30개의 요인들로 다시 최종 의사결정나무를 구성해 본 결과, 10대 남성의 경우, 다른 샘플에서 발견되지 않았던 분류기인 인터넷 활동 중 글 쓰기 빈도가 주요 분기점 역할을 하였고, 20대 남성의 경우, 미디어 매체 레퍼토리에서 카오디오 사용 시간, 영화관에서의 지출, 개인의 호기심 정도가 주요 분류기로 사용 되었다. 20대 여성 샘플에서 가장 예측율이 높았던 분류 흐름은 카오디오를 하루에 13분 이상 사용하고, 6만 5천원 요금제 이상을 사용하고 있을 때 70% 확률로 TV홈쇼핑에 입문하였다. 30대 남성은 가정용TV 시청 시간이 하루 평균 108분 이상, 월 평균 소득 수준이 350만원 이상, 연 평균 영화관 지출이 3만원 이상일 때 TV홈쇼핑 입문 확률이 가장 높았고, 30대 여성은 소득이 250~300만원 미만, 연 평균 영화관 지출 3만원 이상, 삶의 만족도가 평균보다 높지만 아주 높지 않을 경우 TV홈쇼핑 구매로 첫 발을 들일 확률이 높았다.

40대 여성은 단 두개의 노드로 샘플의 23% 에서 68%를 예측하는 우수한 분류기가 존재했는데, 이 때 사용된 분류기는 영화관 지출과 가정용TV 미디어의 연결 시간이었다. 50대 남성은 TV홈쇼핑 구매 입문으로의 비율 자체가 낮고 소수의 구매 입문자들이 각각 서로 다른 미디어 레퍼토리를 지나다보니 의사결정나무가 조금 복잡했다. 이 때 사용된 분류기는 극장에서의 소비, 개인정보에 대한 우려, 핸드폰과 스마트폰 매체를 사용하는 시간, 모바일 요금제, 구매 특성, 삶의 만족도, 쇼핑 빈도로 총 8개였다. 50대 여성의 의사결정나무는 보다 간결했다. 극장에서의 소비가 연 2만원 이상이며, 개인정보 침해에 대한 우려가 아주 낮지 않고, 미디어 연결 레퍼토리에서 케이블TV에 연결된 시간이 하루 평균 243분이 넘어가면 76%확률로 TV홈쇼핑 구매 입문을 예측할 수 있었다. 60대 이상 남성의 경우, 분류기로는 극장 관람 지출, 호기심, 그리고 서울과 수도권 거주 여부가 사용되었고, 60대 이상 여성의 경우는 극장 관람 지출이 2만원 이상이며 SNS 활용지수가 높으면 79% 확률로 TV홈쇼핑에 입문하였다.

본 연구는 기존 연구에서 주로 수행 되어진 이론과 가설 주도의 요인 분석이 아닌 데이터를 통해 알고리즘이 학습하여 주요 변수들을 추천하는 방식을 주로 채택하였다. 물론 이 과정에서 기존의 이론들을 바탕으로 데이터와 잠재 변수들

을 채택하였다. 최종적으로는 연구자의 주관에 따라 추천된 변수들 중 상위 몇 개만을 추려서 다시 의사결정나무를 최적화 하는 과정을 거쳤다. 인공지능 시대의 도래에도 불구하고 학자들의 연역적 사고를 통한 이론의 확장은 당연히 학문의 큰 축을 담당할 것이다. 하지만 이론과 가설 주도의 연구가 알고리즘의 도움을 받을 수 있다면 학문의 발전 속도는 비약적으로 빨라질 것이라 생각한다. 본 연구는 그런 맥락에서 이론주도와 데이터주도의 방식을 혼합하여 사회과학 연구가 앞으로 나아갈 방향을 제시하고 있다는 데에 학문적 의의가 있다. 또한 성별, 연령별로 각기 다른 샘플에서 TV홈쇼핑으로의 구매 입문 확률을 예측해 봄으로써 특정 샘플에서 어떤 이용자가 더 잠재적으로 구매로 이어질 확률이 큰지에 대한 마케팅적 시사점을 제공한다.

## [참고문헌]

- 곽승준 · 최은철 · 이주석 (2018), “국내 TV 홈쇼핑 산업 현황과 시장규모 예측,” 『Journal of The Korean Data Analysis Society』, 20(3), 1369-1380.
- 김주영 · 구양숙 (1997), “케이블 TV 홈쇼핑을 통한 의복구매시 소비자 반응과 추구이점 및 위험지각,” 『한국의류학회지』, 21(6), 1082-1093.
- 김하나 · 김숙응 (2019), “TV 홈쇼핑의 실버층 구매선택요인과 만족도 분석,” 『기업경영연구』, 26(1), 155-186.
- 박광희 (1995), “중독적 구매행동에 관한 연구,” 『소비자학연구』, 6(1), 85-98.
- 박웅기 · 한예진 (2014), “Q 방법론을 이용한 TV 홈쇼핑 구매 선호자 유형연구,” 『한국방송학보』, 28(4), 117-156.
- 박정우 · 이영주 (2014), “디지털유료방송가입자의 TV 홈쇼핑 채널의 시청행위와 구매동기가 반복 구매에 미치는 영향에 관한 연구,” 『한국언론학보』, 58(3), 36-58.
- 배지양 (2018), “TV 홈쇼핑 소비자의 지각된 위험과 TV 홈쇼핑기업의 관계유지전략의 효과: 만족, 불만족, 불평, 재구매의 도를 중심으로,” 『광고학연구』, 29(8), 33-69.
- 손승혜 · 최윤정 · 황하성 (2011), “기술수용모델을 이용한 초기 이용자들의 스마트폰 채택 행동 연구,” 『한국언론학보』, 55(2), 227-251.
- 심미영 (2002), “홈쇼핑 소비자의 상품특성별 구매결정요인 및 소비자만족도 분석-Cable TV 를 중심으로,” 『소비자학연구』, 13(4), 169-199.
- 유수현 · 차경옥 (2004), “인터넷쇼핑과 TV 홈쇼핑 소비자의 구매만족도 비교분석,” 『한국가족자원경영학회지』, 8(1), 61-78.
- 윤덕환 · 채선애 · 송으뜸 · 김윤미 (2017), “2017 TV 홈쇼핑 이용 및 전반적 인식 조사,” 『리서치보고서』, 2017(6), 1-21.
- 윤태진 · 옥미애 (2007), “시청과 소비, 그리고 즐거움: 홈쇼핑 채널 시청자들의 일상에 관한 연구,” 『커뮤니케이션학 연구』, 15(4), 25-45.
- 이광훈 (2007), “인터넷 및 TV 홈쇼핑 채널 간의 소비자 선호 결정 요인,” 『유통연구』, 12(4), 27-47.
- 이상봉 (2011), “TV 홈쇼핑 채널의 이용과 충족연구,” 『한국콘텐츠학회논문지』, 11(12), 241-249.

- 이지은 (2006), “TV 홈쇼핑에서 판매 촉진 유형이 소비자 구매 의도에 미치는 영향: 의류 제품을 중심으로,” 『중앙대학교 대학원 석사학위논문』, 42-44.
- 전현진 · 임채이 (2016), “TV 홈쇼핑 채널특성이 만족도 및 재구매 의도에 미치는 영향-뷰티상품을 중심으로,” 『한국인체미용예술학회지』, 17(3), 79-92.
- 정순희 · 이초인 · 차경옥 (2004), “전업주부들의 TV 홈쇼핑 중독구매성향에 관한 연구,” 『한국가족자원경영학회지』, 8(3), 113-125.
- 한성희 · 양세정, (2009), “TV 홈쇼핑 이용현황 및 소비자 만족도에 관한 연구,” 『소비자정책교육연구』, 5(4), 111-132.
- Cortese, J. and A.M. Rubin (2010), “Uses and gratifications of television home shopping,” *Atlantic Journal of Communication*, 18(2), 89-109.
- Darian, J.C. (1987), “In-home shopping: are there consumer segments?,” *Journal of retailing*.
- James, E.L. and I.C. Cunningham (1987), “A profile of direct marketing television shoppers,” *Journal of Direct Marketing*, 1(4), 12-23.
- Lian, J.W. and D.C. Yen (2013), “To buy or not to buy experience goods online: Perspective of innovation adoption barriers,” *Computers in Human Behavior*, 29(3), 665-672.
- Rubin, A.M. (1994), “Media uses and effects: A uses-and-gratifications perspective.”

## [부록]

〈표-A1〉 사용 변수 목록

대분류	소분류
미디어 특성	스마트폰 보유대수
	스마트폰 3G
	스마트폰 LTE
	스마트폰 LTE-A
	가입 이동 통신사 KT
	가입 이동 통신사 SKT
	가입 이동 통신사 LG U+
	가입 이동 통신사 MVNO
	웨어러블 기기 보유 여부
	월평균 휴대폰 이용 총 금액
	월평균 기기 할부금
	온라인 디지털 콘텐츠 월평균 지출금액 - 신문/잡지/책(웹툰 포함)
	온라인 디지털 콘텐츠 월평균 지출금액 - 동영상/영화
	온라인 디지털 콘텐츠 월평균 지출금액 - TV 방송 프로그램
	온라인 디지털 콘텐츠 월평균 지출금액 - 음악
	온라인 디지털 콘텐츠 월평균 지출금액 - 게임
	유료 애플리케이션 구입 경험여부
	유/무료 애플리케이션 내 아이템 및 확장기능 구입 경험여부
	애플리케이션 관련 지출 금액(리코드)
	극장에서의 영화 관람 지출 금액
	공연 관람 지출 금액
	자주 이용하는 스마트 기기 애플리케이션=쇼핑
	지난 3개월 동안 인터넷 동호회/카페/클럽 글 읽기 활동 빈도
	지난 3개월 동안 인터넷 동호회/카페/클럽 댓글 달기 활동 빈도
	지난 3개월 동안 인터넷 동호회/카페/클럽 게시물 스크랩 활동 빈도
	지난 3개월 동안 인터넷 동호회/카페/클럽 글 쓰기 활동 빈도
	지난 3개월 동안 인터넷 뉴스/토론 게시판 댓글, 글 쓰기 활동 빈도
	지난 3개월 동안 인터넷 뉴스/토론 게시판 게시물 스크랩 활동 빈도
	지난 3개월 동안 온라인 투표 참여 빈도
	지난 3개월 동안 온라인 추천, 평점 주기 기능 활동 빈도
	지난 3개월 동안 인터넷 지식 서비스 질문 글 쓰기 활동 빈도
	지난 3개월 동안 인터넷 지식 서비스 답변 글 쓰기 활동 빈도
	지난 3개월 동안 온라인 상 유용한 정보 등록 활동 빈도



대분류	소분류
	지난 3개월 동안 SNS 답변, 글 쓰기 활동 빈도
	지난 3개월 동안 SNS 정보 공유 활동 빈도
	지난 3개월 동안 SNS 추천, 평점 주기 기능 활동 빈도
	프라이버시 침해에 대한 우려 정도
성격과 라이프 스타일	여가활동으로서의 쇼핑활동 빈도
	제품 구매 행태 - 최신 상품인지가 중요하다
	제품 구매 행태 - 남들이 잘 쓰지 않는 상품이 더 매력적이다
	제품 구매 행태 - 유행에 민감하다
	제품 구매 행태 - 주위 사람이나 인터넷에서의 평판이 구입에 있어서 중요하다
	제품 구매 행태 - 최소한의 기능으로 만족하며 가격이 저렴한 것이 중요하다
	제품 구매 행태 - 유명 브랜드 여부가 중요하다
	제품 구매 행태 - 상품을 직접 눈으로 확인하고 가격을 흥정하기 위하여 관련 매장을 돌아다니며 구매한다
	제품 구매 행태 - 내가 사용할 상품은 내가 직접 구입한다
	제품 구매 행태 - 상품에 대해서 알아보고 저렴하게 구입하기 위해서 인터넷을 활용한다
	나는 어떤 문제에 대해 새로운 해결방법을 배울 때 즐겁다
	나는 복잡한 문제를 보면 왠지 그것을 분석해 보고 싶은 생각이 든다
	나는 잘 모르는 것에 대해서 더 알고 싶어 한다
	나는 어렵고 시간이 많이 드는 문제도 가능한 한 끝까지 풀기위해 노력한다
	개인적 측면 삶의 만족도
	관계적 측면 삶의 만족도
	집단적 측면 삶의 만족도
	나는 나의 미래에 대해서 낙관적이다
	건강상태_불안/우울
인구통계학적 요인	서울, 수도권 거주 여부
	광역시 거주 여부
	가구원 수
	가족 구성=3세대 동거
	성별
	나이
	최종학력
	결혼 여부
	개인 월평균 소득
	직업 유무
	종교 유무
	반려동물 보유 여부

대분류	소분류
미디어 레퍼토리 (매체)	신문/책/잡지
	그림/사진/편지
	가정용TV
	전광판/옥외TV
	휴대용TV
	차량용TV
	데스크탑PC
	일반노트북PC
	넷북
	태블릿PC
	PDA
	내비게이션
	키오스크
	공중전화기
	일반전화기
	인터넷전화기
	일반휴대폰
	PDA폰
	스마트폰
	디지털카메라
	비디오녹화기기
	일반라디오
	가정용오디오
	카오디오
	오디오레코더
	휴대용오디오
	VCR
	DVD플레이어
	PVR/DVR/DivX플레이어
	휴대용비디오재생기
	휴대용게임기
	가정용게임기
	영화관
	노래방
	멀티미디어방
	공연장

대분류	소분류
	갤러리
	박물관
	스포츠경기장
	홈CCTV/블랙박스
	VR/AR(가상/증강현실)
	드론기기
미디어 레퍼토리 (행위)	지상파TV(실시간)
	지상파TV(VOD)
	비지상파TV(실시간)
	비지상파TV(VOD)
	TV테이터방송프로그램
	라디오방송/음악채널프로그램(실시간)
	라디오방송/음악채널프로그램(AOD)
	영화/동영상
	UCC(실시간)
	UCC(다시보기/듣기)
	음악/음원
	사진(그래픽이미지)
	신문기사
	책
	잡지
	통화
	문자메시지
	이메일
	채팅/메신저
	정보콘텐츠
	SNS
	온라인상거래
	게임
	문서작업프로그램
	그래픽프로그램
	그림
	사진(앨범)
	편지
	옥외/지하철방송프로그램
	영화관-영화관람

대분류	소분류
	노래방-노래부르기
	멀티미디어방-DVD시청/게임하기
	공연장-연극/뮤지컬관람
	갤러리-전시회관람
	박물관-박물관관람
	스포츠경기장-스포츠경기관람
	종합편성채널TV방송프로그램(실시간)
	종합편성채널TV방송프로그램(VOD)
	홈CCTV/블랙박스확인
	원격제어
미디어 레퍼토리 (연결)	케이블TV방송서비스
	IPTV방송서비스
	위성방송서비스
	지상파DMB서비스
	위성DMB서비스
	지상파안테나
	유선전화서비스
	이동통신전화서비스
	유선인터넷
	이동통신무선인터넷(2G,3G,LTE)
	와이파이존-유무선공유기이용
	와이파이-에그/브릿지/단비
	와이브로서비스
	패키지미디어를해당기기에재생
	자체기능/저장된파일
	다른미디어기기에연결
	정기구독및우편/택배/배달
	서점/가판대구입및도서관
	기타(출처불분명)
	공간미디어
	OTT단말기

# Factors of newly entering into the purchase on TV home-shopping: Random Forest Analysis based on users' media repertoire data

Changjun Lee\* · Hyungjin Woo\*\* · Sungbok Park\*\*\*

---

## Abstract

The purpose of this study is to analyze the media repertoire, which consists of media device, connectivity, activity of users who had no experience of TV home-shopping but newly entered into the purchase on it, and find the factors of the first purchase by using machine learning algorithms like decision tree and random forest. Firstly, data of 171 potential variables in the total sample (N=12,876) is used for random forest learning algorithm to find the importance of variables. Previously confirmed variables like income-level and employment status were still valid, but the result also shows that privacy concern, expense in the theater, active information seekers were also found as new and important factors of the first purchase on TV home-shopping. The analysis was conducted in 12 sub-sample based on the users' gender and their age group. The most important 30 variables were selected to enter into the decision tree algorithm in each sub-sample. This paper mixed the data-driven method on the existing theory-driven method to find the factors of the media service, and this shows the future direction of this kinds of research. Marketing implications are also given by predicting the possibility of the first purchase on TV home-shopping in the 12 different segments of users,

**Key Words ;** TV Home-shopping, Factor analysis, Media repertoire, Random Forest, Decision Tree

---

\* Assistant Professor, Media & Social Informatics, Hanyang University ERICA, changjunlee@hanyang.ac.kr

\*\* Corresponding author, Professor, Media & Social Informatics, Hanyang University ERICA, hyungjinw@hanyang.ac.kr

\*\*\* Associate Professor, Media & Social Informatics, Hanyang University ERICA, octobernine@hanyang.ac.kr

