

인공신경망 모형과 로지스틱 회귀모형을 이용한 프로야구 관전자 분석 및 예측

정승훈* (경희대학교)

이 연구의 목적은 첫째, 다양한 연구 분야에서 시장세분화의 기준 변수로 측정된 프로야구 관전자의 인구통계학적 특성과 라이프스타일을 측정하여 프로야구 관전자를 세분화하는데 있다. 둘째, 각 세분화된 집단을 인공신경망 및 로지스틱 회귀분석 모델 적용시켜 목표변수(관전만족과 미래 소비행동)에 가장 높은 가능성(분류적중률)을 가진 프로야구 관전자 집단을 분석 및 파악하여 그에 따른 마케팅 전략을 제시하는데 그 목적이 있다. 이를 위해 SPSS 18.0, AMOS 18.0 및 Modeler 14.1을 이용하였다. 먼저 인구통계학적 특성 및 문항의 타당성을 측정하기 위하여 빈도분석, 탐색적 요인분석, 확인적 요인분석 및 신뢰도 분석을 실시하였다. 또한 계층적 군집과 K-means 군집을 병행한 2단계 군집분석을 통해 군집을 세분화 한 후 분류된 집단을 인공신경망 모형과 로지스틱회귀 모형에 적용하여 다음과 같은 결론을 도출하였다. 첫째, 프로야구 관전자를 세분화 한 결과 5개의 군집으로 분류하였다. 둘째, 인공신경망 모형을 이용하여 관전만족과 미래 소비행동이 가장 높은 가능성을 가진 집단을 파악한 결과 군집5로 나타났다(관전만족: 71.3%, 미래 소비행동: 99.3%). 셋째, 로지스틱 회귀모형을 이용하여 관전만족과 미래 소비행동이 가장 높은 가능성을 가진 집단을 파악한 결과 로지스틱 회귀모형 역시 군집5(관전만족: 71.8%, 미래 소비행동: 96.5%)가 가장 높게 나타났다. 넷째, 인공신경망 모형과 로지스틱 회귀모형을 비교하여 관전만족과 미래소비행동이 가장 높은 분류적중률을 가진 집단을 파악하였으며 '20대 미혼 남성 타인 지향적 소비를 가진 정보선호 집단'으로 명명하였다.

주요어: 인공신경망 모형, 로지스틱 회귀모형, 프로야구 관전자, 예측

서 론

예측(prediction)이론의 종류에는 여러 가지가 있는데 그 중에 기계학습에 의한 예측기법이 있고 통계를 바탕으로 한 예측기법이 있다. 기계학습에 의한 예측기법으로는 인공신경망(artificial neural networks)을 이용한 방법과 유전자 알고리즘(genetic algorithm)을 이용한 방법 등이 있으며, 통계기반의 예측기법은 로지

스틱 회귀분석(logistic regression analysis)과 시계열분석(time series analysis) 등으로 구분할 수 있다(김영신, 2003; 박원규, 2002; 정혜원, 2004; 이정학과 정승훈, 2012 재인용).

먼저 기계학습의 대표적인 예측 기법인 인공신경망은 산업, 생산 프로세스의 통제와 최적화, 예측, 패턴인식 등을 위해 광범위하게 사용되며(Aschenwald et al., 2001), 인간의 신경체계와 유사한 성능과 특성을 갖는 정보처리 시스템의 하나로써 뇌의 뉴런들이 상호작용하고 경험을 통해 배우는 생물학적 활동을 모형화한 것이다(신택수와 홍태호, 2006). 즉, 인간의 뇌가 대량의 데이터를 효율적, 병렬적으로 처리 및 학습할 수 있다는 사실에 근거하여, 인간의 생물학적 신경세포를 모델링하여 구현하는 방식으로서 신호처리, 제어, 패턴인식, 음성,

논문 투고일: 2015. 02. 16.

논문 수정일: 2015. 03. 11.

게재 확정일: 2015. 03. 19.

* 저자 연락처: 정승훈(hoon@khu.ac.kr).

* 이 논문은 2012년 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (NRF-2012S1A5B5A07036303)

문자인식, 예측 등의 분야에서 널리 응용되고 있다 (Fausett, 1994; Haykin, 1994; Nelson & Illingworth, 1991). 뿐만 아니라 경영학 분야에서도 기존의 통계 방법과 비교 및 분석을 통하여 반복적인 학습을 거쳐 데이터에 대한 패턴을 찾아내고 이를 일반화함으로써 향후를 예측하는데 있어 유용하게 이용되고 있다(이종진, 2003).

반면, 통계를 기반으로 한 대표적인 예측 방법인 로지스틱 회귀분석은 회귀분석 중 비선형회귀분석(nonlinear regression analysis)에 속하는 방법으로 일반 회귀분석(예: 단순회귀분석 및 다중회귀분석; 독립변수의 개수에 따라 달라짐)과 달리 한 변수가 다른 변수에 의해 설명(explanation)이나 예측되는지를 분석하기 위하여 자료를 함수식으로 표현하여 분석하는 방법을 말한다(강현철 등, 2001). 또한 이는 두 개 또는 그 이상의 값을 가지는 변수를 예측하는데 사용될 수 있으나 일반적으로 두 개의 값만을 가지는 변수에 대한 예측모형을 주로 설계하면 된다. 즉, 여러 입력 변수로부터 두 범주만을 가지는 종속변수를 예측하는데 사용하는 로지스틱 회귀분석은 모형구조에 의해 연관성 및 상호작용 유형을 설명할 수 있으며, 예측확률을 바탕으로 판별분석과 같은 판별 및 예측분석의 기법으로도 사용할 수 있는데 로지스틱 회귀분석은 판별분석에 비해 관련변수들이 정규 분포이어야 한다는 가정이 전제될 필요가 없다(김동성, 2004).

이상과 같이 예측에는 다양한 분석방법이 이용될 수 있다. 이해경(2001)은 소비자행동에 대한 예측이란 자료를 이용하여 학습을 통해 모델을 설정하고 새로운 데이터에 대하여 예상되는 값을 발견하는 것이라고 하였다. 따라서 대표적인 예측 분석 간의 비교를 통하여 스포츠 소비자를 예측할 수 있으며, 스포츠 마케팅 분야에서 적합한 예측 방법을 선택하여 예측력을 비교하고 이를 적용하는 것은 매우 의미 있는 일이라고 할 수 있다(박진기, 2003). 즉, 스포츠 분야에서 소비자의 특성을 파악하고 이러한 소비자 및 관전자가 어떠한 패턴으로 이루어져 있는지를 파악하는 것은 스포츠 소비자행동 연구에서 매우 중요한 부분이다(박진기와 김장환, 2003).

스포츠 분야의 예측과 관련된 선행연구를 살펴보면, 박진기와 김장환(2003)은 스포츠센터 월 회원을 대상으로 이탈유무에 따른 세분화와 이탈 가능성을 예측하고자 의사결정나무, 인공신경망 및 로지스틱 회귀분석을

통하여 연령, 직업, 이용형태 및 불평행동 등이 중요한 예측의 역할을 하는 것을 밝혀내었다. 조용찬과 유재구(2013)는 프로축구 소비자의 소비지출을 기준으로 고, 중, 저 소비 팬으로 시장을 세분화하여 이들의 소비지출을 바탕으로 한 소비행동 예측이 가능한 프로파일도 도출하였다. 즉, 프로축구 소비자의 개인적인 특성은 명확한 분류 및 예측이 어렵지만 행위적 특성인 소비행동은 분류 및 예측이 가능하여 소비자행동 예측에 대한 중요성을 강조하였다. 뿐만 아니라 다양한 스포츠 분야에서 예측과 관련된 연구가 이루어지고 있으며(예: 고재곤과 나상준, 2001; 오광모와 이장택, 2003 등), 스포츠 소비자를 예측하기 위한 방법으로는 세분화(군집분석), 인공신경망, 의사결정나무 및 로지스틱 회귀분석 등이 주로 이용되고 있다.

이에 따라 이 연구는 기계 학습적 시뮬레이션 방법인 인공신경망과 통계 방법인 로지스틱 회귀분석을 이용하여 목표변수인 관전만족과 미래소비행동에 대한 예측률이 높은 프로야구 관전자 집단을 분석하고자 한다. 그러나 지금까지 인공신경망과 관련된 선행연구들을 살펴보면, 인공신경망은 예측에 대한 결과를 제공할 뿐 어떤 변수가 종속변수에게 중요한 영향을 미쳤는지, 어떠한 상호작용에 의해 결과물이 도출되었는지를 나타내지 못함으로써 각 변수에 대한 설명력 보다는 정확한 예측이 필요할 때 주로 이용된다는 단점이 있다(Berry & Linoff, 1997).

뿐만 아니라 로지스틱 회귀분석 역시 인공신경망과 마찬가지로 다음과 같은 특징과 그에 따른 한계점을 가지고 있다. 먼저 가능한 많은 입력 변수를 분석에 포함시키는 것이 예측을 위해 좋은 기준이 되며, 입력 정보가 많을수록 추정이 좋아진다. 그러나 부적절하거나 관련성이 없는 입력변수를 포함시키는 것은 일반화에 역기능적인 효과를 가져 올 수 있고 이는 곧 모델의 불안정성의 원인이 될 수 있기 때문에 사전에 충분한 탐색을 통해서 변수들을 제한하거나, 변수의 선택 및 분석 방법 등과 같은 다양한 방법을 이용하여 이러한 문제점을 극복하기 위해 노력을 기울여야 한다(강현철 등, 2001).

따라서 예측과 관련된 선행연구들은(김영신, 2003; 이창환 등, 2009; 이호영 등, 2006; Balakrishnan et al., 1996; Berry & Linoff, 1997) 다양한 분석과

결과를 통해 이러한 문제점을 극복하고자 다음과 같은 방법을 제안하였다. 즉, 단순히 인공신경망을 이용하여 소비자의 특성을 예측하는 방법 보다는 1차적으로 각 집단을 분류하고 이러한 집단을 세분화하여 예측하는 것이 분류적중률(hit ratio)을 높일 수 있는 보다 좋은 방법이며, 인공신경망의 단점으로 지적된 어떤 변수가 중요한 변수인지를 파악하지 못하는 부분에 대하여 각각 집단의 분류적중률을 통해 목표변수에 대한 예측 가능성이 높은 집단의 특성을 파악할 수 있다고 하였다. 뿐만 아니라 군집분석의 보완 방법으로 인공신경망을 제시함으로써 보다 좋은 예측성과 및 결과를 얻을 수 있으며, 인공신경망 모형은 마케팅 분야에서 소비자행동에 가장 유망한 분야라고 주장 하였다(이정학과 정승훈, 2012; Balakrishnan et al., 1996; Vellido et al., 1999).

한편, 국내 프로야구는 2011년 700만 관중을 돌파하였고 이를 계기로 2012년 약 753만을 기록하여 대한민국 최고의 프로스포츠로 자리매김하였다. 그러나 2013년 약 674만으로 관중 수가 감소하였고, 2014년 역시 약 675만으로 2011년과 2012년에 비해 관중 수는 감소하고 있는 실정이다(문화체육관광부, 2014; 한국야구위원회, 2014).

정승훈(2007)은 프로야구와 같은 프로스포츠에서는 관중 수의 추이가 시장을 유지·발전시키는데 있어 일차적 변수이므로 투자의 개념인 단기적인 마케팅전략 보다는 가치의 극대화라는 장기적 안목을 갖고 프로스포츠 관전자에 대한 정보를 체계적이고 지속적으로 수립할 필요성이 있다고 주장하였다. 이에 따라 체계적 및 과학적으로 마케팅 전략을 수립하여 관전자를 유인할 수 있는 마케팅 전략의 수립이 필요하다고 할 수 있다(Gladden et al., 2001).

이정학과 정승훈(2007)은 프로스포츠 관전자들의 다양한 욕구를 만족시키기 위해 각 구단들은 지속적으로 관전자의 인구통계학적 특성과 소비자행동을 파악하고 각종 데이터를 토대로 소비자의 특성에 맞는 세분화된 목표시장을 선정하여 이를 통한 다양한 마케팅전략을 펼쳐야 한다고 주장하였다. 여기서 말하는 세분화된 목표 시장은 한 시장에서의 고객을 여러 그룹으로 나누는 과정을 말하며, 세분화된 시장은 유사한 성향이나 경험을 가지고 있고 시장 자극에 대해 동일하게 반응하는 고객

그룹이라고 정의 할 수 있다(Bonoma & Shapiro, 1983; Sööllner & Rese, 2001).

이에 따라 시장세분화 기준으로 널리 활용되고 있는 라이프스타일은 심리적·행동적 기준을 포함하는 변인으로서 포괄적으로 소비자를 이해하는데 유용하게 사용될 수 있으며(Solomon, 1999), 라이프스타일을 구성하는 체계는 태도 보다 안정적이고 지속적인 속성을 지니고 있어 특정한 대상이나 상황에 국한되지 않기 때문에 소비자의 전반적인 소비행동에 대한 예측 변인으로서 그 효용성이 높다(Kamakura & Novak, 1992).

또한 라이프스타일 유형 분류 체계는 사이코 그래픽 시장세분화 내에 가장 효율적인 시장세분화 간 근거라고 할 수 있다(Lee & Sparks, 2007). 따라서 점차 다양해져가는 소비자들을 이해하기 위해서는 인구통계학적 변수만을 이용하여 시장을 세분화 하는 것은 다소 부족하다고 판단되므로(Sharma & Lambert, 1994; Solomon & George, 1977), 무엇보다 성공적인 마케팅전략을 위해서는 목표 고객층을 명확히 파악하는 것이 중요하며, 시장조사 및 시장세분화를 위해 인구통계학적 특성과 라이프스타일과 같은 변수들을 사용하는 것이 중요하다(Sobel, 1983).

이와 같이 다양한 변수를 기준으로 프로스포츠 관전자들을 세분화하고 각 집단을 분석 하는 것은 프로스포츠 관전자의 소비자행동인 재관전, 추천의사 및 구매의도 등에 대한 소비 및 행동을 예측하는데 중요한 의미가 있으며, 스포츠 소비자의 행동 등에 영향을 미치는 다양한 소비자행동 변수를 분류 및 예측하기 위해서는 먼저 집단을 세분화하는 것이 마케팅의 기본이 된다고 하였다(이정학과 정승훈, 2012; 조용찬과 유재구, 2013).

시장세분화와 함께 소비자를 예측하기 위해서는 목표 변수가 필요한데, 대표적 목표변수인 관전만족은 소비자들의 필요, 욕구, 기대가 제품이나 서비스 활동을 통해서 충족되어지거나 또는 그 수준을 초과하여 충족되어질 때 재구매와 충성도가 형성되어지는 마음의 상태를 의미하며(Anton, 1996), 오랫동안 스포츠 소비자행동을 연구하는데 목표변수로 이용되어 왔다(허진과 이계석, 2004; Trail et al., 2003). 뿐만 아니라 미래 소비행동은 경기를 관전한 후 갖게 되는 재관전의도, 구전의도, 경기관련 제품구매 등을 포함하는 관전자의 행동의도로(정승훈,

2007) 다양한 분야에서 관전만족과 함께 목표변수로 이용되고 있다(Cronin et al, 2000; 정승훈 등, 2011).

따라서 이 연구는 첫째, 다양한 연구 분야에서 시장세분화의 기준 변수로 측정된 프로야구 관전자의 인구통계학적 특성과 라이프스타일을 측정하여 프로야구 관전자를 세분화하는데 있다. 둘째, 각 세분화된 집단을 인공신경망 및 로지스틱 회귀분석 모델을 적용시켜 목표변수(관전만족과 미래 소비행동)에 가장 높은 가능성(분류적중률)을 가진 프로야구 관전자 집단을 파악하는데 있다. 셋째, 목표변수가 가장 높은 집단 및 세분화된 집단의 특성을 분석하고 그에 따른 마케팅 전략을 제시하는데 그 목적이 있다(그림 1).

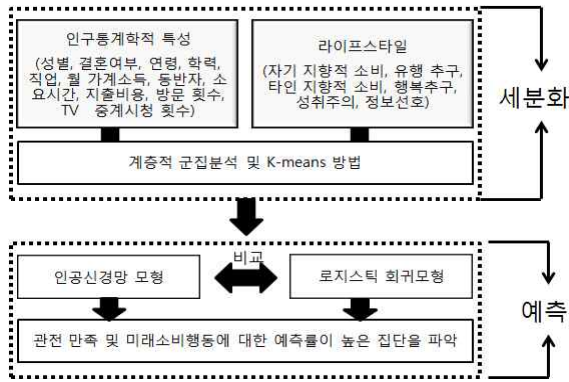


그림 1. 연구 설계 및 목적

연구방법

연구 대상

이 연구의 대상은 2013년 팔도 프로야구 홈팀 롯데(사직), 삼성(대구), SK(문학), 두산(잠실) 및 LG(잠실) 구단의 직접 관전자를 대상으로 각 구단 마케팅 담당자의 사전 동의와 협조를 구한 후 편의표본추출법(convenience sampling method)에 의해 연구대상을 선정하였다. 다음으로 선정된 연구대상에게 연구자를 포함한 보조연구원 4명과 함께 2인 1조로 경기장에서 연구 목적을 설명한 후 설문지를 이용하여 각 구단별 300부씩 총 1,500부의 설문지를 배포하여 회수하였다. 이 중 회수된 설문지는

1,431(95.4%)부였으며, 수집된 자료 중 응답내용이 부실하거나 연구목적에 부합하지 않는 116부를 제외한 최종 1,315(91.9%)부를 결과 분석에 사용하였다. 연구대상자의 인구통계학적 특성은 다음 <표 1>과 같다.

표 1. 연구대상자의 인구통계학적 특성

변수	구분	N	%
성별	남	751	57.1
	여	564	42.9
결혼 유무	결혼	327	24.9
	미혼	988	75.1
연령	20대	830	63.1
	30대	314	23.9
	40대 이상	171	13.0
학력	고졸 이하	200	15.2
	대학 재학	420	31.9
	대학교 졸업	609	46.3
	대학원 이상	86	6.5
직업	학생	454	34.5
	회사원	454	34.5
	공무원	70	5.3
	서비스업	110	8.4
	자유업	82	6.2
	가정주부	52	4.0
	기타	93	7.1
(월) 가계소득	100만원대	112	8.5
	200만원대	325	24.7
	300만원대	272	20.7
	400만원대	182	13.8
	500만원대	170	12.9
	600만원 이상	254	19.3
관전 동반자	가족	336	25.6
	친구	492	37.4
	연인	329	25.0
	동료 및 선후배	123	9.4
	기타	35	2.7
경기장까지 소요시간	30분 미만	293	22.3
	30분 이상-60분 미만	567	43.1
	60분 이상-90분 미만	294	22.4
경기장에서 개인당 지출한 비용 (입장권 제외)	90분 이상	161	12.2
	1만원 이하	136	10.3
	1만원-1만5천원 미만	303	23.0
	1만5천원-2만원 미만	315	24.0
	2만원-2만5천원 미만	155	11.8
프로야구 관전 횟수(1년)	2만5천원-3만원 미만	118	9.0
	3만원-3만5천원 미만	125	9.5
	3만5천원 이상	163	12.4
프로야구 관전 횟수(1년)	1회(처음 방문)	154	11.7
	2회-5회 이하	635	48.3
	6회-9회 이하	217	16.5
프로야구 관전 횟수(1년)	10회 이상	309	23.5
	1회 이하	67	5.1
	2회-5회 이하	156	11.9
프로야구 관전 횟수(1년)	6회-9회 이하	115	8.7
	10회 이상	977	74.3
	합계	1,315	100

조사 도구

시장세분화 변수로 선정된 라이프스타일 문항은 Plummer(1971)가 개발하고 국내 연구에서 박찬민(2004), 유재구(2009), 황혜선(2006), 허성수와 전태준(2011) 및 정승훈(2011)의 연구에서 사용된 설문 문항을 사용하였다. 다음으로 목표변수인 관전만족은 정승훈 등(2011)의 연구에서 사용된 문항을 사용하였으며, 미래 소비행동은 이정학과 정승훈(2007)의 연구에서 사용된 문항을 사용하였다. 이와 같이 선행연구를 통해 선택된 설문 문항은 먼저 연구자가 본 연구의 목적에 맞도록 설문지를 수정 및 보완하였으며, 각각의 항목에 대한 전문가의 검토(3인의 스포츠 경영학 교수)를 통해 예비 설문지 초안을 완성하였다. 다음으로 설문 문항의 내용을 프로야구 직접 관전자에게 검증하기 위하여 2013년 5월 3일 LG와 두산의 잠실 경기에서 100부의 설문지를 배포하여 회수하였다. 이상과 같은 예비조사를 통해 내용의 적합성과 적용 가능성을 검토하였으며, 이 과정에서 특정 문항이 삭제되었다(예: 라이프스타일 문항 중 혼전순결에 대한 문항). 다음으로 기초적인 통계 분석(탐색적 요인분석 및 신뢰도 분석)을 통해 신뢰도가

낮은 문항들을 재검토 하였으며, 최종 설문지의 내용 타당도를 분석하기 위하여 5명의 전문가(스포츠 경영학 교수 3명, 스포츠 사회학 교수 1명, 프로야구단 마케팅 담당자 1명) 회의를 거쳐 최종적으로 라이프스타일 21 문항, 관전만족 및 미래 소비행동 8문항, 인구통계학적 특성 11문항 등 총 40문항을 확정하여 연구를 진행하였다. 또한 설문지의 척도는 인구통계학적 특성을 제외한 모든 문항은 5점 Likert 척도를 사용하였다.

탐색적 요인분석, 확인적 요인분석 및 신뢰도 분석

설문 문항의 내용 타당성을 검증하고자 탐색적 요인 분석 및 확인적 요인분석을 실시하였다(표 2, 표 3). 탐색적 요인분석은 배리맥스 방법으로 고유치(eigenvalue)가 1.0 이상, 요인 적재치(factor loading)가 .5 이상인 문항들만 선택하여 각 요인에 대한 타당성을 확보하였다(Hair et al., 2006). 또한 Cronbach's α 계수를 이용하여 각 변인별 신뢰도를 분석한 결과 선행연구에서 제시한 기준치를 충족하는 것으로 나타났다(Nunnally & Bernstein, 1994).

다음으로 탐색적 요인분석 및 신뢰도 분석 결과를 토대로

표 2. 프로야구 관전자 세분화를 위한 라이프스타일의 탐색적, 확인적 요인분석 및 신뢰도 분석 결과

변인	문항	EFA 적재치	EFA 고유치(분산%)	CFA 적재치	측정문항 오차	개념 신뢰도	AVE	신뢰도 (α)
자기 지향	다른 모습 개발	.806	3.372 (16.058)	.816	.261	.892	.624	.865
	자신감 있는 이미지	.806		.839	.252			
	소중한 존재	.793		.716	.373			
	좋은 모습을 위해 소비	.766		.727	.354			
	나의 이미지	.688		.650	.467			
유행 추구	유행을 빨리 받아들임	.813	2.450 (11.669)	.791	.324	.854	.599	.777
	유행하는 옷을 구매	.774		.646	.323			
	유행하는 기사나 이야기	.759		.765	.301			
	패션표현	.587		.542	.339			
타인 지향	소유품 사회적 배경	.779	2.195 (10.452)	.607	.206	.863	.612	.724
	지위추족	.763		.677	.266			
	제품이나 브랜드에 민감	.601		.602	.240			
행복 추구	비슷한 소비	.593	1.918 (9.133)	.635	.298	.744	.535	.768
	친구로부터 소외되지 않음	.783		.555	.361			
	진정한 행복 가능	.762		.693	.337			
성취 주의	잘됨	.705	1.517 (7.225)	.652	.357	.783	.550	.742
	관습지킴	.754		.517	.279			
	생활 만족	.663		.688	.266			
정보 선호	어려움 극복	.630	1.431 (6.814)	.640	.398	.812	.687	.764
	가격비교	.812		.778	.211			
	주변사람에게 정보제공	.793		.605	.232			
적합도	$\chi^2/df=3.656$, CFI=.905, GFI=.936, AGFI=.914, SRMR=.049, RMR=.043, RMSEA=.058							

확인적 요인분석을 실시하여 모형의 적합성 및 구성 타당성을 검증하였다. 우선 세분화 기준변수로 사용될 라이프스타일 모형의 적합도는 $X^2=636.085$ ($df=174$, $p=.000$), $X^2/df=3.656$, CFI는 .905, GFI는 .936, AGFI는 .914, SRMR은 .049, RMSEA는 .058로 나타났다. 목표변수인 관전만족도와 미래 소비행동에 대한 모형의 적합도는 $X^2=64.447$ ($df=19$, $p=.000$), $X^2/df=3.391$, CFI는 .951, GFI는 .959, AGFI는 .923, SRMR은 .056, RMR는 .050, RMSEA는 .080으로 나타나 기준치를 충족하는 것으로 나타났다. 다음으로 측정항목의 집중타당성을 검증하기 위하여 개념 신뢰도 (construct reliability: CR)와 평균분산추출지수(average variance extracted: AVE)를 산출한 결과 각각의 값은 모두 선행연구에서 제시한 기준치를 충족하는 것으로 나타나 각 변인은 집중타당성이 확보되었다(Bagozzi & Dholakia, 2002; 김계수, 2007). 따라서 프로야구 관전자의 라이프스타일은 타당성이 확보된 6가지 요인(자기 지향적 소비, 유행추구, 타인 지향적 소비, 행복추구, 성취주의 및 정보선호)으로 분석 되었으며, 관전 만족과 미래소비행동 역시 타당성과 신뢰성이 확보되었다.

자료처리방법

이 연구는 SPSS 18.0, AMOS 18.0 및 Modeler 14.1을 이용하여 다음과 같이 자료를 처리하였다. 첫째, 프로야구 관전자의 인구통계학적 특성을 분석하고자 빈도분석을 실시하였다. 둘째, 설문문항의 타당성 및 신뢰성을 분석하기 위하여 탐색적, 확인적 요인분석 및 신뢰도 분석을 실시하였다. 셋째, 프로야구 관전자의 인구통계학적 특성 및 라이프스타일을 기준으로 시장을 세분화하기 위하여 타당성이 확보된 문항을 표준화된 점수로 변환(Z score)하였으며, 1단계 계층적 군집(hierarchical clustering)과 2단계 K-means 군집을 병행한 군집분석(cluster analysis)을 실시하여 프로야구 관전자를 군집화(세분화)하였다. 넷째, 목표변수인 관전만족도와 미래 소비행동에 대한 가장 분류적중률이 높은 집단을 파악하고자 세분화된 각각의 집단을 인공신경망 모형과 로지스틱회귀모형에 적용시켜 분석하였다. 다섯째, 목표변수에 가장 높은 분류적중률을 보인 집단의 특성을 파악하고자 교차분석(cross-tabulation analysis) 및 일원분량분석(one-way ANOVA)을 실시하였으며, 세분화된 집단 간의 유의한 차이를 검증하고자 Scheffe의 사후검증을 실시하였다.

표 3. 관전만족 및 미래소비행동에 대한 탐색적, 확인적 요인분석, 기술 통계분석 및 신뢰도 분석 결과

변인	문항	EFA 적재치	EFA 고유치(분산%)	CFA 표준화 계수	측정문항 오차	개념 신뢰도	AVE	신뢰도 (α)	평균 (표준편차)
관전 만족도	응대 및 편의에 대한 만족	.837	2.750	.743	.388	.891	.672	.842	3.28 (.753)
	지불한 금액에 대해 만족	.832	(34.376)	.805	.269				
	전반적인 만족도	.809		.817	.219				
	경기장 및 시설만족	.795		.680	.261				
미래 소비 행동	다음 경기도 보러움	.859	2.543	.837	.250	.884	.664	.795	3.85 (.754)
	다른 사람에게 권유함	.836	(31.788)	.834	.255				
	관전한 팀을 지지함	.732		.568	.267				
	유니폼 등을 구매함	.691		.583	.268				
모형 적합도	$\chi^2=64.447(df=19, p=.000)$, $\chi^2/df=3.391$, CFI=.951, GFI=.959, AGFI=.923, SRMR=.056, RMR=.050, RMSEA=.080								

결과 및 논의

군집분석

선행연구를 검토한 결과 군집분석의 경우 한 가지 방

법을 선택하여 결과를 도출하는 것보다 계층적 방법에 의해 군집화 하여 적절한 군집의 수를 추정한 후 다시 비계층적 방법을 이용하여 군집의 수를 최종적으로 결정하는 방법이 제안되고 있다(민준영, 1995; 안광호 등, 2009; 이정학 등, 2010; 이학식, 2012; 정승훈, 2011; Berry & Linoff, 1997; Lloyd, 1982). 따라

서 이 연구는 프로야구 관전자의 인구통계학적 특성 및 라이프스타일을 군집의 기준변수로 선정하여 계층적 방법을 실시한 후 다시 비계층적 방법을 사용하여 프로야구 관전자를 세분화 하였다. 이는 최초 군집의 수를 모를 경우 비계층적 방법을 적용하기 어렵기 때문에 군집의 수를 찾기 위해 계층적 군집을 먼저 실행하기 때문이다(정승훈, 2011).

이에 따라 이 연구는 기준 변수인 인구통계학적 특성(성별, 연령, 학력, 직업, 월 가계소득, 관전 동반자, 경기장까지 소요시간, 경기장에서 지출한 비용, 년 간 프로야구 관전 횟수 및 년 간 프로야구 TV시청 횟수)을 표준화 점수(Z score)로 변환하였고 라이프스타일(자기 지향적 소비, 유행추구, 타인 지향적 소비, 행복추구, 성취주의, 정보선호) 변인에 대한 탐색적 요인분석 결과를 통해 요인점수(factor score)를 얻은 후 이를 다시 표준화 점수로 변환하여 군집분석을 실시하였다. 먼저 비계층적 군집분석은 덴드로그램을 분석하여 군집 간의 거리와 평균을 고려하였으며, 이에 따라 최종적으로 군집의 수를 4개~6개 정도의 범위에서 군집을 결정하는 것이 적절하다고 판단하였다.

다음으로 계층적 군집분석 결과를 토대로 분류된 4개~6개 군집을 대상으로 비계층적 군집분석 방법인

K-means 군집분석을 실시하였다. K-means 군집분석 방법은 연구자가 군집의 기준변수와 군집의 수를 미리 지정함으로써 규모가 큰 자료를 처리하는데 상대적으로 용이하기 때문에(안광호 등, 2009), 본 연구에서는 군집을 각각 4개, 5개, 6개로 지정하여 분석하였다. 분석 결과 군집을 4개로 지정하였을 경우 학력에서 군집의 분류가 유의하지 않게($F = .320, p > .05$) 나타나 4개의 군집은 적절하지 않다고 판단되었다. 다음으로 5개의 군집으로 지정하였을 경우 모든 항목에서 유의하게 나타났으나 군집별 분류된 사례 수(군집1: 202명, 군집2: 110명, 군집3: 193, 군집4: 401, 군집5: 409명)가 차이가 나타나 5개의 군집으로 결정하는 것을 보류한 후 6개의 군집으로 군집수를 지정하여 분석하였다. 6개의 군집으로 설정하여 분석한 결과 5개의 군집과 마찬가지로 모든 항목에서 유의하게 나타났으나 군집별 분류된 사례 수(군집1: 119명, 군집2: 401명, 군집3: 188, 군집4: 164명, 군집5: 150명, 군집6: 293명)가 군집을 5개로 지정하였을 때 보다 각 군집의 특성이 명확하게 나타나지 않는 문제점이 발견되었고 각 군집별 중심간 거리 역시 군집을 5개로 지정하였을 때가 더욱 안정적으로 나타났으므로 최종 5개의 군집으로 결정하였다(표 4).

표 4. 군집분석 결과

세분화 기준변수	최종 군집 중심					평균 제곱	평균 오차	F	p	
	군집1	군집2	군집3	군집4	군집5					
인구 통계 학적 특성	성별	-.146	-.426	-.469	.876	-.451	114.509	.653	175.249	.000***
	결혼	-1.703	-1.738	.503	.540	.541	300.968	.084	3580.101	.000***
	연령	1.304	1.541	-.321	-.517	-.401	199.367	.394	505.625	.000***
	학력	.856	-1.636	-.177	-.001	.102	113.121	.658	172.008	.000***
	직업	.518	.359	-.350	.084	-.270	31.148	.908	34.306	.000***
	월 가계소득	.597	-.097	-.345	-.386	.273	46.571	.861	54.099	.000***
	관전 동반자	-.762	-.450	.291	.168	.196	45.701	.864	52.924	.000***
	경기장까지 소요시간	-.257	-.126	.291	-.092	.114	10.034	.972	10.318	.000***
	경기장에서 지출한 비용	.408	-.022	-.392	-.326	.309	36.232	.892	40.600	.000***
	년 간 프로야구 관전 횟수	-.044	-.270	.159	-.489	.498	52.717	.842	62.604	.000***
년 간 프로야구 TV시청 횟수	.325	.128	.397	-.886	.487	116.355	.648	179.625	.000***	
라이프 스타일	자기 지향적 소비	-.295	-.515	.182	.163	.039	16.101	.954	16.879	.000***
	유행추구	.001	-.265	-.530	.171	.154	20.833	.939	22.175	.000***
	타인 지향적 소비	-.042	-.198	-.202	-.119	.286	12.927	.964	13.415	.000***
	행복추구	.037	-.319	-.060	.036	.060	3.539	.992	3.567	.007**
	성취주의	.352	.100	-.197	-.281	.168	19.245	.944	20.380	.000***
	정보선호	-.057	-.293	-.968	.124	.442	69.267	.792	87.508	.000***
군집별 케이스 수	202	110	193	401	409	df = 1310				
다중공선성 검증을 위한 각 군집별 구성 변인에 대한 상관관계 분석	.347~	.275~	.386~	.330~	.247~	p < .01에서 모두 유의함				
	.713	.689	.680	.670	.635					

***p < .001, **p < .01

인공신경망 모형

인공신경망 모형의 적용에는 다음과 같은 단계로 진행되었다. 첫째, 알고리즘은 예측을 위한 수식을 적용하였다. 둘째, 본 연구에서는 선행연구의 방법론(허명희와 이용구, 2008)을 참조하여 파라미터 추정을 훈련 50% 대 검증 50%의 비율로 구성하여 연구를 진행하였다. 셋째, 훈련방법은 Quick방법인 시그모이드(sigmoid) 함수를 사용하였으며, 가중치는 .9로 지정하였다. 넷째, 학습률 에타(eta)는 적응해야 할 방향을 찾고 인공신경망 모형이 반복되어 목표변수를 찾는 과정에서 모형에 수정되는 가중치를 조절하는 역할을 하는데, 본 연구는 가장 일반적으로 사용하는 에타 값인 .3으로 고정하여 연구를 진행하였다. 다섯째, 결과를 통한 은닉층의 뉴런 수 결정이다. 은닉층은 노드 수는 각 군집별 그리고 노드 수에 따라 결과가 다르기 때문에 본 연구에서는 은닉층의 노드 수를 1개, 2개, 3개, 4개, 8개, 16개, 32개 등 다양하게 적용하여 결과를 비교하였다. 그 결과 가장 적합한 은닉층의 수는 3개로 파악되어 모든 군집에 대하여 최종 3개의 은닉층으로 지정하여 연구를 진행하였다.

이와 같은 다섯 단계를 적용하여 각각의 군집별 인공신경망 모형을 분석하였으며, 분석결과 관전만족에 대한 군집별 분류적중률은 군집1(66.7%), 군집2(77.2%), 군집3(71.0%), 군집4(61.1%), 군집5(66.9%)로 나타나 군집2가 가장 높게 나타났다. 다음으로 미래 소비행동에 대한 군집별 분류적중률은 군집1(77.2%), 군집2(80.8%), 군집3(70.5%), 군집4(61.7%), 군집5(69.1%)로 나타나 군집2가 가장 높게 나타났다. 그러나 이 분류적중률은 관전만족과 미래 소비행동에 대한 낮은 가능성과 높은 가능성을 모두 포함하고 있고 훈련 표본 및 검증 표본에 대한 전체 평균 분류적중률을 나타내는데 그치고 있다. 그러므로 본 연구의 목표변수인 관전만족과 미래 소비행동을 앞으로도 지속할 가능성이 가장 높은 집단을 찾기 위하여 훈련 표본 및 검증 표본을 검토한 결과 관전만족과 미래 소비행동이 가장 높은 가능성을 가진 집단을 파악하였다. 분석 결과 군집5(관전만족이 높은 가능성 71.3%, 미래 소비행동 높은 가능성은 99.3%)가 목표변수인 관전만족과 미래소비행동이 가장 높은 가능성을 가진 집단임을 밝혀내었다(표 5).

표 5. 인공신경망 모형을 통한 군집별 관전만족 및 미래 소비행동에 대한 예측 확률(분류 정확도)분석

구분		군집1		군집2		군집3		군집4		군집5	
		관전 만족	소비 행동	관전 만족	소비 행동	관전 만족	소비 행동	관전 만족	소비 행동	관전 만족	소비 행동
훈련 (50%)	낮은 가능성	90.4%	86.2%	100.0%	93.5%	87.5%	65.0%	80.0%	82.1%	60.7%	1.1%
	높은 가능성	45.2%	71.1%	27.8%	75.0%	53.7%	76.0%	30.1%	35.7%	70.4%	99.3%
	전체	70.2%	79.6%	75.9%	88.7%	73.2%	71.1%	59.1%	60.3%	65.5%	67.0%
검증 (50%)	낮은 가능성	76.5%	76.1%	100.0%	90.6%	80.7%	61.8%	90.0%	80.3%	64.8%	3.4%
	높은 가능성	40.0%	73.6%	6.3%	37.5%	51.3%	79.2%	31.2%	37.8%	71.3%	99.3%
	전체	63.0%	74.7%	73.2%	72.9%	68.8%	69.9%	63.1%	63.1%	68.3%	71.1%
평균 분류적중률		66.6%	77.2%	74.6%	80.8%	71.0%	70.5%	61.1%	61.7%	66.9%	69.1%

로지스틱 회귀모형 적용

인공신경망 모형과 함께 로지스틱 회귀분석을 실시하여 프로야구 관전자의 각 군집별 관전만족과 미래 소비행동에 대한 분류적중률을 분석하였다. 본 연구에서는

관전만족 및 미래 소비행동(낮은 집단=0, 높은 집단=1)을 기준으로 이분형 변수(binary variable)로 설정하였기 때문에 일반적인 회귀분석에서와 같이 정규분포를 따르는 것이 아니라 이항분포를 따르게 된다. 이러한 로지스틱 회귀분석은 인공신경망 모형과 마찬가지로

관전만족 및 미래 소비행동이 부정적인지, 긍정적인지를 직접 예측하는 것이 아니라 낮은 집단과 높은 집단에 따라 얼마나 정확하게 예측되는지의 확률을 말한다. 로지스틱 회귀분석의 결과는 -2 Log 우도를 검증(낮을수록 좋음), Cox and Shell R²(0에 가까울수록 좋음), 표준오차(낮을수록 좋음), Homer and Lemeshow 검정(유의하지 않아야 좋은 모델)을 통하여 적합성을 평가하였고 최종 분류적중률을 분석하였다.

분석결과 관전만족에 대한 군집별 분류적중률은 군집1(57.4%), 군집2(70.9%), 군집3(62.7%), 군집4(63.6%), 군집5(68.9%)로 나타나 로지스틱 회귀분석에서도 군집2가 가장 높게 나타났다. 다음으로 미래 소비행동에 대한 군집별 분류적중률은 군집1(63.4%), 군집2(76.4%), 군집3(68.4%), 군집4(64.3%), 군집

5(68.9%)로 나타나 군집2가 가장 높게 나타났다. 그러나 로지스틱 회귀분석 결과 역시 인공신경망 모형과 마찬가지로 관전만족과 미래 소비행동이 낮은 가능성과 높은 가능성에 대한 전체 분류적중률의 결과이므로 본 연구의 목표변수인 관전만족과 미래 소비행동에 대한 높은 가능성을 가진 군집을 분석하였다. 분석결과 군집5(관전만족 높은 가능성 71.8%, 미래 소비행동 높은 가능성 96.5%)가 가장 적절한 것으로 나타났다. 따라서 기계학습에 의한 예측 기법인 인공신경망 모형과 통계기반의 예측기법인 로지스틱 회귀분석 결과 모두 군집5가 관전만족 및 미래 소비행동이 가장 높은 가능성을 가진 집단으로 분석되었다(표 6).

표 6. 로지스틱 회귀모형을 통한 군집별 관전만족 및 미래 소비행동에 대한 예측 확률(분류 정확도)분석

구분		군집1		군집2		군집3		군집4		군집5	
검증방법	기준치	만족	행동	만족	행동	만족	행동	만족	행동	만족	행동
-2 log 우도	낮을수록 좋음	265.24	253.16	131.30	120.18	245.56	238.49	529.70	504.96	521.42	485.56
Cox & Shell R ²	0에 가까워야 좋음	.037	.124	.042	.107	.081	.140	.049	.104	.105	.051
표준오차	낮을수록 신뢰함	.143	.141	.119	.210	.146	.144	.101	.101	.099	.107
Homer & Lemeshow 검정	p>.05(유의하지 않을수록 좋음)	p=.360	p=.003	p=.467	p=.005	p=.059	p=.419	p=.088	p=.082	p=.066	p=.071
각 군집별 관전만족 및 미래 소비행동에 대한 예측(분류적중률)	낮은 가능성	86.7%	69.2%	97.4%	97.4%	80.5%	69.5%	82.2%	78.4%	61.5%	7.9%
	높은 가능성	14.6%	57.1%	11.8%	25.0%	37.5%	67.3%	39.8%	46.0%	71.8%	96.5%
	전체	57.4%	63.4%	70.9%	76.4%	62.7%	68.4%	63.6%	64.3%	66.7%	68.9%

군집5의 특성 파악

관전만족과 미래 소비행동이 가장 높은 가능성을 가진 군집5의 특성을 분석하고자 다른 군집과의 교차분석(X² 검정)과 일원변량을 분석을 실시하여 군집의 특성을 파악하였다. 이를 자세히 살펴보면 성별, 결혼 등 모든 부분에서 유의한 차이가 나타났다(p<.001). 따라서 이러한 특성을 토대로 군집5에 일반적인특성을 분석한

결과 남자가 79.5%, 미혼이 98.5%, 20대가 79.2%로 나타났다. 다음으로 관전 동반자로는 78.3%가 친구 또는 연인, 전체 44.0%가 경기장까지의 소요시간이 30~60분으로 나타났다. 또한 전체 구성 비율 중 가장 많은 41.3%가 년 간 경기장에서 관전하는 횟수가 9회 이상으로 나타났고 95.6%가 TV로 9회 이상 프로야구를 시청하는 것으로 나타났다(표 7).

표 7. 군집별 인구통계학적 특성에 따른 교차분석 결과

구분	군집1	군집2	군집3	군집4	전체	χ^2	***p<.001	
	(명/ 집단 중%)	(명/집단 중%)	(명/집단 중%)	(명/집단 중%)	(명/집단 중%)			
성별	남	130(64.4%)	86(78.2%)	155(80.3%)	55(13.7%)	325(79.5%)	458.386	.000***
	여	72(12.8%)	24(21.8%)	38(19.7%)	346(86.3%)	84(20.5%)		
결혼	결혼	199(98.5%)	110(100%)	6(3.1%)	6(1.5%)	6(1.5%)	1204.789	.000***
	미혼	3(1.5%)	0	187(96.9%)	395(98.5%)	403(98.5%)		
연령	20대	8(4.0%)	4(0.5%)	143(74.1%)	349(87.0%)	326(79.7%)	827.780	.000***
	30대	99(49.0%)	36(32.7%)	48(24.9%)	52(13.0%)	79(19.3%)		
	40대 이상	95(47.0%)	70(63.6%)	2(1.0%)	0	4(1.0%)		
학력	고졸 이하	2(1.0%)	104(94.5%)	18(9.3%)	50(12.5%)	26(6.4%)	812.738	.000***
	대학 재학	2(1.0%)	2(1.8%)	102(52.8%)	136(33.9%)	178(43.5%)		
	대학교 졸업	162(80.2%)	4(3.6%)	71(36.8%)	203(50.6%)	169(41.3%)		
	대학원 이상	36(17.8%)	0	2(3.0%)	12(3.0%)	36(8.8%)		
직업	학생	0	2(1.8%)	100(51.8%)	158(39.4%)	194(47.4%)	441.535	.000***
	회사원	78(38.6%)	50(45.5%)	66(34.2%)	118(29.4%)	142(34.7%)		
	공무원	28(13.9%)	2(1.8%)	6(3.1%)	18(4.5%)	16(3.9%)		
	서비스업	22(10.9%)	26(23.6%)	8(4.1%)	34(8.5%)	20(4.9%)		
	자유업	28(13.9%)	12(10.9%)	6(3.1%)	18(4.5%)	18(4.4%)		
	가정주부	34(16.8%)	16(14.5%)	0	2(.5%)	0		
월 가계 소득	기타	12(5.9%)	2(1.8%)	7(3.6%)	53(13.2%)	19(4.6%)	285.869	.000***
	100만원대	0	4(3.6%)	44(22.8%)	44(11.0%)	20(4.9%)		
	200만원대	17(8.4%)	24(21.8%)	47(24.4%)	167(41.6%)	70(17.1%)		
	300만원대	28(13.9%)	38(34.5%)	38(19.7%)	76(19.0%)	92(22.5%)		
	400만원대	51(25.2%)	18(16.4%)	22(11.4%)	40(10.0%)	51(12.5%)		
	500만원대	42(20.8%)	20(18.2%)	14(7.3%)	36(9.0%)	58(14.2%)		
관전 동반자	600만원 이상	64(31.7%)	6(5.5%)	28(14.5%)	38(9.5%)	118(28.9%)	561.427	.000***
	가족	156(77.2%)	72(65.5%)	14(7.3%)	56(14.0%)	38(9.3%)		
	친구	20(9.9%)	16(14.5%)	95(49.2%)	159(39.7%)	202(49.4%)		
	연인	7(3.5%)	0	52(26.9%)	152(37.9%)	118(28.9%)		
	동료 및 선후배	13(6.4%)	16(14.5%)	26(13.5%)	24(6.0%)	44(10.8%)		
경기장까지 소요시간	기타	6(3.0%)	6(5.5%)	6(3.1%)	10(2.5%)	7(1.7%)	57.610	.000***
	30분 미만	61(30.2%)	38(34.5%)	30(15.5%)	92(22.9%)	72(17.6%)		
	30분 이상-60분 미만	93(46.0%)	38(34.5%)	71(36.8%)	185(46.1%)	180(44.0%)		
	60분 이상-90분 미만	34(16.8%)	16(14.5%)	54(28.0%)	92(22.9%)	98(24.0%)		
경기장에서 개인당 지출한 비용 (입장권 제외)	90분 이상	14(6.9%)	18(16.4%)	38(19.7%)	32(8.0%)	59(14.4%)	196.917	.000***
	1만원 이하	20(9.9%)	18(16.4%)	26(13.5%)	58(14.5%)	14(3.4%)		
	1만원-1만5천원 미만	24(11.9%)	22(20.0%)	65(33.7%)	118(29.4%)	74(18.1%)		
	1만5천원-2만원 미만	36(17.8%)	18(16.4%)	46(23.8%)	109(27.2%)	106(25.9%)		
	2만원-2만5천원 미만	18(8.9%)	14(12.7%)	32(16.6%)	50(12.5%)	41(10.0%)		
프로야구 관전 횟수(1년)	2만5천원-3만원 미만	22(10.9%)	18(16.4%)	12(6.2%)	25(6.2%)	41(10.0%)	254.044	.000***
	3만원-3만5천원 미만	41(20.3%)	6(5.5%)	6(3.1%)	19(4.7%)	53(13.0%)		
	3만5천원 이상	41(20.3%)	14(12.7%)	6(3.1%)	22(5.5%)	80(19.6%)		
	1회(처음 방문)	18(8.9%)	10(9.1%)	10(5.2%)	108(26.9%)	8(2.0%)		
프로야구 TV시청 횟수(1년)	2회-5회 이하	108(53.5%)	72(65.5%)	93(48.2%)	211(52.6%)	151(36.9%)	484.211	.000***
	6회-9회 이하	38(18.8%)	18(16.4%)	40(20.7%)	40(10.0%)	81(19.8%)		
	9회 이상	38(18.8%)	10(9.1%)	50(25.9%)	42(10.5%)	169(41.3%)		
	1회 이하	2(1.0%)	2(1.8%)	0	63(15.7%)	0		
프로야구 TV시청 횟수(1년)	2회-5회 이하	9(4.5%)	10(9.1%)	4(2.1%)	133(33.2%)	0	484.211	.000***
	6회-9회 이하	14(6.9%)	14(12.7%)	16(8.3%)	53(13.2%)	18(4.4%)		
	9회 이상	177(87.6%)	84(76.4%)	173(89.6%)	152(37.9%)	391(95.6%)		

다음으로 각 군집에 따른 라이프스타일에 대한 차이를 파악하고자 일원변량분석을 실시하였으며, 유의한 차이가 파악된 경우($p < .05$), Scheffe의 사후검증을 실시하였다. <표 8>은 5개 군집에 따른 각 변인에 대한 평균 및 표준편차를 나타내고 있다. 군집5의 특성을 살펴보면 타인 지향적 소비 및 정보선호가 다른 군집보다 유의하게 높은 것으로 나타났으며, 다른 군집과의 비교 결과와 인구통계학적 특성 및 라이프스타일 변인에 대한 특성 분석을 통해 '20대 미혼 남성 타인 지향적 소비를 가진 정보선호 집단'으로 명명하였다. 뿐만 아니라 각 군집별 특성을 요약한 결과 군집1은 대졸 기혼 남자 집단으

로 성취주의가 가장 높은 집단으로 나타나 '성취주의 기혼 남성 집단'으로 명명 하였다. 군집2의 경우 남자, 기혼, 고졸 40대 이상으로 나타났으며, 라이프스타일의 경우 전반적으로 낮게 나타나 '40대 기혼 남성 집단'으로 명명하였다. 군집3의 경우 미혼 남성 학생 집단으로 유행추구가 가장 낮은 집단으로 나타나 '비유행추구 남학생 집단'으로 명명하였다. 군집4의 경우 20대 미혼 여성으로 자기 지향적 소비가 가장 높게 나타나 '20대 미혼 여성 자기 지향적 소비 집단'으로 명명하였다. 이상과 같이 각 군집별 특성을 고려하여 군집을 명명하였다(<표 9>).

표 8. 군집에 따른 차이검증

변인	구분	N	M	SD	SS	df	MS	F	p	Scheffe의 사후검증
자기 지향적 소비	군집1	202	3.10	.719	40.037	4	10.009	20.592	.000***	IV, V > I, II III > II
	군집2	110	2.85	.803	636.738	1310	.486			
	군집3	193	3.27	.698	676.775	1314				
	군집4	401	3.41	.686						
	군집5	409	3.41	.666						
	합계	1315	3.29	.718						
유행추구	군집1	202	3.09	.733	66.066	4	16.516	33.324	.000***	IV, V > I, II, III
	군집2	110	2.85	.651	649.279	1310	.496			
	군집3	193	2.71	.750	715.345	1314				
	군집4	401	3.28	.639						
	군집5	409	3.32	.740						
	합계	1315	3.14	.738						
타인 지향적 소비	군집1	202	2.95	.676	38.975	4	9.744	19.550	.000***	V > I, II, III, IV
	군집2	110	2.80	.789	652.919	1310	.498			
	군집3	193	2.82	.701	691.894	1314				
	군집4	401	3.00	.730						
	군집5	409	3.27	.675						
	합계	1315	3.03	.726						
행복추구	군집1	202	3.91	.617	5.829	4	1.457	2.988	.018*	V > II
	군집2	110	3.70	.636	638.943	1310	.488			
	군집3	193	3.84	.772	644.772	1314				
	군집4	401	3.89	.644						
	군집5	409	3.94	.764						
	합계	1315	3.89	.700						
성취주의	군집1	202	3.70	.515	27.429	4	6.857	19.608	.000***	I > II, III, IV V > III, IV
	군집2	110	3.47	.634	458.121	1310	.350			
	군집3	193	3.35	.605	485.549	1314				
	군집4	401	3.33	.614						
	군집5	409	3.58	.585						
	합계	1315	3.48	.608						
정보선호	군집1	202	3.26	.735	197.941	4	49.485	103.731	.000***	V > I, IV > II, III
	군집2	110	2.97	.931	624.940	1310	.477			
	군집3	193	2.52	.732	822.881	1314				
	군집4	401	3.42	.637						
	군집5	409	3.69	.619						
	합계	1315	3.31	.791						

*** $p < .001$, * $p < .05$

I : 군집1 , II : 군집2 , III : 군집3 , IV : 군집4, V : 군집5

표 9. 군집에 대한 특성 비교 및 요약

구분	군집1	군집2	군집3	군집4	군집5
명명	성취주의 기혼 남성 집단	40대 기혼 남성 집단	비유행추구 남학생 집단	20대 미혼 여성 자기 지향적 소비 집단	20대 미혼 남성 타인 지향적 소비를 가진 정보선호 집단
성별 및 결혼	64.4%가 남자 98.5%가 결혼	78.2%가 남자 100%가 결혼	80.3%가 남자 96.9%가 미혼	86.3%가 여자 98.5%가 미혼	79.5%가 남자 98.5%가 미혼
연령	집단 내에서 명확하게 구분되지 않음	63.6%가 40대 이상	74.1%가 20대	87%가 20대	79.7%가 20대
학력	80.2%가 대졸	94.5%가 고졸 이하	52.8%가 대학재학	50.6%가 대졸	집단 내에서 명확하게 구분되지 않음
직업	집단 내에서 명확하게 구분되지 않음	집단 내에서 명확하게 구분되지 않음	51.8%가 학생	집단 내에서 명확하게 구분되지 않음	집단 내에서 명확하게 구분되지 않음
관전 동반자	77.2%가 가족	65.5%가 가족	49.2%가 친구	77.6%가 친구 또는 연인	78.3%가 친구 또는 연인
경기장 소요시간	76.2%가 30분 미만 또는 30~60분 미만	69%가 30분 미만 또는 30~60분 미만	64.8%가 30~60분 또는 60~90분 미만	46.1%가 30~60분 미만	44.0%가 30~60분 미만
관전횟수	53.5%가 2회~5회 이하	65.5%가 2회~5회 이하	집단 내에서 명확하게 구분되지 않음	52.6%가 2회~5회 이하	41.3%가 9회 이상
TV시청 횟수	87.6%가 9회 이상	76.4%가 9회 이상	89.6%가 9회 이상	집단 내에서 명확하게 구분되지 않음	95.6%가 9회 이상
라이스타일 특성	성취주의가 가장 높고 행복추구가 비교적 높은 집단	자기 지향적, 타인 지향적 및 행복추구가 가장 낮은 집단	유행추구가 가장 낮은 집단	자기 지향적이 가장 높은 집단	타인 지향적 소비 및 정보선호가 가장 높은 집단

논의

이정학과 정승훈(2012)은 프로스포츠 관련 연구에서 시장세분화와 인공신경망 모형을 병행하여 연구를 진행하는 것이 각 군집의 특성을 구체적으로 파악할 수 있으며, 목표변수에 대한 보다 정확한 분류정확도를 얻을 수 있을 것이라고 하여 인공신경망을 이용하여 스포츠 관전자를 분석하였다. 뿐만 아니라 다수의 연구를 통해 보다 구체적으로 소비자를 분석하기 위해서는 먼저 소비자를 분류하고 각각의 분류된 집단 중 특정변수에 대한 예측률이 높은 집단을 파악하는 것이 소비자행동을 예측하는 가장 좋은 방법이라고 하였다(김영신, 2003; 이정학과 정승훈, 2012; 이창환 등, 2009; 이호영 등, 2000; Balakrishnan et al., 1996; Berry & Linoff, 1997). 따라서 이 연구는 선행연구의 연구과정과 결과 등(안광호 등, 2009; 이정학 등, 2010)을 검토하여 프로야구 관전자를 분류하였고 이를 인공신경망 모형과 로지스틱 회귀모형에 적용하여(허명희와 이용구, 2008) 다음과 같은 결과를 얻었다.

첫째, 각 군집의 분류 결과가 통계적으로 유의하게 나타났으며, 각각의 특성이 잘 나타난 것으로 파악되어 세분화된 집단을 인공신경망 모형과 로지스틱 회귀모형에 적용하여 목표변수인 관전만족 및 미래소비행동에 대한 분류적중률을 비교한 결과 전반적으로 인공신경망 모형이 높게 나타났다. 이와 같은 결과는 송수섭과 이의훈(2001)의 연구에서 인공신경망 모형이 로지스틱 회귀모형 및 판별분석 보다 분류적중률이 높게 나타난 결과와 일치하고 있으며, 인공신경망 모형은 전통적 통계 모형 보다 우수한 성과와 이론적 타당성을 제공하고 있다고 한 그들의 의견을 지지하고 있다. 뿐만 아니라 Fish et al.(1995)은 산업 마케팅에서 대상이 되는 시장을 두 그룹 및 세 그룹으로 분류하는 문제에 대한 연구에서 기존 통계 방법인 판별분석, 로지스틱 회귀분석, 인공신경망을 이용하여 분류적중률을 비교 분석하였다. 이 분석은 사실적 데이터와 인구통계학적 데이터를 입력 패턴으로 사용하여 인공신경망, 로지스틱 회귀분석, 판별분석을 비교하였다. 결과적으로 인공신경망 74%, 로지스틱 회귀분석 72%, 판별분석 68%의 결과를 얻어 인공

신경망의 분류적중률이 가장 우수한 것으로 나타나 본 연구의 결과와 일치하는 것으로 나타났다.

Dasgupta et al.(1994)은 설문조사에 의하여 수집된 개인 고객의 라이프스타일 변수를 사용하여 고객을 금융적 위험에 대응하는 태도에 따라 고위험(high risk) 그룹과 저위험(low risk) 그룹으로 나누어 두 그룹의 분류적중률을 분석하였다. 결과를 살펴보면 인공신경망 58.9%, 판별분석 55.5%, 로지스틱 회귀분석 55.2%로 나타나 모든 모델에서 전반적으로 저조한 적중률을 보였지만, 결론적으로 인공신경망이 다른 분석 방법에 비하여 가장 우수한 것으로 나타나 본 연구의 결과를 지지하고 있는 것으로 나타났다.

뿐만 아니라 West et al.(1997)은 성격이 유사한 미국의 양판점인 Kmart, Montgomery Ward 및 Sears Reback에 대하여 점포 선택에 대한 설문데이터 및 점포별 특성 데이터를 입력변수로 하여 분류적중률을 비교하였는데, 그 결과 인공신경망 79.3%, 판별분석 70.38%, 로지스틱 회귀분석 70.1%로 나타나 인공신경망 모형이 다른 두 모형에 비하여 훨씬 우수한 것으로 나타났다. 따라서 앞서 언급한 선행연구들과 더불어 다양한 연구를 통해 소비자행동 예측에 있어 기존의 통계적 방법보다 인공신경망 모형이 우수하다는 것이 검증되었다(김동성, 2004; 안진균, 2002; 이정학과 정승훈, 2012; 전성무, 2001; 최형준과 김주학, 2006; Coakley & Brown, 1993; Dasgupta et al., 1994; Fish et al., 1995; Galindo, 1998; Lajbcygier et al., 1995; Raghupathi, et al., 1993; Sharda, 1994; Treigueiros & Berry, 1991; West et al., 1997). 따라서 이 연구를 통해 프로야구 관전자의 특정 부분(본 연구에서 관전만족 및 미래 소비행동이 가장 높은 집단)에 대하여 보다 구체적인 예측과 다양한 분석 방법이 적용될 있으며, 이상과 같은 방법을 통해 앞으로 소비자행동을 예측하는데 있어 기초 자료로 활용될 수 있을 것으로 판단된다.

둘째, 세분화된 집단을 모두 인공신경망 모형과 로지스틱 회귀모형에 적용하여 목표변수인 관전만족과 미래 소비행동에 가장 높은 분류적중률을 보인 군집을 분석한 결과 군집5로 나타났으며, 해당 군집에 대한 특성을 파악하고자 교차분석과 일원변량분석을 실시하여 '20대

미혼 남성 타인 지향적 소비를 가진 정보선호 집단'으로 명명하였다. 이와 같은 결과는 프로야구 관전자의 시장 세분화와 관련된 장현철(1999)의 연구를 지지하고 있는 것으로 나타났는데, 그는 프로야구 관전자의 성별 및 연령에 따른 시장세분화는 장기적인 관점으로 볼 때, 스포츠 산업의 성장에 있어서 커다란 효과가 있다고 제시하여 성별과 연령이 모두 포함된 시장세분화의 필요성을 강조하였다.

또한 이정학 등(2013)은 월드컵 거리응원 단체관람 참여자를 대상으로 일반적인특성, 스포츠 가치관, 소비문화 및 스포츠태도를 기준변수로 4개의 군집으로 시장을 세분화하였는데, 분석결과 특정 군집의 경우 모든 변인들이 전반적으로 다른 군집에 비해 높게 나타났으며, 각 군집별 구분된 특성의 차이로 인하여 각각 군집의 특성을 파악할 수 있다고 하였다. 따라서 본 연구 역시 군집5의 명확한 특성을 기준으로 군집의 특징에 맞는 마케팅 전략을 실행 할 수 있다. 예를 들어, 이 군집은 경기장까지의 소요시간이 짧고 관전 횟수나 TV시청 횟수가 높으며, 타인 지향적 소비의 특성인 소유품을 통한 사회적 배경, 브랜드에 민감, 속하고 싶은 부류와 비슷한 소비 그리고 가격을 비교하고 주변사람에게 정보를 제공하는 특성을 가지고 있는 것으로 나타났다. 이에 따라 프로야구 홈 팀 경기장 주변의 남성 브랜드 매장과 연계하여 특정 제품 구매시 티켓 할인, 동반자와 함께 방문할 경우 식·음료 쿠폰 제공, 프로야구단 소식 등을 소셜네트워크 서비스(SNS)로 다른 사람에게 전달할 경우 다양한 혜택들을 제공(구단 락커룸 방문, 선수들과 만남, 구단의 공식적인 행사 및 비공식적인 행사 초대)하고 이를 통해 지속적으로 관전만족과 미래 소비행동을 높게 유지하도록 한다면 앞으로도 이 집단은 충성도 높은 관전자로 유지할 수 있을 것이다. 즉, 이 집단을 이용한 직·간접적인 마케팅을 통해 이 군집이 프로야구단에 대한 적극적인 홍보를 간접적으로 유발 하도록 유도 한다면 이를 통한 구단의 가치 창출 및 홍보(동반자를 통한 관전자 증가, SNS 등을 통한 구단 제품 또는 팀에 대한 간접적인 홍보 효과 증가 등)는 더욱더 효과가 있을 것이며, 이는 곧 프로야구 관전자의 증가와 더불어 각 구단의 이익 창출에 이바지 할 것이다.

또한 군집1에서 군집4는 인공신경망과 로지스틱회귀

모형을 통해 비록 관전만족과 미래 소비행동에 대한 분류적중률이 군집5에 비해 낮게 나타났지만 각 집단의 특성에 맞는 마케팅을 통해 프로야구 관전자의 특성을 다음과 같이 분석할 수 있다. 군집1(성취주의 기혼 남성 집단)의 경우 성취주의와 행복추구는 높으나 직접 관전 횟수가 비교적 작게 나타났으므로 경기장 방문을 통한 개인과 구단의 성취와 행복을 느낄 수 있도록 구단의 CSR활동 즉, 기업의 사회적 활동(NC 다이노스 구단의 연탄 봉사 등) 등을 경기장 전광판이나 티켓 등에 제시하여 구단과 함께 직·간접적으로 성취주의와 행복추구가 높아지도록 한다면 경기를 통한 행복추구 등이 높아질 것이다. 다음 군집2(40대 기혼 남성 집단)의 경우 라이프스타일 요인에서는 전반적으로 다른 집단에 비해 뚜렷한 특성이 나타나지는 않았으나 인구통계학적 특성에서는 40대 이상 기혼으로 나타났으며, 관전 횟수와 TV 시청 횟수 모두 작게 나타나 앞으로 성장 가능성이 가장 높은 집단으로 볼 수 있다. 따라서 이 집단의 경우에는 부부 동반시 할인 혜택 제공, 가족 동반시 추가 할인 혜택 등과 같은 주변사람과 함께 즐길 수 있는 마케팅을 이용하여 가족 및 주변사람과 함께 경기를 관전할 수 있는 방안이 제공된다면 경기장을 찾는 횟수는 보다 증가할 것이다. 군집3(비유행추구 남학생 집단)의 경우 대부분이 학생으로 구성되어 있고 약 50%가 친구와 함께 경기장을 방문하는 것으로 나타나 친구 패키지(친구와 함께 시즌권 구매시 양쪽 모두 할인, 친구 수에 따른 할인을 제공 등) 등과 같은 다양하고 실현 가능한 마케팅을 이용하여 친구와 함께 경기장을 방문할 수 있는 혜택을 제공해야 한다. 군집4(20대 미혼 여성 자기 지향적 소비 집단)의 경우 연인과 경기장을 방문하는 경우가 높고 경기장에 비교적 가까운 곳에 거주하는 집단으로 나타났으므로 프로야구 경기를 연인과의 데이트에 이용할 수 있는 방안이 마련되어야 한다. 예를 들면 여성 시즌권 구입 고객에게는 남성 유니폼 등을 선물로 제공하여 여성 관전자를 통해 남성 관전자를 유입시키도록 하는 전략이 필요하다. 또한 이 여성 집단은 자기 지향적 소비가 높으므로 여성 관전자 개인의 개성과 이미지를 노출시킬 수 있는 유니폼 등을 개발하고 온라인을 통하여 여성 관전자가 직접 제품을 디자인 할 수 있게 하는 등 차별화된 서비스를 제공하여 개인의 이미지를 표출하도록 장소와

제품을 제공한다면 이를 통한 직접 관전은 높아질 것이다. 이와 같이 세분화된 관전자의 결과를 통하여 프로야구 관전자의 특성을 이해할 수 있으며, 이에 따라 각 집단에 맞는 다양한 대응 방안이 요구된다. 따라서 다양한 방법을 이용하여 프로스포츠 관전자의 특성을 이해하고 분석하는 것은 프로야구 구단의 마케팅 전략을 집중할 수 있는 계기가 되며, 더 나아가 프로스포츠 관중 증가를 위한 발판이 된다고 할 수 있다.

결론

이 연구는 첫째, 다양한 연구 분야에서 시장세분화의 기준 변수로 측정된 프로야구 관전자의 인구통계학적 특성과 라이프스타일을 측정하여 프로야구 관전자를 세분화하는데 있다. 둘째, 각 세분화된 집단을 인공신경망 및 로지스틱 회귀분석 모델에 적용시켜 목표변수(관전만족과 미래 소비행동)에 가장 높은 가능성(분류적중률)을 가진 프로야구 관전자 집단을 파악하는데 있다. 셋째, 목표변수가 가장 높은 집단을 및 세분화된 집단의 특성을 분석하여 그에 따른 마케팅 전략을 제시하는데 그 목적이 있다. 결과와 논의를 통해 다음과 같은 결론을 도출하였다.

첫째, 프로야구 관전자를 세분화 한 결과 5개의 군집으로 분류하였다. 둘째, 인공신경망 모형을 이용하여 관전만족과 미래 소비행동이 가장 높은 가능성을 가진 집단을 파악한 결과는 군집5로 나타났다(관전만족: 71.3%, 미래 소비행동: 99.3%). 셋째, 로지스틱 회귀모형을 이용하여 관전만족과 미래 소비행동이 가장 높은 가능성을 가진 집단을 파악한 결과 군집5(관전만족: 71.8%, 미래 소비행동: 96.5%)가 가장 높게 나타났다. 넷째, 인공신경망 모형과 로지스틱 회귀모형을 비교하여 관전만족과 미래소비행동이 가장 높은 분류적중률을 가진 집단을 분석한 결과 '20대 미혼 남성 타인 지향적 소비를 가진 정보선호 집단'으로 나타났다. 따라서 이 집단을 유지 및 관리하기 위해서는 프로야구 구단별 홈팀 경기장 주변의 남성 브랜드 매장과 연계한 마케팅을 통하여 특정 제품 구매시 티켓 할인 등과 같은 제휴 마케팅을 활용하고 더 나아가 구단의 소셜네트워크 서비스

(SNS) 등을 이용하여 이 집단에 포함된 관전자가 타인에게 다양한 정보를 제공할 수 있도록 다양한 혜택을 제공한다면 이 집단의 관전만족과 미래 소비행동은 앞으로 높게 유지될 수 있을 것이다. 또한 이러한 관전자와 소비자, 관전자와 관전자의 정보 전달 방법을 통하여 직·간접적인 구전 마케팅이 이루어질 수 있도록 노력을 기울인다면 구단의 관중 증가와 마케팅 전략에 도움이 될 수 있을 것이다.

높아지는 야구의 열기 속에서 프로야구 관전자의 인구통계학적 특성과 라이프스타일 등을 기준변수로 관전자를 분석하는 것은 각 프로야구단의 마케팅 전략을 위한 기초자료가 되며, 이러한 집단 중 각 구단의 관심사인 관전 만족과 미래 소비행동이 가장 높은 집단을 예측하여 관리 및 유지하는 것은 향후 프로야구 구단의 관중 증가에 직·간접적 도움이 될 수 있다. 따라서 본 연구는 먼저 관전자를 분류하고 세분화된 각 군집을 인공지능망 모형과 로지스틱 회귀모형에 적용하여 관전만족과 미래 소비행동이 가장 높은 가능성이 있는 집단을 파악하였다. 이러한 과정을 통하여 인공지능망 모형의 우수성과 예측 가능성이 높은 집단을 분석하였으며, 이에 따라 지금까지 스포츠 분야에서 단순히 승·패 예측을 위한 분석방법으로 사용되는데 그쳤던 인공지능망 모형을 이용해 예측이 힘든 프로야구 관전자의 소비자행동을 직·간접적으로 유추할 수 있는 기초자료를 제공하였다. 뿐만 아니라 스포츠마케팅 분야에서 인공지능망 모형과 로지스틱 회귀모형과 같은 다양한 통계방법의 적용과 비교를 통해 후속연구를 위한 발판을 제공하였으며, 더 나아가 이 연구의 결과를 통해 지금까지 진행되어 온 시장세분화 연구의 지속성과 한계를 보완할 수 있는 대안적 기초자료를 마련하였다.

그러나 본 연구 역시 연구를 수행하는데 있어 다음과 같은 한계점을 발견하였으며, 후속연구를 위해 다음과 같은 제안을 하고자 한다. 첫째, 인공지능망 모형과 로지스틱 회귀모형에 대한 스포츠분야의 선행연구 부족으로 인하여 충분한 논의가 이루어지지 않은 한계점과 라이프스타일 이외에 다양한 변인을 연구에 포함시키지 못한 제한점이 노출되었다. 따라서 후속연구에서는 프로야구 관전자의 다양한 변인(관전자의 가치 및 동기 등과 같은 관전을 유발시키는 선행변수)을 포함시켜 연구를

진행한다면 보다 다양하고 세분화된 특성을 분석할 수 있을 것이다.

둘째, 연구 대상을 선정하는데 있어 프로야구 5개의 홈팀 관전자로 한정하여 연구가 이루어졌기 때문에 이를 전체 프로야구 관전자로 일반화하기에는 다소 무리가 따른다. 따라서 후속연구에서는 보다 광범위한 데이터를 수집하여 프로야구 관전자의 소비자행동을 예측한다면 더욱 세분화되고 구체적인 집단을 파악할 수 있을 것이다. 다시 말해 시장세분화(군집분석)의 장점은 광범위한 데이터를 축소할 수 있고 공통된 특성을 기준으로 시장을 세분화 할 수 있기 때문에 이러한 분석방법을 이용하여 단순히 연구자가 직접 수집한 데이터가 아니라 프로야구 각 구단에서 티켓 구매나 제품 구매시 파악할 수 있는 광범위한 정보와 데이터를 활용하여 후속 연구가 이루어진다면 보다 깊이 있는 연구 결과를 얻을 수 있을 것이다.

참고문헌

- 강현철, 한상태, 최종후, 김은석, 김미경(2001). (SAS Enterprise Miner 4.0을 이용한)데이터마이닝: 방법론 및 활용. 서울: 자유아카데미.
- 고재근, 나상준(2001). 지역별에 따른 종목별 스포츠 참여인구수 예측. 한국체육학회지, 40(4), 371-383.
- 김계수(2007). New AMOS 7.0 구조방정식 모형분석. 서울: 한나래 출판사.
- 김동성(2004). 인공지능망을 이용한 무선인터넷 서비스사용자 예측에 관한 연구. 미간행 석사학위논문. 세명대학교 대학원.
- 김영신(2003). 클러스터별 인공지능망 구축을 통한 데이터마이닝 모델의 성능 향상. 미간행 석사학위논문. 아주대학교 대학원.
- 문화체육관광부(2014). 2013 체육백서.
- 민준영(1995). 신경망 클러스터리의 성능평가: GLVQ와 k-means 알고리즘을 중심으로. 미간행 박사학위논문. 성균관대학교 대학원.
- 박원규(2002). 인공지능회로망 기법을 이용한 수요예측에 관한 연구: 안경원의 안경렌즈 수요예측을 중심으로. 미간행 석사학위논문. 연세대학교 대학원.
- 박진기(2003). 데이터 마이닝 기법간의 스포츠센터 고객 이탈가능 예측성 평가. 한국체육학회지, 42(5), 369-377.

- 박진기, 김장환(2003). 데이터 마이닝 기법을 활용한 스포츠 센터 고객 이탈 가능성 예측 모형개발. 한국스포츠리서치, 14(4), 157-171.
- 박찬민(2004). 요가프로그램 참여자의 라이프스타일 유형에 따른 참여 동기: 하타 요가를 중심으로. 미간행 석사학위논문. 연세대학교 대학원.
- 송수섭, 이의훈(2001). 인공신경망을 이용한 소비자 선택 예측에 관한 연구. 한국경영과학회지, 26(4), 55-70.
- 신택수, 홍태호(2006). 자기조직화 신경망과 계층적 군집화 기법(SONN-HC)을 이용한 인터넷 बैं킹의 고객세분화 모형구축. *Asia Pacific Journal of Information Systems*, 16(3), 49-65.
- 안광호, 임병훈, 이영효(2009). 최적 시장세분화 결과 도출을 위한 기준변수와 군집분석기법의 선정에 관한 연구. 마케팅관리연구, 14(3), 157-176.
- 안진근(2002). 사례기반추론, 인공신경망, 통계적 기법 모형을 이용한 신용카드 고객 신용도 예측 연구. 미간행 석사학위논문. 한국과학기술원.
- 오광모, 이장택(2003). 데이터마이닝을 이용한 한국프로야구 선수들이 연봉에 관한 모형연구. 한국스포츠사회학회지, 16(2), 295-309.
- 유재구(2009). 대학생의 사이코그래픽스 배경에 의한 스포츠 정보탐색 모형 검증. 미간행 박사학위논문. 한양대학교 대학원.
- 이정학, 정승훈(2007). 프로축구 관전자의 관전결정요인과 미래 소비행동에 따른 시장세분화. 한국체육학회지, 46(5), 365-376.
- 이정학, 정승훈(2012). 프로축구 관전자의 특성에 따른 시장세분화 및 미래 소비행동 분석을 통한 소비자 예측. 한국체육학회지, 51(1), 219-236.
- 이정학, 정승훈, 고용재, Connaughton, D. P. (2013). 월드컵 거리응원 단체관람 참여자의 특성과 스포츠 가치관, 소비문화 및 스포츠태도에 따른 시장세분화. 체육과학연구, 24(1), 86-104.
- 이정학, 정승훈, 이민섭, 황철상(2010). 골프연습장 이용객의 특성 및 라이프스타일 분석을 통한 시장세분화. 체육과학연구, 21(1), 1037-1050.
- 이종진(2003). 도산예측모형의 예측력 비교: 판별분석모형, 로지모형, 인공신경망 모형을 중심으로. 미간행 석사학위논문. 이화여자대학교 대학원.
- 이창환, 안재현, 이주현, 김태웅(2009). 인공신경망모형과 군집분석을 이용한 교각 세굴심 예측. 대한토목학회논문집 B, 29(2), 111-120.
- 이학식(2012). 마케팅 조사. 서울: 집현재.
- 이혜경(2001). 데이터 마이닝을 이용한 보험시장에서의 고객 관계관리(CRM)에 관한 연구: S 생명보험사 사례를 중심으로. 미간행 석사학위논문. 전주대학교 대학원.
- 이호영, 김성재, 이훈영(2006). 시장 세분화와 모형 결합을 활용한 효과적인 생명보험 고객이탈예측모형에 관한 연구. 보험학회지, 74, 33-58.
- 장현철(1999). 프로야구 관중의 시장세분화에 따른 표적마케팅 전략. 미간행 박사학위논문. 경기대학교 대학원.
- 전성무(2001). 기업부실 예측모형 간 성과차이 분석: 인공신경망모형을 중심으로. 미간행 석사학위논문. 홍익대학교 대학원.
- 정승훈(2007). 프로축구 관전자의 관전결정요인과 미래 소비행동에 따른 시장세분화 전략. 미간행 석사학위논문. 경희대학교 대학원.
- 정승훈(2011). 인공신경망을 이용한 프로스포츠 소비자행동 분석 및 예측. 미간행 박사학위논문. 경희대학교 대학원.
- 정승훈, 이정학, 김희정(2011). 2010 남아공월드컵 거리응원 참여자의 관람동기가 응원태도, 관람만족 및 미래 소비행동에 미치는 영향. 한국스포츠산업경영학회지, 16(3), 29-44.
- 정혜원(2004). 통계적 기법과 데이터마이닝 기법을 이용한 이동통신 VAS 가망고객 scoring 모형 비교 연구. 미간행 석사학위논문. 연세대학교 대학원.
- 조용찬, 유재구(2013). 프로파일 분석에 따른 서비스품질, CSR, 행위인도 인식 연구: 축구팬 소비지출 세분화 기준에 대한 타당성 논의. 한국체육학회지, 52(5), 465-480.
- 최형준, 김주학(2006). 인공신경망(Artificial Neural Network)을 이용한 2005년도 영국 워imbledon 테니스 대회의 경기결과 예측에 관한 연구. 한국체육학회지, 45(3), 459-467.
- 한국야구위원회(2014). 관중현황-패턴트레이스. <http://www.koreabaseball.com/Record/Spectator.aspx>
- 허명희, 이용구(2008). 데이터 마이닝 모델링과 사례. 서울: 한나래 출판사.
- 허성수, 전태준(2011). 생활체육 참가자들의 상징적 소비성향과 운동참가수준, 스포츠자신감 및 운동 중독의 관계. 한국체육학회지, 50(3), 189-201.
- 허진, 이계석(2004). 프로야구 관중의 관람동기와 관람만족, 팀동일시 및 재관람의도에 관한 연구. 한국스포츠산업경영학회지, 9(4), 105-121.
- 황혜선(2006). 현대 소비자의 상징적 소비성향: 20-30대 소비자를 중심으로. 미간행 석사학위논문. 성균관대학교 대학원.

- Anton, J. (1996). *Customer relationship management: making hard decisions with soft numbers*. NJ: Prentice Hall.
- Aschenwald, J., Fink, S., & Tappeiner, G. (2001). Brave new modeling: Cellular automata and artificial neural networks for mastering complexity in economics. *Complexity*, 7(1), 39-47.
- Bagozzi, R. B., & Dholakia, U. M. (2002). Intentional social action in virtual communities. *Journal of Interactive Marketing*, 16(2), 2-21.
- Balakrishnan P. V., Cooper M. C., Jacob, V. S., & Lewis, P. A. (1996). Comparative performance of the FSCL neural net & K-means algorithm for market segmentation. *European Journal of Operational Research*, 93(2), 346-357.
- Berry, M. J. A., & Linoff, G. (1997). *Data mining techniques for marketing, sales and customer support*. NY: Wiley.
- Bonoma, T. V., & Shapior, B. P. (1983). *Segmenting the industrial market*. MA: Lexington Books.
- Coakley, J. R., & Brown, C. E. (1993). Artificial neural networks applied to ratio analysis in the analytical review process. *Intelligent Systems in Accounting, Finance and Management*, 2(1), 19-39.
- Cronin, J. J. Jr., Michael, K. B., & Tomas, M. H. (2000). Assessing the effects of quality, value, and customer satisfaction on consumer behavioral intentions in service environments. *Journal of Retailing*, 76(2), 193-218.
- Dasgupta, C. G., Dispensa, G. S., & Ghose, S. (1994). Comparing the predictive performance of a neural network model with some traditional market response models. *International Journal of Forecasting*, 10(2), 235-244.
- Fausett, L. V. (1994). *Fundamental of neural networks: architectures, algorithms, & applications*. NJ: Prentice-Hall.
- Fish, K. E., Barnes, J. H., & Aiken, M. W. (1995). Artificial neural networks: A new methodology for industrial market segmentation. *Industrial Marketing Management*, 24, 431-438.
- Galindo, J. (1998). *A framework for comparative analysis of statistical and machine learning methods: An application to the Black-Scholes option pricing equation*. Technical report Banco de Mexico, Mexico, DF, 4930.
- Gladden, J. M., Irwin, R. L., & Sutton, W. A. (2001). Managing North American major professional sport teams in the new millenium: A focus on building brand equity. *Journal of Sport Management*, 15(4), 297-317.
- Hair, J. F. Jr., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2006). *Multivariate data analysis(6th ed)*. NJ: Prentice-Hall.
- Haykin, S. (1994). *Neural networks: A comprehensive foundation*. NY: Macmillan College Publishing.
- Kamakura, W. A., & Novak, T. P. (1992). Value-system segmentation: exploring the meaning of LOV. *Journal of Consumer Research*, 19(6), 119-132.
- Lajbcygier, P., Boek, C., Palaniswami, M., & Flitman, A. (1995). *Neural network pricing of all ordinaries SPI options on futures*. Presented at the 3rd Conference on Neural Networks in the Capital Markets.
- Lee & Sparks(2007). Culture influence on travel lifestyle: A comparison of Korean Australians and Koreans in Korea. *Tourism Management*, 28(2), 508-518.
- Lloyd, S. P. (1982). Least squares quantization in PCM. *Transactions on Information Theory*, 28(2), 129-137.
- Nelson, M. M., & Illingworth W. T. (1991). *A practical guide to neural nets*. Boston, MA: Addison-Wesley Longman Publishing Co., Inc.
- Nunnally, J. C. & Bernstein, I. H. (1994), *Psychometric theory(3rd ed)*. NY: McGrawHill.
- Nunnally, J. C. (1978). *Psychometric theory(2nd ed)*. NY: McGrawHill.
- Plummer, J. T. (1971). Lifestyle patterns & commercial bank credit card usage. *Journal of Marketing*, 35(2), 35-42.
- Raghupathi, W., Schkade, L., & Raju, B. (1993). *A neural network approach to bankruptcy prediction*. CHI: Probus Publishing Company.
- Sharda, R. (1994). Neural networks for the OR/MS analyst: An application bibliography. *Interfaces*, 24(2), 116-130.
- Sharma, A., & Lambert, D. M. (1994). Segmentation of markets based on customer service. *International Journal of Physical Distribution & Logistics Management*, 24(4), 50-58.
- Sobel, M. E. (1983). Lifestyle Expenditures in Contemporary America Relations Between Stratification and Culture. *American Behavioral Scientist*, 26(4), 521-533.
- Solomon, M. R. (1999). *Consumer behavior: Buying, having, & being(4th ed)*. NJ: Prentice Hall.
- Solomon, P. J., & George, R. W. (1977). The bicentennial traveler: A life-style analysis of the historian segment. *Journal of Travel Research*, 15(3), 14-17.

- Sööllner, A., & Rese, M. (2001). Market segmentation and the structure of competition: Applicability of the strategic group concept for an improved market segmentation on industrial markets. *Journal of Business Research*, 51(1), 25-36.
- Trail, G., Fink, J. S., & Anderson, D. F. (2003). Sport spectator consumption behavior. *Sport Marketing Quarterly*, 12(1), 8-17.
- Treigueiros, D., & Berry, R. (1991). *The application of neural network based methods to the extraction of knowledge from accounting reports*. In System Sciences, 1991. Proceedings of the Twenty-Fourth Annual Hawaii International Conference on (Vol. 4, pp. 136-146). IEEE.
- Vellido, A., Lisboa, P. J. G., & Vaughan, J. (1999). Neural networks in business: Survey of applications(1992-1998). *Expert Systems with Applications*, 17, 51-70.
- West, P. M., Brockett, P. L., & Golden, L. L. (1997). A Comparative Analysis of Neural Networks and Statistical Methods for Predicting Consumer Choice. *Marketing Science*, 16(4), 370-391.

Professional Baseball Spectator's Analysis and Prediction by Using Artificial Neural Networks Model and Logistic Regression Model

Seung-hoon Jeong
Kyunghee University

This study classified and analyzed groups of spectators of professional baseball through market segmentation and predicted the sports consumer behavior by using artificial neural networks model and logistic regression model. The results of hierarchical cluster analysis, K-means cluster analysis, cross-tabulation analysis and one-way ANOVA using PASW 18.0 and AMOS 18.0 suggest five clusters of consumer segments and by using Modeler 14.1, artificial neural networks model was made to predict the data. By using artificial neural networks model and logistic regression model, hit ratio was grasped about the spectator satisfaction and future consumption behavior. The results are as follow: The hit ratio were high in 'cluster 5' for artificial neural networks model(spectator satisfaction: 71.3%, future consumption behavior: 99.3%) and logistic regression(spectator satisfaction: 71.8%, future consumption behavior: 96.5%). Furthermore, cross-tabulation and one-way ANOVA was performed to understand the cluster's characteristic which had highest hit ratio about the spectator satisfaction and future consumption behavior. And through this marketing strategy was suggested.

Key Words: Professional Baseball, Spectator, Prediction, Artificial Neural Networks, Logistic Regression 