

제품유형에 따른 웹쇼핑 소비자의 조절초점성향 분류

백 종 범[†] · 한 정 석[†] · 장 은 영^{††} · 김 용 범^{††} · 최 자 영^{†††} · 이 수 원^{††††}

요 약

소비자 행동이론에 따르면 사람의 성향은 향상초점과 예방초점이라는 두 가지 조절초점 유형으로 나누어지며, 이 두 가지 성향은 다양한 영역에 있어서 소비자의 의사결정에 많은 영향을 미치는 것으로 알려져 있다. 본 연구에서는 개인화 추천에서 Cold Start 문제의 최소화 및 추천 알고리즘 성능 개선을 위하여 조절초점이론을 적용한다. 이를 위하여, 웹쇼핑 로그로부터 소비자 별 행동변수, 정보탐색활동성 지수를 추출하고 이를 활용한 소비자 조절초점성향 분류 방법을 제안한다. 본 연구는 사회과학/IT 융합 연구로서 소비자행동 이론의 시스템화 가능성을 입증하였다는 점에 있어서 의의를 지니며, 향후 다양한 분야의 이론들을 적용한 IT 서비스에 대한 연구로 확장하고자 한다.

키워드 : 조절초점이론, 제품유형, 분류, 개인화, 웹쇼핑로그, 사용자 모델링, 정보탐색활동성 지수

Regulatory Focus Classification for Web Shopping Consumers According to Product Type

Jongbum Baik[†] · Chungseok Han[†] · Eunyoung Jang^{††} · Yongbum Kim^{††} · Jayoung Choi^{†††} · Soowon Lee^{††††}

ABSTRACT

According to consumer behavior theory, human propensity can be divided into two regulatory focus types: promotion and prevention. These two types have much influence on the consumer's decision in many diverse areas. In this research, we apply regulatory focus theory to personalized recommendation to minimize the cold start problem and to improve the performance of recommendation algorithms. To achieve this goal, we extract the consumer behavior variables and information exploration activity index from web shopping logs. We then use them for classifying regulatory focus of the consumer. This research has the contribution to show the possibility of systematization of consumer behavior theory as an interdisciplinary research tool of social science and information technology. Based on this attempt, we will extend the research to IT services adapting theories on other areas.

Keywords : Regulatory Focus Theory, Product Type, Classification, Personalization, Web Shopping Log, User Modeling, Information Exploration Activity Index

1. 서 론

소비자행동(Consumer Behavior)이라는 사회과학 분야의 학문에서는 소비 행동을 일으키는 이유(원인)를 설명하는 다양한 이론을 제시한다. 소비자행동에서는 소비자(혹은 사용자)의 행동 및 의사결정에 영향을 미치는 요인을 내적 요인(Internal Influences)과 외적 요인(External Influences)으

로 나누어 설명한다[1]. 내적 요인으로는 인구통계정보(Demographics), 생활방식(Lifestyle), 동기(Motivation), 개성(Personality), 지식(Knowledge), 태도(Attitude), 신념(Belief), 느낌(Feeling) 등이 있으며, 외적 요인으로는 문화(Culture), 지역(Locality), 과거 경험 참조 집단(Past Experience Reference Group) 등이 있다.

이와 같은 소비자행동 요인들은 이미 상당부분 개인화 서비스 관련 데이터마이닝 연구들[2-6]에서 고려되고 있다. 그러나 내적 요인 중 동기, 개성, 태도, 신념, 느낌 등 다소 추상적인 개념들에 대해서는 대략적으로 2009년도부터 개인화 서비스에 적용하기 위한 시도가 이루어진 것으로 조사되었으나 아직은 연구의 초기단계인 것으로 판단된다[2,3].

사람의 개성을 개인화 서비스에 반영하기 위한 연구 중, Hu와 Pu의 연구 결과[3]에 따르면 사람의 개성(혹은 성향)

* This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MEST) (No.2011-0027668).

† 준희원: 숭실대학교 컴퓨터학과 박사과정

†† 준희원: 숭실대학교 벤처중소기업학과 석사과정

††† 정희원: 숭실대학교 경영대학 벤처중소기업학과 부교수

†††† 정희원: 숭실대학교 컴퓨터학부 교수

논문접수: 2012년 3월 20일

수정일: 1차 2012년 5월 3일

심사완료: 2012년 5월 3일

* Corresponding Author: Soowon Lee(swlee@ssu.ac.kr)

을 추천 시스템에 반영하였을 경우, Cold Start 문제로 인한 성능 저하가 최소화되었을 뿐 아니라 기존의 협력적 여과(Collaborative Filtering) 추천 기법의 성능도 향상된 것으로 나타났다. 그러나 Hu와 Pu의 연구에서 사람의 개성을 분류하기 위하여 이용한 심리학 분야 개성 판별 도구의 경우, 분류된 사람의 개성과 실제 구매 행동 간의 상관성이 겹침이 되지 않았다는 점에 있어서 개인화 서비스에 바로 적용하기에 다소 미흡한 요소가 존재한다.

이에 반해 소비자행동에서는 사람의 내적 요인, 특히 개성에 속하는 사람의 성향(Propensity)을 소비자의 의사결정 및 행동에 영향을 미치는 중요한 영향요인(Influential Factor)으로 제시한다[7]. 사람의 성향에 대한 소비자행동의 대표적 이론 중 하나인 조절초점이론(Regulatory Focus Theory)[8]에 따르면 사람이 목표를 추구함에 있어서 이상과 희망, 열망 등을 조절기제(Control Mechanism)로 가지는 향상초점(Promotion Focus)과 의무, 책임감을 조절기제로 가지는 예방초점(Prevention Focus)이라는 두 가지 유형으로 사람의 성향이 나누어진다고 한다. 또한 이와 같은 두 가지 성향은 정보의 탐색 및 처리 과정, 대안의 평가 방식 등 다양한 영역에 있어서 소비자의 의사결정에 많은 영향을 미치는 것으로 알려져 있다[8,9].

본 연구에서는 이와 같은 조절초점이론에 의하여 분류된 사람의 성향이 개인의 선호도(Preference)를 형성하는 데 있어서 중요한 요인이 될 수 있다는 가정 하에 소비자행동 이론의 시스템화 가능성을 보여준다. 이를 위하여, 웹쇼핑 로그로부터 소비자 별 행동변수를 추출하고, ‘정보탐색활동성 지수’를 정의한다. 또한, 소비자 별 행동변수, 정보탐색활동성 지수를 입력변수로, 설문에 의한 소비자 조절초점성향 분류 결과(향상/예방 초점)를 목표 변수로 하는 소비자 조절초점성향 분류 방법을 제안한다.

또한 소비자행동 이론에서는 제품 유형 분류 중, 쾌락재(Hedonic)와 실용재(Utilitarian)로 나누어 많은 연구를 수행하고 있다[10,11]. 쾌락재는 즐거움, 기쁨 그리고 긍정적 홍분 등과 같은 감정들을 제공하는 제품을 의미하며 꽃, 의류, 음악, 스포츠카, 고급 시계, 초콜릿 등과 같은 제품들이 있다. 반면에 실용재는 제품의 기능적 요소가 중요시되는 제품을 의미하며 전자레인지, 미니밴, 주택 보안 시스템, 개인 컴퓨터 등과 같은 제품들이 있다[11].

이와 같은 제품유형은 조절초점이론과 서로 상관관계가 깊은 요소들로서 이를 함께 고려하여 연구를 수행한 사례들이 다수 존재한다[10,11]. 그러므로 본 연구에서도 소비자행동 분야 공동연구자들이 ‘웹쇼핑로그’에 등장한 제품들을 분류한 결과 중 쾌락재에 해당하는 제품들을 ‘소비자 별 정보탐색활동성 지수’를 추출하는데 활용한다.

본 논문의 구성은 다음과 같다. 먼저 2장에서는 본 연구에서 제안하는 소비자 조절초점성향 자동 분류 방법을 설명한 후, 3장에서는 제안방법에 대한 실험결과를 논의한다. 마지막으로 4장에서는 결론을 기술하고 향후 연구를 제시한다.

2. 소비자 조절초점성향 자동 분류

2.1 소비자 별 행동변수 추출

본 연구에서 수집된 ‘웹쇼핑로그’는 소비자 ID, 소비자가 탐색한 제품의 ID, 해당 제품 페이지 내에서 발생한 클릭이벤트, 해당 웹페이지(URL)에 접속한 시점 등의 정보로 구성된다. 본 연구에서는 소비자 c가 특정 제품페이지 p내에서 발생시킨 ‘제품페이지 내 클릭이벤트 로그’를 $T^{c,p}$ 로 정의한다. 또한 웹쇼핑로그에서 추출된 소비자 별 $T^{c,p}$ 를 가공하여 소비자 별 행동변수를 추출한다.

소비자 별 행동변수는 크게, ‘제품 정보 탐색’ 및 ‘대안 탐색’ 변수로 구성된다. 제품 정보 탐색 변수 추출을 위해서는 제품별 온라인 구전 참조(상품리뷰, 상품리뷰 등급, 일반리뷰, 일반리뷰 등급 등), 구매 의도(장바구니, 구매하기, 찜하기 등) 등의 행동 유형에 대하여 클릭 수, 클릭 여부, 소요시간 등의 행동변수들을 추출한 다음, 이를 행동변수와 제품페이지 내 총 클릭 수 및 소요시간 등에 대한 총계, 평균, 최대값, 최소값, 표준편차를 구하여 제품정보 탐색 변수를 생성한다. 또한 ‘대안탐색’ 변수는 ‘소비자가 접속한 전체 제품페이지의 수’로 정의된다.

2.2 소비자 별 정보탐색활동성 지수 추출

본 연구에서는 “한 개인의 내면에는 일반적으로 두 가지 조절초점성이 모두 내재되어 있으며 상황적인 요인에 의해 이 중 하나의 조절초점 상태로 일시적인 점화가 일어나기도 한다”라는 Higgins의 연구 결과[8]와 조절초점성향에 따른 정보탐색행동의 차이를 언급한 연구 결과[8-10]에 기반하여, 소비자가 제품 정보 탐색 행동을 활발히 수행하는 정도를 추론하는 식 (1)을 제안한다. 식 (1)에서 $Activated(T^{c,p})$ 는 소비자 c가 제품 페이지 p 내에서 활발한 제품 정보 탐색 행동을 수행하였는지 여부를 판단하는 이진 함수이며 q는 소비자 c가 탐색한 제품 중 쾌락재에 속하는 제품 페이지의 개수를 의미한다. 또한 식 (1)에서 쾌락재에 속하는 제품 페이지지만 이용하는 이유는 실용재에 속하는 제품 페이지 내에서 정보탐색활동에 있어서 별다른 차이가 나타나지 않았기 때문이다.

$$ExplorationActivityIndex(c)$$

$$= \frac{\sum_{p=1}^q (Activated(T^{c,p}) - (1 - Activated(T^{c,p})))}{q}$$

$$= \frac{\sum_{p=1}^q (2 \times Activated(T^{c,p}) - 1)}{q} \quad (1)$$

$$Activated(T^{c,p}) \in \{0, 1\}$$

소비자 별 정보탐색활동성 지수

Table 1은 소비자 별로 식 (1)을 계산한 결과의 일부이다. 소비자-104번의 경우에는 활동적인 행동을 수행한 탐색 대안(제품 페이지 p)의 개수가 세 건, 활동적이지 않은 행동을 수행한 탐색 대안의 개수가 한 건이므로 식 (1)에 따라 $0.5 = (3-1)/4$ 정도의 정보탐색활동성을 지닌 것으로 계산된다. 반면 소비자-1은 비활동적인 행동만을 수행한 소비자로 식 (1)에 따라 -1 정도의 저조한 정보탐색활동성을 지닌 것으로 계산된다.

표 1. 소비자 별 정보탐색활동성 지수 계산 결과
Table 1. Example of calculation of Exploration Activity Index

ID	# of Activated($T^{c,p}$)	# of I-Activeated ($T^{c,p}$)	q	Result of (1)
1	0	2	2	-1
10	2	2	4	0
104	3	1	4	0.5

식 (1)을 계산하기 위해서는 먼저 활발한 제품 정보 탐색 행동이 나타난 제품 페이지와 그렇지 않은 제품 페이지를 분별하는 이진함수 값 $Activated(T^{c,p})$ 를 추출하는 과정이 필요하다. 본 연구에서는 이를 군집화 문제로 정의하고 제품별 온라인 구전 참조에 대한 클릭 수 및 클릭 여부 등의 행동변수, 제품페이지 내 총 클릭 수, 소요시간 등을 입력변수로 하여 k-평균 군집화 알고리즘을 적용한다. 입력변수의 선정은 소비자행동 분야 공동 연구자들이 수행한 조절초점성향에 따른 행동 차이 관련 가설에 대한 검증 실험 결과 [12]에서 유의한 것으로 나타난 측정대상(전체리뷰, 제품페이지 내부)에 속하는 변수들 중, 군집 알고리즘 성능 평가 척도인 실루엣(Silhouette)[13] 수치가 높게 나타나는 변수들을 선택하는 방식으로 이루어졌다.

군집화 결과에 따라 온라인 구전 참조, 제품페이지 내 총 클릭 수 및 소요시간의 수치가 평균적으로 높은 군집을 ‘활발한 제품 정보 탐색 행동 군집’(Cluster-1)으로 정의하고 나머지 군집을 ‘활동성이 저조한 제품 정보 탐색 행동 군집’(Cluster-2)으로 정의하였다.

2.3 소비자 조절초점성향 분류

본 연구에서는 2.1절에서 추출한 ‘소비자 별 행동 로그’와 2.2절에서 추출한 ‘소비자 별 정보탐색활동성 지수’를 결합하여 소비자 조절초점성향 분류기를 학습 및 검증하기 위한 CV(Cross Validation)데이터를 구성한다. 그러나 2.1절에서 추출한 ‘소비자 별 행동 로그’는 변수의 개수가 총 196개로 모든 변수를 분류기 학습에 이용할 경우, 의미적으로 중복되는 변수들 간의 충돌이 발생함으로 인하여 오히려 분류 정확도가 저하되는 것으로 나타났다(Fig. 2 참조). 그러므로 본 연구에서는 F-Statistic에 기반한 p-value[14]를 계산하여 각 속성변수에 대한 중요도를 추출한 후, 중요도가 0.9 이상

인 변수들을 의사결정나무(C5.0) 학습 알고리즘의 입력변수로 이용한다. Fig. 1은 중요도가 0.9 이상인 변수 17개를 입력변수로 이용하여 학습된 의사결정나무의 모형이다. Fig. 1에서 WishListClickBin_Min은 소비자 c가 방문한 각 제품페이지 p 내에서 장바구니 혹은 찜하기 버튼을 클릭하였는지 여부에 대한 행동변수의 최소값을 의미한다. 즉, 소비자 c가 방문한 모든 제품에 대하여 장바구니 혹은 찜하기를 클릭하였을 경우에만 1이 되며, 그렇지 않은 경우 0이 된다. 또한 AlternativeProductPageCnt는 사용자 c가 탐색한 대안의 개수를 의미한다.

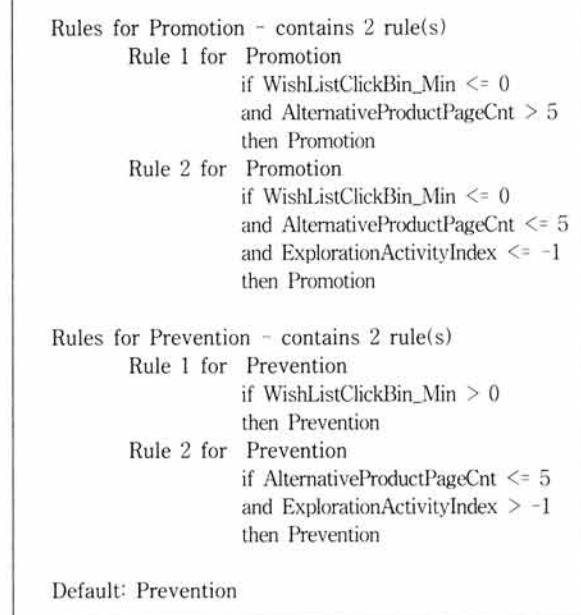


그림 1. 학습된 소비자 조절초점성향 분류 규칙
Fig. 1. Rules for classifying regulatory focus of consumer's learned by C5.0

3. 실험 및 평가

3.1 웹쇼핑로그 수집

본 연구에서는 웹쇼핑로그 수집을 위하여 국내 종합쇼핑몰 사이트 중 하나인 11번가(<http://www.11st.co.kr/>)를 분석하여 소비자의 페이지 이동 경로 및 클릭정보 등을 수집할 수 있는 실험용 웹브라우저(이하 웹쇼핑로그 수집기)를 개발하였다.

개발된 ‘웹쇼핑로그 수집기’를 이용하여 총 163명의 실험 참가자들로부터 총 7,078건의 웹쇼핑로그를 수집하였으며, 설문조사를 통하여 실험 참가자들의 조절초점성향을 측정하였다. 조절초점성향 측정 결과, 64명이 향상 성향, 99명이 예방 성향을 지닌 소비자인 것으로 나타났다. 이와 같이 두 그룹의 비율 차이가 심한 경우, 한쪽 그룹에 치우친 분류 모델이 생성되는 문제가 있으므로 본 연구에서는 데이터의 균형을 맞추기 위하여 63명의 예방 성향 소비자와 61명의 향상 성향 소비자를 임의로 추출하여 총 124명의 소비자에 대한 ‘웹쇼핑로그’ 5,173건을 이용하여 실험을 수행하였다.

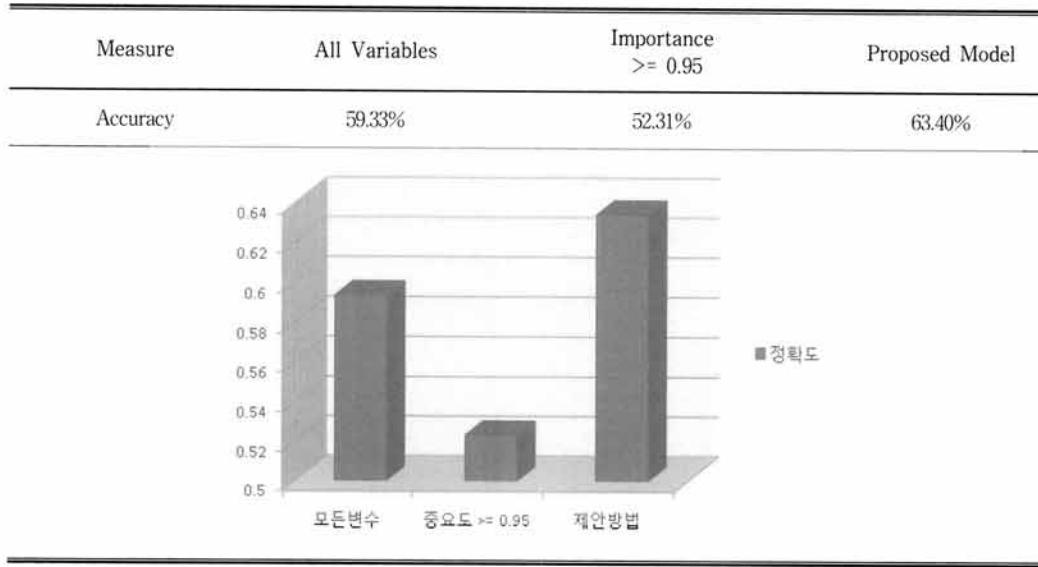


그림 2. 변수조합에 따른 분류 성능 차이 비교
Fig. 2. Comparison result according to variable combinations

3.2 실험 방법

제안방법의 성능을 평가하기 위하여 의사결정나무(C5.0)의 입력변수 조합을 비교하는 실험을 수행하였다. 평가 방법은 10-fold Cross Validation을 이용하였으며, 평가 척도는 분류 정확도(Accuracy)를 이용하였다. 단, 10-fold Cross Validation 수행 시 fold 구성에 따라 성능 평가 결과에 차이가 발생하는 문제가 확인되어 10-fold Cross Validation을 10회 반복 수행한 평균 정확도를 이용하여 시스템의 성능을 평가하였다.

3.3 실험 결과

Fig. 2는 변수조합에 따른 분류 성능의 차이를 비교한 결과이다. Fig. 2에 따르면 제안방법(중요도 ≥ 0.95)이 중요도가 0.95 이상인 변수들만 이용하여 분류기를 학습한 경우보다 평균적으로 약 11% 포인트 더 높은 성능을 보이는 것으로 나타났으며, 모든 변수를 이용하여 분류기를 학습한 경우보다 평균적으로 약 4% 포인트 더 높은 성능을 보이는 것으로 나타났다.

중요도가 0.95 이상인 변수들만 이용하였을 경우의 성능이 가장 저조한 이유는 정보탐색활동성 지수(Exploration ActivityIndex)의 중요도(0.92)가 상대적으로 낮음으로 인하여 입력변수로 선정되지 않기 때문인 것으로 판단된다. 이에 따라 중요도가 0.95 이상인 변수에 정보탐색활동성 지수를 추가하여 의사결정나무를 학습한 결과, Fig. 1과 동일한 형태의 분류 규칙이 생성되는 것으로 나타났다. 또한 모든 변수를 이용하였을 경우에 성능이 저조한 이유는 비슷한 의미의 변수들이 다수 포함되어 서로 간에 간섭을 받기 때문인 것으로 판단된다.

4. 결론 및 향후연구

본 연구에서는 웹쇼핑로그를 이용하여 소비자 조절초점성향을 자동으로 분류하는 시스템을 개발하였다. 실험 결과에 따르면 제안방법이 약 63% 정도의 정확도를 지니는 것으로 나타났다. 또한 제품 정보 탐색 로그를 이용하여 소비자의 제품 정보 탐색활동의 정도를 측정할 수 있는 '소비자 별 정보탐색활동성 지수'를 추출하는 방법을 제안하고 적용함으로써 평균적으로 약 11% 정도 정확도가 향상된 것으로 나타났다.

본 연구는 사회과학/IT 융합 연구로서 소비자행동 분야 이론의 시스템화를 시도하였다는 점에 있어서 의의를 지닌다. 이러한 시도를 토대로 향후 소비자행동 분야뿐만 아니라 보다 다양한 분야의 이론들을 IT 시스템 혹은 서비스로서 제공하기 위한 연구로 확장하고자 한다.

그러나 본 연구에서는 분류기 학습 알고리즘에 이용할 입력변수를 선정하는 과정에서 F-Statistic을 이용하여 계산된 변수중요도 값만으로는 최적의 변수조합을 찾는데 한계가 있는 것으로 나타났다. 향후에는 이와 같은 입력변수 선정 과정을 보다 정교하게 수행할 수 있는 방법을 연구할 필요가 있다. 또한 본 연구에서 생성한 소비자 조절초점성향 분류 모델이 실제로 의미 있는 규칙으로 구성되어 있는지에 대한 소비자행동 분야의 교차검증도 필요하다.

참 고 문 헌

- [1] Richard L. Sandhusen, *Marketing*, Barron's Business Review, 2000.
- [2] Dunn, G., Wiersema, J., Ham J., Aroyo, L. "Evaluating

- Interface Variants on Personality Acquisition for Recommender Systems," *Proc. of UMAP 2009*, Vol.5535, pp.259-270, 2011.
- [3] Rong Hu, Pearl Pu, "Enhancing Collaborative Filtering Systems with Personality Information," *Proc. of the fifth ACM conference on Recommender systems*, pp.197-204, 2011.
- [4] Jill Freyne, Shlomo Berkovsky, Nilufar Baghaei, Stephen Kimani, Gregory Smith, "Personalized Techniques for Lifestyle Change," *Proc. of the 13th conference on Artificial intelligence in medicine*, pp.139-148, 2011.
- [5] Suk Jung Yong, Hyung Do Lee, Han Ku Yoo, Hee Yong Youn, Ohyoung Song, "Personalized Recommendation System Reflecting User Preference with Context-awareness for Mobile TV," *Proc. of Ninth IEEE International Symposium on Parallel and Distributed Processing with Applications Workshops (ISPAW)*, pp.232-237, 2011.
- [6] Xiaoyuan Su, Taghi M. Khoshgoftaar, "A survey of collaborative filtering techniques," *Advances in Artificial Intelligence*, 2009.
- [7] Michael R. Solomon, *Consumer Behavior: Buying, Having, and Being*, 9th Ed., McGraw-Hill, New York, 2010.
- [8] Crowe, Ellen, E. Tory Higgins, "Regulatory Focus and Strategic Inclinations : Promotion and Prevention in Decision Making," *Organizational Behavior and Human Decision Processes*, Vol.69, Issue.2, pp.117-132, Feb., 1997.
- [9] Pham. Michel Tuan, Hannah H. Chang, "Regulatory Focus, Regulatory Fit, and the Search and Consideration of Choice Alternatives," *Journal of Consumer Research*, Vol.37, pp.626-640, 2010.
- [10] J. Jung, "The Influence of Chronic & Situational Regulatory Focus and Product Type on Consumer's Information Search and Decision-Making Rule," M.S. Thesis, Ewha Univ., Seoul, Korea, 2009.
- [11] Uzma Khan, Ravi Dhar, "A Behavioral Decision Theoretic Perspective on Hedonic and Utilitarian Choice," *Inside Consumption*, 2004.
- [12] J. Choi, E. Jang, "The Effect of Chronic Regulatory Focus on Online Review and Information Search Behavior by Using Web Log Data," *Proc. of 2012 International Conference on Economics, Business and Marketing Management*, Vol.29, pp.163-167, 2012.
- [13] Peter J. Rousseeuw, "Silhouettes: a Graphical Aid to The Interpretation and Validation of Cluster Analysis," *Journal of Computational and Applied Mathematics*, Vol.20, pp.53-65, 1987.
- [14] SPSS, "Feature Selection Algorithm," SPSS Korea, 2007.



백종범

e-mail : jongdal100@gmaill.com
 2006년 학점은행제(한국IT전문학교) 전자
 계산학(학사)
 2009년 숭실대학교 컴퓨터학과(석사)
 2009년~현 재 숭실대학교 컴퓨터학과
 박사과정

관심분야: 데이터 마이닝, 정보 검색, 패턴 인식, AI, 소비자
 행동 등



한정석

e-mail : jshan97@mining.ssu.ac.kr
 2004년 관동대학교 컴퓨터공학과(학사)
 2007년 숭실대학교 일반 컴퓨터학과(석사)
 2007년~현 재 숭실대학교 컴퓨터학과
 박사과정

관심분야: 데이터마이닝, 인공지능,
 기계학습 등



장은영

e-mail : young_good@naver.com
 2011년 숭실대학교 벤처중소기업학과
 (학사)
 2011년~현 재 숭실대학교 벤처중소기업
 학과 석사과정
 관심분야: 소비자 심리, 온라인 마케팅,
 SNS마케팅 등



김용범

e-mail : yb@ssu.ac.kr
 2008년 숭실대학교 벤처중소기업학과
 (학사)
 2010년~현 재 숭실대학교 벤처중소기업
 학과 석사과정
 관심분야: 소비자 행동, 마케팅,
 브랜드 관리 등



최자영

e-mail : choi181@ssu.ac.kr
1994년 한양대학교 소비자학과(학사)
1996년 한양대학교 소비자학과(석사)
2001년 The Ohio State University
Marketing(박사)
2001년~2002년 The Ohio State University
Full-time Lecturer
2002년~2006년 삼성전자 한국총괄 마케팅팀 과장
2006년~현 재 숭실대학교 경영대학 벤처중소기업학과 부교수
관심분야: 소비자 심리, 온라인 마케팅, 하이테크 마케팅 등



이수원

e-mail : swlee@ssu.ac.kr
1982년 서울대학교 자연과학대학
계산통계학과(학사)
1984년 한국과학기술원 전산학과(석사)
1994년 University of Southern California
전산학과(박사)
1995년~현 재 숭실대학교 컴퓨터학부 교수
2008년~현 재 한국BI데이터마이닝학회 부회장
2008년~2009년 한국정보과학회논문지(SA) 편집위원장
관심분야: Data Science, 기계학습, 인공지능 등