

온·오프라인 댓글 분석이 활용된 Word2Vec 기반 상품기획 모델연구: 버티컬 무소음마우스 사용자를 중심으로

안영휘
국립공주대학교 컴퓨터공학과 박사수료

A Study on the Product Planning Model based on Word2Vec using On-offline Comment Analysis: Focused on the Noiseless Vertical Mouse User

Yeong-Hwi Ahn
ph.D Student, Division of Computer Engineering, Kongju National University

요약 본 논문에서는 버티컬 무소음 마우스 10,000건에 대한 웹크롤링을 통해 수집된 정형화된 데이터셋을 Word2Vec을 이용하여 단어 간 유사도분석을 시행하고 컴퓨터공학과 대학생 92명에게 5일 동안 제시된 상품을 사용하게 하고 자가보고식 설문 분석을 시행하도록 하였다. 설문 분석은 서술식 형태로 수집하여 단어빈도 분석과 단어 간 유사도분석에서 추출된 상위 50개 단어를 제시하고 선택하는 방식으로 이루어졌다. 전자상거래 사용자 상품평 유사도 분석결과 내용 중 클릭 키워드에 대한 장점으로 통증(.985), 디자인(.963)가 분석되었으며 단점은 가볍다(.952), 적응(.948)이었다. 서술식 빈도분석에서는 버티컬(123개), 통증(118개)이 가장 많이 선택 되었으며 장/단점 유사단어를 선택에 해당되는 장점에서는 버티컬(83개), 통증(75개) 선택 되었으며 단점에서는 적응(89개), 버튼(72개)이었다. 따라서 본 연구에서 적용한 방식을 상품기획 프로세스의 신상품 개발 및 기존 상품의 검토 전략으로 반영 시 중견기업, 중소기업의 의사결정자와 상품기획자는 의사결정에 중대한 자료로 활용 할 수 있을 것으로 기대된다.

주제어 : 상품기획, 유사도, 전처리, 워드투백, 기법, 웹크롤링

Abstract In this paper, we conducted word-to-word similarity analysis of standardized datasets collected through web crawling for 10,000 Vertical Noise Mouses using Word2Vec, and made 92 students of computer engineering use the products presented for 5 days, and conducted self-report questionnaire analysis. The questionnaire analysis was conducted by collecting the words in the form of a narrative form and presenting and selecting the top 50 words extracted from the word frequency analysis and the word similarity analysis. As a result of analyzing the similarity of e-commerce user's product review, pain (.985) and design (.963) were analyzed as the advantages of click keywords, and the disadvantages were vertical (.985) and adaptation (.948). In the descriptive frequency analysis, the most frequently selected items were Vertical (123) and Pain (118). Vertical (83) and Pain (75) were selected for the advantages of selecting the long/demerit similar words, and adaptation (89) and buttons (72) were selected for the disadvantages. Therefore, it is expected that decision makers and product planners of medium and small enterprises can be used as important data for decision making when the method applied in this study is reflected as a new product development process and a review strategy of existing products.

Key Words : Product Planning, Similarity, Text Preprocessing, Word2Vec. Techniqu, Web Crawling

*Corresponding Author : Yeong-Hwi Ahn(yhahn02@naver.com)

Received August 31, 2021

Accepted October 20, 2021

Revised September 20, 2021

Published October 28, 2021

1. 서론

2019년도에 발생한 COVID-19 팬데믹은 전자상거래의 이용자 수와 거래액을 증가에 의미 있는 영향을 미치고 있다. 2021년 6월 전자상거래 매출액은 15조 6,558억 원으로 전년동월대비 23.5% 증가하였다[1]. 이러한 소비패턴의 변화에는 ICT 기술의 발전과 전자상거래 플랫폼 사업자의 공격적인 마케팅이 전자상거래의 확산과 발전을 더욱 가속화시키고 있다[2]. 최근에는 시공간의 제약 없이 연속적 상호작용이 가능한 모바일 디바이스 이용이 보편화함에 따라, 상황과 맥락에 적합한 IT디바이스를 선택적으로 활용하며 끊임 없이 다양한 정보를 탐색할 수 있는 환경이 조성되었다. 오늘날의 소비자는 이러한 디바이스 확장으로 개인용컴퓨터와 모바일을 넘나들며 검색 기반으로 필요한 상품의 정보를 직접 검색하고 비교하는가 하면 사용경험을 다른 소비자들과 공유하는 AISCEAS(Attention, Interest, Search, Comparison, Examination, Action, Share)와 같은 구매에 대한 의사결정을 재점검하는 행동 패턴으로 더욱 확산되고 있다[3]. 이런 구매행동패턴을 분석하기 위한 기법으로 텍스트마이닝이 활용되고 있다. 텍스트마이닝은 대용량의 텍스트 데이터에서 가치와 의미가 있는 다양한 정보를 찾아내는 기법으로 전자상거래에서는 대량의 사용자 상품평을 수집하여 딥러닝 학습을 통한 소비자의 유의미한 자료를 분석할 수 있다. 소비자는 전자상거래에서 구매한 상품에 대하여 자신이 느끼는 감정과 생각을 마음껏 표현하는 창구이기도 하다. 상품에 대한 긍정적인 평가는 기업의 수익 향상과 시장에 대한 우위를 점할 수 있으나 반대로 부정적인 평가는 수익 손실과 기업 생존에 위험을 초래할 수 있다. 이에 본 연구에서는 웹 크롤링을 통해 포털사이트 및 전자상거래 내 사용자 상품평을 수집하여 딥러닝 Word2Vec 기법을 통해 단어 간 유사도 분석을 통해 상품에 대한 장단점을 분석하고 자가 보고식 설문 조사를 병행 진행함에 따라 포털사이트 및 전자상거래에서 수집 분석한 자료가 유의미한지 상호 비교 평가를 활용한 상품기획 방안을 제시하고자 한다.

본 논문의 다음과 같이 구성하였다. 2장에서는 본 연구에 사용된 관련된 연구에 관하여 소개한다. 3장에서는 본연구의 시스템 구성 개요, 아키텍처에서는 스택 구조의 레이어별 수행과정이 기술되어 있으며, 4장에서는 제안 모델의 테스트 및 결과를 기술하였다. 5장에서는 본 연구에 대한 결론에 관한 향후 연구 방향을 기술하였다.

2. 관련연구

2.1 Word2Vec

2013년 구글에서 제안된 Word2Vec은 단어 임베딩 모델로서 간편하고 쉽게 접근할 수 있는 가장 많이 적용되고 있으며 구조는 간단하면서도 기능은 우수한 특징이 있다[4]. 이 워드 임베딩 모델은 SG(Skip-gram)와 CBOW(Continuous Bag Of Word) 라는 두 가지 학습 아키텍처를 제안하는데 CBOW는 주변단어를 이용하여 중심 단어를 예측하는 구조이며 입력층(Input layer), 은닉층(Hidden layer), 출력층(Output layer)로 구성된 간단한 인공신경망의 형태이다[5-7]. 즉, CBOW 모델은 문장의 주변 단어를 기준으로 설정하여 그 사이에 존재할 수 있는 단어를 예측하도록 학습시키는데 목적을 둔다.

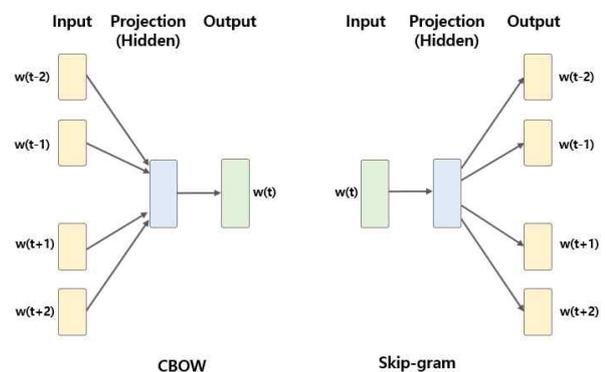


Fig. 1. Two training algorithms of word2vec modeling

예를 들어 ‘그녀는 ___을 먹는다’라는 문장에 대하여 “___”에 들어갈 수 있는 중심단어, 즉“밥”을 예측하도록 학습하는 것을 의미한다[8,9]. 이를 달성하기 위해 입력층에서 은닉층으로 넘길 때의 가중치와 은닉층에서 출력층으로 넘길 때의 가중치를 학습하게 되는 것이다. SG는 CBOW 유사한 구조로 중심단어를 근거로 주변 단어를 예측하는 형태의 인공신경망 모델에서 중심 단어의 입력으로 인해 주변 단어가 출력된다. SG 모델은 CBOW와는 반대로 문장의 중심 단어를 기준으로 앞쪽 또는 뒤쪽에 나타날 수 있는 주변 단어를 예측하도록 학습하는데 목적을 둔다. CBOW와 SG의 비교 시 일반적으로 CBOW 모델이 학습이 빠르고 자주 나타나는 단어들을 잘 표현할 수 있으며 syntactic에서도 SG 모델에 비해 높은 수치를 보인다. SG는 비교적 적은 양의 데이터에 잘 작동하고 빈도가 적은 단어들을 잘 표현하며

semantic에서는 CBOW 보다 높은 수치를 보인다[10, 11]. 따라서 Word2Vec은 인공 신경 모델을 사용하여 대단위 말뭉치 속에 포함된 모든 단어의 벡터값을 구하여 배열함으로써 단어 간의 연관성을 분석하는 기술이다. 이 딥러닝 기술은 함께 쓰인 단어 간의 관계와 계열을 분석할 수 있다.

2.2 유사도

벡터, 공간,점 등으로 표시되는 임의의 두 지점의 거리를 측정하여 유사점을 측정하는 기법으로 코사인 유사도, 자카드 유사도, 유클리디안 유사도, 맨하탄 유사도 등이 있다[12]. 유클리디안 유사도는 거리를 측정하는 공식과 같으며 n차원 공간에서 두 점 사이의 최소 거리를 구하는 접근법이다[13]. 다만, 결과 값이 1보다 큰 값이 나올 수 있으므로 값을 제한하는 과정이 필요하다. 자카드 유사도는 두 문장을 각 단어의 집합으로 생성 한 뒤, 생성된 집합을 통해 유사도를 측정한다[11]. 측정된 값은 0~1 사이이며 1에 가까울수록 유사도가 높음을 의미한다. 맨하탄 유사도는 동일한 방식으로 반복되는 점들의 규칙적인 구조인 사각형 격자로 이루어진 지도에서 출발점에서 도착점까지 가로지르지 않고 갈 수 있는 최소 거리를 구하는 방법이다.

$$\text{similarity} = \cos(\theta) = \frac{\mathbf{A} \cdot \mathbf{B}}{\|\mathbf{A}\| \|\mathbf{B}\|} = \frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2} \sqrt{\sum_{i=1}^n B_i^2}}$$

Fig. 2. Cosine Similarity

코사인 유사도는 내적 공간 내에서 두 벡터 사이의 코사인 각도를 구하는 방법으로 -1~1 사이 값을 가지며, 1에 가까울수록 유사도가 높다[14]. 코사인 유사도는 어떤 개수의 차원에도 적용할 수 있지만 흔히 다차원의 양수 공간에서의 유사도 측정에서 자주 이용된다[15]. 예를 들어 텍스트 마이닝 및 정보 검색 분야에서, 하나 하나의 단어는 각각의 차원을 구성하고 각 문서는 단어가 문서에서 나타나는 회수로 표현된 벡터 값을 가진다[14]. 이러한 다차원 공간 내에서 코사인 유사도는 두 문서의 유사도를 측정하는 매우 효과적인 방법이다. 따라서 본 논문에서는 코사인 유사도를 반 반영하였다(Fig 2.참조).

2.3 제품기획

제품기획이란 회사의 목표로써 개발해야 할 제품 개념을 구체적으로 제시하는 것을 말한다. 제품 개발의 절차로 보면 상품기획은 디자인하기 이전에 진행이 된다. 제품은 회사 비전을 달성하는 수단이며, 치열한 경쟁에서 살아남기 위해서는 판매율이 높은 제품이 반드시 필요하다. 기업의 제품기획 방향은 경영목표로 이어지고 다시 전략목표가 되고 최종적으로는 만들어야 할 제품으로 구체화 된다. 따라서 제품기획은 기업 전략목표 실행의 최전선으로 제품 개념을 구체적으로 만드는 행위이다. 방향이 잘못되면 모든 기획은 엉뚱한 쪽으로 진행이 된다. 결과적으로 중견기업, 중소기업이나 스타트업의 경우는 제품 하나가 기업의 운명을 결정 할 수 있다. 제품기획자는 제품개발 전 프로세스 뿐 아니라 제품 생애 주기 전반에 대한 통찰력을 가져야 한다. 제품이 출시되었다 하더라도 사용자의 사소한 반응이나 클레임 정보를 다각도로 분석하고 종합하여 차시 제품기획의 기초 자료로 활용해야 하기 때문이다. 제품기획자는 일반적으로 개발 업무 절차상 디자인 부서와 처음 업무 협의를 하게 된다. 주로 제품 콘셉트와 시각화를 위해서 디자인 협의하는 일과 제품기획의 특징이 디자인에 잘 적용되었는지 확인하는 게 주된 이슈이다[16].

3. 제안모델

3.1 모델 시스템 구성도

다음은 제안모델의 시스템 구성도(Fig 3. 참조)이다. 제안모델은 단어를 벡터로 표현하는 딥러닝기법인 Word2Vec기법을 활용하여 단어 간 유사도를 적용할 수 있도록 단어의 의미를 벡터화하고 이를 통해 분석하고자 하는 특정 단어와 가까운 거리에 있는 단어를 찾아내 장/단점을 분석 한 상품기획 방법을 제안한 것이다. 제안모델의 기본 구조는 Users, Client Web Application, Similarity Analysis System Server, Database등 4개 부분으로 구성되어 있다.

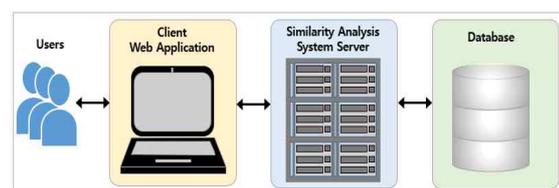


Fig. 3. Proposed Model Configuration

시각화된 화면에서는 상품기획자가 기본 검색 이외에 하부 검색 창에 공지사항, 환경설정, 유사도분석 등의 세부 확인창이 함께 제공되도록 편성되어 의미 있는 자료 검색이 이루어지도록 데이터셋이 관련된다. 사용자는 상품기획자와 상품기획 구성원들을 의미하며 회원가입을 통해 사용자들 간의 필요한 정보를 검색 할 수 있고 분석된 결과를 바탕으로 사용자들 간의 필요한 정보를 공유할 수 있다. 이러한 과정을 통해 가장 적합한 상품을 분석할 수 있는 시스템이 구현되는 구조이다.

3.2 시스템 아키텍처

다음은 제안시스템의 아키텍처(Fig 4. 참조)을 제시하였다. 제안시스템의 아키텍처는 그 역할에 따라 독립된 절차를 수행하는 레이어 및 스택구조로 구성되었으며 단계별로 생성되는 데이터 정보는 다음 단계에서 활용 가능한 형태로 변형된 5단계로 이루어져 있다.

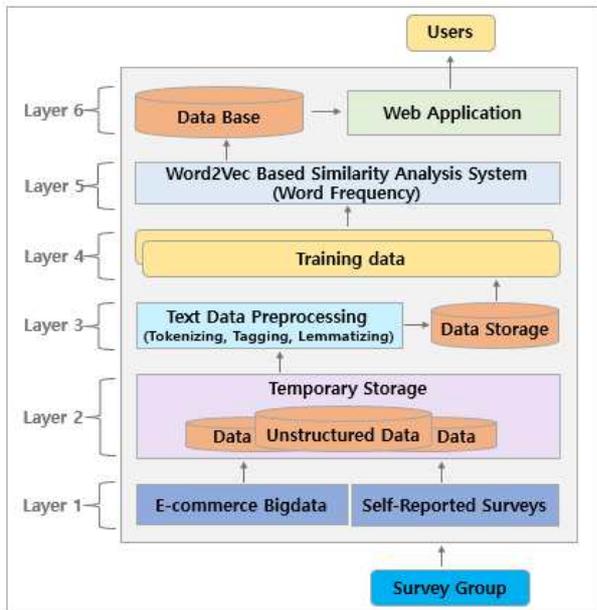


Fig. 4. Proposed System Architecture

첫째, 레이어 1에서는 웹 크롤링을 통해 빅데이터를 수집할 특정 상품군을 선택하여 대량의 자료를 수집하며 자가보고식 설문지 업로드 한다. 둘째, 레이어2에서는 레이어1에서 수집 저장된 비정형 데이터를 임시로 보관한다. 셋째, 레이어3에서는 레이어2에 임시 저장된 데이터를 유사도분석에 적합한 형태인 비정형데이터를 정형데이터로 변환한다. 넷째, 레이어4에서는 레이어3에서 정제 작업이 완료된 정형데이터를 통하여 Word2Vec 임베

딩 학습을 진행한다. 다섯째, 레이어 5에서는 입력받은 데이터를 통해서 제품기획을 위한 단어 간 유사도를 분석한다. 여섯째, 레이어 6에서는 웹 애플리케이션을 통하여 입력된 데이터의 분석 결과를 사용자에게 전달한다.

3.3 시스템 알고리즘

제안모델의 유사도의 정확도를 높이기 위한 알고리즘을(Fig. 5 참조) 제시 하였다. 정확도를 높이기 위해서는 한글, 영문, 숫자만 제시하고 나머지 요소들을 제거하는 과정이 포함되어야 한다. 상품평에 수집된 특수문자나 이모티콘 등이 상품에 대하여 긍정적이거나 부정적인 의미를 내포하는 공식화된 용어로 확정하는데 제한이 따르므로 이런 요소들은 제거하는 과정이 수행되었다. 또한 형태소 분석기로는 뉴스, 영화평, 상품명을 분리 추출, 처리 시간에 적합한 Okt(Open Korea Text)를 사용 하였다.

```
for sentence in online_reivew_list;
temp_X=okt.morphs(sentence, stem=True)
temp_X=[word for word in temp_X if not
word in stopwords]
tokenized_data.append(temp_X)
```

Fig. 5. Proposed Algorithm

다음은 제안모델의 유사도를 분석하기 위한 학습 알고리즘(Fig. 6)을 제시 하였다. 하이퍼파라미터값은 워드 벡터의 특징값으로 임베딩 된 벡터의 차원수를 지정하는 size는 100, 윈도우 크기를 지정하는 windows는 5, 빈도가 적은 단어들은 학습하지 않기 위해 min_count는 5로 설정하였다. 학습을 위한 프로세스 수를 지정하는 workers는 4, 학습의 정확도를 높이기 위해 Skip-gram 선택과 학습횟수는 10,000회로 지정 하였다.

```
model=Word2Vec(sentences=result, size=100,
windows=5, min_count=5, workets=4, sg=1,
iter=1000)
```

Fig. 6. Proposed Algorithm

4. 구현 및 결과

4.1 Online Service Evaluation

다음은 무소음 버티컬 마우스에 대한 온라인 사용자 상품평 10,000건을 적용한 개발환경(Table 1. 참고) 환경이다.

Table 1. General matters of service evaluation

Division	Description
Data	10,000 online user reviews
Target Product	Vertical wire/wireless anti-noise mouse
Development language	Python
Word Embedding	Word2Vec

다음은 온라인 상품평을 수집하여 단어 간 코사인 유사도를 통해 장점, 단점을 분석한 결과(Table 2. 참조)이다.

Table 2. Survey results

Division	Learning results			
	Strength		Weakness	
	Content	Similarity	Content	Similarity
Vertical	Vertical	1.000	Light	.952
	Pain	.985	Adaptation	.948
	Design	.963	Inconvenience	.883
	Wrist	.954	cost-effectiveness	.854
Click	Click	1.000	Hard	.941
	Anti-noise	.966	Game	.844
	Button	.965	Weight	.722
	Sensitivity	.933		

장점에 해당하는 첫 번째 키워드는 ‘통증’이었고 유사도 점수는 .985 이었다. 그 다음은 ‘디자인’ 이었고 유사도는 .963 이었으며 마지막은 ‘손목’ 이었으며 .954 이었다. 단점에 해당하는 첫 번째 키워드와 유사도 점수는 ‘가볍다’ 이었으며 .952 이었고 두 번째는 ‘적응’ 이었으며 .948 이었다. 그 다음은 ‘불편’과 ‘가성비’ 이었으며 유사도 점수는 .883, .854 이었다.

4.2 Off-line Service Evaluation

다음은 무소음 버티컬마우스에 대한 자가보고식 설문지를 컴퓨터 사용량이 많은 대학생 92명을 대상으로 진행하였다. 사용기간을 5일로 설정, 사용 후기를 서술식으로 제출하는 방식과 제시된 단어를 보고 장점, 단어에 유사한 단어를 선택하도록 온라인으로 수집 하였다 (Table 3. 참조).

Table 3. General Matters of Service Evaluation

Division	Content
Survey Object	92 College Students
Target Product	Vertical Wired and Wireless Noiseless Mouse
Use Period	5 Days
Survey Method	Online Form
Survey Form	Self-Reported Survey

다음은 자가보고식 설문지를 서술식으로 수집하여빈도분석한 결과(Table 4. 참조) 이다. 상위랭크된 단어는 ‘버티컬’, ‘통증’, ‘무소음’ 이었으며 각각 123개, 118개, 86개의 단어가 사용 되었다.

Table 4. Survey Results

Division	Frequency Number	Frequency(%)
Vertical	123	134%
Pain	118	128%
Anti-noise	86	93%
Adaptation	82	89%
Wrist	70	76%
Design	67	73%
Button	67	73%
Hard	55	60%
Cost-effectiveness	50	54%
Inconvenience	42	46%
Sensitivity	41	45%
Game	38	41%
Vdt	35	38%

Table 5. Survey Results

Division	Survey results	
	Description	Select Amount
Strength	Vertical	83
	Pain	75
	Anti-noise	68
	Wrist	58
Weakness	Adaptation	89
	Button	72
	Hard	67
	Cost-effectiveness	45

다음은 사전에 학습된 온라인 리뷰 10,000건 중 상위 50개 단어를 제시 후 장점에 유사한 단어인지, 단점에 유사한 단어인지 설문대상자가 중복선택 할 수 있도록 하

여 수집 하였다. 설문에 참여한 대학생 92명이 선택한 결과로(Table 5. 참조)는 장점에 해당하는 첫 번째 키워드는 '버티컬' 이었고 선택 수는 83명 이었다. 그 다음은 '통증' 이었고 선택 수는 75명 이었으며 마지막은 '무소음' 이었으며 68명 이었다. 단점에 해당하는 첫 번째 키워드는 '적응' 이었고 선택 수는 89명 이었다. 그 다음은 '버튼' 이었고 선택 수는 72명 이었으며 마지막은 '힘듦'이 이었으며 선택 수는 67명 이었다.

5. 결론

다음은 본 연구에서는 포털사이트와 전자상거래에서 버티컬마우스 구매자에 대한 사용자 상품평을 수집하여 단어간 코사인 유사도 분석과 자가보고식 설문지를 통해 단어빈도분석, 단어를 제시하고 복수선택하는 방법을 통해 장점, 단점을 분석 진행 하였다. 유사도 분석결과 버티컬 마우스의 장점에 해당하는 키워드는 '통증', '디자인', '손목' 이었으며 유사도 점수는 각각 .985, .963, .954로 1에 가까워 신뢰할 만한 수준을 나타내었다. 또한 자기보고식 단어빈도 분석결과 상위 랭크된 단어는 '버티컬', '통증', '무소음' 이었으며 각각 123개, 118개, 86개가 사용되었으며 단어를 제시 후 선택하는 설문결과로는 장점에 '버티컬', '통증', '무소음'이 있으며 각각 83명, 75명, 68명이 선택하였다. 이는 포털사이트와 전자상거래에서 수집한 사용자 상품평과 자가보고식 설문조사 방식과 유사한 결과가 도출 되었으므로 포털사이트와 전자상거래에서의 사용자리뷰는 유의미한 자료라 증명할 수 있다. 마우스의 장시간 사용은 손목터널증후군의 증세 중에 해당하는 통증을 유발한다. 이러한 문제를 해결하기 위해 여러 형태의 마우스가 개발되어 상품화되고 있으며 특히 버티컬 마우스는 손목 통증을 감소시키고자 개발된 상품으로 알려져 있다. 본 연구에서 확인된 결과를 통해 버티컬 마우스가 기존 일반마우스의 사용에 비해 손목 통증이 적고 디자인에 대한 긍정적인 피드백도 제시되고 있음을 추론할 수 있었다. 포털사이트와 전자상거래에서의 단점에 해당되는 키워드는 '가볍다', '적응', '불편', '가성비' 이었으며 유사도 점수는 각각 .952, .948, .883, .854 이었다. 이러한 결과를 통해 버티컬 마우스는 손에 쥐었을 때 가볍게 느껴지고 사용방법을 익히기 위한 적응이 요구되어 불편할 수 있으며 가격 대비 상품의 질이 다소 저하된다는 소비자의 의견을 추론할 수 있었다. 본 연구에서 제시된 키워드는 긍정보다는 부정의 의미를 가진

단어가 다소 많이 제시되었지만 사용자리뷰에 대한 유사도 분석을 통해 소비자들의 상품에 대한 솔직한 의견을 파악할 수 있다는 점에서 상품 개발 및 개선에 유의미한 정보를 제시하였다고 볼 수 있다. 따라서 본 연구에서 제시하는 상품기획 프로세스를 신상품 개발 및 기존 상품의 품질개선 전략으로 적용할 수 있겠다. 본 연구에서 제시한 구현을 통해 다양한 상품에 대한 피드백을 검증하는 후속 연구가 필요하다. 더 나아가 Collaborate, Content-based filtering 추천시스템과-딥러닝 조합모델 구현을 통한 상품기획 절차 연구도 제안한다.

REFERENCES

- [1] Statistics Korea, Trend of online shopping[Internet] Seoul: Statistics Korea [cited 2024 August 14]. Available from: https://kostat.go.kr/portal/korea/kor_nw/1/1/index.board?bmode=read&aSeq=391169
- [2] Q. Y. Hao, S. J. Lee & K. R. Lee. (2015). The Acceptance of Customer Reviews in Taobao. *Journal of the Korea Convergence Society*, 6(4), 205-212. DOI:10.15207/jkcs.2015.6.4.205
- [3] D. J. Kim, D. in. Park & J. S. Park. (2018). Study on the Change of Marketing Strategy through Data Mining Technique. *Korea Business Review*, 22(2), 177-194. DOI:10.17287/kbr.2018.22.2.177
- [4] D. Lee, J. C. Jo & H. S. Lim (2017). User Sentiment Analysis on Amazon Fashion Product Review Using Word Embedding. *Journal of the Korea Convergence Society*, 8(4), 1-8. DOI:10.15207/jkcs.2017.8.4.001
- [5] S. Lim, H. Lee & Y. Yoon (2020). Prediction of New Drug-Side Effect Relation using Word2Vec Model-based Word Similarity. *The Journal of Korean Institute of Information Technology*, 18(11), 25-33. DOI:10.14801/jkiit.2020.18.11.25
- [6] D. Y. Kim & Y. G. Lee. (2018). News based Stock Market Sentiment Lexicon Acquisition Using Word2Vec. *The Korea Journal of BigData*, 3(1), 13-20. DOI:10.36498/kbigdt.2018.3.1.13
- [7] J. N. Yun, G. H. Kim. & D. K. Ryoo. (2018). A Study on Early Warning Model in the Dry Bulk Shipping Industry by Signal Approach. *Journal of Navigation and Port Research*, 42(1), 57-66.
- [8] J. E. Ahn, Y. J. Kim, H. S. Kim, W. J. Kim, & Lee, Y. (2017). A study on medical herb recommendation system using word2vec. In *Proceedings of the Korean Society of Computer Information Conference*. 83-85
- [9] H. Kang. & J. Yang (2019). Optimization of Word2vec Models for Korean Word Embeddings. *Journal of*

Digital Contents Society, 20(4), 825-833.
DOI:10.9728/dcs.2019.20.4.825

- [10] T. Mikolov, K. Chen, G. Corrado & Dean, J. (2013). Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*.
- [11] K. Heo, J. Yang, D. Kim, K. Bok & J. Yoo. (2019). Design and Implementation of Minutes Summary System Based on Word Frequency and Similarity Analysis. *The Journal of the Korea Contents Association*, 19(10), 620-629.
- [12] Y. H. Ahn & K. R. Park. (2021). Product Planning using Similarity Analysis Technique Based on Word2Vec Model. *In Proceedings of the Korean Society of Computer Information Conference* 11-12.
- [13] J. C. Yun, H. C. Kim, J. J. Kim. & S. D. Youn. (2010). A Study on the CBR Pattern using Similarity and the Euclidean Calculation Pattern. *The Journal of the Korean Institute of Information and Communication Engineering*, 14(4), 875-885.
DOI:10.6109/jkiice.2010.14.4.875
- [14] Y. S. Jeon (2020). *Management for Reducing Emissions in the Seoul Metropolitan Area during Heat Wave Episodes using CMAQ*. Master's dissertation. KonKuk University, Seoul.
- [15] J. W. Choi & D.-K. An (2016). A Study on the Data Analysis of the Written Comments in Lecture Evaluation. *Journal of Digital Convergence*, 14(11), 101-106.
DOI:10.14400/jdc.2016.14.11.101
- [16] L. H. Chen, W. C. Ko. & F. T. Yeh. (2017). Approach based on fuzzy goal programing and quality function deployment for new product planning. *European Journal of Operational Research*, 259(2), 654-663.

안 영 휘(Hwi-Yeong Ahn)

【장학원】



- 2005년 2월 : 국립한밭대학교 컴퓨터 공학과(공학사)
- 2007년 2월 : 국립한밭대학교 컴퓨터 공학과(공학석사)
- 2019년 3월 ~ 현재 : 국립공주대학교 일반대학원 컴퓨터공학과(박사수료)
- 관심분야 : 인공지능, 데이터 사이언스,

IT 컨버전스, 컴퓨터비전, 상품기획

· E-Mail : yhahn02@naver.com