

# 헬스케어 서비스 리뷰를 활용한 서비스 품질 차원 별 중요 단어 파악 방안<sup>1</sup>

## Keyword identifications on dimensions for service quality of Healthcare providers

이 홍 주 (Hong Joo Lee) 가톨릭대학교 경영학부 교수<sup>2</sup>

### ABSTRACT

Studies on online review have carried out analysis of the rating and topic as a whole. However, it is necessary to analyze opinions on various dimensions of service quality. This study classifies reviews of healthcare services into service quality dimensions, and proposes a method to identify words that are mainly referred to in each dimension. Service quality was based on the dimensions provided by SERVQUAL, and patient reviews have collected from NHSChoice. The 2,000 sentences sampled were classified into service quality dimension of SERVQUAL and a method of extracting important keywords from sentences by service quality dimension was suggested. The RAKE algorithm is used to extract key words from a single document and an index is considered to consider frequently used words in various documents. Since we need to identify key words in various reviews, we have considered frequency and discrimination (IDF) at the same time, rather than identifying key words based only on the RAKE score. In SERVQUAL dimension, we identified the words that patients mentioned mainly, and also identified the words that patients mainly refer to by review rating.

*Keywords: SERVQUAL, Keyword extraction, Healthcare, Customer review*

<sup>1</sup> 이 논문은 2017년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임 (NRF-2017S1A5A2A01025690).

논문접수일: 2018년 9월 30일; 1차 수정: 2018년 11월 21일; 게재확정일: 2018년 11월 21일

<sup>2</sup> 제 1저자 (hongjoo@catholic.ac.kr)

## 1. 서론

고객 연구를 위한 디지털 텍스트의 양이 급증하였다. 인터넷 게시판을 통한 토론, 제품 리뷰, 신문 기사, 댓글 등이 고객의 제품과 서비스에 대한 태도 및 중요 속성을 파악하거나 고객들의 행태를 파악하기 위한 주요한 원천이 되었다(박명석 등 2018; 크루즈·이홍주 2014; Humphreys and Wang 2018). 텍스트 마이닝 기반의 고객 리뷰 분석을 통해 제품의 기능이나 속성이 제품 평가에 미치는 영향을 파악하거나(Cao et al. 2011; Humphreys and Wang 2017), 속성이나 기능에 해당하는 단어들을 정하고 각 단어들에 대한 감성분석을 수행하여 분석하였다(Cao et al. 2011; Ghose et al. 2012). 유사한 제품들의 시장 경쟁 상황을 가시화하거나 랭킹을 부여하는 방안도 제시되었다(Chen et al. 2015; Netzer et al. 2012). 탐색제나 경험제에 대한 고객 리뷰가 분석되어 고객에 대한 이해를 높인 것과 마찬가지로, 헬스케어 서비스와 같은 신뢰재(김상만·이연주 2010; 김상만 등 2010; Dulleck and Jerschbamer 2006)에 대한 리뷰를 통해 환자들이 경험한 서비스에 대해 분석할 필요가 있다.

Greaves et al.(2012)은 트위터와 블로그의 많은 메시지와 리뷰는 새로운 의료 서비스 품질에 대한 자료를 제공하며, 환자 중심의 의료를 앞당길 수 있는 기회로 여기고 있다.

트위터에서 의료 서비스 경험에 대한 트윗을 수집하여 내용을 진료 품질, 홍보, 후원모금, 의료 정보, 의료기관 정보, 환자 후원으로 구분하여 분석하였다(Greaves et al. 2014). 트윗의 감성분석 순위와 설문문을 통한 의료기관의 환자경험 점수 순위와의 상관관계는 0.14(p-value = 0.3)이었고, 사망률과의 상관계수는 0.15(p-value = 0.24)로 분석되어 유의한 관계를 보이지는 않았다.

현재는 온라인 리뷰에 대해 평점별 분석이나 텍스트

전체를 활용한 토픽 분석 등의 연구들이 수행되어 왔으나, 다양한 서비스 품질 차원에 대한 의견을 분석할 필요가 있다. 헬스케어 서비스에 대한 의견을 분류하여 분석한 연구(Greaves et al. 2014)들이 있었지만, 서비스 품질을 체계적으로 구분하고 서비스를 포괄하는 하위 차원으로 분류하여 연구하지는 않았다. 서비스 품질에는 다양한 차원들이 존재하므로, 다양한 서비스 차원별로 환자들이 어떠한 사항들을 중요하게 언급하고 있는지를 파악하는 것이 필요하다.

서비스 품질을 측정하기 위한 도구를 만들기 위한 많은 연구들이 진행되어 왔으며, SERVQUAL(Parasuraman et al. 1985, 1988)은 다양한 분야의 산업 및 영역에서 서비스 품질을 측정하는 데 사용되었다(Ladhari 2009). 의료 분야에서도 SERVQUAL을 통한 서비스 품질 측정에 대한 연구가 많이 진행되어왔다(Akter et al. 2010; Babakus and Mangold 1992; Dagger et al. 2007; Dean 1999; Lam 1997). 보건 의료 산업에 대한 SERVQUAL(Lee et al. 2000)의 확장을 제안하고 다양한 측정 방법을 제시하였다 (Park and Agarwal 2016; Sofaer and Firminger 2005). 설문을 통한 서비스 품질을 측정하여 의료 기관의 다른 성과와 비교하는 연구들이 진행되었으며, 설문을 통해 헬스케어 서비스에 맞는 추가 서비스 품질 영역을 찾는 연구들도 많이 진행되었다.

설문이 아니고 소셜 미디어 데이터와 온라인 리뷰를 분석하여 고객이 언급한 내용이 서비스 품질의 어떤 영역에 해당하는지 파악하는 연구들도 많이 진행되었다. Palese and Usai(2018)은 전자제품에 대한 고객 리뷰를 토픽 모델링을 통해 SERVQUAL 차원으로 파악하는 방안을 제안하였다.

본 연구는 헬스케어 서비스에 대한 리뷰를 서비스 품질 차원으로 분류하여 각 차원별로 주요하게 언급되는 단어를 파악하는 방안을 제안한다. 서비스 품질의 차원으로는 SERVQUAL(Parasuraman et al. 1985,

1988)에서 제시한 차원들을 활용하였으며, 영국 의료 보험인 NHS를 활용한 환자들이 자신들이 경험한 서비스에 대한 리뷰를 남기는 NHSChoice 사이트의 리뷰를 수집하였다. 샘플링을 통해 936개의 리뷰를 선정하였으며, 리뷰를 문장단위로 분할하여 대략 2,000개의 문장을 확보하였다. 각 문장을 SERVQUAL의 서비스 품질 차원에 할당하였으며, 서비스 품질 차원별로 할당된 문장들에서 주요한 키워드를 추출하는 방안을 제시하였다. 하나의 문서에서 주요한 단어를 추출하는 RAKE 알고리즘을 활용하였으며, 다양한 문서에서 빈번히 활용되는 단어들을 고려하기 위한 지표를 고려하였다. SERVQUAL 차원별로 고객들이 주요하게 언급한 단어들을 파악하여 분석하였으며, 마찬가지로 리뷰 평점별로 고객들이 주요하게 언급한 단어들을 파악하여 분석하였다.

2장에서는 관련 연구를 정리하였으며, 본 연구에서 활용한 데이터와 연구방안에 대한 설명은 3장에 기술하였다. 4장에는 3장에서 제안한 방안을 통해 파악된 서비스 품질 차원별 중요 단어에 대해 분석하였다. 5장에서는 분석결과에 대해 토의하고 연구의 한계를 논하였다.

## 2. 관련 연구

### 2.1 서비스 품질

서비스 품질은 경쟁 업체와 차별화하려는 조직의 노력에서 가장 중요한 성공 요인 중 하나이다(Ladhari 2009). Zeithaml(2000)은 서비스 품질과 이익 간의 관계에서 직접적(수익성 향상), 공격적(신규 고객 확보), 방어적(고객 유지), 구매 의도 증가, 고객 세그먼트 구별 및 중요 요인 식별을 중요한 여섯 가지 요인으로 파악하였다.

서비스에 대한 고객의 인식을 측정하기 위한 다양한

측정 지표가 제시되어왔다(Cronin and Taylor 1992; Grönroos 1984; Lehtinen and Lehtinen 1982; Parasuraman et al. 1985). 그중에서 SERVQUAL(Parasuraman et al. 1985; 1988)은 서비스 품질 측정에서 가장 잘 알려졌으며, 일반적으로 사용되는 척도이다(Palese and Usai 2018). 몇 가지 업데이트를 한 후, SERVQUAL은 서비스 품질 측정 차원으로 유형(Tangibles, 물리적 시설, 장비 및 인원의 외양), 신뢰성(Reliability, 약속 된 서비스를 신뢰성 있고 정확하게 수행 할 수 있는 능력), 응답성(Responsiveness, 고객을 도우려하고 즉각적인 서비스를 제공하려는 의지), 보증(Assurance, 직원의 지식과 확신을 불리일으키는 능력), 그리고 공감(Empathy, 회사가 고객에게 제공하는 보살핌과 개별적인 주의)을 제안하였다.

유효성, 신뢰성 및 적용 가능성을 포함하여 SERVQUAL의 다양한 측면에 대한 논쟁이 있어왔으며(Ladhari 2009), Cronin and Taylor(1992)는 고객의 인식만을 측정하는 경쟁 SERVPERF 모델을 제안하였다. 서비스 품질 측정 차원에 대한 논의는 이 연구의 목표를 넘어선 것이며, 본 연구는 대표적인 서비스 품질 측정 차원인 SERVQUAL을 사용하였다.

### 2.2 중요 단어 파악 방안

문서에서 중요한 키워드를 파악하는 수동적인 방안은 분야 전문가들이 키워드 체계(Taxonomy)를 활용하여 할당하거나 저자들이 키워드로 제시한 것을 활용하여 왔다. 문서에서 자동으로 키워드를 자동으로 파악하는 방안은 기본적으로 단어의 빈도를 고려하여 왔다. 대표적으로 TF-IDF 지수와 KeyGraph(Ohsawa et al. 1998)가 단어 빈도와 단어가 함께 출현하는 정도를 고려하여 키워드를 자동으로 추출하였다. Andrade and Valencia(1998)은 특정 문서의 단어 빈도를 비교 대상이 되는 문서들의 단어 빈도와 비교하여 문서별로

중요한 키워드를 추출하는 방안을 제시하였다. 이러한 접근방안은 여러 문서에 등장하는 단어들은 문서를 구분하는 데 영향을 덜 미치기 때문에, 키워드로 선정되지 못하는 경우가 많았다. 또한 단어를 활용하는 방안들은 하나의 단어들을 기준으로 분석하기 때문에 다양한 상황에서 사용되는 단어들은 중요 키워드로 뽑히지 못하는 경우가 많았다.

이러한 문제점을 극복하기 위해 여러 문서의 비교를 통한 키워드 자동 추출이 아니라 하나의 문서에서 키워드를 추출하는 방안들이 제시되어왔다. 제시된 방안들은 대부분 자연어 처리와 품사 태그를 활용하여 지도학습, 기계학습 알고리즘과 통계 분석방안을 적용하였다(Rose et al. 2010).

Hulth(2003)은 단어를 추출하는 NP chunks, n-grams, POS tags 방안이 키워드 추출 성과에 미치는 영향을 비교하였다. Mihalcea and Tarau(2004)는 단어가 함께 출현하는 빈도를 활용하여 키워드를 추출하는 필터링 방안을 제안하였으며, 그래프 기반의 랭킹 알고리즘인 TextRank를 제시하였다. 동시 출현을 기준으로 그래프를 작성하여 연관정도를 기준으로 순위를 매겨, 높은 순위를 단어들을 키워드로 추출하였다.

RAKE(Rose et al. 2010)는 도메인과 언어에 무관한 비지도 학습방안으로 개별 문서별로 중요 키워드를 추출하는 방안이다. RAKE는 키워드가 대부분 여러 단어로 구성되며, stop words들은 거의 포함되지 않는다는 관찰에서 출발하였다. 문서에서 stop words를 제외한 단어들을 추출하여 문장단위에서 단어 간의 동시 출현 빈도를 계산한다. 동시 출현 그래프를 통해 출현 빈도, 연결 정도, 연결 정도/출현 빈도의 세 가지 스코어를 계산한다. 여러 단어로 구성된 키워드는 각 단어들의 스코어를 합하여 최종 스코어를 계산한다. 기본적으로 추출된 콘텐츠 단어의 1/3에 해당하는 수만큼의 키워드를 추출한다. RAKE가 TextRank나 Hulth(2003)이 제안한 방안보다 Precision과 Recall을 함께 고려하는 F1

지표의 값이 더 우수하였다. Word2Vec이나 Doc2Vec은 단어들간의 유사도나 문서간의 유사도 파악에 중점을 두고 있으며, 문서에서 중요한 단어파악을 통해 문서를 이해하는 방안과는 다른 접근 방안이다. 본 연구에는 RAKE 알고리즘을 활용하여 리뷰별로 중요 키워드를 추출하고, 여러 리뷰에서 추출되는 키워드를 고려하여 서비스 품질 차원별 키워드 리스트를 제안하는 방안을 제시하였다.

### 3. 연구 방법 및 데이터

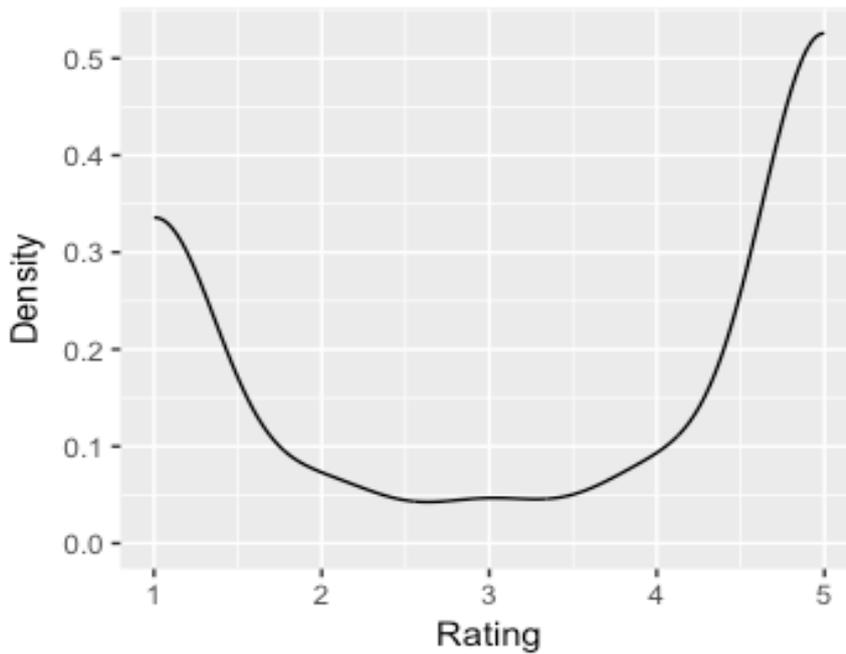
#### 3.1 연구 데이터

영국의 NHSChoice 사이트에서 헬스케어 서비스 경험에 대한 고객 리뷰를 수집하였다. NHS는 영국의 공적 의료보험 서비스이며, NHSChoice는 GP(General Practice)와 병원(Hospital)에 대한 정보제공과 고객들의 의료서비스 경험을 공유하는 공간이다.

GP와 병원을 무작위로 총 7,838개를 선정하였으며, 해당 GP와 병원에 대한 고객 리뷰를 수집하였다. 수집된 리뷰는 총 132,636개이며, 평균 평점은 5점 만점에 3.380이고 표준편차는 1.818이다. <그림 1>은 평점별 고객 리뷰 분포에 대한 밀도(density) 그래프이다. 전자상거래에서 평점에 따른 리뷰 분포는 J 모양을 따르는 것으로 알려져 있으나(Hu et al. 2009), 수집된 리뷰의 분포는 일반적인 J 모양 보다는 5점과 1점으로 양분된 형태를 띠고 있다.

<표 1>은 수집된 전체 리뷰에서 가장 빈번하게 등장하는 30개 단어이다. 단어를 파악할 때 어간추출(stemming)을 수행하였기에 surgeri, practic, nurs 같은 어간의 형태로 표기되었다.

헬스케어 제공자에대한 리뷰이므로 당연히 doctor, staff, nurs, patient, receptionist 같은 역할에 대한 단어가 빈번히 등장하고, surgeri, care, gp, practic 같은



<그림 1> 평점에 따른 리뷰 분포 (Note. Rating은 연속형 변수가 아니고 명목 변수이며, 밀도 그래프의 시각화에 있어서 옵션을 주어 각 평점의 빈도 밀도 값이 부드럽게 연결되도록 하였다.)

<표 1> 가장 빈번히 등장하는 단어 30개

doctor	appoint	surgeri	get	staff
time	always	practic	care	help
patient	gp	need	see	one
servic	year	nurs	receptionist	day
can	wait	call	recept	phone
go	good	never	told	week

의료 기관이나 행위에 대한 단어들이 빈번히 출현한다. 예약을 위한 appoint, phone, call, wait이 빈번히 등장하여 예약과 관련된 고객 경험이 많은 것으로 유추할 수 있다.

<표 1>이 하나의 단어를 기준으로 가장 빈번히 등장하는 단어들을 보여주었다면, <표 2>는 연속된 두 단어의 쌍을 기준으로 가장 빈번히 등장하는 단어쌍을 보여준다.

“get appoint”, “book appoint”, “make appoint”, “call back”, “recept staff”, “appoint doctor”, “tri get”, “appoint day”, “appoint doctor” 와 같은 진료 예약을 잡는 것에 대한 단어쌍이 많았으며, “see doctor”, “see gp”, “doctor nurs”, “staff always”, “doctor surgeri”, “doctor always”, “gp surgeri”와 같은 진료 행위에 대한 단어쌍도 빈번하게 등장하였다.

수집한 전체 데이터 중에서 936개의 리뷰를 무작

<표 2> 가장 빈번히 등장하는 단어쌍 30개

get appoint	recept staff	see doctor	book appoint	make appoint
doctor nurs	call back	staff always	tri get	blood test
doctor surgeri	see gp	surgeri year	appoint day	doctor always
repeat prescript	recommend surgeri	abl get	mani year	one doctor
friend help	gp surgeri	made feel	never get	always found
member staff	medic centr	staff help	appoint doctor	staff friend

위로 선정하여 리뷰를 문장으로 나누었다. 문장은 “:”, “?”, “!” 기호가 등장하는 경우를 기준으로 나누었으며, 나누어진 문장의 총 수는 2,035개이다. 문장을 감성에 따라 매우 긍정, 긍정, 중립, 부정, 매우 부정으로 나누었으며, 또한 관련된 SERVQUAL의 한 차원에 할당하였다. 관련된 차원이 없는 경우에는 기타로 분류하였으며, 문장의 분류에는 SERVQUAL에 대해 이해하고 있는 영어가 모국어인 석사학위 소지자와 박사학위 과정생이 참여하였다. 문장 분류는 분류 목적에 대해 소개한 후에 분류자가 각각 문장에 대해 감성과 SERVQUAL 차원으로 분류하였다. 분류된 결과를 가지고 의논하여 상이한 분류 결과를 동일하게 재분류하였다. <표 1>이 문장을 감성과 SERVQUAL 차원으로 분류한 결과의 빈도표이다.

<표 3> SERVQUAL 차원에 따른 문장 빈도

SERVQUAL 차원	빈도
Assurance	112
Empathy	288
Reliability	487
Responsiveness	296
Tangibles	97
Other	755

### 3.2 연구 방법

본 연구는 RAKE 알고리즘(Rose et al. 2010)을 활용하여 하나의 리뷰나 문장에 포함된 단어의 중요도를 계산하였다. R의 slowraker 패키지(Baker 2017)를 이

용하여 RAKE 스코어를 계산하였다. 중요도 계산에서 POS(Part-of-Speech) 태그가 동사(VB, VBD, VBG, VBN, VBP, VBZ)와 관련된 단어는 제외하였으며, 3글자 이상의 단어만 계산에 포함하였다. RAKE 스코어는 하나의 리뷰나 문장에서의 중요도를 계산하기 때문에 한 단어가 여러 리뷰나 문장에 포함된 경우에는 각 리뷰나 문장에서 계산된 스코어의 평균을 해당 단어의 중요도 스코어로 삼았다.

SERVQUAL 차원으로 구분된 문장을 모아서 차원별로 RAKE 스코어를 계산하였다. 한 단어의 RAKE 스코어가 리뷰나 문장별 스코어의 평균이기 때문에, 해당 단어의 등장 빈도를 고려하여 RAKE 스코어 \* 등장 빈도(Term frequency) 값을 구하였다. RAKE 스코어는 여러 단어로 구성된 경우 더 큰 값을 갖기 때문에, 짧은 단어의 경우에는 등장 빈도를 고려하여 보정하는 것이 필요하다. 또한, TF-IDF 지표(Baeza-Yates and Ribiero-Neto 2010)와 동일하게 RAKE 스코어 \* TF-IDF 값을 계산하였으며, 해당 단어가 너무 많은 리뷰나 문장에 등장하는 경우에는 문장이나 리뷰를 구분하는데 미치는 영향이 작기 때문에 이를 보정한다고 볼 수 있다.

<표 4>는 SERVQUAL 차원 중에서 Assurance에 해당하는 문장에서 세 가지 방안으로 추출한 중요단어이다. 단어의 단위는 가지치기(stemming)된 어근을 활용하였다.

RAKE 스코어는 기본적으로 많은 단어로 이루어진

<표 4> 계산 방안별 중요 단어

RAKE 스코어 기준 상위 단어	빈도(Term Frequency)	RAKE 스코어	
friendli help respect understand	1	16	
high depend clinic care	1	16	
absolut brilliant servic	1	9	
compassion nurs staff	1	9	
elect care centr	1	9	
fist nearest pharmacist	1	9	
health centr manageress	1	9	
small fee payabl	1	9	
son health problem	1	9	
total knee replac	1	9	
RAKE 스코어 * TF 기준 상위 단어	빈도(Term Frequency)	RAKE 스코어	RAKE 스코어 * TF
doctor	30	1.016	30.5
profession	20	1.025	20.5
staff	16	1.041	16.6666667
friendli help respect understand	1	16	16
high depend clinic care	1	16	16
practic manag	3	4	12
consult	11	1	11
absolut brilliant servic	1	9	9
compassion nurs staff	1	9	9
elect care centr	1	9	9
RAKE 스코어 * TF-IDF 기준 상위 단어	빈도(Term Frequency)	RAKE 스코어	RAKE 스코어 * TF-IDF
friendli help respect understand	1	16	99.7505716
high depend clinic care	1	16	99.7505716
doctor	30	1.025	86.413007
profession	20	4	66.3929083
practic manag	3	1.041	61.6295812
staff	16	9	57.6970334
absolut brilliant servic	1	9	56.1096965
compassion nurs staff	1	9	56.1096965
elect care centr	1	9	56.1096965
fist nearest pharmacist	1	9	56.1096965

경우에 스코어 값이 커지기 때문에 RAKE 스코어만으로는 여러 단어로 구성된 단어들이 많이 선택되었다. RAKE 스코어에 단어의 등장 빈도를 곱한 경우에는 여러 단어로 구성된 중요 단어가 선택되는 경우가 줄었으며, 빈번히 등장하는 단어들이 포함되었다. RAKE 스코어에 TF-IDF 값을 곱한 경우에는 등장 빈도가 높은 단어들의 스코어가 좀 줄고 RAKE 스코어가 높고 많

이 등장하지 않는 단어들이 포함되었다. TF, TF-IDF를 고려한 경우가 그렇지 않은 경우보다 중요단어로 선택된 리스트들이 의미를 파악하는 데 효과적으로 보이며, TF, TF-IDF는 순위에 약간의 차이가 있으나 큰 변화는 없었다.

### 4. 분석 결과

<표 5>는 리뷰 평점별 중요 키워드 30개이다. 순위는 RAKE 스코어 \* TF-IDF 값을 기준으로 정하였다.

평점 1부터 4까지 상위 5위까지의 단어는 순서는 상이하지만 appoint, doctor, receipt staff, time, surgeri이다. 평점 5의 경우에도 staff가 4위를 차지한 것을 제외하고는 1위부터 6위까지 동일한 단어였다. doctor와 surgeri(수술의 뜻도 있지만 병원의 이름으로 빈번하게 사용됨)는 리뷰에 빈번하게 등장하기 때문에 상

위 순위에 위치하고 있으며, 예약(appoint), 접수 직원(recept staff)과 시간(time) 또한 상위 순위에 위치하고 있다. 평점과 무관하게 예약, 응대 직원, 시간이 리뷰의 주요한 내용임을 알 수 있다. 예약을 잡고 의사를 만나러 왔다는 표현이 많이 쓰이기도 하고, 예약 잡기가 어렵다는 사항이 주요 불만사항이기도 하다. 응대 직원의 친절여부가 좋은 평점을 주느냐 그렇지 않느냐를 좌우하기도 한다. 평점 1점의 경우에는 접수 직원(recept staff)와 함께 순위 7위와 10에 위치한 접수 직원(receptionist)과 무례함(rude)이 함께 많이 사용되었다.

<표 5> 평점별 중요 단어

순위	평점 5	평점 4	평점 3	평점 2	평점 1
1	recept staff	recept staff	recept staff	appoint	appoint
2	doctor	doctor	appoint	recept staff	doctor
3	surgeri	appoint	doctor	doctor	recept staff
4	staff	surgeri	time	surgeri	surgeri
5	appoint	time	surgeri	time	time
6	time	staff	patient	dai	patient
7	care	patient	receptionist	patient	receptionist
8	year	help	dai	week	week
9	practic	practic	week	receptionist	dai
10	patient	year	phone	phone	rude
11	help	nurs	practic	practic	phone
12	nurs	receptionist	staff	repeat prescript	practic
13	friendli	dai	nurs	year	staff
14	excel servic	care	year	staff	year
15	servic	problem	good	rude	repeat prescript
16	dai	good	repeat prescript	nurs	emerg appoint
17	profession	servic	problem	book	back
18	receptionist	blood test	appoint system	appoint system	blood test
19	problem	friendli	rude	blood test	book
20	excel care	good servic	book	back	hour
21	hospit	week	help	problem	practic manag
22	treatment	phone	blood test	emerg appoint	nurs
23	famili	emerg appoint	emerg appoint	good	peopl
24	good work	repeat prescript	servic	servic	servic
25	excel	happi	peopl	peopl	problem
26	kind	profession	wait room	hour	phone call
27	year ago	excel servic	back	wait room	month
28	good	dai appoint	recept	appoint time	minut
29	blood test	practic nurs	care	minut	wait room
30	practic nurs	year ago	phone call	recept	extrem rude

평점 2와 평점 3에서도 두 단어가 각각 20위안에 위치하고 있지만, 평점 4, 5에서는 접수 직원(receptionist)은 20위 내에 위치하지만 무례함(rude)은 30위 내에 존재하지 않았다. 시간(time) 또한 일(dai), 주(week), 시(hour), 분(minut)와 함께 평점 1, 2 리뷰에 많이 등장하였다. 예약이나 진료를 위해 기다린 시간을 표현하기 위해 사용된 단어로 보인다. 평점 3 리뷰에는 시와 분은 등장하지 않았으며, 평점 4에서는 일과 주만 등장하였고, 평점 5에서는 일만 등장하였다. 평점이 낮은 리뷰의 많은 불만은 대기 시간에 대한 것으로 보인다. 특히 전화로 예약하는 것이 오래 걸리는 불만이 많아 전화(phone)이라는 단어도 평점 1, 2, 3인 리뷰에만 상위 30위 안에 등장하였다. 오래 걸려 겨우 접수 직원과 통화를 하였는데 접수 직원이 불친절 하거나 예약을 상당

히 오랜 기간 후에나 잡을 수 있는 경우가 불만의 많은 부분을 차지하였다.

반대로 평점 5, 4 리뷰에는 care, excel care, help, friendli, famili, excel servic, good work, excel, kind, good과 같은 감성적인 단어들 많이 등장하였다. 친절하고 친밀한 서비스와 간호에 감사하는 내용이 많이 표현되어 있다. 평점 3에는 위 단어중에서 help와 care만 순위에 등장하였으며 평점 1, 2 리뷰에는 이러한 단어들 순위에 포함되지 않았다. 평점을 기준으로 구분한 리뷰에서는 긍정적인 요소와 불만사항을 쉽게 파악할 수 있었다.

<표 6>는 SERVQUAL 차원별 중요 키워드 30개이다. 순위는 RAKE 스코어 \* TF-IDF 값을 기준으로 정하였다.

<표 6> SERVQUAL 차원별 중요 단어

Assurance	Empathy	Reliability	Responsiveness	Tangibles	Others
friendli help respect understand	ey dept emerg phone line yesterdai	doctor	emerg appt fridai morn 8am final	room	hospit
high depend clinic care	recept staff	staff	appoint	appoint system	recent total hip replac treatment
doctor	staff	recept staff	doctor	air condit fan	surgeri
profession	amaz friendli funni put	surgeri	time	chester ward desper	nh patient
practic manag	extrem help pleasant radiograph	bmi alexandra hospit	week	croft medic centr	appoint
staff	mental health servic user	acorn surgeri	sever sleep apnoea osteoarthr	dai case unit	consult
absolut brilliant servic	xrai dept love aswel	nurs	stiflingli hot wait room	easi pram access	doctor
compassion nurs staff	care	appoint	today mondai 5th june	excel menu choic	time

elect care centr	doctor	craig croft medic centr	surgeri	limit disabl bai	staff
fist nearest pharmacist	friendli	earli morn emerg visit	hour	onlin book system	total hip replac
health centr manageress	help	excel minor injuri unit	emerg appoint	open plan recept	5am good fridai morn
small fee payabl	receptionist	hospit till late beri	phone	pleasant relax atmospher	blackheath hospitalon 27th march
son health problem	rude	long stand health condit	dai	poor eyesight imposs	overli aggres park enforc
total knee replac	time	manag due till ill	doctor doesn't releas appt	poor facil manag	privat patient
whilst phone ring	nurs	nh choos book scheme	morn	proper treatment room	patient
medic review time	patient	prescript request staff moral	long wait	wait room big	meniscu knee repair surgeri
excel care	extrem rude	recept staff performig triag	recept staff	cold clinic hospit	offer health care advic
recept staff	appoint	consult	urgent appoint	fairli dirti environ	dai
safe hand	access friendli practic	good servic	consult	main entranc open	blood test
consult	complet unprofession behaviour	recoveri room staff excel	amaz paramed duo	book appoint	back
time	entir wait room	time	bad sinu infect	car park	week
medic issu	excel nurs care	servic	bank holiday weekend	excel care	year
treatment	fantast ent depart	daughter chicken pox vaccin	blood test request	onlin facil	nh
bad attack	friendli help staff	care	call back tomorrow	privat hospit	acorn surgeri
bad impress	friendli recept staff	hospit	croft med centr	rude receptionist	blackheath hospit
bodi wash	front desk staff	patient	dai case patient	ten minut	bmi hospit
care team	full wait room	dai	debilit uterin cramp	privat patient	heart attack
christma period	great bedsid manner	help	depth pre consult	recept area	local hospit

conceiv method	great person care	practic	effici appoint linel	hospit	long time
difficult new	grumpi recept staff	theatr staff	full hip replac	clean	recept area

보증(Assurance)에는 의료인들의 능력을 신뢰하는 단어들이 포함되어 있다. respect, depend, brilliant, excel care, safe hand와 같은 단어들이 여기에 포함된다. 공감(Empathy)은 긍정적인 단어인 amaz(ing), friendli, extrem help, pleasant, love, excel, fantast, great 가 등장하였으며, 부정적인 단어로는 rude, extrem rude, unprofession, grumpi가 등장하였다. 신뢰성(Reliability)에는 신뢰성과 관련된 단어가 많이 발견되지 않았다. excel, good servic 같은 단어들이 있지만 다른 서비스 품질 차원에도 등장하는 단어이다. 응답성(Responsiveness)은 시간과 관련된 단어가 많이 포함되어 있다. time, week, hour, dai, morn(ing) 과 같은 시간에 대한 단어가 있으며, 이와 함께 long wait, urgent appoint, phone 같은 예약이나 진료의 대기 시간에 대한 단어들이 순위에 포함되어 있다. 예약이나 진료를 위한 대기 시간과 관련된 내용이 응답성에 많이 포함된 것으로 보인다. 유형(Tangibles)은 병실(room, ward), 예약시스템(appoint system, onlin book system, onlin facil), 병원시설(car park, wait room, recept area)과 관련된 단어가 많았으며 유형의 시설에 대한 평가가 포함된 clean, dirti 등의 단어가 함께 사용되었다.

## 5. 토의 및 결론

서비스 품질의 차원별로 주요하게 언급되는 단어를 파악하기 위하여 RAKE 알고리즘과 단어의 빈도와 변별성을 고려한 방안을 제안하였으며, 제안한 방안의 의

해 파악된 서비스 품질 차원별 주요 단어를 기술하고 설명하였다. 여러 리뷰에 존재하는 주요 단어를 파악하여야 하기에 RAKE 스코어만을 기반으로 주요 단어를 파악하는 것 보다는, 빈도와 변별성(IDF)을 동시에 고려하는 것이 필요한 것으로 보인다.

수집된 리뷰를 SERVQUAL의 서비스 차원으로 분류하였을 때 신뢰성에 가장 많이 포함되었고, 그 다음으로 감성과 응답성에 많이 분류되었다. 보증과 유형에 대한 내용은 적게 언급되었다. 신뢰성에 대해서는 다른 차원들과 차이가 많은 부분은 크지 않았으며 일반적인 헬스케어 서비스에 대한 단어들이 많이 포함되었다. 평점을 기준으로 긍정적인 경험과 부정적인 경험을 나누는데 많이 포함되는 서비스 차원은 감성과 응답성 차원이라고 볼 수 있다. 접수 직원을 비롯한 직원(staff)의 무례함이 부정적인 감성의 주요 요인이며, 친절하고 친절할 도움을 받은 경우가 긍정적인 감성의 주요 내용이다. 마찬가지로 응답성 측면의 주요한 부정적인 요인은 예약 및 진료의 긴 대기시간이다.

수집한 데이터가 GP와 병원에 대한 리뷰여서 간단한 진료에 대한 내용이 많기 때문이기도 하지만, 헬스케어 서비스에 대한 리뷰에 주로 감정적인 표현이나 대기 시간과 같은 환자들이 쉽게 느끼고 판단할 수 있는 부분이 많았다. 헬스케어 서비스에서 병이나 처방과 같은 부분에 대한 전문적인 지식에 대한 영역에는 많은 내용이 포함되어 있지 않았다. 시설에 대한 부분도 병실이나 병원의 의료장비에 대한 언급은 거의 없었으며, 청결상태나 병실 식사에 대한 부분이 많이 언급되었다.

본 연구에서는 문장 단위로 서비스 품질의 차원에 할당하였다. 따라서, 문장에서 단어별 RAKE 스코어를

계산하였기에 매우 짧은 문장에서 스코어를 구하였다. 이로 인해 빈번히 사용되지 않은 단어들이 높은 스코어를 갖게 되는 경우가 많았다. 각 문장에서는 주요한 역할을 하지만 다른 문장들에는 나타나지 않는 경우가 많았다. 짧은 문장에서 단어의 중요성을 파악하는 방안을 더 고려할 필요가 있다.

## 참 고 문 헌

### [국내 문헌]

1. 김상만, 이연주 2010. “의료서비스산업에서의 고객 지식 획득과 활용방안: 기대 불일치 이론을 중심으로,” *지식경영연구* (11:3), pp. 59-76.
2. 김상만, 엄기현, 오재영 2010. “Online 의료웹서비스 품질과 지식제공성과의 관계 연구,” *지식경영연구* (11:1), pp. 1-17.
3. 박명석, 권영진, 이상용 2018. “댓글이 음원 판매량에 미치는 차별적 영향에 관한 텍스트마이닝 분석,” *지식경영연구* (19:2), pp. 91-108.
4. 루스 안젤리 크루즈, 이홍주 2014. “브랜드 개성 효과: 트위터 상의 브랜드 개성 전달이 온라인 커뮤니티 참여에 미치는 영향,” *지능정보연구* (20:1), pp. 67-101.

### [국외 문헌]

1. Akter, S., D'Ambra, J., and Ray, P. 2010. “Service quality of mHealth platforms: development and validation of a hierarchical model using PLS,” *Electronic Markets* (20:3-4), pp. 209-227.
2. Andrade M., and Valencia, A. 1998. “Automatic extraction of keywords from scientific text: application to the knowledge domain of protein families,” *Bioinformatics* (14:7), pp. 600-607.
3. Babakus, E., and Mangold, W.G. 1992. “Adapting the SERVQUAL scale to hospital services: an empirical investigation,” *Health Services Research* (26:6), pp. 767-786.
4. Baeza-Yates, R., and Ribiero-Neto, B. 2010.

- Modern Information Retrieval: The Concepts and Technology Behind Search* (2nd ed.), Addison Wesley.
5. Cao, Q., Duan, W., and Gan, Q. 2011. "Exploring determinants of voting for the "helpfulness" of online user reviews: A text mining approach," *Decision Support Systems* (50:2), pp. 511-521.
  6. Chen, K., Kou, G., Shang, J., and Chen, Y. 2015. "Visualizing market structure through online product reviews: Integrate topic modeling, TOPSIS, and multi-dimensional scaling approaches," *Electronic Commerce Research and Applications* (14:1), pp. 58-74.
  7. Christopher Baker, "slowraker: A Slow Version of the Rapid Automatic Keyword Extraction (RAKE) Algorithm," R package Version 0.1.1, <https://CRAN.R-project.org/package=slowraker>
  8. Cronin, J. J. Jr. and Taylor, S. A. 1992. "Measuring service quality: a reexamination and extension," *Journal of Marketing* (56:3), pp. 55-68.
  9. Dagger, T. S., Sweeney, J. C., and Johnson, L. W. 2007. "A hierarchical model of health service quality: scale development and investigation of an integrated model," *Journal of Service Research* (10:2), pp. 123-142.
  10. Dean, A. M. 1999. "The applicability of SERVQUAL in different health care environments," *Health Marketing Quarterly* (16:3), pp. 1-21.
  11. Ghose, A., and Ipeirotis, P. G. 2011. "Estimating the Helpfulness and Economic Impact of Product Reviews: Mining Text and Reviewer Characteristics," *IEEE Transactions on Knowledge and Data Engineering* (23:10), pp. 1498-1512.
  12. Greaves, F., Pape, U. J., King, D., Darzi, A., Majeed, A., Wachter, R. M., and Millett, C. 2012. "Associations between Internet-based patient ratings and conventional surveys of patient experience in the English NHS: an observational study," *BMJ Quality & Safety* (21:7), pp. 600-605.
  13. Greaves, F., Ramirez-Cano, D., Millett, C., Darzi, A., and Donaldson, L. 2013. "Harnessing the cloud of patient experience: using social media to detect poor quality healthcare," *BMJ Quality & Safety* (22:3), pp. 251-255.
  14. Greaves, F., Lavery, A. A., Cano, D. R., Moilanen, K., Pulman, S., Darzi, A., and Millett, C. 2014. "Tweets about hospital quality: a mixed methods study," *BMJ Quality & Safety* (23:10), pp. 838-846.
  15. Grönroos, C. 1984. "A service quality model and its marketing implications," *European Journal of Marketing* (18:4), pp. 36-44.
  16. Hu, N., Pavlou P. A. and Zhang, J. 2009. "Overcoming the J-shaped Distribution of Product Reviews," *Communications of the ACM* (52:10), pp. 144-147.
  17. Hulth, A. 2003. "Improved automatic keyword extraction given more linguistic knowledge," in Proceedings of the 2003 Conference on Empirical Methods in Natural Language Processing.

18. Humphreys, A., and Jen-Hui Wang, R. 2018. "Automated Text Analysis for Consumer Research," *Journal of Consumer Research* (44:6), pp. 1274-1306.
19. Ladhari, R. 2009. "A review of twenty years of SERVQUAL research," *International Journal of Quality and Service Sciences* (1:2), pp. 172-198.
20. Lam, S. S. 1997. "SERVQUAL: A tool for measuring patients' opinions of hospital service quality in Hong Kong," *Total Quality Management* (8:4), pp. 145-152.
21. Lee, H., Delene, L. M., Bunda, M. A., and Kim, C. 2000. "Methods of measuring health-care service quality," *Journal of Business Research* (48:3), pp. 233-246.
22. Lehtinen, U., and Lehtinen, J. R. 1982. *Service quality: a study of quality dimensions*, Service Management Institute.
23. Netzer, O., Feldman, R., Goldenberg, J., and Fresko, M. 2012. "Mine Your Own Business: Market-Structure Surveillance Through Text Mining," *Marketing Science* (31:3), pp. 521-543.
24. Mihalcea, R. and Tarau, P. 2004. "TextRank: Bringing order into texts," in *Proceedings of EMNLP 2004* (ed. Lin D and Wu D), Barcelona, Spain.
25. Ohsawa, Y., Benson, N. E., and Yachida, M. 1998. "Keygraph: Automatic indexing by co-occurrence graph based on building construction metaphor," in *Proceedings of ADL'98*.
26. Palese, B., and Usai, A. 2018. "The relative importance of service quality dimensions in E-commerce experiences," *International Journal of Information Management* (40), pp. 132-140.
27. Parasuraman, A., Zeithaml, V. A., and Berry, L. L. 1985. "A conceptual model of service quality and its implications for future research," *Journal of Marketing* (49:4), pp. 41-50.
28. Parasuraman, A., Zeithaml, V. A., and Berry, L. L. 1988. "SERVQUAL: A Multiple-Item Scale for Measuring Consumer Perceptions of Service Quality," *Journal of Retailing* (64:1), pp. 12-40.
29. Park, G.-W., Kim, Y., Park, K., and Agarwal, A. 2016. "Patient-centric quality assessment framework for healthcare services," *Technological Forecasting and Social Change* (113), pp. 468-474.
30. Stuart R., Engel, D., Cramer, N. and Cowley, W. 2010. "Automatic keyword extraction from individual documents," *Text Mining: Applications and Theory* (edited by M. W. Berry and J. Kogan), John Wiley & Sons.
31. Sofaer, S., and Firminger, K. 2005. "Patient perceptions of the quality of health services," *Annual Review of Public Health* (26), pp. 513-559.
32. Zeithaml, V. A. 2000. "Service quality, profitability, and the economic worth of customers: what we know and what we need to learn," *Journal of the academy of marketing science* (28:1), pp. 67-85.

---

● 저 자 소 개 ●

---



**1. 이 홍 주 (Hong Joo Lee)**

현재 가톨릭대학교 경영학부 교수로 재직 중이다. KAIST 테크노경영대학원에서 경영학 박사 학위를 취득하였다. 주요 관심분야는 데이터 분석, 지능형 정보시스템, 온라인 사용자들의 상호작용 등이다. 지금까지 Journal of Electronic Commerce, Cyberpsychology, Behavior, and Social Network, Technological Forecasting & Social Change, International Journal of Electronic Commerce 등 주요 학술지에 논문을 발표하였다.