

밀키트 구독서비스의 리뷰 마이닝을 통한 서비스 기회 발굴

정재연 · 이지호 · 윤장혁[†]

건국대학교 산업공학과

Service Opportunity Discovery Via Review Mining of Meal Kit Delivery Service

Jeayeon Jung · Jiho Lee · Janghyeok Yoon

Department of Industrial Engineering, Konkuk University

The meal kit delivery service market has grown rapidly in line with social changes such as the increase in single-person households and dual-income households. Prior studies were conducted to analyze this attractive market, but little attention has been paid to quantifying service opportunity analysis of the meal kit delivery service. Therefore, this study aims to identify customer needs and their opportunity potential via review mining of meal kit delivery service. First, we collect 5,842 customer reviews of a meal kit delivery service firm “HelloFresh”. Next, we apply topic modeling to identify service topics and their importance from the target firm’s reviews and apply sentiment analysis to calculate the satisfaction level for each topic. Finally, a service opportunity map based on topic importance and satisfaction scores is generated to discover service opportunities for the firm. This study quantifies the process of discovering service opportunities in meal kit delivery service and the research findings can be used to formulate marketing and planning strategies for competitive advantages in the relevant market.

Keywords: Meal Kit Delivery Service, Review Mining, Customer Analysis, Sentiment Analysis, Service Opportunity

1. 서론

현대 식품 시장은 가구 형태의 변화, 맞벌이 가구의 증가와 같은 사회 변화로 인해 전환기를 맞이하였다(Buckley *et al.*, 2007). 1인 가구의 증가는 가장 두드러지는 현대 사회의 변화로, 1인 가구가 간편한 식사를 선호함에 따라 기업들은 이러한 니즈를 반영한 제품들을 선보였다(Daniels and Glorieux, 2015; Jo, 2016). 특히, 손질된 재료와 레시피 카드가 들어있는 밀키트를 구독자들에게 정기적으로 배송하는 밀키트 구독서비스가 새로운 비즈니스 모델로 등장했다. 밀키트 구독서비스는 빠른 성장세를 보이고 있으며, 미국의 리서치 회사 ‘Grand

View Research’의 자료에 따르면 2021년 기준 세계 밀키트 구독서비스 시장의 규모는 150억 달러 규모이며 향후 5년간의 연평균 성장률은 17.4%로 전망된다(Heard *et al.*, 2019).

밀키트 구독서비스의 성장에 발맞추어, 최근 밀키트 구독서비스를 대상으로 한 연구들이 활발히 수행되고 있다. 설문조사를 통해 소비자의 구매 및 만족도에 영향을 미치는 요인을 파악한 연구(Chung and Kim, 2020), 성별과 연령에 따른 소비자 유형이 재구매에 미치는 영향을 파악한 연구가 수행되었다(Lee and Kim, 2022). 또한, 밀키트 관련 소셜 데이터를 수집하고 키워드 분석을 통해 밀키트에 대한 소비자들의 전반적인 인식을 분석한 연구들이 수행되었다(Yi and Choi, 2019; Lee,

본 연구는 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행한 연구 과제입니다(NRF-2021R1A2C1010027). 본 연구는 2020년도 산업통상자원부(MOTIE)와 한국에너지기술연구원(KETEP)의 지원을 받아 수행한 연구 과제입니다(No. 20204010600220).

[†] 연락저자 : 윤장혁 교수, 05029 서울특별시 광진구 능동로 120 건국대학교 산업공학과, Tel : 02-450-0453, Fax : 02-450-3525,

E-mail : janghyoon@konkuk.ac.kr

2022년 10월 4일 접수; 2022년 11월 19일; 2022년 12월 19일 수정본 접수; 2022년 12월 26일 게재 확정.

2021). 선행연구들을 통해 밀키트에 대한 소비자의 인식과 밀키트 구매에 영향을 미치는 요인들이 제시되었지만, 선행연구에는 몇 가지 한계점이 존재한다.

먼저, 밀키트 구독 서비스는 고객과의 정기적인 구매관계를 유지해야 하는 비즈니스 모델로, 고객니즈를 구체적으로 파악하는 것이 중요하다(Janzer, 2020; Baek, 2021). 선행연구들은 등장빈도가 높은 키워드를 활용하여 밀키트 소비자의 전반적인 인식과 니즈를 확인하였다. 하지만, 복잡한 고객니즈를 파악하기 위해서는 문서 주제, 키워드가 활용된 맥락과 같은 추가적인 정보를 활용해야 한다. 이에 문서 내 다양한 정보를 활용하여 밀키트 구독서비스의 고객니즈를 식별하는 연구가 필요하다. 다음으로, 선행연구들은 밀키트 구독서비스 고객의 니즈를 확인하였으나 서비스 개선방향 도출을 위해 고객니즈를 평가할 수 있는 지표는 활용하지 않았다. 따라서, 만족도, 중요도 등 선행연구를 통해 제시된 고객니즈 평가지표를 활용하여 밀키트 구독서비스 기업의 서비스 개선방향을 도출할 수 있는 방법이 필요하다(Chen et al., 2018; Nam and Lee, 2019). 마지막으로, 선행연구들은 밀키트 구독서비스의 서비스 기회를 분석하지 못하였다. 선행연구들은 밀키트 구매에 영향을 주는 제품 특성을 파악하고 등장 빈도가 높은 키워드를 통해 밀키트 소비자들을 분석하였으나, 밀키트 구독서비스의 개선방향을 제시하지는 못하였다. 따라서, 식별된 고객니즈를 활용하여 밀키트 구독서비스의 서비스 기회를 발굴하는 연구가 필요하다.

본 연구는 밀키트 구독서비스 기업의 서비스 기회 발굴을 리뷰 마이닝 기반으로 수행하고자 한다. 분석을 통해 아래 세 가지 연구질문(Research Question; RQ)을 해결할 수 있을 것으로 기대한다.

- 밀키트 구독서비스의 고객들은 어떤 요구를 하는가?
- 밀키트 구독서비스 기업의 서비스 개선방향은 무엇인가?
- 고객은 밀키트 구독서비스의 어떤 요소에서 만족감을 느끼는가?

본 연구는 밀키트 구독서비스 기업의 서비스 기회 발굴을 위해 Jeong et al.(2019)이 제시한 기회 알고리즘 기반의 제품 기회 발굴 방법론을 활용한다. 구체적인 연구절차는 다음과 같다. 1) 밀키트 구독서비스 기업의 리뷰 데이터를 수집하고, 리뷰 내 서비스 관련 키워드를 추출한다. 2) 토픽 모델링을 통해 서비스 토픽을 정의하고 각 토픽의 중요도를 산출한다. 3) 감성분석을 활용하여 각 서비스 토픽에 대한 만족도를 산출한다. 4) 최종적으로, 중요도와 만족도를 축으로 하는 Service Opportunity Map을 작성하여 서비스 기회를 발굴한다.

본 연구의 기여는 다음과 같다. 첫째, 본 연구는 밀키트 구독서비스 기업의 리뷰를 활용하여 서비스 토픽을 식별한다. 서비스 토픽 식별을 위해 활용되는 토픽 모델링 기법은 리뷰 내 키워드들의 조합을 통해 문서 내 주제들을 더욱 객관적이고 명확

하게 분석할 수 있는 장점이 존재하기에(Park, 2019), 분석된 고객니즈는 밀키트 구독서비스 소비자들에 대한 심층적인 이해에 도움을 줄 것으로 기대된다. 둘째, 본 연구는 밀키트 구독서비스 기업의 서비스 기회를 발굴한다. 고객관점에서 서비스 개선방향을 분석한 연구들은 중요도, 만족도 등 여러 지표를 활용하였다(Chen et al., 2018; Nam and Lee, 2019). 선행연구에서 중요도는 고객이 집중하는 서비스 요소를 파악하고자 활용되었으며(Oh, 2001), 만족도는 서비스요소에 대한 고객의 만족도를 측정하기 위해 활용되었다. 서비스 개선은 고객의 만족을 목표로 하기 때문에(Mahmoud et al., 2017), 고객에게 중요하지만 만족되지 않은 서비스 요소를 확인함으로써 서비스 개선방향을 도출할 수 있다(Ulwick, 2005). 따라서 본 연구에서는 서비스 토픽의 중요도와 만족도를 산출하고 기회알고리즘을 활용하여 서비스 기회를 발굴한다. 지표를 통해 고객리뷰 내 기업의 서비스 기회를 정량적으로 도출함으로써 고객 관점의 서비스 개선방향을 파악할 수 있다. 셋째, 본 연구결과에서 발굴된 서비스 기회는 기업의 전략수립 시 활용될 수 있다. 리뷰 마이닝을 통해 발굴된 서비스 기회는 기업의 고객 중심 마케팅 및 기획과정에 도움을 줄 것으로 기대된다. 본 논문은 제2장에서 밀키트 구독서비스 및 비즈니스 기회 발굴과 관련된 선행연구들을 설명하고, 제3장에서는 본 연구의 각 절차의 결과물을 제시한다. 다음으로 제4장에서는 앞서 제시된 연구결과를 통해 세 가지 RQ를 해결하고 서비스 기회를 분석한다. 마지막으로, 제5장에서는 본 연구의 결론 및 한계점에 대하여 서술한다.

2. 배경연구

2.1 밀키트 구독서비스 연구

밀키트 구독서비스는 1인 가구의 편리한 식사와 관련된 니즈를 반영한 비즈니스 모델로, 최근 식품 산업 내 급격한 성장세를 보이고 있다(Khan and Sowards, 2018). 밀키트 구독서비스의 성장을 기반으로, 이와 관련된 연구들이 활발하게 수행되고 있다(Lee et al., 2021; Park and Hwang, 2021; Ahn and Kim, 2022). 선행연구들은 밀키트 구매에 영향을 주는 변수를 식별하고, 밀키트에 대한 소비자 인식을 파악하고자 했으며, 크게 설문 기반, 데이터 기반의 두 가지 유형으로 분류할 수 있다.

설문 기반 연구의 경우, 밀키트 소비자들을 대상으로 한 설문조사 결과를 활용하여 고객만족 및 구매에 영향을 미치는 변수를 식별했다. Yang et al.(2019)은 설문조사를 통해 소비자 만족도에 영향을 미치는 밀키트의 제품 속성을 조사하였다. 또한, Choi et al.(2020)은 소비자들을 대상으로 설문조사를 시행하고 상관분석을 통해 맛, 품질과 같은 제품 속성이 밀키트 소비자의 만족도와 제품 구매에 미치는 영향력을 분석하였다. Lee and Park(2021)은 코로나-19 유행 시기에 밀키트를 구매한 소비자들을 대상으로 설문조사를 수행하였고, 해당 기간 중

밀키트의 제품 속성이 구매와 만족도에 미치는 영향을 분석하였다. 하지만, 설문 기반의 연구방법은 데이터 수집에 비교적 오랜 시간이 소요되며, 데이터의 규모를 확장하는데 한계가 존재한다.

이와 같은 설문 기반 연구들의 한계점을 극복하고자 데이터 기반의 연구들이 수행되었다. Yi and Choi(2019)는 다양한 소셜미디어로부터 밀키트와 관련된 데이터를 수집하여 키워드 등장빈도 분석, 키워드 네트워크 분석을 수행하였다. 키워드 네트워크를 통해 밀키트 메뉴, 브랜드 그리고 특징 3개의 군집을 제시하여 밀키트 시장의 소비자 인식을 전반적으로 파악하였다. 그리고, Lee(2021)는 밀키트와 관련된 키워드를 수집하여 키워드 빈도 분석과 키워드 네트워크 분석을 수행하였다. 등장빈도 상위 80개의 키워드를 확인하고 등장빈도가 높은 키워드를 활용하여 하여 코로나19 이후 밀키트 소비자의 인식을 파악하였다. Choi and Yeon(2022)은 밀키트 제품리뷰 데이터를 수집하여, 워드 클라우드 분석과 감성분석을 수행했다. 밀키트 메뉴(간식/디저트, 구이/찜 등)에 따라 등장빈도가 높은 키워드를 제시하고, 카테고리 별로 메뉴, 재료, 기타 세 가지 하위 항목에 따른 긍·부정 키워드를 파악하여 제품 구매에 영향을 주는 키워드를 제시하였다.

데이터 기반의 선행연구들이 밀키트 시장 전반에 대한 소비자의 인식을 성공적으로 분석하였지만, 분석 결과로 키워드들만 제시되었다. 하지만, 고객니즈는 단일 키워드가 아닌 키워드의 조합으로 표현되기 때문에, 고객이 작성한 리뷰 주제와 키워드의 활용 맥락을 고려한 구체적인 분석이 필요하다. 또한, 고객 관점의 서비스 개선방향을 제시하기 위해서는 고객니즈를 평가할 수 있는 정량적인 지표가 필요하다.

따라서 본 연구는 밀키트 구독서비스 리뷰 데이터를 활용하여 밀키트 구독서비스 기업의 서비스 토픽을 식별하고 서비스 기회를 발굴하고자 한다. 리뷰 데이터는 수집이 용이하고 서비스 및 제품과 관련된 소비자의 다양한 의견이 담겨있기에 본 연구 목표에 적합한 데이터이다. 토픽 모델링을 활용하여 밀키트 구독서비스 리뷰데이터로부터 서비스 토픽을 식별하고 구체적인 고객니즈를 파악한다. 식별된 서비스 토픽에 대한 정량적인 지표를 산출하여 서비스 기회를 발굴하고, 최종적으로 밀키트 구독서비스에 대한 고객관점의 서비스 개선방향을 제시한다.

2.2 리뷰 마이닝 기반 제품 기회 발굴 연구

현대 사회에서 기업의 지속적인 성장을 위해서는 고객의 목소리를 수렴하는 것이 중요하다(Goffin *et al.*, 2012). SNS, 커뮤니티, 블로그와 같은 소셜미디어는 다양한 고객의 의견이 존재하는 데이터 원천으로, 이를 활용하여 제품 및 서비스와 관련된 데이터를 수집하고 데이터로부터 유의미한 정보를 찾고자 하는 다양한 연구들이 수행되었다(Gundecha and Liu, 2012; Liu *et al.*, 2019). 구체적으로, 소셜미디어에 포함된 고객의 의

견을 활용하여 제품 개선방향을 제시한 연구(Qi *et al.*, 2016; Wang *et al.*, 2020; Yang *et al.*, 2021), 기업의 서비스 개선 방향을 제시한 연구(Ibrahim and Wang, 2019; Kim and Lim, 2021; Zhan *et al.*, 2021), 비즈니스 기회를 발굴하는 연구가 수행되었다(Ko *et al.*, 2017; Wu *et al.*, 2018; Choi *et al.*, 2020).

그중, 리뷰 데이터는 고객의 제품 및 서비스에 대한 의견이 담긴 데이터로, 고객관점에서 기업의 제품 및 서비스 개선방향을 탐색하는데 적합한 데이터로 알려져 있다. 실제로, Ko *et al.*(2017)은 스마트폰에 대한 온라인 고객 리뷰를 통해 제품 기회를 분석하는 체계적인 방법론을 제안했다. 리뷰 데이터에 토픽 모델링을 적용하여 제품 토픽을 확인하고 제품 토픽 간의 관계를 파악하고자 키 그래프를 생성하였다. 생성된 키 그래프에 Chance discovery theory를 적용하여 제품 기회를 식별하였다. 그리고 Wu *et al.*(2018)은 스마트폰에 대한 온라인 고객 리뷰에 감성분석, 동적 IPA 분석 그리고 의사결정 트리를 활용하여 제품 및 서비스 개선 우선순위를 발굴하는 방법론을 제시하였다. 감성분석과 의사결정 트리를 활용하여 제품 및 서비스 요소에 대한 만족도를 산출하였으며, 카노 모형을 활용하여 타겟 기업의 제품 및 서비스 기능 개선의 우선순위를 파악하였다. 또한, Choi *et al.*(2020)은 스마트 스피커에 대한 고객 리뷰에 이벤트 탐지 알고리즘과 감성분석 방법을 활용하여 시간에 따라 진화하는 토픽들을 분석한 뒤, 제품 기회를 식별하고 평가하는 방법론을 제시했다. 이벤트 탐지 알고리즘을 통해 시간이 변화함에 따라 등장하는 제품 이벤트를 확인하고 시간 구간에 따른 제품 기회를 식별하고 변화 양상을 통해 제품 기회를 평가했다.

이러한 연구 가운데, Jeong *et al.*(2019)은 소셜미디어 리뷰에 토픽 모델링을 적용하여 제품 토픽을 식별하고, 두 가지 지표와 기회 알고리즘을 활용하여 식별된 토픽 가운데 제품 기회를 발굴하는 방법론을 제시했다. 또한, 고객 관점에서의 중요도, 만족도를 측정하는 두 가지 지표를 통해 각 제품 토픽의 기회점수를 산출하였으며 이를 통해 제품 기회를 발굴했다. 해당 방법론은 리뷰 데이터를 활용하여 토픽을 도출함으로써 고객의 요구사항의 파악이 가능하며, 지표를 활용하여 고객 관점에서의 제품 및 서비스 기회의 발굴이 가능하다. 이러한 장점에 입각하여, 본 연구에서는 Jeong의 제품 기회 발굴 방법론을 활용하여 밀키트 구독서비스의 고객니즈를 식별하고 서비스 기회를 발굴하고자 한다. 기존 밀키트 구독서비스 분야 선행연구들과 달리, 본 연구에서는 토픽 모델링을 활용하여 식별된 고객니즈를 분석한다. 이에 그치지 않고, 정량적인 지표를 활용하여 고객니즈를 평가함에 따라 고객니즈를 세분화할 수 있으며 고객관점의 서비스 개선방향을 도출한다.

3. 분석절차

본 연구의 절차는 Jeong의 방법론에 기반하여 총 4단계로 구성된다(<Figure 1>).

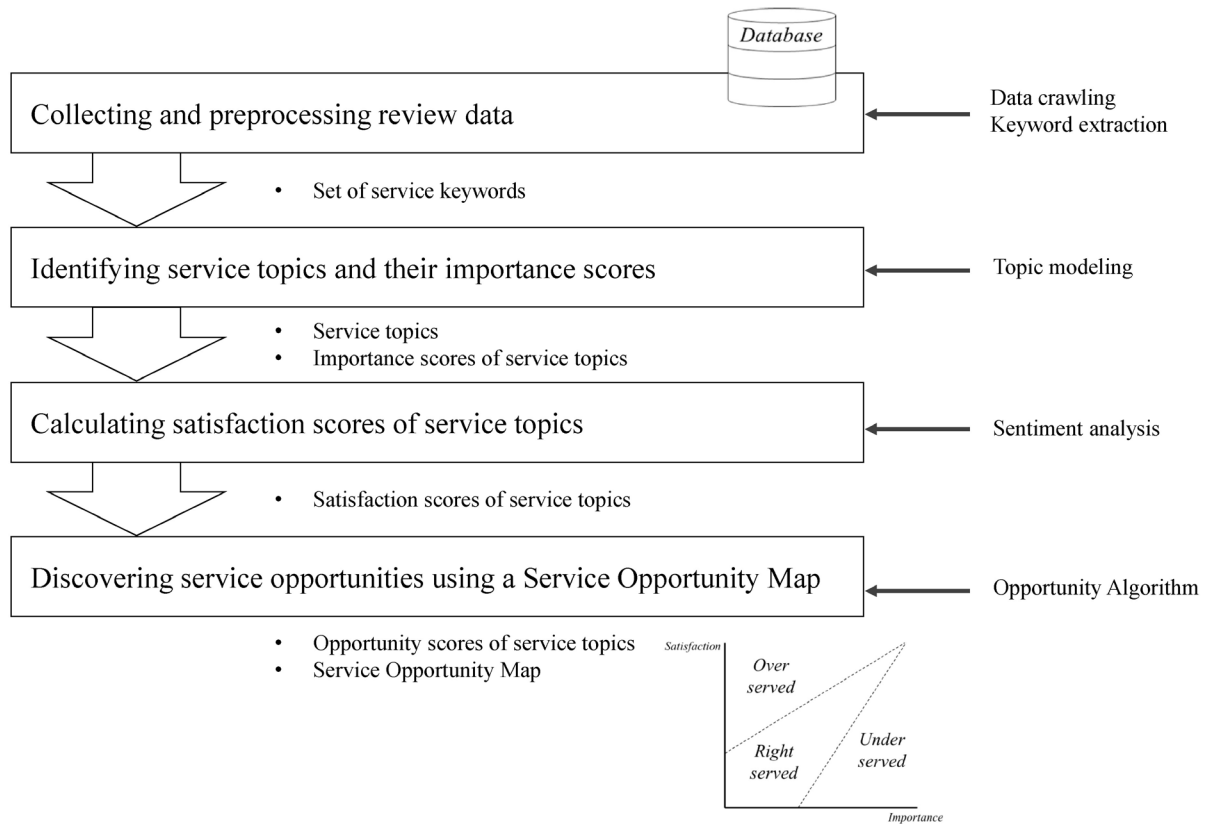


Figure 1. Research Procedure

1) 먼저, 리뷰 데이터를 수집하고 전처리 과정을 거쳐 서비스 키워드를 선정한다. 2) 다음으로, 토픽 모델링 기법인 LDA (Latent Dirichlet Allocation)을 통해 리뷰 내 서비스 토픽을 식별하고 각 서비스 토픽의 중요도를 산출한다. 3) 감성분석을 통해 서비스 키워드의 감성점수를 계산하고 이를 활용하여 서비스 토픽의 만족도를 산출한다. 4) 마지막으로, 기회 알고리즘을 통해 Service Opportunity Map을 구축하고 서비스 기회를 발굴한다. 본 연구에서는 밀키트 구독서비스 기업인 HelloFresh를 타겟 기업으로 선정했다. 다국적 기업인 HelloFresh는 소비자에게 다양한 메뉴 및 식단 옵션을 제공하기 때문에, 밀키트 구독서비스에 대한 소비자의 니즈를 식별하고 서비스 기회를 발굴하기 적합하다. 본 장에서는 HelloFresh의 고객 리뷰에 네 단계로 구성된 연구절차를 적용한 결과를 제시한다.

3.1 리뷰 데이터 수집 및 키워드 추출

본 연구단계에서는 밀키트 구독서비스 기업의 고객니즈 및 서비스 기회를 분석하고자 고객 리뷰를 수집하고 리뷰 내 키워드를 추출한다. 소셜미디어 리뷰 데이터의 수집방안으로는 기업에서 제공하는 Open API(Application Programming Interface)를 활용하거나, 프로그래밍 언어를 통해 웹사이트 내 정보를 직접 수집하는 웹 크롤링 방법 등이 있다. 본 연구는 ConsumerAffairs (<https://www.consumeraffairs.com/>) 페이지 내 HelloFresh의 고객

리뷰를 수집했다. ConsumerAffairs는 다양한 제품 및 서비스에 대한 고객 리뷰를 확인할 수 있는 소비자 리뷰 플랫폼으로, 제공되는 고객 리뷰에는 작성날짜, 본문, 별점, 작성자 등의 내용이 포함되어 있다. 본 연구에서는 수집 가능한 전체 기간인 2016년 8월 16일부터 2022년 1월 24일까지의 5,922개의 리뷰를 Python 라이브러리인 selenium과 beautifulsoup을 활용하여 크롤링한 후 데이터베이스화 하였다. 수집된 리뷰 길이의 평균은 429이고 표준편차는 188이었다. 가장 짧은 리뷰의 길이는 6이고, 가장 긴 리뷰의 길이는 1,023이었다.

다음으로, 서비스 키워드를 정의하고자 리뷰 데이터로부터 키워드를 추출한다. 본 연구에서는 리뷰 내 존재하는 다양한 키워드를 활용하고자 단일어와 복합어를 모두 서비스 키워드로 사용한다. 정의된 서비스 키워드들을 활용하여 리뷰 데이터를 재표현하고, 서비스 키워드가 포함되지 않은 리뷰 데이터는 분석과정에서 제외한다. 키워드 추출을 위해 RAKE, YAKE와 같은 라이브러리를 활용하거나 키워드 추출 API를 사용할 수 있다. 본 연구에서는 Python 라이브러리 중 개별 문서들로부터 키워드를 추출하는 비지도, 도메인-독립적 그리고 언어-독립적인 라이브러리인 RAKE를 사용하여 리뷰 내 24,337개의 키워드를 추출했다(Rose et al., 2010). 추출된 키워드들 중 불필요한 키워드를 제거하는 전처리 과정을 통해 서비스 키워드를 확보하였다. 전체 24,337개의 키워드 가운데, 숫자, 특수기호, 날짜(2016, \$\$, 08-12)등을 제거하여 약 6,300

개의 키워드를 선별하였고, 다음으로 전체 리뷰 내 등장빈도가 1이거나, 키워드로 활용하기 어려운 형용사, 부사(beautiful, supposedly)를 제거했다. 마지막으로 밀키트 구독서비스와 관련된 키워드만을 선별하여 2,161개의 키워드를 서비스 키워드로 정의하였다. 최종적으로 서비스 키워드가 포함되지 않은 리뷰를 제외하여 5,842개의 리뷰를 분석에 활용하였다.

3.2 서비스 토픽 식별 및 중요도 산출

본 연구단계에서는 토픽 모델링 기법인 LDA를 활용하여 리뷰 내 서비스 토픽을 식별한다. LDA는 주어진 문서들의 단어 분포를 활용하여 각 문서별로 어떤 주제가 존재하는지 도출하는 확률적 토픽 모델링 기법이다(Blei *et al.*, 2003). LDA는 문서-키워드 매트릭스와 하이퍼파라미터인 α , β , k 를 입력 받아 문서-토픽 매트릭스와 토픽-키워드 매트릭스를 결과로 출력한다. 하이퍼파라미터 α , β 는 문서에 대한 키워드의 기여

도 분포에 영향을 주며, k 는 토픽의 개수를 의미한다. 토픽 모델링을 통해 도출된 토픽들이 해석 가능하고, 전체 토픽이 모든 문서의 주제를 포괄하기 위해서는 최적의 토픽 수 결정이 요구된다(Zhao *et al.*, 2014; Zhao *et al.*, 2015). 최적의 토픽개수는 토픽 간의 코사인 유사도를 바탕으로 한 Elbow method를 활용하여 탐색할 수 있다(Wang *et al.*, 2014). Elbow method는 군집화 분석 시 클러스터의 최적 개수를 정하기 위해 사용되는 기법이다(Bholowalia and Kumar, 2014; Syakur *et al.*, 2018).

키워드들로 표현된 각 리뷰들은 문서-키워드 빈도 매트릭스의 형태로 LDA의 입력 값으로 사용된다. 본 연구에서는 Elbow method를 활용하여 LDA의 하이퍼파라미터인 토픽 개수를 선정하였다. 토픽 개수에 따른 토픽 간 평균 코사인 유사도의 변화를 관찰하여 유사도의 하향세가 안정화되는 Elbow 지점을 토픽의 개수(k)로 활용할 수 있다(<Figure 2>). 본 연구는 Elbow 지점으로 판단 가능한 16, 24, 29 등과 안정화 이후 지점들로 k 값을 변경하며 토픽들을 레이블링하는 반복실험

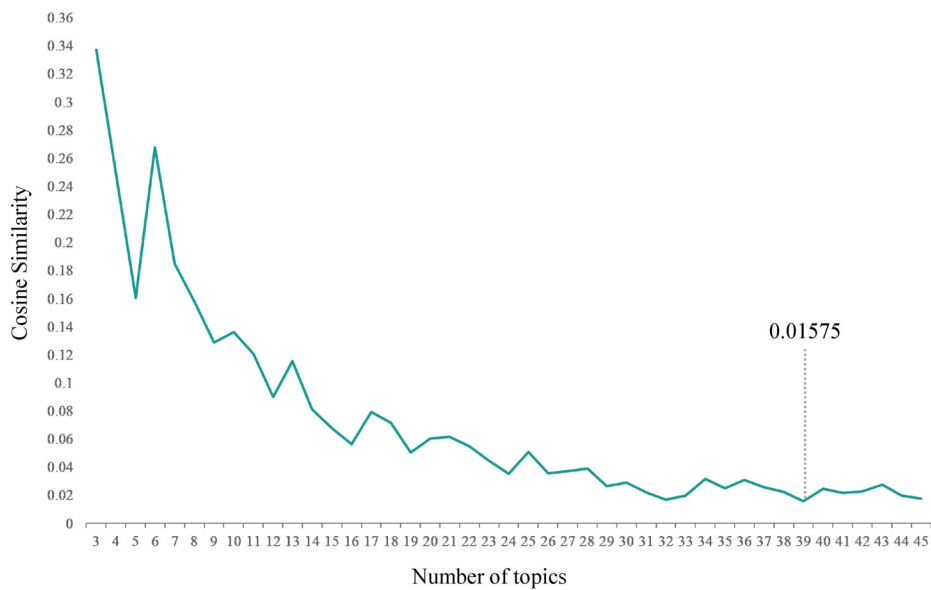


Figure 2. Average of Cosine Similarities of Pairs of Service Topics

Table 1. Part of Identified Service Topics Using Topic Modeling

Index	Topic	Major Keywords
1	Price promotion	code(0.138), promo(0.100), protein(0.099), fresh box(0.072), kids(0.060), HelloFresh box(0.056), Groupon(0.047), promotion(0.045)
2	Fresh vegetables	veggie(0.250), veggies(0.165), sauce(0.111), sauces(0.062), meat(0.046), freshness(0.034), fresh veg(0.022), herb(0.016)
3	Home-cooked meal	dinner(0.603), FedEx(0.042), great service(0.029), families(0.020), subscribing(0.020), taste(0.020), saving money(0.016), money(0.015)
4	Shipment information	ship(0.419), shipment(0.155), shipping(0.132), shipments(0.047), account(0.015), cost(0.014), free shipping(0.012), meal kit service(0.010)
5	Meal selection	selection(0.217), selections(0.100), restaurant(0.080), food delivery(0.050), delivery service(0.043), edible(0.040), meal selection(0.038), services(0.034)

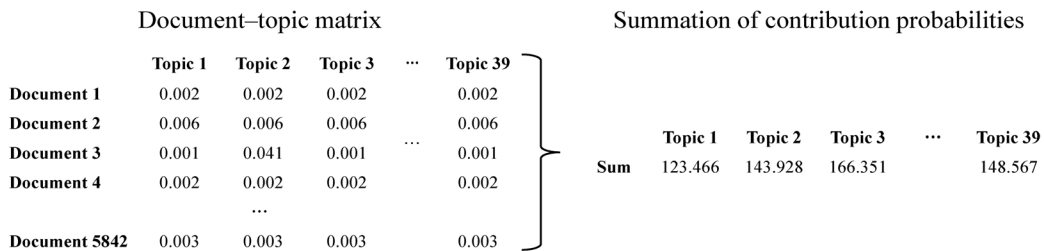


Figure 3. Computing Importance Scores of Service Topics

을 수행하였다. k의 변화에 따라 레이블링된 토픽들을 연구자들이 검토하였고, 가장 이상적인 k로 판단된 39를 최종 토픽수로 정의하였다.

본 연구에서는 Python 라이브러리인 genism을 활용해 LDA를 수행하였으며, 도출된 39개의 서비스 토픽들은 토픽-키워드 매트릭스의 값인 토픽 별 키워드 기여확률을 활용하여 레이블링되었다(<Appendix 1>). <Table 1>에서 본 연구가 도출한 토픽 예시를 확인할 수 있다.

다음으로, 도출된 토픽이 고객들에게 얼마나 중요한 토픽인지를 나타내는 중요도 지표를 산출한다. 문서-토픽 매트릭스의 토픽에 대한 각 문서의 기여도를 활용하여 고객 관점에서의 토픽의 중요도를 산출한다. 리뷰에서 빈번하게 언급되는 토픽은 소비자 관점에서 중요한 서비스 요소이므로(Rai, 2012), 문서-토픽 매트릭스에서 토픽이 모든 리뷰 문서에서 갖는 기여확률을 합하여 중요도를 계산한다(<Figure 3>). 본 연구의 지표들은 기계 알고리즘을 적용한 기회점수의 계산 과정에서 활용되므로 정규화가 필요하다. 따라서, 계산된 중요도 지표는 최대-최소 정규화를 통해 0과 10 사이의 값으로 변환하여 사용한다. 서비스 토픽의 평균 중요도는 3.432 이며, 가장 높은 중요도를 가지는 서비스 토픽은 “Cancellation and refund policy”였다.

3.3 서비스 토픽의 만족도 산출

본 연구단계에서는 감성분석과 토픽-키워드 매트릭스를 활용하여 서비스 토픽의 만족도를 산출한다. 감성분석은 자연어

처리의 한 분야로, 텍스트 데이터 내의 작성자의 의견, 태도, 감정과 같은 주관적인 견해를 긍·부정의 척도로 파악하는 기법이다(Medhat *et al.*, 2014). 감성분석의 결과는 -1과 1사이의 점수로 나타나며, -1은 매우 부정, 1은 매우 긍정으로 해석할 수 있다(Prabowo and Thelwall, 2009; Bonta and Janardhan, 2019). 선행연구들은 감성분석을 활용하여 제품에 대한 고객의 선호도를 분석하고(Bhatt *et al.*, 2015), 영화에 대한 관객의 견해를 파악하는(Singh *et al.*, 2013) 등의 고객 분석 연구를 수행했다.

본 연구에서는 감성분석을 활용하여 서비스 키워드의 감성 점수를 산출하고 산출된 감성점수를 통해 서비스 토픽의 만족도를 계산한다. 문장 내 키워드는 주변 단어, 즉 키워드를 수식하는 단어에 따라 다양한 의미를 가지기 때문에, 같은 키워드라도 문장마다 다른 감성점수를 갖는다(Zhong *et al.*, 2018). 따라서 본 연구에서는 문맥에 따른 키워드의 상이한 감성점수를 반영하고자 특정 키워드가 포함된 문장의 평균 감성점수를 통해 키워드의 감성점수를 산출한다. 이와 같은 방식으로 키워드의 감성점수를 정의하기 위해 우선, 수집한 리뷰데이터 내 모든 문장의 감성점수를 산출한다. 다음으로, 특정 키워드가 포함된 문장들을 찾은 후, 문장들의 감성점수의 평균을 계산하여 해당 키워드의 감성점으로 정의한다.

본 연구는 Google cloud platform이 제공하는 감성분석 패키지인 Natural Language API를 활용해 리뷰 내 모든 문장 33,607개의 감성점수를 계산했다. 키워드의 감성점수를 계산하기 위해 키워드가 포함된 문장들의 평균 감성점수를 산출하여 2,161개의 서비스 키워드의 감성점수를 계산했다. 서비스 토픽

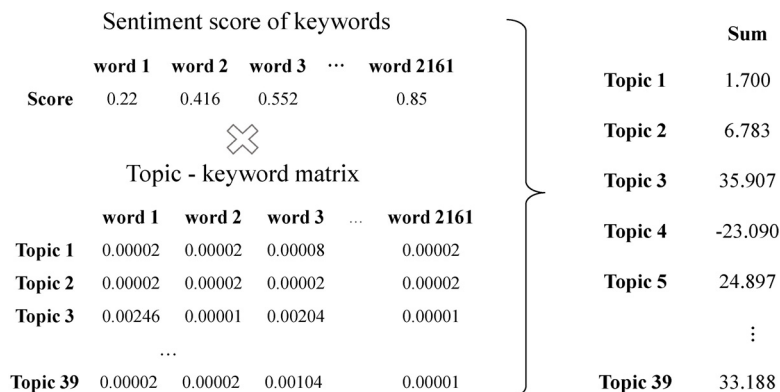


Figure 4. Computing Satisfaction Scores of Service Topics

픽은 기여확률이 높은 주요 키워드들로 정의되기에, 토픽의 만족도에 각 키워드의 기여도를 반영하고자 키워드 감성점수와 토픽-키워드 매트릭스를 곱한다. 도출된 매트릭스는 키워드 감성점수에 토픽 기여확률을 곱한 값들로 이루어진 매트릭스로, 고객들이 정의한 서비스 토픽 별 키워드 가중치가 반영된다. 매트릭스의 토픽 별 모든 키워드의 감성점수를 합하여, 서비스 토픽의 만족도를 계산한다(<Figure 4>). 중요도와 마찬가지로 최대-최소 정규화를 적용하여 각 서비스 토픽의 최종적인 만족도 점수를 산출했으며 서비스 토픽의 평균 만족도 점수는 5.782이고 가장 높은 만족도를 가지는 토픽은 “Fresh meal” 이었다.

3.4 Service Opportunity Map 구축 및 서비스 기회 발굴

본 연구단계에서는 기회 알고리즘을 활용해 서비스 토픽의 기회점수를 산출하고 Service Opportunity Map을 구축하여 밀키트 구독서비스 기업의 서비스 기회를 발굴한다. 기회 알고리즘이란 고객의 니즈를 순위화 하여 비즈니스 기회를 포착하고자 하는 알고리즘으로, 중요하지만 만족되지 않은 고객니즈를 기회로 정의한다(Ulwick, 2005). 따라서 중요도가 높고 만족도가 낮은 토픽을 기회가 존재하는 서비스 토픽으로 정의할 수 있으며, 기회 알고리즘의 수식은 식 (1)과 같다.

$$Opportunity = Importance + \max(Importance - Satisfaction, 0) \quad (1)$$

식 (1)을 활용하여 산출한 각 서비스 토픽의 기회점수 평균은 4.059이고, 가장 높은 기회점수를 가진 토픽은 “Cancellation and refund policy”이었다. 도출된 전체 서비스 토픽의 기회점수는 <Appendix 2>에서 확인할 수 있다.

Service Opportunity Map은 중요도와 만족도를 각 축으로 구축되며, 축들의 평균값을 기준으로 Over-served, Right-served 그리고 Under-served, 세 영역으로 나누어진다. 우선 Over-served 영역에는 중요도는 낮지만, 만족도는 높은 토픽들이 위치한다. 해당 영역의 서비스 토픽들은 고객들이 자주 언급하지는 않지만 높은 만족감을 느끼는 서비스 요소들이다. 그리고 Right-served 영역은 평균에 근접한 중요도와 만족도 점수를 가진 서비스 토픽들이 위치하고 있는 영역으로, 해당 토픽들은 고객들에게 적절한 수준으로 충족된 니즈들이다. 마지막으로, Under-served 영역은 중요도가 높으나 만족도가 낮은 서비스 토픽들이 위치한다. 해당 영역의 서비스 토픽들은 고객들에게 자주 언급되는 토픽들로 중요한 서비스 토픽이지만 고객들의 만족도가 낮은, 즉 충족되지 않은 고객니즈이며 기회점수가 높은 서비스 토픽이기 때문에 타겟 기업의 서비스 기회에 해당하는 토픽들이다.

식별된 서비스 토픽들과 각 토픽의 중요도, 만족도 그리고 기회점수를 통해 HelloFresh의 Service Opportunity Map을 구축했다(<Figure 5>).

Service Opportunity Map의 각 원은 서비스 토픽들이며, 원의 크기는 기회점수에 비례한다. Under-served 영역에는 토픽의 중요도에 비해 만족도가 낮은, 즉 서비스 기회에 해당하는 서

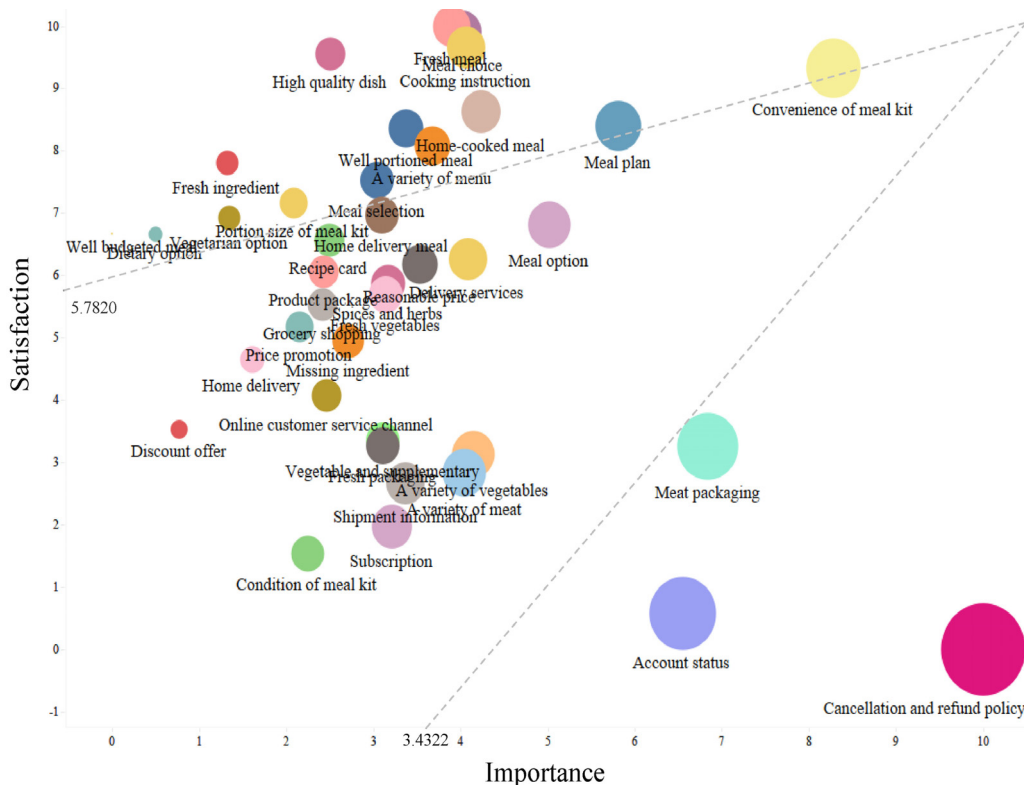


Figure 5. Service Opportunity Map of HelloFresh

비스 토픽들이 분포되어 있으며, 해당 영역에는 세 개의 토픽이 존재한다. 이들은 각각 “Meat packaging”, “Account status” 그리고 “Cancellation and refund policy”이다. 이 가운데 기회점수가 가장 높은 서비스 토픽은 “Cancellation and refund policy”으로, 서비스의 취소 및 환불 정책에 해당하는 토픽이었다.

4. 분석결과 및 토의

본 연구는 온라인 고객 리뷰 마이닝을 통해 밀키트 구독서비스 기업의 고객니즈 및 서비스 기회를 분석했다. 본 장에서는 3장에서 도출한 연구결과를 바탕으로, 서론에서 제시한 세 가지 연구질문에 대한 답을 제시하고자 한다.

4.1 RQ 1 : 밀키트 구독서비스 기업의 고객들은 어떤 요구를 하는가?

첫번째 연구질문은 LDA로 도출된 서비스 토픽들을 분석함으로써 해결할 수 있다. 식별된 39개의 서비스 토픽들은 주요 키워드를 통해 레이블링 되었고 레이블링 된 서비스 토픽들을 통해 타겟 기업에 대한 고객의 요구를 확인할 수 있었다 (<Table 2>).

중요도가 높은 서비스 토픽들을 살펴보면, “Cancellation and refund policy”, “Account status”와 같은 계정 상태 및 지불 관리에 대한 고객들의 니즈를 확인할 수 있었다. 구독서비스를 이용하는 소비자의 경우, 기업과 반복 지불 구조를 갖게 된다 (Tzuo and Weisert, 2018). 소비자는 일정 기간마다 자동으로 지불하게 되며, 구독의 해지를 원할 경우 소비자가 직접 구독 상태를 관리해야 한다. 마찬가지로, 자동으로 결제된 이후 서비스의 중단을 요구할 경우, 환불 요청을 해야한다. 이러한 구

Table 2. Part of Identified Service Topics of Meal Kit Delivery Service

Rank	Topic	Importance score
1	Cancellation and refund policy	10.000
2	Convenience of meal kit	8.278
3	Meat packaging	6.843
4	Account status	6.547
5	Meal plan	5.815
	...	
9	Delivery services	4.081
	...	
35	Vegetarian option	1.343
	...	
38	Dietary option	0.496
39	Well budgeted meal	0.000

독서비스가 갖는 독특한 결제구조로 인해 관련 니즈들이 리뷰로부터 식별된 것으로 파악된다.

또한, “Meat packaging”과 “Delivery services”를 통해 밀키트의 배송과 포장에 대한 소비자의 니즈를 파악할 수 있었다. 밀키트 구독서비스는 단순히 밀키트를 제공하는 것 외에도 식재료를 포장하고 정기적으로 배송하는 서비스를 의미한다. 따라서 기업은 신선식품의 신선도, 육류 포장, 아이스팩, 정시 배송 등과 같은 제품의 배송과 포장과 관련된 다양한 서비스 요소를 고려해야한다. 식별된 서비스 토픽들을 통해, 소비자들은 밀키트의 포장과 배송에 관련된 니즈가 존재하는 것을 확인할 수 있었다.

이외에도 서비스 토픽을 살펴보면, “Vegetarian option”을 통해 고객들이 건강한 식단과 채식 식단을 언급하는 것을 확인할 수 있었다. 실제로, 최근 몇 년 동안 환경 및 건강문제로 인해 채식 식단의 인기가 증가했다(Saari *et al.*, 2021). 소비자들의 선호를 반영하여 기업들은 다양한 채식 제품을 선보였으며, 타겟 기업 또한 채식 및 페스케테리언 옵션을 제공하고 있다.

본 연구질문은 토픽 모델링을 통해 식별된 서비스 토픽들을 분석함에 따라 해결되었다. 서비스 토픽들은 소비자들이 언급하는 키워드에 의해 정의된 것으로, 소비자들의 니즈라고 해석 가능하다. 연구결과를 통해 구독관리, 배송 및 포장 그리고 식단과 같은 타겟 기업의 소비자들의 구체적인 니즈를 확인할 수 있었다. 남은 연구질문들은 파악된 고객니즈들을 기반으로 구축된 Service Opportunity Map, 기회점수 그리고 만족도를 활용하여 해결된다.

4.2 RQ 2 : 밀키트 구독서비스 기업의 서비스 개선방향은 무엇인가?

고객 관점에서 탐색된 서비스 개선방안을 기업의 의사결정에 활용할 경우, 기업은 경쟁우위를 확보할 수 있다(Park and Yoon, 2015). 밀키트 구독서비스 기업의 서비스 개선방향은 Under-served 영역에 위치한 서비스 토픽을 통해 제시할 수 있다. Under-served 영역에 존재하는 토픽들은 고객들에게 자주 언급되지만 충족되지 않은 니즈들이기 때문에 기회가 존재하는 니즈로 해석할 수 있다. 실제로, <Table 3>는 Under-served 영역에 존재하는 서비스 토픽으로, 본 연구가 발굴한 서비스 기회이다. 발굴한 서비스 기회에 대한 고객관점의 서비스 개선방향을 찾기 위하여 관련 리뷰를 활용한다. 문서-토픽 매트릭스의 기여확률을 통해 토픽에 대한 관련 리뷰를 찾을 수 있어, 이를 통해 소비자의 구체적인 불만사항을 확인할 수 있다.

서비스 기회에 해당하는 토픽은 “Cancellation and refund policy”, “Account status” 그리고 “Meat packaging”이다. 이중, 기회점수가 가장 높은 토픽은 “Cancellation and refund policy”로, 구독 취소 및 환불 정책에 대한 토픽이었다. 해당 토픽과 관련 있는 실제 리뷰를 확인하니, “... I had been using HelloFresh for a few weeks when I decided to put my subscription

Table 3. Service Topics in the Under-served Area

Cancellation and refund policy		Account status		Meat packaging	
Opportunity score	20.000	Opportunity score	12.517	Opportunity score	10.426
keyword	sentiment score	keyword	sentiment score	keyword	sentiment score
cancel	-0.629	charge	-0.624	meat	-0.118
refund	-0.534	account	-0.492	meats	0.105
account	-0.492	info	-0.470	open	-0.316
money	-0.119	credit card	-0.600	leak	-0.600
charge	-0.624	date	-0.317	tomato	-0.448
cancelled	-0.696	information	-0.482	package	-0.047
canceled	-0.640	website	-0.296	leaking	-0.618
refunded	-0.314	charges	-0.566	tomatoes	-0.447
cancelling	-0.650	payment	-0.490	fridge	-0.016
money back	-0.669	mistake	-0.362	meat pack	-0.472

on pause using the app. Then, a few days ago they resumed my subscription and charged my credit card \$60. ...” 처럼 서비스의 구독을 취소했음에도 불구하고 지속적으로 상품이 배달왔다는 리뷰를 확인할 수 있었다. 또한 다른 리뷰에서는 서비스 금액이 이중청구 되어 이를 해결하기 위해 서비스 센터에 전화했지만 환불받지 못한 경험이 언급되었다. 서비스 기업의 서비스 품질은 고객 유지에 있어 중요한 측면 중 하나이다 (Venetis and Ghauri, 2004). 밀키트 구독서비스 기업은 서비스 품질을 향상시킴에 따라 고객유지율을 증진시킬 수 있을 것이다. 따라서, 구독 취소 및 환불 정책에 대한 서비스 개선이 필요할 것이다.

또다른 Under-served 영역에 위치한 토픽인 ”Meat packaging”의 주요 키워드는 meat, open, leak 등으로 육류 포장에서 육즙이 새거나, 포장이 열려서 배송되는 것을 확인할 수 있다. 토픽과 관련된 고객 리뷰를 살펴보면, “... ongoing issues with the chicken leaking all over everything. My last box was a complete waste as 3 packages of chicken broke open and saturated everything in rotten chicken juices ...”와 같이 닭고기 육즙이 새어 다른 식재료에 묻어있었고, 이러한 문제가 반복적으로 발생했다는 리뷰를 확인할 수 있었다. 고기의 포장이 열려있었거나 상해있었다고 언급된 리뷰가 다양하게 존재했는데, 실제로 육류 포장과 관련된 고객의 불만사항이 존재함을 파악할 수 있었다. 따라서 기업은 육류 포장과 관련된 기술을 확보하거나, 새로운 배송방법을 개발하는 등의 노력이 필요할 것이다.

이처럼 본 연구가 도출한 Service Opportunity Map의 Under-served 영역에 위치한 서비스 토픽을 통해 기업의 서비스 개선방향을 제시할 수 있었다. 서비스 토픽과 리뷰를 통해 소비자들의 직접적인 불만을 확인할 수 있었으며 기회점수를 바탕으로 타겟 기업의 서비스 개선의 순위화가 가능했다. 본 연구에서 분석된 서비스 기회, “Cancellation and refund policy”,

“Account status” 그리고 “Meat packaging”은 기업의 서비스 개선 전략을 수립할 시 가장 우선적으로 고려되어야 할 서비스들이다.

4.3 RQ 3 : 고객은 밀키트 구독서비스의 어떤 요소에서 만족감을 느끼는가?

해당 연구질문에 대한 답은 Service Opportunity Map의 Over-served 영역에 위치한 서비스 토픽들을 활용하여 제시할 수 있다. Over-served에 위치한 서비스 토픽들의 경우, 중요도에 비해 고객 만족도가 높은 토픽들로 파급적 혁신의 가능성을 가진 서비스 요소들이다(Ulwick, 2009). 또한, 해당 서비스 토픽들을 통해 타겟 기업의 소비자들이 만족감을 느끼는 서비스 요소들을 확인할 수 있으며 기업들은 이러한 서비스 요소들을 지속적으로 제공함으로써 로열티가 높은 고객들을 유지할 수 있다(Choi et al., 2020). 본 장에서는 Over-served 영역에 위치한 서비스 토픽 15개 중 일부를 선택하여 연구질문에 대한 답을 제시하고자 한다(<Table 4>).

가장 높은 중요도 점수를 가진 “Fresh meal”의 주요 키워드를 살펴보면 fresh meal, delicious meal, great experience 등으로 타겟 기업에 대한 긍정적인 키워드들을 확인할 수 있다. 실제 리뷰를 살펴보니, “...Wonderful selections, fresh food, great experience. ... My husband and I have thoroughly enjoyed preparing the meals together and look forward to their arrival!”와 같은 리뷰를 확인할 수 있었는데, 소비자가 밀키트 메뉴, 신선한 음식 등 다양한 서비스 요소에 만족하고 있음을 확인할 수 있었다.

다음으로, “Cooking instruction”은 밀키트와 함께 배송되는 안내서에 대한 서비스 토픽이다. 실제로 “... Easy to follow step by step instructions. And have learned a few cooking strategies along the way...” 와 같은 긍정적인 리뷰를 확인할 수 있었다. HelloFresh가

Table 4. Examples of Service Topics in the Over-served Area

Fresh meal		Cooking instruction		Convenience of meal kit	
Satisfaction score	10.000	Satisfaction score	9.655	Satisfaction score	9.329
keyword	sentiment score	keyword	sentiment score	keyword	sentiment score
fresh meal	0.573	instruction	0.581	cooking	0.552
fresh meals	0.606	instructions	0.604	convenience	0.514
calorie	0.084	ratio	0.321	packaging	0.096
delicious meal	0.794	preparation	0.593	enjoy cooking	0.735
delicious meals	0.779	meal prep	0.341	taste	0.55
great experience	0.775	cooking	0.552	recyclable	0.338
great meal	0.769	step instruction	0.852	groceries	0.267
great meals	0.724	step instructions	0.861	recycle	-0.106
absolutely delicious	0.870	lettuce	-0.432	inconvenience	-0.264
recommendation	0.235	meal preparation	0.568	money	-0.119

제공하는 안내서는 단계별 요리 방법, 과정 별 사진 그리고 필요한 재료에 대한 용량이 자세하게 작성되어 있어 소비자들은 해당 서비스 요소에 만족감을 느끼고 있음을 파악했다. 또한, “Convenience of meal kit”는 밀키트의 편리함에 대한 서비스 토픽으로 Over-served 영역에 위치한 토픽 중 가장 높은 중요도를 가진다. 토픽의 높은 만족도를 통해 밀키트 구독서비스는 현대인들의 편리한 식사에 대한 니즈를 충족시키는 식품 서비스라 해석할 수 있었다. 더불어 토픽의 높은 중요도는 리뷰 내 언급이 많다는 것을 의미하므로, 밀키트 구독서비스를 구독함으로써 느끼는 편리함은 소비자들에게 중요한 요소라 해석된다.

이처럼 Over-served 영역 내 위치한 서비스 토픽들을 활용하여 연구질문에 대한 답을 제시할 수 있었다. 따라서 본 연구결과를 통해 기업은 소비자들이 밀키트 구독서비스의 만족감을 느끼는 서비스를 확인할 수 있으며, 고객이 밀키트 구독 서비스를 지속적으로 사용하는 이유를 파악할 수 있다.

5. 결론

본 연구는 리뷰 마이닝을 통해 밀키트 구독서비스 기업의 서비스 기회를 발굴하였다. 먼저, HelloFresh의 온라인 리뷰를 수집하여 토픽 모델링을 수행하였고 39개의 서비스 토픽을 정의하였다. 정의된 39개의 서비스 토픽을 통해 밀키트 구독서비스 고객들의 니즈를 식별하였다. 다음으로, 기회 알고리즘을 활용하여 Service Opportunity Map을 구축하였고, 중요하지만 만족도가 낮은 서비스 토픽이 포함되어 있는 Under-served 영역을 분석하여 기업의 서비스 개선방향을 제시하였다. 마지막으로, 중요도는 낮지만 만족도가 높은 서비스 토픽이 포함되어 있는 Over-served 영역을 분석하여 고객들이 만족하고 있는 기업의 서비스를 파악하였다.

본 연구는 다음과 같은 기여점을 가진다. 먼저, 본 연구는 토픽 모델링을 통해 리뷰에 포함된 서비스 토픽을 식별했다. LDA를 통해 “Price promotion”, “Fresh vegetables” 등과 같은 다양한 서비스 토픽을 식별함에 따라, 고객이 작성한 리뷰 내 존재하는 다양한 고객니즈를 확인하였다. 이를 통해 기업의 심도있는 고객 이해에 도움을 줄 것으로 기대된다. 다음으로, 본 연구는 밀키트 구독서비스의 리뷰 마이닝을 통해 고객니즈를 분석하고 서비스 기회를 계량적으로 발굴했다는 측면에서 기여점이 존재한다. 만족도와 중요도, 두 개의 지표를 통해 고객니즈를 보다 다양한 관점에서 해석할 수 있었다. 또한, 지표를 통해 고객니즈를 평가하여 서비스 기회를 도출하였기에, 밀키트 구독서비스 소비자에 대한 정량적인 분석 결과를 제시한다. 마지막으로, 본 연구에서 발굴된 서비스 기회는 기업에게 고객 관점의 서비스 개선 방향을 제시한다. 분석된 서비스 기회들은 고객 관점에서 우선적으로 개선되어야 할 서비스들로, 기업의 기획 및 마케팅 프로세스에 활용될 것으로 기대된다.

하지만, 본 연구에도 몇 가지 한계점이 존재한다. 우선, 본 연구에서는 밀키트 구독서비스 기업의 서비스 기회 발굴 과정에서 리뷰 데이터만을 분석에 활용했다. 소셜미디어 데이터 원천으로는 블로그, 소셜 네트워크 서비스 등의 다양한 매체가 존재한다. 여러 매체로부터 데이터를 수집하여 연구에 활용할 경우, 밀키트 구독서비스 기업의 소비자들의 보다 다양한 의견을 수렴할 수 있을 것으로 기대된다. 다음으로, 본 연구는 리뷰의 텍스트 정보만을 활용했다. 리뷰 데이터 작성자의 성별, 나이 등과 같은 인적정보를 분석 과정에서 활용하면 기업에게 더욱 풍부한 분석결과를 제시할 수 있을 것이다. 마지막으로, 본 연구에서 활용된 두 가지 지표, 만족도와 중요도는 개선의 여지가 존재한다. 만족도와 중요도 외에 추가적인 지표를 활용하여 서비스 개선방향을 도출한다면 더욱 풍부한 결과를 얻을 수 있을 것으로 기대된다.

참고문헌

- Ahn, D.-K. and Kim, S.-I. (2022), A study on the buying behavior of meal kits according to the lifestyle of the MZ generation, *Journal of Digital Convergence*, **20**(2), 367-373.
- Baek, N. G. (2021), A Study on Effect of The Subscription?Based Online Service on Consumer Perception Value, Repurchase Intention, *The Academy of Customer Satisfaction Management*, **23**(2), 49-71.
- Bhatt, A., Patel, A., Chheda, H., and Gawande, K. (2015), Amazon review classification and sentiment analysis, *International Journal of Computer Science and Information Technologies*, **6**(6), 5107-5110.
- Bholowalia, P. and Kumar, A. (2014), EBK-means: A clustering technique based on elbow method and k-means in WSN, *International Journal of Computer Applications*, **105**(9).
- Blei, D. M., Ng, A. Y., and Jordan, M. I. (2003), Latent dirichlet allocation, *Journal of Machine Learning Research*, **3**(Jan), 993-1022.
- Bonta, V. and Janardhan, N. K. N. (2019), A comprehensive study on lexicon based approaches for sentiment analysis, *Asian Journal of Computer Science and Technology*, **8**(S2), 1-6.
- Buckley, M., Cowan, C., and McCarthy, M. (2007), The convenience food market in Great Britain: Convenience food lifestyle (CFL) segments, *Appetite*, **49**(3), 600-617.
- Chen, K. J., Yeh, T. M., Pai, F. Y., and Chen, D. F. (2018), Integrating refined kano model and QFD for service quality improvement in healthy fast-food chain restaurants, *International Journal of Environmental Research and Public Health*, **15**(7), 1310.
- Choi, H. and Yeon, K. (2022), A Case Study on Text Analysis Using Meal Kit Product Review Data, *Journal of the Korea Contents Association*, **22**(5), 1-14
- Choi, J., Oh, S., Yoon, J., Lee, J. M., and Coh, B. Y. (2020), Identification of time-evolving product opportunities via social media mining, *Technological Forecasting and Social Change*, **156**, 120045.
- Choi, T., Lee, M., and Kim, D. (2020), The Effect of Selection Attributes of Meal-Kit Home Replacement Food (HMR) on Satisfaction and Repurchase Intention, *Culinary Science & Hospitality Research*, **26**(5), 119-128.
- Chung, H. and Kim, C. (2020), A Study on the Meal Kit Product Selection Attributes on Purchasing Behavior and Satisfaction, *Journal of the Korea Contents Association*, **20**(6), 381-391.
- Daniels, S. and Glorieux, I. (2015), Convenience, food and family lives. A socio-typological study of household food expenditures in 21st-century Belgium, *Appetite*, **94**, 54-61.
- Goffin, K., Varnes, C. J., van der Hoven, C., and Koners, U. (2012), Beyond the voice of the customer: Ethnographic market research, *Research-Technology Management*, **55**(4), 45-53.
- Gundecha, P. and Liu, H. (2012), Mining social media: A brief introduction, *New Directions in Informatics, Optimization, Logistics, and Production*, 1-17.
- Heard, B. R., Bandekar, M., Vassar, B., and Miller, S. A. (2019), Comparison of life cycle environmental impacts from meal kits and grocery store meals, *Resources, Conservation and Recycling*, **147**, 189-200.
- Ibrahim, N. F. and Wang, X. (2019), A text analytics approach for online retailing service improvement: Evidence from Twitter, *Decision Support Systems*, **121**, 37-50.
- Janzer, A. (2020), *Subscription marketing: Strategies for nurturing customers in a world of churn*, Cuesta Park Consulting.
- Jeong, B., Yoon, J., and Lee, J. M. (2019), Social media mining for product planning: A product opportunity mining approach based on topic modeling and sentiment analysis, *International Journal of Information Management*, **48**, 280-290.
- Jo, P. K. (2016), The effects of the economic characteristics of single-person households on the food service industry, *Korean Journal of Community Nutrition*, **21**(4), 321-331.
- Khan, S. A. and Sowards, S. K. (2018), It's not just dinner: Meal delivery kits as food media for food citizens, *Frontiers in Communication*, **3**, 39.
- Kim, J. and Lim, C. (2021), Customer complaints monitoring with customer review data analytics: An integrated method of sentiment and statistical process control analyses, *Advanced Engineering Informatics*, **49**, 101304.
- Ko, N., Jeong, B., Choi, S., and Yoon, J. (2017), Identifying product opportunities using social media mining: Application of topic modeling and chance discovery theory, *IEEE Access*, **6**, 1680-1693.
- Lee, G. (2021), A study on the consumer perceptions of meal-kits using big data: Anticipating after the COVID-19 pandemic ends, *International Journal of Tourism and Hospitality Research*, **35**(11), 227-239.
- Lee, K. and Park, K. (2021), Effects of Selection Attributes of HMR's Meal Kit Products on the Purchasing Behavior and Consumer Satisfaction by the COVID-19, *Food Service Industry Journal*, **17**(3), 277-289.
- Lee, S. and Kim, D. (2022), A Study on the Consumption of Meal-Kit by Dietary Lifestyle Types: Focused on Moderating Effect of Gender and Age, *Journal of Tourism and Leisure Research*, **34**(1), 389-411.
- Lee, Y. N., Jang, H. J., Choi, Y. W., Choi, Y. S., and Oh, J. E. (2021), A study on the consumer perception and importance-performance analysis of the vegetarian meal-kit development, *The Journal of the Korea Contents Association*, **21**(3), 324-335.
- Liu, Y., Jiang, C., and Zhao, H. (2019), Assessing product competitive advantages from the perspective of customers by mining user-generated content on social media, *Decision Support Systems*, **123**, 113079.
- Mahmoud, M. A., Hinson, R. E., and Anim, P. A. (2017), Service innovation and customer satisfaction: the role of customer value creation, *European Journal of Innovation Management*.
- Medhat, W., Hassan, A., and Korashy, H. (2014), Sentiment analysis algorithms and applications: A survey, *Ain Shams Engineering Journal*, **5**(4), 1093-1113.
- Nam, S. and Lee, H. (2019), A text analytics-based importance performance analysis and its application to airline service, *Sustainability*, 6153.
- Oh, H. (2001), Revisiting importance-performance analysis, *Tourism Management*, **22**(6), 617-627.
- Park, H. and Yoon, J. (2015), A chance discovery-based approach for new product-service system (PSS) concepts, *Service Business*, **9**(1), 115-135.
- Park, J. (2019), A Study on Issue Tracking on Multi-cultural Studies Using Topic Modeling, *Journal of the Korean Society for Library and Information Science*, **53**(3), 273-289.
- Park, S. and Hwang, J. (2021), A study on the effect of Meal-Kit selection properties on consumers' purchasing, *The Journal of Social Science*, **28**(1), 157-175.
- Prabowo, R. and Thelwall, M. (2009), Sentiment analysis: A combined approach, *Journal of Informetrics*, **3**(2), 143-157.
- Pyon, C. U., Woo, J. Y., and Park, S. C. (2011), Service improvement by business process management using customer complaints in financial

- service industry, *Expert Systems with Applications*, **38**(4), 3267-3279.
- Qi, J., Zhang, Z., Jeon, S., and Zhou, Y. (2016), Mining customer requirements from online reviews: A product improvement perspective, *Information & Management*, **53**(8), 951-963.
- Rai, R. (2012), Identifying key product attributes and their importance levels from online customer reviews, *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, Vol. 45028. American Society of Mechanical Engineers.
- Rose, S., Engel, D., Cramer, N., and Cowley, W. (2010), Automatic keyword extraction from individual documents, *Text Mining: Applications and Theory*, **1**(1-20), 10-1002.
- Saari, U. A., Herstatt, C., Tiwari, R., Dedehayir, O., and Mäkinen, S. J. (2021), The vegan trend and the microfoundations of institutional change: A commentary on food producers' sustainable innovation journeys in Europe, *Trends in Food Science & Technology*, **107**, 161-167.
- Singh, V. K., Piryani, R., Uddin, A., and Waila, P. (2013, March), Sentiment analysis of movie reviews: A new feature-based heuristic for aspect-level sentiment classification. In *2013 International mutli-conference on automation, computing, communication, control and compressed sensing (imac4s)*, IEEE, 712-717.
- Syakur, M. A., Khotimah, B. K., Rochman, E. M. S., and Satoto, B. D. (2018, April), Integration k-means clustering method and elbow method for identification of the best customer profile cluster. In *IOP Conference Series: Materials Science and Engineering* (Vol. 336, No. 1, p. 012017), IOP Publishing.
- Tzuo, T. and Weisert, G. (2018), *Subscribed: Why the subscription model will be your company's future-and what to do about it*, Penguin.
- Ulwick, A. (2005), *What customers want* (Vol. 8), New York, NY: McGraw-Hill Professional Publishing.
- Ulwick, A. W. (2009), *What Is Outcome-Driven Innovation®(ODI)?*, White Paper.
- Venetis, K. A. and Ghauri, P. N. (2004), Service quality and customer retention: building long-term relationships, *European Journal of Marketing*.
- Wang, A., Zhang, Q., Zhao, S., Lu, X., and Peng, Z. (2020), A review-driven customer preference measurement model for product improvement: sentiment-based importance-performance analysis, *Information Systems and e-Business Management*, **18**(1), 61-88.
- Wang, B., Liu, S., Ding, K., Liu, Z., and Xu, J. (2014), Identifying technological topics and institution-topic distribution probability for patent competitive intelligence analysis: A case study in LTE technology, *Scientometrics*, **101**(1), 685-704.
- Wu, J., Wang, Y., Zhang, R., and Cai, J. (2018), An approach to discovering product/service improvement priorities: Using dynamic importance-performance analysis, *Sustainability*, **10**(10), 3564.
- Yang, C., Wu, L., Tan, K., Yu, C., Zhou, Y., Tao, Y., and Song, Y. (2021), Online user review analysis for product evaluation and improvement, *Journal of Theoretical and Applied Electronic Commerce Research*, **16**(5), 1598-1611.
- Yang, Y., Kwon, Y., Park, Y., and Yun, Y. (2019), Importance-Performance Analysis Regarding Selective Attribution of Meal-Kit Products, *Journal of the East Asian Society of Dietary Life*, **29**(6), 519-528.
- Yi, N. and Choi, J. (2019), A Study of the Consumer Perception of Meal Kit Using Big Data Analysis, *Food Service Industry Journal*, **15**(4), 211-222.
- Zhan, Y., Han, R., Tse, M., Ali, M. H., and Hu, J. (2021), A social media analytic framework for improving operations and service management: A study of the retail pharmacy industry, *Technological Forecasting and Social Change*, **163**, 120504.
- Zhao, W., Chen, J. J., Perkins, R., Liu, Z., Ge, W., Ding, Y., and Zou, W. (2015, December), A heuristic approach to determine an appropriate number of topics in topic modeling, In *BMC bioinformatics* (Vol. 16, No. 13, pp. 1-10). BioMed Central.
- Zhao, W., Zou, W., and Chen, J. J. (2014, December), Topic modeling for cluster analysis of large biological and medical datasets. In *BMC bioinformatics* (Vol. 15, No. 11, pp. 1-11). BioMed Central.
- Zhong, Q., Liang, S., Cui, L., Chan, H. K., and Qiu, Y. (2018), Using online reviews to explore consumer purchasing behaviour in different cultural settings, *Kybernetes*, **48**(6), 1242-1263.

<Appendix>

Appendix Table 1. Service Topics and Major Keywords

Index	Topic	Major Keywords
1	Price promotion	code(0.138), promo(0.100), protein(0.099), fresh box(0.072), kids(0.060), HelloFresh box(0.056), Groupon(0.047), promotion(0.045), carb(0.040), proteins(0.033)
2	Fresh vegetables	veggie(0.250), veggies(0.165), sauce(0.111), sauces(0.062), meat(0.046), freshness(0.034), fresh veg(0.022), herb(0.016), lacking(0.014), carb(0.012)
3	Home-cooked meal	dinner(0.603), FedEx(0.042), great service(0.029), families(0.020), subscribing(0.020), taste(0.020), saving money(0.016), money(0.015), spinach(0.013), make dinner(0.012)
4	Shipment information	ship(0.419), shipment(0.155), shipping(0.132), shipments(0.047), account(0.015), cost(0.014), free shipping(0.012), meal kit service(0.010), fresh recipe(0.010), management(0.009)
5	Meal selection	selection(0.217), selections(0.100), restaurant(0.080), food delivery(0.050), delivery service(0.043), edible(0.040), meal selection(0.038), services(0.034), food delivery service(0.033), competitor(0.024)
6	Meal plan	plan(0.407), meal plan(0.104), planning(0.094), meal planning(0.042), shopping(0.032), healthy meal(0.028), healthy meals(0.020), cooking(0.018), meal plans(0.018), nutritious(0.015)
7	Delivery services	items(0.254), services(0.113), delivery service(0.099), meal delivery(0.074), meal delivery service(0.046), missing item(0.044), meal service(0.042), missing items(0.036), delivery services(0.030), meal delivery services(0.018)
8	Home delivery	house(0.155), cheese(0.100), delay(0.081), called customer service(0.072), delayed(0.054), household(0.034), days late(0.033), delicious food(0.031), free trial(0.029), staff(0.027)
9	Fresh ingredient	dishes(0.240), fresh ingredient(0.124), fresh ingredients(0.122), members(0.053), great food(0.043), membership(0.035), steaks(0.033), errors(0.032), weekly box(0.019), packing(0.018)
10	Meal choice	choice(0.255), flavor(0.236), flavors(0.099), meal choice(0.049), meal choices(0.042), taste(0.033), disappointment(0.016), fresh subscription(0.015), freshness(0.012), great tasting(0.012)
11	Discount offer	discount(0.221), discounts(0.041), PayPal(0.032), refunds(0.032), roast(0.032), discounted(0.031), accounts(0.029), good meal(0.025), avocado(0.024), customers(0.023)
12	Dietary option	diet(0.123), gourmet(0.100), salad(0.075), dietary(0.039), good variety(0.031), gourmet meal(0.031), dinner time(0.031), salads(0.023), restrictions(0.022), subscribe(0.021)
13	Portion size of meal kit	deliveries(0.242), portion(0.092), portion size(0.086), sizes(0.063), portion sizes(0.053), great customer(0.027), great customer service(0.026), home delivery(0.025), good food(0.024), quality ingredients(0.016)
14	Subscription	script(0.273), subscription(0.266), cancel(0.111), free box(0.078), cancelled(0.024), canceled(0.022), prices(0.018), free boxes(0.016), cost(0.014), coupon(0.012)
15	Condition of meal kit	rotten(0.154), spoil(0.152), address(0.105), pepper(0.070), open(0.035), spicy(0.032), fresh produce(0.019), rotten produce(0.019), onion(0.017), rotten food(0.015)
16	Fresh packaging	packs(0.077), ice pack(0.065), program(0.063), meat(0.057), ice packs(0.053), packaging(0.048), juice(0.045), leak(0.042), paper(0.041), bags(0.038)
17	Account status	charge(0.183), account(0.137), info(0.091), credit card(0.089), date(0.056) information(0.039), website(0.035), charges(0.026), payment(0.024), mistake(0.024)
18	Product package	package(0.351), packaged(0.162), home chef(0.033), date(0.020), supplement(0.018), dates(0.018), menu selection(0.016), rotted(0.015), money(0.014), subscription service(0.014)
19	Missing ingredient	missing ingredient(0.078), bags(0.075), missing ingredients(0.071), recommend HelloFresh(0.066), plastic(0.057), taste(0.050), good quality(0.045), customer service rep(0.045), great recipes(0.027) service representative(0.025)

Index	Topic	Major Keywords
20	Spices and herbs	spice(0.118), schedule(0.102), spices(0.090), burger(0.084), herb(0.062), scheduled(0.049), prep time(0.039), taco(0.036), fresh herbs(0.025), hamburger(0.020)
21	Cooking instruction	instruction(0.239), instructions(0.225), ratio(0.128), preparation(0.092), meal prep(0.046), cooking(0.033), step instruction(0.019), step instructions(0.019) lettuce(0.018), meal preparation(0.016)
22	Fresh meal	fresh meal(0.146), fresh meals(0.123), calorie(0.112), delicious meal(0.085), delicious meals(0.067), great experience(0.053), great meal(0.040), great meals(0.026), absolutely delicious(0.024), recommendation(0.019)
23	A variety of vegetables	potato(0.134), potatoes(0.112), green(0.107), carrot(0.079), carrots(0.068), beans(0.054), green bean(0.047), green beans(0.045), onion(0.027), sweet potato(0.027)
24	Well budgeted meal	budget(0.115), schedule(0.063), good thing(0.049), picky eater(0.045), chopping(0.043), wide variety(0.041), premium meal(0.036), return(0.033), busy schedule(0.032), schedules(0.032)
25	Meal option	option(0.347), options(0.265), fish(0.036), gluten(0.030), meal options(0.026), gluten free(0.014), menu options(0.014), seafood(0.013), family member(0.013), vegan(0.012)
26	Vegetarian option	vegetarian(0.202), option(0.088), vegetarian option(0.044), options(0.042), cancellation(0.038), skip week(0.035), skip weeks(0.034), nutrition(0.033), vegetarian options(0.031), weeknight(0.025)
27	Convenience of meal kit	cooking(0.337), convenience(0.123), packaging(0.103), enjoy cooking(0.035), taste(0.028), recyclable(0.028), groceries(0.023), recycle(0.020), inconvenience(0.019), money(0.019)
28	Grocery shopping	gift(0.163), shopping(0.145), grocery(0.108), grocery shop(0.103), gift card(0.087), grocery shopping(0.075), canceling(0.037), Christmas(0.021), lifesaver(0.017), line chat(0.017)
29	Recipe card	cards(0.189), recipe card(0.175), recipe cards(0.126), coupon(0.114), subscribe(0.050), coupons(0.028), customers(0.021), codes(0.014), customer service team(0.013), preparation time(0.012)
30	Vegetable and supplementary	vegetable(0.287), vegetables(0.246), meat(0.034), herb(0.021), rotten(0.019), subscriptions(0.019), meat packaging(0.013), partial(0.012), open(0.011), cooking fun(0.011)
31	High quality dish	health(0.401), cooking(0.109), high quality(0.056), food waste(0.026), asparagus(0.024), plan meals(0.023), amounts(0.021), cooking skill(0.020), cooking skills(0.019), cost(0.019)
32	Reasonable price	price(0.362), cost(0.075), lunch(0.059), reasonable(0.044), fresh food(0.042), pricey(0.028), organic(0.027), full price(0.026), single person(0.022), life easier(0.020)
33	Home delivery meal	grocery(0.282), grocery store(0.221), free meal(0.077), free meals(0.054), pandemic(0.037), money(0.035), covid(0.034), convenience(0.022), food quality(0.017), making dinner(0.013)
34	Online customer service channel	website(0.134), blue apron(0.117), chat(0.090), contacted customer service(0.068), choice(0.066), response(0.060), notification(0.036), free week(0.035), condition(0.028), pricing(0.024)
35	A variety of menu	menu(0.343), menus(0.109), seasoning(0.043), vegetarian meal(0.029), vegetarian(0.029), day late(0.029), favor(0.027), vegetarian meals(0.024), seasonings(0.024), cooking technique(0.023)
36	Meat packaging	meat(0.225), meats(0.066), open(0.064), leak(0.053), tomato(0.047), package(0.040), leaking(0.037), tomatoes(0.037), fridge(0.031), meat pack(0.029)
37	A variety of meat	chicken(0.292), beef(0.105), steak(0.086), meat(0.056), customers(0.052), breast(0.034), ground beef(0.024), fish(0.017), loyal customer(0.016), breasts(0.014)
38	Cancellation and refund policy	cancel(0.238), refund(0.150), account(0.087), money(0.076), charge(0.065), cancelled(0.064), canceled(0.050), refunded(0.026), cancelling(0.011), money back(0.011)
39	Well portioned meal	portion(0.338), portions(0.199), ripe(0.054), portioned(0.051), fresh experience(0.025), apple(0.017), complained(0.014), pantry(0.013), taste(0.013), overripe(0.013)

Appendix Table 2. Opportunity Scores of 39 Service Topics

Index	Topic	Importance score	Satisfaction score	Opportunity score	Index	Topic	Importance score	Satisfaction score	Opportunity score
1	Price promotion	2.146	5.171	2.146	15	Condition of meal kit	2.243	1.533	2.952
2	Fresh vegetables	3.146	5.685	3.146	16	Fresh packaging	3.110	3.267	3.110
3	Home-cooked meal	4.241	8.631	4.241	17	Account status	6.547	0.577	12.517
4	Shipment information	3.372	2.664	4.080	18	Product package	2.433	6.060	2.433
5	Meal selection	3.040	7.517	3.040	19	Missing ingredient	2.709	4.938	2.709
6	Meal plan	5.815	8.404	5.815	20	Spices and herbs	3.167	5.873	3.167
7	Delivery services	4.081	6.265	4.081	21	Cooking instruction	4.064	9.655	4.064
8	Home delivery	1.606	4.647	1.606	22	Fresh meal	3.893	10.000	3.893
9	Fresh ingredient	1.320	7.797	1.320	23	A variety of vegetables	4.145	3.133	5.157
10	Meal choice	4.025	9.901	4.025	24	Well budgeted meal	0.000	6.668	0.000
11	Discount offer	0.770	3.525	0.770	25	Meal option	5.022	6.806	5.022
12	Dietary option	0.496	6.664	0.496	26	Vegetarian option	1.343	6.924	1.343
13	Portion size of meal kit	2.082	7.156	2.082	27	Convenience of meal kit	8.278	9.330	8.278
14	Subscription	3.209	1.974	4.445	28	Grocery shopping	2.419	5.529	2.419
29	Recipe card	2.490	6.572	2.490	35	A variety of menu	3.679	8.069	3.679
30	Vegetable and supplementary	3.109	3.347	3.109	36	Meat packaging	6.843	3.260	10.426
31	High quality dish	2.505	9.559	2.505	37	A variety of meat	4.045	2.834	5.256
32	Reasonable price	3.537	6.173	3.537	38	Cancellation and refund policy	10.000	0.000	20.000
33	Home delivery meal	3.095	6.968	3.095	39	Well portioned meal	3.372	8.356	3.372
34	Online customer service channel	2.459	4.067	2.459		Average	3.432	5.782	4.059

저자소개

정재연 : 건국대학교 응용통계학과에서 2022년 학사학위를 취득하고 건국대학교 산업공학과 석사과정에 재학 중이다. 주요 연구관심분야는 Social media mining for business opportunities, Natural language processing for business intelligence이다.

이지호 : 건국대학교 산업공학과, 컴퓨터공학과에서 2019년 학사학위를 취득하고 건국대학교 산업공학과 석박사통합과정에 재학 중이다. 주요 연구관심분야는 Machine learning-based prediction/decision system, Computational customer analysis,

Computational patent analysis, Natural language processing for business intelligence이다.

윤장혁 : POSTECH 산업공학과에서 학사, 석사 학위를 취득한 후, LG CNS에서 4년 간 재직하였으며, POSTECH 산업경영공학과에서 박사 학위를 취득하였다. 한국지식재산연구원을 거쳐 현재는 건국대학교 산업공학과 정교수로 재직 중이다. 주요 연구분야는 대량 데이터 분석 기반의 Business intelligence, Patent analytics, Social media analytics, Industrial artificial intelligence이다.