JKIICE Journal of the Korea Institute of Information and Communication Engineering

한국정보통신학회논문지 Vol. 29, No. 6: 745~754, June 2025

자연어 처리 기반 고객 불만 분류 및 AS 카테고리 선별 방법 연구

원종수[1] · 나인[2] · 배장원[3*]

NLP-based Classification of Customer Complaints and Categorization for After-Sales Services

Jong-Su Won[1] · In Na[2] · Jang Won Bae[3*]

[1]Ph.D Student, Department of Industrial Management, Korea University of Technology and Education, Cheonan 31253 Korea
[2]Undergraduate Student, Department of Industrial Management, Korea University of Technology and Education, Cheonan 31253 Korea
[3]*Associate Professor, Department of Industrial Management, Korea University of Technology and Education, Cheonan 31253 Korea

요 약

기업에서는 다양한 고객 불만 데이터를 수집하고 있고, 이를 통해서 서비스 품질 관리와 고객 만족도 향상을 위한 활용에 대한 필요성이 높아지고 있다. 이를 위해 본 연구에서는 자연어 처리 방법을 활용하여 고객 불만 데이터를 효과적으로 분석하는 방법을 제안한다. 구체적으로 고객 불만 데이터를 일반 문의(Inquiry)와 서비스 불만(Claim) 민원을 분류하고, 특히, AS와 같은 추가 관리가 필요한 불만 민원 데이터는 카테고리를 분류하여 효과적인 민원 대응이 가능하도록 지원한다. 또한, 제안한 방법은 BERT 및 KoBERT 기반의 자연어 처리 모델로 구현되었으며, 실험 결과 KoBERT 모델이 AS 카테고리 선별에서 F1 Score 0.7046으로 BERT 모델(0.7017)보다 우수한 성능을 보였다. 제안한 방법을 실제 통신 기업의 민원 데이터에 적용한 결과, 인터넷 품질(47.3%)과 TV 품질(30.8%) 관련 불만이 전체의 78.1%를 차지하는 것을 확인하였고, 제안한 방법이 기업의 주요 서비스 품질 문제 분석에 효과적으로 활용될 수 있는 것을 보여주었다.

ABSTRACT

Companies collect various customer complaint data, increasing the need to utilize this information for service quality management and customer satisfaction improvement. To address this need, this study proposes an effective method for analyzing customer complaint data using natural language processing techniques. Specifically, customer complaints are classified into general inquiries and service-related claims, with a focus on categorizing claims that require additional management, such as after-sales services, to support efficient complaint handling. Furthermore, the proposed method was implemented using NLP-based BERT and KoBERT models, and experimental results showed that the KoBERT model outperformed the BERT model in AS category classification, achieving an F1 Score of 0.7046 compared to BERT's 0.7017. When applied to real-world customer complaint data from a telecommunications company, the proposed approach identified that complaints related to internet quality (47.3%) and TV quality (30.8%) accounted for 78.1% of total complaints. These findings demonstrate that the proposed method can be effectively utilized for analyzing key service quality issues in enterprises.

키워드 : 고객 불만, 자연어 처리, BERT, KoBERT

Keywords : Voice of Customer, Natural Language Processing, BERT, KoBERT

Received 7 April 2025, **Revised** 23 April 2025, **Accepted** 20 May 2025

*** Corresponding Author** Jang Won Bae(E-mail:jangwon_bae@koreatech.ac.kr, Tel:+82-41-560-1443)
Associate Professor, Department of Industrial Management, Korea University of Technology and Education, Cheonan, 31253 Korea

Open Access http://doi.org/10.6109/jkiice.2025.29.6.745 pISSN:2234-4772

Ⅰ. 서 론

정보통신기술(Information and Communication Technology; ICT) 기업은 고객 간에 서비스수준 협약서(Service Level Agreement; SLA)를 통해 서비스 품질에 대한 성능과 수준을 명시하는 계약을 체결한다[1]. 서비스 수준 협약을 준수하는 것은 기업의 생존을 위한 경쟁력 확보와 성장에 매우 중요하기 때문에 품질경영이 기업의 주요 KPI(Key Performance Indicator)로 관리 되고 있다. 또한 기업들은 고객만족도 전문 조사 기관인 국가 고객만족도(NCSI), 한국서비스품질지수(KS-SQI) 등을 통해 객관적 설문과 측정을 기반으로 서비스 품질개선과 고객 만족도를 평가하고 있다[2].

ICT 기업은 서비스 품질 확보와 구매 후 고객 경험을 통해 충성도 향상에 노력하고 있으며, 더 나아가 고품질 서비스와 긍정적 고객 경험이 기업의 시장 점유율 증대와 영업이익 증가에 중요한 핵심 요인으로 확인되었다[3].

서비스 품질 향상과 고객 충성도를 높이기 위한 경영활동으로 품질관리시스템 QMS(Quality Management System), 망 관리시스템 NMS(Network Management System), 고객 불만 상담시스템 VoC(Voice of Customer) 개발과 구축, 유지관리 비용이 증가하고 있다. ICT 기업의 애프터서비스(AS) 발생에 대한 품질비용 분석 결과, 매출액 대비 20~30% 비용이 소요되고 있으며, 이 비용은 품질 실패비용에 대한 구체적인 원인 분석을 통해 예방할 수 있다고 분석하였다[4].

그중 고객 불만(VoC)은 고객으로부터 서비스 품질에 대한 의견을 직접적으로 유입, 청취하는 중요한 정보이다. 서비스 품질에 대한 불만과 고객에게 판매나 임대한 단말, 리모컨, 케이블과 같은 소모품까지 고객 불만 내용을 비정형 형태로 수집한다. 예를 들어, ICT 기업은 수집한 고객 불만 데이터를 광대역 인터넷, IPTV, VoIP 전화 등 서비스 유형별 고장 상황을 카테고리로 분류하여 관리 및 저장되며, 이는 서비스 품질 향상과 마케팅 활용을 위해 사용되고 있다.

기존의 고객 불만(VoC) 처리 방식은 상담원이 수동으로 민원을 접수하고 해결하는 방식으로 고객이 불만사항을 제기하면 상담원이 내용을 듣고 고객의 불만 사항을 적절히 분류한 후, 관련 부서로 전달하여 처리하는 구조이다. 이러한 방식에는 여러 가지 문제가 있다. 상담원의 경우 대부분 기술 전문성이 부족하고 이에 따라 불완전한 대응 및 불만 요인 해소 지연으로 인한 처리비용이 증가한다. 특히, 상담원은 고객의 불만을 듣고 해석하여 상담원 수준에서 해결할 수 있는 불만과 AS 처리가 필요한지를 판단하게 되고, 이 분류에 따라 이후 처리 과정의 비용 차이가 클 수 있다. 예를 들어, 상담원의 분류가 부정확할 경우 그 비용이 크게 증가할 수 있다. 또한 일관성 없는 고객 응대와 반복된 AS 방문 협의 역시 기업의 비용 증가에 원인이 된다.

본 연구의 목적은 고객 불만 데이터를 자연어 처리 방법으로 자동 분류하여 기업의 서비스 품질 향상과 운영 비용을 절감하는 데 있다. 구체적으로, ICT 기업의 고객 불만 데이터를 활용하여 일반 문의(Inquiry)와 서비스 불만(Claim)으로 효과적으로 분류하고, 서비스 불만 분류의 경우에는 AS 카테고리로 세분화하여 서비스 품질 관리의 처리비용을 줄이고자 한다.

기존의 연구가 일반적인 감정 분석, 주제 분류, 문서 분류 등과 같이 고객 반응이나 감정을 파악하거나 문서의 주제를 자동으로 분류하는 데 초점을 맞춘 것과 차별화된다. 또한, 특정 산업(ICT)에서 발생하는 고객 불만 데이터를 활용하여 구체적 문제 영역인 인터넷 품질, TV 품질 등 AS 카테고리에 집중하여 서비스 품질 개선 및 운용 비용 절감 기여에 목적이 있다.

제안한 방법을 적용한 실험 결과, 일반 문의(Inquiry)와 서비스 불만(Claim) 민원 분류 모델인 BERT[5]는 정확도 0.9200, 정밀도 0.9233, 재현율 0.9200, F1 Score 0.9205의 성능을 보였으며, AUC-ROC 분석 결과 0.92로 측정되어 모델의 성능이 우수함을 확인할 수 있었다. 또한, Binary Cross-Entropy Loss가 0.2565로 낮아 모델 학습이 안정적으로 이루어졌음을 확인했다.

서비스 불만(Claim) 민원 AS 카테고리 선별 모델에서는 KoBERT[6] 모델이 F1 Score 0.7046으로 BERT 모델(0.7017)보다 우수한 성능을 보였고, Cross-Entropy Loss도 0.6292로 BERT 모델(0.6840)보다 낮아 더 안정적인 학습 성능을 나타냈다. 이는 KoBERT가 AS 카테고리 선별 작업에서 더 적합한 모델임을 확인하였다. 이러한 결과는 ICT 기업이 고객 불만 데이터를 활용하여 서비스 품질을 개선하고 운용 비용을 절감하는 데 기여할 수 있으며, 다른 산업 분야에도 응용 가능성을 시사한다.

Ⅱ. 관련 연구

정보화 시대에 고객이 제품이나 서비스 상호작용을 통해 인식된 정보는 품질 향상을 위한 핵심 요소로써 중요한 정보이다. Ben Rahman et al. (2024)[7]은 고객 만족도 향상을 위한 분석 방법으로 머신 러닝 기반 감정 분석 모델 6개를 비교 분석하였다. BERT 모델이 정확도 95%와 처리 속도에 우수한 성능을 보였다. 이를 통해 서비스 품질 개선과 고객 충성도 향상에 활용, 기업 매출 최대 15% 증대 효과를 기대했다.

온라인 쇼핑 리뷰의 유용성을 분류하기 위한 연구[8]에서는 BERT 모델 파인튜닝을 통해 분류 효과를 평가하였다. 기존의 Bag-of-Words 기반 머신러닝과 성능 비교 결과 파인튜닝된 BERT 모델이 기존 방법보다 높은 성능을 나타냈으며, 온라인 리뷰 분석 활용에 적합함을 보여주었다.

기존 자연어 처리(NLP)에서 사전 학습된 BERT 모델은 일반적인 도메인 Wikipedia, BooksCorpus 같은 데이터를 학습하였다. 반면 CS-BERT 모델[9]은 고객 서비스 도메인에 최적화된 데이터 학습으로 고객 응대 상황과 맥락을 이해하는데 더 우수한 성능을 나타냈다. 도메인 맞춤 학습이 자연어 처리 모델 성능 향상에 중요하다는 것을 알 수 있다.

그러나, 실제 도메인 응용에서 BERT 사전 학습은 코퍼스(Corpus) 양과 특정 도메인의 고유 명사 의미에 따라 분석 정확도가 낮아지는 문제도 있다. 정확도 개선을 위해 도메인 사전 마스크 학습 방식이 적용된 BERT 모델은 평균 12.12% 향상된 성능을 보였다[10].

또 다른 응용 연구인 인간과 상호작용 작업을 위한 챗봇 기술에도 복잡한 문장과 문맥 처리 적정성에 한계가 존재하며, 특히 고객 지원(Customer Support) 시나리오 제공에 기존 머신러닝 학습은 한계가 있다. BERT 모델과 심층 강화 학습(Deep Reinforcement Learning; DRL)을 결합하여 챗봇 응답 성능을 향상할 수 있었다[11].

구글에서 공개한 BERT는 다국어를 지원하지만, 한국어에 특화되어 있지 않아 분석에 한계가 있다. 한국어 기술 문서 분석을 위해 BERT의 구조와 학습 방식을 사용하면서도 한국어 언어적 특성을 반영하기 위한 교착어 특성 및 어순의 유연성을 갖도록 설계된 KoBERT 모델을 이용 기존의 다국어 BERT에 비해 한국어 작업에 효과적인 활용 방법을 제안하였다[12].

기업관련 신문기사 감성 분류 연구에서도 한국어 문서 선행학습을 통해 회계·재무 전문 분야의 감성분류 성과가 다른 모델들에 비해 우수함을 확인하였다[13].

KoBERT는 다른 언어보다 의미 해석이 어려운 한국어에 최적화된 자연어 처리를 위해 SKT Brain 팀이 공개했다. 사전 학습을 위한 다양한 한국어 데이터셋을 연구하여 한국어 문법과 어휘 분석에 성능을 높였다[14].

소규성외 3인 (2022)[15]은 KoBERT 모델을 기반으로 불규칙한 뉴스 데이터 주제를 분리하는 KoBERTSEG 모델을 제안하여 높은 성능을 확인하고, 응용할 수 있도록 제안했다.

표 1은 BERT와 KoBERT 모델의 주요 특징을 비교하였다.

Table. 1 Comparison Table : BERT vs KoBERT

Category	BERT	KoBERT
Language Support	Multilingual support (including mBERT)	Specialized for Korean
Pre-training Data	English-centric, includes multilingual corpora	Korean-centric data
Performance	Relatively lower performance in Korean tasks	Higher performance in Korean tasks
Tokenizer	Sentencepiece (multilingual)	Sentencepiece (optimized for Korean)

관련 연구는 주로 BERT 모델의 감정 분석, 주제 분류의 정확도와 처리 성능에 관한 연구가 진행되었고, 고객 서비스 도메인에 최적화된 데이터 학습을 통해 성능을 향상시킬 수 있는 방법을 제안했다.

본 연구에서는 서비스 품질 관리와 운용 비용 절감을 목적으로 특정 산업(ICT)의 고객 불만 데이터를 활용하여 일반 문의(Inquiry)와 서비스 불만(Claim) 민원 분류 및 AS 카테고리 선별에 특화된 딥러닝 모델인 BERT와 KoBERT 성능을 비교 평가한다. 또한, 단순한 감정 분석이나 주제 분류를 넘어 AS 카테고리 선별을 위한 세분화된 모델 평가를 수행하여 기업이 고객 불만 AS 카테고리별 효과적인 대응 전략을 수립할 수 있도록 한다. 이를 통해 기업의 품질 관리 및 운영 효율화에 실질적인 도움을 줄 수 있는 접근 방안을 제시한다.

Ⅲ. 연구 방법

3.1 데이터 분석을 위한 연구 모형

기존의 고객 불만(Voice of Customer, VoC) 처리 방식은 상담원이 수동으로 민원을 접수하고 해결하는 구조로 운영된다. 그림 1은 고객 불만 데이터를 처리하고 분석하기 위한 전체적인 구조와 단계를 나타낸다.

기존 프로세스에서 ①고객 불만 분류(Classification of Complaints) 단계는 비전문 상담원(일반 상담원)이 고객 불만 데이터를 일반 문의(Inquiry)와 서비스 불만(Claim)으로 수작업 분류한다. 일반 문의는 요금, 약정기간, 해지 신청 등 단순한 정보 요청으로 상담원이 바로 해결 종료한다. 반면, ②서비스 불만(Claim)으로 분류된 경우는 AS로 접수(After-Sales Services Request)된다. 이 과정에서 상담원의 기술적 전문성 부족으로 인해 서비스 불만 분류와 실제 문제 원인이 일치하지 않는 문제가 발생할 수 있다. 이러한 수작업 기반 처리 방식은 다음과 같은 문제점을 초래한다. 첫째, 판단 오류와 일관성 부족이다. 상담원의 기술적 전문성 부족으로 인해 분류 기준이 일관되지 않아 문제 해결이 지연된다. 둘째, 근본 원인 분석에 어려움이 있다. 정보 확인 및 분석 부족으로 근본적인 원인 판단이 어려워 불완전한 대응이 이루어진다. 셋째, 고객 불만 원인 분석에 상담원이 직접 정보를 수집하고 여러 부서와 협의하는 과정에서 민원 해결까지 오랜 시간이 소요된다. 넷째, 민원 처리 프로세스가 체계적으로 정립되지 않고 상담원마다 대응 방식이 달라 고객 만족도가 낮아진다. 다섯째, 반복적인 고객 서비스 불만에 따른 지속되는 상담과 AS 요청으로 품질 관리 및 운영 비용이 증가한다.

이러한 문제를 해결하기 위해 본 연구에서는 자연어 처리(NLP) 기반 딥러닝 모델(BERT와 KoBERT)을 적용하여 프로세스를 개선하였다.

기존 프로세스 중 ①고객 불만 분류(Classification of Complaints) 단계에서 고객 불만 데이터를 일반 문의(Inquiry)와 서비스 불만(Claim)으로 분류함으로써 분류 작업의 정확도와 일관성을 높이고자 한다. 일반 문의는 단순 문의나 정보 요청을 포함한 상담원 직접 응대만으로 종료되고, 서비스 불만은 서비스 품질 문제나 기술적 이슈 등이 포함되어 2단계 분석에 주요 데이터로 활용된다. ②AS 카테고리 선별(Classification into Eight After-Sales Categories) 단계에서도 서비스 불만으로 분류된 데이터를 인터넷 품질, TV 품질 등 8개의 AS 카테고리로 선별하는 작업을 진행한다. 서비스 불만(Claim) 민원 데이터를 기반으로 AS가 필요한지 판단하고, AS가 필요한 경우 8가지 세부 카테고리로 선별하여 체계적인 서비스와 최적의 대응 방안을 수립하는 데 활용한다. 이후 축적된 고객 불만 데이터를 바탕으로 유사한 민원 발생 시 보다 신속하고 효율적인 대응이 가능하도록 개선한다.

연구 모형의 구조는 고객 불만 데이터는 고객상담센터(Customer Support Center)를 통해 접수되며, 원격 기술 분석(Remote Technology Diagnosis)을 통해 단말 신호 상태 확인(Terminal Signal Status Monitoring), 가입자 포트 리셋(Subscriber Port Reset), 고객 단말 전원 리부팅(CPE Device Power Cycling) 등 초기 진단 처리가 이루어진 데이터를 활용한다. 이후 데이터는 BERT 기반 분류 및 평가를 위해 데이터 전처리, 토큰화 및 라벨 인코딩, 데이터 분할, 파인튜닝, 성능 분석 및 결과 도출로 진행된다.

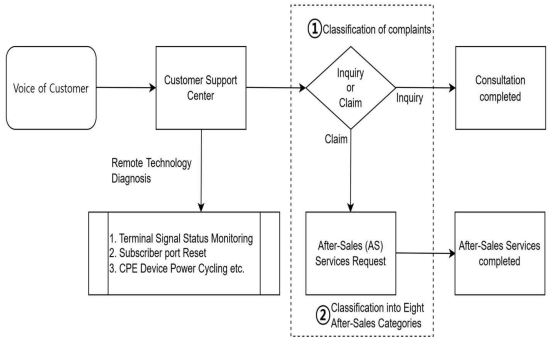

Fig. 1 Classification Process of Customer Complaint Data

3.2 데이터셋

BERT 모델의 성능 비교 및 분석을 위해 ICT 기업의 고객 불만 상담 데이터셋이 사용되었다. 데이터는 고객 상담을 진행한 후 등록된 상담 내용과 처리 결과를 포함하고 있으며, 전체 5,276,143건으로 각 건은 개별 고객 불만 상담 건수를 의미한다.

전체 데이터의 크기는 5.04GB이며, 데이터 필드는 88개로 이루어졌고, 분석에 사용된 주요 변수 필드는 장애상담분류, 불만구분, 상담내용, 장애원인분류 이다. 데이터의 유형은 정해진 형식과 구조로 저장된 정형화 형태이나, 분석을 위한 주요 필드는 텍스트 데이터 형태

로 이루어져 있다. 그림 2는 본 연구에 사용된 데이터셋의 예시를 보여준다.

데이터 전처리 과정으론 주요 변수 필드에 결측값이 있는 데이터를 제거하였고, 이외에 별도의 전처리는 수행하지 않았다. 그리고 BERT 모델이 요구하는 형식에 맞는 SentencePiece 기반으로 한국어에 최적화된 토크나이저를 사용 토큰화했고, 불만 민원 구분을 위해 모델 학습에 적합한 정수형 라벨 구조로 변환하였다. 데이터 분할은 전체 데이터를 학습용(Training)과 테스트용(Test)으로 나누었으며, 각 분할은 랜덤 샘플링을 통해 분할했다. 위와 같은 과정을 통해 고객 불만 데이터를 BERT 모델 학습에 적합한 형태로 변환하였다.

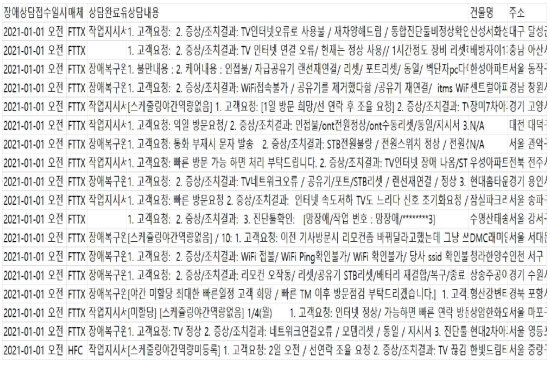

Fig. 2 Customer Complaint Dataset

3.3 일반 문의와 서비스 불만 민원 분류 모델

일반 문의(Inquiry)와 서비스 불만(Claim) 민원 분류 성능 평가를 위해 BERT 모델을 사용하였다. 또한, Hugging Face의 Transformers 라이브러리에서 제공하는 'bert-base-multilingual-cased' 사전 학습 모델을 사용하여 텍스트 분류 작업에 사용하였다. 일반 문의와 서비스 불만 민원 분류 절차는 그림 3과 같다.

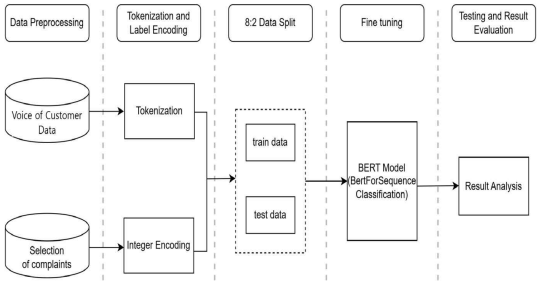

Fig. 3 Classification Process of Customer Complaints

데이터 전처리는 상담 민원 텍스트 데이터 벡터화를 위해 토큰화를 수행하였고, 사전 학습된 BERT 토크나이저를 사용 문장을 단어 단위로 분할한다. 불만 민원은 라벨링 작업을 통해 일반 문의(0)와 서비스 불만(1)으로 이진 분류를 수행하였고, 이를 위해 라벨 인코딩을 적용하여 정숫값을 부여하였다.

데이터의 분할은 Train(80%), Test(20%) 비율로 학습 및 평가를 진행하였고, 학습 데이터와 테스트 데이터의 라벨 비율이 유사하도록 유지하였다.

표 2와 같이 파인튜닝을 통해 높은 정확도와 분류 성능을 평가하고자 하였다. 최대 토큰 길이는 고객 불만 상담 데이터가 긴 문장을 포함하여 문맥 정보를 손실 없이 처리할 수 있도록 BERT 모델의 최대 입력 길이인 512로 설정하였다. 또한, 배치 크기(256), 학습률(5e-5), Epoch(5), 최적화 알고리즘(AdamW)로 적용하였다.

Table. 2 Hyperparameter Settings for Classification of Inquiry and Claim Complaints

Hyperparameter	Value
Maximum Token Length	512
Batch Size	256
Learning Rate	5e-5
Epochs	5
Optimization Algorithm	AdamW
Evaluation Metrics	Accuracy, Precision, Recall, F1 Score

3.4 AS 카테고리 선별 모델

서비스 불만(Claim)으로 분류된 데이터를 기반으로 다음과 같은 추가 전처리 작업을 진행하였다. 첫째, 불만 여부가 명확하지 않은 상담 데이터를 제외하였다. 둘째, 서비스 불만으로 분류되었으나 고장 원인에 대분류·중분류·소분류 데이터가 없어 AS가 불필요한 CS(Customer Service)로 규정하고 분류된 데이터를 제거하였다. 이러한 전처리 과정을 거쳐 최종적으로 분석에 활용할 621,649건의 데이터를 선정하였다.

서비스 불만 민원의 AS 카테고리 선별 성능을 평가하고 비교 분석하기 위해, 한국어 텍스트 분석에 적합한 KoBERT와 BERT 모델을 통해 그림 4와 같이 비교하였다. BERT 모델은 3.3절에서 사용한 모델과 동일하며,

한국어 문장 처리의 최적화를 위해 NAVER 한글 말뭉치를 기반으로 사전 학습된 KoBERT 모델(get_pytorch_kobert_model)을 사용하였다.

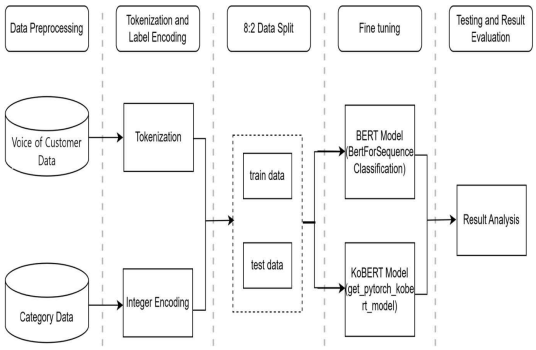

Fig. 4 Classification Process of Categorization of After-Sales Services

일반 문의(Inquiry)는 AS 처리가 불필요한 CS 접수건으로 매핑하고, 서비스 불만(Claim) 민원은 8개의 세부 AS 카테고리로 선별하는 기준을 제시하였다. 제시한 기준은 사용한 데이터 도메인의 특성을 반영하여 설정하였다. 구체적으로 설정한 8개의 세부 AS 카테고리는 다음과 같다. AS 접수 후 정상(Normal Status After After-Sales Service Requests), TV 품질(TV Quality), 가입자 소모품(Subscriber Consumables), 고객 PC(Customer PC), 고객 단말 장비(Customer Premises Equipment), 고객 불만(Customer Complaints), 유선 전화 품질(Fixed-Line Telephone Quality), 인터넷 품질(Internet Quality) 이를 통해 주요 서비스 품질 문제를 보다 명확히 분석하고, AS 카테고리별 맞춤 대응 방안을 수립할 수 있도록 하였다.

데이터의 분할은 Train 497,319건(80%), Test 124,330건(20%) 비율로 학습 및 평가를 진행하였다.

표 3과 같이 한국어 텍스트 분석에 적합한 KoBERT와 BERT의 성능 비교를 위해 두 모델에 동일한 환경으로 파인튜닝을 진행하였다. 표 2와 다르게 설정된 하이퍼파라미터 값은 분석 데이터양이 많아 배치 크기(128), Epoch(3)으로 설정하였고, 모델의 성능 평가도 동일한 정확도, 정밀도, 재현율, F1-Score로 평가했다.

Table. 3 Hyperparameter Settings for Performance Comparison of KoBERT and BERT

Hyperparameter	Value
Maximum Token Length	512
Batch Size	128
Learning Rate	5e-5
Epochs	3
Optimization Algorithm	AdamW
Evaluation Metrics	Accuracy, Precision, Recall, F1 Score

Ⅳ. 실험 결과

본 연구에서는 고객 불만 데이터를 BERT 모델을 활용하여 일반 문의(Inquiry)와 서비스 불만(Claim) 민원을 분류하고, 서비스 불만 민원 데이터를 바탕으로 AS 카테고리 선별 모델을 비교 분석하였다. 4.1절에서는 일반 문의와 서비스 불만 민원 분류 모델의 성능을 평가했으며, 4.2절에서는 KoBERT와 BERT 모델을 활용한 AS 카테고리 선별 성능을 비교하였다. 4.3절에서는 다중 클래스 분류(Multi-class Classification) 성능을 분석하였다.

4.1 일반 문의와 서비스 불만 민원 분류 모델 성능 분석

일반 문의(Inquiry)와 서비스 불만(Claim) 민원 분류 모델의 성능을 평가한 결과는 표 4와 같다. 정확도(Accuracy) 0.9200, 정밀도(Precision) 0.9233, 재현율(Recall) 0.9200, F1 Score 0.9205의 성능을 나타냈다. 이는 모델이 고객 불만 민원을 높은 신뢰도로 분류할 수 있음을 의미한다. 연구에 사용된 이진 분류 모델 손실 함수는 Binary Cross-Entropy Loss이며, 총 손실률(Loss Rate)은 0.2565로 비교적 낮아 모델 학습이 안정적으로 이루어진 것을 확인할 수 있다. 이러한 결과는 일반 문의와 서비스 불만 민원을 분류하는 데 있어 제안된 BERT 모델이 높은 신뢰도로 예측할 수 있음을 의미한다. 또한, 데이터셋의 서비스 불만 비율은 39.8%, 일반 문의 비율은 60.2%로 분석되었으며, 모델은 이 분포를 기반으로 높은 F1 Score(0.9205)와 낮은 손실률(0.2565)을 기록하여 신뢰도 높은 분류 성능을 보였다.

Table. 4 Prediction Performance Results on Customer Complaint Dataset

result	BERT
Acc	0.9200
Prec	0.9233
Rec	0.9200
F1 score	0.9205
Loss	0.2565

추가적인 모델의 분류 성능을 평가하기 위해 그림 5와 같이 AUC-ROC 분석을 수행하였다. 분석 결과 0.92로 측정되었고, 높은 TPR(True Positive Rate)과 낮은 FPR(False Positive Rate)을 보임으로써 모델의 성능이 우수함을 확인할 수 있었다.

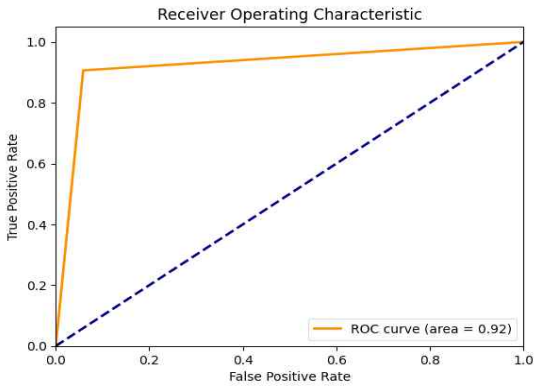

Fig. 5 Receiver Operating Characteristic with Area Under the Curve

4.2 AS 카테고리 선별 모델 성능 분석

AS 카테고리 선별 모델에서는 BERT와 KoBERT 모델을 비교하여 성능을 평가하였다.

표 5와 같이 최종적으로 학습된 BERT와 KoBERT 모델의 성능을 비교한 결과, KoBERT 모델이 Epoch 3에서 정확도(Accuracy) 0.7441, 정밀도(Precision) 0.6927, 재현율(Recall) 0.7441, F1 Score 0.7046의 성능을 보였다. 그러나, 두 모델 모두 Epoch가 증가함에 따라 성능 개선 폭이 점차 감소하였다. KoBERT는 모든 Epoch에서 BERT 모델 보다 높은 F1 Score를 보여, 한국어 고객 불만 데이터를 활용한 AS 카테고리 선별에서 더 적합한 모델임을 확인하였다. 이는 KoBERT가 한국어 문장 구조 및 의미를 보다 잘 반영하기 때문으로 판단된다.

모델의 학습 안정성과 예측 신뢰도를 평가하기 위한 Cross-Entropy Loss 분석 결과, KoBERT 모델의 손실값은 0.6292로 BERT 모델의 손실값(0.6840)보다 낮게 나타났다. 이는 KoBERT 모델이 학습 과정에서 더 효과적이며, AS 카테고리 선별 작업에서 보다 정교한 예측을 수행할 수 있음을 의미한다.

Table. 5 Performance Comparison of KoBERT and BERT for Korean Text Analysis

result	KoBERT	BERT
Acc	0.7441	0.7438
Prec	0.6927	0.6904
Rec	0.7441	0.7438
F1	0.7046	0.7017
Cross entropy loss	0.6292	0.6840

Epoch에 대한 학습 성능 변화는 그림 6과 같이 초기 Epoch 구간에서는 급격한 성능 향상이 나타났으나, 이후 Epoch 3에서는 성능 증가율이 점차 둔화되는 성능 포화(performance saturation) 현상을 보였다. 이는 고객 불만 데이터 내에 학습할 수 있는 새로운 패턴이 제한적이거나, 모델이 이미 주요 정보를 충분히 학습하여 추가적인 학습이 성능 향상에 실질적인 기여를 하지 못함을 의미한다. 이러한 결과는 Epoch가 증가함에 따라 학습 성능이 점차 수렴하는 경향을 보이고, 동시에 과적합(overfitting) 없이 안정적으로 학습이 진행되었음을 알 수 있다.

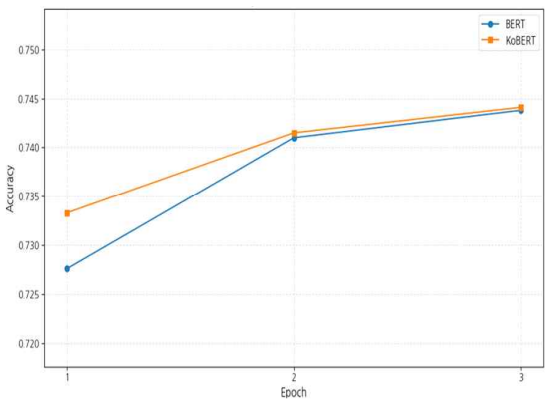

Fig. 6 Model Accuracy over Epochs

4.3 다중 클래스 분류(Multi-class Classification) 성능 분석

AS 카테고리 선별은 다중 클래스 분류 문제로 단일 성능 지표만으로 모델의 성능을 설명하기 어렵다. 따라서 카테고리별 정밀도(Precision), 재현율(Recall), F1 Score를 그림 7과 같이 분석하여 성능 편차를 분석하였다. 분석 결과, 인터넷 품질, TV 품질과 같은 주요 클래스에서 높은 성능을 나타냈으나, 유선 전화 품질, 고객 단말 장비 클래스에서는 상대적으로 낮은 성능을 보였다. 이러한 성능 편차는 데이터의 불균형(imbalance data) 문제로 인터넷 품질(294,009건)과 TV 품질(191,556건)이 차지하는 비중이 높아 해당 클래스는 충분한 학습을 수행할 수 있었다. 반면 낮은 빈도 클래스에 대해서는 학습 데이터가 제한되어 모델이 일반화된 분류 기준을 형성하기 어려웠다. 또한, 클래스 간 텍스트 내용이 유사하거나 중복되는 클래스의 경우 경계를 명확히 구분하기 어려운 것으로 판단된다. 이로 인해 정밀도와 재현율 중 하나가 불균형하게 낮게 나타나는 현상을 보였다.

결론적으로, 다중 클래스 분류 성능 분석을 통해 단일 지표 평가로는 파악하기 어려운 분류 성능의 편차와 한계를 확인할 수 있었다. 이는 향후 실무 적용 시 데이터 불균형 해소, 클래스별 데이터 증강, 낮은 성능 클래스에 대한 보완 학습 필요성이 제기된다.

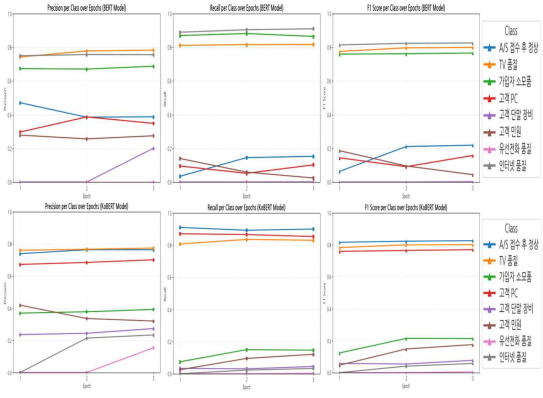

Fig. 7 Multi-class Classification Performance

서비스 품질 문제를 체계적으로 분석하고, AS 대응 방안을 수립하기 위해 KoBERT 모델을 활용하여 총 621,649건의 고객 불만 데이터를 AS 카테고리로 선별한 결과, 표 6과 같은 분포를 나타냈다. 특히 인터넷 품질(47.3%), TV 품질(30.8%) 관련 고객 불만이 전체의 약 78.1%를 차지하여 주요 서비스 품질 문제로 나타나고 있다. 이러한 결과를 바탕으로 ICT 기업은 주요 서비스 불만 AS 카테고리에 대해 우선적으로 개선할 전략을 수립하고, 인터넷과 TV 품질 문제 해결을 위한 예방적 유지관리 및 AS 카테고리에 적합한 대응 정책 수립에도 활용할 수 있는 체계적인 접근 방식을 제시한다.

Table. 6 Classification of customer complaint data into AS categories

AS Category	classification count
Normal Status After After-Sales Service Requests	53,088 (8.5%)
TV Quality	191,566 (30.8%)
Subscriber Consumables	30,448 (4.9%)
Customer PC	17,632 (2.8%)
Customer Terminal Equipment	4,893 (0.8%)
Customer Complaints	26,667 (4.3%)
Fixed-Line Telephone Quality	3,346 (0.5%)
Internet Quality	294,009 (47.3%)

Ⅴ. 결 론

본 연구의 목적은 고객 불만 데이터를 기반으로 한 자동화된 분류 시스템을 통해 서비스 품질 향상과 운영 비용을 절감하는 데 있다. 이를 위해 특정 산업(ICT)의 고객 불만 데이터를 활용하여 일반 문의(Inquiry)와 서비스 불만(Claim) 민원 분류 및 AS 카테고리 선별 과정과 이를 효과적으로 수행할 수 있는 자연어 처리(NLP) 모델을 평가하였다. 또한, 기존의 감정 분석이나 주제 분류를 넘어 AS 카테고리 선별을 위한 세분화된 평가를 수행하여 기업이 고객 서비스 불만 AS 카테고리별로 효과적인 대응 전략 수립을 지원할 수 있다.

BERT 모델은 일반 문의와 서비스 불만 민원 분류에서 높은 성능을 보이며 안정적인 학습을 확인할 수 있었다. 또한, AS 카테고리 선별 성능 비교에서 KoBERT 모델이 한국어 문법과 어휘적 특성을 효과적으로 반영하여 BERT보다 우수한 성능(F1 Score 0.7046)을 나타냈다. 이는 한국어 기반 고객 불만 데이터 분석에 있어 KoBERT 모델이 보다 적합하며, 한국어 데이터 특성을

고려한 모델 설계가 중요함을 시사한다.

KoBERT 모델을 활용하여 고객 불만 데이터를 AS 카테고리로 선별한 결과, 인터넷 품질(47.3%), TV 품질(30.8%) 관련 고객 불만이 전체의 약 78.1%를 차지하는 것으로 나타났다. 이는 고객 불만이 특정 서비스 품질 문제에 집중되어 있으며, ICT 기업이 이러한 주요 고객 서비스 불만 AS 카테고리를 우선적으로 해결하는 전략을 수립해야 함을 시사한다. 또한, AS 카테고리 패턴 분석을 통해 반복적으로 발생하는 문제를 예측하고, 해결할 수 있는 자동화된 대응 시스템 개발이 요구된다. 이를 바탕으로 ICT 기업은 AI 기반 고객 불만 분석 시스템을 활용하여 서비스 품질 개선과 AS 처리에 대한 효율성을 높일 수 있다.

본 연구는 기존의 자연어 처리 기반 고객 불만 분석 연구들과 다음과 같은 차별성을 가진다. 첫째, 기존 연구들이 주로 감정 분석, 주제 분류, 또는 문서 분류 등 일반적인 분류 작업에 초점을 맞춘 반면, 본 연구는 실제 ICT 산업 현장에서 수집된 대규모 고객 불만 데이터를 활용하여, 실무적으로 중요한 'AS 카테고리' 선별이라는 구체적이고 실질적인 문제 해결에 초점을 두었다. 둘째, 한국어에 특화된 KoBERT 모델을 적용·비교함으로써, 한국어 고객 불만 데이터 분석에서의 성능 향상과 실무 적용 가능성을 실증적으로 제시했다. 셋째, 단순히 분류 성능 평가에 그치지 않고, 실제 기업의 서비스 품질 관리와 운영 비용 절감에 직결되는 고객 불만 AS 카테고리 패턴 분석 및 대응 전략 수립을 지원함으로써, 산업 현장에서의 실질적 활용 가치를 확인하였다.

이러한 점에서 본 연구는 기존의 자연어 처리 기반 고객 불만 분석 연구와 차별화되며, ICT 기업의 서비스 품질 개선, 고객 만족도 제고, 운용 비용 절감을 위한 데이터 기반 의사결정 지원에 실질적으로 기여한다.

연구의 한계점으로는 Cross-Entropy Loss 값이 0.6292로 다소 높게 나타났다. 이는 분류 대상이 8개 카테고리로 다소 복잡하다는 점을 감안하면 허용 가능한 수준이지만, 보다 정확한 AS 카테고리 선별을 위해 하이퍼파라미터 조정 등으로 추가적인 개선이 필요하다. 또한, 사용된 데이터가 특정 ICT 기업의 고객 불만 데이터로 한정되어 다른 기업 및 산업의 데이터와 비교 검증이 이루어지지 않았다. 아울러, BERT와 KoBERT 외 다른 최신의 자연어 처리 모델들과의 성능 비교가 이루어지지 않았다. 향후 연구에서는 다양한 산업 분야의 수집된 데이터를 활용하고, 더 많은 자연어 처리 모델의 성능을 비교 분석함으로써 서비스 품질관리와 운용 비용 절감에 기여할 수 있는 방안 모색이 필요하다.

REFERENCES

[1] M. Langer and M. Nerb, "Customer service management: An information model for communication services," in *Proceedings of the Third International IFIP/GI Working Conference*, Munich: DE, pp. 244-257, 2000. DOI: 10.1007/10722515_20.

[2] S. B. Cho and K. Y. Kwang, "A Comparative Study on the KS-SQI, NCSI, and KCSI," *Korean Journal of Hospitality and Tourism*, vol. 17, no. 3, pp. 213-227, Jun. 2008.

[3] T. I. Kim and M. S. Kim, "Sources of the customer loyalty in the high-speed Internet service," *Journal of the Korea Academia-Industrial cooperation Society*, vol. 12, no. 2, pp. 728-735, Feb. 2011. DOI: 10.5762/KAIS.2011.12.2.728.

[4] G. H. Hwang, "A Case Study on the Quality Costs in a ICT Industry," *Journal of the Korean society for Quality Management*, vol. 40, no. 2, pp. 106-116, Jun. 2012. DOI: 10.7469/JKSQM.2012.40.2.106.

[5] J. Devlin, M. W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," in *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologie*s, Minnesota: USA, pp. 4171-4186, 2019. DOI: 10.18653/v1/N19-1423.

[6] SK telecom. Korean BERT (Bidirectional Encoder Representations from Transformers) [Internet]. Available: http://sktelecom.github.io/project/kobert/.

[7] B. Rahman and Maryani, "Optimizing Customer Satisfaction through Sentiment Analysis: A BERT-based Machine Learning Approach to Extract Insights," *IEEE Access*, vol. 12, pp. 151476-151489, Oct. 2024. DOI: 10.1109/ACCESS.2024.3478835.

[8] M. Bilal and A. A. Almazroi, "Efectiveness of Fine tuned BERT Model in Classifcation of Helpful and Unhelpful Online Customer Reviews," *Electronic Commerce Research*, vol. 23, no. 4, pp. 2737-2757, Apr. 2022. DOI: 10.1007/s10660-022-09560-w.

[9] P. Wang, J. Fang, and J. Reinspach, "CS-BERT: a pretrained model for customer service dialogues," in *Proceedings of the 3rd Workshop on Natural Language Processing for*

Conversational AI, Online, pp. 130-142, 2021. DOI: 10.18653/v1/2021.nlp4convai-1.13.

[10] X. Cao, H. Xiao, and W. Jiang, "Knowledge enhancement BERT based on domain dictionary mask," *Journal of High Speed Networks*, vol. 29, no. 2, pp. 121-128, Jan. 2023. DOI: 10.3233/JHS-222013.

[11] K. R. Praneeth, T. S. Ruprah, J. N. Madhuri, A. L. Sreenivasulu, S. Shareefunnisa, and V. S. Rao, "Optimizing Customer Interactions: A BERT and Reinforcement Learning Hybrid Approach to Chatbot Development," *International Journal of Advanced Computer Science & Applications (IJACSA)*, vol. 15, no. 9, pp. 569-578, 2024. DOI: 10.14569/ijacsa.2024.0150958.

[12] S. H. Hwang and D. H. Kim, "BERT-based Classification Model for Korean Documents," *The Journal of Society for e-Business Studies*, vol. 25, no. 1. pp. 203-214, Feb. 2020. DOI: 10.7838/jsebs.2020.25.1.203.

[13] J. W. Hyeon, J. N. Lee, H. K. Cho, "Sentiment Analysis of News on Corporation Using KoBERT," *Korean Accounting Review*, vol. 47, no. 4, pp. 33-54, Jul. 2022. DOI: 10.24056/KAR.2022.08.002.

[14] H. W. Ko, K. C. Yang, M. H. Ryu, T. K. Choi, S. M. Yang, J. W. Hyun, S. H. Park, and K. B. Park, "A Technical Report for Polyglot-Ko: Open-Source Large-Scale Korean Language Models," *arXiv preprint arXiv:2306.02254*, Jun. 2023. DOI: 10.48550/arXiv.2306.02254.

[15] K. S. So, Y. S. Lee, E. S. Chung, and P. S. Kang, "KoBERTSEG: Local Context Based Topic Segmentation Using KoBERT," *Journal of the Korean Institute of Industrial Engineers*, vol. 48, no. 2, pp. 235-248, Apr. 2022. DOI: 10.7232/JKIIE.2022.48.2.235.

원종수(Jong-Su Won)

2015년 한국기술교육대학교 IT융합과학경영학과 과학경영석사
2022년~현재 한국기술교육대학교 산업경영학과 박사과정
※관심분야 : 자연어 처리, ICT, 기술경영

나인(In Na)

2025년 한국기술교육대학교 산업경영학부 학사과정
※관심분야 : 자연어 처리, LLM, 텍스트 분석

배장원(Jang Won Bae)

2007년 고려대학교 전기전자전파공학 공학사
2009년 한국과학기술원 전자공학과 공학석사
2015년 한국과학기술원 산업및시스템공학과 공학박사
2020년~현재 한국기술교육대학교 산업경영학과 부교수
※관심분야 : 모델링, 시뮬레이션, 빅데이터, AI