석 사 학 위 논 문

Master Thesis

안드로이드 및 윈도우즈 환경에서 카카오톡 메시지의 아티팩트 분류 및 심층학습 기반 감성 분석 연구

A Study on Artifacts Classification and Sentiment Analysis by Deep Learning in KakaoTalk Messages under Android and Windows Environments

2020

이 나 비 (Lee, Na-bi)

한 국 과 학 기 술 원

Korea Advanced Institute of Science and Technology

석 사 학 위 논 문

안드로이드 및 윈도우즈 환경에서 카카오톡 메시지의 아티팩트 분류 및 심층학습 기반 감성 분석 연구

2020

이 나 비

한 국 과 학 기 술 원

전산학부 정보보호대학원

안드로이드 및 윈도우즈 환경에서 카카오톡 메시지의 아티팩트 분류 및 심층학습 기반 감성 분석 연구

이 나 비

위 논문은 한국과학기술원 석사학위논문으로
학위논문 심사위원회의 심사를 통과하였음

2019년 12월 16일

심사위원장 김 광 조 (인)

심 사 위 원 김 흥 택 (인)

심 사 위 원 신 인 식 (인)

A Study on Artifacts Classification and Sentiment Analysis by Deep Learning in KakaoTalk Messages under Android and Windows Environments

Nabi Lee

Advisor: Kwangjo Kim

A dissertation submitted to the faculty of
Korea Advanced Institute of Science and Technology in
partial fulfillment of the requirements for the degree of
Master of Science in Computer Science (Information Security)

Daejeon, Korea
December 16, 2019

Approved by

Kwangjo Kim
Professor of Computer Science

The study was conducted in accordance with Code of Research Ethics[1)].

1) Declaration of Ethical Conduct in Research: I, as a graduate student of Korea Advanced Institute of Science and Technology, hereby declare that I have not committed any act that may damage the credibility of my research. This includes, but is not limited to, falsification, thesis written by someone else, distortion of research findings, and plagiarism. I confirm that my dissertation contains honest conclusions based on my own careful research under the guidance of my advisor.

MIS
20183811

이나비. 안드로이드 및 윈도우즈 환경에서 카카오톡 메시지의 아티팩트 분류 및 심층학습 기반 감성 분석 연구. 전산학부 (정보보호대학원). 2020년. 34+iv 쪽. 지도교수: 김광조. (한글 논문)

Nabi Lee. A Study on Artifacts Classification and Sentiment Analysis by Deep Learning in KakaoTalk Messages under Android and Windows Environments. School of Computing (Graduate School of Information Security). 2020. 34+iv pages. Advisor: Kwangjo Kim. (Text in Korean)

초록

카카오톡은 국내에서 압도적인 점유율을 가진 채팅 애플리케이션으로 텍스트, 사진, 음성, 동영상, 화상통화 등 다양한 콘텐츠로 소통할 수 있는 대화 창구이다. 모든 송·수신 메시지는 사용자의 스마트폰과 PC에 데이터베이스 (DB) 형태로 저장되기 때문에 포렌식 연구에서 사용자의 스마트폰 또는 PC에 잔존하는 흔적들 (artifacts)에 대한 분석은 대단히 중요하다. 본 논문에서는 두 단계로 나누어 분석을 진행하였다. 우선 안드로이드 6 환경의 모바일용 카카오톡과 윈도우즈 10 환경의 PC용 카카오톡에서 각각 아티팩트를 추출한 뒤 공통점과 차이점을 분석하였다. 아티팩트는 5개 유형, 54개 세부 속성으로 구분되며 이 중 18개 (33.3%)의 아티팩트가 모바일용 카카오톡 및 PC용 카카오톡에서 공통적으로 확인되었고, 36개 (66.7%)의 아티팩트는 모바일용 카카오톡 또는 PC용 카카오톡 한쪽에서만 획득 가능하였다. 이는 실제 수사환경에서 수사관이 모바일용 카카오톡 아티팩트와 PC용 카카오톡 아티팩트를 상호 보완하여 활용하는 것이 유리함을 의미한다. 두 번째 단계에서는 첫 번째 단계에서 획득한 아티팩트들 중 사용자의 사상이나 감정이 담겨 있어 추가적인 분석이 요구되는 유일한 아티팩트, 카카오톡 메시지를 심층학습 기반 자연어 처리 기법을 활용하여 감성 분석하였다. 이는 사용자의 주관 (최근 관심사, 심리적 상태, 채팅 참여자와의 밀접도 등) 파악에 상당히 유의미한 정보를 제공한다.

핵심낱말 카카오톡, 아티팩트, 포렌식, 심층학습, 감성 분석

Abstract

KakaoTalk is a chat application with an overwhelming share in South Korea and is a dialogue channel that allows users to communicate with various contents such as text, photos, voice, video and calls. Since all transmission and reception messages are stored in database (DB) format on a user's smartphone and PC, an analysis of the remaining traces (artifacts) on the user's smartphone or PC is critical in a forensics study. In this paper, the analysis was carried out in two stages. First, we extracted artifacts from KakaoTalk for mobile devices in Android 6 environment and KakaoTalk for PC in Windows 10 environment each, and then analyzed the similarities and differences between them. Artifacts were divided into five types and 54 detailed attributes, 18 of which (33.3 percent) were commonly identified on KakaoTalk for mobile devices and Kakao Talk for PC, while 36 (66.7 percent) were only available on either KakaoTalk for mobile devices or KakaoTalk for PC. It is advantageous for investigators to complement and utilize KakaoTalk artifacts for mobile devices and KakaoTalk artifacts for PCs in the actual investigation environment. In the second stage, the content of KakaoTalk messages, the only artifact acquired in the first stage that contained the user's thoughts or feelings and required further analysis, was analyzed by sentimental analysis using Deep-learning-based natural language processing techniques. This provided meaningful information that can help investigators figure out the truth of case.

Keywords KakaoTalk, arfitacts, forensics, deep learning, sentiment analysis

차 례

표 차례

그림 차례

제 1 장 머리말

카카오톡은 국내 스마트폰 사용자 중 97%가 애용하는 대표적인 채팅 애플리케이션 (이하 '앱') 이다 [1]. 2010년 안드로이드 OS용, 2013년 윈도우즈 OS용 카카오톡이 출시되면서 사용자는 스마트폰과 PC를 구분하지 않고 텍스트, 사진, 음성, 동영상, 화상통화 등 다양한 콘텐츠로 소통하고 있다. 모든 송·수신 메시지는 사용자의 스마트폰과 PC에 데이터베이스 (이하 DB) 형태로 저장되기 때문에 포렌식 연구에서 사용자의 스마트폰 또는 PC에 잔존하는 카카오톡 메시지의 흔적 (아티팩트) 분석은 대단히 중요하다. 하지만 카카오톡은 WhatsApp, Facebook Messenger 등과 달리 [2] 국내에 그 인기가 국한되어 있어 상대적으로 연구가 활발하지 않고, 모바일용과 PC용으로 구분되어 있음에도 이를 동시에 다루는 논문을 찾아보기 힘들다. 특히 안드로이드 환경에서의 아티팩트 분석방법을 기술한 연구들은 모두 카카오톡 4점대 버전을 분석하여 최근 사용되는 카카오톡 8점대 분석에는 일부 적합하지 않다.

본 논문에서는 카카오톡 메시지를 두 단계로 나누어 분석하였다. 우선 첫 번째 단계에서는 안드로이드 6 (마시멜로) 환경의 모바일용 카카오톡 앱과 윈도우즈 10 환경의 PC용 카카오톡 앱에서 메시지의 아티팩트를 추출한 뒤 모바일용, PC용 각각의 아티팩트 유형과 그 속성 정보를 비교 분석하였다. 국내 OS 시장에서 안드로이드와 윈도우즈의 점유율 [3]이 압도적이기 때문에 본 논문에서는 iOS 및 MAC PC용 카카오톡 분석은 제외하였다. 지금까지 기존 연구들은 모바일용 카카오톡과 PC용 카카오톡에서 획득 가능한 아티팩트에 일부 차이점이 존재함을 간과하였다. 즉, 분석한 카카오톡 아티팩트가 모바일용 카카오톡에서 획득한 것인지 PC용 카카오톡에서 획득한 것인지 명확히 구분 짓지 않고 '카카오톡 아티팩트'로 모호하게 지칭하는 경향이 있다. 그래서 본 논문에서는 모바일용 카카오톡과 PC용 카카오톡에서 획득할 수 있는 아티팩트 유형을 정확히 분류하고 공통점·차이점을 명시함으로써 일선 수사환경에서 모바일용 카카오톡 아티팩트와 PC용 카카오톡 아티팩트 정보를 상호 보완하여 활용할 수 있음을 보였다.

두 번째 단계에서는 아티팩트 분석 단계에서 획득한 카카오톡 메시지를 심층학습 (Deep Learning) 기법을 활용하여 감성 분석 (Sentiment Analysis) 하였다. 사용자의 사상 및 감정이 담긴 메시지는 다른 아티팩트와 달리 추가적인 분석이 요구된다. 메시지에 등장하는 주요 키워드 분석, 단어의 출현 빈도 분석 등 통계적 기반의 분석들은 종종 수사환경에 활용되어왔으나 메시지에 담긴 사용자의 감성 (sentiment), 정서 (affect), 주관 (subjectivity), 감정 (emotion)을 예측하는 심층학습 기반 감성 분석 [4]은 지금까지 시도된 바가 없다. 트위터, 인스타그램 등 인기있는 앱들에 대한 감성 분석 사례가 있긴 하지만 대부분 상업적인 목적으로 시장의 최신 트렌드를 연구하는데 활용되어왔기 때문에 온라인 메시지의 감성 분석을 수사에 활용하려는 시도는 최초이다. 본 논문에서는 카카오톡 메시지 감성 분석을 위한 카카오톡 메시지 말뭉치 (corpus, 또는 데이터셋) 생성부터 다양한 심층학습 모델을 활용한 데이터셋 학습 및 카카오톡 메시지 감성 분석에 적합한 모델을 선별하는 일련의 과정을 다루었다.

본 논문은 총 6장으로 구성되었다. 2장에서 관련 연구들과 배경지식을 설명하며, 3장에서는 카카오톡 메시지의 아티팩트 분류, 4장에서는 심층학습 기반 자연어 처리 기법을 활용한 카카오톡 메시지 감성 분석에 대해 각각 소개하였다. 5장에서는 3장, 4장의 분석 결과를 토대로 카카오톡과 같은 채팅 앱을 효과적으로 포렌식하는 절차를 제안하였으며, 6장에서는 본 논문이 지니는 의의와 향후연구에 대해 기술하였다.

제 2 장 관련 연구 및 배경지식

2.1 관련 연구

2.1.1 카카오톡 메시지 아티팩트 분류

채팅 앱에 대한 기존 연구들은 모바일용 앱 (iOS OS, 안드로이드 OS)에 대한 연구가 주를 이루었다. 그러나 최근에는 데이터 구조를 분석하기에 상대적으로 용이한 윈도우즈 OS 환경의 PC용 버전에 대한 연구가 증가하고 있다. Choi, et al. [5]은 채팅 앱 중 윈도우즈 OS용 카카오톡, 네이트온, QQ 메신저에서 각각 DB 파일들이 어떤 과정을 거쳐 암호화되는지 그 과정을 단계별로 분석하였다. 카카오톡 메시지 송·수신을 위한 키 생성 및 메시지 암호화 프로토콜에 대한 연구는 이 연구가 유일하다. 저자는 윈도우즈 7 환경에서 동작하는 PC용 카카오톡 (버전 2.0.8.990)이 디버거 (debugger)로부터 프로그램을 보호하기 위하여 Themida라는 안티디버깅 도구 (antidebugging tool)를 사용했음을 발견하고 이를 우회하기 위해 범용 도구인 Ollydbg (version 1.10)를 사용하였다. 또한 메시지 암호화·복호화 알고리즘으로 AES-128 CBC mode가 사용되었음을 확인하였다.

카카오톡 암호화 알고리즘 관련 연구에 반해 가시적인 분석이 가능한 아티팩트 분석은 상대적으로 다양하다. Azfar, et al. [6]은 30개의 인기 있는 안드로이드 채팅 앱을 선별하여 각각의 앱이 우선 Instant Messenger (IM), Voice over IP (VoIP), Augmentative and Alternative Communication (AAC) 이 세가지 특성 중 어디에 속하는지 구분한 뒤 XRY (모바일 포렌식 도구)를 통해 30개의 앱으로부터 DB를 논리적으로 추출하고 저자가 분류한 아티팩트 카테고리 (총 20개)에 각각의 앱이 어느 정도 부합하는지 연구하였다. 이 연구에서는 카카오톡 모바일 버전 4.7.6.을 사용하였는데 메시지를 통해 송·수신자, 메시지 및 채팅 유형, 메시지 내용, 차단친구 및 숨김친구 목록 등의 확인이 가능함을 보였다. 단, 비밀채팅 메시지는 일부만 확인이 가능하였고 그룹채팅에 참여하는 구성원이 누구인지 확인하는 것도 제한되었다. 윤종철 등 [7]은 안드로이드 스마트폰 3대를 루팅 (rooting)하여 카카오톡 모바일 버전 4.8.2. DB 파일의 저장 위치를 찾고 SQLite browser를 통해 각 파일의 내용을 확인하였다. 채팅 메시지 이력 (송·수신 여부), 메시지 내용 및 유형, 채팅방 구성원 등에 대해 확인할 수 있음을 보였다. 그런데 앞서 언급된 두 개의 아티팩트 분석 관련 논문이 모바일용 카카오톡 아티팩트에 대해서만 언급하고 있고, 4점대 버전이라 최근 사용되고 있는 8점대 버전에 적용하여 분석 및 이해하는데 일부 제한이 있었다.

한편, 카카오톡 메시지가 삭제되었을 경우 복구하는 방법을 다룬 연구 또한 그 수가 많지는 않으나 윤종철 등 [7]이 메시지를 삭제하거나 특정 채팅방을 나가더라도 백업 파일이 존재한다면 이를 이용해 메시지를 복원할 수 있음을 보였고, Choi, et al [5] 역시 사용자가 카카오톡 대화내용을 백업할 때 단순한 패턴의 패스워드를 사용했다면 패스워드를 우회하여 채팅 메시지를 알아낼 수 있다고 하였다. 또한 채팅 앱 이용 시 암호화된 채널 (SSL/TLS)을 이용하더라도 네트워크 패킷 수집 / 분석을 통해 카카오톡 사용자의 행위를 높은 확률로 유추해낼 수 있다는 Park, et al [8]의 연구도 있었다. 저자는 지도 학습 (supervised machine learning)을 통해 통상적인 카카오톡 사용자의 행위를 학습시킨 뒤 수집한 패킷을 통하여 평균 99.7%의 확률로 카카오톡 사용자의 행위를 분류하였다.

2.1.2 카카오톡 메시지 감성 분석

임좌상 등 [9]은 트위터에서 한국어로 작성된 트윗의 긍정·부정 여부를 통계적 기반 기계학습을 통해 분류하였다. 감정 분류 대상인 텍스트에서 단어를 추출하고 그 단어에 해당하는 사전에 정의된 감정값을 추출하여 문장의 감정 정도를 계산한 뒤 그 결과값으로 긍정·부정을 분류하였다. 실험결과 약 76%의 분류 정밀도 (Precision, 긍정이라고 판정한 것 중 실제 긍정 수)를 보이는 반면 민감도 (Sensitivity, 전체 긍정 수 중 판정된 긍정 수)가 낮아 학습 방법의 업데이트가 요구되었다. 김세진 등 [10]은 인터넷 용어 중 자주 사용되는 비속어, 줄임말, 은어를 해석하기 위한 감성 사전을 Word2Vec을 활용하여 구축하고 감성 사전 구축 전과 후의 정확도를 비교하였다. 비속어 등의 감성 사전을 구축한 경우 (96.1%)가 구축하지 않은 경우 (87.2%) 보다 '헐', '미친', '걸크', '꿀팁' 등 비속어, 줄임말, 은어에 대한 긍정, 부정 등의 극성을 잘 판별해내는 것을 보였다. 박천음 등 [11]은 문맥 표현 중 하나인 ELMo (Embedding from Language Models) 기법을 소개하고 한국어 영화평 감성 분석을 위하여 NSMC (Naver Sentiment Movie Corpus)와 다음카카오 (DaumKakao) 데이터를 셀프 어텐션 (self attention)과 ELMo를 이용한 순환 신경망 (Recurrent Neural Network, RNN)으로 학습시키는 모델을 제안하였다. 기존 모델에 비해 성능이 좋음을 입증하였다.

위 연구들과 같이 특정 대상에 대한 의견을 수집하고 분류하는 과정인 감성 분석은 주로 마케팅 영역에서 활용되어왔다. 설문, 포럼, 블로그, 웹사이트 등에서 여론을 수집하고 그 흐름을 분석하여 시장의 요구에 기민하게 대응하기 위하여 감성 분석을 활용하였다. 한국어 대상 연구도 활발하다. 하지만 상업적인 목적 외에는 연구가 더딘 실정이며, 특히 카카오톡 등 일부 소셜 네트워크 서비스 (이하 SNS)에 대한 감성 분석은 연구가 희박하다. 데이터셋 구성에 어려움이 있기 때문이다. 상품평, 영화평점 및 리뷰는 사용자가 평가할 대상 (상품, 영화 등)이 존재하기 때문에 긍정·부정 등 사용자의 주관이 담긴 데이터를 웹, 블로그, 트위터 등으로부터 수집하는데 어려움이 없다. 반면 카카오톡은 다양한 주제로 개인의 의견을 대화 형식으로 표현하는 사적 영역의 창구이므로 상업적인 목적으로 활용하기에는 제한이 있어 각각의 의견의 극성이 무엇인지, 그 극성이 무엇으로부터 기인했는지 주목하지 않는다. 하지만 수사환경에서는 메시지 하나하나가 사용자의 상태를 유추할 수 있는 중요한 실마리가 될 수 있기 때문에 다양한 관점의 연구가 필요하다.

2.2 배경지식

2.2.1 디지털 포렌식

디지털 포렌식 (Digital Forensics)은 컴퓨터, 모바일, 네트워크를 대상으로 물리적 증거 및 전자적 증거 (Electronic evidence)를 식별, 수집, 분석하고 정보를 특정하여 법정에 증거로 제출하는 등 검증 가능한 형태로 자료를 준비하는 절차를 말한다. 디지털 포렌식에서 활용하는 증거는 생성 증거와 보관 증거로 분류할 수 있다. 생성 증거는 시스템이나 앱이 자동으로 생성한 레지스트리, 로그 등의 데이터로 통상 아티팩트 (artifact)라 함은 이 생성 증거를 의미한다. 보관 증거는 사람의 사상이나 감정을 표현하기 위해 작성한 데이터로 블로그 및 소셜 네트워크 게시글, 메일 등이 해당된다. 최근에는 컴퓨터, 모바일 기타 디지털 저장매체 등에 저장되어 있는 자료와 네트워크를 통해 전송 중인 자료를 포함하여 포렌식 조사 및 수사 업무에 필요한 증거들 (흔적들)을 아티팩트로 통칭하기도 한다. 본 논문에서는 카카오톡으로부터 얻을 수 있는 모든 증거를 아티팩트라고 하였다. 수사환경에서 증거 수집 시에

반드시 고려해야 할 점은 보관 증거는 사용자의 사상이나 감정 즉 주관과 고의가 들어간 데이터이기 때문에 전문 법칙이 적용된다는 점이다.

2.2.2 디지털 포렌식 절차

ISO / IEC 27043: 2015, "Information Technology-Security Techniques-Investigation principles and processes [12]" 는 국제표준화기구 (ISO / IEC: International Organization for Standardization / International Electrotechnical Commission Joint Technical Committee)에 의해 제정된 정보기술, 보안 기술, 사고 조사 원칙 및 절차의 국제표준규격이다. 이 중 디지털 포렌식 절차 표준은 5단계인 준비 단계 (Readiness process), 초동 단계 (Initialization process), 획득 단계 (Acquistive process), 조사 단계 (Investigative process), 병합 단계 (Concurrent process)로 구분된다. 준비 단계는 디지털 포렌식 예비 단계로 조사 기준을 수립하고 효과적인 포렌식 절차를 고민하는 단계이며, 초동 단계에서는 사건 발생 초기 실질적인 조사계획을 수립한다. 획득 단계에서 수사관은 디지털 기기 유무를 확인 (identify), 확보 (acquire), 획득 (collect), 이동 (transport) 시키며, 조사 단계에서는 수집된 디지털 데이터를 분석한 뒤 조사를 마무리한다. 병합 단계는 앞의 4단계와 동시에 진행되는 단계로 모든 절차를 문서화하고 디지털 증거들이 법적 증거로 활용될 수 있도록 무결성 확보에 집중한다.

모바일 포렌식 절차 표준으로는 NIST SP 800-101 Revision1, Guidelines on Mobile Device Forensics [13]이 2006년에 처음 제안되었고 새로운 모바일 기기 등장에 따라 2014년 5월 마지막으로 수정되었다. 위 문서에서는 모바일 포렌식을 보존 (Preservation), 획득 (Acquisition), 점검 및 분석 (Examination & Analysis), 보고 (Reporting)의 4단계로 구분하여 설명하고 있다. 보존 단계는 디지털 데이터들이 전파, 네트워크 간섭 등을 통해 변조되는 위험을 막고 원래의 상태를 유지하도록 하는 단계로 수사관이 사건 현장에서 발견한 모바일 기기의 전원을 우선 차단하는 것 등이 이에 해당한다. 획득 단계에서는 모바일 기기를 수집하여 포렌식 도구를 통해 내부 데이터를 추출하고, 점검 및 분석 단계에서 수집된 모바일 데이터 목록을 작성한 뒤 이를 포렌식 분석 도구를 활용하여 분석한다. 마지막 보고 단계에서는 포렌식 결과에 대한 보고서를 작성한다. ISO/IEC와 NIST에서 제안한 모델들은 단계 분류에 일부 차이가 있으나 핵심 내용들은 매우 유사함을 알 수 있다.

국내에서는 2007년 한국정보통신기술협회 (Telecommunications Technology Association, TTA)에서 정보통신단체표준 (TTAS)으로 이동 전화 포렌식 가이드라인 [13]을 제시하였다. 디지털 포렌식 절차는 수사 준비, 증거 수집, 증거 보관 및 이송, 증거 분석, 보고서 작성 등 5단계로 구성된다. 수사 준비 단계에서는 포렌식 도구의 구비 및 장비 점검, 수사관에 대한 교육을 하며, 증거 수집 단계에서는 현장 사진을 촬영하고 압수수색 영장이 허용하는 범위에서 하드웨어, 소프트웨어, 보조기억장치 등을 수집한다. 증거 보관 및 이송단계에서는 디지털 증거물이 이송 및 보관 과정에서 손상되지 않도록 쓰기 방지 조치 등을 취하며, 증거 분석 단계에서는 상황에 따라 적절한 포렌식 도구를 이용하여 디지털 증거를 검색·복구한 뒤 최종 보고서를 작성한다. 모바일 포렌식 절차는 디지털 포렌식 절차를 기반으로 세분화하여 사전 준비, 초기 대응, 증거 수집, 기기 포장 및 운송 후 보관, 조사 및 분석, 보고서 작성 등 총 6단계로 구분하였다. 초기 대응 과정에서 기기가 정상 작동을 하는 경우, 전원이 꺼져 있는 경우, 외관상 파손 등 비정상적인 상태인 경우 등의 상황을 고려하여야 하며, 증거 수집 역시 물리적 방법과 논리적 방법으로 구분하여 누락되는 증거가 없도록 할 것을 당부하고 있다.

최근에는 소셜 네트워크 서비스 (SNS)에 대한 포렌식 절차 연구도 활발하다. Jang, et al. [14]은 SNS에 대한 조사 준비, 초기 대응, 디지털 증거 보존 및 수집 등 3단계로 나누어 PC 및 모바일 기기에 대한 포렌식 절차를 제안하였고, Yusoff, et al. [15]은 Firefox OS 환경에서 모바일용 SNS 앱에 대한 포렌식 절차를 준비 및 보존 (Preparation and Preservation), 획득 (Acquisition), 점검 및 분석 (Examination and Analysis)의 3단계로 나누어 설명하였다.

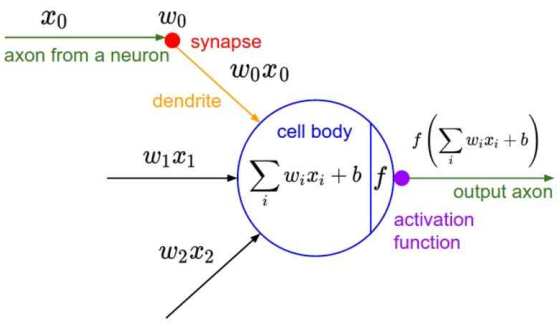
본 논문에서는 모바일용 카카오톡 및 PC용 카카오톡 아티팩트 분석 과정과 심층학습 기반 감성 분석을 통해 카카오톡 메시지 작성자의 심리상태를 추정하는 과정, 그리고 위 포렌식 절차 표준들과 국내·외 SNS 포렌식 절차 연구들을 기반으로 채팅 앱 전반에 대한 효과적인 포렌식 절차를 제안하였다.

2.2.3 기계학습

기계학습 (Machine Learning)은 인공 지능 (Artificail Intelligence)의 한 분야로 인공 지능 체계가 주어진 데이터로부터 스스로 규칙성을 찾는 학습 (training, 또는 훈련) 과정을 거쳐 새로운 지식을 얻어내는 분야이다. 기계학습은 이미지 인식, 영상 처리, 알파고와 같은 분야뿐만 아니라 번역, 감성 분석 등 자연어 처리에 있어서도 유용하게 쓰인다. 기계학습의 역사는 인공 신경망 (Artificial Neural Network) 모델이 등장한 40년대 후반부터 시작되었지만 낮은 컴퓨터 성능, 모델의 한계 등으로 80년대까지 답보상태이다가 1986년 다층 퍼셉트론 (Multi-Layer Perceptron, MLP)의 지도 학습을 위한 역전파 알고리즘 (backpropagation algorithm)과 2000년대 중반 인공 신경망을 복잡하게 쌓아 올린 심층학습 (Deep Learning)의 제안으로 다시 활성화되었다.

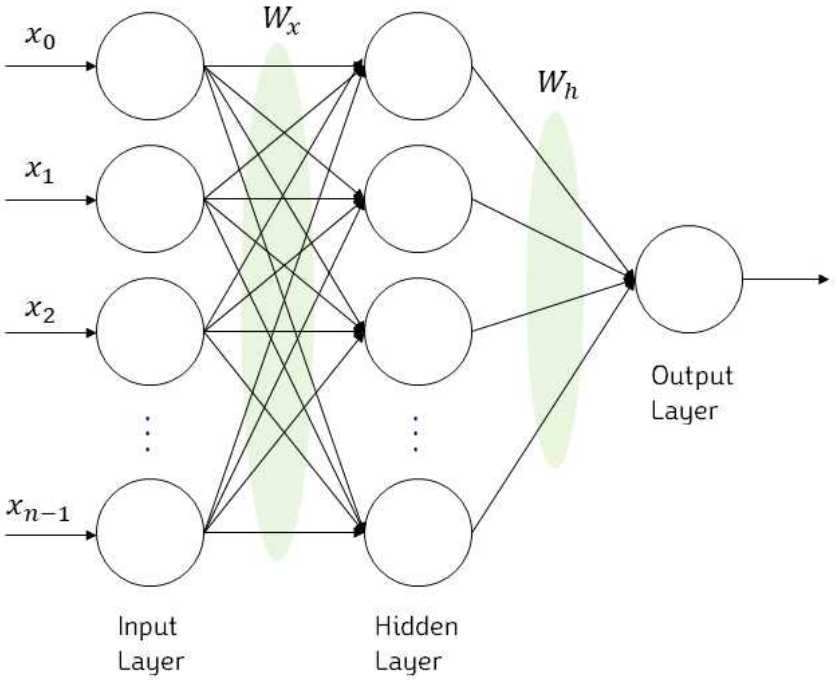
인공 신경망은 인간의 뇌의 학습방법을 수학적으로 모델링한 알고리즘이다. 그림 2.1과 같이 신경계에서 뉴런 (neuron)은 신경세포체 (cell body), 가지돌기 (dendrite), 축삭돌기 (axon)로 구성되며 뉴런과 뉴런의 연결은 시냅스 (synapse)가 담당한다. 먼저 가지돌기 (dendrite)가 이전 뉴런의 축삭돌기 말단으로부터 신호를 받아들이고 이 신호가 일정치 이상의 크기를 가지면 축삭돌기를 통해서 다음 뉴런으로 신호를 전달하게 된다. 신호 전달의 연접 부위인 시냅스의 결합 세기에 따라 뉴런의 성능이 좌우된다. 수식으로 표현해보면 그림 2.1에서 x는 입력값 (input), w는 가중치 (weight)를 의미하는데 시냅스의 결합 세기를 인공 신경망에서는 가중치 w가 대신한다. 각각의 인공 뉴런에서 보내진 입력값 x는 각각의 가중치 w와 곱해져서 종착지인 인공 뉴런에 전달되고 각 입력값과 그에 해당되는 가중치의 곱의 전체 합은 특정한 함수와 만나 인공 뉴련의 출력 신호 $y = f(\sum w_i x_i + b)$를 출력하게 된다 (b: bias). 이렇게 뉴런에서 출력값을 변경시키는 함수 f를 활성화 함수 (activation function)라고 하며, 이때 가중치의 값이 크면 클수록 해당 입력 값이 중요하다는 것을 의미한다.

그림 2.1: 인공 신경망의 뉴런 [16]



일반적으로 사용되는 기본적인 인공 신경망 알고리즘인 다층 퍼셉트론 (MLP)을 살펴보면 그림 2.2와 같이 입력층 (input layer), 은닉층 (hidden layer), 출력층 (output layer)으로 구성된다. 각 층의 노드들은 서로 연결되어 있어 만약 n개의 입력값이 있다면 입력층은 n개의 노드를 가지게 된다. 은닉층은 모든 입력 노드로부터 입력값 $x_{0:n-1}$을 받아 각각의 가중합을 계산하고 이 값을 활성화 함수 f에 적용하여 출력층에 전달하게 된다. 층별 각각의 가중치 w (입력값에 대한 가중치 집합 W_x, 은닉층의 출력값에 대한 가중치 집합 W_h)는 연결 강도로 표현되며 초기에는 랜덤으로 주어졌다가 예측값을 가장 잘 맞추는 값으로 조정되게 된다. 활성화 함수는 비선형 함수를 사용하므로 인공 신경망이 비선형 모델로서 역할을 할 수 있다.

그림 2.2: 은닉층이 1개인 다층 퍼셉트론 (MLP)



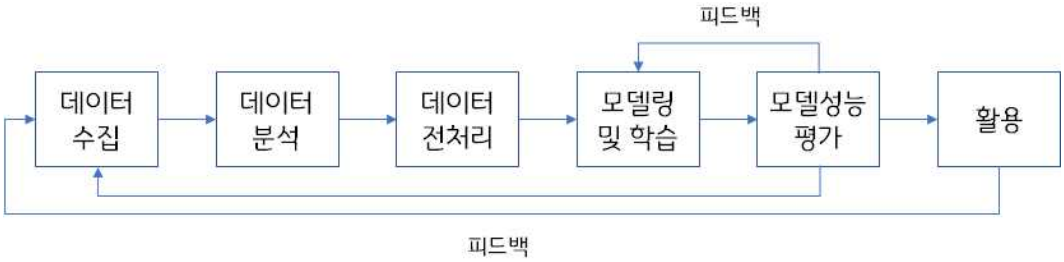
다층 퍼셉트론 (MLP) 이외에도 순환 신경망 (Recurrent Neural Network, RNN), 장·단기 메모리 (Long-Short-Term-Memory, LSTM), 합성곱 신경망 (Convolutional Neural Network, CNN) 등이 사용빈도가 높은 인공 신경망 알고리즘들로 본 논문에서 카카오톡 메시지 데이터셋을 학습시키는데 활용하였다.

2.2.4 기계학습 절차

기계학습은 그림 2.3의 기계학습 절차를 따라 진행된다. 우선 학습시켜야 할 데이터를 수집하는 것 (Data Acquisition)이 중요한데 자연어 처리의 경우 자연어 데이터셋을 말뭉치 또는 코퍼스 (corpus)라고 부르며 웹 크롤링 (crawling) 등을 통해 수집된 SNS 메시지, 영화 리뷰 등이 이에 해당한다. 파일 형식은 txt 파일, csv 파일, xml 파일 등 다양하다. 수집된 데이터는 구조, 노이즈 등을 확인하고 기계학습을 위해 데이터를 어떻게 전처리할 것인가에 대해 분석 (Data Analysis) 한 뒤 빠르고 정확한 학습을 위하여 토큰화, 정제, 정규화, 불용어 제거 등의 전처리 (Data Preprocessing)를 한다.

특히 한국어 토큰화에서는 형태소 (morpheme) 개념을 반드시 이해하고 형태소 분석 및 태깅을 하여야 한다. 기계학습에 대한 코드를 작성하는 단계인 모델링 (Modeling) 단계에서는 기존의 기계학습 알고리즘을 선택하거나 또는 새로운 모델을 고안한 뒤 전처리가 완료된 데이터를 학습시킨다. 통상 학습용 데이터셋을 활용하며 학습 후에는 검증용 데이터셋으로 학습 정도 (정확도, 손실률 등)를 검증하고 이를 기준으로 모델의 성능에 영향을 주는 초매개변수 (hyper parameter)인 은닉층의 수, 뉴런의 수, 드롭아웃 (dropout) 비율 등을 조정한다. 테스트 데이터셋은 모델의 최종 성능을 평가하는 데이터로 평가 방법은 모델이 예측한 데이터가 테스트용 데이터셋의 실제 정답과 얼마나 가까운지를 측정하게 된다. 이진분류에서는 성능 평가에 주로 혼동 행렬 (Confusion Matrix)을 활용한다. 모델이 성공적으로 훈련되었다면 감성 분석, 기계 번역, 음성 인식, 텍스트 분류 등의 자연어 처리 작업에 활용되도록 배포하게 된다.

그림 2.3: 기계학습 절차



2.2.5 자연어 처리

자연어 (Natural Language)는 우리가 일상생활에서 사용하는 언어로, 자연어 처리 (Natural Language Processing, NLP)란 이러한 자연어의 의미를 분석하여 컴퓨터가 처리할 수 있도록 하는 기계학습 기법 중 하나이다. 음성 인식, 내용 요약, 번역, 감성 분석, 텍스트 분류 작업 (스팸 메일, 뉴스 카테고리 분류 등), 챗봇 구축 등에 사용된다. 본 논문에서는 카카오톡 메시지 감성 분석을 위하여 우선 긍정 (1), 부정 (0)으로 라벨링 된 독립적인 카카오톡 메시지 말뭉치를 생성하고 적합한 인공 신경망 구조를 설계하여 예측값 (1 또는 0)을 구한다. 이 과정이 순전파 (feedforward)이다. 그런 다음 예측값을 목표치인 실제 라벨링과 비교하여 오차를 계산하고 오차로부터 가중치와 편향을 업데이트하는 학습 단계를 거친다. 이를 역전파 (backpropagation)라 한다.

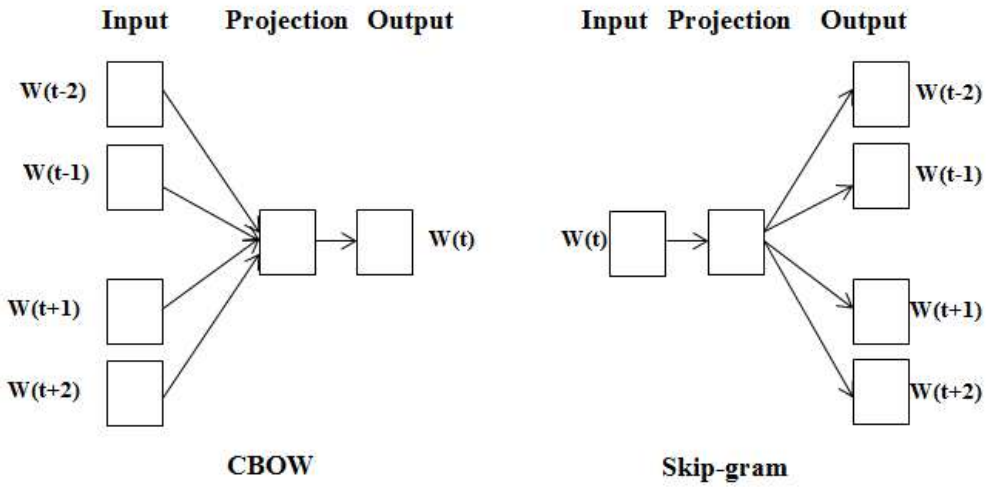
그런데 감성 분석에 사용할 '적절한' 인공 신경망을 선택하는 기준은 별도로 제시된 바가 없다. 고정된 길이의 입력값을 갖는 다층 퍼센트론 (MLP)의 단점을 순환 신경망 (RNN)이 보완하였고, 순환 신경망 (RNN)의 메모리 문제를 장·단기 메모리 (LSTM)가 보완하였다고 해서 장·단기 메모리 (LSTM)가 가장 우수한 알고리즘이라고 단정 지어 말할 수 없다. 또한 합성곱 신경망 (CNN)이 이미지 인식에 좋은 성능을 보이지만 자연어 처리에서도 꾸준히 활용되고 있다. Yin et al. [17]은 순환 신경망 (RNN)과 합성곱 신경망 (CNN)을 통해 감성 분석, 품사 태깅 등 자연어 처리를 진행한 결과 둘 중 어느 알고리즘이 우위에 있다고 결론 내릴 수 없었다고 하였다.

따라서 본 논문에서는 다양한 인공 신경망 알고리즘을 학습에 활용하고 성능을 평가하여 카카오톡 메시지 감성 분석에 적합한 모델을 식별하고자 하였다. 단순히 알고리즘의 종류만 달리하지 않고 각각의 인공 신경망 알고리즘을 단독 혹은 결합하여 학습에 활용하였다. 또한 2013년 구글 연구팀에서 제안한 뒤 지속적인 인기를 누리고 있는 단어 임베딩 기법인 Word2Vec, 2018년 제안된 문맥 표현 기법인 BERT로 인해 더욱 주목받기 시작한 어텐션 (Attention) 기법을 인공 신경망 알고리즘에 적용하여 학습 성능을 각각 평가하였다.

2.2.5.1 Word2Vec

2013년 구글 연구팀인 Milokov et al. [18]이 제안한 Word2Vec은 고품질의 단어 벡터를 효율적으로 구축하기 위해 CBOW (Continuous-Bag-Of-Words)와 Skip-gram 모델을 제안했다.

그림 2.4: Word2Vec 두 가지 모델 [18]



CBOW는 k개만큼의 주변 단어가 주어졌을 때 중심 단어의 조건부확률을 계산한다. Skip-gram 모델은 CBOW와 정반대로 중심 단어가 주어졌을 때 주변 단어를 예측한다. 방향이 다를 뿐 기본적인 매커니즘은 같다. 둘 중 상대적으로 성능이 좋다고 평가받는 Skip-gram에 대해 살펴보면 중심 단어 t와 주변 단어 c 쌍이 주어졌을 때 (W: Word, 단어) t와 c의 연관성이 높으면 긍정 샘플 (positive), 연관성이 낮으면 부정 샘플 (negative)로 이진 분류하는 과정에서 학습된다. Milokov et al.은 말뭉치에 자주 등장하지 않는 희귀한 단어가 네거티브 샘플로 더 잘 뽑힐 수 있도록 Skip-gram을 설계하였다. 이렇게 학습하는 기법을 네거티브 샘플링 (negative sampling)이라고 한다. 네거티브 샘플 확률은 아래 수식과 같다.

$$P_{negative}(w_i) = \frac{f(w_i)^{3/4}}{\sum_{j=0}^{n} f(w_j)^{3/4}}$$

$f(w_i)$란 해당 단어가 말뭉치에서 차지하는 비율 (해당 단어 빈도 / 어휘 집합 크기)을 의미한다.

Milokov et al.은 Word2Vec에 네거티브 샘플링과 함께 자주 등장하는 단어를 학습에서 제외하는 서브 샘플링 (sub sampling) 기법도 적용하였다. Skip-gram 모델은 말뭉치로부터 학습 데이터 쌍을 많이 만들어낼 수 있기 때문에 출현 빈도가 높은 단어의 경우 등장 횟수만큼 모두 학습시키는 것이 비효율적이라고 판단한 것이다. 서브 샘플 확률은 아래 수식과 같다.

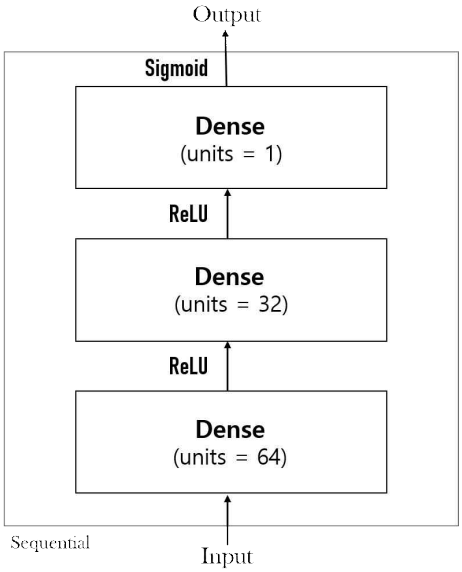
$$P_{subsampling}(w_i) = 1 - \sqrt{\frac{t}{f(w_i)}}$$

t는 앞서 설명한 바와 같이 중심 단어 즉 타깃 단어이다. 서브 샘플링은 학습량을 효과적으로 줄여 계산량을 감소시킨다.

2.2.5.2 다층 퍼셉트론 (MLP)

입력층과 출력층으로 이루어진 단층 퍼셉트론의 형태에서 은닉층이 1개 이상 추가된 신경망을 다층 퍼셉트론 (MLP)이라고 한다. 그림 2.2와 같이 다층 퍼셉트론 (MLP)은 입력층에서 출력층 방향으로 연산이 전개되는 순방향 신경망 또는 피드 포워드 신경망 (Feed-Forward Neural Network, FFNN)의 가장 기본적인 형태이다. 또한 은닉층과 출력층에 있는 모든 뉴런이 바로 이전 층의 모든 뉴런과 연결돼 있는 전결합층 (Fully-Connected layer, or Dense layer)만으로 구성되어 있어 전결합 피드 포워드 신경망 Fully-Connected FFNN)이라고도 한다. 케라스에서는 전결합층을 구현할 때 Dense() 함수를 사용한다. 본 논문에서 설계한 다층 퍼셉트론 (MLP)은 그림 2.5와 같이 은닉층이 2개인 심층 신경망 (Deep Neural Network)이며 노드 (units)는 각각 64개, 32개로 설정하였다. 활성화 함수로는 ReLU와 이진 분류를 위해 Sigmoid 함수를 활용하였다.

그림 2.5: 은닉층이 2개인 다층 퍼셉트론 (MLP) 모델



2.2.5.3 순환 신경망 (RNN)

순환 신경망 (Recurrent Neural Network, RNN)은 그림 2.6과 같이 시점 (time step)이 현 시점 t를 기준으로 일정 구간 연속된다고 가정할 때 은닉층의 노드에서 활성화 함수 f를 통해 나온 결과값 h_{t-1}을 출력층 방향으로 보내는 동시에 은닉층 노드의 다음 계산의 입력으로 보내는 특징을 갖고 있어 다양한 길이의 입력값과 출력값을 받아들일 수 있는 네트워크 구조이다.

그림 2.6: 기본 순환 신경망 구조 (simple RNN)

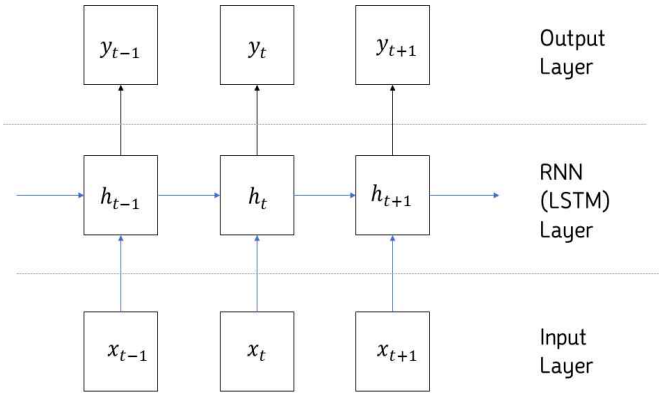
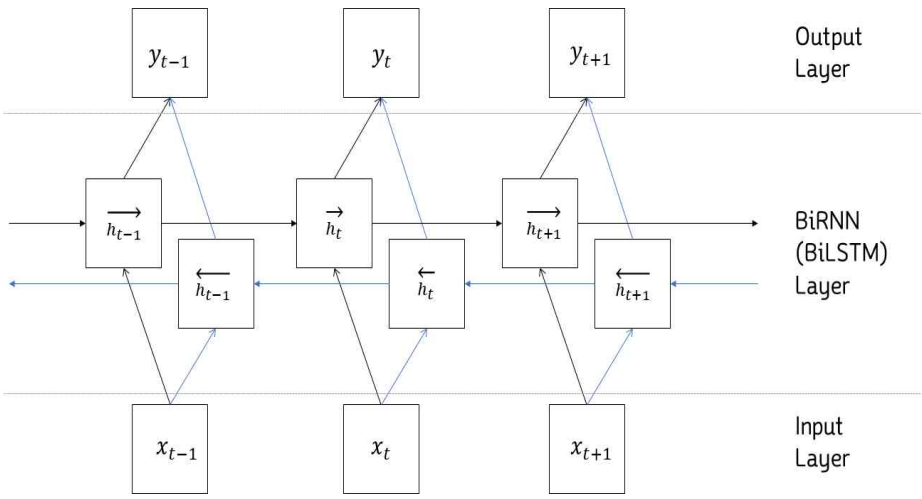


그림 2.6 기본 순환 신경망 구조 (simple RNN)를 수식으로 나타내면 아래와 같다.

$$h_t = f(W_{xh}x_t + W_{hh}h_{t-1} + b_h)$$
$$y_t = W_{hy}h_t + b_y$$

입력값 x_t와 h_{t-1}을 각각의 가중치 W_{xh}와 W_{hh}와 곱하여 더한 값이 은닉층 활성화 함수 f의 입력값으로 활용되어 은닉층의 출력값인 은닉상태 h_t가 됨을 보여준다. 출력층을 거치며 가중치 W_{hy}를 반영한 h_t의 최종 결과값은 y_t이며 각 층의 편향 (bias)은 b로 표시하였다.

그림 2.7: 양방향 순환 신경망 구조 (BiRNN)



양방향 순환 신경망 (Bidirectional RNN, BiRNN)은 그림 2.7과 같이 전방향 상태 (forward states) 정보를 가지고 있는 은닉층과 후방향 상태 (backward states) 정보를 가지고 있는 은닉층 2개로 구성되어 있으며 이 둘은 서로 연결되어 있지 않다. 현 시점 t (time step)를 기준으로 전방향 상태는 입력값을 t=1부터 순차적으로, 후방향 상태는 입력값을 마지막 시점부터 역순으로 주고 학습을 하게 된다. 시점마다 양방향에서 나온 2개의 출력은 각각 학습된 가중치 W_{xh}와 W_{hh}를 통해 하나의 은닉 벡터 h로 만들어지며 출력층을 거쳐 가중치 W_{hy}를 반영한 최종 출력값 y을 계산한다. 전방향 은닉상태를 표시한 아래 수식에서 알 수 있듯이 벡터 방향을 표시한 것 외에는 순환 신경망 (RNN)과 수식이 동일하다.

$$\vec{h}_t = f(W_{x\vec{h}}x_t + W_{\vec{h}\vec{h}}\vec{h}_{t-1} + b_{\vec{h}})$$

2.2.5.4 장·단기 메모리 (LSTM)

순환 신경망 (RNN)은 다양한 길이의 입력값과 출력값을 받아들일 수 있는 네트워크 구조이지만 그 길이가 길어질수록 역전파 시 학습 능력이 크게 저하된다 (vanishing gradient problem). 이를 해결하기 위해 1997년 S.Hochreiter와 J.Schmidhuber가 장·단기 메모리 (Long-Short Term Memory, LSTM)를 제안하였다 [19]. 장·단기 메모리 (LSTM)는 그림 2.6, 그림 2.7의 순환 신경망 (RNN) 및 양방향 순환 신경망 (BiRNN)과 기본 구조는 같으나 그림 2.8과 같이 은닉층에서의 동작이 순환 신경망 (RNN)에 비해 세분화된 것이 특징이다. 장·단기 메모리 (LSTM)는 은닉층의 메모리 셀에 입력 게이트 (input gate), 망각 게이트 (forget gate), 출력 게이트 (output gate)를 추가하여 불필요한 기억을 지우고 기억해야 할 것들을 정한다.

그림 2.8: 순환 신경망 (RNN)과 장·단기 메모리 (LSTM)의 은닉층 구조 [20]

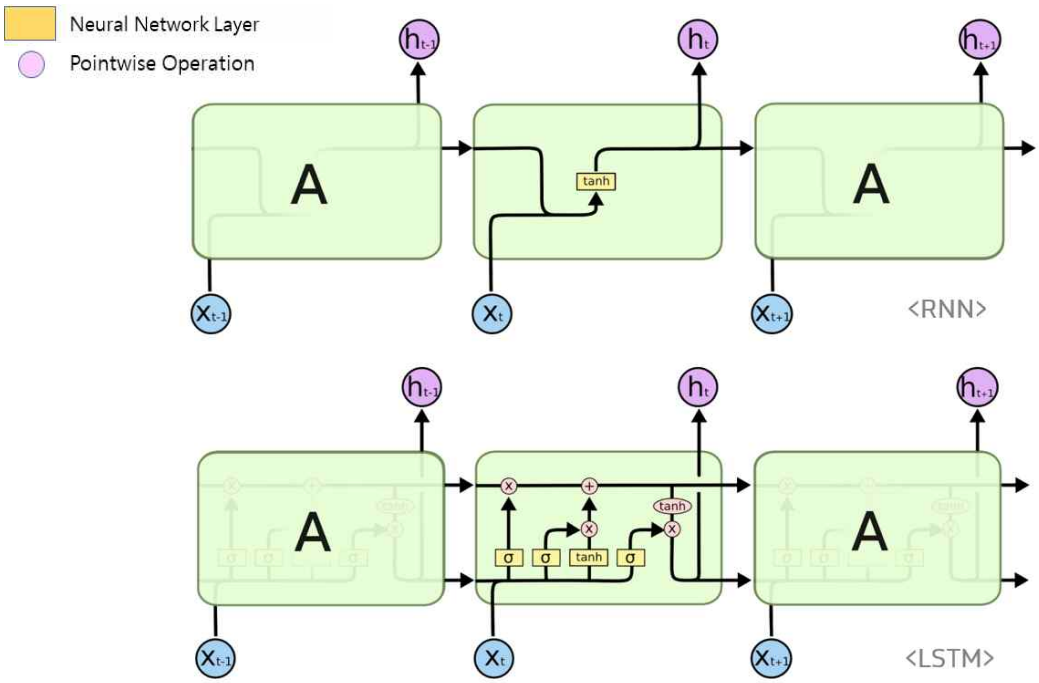
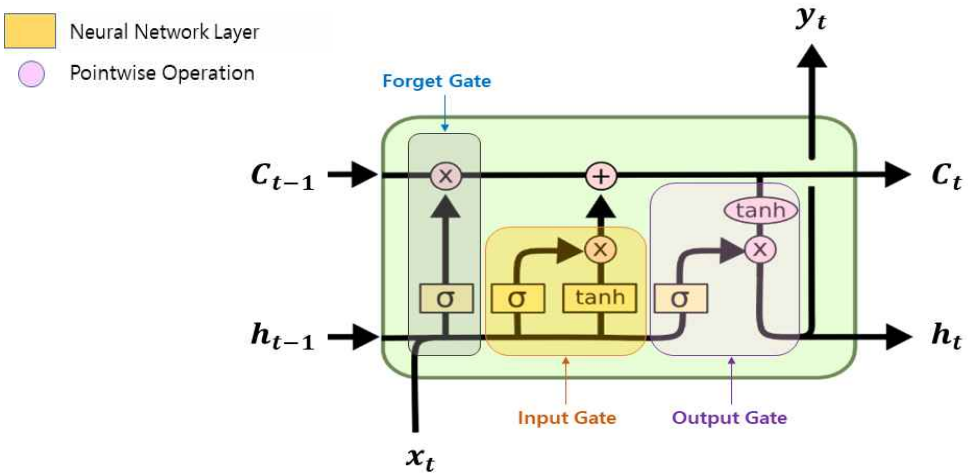


그림 2.9에서 좀 더 자세히 살펴보면 LSTM 셀에서는 상태 (state)가 두 개의 벡터 h_t와 c_t로 나뉜다. h_t는 단기 상태 (short-term state), c_t는 장기 상태 (long-term state)이다. LSTM의 핵심은 네트워크가 장기 상태 (c_t)에서 기억할 부분, 삭제할 부분, 읽어 들일 부분을 학습하는 것이다. 장기 기억 c_{t-1}은 셀의 왼쪽에서 오른쪽으로 이동 시 망각 게이트 (σ: sigmoid)를 지나면서 일부 기억 (정보)을 잃고, 다음 덧셈(+) 연산 시에 입력 게이트로부터 새로운 기억 일부를 받는다. 이는 장기 기억 c_t로 출력된다. c_t는 시점 (time step)마다 일부의 기억을 삭제하고 추가하는 과정을 반복하게 된다. 또한 c_t는 덧셈 연산 후 출력 게이트의 활성화 함수 (하이퍼볼릭탄젠트 함수, tanh)로도 전달되어 단기 상태 h_t와 셀의 출력인 y_t를 만든다. 세부적인 수식은 [20]으로 갈음한다. 양방향 장·단기 메모리 (BiLSTM)는 양방향 순환 신경망 (BiRNN)과 같이 전방향, 후방향 상태를 가지게 된다.

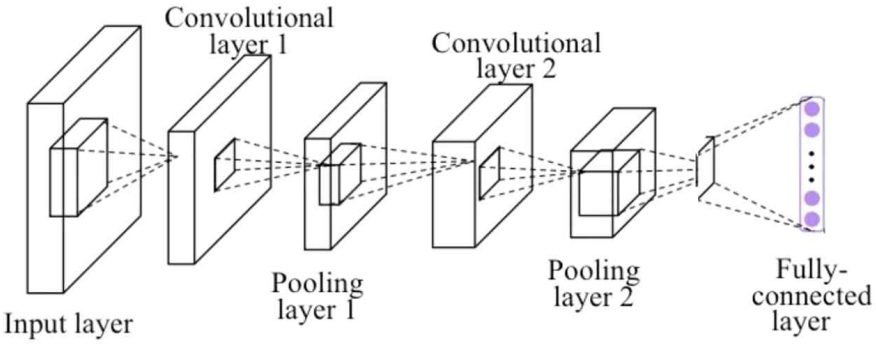
그림 2.9: 장·단기 메모리 (LSTM) 은닉층의 세부 구조 [19]



2.2.5.5 합성곱 신경망 (CNN)

합성곱 신경망 (Convolutional Neural Network, CNN)은 생물의 시신경이 동작하는 원리에서 영감을 얻어 이미지 데이터의 처리에 적합한 구조로 만들어진 신경망이다. CNN은 그림 2.10과 같이 입력된 이미지로부터 계층적 구조의 자질 (feature)을 추출하는 역할을 하는 합성곱층 (convolution layer) 및 풀링층 (pooling layer), 추출된 자질을 입력받아 목표치로 분류하는 역할을 하는 전결합층 (fully connected layer)으로 구성된다. 합성곱층에서는 행렬 형태로 표현된 이미지의 가중치 영역에 해당 이미지보다 사이즈가 작은 필터 (filter) 행렬을 이동시키면서 대응되는 원소끼리 곱한 후 합을 구하는 합성곱 연산을 한다. 풀링층에서는 합성곱으로 얻어진 값들에 대해 특정 사이즈 영역별로 최대값, 평균값 등을 추출하는 연산을 한다. 이러한 연산을 통해 입력된 이미지로부터 유용한 자질들이 계층적으로 추출되며 추출된 자질들은 하나 이상의 전결합층을 통해 입력 데이터를 목표치에 부합하게 분류하는 작업에 활용된다 [21]. 자연어 처리에 CNN을 적용하면 필터 (filter)의 윈도우 사이즈에 따라 지역적인 n-그램 (n-gram) 자질들을 추출할 수 있고 이를 기반으로 상위 계층의 자질들까지도 추출이 가능하여 신뢰도 높은 단어의 의미 유사도 계산 및 학습이 가능하다.

그림 2.10: 합성곱층과 풀링층 쌍이 2번 반복되는 합성곱 신경망 (CNN) 구조 [22]



2.2.5.6 어텐션 (Attention)

문장의 긍정, 부정을 판단하기 위하여 문장의 모든 단어가 동일하게 중요하지는 않다. 어텐션 (Attention) 기법은 카카오톡 메시지의 긍정, 부정을 분류함에 있어 매 시점 (time step)마다 전체 입력 문장을 집중해서 살펴보고 연관이 있는 단어들의 가중치 (attention weights or score)를 조정해 최종 분류에 활용하는 방법이다. 그림 2.11은 어텐션 기반 양방향 장·단기 메모리 (BiLSTM)의 구조를 나타내며 아래 식은 문서 길이가 n인 입력값에 대해 어텐션 기법을 적용한 양방향 장·단기 메모리 (BiLSTM)의 동작 과정을 간단하게 보여준다.

$$f(h_i) = \tanh(Wh_i + b)$$

먼저 f는 은닉층의 출력 벡터 h_i를 입력으로 받아 해당 출력의 감정 점수를 -1~1 사이로 부여한다.

$$a_i = \frac{e^{f(h_i)}}{\sum_{j=1}^{n} e^{f(h_j)}}$$

그리고 은닉층의 마지막 단계에서 입력값을 1부터 n까지 집중해서 살펴보고 어텐션 (Attention) 가중치 a_i를 계산한다. a_i는 f의 결과 $f(h_{1:n})$을 활성화 함수 (softmax)를 통해 확률 변수로 변환한 결과이다.

$$c = \sum_{i=1}^{n} a_i \times h_i$$

은닉층 출력 $h_{1:n}$과 어텐션 가중치 $a_{1:n}$의 가중합을 임의의 문서 벡터 c라고 하면, 문서 벡터 c는 어텐션 가중치가 높은 은닉층 출력에 높게 반영되어 최종 긍정, 부정 분류에 영향을 미침을 알 수 있다. 가령 “너를 좋아해.”라는 문장이 입력으로 들어오는 경우 ‘너’, ‘를’ ‘좋아’, ‘하다’로 토큰화되고, 학습 후 어텐션 네트워크를 거치며 ‘좋아’의 가중치가 가장 큰 값을 갖게 되어 문서가 ‘긍정’으로 분류된다.

그림 2.11: 어텐션 (Attention) 기반 양방향 장·단기 메모리 (BiLSTM) 구조

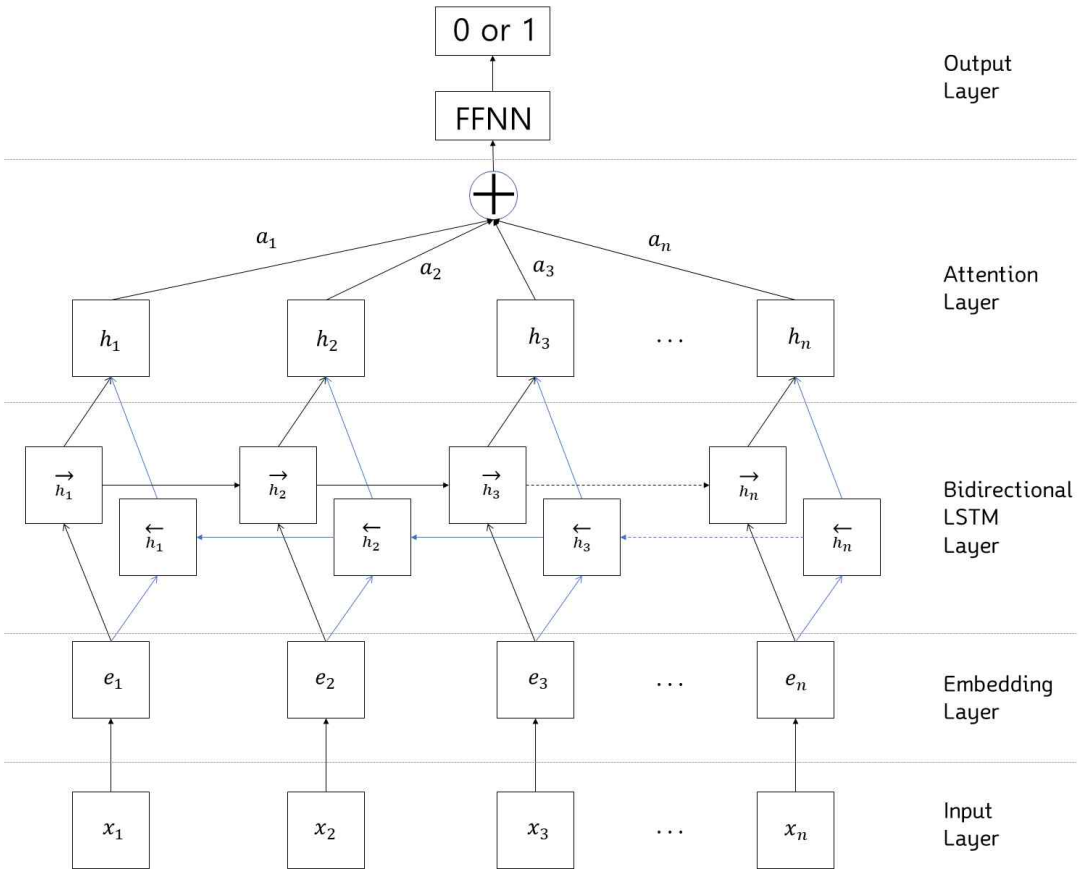


그림 2.11의 출력층에 있는 FFNN (Feed-Forward Neural Network)은 전결합층으로 앞서 2.2.5.2 다층 퍼셉트론 (MLP) 및 2.2.5.5 합성곱 신경망 (CNN) 부분에서 설명했듯이 문서 벡터 c를 긍정 또는 부정으로 분류하는 역할을 담당한다.

제 3 장 카카오톡 메시지 아티팩트 분류

3.1 카카오톡의 기본 기능

카카오톡은 다양한 콘텐츠를 송·수신할 수 있다. 세부 유형은 일반적인 메시지 (text), 이모티콘, 사진, 동영상, 지도, 음성메시지, 파일 등이다. 하지만 표 3.1.에서와 같이 채팅방에 따라 이용할 수 있는 기능에 차이가 있다. 일반채팅의 경우 모바일 버전에서는 카카오톡에서 제공하는 모든 기능을 활용할 수 있다. 그러나 일반채팅을 포함한 모든 채팅의 PC 버전의 경우 카메라 앱으로 실시간 사진이나 동영상을 촬영하여 메시지로 송신하는 것은 제한되며 선물하기, 송금, 영화예매, 음악, 지도 (주소), 음성메시지, 연락처 (프로필, 연락처 모두), 클라우드에 저장된 파일 송·수신이 제한되다. 비밀채팅의 경우 PC용 카카오톡에서는 해당 기능을 제공하지 않기 때문에 모바일 버전만 살펴보면 텍스트, 이모티콘, 저장된 사진 또는 실시간 촬영한 사진이나 동영상까지는 송·수신이 가능하지만 그 외의 기능은 이용할 수 없다. 오픈채팅의 경우 인가된 사용자만 입장 가능한 경우든 불특정 다수가 입장 가능한 경우든 모바일용 카카오톡에서 보이스톡, 페이스톡을 사용할 수 없고 송금, 영화예매 등의 금전거래가 제한된다. 플러스채팅은 모바일용 및 PC용 모두 이용할 수 있는 기능이 다른 채팅들에 비해 상대적으로 적다. 광고 목적의 사업체 계정이 다수를 차지하기 때문에 모바일용 카카오톡에서는 선물하기, 연락처 전송이 제한되고 모바일에 저장된 파일도 첨부할 수 없다. PC용 카카오톡에서도 모든 채팅방에서 공통적으로 제한되는 기능에 더하여 PC에 저장된 파일을 첨부할 수 없다.

표 3.1: 채팅 유형별 송·수신이 가능한 콘텐츠 유형

구분		일반채팅		비밀채팅		오픈채팅				플러스채팅	
						비밀번호 설정		비밀번호 미설정			
		모바일	PC	모바일	PC	모바일	PC	모바일	PC	모바일	PC
text		○	○	○	×	○	○	○	○	○	○
이모티콘		○	○	○	×	○	○	○	○	○	×
저장된 사진	일반	○	○	○	×	○	○	○	○	○	○
	고화질	○	○	○	×	○	○	○	○	○	○
	원본	○	○	○	×	○	○	○	○	○	○
카메라	사진	○	×	○	×	○	×	○	×	○	×
	동영상	○	×	○	×	○	×	○	×	○	×
	치즈 움짤	○	×	○	×	○	×	○	×	○	×
선물하기		○	×	×	×	○	×	○	×	×	×
무료 통화	보이스톡	○	○	×	×	×	×	×	×	×	×
	페이스톡	○	○	×	×	×	×	×	×	×	×
송금		○	×	×	×	×	×	×	×	×	×
뮤 직		○	×	×	×	○	×	○	×	×	×
지도 (주소)		○	×	×	×	○	×	○	×	×	×
캡쳐	프로필	○	○	×	×	○	○	○	○	○	○
	프로필 가리기	○	○	×	×	○	○	○	○	○	○
음성 메시지		○	×	×	×	○	×	○	×	○	×
연락처	프로필	○	×	×	×	○	×	○	×	○	×
	연락처	○	×	×	×	○	×	○	×	×	×
영화예매		○	×	×	×	×	×	×	×	×	×
파일 (클라우드)		○	×	×	×	○	×	○	×	×	×
파일 (기기저장)		○	○	×	×	○	○	○	○	×	×

3.2 분석환경

본 논문에서 사용한 장치와 카카오톡 버전은 표 3.2와 같다. 카카오톡 계정 사용자의 ID와 PW는 사전 획득이 가능하다고 가정하였다. 카카오톡을 통해 생성된 데이터들은 /data/com.kakao.talk/databases와 /data/com.kakao.talk/files 경로 아래 'KakaoTalk.db'와 'KakaoTalk2.db' 파일에 구분되어 저장되어 있다. 카카오톡에 대한 메시지 추출(이미징)과 분석은 각각 상용 포렌식 도구인 MAGNET AXIOM 버전 3.3.1.14874의 Process (이미징)와 Examine (분석)을 활용하였고 범용 포렌식 도구인 Autopsy와 SQLite browser를 활용하여 상호 검증하였다. 대상 데이터는 PC와 모바일에 카카오톡을 설치하고 '19. 7. 1.~'19. 9. 30. 3개월 동안 15명이 송·수신한 문자메시지, 동영상, 음성파일, 보이스톡 등 다양한 콘텐츠를 분석하였다.

표 3.2: 카카오톡 아티팩트 분석에 활용한 PC, 모바일 기기 및 카카오톡 앱 버전

	분석 기종	OS	카카오톡 버전
PC	Intel(R) Xeon(R) CPU E3-1230 v3	Window 10 Pro (64-bit)	3.0.4.2212
모바일	Samsung Galaxy Note5 (SM-N920)	안드로이드 6.0 마시멜로	8.5.4

3.3 분석결과

3.3.1 카카오톡 아티팩트 유형

카카오톡 아티팩트는 표 3.3과 같이 5가지 유형으로 나뉜다. 모바일용 카카오톡과 PC용 카카오톡 모두에서 확인 가능한 아티팩트는 채팅 방 (Chat Rooms), 메시지 (Messages), 친구 (Friends or Contacts) 이상 3개이며, 모바일용 카카오톡에서만 확인 가능한 아티팩트는 통화 (Calls)이고 PC용 카카오톡에서만 확인 가능한 아티팩트는 공유된 사진 (Shared Pictures)이다. 각각의 아티팩트는 다양한 세부 속성을 포함하며 모바일용 카카오톡과 PC용 카카오톡에 동일하게 분류된 아티팩트라 할지라도 세부 속성이 모두 일치하는 것이 아니기때문에 정보를 상호 보완하여 활용하면 카카오톡 사용자에 관한 신뢰도 높은 정보를 얻을 수 있다. 또한 아티팩트별 세부 속성 중 사용자 ID, 채팅 방 ID 등은 채팅 방, 메시지, 친구 등에 반복적으로 나타나고 있어 ID를 기준으로 아티팩트의 속성 정보를 분류하여 비교 분석이 가능하다.

표 3.3: 카카오톡 아티팩트 유형

아티팩트 유형	모바일	PC
카카오톡 채팅 방 (Chat Rooms)	○	○
카카오톡 메시지 (Messages)	○	○
카카오톡 친구 (Friends or Contacts)	○	○
카카오톡 통화 (Calls)	○	×
카카오톡 공유된 사진 (Shared Pictures)	×	○

3.3.2 카카오톡 채팅 방 (Chat Rooms)

카카오톡 채팅 방의 세부 속성 중에서 모바일용과 PC용 모두에서 확인 가능한 아티팩트는 표 3.4에서와 같이 채팅 방 ID (여러 개의 채팅 방을 구분하는 숫자), 채팅 유형, 마지막 메시지, 마지막 메시지의 송·수신 일시이다. 그중 채팅 유형은 표 3.5에서 진하게 표시된 Direct, Multi, OD, OM, Plus, SDirect, Memo로 구분된다. 모바일용 카카오톡에서만 추출되는 속성으로는 사용자 이외의 참가자 현황 (참가자 ID), 작성하였으나 발송하지 않은 메시지 등이 있다. PC용에서는 채팅 방 이름을 확인할 수 있으며 사용자 이외 참가자 수와 연결된 ID는 값이 표시되나 정확하지 않은 값으로 수사에 활용하기에 제한된다.

표 3.4: 카카오톡 채팅 룸 세부 속성

아티팩트 속성	모바일	PC
채팅 방 ID (Chat ID)	○	○
채팅 방 이름 (Room Name)	×	○
기타 참가자 (Other Participants)	○	×
참가자 수 (Number of participants)	×	×
채팅 유형 (Chat Type)	○	○
마지막 메시지 (Last Message)	○	○
마지막 메시지 일시 (Last Message Date/Time)	○	○
보내지 않은 메시지 (Unsent Message)	○	×
연결된 ID (Link ID)	×	○

표 3.5: 채팅 유형

채팅 유형 (Chat Type)		모바일	PC
일반 채팅	**Direct**Chat (1:1)	○	○
	MultiChat (1:N)	○	○
오픈 채팅	**OD** (OpenChat-Direct) (1:1)	○	○
	OM (OpenChat-Multi) (1:N)	○	○
플러스 채팅 (**Plus**Chat)		○	○
비밀 채팅 (**SDirect**Chat, SecretChat)		○	×
나와의 채팅 (**Memo**Chat)		○	○

3.3.3 카카오톡 메시지 (Messages)

카카오톡 메시지 아티팩트의 세부 속성 중에서 모바일용과 PC용 모두에서 확인 가능한 정보는 표 3.6에서 확인할 수 있듯이 같이 보낸 사람의 ID (사용자들을 구분하기 위한 숫자), 메시지 내용, 메시지 유형, 첨부파일 유무이고, 모바일용에서만 확인 가능한 정보는 채팅 방 ID, 보낸 사람 이름, 메시지 생성 일시, 메시지 방향 (송신 또는 수신), 메시지 삭제 일시였다. 위도와 경도는 첨부파일이 지도이면 표시가 된다. PC용에서만 확인 가능한 정보는 메시지의 ID (메시지들을 구분하기 위한 숫자), 메시지 송·수신 일시, 메시지 삭제 여부이다.

표 3.6: 카카오톡 메시지 세부 속성

아티팩트 속성	모바일	PC
보낸 사람 ID (Sender ID)	○	○
보낸 사람 이름 (Sender Name)	○	×
채팅 방 ID (Chat ID)	○	×
메시지 (Message)	○	○
메시지 ID (Message ID)	×	○
메시지 유형 (Message Type)	○	○
메시지 방향 (Message Direction)	○	×
메시지 생성 일시 (Created Date/Time)	○	×
메시지 발송 일시 (Message Sent Date/Time)	×	○
메시지 삭제 여부 (Deleted : Yes or No)	×	○
메시지 삭제 일시 (Deleted Date/Time)	○	×
첨부파일 (Attachment, Additional Information)	○	○
위도 (Latitude)	○	×
경도 (Longitude)	○	×

이 중 메시지 유형은 2016년에 카카오톡 모바일 버전 4.8.2. DB 파일을 SQLite browser로 확인하여 메시지 유형을 분석한 윤형철 등 [7]의 연구가 있어 본 논문의 버전 8.5.4.와 세부적으로 비교 분석하였다. PC용 카카오톡의 메시지 유형에 대한 기존 연구는 없어 참고 목적으로 함께 제시하였다. 표 3.7과 같이 채팅 방 초대 및 퇴장 관련 메시지 - “00님이 00님을 초대했습니다.”, “00님이 나갔습니다.” - 가 기존 윤형철 등의 연구에서는 단순히 ‘0’으로 표시되었다면 최근에는 ‘0’과 함께 feedType (1: invited, 2: left 등)이 표시되어 채팅 방의 상황을 이전보다 상세히 확인할 수 있다. PC용 카카오톡에서는 채팅 방 초대 및 퇴장 관련 메시지가 ‘Attachment’로 나타났다. 텍스트, 사진, 동영상, 연락처 등은 기존 연구와 동일한 숫자로 나타났으나 PC 로그인 메시지의 경우 기존 연구에서는 ‘9’로, 본 연구에서는 ‘71’로 확인되었다. PC용 카카오톡에서도 동일하게 ‘71’이 PC 로그인 메시지를 의미하였다. 또한 모바일 버전 4.8.2.와 비교 시 버전 8.5.4.에서 사용 가능한 기능이 상당히 증가했음을 알 수 있는데 삭제된 메시지 알림 - “삭제된 메시지입니다.” - 이 ‘0’과 함께 feedType 14 (hidden)로 분류되고 일정, 투표, 검색어, 답장, 사진 묶어보내기 등의 기능이 추가되었다. 알림 (notification) 기능은 광고, 선물 메시지, 은행업무 메시지, Link 등을 의미하는데 모바일용 카카오톡에서는 ’Nofitication’, PC용에서는 ‘9’로 분류되었다. 이밖에도 투표, 플러스 채팅 메시지 등 다양한 유형의 메시지들이 숫자 및 문자 (단어)로 구분되어 있음을 확인하였다.

3.3.4 카카오톡 친구 (Friends)

카카오톡 친구 (또는 연락처) 아티팩트의 세부 속성 중 모바일용과 PC용 모두에서 확인 가능한 정보는 카카오톡 사용자의 ID, 카카오톡 프로필 이름 (모바일용보다 PC용에서 더 많은 정보 확인이 가능), 카카오톡 프로필 상태 메시지, 별칭 (카카오톡 프로필 이름), 즐겨찾기 추가 여부 (특정 채팅 방에 대한 즐겨찾기 추가 여부), 숨김친구 여부 (yes or no), 핸드폰 번호 일부 (모바일용보다 PC용에서 더 많은 정보 확인이 가능), 프로필 사진 (이미지) URL 등으로 표 3.8에서 확인할 수 있다. 모바일용에서만 확인 가능한 정보는 사용자의 카카오톡 계정 ID, 연락처에 저장된 친구들의 이름 일부, 계정 생성일, 채팅 방 ID이며, PC용에서만 확인 가능한 정보는 계정 유형, 연락처와 연동된 친구들의 ID인데 계정 유형은 표 3.9와 같이 친구, 차단친구 등의 정보 획득이 가능하나 일부 정보들이 부정확하였다. 연결된 ID는 오픈채팅방에서 주고 받은 내용만 확인되었다.

표 3.7: 메시지 유형

메시지 유형 (Message Type)	윤종철 등 [7] (모바일 ver. 4.8.2.)	모바일 (ver. 8.5.4.)	PC (ver. 3.0.4.2212)
채팅방 초대, 퇴장 알림 (Notice of Invited / left someone to / from Chat)	0	0 (feed Type 1, 2, 4 : invited, left, invited to Open Chat)	Attachment
텍스트 (Text)	1	Chat	Message
사진 (Photo)	2	Photo	Photo
동영상 (Video)	3	Video	Video
연락처 (Contact)	4	4	4
보이스 노트 (Voice Note)	5	Voice Note	Audio
PC login message	9	71	71
정적인 Emoticon (Static)	0	6 (.gif)	6 (.gif)
	12	12 (.png)	
동적인 Emoticon (Animated)	·	20 (.webp)	Emoticon (.png, .webp, 'XCON')
	·	25 (.webp, 'XCON')	
일정 (Event)	16	13	Event (sub Type 1, 2, 3, 4 : Event, Changed, Deleted, Alarmed)
투표 (Poll)	17	14	Poll (subType 1, 2, 4 : Poll, Changed, Deleted)
'지도' 선택/전송 (Location)	18	Location (longitude, latitude, address)	16
카카오톡 ID (Profile)	51	Profile	·
메시지 삭제 알림 Notice of Hidden Message ("This message was deleted.")	·	0 (feed Type 14 : hidden)	Attachment
파일 (File) : pdf, xlsx, mp3 등	·	File Transfer	18
보이스 톡 (Voice Talk)	·	Call (IP4, IP6, Port, call ID, duration)	Call (IP4, IP6, Port, call ID, duration)
공지 (Board Post)	·	Board Post	Announcement
알림 (Notification)	·	Notification	9
#검색어 (#Searching)	·	23	23
답장 (Reply)	·	26	Reply
사진 묶어 보내기 (Grouped Picture)	·	27	27
라이브 톡 (Live Talk)	·	52	52
플러스 채팅 메시지 (Plus Chat from non Plus Chat Friends)	·	71	71
플러스 채팅 메시지 (Plus Chat from Plus Chat Friends)		72	72
삭제된 메시지 (Deleted Message with crypto key)	·	·	16385
삭제된 사진 (Deleted Picture)	·	16386	16386
삭제된 일정 (Deleted Event)	·	16397	16397
삭제된 투표 (Deleted Poll)	·	16398	16398
삭제된 메시지 (Deleted Message)		16402, 16408	16402, 16408

표 3.8: 카카오톡 친구 (연락처) 세부 속성

아티팩트 속성	모바일	PC
카카오톡 로그인 ID (User ID for log-in)	○	○
채팅 방 ID (Chat ID)	○	×
카카오톡 계정 ID (User ID)	○	○
카카오톡 프로필 이름 (Screen Name)	○	○
연락처에 저장된 사용자 이름 (Contact Name)	○	×
계정 유형 (Account Type)	×	○
상태 메시지 (Status Message)	○	○
생성 일자, 친구추가 일자 (Created Date/Time)	○	×
즐겨찾기 (Favorite)	○	○
별칭 (Nickname)	○	○
숨김 (Hidden: Yes or No)	○	○
전화번호 (Phone Number)	○	○
프로필 사진 URL (Profile Picture/Image URL)	○	○
연결된 ID (Link ID)	×	○

표 3.9: 계정 유형

계정 유형 (Account Type)	PC
Friend	친구
Blocked (Group Chat)	그룹 채팅 참여자 중 친구가 아닌 경우
Blocked	차단친구
Group Chat Member	오픈 채팅 참여자
Various Numbers	상호친구였던 2인이 일방 또는 상호 간 연락처 삭제/변경 등의 이유로 정상적인 친구추가(동기화) 등이 이루어지지 않은 상태

3.3.5 카카오톡 통화 (Calls)

카카오톡 사용자가 모바일용 카카오톡에서 보이스톡 및 페이스톡을 사용한 경우 카카오톡 통화 아티팩트를 통해 통화 상태 (시작, 종료 여부), 통화 지속시간, 통화 대상자 ID 및 이름, 해당 통화가 시도된 채팅 방 ID, 통화 유형, 통화 방향 (송신 또는 수신), 통화 생성 및 삭제 일시를 확인할 수 있다. PC에서도 일반채팅의 경우 보이스톡과 페이스톡 모두 가능하지만 관련된 아티팩트는 카카오톡 메시지 아티팩트의 메시지 유형 (3.2.2)에서만 IP, Port, 통화 대상자 ID, 통화 지속시간의 확인이 가능하다. 따라서 통화 관련된 정보는 카카오톡 메시지 아티팩트의 메시지 유형과 카카오톡 통화 아티팩트의 세부 속성을 함께 고려하여 획득하는 것이 상호 보완적이다.

3.3.6 카카오톡 공유된 사진 (Shared Pictures)

카카오톡 공유된 사진은 PC용 카카오톡에서만 확인이 가능한 아티팩트로 해당 메시지 ID, 채팅 방 ID, 보낸 사람 ID, 보낸 일시, 파일의 크기 (bytes), 썸네일 URL, 파일 확장자, MIME 유형, 다운로드 위치 및 다운로드 완료 일시에 대한 정보를 제공한다. 그러나 이 중 MIME 유형과 다운로드 완료 일시에 대한 정보가 일부 정확하지 않아 수사환경에 적용이 제한된다.

제 4 장 카카오톡 메시지 감성 분석

4.1 분석환경

인공 신경망 (Artificial Neural Network)을 구현하는 여러 가지 Neural Network 생성용 library 중 본 연구에서는 Tensorflow와 Keras를 사용하였다. 분석에 활용한 장비는 표 4.1.과 같고 데이터는 카카오톡 메시지 4,000건 중 training dataset과 validation dataset을 8:2 비율로 각각 3200건, 800건으로 나누어 활용하였다. 세부 절차는 기계학습 절차 (그림 2.3)에 따라 진행하였다.

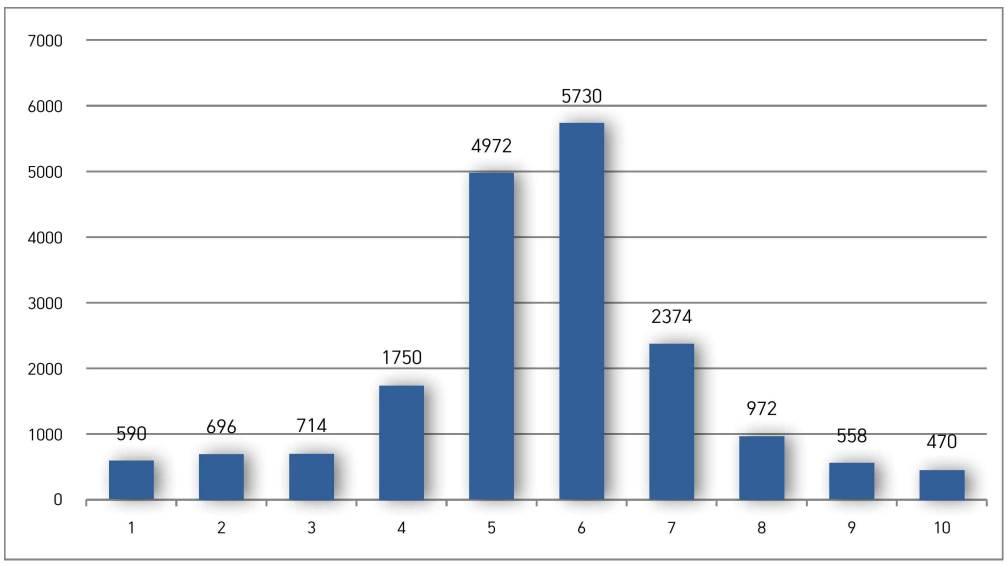
표 4.1: 카카오톡 메시지 감성 분석에 활용한 PC 환경

CPU	OS	RAM	GPU
Intel(R) Core(TM) i7-8700 CPU	Window 10 Enterprise	32.0GB	NVIDIA GeForce GTX 1060 6GB

4.2 데이터 수집

카카오톡 메시지를 감성 분석하려면 각각의 메시지는 긍정 또는 부정에 대한 정보를 포함하고 있어야 한다. 본 논문에서는 2019. 7. 1. ~ 2019. 9. 30.까지의 카카오톡 메시지 18,826건을 수집한 뒤 메시지 작성자 15명에게 메시지 작성 당시의 감정 상태를 수렴하여 1점~10점까지 (1에 가까우면 부정인 상태, 10에 가까우면 긍정인 상태) 점수화하였다. 1점~10점까지 점수화된 메시지 중 연구에 필요한 것은 긍정 또는 부정의 극성을 비교적 명확하게 띄는 메시지이다. 따라서 그림 4.1과 같이 4점~7점 구간은 긍정 또는 부정의 극성이 약하거나 정보전달 목적의 메시지들로 극성을 전혀 띄고 있지 않다고 판단하여 제외하고 1점~3점 구간을 부정의 상태인 '0', 8점~10점 구간을 긍정의 상태인 '1'로 재분류하였다. 이렇게 분류된 메시지는 총 4,000건으로 긍정 2,000건 부정 2,000건 동일한 비율이다.

그림 4.1: 카카오톡 메시지 감정 분포



4.3 데이터 분석

자연어를 컴퓨터가 이해하고, 효율적으로 처리하게 하기 위해서는 컴퓨터가 이해할 수 있도록 자연어를 적절히 변환할 필요가 있다. 본 논문에서 수집한 카카오톡 메시지 18,826건 중 약 80% (15,060건)가 제3장 카카오톡 아티팩트 분석에서 수집한 csv 파일이고, 나머지는 카카오톡 메시지 내보내기 기능을 통해 수집한 txt 파일이다. 이를 Tab으로 구분한 txt 파일 형태로 통일하였다. 또한 데이터는 user, date, contents, label 네 개의 열로 이루어져 있는데 user는 메시지 작성자, date는 메시지 작성 일시, contents는 메시지 내용, label은 긍정(1), 부정(0)인지를 나타낸다.

카카오톡은 일상생활 대화부터 업무, 동호회, 관심사별 오픈채팅 등 사용자에 따라 그 내용과 유형이 다양하고 외부에 유출될 경우 사생활이 침해받을 수 있는 내용도 다수 있기 때문에 특정 개인·집단을 직·간접적으로 지칭 및 유추 가능할 수 있는 내용에 대한 정제가 필요하다. facebook 메신저, 인스타그램 등 여느 SNS와 같이 이모티콘, 띄어쓰기, ‘ㅋㅋ’ 등 분리된 자음·모음 등에 대한 선별적인 정제도 요구된다.

4.4 데이터 전처리

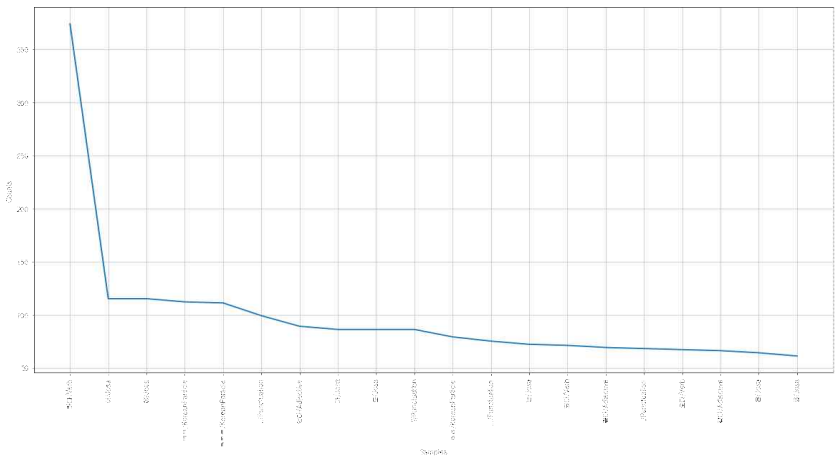
4.4.1 형태소 분석 및 품사 태깅

데이터는 KoNLPy 라이브러리의 okt 클래스를 이용하여 형태소 분석 및 품사 태깅을 하였다 [5]. 카카오톡 메시지는 4.3에서 설명한대로 특성상 맞춤법이나 띄어쓰기 등이 제대로 되어있지 않은 경우가 많기 때문에 정확한 분류를 위하여 띄어쓰기, 맞춤법 등을 정제해주는 KoNLPy를 이용하였다. KoNLPy 라이브러리를 통해 분석된 형태소 하나하나를 데이터의 토큰 (token)이라 한다.

4.4.2 중복토큰 분석

분석한 데이터의 토큰 중 일부는 중복된 경우가 있을 수 있으므로 nltk 라이브러리를 통해서 전처리를 하였다. 가령 카카오톡 메시지 4,000건을 okt로 형태소를 분석해보면 전체 토큰의 개수는 24,496개이지만 중복을 제외한 토큰의 개수는 4,910건이다. 그림 4.2에서 볼 수 있듯 출현 빈도가 높은 상위 토큰 20개는 주로 ‘이’, ‘에’ 같은 조사 또는 ‘ㅋ’ 였다.

그림 4.2: 출현 빈도가 높은 상위 토큰 20개



4.4.3 데이터 벡터화

데이터를 벡터화하는 방법에 따라서 자연어 처리의 성능이 크게 달라진다. 앞서 분석된 토큰을 원-핫 인코딩 (One-hot encoding) 같은 희소 벡터 (sparse vector)로 표현하는 것보다 사용자가 설정한 값으로 벡터 표현의 차원을 맞추는 밀집 벡터 (dense vector)로 표현하는 것이 적어도 처리 시간을 단축시켜준다. 단어 (토큰)을 밀집 벡터로 표현하는 것을 단어 임베딩 (word embedding)이라고 하며 이 밀집 벡터를 임베딩 벡터 (embedding vector)라고 한다. 단어 임베딩을 위해 우선 단어 (토큰)를 고유한 정수에 맵핑 (mapping)시키는 과정인 정수 인코딩 (encoding)을 하여 분석된 토큰을 단어 빈도수가 높은 순으로 낮은 정수 인덱스를 부여한 뒤 이를 카카오톡 메시지 데이터셋에 적용하였다. 정수를 밀집 벡터 (임베딩 벡터)로 맵핑하는 과정인 단어 임베딩은 케라스 임베딩 층을 활용하는 경우와 Word2Vec을 활용하는 경우 2가지로 구분하여 인공 신경망을 학습시켰다.

4.5 모델링 및 학습

본 논문에서는 4.1.3.3에서 언급한 바와 같이 카카오톡 메시지 데이터셋을 Word2Vec을 활용하여 임베딩했는지 여부에 따라 MLP, RNN, LSTM, CNN 등 다양한 알고리즘을 개별 또는 결합하여 학습시켰다. 그런데 위 두 유형의 학습 모델들의 정확도가 70~80% 수준에 머무는 등 답보상태인 것을 확인하고 추가적인 성능 개선을 기대하며 어텐션 기반 알고리즘의 학습도 함께 진행하였다. 따라서 학습에 사용한 모델은 크게 3가지 유형으로 분류한 18개 모델이다. 모델에 따라 은닉층 활성화 함수는 'ReLU (Rectifier Function)' 또는 'tanh' (하이퍼볼릭탄젠트)를 사용하였으며 Adam 옵티마이저 (optimizer)를 통하여 오차를 최소화하였다. 손실함수 (loss function)는 binary_crossentropy이며 출력층 (output layer)에는 활성화 함수로 이진 분류에 적합한 sigmoid 함수를 사용하였다.

표 4.2: 학습에 사용된 모델 분류

기 본	Word2Vec 적용	어텐션 적용
MLP	MLP	MLP
RNN, BiRNN	LSTM	LSTM
LSTM, BiLSTM	BiLSTM	BiLSTM
CNN	CNN	CNN
CNN + LSTM	CNN + LSTM	CNN + LSTM
CNN + BiLSTM		

4.5.1 기본 모델

카카오톡 메시지 데이터셋의 벡터를 입력값 (input)으로 받아 2개의 은닉층에서 학습하는 MLP 모델은 그림 2.5와 같이 2개 은닉층 모두 전연결층으로 연결되었으며 활성화 함수로 Relu를 사용하였다. 또한 RNN, BiRNN, LSTM, BiLSTM, CNN, CNN과 LSTM, CNN과 (Bi)LSTM을 함께 모델링한 CNN+LSTM, CNN+BiLSTM 알고리즘을 활용해 각각 데이터셋을 학습시켰다. 자체 구축한 카카오톡 메시지 데이터셋이 케라스 Embedding()으로 감성 분석에 충분히 특화된 임베딩 벡터를 만들어내는지 여부와 다양한 모델 중 과연 어떤 모델이 감성 분석에 적합한 성능을 보이는지 확인하기 위함이다.

4.5.2 Word2Vec을 적용한 모델

한글 위키백과 300만 단어를 사전 학습한 Word2Vec을 활용하여 카카오톡 메시지 데이터셋을 벡터화한 뒤 이를 입력값으로 가지는 MLP, LSTM, BiLSTM, CNN, CNN+LSTM (4.5.1의 9개 모델 중 성능이 우수하고 구조가 특이한 5개 모델 선별)으로 학습시켰다. 통상 자연어 처리에서 훈련 데이터가 적은 경우 케라스 Embedding()으로 해당 문제에 충분히 특화된 임베딩 벡터를 만들어내는 것이 쉽지 않기 때문에 Word2Vec, FastText, GloVe 등 사전 학습된 임베딩 벡터 (Pre-Trained Word Embedding)를 사용하여 성능 개선을 기대한다. 카카오톡 메시지 감성 분석의 경우에도 사전 학습된 임베딩 벡터의 활용이 성능 개선을 가져오는지 여부를 확인하고자 Word2Vec을 사용하였다.

4.5.3 어텐션 기법을 적용한 모델

Word2Vec의 적용이 입력층에서 입력값의 벡터화 방법에 변화를 주어 다양한 인공 신경망 알고리즘의 성능을 측정하고자 하는 목적이었다면, 어텐션 기법의 적용은 은닉층 마지막 단계에서 출력의 신뢰도를 높이고자 입력 단어들을 재차 살펴보고 가중치를 조정하는 과정인 어텐션 네트워크가 과연 데이터셋 규모가 작은 카카오톡 메시지 데이터셋 학습에도 유의미한 영향을 미치는지 여부를 확인하기 위함이었다. 본 논문에서는 학습 마지막 단계에서 어텐션 네트워크를 추가하여 MLP, LSTM, BiLSTM, CNN, CNN+LSTM (4.5.1의 9개 모델 중 성능이 우수하고 구조가 특이한 5개 모델 선별)에 대한 성능을 확인하였다.

4.6 성능 평가

4.6.1 성능 평가 지표

긍정, 부정 이진 분류 (binary classifier system)에 대한 성능 평가를 위하여 혼동행렬 (Confusion matrix)을 사용하였다. 혼동행렬에서 평가 지표인 정확도 (Accuracy, ACC), 정밀도 (Precision), 민감도 (Sensitivity), 특이도 (Specificity)를 계산하여 성능을 확인했으며, ROC (Receiver Operating Characteristic) 그래프로 민감도와 특이도의 상관관계를 확인하였다.

표 4.3: 혼동 행렬 (confusion matrix)

지표 (indicator)		**예측값 (Predicted)**	
		긍정 (Positive)	부정 (Negative)
실제값 (Orga-nized)	긍정 (Positive)	TP (True Positive)	FN (False Negative)
	부정 (Negative)	FP (False Positive)	TN (True Negative)

정확도 (Accuracy)는 전체 데이터 중 실제 긍정 및 부정의 감정을 담은 메시지를 제대로 분류해낸 비율을 말하며 수식은 아래와 같다.

$$\text{정확도 (Accuracy)} = \frac{TP + TN}{TP + FP + FN + TN}$$

정밀도 (Precisoin)는 긍정으로 예측한 메시지 중 실제 긍정인 메시지의 비율을 뜻한다. 수식은 아래와 같다.

$$\text{정밀도 (Precision)} = \frac{TP}{TP+FP}$$

민감도 (Sensitivity, Recall, True Positive Rate)는 재현율 (Recall)이라고도 하는데 실제 긍정인 메시지 중에서 제대로 긍정적이라고 분류된 메시지의 비율을 말한다. 수식은 아래와 같다.

$$\text{민감도 (Sensitivity)} = \frac{TP}{TP+FN}$$

특이도 (Specificity)는 실제 부정인 메시지 중에서 제대로 부정적이라고 분류된 메시지의 비율이며 수식은 아래와 같다.

$$\text{특이도 (Specificity)} = \frac{TN}{FP+TN}$$

ROC 그래프는 가로축을 1-특이도, 세로축을 민감도로 하여 시각화한 그래프이다. ROC 그래프의 성능이 좋다는 것은 민감도와 특이도 모두 1에 가깝다는 것을 의미하므로 x=0, y=1인 경우 모델이 최적의 성능을 가진다. ROC 그래프에서 AUC (area under the curve)는 전체적인 민감도와 특이도의 상관관계를 보여줄 수 있어 매우 편리한 성능 측정 기준이다. 통상 의료현장에서 '암'을 판정하는 등의 이진 분류에서는 일단 '암'을 판정 (Positive)하는 것만 중요하다고 생각할 수 있지만 '암'이 아닌 환자를 아니라고 판별 (Negative)해내는 것도 중요하다. 본 논문에서도 카카오톡 메시지 감성 분석 시 긍정뿐만 아니라 부정의 감정을 담은 메시지도 잘 판별해내는 것이 중요하다. 수사 측면에서 본다면 되려 부정적인 메시지를 잘 분류하는 것이 긍정적인 메시지 분류 보다도 중요할 수 있다. 따라서 모델의 전반적인 성능을 평가하기 위해서는 위 지표들을 복합적으로 고려해야 한다.

표 4.4: 학습 모델 성능 평가 결과

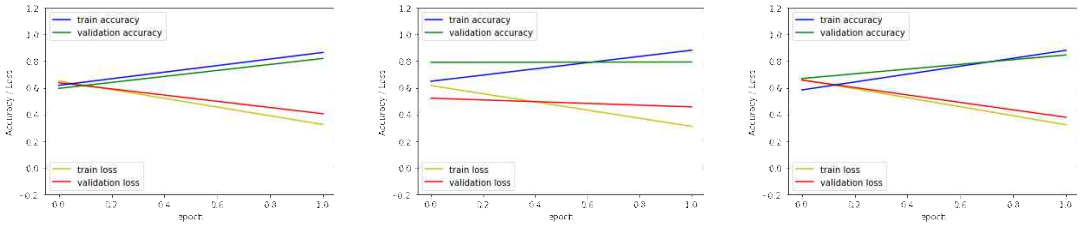
Model	Accuracy	Loss	Precision	Sensitivity	Specificity
MLP	0.8818	0.2495	0.7500	0.8212	0.7263
RNN	0.8227	0.5083	0.6667	0.5888	0.7056
BiRNN	0.8455	0.3816	0.7000	0.8423	0.6390
LSTM	0.8682	0.2735	0.7475	0.7170	0.7578
BiLSTM	0.8466	0.3035	0.7141	0.8541	0.6580
CNN	0.8318	0.2909	0.6814	0.7821	0.6343
CNN + LSTM	0.8727	0.2524	0.7154	0.8090	0.6782
CNN + BiLSTM	0.8313	0.3366	0.7386	0.6959	0.7537
Word2Vec + MLP	0.6864	0.5911	0.6151	0.5876	0.6323
Word2Vec + LSTM	0.7136	0.5227	0.6190	0.6613	0.5930
Word2Vec + BiLSTM	0.7227	0.5237	0.6317	0.6560	0.6175
Word2Vec + CNN	0.7134	0.4958	0.6191	0.6614	0.5929
LSTM + **Attention**	0.8842	0.7149	0.9982	0.7697	0.9987
BiLSTM + **Attention**	0.9256	0.3052	0.9987	0.8523	0.9989
CNN + **Attention**	0.8754	0.4703	0.9987	0.7517	0.9991

4.6.2 기본 모델

표 4.4와 같이 MLP의 정확도가 0.8818로 가장 높았으며 CNN+LSTM, LSTM이 0.85 이상의 정확도를 보였다. RNN은 양방향 (Bidirectional) 모델인 BiRNN의 정확도가 더 높았던 반면 LSTM은 BiLSTM 보다 정확도가 더 높아 양방향 모델이 항상 단방향 모델보다 우수하다고 단정할 수 없음을 보였다. 단, RNN 모델의 장기 의존성 문제(the problem of Long-Term Dependencies)를 보완한 LSTM 모델이 단방향, 양방향 모두 상대적으로 높은 성능을 나타내었다. 정확도가 우수한 MLP, CNN+LSTM, LSTM 모델 (그림 4.3)의 긍정, 부정 메시지 분류 성능을 세부적으로 살펴보면 민감도 (Sensitivity)는 MLP, CNN+LSTM, LSTM 순이며 특이도 (Specificity)는 LSTM, MLP, CNN+LSTM 순이었다. 이는 메시지 감성 분석에 있어 긍정 또는 부정 메시지 분류에 좀더 우수한 성능을 보이는 모델을 상호 보완하여 활용하는 것이 신뢰도 높은 정보 획득에 도움이 된다는 것을 알 수 있다.

그림 4.3: 기본 모델의 학습 정확도 및 손실률

(a) MLP (b) LSTM (C) CNN+LSTM

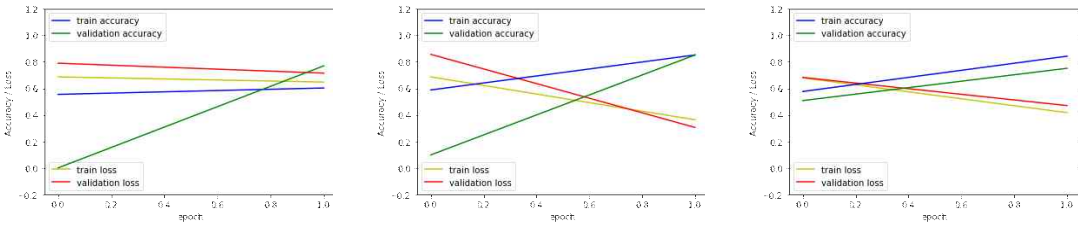


4.6.3 Word2Vec을 적용한 모델

Word2Vec을 사용한 모델은 벡터화 방법이 단어의 유사도를 나타내기 때문에 결과에 긍정적인 영향을 줄 것이라고 판단하였으나 예상 외로 Word2Vec을 사용한 모델은 표 4.4와 같이 낮은 성능을 보였다. 특히 CNN+LSTM은 손실률이 1을 넘는 등 유의미한 결과가 나오지 않아 표에서 제외하였다. 위키백과를 사전 학습한 Word2Vec을 사용하였기 때문에 구어체인 카카오톡 메시지의 의미를 제대로 짚어내지 못했을 가능성과 긍정, 부정 등 감정을 나타내는 단어 유형이 카카오톡과 많은 부분 일치하지 않았을 가능성이 높다. 카카오톡은 실시간 대화가 가능한 채팅 앱이기 때문에 트위터, 인스타그램, 블로그 등과 달리 그 자체의 독특한 언어 습관이 있다. 따라서 카카오톡 메시지 감성 분석에는 카카오톡 메시지로 사전학습한 Word2Vec을 사용하는 것이 성능 개선에 주효할 것으로 판단된다.

그림 4.4: Word2Vec을 적용한 모델의 학습 정확도 및 손실률

(a) LSTM (b) BiLSTM (C) CNN

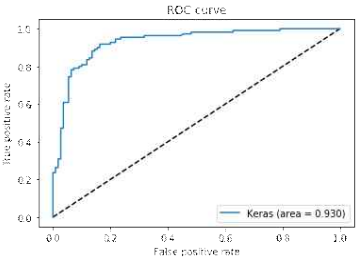


4.6.4 어텐션 기법을 적용한 모델

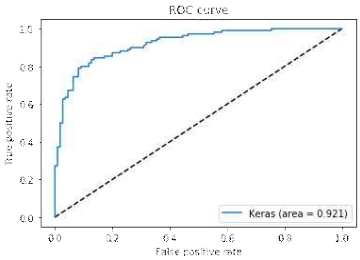
어텐션을 적용한 모델들은 특이도가 1에 가깝다. 즉, 부정적인 감정이 담긴 메시지에 대한 분류를 거의 완벽하게 한다는 의미이다. 긍정 메시지도 분류하는 과정에서 일부 누락은 있으나 일단 긍정 메시지로 분류하면 실제로 긍정 메시지일 확률이 매우 높다. 카카오톡 감성 분석은 긍정, 부정 모두 분류를 잘하는 모델이 이상적이므로 각각의 성능 평가 지표를 제대로 해석하는 것이 중요하다. 따라서 ROC 커브에서 AUC의 면적을 비교하였다. 그림 4.5와 같이 (a) BiLSTM이 0.930, (b) CNN이 0.921, (c) LSTM이 0.804 순으로 BiLSTM에 어텐션을 적용한 기법의 성능이 가장 우수함을 알 수 있다.

그림 4.5: 어텐션을 적용한 모델

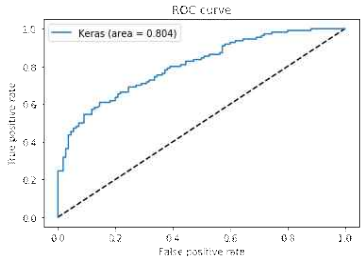
(a) BiLSTM



(b) CNN



(c) LSTM



MLP와 CNN+LSTM은 구현 과정에서 오류가 발생하거나 평가 지표가 제대로 도출되지 않아 최종 결과에서 제외하였다. 어텐션과의 결합에서 발생한 문제인지 자체 알고리즘 (코딩) 문제인지 추후 확인할 예정이다.

4.6.5 평가

본 논문에서 사용한 카카오톡 메시지 데이터셋은 한 문장이 평균 30개의 음절로 이루어져 있으며 채팅 상대와 실시간 주고 받는 온라인상 대화이므로 이모티콘의 사용, 분절된 자음 및 모음의 사용 등의 대표적인 특징을 배제하면 오프라인상 대화와 거의 유사한 수준이다. 이는 트위터나 인스타그램 등 여느 SNS와도 차별되는 형태이다. 따라서 웹 크롤링 등을 통해 위키백과나 트위터, 인스타그램 메시지들을 사전 학습한 Word2Vec 대신 다량의 카카오톡 메시지 데이터셋을 임베딩한 자체 Word2Vec 구축이 유의미한 성능을 얻는데 도움이 될 것으로 기대된다. 또한 어텐션 기법의 적용이 90% 이상의 높은 성능을 도출하는데 역할을 하였으므로 2018년 Google에서 제시한 BERT와 같이 어텐션 네트워크를 2개 이상 연달아 사용한 모델들에 대한 설계 및 성능 확인도 필요할 것으로 판단된다.

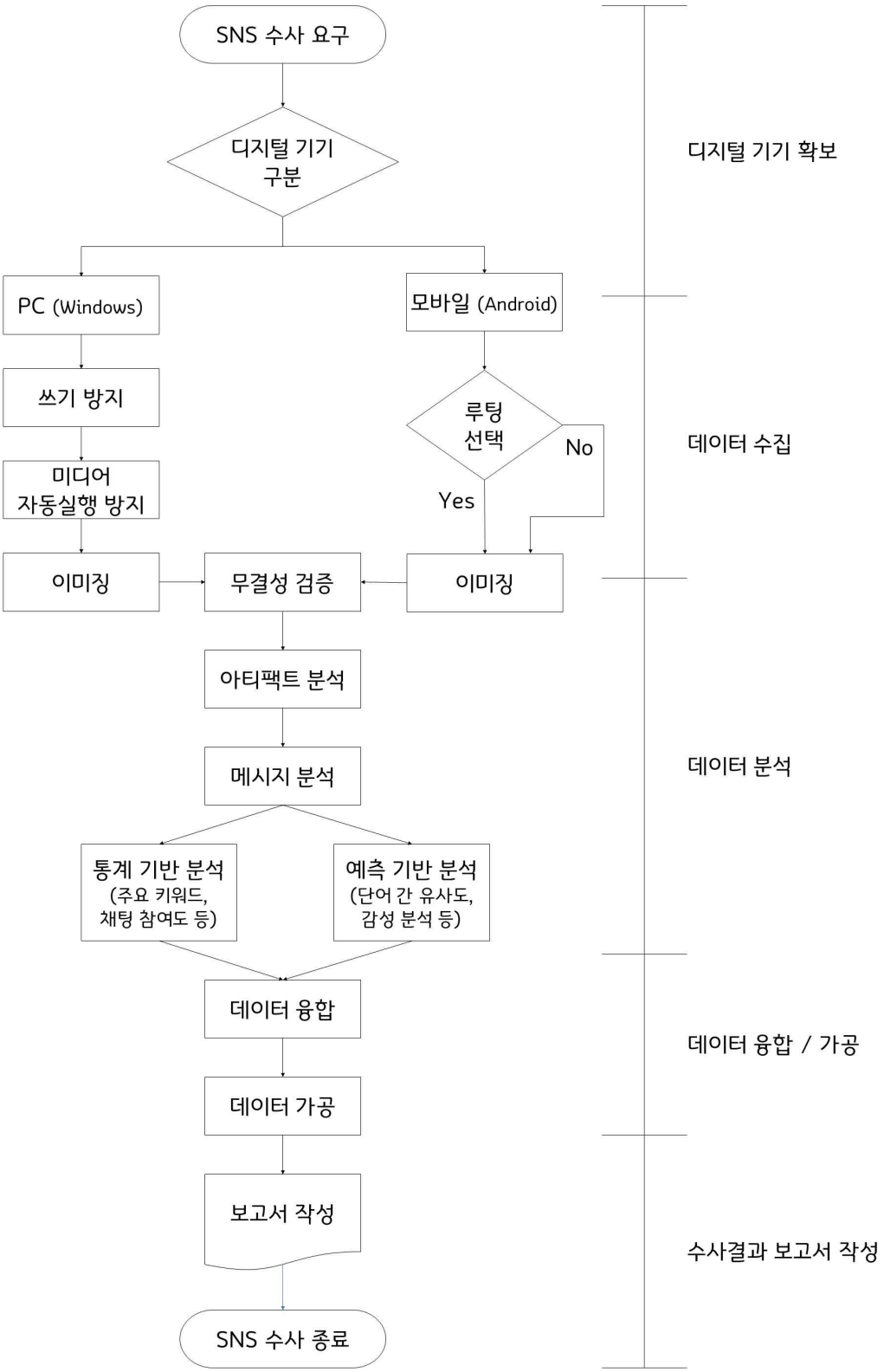
제 5 장 효과적인 채팅 앱 포렌식 절차 제안

실제 수사환경에서는 사고발생 접수를 받는 순간부터 수사종결이 결정되는 시점까지 수사관 개인별·팀별 철저한 임무분장이 이루어진 상태에서 다양한 수사행위를 동시다발적으로 진행한다. 채팅 앱에 대한 분석은 채팅 사용자의 정보를 확인하는 과정에서 그 필요성이 요구되지만 사용자에 대한 정보확인이 항상 전제되는 것은 아니다. 사고현장에 있던 PC, 모바일 등 디지털 기기에서 특정 채팅 앱에 대한 정보를 획득하고 이를 이용하여 채팅 앱 사용자에 대한 정보를 획득하는 등 다양한 경우의 수가 존재한다. 또는 사고현장에 있던 디지털 기기의 전원이 꺼져 있어 별도의 행정적인 절차 (통산자료확인제공요청 등)를 통해 통신사, 채팅 앱 운영사 등으로부터 사용자의 등록 정보를 획득한 뒤 채팅 앱을 선별하여 수사할 수도 있다. 따라서 이 장에서 제시하는 채팅 앱 포렌식 절차는 수사관이 사고발생 접수를 받고 현장에 출동하여 현장에 있던 PC, 모바일 등 디지털 기기로부터 휘발성 데이터를 획득한 뒤 분석을 위한 별도의 공간으로 디지털 기기를 이동시킨 이후의 포렌식 절차를 설명하였다.

채팅 앱 포렌식 절차는 그름 5.1과 같이 디지털 기기 확보, 데이터 수집, 데이터 분석, 데이터 융합 및 가공, 수사결과 보고서 작성 등 5단계로 구분된다. 디지털 기기 확보 단계에서는 사고 현장에서 수집한 디지털 기기를 분석을 위한 별도의 공간으로 이동시킨 뒤 무결성을 검증하고 기기의 OS를 확인한다. 데이터 수집 단계에서는 디지털 증거의 이미징을 위한 무결성 확보를 위해 PC의 경우 쓰기방지장치를 사용하거나 쓰기방지장치의 사용이 제한되는 경우 자동마운트 기능을 해제한다. 이동식 저장장치의 경우는 자동마운트 관련 명령이 적용되지 않으므로 레지스트리 키를 변경하여 쓰기방지기능을 설정할 수 있다 [23]. 그러나 요즘은 다양한 저장 매체가 존재하므로 모든 미디어 및 장치의 자동실행을 방지하기 위해 '모든 미디어 및 장치에 자동 실행 사용' 기능을 해제하여야 한다. 모바일의 경우 루팅(rooting)을 하여야 물리적 이미징이 가능하지만 잠금 기능의 해지 가능 여부에 따라 루팅이 제한되어 논리적 이미징만 실시하는 경우도 있다. 각각의 기기의 디스크 이미징을 통해 디지털 증거를 확보한 뒤에는 다양한 포렌식 도구를 활용하여 채팅 앱의 아티팩트와 메시지를 분석한다. 아티팩트 분석 시 PC 및 모바일에서 획득 가능한 정보들의 공통점과 차이점을 확인하고 상호 보완하여 수사에 활용하여야 한다. 채팅 앱 메시지 분석을 위한 사용자의 실제 메시지는 아티팩트 분석 과정과 채팅 앱 운영사로부터 획득이 가능하며 단순한 통계 기반 분석을 통해 주요 키워드 분석, 채팅 앱 사용자의 채팅 방별 채팅 참여도 등을 확인할 수 있고 본 논문에서 언급한 예측 기반 기계학습을 활용하면 주요 키워드들 간의 의미의 유사도, 메시지에 담긴 사용자의 감정을 예측해 볼 수도 있다. 데이터 융합은 사고발생 접수부터 데이터 분석에 이르기까지 수사의 모든 단계에서 지속적으로 이루어져야 하며, 데이터 분석 단계를 거쳐 최종적으로 수사 관련 데이터를 종합한 뒤 법정에서 증거로 활용될 수 있도록 논리를 갖춘 정보로 가공 후 문서화한다.

모든 포렌식 과정에서 반드시 고려해야 할 점은 우선 수집된 증거가 위변조 되지 않았음 (무결성, integrity)이 증명되어야 하고, 동일한 환경과 조건 하에서는 동일한 분석 결과가 도출되어야 하며, 증거물 수집, 이동, 분석 등 각 단계에서 담당자 및 책임자를 명확히 하고 인수인계 과정을 철저히 관리하는 등의 노력 (연계보관성, chain of custody)도 수반되어야 한다. 또한 증거를 수집할 때는 적법한 절차에 의거하여야 추후 법적으로 증거능력이 부여되며, 휘발성 데이터 등은 수사관이 지체없이 신속하게 획득하여야 한다.

그림 5.1: 채팅 앱 포렌식 절차



제 6 장 맺음말

본 논문에서는 안드로이드 OS용 (모바일) 카카오톡과 윈도우즈 OS용 (PC) 카카오톡에서 각각 아티팩트를 추출하여 5가지 유형으로 구분한 뒤 유형별 세부 속성을 비교 분석하였다. 모바일용 카카오톡과 PC용 카카오톡에서 획득 가능한 아티팩트 속성 정보는 총 54개 (중복 제외: 채팅 방 ID)로 이 중 18개 속성 (33.3%)이 모바일용 카카오톡과 PC용 카카오톡에서 공통적으로 추출되었고, 36개 속성 (66.7%)이 모바일용 카카오톡 (21개) 또는 PC용 카카오톡 (15개) 한쪽에서만 추출되었다. 이는 모바일용 카카오톡과 PC용 카카오톡의 아티팩트 유형과 그 속성에 대한 명확한 인식이 없는 상태에서 둘 중 한쪽에 대한 분석만 진행한다면 모바일용 카카오톡과 PC용 카카오톡 모두를 분석한 경우 대비 약 35% 상당의 정보 누락을 감수하여야 함을 의미한다. 따라서 모바일용 카카오톡과 PC용 카카오톡의 아티팩트 속성을 제대로 알면 공통적으로 추출되는 부분은 신뢰도 높은 정보로 활용하고 각각 추출되는 속성들은 상호 보완하여 수사환경에 효과적으로 활용할 수 있다,

위 아티팩트 추출 및 분류 과정을 통해 획득한 카카오톡 메시지는 기존 통계 기반 키워드 추출 등 단순 분석에서 벗어나 메시지에 담긴 사용자의 사상이나 감정을 분석하기 위하여 심층학습 기반 자연어 처리 기법을 통해 감성 분석을 실시하였다. 단어 임베딩 기법인 Word2Vec을 사용한 모델, 어텐션 기법을 은닉층에 적용한 모델 등 총 18개의 인공 신경망 모델을 학습시키고 성능 평가를 통해 카카오톡 메시지 감성 분석에 적합한 모델이 어텐션 기법을 적용한 BiLSTM (정확도 92/%, AUC 0.93)임을 확인하였다. 사용자의 동의를 얻어 각각의 메시지를 수집하고 메시지에 담긴 감정을 점수화하는 등의 과정에 많은 시간과 노력이 소요되었지만 이는 실제 수사상황이라면 사건 관계자들의 메시지 수집이 합법적으로 가능하므로 일부 해소될 수 있는 부분이며, 이후 상당한 데이터가 축적되어 학습 모델의 성능이 상향 평준화되었을 때 지도 학습이 아닌 비지도 학습으로도 메시지에서 수사에 유의미한 정보를 찾을 수 있을 것으로 기대한다.

마지막으로 본 논문에서는 2장 관련 연구에서 살펴본 국내·외 디지털 포렌식 및 모바일 포렌식 절차 표준과 3장 카카오톡 메시지 아티팩트 분류, 4장 카카오톡 메시지 감성 분석을 통하여 효과적인 채팅 앱 포렌식 절차를 제안하였다. PC, 모바일 각각의 디지털 기기로부터 수집된 데이터를 분석하고 융합을 통한 활용까지의 과정을 도식화하였다. 최근에는 획득 가능한 데이터의 양이 방대해지면서 오히려 수사관의 능력과 재량에 따라 활용되는 정보의 수준과 양에 격차가 발생하고 있는 만큼 적어도 본 논문에서 제안한 포렌식 절차가 데이터를 확보 및 분석하고 효과적으로 융합하여 적재적소에 활용하게 하는 데 일조하길 바란다.

향후에는 카카오톡 메시지 감성 분석의 신뢰도 향상을 위하여 최소 1만건 이상의 카카오톡 메시지를 사전 학습한 Word2Vec을 구축할 예정이다. 카카오톡 메시지는 메시지 한 개가 평균 30개의 음절로 이루어진 짧은 문장이며 채팅 상대와 실시간 주고받는 온라인상 대화이기 때문에 이모티콘의 사용, 분절된 자음 및 모음의 사용 등의 대표적인 특징을 배제하면 오프라인상 대화와 거의 유사한 수준이라 할 수 있다. 이는 트위터나 인스타그램 등 여느 SNS와도 차별되는 형태이다. 따라서 카카오톡 메시지 기반 Word2Vec 구축은 카카오톡 메시지 감성 분석에 유의미한 성능 개선을 가져올 것으로 판단된다.

참 고 문 헌

[1] https://www.statista.com/statistics/984645/south-korea-kakaotalk-usage-by-age/. [Online; accessed 20-Aug.-2019].

[2] https://gs.statcounter.com/os-market-share/all/south-korea/. [Online; accessed 20-Aug.-2019].

[3] https://www.statista.com/statistics/258749/most-. [Online; accessed 20-Aug.-2019].

[4] 김유영, 송민. "영화 리뷰 감성분석을 위한 텍스트 마이닝 기반 감성 분류기 구축." 지능정보연구 22.3 (2016): 71-89.

[5] Choi, Jusop, et al. "Digital forensic analysis of encrypted database files in instant messaging applications on 윈도우즈 operating systems: Case study with KakaoTalk, NateOn and QQ messenger." Digital Investigation 28 (2019): S50-S59.

[6] Azfar, Abdullah, Kim-Kwang Raymond Choo, and Lin Liu. "An 안드로이드 communication app forensic taxonomy." Journal of forensic sciences 61.5 (2016): 1337-1350.

[7] 윤종철, 박용석. "KakaoTalk의 채팅 메시지 포렌식 분석 연구 및 WhatsApp의 Artifacts 와의 비교 분석." 한국정보통신학회논문지 20.4 (2016): 777-785.

[8] Park, Kyungwon, and Hyoungshick Kim. "Encryption Is Not Enough: Inferring user activities on KakaoTalk with traffic analysis." International Workshop on Information Security Applications. Springer, Cham, 2015.

[9] 임좌상, 김진만. "한국어 트위터의 감정 분류를 위한 기계 학습의 실증적 비교." 멀티미디어학회논문지 17.2 (2014): 232-239.

[10] 김세진, 등. "인터넷 용어의 감성 분석을 통한 동영상 광고 효과 분석 시스템 설계." 정보과학회논문지 46.9 (2019): 919-925.

[11] 박천음, 등. "문맥 표현과 셀프 어텐션을 이용한 한국어 영화평 감성 분석." 정보과학회논문지 46.9 (2019): 901-908.

[12] "Information Technology-Security Techniques-Investigation principles and processes", international standard, 2015.

[13] “Guidelines on Mobile Device Forensics”, NIST SP 800-101 Revision1, 2014.

[14] “이동 전화 포렌식 가이드라인”, 한국정보통신기술협회 TAA, 2007.

[14] Jang, Yu-Jong, and Jin Kwak. "Digital forensics investigation methodology applicable for social network services." Multimedia Tools and Applications 74.14 (2015): 5029-5040.

[15] Yusoff, Mohd Najwadi, et al. "Advances of mobile forensic procedures in Firefox OS." International Journal of Cyber-Security and Digital Forensics (IJCSDF) 3.4 (2014): 183-199.

[16] http://www.yoonsupchoi.com/2017/08/08/ai-medicine-4/.
[Online; accessed 11-Sep.-2019].

[17] Yin, Wenpeng, et al. "Comparative study of CNN and RNN for natural language processing." arXiv preprint arXiv:1702.01923 (2017).

[18] Mikolov, Tomas, et al. "word2vec." URL https://code. google. com/p/word2vec (2013).

[19] Hochreiter, Sepp, and Jürgen Schmidhuber. "Long short-term memory." Neural computation 9.8 (1997): 1735-1780.

[20] Olah, Christopher. "Understanding lstm networks." (2015).

[21] Park, Hyun-jung, Min-chae Song, and Kyung-shik Shin. "Sentiment Analysis of Korean Reviews Using CNN: Focusing on Morpheme Embedding." Journal of Intelligence and Information Systems 24.2 (2018): 59-83.

[22] https://www.researchgate.net/figure/Architecture-of-our-unsupervised-CNN-Network-contains-three-stages-each-of-which_283433254. [Online; accessed 11-Sep.-2019].

[23] 디저털 포렌식 (Digital forensics), 노명선, 고시계사, 2018.

사 사

2018년 3월부터 2019년 12월 현재에 이르기까지 많은 분들로부터 다양한 도움을 받았습니다. 연구의 방향과 연구자의 태도에 대해 아낌없는 조언과 지도를 해주신 지도교수님 김광조 교수님께 감사드립니다. 또한 바쁘신 와중에도 학위논문심사에 참여하셔서 진심어린 조언을 해주신 신인식 교수님과 김흥택 교수님께도 감사드립니다.

내 자신을 끊임없이 의심하고, 의미없이 보낸 시간들을 자책하며 아까워할 때마다 이유를 묻지 않고 지지 및 격려를 해준 석사 과정 동기생 홍동연, 한성호에게 고마움을 표하고 싶습니다. 또한 2개 학기 동안 같은 과목을 수강하면서 부족한 나와 프로젝트를 하느라 몇 배로 힘들었을, 하지만 겉으로 드러내지 않고 오직 자신의 실력으로 그 난관을 극복한 백승근에게 고마움과 미안함을 전하고 싶습니다. 이외에도 연구실의 터줏대감 최락용, 이지은, 한 학기 먼저 석사 학위를 취득하고 이제는 어엿한 직장인이 된 최낙준도 나에게 큰 힘이 되어주었습니다.

카이스트에서 보낸 2년의 시간은 2003년 육군사관학교에 들어가 뒤돌아보지 않고 달리던 내게 처음으로 '늦잠'의 즐거움을 느끼게 해 준 동시에 현 시대 젊은이들의 '무한경쟁' 정글을 경험해볼 수 있었던 놀랍고도 특별한 경험이었습니다. 내가 살아가고 있는 시간의 옆과 뒤를 돌아보면서 앞으로 어떻게 살아갈 것인가를 그 어느 때보다 진지하고 치열하게 고민할 수 있었습니다.

지금이 나의 인생에서 어디쯤인지 알 수는 없지만 이 축복이자 행운의 시간을 자양분으로 또 하루하루를 기뻐하며, 고민하며, 즐기며, 의심하며 살아가겠습니다. 다시 한번 격려해주신 모든 분께 진심을 담아 감사함을 전합니다.

약 력

이 름: 이 나 비
생 년 월 일: 1984년 11월 29일
전 자 주 소: butterfly2@kaist.ac.kr

학 력

2001. 3.	-	2003. 1.	고등학교
2003. 1.	-	2007. 3.	육군사관학교 심리학과 (B.S.)
2018. 3.	-	2020. 2.	한국과학기술원 정보보호대학원 (M.S.)

경 력

2007. 8.	-	2008. 12.	5군단 헌병대 소대장
2009. 1.	-	2010. 12.	5군단 헌병대 범죄예방장교
2011. 1.	-	2012. 6.	육본 헌병실 중앙수사대 수사장교
2012. 7.	-	2013. 6.	육본 헌병실 범죄정보분석실 범죄정보분석장교
2013. 7.	-	2013. 12.	헌병 고군반
2014. 12.	-	2015. 3.	수방사 헌병단 수사장교
2015. 4.	-	2015. 9.	국방어학원 영어반
2015. 9.	-	2017. 12.	국방부 조사본부 범죄정보분석실 범죄정보분석장교

연구업적

1. 한성호, 홍동연, 최낙준, **이나비**, 김광조, "(D)PoS 기반 블록체인의 거래 및 합의 방식 분석 ", 한국정보보호학회 하계학술대회 (CISC-S'18), 동신대학교, 나주, 2018. 6. 21.
2. Seunggeun Baek, **Nabi Lee**, and Kwangjo Kim, “Generic Analysis of E-voting Protocols by Simplified Blockchain“, 2019 Symposium on Cryptography and Information Security, Session 2G3-5 (SCIS 2019), Jan., 22-25, 2019, Otsu, Japan.
3. **이나비**, 김광조. ”디지털 포렌식을 위한 안드로이드 및 윈도우즈 환경에서 카카오톡 메시지의 아티팩트 분석 (Ⅰ)“, 한국정보보호학회 충청지부 학술논문발표회, 한남대학교, 대전. 2019. 10. 18.
4. **이나비**, 김광조. ”디지털 포렌식을 위한 안드로이드 및 윈도우즈 환경에서 카카오톡 메시지의 아티팩트 분석 (Ⅱ)“, 한국정보보호학회 동계학술대회 (CISC-W‘19), 중앙대학교, 서울, 2019. 11. 30.