

단어그룹 확장 기법을 활용한 순환신경망 알고리즘 성능개선 연구

박대승¹, 성열우², 김정길^{3*}

¹남서울대학교 컴퓨터소프트웨어학과 석사과정, ²남서울대학교 컴퓨터소프트웨어학과 학부생,
³남서울대학교 컴퓨터소프트웨어학과 교수

A Study on Performance Improvement of Recurrent Neural Networks Algorithm using Word Group Expansion Technique

Dae Seung Park¹, Yeol Woo Sung², Cheong Ghil Kim^{3*}

¹Master's Course, Dept. of Computer Science, Namseoul University

²Student, Dept. of Computer Science, Namseoul University

³Professor, Dept. of Computer Science, Namseoul University

요약 최근 인공지능(AI)과 딥러닝 발전으로 대화형 인공지능 챗봇의 중요성이 부각되고 있으며 다양한 분야에서 연구가 진행되고 있다. 챗봇을 만들기 위해서 직접 개발해 사용하기도 하지만 개발의 용이성을 위해 오픈소스 플랫폼이나 상업용 플랫폼을 활용하여 개발한다. 이러한 챗봇 플랫폼은 주로 RNN (Recurrent Neural Network)과 응용 알고리즘을 사용하며, 빠른 학습속도와 모니터링 및 검증의 용이성 그리고 좋은 추론 성능의 장점을 가지고 있다. 본 논문에서는 RNN과 응용 알고리즘의 추론 성능 향상방법을 연구하였다. 제안 방법은 RNN과 응용 알고리즘 적용 시 각 문장에 대한 핵심단어의 단어그룹에 대해 확장학습을 통해 데이터에 내재된 의미를 넓히는 기법을 사용하였다. 본 연구의 결과는 순환 구조를 갖는 RNN, GRU (Gated Recurrent Unit), LSTM (Long-short Term Memory) 세 알고리즘에서 최소 0.37%에서 최대 1.25% 추론 성능향상을 달성하였다. 본 연구를 통해 얻은 연구결과는 관련 산업에서 인공지능 챗봇 도입을 가속하고 다양한 RNN 응용 알고리즘을 활용하도록 하는데 기여할 수 있다. 향후 연구에서는 다양한 활성화 함수들이 순환신경망 알고리즘의 성능 향상에 미치는 영향에 관한 연구가 필요할 것이다.

키워드 : 인공지능, 챗봇, 딥러닝, 순환신경망 알고리즘, 순환신경망

Abstract Recently, with the development of artificial intelligence (AI) and deep learning, the importance of conversational artificial intelligence chatbots is being highlighted. In addition, chatbot research is being conducted in various fields. To build a chatbot, it is developed using an open source platform or a commercial platform for ease of development. These chatbot platforms mainly use RNN and application algorithms. The RNN algorithm has the advantages of fast learning speed, ease of monitoring and verification, and good inference performance. In this paper, a method for improving the inference performance of RNNs and applied algorithms was studied. The proposed method used the word group expansion learning technique of key words for each sentence when RNN and applied algorithm were applied. As a result of this study, the RNN, GRU, and LSTM three algorithms with a cyclic structure achieved a minimum of 0.37% and a maximum of 1.25% inference performance improvement. The research results obtained through this study can accelerate the adoption of artificial intelligence chatbots in related industries. In addition, it can contribute to utilizing various RNN application algorithms. In future research, it will be necessary to study the effect of various activation functions on the performance improvement of artificial neural network algorithms.

Key Words : AI, Chatbot, Deep Learning, Neural Network Algorithms, Recurrent Neural Network

This work was supported by the National Research Foundation of Korea Grant funded by the Korean Government (NRF-2021R111A4A01049755) and the MSIT (Ministry of Science and ICT), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2020-0-01846) supervised by the IITP (Institute of Information & Communications Technology Planning & Evaluation).

*Corresponding Author : Cheong Ghil Kim(cgkim@nsu.ac.kr)

Received March 16, 2022

Revised April 19, 2022

Accepted April 20, 2022

Published April 28, 2022

1. 서론

최근 다양한 분야에서 대화형 인공지능 챗봇이 도입되고 있다[1-3]. 챗봇을 만들기 위해 직접 개발하기도 하지만, 개발의 용이성과 속도를 높이기 위해 챗봇 플랫폼을 활용하기도 한다. 챗봇 플랫폼에는 오픈소스 플랫폼을 비롯하여 상업용 플랫폼에 이르기까지 다양한 형태의 플랫폼이 존재한다. 대표적으로 Microsoft의 BotFramework, Google의 Dialogflow, Opensource Chatbot Rasa와 같은 플랫폼이 있다[4,5]. 이러한 챗봇 플랫폼에서는 순환신경망 알고리즘인 RNN (Recurrent Neural Network)과 응용 알고리즘을 사용한다[6-8].

순환신경망은 정해지지 않은 길이의 배열 그리고 시계열 데이터 (Sequential Data)의 해석과 처리에 적합한 모델로서 RNN 및 응용 알고리즘 모델은 사전적 의미에서 나아가 앞뒤 문장의 단어들을 종합적으로 고려하여 의미를 파악하고 해석하여 대화의 흐름을 파악하는데 용이하다[6,7]. 또한, 문장의 내재된 의미를 해석하는데 효과적이어서 챗봇 기술에서 문장처리에 많이 사용된다. RNN은 지속적으로 발전되어 LSTM (Long Short Term Memory), GRU(Gated Recurrent Unit)과 같은 응용 RNN이 만들어져 사용되고 있으며 언어모델링, 텍스트 감정분석, 기계번역 등 다양한 분야에 사용되고 있다[9,10].

또한, 앞서 설명한 RNN의 장점과 더불어 이러한 순환신경망 알고리즘은 챗봇 환경에서 학습속도가 빠르고 학습 과정을 모니터링하고 빠르게 검증할 수 있으며, 좋은 추론 성능을 제공하는 장점[8]으로 인해 챗봇에 사용되고 있다. 그러나, 순환신경망의 추론 성능은 도메인, 환경 등 다양한 요소에 따라 추론 성능의 편차가 심하며, 균일하고 높은 추론 성능을 제공하는 데 한계가 있다[11,12].

본 논문에서는 빠른 학습속도로 개발의 용이성을 제공하고 모니터링과 좋은 추론 성능을 제공하여 챗봇에서 사용되는 RNN과 응용 알고리즘에 대해 단어그룹을 활용한 의도확장학습을 통해 성능을 개선하여 추론 성능의 향상방법을 제안하고 검증하고자 한다.

2. 기존연구

2.1 배경기술

2.1.1 챗봇의 아키텍처와 종류

챗봇(Chatbot)은 음성 명령이나 텍스트를 통해서 컴퓨터와 대화하는 프로그램이다. Fig. 1은 챗봇의 동작구조를 보여준다. 사용자가 질의 메시지를 전송하면, 챗봇은 전송된 질의 메시지의 대화 의도를 분석한다. 챗봇 서버가 학습된 모델을 통해 메시지에서 의도와 필요한 변수를 추출한 후, 이 전의 대화 흐름을 통하여 응답 메시지를 선택하거나 생성한다[1,13,14].

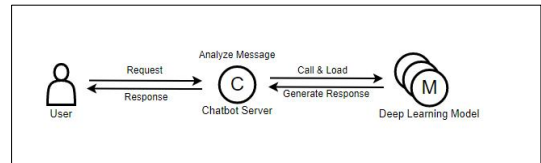


Fig. 1. General Chatbot Architecture

이때, 챗봇은 응답을 생성하는 방식에 따라 규칙 기반, 검색 기반, 생성 기반 모델로 나뉘게 된다[15,16].

규칙 기반(Rule Based) 모델은 정해진 규칙에 따라서 사전에 정의한 응답을 전송하는 가장 기본적인 방법이다. 입력받은 메시지에 응답하는 메시지가 명확하게 정의되어 있는 방식이다[15,17].

검색 기반(Retrieval Based) 모델은 자연어 처리(Natural Language Processing) 방식을 활용하여 머신러닝 등을 통해서 학습된 데이터를 기반으로 사용자의 입력 메시지의 대화 의도와 필요한 파라미터를 추출한다. 미리 학습된 데이터를 기반으로 사용자의 입력을 해석하고, 해석된 정보를 기반으로 가장 적합한 응답 메시지를 전송한다[15,18,19].

생성 기반(Generative Based) 모델은 기존에 학습된 텍스트에 의존하지 않고, 사용자의 입력에 따라서 고유한 응답 메시지를 생성한다. 딥러닝을 사용하여 어떤 응답 메시지를 생성할지 결정한다. 미리 정의된 규칙이나 학습 데이터에만 국한되지 않고, 응답 메시지를 생성한다. 하지만, 구현하기 위해서는 많은 양의 데이터가 필요하며, 응답 메시지를 생성하기 때문에 무의미한 응답 메시지가 생성될 가능성도 존재한다[15,20].

또한, 챗봇은 대화 주제의 범위에 따라 분류할 수도 있으며, 도메인의 범위를 한정하는 Close 도메인 방식과 다양한 주제를 가지고 대화할 수 있는 Open 도메인 방식이 있다.

Close 도메인 챗봇은 특정 영역에 대해서만 응답한다. 식당 예약, 상담 등 특정한 목표에 초점을 맞추고

설계되어 있다[15,21].

Open 도메인 챗봇은 사람과 실제로 대화하는 것처럼 다양한 주제를 가지고 대화를 할 수 있다. 실제로 사람과 대화하는 것처럼 느껴지지만, Close 도메인 챗봇보다 구현하기 어렵다[15,22].

2.1.2 자연어처리

자연어처리는 인공지능 또는 AI의 한 분야를 말하며, 인간이 할 수 있는 것과 거의 같은 방식으로 텍스트와 말을 이해할 수 있는 능력을 컴퓨터에 제공하는 것과 관련이 있다[23].

이러한 자연어처리는 계산 언어학과 통계, 기계 학습 및 딥러닝 모델을 결합한다. 이러한 기술을 함께 사용하면 컴퓨터가 텍스트 또는 음성 데이터의 형태로 인간 언어를 처리하고 화자 또는 작가의 의도와 감정으로 완전한 의미를 이해할 수 있다[23].

이러한 자연어처리 분야는 챗봇과 밀접한 연관이 있다. 학습된 챗봇모델과 대화를 하는 데 있어 사용자가 입력한 질의에 대한 의도 분석을 필수로 필요로 하기 때문이다. 또한, 자연어처리는 시스템 및 서비스 운영을 간소화하고 생산성을 높이며 프로세스를 단순화하는데 중요한 역할을 하고 있다[23].

2.2 인공신경망 알고리즘

2.2.1 RNN (Recurrent Neural Network)

RNN은 인공신경망의 한 종류로서 음성인식, 언어인식과 같은 정보처리를 목적으로 고안되었다. RNN은 입력과 출력을 시퀀스 단위로 처리하는 것이 특징이며, 순환구조를 통해 입력 데이터의 길이에 관계없이 데이터를 처리할 수 있는 유연한 구조를 갖는 것이 특징이다.

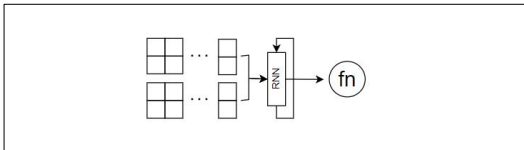


Fig. 2. RNN Structure

Fig. 2와 같이 출력이 다음 계산에 사용되기에 반복 순환되는 구조를 갖게 된다[24]. 앞서 설명한 유연한 구조 덕분에 기계번역, 감성 분석 등 다양한 분야에 활용 가능한 구조를 손쉽게 만들 수 있는 장점이 있다.

2.2.2 LSTM (Long Short-term Memory)

LSTM은 RNN의 응용 알고리즘으로써, RNN에서 학습이 길어져 학습시점과 추론시점이 멀어지는 경우 처음 학습한 내용을 잊어버리게 되는 장기 의존성 문제 (Vanishing Gradient Problem)를 해결하기 위해 고안되었다[24,25].

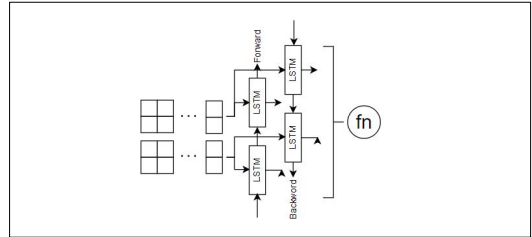


Fig. 3. LSTM Structure

Fig. 3과 같이 LSTM은 앞서 설명한 긴 의존 기간을 요구하는 학습모델을 위해 역전파를 위한 통로를 별도로 만들어 장기 의존성 문제를 해결할 수 있도록 고안되었다[25].

2.2.3 GRU (Gated Recurrent Units)

GRU는 게이트 메커니즘을 활용한 RNN의 응용 알고리즘이다[26]. LSTM의 경우 RNN의 의존성 문제를 해결했으나, 파라미터가 많아지게 되어 학습 데이터가 충분하지 않은 경우 Overfitting 이 발생하게 되는 단점이 있다[25,26]. GRU는 LSTM과 유사한 부분도 있으며 더욱 간략한 구조를 갖고 앞서 설명한 문제를 개선하기 위해 더욱 간략한 구조를 갖는 것이 특징이다.

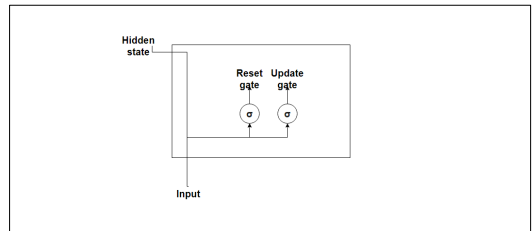


Fig. 4. GRU Structure

Fig. 4와 같이 GRU는 LSTM과 비교해 간결한 구조를 갖으면서 LSTM과 비교해 큰 성능의 변화가 없는 것이 특징이다[26].

3. 제안시스템

3.1 알고리즘 개선 이론과 추론성능 향상방법

본 논문에서 제안하는 문장변조식 단어그룹 확장학습 추론성능 향상 방법은 문장변조식 단어그룹 확장학습 방법이다. 이 방법은 문장 내 핵심단어들에 대해 비슷한 의미를 갖는 단어 그룹을 사용하여 해당 문장과 함께 학습하는 방식이다.

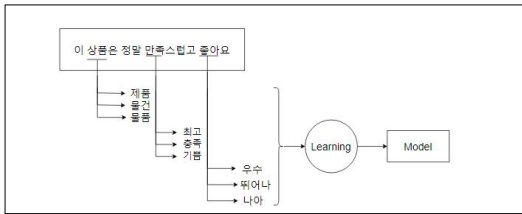


Fig. 5. Sentence Modification Architecture

본 논문에서 제안하는 문장변조식 단어그룹 확장학습은 Fig. 5와 같은 과정을 갖는다. 먼저, 학습할 단어를 선택하기 위한 과정을 거친다. 따라서, 문장에서 단어만을 추출하기 위해 입력된 학습데이터에서 불용어를 제거하는 과정을 거친다. 이후 남은 단어에서 핵심어를 선택하고 추출 하게되는데, 핵심어는 학습데이터의 도메인에 따라 달라지게 된다. 예를 들어, 본 논문에서 구현하는 프로토타입 실험모델에서 사용하는 한국어대화 데이터에서 감정을 분석하고자 하는 경우에는 기분, 좋다, 나쁘다와 같은 단어가 핵심어가 될 수 있으며, 만약 상품의 평가와 관련된 도메인이라면 상품, 좋다, 나쁘다와 같은 핵심어를 구성할 수 있다. 사전에 구성된 핵심어는 표준어국어대사전을 바탕으로 비슷한 의미를 가진 유의어를 바탕으로 단어그룹을 구성하고, 모델을 구성하는 과정에서 데이터와 단어그룹을 같이 학습해 데이터가 포함하는 단어의 범위를 확장해 학습을 하게된다. 이처럼, 문장변조식은 단어그룹만 구성하면 되기에 적용이 빠르고 간편하며, 전체 데이터셋을 기반으로 적용하는 것이 아니므로, 자주 변동하는 데이터에도 쉽게 적용이 가능한 장점이 있다. 또한 문장변조식은 단어그룹의 수를 늘림으로써 지속적으로 성능향상을 시도할 수 있는 장점이 있다.

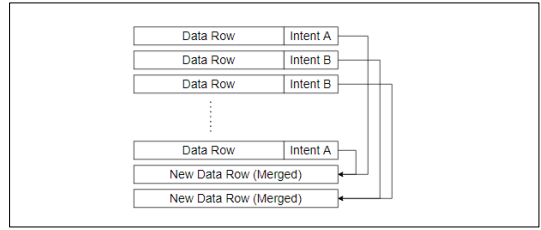


Fig. 6. Sentence Merge Architecture

일반적으로 문장변조식 구조는 문장병합식 또는 가중치강화식 구조와 비교된다. 문장병합식 구조는 Fig. 6과 같이 비슷하거나 동일한 의미의 문장을 병합하여 새로운 문장을 생성한다[27]. 해당 문장을 분리된 문장의 의미와 동일한 의미로써 지도학습을 하여 관련된 문장이 입력되었을 때 보다 나은 추론 성능을 기대할 수 있게 된다. 그러나, 문장병합식의 적용을 위해서는 같은 의미를 갖는 데이터를 실시간으로 파악해 반영해야 하므로 본 논문에서 제안하는 문장변조식에 비해 자주 변동하는 데이터셋에 적용하는 데 어려운 단점이 있다.

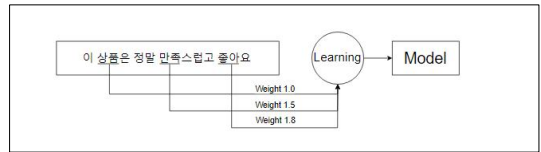


Fig. 7. Weight Reinforcement Architecture

다음으로 가중치강화식 구조는 문장 내 각 핵심단어들이나 주요 의미를 갖는 단어들에 대해 가중치를 부여해 학습하는 방법이다[28,29]. Fig. 7과 같이 핵심단어별 가중치를 부여한다. 학습에 각 가중치가 연산 되어 반영된다[29]. 다만, 가중치강화식은 본 논문에서 제안하는 문장변조식에 비해 일반적인 신경망 알고리즘 구조에 적용하기 어렵고, 학습 과정에서 데이터에 대한 가중치 데이터를 추가로 계산해야 해 연산과정이 복잡하다는 단점이 있다.

3.2 제안시스템 구현

앞서 제안한 문장변조식 알고리즘 개선 방법을 적용하기 위해서 Fig. 8과 같이 기존에 입력된 데이터셋을 전처리 후 토큰화를 진행하는 일반적인 딥러닝 과정 사이에 문장변조식 과정을 추가함으로써 구현하였다.

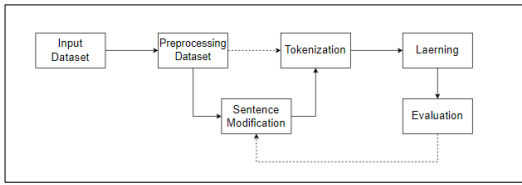


Fig. 8. Implementation System

문장변조식은 식 1과 같이 각 핵심단어에 대해 같은 의미를 갖는 단어그룹을 생성하고 이를 매칭하여 학습할 데이터의 양을 늘려주게 된다.

$$L(a, b, c \dots f) \quad (1)$$

$$Lx = \begin{cases} nx & (n = 1) \\ \lim_{n \rightarrow \infty} nx + x & (n > 2) \end{cases} \quad (2)$$

또한, 단어 그룹내에 속한 단어가 다시 또 다른 단어 그룹과 연관되어 있는 경우 연관된 단어를 활용한 데이터의 수가 식 2와 같이 늘어나는 구조를 갖는다.

4. 실험

본 실험을 시행하고 검증하기 위해 앞서 설명한 구조를 Google Colab 환경에서 Python을 기반으로 Tensorflow와 Keras를 활용해 실험용 모델을 구현하였다.

Table 1. Test Environment

Parameters	Value
Embedding Layer	256
Hidden Layer	256
Batch Size	512
Learning Validation	20%
Early Stopping	Y
Model Checkpoint	Y

각 모델별로 동일한 실험 환경을 구축하기 위해 실험환경을 Table 1과 같이 구성했으며, 활성 함수는 Sigmoid를 기준으로 하였다.

본 실험에서 활용하는 데이터세트는 AIHub의 한국어대화 50만건으로 실험한다. 본 실험에서 검증하는 알고리즘은 RNN을 비롯하여 순환구조를 갖는 RNN의 응용 알고리즘인 GRU와 LSTM을 구현해 실험모델을 구축하고 추론 정확도를 비교를 통해 검증하였다.

4.1 RNN 계열 알고리즘 기반 단어 확장학습

Fig. 9는 LSTM 알고리즘을 통해 테스트 데이터세트를 추론하여 각 데이터의 결과를 1/100 스케일한 분포 그래프이다. LSTM을 비롯한 인공신경망 알고리즘을 통해 질의를 추론하는 경우 Fig. 9와 같이 결과 수치가 0 에서 1 사이에 고르게 분포한다.

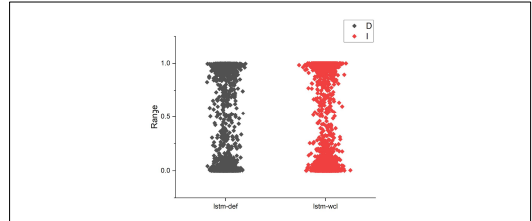


Fig. 9. Inference Distribution 1/100 Graph

이 중 0 또는 1에 가까운 결과는 높은 정확도를 가진 결과로 볼 수 있으며, 0.5 내외의 근접한 수치를 가진 결과 또는 이에 근접한 결과들은 추론이 정확하지 않을 가능성이 크다. 그러므로, 실험에서 정확도가 낮은 데이터의 추론을 향상하기 위해 추론 시 데이터의 의미를 비슷한 단어그룹으로 확장하여 학습하고 결과모델을 만들었다. 이를 통해 데이터가 단어그룹의 의미를 포함하게 함으로써 추론 성능을 개선할 수 있도록 유도하였다.

4.2 알고리즘 개선 실험과 결과

본 연구에서 제안한 단어그룹 확장 기법을 적용하여 Fig. 10과 같이 RNN 알고리즘을 비롯한 GRU, LSTM에서 최소 0.37% 에서 최대 1.25% 의 추론 성능이 향상됨을 검증하였다.

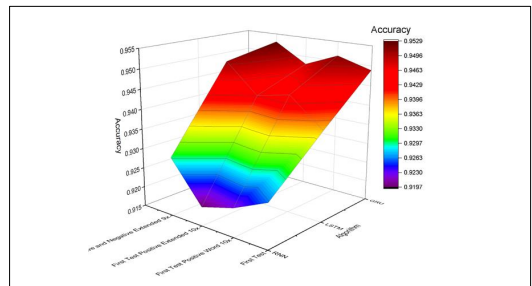


Fig. 10. Test Result

특히 LSTM 알고리즘에서 Fig. 11과 같이 큰 폭의 성능 향상을 보여주었다. 또한, 소규모의 단어그룹을 활용한 실험에서는 추론성능이 감소하는 현상을 보였다.

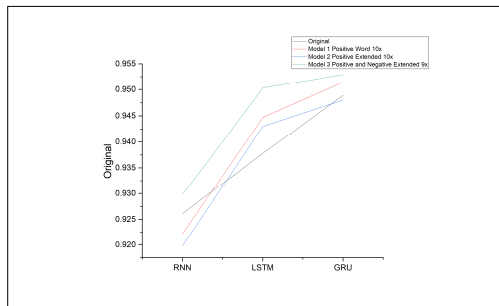


Fig. 11. Experimental Model Comparison Graph

그러나 본 실험에서 Sigmoid 활성화함수를 사용하여 추론 결과값을 압축함으로써 역전파시 Gradient가 점차 줄어들어 학습 능력이 저하됨에 따라 발생하는 Vanishing Gradient Problem과 초기실험에서의 소규모 단어그룹 구성으로 인해 RNN, GRU 알고리즘에서 0.37% ~ 0.4%의 추론 성능 향상으로 LSTM 대비 약 70% 낮은 성능향상을 보였다. 하지만, 본 실험에서 증명한것과 같이 소규모의 단어 그룹에 대한 확장이 아닌 보다 다양한 넓은 범위의 단어 그룹을 사용하면 더 높은 추론 성능 개선이 가능할 것으로 기대한다.

5. 결론

본 논문은 챗봇 플랫폼에서 선호도가 높은 RNN과 응용 알고리즘들을 개선하여 추론 성능을 향상하는 방법을 제안했다. 1:3, 3:9 단어그룹 확장 기법을 활용한 신경망 알고리즘의 추론 성능향상을 구현하여 실험 및 검증하였다. 이를 통해, RNN과 응용 알고리즘에서 최소 0.37%에서 최대 1.25%의 성능향상을 확인하여 인공지능 알고리즘에서 유의미한 결과를 확인하였다. 본 연구를 통해 얻은 이러한 연구성과는 관련 산업에서 인공지능 챗봇 도입을 가속하고 RNN과 응용 알고리즘을 활용하도록 하는데 기여할 수 있다. 향후 연구에서는 본 실험에서 Sigmoid 기반 활성화함수를 기반으로 한 Vanishing Gradient Problem이 발생하였고 이 문제를 최소화하기 위해 ReLU (Rectified Linear Unit) 활성화함수를 기반으로 Vanishing Gradient Problem

을 다루려는 연구들이 진행되고 있으며 다양한 활성화함수들이 발표되고 있다[30,31]. Sigmoid 활성화함수의 경우 0과 1사이의 값만 다루기에 각 Layer에서 값을 계속 곱하면 0에 수렴하기에 Vanishing Gradient Problem과 같은 문제를 갖고 있는데, ReLU 활성화함수는 입력값이 0보다 작으면 음수가 아닌 0으로 출력하여 값이 1보다 작아지지 않게해 Vanishing Gradient Problem을 최소화 할 수 있는 장점이 있다. 다만, ReLU 활성화함수도 여러 단점이 존재해 이를 보완하기 위한 응용 ReLU 함수들이 발표되어 향후 연구에서는 이와 같은 다양한 활성화함수들이 인공지능경망 알고리즘 향상 방법에 미치는 영향에 대한 연구가 필요할 것이다.

REFERENCES

- [1] Rahman, A. M., Al Mamun, A. & Islam, A. (2018). Programming challenges of chatbot: Current and future prospective. *5th IEEE Region 10 Humanitarian Technology Conference 2017, R10-HTC 2017*, 2018-Janua, 75-78. DOI : 10.1109/R10-HTC.2017.8288910
- [2] Hristidis, V. (2019). Chatbot Technologies and Challenges. *2018 First International Conference on Artificial Intelligence for Industries (AI4I)*, 126-126. DOI : 10.1109/ai4i.2018.8665692
- [3] Chen, L., Chen, P. & Lin, Z. (2020). Artificial Intelligence in Education: A Review. *IEEE Access*, 8, 75264-75278. DOI : 10.1109/ACCESS.2020.2988510
- [4] Daniel, G. & Cabot, J. (2021). The Software Challenges of Building Smart Chatbots. *Proceedings - International Conference on Software Engineering*, 324-325. DOI : 10.1109/ICSE-COMPANION52605.2021.00138
- [5] Nguyen, P. A. H. (2021). Evaluating AI chatbot platforms by a fuzzy AHP approach. *Proceedings of 2021 International Conference on System Science and Engineering, ICSSE 2021*, 206-209. DOI : 10.1109/ICSSE52999.2021.9538447
- [6] Prasetyo, A. & Santoso, H. A. (2021). Intents Categorization for Chatbot Development Using Recurrent Neural Network (RNN) Learning. *2021 7th International Conference on Advanced Computing and Communication Systems, ICACCS 2021*, 551-556. DOI : 10.1109/ICACCS51430.2021.9441947

- [7] Majid, R. & Santoso, H. A. (2021). Conversations Sentiment and Intent Categorization Using Context RNN for Emotion Recognition. *2021 7th International Conference on Advanced Computing and Communication Systems, ICACCS 2021*, 46-50. DOI : 10.1109/ICACCS51430.2021.9441740
- [8] Anki, P., Bustamam, A., Al-Ash, H. S. & Sarwinda, D. (2020). High Accuracy Conversational AI Chatbot Using Deep Recurrent Neural Networks Based on BiLSTM Model. *2020 3rd International Conference on Information and Communications Technology, ICOIACT 2020*, 382-387. DOI : 10.1109/ICOIACT50329.2020.9332074
- [9] Mou, L., Ghamisi, P. & Zhu, X. X. (2017). Deep recurrent neural networks for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 55(7), 3639-3655. DOI : 10.1109/TGRS.2016.2636241
- [10] Li, S., Li, W., Cook, C., Zhu, C. & Gao, Y. (2018). Independently Recurrent Neural Network (IndRNN): Building A Longer and Deeper RNN. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 5457-5466. DOI : 10.1109/CVPR.2018.00572
- [11] Williams, R. J. & Zipser, D. (1989). A Learning Algorithm for Continually Running Fully Recurrent Neural Networks. *Neural Computation*, 1(2), 270-280. DOI : 10.1162/neco.1989.1.2.270
- [12] Donahue, J., Hendricks, L. A., Guadarrama, S., Rohrbach, M., Venugopalan, S., Darrell, T. & Saenko, K. (2015). Long-term recurrent convolutional networks for visual recognition and description. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 07-12-June*, 2625-2634. DOI : 10.1109/CVPR.2015.7298878
- [13] Lebeuf, C., Storey, M. A. & Zagalsky, A. (2017). Software Bots. *IEEE Software*, 35(1), 18-23. DOI : 10.1109/MS.2017.4541027
- [14] Daniel, G., Cabot, J., Deruelle, L. & Derras, M. (2020). Xatkit: a Multimodal Low-Code Chatbot Development Framework. *IEEE Access*, 8, 15332-15346. DOI : 10.1109/aACCESS.2020.2966919
- [15] Lokman, A. S. & Ameen, M. A. (2019). Modern chatbot systems: A technical review. *Advances in Intelligent Systems and Computing*, 881, 1012-1023. DOI : 10.1007/978-3-030-02683-7_75
- [16] Hussain, S., Ameri Sianaki, O. & Ababneh, N. (2019). A Survey on Conversational Agents/Chatbots Classification and Design Techniques. *Advances in Intelligent Systems and Computing*, 927, 946-956. DOI : 10.1007/978-3-030-15035-8_93
- [17] Mellado-Silva, R., Faundez-Ugalde, A. & Lobos, M. B. (2020). Learning tax regulations through rules-based chatbots using decision trees: A case study at the time of COVID-19. *Proceedings - International Conference of the Chilean Computer Science Society, SCCC, 2020-Novem*. DOI : 10.1109/SCCC51225.2020.9281267
- [18] Shukla, V. K. & Verma, A. (2019). Enhancing LMS Experience through AIML Base and Retrieval Base Chatbot using R Language. *2019 International Conference on Automation, Computational and Technology Management, ICACTM 2019*, 561-567. DOI : 10.1109/ICACTM.2019.8776684
- [19] Dutt, V., Sasubilli, S. M. & Yerrapati, A. E. (2020). Dynamic Information Retrieval with Chatbots: A Review of Artificial Intelligence Methodology. *Proceedings of the 4th International Conference on Electronics, Communication and Aerospace Technology, ICECA 2020*, 1299-1303. DOI : 10.1109/ICECA49313.2020.9297533
- [20] Varghese, E. & Pillai, M. T. R. (2018). A Standalone Generative Conversational Interface Using Deep Learning. *Proceedings of the International Conference on Inventive Communication and Computational Technologies, ICICCT 2018*, 1915-1920. DOI : 10.1109/ICICCT.2018.8473211
- [21] Nigam, A., Sahare, P. & Pandya, K. (2019). Intent Detection and Slots Prompt in a Closed-Domain Chatbot. *Proceedings - 13th IEEE International Conference on Semantic Computing, ICSC 2019*, 340-343. DOI : 10.1109/ICOSC.2019.8665635
- [22] Roller, S. et al. (2021). Recipes for building an open-domain chatbot. *EACL 2021 - 16th Conference of the European Chapter of the Association for Computational Linguistics, Proceedings of the Conference*, 300-325. DOI : 10.18653/v1/2021.eacl-main.24
- [23] Young, T., Hazarika, D., Poria, S. & Cambria, E. (2018). Recent trends in deep learning based natural language processing [Review Article]. In *IEEE Computational Intelligence Magazine* (Vol. 13, Issue 3, pp. 55-75). Institute of Electrical and Electronics Engineers Inc. DOI : 10.1109/MCI.2018.2840738
- [24] Ruinehart, D. E., Hinton, G. E. & Williams, R. J. (1986). Learning Internal Representations Error Propagation. In *Cognitive Science* (Vol. 1, Issue V). Retrieved from : <https://apps.dtic.mil/docs/citations/ADA164453>

- [25] Hochreiter, S. & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, 9(8), 1735-1780. DOI : 10.1162/neco.1997.9.8.1735
- [26] Chung, J., Gulcehre, C., Cho, K. & Bengio, Y. (2014). Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling. *Presented in NIPS 2014 Deep Learning and Representation Learning Workshop*. DOI : 10.48550/arxiv.1412.3555
- [27] Shen, Y., Zhang, Q., Zhang, J., Huang, J., Lu, Y. & Lei, K. (2019). Improving medical short text classification with semantic expansion using word-cluster embedding. *Lecture Notes in Electrical Engineering*, 514, 401-411. DOI : 10.1007/978-981-13-1056-0_41
- [28] Cao, W., Song, A. & Hu, J. (2017). Stacked residual recurrent neural network with word weight for text classification. *IAENG International Journal of Computer Science*, 44(3), 277-284.
- [29] Kim, S., Na, I., Shin, J. & Society, K. M. (2019). A Method on Associated Document Recommendation with Word Correlation Weights. *Journal of Korea Multimedia Society*, 22(2), 250-259. DOI : 10.9717/kmms.2019.22.2.250
- [30] Hochreiter, S. (1998). The vanishing gradient problem during learning recurrent neural nets and problem solutions. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 6(2), 107-116. DOI : 10.1142/S0218488598000094
- [31] Talathi, S. S. & Vartak, A. (2015). Improving performance of recurrent neural network with relu nonlinearity. *arXiv preprint arXiv:1511.03771*. DOI : 10.48550/arxiv.1511.03771

박 대 승(Dae Seung Park) [학생회원]



- 2015년 2월 : 남서울대학교 컴퓨터학과(공학사)
- 2020년 9월~현재 : 남서울대학교 컴퓨터소프트웨어학과(석사과정)
- 관심분야 : 임베디드, 딥러닝, 컴퓨터비전, 바이오인포매틱스
- E-Mail : easonox@gmail.com

성 열 우(Yeol Woo Sung) [학생회원]



- 2017년 2월~현재 : 남서울대학교 컴퓨터소프트웨어학과 학생
- 관심분야 : 인공지능, 딥러닝
- E-Mail : tjddufdn2@naver.com

김 정 길(Cheong Ghil Kim) [정회원]



- 2003년 2월 : 연세대학교 컴퓨터공학(공학석사)
- 2006년 2월 : 연세대학교 컴퓨터공학(공학박사)
- 2008년 3월~현재 : 남서울대학교 컴퓨터소프트웨어학과 교수

- 관심분야 : 임베디드 시스템, 3D Contents, AR, VR
- E-Mail : cgkim@nsu.ac.kr