

봇 프레임워크를 활용한 챗봇 구현 방안

김기영*

Method of ChatBot Implementation Using Bot Framework

Ki-Young Kim*

요약 본 논문에서 챗봇에서 사용하는 AI알고리즘과 자연어처리 방법을 분류하고 제시하고 챗봇 구현에 사용할 수 있는 프레임워크에 대해서도 기술한다. 챗봇은 사용자 인터페이스를 대화방식으로 구성하여 입력된 문자열을 해석하고 입력된 문자열에 적절한 답을 학습된 데이터에서 선택하여 출력하는 구조의 시스템이다. 최근 콜센터와 주문 업무에 적용하여 인건비를 감소하고 정확한 업무를 할 수 있는 장점이 있다. 하지만 질문에 대한 적절한 답변 집합을 생성하기 위해 학습이 필요하며 이를 위해 상당한 계산 기능을 갖는 하드웨어가 필요하다. 개발을 하는 업체는 물론 AI분야 개발을 학습하는 학생들의 실습은 한계가 있다. 현재 챗봇은 기존의 전통적인 업무를 대체하고 있으며 시스템을 이해하고 구현하는 실습과정이 필요한 실정이다. 정형화되어 있는 데이터에 대해서만 응답을 하는 수준을 넘어 딥러닝 등의 기술을 적용하여 비정형 데이터를 학습시켜 질문에 대한 응답의 정확성을 높이기 위해 RNN과 Char-CNN 등을 사용해야한다. 챗봇을 구현하기 위해서는 이와 같은 이론을 이해하고 있어야한다. 본 논문에서는 단기간에 챗봇 코딩교육에 활용할 수 있는 방안과 기존 개발자, 학생들이 챗봇 구현을 할 수 있는 플랫폼을 활용하여 학생들이 전체시스템을 구현 예를 제시하였다.

Abstract In this paper, we classify and present AI algorithms and natural language processing methods used in chatbots. A framework that can be used to implement a chatbot is also described. A chatbot is a system with a structure that interprets the input string by constructing the user interface in a conversational manner and selects an appropriate answer to the input string from the learned data and outputs it. However, training is required to generate an appropriate set of answers to a question and hardware with considerable computational power is required. Therefore, there is a limit to the practice of not only developing companies but also students learning AI development. Currently, chatbots are replacing the existing traditional tasks, and a practice course to understand and implement the system is required. RNN and Char-CNN are used to increase the accuracy of answering questions by learning unstructured data by applying technologies such as deep learning beyond the level of responding only to standardized data. In order to implement a chatbot, it is necessary to understand such a theory. In addition, the students presented examples of implementation of the entire system by utilizing the methods that can be used for coding education and the platform where existing developers and students can implement chatbots.

Key Words : Chatbot, NLP, BotFrame, RNN, Cloud

1. 서론

4차 산업혁명의 주요기술은 IoT, Cloud Computing, Mobile, BigData로 정의할 수 있으며 국내에서는 ICBM로 정의하기도 한다. 최근에는 AI기

술을 적용하는 AIoT 관련연구도 활발해지고 있다. ICBM도 AI 연계한 다양한 서비스가 형태로 확장되고 있으며 챗봇은 인공지능, 클라우드 컴퓨팅, 모바일과 연동하여 동작하는 대표적인 서비스이다. 인터페이스

The present research has been conducted by the Research Grant of Seoil University

*Department of Software Engineering, Seoil University

Received February 08, 2022

Revised February 11, 2022

Accepted February 21, 2022

를 대화방식으로 구성하여 기존의 콜센터의 업무를 대체하는 서비스로 활용되며 정형화되어 있는 데이터에 대해서만 응답을 하는 것을 넘어 딥러닝 등의 기술을 적용하여 비정형 데이터를 학습시켜 질문에 대한 응답의 정확성을 높이려는 시도가 지속되고 있다[1-2]. 챗봇은 메신저 기반 환경에서 사용자들의 요구나 질문에 대하여 자동으로 응답을 제공해주는 에이전트 서비스로 초기에는 키워드에 대한 정적인 답변을 제공하는 단순한 형태였으나 최근에는 사용자의 질문을 분석하여 자동으로 답변을 제공하기 위한 연구도 진행되고 있다. 웹이나 애플리케이션을 따로 실행하지 않고도 대화하듯 정보를 편리하게 얻을 수 있어 인터넷 쇼핑 주문, 호텔 예약, 정보, 뉴스 검색 등 사용범위가 점차 확대되고 있다. ICT를 선도하는 기업들인 마이크로소프트, 페이스북, 구글 등은 메신저 기반의 챗봇에 플랫폼 개발을 진행하고 있으며 실생활에 적용할 수 있는 SaaS, PaaS 기반의 챗봇 서비스를 제공하고 있다. 특히 페이스북은 메신저기반의 챗봇 플랫폼을 제공하고 있으며, 구글은 챗봇 기술로 사용자의 취향이나 기호를 분석하여 맞춤형 서비스를 제공하는 메신저 애플리케이션 알로(Allo)를 2017년 8월에 출시했으며 2015년 3월 Build 2016 컨퍼런스에서 마이크로소프트도 마이크로소프트 봇 프레임워크를 발표했다. ICT 선도기업이 시장을 형성하자 챗봇 서비스 개발에 수요가 증가하였다 하지만 챗봇 개발은 AI, 자연어처리, 인터페이스 처리 기술, 통신 기술을 필요하는 융합기술개발 분야로 모든 부분을 개발하기에는 한계가 있다. 따라서 본 논문에서는 챗봇을 대학의 한 학기에 해당하는 기간에 인공지능 기반의 챗봇 코딩교육에 활용할 수 있는 플랫폼 활용방안과 구현 예를 제시한다. 이를 위해 챗봇에서 사용되는 기술과 동향에 대해 조사를 통해 챗봇 개발자 학습을 위한 개발자와 학생들이 챗봇 구현을 할 수 있는 플랫폼에 대해 조사하고 구현 방안에 대한 내용을 다룬다. 1장에서는 챗봇기술 개요 2장 관련기술 3장 챗봇 플랫폼, 4장 구현 예, 5장 결론으로 구성하였다.

2. 관련연구

2.1 챗봇 서비스

챗봇(Chatbot)의 개념은 인터넷 보급이 시작되던 1990년도부터 개발 및 서비스가 시도되었다. 당시의 챗봇은 단순한 대화를 할 수 있어 사람들의 흥미를 유발시키거나 재미삼아 사용하는 것에 불과 했다. 예상된 질문에 대한 답변 문장을 프로그램 내에 저장하고 해당 질문의 문장이 들어오면 단어와 간단한 문장 분석을 통해 매칭된 답변 문장을 출력하는 방식으로 비정형 대화에 대한 답변을 할 수 없는 한계로 인해 얼마가지 않아 사용자들의 관심에서 멀어진 서비스가 되었다. 2010년에 설립한 딥마인드를 2014년도에 인수한 구글이 인공지능 기술개발을 진행하여 2016년 알파고를 발표했다. 일반인들은 알파고를 통해 인공지능의 발전을 확인하고 관심을 갖게 되었고 AI분야는 보편적 기술로 인식되면서 기존 서비스에 접목시키는 시도가 폭넓게 진행되었다. 그 결과 AI 마켓이 확장되었고 AI에 사용 되던 딥러닝(Deep Learning)을 챗봇(Chatbot) 서비스에 적용하게 되었다. 기존에 단순한 대화를 하던 챗봇(Chatbot)은 웹사이트의 콘텐츠와 주문, 홈페이지의 Q&A의 데이터를 학습하여 사용자가 원하는 정보를 해석하고 적절한 답변을 제공할 수 있게 되었다. 현재의 챗봇 서비스는 채팅과 로봇의 합성어로 인간과 문자 대화를 통해 질문에 적절한 답 또는 연관 정보를 제공하는 인공지능 기반 소프트웨어를 의미한다. 딥러닝의 알고리즘과 데이터 셋의 수집과 전처리는 클라우드 서비스를 활용하여 학습에 필요한 자원의 제약을 해결할 수 있다[3].

사용자의 단말을 통해 입력된 문자열들을 텐서플로어, 케라스 등의 딥러닝 플랫폼 서버에 챗봇 서버를 구성하고 사용자가 메신저를 통해 챗봇 서버로 송신한 문자열들을 해석하고 답변을 사용자 단말에 전송하도록 하였다. 딥러닝(Deep Learning)과 자연어처리(Natural Language Processing) 기술은 채팅 문자열을 해석하여 능동적인 응답을 사용자에게 제공하는 것을 가능하게 하였다. 챗봇의 기본 기능은 사용자가 입력한 문자열 형식의 언어를 해석하고 적절한 답변을 문자열로 출력해주는 구조로 언어처리

에 대한 기술이 필요하며 국가별 언어를 인식하여야 글로벌한 서비스가 가능하다. 현재 언어의 인식률의 정확성은 영어, 중국어 순으로 축적된 데이터의 따라 결정된다. 챗봇 서비스는 운영하면서 지속적으로 데이터의 수집을 진행하고 수집된 데이터를 학습 데이터로 사용하여 정확성을 향상시킬 수 있다. 이와 같은 특성으로 인해 한국어는 영어, 중국어에 비해 상대적으로 축적된 데이터의 양이 적어 정확도가 낮은 상태였지만 학습에 필요한 데이터가 축적되면서 영어, 중국어와 유사한 정확성을 나타낼 수 있다. 이와 같은 이유로 챗봇 서비스는 단독으로 구축하는 것보다 클라우드 서비스를 통해 구축하는 것이 서비스 운영을 하면서 정확성을 높일 수 있다[4-5].

2.2 자연어 처리 기술

챗봇 구현에서 필요 기술은 자연어 처리(Natural Language Processing)이다. 컴퓨터와 인간 언어 사이의 상호 작용하는 기술로 인공지능의 핵심 기능 중 하나이며 자연어처리 기술은 인간 언어정보처리 원리와 이해를 위한 언어학과 뇌인지 언어정보처리 분야까지 핵심요소이며 기계번역, 대화체질의응답 시스템 대화시스템, 정보검색, 시맨틱웹, 딥러닝, 빅데이터 분석 분야활용되고 있다[6].

1950년대부터 기계 번역과 같은 자연어 처리 기술이 연구되기 시작되었고 2010년 딥러닝 기술의 활용으로 자연어처리를 통한 서비스의 상용화가 가능해졌다. 1990년대 이후에는 대량의 데이터를 활용하는 기계학습과 통계적 자연어 처리 기법을 주로 사용하였다. 딥러닝과 딥러닝기반의 자연어처리가 방대한 텍스트로부터 의미 있는 정보를 추출하고 활용할 수 있는 언어처리 연구 개발은 최근 전 세계적으로 활발히 진행되고 있다[8].

NLP에는 자연어 분석, 자연어 이해, 자연어 생성 등의 기술이 사용된다. 문장을 해석하는 시멘틱 분석과 문장의 의미를 결정하는 실용 분석 등으로 구분할 수 있으며 NLP 솔루션은 전 산업에서 활용되고 있다. 은행, 금융 서비스 등에서 고객을 유지하고 비용을 절감하기 위해 적극 활용 중에 있어 관련 개발자의 수요가 높아지고 있다. NLP 솔루션은 기업의

정보 검색, 사용자의 의도 분석, 고객 서비스 등과 규칙 준수 프로세스 자동화 및 응용 프로그램을 효율적으로 수행하고 발생할 수 있는 위험을 완화하도록 설계, 구축하고 있다. NLP 솔루션을 사용해 프로세스를 자동화하면 은행, 금융 서비스 및 보험 등의 기업은 생산성 향상이 가능하며 해석 프로세스에서 발생할 수 있는 오류율을 매우 낮은 수준으로 유지할 수 있어 회사가 구조화되지 않은 데이터를 처리하는 방식을 최적화할 수 있다. 이와 같은 기술은 바이트의 NLP 프레임 워크 ERNIE 2.0를 활용하면 개발 시간을 효율적으로 단축할 수 있다. 그 외에 텐서플로우, 케라스, AWS, MS의 Azure의 활용을 통한 효율적 개발이 가능하다.

2.3 챗봇의 주요 핵심기술

챗봇을 인공지능을 활용하여 구현하기 위해서 언어를 처리하고 추론기능을 활용 할 수 있어야 하고 요청한 정보를 비정형 데이터에서 검색하고 현재 상황을 정보화하고 응용할 수 있어야 한다. 이와 같은 요구사항을 해결하기 위해서 자연어처리(Natural Language Processing)가 가능하여야 한다. 예를 들면 통역, 질의응답, 자동번역 등이 포함된다. 또한 문자, 도형, 음성 등을 식별이 가능한 패턴인식이 가능해야 하고 정보자원의 의미를 이해하고 처리할 수 있는 추론기능을 지원하여야 한다. 추론 처리에 입력값으로 사용할 비정형 형식으로 저장되어 있는 텍스트 데이터에서 유의미한 정보를 검색할 수 있는 기능이 가능하여야 한다.

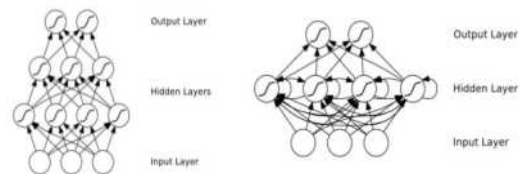


그림 1. FNN과 RNN
Fig. 1. FNN and RNN

챗봇에서 널리 활용되는 알고리즘은 그림 1과 같으며 챗봇 서비스에서는 FNN보다 RNN의 정확도

가 높은 편이어서 활용을 많이 하고 있다.

RNN을 서비스에 포함시키는 세부 구현 방법 및 처리방식에 대해서는 다음 장에서 설명한다.

2.4 챗봇 클라우드 서비스

챗봇 서비스는 일반적으로 클라우드 시스템 기반으로 구성된다. 그 형태는 그림 2와 같다. 대표적인 클라우드 서비스는 AWS, Azure가 있다.

SaaS	•Email, ERP •Yahoo, Gmail, Facebook, twitter
PaaS	•Application Development •Google Apps Engine, Microsoft Azure
IaaS	•System Management, Networking •AWS EC2, Joyent, Rackspace

그림 2. 클라우드 서비스
Fig. 2. Cloud Services

현재 Azure는 봇 프레임워크를 통해 챗봇을 지원하고 있으며 PaaS로 지원되고 있다. IaaS 방식의 AWS와 구분될 수 있지만 다양한 기능이 추가되면서 클라우드 서비스는 PaaS와 IaaS 중에 한가지 형태로 구분하기에는 모호해지고 있으며 양쪽 서비스 형식을 갖추고 있다.

3. 챗봇 프레임워크

3.1 챗봇 알고리즘

챗봇을 프레임워크에서 사용할 수 있는 알고리즘은 FNN, RNN등을 고려할 수 있다. 챗봇의 적정한 응답을 위한 모델은 RNN(Recurrent Neural Networks)이 효율성이 높다. 순차 데이터(Sequential Data)에 효율적이다. 동작절차는 다음 그림 3과 같이 입력 값을 순차적으로 처리하고 단어의 순서를 보존하는 형태로 진행된다.

하지만 RNN은 입력된 문자열의 단어 사이가 멀어질수록 단어 간의 상관관계를 의미를 기억하지 못해 관계있는 단어의 사이가 먼 문장에서는 문맥을 파악

하기가 어려운 구조이다. 거리가 먼 단어 간의 의미를 기억할 수 있는 기법에 대해 연구가 진행되었으며 단점을 해결한 것이 RNN의 특수한 경우에 해당하는 것이 LSTM(Long Short Term Memory networks)이다. 텐서플로우 등에서 챗봇을 구현할 때 LSTM Cell을 기본으로 사용하는 것이 일반적이다. 또 다른 방식은 Char-CNN으로 입력된 문장에서 의도에 해당하는 단어를 추출할 수 있으며 입력된 언어를 분류와 추론의 흐름은 그림 3과 같이 동작한다.

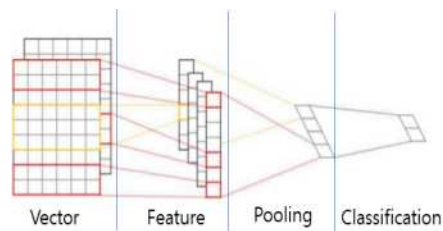


그림 3. Char-CNN 흐름도
Fig. 3. Char-CNN Flow

Char-CNN은 입력된 문자열에 주요 단어를 추출할 수 있는 모델이다. 입력된 Word2Vec 또는 GloVe를 사용하여 n개의 단어로 구성된 문장을 단어별로 k차원의 행벡터로 변환할 수 있다. 2*3크기의 CNN 필터이며 필터 개수만큼 형식지도(Feature Map)를 만들고 Max-Pooling을 통해 결과 값을 출력하는 구조를 갖는다.

챗봇을 구현하기 위해서는 자연어처리 기술, 학습 모델의 이해와 구현이 필요하다. 현업의 개발자들도 이와 같은 기술을 직접 구현하는 것은 한계가 있다. 일반적으로 오픈소스로 제공되는 텐서플로어, 케라스, 파이토치 등을 활용하여 구축하고 있지만 개발환경 구축이 쉽지 않고 최적화의 난이도가 높아 개발기간의 길어지게 된다. 다음 장에서는 이와 같은 제약을 해결할 수 있는 봇 프레임으로 챗봇 구현 실증을 다음 장에서 제시한다.

4. 구현 예

구현에 적용한 프레임은 MS사의 Azure에서 지원

하는 BoT Framework-NET Core3.1 버전을 사용하였다. 클라우드 서비스 기반으로 챗봇 서비스 구현에 적합한 프레임으로 테스트 단계에서 에뮬레이터를 제공하여 결과를 확인하고 디버깅이 가능하도록 지원하고 있어 챗봇 구현시간을 단축시킬 수 있다. 또한 다양한 NLP의 알고리즘에서 최적화된 알고리즘으로 데이터 셋과 학습을 진행시켜준다.

```

class EchoBot : ActivityHandler
{
    protected override async Task OnMessageActivityAsync(ITurnContext<MessageActivity> turnContext)
    {
        var replyText = $"Echo: {turnContext.Activity.Text}";
        await turnContext.SendActivityAsync(MessageFactory.Text(replyText, replyText));
    }

    protected override async Task OnMembersAddedAsync(IList<ChannelAccount> membersAdded, ITurnContext turnContext)
    {
        var welcomeText = "Hello and welcome!";
        foreach (var member in membersAdded)
        {
            if (member.Id != turnContext.Activity.Recipient.Id)
            {
                await turnContext.SendActivityAsync(MessageFactory.Text(welcomeText, welcomeText));
            }
        }
    }
}
    
```

그림 4. 봇 프레임 구성 소스
Fig. 4. Source of Bot Frame

봇 프레임의 구성은 EchoBot.cs에서 제공하는 메소드의 OnMessageActivityAsync()는 입력된 메시지를 Echo를 붙여서 회신하도록 되어 있고 OnMembersAddedAsync()는 사용자가 처음 접속했을 때 동작하도록 구성되어 있고 소스의 일부는 그림 4와 같다.

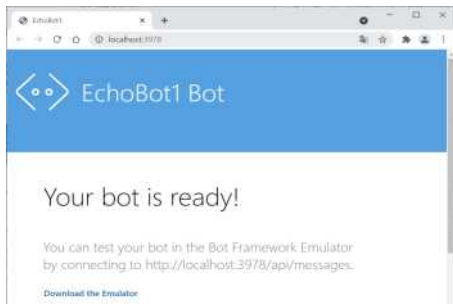


그림 5. 봇 프레임 클라우드 서비스 연결
Fig. 5. Connection of Bot Frame Cloud

클라우드 시스템과 연동이 되면 클라우드 서비스에서 그림 5와 같은 응답을 수신하여 웹으로 확인이 가능하다. 클라우드 서비스의 봇 프레임과 연동이 된

후 동작은 그림 6과 같다. 우측 하단에 문자열을 입력하면 해당하는 응답이 출력되는 것을 좌측하단과 같이 봇의 실행결과를 확인할 수 있다.

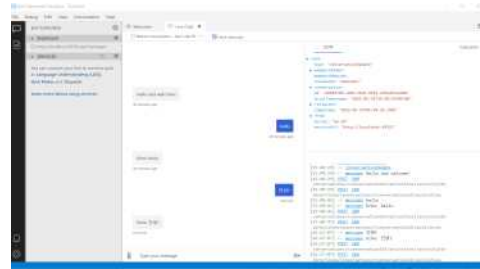


그림 6. 출력결과
Fig. 6. Result of Output

5. 결론

본 논문에서는 챗봇 구현에 필요 기술과 알고리즘에 대한 분류를 하였고 AI 기술을 적용한 응용서비스 개발 교육에 적합한 봇 프레임을 소개하고 이를 이용한 챗봇 서비스 구현 방안을 제시하였다. MS사의 봇 프레임은 텐서플로우, 케라스, 파이토치보다 상대적으로 챗봇 개발자들이 인공지능 알고리즘을 서비스에 적용하여 최종 결과물을 도출하여 상용화 수준의 정확성과 응답성을 갖는 챗봇을 구현할 수 있다. 이는 Azure기반의 클라우드 시스템을 활용할 수 있어 가능 봇 프레임을 활용한 챗봇은 AI의 알고리즘과 자연어 처리 기술인 NLP를 지원하고 있어 챗봇 개발 학습자가 학습모델 알고리즘을 구현하지 않고 이론적으로 이해하고 활용하는 방식으로 서비스 구현이 가능하다. 또한 클라우드 서비스와 연동하여 전체 서비스를 구현하기에 용이하며 개발자는 물론 AI기술을 학습하고 응용서비스를 구현하는 학습자가 이론을 단기간을 학습을 통해 서비스로 구현해 볼 수 있는 효율적인 프레임이라고 할 수 있다. 제약사항으로는 클라우드 서비스 기반의 PaaS 형식으로 제공하고 있어 회사, 학교 등의 서버 등으로 활용한 운영하는 것은 한계가 있다는 것이다. 하지만 챗봇에서 필수적으로 필요한 AI 학습모델과 자연어 처리를 지원하기 때문에 챗봇 서버 구축에 필요한

하드웨어 구입, 유지보수 비용을 절감할 수 있다는 장점이 있다. 향후 다양한 플랫폼 상의 챗봇 구축 실습과 상용화 버전 개발을 위해 봇 프레임워크 서비스를 연동하여 챗봇 기능을 확장하고 타 플랫폼과의 연동 상호 연동된 서비스 구현에 관한 연구가 필요하다.

REFERENCES

- [1] Falguni Patel, Riya Thakore, Ishita Nandwani, Santosh kumar Bharti, "Combating Depression in Students using an Intelligent Chat Bot: A Cognitive Behavioral Therapy", IEEE, 2019.
- [2] Jack Cahn, "CHATBOT: Architecture, Design, & Development", University of Pennsylvania School of Engineering and Applied Science Department of Computer and Information Science April 26, 2017.
- [3] Bayu Adhi Tama and Kyung-Hyune Rhee, "A Comparative Study of Phishing Websites Classification Based on Classifier Ensembles. Journal of Multimedia Information System VOL. 5, NO. 2, pp. 99-104, June 2018.
- [4] A Platform for Human-Chat bot Interaction Using Python
- [5] Nitiraj Singh Sandu, Ergun Gide, "Adoption of AI- Chat bots to Enhance Student Learning Experience in Higher Education in India", IEEE, 2019.
- [6] Tussanai Parthornratt, Pasd Putthapipat, Dollachart Kitsawat, Prapap Koronjaruwat, "A Smart Home Automation via Facebook Chat bot and Raspberry Pi", IEEE, 2018.
- [7] Parth Thosani, Manas Sinkar, Jaydeep Vaghasiya, Radha Shankarmani, "A Self Learning Chat-Bot from User Interactions and Preferences", IEEE, 2020.
- [8] Bhavika R. Ranoliya, Nidhi Raghuvansh and Sanjay Singh, "Chatbot for University Related FAQs", IEEE, 2017.

저자약력

김기영(Ki-Young Kim)

[중심회원]



- 2004년 3월~현재 : 서일대학교 소프트웨어공학과 부교수

<관심분야>

빅데이터, IoT, ITS, 네트워크보안