

EP-Bot: Empathetic Chatbot Using Auto-Growing Knowledge Graph

SoYeop Yoo and OkRan Jeong*

Gachon University, Seongnam-si, 13120, Korea

*Corresponding Author: OkRan Jeong. Email: orjeong@gachon.ac.kr

Received: 16 November 2020; Accepted: 29 December 2020

Abstract: People occasionally interact with each other through conversation. In particular, we communicate through dialogue and exchange emotions and information from it. Emotions are essential characteristics of natural language. Conversational artificial intelligence is an integral part of all the technologies that allow computers to communicate like humans. For a computer to interact like a human being, it must understand the emotions inherent in the conversation and generate the appropriate responses. However, existing dialogue systems focus only on improving the quality of understanding natural language or generating natural language, excluding emotions. We propose a chatbot based on emotion, which is an essential element in conversation. EP-Bot (an Empathetic PolarisX-based chatbot) is an empathetic chatbot that can better understand a person's utterance by utilizing PolarisX, an auto-growing knowledge graph. PolarisX extracts new relationship information and expands the knowledge graph automatically. It is helpful for computers to understand a person's common sense. The proposed EP-Bot extracts knowledge graph embedding using PolarisX and detects emotion and dialog act from the utterance. Then it generates the next utterance using the embeddings. EP-Bot could understand and create a conversation, including the person's common sense, emotion, and intention. We verify the novelty and accuracy of EP-Bot through the experiments.

Keywords: Emotional chatbot; conversational AI; knowledge graph; emotion

1 Introduction

Artificial intelligence technology is growing rapidly with countless big data and the development of technologies that can compute it. In particular, deep learning technology shows high performance in areas such as images and visions, showing various possibilities. Recently, with the advent of language models, the natural language processing field is also under active study [1,2].

Conversational artificial intelligence (AI) technology is emerging as research in various fields is active and good results are achieved in each area. It is not one research field or technology, but a combination of technologies that allow chatbots and voice assistants to communicate as if they



This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

were humans. A large amount of data, as well as natural language processing(NLP) technology, is used to enable computers to interact with people [1,3–5].

Natural language processing research consists of various tasks, such as named entity recognition, summarization, and sentence classification. It mainly consists of the natural language understanding (NLU) and the natural language generation (NLG) that processes natural language like a human being. NLP research is a crucial area in conversational AI research. But it is classified as a difficult task because computers have to learn all parts, unlike people who can naturally communicate through social and cultural learning [6–9].

A person communicates through conversation with others. Conversations that people send and receive include information that is ostensibly revealed, but also that is nested within sentences, such as the intention or emotion of the utterance. Emotions are a fundamental characteristic, especially since people can communicate with emotions through natural language. However, many conversational AI studies focus on improving the performance to understand natural language and generate answers [10,11].

The purpose of the conversational AI study is to enable computers to communicate like humans, after all. Emotional characteristics are essential for this. Computers should be able to understand the emotions that are implicit in the conversation and generate appropriate answers based on them. Dialogue systems that are not learned based on emotional information sometimes generate inappropriate responses. If we hear “I failed the exam” from a friend, how do we usually answer? There is no direct word for emotion in the sentence, but we can grasp the context and see that the friend feels bad or sad. A dialogue system that does not learn these emotions can give a wrong answer or, based on the similarity of sentences, give the same response as the answer to “I passed the exam” like “Congratulations.”

We propose a dialogue system based on emotion, which is an essential factor in conversation. To identify emotions from the utterance and generate appropriate answers based on the identified emotions. At this time, knowledge graphs are used to understand a person’s utterance better. Knowledge graphs are a graph that represents a person’s knowledge, including WordNet [12], YAGO [13], ConceptNet [14], etc. They can help computers learn a person’s common sense. However, the existing knowledge graphs are limited to certain languages, such as English, and there are limitations in not being able to respond to new words or new knowledge.

In this paper, we make it possible to respond to newly generated words or knowledge by utilizing PolarisX [15], the automatically expanding knowledge graph. To overcome the limitations of the existing knowledge graphs, PolarisX crawls and analyzes news and social media in real-time. It then extracts new relationship information and expands the knowledge graph. By using the automatically extended knowledge graph, it is possible to improve the performance of understanding the utterance’s intention and emotions and respond to new words or knowledge.

We propose EP-Bot (an Empathetic PolarisX-based chatbot), an empathetic chatbot using an auto-growing knowledge graph. It utilizes an automatically extended knowledge graph. It also utilizes PolarisX, an auto-growing knowledge graph, to understand the emotions and intentions inherent in the conversation more like a human being. It also generates appropriate answers to the utterance by using the extracted emotions and intentions.

2 Related Work

Various conversational artificial intelligence technologies are already being used in our daily lives, including Google Assistant and Apple Siri. With the rapid development of language-related

technology, agents based on conversational AI technology are playing a role in communicating with people to understand their intentions and provide them with necessary information. Among many technologies, especially in natural language processing, the development of language models has enabled much research in various fields, from natural language processing to natural language understanding, to natural language generation.

2.1 Neural Network-Based Language Models

A language model refers to a model that assigns a probability to a sentence. It is mostly divided into models based on statistical techniques and those based on the neural network. The recent emergence of language models based on the neural network has enabled natural language processing-related technologies' rapid growth. Models such as ELMo [16], BERT [17], and OpenAI GPT [18] are being used in various language processing, showing much better performance than conventional techniques.

The language model achieved the best performance in major benchmarks with the emergence of a pre-learned deep learning-based language model in 2018. ELMo [16] proved that the language model, which first pre-learned vast amounts of data, can perform well in various NLU tasks. ELMo uses the Bi-LSTM architecture to predict which word follows a given word sequence or which word precedes it. However, RNN models such as LSTM feature language models based on the Transformers structure because of the limitation that the longer the sentence, the slower the calculation speed, and the longer the distance, the less accurate the relationship can be expressed. GPT-1 [18] presents a forward language model using the decoder of Transformers [19].

Both ELMo and GPT were shown to be the best performance on the benchmark at the time of their appearance, but the forward or reverse model's structural nature is highly likely to make false predictions. BERT, a two-way model that simultaneously looks at the targeted word's front and back words, wants to solve this problem. The performance was greatly improved by proposing a two-way pre-trained language model using only the encoder of Transformers. With the unveiling of deep learning-based language models such as ELMo, GPT, and BERT, the natural language processing field is developing rapidly [20–22].

2.2 Conversational AI

Conversational AI is an area encompassing many areas of natural language processing. The combination of NLU, which can understand a person's language, and several technologies in the NLG field that create a new natural language can be used to implement interactive artificial intelligence that can interact with people. With the rapid development of deep learning-based pre-learning language models, research on interactive artificial intelligence technology is also actively carried out [1,3,7].

A chatbot is a model that can interact with people most by applying interactive artificial intelligence technology. Recently, they have also learned the knowledge base or knowledge graph to improve chatbot performance. It is used to understand a person's intentions and to learn necessary information [6,7].

However, many existing chatbots are based on specific purposes such as delivering information and performing missions, so they often fail to respond appropriately to human speech's emotional information. In conversation, emotions are used as a significant factor. There is a limitation that even the same sentence often implies different meanings depending on the emotion, so if the dialogue system can't find emotional information, it can't give a proper answer [11,23–26].

2.3 PolarisX

A knowledge graph is a graph that shows the relational meaning of a word by linking words to words. Knowledge graphs are relationship-based, word-based, and can be used as a technology to help computers deduce words as humans perceive them in different contextual meanings [27,28]. There are many different types of knowledge graphs, such as WordNet [12], YAGO [13], and ConceptNet [14].

However, many of the existing knowledge graphs are limited to not being able to continue to expand because they are based only on data for specific languages, such as English, or on existing wiki data. PolarisX [15] is an auto-expanded knowledge graph that improves these thresholds. Continue to collect new data, extract new relationships from the collected data, and then expand them. We want to use PolarisX to improve the ability to identify people's intentions and emotions in chatbots.

We propose an emotion-based chatbot using a pre-trained language model and PolarisX, an automatically extended knowledge graph. To improve the limitations of models that utilize only existing language models or emotional information, PolarisX is used to propose chatbot models that can communicate based on human emotions and actions.

3 EP-Bot

Since people acquire various information through dialogue, artificial intelligence technology must also be able to create the next conversation using the information they understand and grasp to learn by themselves. A person can help his or her work as a personal secretary or wants to talk naturally like a friend, mainly through conversations with conversational AI. However, many existing studies research models by classifying them as secretary and daily conversation.

Conversations to people are not just for information acquisition but also have various meanings such as emotional exchange and interaction. Since conversational AI permeates everyday life for functional performance, chatbots may not understand conversations that are not goals to be carried out or give out incorrect answers. As shown in Fig. 1, the existing language model-based chatbot generates an inappropriate response when a person attempts an emotional conversation. It is only learned based on a typical talk or generates a reply using the similarity between sentences. However, people often have emotional conversations. To this end, we would like to propose a chatbot that enables emotion-based conversations.

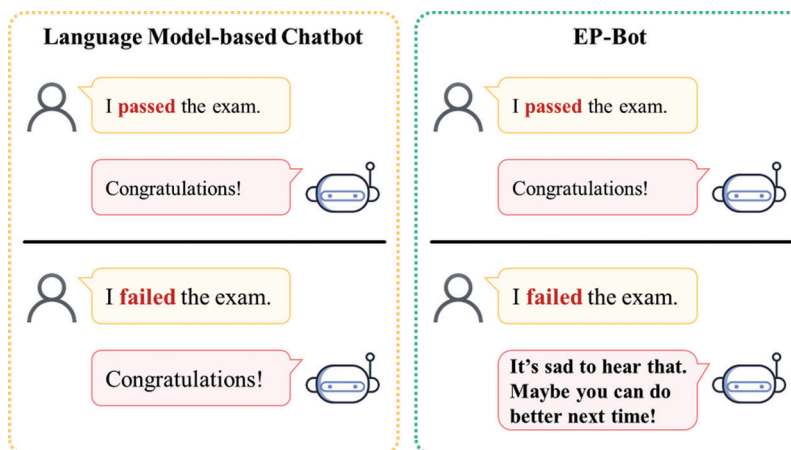


Figure 1: Motivating example

We propose a chatbot model that serves as a secretary to obtain the necessary information through dialogue and enables everyday conversation. Using PolarisX, a knowledge graph that automatically expands, it delivers accurate information to users. It suggests an emotional-based conversation creation model, EP-Bot, utilizing the information inherent in the conversation, such as emotion and intent of the conversation.

3.1 Overview

EP-Bot is an emotion-based chatbot that enables both information acquisition and daily conversation. The various elements inherent in the conversation should be identified and utilized to provide information and to generate appropriate answers according to one’s emotions. To this end, we use PolarisX, an automatically extended knowledge graph, and build a model that detects emotions and intentions, and finally produces the next utterance.

Fig. 2 shows the overall structure of the EP-Bot. It extracts extractable knowledge information from the user’s utterance using PolarisX, an automatically extended knowledge graph. The appraisal extracted from the emotion-act extraction model and knowledge information embedding is concatenated in act embedding and used as input of the model for generating the next utterance. Open AI GPT model is used to create conversation.

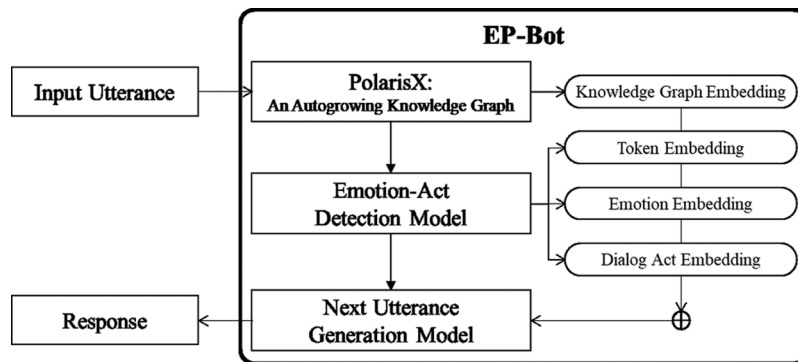


Figure 2: The overall architecture of EP-bot

3.2 Emotion-Act Detection Model Using PolarisX

In order to use the information inherent in the conversation to give an appropriate answer, the model must first be able to extract the information. There is a lot of information embedded in the conversation, but we selectively figure out and utilize some information. We extract information about relationships with entities in sentences to provide information that we want to get from an utterance of a person. We also extract dialogue action and emotion information to respond appropriately to a person’s intentions and help identify their emotions.

A person’s common sense can be expressed as relationships between different objects. The link between the two entities and the relationship information between them allows knowledge to be represented graphically. Many existing knowledge bases have linked each object to a relationship to create a knowledge graph. Still, they have limitations because they have a lot of data in a particular language or cannot respond to new knowledge. PolarisX is an automatically extended knowledge graph that extracts relationships from news data and links them to knowledge graphs to cope with new words as well as various languages. Automatically extended knowledge graphs enable the delivery of more meaningful knowledge information to people.

Fig. 3 shows the input representation used in the model. Token embedding using the existing WordPiece, PolarisX embedding, emotion embedding, and act embedding are all concatenated and used as input for the model. Auto-growing knowledge graph, PolarisX is used as an embedding containing knowledge in EP-bot. After tokenization of the user’s utterance, we use the corresponding embedding value from PolarisX. ALBERT [29] model is utilized for the training model.

Input Utterance	[CLS]	I	am	worried	about	something	[SEP]	What	is	that	[SEP]
KG Embedding	E _{UNK}	E _I	E _{am}	E _{worried}	E _{about}	E _{something}	E _{UNK}	E _{what}	E _{is}	E _{that}	E _{UNK}
Emotion Embedding	E _{UNK}	E _{fear}	E _{fear}	E _{fear}	E _{fear}	E _{fear}	E _{UNK}	E _{happy}	E _{happy}	E _{happy}	E _{UNK}
Dialog Act Embedding	E _{UNK}	E _{inform}	E _{inform}	E _{inform}	E _{inform}	E _{inform}	E _{UNK}	E _{question}	E _{question}	E _{question}	E _{UNK}
Token Embedding	E _{UNK}	E _I	E _{am}	E _{worried}	E _{about}	E _{something}	E _{UNK}	E _{what}	E _{is}	E _{that}	E _{UNK}

Figure 3: Input representation

3.3 Next Utterance Generation Model

In EP-Bot, a model that utilizes all of the previous extracts of PolarisX embedding, emotion embedding, and act embedding to generate the next appropriate utterance is very important. We use the OpenAI GPT model [18] to create the next utterance.

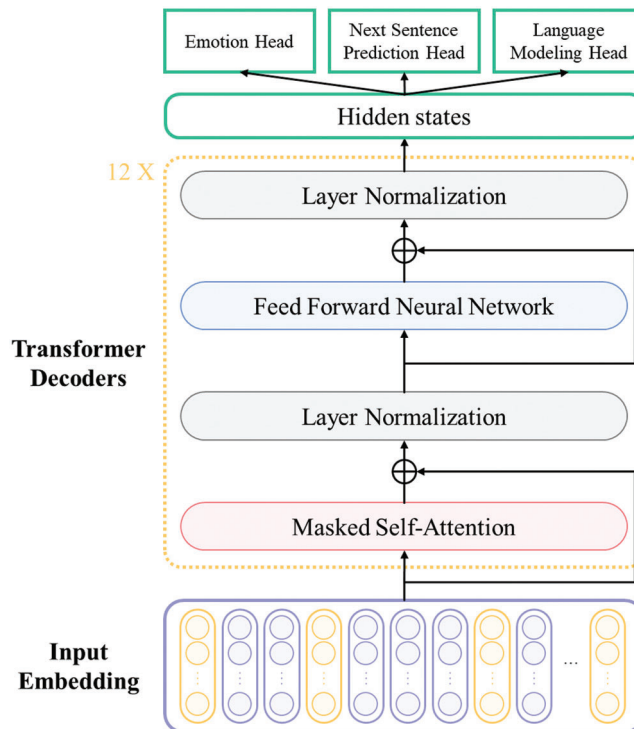


Figure 4: Architecture of next sentence generation model

OpenAI GPT model is a compelling pre-trained language model proposed by OpenAI that can perform various NLP tasks. It is proposed to address the limitations of the prior state-of-the-art NLP models trained using supervised learning. It uses a multi-layer Transformer decoder [19] and applies a multi-head self-attention and feed-forward neural network layers for unsupervised pre-training. It shows the advantage of generative pre-training [18].

Fig. 4 is the architecture of an Open AI GPT-based model to generate the next sentence. Input embedding uses knowledge graphs, emotions, dialog act, token embedding, and embedding for segmentation to distinguish the order of conversations, which is used by concatenating each token location. We use the decoder of Transformers [19] to generate the next sentence. We calculate the result by stacking 12 decoder blocks of Transformers, obtain the loss value for each head for the next emotion classification, head for predicting the next sentence, and head for language modeling, and then obtain the final loss value.

4 Experiments

We propose EP-Bot, an emotion-based chatbot. EP-Bot utilizes PolarisX, an automatically extended knowledge graph, to better understand the emotions and intentions inherent in the conversation. We use extracted embedding values to generate the next utterance. We experiment with the proposed model to verify its performance.

To verify the proposed model, we use the server with Ubuntu 18.04, AMD Ryzen Threadripper 1950X 16-Core processor, and 125 GB RAM. We also experiment on the GPU of NVIDIA GeForce RTX 2080 SUPER for training the model.

4.1 Dataset

Implementing and experimenting with emotion-based chatbots requires datasets that contain emotional data, not just conversational data. We utilize DailyDialog [30] dataset, and it consists of 13,118 multi-turn conversations. Each conversation has two to seven turns of utterances talked by two speakers. Each turn contains the said utterance, emotion, and act data. The original DailyDialog dataset is provided by creating different files for each conversation set, including emotion, act, and an interception. However, we preprocess and use each information as single data for the convenience of implementation and experimentation of the model.

4.2 Evaluation of Emotion-Act Detection Model

EP-bot analyzes emotion and intention for a given utterance and thus generates the next dialogue. For this purpose, when an utterance is given, it extracts the given utterance's emotion and intention first. Therefore, it should be possible to classify the emotions and intentions of the speaker more accurately. We conduct an accuracy experiment on the emotion-act detection model used in the EP-bot. We train the model with the train set of the DailyDialog dataset and verify it using the validation set. Also, to validate knowledge graph embedding, a model without knowledge graph and a model based on knowledge graph embedding have been experimented with.

Tab. 1 shows the results of a comparative experiment. We have verified the validity of knowledge graphs and conducted experiments for comparison between models. We can see that the knowledge graph embedding appears relatively more accurate when used than when not used. Although the results of the emotion and act vary slightly depending on the model used, it can be seen that the use of knowledge graph embedding improves performance when compared with the mean. We extract the Emotion and Act by using the model with the best results.

Table 1: Results Of emotion-act detection model

Model	Emotion-F1	Act-F1	Average-F1
Baseline			
Electra-small	83.12	82.95	83.03
BERT-base	83.44	82.48	82.96
ALBERT-base	83.42	82.70	83.06
With KG Embedding			
Electra-small	83.51	84.07	83.79
BERT-base	84.09	84.04	84.06
ALBERT-base	85.64	84.65	85.14

4.3 Evaluation of Next Utterance Generation Model

The EP-bot we propose utilizes knowledge graphs, emotions, and dialog act. Since the EP-bot is a conversational chatbot, the performance of the model that generates the next conversation about a given utterance is essential. However, it is challenging to select evaluation criteria because evaluating dialog generation models is ambiguous. We use hit ratio, perplexity, F1-score, and BLEU score, commonly used by existing dialog generation models as evaluation metrics.

Hit@1: It is a metric indicating top 1 accuracy.

Perplexity (PPL): It is a metric to assess how well the model predicted the correct dataset's next utterance. PPL can represent the reciprocal of the probability for the test data normalized by the number of words, indicating the degree of confusion. Thus, the lower the PPL value, the better the performance of the language model.

F1-score (F1): We calculate F1-score at token-level. We measure the F1 core values based on the predicted and actual values in token units.

BLEU score (BLEU): BLEU, short for Bilingual Evaluation Understudy, is a metric that measures performance by comparing how similar machine-generated text is to man-generated golden labels. Based on N-gram, higher performance indicates better performance.

We conduct comparative experiments with existing studies to evaluate the performance of the proposed model. The experiment was carried out using a model using Seq2Seq [31], Transformer [19]-based models, and OpenAI GPT [18] mostly used in NLG. These models are models that exclude emotional information and simply generate dialogue. Although there are not many models that generate conversations, including emotions, we have used EmpTransfo [32], which is most similar to the model we propose, in comparison experiments.

Tab. 2 shows the results of the comparative experiments of the dialog generation model. The higher the Hit@1, F1, and BLEU scores, the lower the PPL, the better the model. In the comparative experiment, we would like to include the knowledge graph and the final proposed model EP-bot, except for dialog act and knowledge graph in the EP-bot, to demonstrate that the embedding values of the knowledge graph, emotion, and dialog act affect performance improvement. Besides, we can see that the Hit@1 and PPL evaluations perform better than previous studies.

The next utterance generation model in EP-bot shows some improvement over existing studies by utilizing knowledge graphs, emotions, and actions. However, F1 scores and BLEU scores show

somewhat lower results than previous studies. Because the existing Seq2Seq and Transformer models focus only on giving the correct answer, regardless of one’s feelings, indicators such as F1 and BLEU can show good results. EP-bot is more focused on people’s feelings, so evaluating the accuracy of the sentences may be somewhat lower. Still, it is meaningful enough to show high performance in perplexity, which is used as a key evaluation index.

Table 2: Results on next utterance generation model

Model	Hit@1	PPL	F1	BLEU
Models without emotion				
Seq2Seq + Attention	9.41	129.3	10.22	5.58
Transformer ranker	17.20	–	26.37	15.79
OpenAI GPT without emotion	75.01	10.19	18.2	3.755
Models with emotion				
EmpTransfo	77.25	10.63	19.39	3.99
EmpTransfo + action + topic	78.47	9.04	17.27	2.45
EP-bot without KG, act	71.49	9.90	14.43	3.70
EP-bot	78.69	8.22	18.52	4.23

5 Conclusion

Conversational AI is a convergence technology that reduces the gap between computer and human interaction so that computers can communicate like humans. With advances in technologies such as big data, machine learning, and natural language processing, conversational AI is also developing rapidly. However, existing chatbots are mainly designed to carry out specific tasks, so there is a limitation in general conversations that they do not understand at all what is inherent in a person’s speech or give a wrong answer. For a chatbot to talk more like a person, it is necessary to identify and utilize the emotions and intentions contained in the dialogue.

We propose a chatbot based on the emotions inherent in a person’s utterance. In particular, PolarisX, a knowledge graph that automatically scales to improve the ability to identify emotions and intentions. Knowledge graphs can help computers understand the commonsense contained in conversations. EP-bot, a model that extracts emotions and intentions inherent in ignition based on knowledge graphs, and produces the following sentences based on them, was proposed and verified by experiments.

Computers need to understand people more and communicate like humans to learn their emotions and intentions. The EP-bot, which we propose, is expected to be used in various ways as a model that uses knowledge graphs to enable emotional and intentional conversations.

Still, conversations contain a variety of information and the relationship, emotion, and action information we use. In future work, we would like to extract the various information contained in the conversation and find the features that affect the performance of the dialogue model to improve the performance of the EP-Bot proposed. We also establish an environment to enable peoples to communicate with the EP-Bot in a real dialogue environment.

Funding Statement: This research was supported by Basic Science Research Program through the NRF (National Research Foundation of Korea), the MSIT (Ministry of Science and ICT), Korea, under the National Program for Excellence in SW supervised by the IITP (Institute for

Information & communications Technology Promotion) and the Gachon University research fund of 2019 (Nos. NRF2019R1A2C1008412, 2015-0-00932, GCU-2019-0773).

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

- [1] A. Ram, R. Gabriel, M. Cheng, A. Wartick, R. Prasad *et al.*, “Conversational AI: The science behind the Alexa prize,” *arXiv preprint*, arXiv:1801.03604, pp. 1–18, 2018.
- [2] A. P. Chaves and M. A. Gerosa, “How should my chatbot interact? A survey on human-chatbot interaction design,” *arXiv*, arXiv:1904.02743, pp. 1–44, 2019.
- [3] J. Gao, M. Galley and L. Li, “Neural approaches to conversational AI,” *Foundations and Trends® in Information Retrieval*, vol. 13, no. 2–3, pp. 127–298, 2019.
- [4] S. Poria, N. Majumder, R. Mihalcea and E. Hovy, “Emotion recognition in conversation: Research challenges, datasets, and recent advances,” *arXiv*, vol. 7, pp. 100943–100953, 2019.
- [5] B. Bogaerts, E. Erdem, P. Fodor, A. Formisano, G. Ianni *et al.*, “Open domain question answering and commonsense reasoning,” *Electronic Proceedings in Theoretical Computer Science*, vol. 306, pp. 396–402, 2019.
- [6] R. Yan, “Chitty-chitty-chat bot’: Deep learning for conversational AI,” *IJCAI Int. Joint Conf. on Artificial Intelligence*, vol. 2018, no. July, pp. 5520–5526, 2018.
- [7] M. Zaib, Q. Z. Sheng and W. Emma Zhang, “A short survey of pre-trained language models for conversational AI—A new age in NLP,” *ACM Int. Conf. Proc. Series*, Melbourne, VIC, Australia, pp. 1–4, 2020.
- [8] K. R. Chowdhary, “Natural language processing,” *Fundamentals of Artificial Intelligence*, New Delhi: Springer, pp. 603–649, 2020.
- [9] J. Hirschberg and C. D. Manning, “Advances in natural language processing,” *Science (80-.)*, vol. 349, no. 6245, pp. 261–266, 2015.
- [10] B. G. Smith, S. B. Smith and D. Knighton, “Social media dialogues in a crisis: A mixed-methods approach to identifying publics on social media,” *Public Relations Review*, vol. 44, no. 4, pp. 562–573, 2018.
- [11] E. W. Pamungkas, “Emotionally-aware chatbots: A survey,” *arXiv*, arXiv:1906.09774, 2019.
- [12] G. A. Miller, “WordNet: A lexical database for English,” *Communications of the ACM*, vol. 38, no. 11, pp. 39–41, 1995.
- [13] F. M. Suchanek, G. Kasneci and G. Weikum, “Yago: A core of semantic knowledge,” *Proc. of the 16th Int. Conf. on World Wide Web-WWW’07*, Banff, Alberta, Canada, pp. 697–706, 2007.
- [14] C. Havasi, R. Speer and J. Alonso, “Conceptnet: A lexical resource for common sense knowledge,” *Recent Advances in Natural Language Processing V: Selected Papers from RANLP 2007*, Amsterdam, Netherlands: John Benjamins Publishing, pp. 269–280, 2009.
- [15] S. Y. Yoo and O. R. Jeong, “Automating the expansion of a knowledge graph,” *Expert Systems with Applications*, vol. 141, pp. 112965, 2020.
- [16] M. Peters, M. Neumann, M. Iyyer, M. Gardner, C. Clark *et al.*, “Deep contextualized word representations,” *Proc. of the 2018 Conf. of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, New Orleans, Louisiana, vol. 1, pp. 2227–2237, 2018.
- [17] J. Devlin, M. W. Chang, K. Lee and K. Toutanova, “BERT: Pre-training of deep bidirectional transformers for language understanding,” *NAACL HLT, 2019 Conf. of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies-Proc. of the Conf.*, vol. 1, pp. 4171–4186, 2019.
- [18] M. E. Peters, M. Neumann, M. Iyyer, M. Gardner, C. Clark *et al.*, “Improving language understanding by generative pre-training,” *Open AI*, pp. 1–10, 2018. [Online]. Available: https://cdn.openai.com/research-covers/language-unsupervised/language_understanding_paper.pdf.

- [19] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones *et al.*, “Attention is all you need,” *Advances in Neural Information Processing Systems*, vol. 2017, no. December, pp. 5999–6009, 2017.
- [20] D. W. Otter, J. R. Medina and J. K. Kalita, “A survey of the usages of deep learning for natural language processing,” *IEEE Transactions on Neural Networks and Learning Systems*, pp. 1–21, 2020.
- [21] K. Jing and J. Xu, “A survey on neural network language models,” *arXiv*, arXiv:1906.03591, 2019.
- [22] F. Almeida and G. Xexéo, “Word embeddings: A survey,” *arXiv*, arXiv:1901.09069, 2019.
- [23] L. Zhou, D. Li, J. Gao and H. Y. Shum, “The design and implementation of XiaoIce, an empathetic social chatbot,” *arXiv*, arXiv:1812.08989, 2018.
- [24] M. Karna, D. S. Juliet and R. C. Joy, “Deep learning based text emotion recognition for chatbot applications,” *Proc. of the 4th Int. Conf. on Trends in Electronics and Informatics, ICOEI 2020*, Tirunelveli, India, pp. 988–993, 2020.
- [25] M. C. Lee, S. Y. Chiang, S. C. Yeh and T. F. Wen, “Study on emotion recognition and companion Chatbot using deep neural network,” *Multimedia Tools and Applications*, vol. 79, no. 27–28, pp. 19629–19657, 2020.
- [26] C. H. Kao, C. C. Chen and Y. T. Tsai, “Model of multi-turn dialogue in emotional chatbot,” *Proc.—2019 Int. Conf. on Technologies and Applications of Artificial Intelligence, TAAI 2019*, Kaohsiung, Taiwan, pp. 5–9, 2019.
- [27] Q. Wang, Z. Mao, B. Wang and L. Guo, “Knowledge graph embedding: A survey of approaches and applications,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 29, no. 12, pp. 2724–2743, 2017.
- [28] E. Lisa and W. Wolfram, “Towards a definition of knowledge graphs,” *Joint Proc. of the Posters and Demos Track of 12th Int. Conf. on Semantic Systems—SEMANTiCS2016 and 1st International Workshop on Semantic Change & Evolving Semantics (SuCCESSI6)*, Leipzig, Germany, pp. 1–4, 2016.
- [29] Z. Lan, M. Chen, S. Goodman, K. Gimpel, P. Sharma *et al.*, “ALBERT: A lite BERT for self-supervised learning of language representations,” *arXiv*, arXiv:1909.11942, 2019.
- [30] Y. Li, H. Su, X. Shen, W. Li, Z. Cao *et al.*, “DailyDialog: A manually labelled multi-turn dialogue dataset[†],” *Proc. Eighth Int. Joint Conf. on Natural Language Processing*, Taipei, Taiwan, pp. 986–995, 2017.
- [31] I. Sutskever, O. Vinyals and Q. V. Le, “Sequence to sequence learning with neural networks,” *Proc. of the 27th Int. Conf. on Neural Information Processing Systems*, vol. 2, pp. 3104–3112, 2014.
- [32] R. Zandie and M. H. Mahoor, “EmpTransfo: A multi-head transformer architecture for creating empathetic dialog systems,” *arXiv*, arXiv:2003.02958, 2020.