

Generative Adversarial Network 학습을 통한 감정 텍스트 생성에 관한 연구

김우성*, 김현철*

*고려대학교 컴퓨터학과

e-mail:adventure2165@korea.ac.kr

A Study on the Emotional Text Generation using Generative Adversarial Network

Woo-seong Kim*, Hyeoncheol Kim*

*Dept. of Computer science and Engineering, Korea University

요 약

GAN(Generative Adversarial Network)은 정해진 학습 데이터에서 정해진 생성자와 구분자가 서로 각각에게 적대적인 관계를 유지하며 동시에 서로에게 생산적인 관계를 유지하며 가능한 긍정적인 영향을 주며 학습하는 기계학습 분야이다. 전통적인 문장 생성은 단어의 통계적 분포를 기반으로 한 마르코프 결정 과정(Markov Decision Process)과 순환적 신경 모델(Recurrent Neural Network)을 사용하여 학습시킨다. 이러한 방법은 문장 생성과 같은 연속된 데이터를 기반으로 한 모델들의 표준 모델이 되었다. GAN은 표준모델이 존재하는 해당 분야에 새로운 모델로써 다양한 시도가 시도되고 있다. 하지만 이러한 모델의 시도에도 불구하고, 지금까지 해결하지 못하고 있는 다양한 문제점이 존재한다. 이 논문에서는 다음과 같은 두 가지 문제점에 집중하고자 한다. 첫째, Sequential 한 데이터 처리에 어려움을 겪는다. 둘째, 무작위로 생성하기 때문에 사용자가 원하는 데이터만 출력되지 않는다. 본 논문에서는 이러한 문제점을 해결하고자, 부분적인 정답 제공을 통한 조건별 생산적 적대 생성망을 설계하여 이 방법을 사용하여 해결하였다. 첫째, Sequence to Sequence 모델을 도입하여 Sequential 한 데이터를 처리할 수 있도록 하여 원시적인 텍스트를 생성할 수 있게 하였다. 둘째, 부분적인 정답 제공을 통하여 문장의 생성 조건을 구분하였다. 결과적으로, 제안하는 기법들로 원시적인 감정 텍스트를 생성할 수 있었다.

1. 서론

생산적 적대 신경망(Generative Adversarial Network)은 정해진 학습 데이터에서 정해진 생성자와 구분자가 서로 각각에게 적대적인 관계를 유지하며 동시에 서로에게 생산적인 관계를 유지하며 학습하는 기계학습 분야이다. 학습 과정이 기존의 군집화와 다르게 정해진 학습 데이터 안에서 생성자와 구분자의 생산적 적대 경쟁의 균형을 맞추는 것이 추가 된다는 것이 기존의 비지도 학습과 다르게 비지도 학습의 새로운 분야로 구분되며 이러한 점 때문에 GAN은 이미지 변환, 음성신호, 자연어처리 등 다양한 분야에 성과를 내고 있다.

하지만 GAN 모델은 문장 생성에 한해서 모델이 가지는 몇 가지 문제점이 존재한다.

1. 문장과 같은 Sequential한 데이터에는 적합하지 않다.
2. 무작위로 생성하기 때문에 사용자가 원하는 데이터만 출력되지 않는다.

본 논문에서는 GAN 모델에 Sequence to Sequence[1] 모델 적용 및 Conditional 한 데이터 생성 알고리즘을 적용하여 GAN 모델을 사용하여 학습할 때 발생할 수 있는 2가지 문제를 해결하고자 한다. Sequence to Sequence 모

델은 문장 생성과 같이 Sequential 한 데이터에 적합하기 때문이다. 각각의 문제점을 해결하기 위해 본 논문에서는 다음의 방법을 제안한다.

1. Recurrent Neural Network 기반의 Seq2Seq이 적용된 SeqGAN 모델을 사용하여 Sequential 한 데이터에서 적용될 수 있도록 한다.
2. 학습 시에 정답 또한 제공함으로써 어느 하나의 답에만 치우치게 하는 것이 아니라 주어진 답에 맞추어서 생성할 수 있도록 Conditional 하게 설계된 Conditional SeqGAN 모델을 사용한다.

본 논문에서는 이러한 제안 방법을 확인할 수 있는 원시 문장 생성 환경을 구축하였고, 각각에 제안 방법에 대해서 실험하였다. 실험 결과를 통한 기여는 다음과 같이 요약할 수 있다.

- Conditional SeqGAN을 통한 특정 감정세기에 맞는 원시 텍스트 생성 모델 개발

본 논문의 구성은 다음과 같다. 제 2장에서는 관련 연구에 대해서 다루며 제 3장에서는 제안하는 방법 및 모델에 대

해서 설명한다. 제 4장에서는 실험 설계 및 과정에 대해 다루었다. 제 5장에서는 실험 결과를, 마지막으로 제 6장에서는 결론 및 향후 연구에 대해 기술하였다.

2. 관련 연구

GAN은 Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair*, Aaron Courville, Yoshua Bengio[2]가 최초로 제안한 개념으로, Generator와 Discriminator가 서로 적대적인 관계를 유지하면서 생산적인 관계 또한 유지하며 서로 경쟁을 통해 학습을 진행하는 신경망을 의미한다. 여기서 Generator는 최대한 Discriminator를 속여 자신이 직접 만든 결과물을 최대한 Discriminator가 자연스럽다고 여길 수 있도록 출력하는 것을 목표로 학습을 진행한다. 반대로 Discriminator는 최대한 Generator가 생산한 결과물을 부자연스럽다고 판단할 수 있도록 검출해 내는 것을 목표로 학습을 진행한다. 이러한 적대적이면서도 생산적인 관계를 통하여 GAN은 최대한 자연스러운 결과물을 학습하여 출력하도록 설계되어 있다.

GAN에 Sequence한 데이터를 사용할 수 있도록 적용하는 초기 연구 중 하나는 Lantao Yu, Weinan Zhang, Jun Wang, Yong Yu[3]가 Generator를 LSTM과 강화학습을 사용하여 대처한 SeqGAN연구이다. 기존의 GAN은 Continuous한 데이터를 대상으로 하였으며 이는 즉 문장과 같이 Discrete한 데이터는 사용하기 어렵다는 단점을 해결하기 위하여 Generator에는 LSTM 구조를 사용하여 Discrete한 Token Sequence를 생성하도록 설정하였고, Discriminator는 CNN 구조를 사용하여 각 만들어진 sequence에 대해 판별을 진행하여 reward를 주고 이를 Policy Gradient를 통해서 LSTM을 업데이트 하도록 학습을 진행함으로써 문제를 해결하였다.

Mehdi Mirza, Simon Osindero[4]이 제안한 Conditional Generative Adversarial Nets는 기존의 Generator가 Random Noise를 입력 받아 결과 값을 생성하던 GAN의 단점을 보완한 Conditional GAN을 제안하였다. 이는 학습시 Condition이 없는 Unsupervised Learning에 특정 Condition 데이터를 추가로 학습시켜주는 방법으로 기존의 GAN의 방식을 변경하였다. 이를 통하여 사용자는 원하는 특정 Condition에 부합하는 결과 값을 얻게 되었다. 이 연구를 기반으로 하여 Yi-Lin Tuan and Hung-Yi Lee[5]이 제안한 Conditional Sequence Generative Adversarial Nets는 Conditional GAN의 기법을 SeqGAN에 적용한 모델을 제안하였다.

3. 제안 방법 및 모델

본 장에서는 GAN에서 해결하지 못한 2가지 문제점을 해결하기 위한 제안 방법 및 모델의 구조 및 모델 학습에 대해서 설명한다. 제안한 방법 및 모델에서는 다음과 같은

가설을 세웠다.

1. SeqGAN과 Conditional GAN을 합친 Conditional SeqGAN 모델을 감정 텍스트에 적용한다.

2. 감정 세기라는 Condition을 제한하였을 때 Conditional SeqGAN을 통하여 Condition에 맞게 원시적인 문장을 생성할 수 있다.

위와 같은 가정에 기초하여 본 장에서는 Conditional SeqGAN 모델을 감정 텍스트 데이터에 적용하여 적절한 세기에 맞는 원시적인 문장을 생성하도록 한다.

4. 실험 설계 및 과정

본 장에서는 앞에 정의한 Conditional SeqGAN을 사용하여 실전 데이터에 해당 모델을 적용하면서 산출된 원시 문장들을 확인하고 해당 문장이 적절하게 생성되었는지에 대해 조사 및 평가에 관한 내용을 다룬다. 실험 환경은 모두 Python 3으로 구현하였다.

제안한 방법 및 모델에 적용하기 위해 사용한 데이터는 Semeval 2018 Task 1: Affects in Tweet[6]을 사용하였다. 해당 데이터는 4가지 열로 구성되어 있다. 먼저 첫 번째로 “Tweet”에는 2017년 6월과 7월 사이에 등록된 트윗의 문장들을 사용하였으며 “ID”에는 인덱스 ID를, “Affect Dimension”에는 Anger, Fear, Joy, Sadness 4가지 종류로 트위터 문장들을 구분하며, “Intensity Class”에는 “Affect Dimension”에서 구분된 감정들의 세기를 0에서 3까지로 분리하여, 0에서 가장 낮은 세기를, 3에서 가장 높은 세기를 보인 것으로 구분 되어 있다. 본 실험에서는 4가지 감정들 중 “Affect Dimension”상태 중 “Anger”상태인 데이터만을 사용하여 “Anger”감정을 가졌을 때의 트위터 문장을 “Intensity Class” 구분에 맞게 문장을 생성하도록 데이터를 사용하도록 하였다.

학습데이터 중 문장 데이터는 “Tweet” 열에 존재하며 일반 문장들과 달리 140자의 짧은 문장이며, #(해시태그)와 @(멘션)들이 부가적인 기능으로 존재하여 사람들 간의 소셜 네트워크 활동을 돕는다. 멘션의 경우 다른 사람을 호칭하는 부가적인 기능을 가진 고유한 단어이고, 해시태그는 단어의 검색을 할 수 있다는 부가적인 기능을 소지하고 있는 단어이다. 해시태그의 경우 #이 붙여져 있는 특수 단어라는 점에서 #이 붙어있지 않은 동일한 단어가 존재하더라도 같게 보지 않고 다른 고유한 단어로 간주하였으며 멘션의 경우 크게 사용자를 호출하는 목적으로 사용하는 단어이기 때문에 <user> 태그로 변경하여 주었다.

그 외에도 이모티콘과 같이 단어가 아닌 부분들에 대해서는 먼저 제거해주는 전처리를 진행하였다. 이후 문장들

을 공백을 기준으로 하여 단어별로 모두 나누어 주었다. Tweet 옆에는 문장이 띄어쓰기를 기준으로 하여 단어별로 분리해주었다. 또한 Affect Dimension에는 모두 Anger를 기준으로 하였기 때문에 0으로 변환해주었다. 마지막으로 Intensity Class는 설명을 제거하고 숫자만을 남겨두어 주었다.

주어진 학습 데이터를 사용하여 Conditional SeqGAN[7]은 학습 데이터를 사용하여 생성적 적대 신경망을 통하여 감정세기가 분류된 원시 문장을 생성하게 된다. 감정이 “anger”인 상태에서 임의로 0에서 3까지의 감정세기를 무작위로 총 10000번 입력해 주었고 주어진 세기에 따라 총 10000개의 원시 Tweet문장을 생성하였다.

5. 실험 결과

Conditional SeqGAN의 목표인 감정 세기에 따른 원시 문장생성을 위하여 Conditional SeqGAN을 구현하였다. 해당 모델을 이용하여 0에서 3의 범위 내에서 무작위 자연어 수준의 세기를 입력 값으로 하였을 때 그에 해당되는 텍스트가 생성되었으며 학습을 생성자 10번, 구분자 10번, 생산적 적대 신경망 10번을 진행한 뒤 1만개의 원시 문장을 생성해내었다. 생성된 원시 문장들은 학습한 Tweet 문장들을 최대한 흉내 낸 모습을 취하고 있으나, 대부분 문맥이나 문법을 대체로 맞추지 못하는 경우가 많았다. 이는 GAN이 차후 해결해야할 과제로 파악된다. 하지만 일부 텍스트의 경우에는 자연스러운 경우 또한 존재하였다. 표 1은 GAN이 생성된 원시 문장의 일부를 예시로 보여주고 있다.

Intensity Classes	Tweet
0	<user> <user> <user> what a idea #success id a backs i #kikme
0	me i the to fuming a free massage a backs a about hurt roach was massage #worry
1	<user> <user> <user> was their want in was ass game to feels hell both heat <user> crashing
2	<user> i <user> told <user> glowing a most least in talk eat and just the mind and just this #revenge was #mad #upset players
3	<user> <user> so always trusted #anger in by blooded <user> an absolute piece an affront <user> #furious the

<표 1> 실제 생성된 원시 문장 예시

6. 결론 및 향후 연구

본 논문에서는 GAN을 사용하여 원시 감정 텍스트를 생성하고자 했다. 그에 따라 GAN의 무작위 생성을 해결하기 위해 Conditional GAN을 도입하였고, Sequential한 데이터를 사용하기 위하여 SeqGAN을 도입하여 Conditional SeqGAN으로 원시 문장을 생성하는 실험을 설계하였다.

실험 결과에 기초하여 본 논문에서는 실험 내용에 서술한바 와 같이 Conditional SeqGAN을 사용함으로써 Sequential한 문장 데이터를 활용하여 감정 세기와 같이 조건별로 원시 문장을 생성할 수 있었다.

향후 연구에서는 먼저 Conditional SeqGAN을 통해 생성된 원시 문장이 제대로 생성되었는지 평가하는 것을 우선시 할 것이며, 그 외에 문장이 제대로 문법에 맞추어 완전한 감정 텍스트를 생성할 수 있도록 하는 것이 궁극적인 연구의 목적이 될 것이다.

Acknowledgement

“이 논문은 2017년도 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임 (NRF-2017R1A2B4003558)”

참고문헌

[1] Ilya Sutskever, Oriol Vinyals, Quoc V. Le. “Sequence to Sequence Learning with Neural Networks.” arXiv preprint arXiv:1409.3215v3 (2014).
 [2] Ian, Jean, et al. “Generative Adversarial Nets.” arXiv preprint arXiv:1406.2661 (2014).
 [3] Lantao Yu, Weinan Zhang, Jun Wang, Yong Yu. “SeqGAN: Sequence Generative Adversarial Nets with Policy Gradient” arXiv preprint arXiv:1609.05473v6 (2017).
 [4] Mehdi Mirza, Simon Osindero. “Conditional Generative Adversarial Nets” arXiv preprint arXiv:1411.1784v1 (2014).
 [5] Yi-Lin Tuan and Hung-Yi Lee. “Improving Conditional Sequence Generative Adversarial Networks by Stepwise Evaluation” arXiv preprint arXiv:1808.05599 (2018)
 [6] Saif M. Mohammad, Felipe Bravo-Marquez, Mohammad Salameh, and Svetlana Kiritchenko. Semeval-2018 Task 1: Affect in tweets. In Proceedings of International Workshop on Semantic Evaluation (SemEval-2018), New Orleans, LA, USA.(2018)
 [7] “Implementation of Sequence Generative Adversarial Nets with Policy Gradient”, Github, last modified 11 Apr 2018, accessed DEC 21 2019, <https://github.com/LantaoYu/SeqGAN>