

심층 합성곱 생성적 적대 신경망을 이용한 전력 데이터 생성 및 정확도 측정

김태근, 임세현, 윤성국
 송실대학교

Generating Power Data and Performance analyzing Using Deep Convolutional Generative Adversarial Network

Kim-Tae Geun, Lim-Se Heon, Yoon-Sung Guk
 Soongsil University

Abstract - 최근 전력계통에서도 부하예측, 에너지 스케줄링, 전압최적화, 손실 데이터 복원 등에 기계학습 기법들이 활발하게 도입되고 있다. 전력 데이터는 기계학습 기법을 사용하기에 충분하지 않은 양을 가지는 경우가 있다. 본 논문에서는 이러한 문제를 해결하기 위해 심층 합성곱 생성적 적대 신경망(Deep Convolutional Generative Adversarial Network, DCGAN)을 이용하여 가상의 전력 데이터를 생성하는 기법을 제안하였다. 사례연구로서 미국 피칸 스트리트 180일 데이터를 통해 추가 28일의 데이터를 생성하였고 생성된 데이터는 원본 데이터와 높은 수준의 유사성을 가짐을 확인하였다.

데이터를 생성하여 판별기의 입력으로 보내고, 판별기는 실제 데이터와 생성기로부터 만들어진 데이터를 비교하여 진짜 이미지인지 가짜 이미지인지를 구별한다. 생성적 적대 신경망은 이러한 두 개의 신경망이 맞물려 있는 구조를 훈련함으로써, 실제 데이터와 유사한 가짜 데이터를 생성한다. 생성적 적대 신경망은 비교적 적은 데이터를 가지고도 기존의 데이터와 유사한 새로운 데이터를 생성해 낼 수 있는 이점이 있다. [5]

1. 서 론

최근 하드웨어의 발전과 데이터 양의 증가로 인해 여러 분야에서 다양한 인공지능 기법들이 적용하여 성공적인 결과를 얻고 있다. 전력계통 분야에서도 전력 데이터의 실시간 취득 및 저장이 용이해짐으로써, 기계학습 기법을 이용한 부하 예측, 에너지 스케줄링, 전압최적화, 손실 데이터 복원 방법 등의 연구가 진행되었다. [1]

기계학습 기법이 성공적인 결과를 얻기 위해서는 양질의 다량의 데이터가 필수적이다. 그러나 전력 데이터는 간혹 충분하지 않은 데이터 양을 가지고 예측, 최적화 등을 수행해야 하는 경우가 발생한다. 이러한 경우 높은 수준의 결과를 내지 못할 가능성이 있다.

생성적 적대 신경망(Generative Adversarial Network, GAN)은 생성기와 판별기 두 개의 인공지능망을 사용하여 원본 데이터와 유사한 데이터를 생성하는 것을 목표로 하는 기계학습 기법이다. 생성적 적대 신경망을 이용한 선행 연구로는, 생성적 적대 신경망을 이용한 신재생 에너지 발전 시나리오 예측 [2], 생성적 적대 신경망을 이용한 PMU 이상 데이터 측정 [3], 생성적 적대 신경망을 이용한 날씨 및 태양광 발전 예측 [4] 등이 있다.

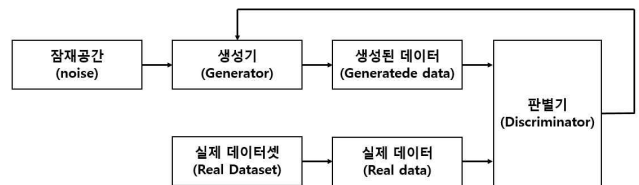
본 논문에서는 최신의 생성적 적대 신경망 기법인 심층 합성곱 생성적 적대 신경망(Deep Convolutional Generative Adversarial Network, DCGAN)[5]을 이용하여 전력 데이터를 생성하였다. 미국 피칸 스트리트[6]의 2019년 5월부터 10월의 평일 180일 데이터(32모선 모델)를 사용하여 28일의 새로운 데이터를 생성하였다. 생성된 데이터의 성능 측정방법으로, 생성적 적대 신경망의 성능 측정 지수인 FID(Frechet Inception Distance)와 상호상관(Cross-Correlation) 지수를 사용하여 생성된 데이터와 기존의 데이터의 유사성을 측정하였고 제안하는 기법을 통해 생성된 데이터는 원본 데이터와 높은 수준의 유사성을 가지는 것을 확인하였다.

2. 본 론

2.1 생성적 적대 신경망(Generative Adversarial Network, GAN)

생성적 적대 신경망(Generative Adversarial Network, GAN)은 생성기와 판별기의 두 개의 신경망으로 구성된 심층 신경망이다. 이 두 개의 신경망이 서로를 향하게 하고, 서로 대항하듯이 훈련하게 함으로써, 결과적으로 생성 모델(Generative model)을 만들어 낸다. 생성기(Generator)는 잠재공간(잠음)으로부터

2.1.1 생성기(Generator)



〈그림 1〉 생성적 적대 신경망의 구조

생성기는 잠재공간으로부터 실제 데이터와 같은 사이즈의 데이터를 만들어내는 신경망으로 생성기에서 생성된 데이터를 판별기의 입력으로 보내고, 생성된 이미지는 판별기에 의해 진짜 데이터인지, 가짜 데이터인지 분류된다. 생성기의 손실 함수는 다음과 같이 정의된다.

$$Loss_G = -E_z[D(G(z))] \quad (1)$$

z 는 잠재공간의 잡음(noise)이고 $G(x)$, $D(x)$ 는 각각 생성기, 판별기의 함수를 의미한다. 생성기는 생성한 데이터를 판별기가 진짜 데이터인지, 만들어진 데이터인지 구분할 수 없는 것을 목표로 학습을 진행한다.

2.1.2 판별기(Discriminator)

판별기는 입력으로 들어온 데이터가 실제 데이터인지, 만들어진 데이터인지 구분해내는 역할을 한다. 입력으로 받은 데이터를 실제 데이터와 비교하여 실제 데이터인지, 만들어진 데이터인지 판별한다. 판별기의 손실함수는 다음과 같다.

$$Loss_D = -E_{real}[D(real)] + E_z[D(G(z))] \quad (2)$$

$real$ 은 실제 데이터를 의미한다. 판별기는 생성기에서 생성한 데이터와 실제 데이터를 잘 구분할 수 있도록 학습한다.

2.1.3 생성기와 판별기의 최대·최소 알고리즘

생성적 적대 신경망에서는 생성기와 판별기 두 개의 신경망이 서로를 향하게 하고, 대항하듯이 훈련한다. 각각의 손실함수를 만족하면서 훈련이 이뤄지기 위하여 다음과 같은 최대·최소 알고리즘을 만족해야 한다.

$$\min_G \max_D V(G, D) = E_{real}[D(real)] - E_z[D(G(z))] \quad (3)$$

$V(G, D)$ 는 생성적 적대 신경망의 가치함수를 의미한다. 최대·최소 알고리즘의 가치함수는 수식(2)에 음수를 곱한 것으로 정의된다. 초기의 훈련단계에서는 $Loss_G$ 와 $V(G, D)$ 의 값은 크고, $Loss_D$ 의 값은 작게 나타난다. 훈련이 점차 진행될수록,

$Loss_G$ 의 값을 최소화하고, $Loss_D$ 의 값을 최대화도록 훈련한다.

2.2 심층 합성곱 생성적 적대 신경망(Deep Convolutional Generative Adversarial Network, DCGAN)

심층 합성곱 생성적 적대 신경망은 기존의 생성적 적대 신경망의 단점인 수렴이 불안정한 문제와 성능 문제를 개선하기 위하여 제안된 기법이다. 기존의 생성적 적대 신경망과 전체적인 구조는 거의 유사하고, 생성기와 판별기의 내부 구조가 다른 특징이 있다. 심층 합성곱 생성적 적대 신경망 기법은 기존의 생성기와 분류기의 레이어를 합성곱 신경망(Convolutional Neural Network, CNN)으로 변경하였고, 생성기와 분류기에서 모두 배치 정규화(Batch-Normalization)를 사용하여 불안정한 수렴과 생성 모델의 정확도를 개선한 기법이다. [5]

2.3 실험 및 결과 평가

본 논문에서는 미국의 5월부터 10월의 180일 평일 유효전력 데이터를 이용하여 28일의 데이터를 생성하고, 생성된 데이터와 실제 데이터간의 각각의 지수를 계산하였다. 실험의 정확도를 도시하기 위하여, 데이터를 군집화한 후 모션별로 분류하여 훈련하였다. 생성된 데이터와 실제 데이터의 유사성을 비교하는 지표로 FID(Frechet Inception Distance)와 상호 상관(Cross-Correlation)을 계산하였다.

2.3.1 Frechet Inception Distance(FID)

FID는 두 데이터 셋 사이의 거리를 계산하는 방법으로 생성적 적대 신경망 기법의 성능평가지표로 사용된다. 실제 데이터와 생성된 데이터의 특징점을 추출한 뒤 두 집합간의 평균(μ)과 공분산(\sum)을 이용하여 데이터 간의 거리를 계산한다. FID 지수의 값이 작을수록 데이터가 비슷하다고 평가된다. FID를 계산하는 수식은 다음과 같다.

$$FID(r, g) = \|\mu_r - \mu_g\|_2^2 + Tr(\sum_r + \sum_g - 2(\sum_r \sum_g)^{\frac{1}{2}}) \quad (4)$$

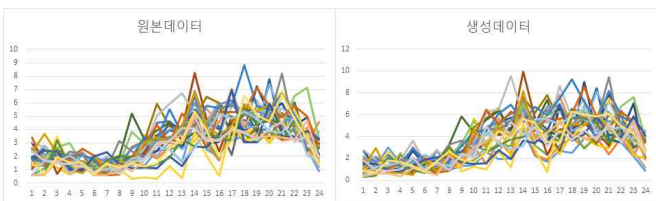
2.3.2 상호상관(Cross-Correlation)

상호상관은 두 변수 간에 어떤 선형적 또는 비선형적 관계를 갖고 있는지를 평가하는 방법이다. 본 논문에서는 피어슨 상관계수를 이용하였다. $cov(X, Y)$ 는 X, Y의 공분산, σ_X, σ_Y 는 X, Y의 표준편차, μ_X, μ_Y 는 X, Y의 평균을 의미한다. 상관계수를 계산하는 수식은 다음과 같다.

$$\text{상관계수}(r_{X,Y}) = \frac{cov(X, Y)}{\sigma_X \sigma_Y} = \frac{E[(X - \mu_X)(Y - \mu_Y)]}{\sigma_X \sigma_Y} \quad (5)$$

일반적으로, 상관계수는 -1에 가까울수록 강한 음적 상관관계, +1에 가까울수록 강한 양적상관관계가 있다고 해석한다.

2.3.3 결과



<그림 2> 분류된 모션 데이터 중 25번 모션의 데이터의 그래프와 훈련결과 그래프

<그림 2>는 25번 모션의 결과를 보여준다. 여기서 x축은 시간, y축은 전력량, 각각의 그래프는 서로 다른 요일을 나타낸다. 시각적으로 판단하기에도, 유사하게 생성됨을 확인할 수 있다.

<표 1>에서는 각 모션의 FID와 상호상관 계수를 도시하였다. FID 지수의 기준을 제시하기 위하여, 군집화한 데이터 중 사용

된 일일 데이터 간의 FID 지수 평균을 계산하였다. FID 지수의 평균값은 15.9925로 25번 모션을 제외하면 <표 1>의 FID 값 보다 훨씬 크다. 즉, 생성적 적대 신경망으로 생성된 데이터는 원본 데이터와 상당히 유사하다는 것을 확인할 수 있다. 상호상관계수 또한 5번 모션을 제외하면 모두 0.85 이상으로 동일하게 생성된 데이터가 기존의 데이터와 매우 유사함을 확인하였다.

<표1>모션별 원본 데이터와 생성된 데이터의 유사성 평가

모션	FID	상호상관계수
1모션	2.3819	0.9212
5모션	15.1288	0.6525
10모션	1.2047	0.8822
15모션	0.2798	0.8680
20모션	1.9363	0.8822
25모션	25.3797	0.9474

3. 결 론

본 논문에서는 미국의 180일 평일 유효전력 데이터로 심층 합성곱 생성적 적대 신경망 기법(DCGAN)을 훈련시켜 새로운 28일의 평일 유효전력 데이터를 생성하고 그 성능을 검증하였다. 생성된 데이터는 실제 데이터와 낮은 FID 값과 1에 가까운 피어슨 상관계수를 가진다. 이는 생성된 데이터는 실제 데이터와 통계적으로 비슷한 특징을 가지는 것을 의미한다. 따라서 심층 합성곱 생성적 적대 신경망 기법을 통해 전력 데이터 부족한 상황에서 추가 데이터를 확보할 수 있음을 확인하였다.

후속 연구로서, 심층 합성곱 생성적 적대 신경망을 이용한 고장 발생데이터, 이상데이터의 분석을 고려할 수 있다.

감사의 글

본 연구는 2020년도 산업통상자원부의 재원으로 한국에너지기술평가원(KETEP)의 에너지인력양성사업으로 지원받아 수행한 인력양성 성과입니다.(No. 20184010201690)

[참 고 문 헌]

- [1] Chao Ren, Student Member, IEEE, and Yan Xu, Senior Member, IEEE, 'A Fully Data-Driven Method Based on Generative Adversarial Networks for Power System Dynamic Security Assessment With Missing Data', IEEE TRANSACTIONS ON POWER SYSTEMS, VOL. 34, NO. 6, NOVEMBER 2019
- [2] Yize chen, Yishen Wang, Daniel Kirschen, Fellow, IEEE, and Baosen Zhang, 'Model-Free Renewable Scenario Generation Using Generative Adversarial Networks', IEEE TRANSACTIONS ON POWER SYSTEMS, VOL 33, NO.3 MAY 2018
- [3] A. Aligholian, A. Shahsavari, E. Cortez, E. Stewart, and H. Mohsenian-Rad, "Event Detection in Micro-PMU Data: A Generative Adversarial Network Scoring Method," in Proc. of the IEEE PES General Meeting, Montreal, QB, Canada, August 2020.
- [4] Fei Wang, Zhanyao Zhang, Chun Liud, Yili Yub, Songling Pange, Neven Duićf, Miadreza Shafie-khah, João P.S. Catalão, "Generative adversarial networks and convolutional neural networks based weather classification model for day ahead short-term photovoltaic power forecasting," Energy Conversion and Management, 2019, 443-462
- [5] Alec Radford, Luke Metz, Soumith Chintala, "Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks", ICLR, 19 Nov 2015 (v1), 2016 ICEE 2019, Hongkong, Jul. 2-6, 2019.
- [6] Pecan Street INC., "Pecan street data." <http://www.pecanstreet.org>.