

<https://doi.org/10.7236/JIIBC.2019.19.5.79>

JIIBC 2019-5-11

# LSTM과 GRU 딥러닝 IoT 파워미터 기반의 단기 전력사용량 예측

## Short-term Power Consumption Forecasting Based on IoT Power Meter with LSTM and GRU Deep Learning

이선민\*, 선영규\*, 이지영\*, 이동구\*, 조은일\*, 박대현\*, 김용범\*, 심이삭\*, 김진영\*\*

Seon-Min Lee\*, Young-Ghyu Sun\*, Jiyoung-Lee\*, Donggu Lee\*, Eun-Il Cho\*,  
Dae-Hyun Park\*, Yong-Bum Kim\*, Isaac Sim\*, Jin-Young Kim\*\*

**요약** 본 연구에서는 Long Short Term Memory (LSTM) 신경망과 Gated Recurrent Unit(GRU) 신경망을 Internet of Things (IoT) 파워미터에 적용하여 단기 전력사용량 예측방법을 제안하고, 실제 가정의 전력사용량 데이터를 토대로 예측 성능을 분석한다. 성능평가 지표로써 Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), Mean Percentage Error (MPE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE)를 이용한다. 실험 결과는 GRU 기반의 모델이 LSTM 기반의 모델에 비해 MAPE 기준으로 4.52%, MPE 기준으로 5.59%만큼의 성능개선을 보였다.

**Abstract** In this paper, we propose a short-term power forecasting method by applying Long Short Term Memory (LSTM) and Gated Recurrent Unit (GRU) neural network to Internet of Things (IoT) power meter. We analyze performance based on real power consumption data of households. Mean absolute error (MAE), mean absolute percentage error (MAPE), mean percentage error (MPE), mean squared error (MSE), and root mean squared error (RMSE) are used as performance evaluation indexes. The experimental results show that the GRU-based model improves the performance by 4.52% in the MAPE and 5.59% in the MPE compared to the LSTM-based model.

**Key Words** : Power Meter, Internet of Things, Deep Learning, LSTM, GRU, Short-term Power Forecasting

### 1. 서론

정확한 전력 예측은 전력 공급용량과 신뢰성이 향상될 수 있다. 동시에 전력공급회사는 전력 관련 건설을 합리적으로 수행하고 예측값에 따라 자원 낭비를 피할 수 있

다<sup>[1]</sup>. 현재 단기 전력 예측은 주로 시계열, 회귀분석, 인공 신경망과 같은 방법을 사용한다<sup>[2, 3]</sup>. 인공 신경망은 비선형 함수에 대한 뛰어난 적합성으로 인해 실용적인 단기 부하 예측방법 중 하나이다.

기본적인 인공 신경망에서 발전한 형태가 순환 신경망

\*준회원, 광운대학교 전자융합공학과

\*\*정회원, 광운대학교 전자융합공학과

접수일자 2019년 9월 2일, 수정완료 2019년 10월 2일

게재확정일자 2019년 10월 4일

Received: 2 September, 2019 / Revised: 2 October, 2019 /

Accepted: 4 October, 2019

\*Corresponding Author: jinyoung@kw.ac.kr

Dept. of Electronic Convergence Engineering, Kwangwoon Univ., Seoul, Korea

이다. 순환 신경망은 매개변수 선택 시 기본적인 인공 신경망과 다르다. 일반적으로, 인공 신경망은 각 계층에서 다른 매개변수를 선택하는 반면, 순환 신경망은 학습 매개변수의 수를 크게 줄일 수 있는 매개변수를 공유한다<sup>[4]</sup>. 그러나 단순한 순환 신경망의 경우 장기 의존성 문제가 있다. 이 문제를 해결할 방안으로 순환 신경망의 변형 구조인 Long Short Term Memory (LSTM) 신경망과 Gated Recurrent Unit(GRU) 신경망이 등장했다. GRU는 LSTM보다 매개변수가 적고 수렴 속도가 빠르며 정확도도 비슷하다<sup>[5]</sup>. 따라서 본 논문은 LSTM 신경망과 GRU 신경망에 대해 알아보고 주커 지역 사회의 단기 전력 예측에 두 가지 신경망을 적용해 성능평가를 진행한다. 본 논문의 구성은 다음과 같다. 2장에서는 딥러닝 알고리즘으로써 순환신경망과 그 순환신경망의 변형 형태인 LSTM 신경망과 GRU 신경망에 대해 알아본다. 이어지는 3장에서는 본 연구에서 진행한 실험 데이터와 실험 파라미터를 알아본다. 또한, 실험을 통해 나온 실험 결과를 분석하고 5가지 지표들로 각각 성능평가를 진행한다. 마지막으로 4장에서는 논문의 내용을 정리하고 진행 가능한 추후 연구에 대해 제시하며 본 논문을 마무리 짓는다.

## II. 딥러닝 알고리즘

### 1. 순환 신경망

순환 신경망은 기존의 전방 전달 신경망에서 내부가 순환하는 구조로 확장된 신경망이다. 순환 구조는 정보가 지속적으로 전달되는 것을 도와서 시계열 데이터 처리에 좋은 성능을 보인다. 순환 신경망의 구조는 그림 1과 같다. 그림 1의 등호 기준으로 왼쪽 그림은 신경망 A가 입력  $x_t$ 을 받아  $h_t$ 을 출력하고 루프를 통해 이전 정보가 네트워크의 다음 단계로 전달되는 순환 신경망의 기본 구조를 의미한다<sup>[6]</sup>. 기본 구조를 시간의 흐름에 따라 펼치면 그림 1의 등호 기준으로 오른쪽 그림처럼 표현된다. 그러나 이런 기본 구조의 순환 신경망은 데이터의 정보가 길어질수록 정보의 연결 방법을 학습할 수 없는 '장기 의존성 문제'가 나타난다<sup>[6]</sup>. 이 문제를 해결한 신경망이 LSTM과 GRU이다.

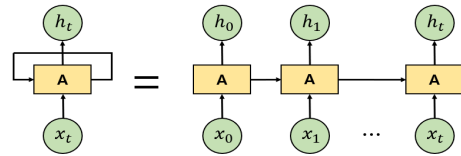


그림 1. 순환 신경망의 구조.  
Fig. 1. Structure of Recurrent Neural Network.

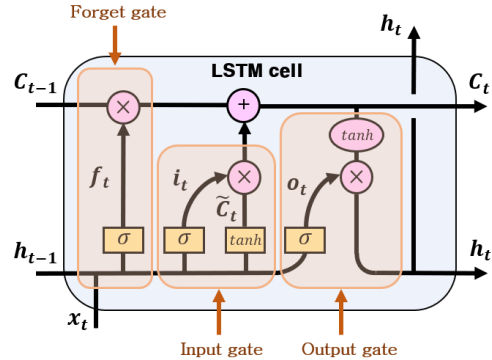


그림 2. LSTM의 구조.  
Fig. 2. Structure of LSTM.

### 2. LSTM

LSTM은 [7]에서 처음 소개되었고, 이후 많은 사람에게 의해 다듬어졌다. LSTM은 망각 게이트, 입력 게이트 그리고 출력 게이트 총합 3개의 게이트를 통해 장기 의존성 문제를 해결했다. 그림 2는 LSTM의 기본 구조이다. 이제 LSTM의 연산과정을 서술하겠다.

첫 번째 단계인 망각 게이트의 연산을 알아보겠다. 셀 상태  $C_{t-1}$ 는  $h_{t-1}$ 과  $x_t$ 을 보고 어떤 정보를 버릴지 결정한다. 이는 수식 1과 같이 표현된다<sup>[6]</sup>. 앞으로 나올 수식에서  $\sigma$ 는 시그모이드 활성화 함수,  $W$ 는 가중치,  $b$ 는 바이어스 값을 나타내며, 아래 첨자는 해당 게이트를 의미한다.

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f). \quad (1)$$

두 번째 단계는 입력 게이트이다. 시그모이드 층은 갱신할 값  $i_t$ 을 결정하고, tanh 층은 셀 상태에 더해질 수 있는 새로운 후보 값들의 벡터  $\tilde{C}_t$ 을 만든다 [6]. 이 과정을 수식으로 나타내면 다음과 같다.

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i). \quad (2)$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C). \quad (3)$$

이제  $i_t$ 와  $\tilde{C}_t$ 을 합쳐서 셀 상태  $C_{t-1}$ 을 새로운 셀 상태  $C_t$ 로 갱신한다<sup>[6]</sup>. 이 연산을 수식으로 표현하면 다음과 같다.

$$C_t = (f_t \times C_{t-1}) + (i_t \times \tilde{C}_t). \quad (4)$$

마지막 단계는 무엇을 출력할지 결정하는 출력 게이트이다. 시그모이드 층은 셀 상태에서 어떤 정보들을 출력할지 결정하고, 값이 -1에서 1 사이 값을 갖도록 셀 상태를 tanh에 넣는다<sup>[6]</sup>. 그리고 결정된 부분만 출력되도록 tanh출력을 다시 시그모이드 게이트 출력과 곱한다<sup>[6]</sup>. 이 과정을 수식으로 나타내면 다음과 같다.

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o). \quad (5)$$

$$h_t = o_t \times \tanh(C_t). \quad (6)$$

### 3. GRU

GRU는 <sup>[8]</sup>에서 처음 소개되었다. GRU는 LSTM에서 변형된 구조로 그림 3과 같은 구조를 가지며 연산과정을 수식적으로 나타내면 다음과 같다.

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t]). \quad (7)$$

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t]). \quad (8)$$

$$\tilde{h}_t = \tanh(W \cdot [r_t \times h_{t-1}, x_t]). \quad (9)$$

$$h_t = (1 - z_t) \times h_{t-1} + z_t \times \tilde{h}_t. \quad (10)$$

GRU는 ‘업데이트 게이트’와 ‘리셋 게이트’ 총 두 개의 게이트만 가지고 있어서 세 개의 게이트를 가지고 있는 LSTM보다 연산속도가 빠르다. 그리고 성능 또한 <sup>[9]</sup>의 논문의 결과처럼 LSTM과 비슷하거나 더 좋은 성능을 낸다. 즉, 구조적으로 LSTM보다 간단하지만, 연산속도와 결과물의 성능이 더 뛰어나다는 점에서 최근에 집중 받고 있는 신경망이다. 본 연구는 이점에 초점을 맞추어 같은 환경에서 GRU와 LSTM 두 가지 학습 모델의 성능을 평가한다.

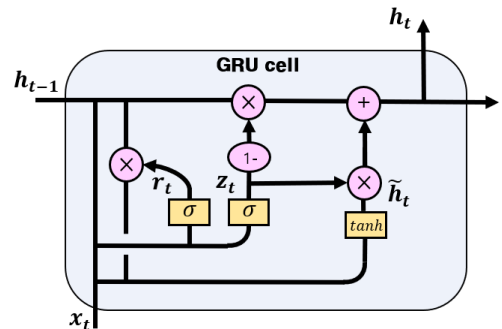


그림 3. GRU의 구조.  
 Fig. 3. Structure of GRU.

## III. 실험 환경 및 결과

### 1. 실험 데이터

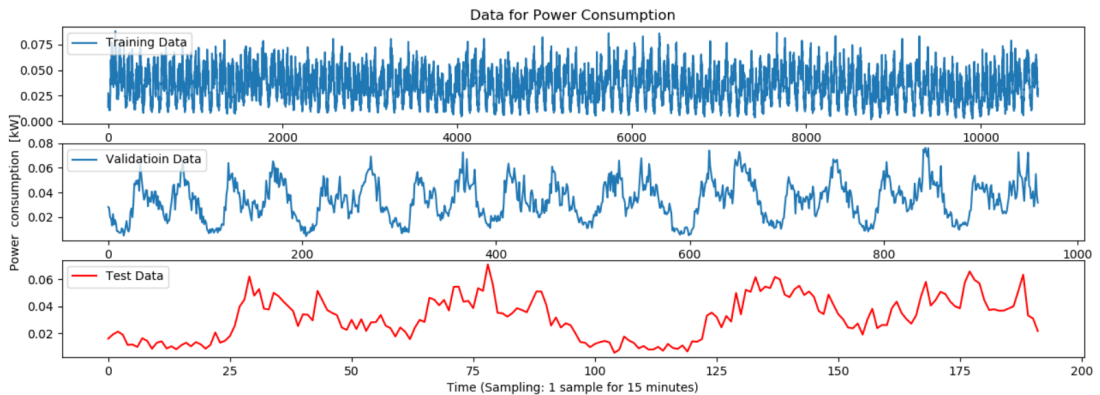


그림 4. 전력소모량에 대한 데이터.  
 Fig. 4. Data for power consumption.

본 논문에서 진행한 실험은 매사추세츠 대학교가 공개적으로 제공하는 가정단위 전력사용량 데이터를 활용하였다<sup>[10, 11]</sup>. 이 데이터는 2014년부터 2016년까지 매사추세츠 서부에 있는 114개 아파트에 대한 전력사용량이다. 본 실험에서는 114개 아파트의 전력사용량을 다 합친 후 15분 주기로 샘플링하여 데이터를 사용하였다. 학습 단계에서는 2016년 5월 24일부터 9월 11일까지의 데이터, 검증 단계에서는 9월 12일부터 21일까지의 데이터 그리고 테스트 단계에서는 9월 23일부터 24일까지의 데이터를 사용하여 실험을 진행하였다. 학습, 검증 그리고 테스트 각각에 대한 데이터의 변화 추이는 그림 4와 같다.

## 2. 실험 파라미터

본 연구에서 진행한 실험은 look\_back, 셀의 개수, 배치 사이즈, 학습률 그리고 epoch로 총합 5개의 파라미터를 조정해가며 실험하였다.

첫 번째로, look\_back의 크기이다. look\_back은 입력층에 들어가는 데이터 셋의 크기이다. 하루 동안의 데이터로 다음 데이터를 예측하기 위해 이 값은 96으로 설정했다. look\_back 값이 96이 되면서 본 모델은 총 10560개 샘플들로 학습한다.

두 번째로, 셀의 개수이다. 이 파라미터는 은닉층에서 사용할 LSTM 또는 GRU 셀이 각각 계층마다 몇 개인지를 의미한다. 본 실험에서 사용하는 데이터는 특징이 '전력사용량' 하나이고 같은 특징을 예측하는 것이므로 신경망의 복잡도가 낮고 매개변수들의 수가 적어 이 파라미터는 2로 설정했다.

세 번째로, 배치 사이즈이다. 배치 사이즈는 한 번에 학습되는 데이터 개수를 의미한다. 다시 말해, 한 번에 신경망 네트워크로 넘겨지는 학습 데이터의 수를 말한다.

이 값은 실험적으로 48로 설정했다.

네 번째로, 학습률이다. 한 번의 학습으로 얼마만큼 학습해야 할지, 즉 매개변수 값의 갱신 정도를 결정하는 것이 학습률이다. 이 값은 실험적으로 0.01로 설정했다.

마지막으로 epoch는 하나의 단위이다. 1 epoch는 학습에서 학습 데이터를 모두 소진했을 때의 횟수이다. 일반적으로 반복 학습을 하면 모델의 성능이 올라가는 경향이 있다. 하지만 이 값을 너무 크게 하면 과적합으로 빠져서 일반화 성능이 떨어질 수 있다. 본 실험의 경우, 만개 이상의 학습 데이터를 모음으로써 일반화 성능을 높였기 때문에 epoch는 2로 지정했다.

표 1은 본 실험에서 설정한 파라미터를 정리한 것이다. 성능 비교를 위해 같은 데이터로 학습시킨 LSTM의 파라미터 또한 표 1에 정리되어 있다.

표 1. 실험 파라미터.

Table 1. Parameters of experiment.

Parameters	LSTM	GRU
look_ahead	96	96
Number of cells	2	2
Batch size	48	48
Learning rate	0.01	0.01
Epoch	2	2

## 3. 실험 결과

그림 5는 각각 LSTM과 GRU로 학습된 모델의 테스트 입력에 대한 예측결과를 보여준다. 그림을 보면 두 모델 모두 실제 전력소모량의 흐름을 전반적으로 잘 예측하지만, GRU로 학습된 모델이 LSTM으로 학습된 모델보다 더 정확히 예측하는 것을 확인할 수 있다. 두 모델

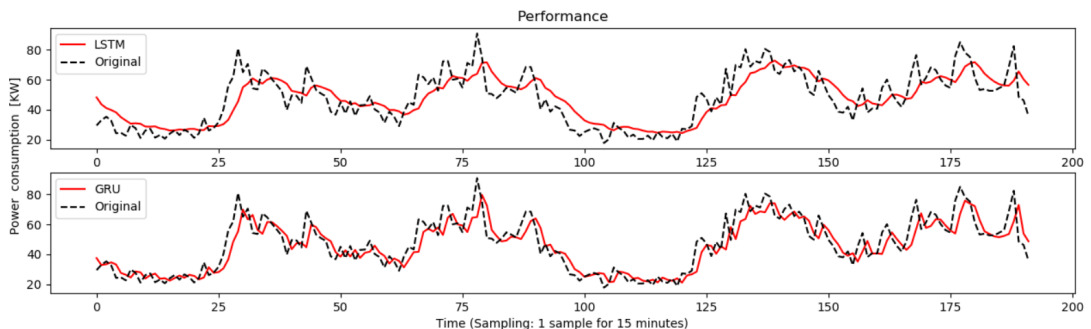


그림 5. LSTM과 GRU의 예측결과.

Fig. 5. Prediction result of LSTM and GRU.

의 예측결과에 대해 성능평가를 Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), Mean Percentage Error (MPE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE)의 5 가지 지표들로 진행하였다. 각 성능평가 방법을 수식으로 표현하면 다음과 같이 나타난다.

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - p_i|. \quad (11)$$

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{y_i - p_i}{y_i} \right|. \quad (12)$$

$$MPE = \frac{100\%}{n} \sum_{i=1}^n \frac{y_i - p_i}{y_i}. \quad (13)$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - p_i)^2. \quad (14)$$

$$SE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - p_i)^2}. \quad (15)$$

수식 11에서 15를 살펴보면,  $y$ 는 실제로 측정된 전력 소모 값,  $p$ 는 예측결과 값 그리고  $i$ 는 테스트 데이터 샘플의 개수를 의미한다. 각 아래 첨자들은 샘플 번호이다. LSTM의 예측결과와 본 논문에서 진행한 실험의 결과의 성능을 정리해보면 표 2와 같이 나타난다. MAE 기준으로 1.58, MAPE 기준으로 4.52%, MPE 기준으로 5.59% 만큼의 성능개선을 확인했다. 또한 MSE와 RMSE 기준을 통해 GRU로 학습하면 LSTM으로 학습한 것보다 오차가 감소함을 확인할 수 있다. 5가지 지표에 대해서 GRU의 실험 결과가 LSTM의 실험 결과보다 성능 개선된 것을 확인할 수 있다.

표 2. LSTM과 GRU의 성능평가.  
 Table 2. Performance evaluation of LSTM and GRU.

Evaluation Method	LSTM	GRU	Improvement
MAE	8.10	6.52	1.58
MAPE	18.38	13.86	4.52
MPE	-6.30	-0.35	5.95
MSE	106.79	77.39	29.4
RMSE	10.33	8.80	1.53

#### IV. 결 론

본 연구에서는 단기 전력 예측을 하나의 특징으로

GRU로 학습된 모델을 통해 예측하였다. UMass Trace 본 연구에서는 단기 전력 예측을 하나의 특징으로 GRU로 학습된 모델을 통해 예측하였다. UMass Trace Repository의 전력사용량 데이터로 구성된 모델을 학습시키고, 실험을 통해 하이퍼파라미터 5개를 설정하였다. 마지막으로 5가지 지표를 통해 성능평가를 진행하였고 GRU가 LSTM보다 연산속도가 빠르며 성능이 더 좋음을 확인하였다.

추후 진행될 수 있는 연구로는, 다양한 특징을 통해 예측하는 것과 본 연구에서 진행한 단기 전력 예측뿐만 아니라, 중기, 장기 전력 예측이 가능한 모델을 생각해 볼 수 있다. 이러한 모델은 복잡한 네트워크를 가지게 될 텐데 LSTM보다 간단한 구조이며 연산속도가 더 빠른 GRU를 사용하면 굉장히 효율적일 것이다. 본 연구의 결과물을 기반으로 전력사용량 예측 관련 연구는 주거 지역의 에너지 수요량을 파악할 수 있게 되고 소비자는 예측된 값을 통해 전력사용량 증가로 나타나는 블랙아웃 현상을 예방할 수 있을 것으로 기대된다.

#### References

- [1] D. Gan, Y. Wang, N. Zhang and W. Zhu, "Enhancing short-term probabilistic residential load forecasting with quantile long-short-term memory," *Journal of Engineering*, vol. 2017, no. 14, pp. 2622-2627, Jan. 2017. DOI: <https://doi.org/10.1049/joe.2017.0833>
- [2] Y. Chen, B. Zhang, J. Wang, B. Mao, R. Fang, C. Mao and S. Duan, "Active control strategy for microgrid energy storage system based on short-term load forecasting," *Power System Technology*, vol. 35, no. 8, pp. 35-40, Aug. 2011.
- [3] P. Zhang, X. Pan and W. Xue, "Short-term load forecasting based on fuzzy clustering wavelet decomposition and BP neural network," in *Proc. 2011 IEEE Asia-Pacific Power and Energy Engineering Conference*, pp. 1-4, Wuhan, China, Mar. 2011. DOI: <https://doi.org/10.1109/appeec.2011.5748523>
- [4] D.C. Park, M. A. El-Sharkawi, R. J. Marks, L. E. Atlas and M. J. Damborg, "Electric load forecasting using an artificial neural network," *IEEE Trans. Power Systems*, vol. 6, no. 2, pp. 442-449, May 1991. DOI: <https://doi.org/10.1109/59.76685>
- [5] J. Zheng, X. Chen, K. Yu, L. Gan, Y. Wang and K. Wang, "Short-term power load forecasting of residential community based on GRU neural network," in *Proc. 2018 IEEE International Conference on Power System Technology (POWERCON '2018)*, pp.4862-4868, Guangzhou, China, Nov. 2018. DOI: <https://doi.org/10.1109/powercon.2018.8601718>

- [6] C. Olah, "Understanding LSTM networks," 2015; Available at <http://colah.github.io/>
- [7] S. Hochreiter and J. Schmidhuber, "Long short-term memory", *Neural Computation*, vol. 9, no. 8, pp. 1735-1780, Nov. 1997.  
DOI: <https://doi.org/10.1162/neco.1997.9.8.1735>
- [8] K. Cho, B. V. Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk and Y. Bengio, "Learning phrase representations using RNN encoder-decoder for statistical machine translation," in *Proc. 2014 Conference on Empirical Methods in Natural Language Processing*, pp. 1724-1734, Doha, Qatar, Oct. 2014.  
DOI: <https://doi.org/10.3115/v1/d14-1179>
- [9] J. Chung, C. Gulcehre, K. Cho and Y. Bengio, "Empirical evaluation of gated recurrent neural networks on sequence modeling," in *Proc. 2014 NIPS Workshop on Deep Learning*, pp.1-9, Montreal, Canada, Dec. 2014.
- [10] M. Liberatore and P. Shenoy, "UMass Trace Respository," 2018; Available at <http://traces.cs.umass.edu/>
- [11] Y. Cheng, C. Xu, D. Mashima, V. L. L. Thing and Y. Wu, "PowerLSTM: Power demand forecasting using long short-term memory neural network," in *Proc. 2017 Advanced Data Mining and Applications*, pp. 727-740, Singapore, Oct. 2017.  
DOI:[https://doi.org/10.1007/978-3-319-69179-4\\_51](https://doi.org/10.1007/978-3-319-69179-4_51)
- [12] K. Seo, "A comparison study on back-propagation neural network and support vector machines for the image classification problems," *Journal of the Korea Academia-Industrial cooperation Society*, vol. 9, no. 6, pp. 1889-1893, Dec. 2008.  
DOI: <https://doi.org/10.5762/KAIS.2008.9.6.1889>
- [13] J. Ahn, S. Park and C. Kim, "A study on neural network model for winter electric power demand prediction," *The Journal of KIIT*, vol. 15, no. 9, pp. 1-9, Sept. 2017.  
DOI: <https://doi.org/10.14801/kiit.2017.15.9.1>

## 저 자 소 개

### 이 선 민(준회원)



- 2016년 3월~현재 : 광운대학교 전자융합공학 학사과정
- 관심분야 : 에너지인터넷, 인공지능, 스마트 그리드, 디지털통신

### 선 영 규(준회원)



- 2018년 2월 : 광운대학교 전자융합공학 학사 졸업
- 2018년 3월~현재 : 광운대학교 전자융합공학과 석박통합과정
- 관심분야 : 인공지능, 무선 에너지 하베스팅, 디지털통신, 전력선 통신

### 이 지 영(준회원)



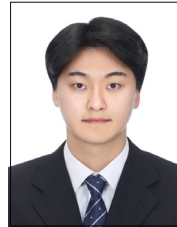
- 2016년 3월~현재 : 광운대학교 전자융합공학 학사과정
- 관심분야 : 인공지능, 디지털통신, 에너지인터넷

### 이 동 구(준회원)



- 2019년 2월 : 광운대학교 전자융합공학 학사 졸업
- 2019년 3월~ 현재 : 광운대학교 전자융합공학 석박통합과정
- 관심분야 : 인공지능, 전력선통신, 스마트 그리드

### 조 은 일(준회원)



- 2014년 3월~현재 : 광운대학교 전자융합공학 학사과정
- 관심분야 : 스마트 그리드, 인공지능, 사물인터넷

박 대 현(준회원)



- 2014년 3월~현재 : 광운대학교 전 자용합공학 학사과정
- 관심분야 : 스마트 그리드, 인공지능, 사물인터넷

김 용 범(준회원)



- 2014년 3월~현재 : 광운대학교 전 자용합공학 학사과정
- 관심분야 : 스마트 그리드, 인공지능, 사물인터넷

심 이 삭(준회원)



- 2016년 2월 : 광운대학교 전자융합공학과 졸업
- 2016년 3월~ 현재 : 광운대학교 전자융합공학과 석박통합과정
- 관심분야 : 전력선 통신, 무선 에너지 하비스팅, Backscatter, 인공지능

김 진 영(정회원)



- 1998년 2월 : 서울대학교 전자공학과 공학박사
- 2001년 2월 : SK텔레콤 네트워크 연구소 책임연구원
- 2001년 3월 ~ 현재 : 광운대학교 전자융합공학과 교수
- 관심분야 : 차세대이동통신, 가시광통신, 전력선통신, 인공지능

※ 본 연구는 과학기술정보통신부 및 정보통신기획평가원의 대학ICT연구센터지원사업의 연구결과로 수행되었음 (IITP-2019-2016-0-00311\*)