논문 2018-1-5

양방향 LSTM을 활용한 전력수요 데이터 예측 기법 연구

고상준*, 윤호영**, 신동명*†

Electronic Demand Data Prediction using Bidirectional Long Short Term Memory Networks

Sang-Jun Ko*, Ho-Yeong Yun**, Dong-Myung Shin*†

요 약

산업 분야에서 전력수요 예측은 예측 전력량보다 실측 전력량이 일정 범위에서 벗어나면 이상 징후로 판별하여 기계 오작동, 전원 꺼짐 등의 이상 현상을 탐지하는데 활용되기도 한다. 전력 데이터는 계절성과 추세성이 강한 시계열 데이터이기 때문에 예측이 어려운 편이다. 본 연구에서는 딥 러닝 기반의 알고리즘인 LSTM과 더불어 최근에 각광받고 있는 알고리즘인 양방향(Bi-Directional) LSTM 알고리즘을 활용한 전력 예측 모형을 수립하였으며, 해당 모형은 4년간의 전력 데이터를 활용하여 학습, 평가, 검증하였다.

Abstract

Power demand prediction is used to detect anomalies such as machine malfunction and power off when the measured power amount deviates from the predicted power amount by a certain range in the industrial sector. Power data is difficult to predict because it is time series data with strong seasonality and trend.

In this paper, a power prediction model based on bi-directional LSTM algorithm, which is a deep learning based algorithm and the most popular algorithm in recent years, has been established along with LSTM. The model was learned, evaluated, and verified using four years of power data.

한글키워드 : 전력 수요 예측, 딥 러닝, 순환 신경망, 양방향 장단기 기억 네트워크

keywords : Electronic Demand Prediction, Deep Learning, Recurrent Neural Network, Bidirectional Long Short Term Memory Networks

1. 서 론

전력 산업은 생산과 동시에 소비가 이루어지는 특성을 가지고 있기 때문에 전력수요 예측 오차는 전력계통의 안정적 운영을 위협할 수 있으므로. 최적 운영을 위해서는 전력 수요에 대한 정확한 예측이 필수적이다[1]. 산업현장에서 전력 수요 예측은 이상 현상을 판별하는데 활용되기도 한다. 여기서 이상 현상은 기계의 오작동, 전원 꺼짐 등을 예로 들 수 있다. 현재 전력량이 예측 전력량의 일정 범위를 초과하거나 미달할 경우

* 엘에스웨어(주)
** 연세대학교 산업공학과
† 교신저자: 신동명(roland@lsware.com)
접수일자: 2018.05.28. 심사완료: 2018.06.15.
게재확정: 2018.06.20.

www.kci.go.kr

이상 현상이 발생했다고 의심할 수 있다. 이를 이상 징후 판별이라고 정의한다.

전력 수요량 데이터는 계절성과 추세성이 강하게 나타나는 시계열 데이터이다. 일, 월, 분기별로 데이터의 특징이 상이하기 때문에 일반적인 데이터와 비교했을 때, 미래 데이터를 예측하는 것이 쉽지 않은 편이다.

BOX 외 1명이 시계열 데이터 분석을 위해 ARIMA(Auto-regressive Integrated Moving Average) 모형을 사용한 후[2], 지난 30년간 가장 인기 있는 선형 모델 중 하나인 ARIMA 모형을 이용한 시계열 예측 연구가 주로 진행되었다[3]. Contreras 외 1명은 2가지 ARIMA 모형을 이용하여 각각 스페인과 캘리포니아 전력 시장에서 시간당 가격을 예측하였으며, 평균 오차율은 스페인과 캘리포니아 각각 대략 10%, 5%의 결과를 나타내었다[4]. 또한 ARIMA 모형의 한계인 비선형 패턴을 쉽게 포착할 수 없는 문제를 해결하기 위해 Pai 외 1명은 ARIMA 모형에 SVM(Support Vector Machine)을 결합하여 주식 시장을 예측하는 하이브리드 방법론을 제시하였다[5].

최근에는 RNN(Recurrent Neural Networks)의 일종인 LSTM(Long Short-Term Memory) 기법을 활용하여 시계열 데이터를 예측하는 연구가 진행되고 있으며, Kong 외 5명은 개별 가정용 가구의 스마트 미터기의 데이터를 사용하여 LSTM 모델로 단기 부하 예측 연구를 진행하였다[6]. 또한 Marino 외 2명은 빌딩 에너지 부하 예측을 위해 두 종류의 LSTM 모델을 이용하여 LSTM 기반 신경 네트워크의 효율성을 연구했고[7], Zheng 외 2명은 SD(Similar days) 선택 및 EMD (Empirical Mode Decomposition) 방법과 하이브리드 된 LSTM 신경망 모델을 제시하여 저주파수 및 고주파수 시계열 데이터에서의 단기간 부하 예측 문제를 연구하였다[8].

그 외에도 Pezzulli 외 6명은 베이지안 계층형 모델을 이용하여 1년 후의 겨울철 전기 피크 수요를 예측하는 연구를 진행하였으며[9], Yuan 외 4명은 단기 풍력 예측을 위해 GSA (Gravitational Search Algorithm)를 이용해서 LSSVM(Least Square Support Vector Machine)의 매개변수를 최적화 시켜 더 높은 정확도를 보여주는 LSSVM-GSA 하이브리드 모델을 제안하는 등 [10], 다양한 시계열 데이터 예측 연구가 계속해서 진행되고 있다. 본 연구는 기존의 LSTM을 개선시킨 Bi-LSTM기법을 활용하여, 전력량 데이터의 예측률을 높이는 모델을 개발하는 것을 목적으로 하였다.

본 연구에서는 4년간의 전력량 데이터를 토대로 예측 모형 연구를 진행하였다. 2장에서는 전력 데이터를 예측하는 기법인 ARIMA기법, RNN, LSTM, Bi-LSTM에 대해 상세하게 서술하였다. 3장에서는 Bi-LSTM 기반의 예측 모형의 학습과정을 5단계로 서술하였다. 4장에서는 결론과 향후 연구 방향을 제시하였다.

2. 예측모형 연구

2.1. ARIMA 모형[11]

ARIMA 모형은 시계열 분석 기법의 한 종류로 과거의 관측값과 오차를 사용해서 현재의 시계열 값을 설명하는 ARMA 모델을 일반화 한 모델이다. 이는 ARMA 모델이 데이터가 안정적(Stationary)인 상태에서만 적용 가능한 것에 비해, 분석 대상이 약간은 불안정적인(Non-Stationary) 특징을 보여도 적용이 가능하다. 하지만 시계열 데이터를 안정적 시계열로 변환하는 것이 일반적인 절차이다.

불안정적인 시계열 데이터를 안정적인 시계열 데이터로 변환하는 방법에는 주로 차분과 로그함

www.kci.go.kr

수를 사용한다. 바로 이전 값과 비교할 때의 예시는 아래와 같다.

$$\triangle Y_t = Y_t - Y_{t-1} = Y - t - BY_t = (1-B)\,Y_t$$

1차 차분을 적용했음에도 비정상적인 경우는 로그함수 혹은 2차 차분을 사용한다. 하지만 로그와 차분을 사용하여 안정적인 데이터로 변환하는 과정은 오히려 시계열 데이터의 특성을 무시하는 통계적 가정이 될 수 있다. 이는 통계적 기법의 단점이기도 하며, 데이터에 기반하여 모수를 추정하기 때문에 지속적인 학습에 어려움이 있어 본 연구에서는 채택하지 않았다.

2.2. 순환신경망 기법[12]

통계적 기법과는 달리 딥 러닝 기법에서는 가정이 요구되지 않는다. 또한 모형이 한 번 완성되면 입력데이터에 따라 쉽게 학습이 이루어지기 때문에 실시간 관리에도 용이하다.

순환 신경망(Recurrent Neural Networks)은 주로 자연어, 유전자, 손 글씨, 음성신호, 센서가 감지한 데이터, 주가 등 배열(Sequence 또는 시계열 데이터)의 형태에 활용된다. 기존 인공신경망(Artificial Neural Network)은 입력층에서 은닉층, 출력까지 진행되는 반면에 순환 신경망은 은닉층(T)의 결과가 다음 은닉층(T+1)의 입력으로 작용된다.

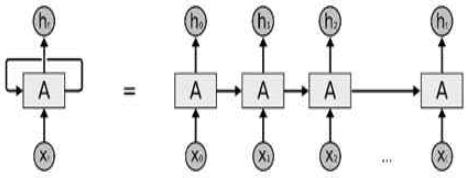

그림 1. 순환신경망의 구조
Fig. 1. Structure of recurrent neural network

입력층에 입력되는 데이터는 입력층과 은닉층을 연결하는 가중치 연산을 거쳐 다음 은닉층으로 전달되게 되며, 은닉층에서는 출력층으로 전달할 결과를 아래의 수식을 통해 생성한다.

$$h_t = \tanh(W_{hh}h_{t-1} + W_{xh}x_t)$$

각 층에는 데이터의 패턴과 이에 해당하는 분류 결과를 학습시키기 위한 행렬이 존재하며, 이 행렬 내의 가중치 값들을 최적화하여 예측 정확도가 가장 높은 가중치 집합을 찾는 방식이다. 은닉층의 결과가 다시 은닉층의 입력으로 작용되기 때문에 시간적 흐름에 따른 데이터, 즉 시계열 데이터의 패턴을 분석하여 특정 시점의 결과 예측에 이전 시점의 데이터들이 관여한다는 특징이 있다.

하지만 특정 시점으로부터 먼 과거의 정보를 예측 결과에 반영하지 못하는 구조적 단점이 있는데 이를 기울기의 사라짐 문제(Gradient Vanishing Problem), 장기 의존성 문제(Long-Term Dependencies) 문제라고 한다. 이러한 문제를 해결을 위해 장단기 기억 네트워크(Long Short Term Memory Network)를 활용한 예측 모형을 설계하였다.

2.3. 장단기 기억 네트워크(LSTM Network)

순환신경망의 장기 의존성 문제를 보완한 장단기 기억 네트워크(이하 LSTM)가 등장하였다. 학습을 할수록 과거의 기억을 잃는 순환신경망의 단점을 LSTM은 3개의 게이트(Forget, Input, Output) 구조를 도입하여 보완하였다. 그림 2는 LSTM의 구조를 뜻한다.

셀 내부 동작은 크게 두 가지로 현 시점의 은닉층 연산을 수행하는 게이트와 과거 시점의 데이터에 대한 유지 연산을 담당하는 셀 상태 정보 연산 게이트가 존재한다. 두 연산의 결과 값은

www.kci.go.kr

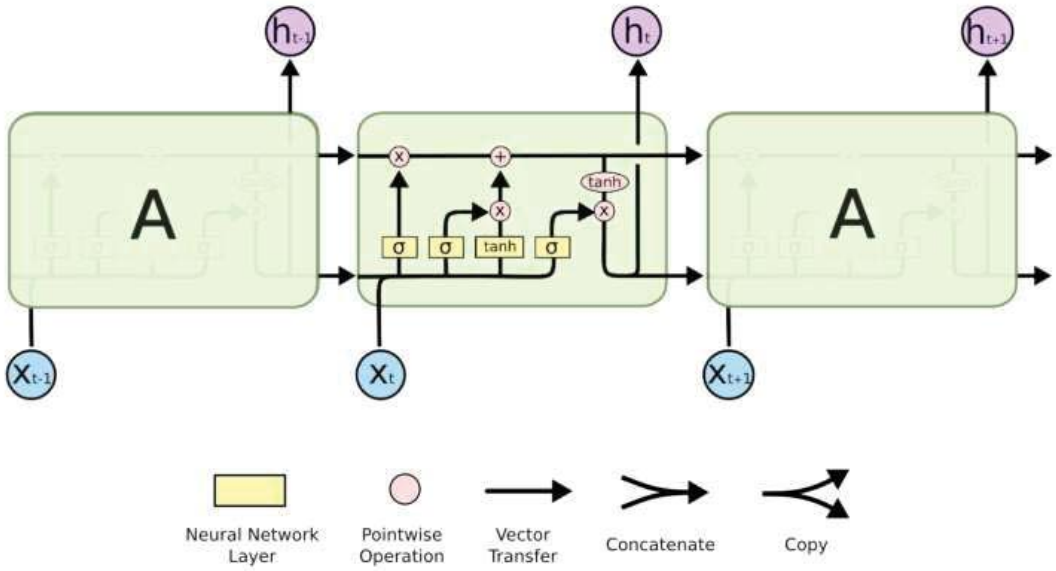

그림 2. 장단기 기억 네트워크(LSTM)의 구조
Fig. 2. Structure of LSTM

다음 시점의 장단기 메모리 셀로 전달되어 연산에 활용된다. 이전 시점에 대한 정보를 Forget 게이트를 통해 정보의 유지를 관리하는데, 이러한 방식으로 먼 과거의 시점의 셀 상태 정보를 은닉층의 연산 결과와 함께 다음 시점의 은닉층으로 전달하게 된다[13].

Forget 게이트 층은 하나의 시그모이드 층에 의해 결정된다. h_{t-1}과 x_t를 보고 셀 상태 C_{t-1}에서의 각 숫자를 위한 0과 1 사이의 값을 출력한다. 1은 정보를 완전히 유지함을 뜻하며, 0은 완전히 제거함을 뜻함으로써 데이터를 장기 기억 여부를 관리한다. Input 게이트의 시그모이드 층은 어떤 값을 갱신할 것인지를 결정하는 게이트이다. 이는 장기데이터를 기억하는 Forget 게이트 층과는 달리 단기 데이터의 기억을 관리한다. tanh 층은 셀 상태에 더해질 수 있는 새로운 후보 값들의 벡터 C_t로 구성된다. Output 게이트는 무엇을 출력할지를 결정하며, 셀 상태에 기반을 두지만 큰 영향을 주지는 않는다.

LSTM은 상대적으로 복잡한 기법으로 최초로 제안된 이후 개선된 알고리즘이 발표되었다. 대표적으로 GRU(Gated Recurrent Unit)이 있는데, 3개의 게이트를 갖는 LSTM과는 달리, GRU는 Reset 게이트(r), Update 게이트(z)를 통해 정보의 전달과 조합 방식을 다르게 구현하기도 하였다. 또 다른 개선 알고리즘으로는 양방향 LSTM (Bi- Directional LSTM)이 좋은 성능을 보이고 있다.

2.4. 양방향 장단기 기억 네트워크(Bi-LSTM)

앞서 설명한 RNN과 LSTM은 은닉 계층에 과거의 데이터 정보를 기억하기 때문에 시계열 예측에 적합하다. 하지만 입력 순서를 시간 순서대로 입력하기 때문에 결과물이 직전 패턴을 기반으로 수렴하는 경향을 보인다는 한계가 존재한다[14]. 이런 단점을 해결하는 것을 목적으로 양방

www.kci.go.kr

향 순환신경망(Bi-RNN)이 제안되었는데, 양방향 순환신경망은 순방향과 역방향의 두 개의 분리된 순환신경망을 통해 학습을 시킨다. 기존의 순방향에 역방향을 추가하여, 은닉층에 추가시킴으로써 성능이 향상됨을 볼 수 있었다.

양방향 순환신경망은 일반적인 순환신경망보다 높은 성능을 보였지만, 데이터의 길이가 길고 층이 깊으면, 과거의 정보가 손실되는 단점이 존재하였다. 마찬가지로 이러한 단점을 극복하기 위해 제안된 알고리즘이 양방향 LSTM이다.

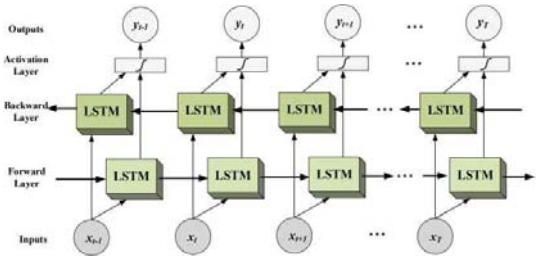

그림 3. Bi-LSTM의 구조[10]
Fig. 3. Structure of Bi-LSTM

양방향 LSTM은 기계번역 분야에서 좋은 성능을 내고 있는데, 출력값에 대한 손실을 최소화하는 과정에서 모든 파라미터들이 동시에 학습되는 종단간 학습이 가능하며, 단어와 구(Phrase)간 유사성을 입력벡터에 내재화해 성능을 개선하였으며, LSTM의 기본적인 성능과 Attention 매커니즘을 도입하여, 데이터의 길이가 길어져도 성능 저하를 막을 수 있었다.

머신러닝 분야에서 발표되고 있는 최신/최고의 연구성과는 모두 양방향 LSTM이 적용된 모델일 정도로 높은 성능의 알고리즘 중 하나이다.

3. 전력수요 예측 모형

본 연구에서 제안한 양방향 LSTM 기반의 전력수요 예측모형은 4년간의 전력 수요량 수집 데이터를 예측하는 것을 목표로 한다. 설계된 예측모형의 구조는 그림 4와 같다.

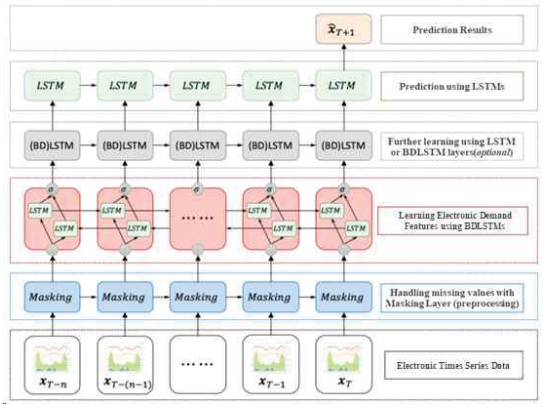

그림 4. 전력수요 예측모형의 구조
Fig. 4. Structure of Power Demand Forecasting Model

3.1. 결측치 대체

데이터의 특성상 완벽하게 수집되는 경우가 많지는 않다. 본 연구의 데이터 또한 4년간의 전력량을 1시간 단위로 수집하였는데, 정상적인 값(0, 공백, ***, #VALUE!값이 아닌 경우)이 91% 정도만 차지하였다. 해당 결측치는 보간법을 활용하여 대체하였다. 보간법은 1, n, 3의 값이 있을 때, n의 값은 1과 3의 평균인 2로 대체되는 결측치 대체 기법으로 간단하게 적용할 수 있다. 보간법 이외에 최대 우도 추정법, 확률변수법 등을 활용할 수 있다.

3.2. 데이터 분류 및 목적함수

연구 데이터는 2013년부터 2016년까지 4년의 전력데이터를 활용하였다. 2013년과 2014년 데이터는 학습용 데이터로 활용하고, 2015년 데이터는 평가용, 2016년 데이터는 검증용 데이터로 분

www.kci.go.kr

류하였다.

예측된 데이터의 오차율 확인은 평균 제곱근 오차(Root Mean Squre Error)를 사용하였으며, 식은 아래와 같다.

$$SE = \sqrt{\frac{1}{n}(\sum_{t=1}^{n}(\text{real Data}_t - \text{prediction Data}_t)^2)}$$

3.3. 모형 학습 및 검증

학습/평가용 데이터를 활용하여 실제값과 예측값의 에러가 적어지도록 LSTM 모형을 활용하여 학습시켰다. 실제값과 예측값의 차이를 입력층, 은닉층의 개수에 따라 실험을 달리하여 최적의 모형을 도출하였다.

모형의 검증은 검증용 데이터를 활용하여 실제값과 예측값의 에러를 계산하는 과정이다. 가중치 업데이트는 하지 않고 Forward 방향으로만 에러의 합을 계산한다. 에러가 충분히 작다고 판단되면 모형을 완료시키고, 에러가 크다면 입력층, 은닉층의 개수를 조절하여 재검증 과정을 거친다.

3.4. 이상 유무 판단

이상치의 기준을 마련하는 것은 여러 가지 방법이 있겠지만, 적용 환경마다 상이하기 때문에 관리자의 권한으로 남겨두었다. 다만 기본값을 제공하기 위해 통계적 기법을 활용하였는데, 예측값과 실제값에 자연로그를 취하여, 등분산을 만족하게끔 변환하고, $\pm 2\sigma$ 내에 포함되지 않으면 이상으로 판단하였다.

4. 실험 및 성능평가

본 장에서는 제안한 전력수요 예측모형을 구현하고 실험한 환경 및 모델의 학습, 테스트, 검증 결과에 관하여 서술하였다.

4.1. 실험 환경

본 모델은 C#을 활용하여 개발하였고, 예측 엔진은 Tensorflow 1.7버전을 활용하였다. Learning rate는 0.01로 설정하였으며, 통계적 유의함을 위해 100회 반복실험을 진행하였다. 실험결과는 기존에 LSTM을 활용하여 개발한 모형과 양방향 LSTM을 활용한 모형을 비교실험하였다.

4.2. 실험 결과

LSTM과 양방향 LSTM의 예측률을 비교 분석한 결과는 표 1과 같다. 100번의 실험 중에 LSTM의 최고 예측률은 약 88%(0.879577), 양방향 LSTM 경우 약 94%(0.939901)로 나타났으며, 최소(worst case) 예측률은 LSTM이 약 84%(0.840132), 양방향 LSTM이 약 89%(0.892095), 평균은 LSTM이 약 86%(0.859743), 양방향 LSTM이 약 92%(0.915001)로 전반적인 수치는 Bi-LSTM이 LSTM보다 높은 성능을 보였다.

표 1. 실험결과 예측률(단위 : %)
Table 1. Results of prediction rate(%)

	LSTM	Bi-LSTM
최고값	0.879577	0.939901
최소값	0.840132	0.892095
평균	0.859743	0.915001
표준편차	0.011723	0.0141

표 2.는 LSTM과 양방향 LSTM의 RMSE값을 비교 분석한 결과이다. LSTM의 경우 RMSE 값은 최소 0.011493821, 최대값은 0.025662053으로 평균 0.019290822였다. 표준편차는 0.004358085이다. 반면 양방향 LSTM의 경우는 최소값 0.009534, 최대값 0.016468, 평균 0.013103, 표준편

www.kci.go.kr

차는 0.002045로 Bi-LSTM이 LSTM보다 전반적으로 좋은 결과를 보였다.

표 2. 실험결과_RMSE
Table 2. Result of RMSE

	LSTM	Bi-LSTM
최소값	0.011493821	0.009534
최대값	0.025662053	0.016468
평균	0.019290822	0.013103
표준편차	0.004358085	0.002045

100번의 실험 중에서 가장 좋은 결과를 보인 실험에 대한 실데이터와 예측데이터를 비교한 값을 그림 5.에 그래프 형태로 표현하였다.

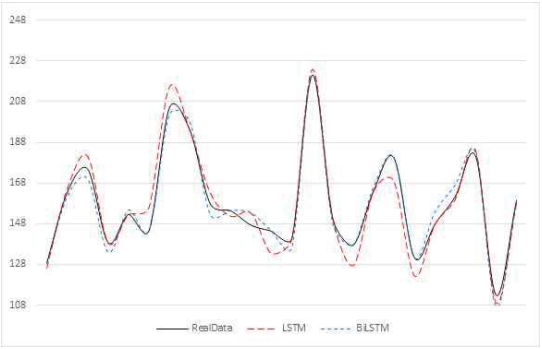

그림 5. LSTM과 Bi-LSTM의 데이터 예측 추이
Fig. 5. Data prediction trend of LSTM and Bi-LSTM

4년간 수집한 전력 수요 데이터 중 30%에 해당하는 약 440일 간의 수요 데이터와 이를 LSTM와 양방향 LSTM으로 학습시킨 뒤 예측한 데이터와의 비교한 결과 LSTM으로 예측한 값보다 양방향 LSTM으로 예측한 값이 실제 데이터와 겹치는 모양을 보였다.

5. 결 론

본 연구에서는 전력 이상징후를 탐지하기 위한 목적으로 전력 수요데이터를 예측하는 모델 개발을 선제적으로 진행하였다. 전력 수요 데이터(시계열 데이터)를 예측하기 위해 장기 의존성 문제를 극복하고, 과거의 패턴에 강하게 종속되지 않고자, 양방향 LSTM을 활용하였다. 본 연구는 전력데이터의 수요 데이터만 고려하였는데, 전력 수요에 영향을 주는 외부적인 요인들을 함께 고려한다면 좀 더 현실성 반영할 수 있는 모형이 될 것으로 사료되며, 추후 연구로 진행하고자 한다.

Acknowledgement

본 연구는 산업통상자원부(MOTIE)와 한국에너지기술평가원(KETEP)의 지원을 받아 수행한 연구 과제입니다. (No. 2016020108280)

참 고 문 헌

[1] 한국전력거래소, “단기 전력수요 예측 기법 및 적용 방안에 관한 연구”, 2011.
[2] Box, George EP, and Gwilym M. Jenkins. Time series analysis: forecasting and control, revised ed. Holden-Day, 1976.
[3] Zhang, G. Peter. "Time series forecasting using a hybrid ARIMA and neural network model." Neurocomputing 50 (2003): 159-175.
[4] Contreras, Javier, et al. "ARIMA models to predict next-day electricity prices." IEEE transactions on power systems 18.3 (2003): 1014-1020.
[5] Pai, Ping-Feng, and Chih-Sheng Lin. "A hybrid ARIMA and support vector machines model in stock price forecasting." Omega 33.6 (2005): 497-505.

www.kci.go.kr

[6] Kong, Weicong, et al. "Short-Term Residential Load Forecasting based on Resident Behaviour Learning." IEEE Transactions on Power Systems 33.1 (2018): 1087-1088.
[7] Marino, Daniel L., Kasun Amarasinghe, and Milos Manic. "Building energy load forecasting using deep neural networks." Industrial Electronics Society, IECON 2016-42nd Annual Conference of the IEEE. IEEE, 2016.
[8] Zheng, Huiting, Jiabin Yuan, and Long Chen. "Short-term load forecasting using EMD-LSTM neural networks with a Xgboost algorithm for feature importance evaluation." Energies 10.8 (2017): 1168.
[9] Pezzulli, Sergio, et al. "The seasonal forecast of electricity demand: A hierarchical Bayesian model with climatological weather generator." Applied Stochastic Models in Business and Industry 22.2 (2006): 113-125.
[10] Yuan, Xiaohui, et al. "Short-term wind power prediction based on LSSVM－GSA model." Energy Conversion and Management 101 (2015): 393-401.
[11] P.J. Brockwell and R.A. Davis, "Time Series: Theory and Methods, Springer Series in Statistics", 1986.
[12] Adam Gibson & Josh Patterson, "Deep Learning : A Practitioner's Approach", O'REILLY & Associates Inc., 2016.
[13] 박성혁 외, "순환신경망 기반의 사용자 의도 예측 모델", 정보과학회논문지 제 45권 4호, 2018.
[14] ÖzalYildirim, "A novel wavelet sequence based on deep bidirectional LSTM network model for ECG signal classification", Computers in Biology and Medicine, Vol.96. pp.189-202, 2018.
[15] T. Robinson. "An application of recurrent nets to phone probability estimation," IEEE Transactions on Neural Networks, vol. 5 no. 2, pp. 298－305, 1994.

저 자 소 개

고상준(Sang-Jun Ko)

2016년 한국산업기술대학교 컴퓨터공학과 학사
2018년 성균관대학교 소프트웨어플랫폼학과 석사
2018년-현재 엘에스웨어(주) 선임
<주관심분야> 코드 유사도, 악성코드 탐지, 블록체인

윤호영(Ho-Yeong Yun)

2012년 한성대학교 산업경영공학과 학사
2016년-2018년 엘에스웨어(주) 선임
2012년- 현 재 연세대학교
산업공학과 박사과정
<주관심분야> 최적화 이론, 알고리즘

신동명(Dong-Myung Shin)

2003년 대전대학교 컴퓨터공학과 박사
2001년-2006년 한국정보보호진흥원
응용기술팀 선임연구원
2006년-2014년 한국저작권위원회
저작권기술팀 팀장
2014년-2016년 한국스마트그리드사업단 보안인증팀 팀장
2016년-현재 엘에스웨어(주) 연구소장/이사
<주관심분야> 오픈소스 라이선스, 시스템/네트워크보안, SG인증/보안, SW취약점분석・감정

www.kci.go.kr