RESEARCH

KEPCO Journal on Electric Power and Energy
p-ISSN 2465-8111, e-ISSN 2466-0124
Volume 7, Number 1, June 2021, pp. 115-123
DOI 10.18770/KEPCO.2021.07.01.115

Development of Comparative Verification System for Reliability Evaluation of Distribution Line Load Prediction Model

배전 선로 부하예측 모델의 신뢰성 평가를 위한 비교 검증 시스템

Haesung Lee, Byung-Sung Lee, Sang-Keun Moon, Junhyuk Kim, Hyeseon Lee

Abstract

Through machine learning-based load prediction, it is possible to prevent excessive power generation or unnecessary economic investment by estimating the appropriate amount of facility investment in consideration of the load that will increase in the future or providing basic data for policy establishment to distribute the maximum load. However, in order to secure the reliability of the developed load prediction model in the field, the performance comparison verification between the distribution line load prediction models must be preceded, but a comparative performance verification system between the distribution line load prediction models has not yet been established. As a result, it is not possible to accurately determine the performance excellence of the load prediction model because it is not possible to easily determine the likelihood between the load prediction models. In this paper, we developed a reliability verification system for load prediction models including a method of comparing and verifying the performance reliability between machine learning-based load prediction models that were not previously considered, verification process, and verification result visualization methods. Through the developed load prediction model reliability verification system, the objectivity of the load prediction model performance verification can be improved, and the field application utilization of an excellent load prediction model can be increased.

Keywords: *Intelligent Power Distribution Grid, Artificial Intelligence, AI, Load Prediction Model, Machine Learning*

I. Introduction

EV, PV 등 신재생에너지의 계통 연계에 따른 복잡성과 불확실성이 더욱 증대된 배전 시스템 환경에서는 빅데이터가 접목된 인공지능 기술이 배전망 운영의 선제적 대응, 새로운 수익 창출, 투자비 감소 등 다양한 측면으로 안정적이고 효율적인 계통운영에 많은 도움이 될 것이다. 특히, 인공지능 기반 부하예측을 통해 향후 증가할 부하를 고려하여 적정 설비 투자액을 추정하거나, 최대 부하를 추정하여 필요한 전력량을 미리 파악하고 추정된 최대 부하를 분산시키기 위한 정책 수립의 기초자료를 제공함으로써 과도한 전력이 생산되거나 계통운영을 위한 불필요한 경제적 투자를 방지할 수 있다 [1]-[3]. 인공지능 기술의 발전으로 배전 선로 부하예측 모델에 적용 가능한 기계학습 알고리즘들은 더욱 다양해지고, 예측 성능은 기계학습 알고리즘의 성능 발전과 더불어 고도화되고 있다. 현재까지 개발된 대부분의 배전 선로 부하예측모델은 실제 측정값을 기반으로 학습 및 추론이 이루어지는 지도학습을 기반으로 배전망 운영을 통해 획득되는 데이터의 특성에 따라 Decision tree, Regression, Random Forest 등 다양한 기계학습 알고리즘이 적용될 수 있으며, 각 기계학습 알고리즘의 성능은 학습에 사용되는 데이터의 전처리 방법이나 학습 데이터 구성 방식에 따라 크게 좌우된다 [4]-[6]. 즉, 부하예측의 성능은 기계학습 알고리즘의 학습 작동 방식과 학습 데이터의 특성에 따라 결정되는 것이다. 배전 선로의 경우, 선로가 위치한 지역적 특성이나 수용가 특성에 따라 부하 패턴이 다르며, 획득되는 데이터 특성 역시 선로별로 달라진다. 즉, 특정 부하 패턴을 가지는 A 선로의 부하예측에 우수한 성능을 보였던 어떠한 부하예측 모델이 또 다른 부하 패턴을 갖는 B 선로의 부하예측에는 성능이 저하될 수 있으며, 학습 데이터의 특성, 학습 횟수, 학습 데이터의 양에 따라 동일한 기계학습 알고리즘을 기반으로 하고 있더라도 부하예측모델 간의 성능은 차이를 보이게 된다. 따라서, 서로 다른 부하 패턴을 가지는 특정 선로에 최적의 예측 성능을 보이는 부하예측 모델의 선별을 통해 현장 적용의 신뢰성을 확보하기 위해서는 부하예측 모델들 간의 다각적인 성능 비교 검증이 반드시 필요하다. 그러나 현장 적용의 신뢰성을 확보하기 위해서는 배전 선로 부하예측 모델들 간의 성능 비교 검증이 반드시 선행되어야 하지만 아직까지 성능 검증 방법, 검증 프로세스 등 배전선로 부하예측 모델 간 비교 성능 검증 체계가 마련되어 있지 않다.

결과적으로, 우수한 성능을 보이는 부하예측모델이라 하더라도 이미 특정 선로에 현장 적용 중인 부하예측 모델이나 다른 기

Article Information

Manuscript Received September 16, 2020, Accepted October 20, 2020, Published online June 30, 2021

The Authors are with KEPCO Research Institute, Korea Electric Power Corporation, 105 Munji-ro Yuseong-gu, Daejeon 34056, Republic of Korea.

Correspondence Author: *Haesung Lee* (*seastar@kepco.co.kr*)

This paper is an open access article licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International Public License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-nd/4.0
This paper, color print of one or more figures in this paper, and/or supplementary information are available at http://journal.kepco.co.kr.

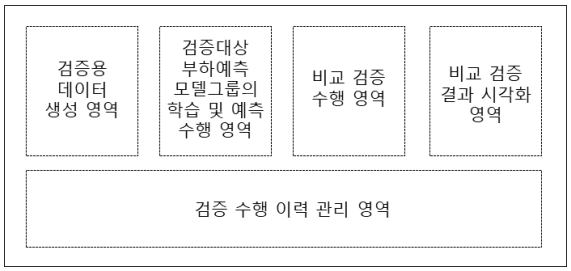


Fig. 1. Configuration of load prediction model performance comparison verification system.

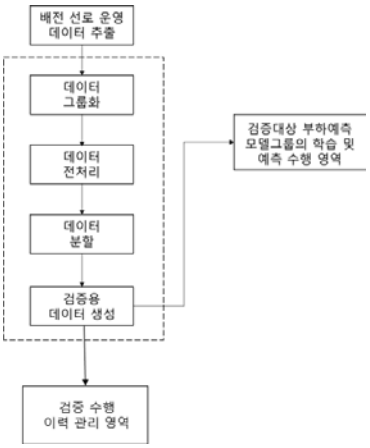


Fig. 2. Process flow diagram of 'generating data for verification.'

계학습 알고리즘을 기반한 부하예측 모델과의 비교 검증 없이는 부하예측 모델 간의 우의를 쉽게 가늠할 수 없어 현장 적용의 신뢰성 판단을 내릴 수 없다. 더욱이 데이터 분석이나 기계학습에 대한 지식 및 경험이 부족한 배전망 운영원이나 배전계획 등을 수행하는 의사 결정자의 경우에는 기계학습 기반 부하예측 모델의 성능 신뢰성에 대한 객관적인 판단을 내리는 것이 불가하다. 때문에, 우수한 성능을 보이는 부하예측 모델이 개발되어도 현장 적용을 위한 신뢰성 확보가 어려워 현장 활용이 매우 어려운 실정이다.

본 논문에서는 기존에 고려되지 않았던 기계학습 기반 부하예측 모델 간 성능 신뢰도 비교 검증 방법을 제안하고 부하예측 모델의 신뢰성 검증 시스템을 개발하여 부하예측 모델 성능 검증에 대한 객관성을 높이고, 우수한 부하예측 모델의 현장 적용 활용도를 높이고자 한다. 본 논문의 구성은 다음과 같다. 2장에서는 기계학습 기반 배전 선로 부하예측 모델의 성능평가 방법에 대해 논하고, 부하예측 모델의 성능 비교 검증 기법을 제안한다. 3장에서는 제안한 선로 부하예측 모델 성능 비교 검증 기법을 기반으로 부하예측 모델 성능 비교 검증 시스템을 개발한다. 마지막으로 4장에서는 결론에 대해 논한다.

II. 기계학습 기반 부하예측 모델 신뢰성 비교 검증

본 논문에서는 기존에 고려되지 않았던 기계학습 기반 부하예측 모델 간 성능 신뢰도 비교 검증 기법을 제안하여 부하예측 모델 성능 검증에 대한 객관성을 높이고, 우수한 부하예측 모델의 현장 적용 활용도를 높이고자 한다. 본 논문의 배전 선로 부하예측 모델 성능 비교 검증 시스템의 구성은 Fig. 1과 같다.

TABLE 1
Example of Grouping Data for Verification

대분류	세분류
기운용 모델의 예측 정확도	정확도 0% 이상 ~ 30% 미만 그룹 정확도 30% 이상 ~ 60% 미만 그룹 정확도 60% 이상 ~ 80% 미만 그룹 정확도 80% 이상 ~ 90% 미만 그룹
선로 부하 패턴	정확도 90% 이상 ~ 100% 그룹 일정 주기 패턴 반복 그룹 불규칙 추세 그룹
절체 패턴	이상 추세 그룹 절체 패턴 포함 그룹
PV 부하 패턴	절체 패턴 미포함 그룹 PV 부하 패턴 포함 그룹
EV 부하 패턴	PV 부하 패턴 미포함 그룹 EV 부하 패턴 포함 그룹 EV 부하 패턴 미포함 그룹
수용가 증감 패턴	수용가 증감 패턴 포함 그룹 수용가 증감 패턴 미포함 그룹

서로 다른 기계학습 알고리즘을 기반으로 하거나 서로 다른 부하 패턴을 갖는 데이터로 학습을 수행한 부하예측모델 간 성능 비교 검증 결과의 신뢰성 확보를 위해서는 예측 정확도에 대한 비교 검증이 동일한 기준과 환경에서 이루어져야 한다. 이를 고려하여, 본 논문에서 제안하는 검증 시스템은 크게 '검증용 데이터 생성 영역', '검증 대상 부하예측 모델의 학습 및 예측 수행 영역', '비교 검증수행 영역' 그리고 '비교 검증 결과 시각화 영역'의 5개 영역으로 구성된다.

'검증용 데이터 생성 영역'에서는 특정 부하 패턴을 갖는 비교 검증용 데이터를 생성하여 검증 대상 부하예측 모델의 학습 및 예측 수행 영역'에서 검증 대상이 되는 부하예측 모델 별로 동일한 검증용 데이터로 학습을 수행하고 예측 결과를 도출한다. '비교 검증수행 영역'에서는 부하예측 모델별로 도출된 예측 결과를 기반으로 본 논문에서 제안하는 검증 지표를 적용하여 각 부하예측 모델별 성능 신뢰도를 산출하고 이를 비교 성능 검증 근거로 활용한다. '비교 검증 결과 시각화 영역'에서는 부하예측 모델 별로 산출된 신뢰도를 다각도로 시각화 하여 데이터 분석이나 기계학습에 대한 지식 혹은 경험이 부족한 배전운영 의사결정자로 하여금 손쉽게 부하예측모델 간 성능 비교 검증이 가능하도록 한다. '검증수행 이력관리 영역'은 검증을 완료한 부하예측 모델의 검증 결과정보를 관리하여 기개발된 부하예측 모델의 성능 개선이나 새로운 부하예측 모델 개발을 목적으로 검증 이력 정보가 활용될 수 있도록 한다.

A. 검증용 데이터 생성

동일한 기계학습 알고리즘을 기반으로 한 부하예측 모델들이라 하더라도 학습을 수행한 데이터의 특성에 따라 부하예측 모델 간 예측 성능에 차이를 보이게 되며, 반대로 동일한 학습 데이터로 학습을 수행하였더라도 기계학습 알고리즘별로 서로 다른 예측 성

능에 차이를 보인다. 때문에 '검증용 데이터 생성' 단계에서는 서로 다른 기계학습 알고리즘을 기반으로 하거나 서로 다른 패턴을 갖는 데이터로 학습을 수행한 부하예측 모델 간 비교 검증 신뢰성을 확보하기 위해, 비교 검증 대상이 되는 부하예측 모델 그룹에 대해 다양한 패턴을 갖는 데이터를 생성하여 검증용 데이터로 제공하게 된다.

Fig. 2는 '검증용 데이터 생성 영역'의 수행 프로세스 흐름도를 보인다. '검증용 데이터 생성 영역'의 '데이터 그룹화' 단계에서는 별도로 구축된 선로 운영 데이터베이스에서 추출된 배전 선로 운영 데이터를 TABLE 1과 같은 그룹으로 데이터 특성에 따라 그룹화한다. 검증용 데이터의 그룹화를 통해 데이터 특성별로 검증용 데이터 그룹을 나누고, 다양한 그룹의 데이터에 대한 부하예측모델 간 성능 비교를 통해 데이터 특성별로 상이한 부하예측 결과에 대한 비교 검증이 가능하다.

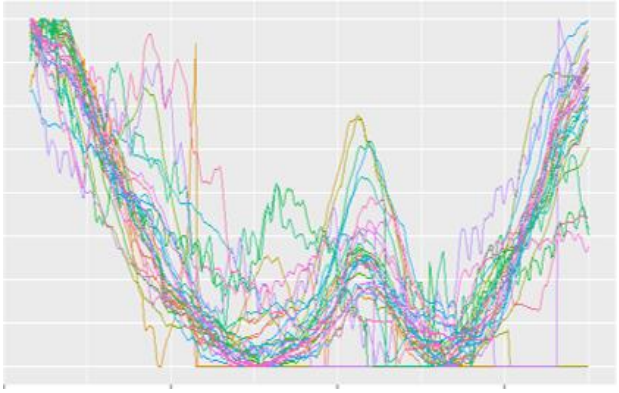
즉, 특정 데이터 패턴에 보다 높은 예측 결과를 보이는 부하예측 모델을 선별해 낼 수 있으며, 비교 검증 결과에 대한 신뢰성을 보다 높일 수 있다. 예를 들어, TABLE 1의 검증용 데이터 그룹화 예시에서 기운용 모델의 예측 정확도에 따른 데이터 그룹은 기운용 중인 부하예측모델의 예측 정확도에 따라 데이터 그룹을 나누어, 학습 데이터의 품질에 따른 부하예측모델 간 성능 비교 검증을 가능하도록 한다.

학습 데이터의 품질에 따른 부하예측 모델의 성능 결과 제공을 통해 상대적으로 우수한 정확도를 가지는 부하예측 모델을 판별해낼 수 있을 뿐만 아니라 학습 데이터 전처리를 통한 부하예측 모델의 예측 성능 개선 방안을 제공할 수 있을 것이다. TABLE 1의 '선로 부하 패턴' 그룹의 경우, 전국 단위의 선로별 부하 데이터를 대상으로 Fig. 3과 같이 부하 패턴별로 그룹화한다.

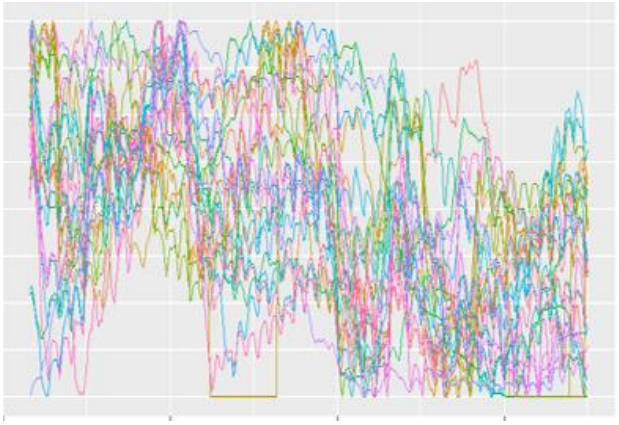
Fig. 3과 같은 선로 부하 패턴별 검증용 데이터 제공을 통한, 부하예측 모델 간 비교 검증을 수행하여 특정 선로 부하 패턴에 상대적으로 우수한 성능을 보이는 부하예측모델을 선별해 낼 수 있다. 예를 들어, 일정 주기로 반복적인 부하 패턴을 보이는 데이터 그룹에서 우수한 학습 성능을 보인 A 부하예측 모델의 경우, 불규칙한 부하 패턴을 보이는 데이터 그룹에서는 낮은 학습 성능을 보일 수 있으며, 일정 주기로 반복적인 부하 패턴을 보이는 데이터 그룹에서 A 부하예측 모델보다 낮은 학습 성능을 보인 B 부하예측 모델의 경우, 불규칙한 부하 패턴을 보이는 데이터 그룹에서는 A 부하예측 모델보다 높은 학습 성능을 보일 수 있다. 이 경우, 불규칙한 부하 패턴을 보이는 선로의 부하예측을 위해서는 B 부하예측 모델이 적합하며, 일정 주기로 반복적인 부하 패턴을 보이는 선로의 부하예측을 위해서는 A 부하예측 모델이 적합한 것으로 비교 검증을 통해 판단될 수 있다.

'검증용 데이터 생성 영역'의 '데이터 전처리' 프로세스 단계에서는 그룹화된 데이터를 검증용 데이터로 변환하기 위한 전처리를 수행한다. 전처리는 데이터 형태에 따라 이상치에 대한 보정과 결측치에 대한 데이터 대체를 수행하게 된다. '데이터 분할' 프로세스 단계에서는 전처리 완료된 데이터를 기계학습용과 테스트용 데이터로 분할한다. 분할된 기계학습용 데이터는 '검증 대상 부하예측 모델 그룹의 학습 및 예측수행 영역'에서 부하예측 모델의 학습을 위해 활용되며, 테스트 데이터는 비교 성능 검증의 판단 기준이 되는 예측 결과를 얻는 데 활용된다.

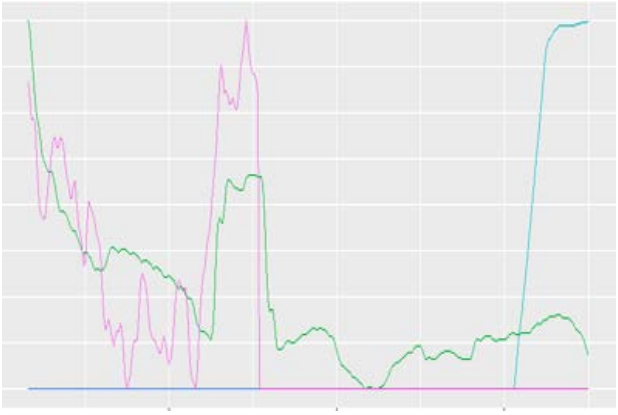
'데이터 분할' 프로세스 단계에서 '학습용 데이터'와 '테스트용 데이터'를 구성하여 최종적으로 검증용 데이터를 생성하고, 생성된 데이터는 '검증 수행 이력관리 영역'에 포함된 데이터 저장소에 저장되게 된다.



(a)



(b)



(c)

Fig. 3. Data grouping according to lone load pattern characteristics. (a) 일정 주기 패턴 반복 그룹 특성. (b) 불규칙 추세 그룹 특성. (c) 이상 추세 그룹 특성.

B. 검증대상 부하예측 모델 그룹의 학습 및 예측 수행

부하예측 모델들 간 예측 성능 비교 검증을 위해서는 '검증용 데이터 생성' 단계에서 생성된 검증용 데이터를 활용하여 부하예측 모델별로 학습을 수행하고, 학습된 부하예측 모델별로 예측을 하여, '비교 검증 수행영역'에서 비교 검증을 위한 부하예측 모델별 예측성능 지표를 산출하여야 한다.

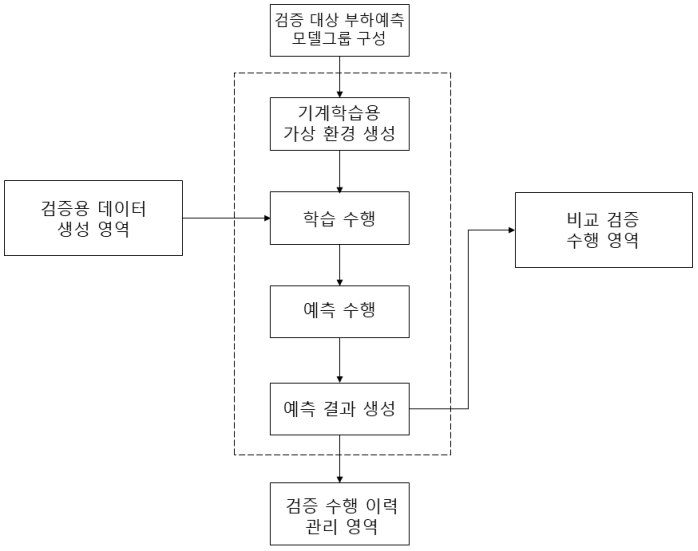


Fig. 4. Process flow chart of 'training and predicting the load prediction model group to be verified.'

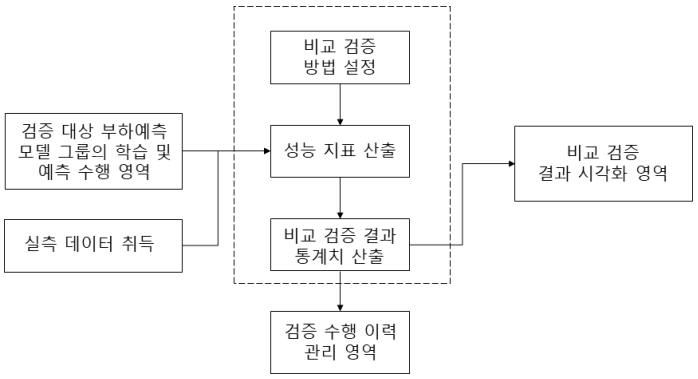


Fig. 5. Process flow chart of 'comparison verification.'

Fig. 4는 '검증 대상 부하예측 모델그룹의 학습 및 예측수행'의 프로세스 흐름도를 보인다. 부하예측 모델 간 예측 성능 비교 검증을 위하여 서로 다른 개발환경에서 개발된 각각의 부하예측 모델들의 학습과 예측을 수행하기 위해서는 각 부하예측 모델별로 학습과 예측을 수행하기 적합한 수행 환경이 제공되어야 한다. 본 논문의 부하예측모델 예측 성능 비교 검증 장치에서는 비교 검증 대상에 속한 각 부하예측 모델의 학습 환경에 적합한 가상환경(virtual environment)을 생성하여, '검증용 데이터 생성 영역'에서 생성된 데이터로 학습과 예측을 수행하여, 비교 검증을 위한 예측 결과를 생성한다. 예를 들어, 검증 그룹 A에 속한 3개의 부하예측 모델의 기계학습 환경이 모두 상이하다면 부하예측 모델별로 3개의 가상환경을 생성하여, 생성된 각각의 가상환경에서 '검증용 데이터 생성 영역'에서 생성된 데이터로 학습을 수행하게 된다. 학습이 완료되면 예측을 수행하고, '검증용 데이터 생성 영역'에서 생성된 '테스트용 데이터'를 바탕으로 예측 결과를 생성하게 된다. 부하예측 모델별 예측 결과는 '비교검증 영역'에서 비교 검증을 위한 기초자료로 활용되고, '검증수행 이력 관리 영역'에 포함된 데이터 저장소에 검증 이력 관리를 위한 목적으로 저장된다.

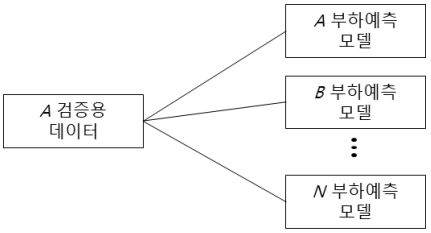


Fig. 6. Performance comparison verification between load prediction models based on specific line data.

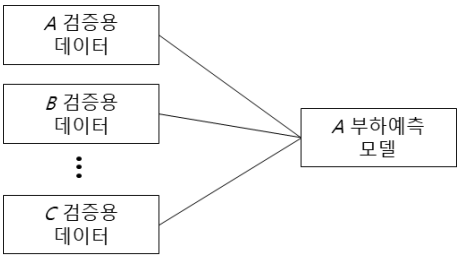


Fig. 7. Comparison of load prediction performance by load pattern.

C. 비교 검증 수행

'비교 검증 수행 영역'에서는 '검증대상 부하 예측 모델 그룹의 학습 및 예측 수행 영역'에서 생성된 부하 예측 모델들의 예측 결과에 대한 성능 지표를 산출하여 비교 검증을 수행한다. Fig. 5는 '비교 검증 수행 영역'의 프로세스 흐름도를 보인다. '비교 검증수행 영역' 프로세스의 '비교 검증 방법 설정' 단계에서는 비교 검증 방법을 결정한다. 서로 다른 기계학습 알고리즘을 기반으로 개발된 부하 예측 모델들이나 서로 다른 부하 패턴을 보이는 데이터로 학습을 수행한 부하 예측 모델들 간 예측 성능 비교 검증의 신뢰성을 보장하고, 특정 배전 선로에 최적의 예측 성능을 보이는 부하 예측 모델을 선별에 내기 위하여, 본 논문에서는 2가지의 부하 예측 모델 성능 비교 검증 방법을 제안한다.

1) 부하예측 모델 성능 비교 검증 기법

① 특정 선로 최적화 부하예측 성능 비교 검증: 서로 다른 기계학습 알고리즘을 기반으로 하는 부하예측 모델 간의 예측 성능을 비교 검증하는 방법으로 특정 선로나 EV 또는 PV 등을 포함하는 특정 부하 패턴에 최적의 성능을 보이는 부하예측 모델 선별이 가능하다. 즉, Fig. 6과 같이 특성 선로의 부하 데이터 한 세트 혹은 특정 부하 패턴을 갖는 데이터 한 세트를 검증용 데이터로 하여 서로 다른 기계학습 알고리즘을 기반으로 하는 *N*개의 부하예측 모델들 간 비교 검증을 수행한다.

② 부하 패턴 별 부하 예측 성능 비교 검증: 여러 배전 선로의 부하 데이터 또는 다양한 부하 패턴 별로 학습을 수행한 단일 부하 예측 모델의 예측 성능을 비교 검증하는 방법으로, 특정 부하 예측 모델이 어떠한 배전 선로 혹은 어떠한 부하 패턴에 유효한 예측 성능을 보이는지 판별 가능하며, 비교 검증 결과는 해당 부하 예측 모델의 성능 개선을 위한 기초 정보로 활용될 수 있다. 즉, Fig. 7과 같이 여러 배전 선로의 부하 데이터 세트들 혹은 다양한

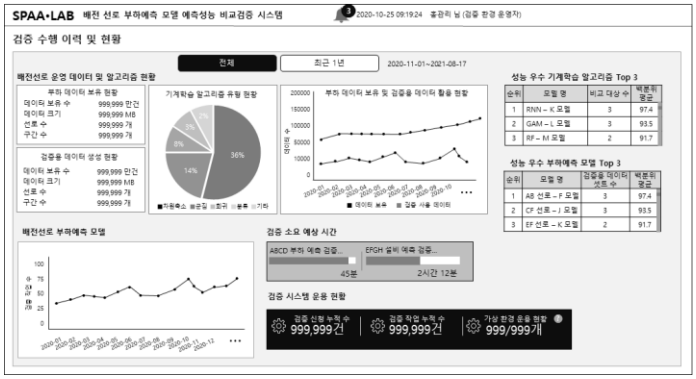


Fig. 8. Visualization of verification performance history and status.

부하 패턴 별 데이터 세트들에 대해 단일의 부하 예측 모델이 학습을 수행하여 다양한 부하 패턴별 부하 예측 모델의 예측 성능에 대한 비교 검증을 수행한다.

'비교 검증 방법 설정'의 다음 단계인 '성능 지표 산출' 단계에서는 '검증 대상 부하 예측 모델 그룹의 학습 및 예측 수행 영역'에서 생성된 부하 예측 결과와 실측 부하 값을 비교하여 비교 검증을 위해 정의한 TABLE 2의 성능 지표 항목을 기반으로 종합적인 신뢰도 지표 값을 산출하게 된다. 본 논문에서는 기계학습 알고리즘의 성능 평가를 위해 사용되는 다양한 평가 방법들을 다각적으로 고려하여 검증 신뢰도를 높이고자 하였다 [7]-[10]. TABLE 2는 본 논문에서 부하예측모델 간 예측 신뢰도 비교 검증을 위해 정의한 검증 지표 항목이다.

비교 검증 대상이 되는 각 부하 예측 모델 별로 지표 값들을 산출하게 되면, '특정 선로 최적화 부하 예측 성능 비교 검증'과 '부하 패턴 별 부하 예측 성능 비교 검증'의 비교 검증 방법에 따라 M개의 검증용 데이터 세트별로 N개의 모델에 대한 성능 검증 지수와 오차에 대한 통계 등을 생성한다. 최종 산출된 성능 지표 값과 통계 값들은 '검증 수행 이력 관리 영역'에 저장되어, 이력 관리를 목적으로 활용될 수 있다.

D. 비교 검증 결과 시각화

'비교 검증 결과 시각화 영역'에서는 부하 예측 모델 별로 산출된 신뢰도를 다각도로 시각화 하여 데이터 분석이나 기계학습에 대한 지식 혹은 경험이 부족한 배전 운영 의사결정자로 하여금 손쉽게 부하예측모델 간 성능 비교 검증이 가능하도록 한다. 비교 검증 결과 시각화 영역'은 크게 검증 이력이나 생성된 검증용 데이터에 대한 종합 정보를 보여주는'검증 수행 이력 및 현황 시각화'와 '비교 검증 결과 시각화'로 구성되며, '비교 검증 영역'에서 산출된 M개의 검증용 데이터 세트에 대한 N개의 부하 예측 모델 간 예측 성능 비교 검증 결과를 다각도로 시각화 하여 표현한다.

1) 검증 수행 이력 및 현황 시각화

Fig. 8은 '검증 수행 이력 및 현황 시각화' 설계 화면을 보인다. '검증 수행 이력 및 현황 시각화'에서는 6개 항목의 검증 수행 이력 및 현황 정보를 시각화 한다.

① 배전선로 운영 데이터 및 알고리즘 현황: '배전 선로 부하 예측 모델 예측 성능 비교 검증 시스템'에서 보유하고 있는 부하

TABLE 2
Load Prediction Model Performance Verification Indicators

검증 지표 항목	지표 산출 방법
MAPE (Mean Absolute Percentage Error)	▌실측 부하 값과 예측 부하 값의 편차(절댓값)를 부하 측정값으로 나누고 그 평균을 계산 ▌값이 클수록 예측 성능이 낮은 것을 의미 ▌MAPE 산출식 $MAPE_{\%} = \frac{1}{N}\sum_{t=1}^{N}\left\lvert\frac{y_t - \hat{y}_t}{y_t}\right\rvert \times 100$ $y_t =$ 실측 부하값, $\hat{y}_t =$ 예측 부하 값 $N =$ 부하데이터 수
NRMSE (Normalized Root Mean Square Error)	▌실측 부하 값과 예측 부하 값의 차이를 정규화 함 ▌NRMSE 산출식 $NRMSE_{\%} = \frac{\sqrt{\frac{1}{N}\sum_{t=1}^{N}(y_t - \hat{y}_t)^2}}{mean(y_t)} \times 100$ $y_t =$ 실측 부하값, $\hat{y}_t =$ 예측 부하 값, $N =$ 부하 데이터 수
M-NRMSE (Modified-Normalized Root Mean Square Error)	▌NRMSE에 최대 부하 측정값 대비 예측 부하 값의 비율을 가중치로 하여 최대 부하 측정값에 근사한 예측 부하 값들에 대한 중요도를 높임 ▌값이 클수록 예측 성능이 낮은 것을 의미
MPE (Mean Percentage Error)	▌산출 값이 양수 일 경우, 예측 부하 값이 실측 부하 값보다 낮게 분포함을 의미하고, 음수일 경우에는 예측 부하 값이 실측 부하 값보다 높게 분포함을 의미 ▌MPE 산출식 $MPE_{\%} = \frac{1}{N}\sum_{t=1}^{N}\frac{y_t - \hat{y}_t}{y_t} \times 100$ $y_t =$ 실측 부하값, $\hat{y}_t =$ 예측 부하 값, $N =$ 부하 데이터 수
Adj. R-Square	▌산출 값이 1에 가까울수록 부하 예측 모델의 학습 성능이 좋은 것으로 판단될 수 있고, 산출 값이 0에 가까울수록 학습 성능이 떨어지는 것으로 판단될 수 있다. ▌Adj. R-Square 산출식 $Adj.R^2 = 1 - \frac{(n-1)(1-R^2)}{n-p-1}$ $R^2 = 1 - \frac{\sum_{t=1}^{N}(y_t - \hat{y}_t)^2}{\sum_{t=1}^{N}(y_t - \bar{y})^2}$ $y_t =$ 실측 부하값, $\hat{y}_t =$ 예측 부하 값, $N =$ 부하 데이터 수
PMR (Predict Model Reliability)	▌검증 지표 항목별 지표 값을 1과 100 사이의 값으로 정규화하고 모든 지표 값들을 산술 평균하여 종합 지표 값으로 산출 ▌정규화 $y_{nom} = \frac{y - \min(y)}{\max(y) - \min(y)} \times 100$ ▌PMR 산출식 $PMR_{\%} = \frac{1}{5}\sum_{i=1}^{5} y_{nom}^{i} \times 100$

Fig. 9. Visualization of performance comparison verification results for specific line load prediction.

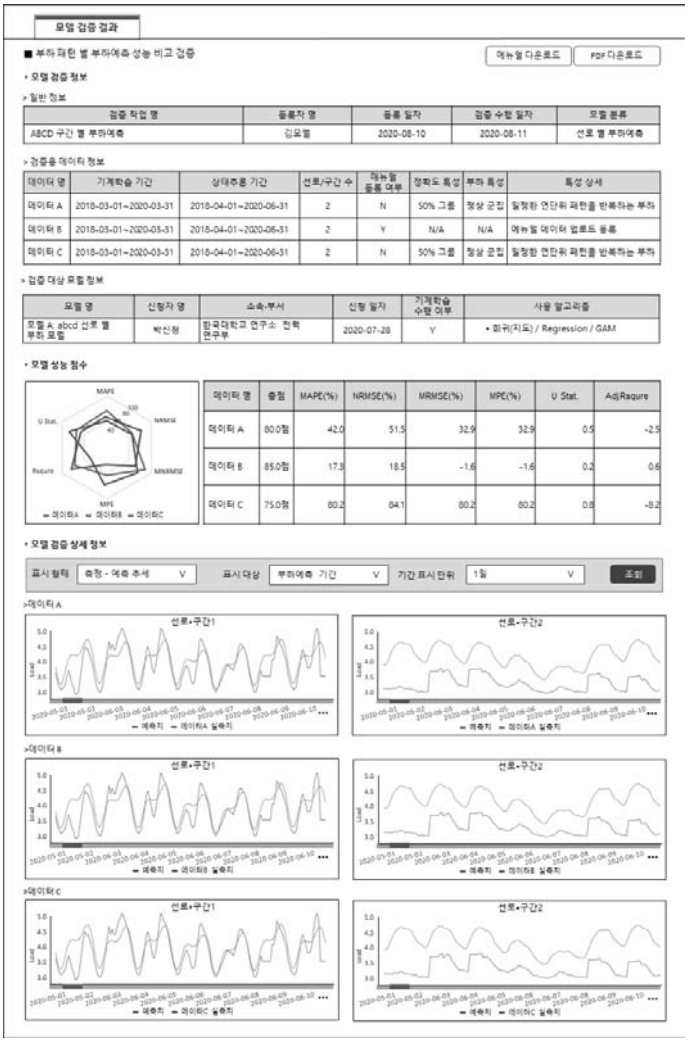


Fig. 10. Visualization of load prediction performance comparison verification by load pattern.

데이터의 구축 현황과 구축된 부하 데이터를 기반으로 생성된 검증용 데이터 현황을 시각화 하여 제공한다. 많은 부하 데이터를 보유하고 있다면, 이를 기반으로 양질의 검증용 데이터를 생성할 수 있고, 검증 결과의 신뢰성을 확보할 수 있을 것이다. 또한, '검증 수행 이력 및 현황 시각화' 화면에서는 기계학습 알고리즘 유형별 부하 예측 모델의 검증 완료 현황을 시각화 하여 기계학습 알고리즘별 활용 현황을 알 수 있다. '부하 데이터 보유 및 검증용 데이터 활용 현황'에서는 월별 부하 데이터 확보 현황과 검증용 데이터 활용 현황을 시각화 하여 시간에 따른 부하 데이터 확보 추이와 검증용 데이터 생성 추이 정보를 알 수 있다.

② 배전선로 부하 예측 모델 검증 수행 현황: '배전 선로 부하 예측 모델 예측 성능 비교 검증 시스템'에서 검증 수행 완료한 현황을 시각화 하여 시간에 따른 배전선로 부하 예측 모델의 검증 수행 완료 현황 추이를 파악 가능할 수 있도록 한다.

③ 성능 우수 기계학습 알고리즘 Top *K*: '배전 선로 부하 예측 모델 예측 성능 비교 검증 시스템'에서 검증을 완료한 부하 예측 모델 중 가장 우수한 성능의 상위 *K*개의 모델과 해당 모델이 기반으로 한 기계학습 알고리즘의 정보를 제공한다. 검증 성능이 우수한 기계학습 알고리즘의 정보 제공을 통해, 특정 배전 선로의 부하 패턴에 비교적 우수한 성능을 보이는 기계학습 알고리즘을 선별할 수 있으며, 이를 기반으로 예측 성능이 우수한 부하 예측 모델 개발이 가능하다.

④ 성능 우수 부하 예측 모델 Top *K*: '배전 선로 부하 예측 모델 예측 성능 비교 검증 시스템'에서 검증을 완료한 부하 예측 모델 중 특정 배전 선로에서 가장 우수한 예측 성능을 보인 상위 *K*개의 부하 예측 모델과 각각의 부하 예측 모델이 예측을 수행한 특정 선로의 정보를 제공한다. 특정 부하 예측 모델이 우수한 예측 성능을 보였던 특정 선로의 정보 제공을 통해, 특정 배전 선로에 대해 최적 예측 성능을 보인 부하 예측 모델을 선별할 수 있으며, 해당 부하 예측 모델의 현장 적용 활용도를 높일 수 있을 것이다.

⑤ 검증 소요 예상 시간 및 검증 시스템 운용 현황: '배전 선로 부하 예측 모델 예측 성능 비교 검증 시스템'에서 가장 마지막으로 검증 수행 중인 검증 대상 그룹의 검증 소요 예상 시간을 시각화 한다. 또한, 검증을 의뢰하거나 신청한 누적 건수, 검증을 완료한 누적 건수, 그리고 검증 대상인 부하 예측 모델의 학습이 수행되는 가상 환경의 운영 건수를 시각화 한다. 이를 통해, '배전 선로 부하 예측 모델 예측 성능 비교 검증 시스템'의 운영 현황을 파악해 낼 수 있다.

TABLE 3
Development environment

프로세스 명	개발 환경
검증 프로세스	Ubuntu 18.0.4, Python 3.7.4
검증 관리 프로세스	Ubuntu 18.0.4, apache tomcat 9, MariaDB 10.3, open JDK 8
시각화 프로세스	Ubuntu 18.0.4, apache tomcat 9, MariaDB 10.3, open JDK 8, ApexCharts

2) 비교 검증 결과 시각화

'비교 검증 결과 시각화'에서는 '특정 선로 최적화 부하 예측 성능 비교 검증'과 '부하 패턴 별 부하 예측 성능 비교 검증'의 두 가지 검증 방법에 따라 배전 선로 부하 예측 모델 간 비교 검증 결과 정보를 시각화 한다. '특정 선로 최적화 부하 예측 성능 비교 검증' 시각화는 하나의 특정 선로를 대상으로 부하 예측 모델 간 예측 성능에 대한 비교 검증 결과를 다각적으로 시각화 하여 사용자로 하여금 손쉽게 특정 선로에 가장 적합한 예측 성능을 보이는 부하 예측 모델을 선별해 낼 수 있도록 한다. '부하 패턴 별 부하 예측 성능 비교 검증' 시각화는 서로 다른 부하 패턴을 보이는 다수의 검증용 데이터 세트를 대상으로 다양한 부하 패턴 별 단일의 부하 예측 모델의 예측 성능을 시각화 하여 특정 부하 예측 모델이 어떠한 부하 패턴에 가장 최적의 성능을 보이는지 평가할 수 있다.

3) 특정 선로 최적화 부하예측 성능 비교 검증

'특정 선로 최적화 부하예측 성능 비교 검증' 시각화에서는 총 3개 항목에 대한 비교 검증 결과를 시각화 한다. Fig. 9는 '특정 선로 최적화 부하예측 성능 비교 검증'시각화 화면을 보인다.

① 모델 검증 정보: 검증을 완료한 검증 대상 그룹에 속한 *N*개의 부하예측 모델들이 사용한 기계학습 알고리즘, 검증용 데이터 세트의 특성에 대한 검증 정보를 시각화 한다. 수행을 완료한 검증에 대한 검증용 데이터의 특성과 검증 대상이 되는 부하예측 모델에 대한 정보를 제공함으로써 검증 결과의 신뢰성을 높일 수 있다.

② 모델 성능 점수: 특정 선로의 부하 데이터에서 생성된 검증용 데이터를 기반으로 검증 대상 그룹에 속한 *N*개의 부하예측 모델 별로 산출된 검증 지표 값들을 시각화 한다. 각 부하예측 모델 별 검증 지표 값들을 단일의 레이다 차트(Radar chart)로 시각화 하여 사용자로 하여금 특정 선로의 가장 최적화된 부하예측 모델을 손쉽게 파악할 수 있도록 한다.

③ 모델 검증 상세 정보: 특정 선로 구간의 특성에 따른 *N*개의 부하예측 모델 별 예측 정확도를 예측 기간에 따라 시각화 하여 제공한다. 구간에 따른 부하예측 모델 간 예측 성능에 대한 비교 검증이 가능하므로, 특정 구간의 부하예측에 가장 적합한 부하예측 모델을 선별해 낼 수 있다.

4) 부하 패턴 별 부하예측 성능 비교 검증

'부하 패턴 별 부하예측 성능 비교 검증' 시각화에서도 총 3개 항목에 대한 비교 검증 결과를 시각화 한다. Fig. 10은 '부하 패턴 별 부하예측 성능 비교 검증' 시각화 화면을 보인다.

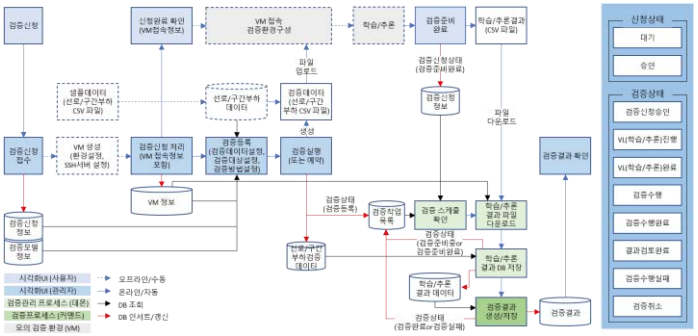


Fig. 11. Flow chart of line load prediction model reliability comparison verification process.

① 모델 검증 정보: 검증을 완료한 단일 부하예측 모델이 사용한 기계학습 알고리즘과 *N*개의 검증용 데이터 세트들의 특성에 대한 검증 정보를 시각화 한다. 수행을 완료한 검증에 대한 검증용 데이터의 특성과 검증 대상이 되는 부하예측 모델에 대한 정보를 제공함으로써 검증 결과의 신뢰성을 높이는 것을 목적으로 한다.

② 모델 성능 점수: *N*개의 선로들의 부하 데이터로 기반으로 생성된 *N*개의 검증용 데이터 세트 별 단일 부하예측 모델의 예측 결과에 예측 성능 검증 지표 값들을 시각화 한다. 각 *N*개의 검증용 데이터 세트 별 검증 지표 값들을 단일의 레이다 차트(Radar chart)로 시각화 하여 사용자로 하여금 특정 기계학습을 기반으로 하는 부하예측 모델에 가장 적합한 배전 선로를 파악할 수 있다.

③ 모델 검증 상세 정보: *N*개의 선로 구간 특성에 따른 부하예측 모델의 예측 정확도를 예측 기간에 따라 시각화 하여 제공한다. 선로 구간의 부하 패턴에 따른 예측 모델의 성능 검증이 가능하므로, 특정 기계학습 알고리즘을 기반으로 부하예측 모델이 어떠한 부하 패턴에 가장 최적의 예측 성능을 보이는지 검증해 낼 수 있다.

III. 선로 부하예측 모델 신뢰성 비교 검증 시스템 개발

본 장에서는 선로 부하예측 모델 간 신뢰성 비교 검증 시스템 개발에 대해 논한다. Fig. 11은 신뢰성 비교 검증 시스템의 업무 프로세스 흐름도를 보인다.

Fig. 11에서 보이는 바와 같이 프로세스는 크게 '시각화 프로세스', '검증 관리 프로세스' 그리고 '검증 프로세스'로 나뉜다. '시각화 UI'는 시스템에서 제공하는 검증 기능을 활용하는 사용자를 위한 사용자 화면과 시스템 관리자 화면으로 구성된다. '검증 관리 프로세스'는 검증 환경 설정과 검증 수행 이력 관리를 수행하며, '검증 프로세스'는 검증에 활용되는 학습 데이터를 생성할 뿐만 아니라 검증 대상이 되는 선로 부하 예측모델들에 대하여 학습과 예측을 수행하고, 예측 결과에 대한 검증을 수행한다. 검증 결과는 '사용자 시각화 프로세스'를 통해 사용자에게 제공된다. TABLE 3은 부하예측 모델 신뢰성 비교 검증 시스템의 개발 환경을 보인다.

본 논문에서 개발한 시스템은 운영체제인 Ubuntu 환경에서 Apache tomcat 기반의 웹 어플리케이션 서버로 구현되었으며, '시각화 프로세스'의 'ApexCharts'은 자바스크립트 기반의 오픈소스

```python
def mape(df_valid):
    y_true = df_valid['measureVal']
    y_pred = df_valid['predictVal']

    mape = np.mean(np.abs((y_true - y_pred) /
                          y_true).replace([np.inf, -np/np.inf], np.nan))
    return round(mape, 3)


def nrmse(df_valid):
    y_true = df_valid['measureVal']
    y_pred = df_valid['predictVal']
    RMSE = mean_squared_error(y_true, y_pred) ** 0.5
    mean_y = np.mean(y_true)

    nrmse = RMSE / mean_y
    return round(nrmse, 3)
```

Fig. 12. Source code of MAPE and NRMSE.

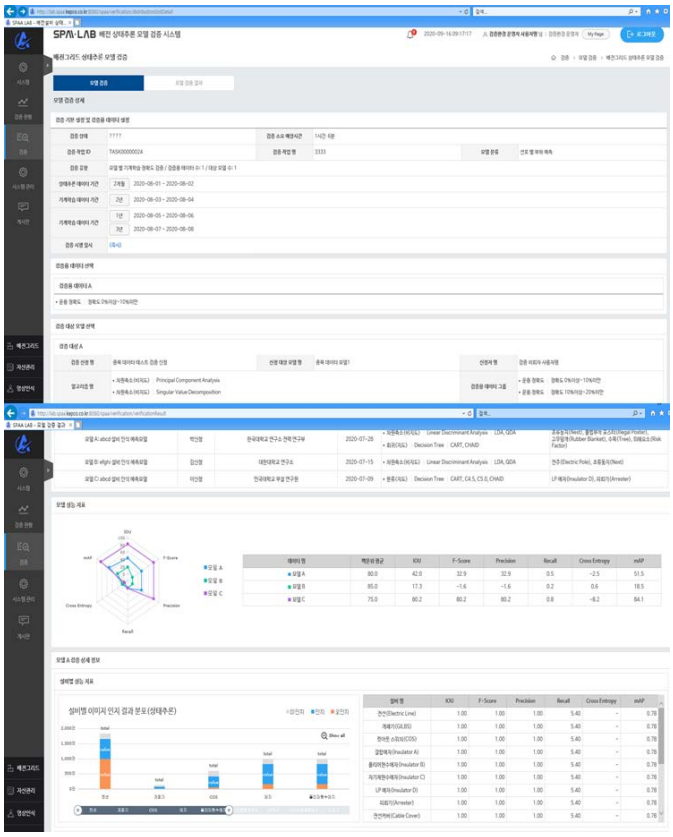


Fig. 13. System for comparing and verifying reliability of distribution line load prediction model.

차트 라이브러리로 본 논문에서는 검증 결과의 시각화 구현을 위해 사용되었다. '검증 프로세스'를 구성하는 단위 프로세스들은 Python을 통해 라이브러리로 개발하였다. Fig. 12는 TABLE 2에서 정의된 지표 항목 중 MAPE와 NRMSE에 해당하는 라이브러리의 소스 코드 일부를 보인다.

본 논문에서 개발된 검증 라이브러리를 통해 다양한 배전 인공지능 시스템에 쉽게 이식되거나 보다 확장성과 호환성을 제공하여 신뢰적인 배전 인공지능 어플리케이션 개발 활성화가 가능하다.

Fig. 13은 배전 선로 부하 예측 모델 검증 시스템 개발 화면을 보인다. 본 논문에서 개발한 배전 선로 부하예측 모델 검증 시스템을 통해 부하예측 모델의 성능 신뢰성을 확보하고, 부하예측 모델의 현장 적용 활용도를 획기적으로 높일 수 있을 뿐만 아니라, 데이터 분석이나 기계학습에 대한 지식 및 경험이 부족한 배전망 운영원이나 배전 계획 등을 수행하는 의사 결정자도 기계학습 기반 부하예측 모델의 성능 신뢰성에 대한 객관적인 판단을 손쉽게 효과적으로 내릴 수 있다.

IV. Conclusion

본 논문에서는 기존에 고려되지 않았던 기계학습 기반 부하예측 모델 간 성능 신뢰도 비교 검증 기법을 제안하고 부하예측 모델의 신뢰성 검증 시스템을 제안하여 부하예측 모델 성능 검증에 대한 객관성을 높이고, 우수한 부하예측 모델의 현장 적용 활용도를 높이고자 하였다. 본 논문에서 제안하는 '기계학습 기반 배전선로 부하예측 모델의 성능 신뢰성 비교 검증 기법'을 통해 부하예측 모델들 간의 다각적인 성능 비교 검증을 효율적으로 수행하여 부하예측 모델의 성능 신뢰성을 확보하고, 부하예측 모델의 현장 적용 활용도를 획기적으로 높일 수 있다. 또한, 본 논문에서 개발한 부하예측 모델 신뢰성 비교 평가 시스템을 데이터 분석이나 기계학습에 대한 지식 및 경험이 부족한 배전망 운영원이나 배전계획 등을 수행하는 의사 결정자도 기계학습 기반 부하예측 모델의 성능 신뢰성에 대한 객관적인 판단을 손쉽게 효과적으로 내릴 수 있다. 때문에, 우수한 성능을 보이는 부하예측 모델의 현장 적용을 위한 객관적 신뢰성 확보를 통해, 현장 활용도를 높일 수 있다. 결과적으로, 우수한 부하예측 모델의 현장 활용성을 높이면, 향후 증가할 부하를 고려하여 적정 설비 투자액을 추정하거나, 최대 부하를 추정하여 필요한 전력량을 미리 파악하고 추정된 최대 부하를 분산시키기 위한 정책 수립의 보다 정확한 기초자료를 제공할 수 있을 것이므로, 과도한 전력이 생산되거나 계통 운영을 위한 불필요한 경제적 투자를 방지할 수 있을 것이다. 본 논문에서는 배전선로 부하예측 모델의 성능 신뢰도 평가를 고려하였지만 향후 연구과제에서는 태양광이나 풍력 등 신재생 에너지 발전량 예측 모델과 EV(Electrical vehicle) 부하 예측 모델의 성능 신뢰도 평가 기법에 대한 연구개발을 진행하도록 하겠다.

Acknowledgment

This paper is the results of "Development of predictive technologies and diagnostic inference engines for distribution system operations" project supported by 2017 R&D project performed by Korea Electric Power Corporation.

본 연구는 한국전력공사 전력연구원의 2017년 자체과제로 수행한 '상태추론 기반 배전설비 예지 기술 및 엔진 개발' 연구과제의 기술개발 결과임.

References

[1] Warren, Chris "Can Artificial Intelligence Transform the Power System?" KEPCO Journal on Electric Power and Energy v.5. no.2, 2019.

[2] Negnevitsky, Michael, Paras Mandai, Anurag K. Srivastava, "Machine learning applications for load, price and wind power prediction in

power systems," 15th international conference on Intelligent System Applications to Power Systems, IEEE, 2009.

[3] Kelp, S. M., Sanjay V. Dudul. "Short-term Maharashtra state electrical power load prediction with special emphasis on seasonal changes using a novel focused time lagged recurrent neural network based on time delay neural network model," Expert System with Applications 38. 3, 2011.

[4] Hall, Mark A, Lloyd A. Smith, "Practical feature subset selection for machine learning," Proceeding of the 21st Australasian Computer Science Conference ACSC'98, 1998.

[5] Batista, Gustavo EAPA, Ronaldo C. Prati, Maria Carolina Monard. "A study of the behavior of several methods for balancing machine learning training data," ACM SIGKDD explorations newsletter 6.1, 2004.

[6] Hullemeier, Eyke. "Fuzzy methods in machine learning and data mining: Status and prospects," Fuzzy sets and Systems, 156 (3), 2005.

[7] De Myttenaere, Arnaud, Colden, B., Le Grand, B., Rossi F., "Mean absolute percentage error for regression models," Neurocomputing, 192. 2016.

[8] Voyant, C., Notton, G., Kalogirous, S., Nivet, M. L, Paoli C., Motte, F., Fouilloy, A., "Machine learning methods for solar radiation forecasting: A review," Renewable Energy, 105, 2017.

[9] Theocharides, Spyros., Makrides, G., Georghiou, G. E., Kyprianous, A., "Machine learning algorithms for photovoltaic system power output prediction," Proceeding of the 2018 IEEE International Energy Conference, 2018.

[10] De Myttenarare, Arnaud, Golden, B., Le Grand, B., Rossi, F, "Mean absolute percentage error for regression models," Neurocomputing, 192, 2016.