

Management Automation Technique for Maintaining Performance of Machine Learning-Based Power Grid Condition Prediction Model

기계학습 기반 전력망 상태예측 모델 성능 유지관리 자동화 기법

Haesung Lee, Byunsung Lee, Sangun Moon, Junhyuk Kim, Heysun Lee
이해성, 이병성, 문상근, 김준혁, 이혜선

Abstract

It is necessary to manage the prediction accuracy of the machine learning model to prevent the decrease in the performance of the grid network condition prediction model due to overfitting of the initial training data and to continuously utilize the prediction model in the field by maintaining the prediction accuracy. In this paper, we propose an automation technique for maintaining the performance of the model, which increases the accuracy and reliability of the prediction model by considering the characteristics of the power grid state data that constantly changes due to various factors, and enables quality maintenance at a level applicable to the field. The proposed technique modeled a series of tasks for maintaining the performance of the power grid condition prediction model through the application of the workflow management technology in the form of a workflow, and then automated it to make the work more efficient. In addition, the reliability of the performance result is secured by evaluating the performance of the prediction model taking into account both the degree of change in the statistical characteristics of the data and the level of generalization of the prediction, which has not been attempted in the existing technology. Through this, the accuracy of the prediction model is maintained at a certain level, and further new development of predictive models with excellent performance is possible. As a result, the proposed technique not only solves the problem of performance degradation of the predictive model, but also improves the field utilization of the condition prediction model in a complex power grid system.

초기 학습 데이터의 과적합으로 인한 전력망 상태예측 모델의 성능 감소를 방지하고 예측모델의 예측 정확도 유지를 통한 지속적인 현장 활용을 위해서는 기계학습 모델의 예측 정확도를 지속적으로 관리할 필요가 있다. 이를 위해, 본 논문에서는 다양한 요인에 의해 끊임없이 변화하는 전력망 상태 데이터의 특성을 고려하여 예측모델의 정확성과 신뢰성을 높이고 현장 적용 가능한 수준의 품질을 유지하기 위한 기계학습 기반 전력망 상태예측 모델의 성능 유지관리 자동화 기법을 제안한다. 제안 기법은 워크플로우 관리 기술의 적용을 통해 전력망 상태예측 모델 성능 유지관리를 위한 일련의 태스크들을 워크플로우의 형태로 모델링하고 이를 자동화하여 업무를 효율화 하였다. 또한, 기존 기술에서는 시도되지 않았던 학습데이터의 통계적 특성 변화 정도와 예측의 일반화 수준을 모두 고려한 예측모델의 성능 평가를 통해 성능 결과의 신뢰성을 확보하고 이를 통해 예측 모델의 정확도를 일정 수준으로 유지관리하고 더욱 성능이 우수한 예측모델의 신규 개발이 가능하다. 결과적으로 본 논문에서 제안하는 전력망 상태예측 모델 성능 유지관리 자동화 기법을 통해 예측모델의 성능 저하 문제를 해결하여 분산자원 연계 등 외부 환경의 변화에 유연한 예측모델 관리를 통해 정확성과 신뢰성이 보장된 예측 모델의 지속적인 활용이 가능하다.

Keywords: Intelligent Power Distribution Grid, Artificial Intelligence, AI, Load Prediction Model, Performance Evaluation, Programming Library

I. Introduction

현재까지 국내의 전력공급 체계는 정전횟수, 정전시간, 배전 손실률 등에서 프랑스, 영국, 미국 등과 비교할 때 매우 우수한 수준이며 안정적으로 전력망을 운영하고 있지만, 기후의 변화, 신재생에너지 공급 의무화 제도(Renewable portfolio standard, RPS)에 의한 막대한 용량의 신재생에너지 수용성, 전기 자동차의 급격한

수요팽창 가능성, 그리고 새로운 전기수요의 창출 등으로 인한 전력 산업의 급격한 변화에 대비하고 안정적인 배전망 운영을 위하여 빅데이터와 AI 기술 도입 필요성이 더욱 증대되었다 [1]. 전력 설비의 효율적인 관리와 정전 위험의 적시 감지, 전력 수요 및 서비스에 대한 효과적인 예측을 위해 빅데이터와 AI 기술이 활용되면서 날로 복잡해지는 전력망을 보다 경제적이고 효율적인 방향으로 운영 가능해진 것이다. 그러나, 시간의 흐름에 따라 특성 분포

Article Information

Manuscript Received September 9, 2020, Accepted September 21, 2020, Published online December 30, 2020

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@kepcoco.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.kepcoco.kr>.

가 계속적으로 변화하고 단시간에 대량으로 발생하는 전력 데이터의 특성으로 인해 전력망 상태예측 모델의 성능은 시간이 지남에 따라 지속적으로 감소하게 된다.

본 논문에서는 다양한 요인에 의해 끊임없이 변화하는 전력망 상태 데이터의 특성을 고려하여 지속해서 예측모델의 정확성과 신뢰성을 높이고 현장 적용 가능한 수준의 품질을 유지하기 위한 기계학습 기반 전력망 상태예측 모델의 성능 유지관리 자동화 방법'을 제안하고, 제안 기술이 적용된 상태예측 모델 성능 유지관리 자동화 시스템을 개발한다. 제안 방법은 워크플로우 관리 기술의 적용을 통해 전력망 상태예측 모델 성능 유지관리를 위한 일련의 태스크들을 워크플로우의 형태로 모델링하고 이를 자동화 시스템으로 구현하고자 하였다. 또한, 모델 검증에 있어 모델 성능 측정 단계에서 얻어진 오차 데이터를 기반으로 학습 데이터의 통계적 특성 변화가 예측 모델 정확도에 미치는 영향도를 평가를 위한 평균오차 변동률 산출식과 모델의 일반화 수준을 검증하기 위한 훈련-테스트 오차비율 변동률 산출식을 제안한다. 제안된 기법을 기반으로 한 상태예측 모델 유지관리과리를 통해 상태예측 모델의 정확성과 신뢰성을 높이고 현장 적용 가능한 수준의 상태예측 결과의 품질을 지속적으로 유지할 수 있다. 본 논문의 구성은 다음과 같다. II장에서는 전력망 상태예측 모델의 성능저하에 대해 논하고, III장에서는 전력망 상태예측 모델 유지관리 자동화 기법을 제안한다. IV장에서는 본 논문에서 제안한 전력망 상태예측 모델 유지관리 자동화 기법을 기반으로 전력망 상태예측 모델 유지관리 수행 워크플로우를 정의한다. V장에서는 결론에 대해 논한다.

II. 전력망 상태예측 모델의 성능 저하

전력망 상태예측을 위한 빅데이터와 AI 기술 대부분은 기계학습(Machine learning)을 기반으로 하고 있다. 기계학습은 크게 지도학습(Supervised learning)과 비지도학습(Unsupervised learning)으로 나눌 수 있다 [2]-[5]. 전력망 상태예측에 있어 예측하고자 하는 목표 값을 얻기 위해서는 예측 값에 대한 비교 평가를 할 수 있는 특성 데이터(패턴 또는 정답)를 포함하는 과거 이력 데이터가 필요하므로 거의 모든 전력망 상태예측 모델은 지도학습을 기반으로 만들어진다. 특히, 전력망 상태예측은 시계열 데이터를 학습하여 연속된 값(Continuous value)을 예측하는 회귀(regression) 문제이다. 3개월이나 1년 후 등 특정 선로의 미래 어느 시기의 최대부하를 예측하는 문제를 예로 들 수 있다. 연속된 값을 예측하는 회귀문제를 다루는 지도학습을 위해 주로 사용되는 방법은 선형회귀(Liner regression), 로지스틱 회귀(Logistic regression), 신경망(Neural network)이 있다.

일반적으로 회귀문제를 다루는 지도학습 기반의 예측모델은 예측한 값의 분포가 학습한 데이터 분포와 비슷할 때 정확도가 높다고 할 수 있다. 또한, 학습한 데이터가 아닌 새로운 데이터에 대해 정확한 예측을 했다면, 새로운 데이터에 대하여 일반화(Generalization)되었다고 한다 [6]. 즉, 새로운 데이터의 특성이 학습데이터의 특성과 비슷한 특성 분포를 가진다면 예측모델의 높은 예측 정확도를 기대할 수 있다. 반대로 새로운 데이터의 특성이 학습데이터의 특성과 다르다면 예측 정확도는 떨어지게 되며, 예측모델이 새로운 데이터에 일반화되지 못한 것이다. 새로운 데이터에 일반화되지 못하는 주된 요인은 학습데이터에 대한 과적합(overfitting)이다. 과적합은 예측모델이 학습데이터를 과하게 학습하여 새로운 데이터에 대해서는 예측 오차가 증가하는 현상으로

기계학습 알고리즘의 예측 정확도를 떨어뜨리는 가장 큰 원인이다. 또한, 예측 모델의 예측 정확도는 개발 초기에는 비교적 높은 성능을 보이다가 시간이 지날수록 정확도가 떨어지게 되는 현상이 나타나는데, 이러한 현상을 '개념 변화(Concept drift)'라 한다 [7]. 즉, '개념변화'는 시간이 지남에 따라 모델의 통계적 특성이 바뀌는 현상을 의미한다. 예측 모델은 모델링 대상의 통계적 특성을 학습하게 되는데, 시간이 지나면서 예측 대상의 통계적 특성이 바뀐다면 변화되기 이전에 특성만을 학습한 예측 모델은 예측 정확도가 떨어질 수 밖에 없는 것이다. 특히, 시간의 흐름에 따라 특성 분포가 계속적으로 변화하고 단시간에 대량으로 발생하는 전력 데이터의 특성으로 인해 전력망 상태예측 모델의 성능은 시간이 지남에 따라 지속적으로 감소하게 된다. 과적합과 개념변화로 인한 전력망 상태 예측모델의 성능 저하 문제의 해결은 다음과 같은 세 가지 이유로 인해 매우 어렵다. 첫째, 전력망 상태 예측모델이 학습하는 데이터는 실제 데이터의 부분 집합이며, 실제 데이터를 모두 수집하기에는 많은 어려움이 따른다. 둘째, 실제 데이터를 모두 수집하여도 모든 데이터를 학습시키기 위해서는 측정 불가한 수준의 시간과 높은 경제적 비용이 요구된다. 마지막으로 학습 데이터만을 가지고, 실제 데이터의 예측 오차가 발생하는 지점을 파악해내는 것은 매우 어렵다. 초기의 학습 데이터로 우수한 예측 정확도를 보인 전력망 상태예측 모델이라도 실제 데이터의 특성 변화로 인해 일반화 정도가 감소하면서 결과적으로 예측 결과에 대한 정확성과 신뢰성을 보장할 수 없으며, 예측모델의 현장 적용은 물론 지속적인 활용에 어려움이 따른다. 따라서 초기 학습 데이터의 과적합으로 인한 전력망 상태예측 모델의 성능 감소를 방지하고 예측모델의 예측 정확도 유지를 통한 지속적인 현장 활용을 위해서는 기계학습 모델의 예측 정확도를 지속적으로 관리할 필요가 있다.

III. 전력망 상태예측 모델 성능 유지관리 기법

전력망 상태예측 모델 성능 유지관리 과정은 신규 데이터에 대해 예측 모델의 품질을 다양한 매트릭스를 이용하여 주기적으로 측정하는 것으로 시작된다. 모델의 유형(classification 또는 regression)에 따라 다른 검증 매트릭스가 적용되며, 측정된 지표는 예측 모델의 개념변화 및 일반화의 검증을 위한 근거 데이터로 사용된다. 이와 같은 검증 단계를 통해 예측 모델의 현장 적용 가능 여부를 진단하고, 현장 적용에 요구되는 수준을 미충족할 시, 모델의 재학습과 리엔지니어링(모델 변경, 특징 재설계, 초매개변수 최적화 등)을 수행한다.

Fig. 1은 본 논문에서 제안하는 전력망 상태예측 모델 성능 유지관리 흐름도를 보인다. 먼저, 기계학습 서버에서 신규 데이터를 대상으로 모델 검증에 필요한 지표들을 측정한다. 예측모델 검증 태스크에서는 모델 별로 측정된 값들을 이용하여 데이터의 의미 분포 변화(concept drift) 발생 유무와 일반화 정도를 계산하고 이를 통해 예측 모델을 검증한다. 다음으로 모델 관리 의사결정 태스크에서는 기계학습 엔지니어가 예측모델의 측정 및 검증 데이터를 검토하여 모델 관리 방법을 결정한다. 마지막으로, 결정된 모델 관리 유형에 따라 예측 모델은 현상유지, 재학습 또는 리엔지니어링 단계를 거치게 된다.

A. 모델 측정

모델 측정 단계에서는 신규 학습 데이터를 대상으로 상태예

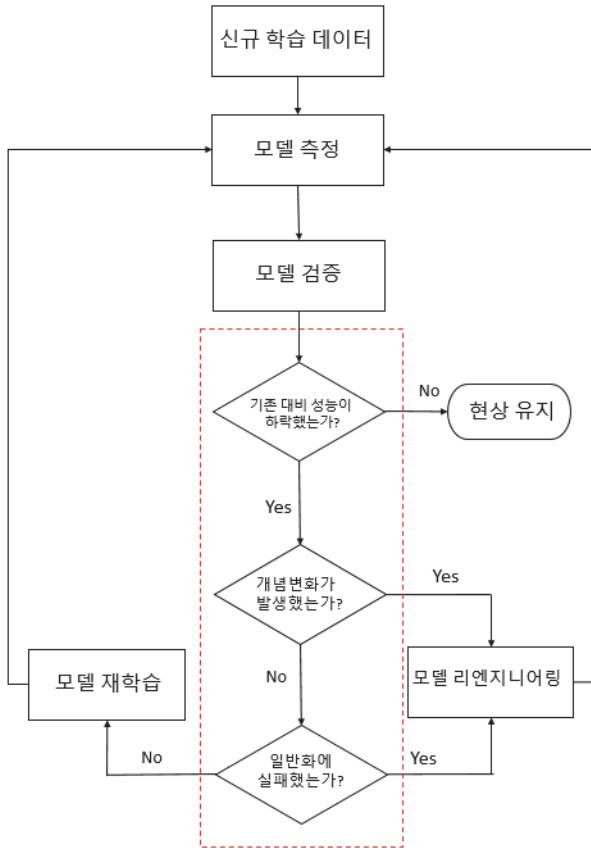


Fig. 1. 상태추론 모델 유지관리 수행 흐름도.

측 모델의 예측 신뢰도 및 정확도를 검증하기 위한 지표들을 측정한다. 모델 측정은 상태예측 모델이 기반으로 하고 있는 기계학습 모델 유형에 따라 다른 지표 산출방식이 적용된다. 기계학습 모델 유형에 따른 성능지표 산출방식은 다음과 같다.

- 분류 모델: 로그 손실(logarithm loss), 오차행렬(confusion matrix), AUC (Area Under Curve)
- 회귀 모델: 평균 절대 오차(MAE), 평균 제곱근 오차(RMSE), 결정계수(R-Squared)

결과적으로 모델 측정 단계에서 위의 정의된 성능 지표 방식을 적용하여 전력망 상태예측 모델의 예측 오차 값을 산출한다.

B. 모델 검증

모델 검증 단계에서는 측정 태스크에서 얻어진 오차 데이터를 이용하여 모델의 '개념변화 발생 여부'와 '일반화 실패 여부'를 판단한다.

1) 개념변화 발생 여부 판단

개념변화가 발생하면 상태예측 모델의 일시적인 성능저하가 아닌 지속적인 성능 저하로 이어지기 때문에 학습할 데이터의 개념변화 여부를 판단하여 상태예측 모델의 개선을 통해 성능을 유지할 수 있어야 한다. 신규로 발생한 학습 데이터의 개념 변화를 감지하기 위해, ω 의 크기를 갖는 윈도우를 사용한다. 특정 시점 t 에

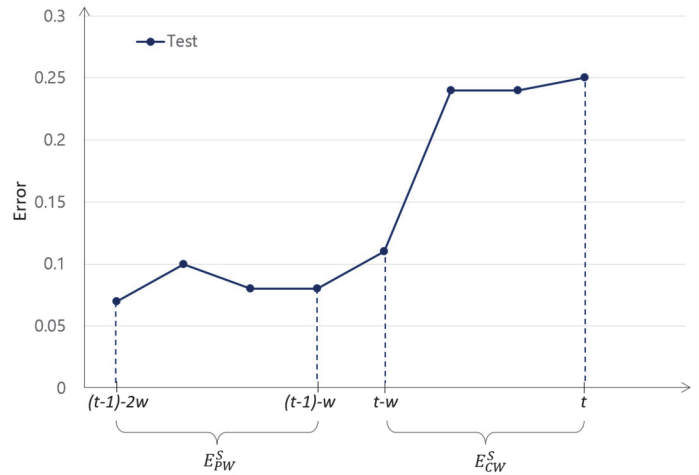


Fig. 2. 상태예측 모델의 오차 변화 추세.

서 측정된 상태예측 모델의 오차를 E_t^S , 윈도우 크기를 ω 라고 정의했을 때, 현재 윈도우 E_{CW}^S 와 직전 윈도우 E_{PW}^S 의 평균오차는 Eq. (1)과 같이 정의된다.

$$E_{CW}^S = \frac{1}{\omega} \sum_{i=t-\omega}^t E_i^S \tag{1}$$

$$E_{PW}^S = \frac{1}{\omega} \sum_{i=t-1-2\omega}^{t-1-\omega} E_i^S$$

평균오차 변동률인 (CME: Change rate of Mean Error)는 현재 윈도우와 직전 윈도우의 평균오차 간의 변동률로 정의되며 Eq. (2)와 같다. CME 값이 특정 임계 값보다 낮을 때 '개념변화'가 발생한 것으로 판단한다.

$$CME = \frac{E_{PW}^S - E_{CW}^S}{E_{PW}^S} \tag{2}$$

Fig 2는 현재 윈도우 E_{CW}^S 와 직전 윈도우 E_{PW}^S 에서의 상태예측 모델의 오차 변화 추세를 보이며, $t-\omega$ 의 시점에서 '개념 변화'가 발생한 것을 알 수 있다.

2) 일반화 실패 여부 판단

예측 모델의 일반화는 훈련 데이터가 아닌 관측되지 않은 향후 데이터에 대해 모델이 얼마나 잘 예측할 수 있는지를 나타낸다. 따라서 모델의 일반화 수준은 테스트 오차(test error, E^S)가 훈련 오차(train error, E^T)에 얼마나 근사하는 지로 정의할 수 있다. 두 오차 간의 차이가 적을수록 일반화 수준이 높은 모델로 간주되고, 반대로 훈련 데이터에 과적합된 모델의 경우 훈련 오차에 비해 테스트 오차가 상대적으로 높게 측정된다. 따라서 예측 모델의 일반화 수준을 판단하기 위한 기준으로 고려되는 훈련-테스트 오차 비율 변동률인 Eq. (3)의 CTTE (Change rate of Train-Test Error)은 현재 시점 t 에서 측정된 훈련-테스트 오차 비율과 직전 측정 시점인 $t-1$ 의 훈련-테스트 오차비율 간의 차이를 나타낸다. 결과적으로 CTTE 값이 특정 임계 값보다 낮을 때 일반화에 실패한 모델로 판단한다.



Fig. 3. 상태추론 모델의 훈련-테스트 오차 변화 추세.

$$CTTE = \frac{E_{t-1}^S}{E_{t-1}^T} - \frac{E_t^S}{E_t^T} \quad (3)$$

Fig. 3은 테스트 오차 E^S 와 훈련 오차 E^T 의 시간 t 의 흐름에 따른 추세를 보이면 $t-1$ 와 t 사이에서 일반화가 실패한 것을 확인할 수 있다.

C. 유지관리 방식의 결정

‘유지관리 방식의 결정’ 단계에서는 모델 측정 및 검증 태스크를 통해 산출된 데이터를 기계학습 엔지니어가 분석하고, 모델의 관리 방법을 결정한다. 예측 모델의 성능이 신규 학습 데이터에 대해서 우수한 성능을 보인다면 현상유지를 하고, 기존대비 약간의 성능 하락이 발생할 경우 모델 재학습을 고려한다. 만약 모델 검증 결과가 ‘개념변화 발생’ 또는 ‘일반화 실패’를 보이면 ‘모델의 리엔지니어링’을 고려한다.

D. 모델 재학습

모델 재학습 태스크는 신규 훈련 데이터에 대해 기존 예측 모델을 재학습하는 과정이다.

E. 모델 리엔지니어링

모델 리엔지니어링 태스크는 예측 모델의 품질에 유의미한 변화가 있을 때, 기존 예측 모델을 수정 및 재설계하는 과정이다. 모델 리엔지니어링에는 다음과 같은 기법들이 있다.

- 초매개변수: 최적화 기계학습 모델을 설계할 때 엔지니어에 의해 수동적으로 설정되는 매개변수를 초매개변수 (hyperparameter)라고 하며 대표적으로 학습률(learning rate), 정규화(regularization) 매개변수, 그리고 심층신경망의 은닉계층(hidden layer) 및 노드 수 등이 있다. 이러한 초매개변수는 예측 모델의 성능에 큰 영향을 미치기 때문에, 모델에 유의미한 성능 저하가 발생할 시 초매개변수 최적화를 고려한다.
- 특징 설계: 초매개변수 최적화 여부와 관계없이 모델의 성능

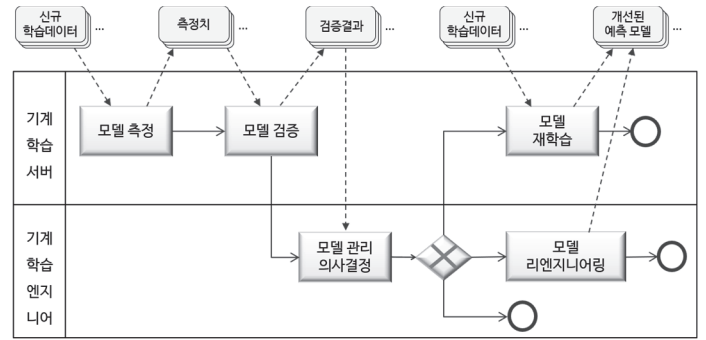


Fig. 4. 상태예측 모델 유지관리 자동화 워크플로우.

이 지속적으로 저하가 될 시 개념변화를 의심해야 하고 이에 대한 조치로서 특징 설계를 고려한다. 새로운 특징 추출, 선택, 그리고 제거를 통해 새로운 특징 집합을 구성하여 예측 모델의 성능을 높인다.

- 모델 선택: 초매개변수 최적화 특징 설계를 수행했음에도 모델의 성능저하가 지속될 시 기존 예측모델의 적용을 중지하고 다른 예측모델로의 교체 또는 새로운 예측모델의 생성을 수행한다.

IV. 전력망 상태추론 모델 성능 유지관리 워크플로우

본 장에서는 전력망 상태예측 모델의 성능을 지속적으로 관리하기 상태예측 모델 성능 유지관리 자동화 시스템의 개발에 대해 논한다. 본 논문에서 제안하는 상태예측 모델 성능 유지관리 자동화 시스템은 상태예측 모델 성능 유지관리에 요구되는 일련의 태스크들을 워크플로우 형태로 정의하고 자동화한다.

Fig. 4는 ‘전력망 상태예측 모델 성능 유지관리’를 위한 워크플로우를 나타내며, 자동화된 태스크를 수행하는 기계학습 서버와 모델 관리 관련 업무 태스크를 수행하는 기계학습 엔지니어 간의 협업을 통해 해당 워크플로우가 수행된다.

워크플로우를 구성하는 필수 태스크들로 모델 측정, 모델 검증, 모델 관리 의사결정, 모델 재학습 그리고 모델 리엔지니어링이 있다. 먼저, 기계학습 서버에서 신규 데이터를 대상으로 모델 검증에 필요한 지표들을 측정한다. 상태예측 모델 검증 태스크에서는 모델 별로 측정된 값들을 이용하여 데이터의 의미 분포 변화 (concept drift) 발생 유무와 일반화 정도를 계산하고 이를 통해 상태예측 모델을 검증한다. 다음으로 모델 관리 의사결정 태스크에서는 기계학습 엔지니어가 상태예측 모델의 측정 및 검증 데이터를 검토하여 모델 관리 방법을 결정한다. 마지막으로, 결정된 모델 관리 유형에 따라 예측 모델은 현상유지, 재학습 또는 리엔지니어링 단계를 거치게 된다.

이 외에도 데이터 전처리, 예측 모델 현황 분석, 모델 교체 및 재배포와 같은 새로운 태스크들이 구현 및 추가되어 전력망 상태예측 모델 관리에 필요한 다양한 워크플로우를 정의 및 운용하는 것이 가능하다. 즉, 본 논문에서는 상태예측 모델 관리 절차를 일련의 태스크(task)과 이들 간의 제어흐름으로 정의되는 워크플로우의 형태로 구현한다. Fig. 5는 상태예측 모델 관리를 위한 워크플로우 관리 시스템의 아키텍처를 나타낸다.

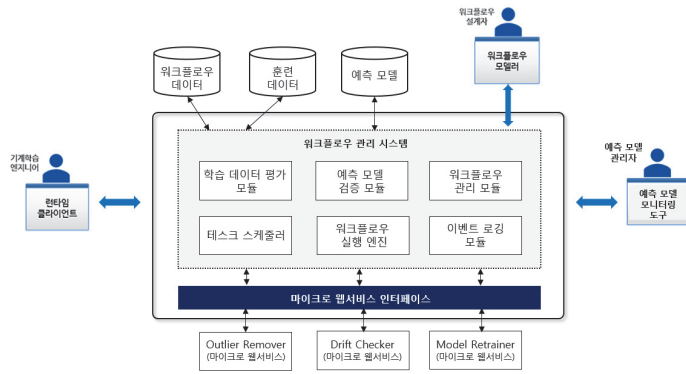


Fig. 5. 예측 모델 관리를 위한 워크플로우 관리 시스템의 아키텍처.

- 예측 모델 관리 워크플로우 정의 및 배치: 워크플로우 설계자는 워크플로우 모델러를 이용하여 예측 모델 관리 워크플로우를 모델링한다.
- 예측 모델 관리 워크플로우 실행: 워크플로우 실행 엔진은 워크플로우 인스턴스의 생성 및 실행을 담당하는 핵심 모듈이다. 자동화된 태스크의 경우 마이크로 웹서비스 인터페이스를 통해 해당 웹서비스를 호출하여 태스크를 수행한다. 업무 태스크의 경우, 마이크로 웹서비스 인터페이스를 통해 웹서비스를 호출하여 태스크를 수행한다. 업무 태스크의 경우, 이를 담당하는 기계학습 엔지니어와 같은 인적자원에 태스크를 할당하여 수행하도록 한다. 또한, 모든 태스크들의 생성, 할당 그리고 실행 스케줄 관리는 태스크 스케줄러에 의해 수행된다. 데이터 측면에서 태스크 수행에 필요한 입출력 데이터와 훈련 데이터에 대한 I/O는 데이터 질의 모듈에서 수행된다. 태스크를 수행하면서 발생하는 이력 데이터는 이벤트 로깅 모듈에 의해 생성된다.
- 예측 모델 모니터링 및 관리: 예측 모델 관리자는 모니터링 도구를 예측 모델의 전체 현황, 모델 별 성능 및 품질에 대한 현황을 분석하며, 예측 모델 관리 모듈을 통해 예측 모델의 생애주기에 해당하는 생성, 배치, 개선 그리고 삭제에 대한 관리 활동을 수행한다.

위에서 정의한 시스템 아키텍처를 기반으로 예측 모델 관리를 위한 워크플로우는 Def. 1과 같이 정형적으로 정의될 수 있다.

Def. 1. A, D, P, S 의 요소 집합이 주어졌을 때, 상태추론 모델 관리 워크플로우 $\Lambda = (\delta, \kappa, \phi, \xi)$ 같이 4-튜플로 구성되며 각각의 요소에 대한 정의는 다음과 같다.

- A 는 예측 모델 관리 워크플로우를 구성하는 단일 태스크들의 집합이다.
- D 는 예측 모델 관리 워크플로우의 입-출력 데이터 객체들의 집합이다.
- P 는 예측 모델 관리 태스크를 수행하는 인적자원 또는 컴퓨팅 자원들의 집합이다.
- S 는 예측 모델 관리 태스크를 구현한 웹서비스들의 집합이다.
- $\delta: A \rightarrow \wp(P)$ 은 예측모델 관리 태스크들과 이를 수행하는 자원과의 연결 관계에 대한 사상 함수이다. 특정 태스크 α_k 를

입력 받아 α_k 를 수행하는 인적자원 또는 컴퓨팅 자원($\phi(\alpha_k)$) 들을 출력한다.

- $\xi: A \rightarrow \wp(S)$ 는 예측모델 관리 태스크들과 이를 구현한 웹서비스와의 연결 관계에 대한 사상 함수이다. 특정 태스크 α_k 를 입력 받아 α_k 를 구현한 웹서비스($\xi(\alpha_k)$)들을 출력한다.

예측모델 검증과 재학습 자동화를 위한 Def. 1의 워크플로우는 모델 측정 태스크, 모델 검증 태스크, 모델 관리 의사결정 태스크, 모델 재학습 태스크, 그리고 모델 리엔지니어링 태스크를 포함하고 있다. 결과적으로 Def. 1을 통해 예측모델 관리 활동을 자동화하고 상태추론 모델 관리를 비즈니스 프로세스 수준에서 운영할 수 있게 되면 상태추론 모델 관리의 유연성을 제공할 수 있게 된다. 결과적으로 본 논문에서 제안한 워크플로우 기반 상태추론 모델유지관리 기법을 통해 얻어지는 효과는 다음과 같다. 첫째, 예측 모델의 성능측정, 검증, 재학습, 그리고 배치와 관련된 관리 활동들을 자동화할 수 있다. 둘째, 표준화된 상태추론 모델 관리 절차를 효율적으로 준수할 수 있도록 상태추론 모델 관리를 비즈니스 프로세스 수준에서 운용 가능하도록 하여 상태추론 모델 관리를 체계적으로 수행할 수 있고 이를 통해 우수한 검증 신뢰도를 확보하게 된다. 예를 들어 BPMN (Business Process Model and Notation) 또는 WS-BPEL (Web Service Business Process Execution Language) 과 같은 워크플로우 모델링 언어를 통해 자동화된 태스크 뿐만 아니라, 모델 관리 관련 의사결정과 같이 사람에 의해 수행되는 업무 태스크를 함께 워크플로우의 형태로 모델링이 가능하다. 표준화된 예측모델 관리 절차 준수를 효율화하여 상태추론 모델 관리를 체계적으로 수행할 수 있다. 워크플로우의 특성상 정의된 절차로 실행을 강제하고, 상황에 따라 실행흐름을 제어한다. 그러므로 워크플로우를 통해 상태추론 모델 관리 절차를 표준화하여 이를 준수하는 형태로 모델 관리를 체계적으로 수행할 수 있게 된다. 셋째, 워크플로우 모델링 기능을 통해 외부 환경의 변화에 대응하기 위한 상태추론 모델 관리의 유연성을 제공한다. 워크플로우 모델링에는 태스크 등록, 제어흐름(control-flow) 정의, 이벤트-트리거링, 외부자원 호출(예: 웹서비스, 레거시 시스템)과 같은 요소들을 포함하므로 다양한 모델 관리 시나리오를 표현할 수 있으며, 이에 대한 워크플로우의 변경 또한 용이하다. 넷째, 워크플로우 관리 시스템 적용을 통해 전력망 예측 분야에서 실시간 모니터링, 자동화, 분석 및 최적화와 같은 기존 워크플로우 관리 기술의 이점을 기대할 수 있다. 다섯째, 모델 유형에 따른 예측 오차 측정 매트릭스를 제안하여 다양한 모델에 범용적인 모델의 예측 성능 측정이 가능하므로 특정 도메인 영역에 최적화된 학습 모델 선정이 가능하다. 마지막으로, 기존 기술에서는 시도되지 않았던 학습 데이터의 통계적 특성 변화 정도와 예측의 일반화 수준을 모두 고려한 상태추론 모델의 성능 평가를 통해 성능 결과의 신뢰성을 확보하고 이를 통해 상태추론 모델의 정확도를 일정 수준으로 유지관리하고 보다 정확도 높은 상태추론 모델의 신규 개발이 가능하다. 결과적으로 본 발명을 통해 지금까지 기계학습 기반의 상태추론 모델이 가지고 있던 현장 적용의 한계점을 해결하고 분산자원 연계 등 외부 환경의 변화에 유연한 상태추론 모델 관리를 통해 정확성과 신뢰성이 보장된 전력망 상태추론 모델의 지속적인 활용을 가능케 할 수 있다.

V. Conclusion

본 논문에서는 다양한 요인에 의해 끊임없이 변화하는 전력

망 상태 데이터의 특성을 고려하여 상태예측 모델의 정확성과 신뢰성을 높이고 현장 적용 가능한 수준의 상태예측 결과의 품질을 유지하기 위한 기계학습 기반 전력망 상태예측 모델의 성능 유지 관리 자동화 기법'을 제안하였다. 제안 기법에서 워크플로우 관리 기술의 적용을 통해 예측 모델 관리를 위한 일련의 태스크들을 워크플로우의 형태로 모델링하고 이를 자동화하였다. 또한, 다양한 상태예측 모델의 검증에 고려하여 예측 모델의 유형을 분류 모델과 회귀 모델로 유형화한 모델의 범용적 예측 오차 측정 매트릭스를 고려하였다. 마지막으로 모델 검증에 있어 모델 측정 단계에서 얻어진 오차 데이터를 기반으로 학습 데이터의 통계적 특성 변화가 예측 모델 정확도에 미치는 영향도를 검증하기 위한 평균오차 변동률 산출식과 모델의 일반화 수준을 검증하기 위한 훈련-테스트 오차비율 변동률 산출식을 제안하여, 모델 성능 저하 정도를 판단하고 저하 요인에 따라 재학습이나 리엔지니어링과 같은 유지관리 방법이 적용될 수 있도록 하였다.

본 논문에서 제안하는 상태예측 모델 유지관리 기법을 통해 얻어지는 효과는 다음과 같다. 첫째, 워크플로우 기술을 통해 예측 모델의 성능측정, 검증, 재학습, 그리고 배치와 관련된 관리 활동들을 자동화할 수 있다. 둘째, 표준화된 예측 모델 관리 절차를 효율적으로 준수할 수 있도록 예측모델 관리를 비즈니스 프로세스 수준에서 운용 가능하도록 하여 예측모델 관리를 체계적으로 수행할 수 있고 이를 통해 우수한 검증 신뢰도를 확보하게 된다. 예를 들어 BPMN (Business Process Model and Notation) 또는 WS-BPEL (Web Service Business Process Execution Language)과 같은 워크플로우 모델링 언어를 통해 자동화된 태스크 뿐만 아니라, 모델 관리 관련 의사결정과 같이 사람에게 의해 수행되는 업무 태스크를 함께 워크플로우의 형태로 모델링이 가능하다. 표준화된 예측모델 관리 절차 준수를 효율화하여 예측모델 관리를 체계적으로 수행할 수 있다. 워크플로우의 특성상, 정의된 절차로 실행을 강제하고 상황에 따라 실행 흐름을 제어한다. 그러므로 워크플로우를 통해 예측 모델 관리 절차를 표준화하여 이를 준수하는 형태로 모델 관리를 체계적으로 수행할 수 있게 된다. 셋째, 워크플로우 모델링 기능을 통해 외부 환경의 변화에 대응하기 위한 예측 모델 관리의 유연성을 제공한다. 워크플로우 모델링에는 태스크 등록, 제어 흐름 (control-flow) 정의, 이벤트-트리거링, 외부자원 호출(예: 웹서비스, 레거시 시스템)과 같은 요소들을 포함하므로 다양한 모델 관리 시나리오를 표현할 수 있으며, 이에 대한 워크플로우의 변경 또한 용이하다. 넷째, 워크플로우 관리 시스템 적용을 통해 전력망 예측 분야에서 실시간 모니터링, 자동화, 분석 및 최적화와 같은 기존 워크플로우 관리 기술의 이점을 기대할 수 있다. 다섯째, 모델 유형에 따른 예측 오차 측정 매트릭스를 제안하여 다양한 모델에 범용적인 모델의 예측 성능 측정이 가능하므로 특정 도메인 영역에

최적화된 학습 모델 선정이 가능하다. 마지막으로 기존 기술에서는 시도되지 않았던 학습 데이터의 통계적 특성 변화 정도와 예측의 일반화 수준을 모두 고려한 예측 모델의 성능 평가를 통해 성능 결과의 신뢰성을 확보하고 이를 통해 예측 모델의 정확도를 일정 수준으로 유지관리하고 보다 정확도 높은 예측모델의 신규 개발이 가능하다. 결과적으로 제안된 기법을 통해 지금까지 기계학습 기반의 예측모델이 가지고 있던 현장 적용의 한계점을 해결하고 분산 자원 연계 등 외부 환경의 변화에 유연한 예측모델 관리를 통해 정확성과 신뢰성이 보장된 예측 모델의 지속적인 활용을 가능케 할 수 있을 것이다.

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] Yeom, Chan-Uk, Keun-Chang Kwak, "Short-term electricity load forecasting using a TSK-based extreme learning machine with knowledge representation," *Energies* 10, p.1613, 2017.
- [2] Yildiz, Baran, Jose I. Bilbao, Allistair B. Sproul. "A review and analysis of regression and machine learning models on commercial building electricity load forecasting", *Renewable and Sustainable Energy Reviews*, 73, 2017.
- [3] Ertugrul, Omer Faruk "Forecasting electricity load by a novel recurrent extreme learning machines approach," *International Journal of Electrical Power & Energy Systems*, 78, 2016.
- [4] Liu, Yang, W. Noradin Ghadimi, "Electricity load forecasting by an improved forecast engine for building level consumers," *Energy*, 139, 2017.
- [5] Yang Seung Kwon, Song Taek Ho, "Stochastic Real-time Demand Prediction for Building and Charging and Discharging Technique of ESS Based on Machine-Learning," *KEPCO Journal on Electric Power and Energy*, vol. 5, no. 3, 2019, DOI
- [6] Dietterich Tom, "Overfitting and undercomputing in machine learning," *ACM computing surveys (CSUR)*, 27(3), 1995.
- [7] Stanley, K. O, "Learning Concept Drift with a Committee of Decision Trees", *Informe tecnico*, "UT-AI-TR-03-302, Department of Computer Sciences, University of Texas at Austing, USA, 2003.