

지능형 Self-Organizing Network를 위한 설명 가능한 기계학습 연구 동향

Trend in eXplainable Machine Learning for Intelligent Self-organizing Networks

권동승 (D.S. Kwon, dskwon@etri.re.kr)

지능형스물셀연구실 책임연구원

나지현 (J.H. Na, jhna@etri.re.kr)

지능형스물셀연구실 책임연구원/실장

ABSTRACT

As artificial intelligence has become commonplace in various fields, the transparency of AI in its development and implementation has become an important issue. In safety-critical areas, the eXplainable and/or understandable of artificial intelligence is being actively studied. On the other hand, machine learning have been applied to the intelligence of self-organizing network (SON), but transparency in this application has been neglected, despite the critical decision-makings in the operation of mobile communication systems. We describes concepts of eXplainable machine learning (ML), along with research trends, major issues, and research directions. After summarizing the ML research on SON, research directions are analyzed for explainable ML required in intelligent SON of beyond 5G and 6G communication.

KEYWORDS 6G, beyond 5G, eXplainable machine learning, self organization network

1. 서론

이동통신은 세대별로 혁신 기술이 도입되어 왔으며, 5G에서는 세 가지 무선액세스 기술(enhanced Mobile Broadband, ultra-Reliable and Low-Latency Communication, massive Machine Type Communication), 다양한 셀룰라 형태(매크로·마이크로·피코·펩토 셀, 클라우딩 기지국, 개방형 무선액세스 네트워크)가 공존하는 이기종 망(HetNets: Heterogeneous Networks), 제어

와 데이터 평면 분리, 그리고 망 가상화·슬라이싱 기술 등으로 세상의 모든 것을 연결시키는 것을 목표로 한다는 점에서 이전 세대와 달리 애플리케이션의 이해 관계자도 많은 관심이 있다. 따라서 이동통신시스템의 핵심성능지표(KPI: Key Performance Indicator)에도 큰 변화가 발생했는데, 운영자 관점에서 KPI인 용량, QoS, CAPEX와 OPEX에 대한 요구가 증가되었고, 사용자 관점에서 추가되는 KPI로는 완벽한 연결성, 시·공간적 서비스 균일성, 거의 무

* DOI: <https://doi.org/10.22648/ETRI.2023.J.380610>

* 이 논문은 2023년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임(No. 2018-0-01659, 5G NR 기반 지능형 오픈 스물셀 기술 개발).



한한 용량에 대한 인식 또는 대기 시간 제로, 서비스 비용 등이 있다. 이런 KPI들을 실현하기 위해 5G의 자가 조직화 네트워크(SON: Self-Organizing Network)는 이동통신망 상태뿐만 아니라 응용 프로그램의 필요에 따라 언제 어디서나 적절한 양의 망 자원을 전환·집중하여 사용자가 원활하고 무제한으로 연결할 수 있도록 이동통신망의 구성 요소들을 지능적으로 자동 구성하는 것이 필요하다[1].

최근 상업, 과학, 산업, 의학 등에서 인공 신경망(ANN: Artificial Neural Network) 기술을 적용한 애플리케이션이 채택되고 있으나, 인공지능(AI: Artificial Intelligence) 기반 시스템 사용 시 주요 장애는 종종 투명성이 부족하다는 것이다. 특히 기계학습(ML: Machine Learning)은 결과와 예측 측면에서 강력하지만, 알고리즘의 내부 작업이 불투명하여 그 결과 또는 예측을 신뢰하기가 어렵다는 이슈가 있다. 따라서 스스로를 설명할 수 없는 시스템에 중요한 결정을 맡기는 것으로 초래되는 위험 문제를 해결하고자 하는 기술이 설명 가능한 ML(XML: eXplainable ML)이다. ML에서 설명 가능성은 ML 기반 결정을 정당화하는 도구이며, 모델 개선과 당면한 문제에 대한 새로운 통찰력을 얻기 위해 예측을 검증하여 보다 신뢰할 수 있는 ML 시스템 개발에 도움을 준다. 따라서 XML은 일반적으로 높은 수준의 예측 정확도를 유지하면서 “인간 사용자가 새로운 세대의 ML 파트너를 이해하고 적절하게 신뢰하고 효과적으로 관리할 수 있도록” 보다 설명 가능한 모델을 생성하는 ML 기술을 만드는 것을 목표로 한다[2].

XML의 핵심은 시스템에서 내린 결정, 권장 사항, 예측 또는 조치와 이 시스템이 만들어지는 프로세스를 설명하는 기능이다. 여기서 설명은 해석 가능성 개념과 밀접하게 관련되는데, 내부 검사를 통해 또는 생성된 설명을 통해 시스템의 작동을 인간이 이해할 수 있는 경우 시스템은 해석 가능하다고

할 수 있다. 그런데 대부분의 ML 모델은 쉽게 해석할 수 없기 때문에 ML 모델의 설명은 종종 어려운 작업이다. 이 작업의 예로는 모델의 정확성 평가에 도움을 주는 예측의 시각화 생성 기술, 해석 불가능한 복잡한 모델의 예측 해석과 그 예측에 대한 정당화 생성, 그리고 본질적으로 해석 가능하고 추론을 통해 설명할 수 있는 모델을 고안하는 해석 가능한 모델 등이 있다[3].

B5G에서 SON은 이동통신시스템에 환경 변화가 발생하더라도, 원하는 목표를 유지할 수 있을 만큼 충분히 확장 가능하고, 안정적이며 민첩하게 동작하는 적응형 자율 네트워크로 정의할 수 있다. 또한, 환경과의 지속적인 상호작용을 통해 특정 작업의 작동 시기와 방법을 독립적으로 결정할 수 있을 뿐만 아니라, 이 시스템에서 수행한 이전 작업을 학습하여 그 성능을 향상할 수 있어야 한다. 또한, 이 SON은 고밀도 네트워크(UDN: Ultra-Dense Network), HetNets, 그리고 다중 RAT·벤더·사업자 환경에서도 네트워크의 단순 최적화 이상으로 자율 동작시킬 수 있어야 한다. 따라서 B5G SON은 기존 세대와 달리 인텔리전스 패러다임으로 진화해야 한다[4].

그동안 인간의 개입을 최소화하면서 이동통신 시스템의 설치·운영·관리를 자동화시키기 위해 다양한 ML을 SON에 적용한 학술적 연구가 있었다. SON에 적용된 ML 기술로는 지도학습, 비지도학습, 컨트롤러, 강화학습(RL: Reinforcement Learning), Markov Models, Heuristics, Dimension Reduction, 전이학습(TL: Transfer Learning) 등이 있다. 그런데 이 ML들에서는 설명 가능성 혹은 해석 가능성을 고려하지 않은 것으로 문헌상에서 조사 분석되었다. 또한, 3G/4G SON 제품 공급업체는 SON 세부 기능을 단일의 SON “블랙박스” 개념으로 구현하였다. 따라서 이동사업자가 신뢰할 수 있는 SON이 5G/B5G

네트워크에 완전히 스며들기 위해서 SON은 자동화 정도를 손상하지 않으면서 어느 정도의 투명성을 제공하도록 설계하는 것이 바람직하다. 즉, SON에 적용되는 ML은 투명성이라는 요구사항을 만족시키는 것이 필요하다.

본고에서는 설명 가능한 ML의 개념, 연구 동향 그리고 이슈와 연구 방향을 기술하였다. 그리고 SON에서 연구된 ML 연구를 간략 정리한 후, B5G/6G의 지능형 SON에 필요한 설명 가능한 ML 연구 방향을 정리하였다.

II. 설명 가능한 기계학습 연구 동향

1. 설명 가능한 기계학습이란

XML이 무엇인가를 정의하기 위해 우선 XML 관련해서 설명 가능성, 해석 가능성, 이해 가능성에 대한 용어 정의는 다음과 같다[5]. 해석 가능성(Interpretability)은 인간이 이해할 수 있는 용어로 의미를 설명하거나 제공하는 능력이다. 설명 가능성(Explainability)은 인간과 의사 결정자 사이의 인터페이스로서 동시에 의사 결정자의 정확한 대리인이자 인간이 이해할 수 있는 설명이라는 개념과 관련된다. 설명 가능성은 내부 기능을 명확히 하거나 자세히 설명하려는 의도로 모델이 취하는 모든 조치는 절차를 의미한다. 이해 가능성(Understandability 혹은 Intelligibility)은 모델이 내부적으로 데이터를 처리하는 알고리즘적 수단이나 내부 구조를 설명할 필요 없이 인간이 모델의 기능(모델 작동 방식)을 이해할 수 있도록 하는 모델의 특성을 나타낸다. 설명 가능성은 해석 불가능한 모델을 설명 가능한 모델로 변환하는 기술을 다루기 때문에 post-hoc 설명 가능성(외부 XML)과 연결된다. 따라서, XML은 청중이 주어졌을 때 기능을 명확하거나 이해하기 쉽게 만드는 세부 사항이나 이유를 생성하는 것으로 정의

할 수 있다.

XML의 한 분류는 설계에 의해 해석할 수 있는 모델(투명한 모델)과 외부 XML 기술을 통해 설명할 수 있는 모델(post-hoc 설명 가능성 모델)로 구분한다. XML의 또 다른 분류는 1) 입력에서 출력으로의 매핑이 사용자에게 보이지 않는 불투명 시스템(불투명 ML); 2) 사용자가 매핑을 수학적으로 분석할 수 있는 해석 가능한 시스템(해석 가능한 ML); 3) 모델이 매핑이 만들어진 배후의 이론적 근거를 이해할 수 있도록 특정 출력과 함께 symbols 또는 규칙을 출력하는 이해 가능한 ML이다.

투명한 ML 모델로는 알고리즘적으로 투명한 모델, 분해 가능한 모델과 시물레이션 가능한 모델이 있다. 알고리즘 투명성은 입력 데이터에서 주어진 출력을 생성하기 위해 모델이 따르는 프로세스를 이해하는 사용자의 능력을 다룬다. 분해 가능성은 모델의 각 부분(입력, 매개변수 및 계산)을 설명하는 능력을 나타낸다. 이 특성은 모델의 동작을 이해, 해석 또는 설명하는 능력을 강화할 수 있지만, 모든 모델이 이 속성을 충족할 수 있는 것은 아니다. 분해 가능성은 모든 입력이 쉽게 해석 가능해야 한다. 시물레이션 가능성은 엄격하게 인간에 의해 시물레이션되거나 생각되는 모델의 능력을 나타내므로 복잡성이 이 클래스에서 중요한 위치를 차지한다. 시물레이션 가능한 모델은 동시에 분해 가능하고 알고리즘적으로 투명한 모델이다.

투명한 ML 모델이란 ML 모델 투명성 수준 중 하나 또는 모든 수준(즉, 시물레이션 가능성, 분해 가능성 및 알고리즘 투명성)에 속할 수 있는 투명한 모델을 말한다. 투명한 모델로는 Linear/logistic regression, Decision trees, K-Nearest Neighbors(K-NN), 규칙 기반 학습, General additive models, Bayesian models 등이 있다.

블랙박스인 ML 모델에 대한 post-hoc 설명 가능

성은 설계상 쉽게 해석할 수 없는 모델이 아니라, 텍스트 설명, 시각적 설명, 로컬 설명, 예에 의한 설명, 단순화에 의한 설명과 기능 관련성 설명과 같은 해석 가능성을 향상시키기 위한 다양한 수단에 의지하는 모델을 대상으로 한다. 텍스트 설명은 모델의 결과 설명에 도움이 되는 텍스트 설명을 생성하는 학습을 통해 모델에 대한 설명 가능성을 제공한다. 시각적 설명은 모델의 동작을 시각화하는 것을 목표로 한다. 로컬 설명은 솔루션 공간을 분할하고 전체 모델과 관련된 덜 복잡한 솔루션 하위 공간에 설명을 제공하여 설명 가능성을 제시한다. 예시 설명은 특정 모델이 생성한 결과와 관련된 데이터 예시를 추출하여 모델 자체를 더 잘 이해할 수 있도록 한다. 단순화에 의한 설명은 설명되어야 할 훈련된 모델을 기반으로 완전히 새로운 시스템을 재구성하는 기술을 집합적으로 나타내는 것이다. 기능 관련성 설명은 관리되는 변수에 대한 관련성 점수를 계산하여 모델의 내부 기능을 명확히 하는 것이다.

ML 모델에 대한 post-hoc 설명 가능성에는 첫 번째, 모델의 내부 처리 또는 내부 표현을 무시하고 모든 ML 모델에 원활하게 적용할 수 있는 post-hoc 설명 가능성을 제공하는 모델 불가지론적(Model-agnostic) 기술, 두 번째, 특정 ML 모델을 설명하도록 맞춤화되거나 특별히 설계된 post-hoc 설명 가능성 기술의 두 가지가 있다. 모델 불가지론적 기술은 예측 절차에서 일부 정보를 추출하려는 의도로 모든 모델에 연결되도록 설계하는 것이다. 단순화 기술을 사용하여 다루기 쉽고 복잡성을 줄이기 위해 선행 작업을 모방하는 프록시 생성과 모델에서 직접 지식을 추출하거나 단순히 시각화하여 동작을 쉽게 해석하는 것에 중점을 두며, 모델 단순화, 특징 관련성 추정과 시각화 기술이 있다. Post-hoc 설명 가능성 기술로는 첫 번째로는 다양한 지도학습 모델

을 다루는 Shallow ML 모델에서 엄격하게 해석 가능한 접근 방식(예: KNN 및 결정 트리)이 있지만, 추가 설명 계층이 필요한 보다 정교한 학습 알고리즘(예: LIME(Local interpretable model-agnostic explanations), SHAP, ICE 등)에 의존한다. 두 번째로는 CNN(Convolutional Neural Network), RNN(Recurrent Neural Network), DNN(Deep Neural Network)과 투명 모델을 포함하는 하이브리드 방식 등에 대한 post-hoc 설명 가능성 기술이 연구되고 있다.

2. 설명 가능한 기계학습 연구 동향

여기서는 DARPA에서 추진 중인 XAI 프로그램, DNN 동작을 이해하기 위한 방법들, 훈련된 ANN에서 규칙 추출 기술, 자가 설명 신경망, 퍼지 모델링과 진화 퍼지 시스템을 설명한다.

2017년 5월 DARPA는 설명 가능하고 매우 정확한 모델 제공을 목표로 하는 프로그램 XAI를 시작했다[6]. 이 XAI 시스템의 목적은 설명을 제공하여 인간이 그 행동을 더 이해하기 쉽게 만드는 것이다. XAI 시스템은 수행한 작업, 현재 수행 중인 작업, 다음에 수행할 작업을 설명해야 하고, 행동하고 있는 중요한 정보도 공개하는 것이다. DARPA가 하려는 것은 높은 수준의 학습 성능을 유지하면서 더 설명 가능한 모델을 생성하는 일련의 ML 기술이다. 이것이 성공할 수 있다고 생각하는 이유는 기존 ML 구조와 달리 XAI가 새로운 학습 프로세스, 설명 가능한 모델, 그리고 설명 인터페이스 구조를 가지기 때문이다. 설명 가능한 모델로는 설명 가능한 특징을 학습하는 수정된 DL 기술인 심화 설명(예: Learning Semantic Association, Learning to Generate Explanations 등), 해석 가능한 모델로 보다 구조화되고, 해석 가능하고, 인과적인 모델을 학습하는 기술(예: Stochastic And-Or-Graph, Bayesian Program Learning 등), 그리고 블

랙박스 같은 모델로부터 설명 가능한 모델을 추론하는 기술인 모델 유도(예: LIME, Bayesian Rule Lists 등)를 고려하고 있다.

DNN 동작을 이해하기 위한 일반적인 세 가지 접근 방식은 다음과 같다[7]. 첫 번째, 네트워크 일부를 투명하게 만드는 것으로, DL 모델의 예측 설명에서 민감도 분석(SA: Sensitivity Analysis)과 계층별 관련성 전파(LRP: Layer-wise Relevance Propagation)를 사용한다. 두 번째, 네트워크 구성 요소의 의미를 학습하는 것으로, 개체의 특정 부분에 의해 자주 활성화되는 경우 그 뉴런은 의미를 가질 수 있다. 세 번째, 현 결정을 뒷받침하는 근본적인 이유를 설명 생성을 통해 알려준다. DNN에서 목표는 쉽게 해석할 수 있는 것(예: 입력 변수) 관점에서 모델이 예측하는 것(예: 범주)을 이해하기 위해 post-hoc 해석 가능성에 중점을 둔다. ML에서 주요 연구는 프로토타입을 구축하여 모델에서 학습한 개념을 해석하는 것과 관련 입력 변수를 식별하여 모델의 결정을 설명하는 것이다.

훈련된 ANN에서 규칙 추출 기술[8]은 훈련된 ANN 내에는 훈련 단계에서 획득한 지식이 다음과 같이 인코딩된다는 기본 인식에서 출발한다: (a) 네트워크 아키텍처 자체(예: hidden 유닛의 수); (b) ANN의 각 hidden과 출력 단위와 관련된 활성화 함수; (c) 일련의 실수 값의 매개변수(가중치). 훈련된 ANN에서 설명(또는 규칙)을 추출하는 작업은 (a), (b) 및 (c)의 집합적 효과를 이해할 수 있는 형태로 해석하는 것이다. 규칙 추출 기술로는 분해적 접근을 사용한 Boolean 규칙 추출, 교육적 접근을 사용한 Boolean 규칙 추출, 그리고 퍼지 룰 추출 기술이 있다. 분해적 접근을 사용한 Boolean 규칙 추출은 훈련된 ANN 내의 개별(hidden 및 출력) 유닛 수준에서 기존의 Boolean 규칙을 검색하고 추출하는 것이다. 교육적 접근을 사용한 Boolean 규칙 추출 기술에는 다양

한 연구가 있지만, 기본적으로는 학습을 통한 규칙을 추출하는 것이다. 퍼지 룰 추출 시스템은 다음 세 가지로 구성된다: 첫 번째, 퍼지 규칙의 형태로 기존 전문 지식을 ANN 구조에 삽입하는 일련의 메커니즘/절차이다(지식 초기화 단계); 두 번째, ANN 훈련 과정에서 훈련 데이터 패턴에 따라 구성된 함수를 조정하는 데 중점을 둔다; 세 번째, 일련의 수정된 구성된 함수의 형태로 포함된 개선된 지식의 분석과 추출을 하는 것이다.

해석 가능성이 학습 중에 이미 구조적으로 내장되고 정규화를 통해 실현되는 복잡한 모델인 자가 설명 신경망(SENN: Self-Explaining Neural Networks)이 있다[9]. 이 모델은 해석 가능성에 대한 세 가지 요구사항인 명시성, 충실성과 안정성을 충족시킨다. 충실성과 안정성은 모델에 특별히 맞춰진 정규화를 통해 실현되고, 안정성은 유사한 입력이 유사한 설명이 생성되도록 보장하는 것이다. SENN은 입력을 해석 가능한 basis 특징의 작은 집합으로 변환하는 개념 인코더, 관련성 점수를 생성하는 입력에 종속되는 parametrizer, 예측을 생성하기 위해 결합하는 aggregation 함수 세 가지 요소로 구성된다.

퍼지 집합 이론(FST: Fuzzy Set Theory)을 사용하여 설명과 예측 모델을 설계하는 퍼지 모델링(FM: Fuzzy Modeling)은 원래 입력과 출력 간 고도로 비선형적인 관계를 나타내는 동시에 단순화된 자연어를 사용하여 그 관계에 대한 지능적 관점을 제공할 수 있는 지식 기반 모델을 개발하는 것이다[10]. 이것은 퍼지 규칙에 의해 달성되었고, 현 퍼지 규칙 기반 모델은 FM에서 일반적인 관행이다. 데이터로부터 만들어진 퍼지 규칙 기반 모델은 주로 정확성을 위해 설계되었지만, 해석할 수 없는 퍼지 규칙 기반 모델은 블랙박스 모델과 유사하다. 해석 가능성은 대략(퍼지) 모델의 지식 기반을 읽고 이해하는 능력으로 정의될 수 있는 퍼지 규칙 기반 모델의 속성이지만

만, FST의 단순한 사용으로 해석 가능성이 부여되는 것이 아니라 추가 방법론이 필요하다. 해석 가능한 퍼지 모델을 설계하려면 모델링 단계에서 다음 세 가지 방법이 있다: 첫 번째, 가용 데이터와 전문가의 지식 모두를 고려하는 것인 효과적인 모델에 결정적이므로, 예측과 설명 능력면에서 균형 잡힌 모델 설계를 위해 유도된 지식을 전문가 규칙과 통합하는 반복적 접근 방식을 사용하는 것이다; 두 번째, 해석 가능성과 정확성이 상충되는 목표임을 인식하고 솔루션의 파레토 프런트 달성을 위해 다중 목표 기술을 채택하는 것이다; 세 번째, Ad-hoc 알고리즘으로 데이터로부터 퍼지 규칙을 유도하는 알고리즘 내에 해석 가능성 제약을 통합하는 것이다.

퍼지 규칙 기반 시스템에 파라미터 최적화를 위해 진화 알고리즘(EA: Evolution Algorithm)을 적용한 진화 퍼지 시스템(EFS: Evolutionary Fuzzy Systems)에 대한 연구가 있다[11]. 사용 가능한 모든 정보를 이용하여 가장 깊은 수준에 도달하려면 지식 영역과 데이터 분석 간의 강력한 시너지가 필요하다. 따라서 데이터 관점에서 데이터와 관련된 품질을 인식해야 하고, 적용 대상인 애플리케이션 영역에서 사용할 수 있는 전문 지식의 표현도 고려해야 한다. 이 시나리오에서 FST는 유용한 도구가 될 수 있다. 학습 관점에서는 입력 특징을 퍼지 회원 함수가 있는 퍼지 변수로 변환하면 어느 정도의 불확실성이 있는 데이터에 잘 적응하는 설명 모델을 얻을 수 있다. 일반적으로 퍼지 집합을 포함하는 퍼지 IF-THEN 규칙으로 구현되는 퍼지 규칙 기반 시스템(FRBS: Fuzzy Rule-Based Systems)은 지식 기반(KB: Knowledge Base)과 추론 엔진 모듈로 구성된다. 지식 기반은 규칙 기반(RB: Rule Base)의 퍼지 규칙과 데이터베이스(DB: Data Base)의 퍼지 집합으로 구성되고, 추론 엔진은 퍼지화 인터페이스, 추론 시스템 및 역퍼지화 인터페이스를 포함한다. FRBS에서 학습 절차는 관

찰을 기반으로 주어진 성능 메트릭에 가장 근접한 모델 검색을 포함한다. 또 FRBS에서 명료도는 주로 추론과 관련되고, 이해도는 인간이 검사하고 이해할 수 있는 방식으로 모델을 인코딩하는 학습 알고리즘 능력으로 정의되며, 퍼지 규칙과 KB와 관련된다. 솔루션 모집단을 기반으로 하는 일종의 메타 휴리스틱 기반 최적화 기술인 EA는 ML 기법으로 특별히 설계된 것은 아니지만 학습 과제를 최적화 문제로 모델링하여 진화를 통해 효과적으로 해결할 수 있다는 장점이 있다. 즉, 복잡하고 잘 정의되지 않은 문제 공간에서 EA의 강력한 검색으로 다양한 ML 및 지식 검색 작업을 성공적으로 실현할 수 있다. 따라서 FRBS의 본질적인 이해 가능성과 설명 가능성에 FRBS를 개선하기 위한 최적화 기술로서 EA를 적용한 것이 EFS이다. EFS는 정확성, 해석 가능성 또는 이 둘을 조합하여 원하는 목적 함수 측면에서 보다 신뢰할 수 있는 솔루션을 얻기 위해 FRBS 구성 요소의 미세 설정에 EA를 활용하는 것이다. EA는 학습 단계로 이어지는 FRBS의 구축 단계에서 사전적으로 적용되거나 혹은 튜닝 단계인 FRBS의 미세 조정을 위해 사후적으로 고려될 수 있다. 또한, 확장성을 위해 여러 설계 기준을 동시에 최적화할 수 있는 다목적 진화 알고리즘(MOEA: Multi-Objective Evolutionary Algorithms)을 사용한 소위 다목적 진화 퍼지 시스템(MOEF: Multi-Objective Evolutionary Fuzzy Systems)을 사용하는 것이 좋다.

3. 설명 가능한 기계학습 연구 방향

여기서는 XAI에서의 이슈와 도전 그리고 연구 방향을 서술하였다. 우선 DARPA에서 제기한 XAI에서의 이슈와 도전은 다음과 같다[6].

첫 번째, 정확성과 해석 가능성의 트레이드오프다. XAI 연구의 핵심은 해석 가능성의 기술과 한계

를 탐구하는 것이다. 해석 가능성은 정확성과 충실도를 포함하는 균형을 고려하고, 정확성, 해석 가능성과 취급 용이성 간의 균형을 유지해야 한다. 특징 회귀, 분류 또는 의사 결정 등의 문제 해결을 위해 어떤 종류의 ML 기술을 사용해야 하는지 결정할 때 고려해야 할 기준은 정확성과 해석 가능성/설명 가능성이다. 정확성에 대한 필요성으로 신경망과 DL 기반의 블랙박스 솔루션을 수용하는 추세이고, 이 시스템들은 복수의 뉴런 계층을 기반으로 하는데 첫 번째 계층은 데이터에서 더 간단한 속성을 추출하고, 이후 계층에서 결합되어 더 복잡하지만 더 대표적인 속성을 형성한다. 정확한 모델을 얻을 수 있는 블랙박스 솔루션의 놀라운 능력에도 불구하고, 종종 더 높은 시스템 복잡성이 따른다. 따라서 설계자는 종종 시스템이 어떻게 작동하는지 이해할 수 없으며 시스템이 특정 출력을 생성하는 이유를 해독할 수도 없다. 해석 기준 ML은 일부 응용 분야에서 인간 사용자에게 자신의 생각과 행동을 설명할 수 없는 기계의 무능력으로 인해 제한될 수 있다. 현재 심층망을 해석하고 설명하는 것이 떠오르는 연구 분야이지만, 대부분의 제안은 post-hoc 해석 가능성에 중점을 두고 있으며, ML 모델을 더 잘 이해하려면 더 많은 노력이 필요하다.

두 번째, 설명을 단순화하기 위해 추상화를 사용하는 것이다. 높은 수준의 패턴은 큰 단계에서 큰 계획을 설명하기 위한 기초이므로, 추상화 발견을 자동화하는 것은 오랫동안 도전 과제였으며, 학습과 설명에서 추상화를 발견하고 공유하는 것은 오늘날 XAI 연구에서 큰 이슈이다.

세 번째, 설명의 효과성을 평가하고 측정하는 것이다. XAI 시스템이 비 XAI 시스템보다 사용자에게 더 이해하기 쉬운지 여부를 측정하는 일반적인 방법은 현재 없다. 이 측정 중 일부는 설명의 명확성과 유용성에 대한 주관적인 평가를 통해 측정할 수 있

는 사용자 만족도와 같이 사용자 관점에서 주관적인 측정이다. 설명의 효율성에 대한 보다 객관적인 척도는 작업의 효과성일 수 있다; 즉, 설명이 사용자의 의사 결정을 향상시키는가? 설명의 효과에 대한 신뢰할 수 있고 일관된 측정은 여전히 오픈 연구 문제다. 물론 XAI 시스템에 대한 평가와 측정에서 평가 프레임워크, 다른 생각과 상호 이해, 상식과 논증도 이슈이다.

규칙 추출 기술에서 제기된 이슈는 다음과 같다 [8].

첫 번째, 주어진 문제 영역의 요구사항에 기술 집합을 일치시키기 위한 일련의 기준을 공식화하는 것이다. 예를 들어, 실용적인 수준에서 아직 등장하지 않은 것은 이산 데이터와 구별되는 실제 값 데이터를 포함하는 응용 문제 영역에 대해 어떤 규칙 추출 기술이 최적인지 결정하는 것이다. 또한, 데이터에서 규칙 추출을 위한 다른 유도 기술에 비해 ANN/규칙 추출 기술의 성능에서 보고된 개선이 모든 문제 영역에 적용되는지 여부도 불확실하다.

두 번째, 애플리케이션 프로그램에서 추출된 규칙 집합이 규칙 집합이 제거되고 훈련된 ANN보다 우수한 성능을 나타내는 것이 ANN 훈련이 완료된 후 출력 장치에 남아 있는 오류를 어느 정도까지 제거해야 하는지에 대한 것이 명확하지 않다는 것이다. 따라서 추출된 규칙 집합이 원래 네트워크보다 더 나은 일반화를 나타내는 조건 집합을 식별하는 것이 이슈이다.

세 번째, 조사된 규칙 추출 기술의 특징이 ANN 훈련 프로세스 완료 후 규칙이 추출된다는 점에서 발생한다. 즉, 최종 규칙 집합의 하위 집합 추출이 훈련 프로세스의 어느 지점인지, 규칙 삽입과 규칙 추출 작업이 ANN 훈련 중에 발생할 수 있는지에 대한 질문이다.

네 번째, 규칙 추출 알고리즘의 복잡성에 대한 추

가 사항으로 충실도가 높은 네트워크를 모방하는 최소 규칙 집합을 찾는 문제가 어려운(아마도 NP-hard) 문제인지 여부를 결정하는 것이다. RULEX 기술을 제외하고, 조사된 규칙 추출 기술은 검색할 샘플 공간의 크기를 제한하기 위해 어떤 형태로든 휴리스틱이 필요하다. 예를 들어, 분해적 접근 방식에서 임계값은 ANN이 생성한 최종 결정에 큰 영향을 미치지 않는 입력을 필터링하는 데 사용된다. 따라서 잠재적인 연구 이슈는 생성된 규칙의 품질과 효율성에 대한 이 휴리스틱의 영향을 평가하는 것이다.

XML의 연구 방향은 다음과 같다[6].

첫 번째, 더 강력한 모델별 접근 방식을 개발하는 것이다. 이 방향의 이점은 결과 접근 방식이 모델의 고유한 특징을 활용하여 설명을 생성할 수 있고 아마도 충실도를 향상시킬 뿐만 아니라 모델의 결과를 설명하는 대신 모델의 내부 작동을 더 잘 분석할 수 있다는 것이다.

두 번째, 견고성 문제를 해결하는 유망한 방법으로 XAI와 통계 간의 연결을 설정하는 추가 방법을 탐색하여 잘 연구된 다양한 도구를 활용하는 것이다.

세 번째, 선형 회귀와 신경망의 결합 방법처럼 불투명 모델의 표현력과 투명 모델의 명확한 의미를 결합한 하이브리드 모델을 설계하는 것이다. 이 방향은 불투명 모델과 투명 모델 사이의 격차를 해소하는 데 도움이 될 뿐만 아니라 설명 가능한 모델을 수행하는 최신 기술 개발에도 도움이 될 수 있다.

네 번째, 설명 가능성 방법들을 결합하는 것이다. 상대적으로 더 강력한 설명을 달성하기 위해 서로 다른 해석 가능성 방법을 결합할 가능성에 대한 접근에는 거의 관심이 없었다. 문헌에는 일부 기술이 다른 기술을 보완하는 데 어떻게 사용될 수 있는지(예: 민감한 분석과 시각화)가 있지만, 서로 다른 해석

가능성 방법들을 새로운 부가가치 기술을 상승적으로 생성할 수 있는 기본과 구성 가능한 구성 요소로 취급하는 방법은 없었다. 따라서 XAI에서 구성 가능성을 활성화하면 이 분야의 최적화 문제를 효과적으로 해결하고 설명 가능성과 정확도를 같은 방향으로 이동시키는 데 잠재적으로 기여할 수 있다. 또한, 궁극적인 조합은 ML 모델에 초점을 맞춘 실제 해석 가능성 방법과 전문가 시스템과 관련된 설명 가능성의 고전적인 솔루션과 관련될 수 있다. 기존의 설명 가능한 전문가 시스템과 현재 해석 가능한 ML 방법의 요소를 정확히 결합하는 방법은 논쟁 주제이다.

III. SON에서의 기계학습 연구 동향

SON에 적용된 ML 연구 동향으로 문헌과 특허에 대한 분석 결과를 제시하였다.

SON에 적용할 ML 알고리즘은 일반적으로 이동 통신환경에 변화가 발생하더라도 원하는 목표를 유지할 수 있을 만큼 충분히 확장 가능하고 안정적이며 민첩하게 동작해야 한다. 그동안 SON에 적용된 ML은 특정 유즈 케이스에 적합한 솔루션 연구가 주로 진행되었다. 즉 SON 세부 기능별로 고유 요구사항이 있고, 다음과 같이 특정 기능에 대해 특정 ML 알고리즘이 우수하게 작동한다는 연구가 있었다 [12,13].

자가 구성에서 동작 파라미터, NCL(Neighbor Cell List)와 무선 파라미터의 설정에는 Q-Learning, Fuzzy Q-Learning 등이 적용되었다. 동작 파라미터, NCL과 무선 파라미터별로 주어진 데이터 형태, 요구 처리 시간, 허용 가능한 복잡도를 고려한 최적 ML 선택이 필요하다.

자가 최적화의 세부 기술인 커버리지와 용량 최적화, 이동성 관리, Mobility Robust Optimization,

로드 밸런싱, 자원 최적화, 그리고 SON 기능 조정에서 연구된 알고리즘을 순서대로 나열하면 Markov Chain/Hidden Markov Chain, Collaboration Filtering, Self-Organizing Maps, Feedback/Fuzzy Logic Controller, 그리고 Q-Learning/Fuzzy Q-Learning이다. 지속적 최적화라는 고유의 특성상 무선 환경과 지속적으로 상호작용하면서 동작되는 알고리즘이 선호된다. 자가 최적화에는 많은 최적화 목표가 있고, 지속적인 최적화를 위해 동적 조정이 필요한 파라미터가 매우 많기 때문에 많은 연구가 필요하다. 특히 제어 대상 파라미터들의 공통성으로 인해 하나의 목표를 다른 목표에 결합하는 것은 여전히 해결해야 할 문제이다. 따라서 동일 파라미터가 자가 최적화의 여러 세부 기능에 공동 최적화 대상인 점을 감안하여 세부 기능별 최적 ML 알고리즘 선택보다는 우선 자가 최적화의 세부 기술별 충돌을 해소하거나 조정하는 전체 Framework 설계 후, 각 세부 기능별 최적 ML 선택이 적절하다.

자가 치유에서 연구된 알고리즘을 순서대로 나열하면 Markov Chain/Hidden Markov Chain, K-NN, Support Vector Machine, Collaboration Filtering, 그리고 Feedback/Fuzzy Logic Controller이다. 자가 치유에서 정보 수집→결함 검출→결함 진단→결함 보상/복구→보고 과정의 각 단계에 ML을 적용한 연구가 있었다. 각 단계별로 최적 알고리즘 선택이 적절하다. 결함을 신속하고 정확하게 자동 검출하는, 즉 False 알람 가능성을 최소화하는 검출 정확도와 빠른 검출 시간 간 최상의 균형이 필요하다. 진단을 위해서 연속 KPI 모델링 방법, 과거 데이터에서 모델 학습 방법, 오류 보상을 위해 인접 셀 파라미터 자동 조정 방법 등의 기술적 이슈를 해결해야 한다. 자가 치유의 주요 연구 문제는 대규모 time scale에 걸쳐 환경과 outage·anomaly를 모델링하고, 시뮬레이션하는 어려움에서 비롯된다. 분석적 접근 방식

을 사용하는 경우 Markov chain 분석 또는 Bayesian 분석을 사용하여 복구 확률을 예측할 수 있지만, 시스템 레벨 시뮬레이터를 통해 확률을 검증하는 것은 연구할 가치가 있는 주제이다.

그동안 SON에 ML을 적용하는 연구가 진행되었지만, 아직도 특정 SON 기능에 적용되지 않은 ML 기술들은 많다. 모든 SON 기능에 모든 알고리즘을 적용하는 연구가 필요한 것은 아니지만, ML의 실제 동작 여부 판단을 위해서 다양한 ML에 대한 추가 연구는 여전히 필요하다. 특히, 5G에서 물리계층 관리, 자동 RAT 선택, 종단 간 연결성, 하이브리드 망 구조, 연결된 머신으로부터 학습, 제어와 데이터 평면 분리, 네트워크 기능 가상화 등에 적용되는 ML에 대한 연구가 필요하다. 그리고 5G와 B5G에서 SON은 인간의 개입 없이 자율적이면서 신뢰할 수 있는 동작을 보장하기 위해 ML 알고리즘의 투명성이 요구된다. 그런데, 현재 검토한 문헌에서는 이 투명성을 수용하거나 또는 지향하는 ML을 SON에 적용한 연구는 파악되지 않고 있다.

ML이 적용된 SON의 특허정보에 대한 거시적 분석과 심층 분석 결과는 다음과 같다[1]. SON 특허정보의 정량적 분석 결과, ML이 적용된 기술은 전체 특허의 27%(46건)로서, 아직은 초기 단계로 파악되었다. 그리고 ML 적용 SON 특허 중 자가 최적화 기술이 24%(41건), 자가 치유 기술이 3%(5건)로 분석되었다. 자가 최적화·치유 기술에 적용된 ML 기술은 강화학습 계열이 15건, ML이 9건, Fuzzy Logic이 7건, Markov 모델 계열이 6건으로 분석되었다. ML 적용 자가 최적화 특허는 HetNets 환경에서 자원할당, 로드 밸런싱과 핸드오버 최적화 세부 기술에서 강화학습, Fuzzy Logic과 Markov 모델 순으로 사용되었다. 그리고 ML이 적용된 자가 치유 특허는 결함 검출과 진단에서 생성적 대립 망과 CNN을 사용했다. SON에 적용된 ML 기술관련 서베이 참고문

현 [2]를 보면 SON의 거의 모든 분야에 ML이 적용되었음을 알 수 있다. 그런데, SON 특허정보의 정량분석 결과, 현재 ML 적용 SON 특허는 SON 유효특허의 27%이고, 그 특허 중 70% 정도가 자원할당과 로드 밸런싱에 ML을 적용했다. 세부적으로는 ML이 SON 요소기술에 국한된 문제 해결을 위해 적용된 특허들이었고, SON 특허의 특허경쟁력지수 최고점을 비교해 보아도 ML 적용 특허가 ML 미적용 특허의 절반 정도로 작았다. 따라서 ML 적용 SON 특허는 아직 초기 연구개발 단계라 할 수 있으므로, SON 프레임워크와 자가 구성·최적화·치유 기술의 세부 분야에 Big Data를 활용한 ML을 적용하는 연구개발이 필요하다. 물론 설명 가능한 또는 해석 가능한 ML이 적용된 SON 관련 특허는 없었다.

IV. 지능형 SON에서 XML 연구 방향

투명하고 지능적인 SON 실현을 위해 해석 가능한 및/또는 설명 가능한 ML 연구개발을 위한 고려 사항은 다음과 같다.

첫 번째, 문헌상에서 SON을 대상으로 한 XML 연구는 없었으므로, 타 응용 분야에서 연구된 XML 기술들을 출발점으로 하는 것이 타당하다. 타 응용 분야에서 연구된 XML 기술들은 다음과 같다. 투명한 ML로는 Linear/logistic regression, Decision trees, K-NN, 퍼지 시스템 기반의 규칙 기반 학습, General additive models, Bayesian models이 있다. 해석 가능성이 학습 중에 이미 구조적으로 내장되고 정규화를 통해 실현되는 자가 설명 모델과 FRBSs의 본질적인 이해 가능성과 설명 가능성에 FRBSs 개선을 위한 최적화 기술로 EA를 적용한 EFS가 있다. DL 등 불투명한 ML 모델에 대한 post-hoc 설명 기술로 단순화에 의한 설명(예: LIME), 특징 관련성 설명(예: SHAP, SA 등), 시각화 설명, 트리 앙상블과 Support Vector

Machine 등이 있다.

두 번째, SON의 세부 기술별 연구된 ML 기술은 다음과 같다: 자가 구성에서는 Q-Learning과 Fuzzy Q-Learning을 적용한 연구가, 자가 최적화에서는 Markov Chain/Hidden Markov Chain, Collaboration Filtering, Self-Organizing Maps, Feedback/Fuzzy Logic Controller, Q-Learning/Fuzzy Q-Learning 등을 적용한 연구가 있다. 자가 치유에서는 Markov Chain/Hidden Markov Chain, K-NN, Support Vector Machine, Collaboration Filtering, Feedback/Fuzzy Logic Controller 등이다. 타 응용 분야에서의 XML 기술과 SON의 세부 기술에서 적용된 ML 기술에 공통적으로 사용된 기술은 Fuzzy 기반 기술, K-NN, DL 등임을 알 수 있다.

세 번째, 이동통신시스템의 실시간 또는 비실시간 제어와 B5G의 분산 구조에 대처하는 글로벌 및/또는 국부 제어를 할 수 있도록 SON 세부 기능(자가 구성, 자가 최적화, 자가 치유)별로 XML 연구를 위한 고려 사항은 다음과 같다. 자가 구성은 고유 특성상 대부분 offline으로 동작하는 것으로 한다. 자가 최적화는 지속적 최적화라는 고유의 특성상 무선 환경과 지속적으로 상호작용하면서 동작하는 알고리즘이 선호되며, 최적화 주기도 자가 최적화의 세부 기능별 요구사항에 따라 비실시간과 준실시간 루프를 동시에 사용하거나 독립적으로 사용할 수 있어야 한다. 자가 치유에서 결함 검출, 결함 진단과 결함 보상/복구는 준실시간 루프를 적용하는 것으로 한다.

네 번째, 이동통신시스템에서 수집되는 불확실성이 내재된 빅데이터와 전문가 지식의 효과적 활용이 필요하다.

상기 사항을 고려하여 제안하는 연구 방향은 다음과 같다.

첫 번째, 이동통신시스템에서 수집되는 빅데이

터와 전문가 지식을 활용할 수 있는 FRBSs와 FRBSs 내의 여러 파라미터를 최적화시키는 ML을 결합한 XML을 연구하는 것이다. 균형 있는 예측과 설명 능력을 갖는 퍼지 모델 설계를 위해서 유도된 지식을 전문가 규칙과 통합하는 반복적 접근 방식인 FRBSs는 이해 가능성과 설명 가능성이 내재되어 있다고 할 수 있지만, SON의 요구사항(확장성, 안정성, 민첩성)을 충족하기에는 부족하다. 한편, ML이 적용된 SON 연구에서도 퍼지 로직과 DL을 결합한 연구가 많았다. 따라서 이동통신시스템에서 수집되는 불확실성이 높은 빅데이터에서 규칙을 추출할 수 있는 FRBSs와 FRBSs의 여러 파라미터 최적화를 위한 진화 알고리즘 혹은 Q-Learning 등 DL을 결합한 연구를 제안한다.

두 번째, 정확성과 적응성은 우수하지만 블랙박스인 DNN 모델을 post-hoc으로 해석하는 것이다. 이는 DNN 모델의 내부 작동을 설명하는 대신에 블랙박스 모델의 동작을 특징짓는 것, 즉 쉽게 해석할 수 있는 것 관점에서 모델의 예측 결과를 이해할 수 있는 post-hoc 해석 가능성 기술을 연구하는 것이다. 이 접근은 타 응용 분야에서는 여러 연구가 있었지만, SON에 적용된 연구는 없었다. 구체적으로는 프로토타입을 구축하여 모델에서 학습한 개념을 해석하는 것과 관련 입력 변수를 식별하여 모델의 결정을 설명하는 것이 있다. DNN 결정을 설명하는 세부 기술은 가장 중요한 입력 특징을 파악하는 민감도 분석, 관련성 점수의 합으로 분해하여 모델의 결정을 설명하는 간단한 테일러 분해, 관련성 전파 기법이 있다.

세 번째, 설명 가능한 모델로서 설명 가능한 특징을 학습하기 위해 수정된 DL 기술인 심화 설명을 연구하는 것이다. 즉, 설명을 생성하기 위해서 학습을 하는 DL을 연구하는 것이다. 예로 DNN을 대상으로는 자가 설명 모델로, 해석 가능성이 학습 중에 이

미 구조적으로 내장되고 정규화를 통해 실현되는 복잡한 모델에 대한 연구가 있다.

네 번째, 장기적으로는 SON에 적용을 목표로 해석 가능성 및/또는 설명 가능성이 내재된 투명한 DNN 모델을 연구하는 것이다. 즉 모델 생성 단계부터 더욱 구조화되고, 해석 가능하면서 인과적인 모델을 학습하는 해석 가능한 모델이면서 SON의 요구사항(확장성, 안정성, 민첩성)을 충족시키는 DNN 기반 XML을 연구하는 것이다. 타 응용 분야에서 연구된 해석 가능한 모델(Linear Models, Decision Trees, Rule based models, Naive Bayes, K-NN, Interactive models, Bayesian Networks 등)과 설계에 의해 해석 가능한 모델(Decision Trees, Decision Rules, Decision Sets, Linear Models, Surrogate Fitting 등)을 참고하여 연구하는 것이다. 물론 선형 회귀와 신경망의 결합 방법과 같이 불투명 모델의 표현력과 투명 모델의 명확한 의미를 결합한 하이브리드 모델도 포함한다.

V. 결론

최근 다양한 분야에서 AI 기술이 보편화 되면서 AI의 투명성이 이슈가 되고 있는데, 특히 안전이 중요한 분야에서는 설명 가능성 및/또는 이해 가능성이 내재된 AI, 즉 X-AI/ML의 필요성이 강조되어 여러 연구가 진행 중이다. 한편 SON의 지능화를 위해 ML을 적용한 연구가 있었지만, 이동통신시스템 동작에 중요한 결정을 하는 SON에 적용하는 ML에서 투명성을 고려한 연구는 없었다. 따라서 본고에서는 설명 가능한 ML의 개념, 연구 동향 그리고 주요 이슈와 연구 방향을 기술하였다. 그리고 SON에서 연구된 ML 연구를 간략 정리한 후, 이를 바탕으로 5G/6G의 지능형 SON에 필요한 설명 가능한 ML 연구 방향을 정리하였다.

약어 정리

AI	Artificial Intelligence
ANN	Artificial Neural Network
CNN	Convolutional Neural Network
DNN	Deep Neural Network
EA	Evolution Algorithm
EFS	Evolutionary Fuzzy Systems
FM	Fuzzy Modeling
FRBSs	Fuzzy Rule-Based Systems
FST	Fuzzy Set Theory
KNN	K-Nearest Neighbors
KPI	Key Performance Indicator
LIME	Local Interpretable Model-agnostic Explanations
LRP	Layer-wise Relevance Propagation
ML	Machine Learning
RAN	Radio Access Network
RIC	RAN Intelligent Controller
RNN	Recurrent Neural Network
SA	Sensitivity Analysis
SENN	Self-Explaining Neural Networks
SON	Self-Organizing Network
XAI	eXplainable AI
XML	eXplainable ML

참고문헌

- [1] 권동승 외, "기계학습에 의한 Self-Organizing Network 특허동향," 한국통신학회논문지, 16권 12호, 2021, pp. 2372-2382.
- [2] A. Adadi et al., "Peeking inside the black-box: A survey on explainable artificial intelligence(XAI)," IEEE Access, vol. 6, 2018, pp. 52138-52160.
- [3] O. Biran and C. Cotton, "Explanation and justification in Machine Learning: A survey," in Proc. IJCAI 2017, (Melbourne, Australia), Aug. 2017.
- [4] Small Cell Forum, Small Cell SON and Orchestration from 4G to 5G, Document 233.10.01, Feb. 2020.
- [5] A.B. Arrieta et al., "Explainable artificial intelligence(XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI," Inf. Fusion, vol. 58, 2020, pp. 82-115.
- [6] D. Gunning et al., "XAI-Explainable artificial intelligence," Sci. Robot., vol. 4, no. 37, 2019.
- [7] F. Xu et al., "Explainable AI: A brief survey on history, research areas, approaches, and challenges," in NLPCC 2019, vol. 11839, Springer, Cham, Switzerland, 2019, pp. 563-574.
- [8] R. Andrews et al., "Survey and critique of techniques for extracting rule from trained artificial neural networks," Nerocomputing research centre of Queensland University of technology in Australia, Jan. 1995.
- [9] D. Alvarez-Melis et al., "Towards robust interpretability with self-explaining neural networks," in Proc. NeurIPS 2018, (Montreal, Canada), Dec. 2018.
- [10] C. Mencar et al., "Paving the way to explainable artificial intelligence with fuzzy modeling," in Fuzzy Logic and Applications, Springer, Cham, Switzerland, 2019.
- [11] A. Fernandez et al., "Evolutionary fuzzy systems for explainable artificial intelligence: Why, when, what for, and where to," IEEE Comput. Intell. Mag., 2019, pp. 69-81.
- [12] 권동승, 나지현, "Self-Organizing Network에서 기계학습 연구동향-I," 전자통신동향분석, 제35권 제4호, 2020, pp. 103-114.
- [13] 권동승, 나지현, "Self-Organizing Network에서 기계학습 연구동향-II," 전자통신동향분석, 제35권 제4호, 2020, pp. 115-134.