

7QCI 정보 제공 환경상의 허위정보 오인 개선을 위한 진실 확률 추정 기반 인공지능 의사결정 최적화 알고리즘에 관한 연구

서 성 현*

호서대학교 빅데이터시학부 초빙교수

Truth Probability Estimation-Based AI Decision Optimization for Reducing Misjudgment Caused by False Information in 7QCI Environments

Sung-Hyun Seo*

Adjunct Professor, Department of Big Data and Artificial Intelligence, Hoseo University, Chungcheongnam-do 31499, Korea

[요 약]

본 연구논문은 7QCI 정보 제공 환경상에서 허위정보 유입으로 인해 발생할 수 있는 인공지능 의사결정 오인 문제를 개선하기 위하여, 진실 확률 추정 기반의 인공지능 의사결정 최적화 알고리즘을 제안한다. 본 연구논문에서 제안된 알고리즘은 데이터 수집, 데이터 예측, 실시간 관측 및 분석, 의사결정의 4계층 구조로 구성해 예측 결과와 사실 정보 간의 차이를 정량적으로 분석하여 예측 신뢰 확률과 사실 신뢰 확률을 산출하고 두 확률을 동적 가중치로 융합시킨 최종 진실 확률이 추출되도록 설계하였다. 모의실험 결과, 7QCI 정보 제공 환경상에서 예측 신뢰 확률이 크게 변동하더라도 최종 진실 확률을 안정적으로 유지시킴으로써, 허위정보로 인한 오인 가능성 완화 및 진실 확률 추정 기반의 인공지능 의사결정 최적화 기준을 확보하였다.

[Abstract]

To mitigate AI decision-making misjudgment caused by false information inflow in a 7QCI(Seven Quality Class Indicator) information provision environment, this study proposes a truth probability estimation-based artificial intelligence decision optimization algorithm. The proposed algorithm consists of four layers: data collection, data prediction, real-time observation and analysis, and decision-making. It derives predictive reliability probability and factual reliability probability by analyzing the differences between predicted results and actual data. These probabilities are integrated through dynamic weighting to estimate the final truth probability. This approach helps overcome the limitations of prediction-based decision-making under conditions of false information and input parameter distortion. Simulation results demonstrate that the final truth probability remains stable even when the predictive reliability probability varies significantly. The proposed method provides a reliable and effective basis for artificial intelligence decision optimization in a 7QCI environment.

색인어 : 인공지능 의사결정, 최종 진실 확률, 사실 신뢰, 예측 신뢰, 7QCI**Keyword** : AI Decision, Final Truth Probability, Factual Reliability, Predictive Reliability, 7QCI<http://dx.doi.org/10.9728/dcs.2026.27.5.1359>

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Received 11 March 2026; Revised 13 April 2026

Accepted 13 April 2026

*Corresponding Author; Sung-Hyun Seo

Tel: E-mail: sshyun1472@nate.com

I. 서론

최근 인공지능(Artificial Intelligence, AI)은 다양한 지능형 시스템에서 핵심 의사결정 엔진으로 활용되고 있다. 딥러닝 기반 모델은 대규모 데이터로부터 복잡한 패턴을 학습하여 예측 및 추론을 수행하며, 이러한 결과는 자동화된 정책 결정과 제어 과정에 직접 반영된다[1]. 그러나, AI 기반 의사결정은 입력 데이터의 품질과 신뢰성에 구조적으로 의존하며, 데이터의 진위 여부가 충분히 검증되지 않은 상태에서 수행되는 추론은 시스템 수준의 오류로 확장될 수 있다.

AI 모델은 확률적 출력(probabilistic output)을 생성하지만, 실제 운용 환경에서는 이를 단일 결정값으로 해석하여 정책 판단에 활용하는 경우가 많다. 이 과정에서 입력 정보가 허위정보임에도 불구하고 진실정보로 오인될 경우, 모델은 통계적으로는 일관된 추론을 수행하더라도 결과적으로 왜곡된 판단을 초래하게 된다. 이러한 문제는 단순한 분류 오류(misclassification)를 넘어, 의사결정 단계에서의 오판(misjudgment)으로 이어질 수 있다.

특히, AI 추론 모델은 학습 데이터의 통계적 특성을 기반으로 설계되며, 입력 정보의 진위 자체를 별도로 검증하는 메커니즘은 일반적으로 포함되지 않는다. 따라서, 허위정보가 모델 입력으로 유입될 경우, AI 추론 모델은 이를 사실로 가정하여 확률 계산을 수행하게 되어 후속 정책 결정에 직접 영향을 미치는 문제점이 있다[2].

또한, 딥러닝 모델은 높은 예측 성능을 제공하지만 내부의 사결정 과정이 불투명한 블랙박스 특성을 가진다[3]. 이로 인해 특정 판단 결과가 도출된 근거를 명확히 설명하기 어렵고, 입력 정보의 신뢰도가 낮은 경우에도 동일한 가중 구조로 추론이 수행된다. 결과적으로 예측 정확도 향상만으로는 허위정보 오인 문제를 근본적으로 해결하기 어렵다.

기존 연구들은 주로 모델 구조 개선이나 학습 알고리즘 최적화를 통해 예측 정확도를 향상시키는 데 초점을 두어 왔다. 그러나, 이러한 접근은 입력 정보의 진실성 검증 문제를 전혀 다루지 않았다. 즉, 모델의 출력 확률은 존재하지만, 입력 정보 자체의 진실 확률(truth probability)을 추정하고 이를 의사결정 과정에 반영하는 접근 방식은 전혀 논의된 바 없다 하겠다[4].

이에 본 연구논문은 허위정보로 인한 AI 오인을 개선시키기 위하여 진실 확률 추정(Truth Probability Estimation, TPE) 기반의 AI 의사결정 최적화 알고리즘을 제안한다.

즉, 7QCI 정보 제공 환경상의 허위정보 오인 개선을 위한 진실 확률 추정 기반 인공지능 의사결정 최적화 알고리즘은 예측 신뢰 확률 $P_{p,i}(t)$ 와 사실 신뢰 확률 $P_{f,i}(t)$ 을 분리하여 산출하고, 두 확률을 동적 가중 결합함으로써 최종 진실 확률 $P_{t,i}(t)$ 을 재정의하는 구조를 제안한다. 이때 동적 가중치 $\beta_i(t)$ 는 예측 정보의 신뢰도를 반영하며, 예측 결과와 사실 정보 간 일관성에 따라 적응적으로 조정된다.

즉, 본 연구논문의 핵심은 예측 확률 자체를 사용하는 것이 아니라, 사실 검증과 예측 불확실성을 반영하여 보정된 진실 확률을 의사결정 입력으로 사용하는 데 있다[4],[5]. 이를 통해 허위정보로 인한 확률 왜곡을 방지하고자 하는 것이라 하겠다.

본 연구논문은 7QCI 유형별 오류 비용의 비대칭성을 고려하여, 진실 추정 임계값을 차등적으로 설정함으로써 허위 정보의 허용(False Positive)과 정상 정보의 차단(False Negative) 간 균형을 최적화하는 인공지능 의사결정 알고리즘을 제안한다. 특히, 현실 환경에서 빈번하게 생성·유통되는 허위정보로 인해 발생할 수 있는 오학습과 왜곡된 추론 문제를 완화하고, 단순한 예측 오차를 넘어 서비스 차단 오류, 자원 배분 왜곡, 정보보안 오판, 금융, IoT, 미디어 서비스의 판단 오류 등 다양한 형태의 잠재적 피해를 최소화하고자 한다. 이를 통해, 불확실성이 높은 정보 환경에서도 보다 신뢰성 있는 인공지능 추론과 안정적인 의사결정 수행이 가능하도록 하는 데 연구의 목적이 있다.

II. 기존 AI 의사결정 알고리즘

인공지능 시스템상의 의사결정 알고리즘은 입력 데이터로부터 특정 행동을 선택하는 핵심 기능을 수행한다. 의사결정 알고리즘은 입력 벡터 x 가 주어졌을 때 최종 결정 $D(x)$ 을 출력하는 함수로 정의할 수 있다.

이러한 의사결정 알고리즘은 입력 정보의 처리 방식과 학습 원리에 따라 여러 유형으로 구분될 수 있으며, 대표적으로는 분류 기반, 확률(베이지) 기반, 및 강화학습 기반의 의사결정 알고리즘으로 구분할 수 있다.

2-1 분류 기반 의사결정 알고리즘

분류 기반 의사결정 알고리즘은 가장 널리 활용되는 인공지능 결정 기법인 것으로, 입력 데이터가 특정 범주에 속할 가능성을 계산하고 그 중 가장 가능성이 높은 선택을 결정으로 채택하는 구조를 갖는다. 이 기법은 지도학습(supervised learning)에 기반하며, 과거 데이터로부터 학습된 모델이 새로운 입력에 대해 확률 또는 점수를 산출하는 형태로 구현된다.

(이진)분류 기반 의사결정 알고리즘은 식 (1)과 같다.

$$D(x) = \begin{cases} 1, & P(y=1|x) \geq \tau \\ 0, & P(y=1|x) < \tau \end{cases} \quad (1)$$

여기서, $P(y=1|x)$ 는 입력 벡터 x 가 특정 클래스(예: 이상, 승인, 스팸 등)에 속할 확률을 의미하며, τ 는 사전에 설정된 임계값이다.

예를 들어, 스팸 메일 필터링 시스템은 E-메일의 제목, 발

신자, 본문 단어 분포 등의 특징을 입력으로 받아 해당 메일이 스팸일 확률을 계산한다. 계산된 확률이 일정 기준 이상이면 스팸으로 분류하고, 그렇지 않으면 일반 메일로 처리한다. 이 과정은 매우 직관적이며, 인공지능 시스템에서도 가장 보편적으로 사용되는 구조이다.

금융 분야에서도 동일한 방식이 활용된다. 고객의 신용 점수, 소득 수준, 거래 이력 등을 입력으로 하여 대출 부도 확률을 계산하고, 그 확률이 기준값을 넘으면 대출을 거절하는 방식이다. 이처럼 분류 기반 의사결정은 “확률이 높으면 선택한다”는 단순한 논리를 따른다.

이 기법은 계산이 빠르고 구현이 용이하며, 다양한 모델(로지스틱 회귀, 의사결정 트리, 신경망 등)에 쉽게 적용할 수 있다 하겠으나, 입력 정보가 왜곡되었거나 환경이 변화한 경우에도 동일한 기준을 적용한다는 한계가 있다. 또한, 임계값이 고정되어 있을 경우, 서로 다른 IT 서비스 환경이나 사용 목적에 따른 유연한 조정이 어렵다 할 것이다.

2-2 확률론적 의사결정 알고리즘

확률론적 의사결정 알고리즘은 통계적 추론 이론을 바탕으로 불확실성을 정량적으로 반영하는 구조를 갖는다. 특히 베이즈 기반 접근은 기존에 알고 있던 확률 개념과 새롭게 관측된 여럿 정보를 결합하여 최종 판단을 내리는 기법을 취한다고 하겠다[5],[6]. 이는 단순히 모델 출력 확률을 사용하는 것을 넘어, 기존 정보와 새로운 데이터를 종합적으로 고려한다는 특징을 가진다.

베이즈 기반 의사결정 알고리즘은 식 (2)와 같다.

$$P(y|x) = \frac{P(x|y) \cdot P(y)}{P(x)} \quad (2)$$

여기서, $P(y)$ 는 사건의 사전 확률(prior), $P(x|y)$ 는 사실 정보의 조건부 확률(likelihood) 및 $P(y|x)$ 는 관측 이후의 사후 확률(posterior)을 의미한다.

예를 들어, 의료 진단 시스템은 특정 질병의 기본 발생 확률이 낮더라도, 검사 결과가 양성으로 나올 경우 해당 질병의 가능성은 증가한다. 그러나, 단순히 “양성이므로 질병이다”라고 판단하지 않고, 기존 발생 확률과 검사 정확도를 함께 고려하여 최종 확률을 계산한다. 이러한 방식은 불확실한 환경에서 보다 합리적인 결정을 가능하게 한다.

이 기법은 확률을 지속적으로 갱신(update)하는 방식으로 새로운 정보가 들어올 때마다 판단이 조정되기 때문에 데이터가 점진적으로 축적되는 환경에서 효과적이다. 또한, 확률 기반이어서 판단의 논리적 근거를 수치로 제시할 수 있다 할 것이다.

그러나, 이 기법은 사전 확률 설정에 따라 결과가 달라질 수 있으며, 복잡한 모델에서는 계산 비용이 증가할 수 있다.

무엇보다도 최종 단계에서는 여전히 확률을 비교하여 결정함으로써, 기준 설정의 문제가 해결된 바 없다 할 것이다.

2-3 강화학습 기반 의사결정 알고리즘

강화학습 기반 의사결정 알고리즘은 단일 시점의 판단이 아니라, 시간에 따라 반복되는 순차적 의사결정 문제를 해결하기 위한 구조를 갖는다. 이 기법은 현재 상태에서 특정 행동을 선택하고, 그 결과로 얻는 보상을 통해 결정 규칙을 점진적으로 개선한다. 설계 목적은 상태-행동-보상 상호작용을 반복적으로 학습함으로써 장기적으로 누적 보상을 최대화하는 것에 있다[7].

누적 보상 의사결정 알고리즘은 식 (3)과 같다.

$$G = \sum_{t=0}^T \alpha^t r_t \quad (3)$$

여기서, r_t 는 시점 t 에서 얻는 보상, $(0 < \alpha < 1)$ 는 미래 보상의 중요도를 조절하는 가중치이다.

예를 들어, 자율주행 차량은 현재 속도와 주변 차량의 위치를 고려하여 가속, 감속, 차선 변경 등의 행동을 선택한다. 이 선택은 단순히 현재 상황만을 고려하는 것이 아니라, 이후 교통 흐름과 안전성에 영향을 미친다. 강화학습은 이러한 장기적인 결과를 고려하여 최적의 정책을 학습한다.

또 다른 예로, 네트워크 자원 관리 시스템을 들 수 있다. 특정 사용자에게 자원을 많이 할당하면 단기적으로는 효율이 높아 보일 수 있지만, 장기적으로는 다른 사용자에게 불리하게 작용할 수 있다. 강화학습은 이러한 시간적 상호작용을 반영하여 전체 시스템 효율을 최적화한다.

이 기법은 동적 환경에 적합하고, 복잡한 정책 학습이 가능하긴 하나, 충분한 학습 과정이 필요하며 상태 공간이 커질수록 계산 부담이 증가한다. 또한, 보상 함수 설계가 부적절할 시, 비합리적인 결정 규칙이 학습될 위험이 존재한다.

III. 진실 확률 추정 기반 인공지능 의사결정 최적화 알고리즘

본 장에서는 7QCI 정보 제공 환경에서 발생할 수 있는 허위정보 오인 문제를 최대한 개선하고, 이에 따른 인공지능의 사결정 오류를 최소화하기 위한 진실 확률 추정 기반 의사결정 최적화 알고리즘을 제안한다.

7QCI 기반 정보 환경에서는 다양한 품질 등급(QoS Class Identifier)에 따라 상이한 트래픽 특성과 서비스 요구 조건이 공존하며, 이 과정에서 외부 이벤트, 비정상적 데이터 변동, 또는 의도적 허위정보 유입으로 인해 입력 데이터의 신뢰성이 저하될 가능성이 존재한다. 기존의 단일 예측 모델 기반

의사결정 방식은 7QCI 기반 서비스 환경의 QoS 이질성과 학습-운영 간 분포 차이(distribution shift), 시간에 따른 개념 변화(concept drift)에 능동적으로 대응하지 못하는 구조적 한계를 가진다[8],[9]. 즉, 실시간 빅데이터 유입과정 중 허위 정보가 포함되어 있음에도 불구하고 이를 적절히 식별하지 못한 채 사실 정보로 간주하여 미래 상태를 그대로 예측하게 됨으로 말미암아 궁극적으로는 시스템 오판의 가능성을 증대시킬 위험이 있다.

이에 따라, 본 연구논문은 7QCI 정보 제공 환경상에서 유입되는 대규모 데이터 중 허위정보로 인해 왜곡될 수 있는 미래 상태 예측값과 현재 시점에서 관측되는 사실 정보 간의 차이를 정량적으로 분석하고, 그 오차의 크기와 시간적 변화 추이를 반영한 진실 확률을 추정하고자 한다. 추정된 진실 확률은 의사결정 계층에서 동적 가중치를 조정 인자로 활용하여 인공지능 예측값과 사실 정보의 반영 비율을 적응적으로 재조정함으로써 최종 판단의 신뢰성과 안정성을 향상시키도록 설계된다[4]. 즉, 본 연구논문에서는 진실 확률 추정을 기반으로 허위정보 오인을 완화하고, 동적 환경 변화에 보다 유연하게 대응할 수 있는 인공지능 의사결정 최적화 알고리즘을 제안한다.

본 연구논문에서 제안하는 알고리즘은 다음과 같은 4계층 모델로 구성된다.

1. 데이터 수집 계층
2. 데이터 예측 계층
3. 실시간 관측 및 분석 계층
4. 의사결정 계층

데이터 수집 계층에서는 인공지능 시스템 상태 및 사실 정보를 통합적으로 확보하고, 데이터 예측 계층에서는 이를 기반으로 미래 상태를 추정한다. 실시간 관측 및 분석 계층에서는 예측 결과와 사실 정보 간의 차이를 정량적으로 분석하여 예측 오차를 산출하고, 오차의 누적 정도와 환경 이탈 수준을 지속적으로 평가함으로써 현재 운용 환경이 학습 당시의 조건으로부터 얼마나 벗어났는지를 판단한다. 또한, 인공지능 예측값과 사실 정보 간의 차이가 설정된 임계값을 초과할 경우, 의사결정 계층에서는 예측 신뢰 확률의 반영 비중을 감소시키고 사실 신뢰 확률의 반영 비중을 증가시키고 동적 가중치를 동적으로 조정하는 메커니즘을 적용한다. 이를 통해, 7QCI 정보 환경에서 발생 가능한 허위정보 오인을 완화시키고 분포 변동 상황에서도 의사결정 편향을 최소화함으로써 동적 네트워크 환경 하에서 안정적이고 신뢰성 있는 인공지능 의사결정 수행이 가능하도록 한다.

3-1 데이터 수집 계층 모델

데이터 수집 계층 모델은 인공지능 시스템 운영 과정에서 발생하는 사실 정보 $x_i(t)$, 이벤트 스트림 $e_{k,i}(t)$, 이력 데이터

집합 $H_i(t)$ 를 지속적으로 수집, 저장하는 기능을 수행한다. 이 정보들($x_i(t)$, $e_i(t)$, $H_i(t)$)은 단순한 상태 기록만이 아니라, 이후 계층에서 규명될 진실 확률 추정(Truth Probability Estimation) 및 신뢰도 평가(Reliability Assessment)를 위한 기초 데이터로 활용된다. 즉, 데이터 수집 계층 모델은 시점 t에서의 사실 정보 $x_i(t)$, 이벤트 정보 $e_i(t)$, 그리고 이력 데이터 집합 $H_i(t)$ 을 통합적으로 관리함으로써, 이후 수행되는 예측, 실시간 분석, 및 의사결정 과정에 필요한 신뢰성 있는 입력 정보들을 제공한다.

• 사실 정보(Factual Information, $x_i(t)$)는 7QCI 정보 제공 환경하에 바라본 시점 t에서의 인공지능 시스템의 실제 운영 상태를 사실적으로 나타낸 정보로, 식 (4)와 같이 n차원 상태 변수로 정의된다. 사실 정보는 이후 계층에서 산출되는 사실 신뢰 확률(Factual Reliability Probability, $P_{f,i}(t)$)를 추정하기 위한 근거 데이터로 활용된다.

$$x_i(t) = [x_{1,i}(t), x_{2,i}(t), x_{3,i}(t), \dots, x_{n,i}(t)]^T \quad (4)$$

여기서, 사실 정보 $x_i(t)$ 는 시점 t에서 측정된 n번째 상태 변수들을 총칭하며, $\sum_1^n x_{n,i}(t)^T$ 는 시스템 부하, 자원 사용률, 처리 지연, 큐 길이 등 인공지능 시스템의 동작 특성을 총량으로 나타내는 지표이다.

각각의 n개 상태 변수들은 실시간으로 관측된 실제값 임으로, 예측 신뢰 확률 $P_{p,i}(t)$ 과 대비되는 사실 신뢰 확률 $P_{f,i}(t)$ 산정의 근거가 된다.

• 이벤트 스트림(Abrupt Change Cause, $e_i(t)$)은 7QCI 정보 제공 환경상의 시점 t에서의 인공지능 시스템 운영 과정에서 발생하는 비연속적(이산적)인 k사건(event)들의 시간적 집합으로 정의되며, 시간 순서에 따라 정렬된 이벤트들의 시계열 스트림으로 표현된다(식 (5) 참조).

$$e_i(t) = \{(e_{k,i}, \tau_k) | \tau_k \leq t\} \quad (5)$$

여기서, $e_{k,i}$ 는 k번째 이벤트를 의미하며, τ_k 는 해당 이벤트의 발생 시점을 나타낸다.

이벤트는 요청 도착, 오류 발생, 결정 규칙 변경, 외부 환경 변화 등과 같은 비연속적 시스템 변화 요인을 포함하며, 상태 변수의 급격한 변동을 유발하는 원인 변수로 작용한다.

이에 따라, 사실 정보의 변화는 식 (6)과 같이 표현될 수 있다.

$$x_i(t) = f(x_i(t-), e_i(t)) \quad (6)$$

사실 정보의 변화는 이벤트 스트림 $e_i(t)$ 가 직전 사실 정보 $x_i(t-)$ 에 영향을 미쳐 새로운 사실 정보 $x_i(t)$ 를 형성함을 나타내며, 이벤트 스트림 $e_i(t)$ 은 인공지능 시스템 상태를 전이

시키는 직접적인 유발 요인으로 작용한다.

• 이력 데이터 집합(Window Data History, $H_i(t)$)은 7QCI 정보 제공 환경 하에 시점 t 기준 과거 일정 구간의 시간 동안의 사실 정보와 이벤트 스트림을 포함하는 누적 빅데이터의 집합으로 정의된다(식 (7) 참조).

$$H_i(t) = \{z_i(\tau) \mid \tau \in [t-T, t]\}, \quad z_i(\tau) = \begin{bmatrix} x_i(\tau) \\ e_i(\tau) \end{bmatrix} \quad (7)$$

여기서, T 는 관측 윈도우 길이를 의미하며, $H_i(t)$ 는 시점 t 이전의 인공지능 시스템의 동작 패턴을 반영한다.

이력 데이터 집합 $H_i(t)$ 는 7QCI 정보 제공 환경 하에 데이터 예측 계층에서 예측 신뢰 확률 $P_{p,i}(t)$ 과 진실 신뢰 확률 $P_{f,i}(t)$ 을 산출하기 위한 학습 입력으로 활용된다.

인공지능 시스템의 운영 과정에서 이벤트 스트림 $e_i(t)$ 이 발생하면, 해당 이벤트는 시스템 상태에 영향을 미치며 사실 정보 $x_i(t)$ 의 변화를 유발한다.

이러한 변화는 시간 축에 따라 누적되어 이력 데이터 집합 $H_i(t)$ 를 구성한다(식 (8) 참조).

$$e_i(t) \rightarrow x_i(t) \rightarrow H_i(t) \quad (8)$$

따라서, 이력 데이터 집합 $H_i(t)$ 는 단순한 데이터의 집합이 아니라, 이벤트 스트림 $e_i(t)$ 의 발생 패턴, 사실 정보 $x_i(t)$ 의 변화 크기 및 방향성, 예측-관측 오차의 누적 특성을 동시에 내포하는 구조적 정보 집합 또는 빅데이터라 하겠다.

3-2 데이터 예측 계층 모델

데이터 예측 계층 모델은 데이터 수집 계층에서 제공된 이력 데이터 집합 $H_i(t)$ 을 입력값으로 삼아 미래 시점 $t+\Delta$ 에서의 인공지능 시스템 상태를 확률적으로 추론한다.

본 계층은 머신러닝 또는 딥러닝 기반 예측 모델을 활용하여 시스템 부하, 수요 변화, 상태 전이 가능성 등 주요 변수에 대한 사전 추론 결과를 생성한다. 예측 결과는 단일 결정값에 국한되지 않고 확률적 형태로 해석되어 미래 상태에 대한 조 건부 분포를 통해 표현된다.

• 인공지능 예측값 $\hat{x}(t)$ 은 시각 t 에서의 예측값으로, 실제로는 직전 시점까지의 이력 데이터 집합 $H_i(t)$ 을 바탕으로 산출된 한 시점 앞의 예측값이라 하겠다. 이는 직전 시점까지의 과거의 이력 데이터 집합 $H_i(t)$ 또는 빅데이터를 입력으로 학습된 인공지능 예측 모델 $F_i(\cdot)$ 를 통해 식 (9)와 같이 산출된다.

이는 과거 상태 변화 패턴과 이벤트 누적 특성을 반영한 미래 상태의 중심 경향을 나타낸 것으로, 주어진 정보 하에서

가장 가능성이 높은 미래 상태를 점 형태로 제시하는 예측 결과이다.

$$\hat{x}_i(t) = f_{\theta}(H_i(t)) \quad (9)$$

7QCI 정보 제공 환경 하에 인공지능 시스템에서 미래 상태는 단일 값으로 고정되어 결정되는 것이 아니라 다양한 외부 요인과 내재적 변동성의 영향을 받는다.

따라서, 본 연구논문에서는 미래 상태를 확률변수로 가정하고, 이를 예측 확률분포 $P(x_i(t) \mid H_i(t))$ 의 형태로 모델링한다.

• 즉, 예측 확률분포 $P_i(\hat{x})$ 는 미래 상태는 단일값이 아니라 확률변수로 간주되며, 과거 이력 $H_i(t)$ 가 주어졌을 때 미래 시점 $t+\Delta$ 의 상태는 평균과 불확실성을 함께 표현한 확률적 예측 구조로 모델링된다[5],[6](식 (10) 참조).

$$P_i(\hat{x}) = P(x_i(t) \mid H_i(t)) = N_i(\hat{x}_i(t+\Delta), \Sigma(t)) \quad (10)$$

N 은 정규분포(Normal distribution)이고, 공분산 행렬 $\Sigma(t)$ 는 상태 변수 간 상관관계 및 예측 오차의 분산을 포함하는 것으로, 예측값 주변에서의 변동 특성 및 미래 상태가 어느 범위 내에서 변화할 수 있는지를 정량적으로 제시한다. 즉, 공분산 값이 클수록 예측의 불확실성이 크며, 반대로 작을수록 예측이 안정적임을 의미한다.

예측 확률분포 $P_i(\hat{x})$ 는 현재 시점 t 에서의 과거 이력 $H(t)$ 가 주어졌을 때, 미래 시점 $t+\Delta$ 의 사실 정보 $x_i(t+\Delta)$ 는 평균이 $\hat{x}_i(t+\Delta)$ 이고 불확실성이 $\Sigma(t)$ 인 정규분포를 따른다.

이는 미래 상태를 단순한 점 추정값으로만 제시하는 것이 아니라 평균과 변동성을 함께 포함하는 확률 분포의 형태로 표현되는 접근이라 할 수 있다.

따라서, 데이터 예측 계층은 미래 상태가 정확히 하나의 값으로 정해져 있는 것이 아니라 일정한 불확실성을 내포한 확률적 범위 내에서 존재한다는 전제하에 모델링되며, 이는 과거 이력 데이터 집합 $H_i(t)$ 을 바탕으로 미래 상태의 평균값과 불확실성을 갖는 정규분포로 모델링된다 하겠다.

3-3 실시간 관측 및 분석 계층 모델

실시간 관측 및 분석 계층 모델은 7QCI 정보 제공 환경 하에 바라본 현재 시점에서의 인공지능 시스템 상태를 즉각적으로 반영하고, 사실 정보를 정량화하는 역할을 수행한다.

본 계층은 확률 지표 변화율 $P_i'(\hat{x})$, 이벤트 감지함수 $E_i(\hat{x})$ 를 활용한 이벤트 발생 여부 등을 분석해 인공지능 시스템의 단기적 동적 특성을 평가하며, 이를 통해 현재 7QCI를 제공하는 불완전한 ICT 환경이 안정적인지 혹은, 급격한 구조적 변화가 발생했는지를 판단하고, 사실 기반 의사결정 근거를 마련한다.

• 확률 지표 변화율 $P'_i(\hat{x})$ 은 일정 구간 Δt 동안 인공지능 시스템의 예측 확률분포 $P'_i(\hat{x})$ 가 얼마나 빠르게 변화하는지를 나타내는 시간적 미분 근사값으로 식 (11)처럼 정의된다. 이는 단순한 상태값의 증감을 나타내는 것이 아니라, 인공지능 시스템의 신뢰도 또는 위험도가 시간에 따라 어떤 방향성과 속도로 이동하고 있는지를 정량적으로 표현하는 동적 안정성 지표라 할 수 있다.

$$P'_i(\hat{x}) = \frac{P_i(x_i(t)) - P_i(x_i(t - \Delta t))}{\Delta t} \quad (11)$$

확률 지표 변화율 $P'_i(\hat{x})$ 는 다음과 같이 해석된다. 즉, $|P'_i(\hat{x})|$ 가 예측오차 임계값보다 작으면 인공지능 시스템이 안정 범위 내에서 운영되고 있음을 뜻하며, $|P'_i(\hat{x})|$ 가 예측오차 임계값보다 크면 급격한 환경 변화 가능성이 존재한다는 것을 의미한다고 하겠다.

• 이벤트 감지함수(Threshold-based Event Function, $E_i(\hat{x})$)는 식 (12)에서 보여지는 바와 같이, 예측 오차 임계값과의 비교를 통해 인공지능 시스템의 비정상 상태 진입 여부를 판별하는 함수이다.

$$E_i(\hat{x}) = \begin{cases} 1, & |P'_i(\hat{x})| > \theta_{e,i} \\ 0, & otherwise \end{cases} \quad (12)$$

여기서, $\theta_{e,i}$ (예측 오차 임계값) = $\frac{C_{FP,i}}{C_{FP,i} + C_{FN,i}}$, $C_{FP,i}$: 허위 정보를 허용했을 시 소요비용, $C_{FN,i}$: 사실 정보를 차단했을 시 소요비용, $E_i(\hat{x})=1$: 이벤트 발생, $E_i(\hat{x})=0$: 정상 상태 유지를 의미한다.

다시 말해, $\theta_{e,i}$ 는 허위정보를 허용했을 때의 비용과 사실정보를 차단했을 때의 비용을 함께 반영하는 비용에 민감한 임계값이라 하겠고, 이는 정보 수용 성향의 보수성 또는 공격성을 조절하는 기준이 된다[10].

$C_{FP,i}$ 가 커져 $\theta_{e,i}$ 이 올라가면, 커트라인이 상승되어 허위 정보 뿐만 아니라 사실 정보도 배제될 가능성이 높아져 정보 유입에 대한 성향이 보수성(보다 높은 신뢰도를 갖는 정보만을 선택적으로 받아들여지게 되므로 허위 정보가 유입될 가능성은 감소되나 사실 정보도 더불어 배제됨)을 띈다.

$C_{FN,i}$ 가 작아 $\theta_{e,i}$ 이 내려가면, 커트라인이 하락되어 사실 정보를 비롯해 허위 정보까지 수용할 가능성이 높아져 정보 유입에 대한 성향이 공격성(보다 많은 정보를 수용하려는 성향을 보이며 사실 정보가 인공지능 시스템에 반영될 가능성은 높아지나 의사결정 과정에서 허위 정보도 포함될 가능성이 존재함)을 갖는다.

3-4 의사결정 계층 모델

의사결정 계층 모델은 예측 신뢰 확률 $P_{p,i}(t)$ 과, 사실 신뢰 확률 $P_{f,i}(t)$ 을 종합하여 최종 진실 확률 $P_{t,i}(t)$ 을 산출하고, 이를 기반으로 7QCI 정보 제공 환경하에 구축된 인공지능 시스템의 제어 결정을 수행할 수 있을 것이다[11].

• 예측 신뢰 확률 $P_{p,i}(t)$ 은 식 (13)과 같이, 인공지능 예측값 $\hat{x}_i(t)$ 과 사실 정보 $x_i(t)$ 의 차이가 작을수록 예측을 더 신뢰할 수 있다. 즉, 두 값의 차이는 단순한 수치 비교를 넘어 현재 운용 환경에서 인공지능 예측이 얼마나 현실에 부합하는지를 판단하는 핵심 기준으로 작용한다.

예를 들어, 특정 클래스 i 에서 $P_{p,i}(t)$ 가 높게 산출되면 인공지능 예측 결과를 적극적으로 반영할 수 있으며, 반대로 $P_{p,i}(t)$ 가 낮게 나타나면 예측 결과에 대한 의존도를 줄이고 사실 정보의 반영 비율을 높이는 방향으로 의사결정 전략을 조정할 수 있다.

$$P_{p,i}(t) = 1 - |\hat{x}_i(t) - x_i(t)| \quad (13)$$

여기서, 7QCI 중 클래스 i 에 대한 인공지능 예측값 $\hat{x}_i(t)$, 7QCI 중 클래스 i 에 대한 사실 정보 $x_i(t)$ 이다.

• 사실 신뢰 확률 $P_{f,i}(t)$ 은 식 (14)와 같이, 사실 정보 $x_i(t)$ 의 크기와 예측 오차 $Er_i(t)$ 의 영향을 동시에 반영하여 산정된다. 즉, 사실 정보 $x_i(t)$ 이 클수록 해당 정보가 반영하는 사실성의 수준은 높아지지만, 예측 오차 $Er_i(t)$ 가 함께 증가할 경우 그 사실 신뢰 확률 $P_{f,i}(t)$ 는 감소하게 된다. 반대로 예측 오차 $Er_i(t)$ 가 작을수록 사실 정보 $x_i(t)$ 은 보다 안정적이고 일관된 사실 정보 $x_i(t)$ 로 간주될 수 있다. 즉, 사실 신뢰 확률 $P_{f,i}(t)$ 은 실제 관측된 사실 정보 $x_i(t)$ 의 유효성과 안정성을 동시에 정량화하는 역할을 한다.

$$P_{f,i}(t) = x_i(t)(1 - Er_i(t)) \quad (14)$$

여기서, 예측 오차 $Er_i(t)$ 는 동일 시각에서의 예측 오차 또는 관측-예측 간 불일치를 나타내는 지표이다.

• 최종 진실 확률 $P_{t,i}(t)$ 은 단일 정보에 의존하지 않고, 두 확률 정보를 융합해 일관성, 안정성 및 적응성을 동시에 확보할 수 있음에 따라, 7QCI 중 클래스 i 에 대해 식 (15)와 같이 정의할 수 있다.

$$P_{t,i}(t) = \beta_i(t)P_{p,i}(t) + (1 - \beta_i(t))P_{f,i}(t) \quad (15)$$

여기서, 동적 가중치 $\beta_i(t) = \frac{P_{p,i}(t)}{P_{p,i}(t) + P_{f,i}(t)}$: 예측 신뢰 확률과 사실 신뢰 확률을 상대적으로 균형있게 조절하기 위해 사용되는 계수로, $0 \leq \beta_i(t) \leq 1$ 범위를 갖는다.

예를 들어, 특정 시점에서 예측 신뢰 확률 $P_{p,i}(t)$ 가 0.8이고 사실 신뢰 확률 $P_{f,i}(t)$ 가 0.2인 경우, $\beta_i(t)$ 는 $0.8/(0.8+0.2)=0.8$ 로 계산된다. 이는 전체 의사결정 과정에서 약 80%의 비중이 인공지능 예측 정보를 기반으로 이루어짐을 의미하며, 인공지능 시스템이 예측 결과를 상대적으로 높은 신뢰도로 판단하고 있음을 나타낸다.

반대로 예측 신뢰 확률 $P_{p,i}(t)$ 가 0.3이고 사실 신뢰 확률 $P_{f,i}(t)$ 가 0.7인 경우에는 $\beta_i(t)$ 가 $0.3/(0.3+0.7)=0.3$ 으로 계산된다. 이 경우 의사결정 과정에서 예측 정보가 차지하는 비중은 약 30%에 불과하며, 인공지능 시스템은 예측 결과보다 실제 관측된 사실 정보에 더 큰 비중을 두어 판단을 수행하게 된다.

이와 같이, $\beta_i(t)$ 는 예측 정보와 사실 정보 사이의 상대적 신뢰도를 반영하여 두 정보의 영향력을 동적으로 조절하는 역할을 하며, 시스템이 특정 정보에 과도하게 의존하지 않도록 균형 잡힌 의사결정을 가능하게 한다.

IV. 진실 확률 추정 기반 인공지능 의사결정 최적화 알고리즘에 관한 모의실험

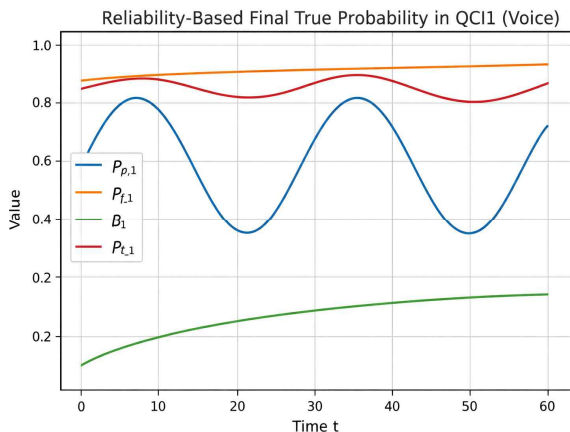


그림 1. QC11(Voice 서비스) 환경에서의 예측(사실) 확률, 동적 가중치 및 최종 진실 확률의 동적 변화를 나타낸 그래프

Fig. 1. Shows the dynamic changes of predictive/factual probability, adaptive weight, and final truth probability in QC11 (voice service)

그림 1은 QC11(Voice) 서비스 환경에서 시간 변화에 따라 예측 신뢰 확률 $P_{p,1}(t)$, 사실 신뢰 확률 $P_{f,1}(t)$, 동적 가중치 $\beta_1(t)$, 그리고 이들을 융합하여 산출한 최종 진실 확률 $P_{t,1}(t)$ 의 동적 변화를 나타낸 것이다.

그림 1을 살펴보면, 예측 신뢰 확률 $P_{p,1}(t)$ 이 크게 하락하는 구간에서도 최종 진실 확률 $P_{t,1}(t)$ 은 급격히 붕괴하지 않고 비교적 안정된 범위 내에서 유지된다는 것이다. 이는 사실 신뢰 확률 $P_{f,1}(t)$ 이 보완적 역할을 수행해 예측 오차 $Er_1(t)$ 확대에 따른 신뢰도 저하를 상쇄하고 있음을 보여준다. 반대로 예측 신뢰 확률 $P_{p,1}(t)$ 이 상승하는 구간에서는 최종 진실

확률 $P_{t,i}(t)$ 도 이에 상응하여 완만하게 상승함으로써, 인공지능 예측값 $\hat{x}_i(t)$ 의 유효성이 높아진 상황에서는 그 신뢰도가 상승될 수 있도록 적절히 반영하고 있음을 확인할 수 있다.

7QCI 중 QC11은 음성 서비스로, 낮은 지연과 연속적인 전송 품질 보장이 핵심 요소라 하겠다. 따라서, 순간적인 예측 오차 $Er_1(t)$ 나 외부 환경 변화에 의해 의사결정이 크게 흔들릴 경우 서비스 품질 저하로 직결될 가능성이 높다. QC11 환경에서는 단순히 평균적인 예측 정확도를 높이는 것만으로는 충분하지 않으며, 순간적인 예측 오차나 외부 환경 변화에 의해 제어 판단이 급격히 변동하지 않도록 안정적인 의사결정 구조를 확보하는 것이 중요하다 할 것이다.

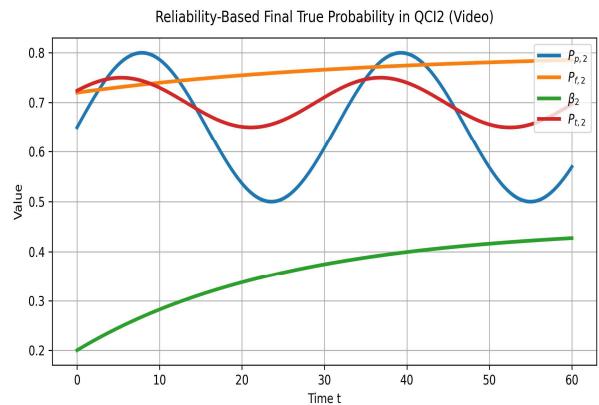


그림 2. QC12(Video 서비스) 환경에서의 예측(사실) 확률, 동적 가중치 및 최종 진실 확률의 동적 변화를 나타낸 그래프

Fig. 2. Shows the dynamic changes of predictive/factual probability, adaptive weight, and final truth probability in QC12 (video service)

그림 2는 QC12(Video) 서비스 환경에서 시간 변화에 따라 예측 신뢰 확률 $P_{p,2}(t)$, 사실 신뢰 확률 $P_{f,2}(t)$, 동적 가중치 $\beta_2(t)$, 그리고 이들을 융합하여 산출한 최종 진실 확률 $P_{t,2}(t)$ 의 동적 변화를 나타낸 것이다.

그림 2를 살펴보면, 예측 신뢰 확률 $P_{p,2}(t)$ 은 시간에 따라 비교적 큰 폭으로 상승과 하락을 반복하는 반면, 사실 신뢰 확률 $P_{f,2}(t)$ 은 전체적으로 더 높은 수준에서 완만하게 증가하는 경향을 보인다. 또한, 동적 가중치 $\beta_2(t)$ 는 초기의 낮은 값에서 시작하여 시간 경과에 따라 점진적으로 증가하는 형태를 나타낸다.

즉, 예측 신뢰 확률 $P_{p,2}(t)$ 이 하락하는 구간에서도 최종 진실 확률 $P_{t,2}(t)$ 은 급격히 붕괴하지 않고 비교적 안정된 범위 내에서 유지된다는 것이다. 이는 사실 신뢰 확률 $P_{f,2}(t)$ 이 보완적 역할을 수행함으로써, 인공지능 예측값 $\hat{x}_2(t)$ 의 일시적 저하가 최종 판단에 과도하게 반영되지 않도록 조정되고 있음을 보여준다. 반대로, 예측 신뢰 확률 $P_{p,2}(t)$ 이 상승하는 구간에서는 최종 진실 확률 $P_{t,2}(t)$ 도 상승하는 경향을 보이나, 그 상승 폭은 예측 신뢰 확률 $P_{p,2}(t)$ 에 비해 상대적으로 제한적으로 나타난다. 특히, 일부 구간에서는 최종 진실 확률

$P_{t,2}(t)$ 이 예측 신뢰 확률 $P_{p,2}(t)$ 보다 높은 수준 또는 낮은 수준으로 조정 형성되는데, 이는 최종 진실 확률 $P_{t,2}(t)$ 이 인공 지능 예측값 $\hat{x}_1(t)$ 을 그대로 추종하는 것이 아니라 사실 신뢰 확률 $P_{f,2}(t)$ 과의 융합을 통해 보다 안정적인 방향으로 보정되고 있음을 의미한다.

7QCI 중 QCI2는 (대화형) 영상 서비스로, 연속적인 데이터 전송과 안정적인 재생 품질 유지가 핵심 요소로 작용한다. (대화형) 영상 서비스는 시간에 따른 상태 변화에 민감하며, 일시적인 예측 오차나 외부 환경 변화가 발생할 경우 버퍼링, 화질 저하, 재생 끊김 등의 형태로 사용자 체감 품질이 저하될 가능성이 높다. 따라서, QCI2 환경에서는 단순히 평균적인 예측 정확도를 높이는 것만으로는 충분하지 않으며, 인공 지능 예측값 $\hat{x}_2(t)$ 이 순간적으로 크게 변동하더라도 최종 판단이 과도하게 흔들리지 않도록 안정적인 의사결정 구조를 유지하는 것이 중요하다 할 것이다.

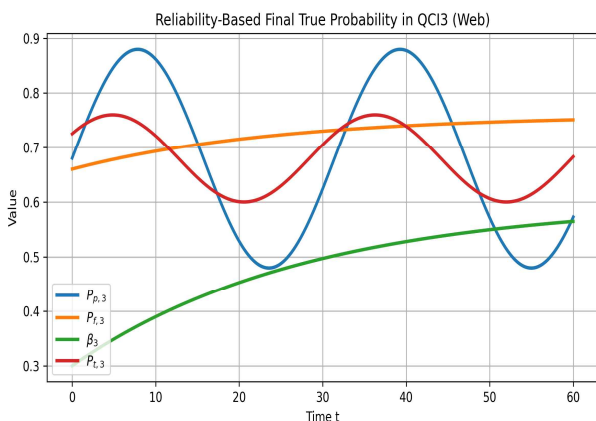


그림 3. QCI3(Web 서비스) 환경에서의 예측(사실) 확률, 동적 가중치 및 최종 진실 확률의 동적 변화를 나타낸 그래프

Fig. 3. Shows the dynamic changes of predictive/factual probability, adaptive weight, and final truth probability in QCI3 (web service)

그림 3은 QCI3(Web) 서비스 환경에서 시간 변화에 따라 예측 신뢰 확률 $P_{p,3}(t)$, 사실 신뢰 확률 $P_{f,3}(t)$, 동적 가중치 $\beta_3(t)$, 그리고 이들을 융합하여 산출한 최종 진실 확률 $P_{t,3}(t)$ 의 동적 변화를 나타낸 것이다.

그림 3을 살펴보면, 예측 신뢰 확률 $P_{p,3}(t)$ 은 시간에 따라 비교적 큰 폭의 주기적 상승과 하락을 반복하고 있으며, 대체로 두 차례 뚜렷한 상승 구간과 두 차례 하락 구간이 나타난다. 반면, 사실 신뢰 확률 $P_{f,3}(t)$ 은 전체적으로 상대적으로 안정된 수준에서 완만하게 증가하는 경향을 보이며, 예측 신뢰 확률 $P_{p,3}(t)$ 에 비해 변동성이 훨씬 작게 나타난다. 또한, 동적 가중치 $\beta_3(t)$ 는 초기의 낮은 수준에서 시작하여 시간 경과에 따라 점진적으로 증가하는 형태를 보인다. 최종 진실 확률 $P_{t,3}(t)$ 은 예측 신뢰 확률 $P_{p,3}(t)$ 의 급격한 변동을 그대로 따르지 않고, 전체적으로 보다 완만하고 안정적인 범위 내에서

변화하는 특성을 나타낸다.

즉, 예측 신뢰 확률 $P_{p,3}(t)$ 이 하락하는 구간에서도 최종 진실 확률 $P_{t,3}(t)$ 은 급격히 붕괴하지 않고 비교적 안정된 범위 내에서 유지된다는 것이다. 이는 사실 신뢰 확률 $P_{f,3}(t)$ 이 보완적 역할을 수행함으로써, 인공지능 예측값 $\hat{x}_3(t)$ 의 일시적 저하가 최종 판단에 과도하게 반영되지 않도록 조정되고 있음을 보여준다. 또한, 예측 신뢰 확률 $P_{p,3}(t)$ 이 중간 구간과 후반 구간에서 상대적으로 큰 폭으로 하락하더라도 최종 진실 확률 $P_{t,3}(t)$ 은 그보다 훨씬 완만한 하락 형태를 보이며 일정 수준 이하로 급격히 낮아지지 않는다. 반대로, 예측 신뢰 확률 $P_{p,3}(t)$ 이 상승하는 구간에서는 최종 진실 확률 $P_{t,3}(t)$ 도 상승하는 경향을 보이나, 그 상승 폭은 예측 신뢰 확률 $P_{p,3}(t)$ 에 비해 상대적으로 제한적으로 나타난다. 특히, 일부 구간에서는 최종 진실 확률 $P_{t,3}(t)$ 이 예측 신뢰 확률 $P_{p,3}(t)$ 보다 높은 수준 또는 낮은 수준으로 조정되어 형성되는데, 이는 최종 진실 확률 $P_{t,3}(t)$ 이 인공지능 예측값 $\hat{x}_3(t)$ 을 그대로 추종하는 것이 아니라 사실 신뢰 확률 $P_{f,3}(t)$ 과의 융합을 통해 보다 안정적인 방향으로 보정되고 있음을 의미한다.

7QCI 중 QCI3는 Web 서비스로, 응답성 확보와 안정적인 데이터 전달이 핵심 요소로 작용한다. Web 서비스는 음성이나 대화형 영상 서비스와 같이 연속적인 실시간 재생 품질이 절대적으로 요구되진 않지만, 네트워크 상태 변화나 순간적인 예측 오차가 누적될 시, 페이지 로딩 지연, 응답 속도 저하, 콘텐츠 전송 지연 등의 형태로 사용자 체감 품질이 저하될 가능성이 높다. 따라서, QCI3 환경에서는 단순히 평균적인 예측 정확도를 높이는 것만으로는 충분하지 않으며, 인공지능 예측값 $\hat{x}_3(t)$ 이 순간적으로 크게 변동하더라도 최종 판단이 과도하게 흔들리지 않도록 안정적인 의사결정 구조를 유지하는 것이 중요하다 할 것이다.

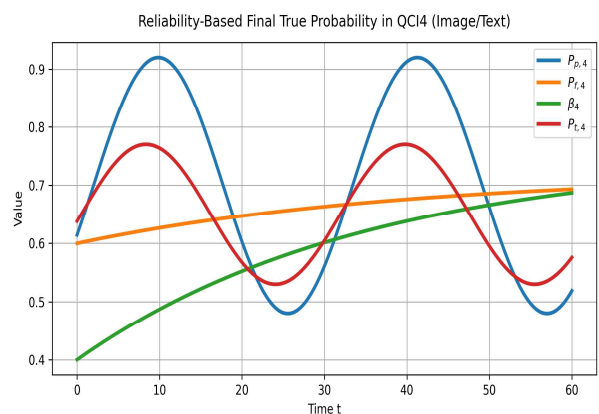


그림 4. QCI4(Image/Text 서비스) 환경에서의 예측(사실) 확률, 동적 가중치 및 최종 진실 확률의 동적 변화를 나타낸 그래프

Fig. 4. Shows the dynamic changes of predictive/factual probability, adaptive weight, and final truth probability in QCI4 (image/text service)

그림 4는 QCI4(Image/Text) 서비스 환경에서 시간 변화에 따라 예측 신뢰 확률 $P_{p,4}(t)$, 사실 신뢰 확률 $P_{f,4}(t)$, 동적 가중치 $\beta_4(t)$, 그리고 이들을 융합하여 산출한 최종 진실 확률 $P_{t,4}(t)$ 의 동적 변화를 나타낸 것이다.

그림 4를 살펴보면, 예측 신뢰 확률 $P_{p,4}(t)$ 은 시간에 따라 비교적 큰 폭의 상승과 하락을 반복하며, 중반부와 후반부에서 뚜렷한 고점과 저점을 형성하는 모습을 보인다. 반면, 사실 신뢰 확률 $P_{f,4}(t)$ 은 전체적으로 완만한 증가 추세를 유지하면서 예측 신뢰 확률 $P_{p,4}(t)$ 에 비해 상대적으로 안정적인 흐름을 나타낸다. 또한, 동적 가중치 $\beta_4(t)$ 는 초기의 낮은 값에서 출발하여 시간 경과에 따라 점진적으로 증가하는 형태를 보인다.

즉, 예측 신뢰 확률 $P_{p,4}(t)$ 이 하락하는 구간에서도 최종 진실 확률 $P_{t,4}(t)$ 은 급격히 붕괴하지 않고 비교적 안정된 범위 내에서 유지된다는 것이다. 이는 사실 신뢰 확률 $P_{f,4}(t)$ 이 보완적 역할을 수행함으로써, 인공지능 예측값 $\hat{x}_4(t)$ 의 일시적 저하가 최종 판단에 과도하게 반영되지 않도록 조정되고 있음을 보여준다. 또한, 예측 신뢰 확률 $P_{p,4}(t)$ 이 중간 구간과 후반 구간에서 상대적으로 큰 폭으로 하락하더라도, 최종 진실 확률 $P_{t,4}(t)$ 은 그보다 완만한 하락 형태를 보이며 일정 수준 이하로 급격히 낮아지지 않는다. 반대로, 예측 신뢰 확률 $P_{p,4}(t)$ 이 상승하는 구간에서는 최종 진실 확률 $P_{t,4}(t)$ 도 상승하는 경향을 보이나, 그 상승 폭은 예측 신뢰 확률 $P_{p,4}(t)$ 에 비해 상대적으로 제한적으로 나타난다. 특히, 일부 구간에서는 최종 진실 확률 $P_{t,4}(t)$ 이 예측 신뢰 확률 $P_{p,4}(t)$ 보다 높은 수준 또는 낮은 수준으로 조정 형성되는데, 이는 최종 진실 확률 $P_{t,4}(t)$ 이 인공지능 예측값 $\hat{x}_4(t)$ 을 그대로 추종하는 것이 아니라 사실 신뢰 확률 $P_{f,4}(t)$ 과의 융합을 통해 보다 안정적인 방향으로 보정되고 있음을 의미한다.

7QCI 중 QCI4는 Image/Text 서비스로, 정보의 정확한 전달과 안정적인 콘텐츠 제공이 핵심 요소로 작용한다. Image/Text 서비스는 음성이나 대화형 영상처럼 연속적인 실시간성이 절대적으로 요구되지는 않지만, 전송 상태의 불안정이나 순간적인 판단 오류가 누적될 경우 이미지 로딩 지연, 텍스트 전달 오류, 콘텐츠 응답 저하 등의 형태로 사용자 체감 품질이 저하될 가능성이 높다. 따라서, QCI4 환경에서는 단순히 평균적인 예측 정확도를 높이는 것만으로는 충분하지 않으며, 인공지능 예측값 $\hat{x}_4(t)$ 이 순간적으로 크게 변동하더라도 최종 판단이 과도하게 흔들리지 않도록 안정적인 의사결정 구조를 유지하는 것이 중요하다 할 것이다.

그림 5는 QCI5(Immersive Content: AR/VR/MR/XR) 서비스 환경에서 시간 변화에 따라 예측 신뢰 확률 $P_{p,5}(t)$, 사실 신뢰 확률 $P_{f,5}(t)$, 동적 가중치 $\beta_5(t)$, 그리고 이들을 융합하여 산출한 최종 진실 확률 $P_{t,5}(t)$ 의 동적 변화를 나타낸 것이다.

그림 5를 살펴보면, 예측 신뢰 확률 $P_{p,5}(t)$ 은 시간에 따라 비교적 큰 폭의 상승과 하락을 반복하며, 중반부와 후반부에

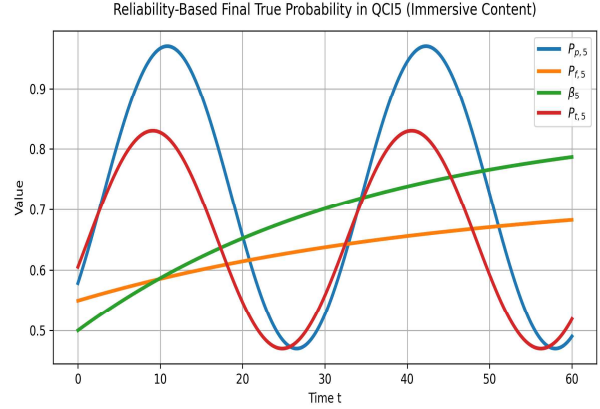


그림 5. QCI5(Immersive 서비스) 환경에서의 예측(사실) 확률, 동적 가중치 및 최종 진실 확률의 동적 변화를 나타낸 그래프

Fig. 5. Shows the dynamic changes of predictive/factual probability, adaptive weight, and final truth probability in QCI5 (immersive service)

서 뚜렷한 고점과 저점을 형성하는 모습을 보인다. 반면, 사실 신뢰 확률 $P_{f,5}(t)$ 은 예측 신뢰 확률 $P_{p,5}(t)$ 과 유사한 흐름을 나타내면서도 전체적으로는 더 완만한 범위에서 변화한다. 동적 가중치 $\beta_5(t)$ 는 초기의 낮은 값에서 출발하여 시간 경과에 따라 점진적으로 증가하는 형태를 보인다.

즉, 예측 신뢰 확률 $P_{p,5}(t)$ 이 하락하는 구간에서도 최종 진실 확률 $P_{t,5}(t)$ 은 사실 신뢰 확률 $P_{f,5}(t)$ 의 보완 작용에 의해 비교적 안정적으로 유지된다 하겠으나, 전반적인 변화 방향에서는 최종 진실 확률 $P_{t,5}(t)$ 이 예측 신뢰 확률 $P_{p,5}(t)$ 을 일정 부분 따라가는 경향도 함께 나타난다. 이에 합당한 이유로는 Immersive Content 서비스가 본질적으로 미래 품질 상태에 대한 선제적 판단을 필요로 하는 성향을 가지고 있기 때문이라 할 것이다. 특히, 중간 구간과 후반 구간에서 예측 신뢰 확률 $P_{p,5}(t)$ 이 크게 하락할 시, 최종 진실 확률 $P_{t,5}(t)$ 역시도 상대적으로 낮은 수준으로 조정되나, 그 하락 폭은 예측 신뢰 확률 $P_{p,5}(t)$ 에 비해 훨씬 완만하게 나타난다. 이는 사실 신뢰 확률 $P_{f,5}(t)$ 이 보완적 역할을 수행함으로써, 인공지능 예측값 $\hat{x}_5(t)$ 의 일시적 저하가 최종 판단에 과도하게 반영되지 않도록 조정되고 있음을 보여준다.

또한, 예측 신뢰 확률 $P_{p,5}(t)$ 이 상승하는 구간에서는 최종 진실 확률 $P_{t,5}(t)$ 도 이에 동조하여 상승하는 경향을 보이나, 그 상승 폭은 예측 신뢰 확률 $P_{p,5}(t)$ 에 비해 제한적이다 하겠다. 일부 구간에서는 최종 진실 확률 $P_{t,5}(t)$ 이 예측 신뢰 확률 $P_{p,5}(t)$ 의 흐름을 반영하면서도 보다 부드럽고 완만한 곡선 형태로 형성되는데, 이는 최종 진실 확률 $P_{t,5}(t)$ 이 인공지능 예측값 $\hat{x}_5(t)$ 을 그대로 추종하는 것이 아니라 사실 신뢰 확률 $P_{f,5}(t)$ 과의 융합을 통해 보다 안정적인 방향으로 보정되고 있음을 의미한다.

7QCI 중 QCI5는 Immersive Content 서비스로, 고해상

도 실감형 콘텐츠, 몰입형 미디어, 대용량 멀티미디어 데이터 전송과 같이 연속적인 품질 유지와 안정적인 사용자 체감 보장이 핵심 요소로 작용한다. Immersive Content 서비스는 순간적인 전송 상태 변화나 예측 오차가 발생할 경우 화면 품질 저하, 몰입감 손실, 응답 지연 등의 형태로 사용자 체감 품질에 직접적인 영향을 줄 가능성이 높다. 따라서, QCI5 환경에서는 단순히 평균적인 예측 정확도를 높이는 것만으로는 충분하지 않으며, 인공지능 예측값 $\hat{x}_5(t)$ 이 순간적으로 크게 변동하더라도 최종 판단이 과도하게 흔들리지 않도록 안정적인 의사결정 구조를 유지하는 것이 중요하다고 할 것이다.

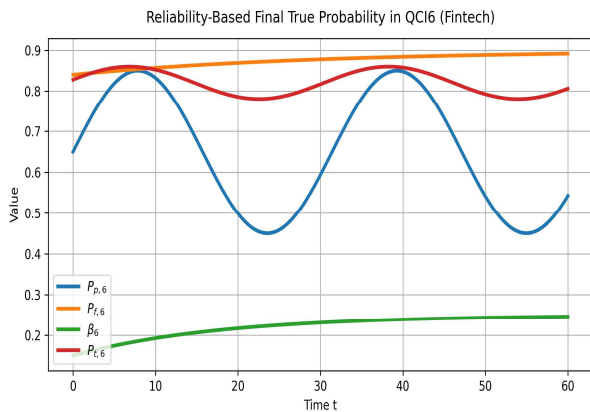


그림 6. QCI6(Fintech 서비스) 환경에서의 예측(사실) 확률, 동적 가중치 및 최종 진실 확률의 동적 변화를 나타낸 그래프

Fig. 6. Shows the dynamic changes of predictive/factual probability, adaptive weight, and final truth probability in QCI6 (fintech service)

그림 6을 살펴보면, 예측 신뢰 확률 $P_{p,6}(t)$ 은 시간에 따라 비교적 큰 폭의 상승과 하락을 반복하며, 초반부와 중반부에서 뚜렷한 고점을 형성한 뒤 그 사이 구간과 후반부에서 낮은 저점을 보이는 형태를 나타낸다. 반면, 사실 신뢰 확률 $P_{f,6}(t)$ 은 전체적으로 높은 수준에서 완만한 진동을 유지하며, 예측 신뢰 확률 $P_{p,6}(t)$ 에 비해 상대적으로 안정적인 흐름을 보인다. 또한, 동적 가중치 $\beta_6(t)$ 는 매우 낮은 값에서 시작하여 시간 경과에 따라 점진적으로 증가하는 형태를 나타내며, 최종 진실 확률 $P_{t,6}(t)$ 은 전 구간에 걸쳐 높은 수준을 유지하면서도 완만한 상승 추세를 보인다.

즉, 예측 신뢰 확률 $P_{p,6}(t)$ 이 하락하는 구간에서도 최종 진실 확률 $P_{t,6}(t)$ 은 사실 신뢰 확률 $P_{f,6}(t)$ 의 보완 작용에 의해 안정적으로 유지된다는 점을 확인할 수 있다. 특히, 중반부와 후반부에서 예측 신뢰 확률 $P_{p,5}(t)$ 은 비교적 큰 폭으로 낮아지지만, 최종 진실 확률 $P_{t,6}(t)$ 은 이에 따라 급격히 하락하지 않고 상위 수준에서 거의 일정하게 유지된다. 이는 사실 신뢰 확률 $P_{f,6}(t)$ 이 전체 구간에서 높은 수준을 유지하고 있기 때

문에 인공지능 예측값 $\hat{x}_6(t)$ 의 일시적 저하가 최종 판단에 과도하게 반영되지 않도록 보장하고 있음을 보여준다.

또한, 예측 신뢰 확률 $P_{p,5}(t)$ 이 상승하는 구간에서는 최종 진실 확률 $P_{t,6}(t)$ 도 상승 방향을 유지하지만, 그 변화 폭은 예측 신뢰 확률 $P_{p,5}(t)$ 에 비해 매우 제한적으로 나타난다. 이는 최종 진실 확률 $P_{t,6}(t)$ 이 예측 신뢰 확률 $P_{p,6}(t)$ 의 변화 방향을 일부 반영하면서도 이를 그대로 따라가지 않고 사실 신뢰 확률 $P_{f,6}(t)$ 과의 융합을 통해 보다 완만하고 안정적인 값으로 조정되고 있음을 의미한다. 또한, QCI6가 그림 1의 QCI1와 그래프 성향이 유사하게 나타나는 이유는 두 서비스가 모두 순간적인 예측 변동보다 안정적이고 신뢰성 있는 최종 판단 유지가 더 중요한 특성을 가지기 때문이다. QCI1은 음성 서비스로서 저지연성과 연속성이 핵심이며, QCI6는 Fintech 서비스로서 정확성과 보수적 안정성이 핵심이다. 서비스 성격은 상이하지만, 두 서비스 모두 인공지능 예측값 $\hat{x}_i(t)$ 의 일시적 흔들림이 최종 제어 판단에 과도하게 반영될 시, 품질 저하 또는 서비스 신뢰 손상으로 직결될 수 있기 때문이라 하겠다.

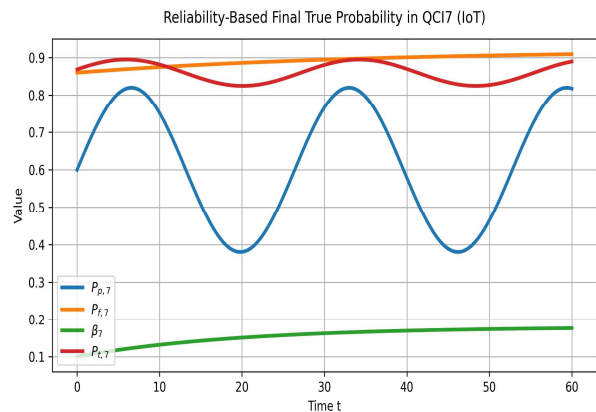


그림 7. QCI7(IoT 서비스) 환경에서의 예측(사실) 확률, 동적 가중치 및 최종 진실 확률의 동적 변화를 나타낸 그래프

Fig. 7. Shows the dynamic changes of predictive/factual probability, adaptive weight, and final truth probability in QCI7 (IoT service)

그림 7을 살펴보면, 예측 신뢰 확률 $P_{p,7}(t)$ 은 시간에 따라 비교적 큰 폭의 상승과 하락을 반복하며, 초반부·중반부·후반부에서 뚜렷한 고점과 저점을 형성하는 주기적 변동 패턴을 보인다. 반면, 사실 신뢰 확률 $P_{f,7}(t)$ 은 전체적으로 높은 수준에서 비교적 완만한 진동을 유지하며, 예측 신뢰 확률 $P_{p,7}(t)$ 에 비해 상대적으로 안정적인 흐름을 나타낸다. 또한, 동적 가중치 $\beta_6(t)$ 는 매우 낮은 값에서 시작하여 시간 경과에 따라 점진적으로 증가하는 형태를 보이며, 최종 진실 확률 $P_{t,7}(t)$ 은 전체 구간에서 높은 수준을 유지하면서도 완만한 상승 추세를 나타낸다.

즉, 예측 신뢰 확률 $P_{p,7}(t)$ 이 하락하는 구간에서도 최종 진

실 확률 $P_{i,7}(t)$ 은 사실 신뢰 확률 $P_{f,7}(t)$ 의 보완 작용에 의해 비교적 안정적으로 유지된다는 점을 확인할 수 있다. 특히 중반부와 후반부의 저점 구간에서 예측 신뢰 확률 $P_{p,7}(t)$ 은 큰 폭으로 낮아지지만, 최종 진실 확률 $P_{i,7}(t)$ 은 이에 비례하여 급격히 하락하지 않고 상위 수준에서 안정적으로 유지된다. 이는 사실 신뢰 확률 $P_{f,7}(t)$ 이 전체적으로 높은 수준에서 비교적 안정된 흐름을 유지하고 있기 때문에 인공지능 예측값 $\hat{x}_7(t)$ 의 일시적 저하가 최종 판단에 과도하게 반영되지 않도록 보정하고 있음을 보여준다.

또한, 예측 신뢰 확률 $P_{p,7}(t)$ 이 상승하는 구간에서는 최종 진실 확률 $P_{i,7}(t)$ 도 상승 방향을 유지하지만, 그 변화 폭은 예측 신뢰 확률 $P_{p,7}(t)$ 에 비해 매우 제한적으로 나타난다. 이는 최종 진실 확률 $P_{i,7}(t)$ 이 예측 신뢰 확률 $P_{p,7}(t)$ 의 변화 방향을 일정 부분 반영하면서도 그대로 따라가지 않고 사실 신뢰 확률 $P_{f,7}(t)$ 과의 융합을 통해 보다 완만하고 안정적인 값으로 조정되고 있음을 의미한다. 이는 최종 판단이 단순한 예측치 중심이 아니라 안정성, 지속성, 신뢰성을 우선하도록 설계되어 있음을 시사한다.

7QCI 중 QCI7은 IoT 서비스로, 다수의 센서와 단말이 지속적으로 정보를 송수신하는 환경에서 연결의 지속성, 데이터 전달의 안정성, 그리고 시스템 전체의 신뢰성 유지가 매우 중요한 서비스 특성을 가진다. IoT 환경에서는 개별 단말의 상태 변화나 순간적인 예측 오차가 빈번하게 발생할 수 있으나, 이러한 단기적 변동이 곧바로 전체 제어 판단의 급격한 흔들림으로 이어질 경우 오동작, 잘못된 상태 인식, 제어 불안정 등의 문제가 발생할 가능성이 크다. 따라서, QCI7 환경에서는 단순히 예측값의 민감한 반응성을 높이는 것보다, 예측 신뢰 확률 $P_{p,7}(t)$ 이 일시적으로 크게 변동하더라도 최종 판단은 안정적으로 유지되도록 하는 의사결정이 더욱 중요하다 할 것이다.

V. 결 론

본 연구논문에서는 7QCI 정보 제공 환경상에서 허위정보 유입으로 인해 발생할 수 있는 인공지능 의사결정 오인 문제를 개선하기 위하여 진실 확률 추정 기반의 인공지능 의사결정 최적화 알고리즘을 제안하였다. 본 연구논문에서 제안된 알고리즘은 데이터 수집 계층, 데이터 예측 계층, 실시간 관측 및 분석 계층 및 의사결정 계층으로 구성되며, 인공지능 예측값과 사실 정보 간의 차이를 정량적으로 분석하여 예측 신뢰 확률과 사실 신뢰 확률을 각각 산출한 후, 이를 동적 가중치 기반으로 융합시킨 최종 진실 확률이 추출되도록 설계하였다. 이를 통해, 단일 예측값에만 의존하였던 기존 방식의 한계를 보완하고, 허위정보 또는 실시간 분포 변동 상황에서도 보다 안정적이고 신뢰성 있는 의사결정이 가능하도록 하였다.

모의실험에서는 QCI1(Voice)부터 QCI7(IoT)까지 각기 다

른 서비스 특성을 갖는 7QCI 환경을 대상으로 제안 알고리즘의 동작 특성을 분석하였다. 그 결과, 예측 신뢰 확률은 시간에 따라 비교적 큰 폭의 변동을 보이는 반면, 최종 진실 확률은 사실 신뢰 확률의 보완 작용과 동적 가중치의 조정 효과에 의해 보다 안정적인 범위 내에서 유지됨을 확인할 수 있었다. 더욱이, QCI1, QCI6, QCI7과 같이 안정성과 신뢰성이 핵심인 ICT 환경에서는 최종 진실 확률이 상위 수준에서 보수적으로 유지되는 특성이 두드러졌으며, QCI2, QCI3, QCI4, QCI5와 같이 연속적 품질 유지와 사용자 체감 품질이 중요한 ICT 환경에서는 예측 신뢰 확률의 변화 방향을 일정 부분 반영하되 그 진폭은 완화하는 형태로 나타났다. 이는 7QCI 유형별 특성을 반영하면서도 공통적으로 허위정보 오인 완화와 의사결정 안정성 확보에 기여할 수 있음을 보여준다 할 것이다.

또한, 본 연구논문에서 제안한 진실 확률 추정 기반 인공지능 의사결정 최적화 알고리즘은 예측 중심 판단 구조가 가지는 취약성을 개선하고, 사실 정보와의 융합을 통해 최종 판단의 신뢰성과 안정성을 향상시킬 수 있음을 확인하였다. 이는 7QCI 정보 제공 환경과 같이 다양한 QoS 특성과 동적 네트워크 조건이 공존하는 ICT 환경에서, 허위정보 유입에 따른 오판 가능성을 줄이고 보다 합리적이며 적응적인 인공지능 의사결정을 수행할 수 있게 되었음을 내포한다고 하겠다.

향후 연구에서는 생성형 AI와 휴먼 로봇을 포함시킨 9QCI 별 QoS 서비스 차원의 민감도를 보다 정교하게 반영한 임계값 및 가중치를 대상으로 보다 확장시킨 최적화 기법을 추가적으로 연구할 계획이다. 또한, 이에 따른 보다 실용적이고 범용적인 의사결정 최적화 모델도 좀 더 구체적으로 구현해 볼 수 있는 것이라 예측되는 바 라 할 것이다.

참고문헌

- [1] Y. LeCun, Y. Bengio, and G. Hinton, "Deep Learning," *Nature*, Vol. 521, No. 7553, pp. 436-444, May 2015. <https://doi.org/10.1038/nature14539>
- [2] K. Shu, A. Sliva, S. Wang, J. Tang, and H. Liu, "Fake News Detection on Social Media: A Data Mining Perspective," *ACM SIGKDD Explorations Newsletter*, Vol. 19, No. 1, pp. 22-36, 2017. <https://doi.org/10.1145/3137597.3137600>
- [3] A. Adadi and M. Berrada, "Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI)," *IEEE Access*, Vol. 6, pp. 52138-52160, 2018. <https://doi.org/10.1109/ACCESS.2018.2870052>
- [4] Y. Li, J. Gao, C. Meng, Q. Li, L. Su, B. Zhao, ... and J. Han, "A Survey on Truth Discovery," *ACM SIGKDD Explorations Newsletter*, Vol. 17, No. 2, pp. 1-16, 2015. <https://doi.org/10.1145/2897350.2897352>
- [5] J. Gawlikowski, C. R. N. Tassi, M. Ali, J. Lee, M. Humt, J. Feng, ... and X. X. Zhu, "A Survey of Uncertainty in Deep

Neural Networks,” *Artificial Intelligence Review*, Vol. 56, No. Suppl 1, pp. 1513-1589, 2023. <https://doi.org/10.1007/s10462-023-10562-9>

- [6] H. Wang and D.-Y. Yeung, “A Survey on Bayesian Deep Learning,” *ACM Computing Surveys*, Vol. 53, No. 5, 108, pp. 1-37, 2020. <https://doi.org/10.1145/3409383>
- [7] M. Ghavamzadeh, S. Mannor, J. Pineau, and A. Tamar, “Bayesian Reinforcement Learning: A Survey,” *Foundations and Trends in Machine Learning*, Vol. 8, No. 5-6, pp. 359-483, 2015. <https://doi.org/10.1561/22000000049>
- [8] P. Ameigeiras, J. Navarro-Ortiz, P. Andres-Maldonado, J. M. Lopez-Soler, J. Lorca, Q. Perez-Tarrero, and R. Garcia-Perez, “3GPP QoS-based Scheduling Framework for LTE,” *EURASIP Journal on Wireless Communications and Networking*, Vol. 2016, 78, 2016. <https://doi.org/10.1186/s13638-016-0565-9>
- [9] J. Gama, I. Žliobaitė, A. Bifet, M. Pechenizkiy, and A. Bouchachia, “A Survey on Concept Drift Adaptation,” *ACM Computing Surveys*, Vol. 46, No. 4, 44, pp. 1-37, March 2014. <https://doi.org/10.1145/2523813>
- [10] S. Viaene and G. Dedene, “Cost-Sensitive Learning and Decision Making Revisited,” *European Journal of Operational Research*, Vol. 166, No. 1, pp. 212-220, 2005. <https://doi.org/10.1016/j.ejor.2004.03.031>
- [11] Q. Li, Y. Li, J. Gao, L. Su, B. Zhao, M. Demirbas, ... and J. Han, “A Confidence-Aware Approach for Truth Discovery on Long-Tail Data,” *Proceedings of the VLDB Endowment*, Vol. 8, No. 4, pp. 425-436, December 2014. <https://doi.org/10.14778/2735496.2735505>



서성현(Sung-Hyun Seo)

2003년 : 성균관대학교 (공학석사-전기전자컴퓨터공학과)

2020년 : 한성대학교 (공학박사-스마트융합학부)

2014년~2023년: 특허법인대한, 휴피아특허사무소, 이사

2013년~현 재: ㈜비즈팻, 대표

2021년~현 재: 한국외국어대학교, 미네르바대학, 외래교수

2023년~현 재: 선문대학교, 전자공학과, 외래교수

2024년~현 재: 호서대학교, 빅데이터AI학부, 초빙교수

2024년~현 재: 평택대학교, 융합소프트웨어학과, 외래교수

2025년~현 재: ㈜4D PnC, 책임연구원

2026년~현 재: 한세대학교 IT학부, 컴퓨터공학전공, 외래교수

※ 관심분야 : 창의공학설계, SW컴퓨팅, AI System, BigData, ICT융합플랫폼, 신호처리및정보보안, 지식재산권 등