

행정-정책 의사결정에서 머신러닝(machine learning) 방법론 도입의 정책적 함의: 기계의 한계와 증거기반 의사결정(evidence-based decision-making)*

김 병 조**
은 종 환***

머신러닝과 딥러닝 등 인공지능 기술의 급속한 발전은 행정-정책 분야에도 영향을 확대하고 있다. 이 연구는 전통적 행정-정책 의사결정 모형인 합리적 모형의 이론적 이상과 현실적 한계에 대한 검토를 바탕으로 최근 연구가 활발히 진행중인 머신러닝 기반 의사결정 사례를 행정-정책 의사결정 관점에서 분석함으로써 인공지능 기술의 행정-정책 분야 도입이 가져올 변화를 논의하고 다음과 같은 두 가지 함의를 도출하였다. 첫째, 분석 대상과 목표가 사전에 잘 수립된 구조화된 문제(structured problems)에서는 기계가 인간의 의사결정에 비해 훨씬 나은 성과를 보이는 것으로 나타나 합리적 의사결정의 구현 가능성이 높아지고 있다. 둘째, 미래를 위한 전략적 의사결정이나 지금까지 경험하지 못한 새로운 환경 변화에 대한 대처 등 기계가 학습할 데이터가 부족하거나 목표가 모호한 비구조화된 문제(unstructured problems)의 경우 인간의 메타(meta) 의사결정이 여전히 중요하다. 따라서 머신러닝 등 인공지능 기술을 행정-정책분야에 성공적으로 도입하려면 다음과 같은 세 가지 사항을 고려해야 한다. 첫째, 해당 분야가 머신러닝으로 구현해도 될 구조화된 문제 영역인지 혹은 인간의 판단이 필수적인 메타(meta) 의사결정 영역인지에 대한 구분이 선행해야 한다. 둘째, 구조화된 문제 영역으로 판단될 경우에도 어떠한 행정-정책 기준(예: 책임성, 효과성, 효율성)으로 알고리즘을 구현할 것인지에 대한 인간의 메타 의사결정이 필요하다. 셋째, 메타 의사결정의 질을 향상하기 위해 의사결정자는 일화적 경험(anecdotal experience)에 의존하는 대신 체계적(systematic)인 의사결정을 내릴 수 있도록 증거기반 의사결정(evidence-based decision-making)을 이해하고 실천해야 한다.

[주제어: 머신러닝, 증거기반 의사결정, 합리적 의사결정]

* 논문에 유익한 심사평을 주신 심사자분들께 감사드립니다.

이 연구는 중소벤처기업 지원정책 효율화를 위한 방안 연구에 대한 자문 목적으로 연구된 내용을 토대로 하여 내용을 추가 발전시킨 것임을 밝힙니다.

** 주저자

*** 교신저자

I. 서론

4차 산업혁명의 도래와 인공지능(Artificial Intelligence) 기술의 발전은 인류의 기회이자 위협으로 속도를 높이며 우리에게 다가오고 있다(Davenport & Kirby, 2015). 인공지능 기술은 사회전체를 창조적으로 파괴하며 정부부문에 일하는 방식의 변화를 가져올 것으로 기대된다(윤상오, 이은미, & 성욱준, 2018; 황종성, 2017). 이에 따라 정부는 지난 2017년 4차 산업혁명 기술의 발전을 전략적으로 수용하기 위해 지능형 정부(intelligent government) 추진 전략을 발표하였다(행정안전부, 2017). 지능형 정부 추진에 있어 핵심적 역할을 하는 것은 인공지능과 관련된 기술이다. 머신러닝 등 인공지능 기술을 통한 행정-정책 의사결정 체계를 지향하는 지능형 정부 개념은 기존 정책 의사결정의 단점이라고 할 수 있는 불확실성, 오류 가능성, 시의성 결여, 부정 부패 가능성, 편견과 오류 가능성 등 문제점을 해결할 수 있는 대안이 될 수 있다(윤상오, 2017). 이러한 낙관적 전망의 배경에는 그간 비록 이상적이지만 현실적으로 불가능하다고 여겨져 왔던 합리적 정책결정 모형을 인공지능 기반 의사결정을 통해 실현할 수 있을 것이라는 기대가 깔려 있다(황종성, 2017). 이 연구는 머신러닝에 기반한 의사결정이 행정-정책 의사결정 과정에 제시하는 이론적 실천적 시사점을 탐구한다.

인공지능과 정부 역할의 관계에 관한 주요 연구 현황은 다음과 같다. 먼저 현재까지 주된 연구들은 인공지능의 발전에 대한 정부의 대응을 탐색하고 인공지능이 가져올 정부 변화에 대한 거시적 전망을 제시하는 시론적 연구다(Eggers, Schatsky, & Viechnicki, 2017; 성욱준 & 황성수, 2017; 윤상오 et al., 2018; 진상기 & 박영원, 2017; 황종성, 2017). 또한 인공지능 기술을 정부 운영에 적용할 경우 어떠한 변화가 발생할 것인지를 탐구하는 사례연구도 등장하고 있다(Ojo, Mellouli, & Ahmadi Zeleti, 2019; 이강윤 & 김준혁, 2016). 그러나 행정학 연구에서 머신러닝 등 인공지능 기술 도입에 관한 이론적 논의와 실증 연구가 초기 단계인 만큼 다음과 같은 점에 대해서는 여전히 더 많은 연구가 요구되고 있다.

첫째, 인공지능이 행정-정책 의사결정에 도입될 경우 나타나게 될 변화에 대한 분석이 부족하다. 특히 사례를 바탕으로 행정-정책 의사결정 방식의 변화를 구체적으로 탐구하는 실증 및 사례 연구는 아직 충분하지 않다. 종래의 연구들은 거시적 차원에서 변화의 방향을 논의하거나(윤상오 et al., 2018) 인공지능의 적용 사례에 한정하여 연구하였다(이강윤 & 김준혁, 2016). 이러한 연구는 인공지능의 적용으로 인해 기존의 정부의 일하는 방식의 어떤 구성요소가 변화하는지, 인공지능과 정부 주체의 상호작용이 어떤 결과를 가져오는지 등에 대한 구체적인 상을 제시하지 못한다는 한계가 있다.

둘째, 인공지능의 발전이 가져오는 정부 의사결정 방식의 변화를 이론적 관점에서 탐구한 연구가 부족하다. 이론은 현실을 보편적으로 설명하는 준거틀을 제시한다는 점을 고려할 때 인공지능이 가져오는 정부의 일하는 방식의 변화도 이론의 관점에서 평가해볼 수 있을 것이다. 특히 행정-정책 의사결정은 다양한 사회적 가치들을 종합적으로 반영하기 때문에 복잡한 환경하에

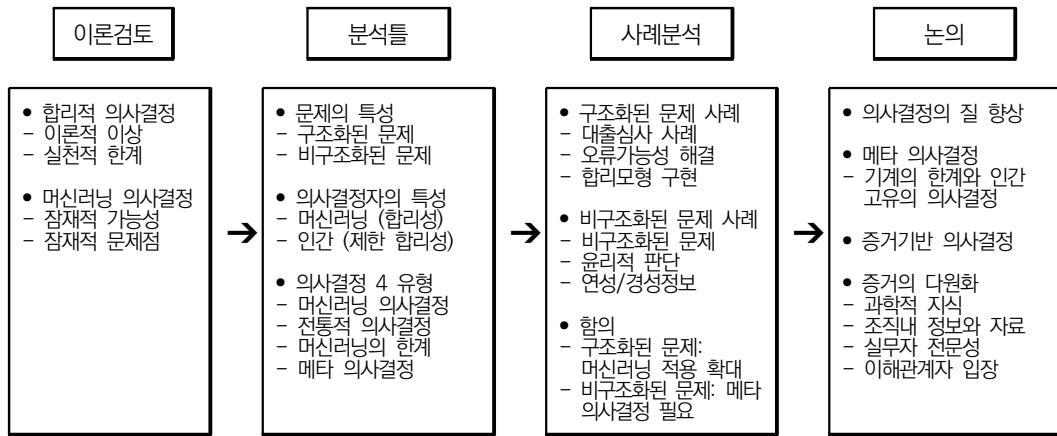
서 의사결정이 이루어지게 된다(Kitzhaber, 1993). 따라서 정책 의사결정 관점에서 인공지능 기술의 도입이 초래할 결과를 탐구하는 것은 시의 적절하며 나아가 인공지능 기술에 대한 행정학의 대응 방향에 대한 실천적 시사점을 제공할 수 있다.

이 연구의 목적은 머신러닝 등 인공지능 기술의 도입이 행정-정책 의사결정에 제시하는 정책적 함의를 전통적인 합리적 의사결정 모형과 머신러닝 기반 의사결정의 비교를 통해 탐구하는 것이다. 합리모형은 이론적으로는 이상적 모형으로 제시되어 왔지만 실천적으로는 인간의 제한된 합리성(bounded rationality)과 정책 의사결정 환경의 한계로 인해 현실에서 완전히 구현되지 못했다(정정길, 최종원, 이시원, 정준금, & 정광호, 2003). 한편 머신러닝 기반 의사결정 기법은 인간의 제한된 합리성을 넘어서 가용한 모든 정보를 분석하여 최적의 해를 도출할 수 있다는 점에서 합리모형의 이상을 현실에서 구현할 가능성을 보여주고 있다. 따라서 이 연구는 머신러닝 기법을 행정-정책 의사결정에 활용한 연구 사례들을 검토함으로써 인공지능 기술의 도입이 기존의 합리적 정책의사결정 모형의 한계를 어떻게 보완할 수 있는지 그리고 인간의 판단이 여전히 필수적인 영역은 어디인지를 탐구한다.

이 연구의 이론적 실천적 함의는 다음과 같다. 첫째, 기존의 합리모형에 대한 비판이 머신러닝과 같은 인공지능 기술의 발전에 힘입어 어떻게 극복될 수 있는지 제시한다. 이는 인공지능에 대한 막연한 낙관론 또는 비판론을 넘어 다가오는 변화의 흐름을 이론적 상황에서 구체적으로 이해하는데 기여할 것이다. 둘째, 인공지능 기술의 일종인 머신러닝 기법을 활용한 행정-정책 의사결정의 가능성과 한계를 실제 연구 사례를 통해 검토함으로써 정책 의사결정 주체로서의 인간의 역할에 대한 시사점을 제시한다. 특히 정치, 경제, 사회적 상황이 복잡하게 얽혀 이루어지는 행정-정책 의사결정에서 머신러닝을 통한 자동화된 의사결정의 한계를 지적하고 인간의 역할인 메타(meta)적 의사결정의 중요성을 밝힘으로써 인공지능 시대에서 정부 역할에 대한 시사점을 도출한다. 셋째, 인간의 메타 의사결정의 퀄리티를 향상하기 위한 구체적 실천방안으로서 증거기반 의사결정(evidence-based decision-making)과 증거의 다원화의 중요성을 제시한다.

이 연구의 전개는 다음과 같다 (<그림 1> 참조). 먼저 이론 검토 단계에서는 합리적 의사결정 모형에 바탕한 행정-정책 의사결정의 이론적 이상과 실천적 한계를 개괄하고 머신러닝 기반 의사결정의 잠재적 가능성과 문제점을 검토한다. 분석틀 구축 단계에서는 문제의 특성과 의사결정자의 특성이라는 두가지 중요한 요소를 활용해 의사결정 유형을 4가지로 구분한다. 사례분석 단계에서는 앞서 도출한 문제의 특성(구조화된 문제 혹은 비구조화된 문제)에 따라 머신러닝 기반 의사결정 기법의 적용 사례들을 심층 분석하고 이를 통해 행정-정책 의사결정 영역 가운데 머신러닝 기반 의사결정의 이점을 활용할 수 있는 영역을 식별해낸다. 마지막으로 논의 단계에서는 머신러닝 기반 의사결정의 한계에 대한 인식을 바탕으로 인간 고유의 의사결정 영역인 메타(meta) 의사결정의 중요성을 확인하고, 이를 실행하기 위한 구체적 방안으로서 증거기반 의사결정과 증거의 다원화 방안을 논의한다.

〈그림 1〉 분석 흐름도
 〈Figure 1〉 Analytical Flowchart



II. 이론 검토 및 분석틀 도출

1. 합리적 행정-정책 의사결정 모형의 한계

합리적 정책 의사결정 연구에서 합리성은 좁은 의미의 합리성, 즉 합리적 선택(rational choice)에 초점을 맞춘 경제적 합리성 내지는 도구적 합리성의 의미로 사용된다(최종원, 1995). 라서 합리적 정책의사 결정은 다음과 같은 조건을 완벽하게 만족할 때 성공적으로 이루어진다(Goodwin, Wright, & Phillips, 2004; 정정길 et al., 2003). 첫째, 정책 의사결정을 통해 달성하려는 목표가 명확하게 정의되어야 한다. 둘째, 목표를 달성할 수 있는 모든 가능성 있는 대안들을 탐색하여야 한다. 셋째, 대안들을 실행하였을 때 나타나는 결과를 완벽히 예측할 수 있어야 한다. 넷째, 각 대안들의 결과를 비교 평가할 수 있어야 하며 이를 바탕으로 문제 해결 혹은 목표 달성을 위해 최선의 대안을 선택할 수 있어야 한다.

합리적 의사결정 모형을 이상적 모형이라고 보는 이유는 다음과 같은 세 가지 전제들이 충족될 때 비로서 완벽한 합리적 의사결정이 가능하기 때문이다. 첫번째 전제는 시간과 공간을 초월한 완벽한 정보를 수집할 수 있어야 한다는 것이다. 대안들을 완벽히 탐색하고 대안 실행의 결과로 발생하는 결과를 예측하기 위해서는 정책의 시행에 따른 변화 및 정책 대상 요소들의 상호작용을 모두 계산해 낼 수 있어야 하며, 이러한 예측은 완벽한 정보로부터 출발할 수 있기 때문이다. 두번째 전제는 주관이 개입되지 않는 객관적 정보의 확보다. 합리적 의사결정을 위해서는 최선의 선택을 하게끔 하는 완벽한 분석이 전제되어야 한다. 완벽한 분석이 가능하기 위한 정보는 객관적이고 하나의 해석으로만 가능한 정보가 필요하다(De Martino, Kumaran, Seymour, & Dolan, 2006). 다양하게 해석될 수 있는 정보는 최선의 합리적 선택을 불가능하게 하기 때문이

다. 세번째 전제는 의사결정의 좋고 나쁨을 판단할 수 있는 대 전제가 되는 하나의 일원화된 가치체계가 선행해야 한다는 것이다. 합리모형은 비용과 편익을 고려할 때 가장 효율적인 대안이 최선의 대안이라는 공리주의적 가정에 바탕한다. 모든 사람의 효용을 동일하게 바라보는 공리주의는 사회 경제적 조건이나 계층에 관계없이 동일한 가중치를 가지고 있다고 전제하고 있기 때문에 최대 다수의 최대 행복이 유일한 정책 의사결정의 가치기준임을 전제하는 것이다 (Gruber, 2005; Weimer & Vining, 2017).

이와 같은 합리모형의 강한 가정들(완전한 정보, 객관적 자료, 단일화된 가치체계)은 이상적 원형(정정길 et al., 2003)일 뿐 현실의 정책 의사결정을 제대로 설명하기 어렵다는 비판이 지속적으로 제기 되었다. 즉, 현실의 세계에서는 미래를 정확하게 예측하기 어려울 뿐만 아니라 상황이 변한다. 과거의 정보 또한 선택적으로 수집되기 때문에 완전한 정보를 구현하기 어렵다. 또한 인간의 인지능력의 한계로 말미암아 일정수준이상의 데이터를 동시에 분석하기도 어렵다. 설명 정보를 완벽히 수집 분석 가능하다고 하더라도 현실에서 수집되는 자료들이 모두 객관적인 것 인지지에 대해 논란이 존재할 수 있으며 합리성에 기반한 의사결정이 정당한 것인지에 대한 논란 또한 존재한다.

이에 대한 대응으로 합리모형에 대한 대안적 의사결정모형이 여러 형태로 등장하였다. 대안 모형들은 합리모형에 대한 세가지 전제들에 대한 비판을 바탕으로 새로운 관점을 제공한다. 만족모형은 인간의 한계로 말미암아 완벽한 정보를 수집 분석 못하기에 제한된 합리성 내에서 절차적 합리성을 구현해야 한다는 입장이다. 만족모형은 객관적 정보와 일원화된 가치체계에 대해서는 명시적 비판을 하지 않았다(March & Simon, 1993). 점증주의 모형은 합리모형의 세가지 전제에 대해 모두 회의적인 입장이다(Lindblom, 1959). 개인이 아닌 집단에서 이루어지는 의사결정을 설명하는 이론도 발전되었다. 사이먼의 제한적 합리성을 계승하며 의사결정자가 다수일 때의 현상을 설명하는 회사모형(Cyert & March, 1963)도 발전되었고, 조직의 의사결정의 비합리성과 우연성을 설명하기 위한 쓰레기통 모형도 소개되었다.

합리모형에 대한 비판으로 등장한 다양한 정책 의사결정 모형을 위의 세 가지 요소에 대한 입장을 중심으로 구분하면 다음과 같다. 인간의 합리성 추구는 지향하지만 인간의 정서적 물리적 한계로 말미암아 제한적 합리성을 지지하는 만족모형(March & Simon, 1993)이 대안적 의사결정모형으로 제시된 바 있다. 점증주의(incrementalism)은 현실의 정책 과정이 하나의 의사결정주체가 아닌 다양한 의사결정주체간 협상과 합의에 기반하는 다원적 정치체제내에서 이루어지며 정책 또한 단절적이 아닌 연속적으로 이루어지는 현상을 바탕으로 대안적 의사결정모형으로 제안되었다(Lindblom, 1959). 이밖에도 제한적 합리성 모형에 바탕을 두고 다양한 정책결정 주체를 참여시키는 회사모형(Cyert & March, 1963), 합리성 자체에 대한 회의적 시각에서 구성된 쓰레기통모형(Cohen, March, & Olsen, 1972) 등이 합리모형에 대한 대안적 의사결정모형으로 제시된 바 있다(정정길 et al., 2003). 합리모형과 이에 대한 비판으로 등장한 다양한 의사결정 모형을 합리모형의 가정과 의사결정의 주요한 특성을 정리하면 아래 <표 1>과 같다.

〈표 1〉 합리모형과 비판적 정책의사결정 모형
 〈Table 1〉 A Comparison of Policy Decision-Making Models

| 구 분 | 합리모형 | 만족모형 | 점증주의 | 회사모형 | 쓰레기통모형 |
|--------|------|------|------|------|--------|
| 정보의 취득 | 완전 | 제한 | 제한 | 제한 | 제한 |
| 정보의 성격 | 객관 | 객관 | 혼합 | 객관 | 혼합 |
| 가치체계 | 단일 | 단일 | 다양 | 단일 | 다양 |
| 의사결정자 | 단일 | 단일 | 다수 | 다수 | 다수 |
| 합리성 | 옹호 | 옹호 | 비판 | 옹호 | 비판 |
| 다양한 견해 | 배제 | 배제 | 반영 | 반영 | 반영 |

2. 머신러닝 기반 행정-정책 의사결정의 가능성

전통적인 합리적 의사결정 모형의 실천적 한계에 대한 인식과 더불어 최근 머신러닝 등 인공지능 기술 발전이 행정-정책 관련 의사결정 방식을 변화시킬 수 있다는 전망이 제시되고 있다(황종성, 2017). 머신러닝 기술은 광범위한 데이터를 바탕으로 최적화된 의사결정을 가능하게 한다. 인공지능 기술의 도입은 정보통신기술에 기반하여 정책 의사결정을 보조적으로 돕던 기존의 전자정부 개념에서 한발 더 나아가 행정-정책 의사결정에서 머신러닝 기반 의사결정 기법의 도입과 같은 더욱 적극적인 역할에 대한 가능성과 기대를 높이고 있다(황종성, 2017). 아래에서는 행정-정책 의사결정에서 머신러닝 기술이 가져올 가능성과 잠재적 문제점을 검토한다.

인공지능 연구의 선구자 중 하나인 Herbert Simon(1956)은 지난 1950년대에 이미 “앞으로 20년 안에 기계는 인간이 할 수 있는 어떤 종류의 일도 해 낼 수 있게 될 것이다”고 예측했다. 또 다른 선구적 연구자인 Marvin Minsky(1967)도 “앞으로 한 세대 안에 인공지능을 창조하는 문제는 상당한 수준으로 해결될 것이다”고 예측했다. 이들의 예측과 달리 21세기가 된 현재에도 인공지능과 머신러닝은 인간의 능력을 100% 대체하지는 못했다. 하지만 머신러닝과 인공지능의 발달 속도는 기하급수적이어서 조만간 인간과 기계의 역량이 역전되는 상황이 발생하지 않으리라는 보장은 누구도 하지 못한다.

머신러닝에 대한 관심이 공공 및 민간 분야 전반으로 확산되면서 최근 들어 다양한 분야에서 머신러닝의 원리를 적용하려는 시도가 늘어나고 있다. 그러한 관심 가운데 하나가 머신러닝의 판단 및 예측 능력의 정확성이다. 기계가 학습할 수 있는 충분한 양의 정보를 제공하고 난 뒤 머신러닝의 판단이 인간의 그것보다 정확히 결과를 예측할 수 있다면 이는 인간의 영역에 도전하는 또 하나의 큰 물줄기가 될 수 있다. 특히 각종 복잡한 양적, 질적 정보들을 유기적으로 고려하여 종합적으로 판단을 내리는 영역에서 인간이 수년, 수십년의 경험을 통해 쌓아 올린 암묵적 지식(예: tacit knowledge; Reber, 1989)과 지혜에 근거한 판단과 예측을 단시간에 대량의 정보를 주입해 학습시킨 기계가 능가할 수 있다면 이는 긍정적인 방향이든 부정적인 방향이든 중요한 변화 임에는 틀림없다.

인간의 종합적 판단에 기계가 도전하는 분야는 행정 민원 정보 수집 및 대응 (미국 보스턴 시 사례), 보건당국의 식당 위생검사 및 식중독 예방 툴의 개발 (미국 네바다 주 사례), 정부 민원 콜 센터의 인공지능 기반 프로그램 구축, 부정부패 사례 적발 프로그램의 개발 (중국 정부 사례) 등 다양한 영역으로 확대되고 있다. 저소득층 지원을 위한 금융대출 의사결정은 대표적인 머신러닝 기반 의사결정 대상 가운데 하나다. 은행 등 금융기관에서는 개별 대출자의 인적 정보, 교육 및 사회적 자본 정보, 재정 정보 등을 종합적으로 고려하고 가중치를 미세조정함으로써 대출의 가부와 대출 규모를 판단하게 된다. 그러나 수십년에 걸쳐 축적된 노하우에도 불구하고 대출금액의 회수에 실패하는 경우는 여전히 나타나며 디폴트를 정확히 예측하는 일은 기관 실적에 직접적으로 영향을 미치게 된다. 이와 같이 대규모 정보가 복잡하게 상호 관련된 사안에 대해 인간의 축적된 경험에 기반한 판단을 머신러닝 기반 의사결정으로 전환하였을 때 의사결정의 정확성(accuracy)과 결과에 대한 예측(prediction)은 어떻게 바뀔 수 있는지는 행정-정책분야에서도 예의주시하고 있는 주요 관심 분야다.

3. 분석틀

머신러닝 기반 의사결정이 행정-정책 분야에 미치는 영향에 관한 기존의 탐색적 연구들은 대체로 기계가 인간의 의사결정을 대체할 수 있는지 여부를 탐구해 왔다. 그러나 머신러닝 기반 의사결정은 모든 정책영역에서 인간의 의사결정을 대체할 수 없다. 연구의 진전을 위해 필요한 다음 단계의 질문은 바로 ‘대체 가능한 영역과 불가능 영역은 어디인가’ 라는 것이다. 이를 구체적으로 살펴보기 위해서 행정-정책 의사결정 과정에서 머신러닝 기반 의사결정과 인간의 의사결정을 비교하기 위한 분석틀을 마련할 필요가 있다.

행정-정책 의사결정의 주요 요소는 정책 문제(대상)와 정책 의사결정자(행위자)를 포함한다. 먼저 정책 문제의 특성은 구조화된 문제(structured problems)와 비구조화된 문제(unstructured problems)로 구별할 수 있다. 구조화된 문제는 목표가 명확하게 정의되어 있고 문제해결을 위한 구성요소가 명확히 밝혀져 있어 목표를 달성하기 위한 의사결정과정에 대한 합의가 이루어진 문제이다. 반대로 비구조화된 문제는 문제해결의 목표가 모호하고, 판단을 위해 필요한 정보의 범위가 정해지지 않고, 따라서 의사결정 과정에 대한 합의가 명확히 구성되어 있지 않은 문제라고 할 수 있다. 하지만 실제 정책 영역에서 구조화된 문제와 비구조화된 문제를 구별하는 것은 쉽지 않다. 공공조직의 목표 모호성 문제는 이론적으로 많은 논의가 되어 왔다(Rainey, 2009; 전영한, 2004). 특히 목표 모호성은 다양한 조직 목표의 경쟁적 해석을 초래할 수 있으므로 (Chun & Rainey, 2005), 구조화된 정책 문제를 식별하기 위해서는 우선 목표 모호성이 낮고 정책 목표에 대해 경쟁적 해석이 아닌 단일한 해석이 가능한 미시적 정책 문제를 중심으로 머신러닝 기반 의사결정을 적용한 뒤 점차 확대해 나가는 접근이 필요하다.

의사결정자의 특성은 머신러닝에 기반한 의사결정과 전통적인 인간 의사결정으로 나누어볼

수 있다. Simon (1972)이 지적한 대로 인간의 의사결정은 한정된 정보를 토대로 한정된 대안을 검토하여 그중 만족도가 높은 대안을 선택하는 제한적 합리성(bounded rationality)으로 특징지어진다. 반면 머신러닝의 의사결정은 대량의 정보를 동시에 처리함으로써 가용한 모든 대안을 비교하여 최적의 해를 구하는 완전한 합리성(rational choice)을 구현할 수 있다.

이처럼 정책문제(대상)와 정책결정자(행위)라는 두 가지 차원에 따라 행정-정책 의사결정 유형을 구분하면 아래 <표 2>와 같이 4가지 유형을 도출할 수 있다. 첫번째 유형은 구조화된 문제에 머신러닝을 활용한 의사결정을 수행하는 ‘머신러닝 의사결정’ 형태이다. 머신러닝 의사결정의 특징은 구조화된 문제에 대한 정보수집과 가능한 모든 대안을 분석한 후 최적의 결과를 도출해낼 수 있다는 것이다. 두번째 유형은 구조화된 문제에 인간이 제한적 합리성을 바탕으로 의사결정을 수행하는 ‘전통적 의사결정’ 방식이다. 인간은 제한적 합리성 하에서 만족할만한 대안을 도출해내며 이는 현실의 정책의사결정과 유사하다. 세번째 유형은 비구조화된 문제에서 머신러닝을 도입하는 경우이다. 이 경우 정책 목표와 의사결정 기준이 불분명하기 때문에 머신러닝의 의사결정은 최적 해를 도출할 수 없다는 점에서 머신러닝의 한계를 보여 준다. 네번째 유형은 비구조화된 문제영역에서 인간이 의사결정을 수행하는 경우이다. 인간은 문제를 무엇인지 숙고하고 이를 정의하고 목표를 수립하는 ‘메타적 의사결정과정’을 통해 문제에 접근해 나아간다. 따라서 메타 의사결정은 머신러닝에 의해 대체되지 않은 인간 고유의 의사결정 영역이라고 할 수 있다.

〈표 2〉 문제의 특성과 의사결정자의 특성에 따른 행정-정책 의사결정 유형 분류
〈Table 2〉 A Types of Policy Decision-Making

| 문제의 특성 | 의사결정자의 특성 | |
|---|---|---|
| | 완전한 합리성 (rationality) 기반 머신러닝 의사결정 | 제한적 합리성 (bounded rationality) 기반 인간 의사결정 |
| 구조화된 문제 (structured problems) | 유형 1. 머신러닝 의사결정 | 유형 2. 전통적 의사결정 |
| <ul style="list-style-type: none"> 문제의 성격이 잘 알려짐 문제해결의 목표 명확 목표 달성을 위한 의사결정 기준 존재 | <ul style="list-style-type: none"> 가용한 모든 정보 완전 활용 사전 제시된 기준에 따라 가능한 대안을 모두 탐색 후 최적 해를 도출 | <ul style="list-style-type: none"> 인지 한계로 인해 가용한 모든 정보의 완전 활용 불가 사전 제시된 기준에 따라 일부 가용한 대안을 비교 후 적정 해를 도출 |
| 비구조화된 문제 (unstructured problems) | 유형 3. 머신러닝의 한계 | 유형 4. 메타 의사결정 |
| <ul style="list-style-type: none"> 문제의 성격이 알려지지 않음 (새로운 문제) 문제해결의 목표 모호 목표 달성을 위한 의사결정 기준 부재 | <ul style="list-style-type: none"> 새로운 문제 발생 시 해결을 위한 목표를 자의적으로 설정 불가 문제해결을 위해 어떤 정보를 활용(또는 비활용) 할지에 대한 판단 불가 목표가 불분명하고 정보 활용(또는 비활용) 기준이 부재하여 해 도출 불가 | <ul style="list-style-type: none"> 새로운 문제 발생 시 해결을 위해 어떤 목표를 정립해야 할지 판단 문제해결을 위해 어떤 정보를 활용(또는 비활용) 할지에 대한 판단 문제 상황에 맞춰 목표 및 활용 정보의 범위를 조정하여 최선의 해를 도출 |

III. 사례 분석

위에서 도출한 4 가지 유형을 토대로 이 장에서는 행정-정책 의사결정 과정에서 머신러닝 기술을 도입한 사례들을 분석한다. 기존의 머신러닝 기반 의사결정 사례 연구들은 개별 사례들을 종합적으로 이해할 수 있는 일관된 분석틀을 제시하지 못했다. 이를 넘어서기 위해 이 연구에서는 의사결정 사안이 구조화된 문제(structured problems)인지 비구조화된 문제(unstructured problems)인지를 기준으로 머신러닝 기반 의사결정 사례들을 구분하여 분석한 뒤 각 상황에서 최적의 의사결정을 내리기 위해 고려해야 할 함의를 도출한다.

1. 구조화된 문제에서 머신러닝 기반 의사결정

1) 기계 대 인간: 누구의 판단이 더 나은가?

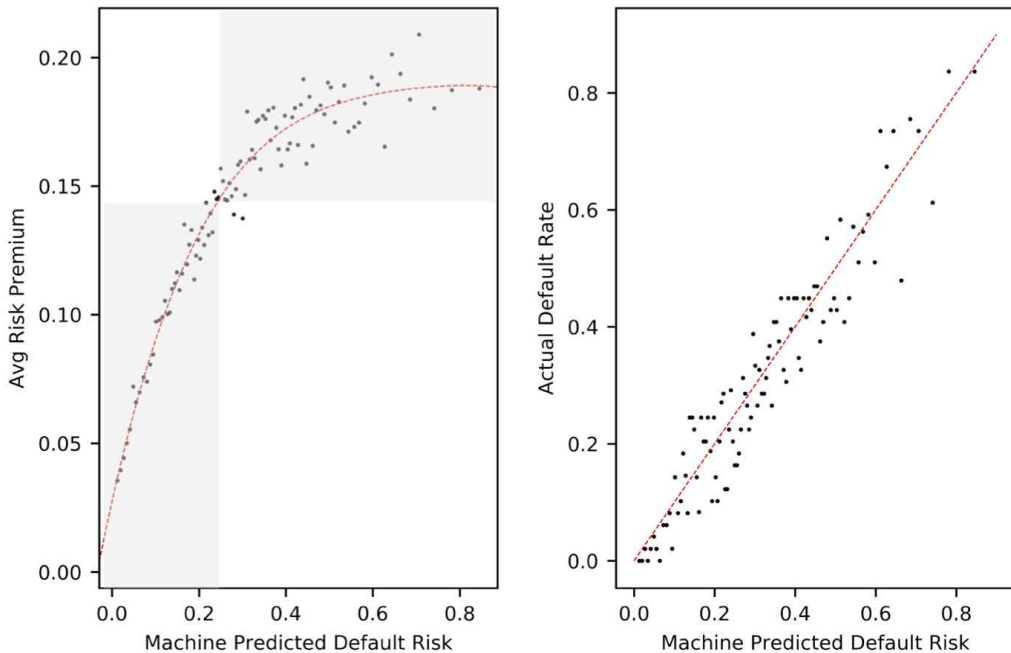
공공정책과 머신러닝의 결합에 대한 관심이 고조되고 있는 가운데 최근 서민 금융대출 자격 평가 및 심사 과정에서 머신러닝을 활용한 의사결정의 정확성과 예측력을 분석한 연구가 실시되었다(Fu, Huang, & Singh, 2018). 이 연구는 특히 머신러닝을 활용한 의사결정의 결과를 인간의 의사결정의 결과와 직접 비교하였다는 점에서 관심을 끌었다. 아래에서 이들의 연구 사례를 검토하면서 구조화된 문제(structured problems) 상황에서 머신러닝 기반 의사결정의 잠재성을 검토한다.

의사결정의 주요한 경로는 미래에 대한 예측(prediction)에 기반한 것이다. 가령 특정 사람이 보유한 기술과 능력에 기반하여 이 사람이 직업을 가질 수 있을 것인가, 혹은 직업을 잃을 것인가의 문제, 특정 사람이 보유한 경제적 사회적 자산에 비추어 이 사람이 대출을 승인 받을 수 있을 것인가의 문제 등 다양한 의사결정의 문제들이 '주어진 조건을 토대로 하여 미래 결과를 예측하고, 그 예측에 기반하여 가부간 의사를 결정한다'는 구조를 띠고 있다. 이러한 종류의 의사결정에서 기계가 인간의 의사 결정을 앞설 수 있을 것인가?

이 연구 질문에 답하기 위해 Fu et al. (2018)은 다음과 같은 연구를 설계하였다. 이들은 먼저 미국의 한 금융기관에서 2007년 3월부터 2008년 10월까지 일반 개인 고객 247,443명을 대상으로 한 대출심사 및 승인 관련 자료를 수집하였다. 이 자료에는 개인의 신용 정보, 재정 상황, 인구학적 정보, 교육 및 직업 정보 등 대출 심사에 필요한 모든 정보들이 포함되어 있었다. 또한 연구팀은 개인고객들에 대한 대출심사 결과 및 대출정보, 상환정보까지 수집하였다. 그 결과 전체 고객 가운데 약 8% 수준인 20,668명이 대출 승인을 받았으며 대출받은 사람들 가운데 13,976명은 무사히 대출금을 상환했으나 32%에 달하는 6,692명이 파산한 것으로 나타났다. 즉 대출을 승인한 의사결정의 결과 가운데 32%가량이 잘못된 의사결정을 내림으로써 개인고객의 파산과 금융기관측의 손실을 초래한 것이었다. 이러한 인간의 의사결정을 기계로 대체할 경우 과연 파산을

이 줄어들 것인가를 확인하기 위해 머신러닝 분석을 실시하였다. 이들은 먼저 전체 데이터 가운데 80%를 기계가 학습을 하는데 사용하도록 하였고, 남은 20%를 실제 의사결정을 테스트하는데 활용하도록 설계하였다. 머신러닝 프로세스는 XGBoost model(Chen & Guestrin, 2016)과 베이지언 최적화(Bayesian optimization)를 활용하였다. 이러한 학습을 통해 기계가 주어진 정보를 토대로 대출자의 파산 위험을 예측하도록 설계하였다.

〈그림 2〉 파산을 예측 결과: 기계 대 인간 비교 (Fu et al., 2018에서 재인용)
 (Figure 2) Results of the Default Rate Prediction (Reprinted from Fu et al. (2018))

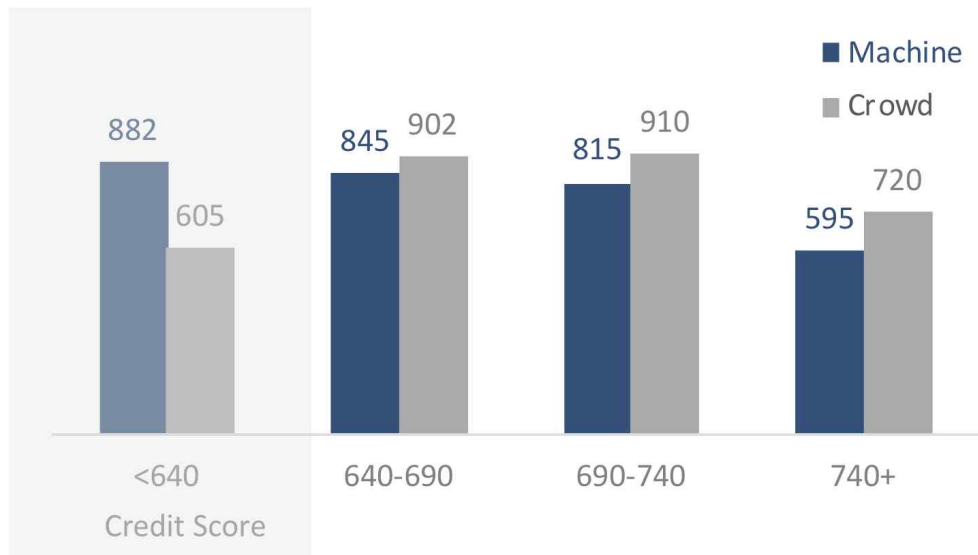


그 결과 기계가 인간의 의사결정보다 훨씬 뛰어난 성과를 보였다. 〈그림 2〉의 오른쪽 그래프는 기계가 예측한 파산위험율과 실제 파산율을 비교한 것이다. 그래프에서 보이는 대로 기계는 파산율을 거의 정확하게 예측하였다. 위 왼쪽 그래프는 기계의 파산위험 예측과 실제 인간의 파산위험 예측을 비교한 것이다. 위에서 볼 수 있듯이 파산위험율이 20% 이하인 구간에서는 기계와 인간의 예측이 거의 유사하다. 그러나 파산위험율이 20%를 넘어서는 구간에서는 기계가 60% 혹은 80%의 위험을 예측한 구간에서조차 인간의 예측은 위험율을 20% 이하로 보고 있다. 즉 파산위험을 기준으로 볼 때 저위험 집단에서는 기계와 인간의 의사결정에서 큰 차이가 나타나지 않을 수 있지만, 다양한 위험 요소를 종합적으로 고려해야 하는 고위험 구간에서는 인간의 파산위험 판단이 기계에 비해 다소 느슨하고 위험을 낮게 평가하는 경향이 나타난다는 것을 보여준다.

뿐만 아니라 인간과 기계가 각각 산출한 파산위험 예측에 따른 대출 승인이 기업에 어떤 결과

를 가져왔는지 분석한 결과 동일한 투자금액에 대해 기계는 항상 인간에 비해 통계적으로 유의미한(statistically significant) 성과 격차를 보여주었다. 즉 기계의 판단 결과가 인간에 비해 항상 더 나은 성과를 보여주었다. 기계는 또한 사람에 비해 실제로 대출이 필요한 집단에게 더욱 많은 기회를 제공하는 것으로 나타났다. 기계는 신용점수가 낮고, 집을 보유하지 않고, 신용 역사(credit history)가 짧은 사람들에게 더욱 관대한 대출을 제공하였다. 더욱 주목할 점은 머신러닝의 예측이 이자율을 정하는데 활용될 경우, 대출자는 더욱 낮은 이자율로 대출을 받을 수 있는 반면 투자자도 동시에 더 높은 이익을 창출할 수 있다는 것이었다(Fu et al., 2018). <그림 3>에서 보여주듯 신용점수가 640점을 넘는 세 구간 모두에 걸쳐 기계는 인간보다 적은 수의 대출을 승인하였다. 그러나 신용점수가 가장 낮은 '640점 미만' 구간에서는 인간이 605건의 대출을 승인한 반면 기계는 882건의 대출을 승인하였다. 이처럼 기계는 대출이 더욱 필요한 저소득층에 더 많은 대출을 승인하면서도 동시에 기업의 관점에서는 최대의 이익을 유지할 수 있는 의사결정을 제시하였다.

〈그림 3〉 신용점수 별 대출 승인 건수: 기계 대 인간 비교(Fu et al., 2018에서 재인용)
 〈Figure 3〉 Loan Approval by Credit Scores(Reprinted from Fu et al. (2018))



2) 오류 가능성과 해결 방안

무궁한 가능성에도 불구하고 머신러닝 기반 의사결정의 오류 가능성은 여전히 존재하며 그 중 하나는 의도하지 않은 차별(discrimination)의 문제다. 머신러닝 기반 의사결정은 인간의 판단을 상회하는 판단 능력과 성과를 보여준다. 그러나 문제는 기계가 차별을 하는 것처럼 보인다는 것이었다. 이러한 차별은 성별(예: 남성 대 여성), 인종(예: 흑인 대 백인), 국적(예: 내국인 대

이주민) 등 다양한 사회집단에 따라 나타날 수 있다. 성별에 따른 차별의 사례를 가상의 수치를 활용해 설명하면 다음과 같다. 가령 머신러닝을 통한 대출자격 심사 결과 남성 중 42%가 대출에 성공한 반면 여성 가운데서는 오직 30%만이 대출에 성공하였다. 이러한 결과는 여성에 대한 차별인가?

머신러닝의 차별에 대해 연구해온 컴퓨터공학자들은 기계가 비가시적으로 차별하는 것을 막기 위하여 하나의 단순하고 간결한 기준을 추가하였다. 그것은 바로 자격을 가진 집단 내에서는 동일한 기회를 가져야 한다는 동일 기회(equal opportunity among the qualified)의 원칙이다. 동일 기회의 원칙이 머신러닝에 적용되는 방식은 다음과 같다. 현대 사회에서 여전히 성별에 따른 임금 격차가 실재 한다는 사실은 무수한 연구를 통해 입증된 사실이다(Weichselbaumer & Winter-Ebmer, 2005). 여성은 남성에 비해 경제적으로 불리한 환경에 처해 있다. 이 때문에 금융기관이 사전에 설정한 대출자격 기준에 도달할 수 있는 여성은 남성에 비해 현저히 적을 수밖에 없다. 예를 들어 여성 중 50%가 자격요건이 되고, 남성 중 70%가 자격이 된다고 하자. 이 경우 지원자 가운데 몇 퍼센트의 여성과 남성이 대출을 받아야 하는가? 올바른 기준은 여성의 50%와 남성의 70%가 아니다. 전술한 동일 기회의 원칙에 따르면 대출자격을 확보한 사람들 중에서는 대출을 받을 수 있는 확률이 동일해야 한다. 가령 대출 확률이 60%라고 하자. 이 경우 남성 가운데 대출자격을 확보한 집단, 즉 남성 중 70%의 '자격집단' 내에서 만약 60%가 대출을 받았다면 여성 가운데 대출자격을 확보한 집단, 즉 여성 중 50%의 '자격집단' 내에서도 60%가 대출을 받아야 한다. 이 경우 외관상 전체 남성 중 42%가 대출에 성공한 반면 전체 여성 중 30%만이 대출에 성공한 것으로 보이지만, 자격집단 내 동일기회 원칙에 따라 남녀 모두 자격 집단 내에서는 60%가 대출에 성공함으로써 머신러닝의 판단이 소수자(minority)를 차별한다는 의혹을 벗어날 수 있게 된다.

3) 합리모형의 구현 가능성

이 사례는 현재 진행되고 있는 머신러닝 연구에 중요한 시사점을 제시하고 있다. 그 중 중요한 시사점은 기계가 인간의 판단과 의사결정을 어느 정도 까지 대체할 수 있는냐에 관한 것이다. 동일한 정보가 제공될 경우 기계는 인간보다 언제나 더 나은 판단을 제시하며 그 결과 기업의 이익에도 큰 기여를 하게 된다. 합리모형이 구상했던 이상적 의사결정의 가능성이 머신러닝을 바탕으로 구현 가능성이 열린 것이다. 머신러닝 기법은 인간의 인지적 한계를 벗어나 데이터수집과 분석을 가능하게 하여 합리적인 의사결정을 가능하게 한다. 또한 머신러닝 기반 의사결정은 개인의 주관과 정서를 배제하기 때문에 객관주의에 입각한 의사결정을 가능하게 한다. 이와 같은 사례가 정부 영역에 도입되면 정책의 공급자(정부)의 오류를 줄여 낭비를 감소시킬 뿐 아니라, 수요자(국민, 여기서는 자금이 필요한 사람)에게 보다 많은 자원을 공급 하기 때문에 이상적인 정책 의사결정에 다가간 것이라 평가할 수 있다. 그렇다면 앞으로 더욱 많은 의사결정과 심사 및

평가 과정이 기계에 의해 대체되어야 할 것인가. 기계의 판단에는 오류나 편의(bias)가 없을 것인가. 이 연구는 기계와 인간의 의사결정 성과(performance) 비교를 시도했다는 점에서 주목을 받았지만 기계 의사결정의 한계라는 중요한 질문에 대한 답은 명확히 제시하지 못하고 있다. 아래에서는 그 가운데 주요한 이슈를 특히 비구조화된 문제 상황에서 머신러닝 기반 의사결정의 한계를 중심으로 검토한다.

2. 비구조화된 문제에서 머신러닝 기반 의사결정

앞서 논의한 대출자격 심사 사례는 구조화된 문제에 대한 머신러닝 기반 의사결정의 효과를 보여주고 있다. 이 밖에도 공공분야에서 인공지능 및 머신러닝 기반 의사결정의 가능성을 탐구하는 연구가 최근 증가하고 있다. 아래에서는 비구조화된 문제(unstructured problems) 상황에서 머신러닝 기반 의사결정 도입 사례들을 검토함으로써 윤리적 판단 및 연성정보(soft information) 처리 문제 등 머신러닝 기반 의사결정의 개선 과제들을 살펴본다.

1) 비구조화된 문제(unstructured problems) 상황

행정-정책 의사결정의 특징 가운데 하나는 행정 및 정책 담당자들이 과거에는 경험하지 못했던 전혀 새로운 종류의 문제들에 직면하는 경우가 많다는 점이다. 급속한 기술 발전에 따른 행정-정책 수요의 변화라든가 예상하지 못한 자연 재해 상황에 적절하게 대처하기 위해서는 기존의 축적된 경험과 자료에만 의존할 수 없는 경우가 있다. 가령 2012년 발생한 동일본 대지진의 경우가 대표적인 사례다. 원자력 발전 관련 규정과 각종 정책을 마련하는 인간은 대규모 지진 발생 상황이나 대규모 쓰나미 발생 상황에 대한 개별 매뉴얼을 철저히 준비했지만 두 가지 상황이 동시에 발생하는 재난 상황에 대한 준비는 부족했고 그 결과 인간은 원자력 발전소의 멜트다운(melt down)이라는 전혀 새로운 재난을 직면해야 했다. 이처럼 문제의 성격이 알려지지 않은 새로운 문제 상황이 발생할 경우 문제해결의 목표가 모호해지고 목표 달성을 위한 의사결정 영역도 불확실해진다. 이 경우 머신러닝 기반 의사결정은 합리적 의사결정이라는 장점을 전혀 발휘할 수 없다. 새로운 문제 해결을 위한 목표를 자의적으로 설정할 수 없으며 문제해결을 위해 어떤 정보를 활용해야 할지에 대한 판단도 머신러닝 의사결정 범위를 넘어선다. 따라서 최적의 해를 도출하는 것은 불가능해진다.

일례로 미국의 주 정부 법무 부서의 문서 검토 및 분류를 위해 도입된 머신러닝 기반 의사결정 프로그램은 개별 문서의 성격을 파악하는 과정에서 전혀 새로운 종류의 문서가 등장할 때마다 제대로 된 판단을 하지 못한다는 것이 드러났다(Cluster, 2018). 가령 변호사와 의뢰인 사이의 비밀 유지 특권에 따라 보호받아야 할 문서들을 분류하는데 있어서 특권 문서를 구성하는 내용을 학습 시키는 과정에서 오류가 발생하는 것이다. 개별 사안마다 구체적인 사건 상황과 등장 인물,

각 인물간의 관계 등이 매우 복잡하게 얽혀있는 법률 문서에서 어떤 문서가 비밀 유지가 필요한 특권 문서인지 일관된 기준을 제공할 수 없기 때문에 머신러닝 기반 의사결정은 특권 문서의 분류라는 최적의 해를 전혀 효과적으로 도출하지 못한 것이다. 결국 법률문서 분류 작업은 변호사가 직접 특권이 적용될 수 있는 문서를 검토 해서 특권 적용에 대한 최종 결정을 내리는 전통적인 방식으로 회귀할 수 밖에 없었다.

2) 윤리적 판단

머신러닝 기반 의사결정 및 판단의 잠재적 문제점 중 하나가 윤리적 판단의 정확성에 관한 것이다. 인간 행동은 컴퓨터 신호와 같은 이진법적인(binary) 방식으로 구분될 수 없는 경우가 많으며 그 가운데 하나가 바로 부패(corruption)에 대한 판단이다. 중국과학원과 정부가 부패 감지 시스템으로 개발한 ‘제로 트러스트(Zero-Trust)’ 시스템은 중국 정부가 보유한 150종 이상의 데이터를 바탕으로 중국 공무원들의 비리를 감시하는 인공지능 시스템이다(국경완, 2019). 빅데이터를 토대로 한 머신러닝 분석을 통해 개별 공무원의 이상 행동(예컨대 공무원의 계좌에서 많은 돈이 한꺼번에 입금 되는 경우, 주변 친인척의 근거 없는 사치품 소비 등)을 탐지하는 방식이다. 중국 전역으로 확대 적용되기 전 시범 운용 과정에서 8,000명 이상의 부패 의심 사례를 탐지해냄으로써 주목을 받았다. 그러나 문제는 이 가운데 실제 부패 사범으로 확정된 경우는 25%에 그치는 것으로 나타났다. 이 과정에서 부수적으로 발생할 수 밖에 없는 인권침해 문제는 차치하더라도 인공지능의 광범위한 데이터 수집은 공공의 부패를 감시한다는 명분에도 불구하고 판단의 정확도는 여전히 개선의 여지가 남아 있다. 특히 부패 등 윤리적 문제에 대한 판단은 정량적인 데이터에 기반한 기계적 판단만으로는 한계가 있음을 보여준다.

3) 경성정보(hard information)와 연성정보(soft information)

의사결정을 위해 처리되는 정보는 경성정보(hard information)와 연성정보(soft information)로 나눌 수 있다. 경성정보는 대상의 특성을 수치로 변환할 수 있는 양적(quantitative) 정보들을 가리키는 반면 연성정보는 수치화 하기 어려운 질적(qualitative) 정보를 일컫는다. 연성정보의 예로는 인간 개개인의 견해, 감정, 혹은 비언어적인 표현 등을 들 수 있다. 최근 급속히 발전하고 있는 머신러닝 기술은 양적인 경성정보는 물론 수치화 하기 어려운 연성정보를 처리하는 기술 발전으로 이어지고 있다. 실리콘 밸리의 주요 기업들은 이미 인터넷 검색시 인간의 표정을 토대로 상품 구매 여부를 예측하는 등 연성정보 해석의 정확도를 높이는데 박차를 가하고 있다. 그럼에도 불구하고 아직까지 연성정보에 대한 완전한 정복은 요원한 것으로 보인다. 가령 자율주행차의 오류로 운전자가 사망한 사고(중앙일보, 2018)나 미국 실리콘밸리에서 경비 로봇이 아이를 다치게 한 사례(CNN, 2018)를 예로 들 수 있다. 조사결과 자율주행차는 태양의 역광 때문에 흰색 트랙을 하늘로 오해한 것으로 나타났으며, 실리콘밸리 사례의 경우 경비 로봇이 16개월 아이를

사람으로 인식하지 못해 사고가 발생한 것으로 알려졌다. 머신러닝 기술이 연성정보를 해석하는데 있어서 아직 한계가 있다는 것을 보여주는 사례다.

이상에서 검토한 바와 같이 머신러닝 기반 의사결정은 비구조화된 문제(unstructured problems) 상황이나 윤리적 판단, 연성정보(soft information)의 처리 등과 같이 데이터 수집이 어렵거나 수집된 정보의 해석이 불확실한 문제 상황에서는 인간의 의사결정을 완전히 대체하지 못하고 있으며 오히려 인간의 의사결정보다 정확도가 떨어지고 오류발생 가능성이 높은 것으로 나타났다.

3. 시사점

머신러닝 기반 의사결정 연구 사례들을 구조화된 문제와 비구조화된 문제 영역으로 구분하여 분석함으로써 두 가지 중요한 시사점을 도출하였다. 첫째, 구조화된 문제 영역에서 합리적 의사결정의 구현 가능성이다. 둘째, 비구조화된 문제 영역에서 기계의 한계와 인간 역할의 중요성이다.

1) 구조화된 문제에서 합리적 의사결정의 구현 가능성

머신러닝을 통한 인공지능의 대출 심사는 인간이 행한 대출심사에 비해 공급자와 수요자 모두에게 보다 나은 편익을 제공한다. 공급자에게는 디폴트 가능성을 낮춤으로서 수익을 올리고 수요자에게는 신용점수가 낮더라도 대출가능성을 높이는 결과를 낳은 것이다. 이러한 인공지능의 활용가능성이 정부의 정책 의사결정에 도입될 경우 정책 의사결정의 이상적 모형으로 제시되었던 합리모형의 실천적 구현 가능성을 보여준다. 머신러닝 기술은 인간에 비해 월등히 높은 데이터를 동시에 빠른 속도로 분석해 낼 수 있다. 완전한 정보와 이에 대한 분석을 적절한 시간 내에 가능하게 하기 때문에 합리모형의 실현가능성을 높인다. 또한 머신러닝 기반 의사결정은 인간의 주관적 감정이 개입되지 않기 때문에 합리모형이 전제하는 객관주의의 실현을 가능하게 한다. 머신러닝에 의한 대출 심사와 인간에 의한 대출 심사의 비교에서 드러나듯이 인간은 위험을 과소평가하는 편향을 보이는 반면 머신러닝은 대출자의 디폴트 가능성을 정확하게 예측해 낼 수 있었다. 머신러닝 기반 의사결정은 주어진 가치체계 하에서 목적을 효과적으로 달성하는데 인간에 비해 우수한 능력을 보여주었다. 대출심사 사례에서는 주어진 정보를 바탕으로 모든 인간을 다른 조건은 모두 동등하다고 전제했을 경우 객관적으로 대출심사를 했던 것이다. 결과적으로 이 사례 분석은 머신러닝 등 인공지능 기술을 매우 잘 구조화된 정책 문제에 적용한다면 인간에 비해 우수한 성과를 낼 수 있으며 합리모형의 이상적 형태를 구현할 수 있다는 가능성을 보여준다.

2) 비구조화된 문제에서 머신러닝의 한계와 인간의 역할

그럼에도 불구하고 인공지능을 활용한 정책 의사결정이 모든 영역에 적용하기에는 한계가 있

다. 첫번째는 메타(meta) 의사결정에 관한 것이다. 즉, 정책결정에 인공지능을 활용할 것인지 아닌 것인지에 대한 결정은 인간의 고유한 영역이라는 것이다. 이러한 메타적 의사결정은 인공지능의 개입여부에 관한 판단 뿐만이 아니라 인공지능이 어떤 기준과 목적으로 정책 의사결정을 할 것인지에 대한 판단과 관련된다. 대출심사 사례에서도 드러나듯이 인공지능은 사회적 구조와 맥락에 대한 조건을 지정하지 않으면 수리적으로 동등한 존재로 취급하게 된다. 사회적 약자에 대한 비율 할당과 같이 사회적 가치에 대한 고려 여부에 대한 의사결정은 인공지능으로 대체할 수 없는 것이다. 이러한 맥락에서 합리모형이 전제하는 일원화된 가치체계에 대한 가정은 모든 정책 의사결정체계에 적용하는 것은 어렵다. 다만 인간이 신중히 판단한 독립적으로 구성 가능한 가치체계와 조건하에서 인공지능에게 합리적 분석을 위임하는 것으로 구성되어야 할 것이다.

두번째 인공지능으로 구성된 합리모형에 정책 의사결정을 맡기기 어려운 이유는 다음과 같다. 인공지능에 기반한 합리모형은 인공지능이 구성한 체계내에서는 완벽한 선택을 구현할 가능성이 높다. 그러나 실제 정책이 펼쳐지는 세계는 단절적 세계가 아니라 끊임없이 상호 연결되어 작용하는 연속적 체계이다. 사례에서는 디폴트 가능성을 낮추는 최선의 선택을 한 것이지만 실제 정책 현장에서의 정책 시행은 단기적 효과 뿐 아니라 장기적 성과와 가능성을 바탕으로 판단되어야 한다. 그러므로, 디폴트 가능성을 낮추는 것과 같은 단기적 효과 뿐 아니라 대출 여부에 인한 국민 경제 전체에 미치는 파급효과를 장기적으로 고려하여 판단할 필요가 있는 것이다. 결국 인공지능이 인식하는 인간사회는 인간이 인식하는 인간사회와는 다르기 때문에 인공지능이 합리모형을 완벽히 구현할 것이라는 것은 아직까지 이상에 가깝다고 볼 수 있다.

끝으로 인공지능에 기반한 정책 의사결정은 인간의 올바른 의사결정 과정과 증거기반 의사결정의 중요성을 환기시킨다. 의사결정이 합리적으로 이루어진다고 하더라도 결과에 대한 판단의 시간 범위를 정하는 방식에 따라 성과의 측정이 달라진다. 그러므로 이제까지의 결과 중심의 판단에서 과정과 절차를 중시하는 의사결정과정 체계를 구축한다면 보다 투명하고 공정한 의사결정체계를 구축할 수 있을 것이다. 또한 절차를 중시하는 의사결정체계를 구축함에 있어서 증거기반 의사결정체계를 구축하는 것은 중요성은 점차 커지고 있다.

IV. 논의

정부의 의사결정은 개별 정책 담당자 차원에서 이루어지는 것도 있지만 부처 등 조직 차원에서 이루어지는 경우가 더욱 많다. 특히 많은 정부 의사결정은 새로운 행정-정책 수요에 부응하기 위한 판단을 필수적으로 수반해야 하는 경우가 빈번하다. 이 경우 행정-정책 의사결정을 인공지능에 맡기는 것이 적절한 것인가 또는 심지어 가능한 것인가에 대한 논의가 필요하다.

1. 메타(meta) 의사결정: 기계의 한계와 인간의 역할

앞서 분석한 사례를 예로 들면, 개별 고객에 대한 대출자격 평가 및 심사에 기반한 대출 여부 의사결정을 아우르는 상위 수준의 의사결정(meta decision)은 바로 “대출자격을 기계가 심사해도 되는가”에 대한 의사결정이다. 대출자격 평가 및 심사를 사람에서 기계로 대체할 것인가에 대한 판단은 기계가 아닌 사람이 직접 수행해야 하는 의사결정이다. 이러한 의사결정에 따라 대출자격 평가 및 심사 과정을 기계로 대체하거나 혹은 기존 관행대로 사람이 계속 수행하는 데 따라 소요되는 비용과 자원의 규모가 달라진다. 또한 중장기 미래 전략의 관점에서 조직내 주요 직무를 기계로 대체할 지에 관한 전략적 의사결정의 결과는 의사결정과 결과 확정 사이의 시간적 간격(decision lag) 문제를 낳기도 한다.

대출자격 심사와 같이 의사결정을 위한 소스(source)와 의사결정의 목적(goal)이 외부에서 명확히 제시된 구조화된 문제(structured problems)의 경우 머신러닝 기반 의사결정이 더욱 정확한 판단을 제시할 가능성이 높다. 반면 의사결정을 위해 필요한 소스에 어떤 종류의 정보를 포함할 것인지(예: 상환 가능성 관련 정보 뿐만 아니라 형평성 관점에서 대출자의 필요 수준을 고려할 것인지 말 것인지) 혹은 의사결정의 최적 목표를 어떤 것으로 설정할 것인지(예: 효과성 혹은 효율성), 나아가 어떠한 종류의 의사결정을 기계에 맡길 것인지 등은 구조화된 의사결정의 범위를 포괄하는 의사결정, 즉 메타(meta) 의사결정의 범위에 속한다. 메타 의사결정은 머신러닝 기반 의사결정 알고리즘의 내용과 방향, 더불어 알고리즘 구현 여부 자체를 결정하는 상위 개념의 의사결정이다. 이러한 메타 의사결정은 머신러닝과 딥러닝 등 다양한 인공지능 기술 기반 의사결정이 확대 강화되는 가운데서도 결코 기계에게 내어주지 못하는 (혹은 기계가 결정하지 못하는) 인간만의 고유한 의사결정 영역으로 남게 된다. 이러한 상황에서 최선의 판단을 내릴 수 있도록 도와주는 수단이 바로 증거기반 의사결정 (evidence-based decision making)이다. 아래에서는 머신러닝에 의해 대체될 수 없는 인간 고유의 의사결정을 수행하기 위한 기법으로 주목받고 있는 증거기반 의사결정의 원리를 검토하고 인공지능 시대에 인간의 역할에 대한 함의를 탐색한다.

2. 증거기반 의사결정의 역할

1) 의사결정의 질(decision quality) 향상: 결과중심 접근에서 과정중심 접근으로

조직의 의사결정(organizational decision making)은 개인의 의사결정에 비해 예산 규모가 크고 의사결정의 영향 범위가 넓고 장기간에 걸쳐 의사결정의 효과가 나타나는 특성이 있다. 이 때문에 잘못된 의사결정은 조직의 자원을 훼손하고 장기적으로 큰 손실을 초래하게 된다 (Yates, 2003). 이러한 의사결정의 특성 때문에, 조직의 의사결정의 질적 수준(quality)을 어떻게

게 측정할 것인가는 연구자들의 오랜 과제였다(O'Reilly, 1983). 전통적인 의사결정 연구자들은 의사결정이 얼마나 올바른 것이었는지 판단하는 방법으로서 의사결정이 초래하는 결과에 주목했다. 즉, 의사결정의 결과가 바람직하거나 기대를 충족할 경우 해당 의사결정은 높은 수준의 의사결정(high-quality decision)이라고 판단하게 되는 것이다. 이러한 “결과 중심의 의사결정 퀄리티(results-oriented decision quality)” 개념은 오랫동안 연구자들이 애용하는 판단 기준을 제공하였다(Yates & Potworowski, 2012). 그러나 이러한 결과 중심의 의사결정 수준 판단은 치명적인 문제를 내재하고 있다. 조직의 의사결정은 많은 경우 의사결정의 시점과 그 의사결정의 결과가 현실화되는 시점 사이에 시간적 간극이 발생한다. 예를 들어 급변하는 환경에 발맞추기 위한 각 조직의 중장기 전략은 그 성패를 판단할 수 있을 때 까지 수년 이상의 시간이 걸릴 수 있다. 특히 대규모의 예산과 자원이 투자되어야 하는 의사결정의 경우 그 결과는 오랜 시간에 걸쳐 점진적으로 나타나는 경향이 있다. 이러한 의사결정-결과산출 간의 간극(decision lag)은 의사결정의 정확성을 결과에 비추어 판단하고자 하는 결과 중심 의사결정 수준(results-oriented decision quality)의 치명적인 약점이다.

이를 극복하기 위해 최근 주목받고 있는 대안이 바로 과정 중심의 의사결정 수준(process-oriented decision quality) 판단 접근이다(Barends, Rousseau, & Briner, 2014; Rousseau, 2012; Yates & Potworowski, 2012). 결과를 근거로 의사결정의 수준을 판단하기에는 시간적 제약이 크다. 이 경우 조직에서 상시적으로 요구되는 각종 의사결정의 수준(decision quality)을 향상하기 위해서는 의사결정의 과정(decision process)에 주목해서 이 과정을 개선해야 한다는 것이 바로 과정 중심의 의사결정 접근법의 요지다. 의사결정 과정을 최선의 과정으로 유지하기 위해서는 두 가지 요소가 필요하다. 첫째, 의사결정에 활용되는 데이터와 근거 자료의 질(quality of evidence)과 둘째, 이러한 수준 높은 자료를 이해하고 당면한 조직 문제에 적절하게 적용할 수 있는 의사결정자의 근거 활용의 질(quality of use)이다. 이와 같이 당면한 의사결정에 최선의 정보를 제공할 수 있는 다양하고 수준 높은 근거자료(high quality evidence)와 이를 의미를 정확히 이해하고 당면한 조직 맥락에 맞게 적절하게 적용할 수 있는 사용의 퀄리티(high quality use)에 기반한 의사결정을 증거기반 의사결정(Rousseau, 2006, 2012)이라고 한다.

2) 증거기반 의사결정 (Evidence-based decision making)

증거기반 의사결정은 1990년대 초반 의학분야에서 시작되었다. 당시 연구자들이 미국의 의료현장에서 진료하는 의사들을 상대로 연구한 결과, 전체의 85%에 달하는 의사들이 수십 년 전 자신이 의대에 다닐 때 배웠던 지식을 여전히 진료 현장에서 사용하고 있었고, 오직 15%의 의사들만이 최신의 의학지식을 진료와 치료에 활용하고 있다는 사실을 알게 되었다(Tranfield, Denyer, & Smart, 2003). 이러한 충격적인 사실에 기반해 의료계 일각에서는 최신의 의학 지식에 기반한 최적의 진료 의사결정을 내려야 한다는 증거기반의료(evidence-based medicine) 운

동이 펼쳐졌고 그에 기반해 코크란 협력체(Cochrane Collaboration)가 설립되었다. 이후 최신의 과학적 지식을 의사결정 전반에서 활용해야 한다는 아이디어는 큰 힘을 얻었고 이러한 운동은 보건, 간호 등 의료계 전반을 넘어 공공정책, 사회복지, 경찰행정, 교정, 경영 등 다양한 분야로 확산되었다.

이러한 토대 위에 확립된 증거기반 의사결정의 원리는 다음과 같다. 조직의 의사결정 사안이 발생할 경우 의사결정자는 과학적 지식, 조직 내 정보와 데이터, 전문가 식견, 이해관계자 입장 등 다양한 증거자료들을 종합적으로 판단해서 최상의 선택을 해야 한다는 것이다(Rousseau, 2012; Rynes, Colbert, & O'Boyle, 2018). 이러한 견해는 일견 너무나 상식적인 것으로 보이지만 보통 사람들이 생각하는 '근거에 기반한 의사결정'은 증거기반 의사결정과 다소 상이하다. 가령 어느 조직의 리더가 구성원들의 직무만족도를 토대로 조직운영의 개선점을 찾고자 하는 상황을 상정한다. 이 상황에서 리더는 두가지 설문결과를 받아보았다. 인사부에서 헤드쿼터 근무자 200명을 상대로 실시한 보수/업무환경 만족도 조사와 해외영업부에서 산하 해외 지사 직원들 150명을 상대로 실시한 일반 직무만족도 조사결과 등 두가지는 서로 상반된 수치를 제시하고 있다. 이 경우 리더는 어떤 자료에 기반해 의사결정을 해야 하는가? 대부분의 조직내 의사결정은 이처럼 다양한 근거의 소스들이 혼재된 상황에서 이 중 어떤 정보를 선택하고, 어떤 정도의 가중치를 두어야 하는지에 의해 결정된다. 어떤 조직 의사결정도 자의적인 기준에 의해 결정되지 않으며 구체적인 수치에 근거해 결정된다. 그러나 증거기반 의사결정은 한발 더 나아가 그러한 근거가 되는 수치의 신뢰성, 즉 증거의 신뢰성(quality of evidence)에 주목하여 의사결정의 퀄리티를 높이하고자 하는 것이다.

3) 증거의 퀄리티와 다원화

조직의 의사결정은 대부분의 경우 객관적인 정보에 근거하고 있다. 어떤 조직도 대규모 예산이 투입되는 중요한 의사결정을 리더의 직관이나 기분에 근거하여 결정하지 않는다. 그렇다면 증거기반 의사결정은 기존의 조직 의사결정과 어떤 면에서 다른가? 증거기반 의사결정의 핵심 요소 중 하나는 의사결정의 근거가 되는 정보와 자료(즉 증거자료)의 신뢰성을 중요하게 고려한다는 것이다(Rousseau & Gunia, 2016). 먼저 의사결정의 토대를 제공하는 근거 자료의 종류를 네 가지로 구분한다(HakemZadeh & Baba, 2016).

첫번째 종류는 과학적 지식(scientific knowledge)이다. 많은 연구자들과 실무자들이 공통적으로 호소하는 과학적 지식의 문제점은 바로 연구와 실무 간의 격차(research-practice gap)라고 할 수 있다. 연구자들은 조직 관리와 조직내 의사결정에서 중요한 요소들과 최적의 의사결정을 위해 필요한 조건들이 무엇인지 과학적 연구를 통해 밝혀왔다(Locke, 2009). 그러나 현장의 실무자들은 이러한 유용한 지식과 팁들이 가용하다는 사실 자체를 알기 어렵고 따라서 현장에서는 기존의 관행대로 조직을 운영하든가 혹은 새로운 시도를 통해 실패를 거듭하며(learning

by doing) 노하우를 축적해간다. 증거기반 의사결정은 실무자들이 조직의 중요한 의사결정을 앞두고 관련된 과학적 증거가 가용한 지 여부를 먼저 확인하고 이를 적극 활용할 필요가 있다고 강조한다.

두번째 종류의 증거는 조직 내 정보와 자료(organizational data)다. 실무자들이 과학적 지식에만 기댈 수 없는 이유 중 하나가 바로 개별 조직은 각자의 조직적 맥락(organizational context)이 있어서 일반화된 과학적 지식을 그대로 적용할 수 없다는 것이다. 이 경우 각 조직은 축적된 조직 내부 자료를 바탕으로 패턴을 찾아내고 미래를 예측할 수 있어야 한다. 따라서 조직 내 자료는 증거기반 의사결정의 중요한 두번째 축이다.

세 번째 종류의 증거는 실무자의 전문성(professional expertise)이다. 앞서 살펴본 대로 기계는 구조화된 문제(structured problems)를 해결하는데 있어 인간보다 뛰어난 성과를 보여준다. 하지만 조직이 당면한 의사결정의 성격이 기존에는 알려지지 않은 완전히 새로운 성격의 것이라면 기계가 의사결정을 학습할 데이터가 존재하지 않게 된다. 이처럼 기존에 자료가 축적되지 않은 새로운 종류의 상황에서는 머신러닝의 원리가 통하지 않게 된다. 예를 들어 대규모 지진과 쓰나미의 결합이 원자력 발전소의 멜트다운(melt down)을 초래할 수 있을 것이라는 사실은 2011년 동일본 대지진 이전에는 누구도 예측하지 못했다. 이러한 새로운 상황에는 인간의 축적된 경험과 노하우가 의사결정을 위한 정보와 근거를 제공할 수 있다.

마지막으로 네 번째 종류의 증거는 이해관계자의 입장(stakeholder perspective)이다. 20여년 전 처음 증거기반의학 개념이 정립될 당시 의학자들이 주목한 지점은 현장의 의료진이 주어진 증상을 치료하는 데만 집중함으로써 환자의 의견을 무시하고 환자의 삶의 질을 의사결정 과정에서 고려하지 않았다는 것이었다. 가령 말기 암 환자에게는 실험적 처치 방법을 적극 받아들여 서 고통스러운 항암 치료를 계속할 것인지 혹은 더 이상의 치료를 중단하고 남은 시간을 주변 사람들과 충분한 시간을 보내며 삶을 정리하는데 사용할 것인지 결정할 권리가 있으며 이러한 환자의 요구는 의학적 의사결정에서 중요한 결정요인이 되어야 하는 것이다. 같은 원리에 의해 대부분의 조직 의사결정은 해당 의사결정에 의해 영향 받는 다양한 이해관계자가 존재하며 그러한 이해관계자의 입장을 충분히 고려하는 것 또한 의사결정의 퀄리티를 결정하는 중요한 요소가 된다. 가령 정부 혹은 공공기관의 의사결정은 다수 국민에게 직접적인 영향을 미치게 되며 이때 의사결정은 단순히 과학적 근거와 조직 내부 자료에 기대어 결정할 수만은 없는 것이다. 기업의 의사결정도 마찬가지다. 대규모 구조조정은 당장 기업에서 고용한 근로자에게 직접적인 영향을 미치게 되며, 구조조정 이후 살아남은 근로자들 또한 심리적 충격에 따른 직무동기 저하와 생산성 하락을 경험할 수 있다. 이런 경우 단기적인 비용절감에 초점을 맞춘 구조조정은 중장기적으로 기업의 경쟁력에 치명적인 타격이 되는 잘못된 의사결정이 될 수 있는 것이다.

이처럼 증거기반 의사결정은 단순히 근거에 기반한 의사결정이 아니라 신뢰할 만한 과학적 지식, 조직내 정보와 자료, 실무자의 전문성, 이해관계자의 입장 등을 종합적으로 고려한 최적의 의사결정체계이다. 이와 같은 의사결정체계를 갖추고 인공지능의 강점인 무한히 확장 가능한

정보수집과 처리, 정확한 분석과 빠른 판단, 객관주의의 확보를 활용한다면 보다 이상적인 정책 의사결정 체계를 갖출 수 있을 것이다.

V. 결론

이 연구는 합리적 모형의 이상과 한계를 검토하고 머신러닝 기반 의사결정 사례를 행정-정책 의사결정 관점에서 분석하여 다음과 같은 함의를 도출하였다. 첫째, 대출자격 평가 및 심사와 같은 구조화된 문제(structured problems)와 인간이 직접 판단해야 할 비구조화된 문제(unstructured problems)를 구분하는 것이 중요하다. 둘째, 인간 고유의 의사결정 영역인 메타(meta) 의사결정을 위해서 정책 담당자 등 의사결정자는 증거기반 의사결정(evidence-based decision making)을 중요성을 인식하고 실제 의사결정에서 적극 실현해야 한다.

참고 문헌

- 강정수 (2015). 인공지능 신화. 이슈&트렌드 2015년 9월, 한국인터넷진흥원.
- 국경완 (2019). 인공지능 기술 및 산업 분야별 적용 사례. 주간기술동향, 1888: 15-27.
- 성욱준·황성수 (2017). 지능정보시대의 전망과 정책대응 방향 모색. 정보화정책, 24(2): 3-19.
- 윤상오 (2017). 지능정보시대의 정책결정: 인공지능 정책결정의 주요 쟁점들. 한국경영과학회 학술대회는 문집, 2447-2470.
- 윤상오·이은미·성욱준 (2018). 인공지능을 활용한 정책결정의 유형과 쟁점에 관한 시론. 한국지역정보학회지, 21(1): 31-59.
- 이강윤·김준혁 (2016). 인공지능 왓슨 기술과 보건의료의 적용. 의학교육논단, 18(2): 51-57.
- 전영한 (2004). 공공조직의 목표모호성: 개념, 측정 그리고 타당화. 한국행정학보, 38(5): 49-65.
- 정정길·최종원·이시원·정준금·정광호 (2003). 정책학원론, 서울: 대명출판사.
- 진상기·박영원 (2017). 제 4 차 산업혁명의 미래전략체계에 관한 연구: AHP 분석을 중심으로. 한국지역정보학회지, 20(3): 31-58.
- 최종원 (1995). 합리성과 정책연구. 한국정책학회보, 4(2): 131-160.
- 황종성 (2017). 인공지능시대의 정부: 인공지능이 어떻게 정부를 변화시킬 것인가. *IT& Future Strategy*, 3.
- 행정안전부 (2017). 지능형정부 기본계획. 3.31.
- Barends, E., Rousseau, D. M., & Briner, R. B. (2014). *Evidence-Based Management: The Basic Principles*, Retrieved from The Center for Evidence-Based Management website: <https://www.cebma.org/>
- Chun, Y. H., & Rainey, H. G. (2005). Goal ambiguity in US federal agencies. *Journal of Public Administration Research and Theory*, 15(1): 1-30.

- Cluster. (Nov.19.2018). Machine Learning Lessons 5 Companies Share their Mistakes. Retrieved from <https://clusterdata.nl/bericht/news-item/machine-learning-lessons-5-companies-share-their-mistakes/>
- Cohen, M. D., March, J. G., & Olsen, J. P. (1972). A garbage can model of organizational choice. *Administrative Science Quarterly*, 17(1): 1-25.
- Cyert, R. M., & March, J. G. (1963). A behavioral theory of the firm. *Englewood Cliffs, NJ*, 2(4): 169-187.
- Davenport, T. H., & Kirby, J. (2015). Beyond automation. *Harvard Business Review*, 93(6): 58-65.
- De Martino, B., Kumaran, D., Seymour, B., & Dolan, R. J. (2006). Frames, biases, and rational decision-making in the human brain. *Science*, 313(5787): 684-687.
- Eggers, W. D., Schatsky, D., & Viechnicki, P. (2017). AI augmented government: using cognitive technologies to redesign public sector work. *Deloitte Center for Government Insights*.
- Fu, R., Huang, Y., & Singh, P. V. (2018). *Crowd, Lending, Machine, and Bias*. <https://doi.org/10.2139/ssrn.3206027>
- Goodwin, P., Wright, G., & Phillips, L. D. (2004). *Decision analysis for management judgment*. Wiley Chichester.
- Gruber, J. (2005). *Public finance and public policy*. Macmillan.
- HakemZadeh, F., & Baba, V. V. (2016). Toward a theory of collaboration for evidence-based management. *Management Decision*, 54(10): 2587-2616.
- Kitzhaber, J. A. (1993). Prioritising health services in an era of limits: the Oregon experience. *BMJ*, 307(6900): 373-377.
- Lindblom, C. E. (1959). The science of muddling through. *Public Administration Review*, 19(2), 79-88.
- Locke, E. A. (2009). *Handbook of Principles of Organizational Behavior*. John Wiley & Sons, Inc.
- March, J. G., & Simon, H. A. (1993). *Organizations. 1958*. NY: Wiley, New York.
- Ojo, A., Mellouli, S., & Ahmadi Zeleti, F. (2019). A Realist Perspective on AI-era Public Management. *20th Annual International Conference on Digital Government Research*, 159-170. ACM.
- O'Reilly, C. A. (1983). The use of information in organizational decision making: A model and some propositions. *Research in Organizational Behavior*, 5: 103-139.
- Rainey, H. G. (2009). *Understanding and Managing Public Organizations*. John Wiley & Sons.
- Reber, A. S. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology. General*, 118(3): 219.
- Rousseau, D. M. (2006). Is There Such a Thing As "Evidence- Based Management?" *Academy of Management Review. Academy of Management*, 31(2): 256-269.
- Rousseau, D. M. (2012). *The Oxford Handbook of Evidence-Based Management*. Oxford University Press.
- Rousseau, D. M., & Gunia, B. C. (2016). Evidence-Based Practice: The Psychology of EBP Implementation. *Annual Review of Psychology*, 67: 667-692.
- Rynes, S. L., Colbert, A. E., & O'Boyle, E. H. (2018). When the "Best Available Evidence" Doesn't Win: How Doubts About Science and Scientists Threaten the Future of Evidence-Based

- Management. *Journal of Management*, 0149206318796934.
- Simon, H. A. (1972). Theories of Bounded Rationality. In C. B. McGuire & R. Radner (Eds.), *Decision and Organization*. North-Holland Publishing Company.
- Tranfield, D., Denyer, D., & Smart, P. (2003). Towards a methodology for developing evidence-informed management knowledge by means of systematic review *. *British Journal of Management*, 14: 207-222.
- Weichselbaumer, D., & Winter-Ebmer, R. (2005). A meta-analysis of the international gender wage gap. *Journal of Economic Surveys*, 19(3): 479-511.
- Weimer, D. L., & Vining, A. R. (2017). *Policy analysis: Concepts and practice*. Routledge.
- Yates, J. F. (2003). *Decision Management: How to Assure Better Decisions in Your Company*. Jossey-Bass.
- Yates, J. F., & Potworowski, G. A. (2012). Evidence-Based Decision Management. In D. M. Rousseau (Ed.), *The Oxford Handbook of Evidence Based Management*, 198-222.

〈English Translation on the Korean References〉

- Kang, J.S. (2015). The Myth of Artificial Intelligence. Issue & Trend. September, 2015. *KISA*.
- Kuk, G.W. (2019). Artificial Intelligence Technology and Application Cases by Industry Sector. *Weekly Technology Trend*, 1888: 15-27.
- Sung, W.J., & Hwang, S.S. (2017). A Review of Intelligent Society Studies: A look on the future of AI and policy issues. *Information Policy*, 24(2): 3-19.
- Yoon, S.O. (2017). Policy Decisions in the Intelligent Information Age: Key Issues in Artificial Intelligence Policy Decisions. *Presented at the Korean Management Science Society*, 2447-2470.
- Yoon, S.O., Lee, E.M., Sung, W.J. An Analysis of the Types and Issues of Policy Decision Making Using Artificial Intelligence. *The Journal of Korean Association for Regional Information Society*, 21(1): 31-59.
- Lee, G.Y., & Kim, J. (2016). Artificial Intelligence Technology Trends and IBM Watson References in the Medical Field. *Korean Medical Education Review*, 18(2): 51-57.
- Chun, Y.H., (2004). Goal Ambiguity in Public Organizations: Concept, Measurement, and Validation. *The Korean Association For Public Administration*, 38(5): 49-65.
- Jung, J.G., Choi, J.W., Lee, S.W., & Jung, K.H. (2003). *Policy Sciences*. Seoul: Daemyung Press.
- Jin, S., & Park, Y. (2017). The Study on the Korean Future Readiness Policy for the 4 th Industrial Revolution. *Journal of Korean Association for Regional Information Society*, 20(3): 31-58.
- Choi, J. (1995). On the Concept of Rationality and Policy Studies. *The Korean Association for Policy Studies*, 4(2): 131-160.
- Hwang, J. (2017). Artificial Intelligence in Government: How AI Will Change Government, *IT & Future Strategy*, 3.

Ministry of the Interior and Safety. (2017). Basic Plan of Intelligent Government. March. 31.

김병조: 미국 카네기멜론대학교에서 정책학 박사학위를 취득하고, 현재 캘리포니아주립대 교수로 재직하고 있다. 주요 관심분야는 이민정책, 보건의료정책, 증거기반 의사결정, 공공종사자 사회연결망 분석 등이다. 주요 논문들은 Academy of Management Discoveries, Journal of Applied Behavioral Science 등 SSCI 저널에 게재되었다.(bkim12@csuchico.edu) **은종환:** 서울대학교에서 정책학 박사학위를 취득하고, 현재 방송통신대학교 공존협력연구소에서 전임연구원으로 재직하고 있다. 주요 관심분야는 미래인재, 정책이론, 사회정체성, 부패, 연구방법론 등이다. 주요논문으로 “Exploring Perceptions of Sustainable Development in South Korea: An Approach Based on Advocacy Coalition Framework’s Belief System”이 있다.(jo.eun3299@gmail.com)

투 고 일: 2019년 10월 27일

수 정 일: 2019년 11월 28일

게재확정일: 2019년 12월 23일

Incorporating Machine Learning into Public Administration: The Role of Evidence-Based Decision-Making

Kim, Byeong Jo & Eun, Jonghwan

The present study explores the policy implications of the introduction of machine learning into public administration. We first briefly overview the theoretical implications and practical limitations of a rational decision-making perspective, focusing on its unrealistic assumptions. We then discuss the capability of machine learning-based decision-making that may overcome the limitations of human decision-making by reviewing recent research comparing the decision performance of machine learning algorithms with human decisions in the context of crowdlanding. Our findings demonstrate that machine learning algorithms may outperform human decisions in structured problems, whereas human decisions are critical in unstructured problems such as meta-decisions (e.g., what to decide and how to decide). We draw three policy implications for the introduction of machine learning into public administration. First, public administrators and policymakers should determine the nature of the decision task; i.e., whether it is a structured problem or an unstructured problem that may require meta-decisions. Second, meta-decisions regarding the decision criteria (e.g., accountability, effectiveness) are required for machine learning algorithms aiming to resolve structured problems. Third, public administrators and policymakers should develop the ability and skills to implement evidence-based decision-making in order to make a high-quality meta-decision for unstructured problems.

Keywords: machine learning, evidence-based decision-making, rational model of decision-making