



캐나다, AI 책임에 관한 원칙

손민숙 연구원

요약

캐나다 금융감독원(OSFI)은 글로벌리스크연구소(GRI)와 함께 AI가 금융서비스 기관에 미치는 윤리적·법적·금융상의 영향에 대한 보고서를 발간함. 해당 보고서는 안전한 AI 개발을 위해 준수하여야 할 AI 책임에 관한 ‘EDGE 원칙’을 제시하였는데, ‘EDGE’는 각각 설명 가능성(Explainability), 데이터(Data), 거버넌스(Governance), 윤리(Ethic)를 의미함

- 캐나다 금융감독원(OSFI)¹⁾은 글로벌리스크연구소(GRI)²⁾와 함께 인공지능(AI) 관련 리딩 커뮤니티(FIFAI)³⁾를 구축하고 AI의 위험관리에 대한 논의를 진행하는 한편, AI가 금융서비스 기관에 미치는 윤리적·법적·금융상의 영향에 관한 보고서⁴⁾를 발간함
 - 현재 캐나다는 데이터의 가용성 및 접근성이 극적으로 향상되고, 모델링 기술이 크게 발전함에 따라 금융기관 전반에서 AI 모델이 적용되고 있음
 - AI 기술의 급속한 발전에 따라 OSFI는 GRI와 협력하여 학계, 규제기관, 은행, 보험, 연금 및 핀테크 등 금융서비스 전문가들로 구성된 AI 리딩 커뮤니티 FIFAI를 운영하고 있음
- 해당 보고서는 안전한 AI 개발을 위해 준수하여야 할 AI 책임에 관한 ‘EDGE 원칙’을 제시하였는데, ‘EDGE’는 각각 설명 가능성(Explainability), 데이터(Data), 거버넌스(Governance), 윤리(Ethic)를 의미함
 - ‘설명 가능성(Explainability)’의 설명 대상은 특정 예측 결과에 대한 각 변수의 중요도이며, 설명 수준은 수신인(데이터 과학자, 사업가, 규제기관, 고객 등) 및 적용되는 사례의 중요성(신용도 결정, 마케팅)에 따라 달라짐
 - AI 모델이 목적에 적합한지 확인하기 위해서는 모델 설계 초기에 설명 가능성이 고려되어야 하며, 모델 성능 검증 전 이론, 데이터, 방법론 및 접근 방식 등 기타 기본적인 요소들이 검토되어야 함
 - 설명 가능성은 AI 모델에 관한 적절한 정보를 제공하는 공시와 관련성이 높으며, 공시의 투명성과 의사결정 등에 대한 설명 가능성은 금융기관과 고객 간의 신뢰를 높이는 데 작용함
 - AI가 다루는 ‘데이터(Data)’에는 금융거래 및 고객 정보와 같은 구조화된 데이터(Structured data), 소셜 미디어 게시물, 이메일 또는 고객 리뷰와 같은 구조화되지 않은 비정형 데이터(Unstructured data), 인공적으로 생성된

1) Office of the Superintendent of Financial Institutions

2) Global Risk Institute

3) Financial Industry Forum on Artificial Intelligence

4) OSFI&GRI(2023. 4. 17), “Financial Industry Forum on Artificial Intelligence: A Canadian Perspective on Responsible AI” (<https://www.osfi-bsif.gc.ca/Eng/osfi-bsif/rep-rap/Pages/ai-ia.aspx>, Modified Date: 2023-06-29)

합성 데이터(Synthetic data)가 있음

- 새로운 데이터 소스 및 유형, 데이터양의 증가, 데이터 생성 및 활용 속도가 증가함에 따라 금융기관의 데이터 위험관리 및 제어 기능의 통합, 표준화가 어려워질 수 있음
- 특히 데이터가 조직 내 저장장치에 존재하거나, 서로 다른 외부 소스에서 들어오는 경우 어려움이 증가하므로 예상치 못한 방식으로 작동할 경우 제재할 수 있는 강력한 거버넌스가 필요함
- 금융기관은 사업 목적 달성을 위해 적절한 데이터를 수집·관리·분석하여야 하며, 데이터 전략을 사업전략에 접목해 데이터 전략이 사업전략과 일치하도록 하여야 함
- ‘거버넌스(Governance)’란 ‘감독’의 의미를 포함하나 감독보다 더 넓은 개념으로 의사결정, 책임, 관리, 위험 모니터링 및 완화, 성과 보고를 위해 조직이 마련한 구조, 시스템 및 관행을 의미함
 - AI 거버넌스 구조는 기관 규모 및 사업 거버넌스 프레임워크의 고도화 등에 따라 금융기관마다 차이를 나타냄
 - 일반적으로 거버넌스는 ① 모든 조직의 수준을 포괄하여야 하고, ② 역할 및 책임이 명확하여야 하며, ③ 위험 선호도의 정의 또한 명확하여야 할 뿐만 아니라, ④ 활용 사례의 위험을 반영하여야 하고, ⑤ 금융기관의 AI 도입 및 변화에 따라 유연한 대처가 가능해야 함
 - 금융기관은 AI 거버넌스를 위해 새로운 역할을 만들거나 기존 팀을 재구성하고, 구성원들이 새로운 기술과 흐름에 적응할 수 있도록 교육의 기회를 제공해야 함
- 사업 분야에서 ‘윤리(Ethic)’는 조직이 결정을 내리는 방식에 관한 도덕적 원칙과 가치를 의미하고, AI 윤리는 AI 기술의 개발 및 사용에서 도덕적 행위를 판단하기 위해 널리 인정되는 옳고 그름의 표준이 되는 일련의 가치, 원칙 및 기술로 정의⁵⁾할 수 있음
 - 정부에 의해 설정되는 법적 기준과는 달리 윤리적 기준은 법 이상의 원칙과 가치에 기초함
 - 윤리적 원칙과 가치는 상대적일 수 있고, 윤리적 기준은 조직 내 특정 용도에 따라 달라질 수 있으며, 시간이 지남에 따라 변화할 수 있음
 - 일반적으로 윤리는 법적 구속력이 없으나 AI 관련 원칙과 가치는 법으로 성문화(예: EU의 일반데이터보호규정; GDPR⁶⁾)되는 추세이며, 이후에도 이와 같은 추세는 계속될 것으로 예상됨

○ AI가 금융기관 및 고객에게 효용을 제공해왔고, 향후에도 금융기관은 더 많은 AI를 상품, 프로세스 및 의사결정에 포함시킬 것으로 예상되는바, 새로운 위험 및 심화된 위험의 통제를 위해 적절한 규제 마련이 요구됨

- AI의 원칙 및 규제에 관해서는 공통적으로 ① 개념이 적정할 것(명확성, 신뢰성, 확고함, 지속가능성), ② AI로 인한 문제(투명성, 책임)를 해결하는 적절한 거버넌스 구조를 갖출 것, ③ 설명 가능성은 중요한 의사결정(특히 고객에게 영향을 미치는 의사결정 포함)에서 필수적으로 고려될 것, ④ 인공지능은 개인과 사회에 해악(예: 편견, 차별, 윤리적 고려사항, 개인정보보호 문제)을 초래하지 않을 것 등의 요건을 요구함
- 규제는 혁신과 균형을 맞추는 것이어야 하며, 동시에 금융기관이 지속적으로 혁신을 추구하고 경쟁력을 유지할 수 있는 방향이어야 함

5) Leslie, David(2019), “Understanding artificial intelligence ethics and safety: A guide for the responsible design and implementation of AI systems in the public sector”, The Alan Turing Institute

6) General Data Protection Regulation