사고(思考)의 자동화 — 기회와 위험성

인공지능의 비즈니스 측면

정치, 과학 및 경제학은 다양한 비즈니스 영역에서 인공지능(AI)의 사용을 더욱 발전시키기 위해 전 세계적으로 수십억 달러를 투자하고 있습니다. 이와 동시에, 이러한 분야에서는 합리적인 기술 처리를 위해 법률, 표준 및 윤리 지침을 활용하여 변수를 개발하고자 하며, 사람들 사이의 거부감을 최소화하고 수용성을 높이기 위해 사회적 이익을 강조합니다. 오늘날 기술은 얼마나 발전했고, 가까운 미래에는 어떤 일들이 벌어질까요? 앞으로 인간은 결정 속도가 점점 더 빨라지고, 진행과정과 병행되는 학습 덕분에 더 다양한 결정을 내릴 수 있는 기계에 의사결정 권한을 넘겨주게 될까요?

목차

인공지능은 인간의 사고 과정과 뇌 안의 시냅스 연결을 모방하여 더욱 인간적인 사고와 이해 능력을 갖춘 알고리즘, 지속적으로 학습하는 알고리즘을 통해 진행되며 인간의 개입 없이도 작업을 완료할 수 있는 머신러닝, 인간과 기계 간의 협력, 로봇에 구축된 AI 시스템 등 다양한 기술을 포괄합니다. AI에 투자하는 기업들은 스스로 미래에 더 높은 수익을 거둘 것으로 기대합니다. 연구 및 개발 작업은 주로 사고, 지식 처리, 계획, 커뮤니케이션 및 인식과 같은 인간의 지능적 행동의 자동화에 초점을 두며, 여기에는 인간이 수행할 수 있는 모든 지적 작업을 완벽히 익힐 수 있는 시스템까지 포함됩니다.

기계, 장치 및 로봇에 탑재된 인간적인 지능에 대한 이 아이디어는 매우 먼 미래의 이야기처럼 들리며 따라서 추측의 성격이 강해 윤리적 및 사회적 측면의 분석을 위한 근거로 사용하기에는 적합하지 않습니다. 결국 기계는 과거의 데이터를 통해서만 학습하며 이러한 방식으로 미래를 상상하기에는 아직 부족한 점이 많습니다. 지능형 시스템은 여전히 이해와 의식을 갖추지 못한 상태입니다. 유연성, 창의성, 감성 지능 및 비판적 판단과 같은 인간의 능력은 (여전히) 기계와 차이를 보입니다. 과학자들은 인공 지능이 인간 수준에 도달하는 데 수십 년 또는 1세기가 걸릴 것으로 추정합니다.

따라서 현재 연구가 보다 먼 미래의 시나리오에 집중하고 있더라도, 인간과 기계 사이의 협력 및 각각의 강점이 이루는 시너지는 기술 사용의 최전선에 있습니다. 미래에 인간과 기계는 미래에 어떤 방식으로 협력할까요? AI는 인간의 의사 결정을 지원할까요, 아니면 스스로 결정을 내릴까요? 앞으로 몇 년 동안 실현될 가능성이 높은 시나리오는 다음과 같습니다. 인공 지능 시스템과 인간 전문가가 함께 일하면 혼자 일하는 경우에 비해 더 나은 결정을 내릴 수 있습니다. 어떤 경우에는 인간의 개입이 사라지게 될 것입니다. 따라서 자율 주행 선박은 독립적으로 위험을 피하고 선장 없이도 복잡한 도킹 조작을 완료할 수 있습니다. 센서와 컴퓨터 기술이 탑재된 지능형 로봇은 센서 데이터를 읽을 뿐만 아니라 해석하고 그 해석에 따라 동작을 수정합니다.

이러한 이유로 인공 지능은 현재 주로 머신러닝과 같은 특정 애플리케이션에서 구현됩니다. 이러한 경우 기계는 이전과 동일한 방식으로 프로그래밍 대신 기술을 학습함으로써 어느 정도의 지능과 자율성을 얻게 됩니다. 머신러닝의 목표는 다양한 소스의 데이터를 지능적으로 결합하고 관련 연결을 식별하며 결론을 도출하고 예측하는 것입니다. 예를 들어, 산업 현장에서는 기계 고장(예측 유지 보수) 또는 전문가 시스템을 사용한 기계 운영자 지원과 같은 특정 상황을 예측할 수 있으므로 운영 프로세스를 최적화할 수 있습니다. 가상 비서, 언어 인식, 자율 주행 자동차 등과 같은 AI 애플리케이션은 모두 사람의 지시에 따라 움직입니다. 우려의 시각도 존재하지만, AI가 제공할 수 있는 기회와 장점은 분명합니다. 그럼에도 불구하고 이렇게 지능적인 기술을 사용하여 과도한 기계 자율성 또는 (데이터 및 의사 결정에 대한) 오용을 어떻게 예방하거나 배제하고 윤리적 조치와 보안은 물론 사적 영역에 대한 보호를 보장하려면 어떻게 해야 할까요?

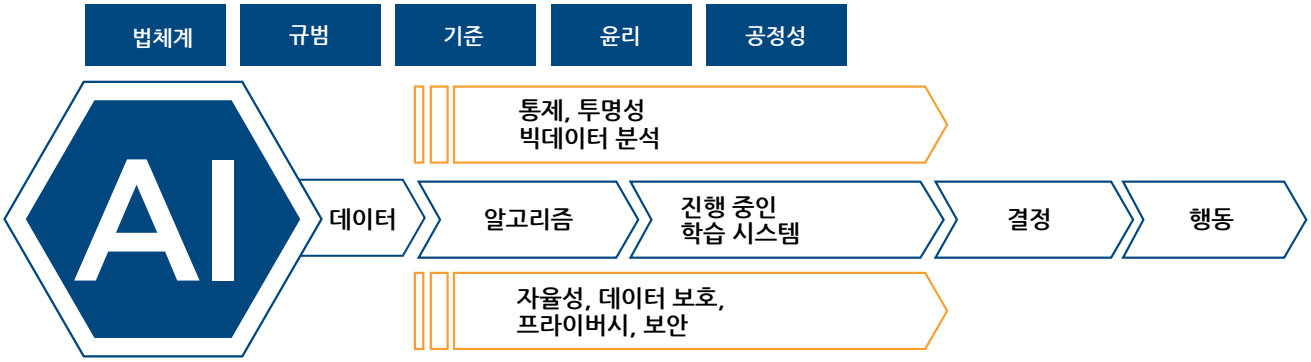
투명성 대 데이터 보호, 통제 대 자율성

불투명한 블랙박스 알고리즘과 자동화된 의사 결정은 많은 사람들의 두려움을 증폭시킵니다. 사람들은 자신에게 즉각적으로 영향을 주거나 관련이 있는 기술적 의사결정 앞에서 무력감을 느낍니다. 자율 무기 시스템 사용, 인간에 대한 불투명한 감시, 신용등급 및 자율 주행 차량의 오작동에 관한 뉴스 보도는 부정적인 예시에 해당합니다. 기차역에서 대규모 안면 인식을 시도한 사례나, Amazon에서 친구를 위해 쇼핑할 경우 자신의 프로필이 변경되고 이를 수정할 수 없는 사례와 같은 오류가 계속 누적된다면 부정적인 뉴스 보도는 더욱 늘어날 것입니다.

이와 대조적으로, 훈련된 지능형 시스템이 초빙된 피부과 전문의 58명 대부분보다 더 나은 결과를 제공하는 의학 분야의 사례 처럼 유용한 애플리케이션도 많이 찾아볼 수 있습니다. 방사선 정보학의 발전 또한 매우 획기적입니다. 방사선 데이터는 분자 생물학적 데이터 및 기타 임상 표지와 상관 관계가 있으며 알고리즘을 통해 해석됩니다. 의학 분야의 인공 지능은 미세 전이 또는 정신병, 우울증, 암, 알츠하이머와 같은 심각한 질병의 조기 진단을 위한 중요한 지원 도구입니다.

이전에는 알고리즘이 의사결정으로 연결되는 데이터의 기본 토대를 형성했습니다. 알고리즘은 인공 지능을 사용하여 의사 결정의 기반이 되는 복잡한 양의 데이터에서 패턴을 찾아내고 높은 확률로 예측을 수행합니다. 알고리즘은 이전에 인간이 발견하지 못했던 새로운 인과관계를 찾을 수 있습니다. 알고리즘은 이미 개인이 현금 지급기에서 돈을 인출할지 여부를 예측하고 있습니다. 또한 알고리즘은 소셜 미디어에서 얼굴을 인식하고 스마트폰에서 작업을 실행하기도 합니다. 소비자 보호를 위한 인공 지능은 어떤 형태일까요? 개인 데이터와 특성이 데이터 처리에 사용되고 연결되는 사례가 증가하고 있습니다. EU 기본 데이터 보호 규정은 의사 결정의 투명성을 요구합니다. 따라서 대출이 허가되지 않은 이유를 물을 때 “기계가 결정을 내렸다”는 대답은 충분한 이유가 될 수 없습니다. 자동화된 결정이 사람이나 물체에 불리한 영향을 끼칠 때 근본적으로 책임 문제가 발생합니다. 이러한 경우에는 AI 시스템에 법적 신분이 할당되지 않고 제조업체가 책임을 져야 합니다.

AI의 자율성의 정도는 주로 사용 목적, 개발자가 정의한 한도, 법적 및 운영 요구 사항, 물리적 프로세스 및 기술 규범에 의해 제한됩니다. 특히 실제 제품에 내장된 기능은 제조업체가 정의한 안전 개념 및 예상된 제품 특성의 틀 안에서만 작동합니다. 따라서 산업 환경에서 윤리적 측면은 상대적으로 부수적인 역할을 수행합니다. 통제가 불가능한 기계, 드론 또는 로봇 상태는 제품을 통제하고자 하고 이에 대한 책임을 지는 제조업체에게도 바람직하지 않습니다. 따라서 통제가 불가능한 독립적인 기계에 대해 프랑켄슈타인과 같은 공포 가득한 시나리오는 지나친 것으로 보입니다. 결과적으로 제한사항은 자동으로 설정되고 동시에 여유 공간이 생성되어 기술과의 합리적인 상호 작용이 가능합니다.



이미지 1: AI를 위한 법적 프레임워크는 반드시 균형을 추구해야 합니다.

법적 및 윤리적 프레임워크의 필요성

규범과 표준뿐만 아니라 법적 프레임워크도 마련되어야 합니다. 2019년, 유럽 집행위원회의 AI 고급 전문가 그룹은 AI에 대한 유럽 전략을 마련하고 공정성, 보안, 투명성, 미래의 작업 환경, 민주주의와 같은 주제에 대한 윤리적 지침을 제시할 계획입니다. 인공 지능을 위한 가치 시스템의 형태는 수많은 국제협회 및 표준화기구로 구성된 자율 및 지능형 시스템 윤리를 위한 개방형 커뮤니티(OCEANIS, https://ethicsstandards.org)에서 논의됩니다. 회원들은 철학, 신학, 심리학, 사회학 전문가들과 함께 윤리적 인공 지능을 만들어낼 요구 사항 목록을 작성할 계획입니다.

이와 유사한 맥락에서, 2017년 5월 유럽 경제 사회위원회(EESC) 성명서는 "(디지털) 단일 시장, 생산, 소비, 고용 및 사회에 대한 인공 지능의 영향"을 집중적으로 다루었습니다. 성명에서 위원회는 인공 지능의 개발, 구현 및 사용을 위한 행동 강령을 명시적으로 홍보합니다. 일부 기업은 이미 이러한 규범을 정의하고 있으며, 외부 전문가로 구성된 모니터링 기구 형태로 윤리위원회를 실현하며 해당 문제를 처리하는 직원을 대상으로 윤리 교육을 실시하고 있습니다. 이러한 경우에는 윤리적 책임을 부담하는 데이터 처리가 가장 우선시됩니다.

윤리적 관점에서 AI 시스템이 인간의 존엄성, 안전, 독립성 및 선택의 자유에 미치는 영향을 규제되어야 합니다. 따라서 AI 기반 안면 인식은 개인 정보 보호 및 표현의 자유와 같은 기본적인 인권에 위협이 될 수 있습니다. 데이터 보호는 특히 데이터가 인간의 결정에 영향을 미치거나 새로운 (비)진실을 사실이라고 주장하며 인간의 행동을 조작하는 데 사용되는 상황에서 특히 중요합니다. 기본 데이터가 정확하고 편견이나 선호도를 반영하지 않는 것은 중요합니다. 보안 측면에는 이러한 기계의 신뢰성과 모든 상황에서의 안전한 기능이 포함됩니다. 또한 AI 시스템의 의사 결정에 대한 투명성과 이해도는 일반적인 수용을 위해 매우 중요합니다. AI 시스템의 운영 원칙, 행동 및 결정은 항상 접근 및 검증이 가능해야 합니다.

시스템은 작동 중에도 학습하며, 시스템이 언제나 특정 학습 결과가 어떻게 생성되었는지 이해하는 것은 아닙니다. 데이터 처리의 결과를 더 이상 예측할 수 없는 경우 그 결과로 인해 손해가 발생하거나 법률 또는 데이터 개인 정보 위반이 발생할 수 있습니다. 또한 시스템 개발자의 의도가 없는 경우처럼 개인을 차별할 수 있는 AI 결정에 대해 공정성 및 형평성을 확보해야 한다는 요구 사항도 존재합니다. 이러한 이유로 인해 사전적 위험 평가를 수행해야 하며 기본 코드 및 알고리즘의 공개 없이 중립적인 제 3자가 진행하는 알고리즘 기반 애플리케이션에 대한 지속적이고 포괄적인 평가가 보장되어야 합니다.

결과

법적 프레임워크와 폭넓은 사회적 수용은 인공 지능의 지속적인 성공을 위한 전제 조건입니다. AI 의사 결정의 기초에 대한 이해도를 높이기 위해 노력하는 해석 가능한 AI에 대한 연구는 신뢰도를 강화할 수 있습니다. 미래에 규범과 표준은 보안, 투명성, 이해도 및 윤리적 정당성과 관련하여 AI 시스템을 확인, 검증 및 통제하기 위해 유연한 규제 프레임워크를 정의하여 소셜 미디어를 통해 선거 결과에 영향을 미치는 것과 같은 지나친 활용을 방지하고, 필요한 경우 초지능형 기계를 비활성화할 수 있도록 해야 합니다. 지침이 되는 정책의 원칙은 인간이 언제나 기계에 대한 통제권을 유지하는 인간 지시형 인공 지능을 위한 기반이 되도록 가능한 한 빨리 국제적 수준에서 수립되어야 합니다.

흥미로운 링크:

OCEANIS:
https://ethicsstandards.org

EESC 입장 관련 문서:
https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52016IE5369&qid=1558609151835&-from=EN

Basler AG
독일, 본사
전화: +49 4102 463 500
팩스: +49 4102 463 599
sales.europe@baslerweb.com
www.baslerweb.com

Basler Asia Pte Ltd.
싱가포르
전화: +65 6367 1355
팩스: +65 6367 1255
sales.asia@baslerweb.com

Basler Korea
대한민국
전화: +82 31 713 3114
sales.korea@baslerweb.com