AI 이노베이션 레포트(21600685\_조예성)

**□ 동영상 강의 요약**

영상에서 말씀하셨듯, 기업들이 AI를 도입해야하는 가장 큰 이유는 앞으로 시장을 이끌어갈 주요 기술과 산업이기 때문일 것이다. 그러한 이유 중 가장 큰 것은 AI는 다른 산업과 달리 축적된 데이터를 기반으로 ‘학습’을 하는 기능, 그리고 그로 인해 인간과 비슷할 만큼의 고도의 정확성을 갖는 판단 능력, 그리고 인간을 훨씬 뛰어넘는 일 처리 속도 때문일 것이다.

많은 사례들을 분석하여 주셨지만, 그 중 중점적으로 의료 분야에 널리 쓰이고 있는 IBM의 왓슨과 자율 주행 자동차 2가지 사례를 언급하고자 한다. IBM왓슨은 패턴 인식 기반의 기술을 이용하여 이미지를 정확하게 판독하는 시스템으로 해당 정확도가 대장암의 경우 98%, 방광암의 경우 97% 등 거의 100%에 육박한 정확도를 보이고 있다고 해 평균적으로 암 진단의 20%의 오진을 범하는 전문 의사들에 비해 훨씬 더 좋은 판단을 내리고 있다. 이러한 기업들이 개발한 인공지능 알고리즘은 대중에도 공시가 되어 이를 이용하여, 해당 산업 분야가 아닌 다른 분야에도 널리 쓰이고 있는 경우가 많다. Baker$Hostertlier의 AI 변호사 Ross의 경우 IBM 왓슨의 알고리즘을 활용한 법률자문 솔루션을 개발한 것이 그 예라고 할 수 있다. 인공지능 알고리즘 자체의 개발도 중요하겠지만 그보다, 어쩌면 그 이상으로 중요한 것은 이미 공시된 알고리즘을 적절히 활용하여 다양한 분야에 접목시키는 것도 현재로서는 중요한 인사이트가 될 수 있다. 하지만, 이렇듯 방대한 양의 데이터를 학습해 인간과 비슷한, 어쩌면 더 뛰어난 정확도와 속도로 분석하고 판단하는 인공지능은 많은 문제에 직면해있는데, 교수님께서 말씀하셨던 5가지의 문제점 중 가장 대표적으로 이야기하고 싶은 문제점은 바로 ‘데이터의 편향성’이다. 데이터과학을 공부하거나 여러 공모전을 나가 직접 머신러닝 모델을 구현하다보면, 가장 어려운 부분 중 하나가 바로 데이터의 전처리, 편향성을 다루는 방법인데, 매우 방대한 데이터를 다룬다 한들 데이터가 편향이 되어있을 지 아닐 지는 쉽게 예단할 수가 없는 법이다. IT 기업에 지원하는 지원자의 특징을 분석하는 데이터에서 대부분의 지원자들이 남성인 경우, 데이터의 양이 많아도 데이터는 필히 편향될 수밖에 없다. 이것이 같는 위험성은 생각보다 거대한데 AI 자율주행자동차의 경우를 예로 들 수 있다. 대개 학습은 대형 도시의 길들에 직접 주행을 해봄으로서 학습을 하겠지만, 만일 이러한 AI 자율주행자동차가 상용화되어, 낮선 시골길이나 오프로드로 들어서게 되면, 이 자율주행 기능은 학습되지 못한 상황에 놓여 위험한 상황을 초래할 수도 있게 된다. 물론 수많은 기업들은 이러한 AI의 이면성에 도전하는 차원으로 여러 시도들을 펼치고 있다. 왓슨의 경우 AI OpenScale을 통해 알고리즘의 투명성을 높여 실제로 오류나 데이터의 편향성을 시시각각 고치려는 노력을 쏟고 있다. 하지만, AI가 온전히 우리 산업의 전반적인 분야에 적용되기 위해서는 데이터의 편향성을 온전히 극복할 수 있는 방법이 요구될 것이다. 또한 아직 사람이 명확히 파악하기 어려운 딥러닝의 작동 방법을 사람이 실시간으로 분석할 수 있는 수학적, 통계학적 백그라운드가 온전히 자리잡혀야만 할 것이다.