

Seminar for Intellectual Exchange

SIE BRIEF

SIE(Seminar for Intellectual Exchange)는 산업정책연구원(IPS)이 정기적으로 주관하여 전 구성원들이 지식을 공유하는 연구 세미나입니다.

Contents

1. 빅데이터·AI·기계학습의 이해
2. 빅데이터 시대의 디지털전환
3. aSSIST AI·빅데이터 MBA 및 박사과정 소개
4. aSSIST 비주얼 파이썬 개발

빅데이터의 이해와 aSSIST AI MBA 소개

김진호 서울과학종합대학원 AI·빅데이터 MBA 교수

1. 빅데이터·AI·기계학습의 이해

빅데이터와 데이터 기반 의사결정

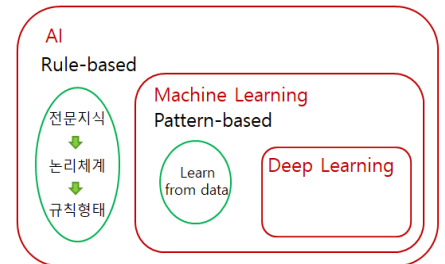
빅데이터의 핵심은 다양한 대규모 데이터를 활용하여 작은 규모의 데이터에서는 추출이 불가능했던 패턴이나 상관성을 파악하고 인사이트를 추출하여 의사결정에 활용하는 것입니다. 데이터 기반 의사결정이란, 지도학습을 바탕으로 최적의 통계/기계학습 모델을 선정·적용하여 사건 발생 가능성을 계산하고 이를 기반으로 의사 결정을 하는 것을 말합니다. 이 과정에서 AI와 기계학습은 데이터 기반 의사결정의 Tool이라고 할 수 있습니다.

인공지능(AI), 기계학습(Machine Learning), 딥러닝(Deep Learning)

가장 넓은 범위인 인공지능(AI)은 Rule-based로, 전문지식을 논리체계를 통해 규칙 형태로 만드는 것입니다.

기계 학습 (Machine Learning) 은 Pattern-based로, 데이터로부터 체계를 학습하여 패턴을 도출하여 문제를 해결하게 됩니다. 기계학습은 명시적으로 프로그래밍 하지 않아도 컴퓨터가 학습할 수 있는 능력을 갖게 하는 것입니다. 기계 학습에서 학습이란 데이터를 분석하여 일반화된 법칙을 이끌어내는 것으로, 데이터가 많을수록 보다 정확한 결과를 도출하게 됩니다.

인공지능-머신러닝-딥러닝 체계



<출처> 김진호, 빅데이터의 이해 발표자료

초기 머신러닝은 단순한 알고리즘인 Shallow 구조로, 연구자가 생각한 명확한 특징 변수를 머신러닝 모델에 입력하는 구조로서 변수 선택과 모델이 분리되어 있었습니다.

딥러닝이란 이러한 기존의 단순한 구조를 여러 층으로 다층화한 '심층 인공 신경망'을 활용하여 고도화된 결과를 도출하는 기술로서, 변수 선택과 머신러닝 모델이 일체화된 형태입니다. 딥러닝의 대표적인 사례로는 자율운행 자동차 기술 등이 있습니다.

2. 빅데이터 시대의 디지털 전환

빅데이터 시대의 차별적 경쟁우위

기업의 경쟁우위는 비즈니스의 다양한 영역에서 경쟁사 대비 1%를 증대시키거나 감소시키는 것입니다. 예를 들어 원가, 효율, 문제 탐지, 새로운 기회 포착 등의 측면에서의 경쟁력이 필요하며, 빅데이터 시대에는 이러한 경쟁우위를 달성할 수 있는 수단이 데이터 분석입니다. 데이터 분석을 통해 고객과 시장에 대한 인사이트를 추출하고, 그것을 바탕으로 현명한 의사결정을 할 때 지속 가능한 경쟁 우위를 창출할 수 있습니다.

하버드의 마이클 포터 교수는 비교적 낮은 비용, 무관심한 틈새시장, 쉽게 모방할 수 없는 차별성을 기업 경쟁력의 3가지 원천으로 지목했는데, 3가지 원천의 바탕 역시 데이터 분석에서 나옵니다. 시장과 고객에 대한 데이터의 정밀한 분석을 통해 비용 절감과 새로운 시장을 발견할 수 있으며, 차별성 역시 데이터 분석의 경험과 기계학습 알고리즘에서 나옵니다. 이처럼 기업은 빅데이터 시대에 대응하기 위해 데이터 분석을 전략적으로 활용하여 새로운 차별적 경쟁우위를 구축해나가야 합니다.

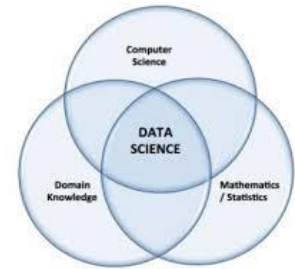
디지털 전환(DT)

디지털 전환(Digital Transformation)이란 소셜, 모바일, 사물인터넷, 클라우드, 빅데이터 등 5대 핵심기술을 도입하여 비즈니스를 혁신하는 것입니다. 빅데이터 시대에는 모바일, 센서, 소셜미디어 등의 데이터가 폭증하고, 이러한 빅데이터를 저장·처리·분석하기 위한 오픈소스 프로그램과 클라우드 서비스 등의 기능 및 접근성이 크게 개선되었습니다. 이러한 환경 변화와 미래 기술을 자신의 사업을 혁신하는 새로운 도구로 활용하는 것이 점차 중요해지고 있습니다.

3. aSSIST AI·빅데이터 교육 과정_ '데이터 사이언티스트' 양성

데이터 분석 프로세스는 문제 정의→데이터 수집·처리→분석/모델링→해석→전달→의사결정 순으로 이루어집니다. 영역별로는 IT 영역, 통계 영역, 현업 영역으로 나누어 볼 수 있으며, 실무적으로 문제를 처리할 때 영역 간의 GAP을 극복하는 것이 중요합니다. 따라서 각 영역을 아우르며 데이터를 다각적으로 이해, 처리, 활용이 가능한 '데이터 사이언티스트(데이터 과학자)'의 역할이 중요합니다.

| IT 영역 | 통계 영역 | 현업 영역 |
|-----------------------|---|--|
| · 필요한 입력 데이터 수집 및 전처리 | · 입력 데이터 지정 · 분석기법 적용 · 최적 모델 선정 · 결과 도출 | · 문제 정의 · 필요한 결과 탐색 · 분석 기법 결정 · 분석 결과 해석 · 활동 전개 및 평가 |
| 데이터 엔지니어 | 데이터 애널리스트, 머신러닝 엔지니어 | 도메인 엑스퍼트 |
| 데이터 사이언티스트 | | |



(출처) 김진호, 빅데이터의 이해 발표자료

서울과학종합대학원(aSSIST)은 차별화된 커리큘럼 및 종합 실습을 통해 데이터 사이언티스트를 양성하는 aSSIST AI·빅데이터 MBA 및 박사과정을 운영하고 있습니다. 본 과정은 데이터 분석 및 코딩을 잘 하며, 도메인 특징을 잘 아는 현업 담당자들과의 협업을 통해 문제를 제대로 정의하고 전체 분석 프로젝트를 리드할 수 있는 실전형 전방위 데이터 사이언티스트 양성을 목표로 하며, 실제 현업에 많은 데이터 전문가를 배출하고 있습니다.

4. aSSIST 비주얼 파이썬 개발

빅데이터 시장이 빠르게 성장하면서, 변화하는 비즈니스 환경에 대처하기 위해서는 실시간으로 데이터 흐름을 수집·분석하는 것이 보다 중요해졌습니다. 이에 기업이 IT조직에 의존하지 않고 데이터 기반 인사이트를 창출하는 '셀프 서비스 BI'의 필요성이 높아짐에 따라, 서울과학종합대학원(aSSIST)에서는 개발자/비개발자 모두가 별도의 학습 없이 직관적으로 사용할 수 있는 GUI 기반 빅데이터 분석 솔루션 'aSSIST Visual python'을 개발하고 있습니다. 오픈소스 SW 파이썬을 기반으로 쉬운 인터페이스와 커뮤니케이션이 용이한 셀프 서비스 BI로서, 향후 기업 솔루션, 빅데이터 분석 교육 도구 등으로 상용화할 예정입니다.

산업정책연구원(IPS)은

각계 각국의 산업정책에 대한 학술연구, 정부에 대한 정책자문, 기업 전략 개발 등 다양한 분야의 연구를 진행하고 있으며, 26년 역사의 지속가능경영, 브랜드 및 디자인경영, 경쟁력 연구를 통해 대한민국의 지속가능한 경쟁력 강화하고 있습니다.