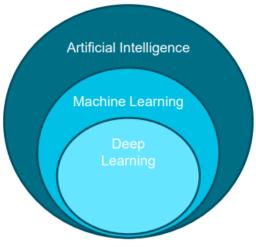
# Al @ Finance

#### 개요

- AI, ML 및 DL 개요
- AI 및 머신 러닝 @ Finance
- 자연어 처리
- 자연어 처리 @ 금융
- 대체 데이터 처리

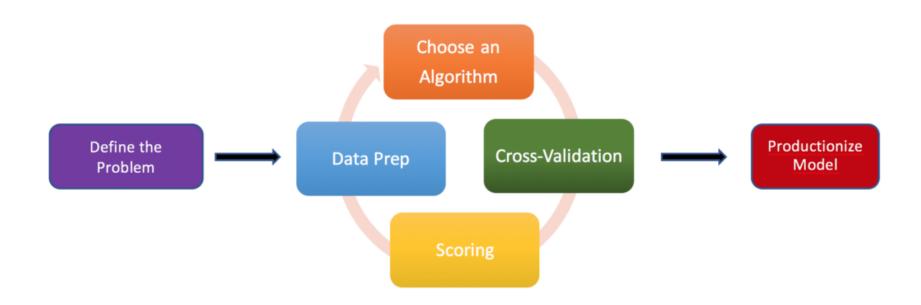
#### AI,ML, and DL



From: Towards Data Science

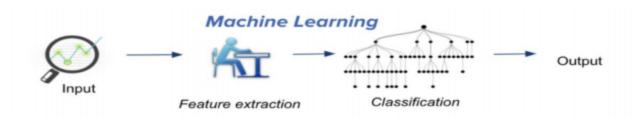
- AI (1950년대): 인간 지능 프로세스 모방학습, 추론 및 자기 수정을 포함한 기계체스 게임으로 응용 프로그램을 시작한 다음다른 모든 AI 젊은이들의 이론적 기반
  - ML(1980년대): 수학과 수학을 통합한 AI의 하위 집합 애플리케이션에서 데이터를 학습할 수 있는 통계 애플리케이션 분류 및 클러스터링
  - Deep Learning(2010년대): ML에 대한 혼합에 다층 신경망의 개념을 추가 응용 프로그램은 세계 챔피언과 바둑 게임으로 시작
     자가 운전을 위한 오디오, 비디오 처리

### **Machine Learning Work Flow**

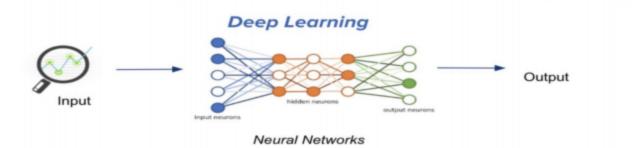


- 1. 데이터와 소스를 준비(기술 준비)
- 2. 모델을 코드화
- 3. 모델을 교육, 평가하고 조정
- 4. 교육받은 모델을 사용
- 5. 모델에서 예측한 것을 가져옴
- 6. Monitor the ongoing predictions ; 진행 중인 예측을 모니터링
- 7. 모델과 버전을 관리

### **Deep Learning**



Traditional machine learning uses hand-crafted features, which is tedious and costly to develop.



Deep learning learns hierarchical representation from the data itself, and scales with more data.

#### ML vs. DL

. ما =	— Marchine Leavine	= Paradamaia	_ ,,,,,
Aa 이름	■ Machine Learning	■ Deep Learning	를 비고
<u>방법</u> ( <u>Methodology)</u>	규칙 기반의 '데이터 중심' 프로세스	규칙 기반 데이터 중심 프로세스로 신경망을 활용	
분석 목표 (Analysis Goal)	최소한의 인적 자원 개입을 통한 의사결정 & 데이터의 패턴 및 관계(Decision making with minimum human intervention Patterns and relationships in data)	최소한의 인적 자원 개입을 통한 의사결정 & 데이터의 패턴 및 관계(Decision making with minimum human intervention Patterns and relationships in data)	
<u>사용/ 강도</u> ( <u>Uses/Strength)</u>	분류, 클러스터링에 적합(Good for classification and clustering)	분류, 클러스터링에 적합, 이미지, 비디오, 오디오 데이터에 대한 분석 개선, 새 데이터를 샘플링/ 생성 가능(Good for classification and clustering Better analysis on images, video, audio data Can sample/create new data points)	
<u>투명성</u> ( <u>Transparency)</u>	보다 투명한 모델, 모델의 의사결정을 쉽게 설명 가능 (More transparent: able to more easily explain decisions made by the model)	더 불투명함, 모델이 결정을 내린 이유를 설명하기 어려움 (More opaque: difficult to explain why the model made a given decision)	
HW 요구사항 ( <u>Hardware</u> Requirements)	표준 하드웨어 사용	큰 데이터를 처리할 수 있는 능력, 고가의 하드웨어(GPU) 필요	
데이터 요구사항 ( <u>Data</u> <u>Requirements)</u>	구조화, 레이블이 지정된 데이터 Use labeled data/Feature	사전에 레이블 되지 않은 데이터	머신러닝 안에 딥러 닝 이라는 분야가 있 는 것

# 금융 분야의 ML/DL 어플리케이션 (ML/DL Applications in Finance field)

Robo Advisor

Al @ Finance 2

자동화 포트폴리오 관리 서비스, 낮은 자문료

· Algorithmic trading

ai를 이용한 모든 주식 거래 결정을 하는 서비스 운영'헤지펀드', 일반 헤지 펀드, 퀀트 펀드보다 우수한 성능

• Fraud Detection

자금세탁 및 대 태러 금융 솔루션, 정확한 위험 결정 및 안전한 신원 확인

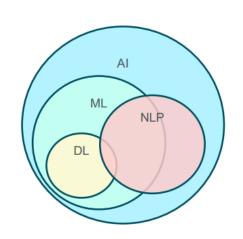
Underwriting

인수 및 신용 점수에 대한 자동화된 작업 수행, 인력 투자 없이 신속하게 신용 모델링

Risk Management

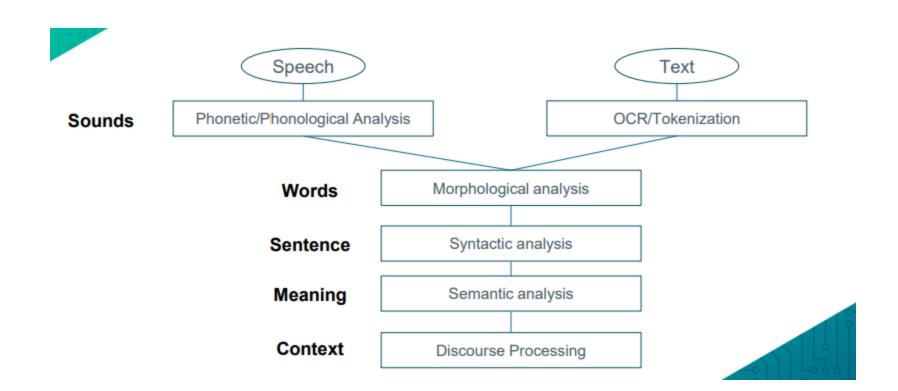
공개적으로 사용가능한 데이터 내에서 발생하는 큰 영향을 주는 이벤트와 새로운 리스크의 초기 신호를 감지, 중요 이벤트 및 위 반 정보를 뉴스 발표 훨씬 전에 검색

### 자연어 처리 (Natural Language Processing)



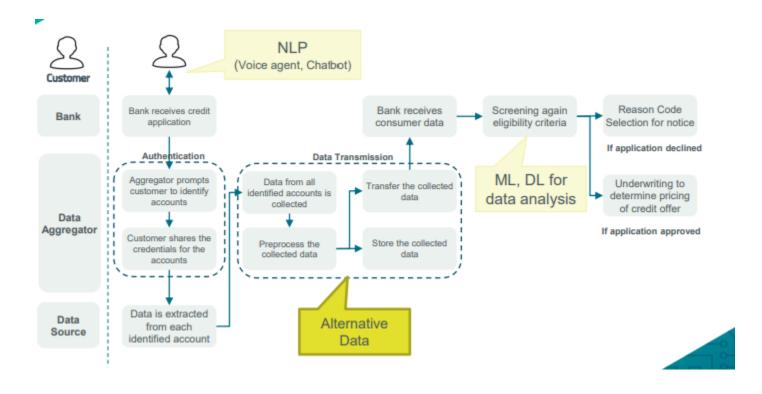
#### 인공지능의 하위 분야

- 인간이 인간언어로 컴퓨터와 대화할 수 있도록 허용
  - 。 용도: 기계번역, 문자구분, 음성 에이전트



#### **Alternative Data**

Al @ Finance 3



### **Conventional Underwriting Cash-flow Data**

Source	Underwriting Data Type		
Bank and Prepaid Accounts (Direct Deposit Accounts,	Financial Data  ➤ Inflows (Paystubs, direct deposits, etc.)  ➤ Outflows (Expenses)  ➤ Fees (Overdraft)	Inferred¹ Financial Data  ➤ Monthly balance / cushion  ➤ Income  ➤ Balance history	
Prepaid Cards)	Non-Financial Data  ➤ Time of day transaction made  ➤ Merchant identity or purchase type	Inferred Non-Financial Data  > Frequency (of deposits, transactions)	
Small Business Sources (Payment processing records, E-Commerce platforms,	Financial Data  ➤ Business credit card transaction data  ➤ Accounts receivable  ➤ Checking account data	Inferred Financial Data  ➤ Gross/Net profit margin  ➤ Income  ➤ Business efficiency	
Accounting Software)	Non-Financial Data ➤ Business inventory	Inferred Non-Financial Data  > Business efficiency (inventory turnover)	

# 대체 데이터의 예(Alternative Data & Concerns)

- 재무상태에 관한 추가 정보 제공
  - 고용 상태, 소득, 고용주, 직업 유형, 근무 이력, 보험, 임대료 및 공공요금 납부 내역 등
- 온라인 활동
  - 앱 사용, 위치, 소셜 미디어 데이터, 온라인 쇼핑 등
  - ∘ IoT 데이터 : 자율 주행 자동차 모니터링, 연습 추적 등
- 우려사항

프라이버시, 공정한 대출과 신용, 정확성 및 설명 가능성 투명성 등 제공