

OCUINSIGHT⁺

Tech Education



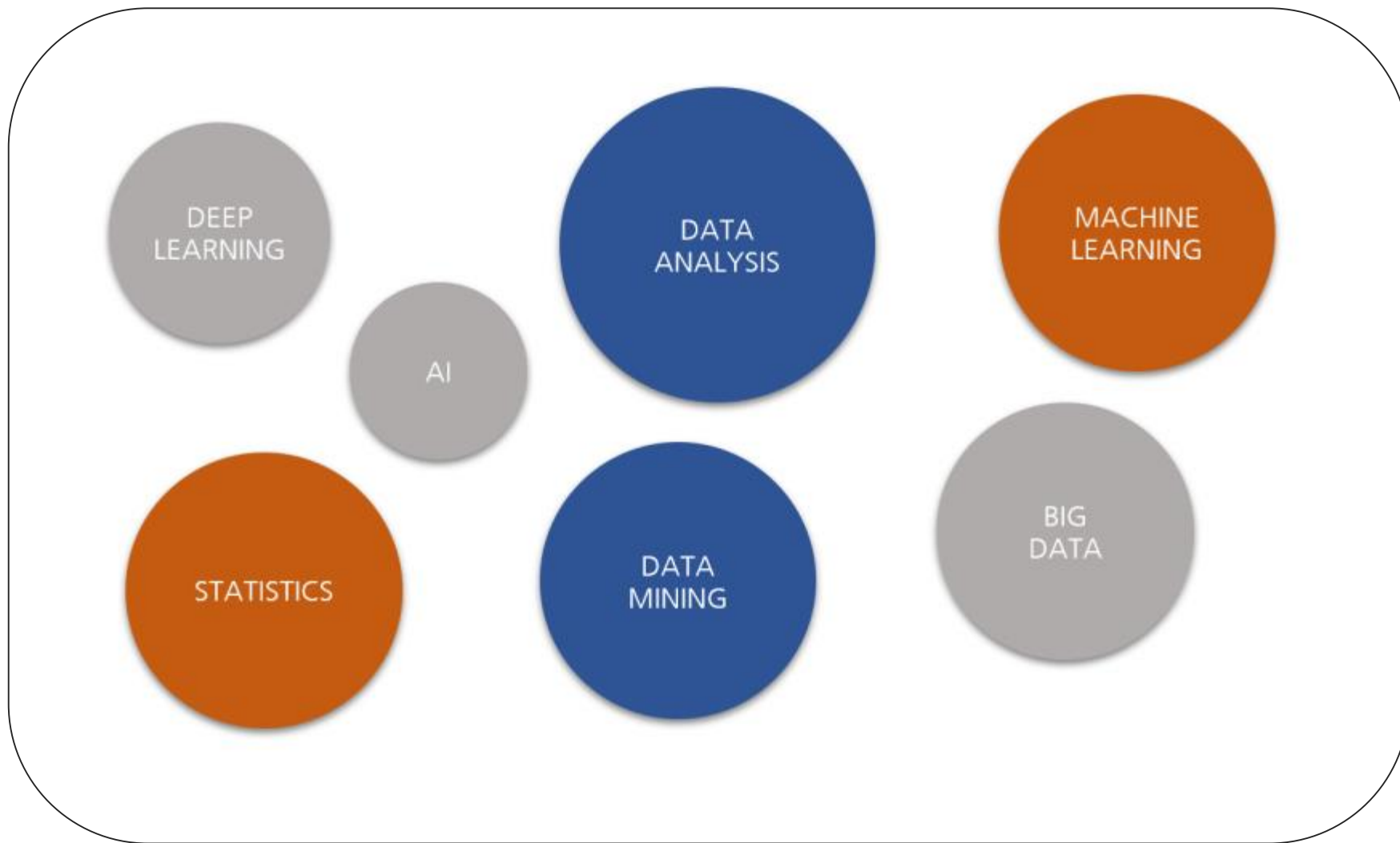
Digital Tech. 센터, AI/Data 그룹
강사 - 김병태매니저
2023.05



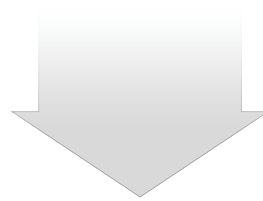
CONTENTS

데이터 분석 및 머신러닝 기초 강의 자료

4차 산업혁명이라는 트렌드에 따라 수많은 뉴스에서 등장하는 단어들



데이터 분석 (Data analysis)은 유용한 정보를 발굴하고 결론 내용을 알리며 의사결정을 지원하는 것을 목표로 데이터를 정리, 변환, 모델링하는 과정. 데이터 분석은 방대한 데이터에서 필요한 정보를 추출하여 과제 대상의 내재된 법칙을 찾아내게 하는 기술



기업에 수익을
극대화

업무 전체의
효율성을 높여줌

기업 리스크를
줄여줌

비용의 지출을
억제하고, 원가
절감 작용

기업에서의 데이터 분석 필요성

프로세스와 서비스에 대한 더 깊이 이해

고객 경험과 고객 문제에 대해 상세한 인사이트를 제공

데이터를 넘어 패러다임으로 전환

개인화된 고객 경험을 만들고 관련 디지털 제품을 구축하고 운영 최적화

더 나은 비즈니스 결정을 내리도록 돕기 위해 사용

업무 효율성을 높이고 비즈니스 리스크 관리를 강화

고객 이해도와 만족도를 높이고 마케팅 최적화

머신러닝 - 과거의 데이터를 학습하여 미래를 예측하는 기술

Human Learning



학생



기출문제



풀이법



시험

Machine Learning



기계



데이터



ML 알고리즘

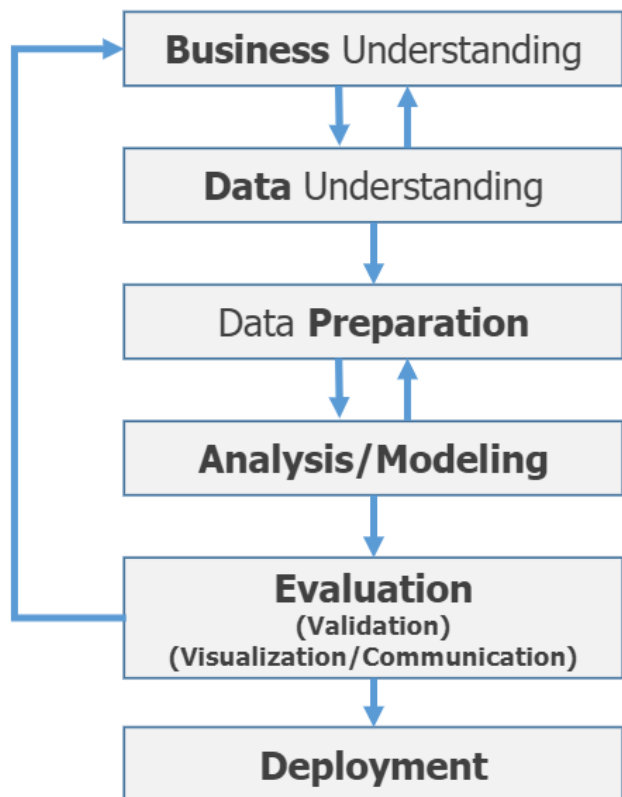


예측

인간 학습과 기계 학습은 유사. 인간이 문제의 풀이법을 익히는 것처럼 머신러닝도 알고리즘을 통해 예측

기업에서의 머신러닝

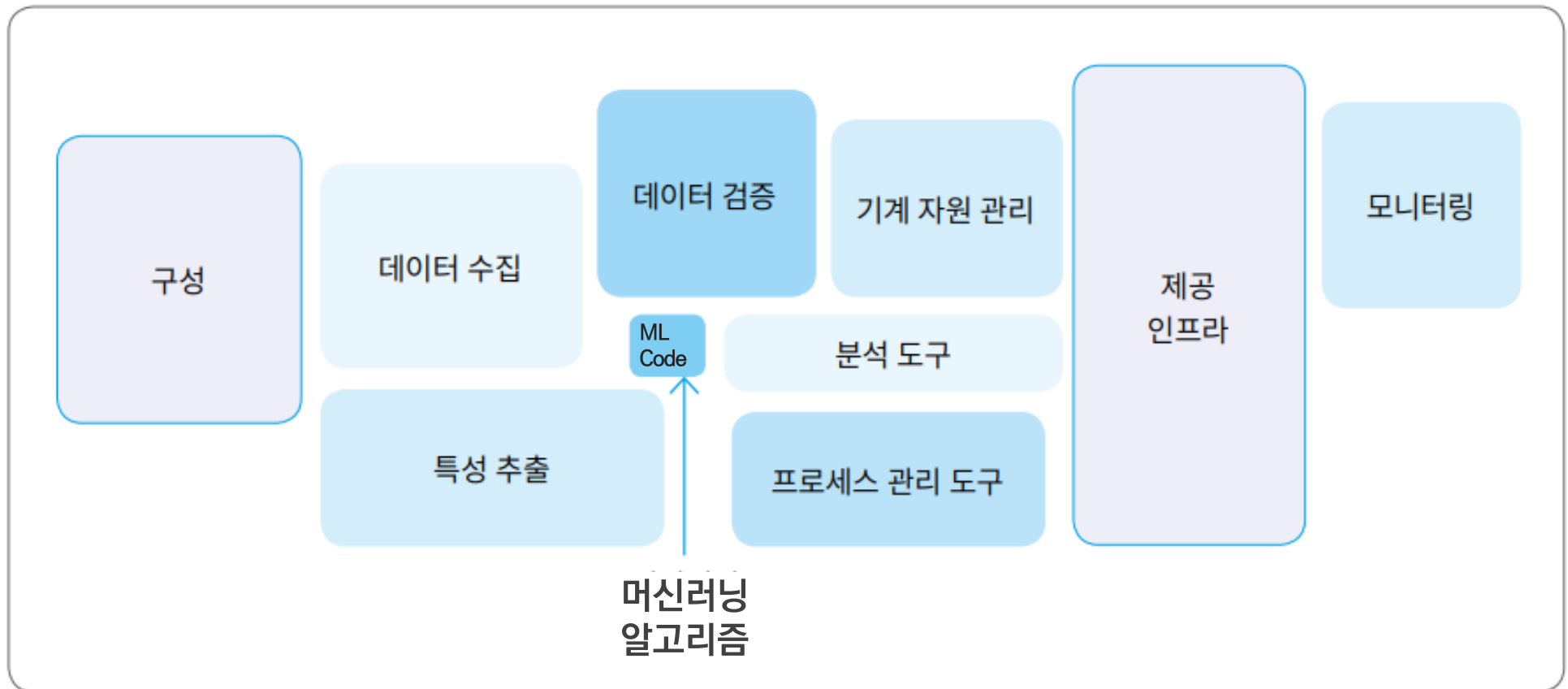
[Data Analytics]



[Biz. 전문가의 역할]

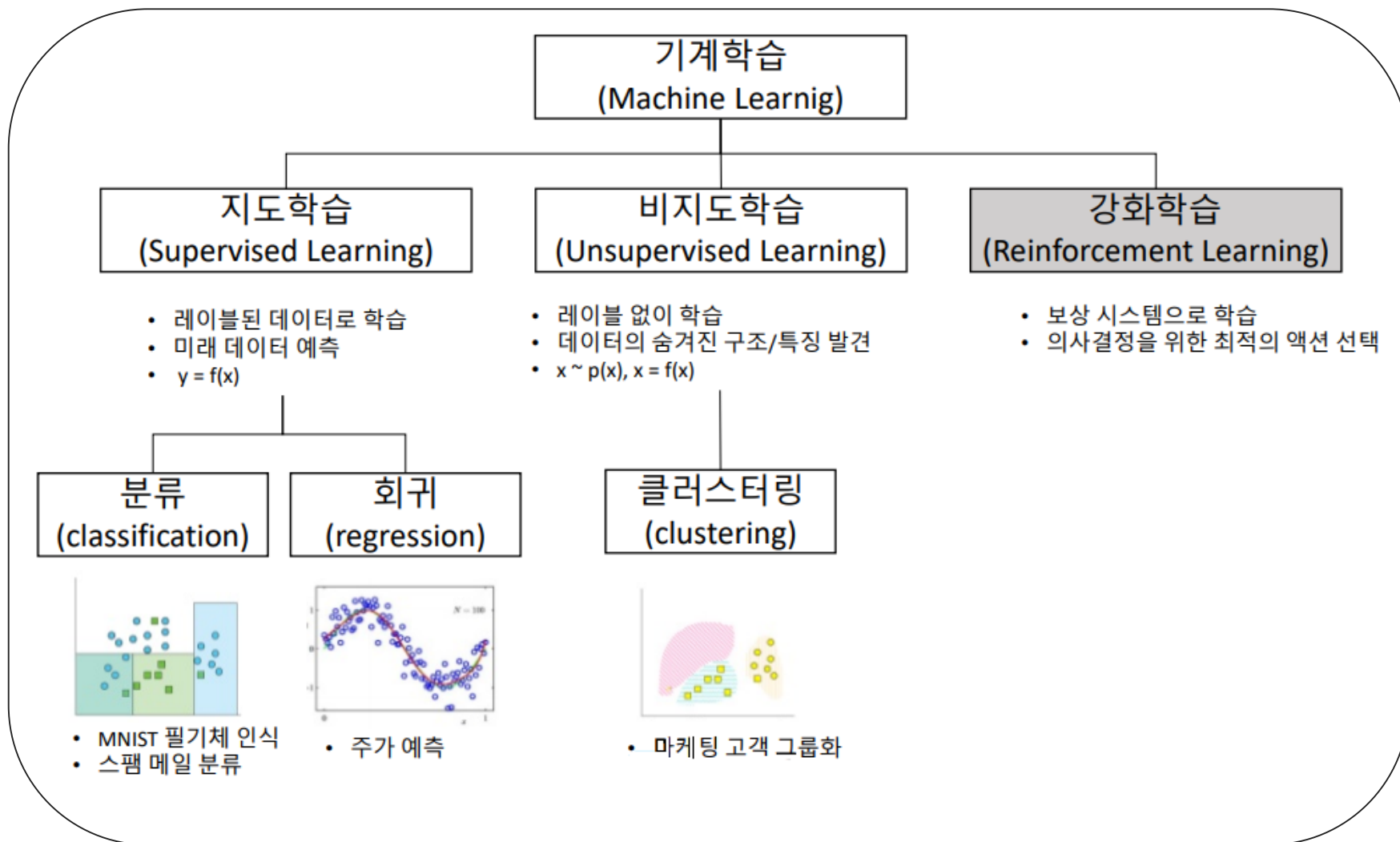
- ✓ 구체적인 Biz. 활용 계획을 Design
- ✓ 모델 활용 효과 & 한계를 정확히 알고 목표 설정
- ✓ 분석 기법/알고리즘에 대한 개괄적인 이해
- ✓ 분석가 수행 프로세스에 대한 이해
- ✓ 보유한 Data에 대한 정확한 이해, 제공 (의미, 기준 등 Biz. Logic)
- ✓ Data 패턴에 대한 의미 해석 시 의견 제시
- ✓ 지속 활용을 위한 모델 운영/관리 방안, 변화관리를 위한 Plan

머신러닝 구성요소



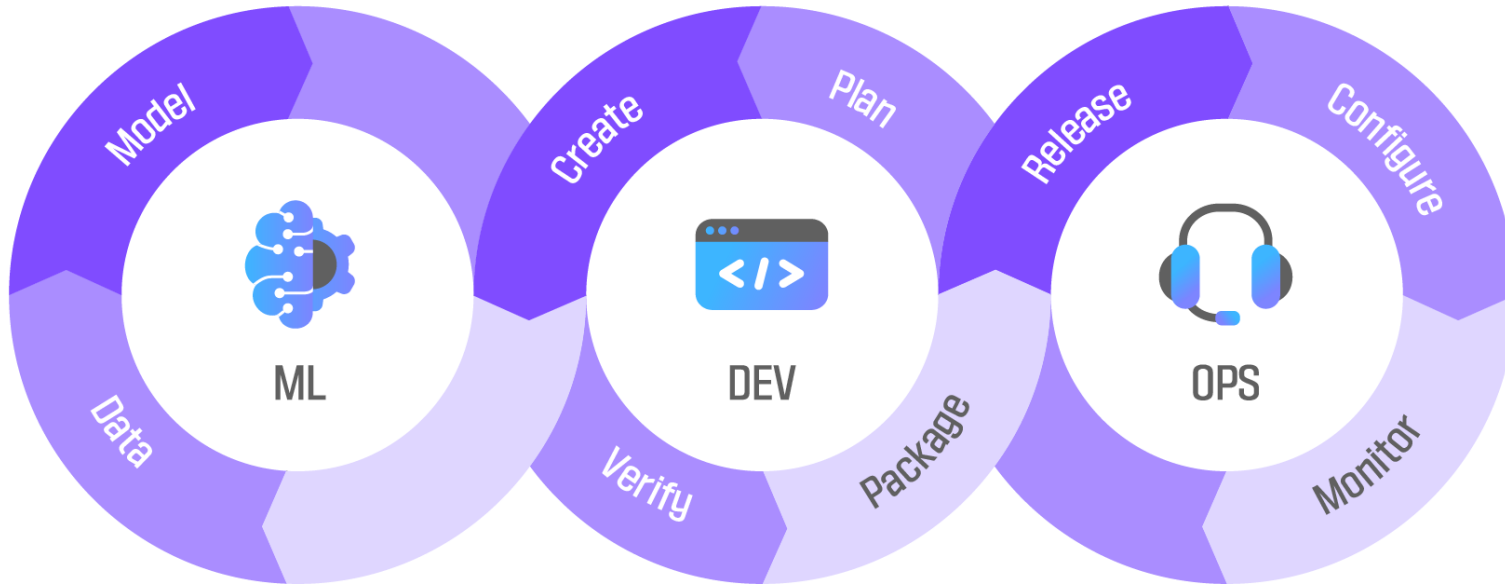
머신러닝 모델이 ML 시스템의 핵심이기는 하지만 모델을 운영하기 위해 기반 데이터와 인프라를 포함한 모든 시스템이 유기적으로 돌아가야 함.

머신러닝 종류



ML Ops

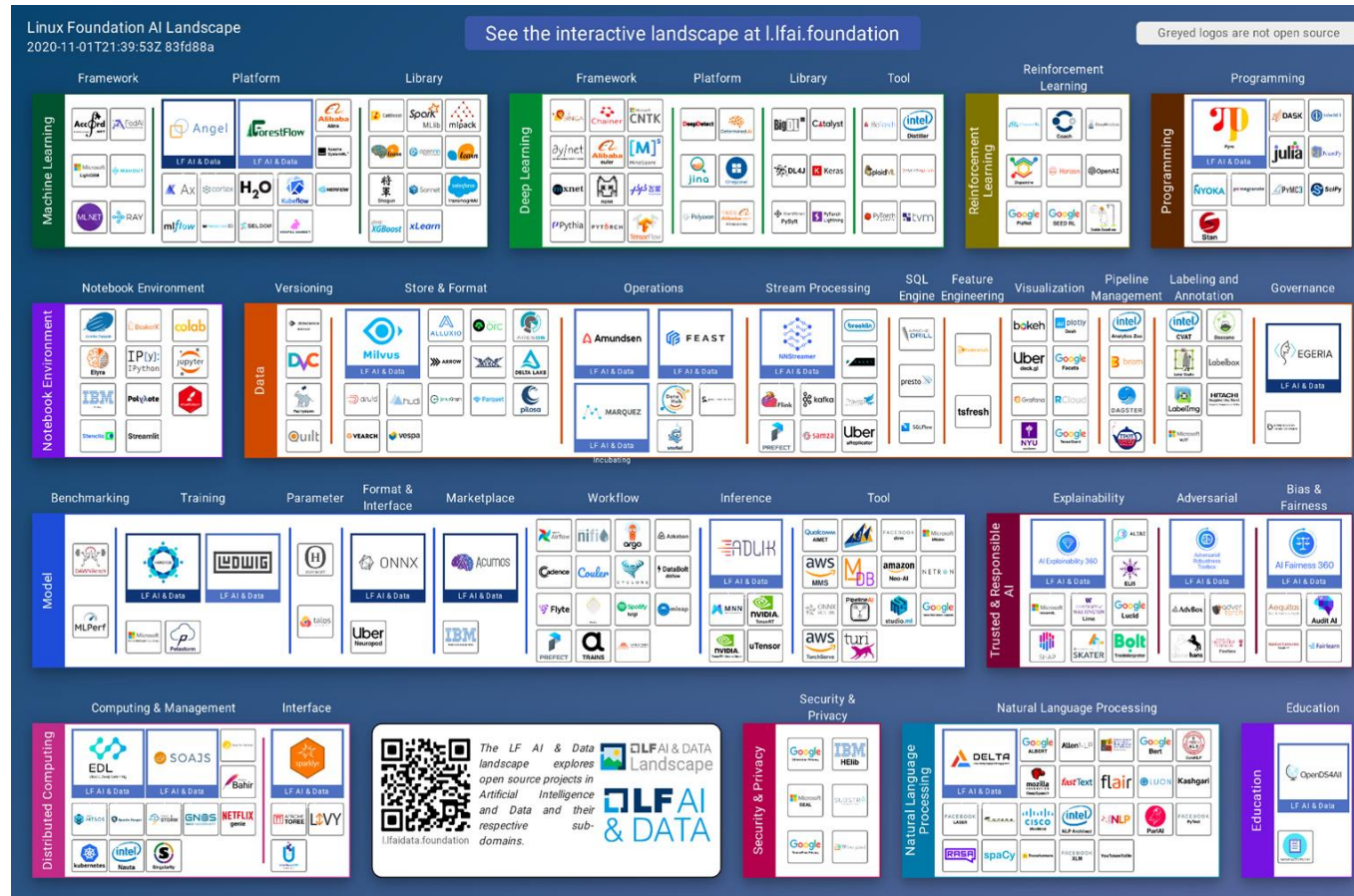
ML Ops란 머신 러닝(Machine Learning)과 개발(Dev),운영(Operations)을 합친 용어로 프로덕션 환경에서 머신 러닝(ML) 모델이 지속적이고 안정적으로 배포되도록 유지, 관리, 모니터링 해주는 것



[프로덕션에서 기계 학습 모델을 안정적이고 효율적으로 배포하고 유지 관리하는 것을 목표로 하는 패러다임]

ML Ops 오픈소스 생태계

고객이 이 모든 영역을 커버링하기란 매우 어렵습니다.



[데이터 관련 기술]

데이터 수집 파이프라인

Airflow, Spark Streaming, Flink, Kafka, Flume, Sqoop, ...

데이터 저장

Hadoop, MySQL, Amazon S3, MinIO, ...

데이터 관리

DVC, Feast, TFDV, Amundsen, ...

[모델 관련 기술]

모델 개발

Docker, Kubeflow, Jupyter Hub, Optuna, Ray, Katib, ...

모델 버전 관리

Git, MLflow, Github Action, Jenkins ...

모델 학습 스케줄링 관리

Kubernetes, Grafana, ...

[서빙 관련 기술]

모델 패키징

Flask, FastAPI, BentoML, Kubeflow, TFServing, Seldon-core, ...

서빙 모니터링

Prometheus, Grafana, Tanos, ...

파이프라인 관리

Airflow, Kubeflow, argo workflows, ...

출처

= <https://ml-ops.org/content/state-of-ml-ops>

= <https://landscape.lfai.foundation/>

Acculnsight+ 가 필요한 이유

Acculnsight+ 는 머신러닝 모델 개발과 운영의 간극을 줄여줍니다.

Acculnsight+ 의 중요 목표

1. 머신러닝 모델 개발(dev)과 머신러닝 모델 운영(ops) 에서 발생하는 문제의 반복을 최소화 하면서 비즈니스 가치를 창출
2. 머신러닝 모델링에 집중할 수 있도록 안정된 인프라를 만들고 자동으로 운영
3. 빠른 시간안에 적은 Risk 로 Production 까지 진행할 수 있도록 기술적 마찰을 감소
4. 데이터 분석을 모르는 현업 사용자들을 위한 쉬운 모델 개발 지원
5. 최신 기술과 ML알고리즘을 제공함으로써 최적의 머신러닝 서비스 운영 지원
6. 기업에서의 필수인 ML Ops 전체를 지원하는 플랫폼

감사합니다.