A review of the ML stack in industry and the problems that ML stakeholders and tool builders must solve to drive AI adoption

SARAH CATANZARO
General Partner
Amplify Partners
Who am I and why am I here?
Agenda

01 A SNAPSHOT
Snapshot of ML in industry (beyond FAANG/MANGA)

02 THE ML STACK
A taxonomy of applications, platforms and infrastructure

03 KEY CHALLENGES
Key challenges ML practitioners face

04 KEY CHALLENGES
Key challenges ML tool builders face
A snapshot of ML in industry
A Snapshot of ML in Industry

High ROI use cases are emerging among mid-market and public companies

ML-driven SaaS applications

- E-commerce
- Marketplaces
- Finance
- Real Estate
A Snapshot of ML in Industry

Structured Data

Embedding
(structured/semi-structured data)

Text and visual data
Mid-market and public companies have hired ML practitioners and plan to expand their team, but are still experimenting with different organizational models.
The ML stack; a taxonomy
The ML stack; a taxonomy

**Model Operations**
- Model Monitoring
- Model Optimization & Deployment
- Model Compliance
- Continuous Learning

**Build & Deploy**
- Data Science Framework
- Experiment Tracking
- Version Control
- Auto ML
- Distributed Training

**Data Management**
- Databases: Cloud Data Warehouse, Data Lake, Vector Database
- Feature Management: Feature Store, Feature/Metrics Layer
- Data Labeling: Labeling Service, Weak Supervision
Data Management

Databases
- DATA WAREHOUSE
  - snowflake
  - Google BigQuery
  - amazon REDSHIFT
  - FIREBOLT
- DATA LAKE
  - DELTA LAKE
  - ICEBERG
  - huDi
- VECTOR DATABASE
  - vespa
  - Milvus
  - Weaviate
  - Pinecone

Feature Management
- FEATURE STORE
  - TECTON
  - FEAST
  - LOGICAL CLOCKS
- FEATURE LAYER/PROXY

Data Labeling
- LABELING SERVICES
  - scale
  - Labelbox
  - Label Studio
- WEAK SUPERVISION
  - snorkel
  - Cleanlab
Build and Deploy

DATA SCIENCE FRAMEWORK:
- Metaflow
- mlflow
- Kubeflow

EXPERIMENT TRACKING:
- Weights & Biases
- comet
- neptune.ml

VERSION CONTROL:
- R
- Verta.ai
- DVC

AUTOML:
- DataRobot
- tangram
- H2O.ai

DISTRIBUTED TRAINING:
- RAY
- DASK
- bodo.ai
Model Operations

- **Model Monitoring**
  - GANTRY
  - arize
  - WHYLABS

- **Model Optimization & Deployment**
  - cortex
  - OctoML
  - modelbit
  - deci.

- **Model Compliance**
  - Arthur
  - comet
  - fiddler
  - truera

- **Continuous Learning**
  - River
  - ALECTIO
  - FAUST

Amplify Partners - Proprietary and Confidential
Challenges that ML practitioners face
SCOPING ML PROJECTS

HOW CAN YOU ANTICIPATE THE:

• Likelihood of success?
• Resources (data, compute, time) needed?

HOW CAN YOU SCOPE ML EXPERIMENTS TO:

• Avoid expending unnecessary effort on projects that aren’t feasible
• Leverage engineering resources effectively
ACCESSING AND ITERATING ON DATASETS

PRACTITIONERS NEED:

• Access to high quality data
• Tools for data-centric ML
• Unstructured data wrangling
BUILDING SOFTWARE APPLICATIONS WITH ML

Practitioners must build software applications for which a prediction service (ML model) is one component

NEED TO INTEGRATE WITH:

- Data management systems
- Other [micro]services (e.g. authentication)
- Developer tools (e.g. CI/CD)
MANAGING INFRASTRUCTURE

• DS should spend time on modeling and analyses NOT managing infrastructure and environments!

If k8s is so hard for R&D, why do you keep trying to use it?

He’s out of line, but he’s right.
INTRODUCING

Modal Labs
Challenges ML Tool Builders Face
Comparing the “Modern Data Stack” and ML Stack

**Data Sources**
- Applications & APIs
- Databases
- Files
- Events

**Ingestion**
- Connectors (Meltano, Singer, Airbyte, Fivetran, Stitch)
- Event Streams (Kafka Connect, Kinesis Firehose, Data Flow)

**Storage / Warehousing**
- Transformation (DBT, Dremio, Spark, Python, R)

**Publication / Serving**
- Analytics & Business Intelligence (Superset, Metabase, Redash, Looker, Tableau)
- Operations (HighTouch, Censys, Retool)
- Data Science & AI/ML (AWS SageMaker, DataRobot, H₂O)

Enable self-service E2E data products

Data Ops principles and tools
NO DOMINANT DESIGN FOR THE ML STACK

- NO system of record
- NO best practices or standards
- NO clear roles and responsibilities
No partnership ecosystem: Why everything becomes an end-to-end platform

BEST OF BREED STACK

ALL IN ONE SUITE
• Different applications may have different requirements

• Different model architectures may have different requirements