The Machine Learning Project Lifecycle

Steps of an ML project

The ML project lifecycle

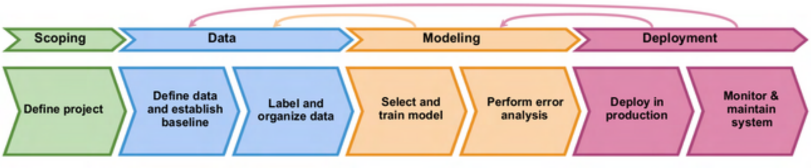


Figure: The ML project lifecycle

- Scoping
- Data
- Mo deling
- Deployment

The Machine Learning Project Lifecycle

Case study: speech recognition

Speech recognition: Scoping stage

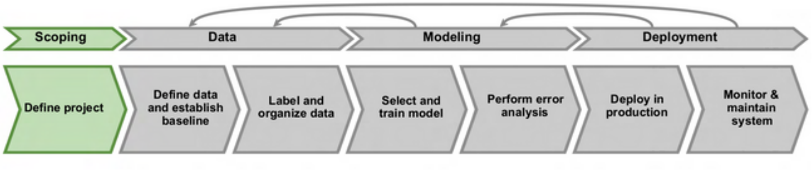


Figure: The ML project lifecycle

- Decide to work on speech recognition for voice search.
- Decide on key metrics:
  - Accuracy, latency, throughput
- Estimate resources and timeline

Speech recognition: Data stage

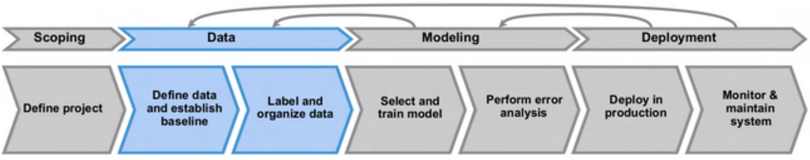


Figure: The ML project lifecycle

Define data:

- Is the data labeled consistently?
- How much silence before/after each clip?
- How to perform volume normalization?

Examples:

- “Um, today’s weather”
- “Um… today’s weather”
- “Today’s weather”

Speech recognition: Modeling stage

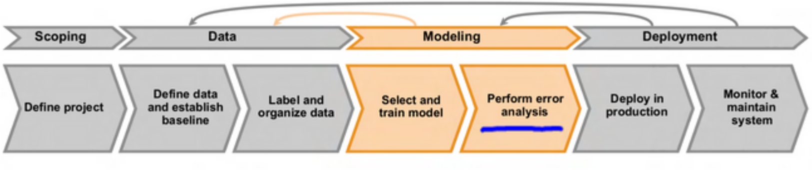
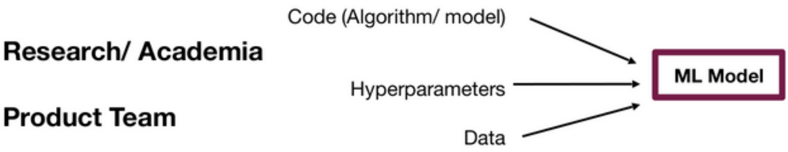


Figure: The ML project lifecycle



Speech recognition: Deployment stage

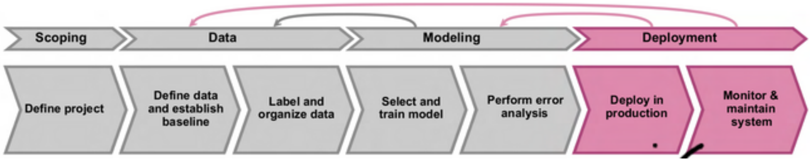
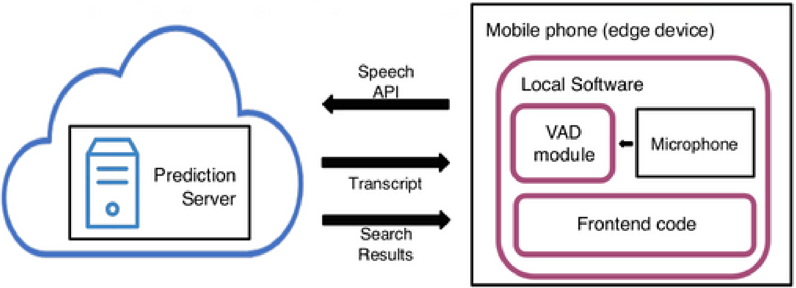


Figure: The ML project lifecycle



Key requirements for an MLOps foundation

AI-driven organizations are transforming industries:

- Leverage data and machine learning to tackle their hardest problems.
- According to McKinsey, companies that adopt AI across workflows by 2025maysee . +120% cash flow growth by 2030

But operationalizing ML is not easy:

- ML systems incur significant technical debt if unmanaged.
- Combine traditional software issues plus ML-specific challenges:
  - Complex hardware/software dependencies Data needs to
  - be validated, not just code Models degrade over time due
  - to changing environments Fail silently;harder to debug
  - and monitor
- Building a model is easy;managing its lifecycle is hard.

Bridging ML and IT: MLOps Meets DevOps

DevOps: Proven practice for large-scale software systems

- Evolved over decades of software engineering experience.
- Benefits:
  - Shorter development cycles
  - Faster deployment velocity
  - Reliable, high-quality releases

MLOps: Extending DevOps to Machine Learning

- MLOps = ML system development (Dev) + ML system operation (Ops)
- Inspired by DevOps, but adapted for the ML lifecycle.
- Challenges:
  - Continuous Integration / Delivery (CI/CD) is harder in ML
  - ML systems involve data, models, code, and environments

CI/CD in ML Systems: Beyond Traditional Software

ML systems redefine CI/CD and introduce new requirements:

- Continuous Integration (CI):
  - Validate not just code and components
  - Also validate data, data schemas, and trained models
- Continuous Delivery (CD):
  - Not just deploying software
  - Deliver end-to-end ML pipelines that deploy model services
- Continuous Training (CT):
  - Unique to ML systems
  - Automate retraining of models for testing and deployment
- Continuous Monitoring (CM):
  - Monitor more than just system errors
  - Track live inference data and model performance degradation