Simple, Yet Scalable MLOps: Bridging the Gap Between Data Science and CI/CD

Sameeksha Garg, CMU
Dr. Sachin Garg, NavankurIT
Mitesh Singh Jat, Independent Consultant

May 18–20, 2026 | Minneapolis, Minnesota

Problem Statement

ML models often work in notebooks but fail in production because:

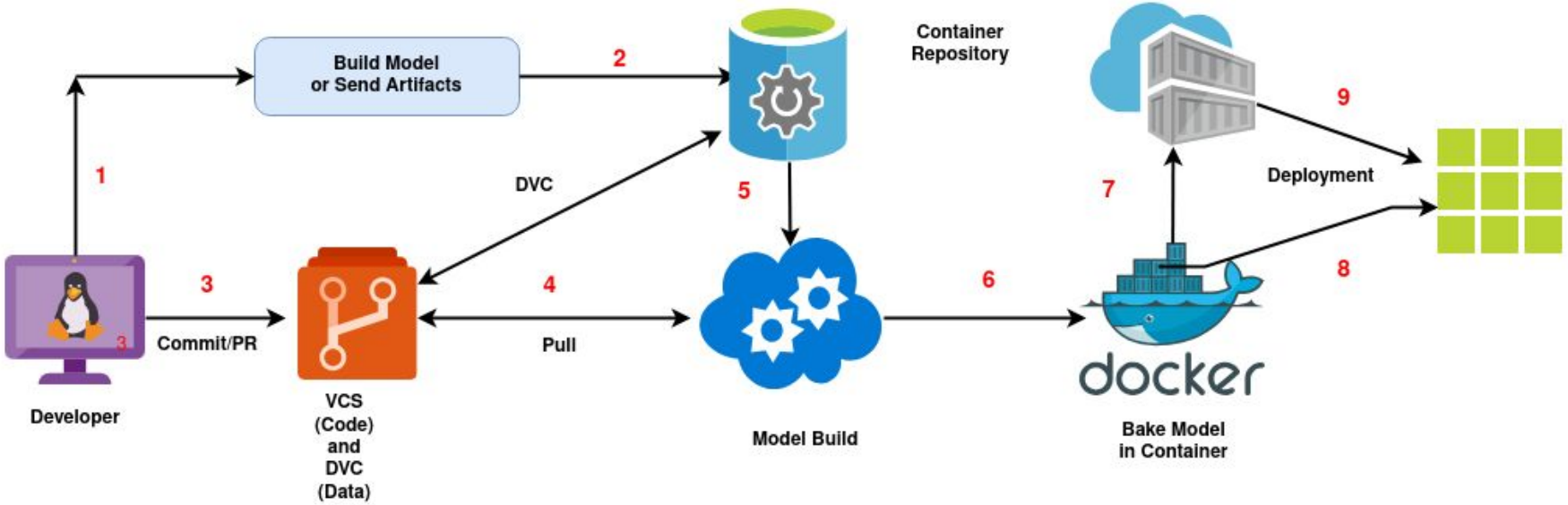
- Code changes
- Training data changes
- Model artifacts change
- Runtime environments change
- Nobody can easily reproduce what was deployed

Traditional DevOps ships code. ML systems ship **code, data, and model behavior**.

MLOps

- DevOps for Machine Learning
- Link between Data Scientists ⇔ Operations/Infrastructure Teams
- Addresses Various Challenges in the ML Lifecycle
  - Version Control (both Data & Code)
  - Model Reproducibility
  - Efficient Deployment with Rollback
  - Scalability
- Tools like MLflow/Google Vertex AI/Amazon SageMaker/Azure ML
- Our Proposal: *Simplify Operations without vendor lock-in*

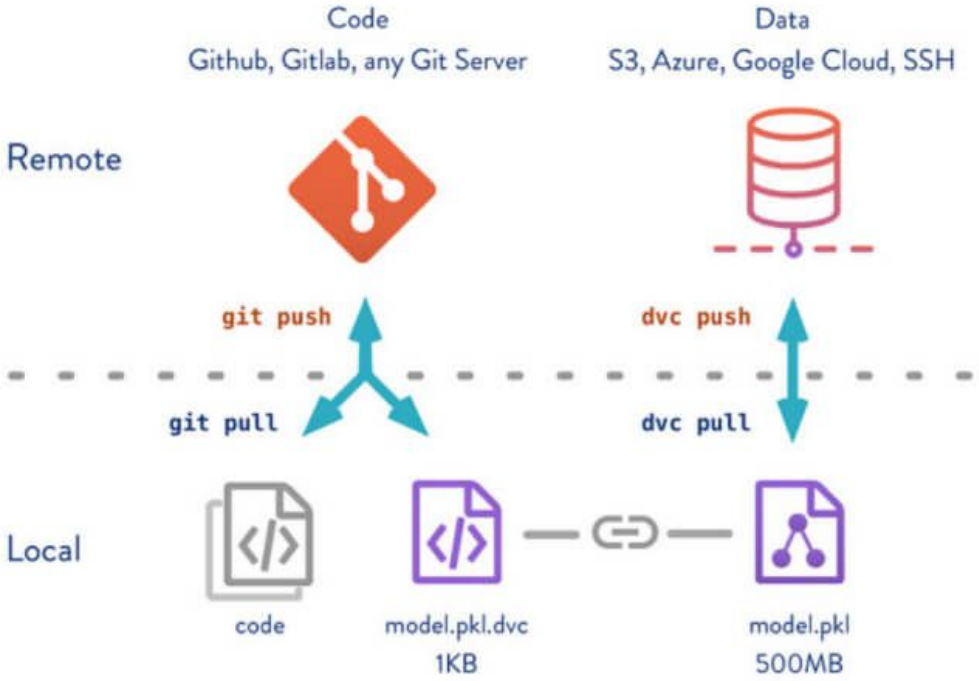
From Model Artifact to Deployable Container



Version Control & Model Reproducibility

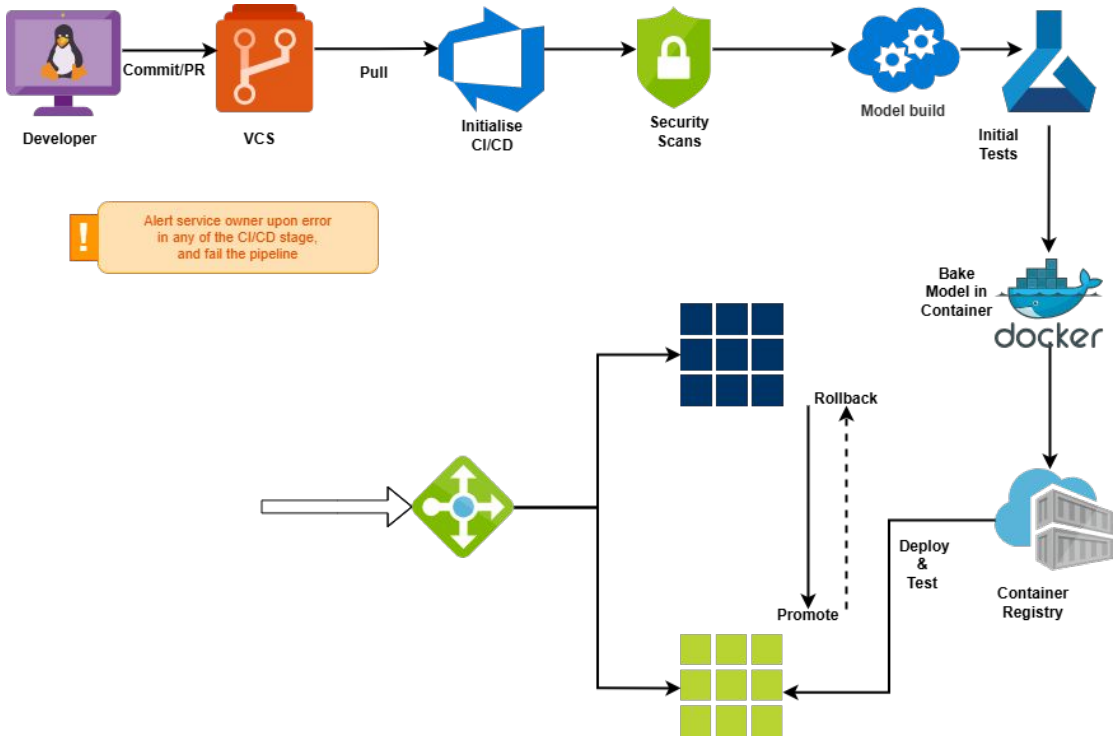
- Version Control Systems (CVS/Git etc.) are designed to manage text
- Training data and model representations can be binary data
- Traditional VCS is not ideal for large binary artifacts
- Model reproducibility requires version controlling both data and code
- Data Version Control (DVC) is the suggested tool
  - DVC manages the metadata in Git
  - Actual data is stored outside of VCS (object stores, file systems etc.)

Version Control & Model Reproducibility (DVC)



MLOps CI/CD: Test, Promote, or Roll Back

- New Feature Development/ Bug Fix
- Build/Update Model
- Test Model
- Containerize Model
- Deploy to Blue/Green environment
- Observe Model
- Promote to Prod / Roll Back



Example Safe Deployment: Blue-Green in Kubernetes

BLUE: Current

GREEN: New

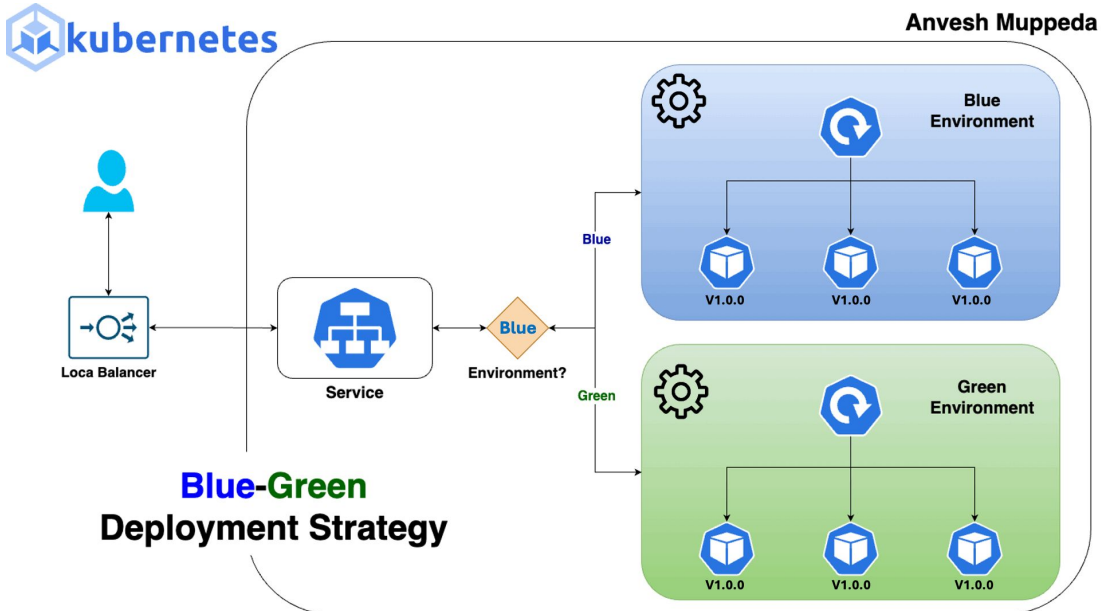


Image: [Blue-Green Deployment in Kubernetes]

Takeaways

Scalable MLOps does not have to be complex

- **Model reproducibility is *key***
  Models are **not** just code**:** *data* the model is trained on is a key artifact
- **Models should be deployable artifacts**
  No “it works on my machine” - Bake the model and dependencies into a consistent container.
- **Promotion should be gated**
  Human in the Loop
- **The UNIX Way: KISS & eschew Complexity**
  Git + DVC + Docker (+ Kubernetes) can provide scalable MLOps without vendor lock-in.

Thank you!

Sameeksha Garg

Sachin Garg

Mitesh Singh Jat