Roadmap to Learn MLOps from Basic to Advanced

---

1. Foundations of MLOps

Goal: Understand the basics of MLOps, why it is needed, and the foundational concepts.

Topics to Learn:

- **What is MLOps?**
  - Importance of MLOps in the ML lifecycle.
  - MLOps vs DevOps.
  - Key components: ML model lifecycle, CI/CD, automation.
- **Basic ML Workflow**
  - Problem formulation, data preparation, model development, and deployment.
- **Programming Fundamentals**
  - Python (mandatory), familiarity with shell scripting.
  - Git and version control systems.

Tools to Explore:

- Python, Jupyter Notebook, Git/GitHub.

Resources:

- Courses: Coursera's "Introduction to MLOps" or Fast.ai
- Book: "Practical MLOps" by Noah Gift.

---

2. Machine Learning Fundamentals

Goal: Strengthen ML knowledge to understand the deployment and monitoring aspects.

Topics to Learn:

- Supervised vs Unsupervised Learning.

- Model evaluation metrics (e.g., accuracy, precision, recall, AUC).
- Hyperparameter tuning.
- Overfitting vs Underfitting.
- Transfer Learning.

Tools to Explore:

- Scikit-learn, TensorFlow/PyTorch, XGBoost, LightGBM.

---

3. MLOps Core Concepts

Goal: Dive into key concepts specific to MLOps.

Topics to Learn:

1. **Model Versioning**
   - Tools like DVC, MLflow, or Weights & Biases (W&B).
   - Tracking experiments, datasets, and hyperparameters.
2. **Model Deployment**
   - What is deployment?
   - Types of deployments (Batch Inference, Real-Time Inference, Edge Deployment).
3. **Model Monitoring and Maintenance**
   - Concept drift, model drift, and retraining pipelines.
   - Monitoring model performance over time.

Tools to Explore:

- MLflow, W&B, DVC.

Resources:

- Video tutorials on DVC and MLflow.
- Hands-on experimentation with Colab or local projects.

---

4. Cloud Platforms for MLOps

Goal: Learn to use cloud services for scalable MLOps practices.

Topics to Learn:

- **Overview of Cloud Providers**
  - AWS, GCP, Azure (pick one and specialize).
- **Managed ML Services**
  - AWS Sagemaker, Google Vertex AI, Azure ML Studio.
- **Cloud Deployment Basics**
  - Docker, Kubernetes, and serverless architectures.
- **Infrastructure as Code (IaC)**
  - Terraform, CloudFormation.

Tools to Explore:

- Docker, Kubernetes, AWS/GCP/Azure.

Resources:

- Free Cloud credits (Google Cloud's free tier, AWS Educate).
- YouTube: "Introduction to Docker and Kubernetes for ML".

---

5. Data Engineering for MLOps

Goal: Learn to manage and preprocess data pipelines effectively.

Topics to Learn:

- **ETL Pipelines** (Extract, Transform, Load).
- **Data Versioning and Lineage.**
- Scalable data processing using **Apache Spark** or **Databricks**.
- Data Warehousing: Snowflake, BigQuery.

Tools to Explore:

- Apache Airflow, Prefect, Spark, Snowflake, BigQuery.

Resources:

- Datacamp courses on data engineering.
- Practical tutorials on Airflow and Spark.

---

6. Continuous Integration / Continuous Deployment (CI/CD)

Goal: Learn CI/CD pipelines specific to ML workflows.

Topics to Learn:

- Building CI/CD pipelines for ML models.
- Automating testing of data, models, and code.
- GitHub Actions, Jenkins, CircleCI.
- Automated retraining pipelines.

Tools to Explore:

- GitHub Actions, Jenkins, Kubeflow Pipelines.

Resources:

- Blog: "MLOps with GitHub Actions".
- Hands-on projects using Kubeflow.

---

7. Advanced Topics in MLOps

Goal: Master the advanced tools and techniques in MLOps.

Topics to Learn:

1. **Feature Stores**
    - What is a feature store? Importance in MLOps.
    - Tools: Feast, Tecton.
2. **Advanced Model Deployment**
    - A/B testing, Canary Deployments, Multi-Model Serving.
3. **Advanced Monitoring**
    - Logging and alerting for ML pipelines.
    - Tools like Prometheus, Grafana.
4. **Distributed Training and Serving**
    - Horovod, Ray for distributed model training.
5. **MLOps for Large Language Models (LLMs)**
    - Fine-tuning, hosting, and monitoring LLMs.

Tools to Explore:

- Feast, Ray, Horovod, Prometheus, Grafana.

Resources:

- Advanced tutorials on LLM fine-tuning with Hugging Face.
- Blogs on feature stores and monitoring.

---

8. Security and Compliance in MLOps

Goal: Learn how to secure ML systems and ensure compliance.

Topics to Learn:

- Data privacy laws (GDPR, CCPA).
- Model explainability and interpretability.
- Secure model deployment practices.
- Adversarial attacks and defenses.

Tools to Explore:

- SHAP, LIME for interpretability.
- Presidio for data anonymization.

---

9. Projects and Portfolio Building

Goal: Build a portfolio showcasing end-to-end MLOps projects.

Suggested Projects:

1. **End-to-End ML Pipeline**
    - Data preprocessing, training, deployment, and monitoring.
2. **Real-Time ML System**
    - Deploy a real-time fraud detection or recommendation system.
3. **Cloud-Based ML Workflow**
    - Use AWS/GCP/Azure for a scalable ML solution.

Tools for Deployment:

- Streamlit, FastAPI, Flask for UI/API.
- Docker, Kubernetes for containerization.

---

10. Networking and Community Engagement

- Contribute to open-source MLOps projects.
- Join communities: MLOps Community Slack, GitHub repos, LinkedIn groups.
- Follow experts: Chip Huyen, Google MLOps blog.