ML Ops

It is a process of building ML models , deploying them , continuously Monitoring them and fixing them .

Devops vs Mlops

Dev ops = Process of building and deploying any software application

ML ops = Process of building and deploying any Machine learning model.

Imagine You are running a movie recommendation website :

1) **Data collection and Pre processing** (Getting the ingredients ready)

2) **Feature engineering** (Selecting the right ingredients)

3) **Model Building**

4) **Model Training** (Teaching the Model)

5) **Model validations** (Testing the Model)

6) **Model Deployment** (Putting the Model to work)

Prediction technique :

-> Real time prediction

-> Batch mode prediction

Deployment technique :

-> Canary deployment (Deploying model for small percentage of users first)

-> Blue-green deployment (switching from an old model to new model without downtime)

-> shadow deployment (Running the new model alongside of old model and compare the result before switching fully)

MLOps pipelines typically uses containerization (Docker) and orchestration tools (Kubernetes) to manage the deployment of models.

7) Monitoring and Maintenance (Keeping an eye on model)

Several factors that can degrade the performance of model :

-> **Data drift** : Changes in the input data distribution as compared to training data

-> **Model drift** : Changes in the accuracy of my model due to environment changes and problem in domain

-> **Concept drift** : The underlying patterns that the model was trained may no longer hold eg (customer behavior changes)

When this will happen you will retrain the model to make it better again .

8) CI /CD Pipelines (Making the system run smoothly)

CI = Continuous Integration (This is about automatically testing your changes every time you improve the model or change its code , It automatically gets tested to make sure it works:

CD = Continuous Deployment (Once the test are passed , the new model or code version is automatically deployed (put into production) , This make sures that the website always used the best model without manual work)

9) Automation and Orchestration (Automating the workflow)

10) Versioning (Tracking everything)

-> Data versioning = You can track which version of data was originally used .

-> Model versioning = Each time you update the model , you save that version , so it new model doesn't work fine you can roll back to previous version

-> Code versioning = You can use version control system (Git) you can keep a record of changes to code .

11) Collaboration = MLOps involves the collaboration between Data scientist (Who will create the model) and Engineers (Who will deploy the model)

MLOps tools you might use :

1) **Data collection and pre processing** : Tools like Apache airflow to automate the data flows

2) **Model training** : scikit-learn , TensorFlow and lots of other libraries ,

3) **Model deployment** : Docker , Kubernetes , to manage and deploy models in containers.

4) **Monitoring and Maintenance** : MLflow or Prometheus you can use to track model performance

5) **CI/CD** : Jenkins or Github or Ansible to automate the testing and deployment