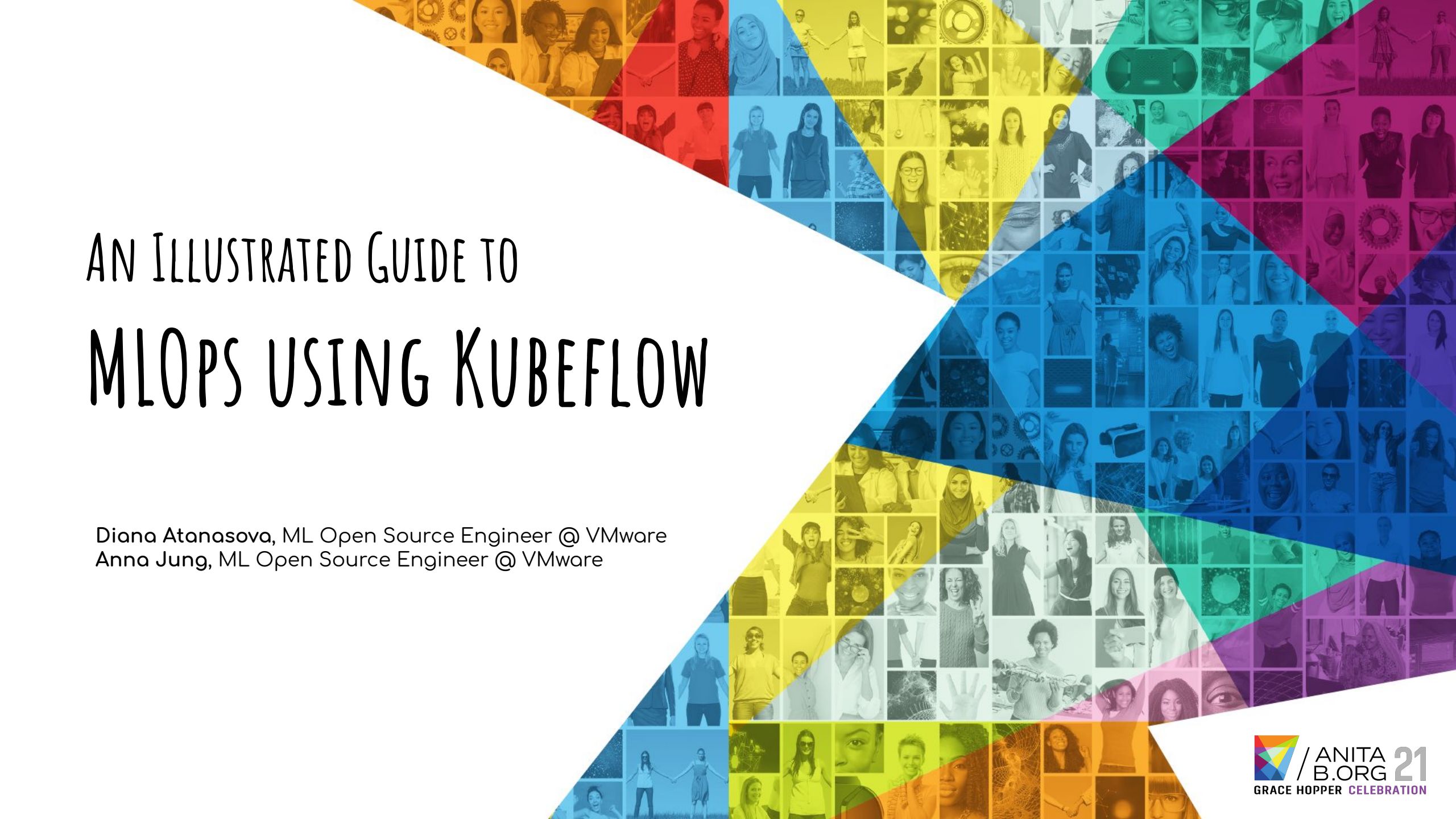
/ ANITA B.ORG 21
GRACE HOPPER CELEBRATION
DARE TO
TRANSFORM



An Illustrated Guide to

MLOps using Kubeflow

Diana Atanasova, ML Open Source Engineer @ VMware
Anna Jung, ML Open Source Engineer @ VMware

ANITA B.ORG 21
GRACE HOPPER CELEBRATION

Diana

Open Source Engineer, ML

Diana Atanasova is an Open Source Engineer at VMWare/Bulgaria. Diana is contributing in ML related projects like Kubeflow and ONNX. She has years of experience working on Network Management and Orchestration Systems. She enjoys singing in a choir.

Anna

Open Source Engineer, ML

Anna Jung is an Open Source Engineer at VMware, contributing to various open source projects related to Machine Learning. She believes in the importance of giving back to the open source community and is passionate about increasing diversity in open source. When away from the keyboard, Anna is often at film festivals supporting independent filmmakers.

Agenda

Machine Learning

Machine Learning Operations

Kubeflow

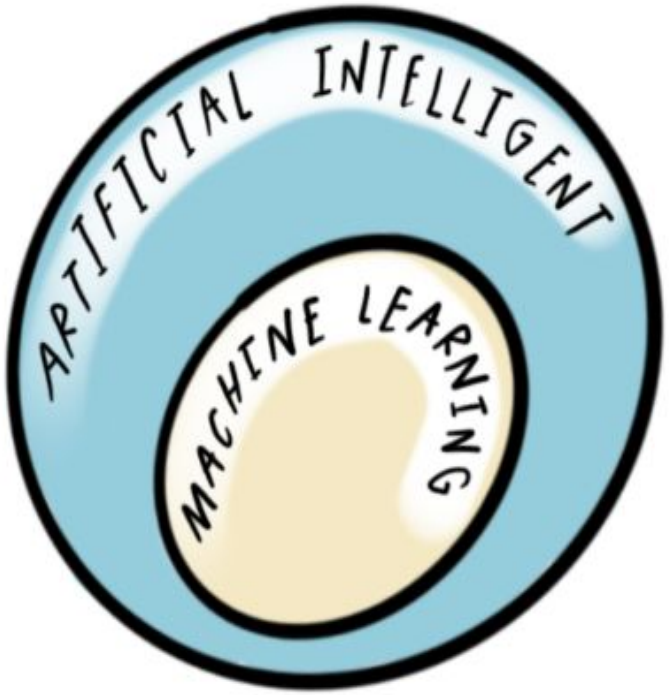
ML + MLOps + Kubeflow Workshop

What is MACHINE LEARNING?

Definition

"field of study that gives computers **the ability to learn without being explicitly programmed**" - Arthur Samuel

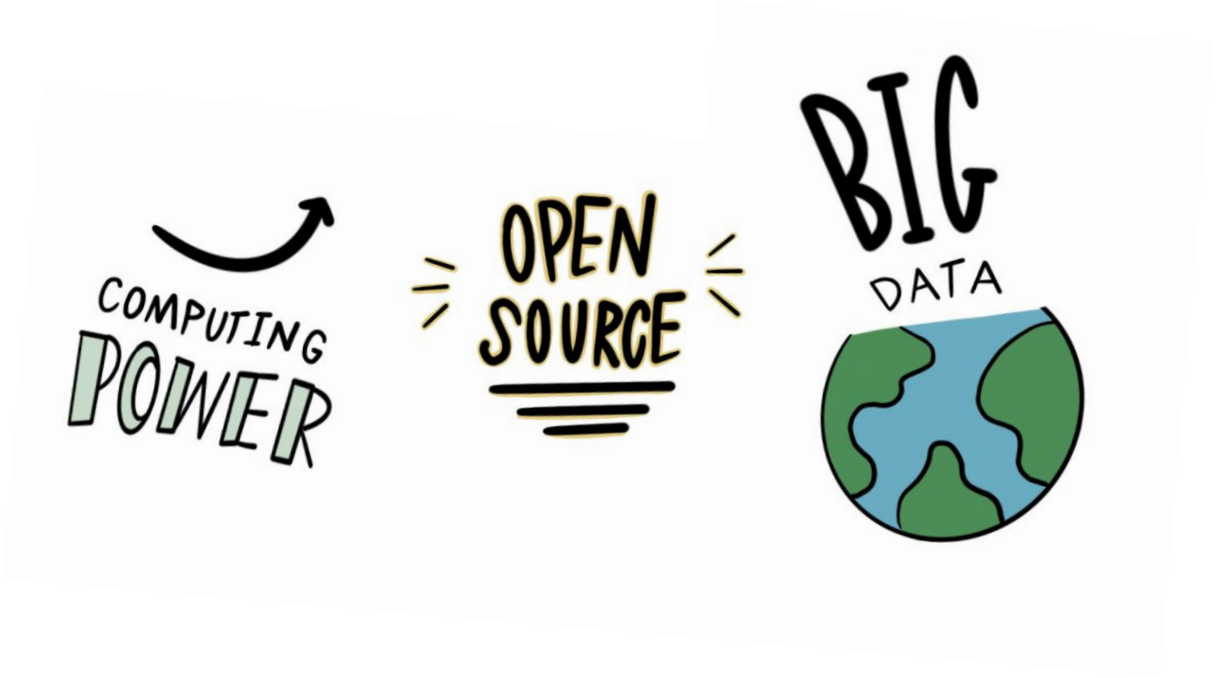
Machine Learning is a **subset** of Artificial Intelligence



How Do Machine **Learn**?

Why Machine Learning?

Reasons Why Machine Learning Is Everywhere

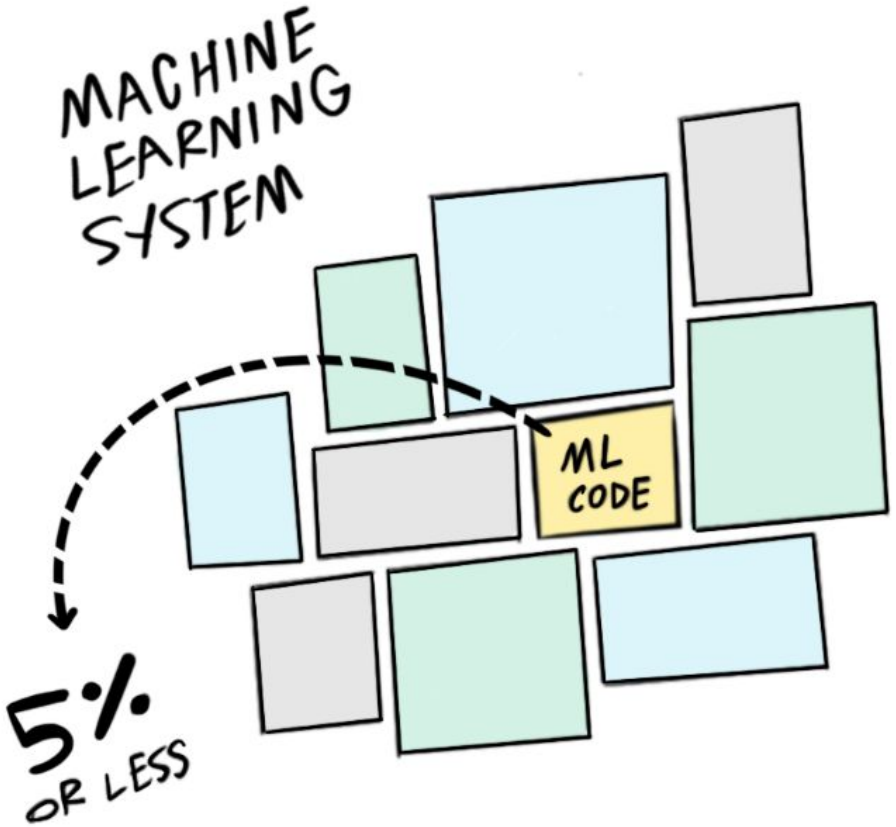


What is MLOps?

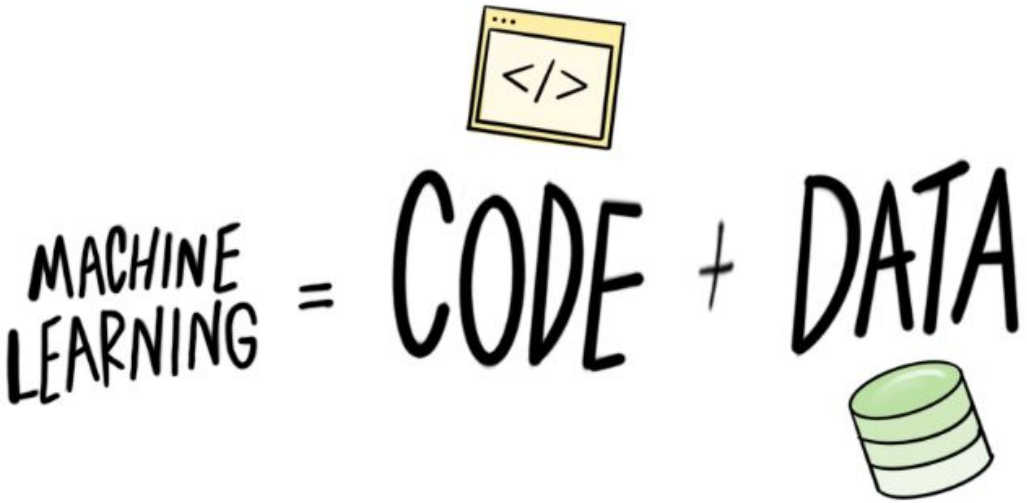
Definition

"A set of practices to build, deploy, and maintain machine learning models"

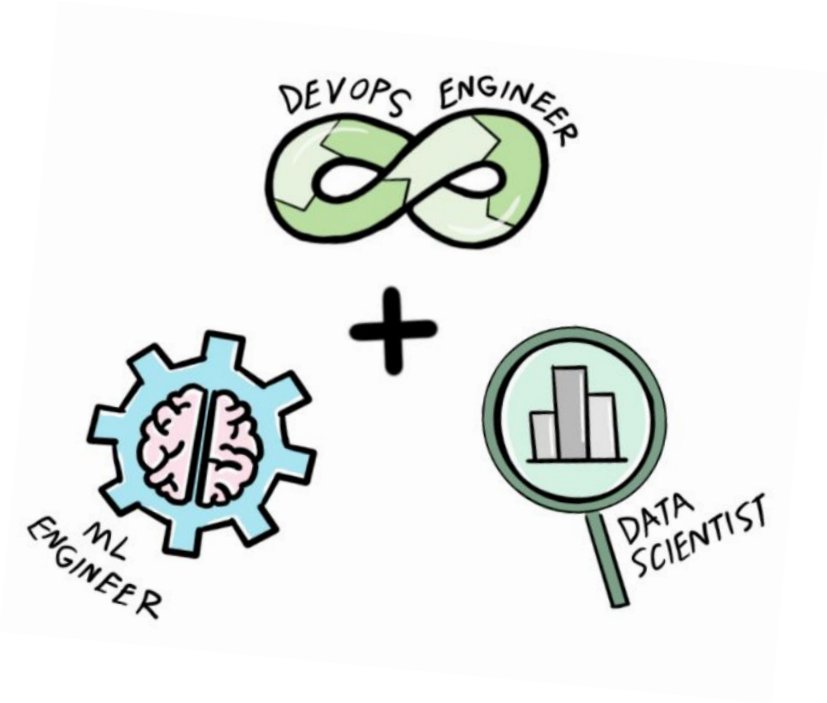
Machine Learning Code Is Important, But It's Not Enough

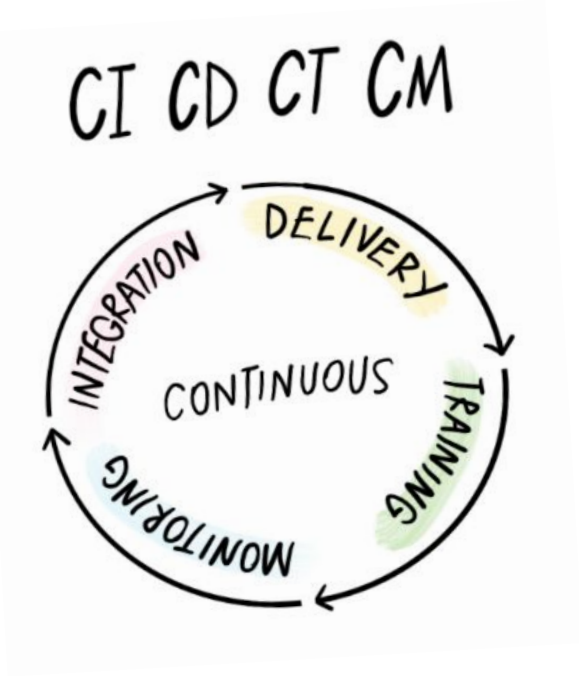


It's Not Just DevOps for ML

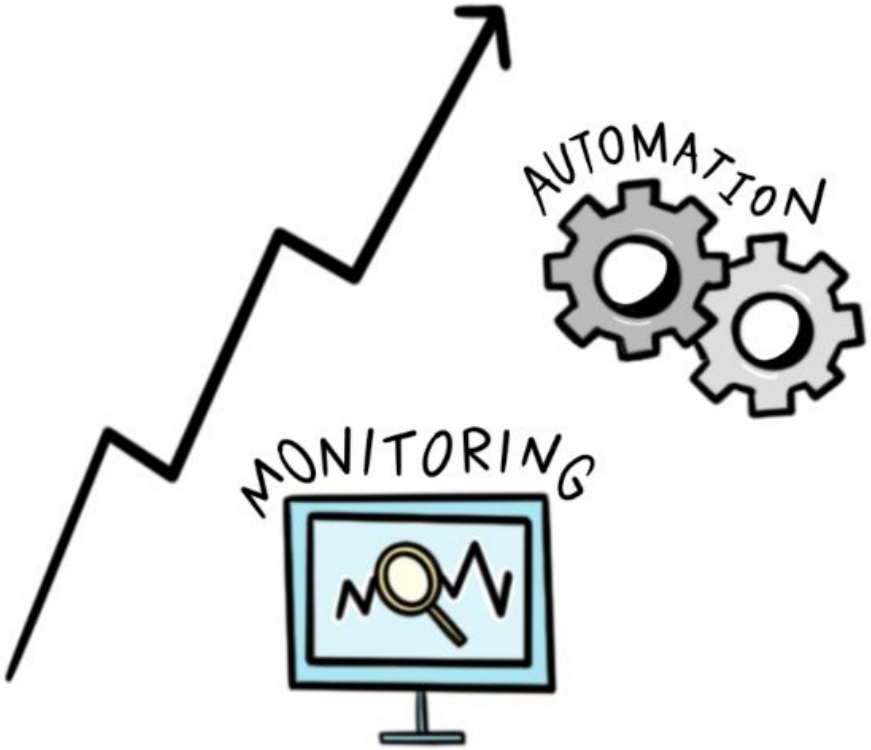


MLOps


DEVOPS ENGINEER
ML ENGINEER
DATA SCIENTIST


CI CD CT CM
INTEGRATION
DELIVERY
CONTINUOUS
TRAINING
MONITORING

What You Need Is a Complete End To End Workflow



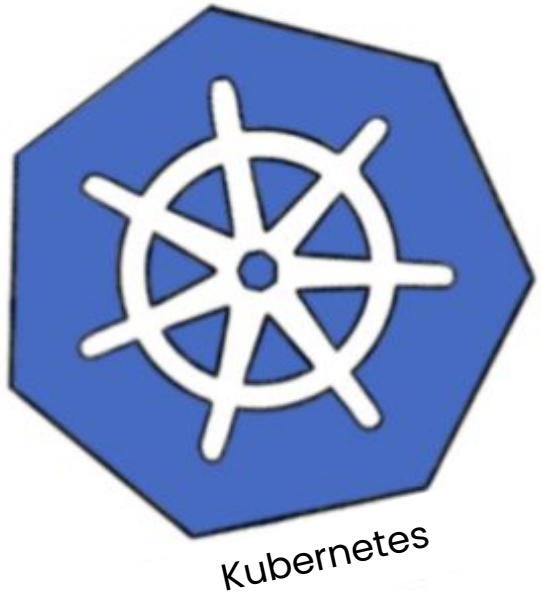
What is KUBEFLOW?

Definition

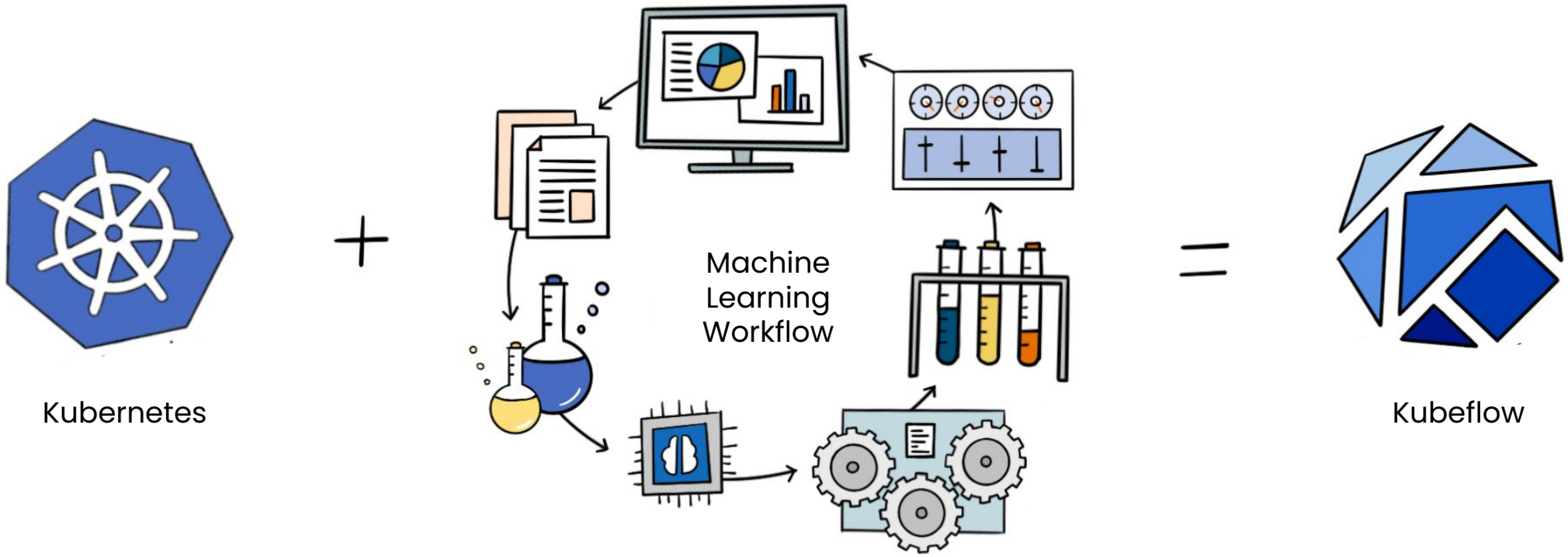
"an **Open Source** Kubernetes-native **platform for** developing, orchestrating, deploying and running scalable and portable **ML workflows**."

Kubernetes

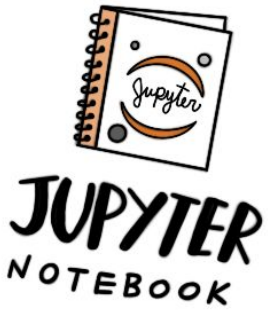
"Kubernetes, also known as K8s, is an **open-source system for automating** deployment, scaling, and **management of containerized applications.**"

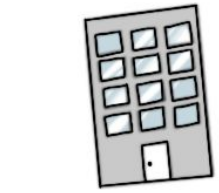


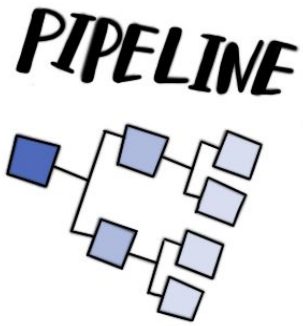
Machine Learning On Kubernetes



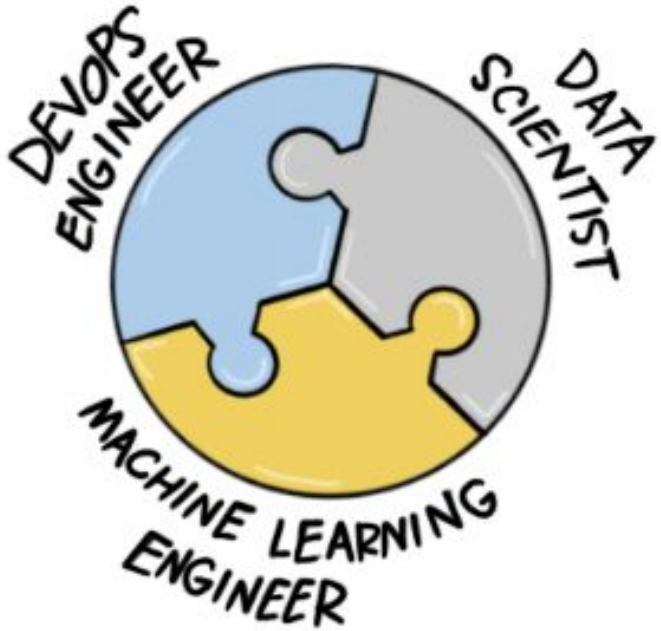
Kubeflow Has **All** You Need







Kubeflow Benefits All

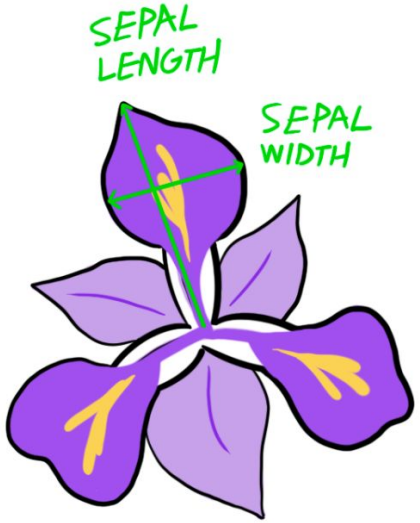


Let's put it all together!

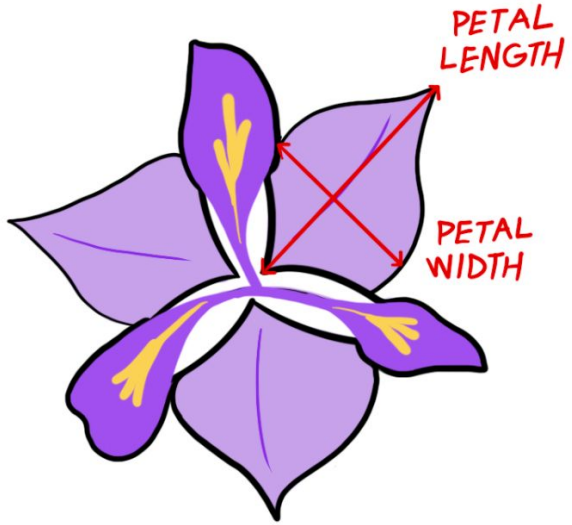
ML + MLOps + Kubeflow

Let's Use Machine Learning To Identify Iris (Dataset Explained)

Setosa



Versicolor



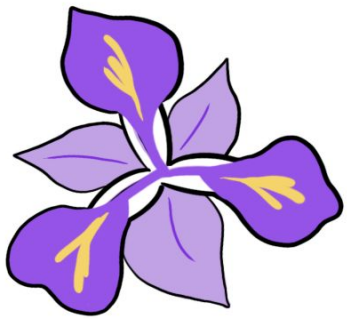
Virginica

Features:

- sepal length (cm)
- sepal width (cm)
- petal length (cm)
- petal width (cm)

Problem Statement

Problem Instance



IRIS FLOWER

FEATURES

SEPAL
WIDTH
HEIGHT

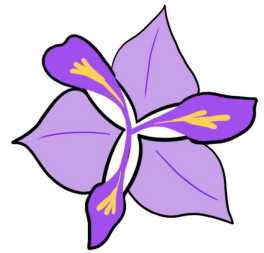
PETAL
WIDTH
HEIGHT

CLASSIFIER

PREDICT

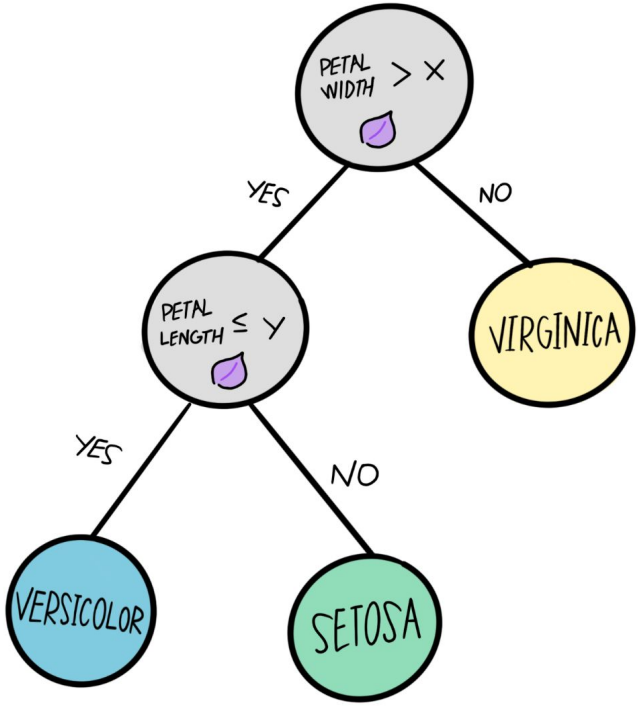
LABELS

Setosa
Versicolor
Virginica

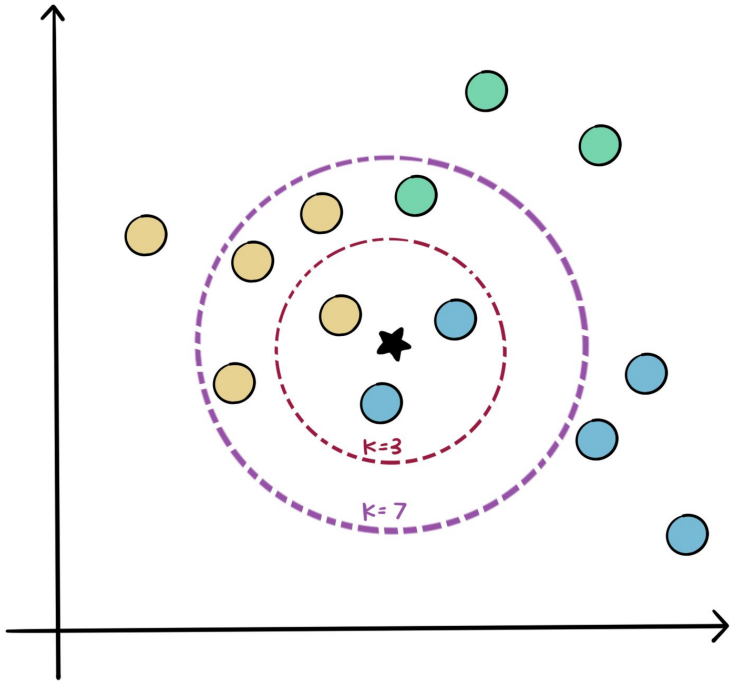


Algorithms Explained

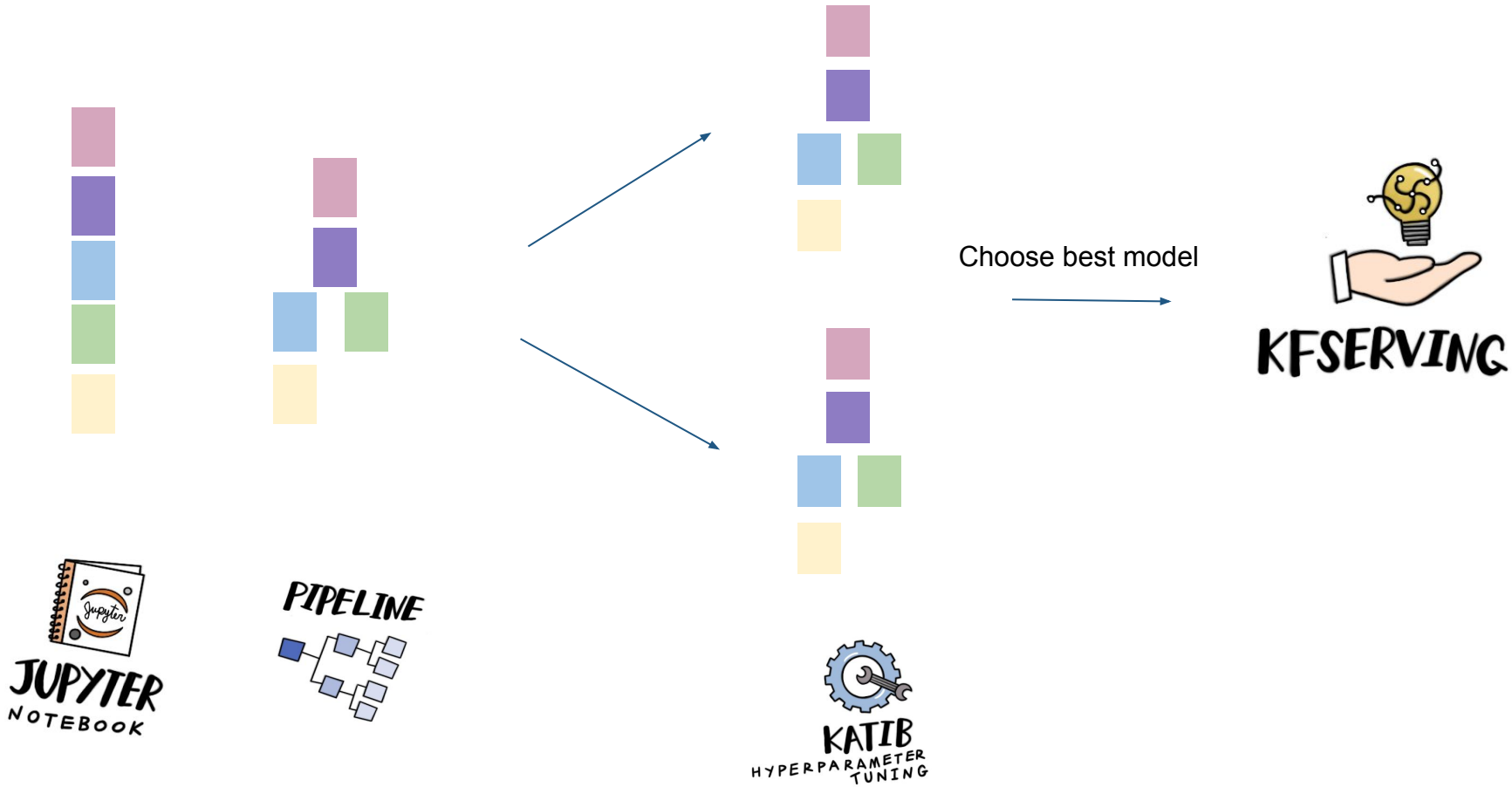
Decision Tree



K-Nearest Neighbours



Kubeflow ML WorkFlow



Hands-On Workshop

Build an automated machine learning pipeline!
Setup Guide: https://bit.ly/mlops-workshop-setup

Step 1

ML code to identify Iris

Step 2

Kubeflow Pipeline

Step 3

(optional) Explore other components of Kubeflow

Key Takeaways

- Open Source is a huge part of ML industry
- MLOps is the **hidden cost** of ML
- Kubeflow is an open source **MLOps platform**
- Try out Kubeflow for your next ML adventure
- Kubeflow **welcomes new contributors!**

Resources

- Workshop
  - Setup Guide https://bit.ly/mlops-workshop-setup
  - ML Code https://bit.ly/mlops-workshop-ml
  - Pipeline Code https://bit.ly/mlops-workshop-pipeline

- Kubeflow
  - Repo https://github.com/kubeflow
  - Docs https://www.kubeflow.org/docs/about/kubeflow/

THANK YOU!

DIANA AND ANNA WILL BE AVAILABLE ON **GHC OSD SLACK** FOR ANY QUESTIONS

CONNECT WITH US!

LINKEDIN

- Diana: https://www.linkedin.com/in/atanasova-diana/
- Anna: https://www.linkedin.com/in/antheajung/

TWITTER

- Anna: https://twitter.com/antheajung