# Deploy ML models

by Pratik Parmar aka Pintudo

GDG Baroda

Hello!

# I AM Pratik Parmar

And I am here to bore you with Machine Learning.

GDG Baroda

- App Idea: Dog Breed classifier

And dogs are like...

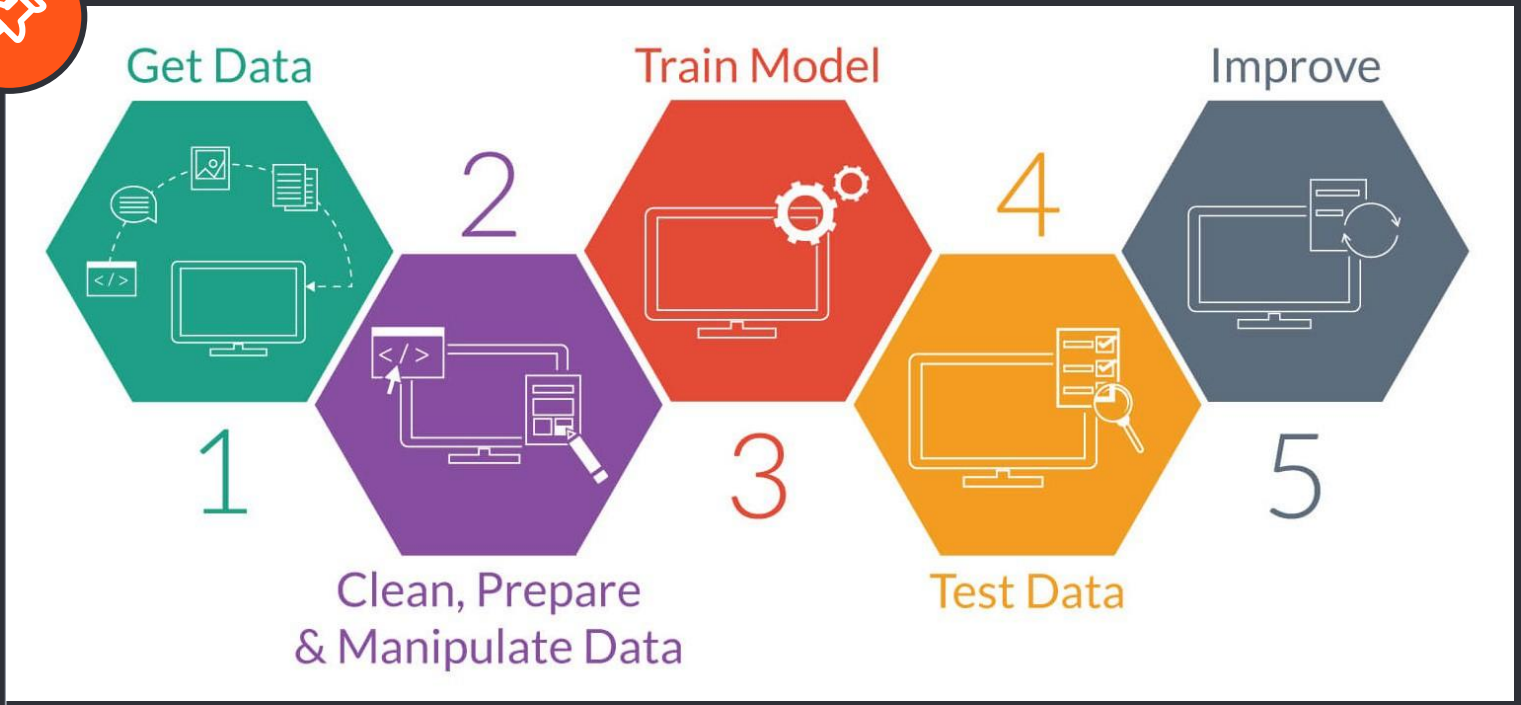
GDG Baroda

1

# Let’s talk about basic Machine Learning workflow

GDG Baroda

# General Machine Learning Workflow




GDG
Baroda

## After evaluation

The workflow for building machine learning models often ends at the evaluation stage:

you have achieved an acceptable accuracy, and "*ta-da! Mission Accomplished.*"


GDG Baroda

## Why ???

Maybe, going the extra mile to put your model into production is not always needed. And even when it is, this task is delegated to a system administrator.


GDG Baroda

*During my course at Coursera I was always asking myself—I have my model, which I can run in Jupyter Notebook and see the result, but what can I do with it? How can other use it?*

GDG Baroda

# Why Demo ??

No matter how silly your project is, demonstrating our work is generally great way to get a wider audience interest

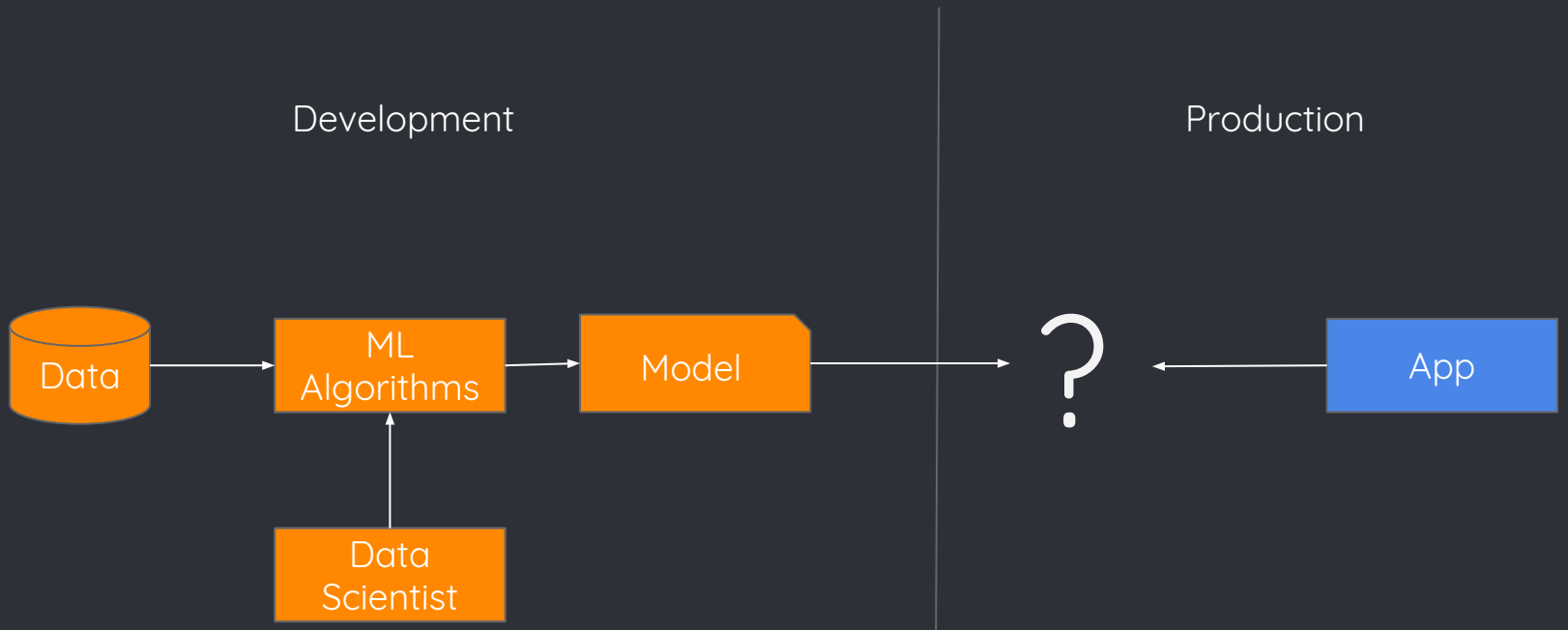

GDG Baroda

2

# How to deploy Machine Learning models?

GDG Baroda

# How to deploy a ML model ?




GDG Baroda

*Any guess*

GDG Baroda

# Serving

Serving is how you *apply* a ML model, after you've trained it

GDG Baroda

# 3 options to implement Machine Learning models

**Rewrite it**

Rewriting the whole code in the language that the software engineering folks work.

**API-first approach**

Create web API for your ML model using any web framework i.e. Flask or Django

**Tensorflow Serving**

TensorFlow Serving makes it easy to deploy new algorithms and experiments, while keeping the same server architecture and APIs. (API-first approach, but only for tensorflow)

GDG Baroda

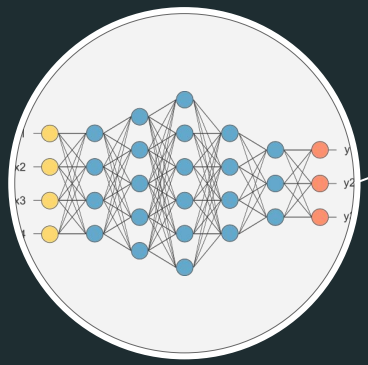
3

# Let’s get our hands dirty with the code

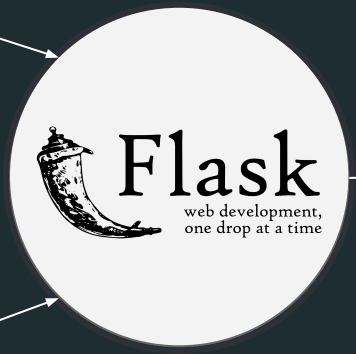
GDG Baroda

# Create simple Web API for Dog Image Classifier

Tensorflow



Pretrained Model




Flask


REST API

GDG Baroda

# Which model to choose?

GDG Baroda

# Pretrained model

For the sake of this session, we're gonna use retrained MobileNet Model

GDG Baroda

## Main steps to to deploy Model are

Train the model  and saving the checkpoints on the disk

Load saved model and test that it works properly

Export model into Protobuf format (*.pb*)

Create the client to issue requests and make an API

GDG Baroda

## Protobuf

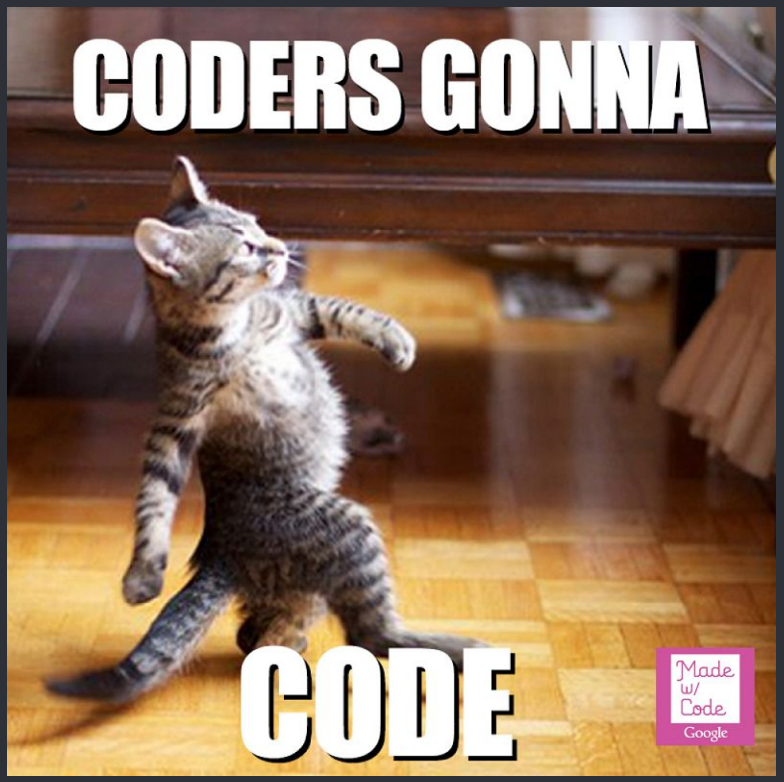
Protocol buffers are Google's language-neutral, platform-neutral, extensible mechanism for serializing structured data – think *XML*, but smaller, faster, and simpler.

## Export the model into Protobuf

Tensorflow serving provides *SavedModelBuild* class to save the model as Protobuf.

GDG Baroda

- Give some space because




GDG Baroda

- Machine Learning for everyone


GDG Baroda

- NO QUESTIONS PLEASE 👽

GDG Baroda

# Thank you for bearing me 

GDG Baroda