

# The Role of Containers in the Development & Deployment of ML models

## MLOps

Khalid El Yaacoub & Simone Bonato

# Our Agenda

1. Brief intro: What is ML and MLOps?
2. Common problems in ML Models deployment !
3. Poll! 📊
4. What is a “Container”? 🏠
5. Docker 🚢
6. Containers and ML models
7. Take-home message 🏠

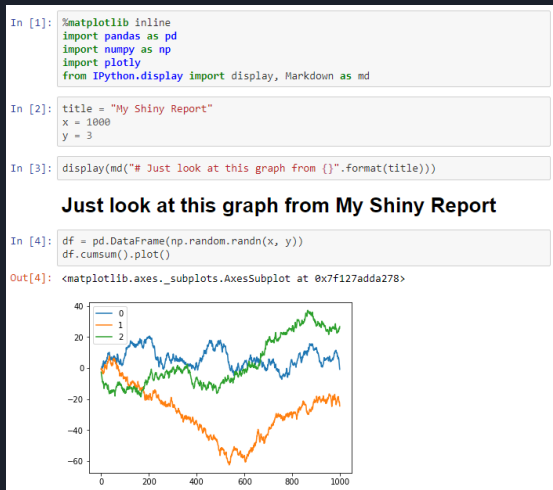


Scan the QR code or go to [menti.com](https://menti.com) and use code:

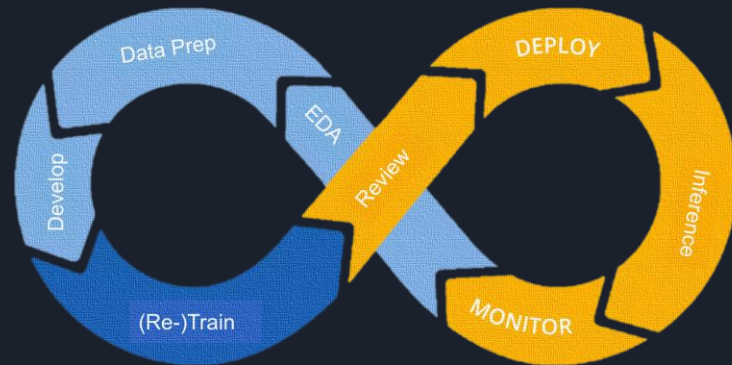
8564 7925

# What is ML and the reason for MLOps

- Get **INSIGHTS** from **DATA**
- **ML Pipeline** = mathematical representation of real-world process
  - Solve problems that are too complex for traditional methods
- The GAP: building models  $\neq$  deploying models in production systems
- **MLOps**: Facilitate the **deployment** of models into production
  - Avoiding issues



## MLOps cycle



# ML models deployment issues !

- **Build a model** on a machine  $\neq$  use model **at scale**, globally on a server
  - The “It works on my machine” paradox
- Model can be written in one language (E.g. Python), but has to interact with other apps written in other languages
  - E.g. data ingestion, data preparation, front-end, etc.
  - Need modularity
  - Docker! (Wait a bit for this)
- Models need to be retrained often:
  - Availability of new data
  - Outdated models
  - New requirements

## Top 20 Replies by Programmers when their programs don't work...

20. That's weird...
19. It's never done that before.
18. It worked yesterday.
17. How is that possible?
16. It must be a hardware problem.
15. What did you type in wrong to get it to crash?
14. There has to be something funky in your data.
13. I haven't touched that module in weeks!
12. You must have the wrong version.
11. It's just some unlucky coincidence.
10. I can't test everything!
9. THIS can't be the source of THAT.
8. It works, but it hasn't been tested.
7. Somebody must have changed my code.
6. Did you check for a virus on your system?
5. Even though it doesn't work, how does it feel?
4. You can't use that version on your system.
3. Why do you want to do it that way?
2. Where were you when the program blew up?
1. It works on my machine.



Poll time!



“What percentage of their time do Data Scientists spend deploying ML models?”



Scan the QR code or go to [menti.com](https://menti.com) and use code: **8564 7925**

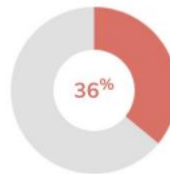
# Poll solutions

“ Although AI budgets are on the rise, only 22 percent of companies that use machine learning have successfully deployed an ML model into production. “

## What percentage of your data scientists' time is spent deploying ML models?



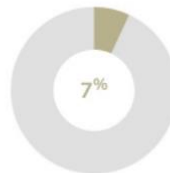
**36%** of survey participants said their data scientists spend **a quarter** of their time deploying ML models



**36%** of survey participants said their data scientists spend **a quarter to half** of their time deploying ML models



**20%** of survey participants said their data scientists spend **half to three-quarters** of their time deploying ML models



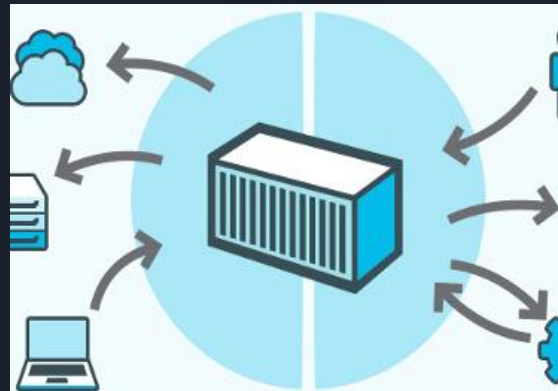
**7%** of survey participants said their data scientists spend **more than three-quarters** of their time deploying ML models

1% of respondents said they were unsure.

# What is a Container?



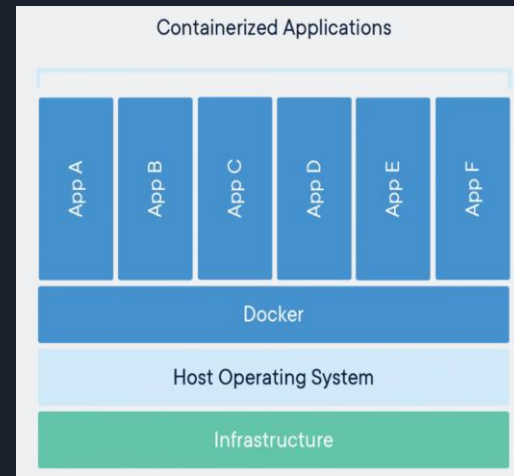
- Unit of software that packages up all the code and dependencies required to run an application
- **Isolate** application from working environment
- Ensure **reproducibility** in different development environments
- Portable, easily shared between developers making the development and deployment processes more efficient.




# Docker



- Docker is an open platform that uses OS-level virtualization to deliver software applications in packages called containers
- Consists of a Docker Engine, a portable lightweight packaging tool, and Docker Hub, a cloud service for sharing applications and automating workflows
- Docker enables applications to be assembled from the required components by separating application dependencies from infrastructure







# How Containers help ML model development & deployment

- Containers provide a way for ML developers to isolate environments from each other
- Advantages:
  - Remove central IT bottlenecks in the MLOps life cycle
  - Better collaboration for ML Engineers when sharing code and research
  - Old projects can be instantly reproduced and rerun
- Provide organizations flexibility and freedom in working across platforms and architectures

```
1 FROM python:3.7
2
3 RUN pip install virtualenv
4 ENV VIRTUAL_ENV=/venv
5 RUN virtualenv venv -p python3
6 ENV PATH="$VIRTUAL_ENV/bin:$PATH"
7
8 WORKDIR /app
9 ADD . /app
10
11 # Install dependencies:
12 RUN pip install -r requirements.txt
13
14 # Run the application:
15 CMD ["python", "app.py"] ]
```



# The take-home message



- Deploying ML models is more complicated than simply developing them.
  - E.g. building model at scale, flexibly update model with new data., etc..
- Containers allow us to package up all the code & dependencies required to run an application in a single unit of software
- Docker is an open platform that uses OS-level virtualization to deliver software applications using containers to eliminate the gap between environments

**“If it works on your machine, then it works on your colleague’s machine too!”**