Deploying your ML models

-Srikar Varanasi

We have a problem!
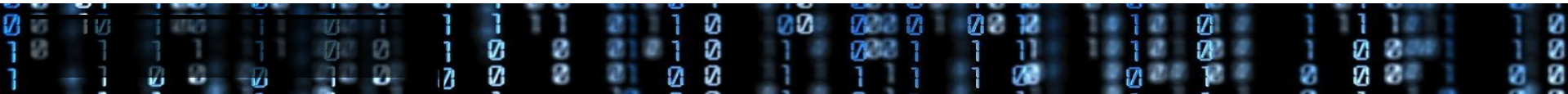
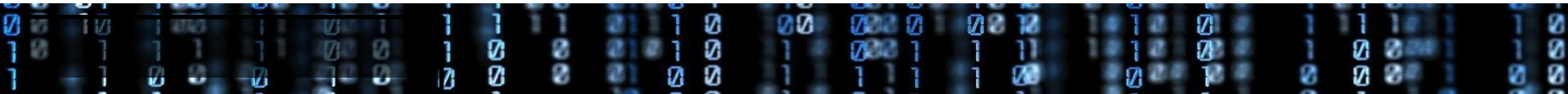# You've built a great Machine learning model. Now What?

Creating API

- API specifies how software components should interact.
- Separate the code and functionality
- Easy to scale
- Interacts with different languages (Python, Java etc.)

RESTful

- Set of constraints to be used for creating web services
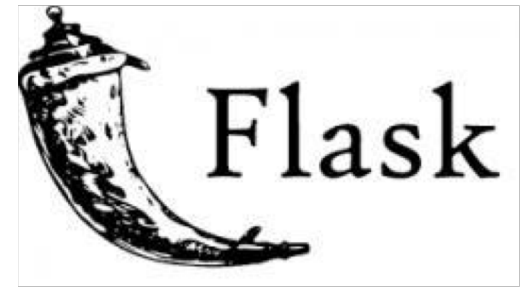- Server and client
- Stateless

GET /POST requests

➢ Way to get data - Usually in JSON or XML
➢ GET requests include all required data in the URL
➢ POST requests data in message body

• E.g. Get:
  c.com/test/demo_form?name1=value1&name2=value2

Using Flask

- Flask is a micro web-framework written in python
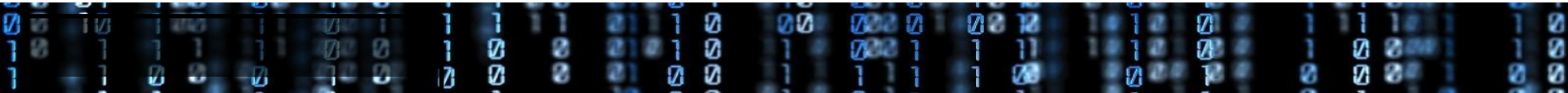- Flask is minimal

Process for building APIs for ML Models

- Build the Machine learning model
- Pickle or save the model
- Create a service
- Create request url(s) for the service
- Process the request and generate outputs using your model
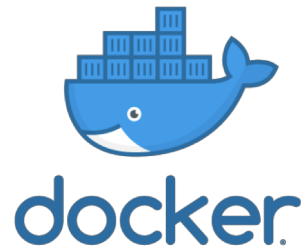- Send the appropriate response
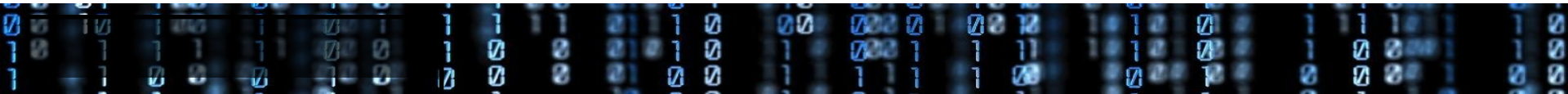- Test your API

Demo

Taking it further

1. Dockerizing your services
2. Additional frameworks - Django, Falcon
3. Amazon lambda

Resources

- Introduction to APIs
- Flask tutorial
- Docker Tutorial
- ML model as API
- Productize the models

Thank you