Documentation:

1. Model code
   1. Python pipeline class for modularization
      1. data prep code (for incoming test data)
         1. chars
         2. datatyping
         3. dummy vars
         4. imputing nas
         5. scaling
      2. Model binary – joblib + gzip
   2. Linting
   3. Testing
2. API
   1. Single row
   2. Batch
3. Containerization
   1. Docker
      1. https://xaviervasques.medium.com/quick-install-and-first-use-of-docker-327e88ef88c7
   2. Shell script
4. Optimization
   1. Scaling up
      1. <https://docs.docker.com/reference/cli/docker/volume/create/>
      2. Multiple containers can use the same VOLUME (store data/code there and spin up new containers for incoming events as needed) <https://towardsdatascience.com/build-and-run-a-docker-container-for-your-machine-learning-model-60209c2d7a7f>
   2. Kubernetes
5. Notes
   1. Why use joblib instead of pickle?
      1. It is a more general-purpose serialization library and is compatible with a wide range of Python objects, including custom objects and classes. **joblib is generally faster than pickle for serializing and deserializing large NumPy arrays, due to its more efficient storage and memory-mapping techniques**.

List of files + file structure:  
python\_model.txt

python\_tests.txt

python\_api.txt

data.pickle

dockerfile

documentation.pdf

requirements.txt

Kubernetes config.yml ??

Pipeline config

run\_api.sh