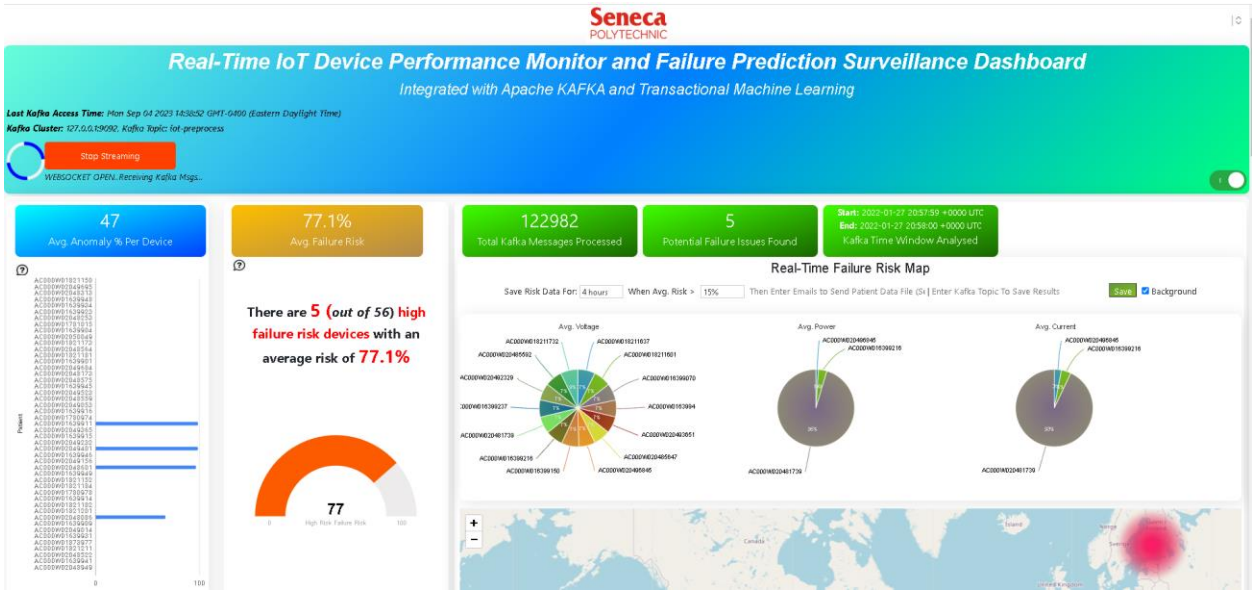
CSP 450 Fall 2023

PROJECT 5: Kubernetes, Docker, Python and TML Real-Time Dashboard Modifications

A. Make Modifications to TML Solution and Dashboard, Rebuild the Container, and Discuss Your New Solution

1. **Run your TML Dashboard:**
    a. **Open a browser i.e. Google Chrome**
    b. **Enter URL:** http://localhost:9005/iot-failure-seneca.html?topic=iot-preprocess2,iot-preprocess&offset=-1&groupid=&rollbackoffset=500&topictype=prediction&append=0&secure=1
2. **You should see:**



3. **Go to your raspberrypi repo**

    a. Open Python Script: **iotsolution-scripts-data/preprocess-iot-monitor-customdata.py** file

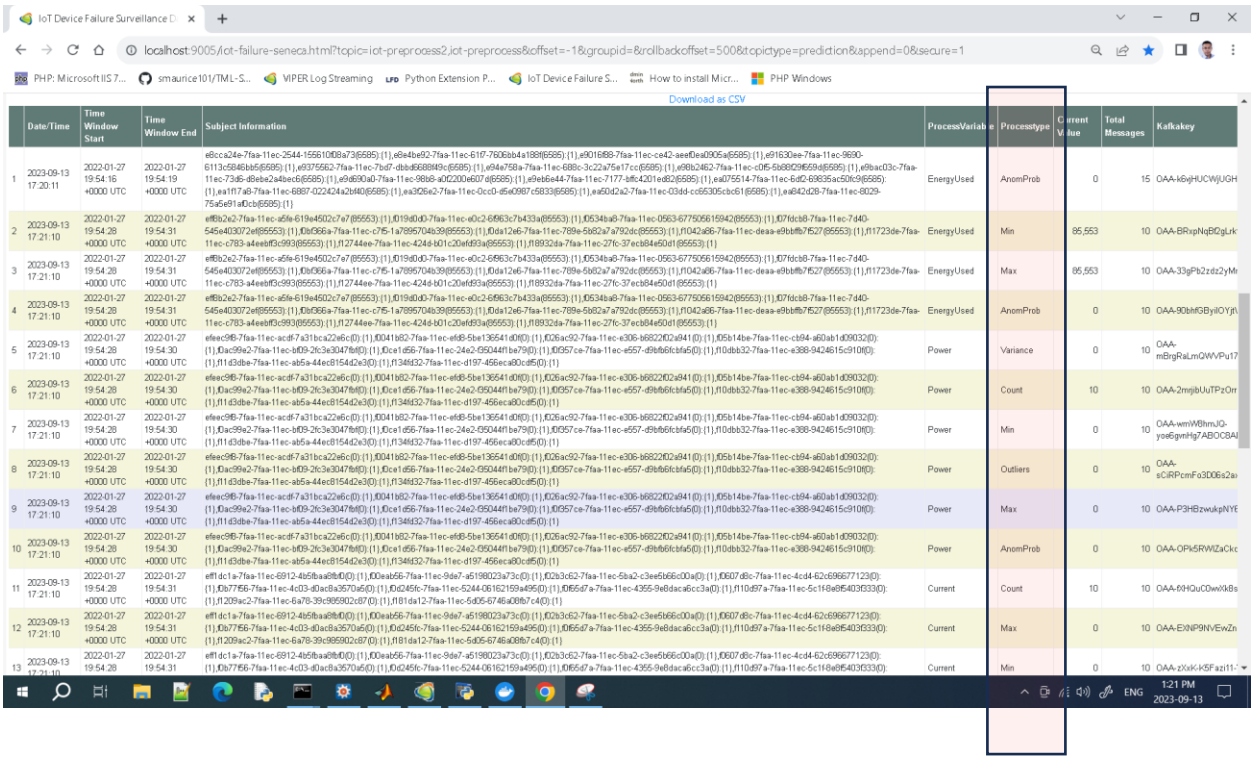
    b. Modify the preprocess types on **line 205 with the following:** MIN, MAX, COUNT, VARIANCE, OUTLIERS, ANOMPROB (shown below – then **Commit Change in Github**): For example:
        i. **preprocesslogic='MIN,MAX,COUNT,VARIANCE,OUTLIERS,ANOMPROB'**

4. **Create a new container and append "_modification" to the name of your container:**
   a. **Build your TML docker container:**
   a. NORMAL BUILD: docker build -t <your dockerhub id>/<container name>_modification --build-arg CHIP=AMD64 --no-cache --network=host .

   (**NOTE:** The '--no-cache' argument - this will build your container FROM SCRATCH (as you are doing now, and **takes about 20 minutes**)

5. **Push your container to Dockerhub:** docker push <your dockerhub id>/<container name>_modification

6. **Run your TML Dashboard and Show the Preprocess types in the TML Dashboard table:** MIN, MAX, COUNT, VARIANCE, OUTLIERS, ANOMPROB. For example:



7. **Explain what the preprocessing is doing to the data; explain the importance of data preprocessing.**

B. Kubernetes: Run Docker container in Kubernetes

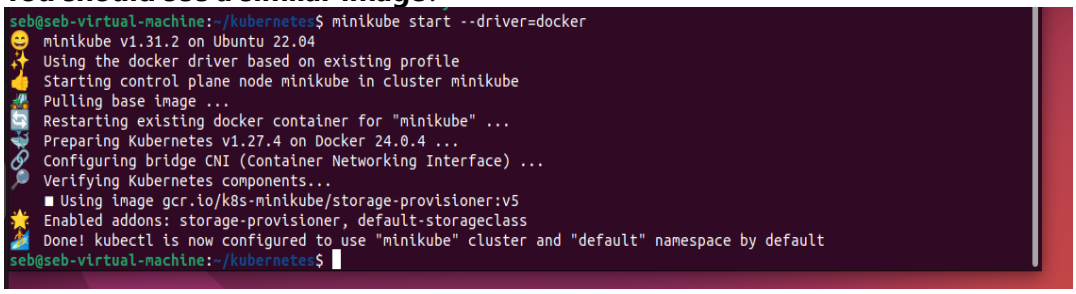
8. **Create a folder in your VM called kubernetes**
    a. **Note minikube is a ONE node Kubernetes cluster – it is the SAME functionality as a production grade Kubernetes cluster**
9. **cd to kubernetes folder**
10. **Go to: https://github.com/smaurice101/raspberrypi/blob/main/kubernetes/yaml**
    a. **Replace LINE 18 with your Docker image: image:** <your dockerhub id>/<container name>_modification
        i. **Commit the change**
    b. **Download yaml file to your kubernetes folder and rename it: senecaiot.yml**
11. **Now install Kubernetes and Kubectl:**
    a. **RUN:** `wget https://storage.googleapis.com/minikube/releases/latest/minikube-linux-amd64`
    b. **RUN: sudo install minikube-linux-amd64 minikube**

12. **Download: kubectl**
    a. `curl -LO https://storage.googleapis.com/kubernetes-release/release/`curl -s https://storage.googleapis.com/kubernetes-release/release/stable.txt`/bin/linux/amd64/kubectl`
    b. **RUN**: `sudo chmod +x kubectl`
    c. **RUN**: `sudo install -o root -g root -m 0755 kubectl /usr/local/bin/kubectl`

13. **Your kubernetes folder should look like:**

```
seb@seb-virtual-machine:~/kubernetes$ ls
kubectl   minikube   minikube-linux-amd64   senecaiot.yml
seb@seb-virtual-machine:~/kubernetes$
```

    a. **RUN Kubernetes: minikube start --driver=docker**
    b. `make sure docker engine is installed. If not run: sudo apt-get install docker.io`
    c. **RUN**: `sudo chmod 666 /var/run/docker.sock`

    d. **You should see a similar image**:
    e. 

14. **Create POD inside Kubernetes running your Docker Container**
    a. **RUN: kubectl apply -f senecaiot.yml**
    b. **RUN: kubectl get pods**

c. **You should see – You are now running your container inside Kubernetes:**

```
seb@seb-virtual-machine:~/kubernetes$ kubectl apply -f senecaiot.yml
deployment.apps/seneca-iot-deployment unchanged
seb@seb-virtual-machine:~/kubernetes$ kubectl get pods
NAME                                     READY   STATUS    RESTARTS      AGE
seneca-iot-deployment-78757d978d-czht5   1/1     Running   2 (3m6s ago)   20h
seb@seb-virtual-machine:~/kubernetes$
```

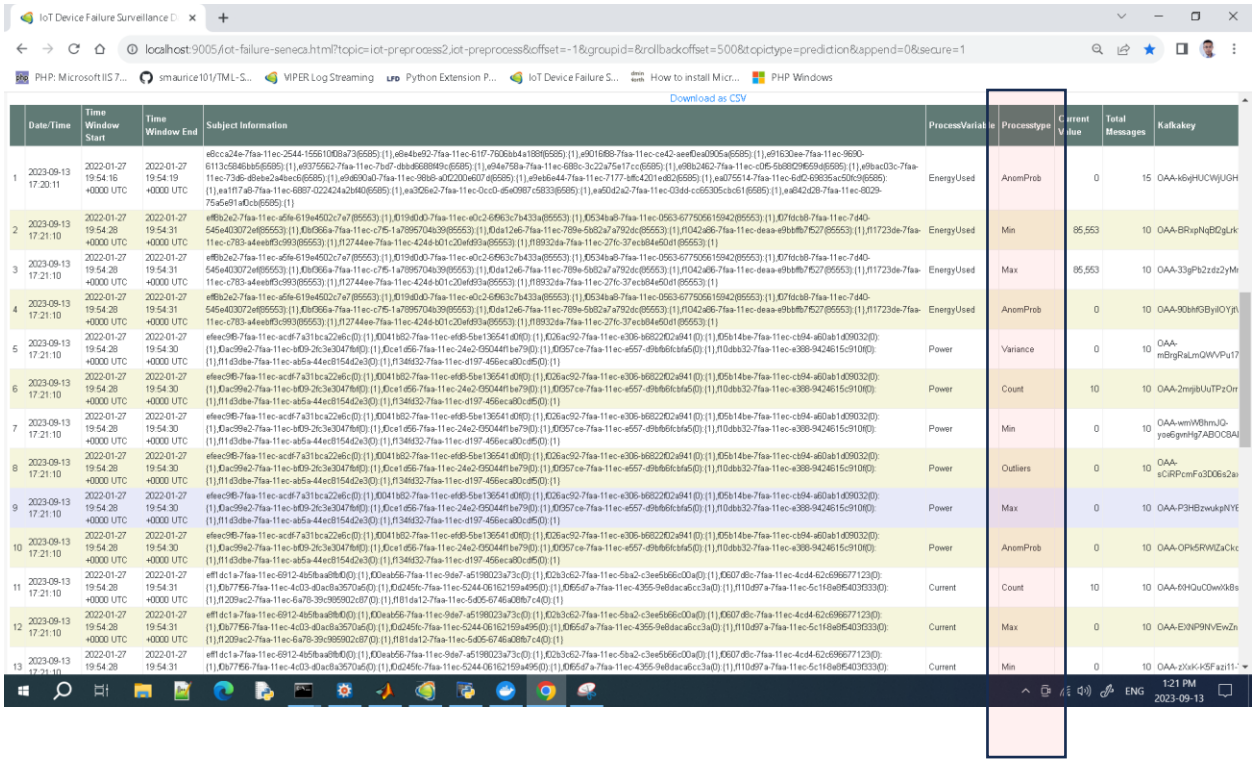
15. PORT Forward 9005:

a. **RUN: kubectl port-forward <pod name> 9005:9005**

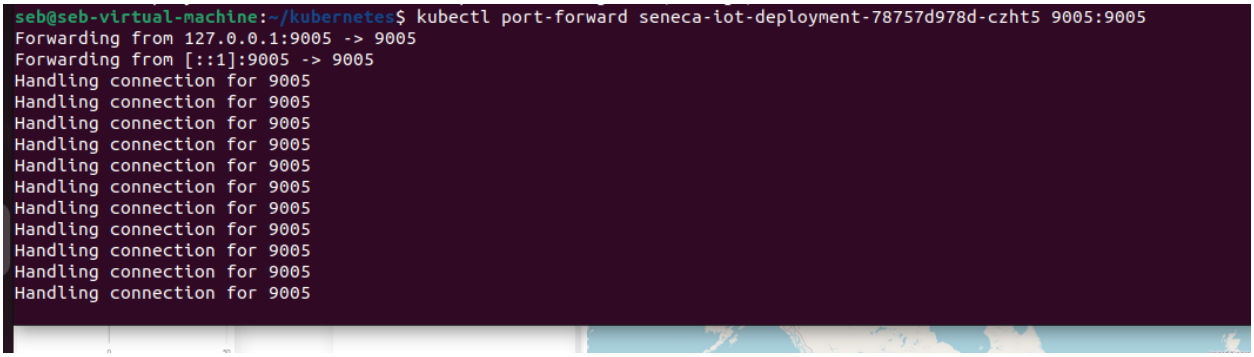
i. **For example: kubectl port-forward seneca-iot-deployment-78757d978d-czht5 9005:9005**

```
seb@seb-virtual-machine:~/kubernetes$ kubectl get pods
NAME                                     READY   STATUS    RESTARTS      AGE
seneca-iot-deployment-78757d978d-czht5   1/1     Running   2 (12m ago)   20h
seb@seb-virtual-machine:~/kubernetes$ kubectl port-forward seneca-iot-deployment-78757d978d-czht5 9005:9005
Forwarding from 127.0.0.1:9005 -> 9005
Forwarding from [::1]:9005 -> 9005
```

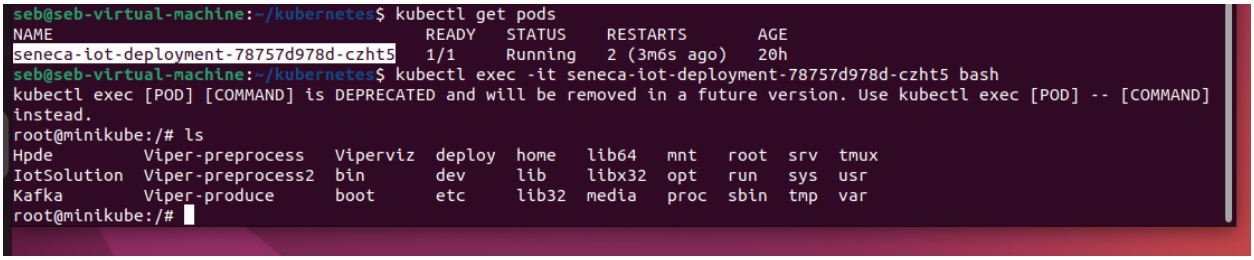
b. **RUN TML Dashboard: SEE SECTION A STEP 1 and take a screenshot as below:**



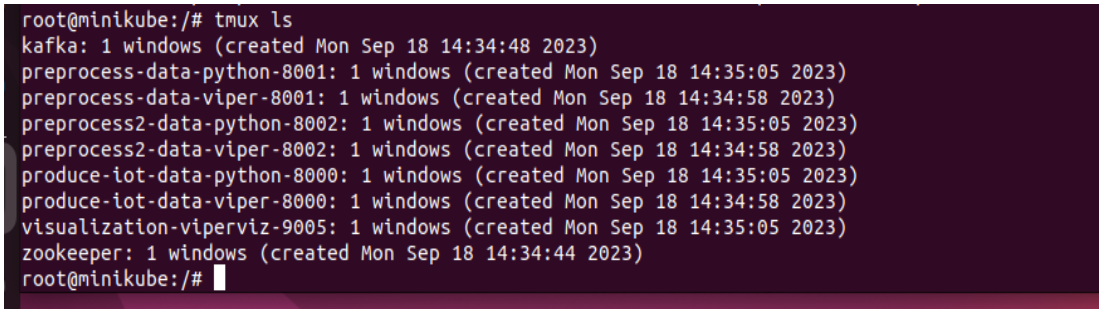
16. **Go to Port Forward terminal:**
    a. **You should see this:**



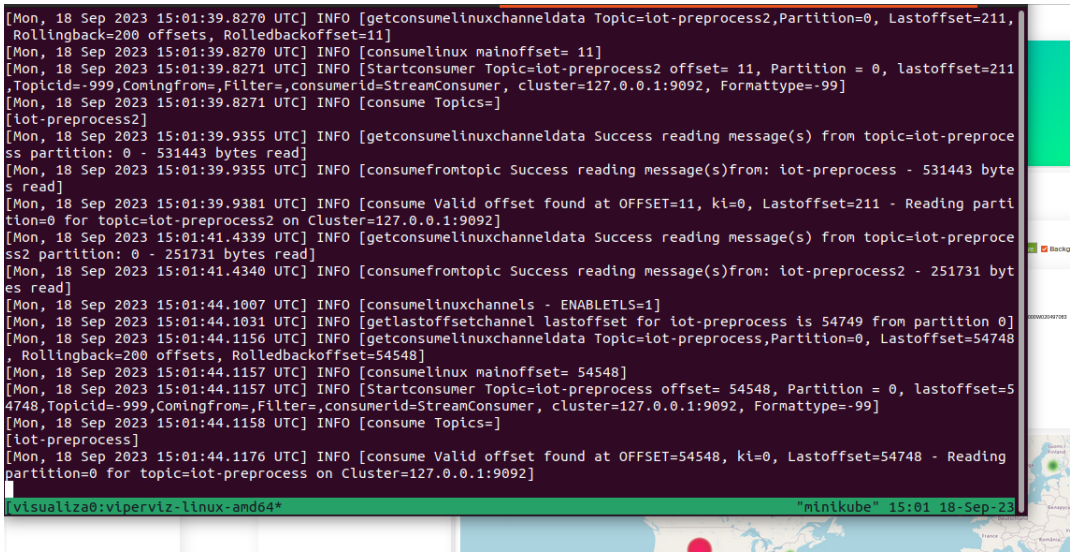
17. **Open new Terminal: Go inside the running container:**
    a. **RUN: kubectl exec -it <NAME> bash**
        i. **For example: kubectl exec -it seneca-iot-deployment-78757d978d-czht5 bash**
        ii. **You should see similar screen:**



    b. **RUN: tmux ls**
    c. **You should see:**



    d. **TMUX into visualization window:**
        i. **RUN: tmux a -t visualization-viperviz-9005**
        ii. **You should see:**



e. Exit out of tmux

i. **RUN: Ctrl+B, then D**

ii. **You should see:**

```
root@minikube:/# tmux a -t visualization-viperviz-9005
[detached (from session visualization-viperviz-9005)]
root@minikube:/#
```

f. Exit from container:

i. **RUN: exit**

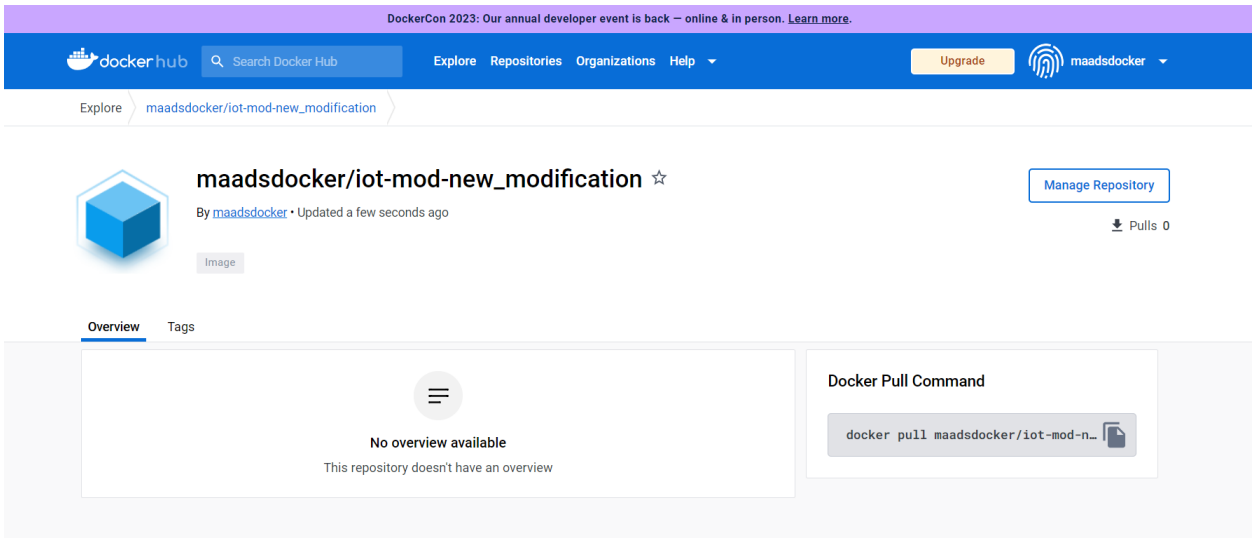
```
root@minikube:/# exit
exit
seb@seb-virtual-machine:~/kubernetes$
```

g. DONE!

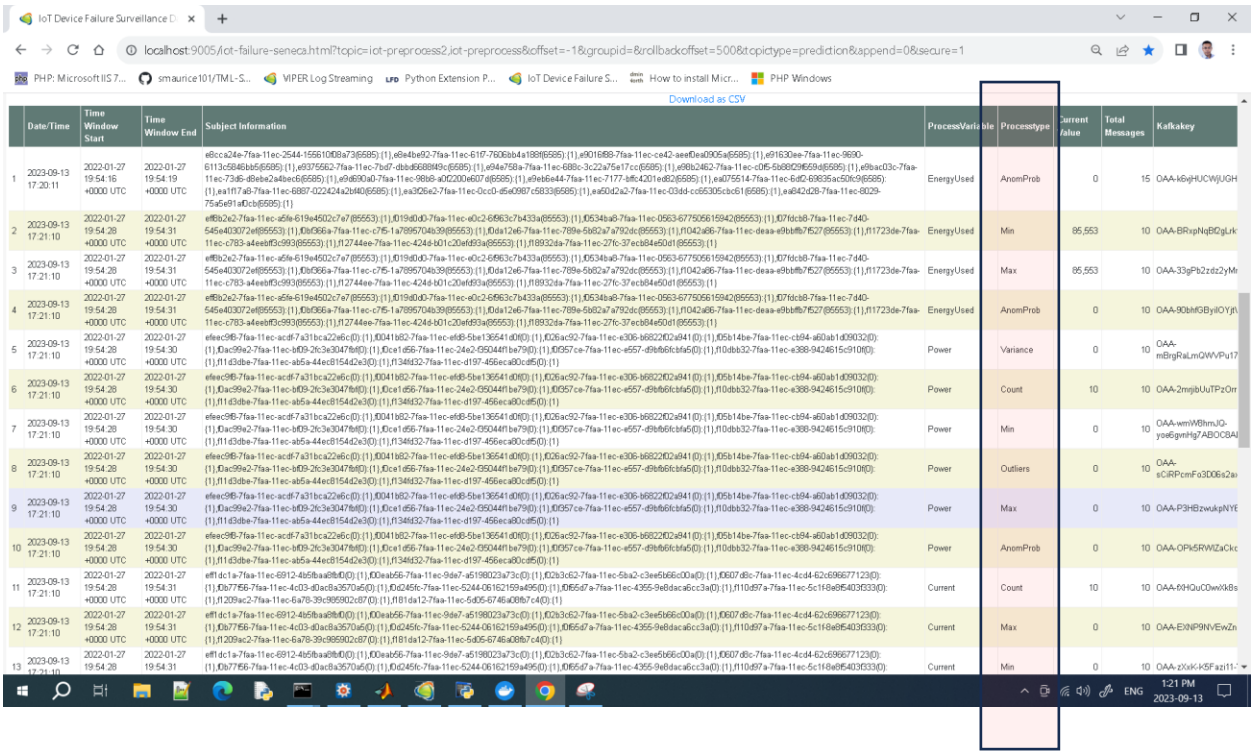
Submit ONE (1) PDF Document to Black Board for Grading containing ALL of the commands/screenshots below:

1. **Provide the Docker Pull command to your modified container from Docker Hub**
    a) For example: **docker pull maadsdocker/iot-mod-new_modification**
    b) **Replace maadsdocker with your DockerHub ID**

2. **Provide a screenshot of your container in Dockerhub: For example**



3. **Provide Screenshot of your dashboard running WITH the table showing each of the preprocess types in YOUR VM and machine browser:**
    a) Explain in **200 words or less** what the preprocess types are doing.



4. **Provide Screenshot of your running container running in YOUR VM or machine that shows your Dockerhub ID in the Docker Run command.** For example:

```
smaurice@DESKTOP-H0DIAMM:/mnt/c/MAADS/DOCKER/TML-Solution/docker/seneca$ docker run -p 9005:9005 maadsdocker/seneca-iot-
tml-kafka-amd64
Stopping MariaDB database server: mariadbd.
Starting MariaDB database server: mariadbd . ..
```

5. **Provide Screenshot of the containers AND Go inside your running container:**
   a) **Run**: docker ps
   b) You should see something similar to:

```
smaurice@DESKTOP-H0DIAMM:/mnt/c/MAADS/DOCKER/TML-Solution/docker/seneca$ docker ps
CONTAINER ID   IMAGE                                  COMMAND                  CREATED          STATUS          PORTS
                 NAMES
eb6bb34904b5   maadsdocker/seneca-iot-tml-kafka-amd64   "/bin/bash -c 'while…"   11 minutes ago   Up 11 minutes   0.0.0.
0:9005->9005/tcp   intelligent_curie
```

   c) **Take note of your Container ID – in the example – it is: eb6bb34904b5**
      1. **Run:** docker exec -it eb6bb34904b5 bash
      2. **Then type:** ls
      3. **Provide screenshot of inside your container - You should see:**

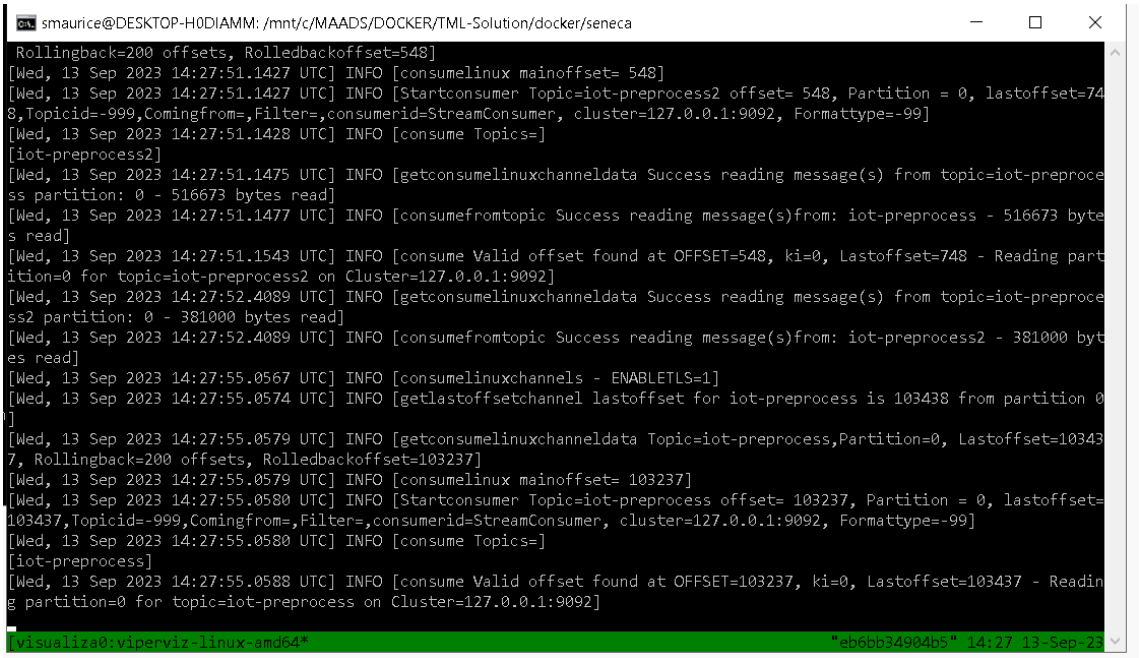
```
smaurice@DESKTOP-H0DIAMM:/mnt/c/MAADS/DOCKER/TML-Solution/docker/seneca$ docker exec -it eb6bb34904b5 bash
root@eb6bb34904b5:/# ls
Hpde          Viper-preprocess   Viperviz  deploy  home   lib64   mnt   root  srv  tmux
IotSolution   Viper-preprocess2  bin       dev     lib    libx32  opt   run   sys  usr
Kafka         Viper-produce      boot      etc     lib32  media   proc  sbin  tmp  var
```

   d) **Now type:** tmux ls
   e) **Provide screenshot of TMUX windows - You should see the following TMUX windows:**

```
root@eb6bb34904b5:/# tmux ls
kafka: 1 windows (created Wed Sep 13 14:05:38 2023)
preprocess-data-python-8001: 1 windows (created Wed Sep 13 14:06:02 2023)
preprocess-data-viper-8001: 1 windows (created Wed Sep 13 14:05:54 2023)
preprocess2-data-python-8002: 1 windows (created Wed Sep 13 14:06:02 2023)
preprocess2-data-viper-8002: 1 windows (created Wed Sep 13 14:05:54 2023)
produce-iot-data-python-8000: 1 windows (created Wed Sep 13 14:05:54 2023)
produce-iot-data-viper-8000: 1 windows (created Wed Sep 13 14:05:46 2023)
visualization-viperviz-9005: 1 windows (created Wed Sep 13 14:06:02 2023)
zookeeper: 1 windows (created Wed Sep 13 14:05:30 2023)
```

f) Provide screenshot of visualization-viperviz-9005 window: Go inside a TMUX window called: visualization-viperviz-9005

1. **Run: tmux a -t visualization-viperviz-9005**
2. **You should see:**



g) Provide screenshot of exit screen: Exit out of the TMUX window by pressing: CTRL+B, THEN D

You should see:

```
root@eb6bb34904b5:/# tmux a -t visualization-viperviz-9005
[detached (from session visualization-viperviz-9005)]
root@eb6bb34904b5:/#
```

h) Provide screenshot of exit container: Exit Out of the Container:

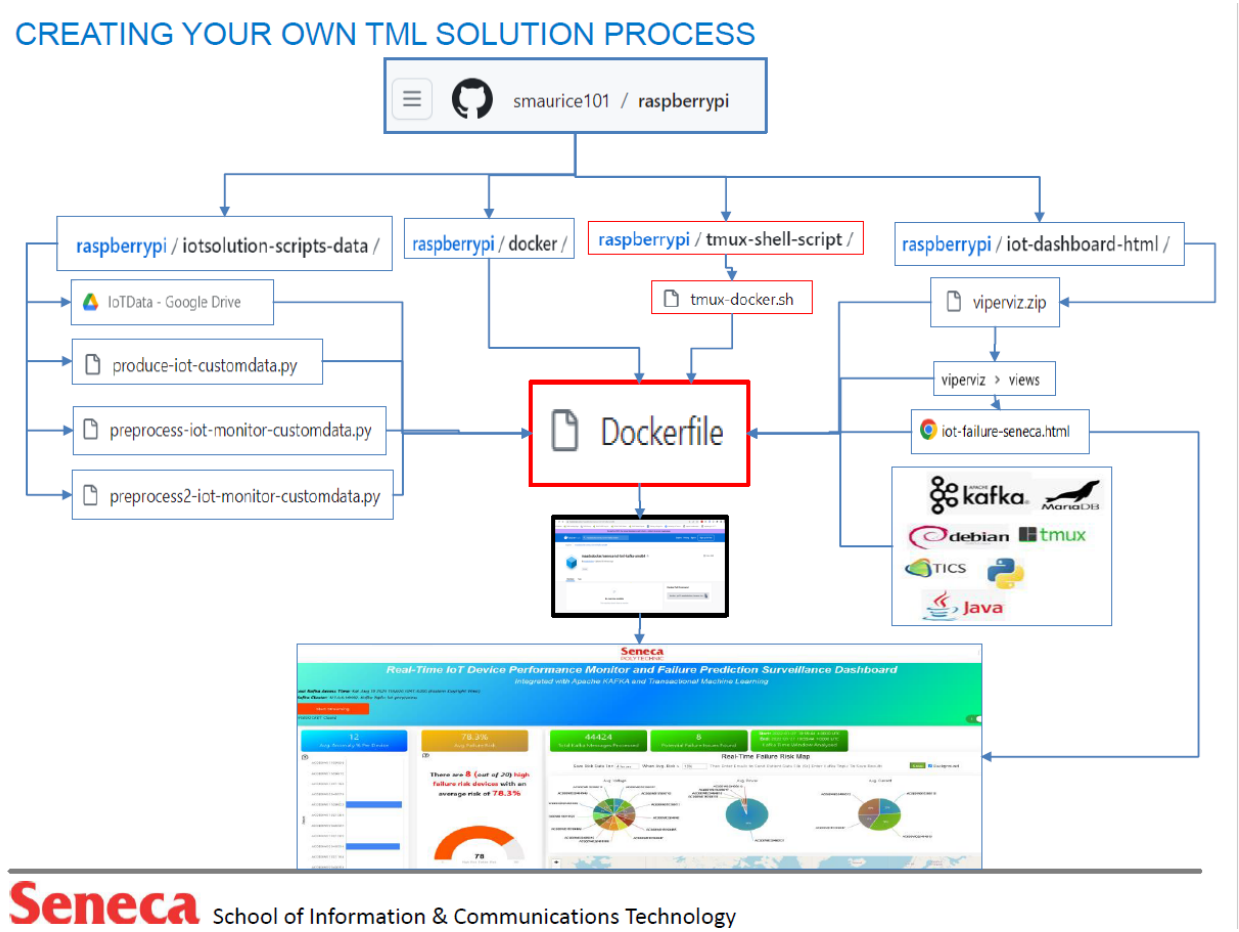
1. **You should see:**

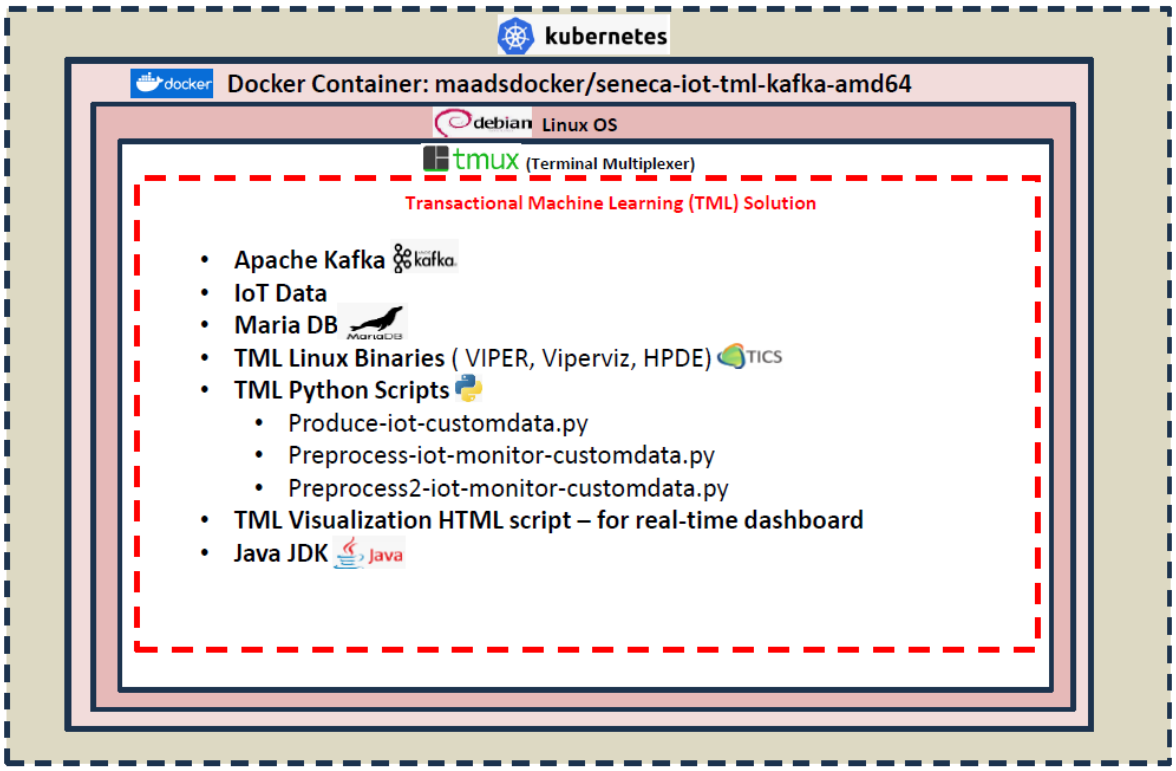
```
root@eb6bb34904b5:/# exit
exit
smaurice@DESKTOP-H0DIAMM:/mnt/c/MAADS/DOCKER/TML-Solution/docker/seneca$
```

6. SUBMIT ALL KUBERNETES STEPS/SCREENSHOTS in SECTION B and ALL STEPS and SUB-STEPS in 1-17

APPENDIX (for your reference)

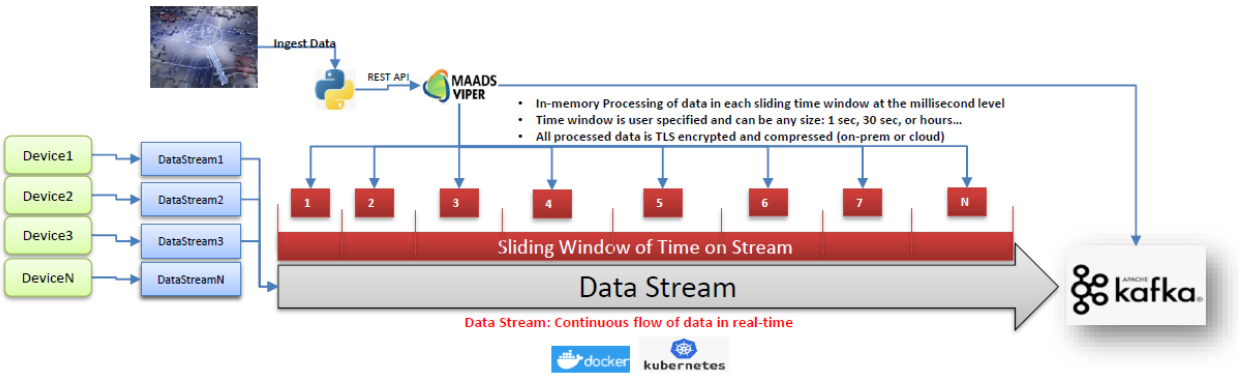
(Source: How TML Works.PDF)




kubernetes
Docker Container: maadsdocker/seneca-iot-tml-kafka-amd64
Linux OS
(Terminal Multiplexer)
Transactional Machine Learning (TML) Solution
Apache Kafka
IoT Data
Maria DB
TML Linux Binaries (VIPER, Viperviz, HPDE)
TML Python Scripts
Produce-iot-customdata.py
Preprocess-iot-monitor-customdata.py
Preprocess2-iot-monitor-customdata.py
TML Visualization HTML script – for real-time dashboard
Java JDK

How TML Preprocessing works?

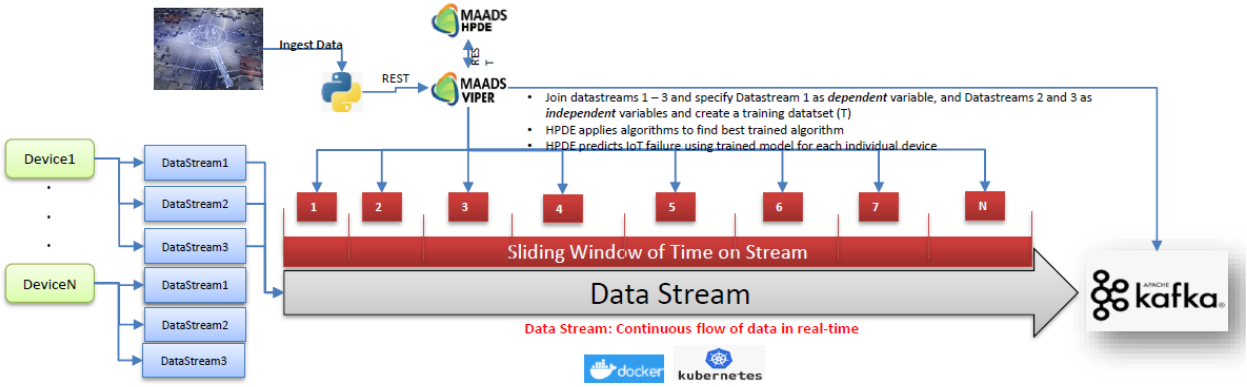
- All data streams from devices flow into Apache Kafka to a Kafka Topic
- TML performs **in-memory** processing of data in the Kafka Topic using TWO components across all **sliding time windows**
  - Python Script that uses the MAADSTML python library functions
  - MAADS-VIPER binary that can run in Linux, Mac, Windows (or other operating systems) on any Chip (32 or 64 bit) architecture (AMD, ARM, PPC, S390x, etc.)
- REST API connect MAADSTML python script to MAADS-VIPER
- 35+ different processing types: min, max, dataage, timediff, variance, anomaly prediction, outlier detection, etc…
- Apache Kafka is the central source of both input and output data – ***no external real-time database needed***
- ***No SQL queries are made for processing and machine learning***
- *Our technology can process unlimited number of devices (billions at high speed)*
- All TML solutions are containerized with Docker and scale with Kubernetes



How TML Machine Learning works?

- All data streams from devices flow into Apache Kafka to a Kafka Topic
- TML performs **in-memory** machine learning of data in the Kafka Topic by **joining data streams** using THREE components across all **sliding time windows**:
  - Python Script that uses the MAADSTML python library functions
  - MAADS-VIPER binary that can run in Linux, Mac, Windows (or other operating systems) on any Chip (32 or 64 bit) architecture (AMD, ARM, PPC, S390x, etc.)
  - MAADS-HPDE binary that can run in Linux, Mac, Windows (or other operating systems) on any Chip (32 or 64 bit) architecture (AMD, ARM, PPC, S390x, etc.)

- REST API connect MAADSTML python script to MAADS-VIPER and MAADS-HPDE
- 5 different algorithm types: logistic regression, linear regression, gradient boosting, neural networks, ridge regression
- Apache Kafka is the central source of both input and output data for estimated parameters – ***no external real-time database needed***
- ***TML auto-creates individual machine learning models for each Device at the "entity" level and joins datastreams 1-3 for each device and user specifies "Dependent" variable streams, and "Independent" variables streams***
- *Our technology can build unlimited machine learning models (billions at high speeds) for unlimited number of devices (billions at high speed)*
- All TML solutions are containerized with Docker and scale with Kubernetes



Additional TML Resources:

1. **TML Crash Course Videos:**
   https://github.com/smaurice101/raspberrypi/tree/main/TML%20Crash%20course/Videos

2. **TML Binaries:**
   a) https://github.com/smaurice101/transactionalmachinelearning

3. **MAADSTML Python Library:**
   a) https://pypi.org/project/maadstml/

4. **TML Blogs:**
   a) **Stream Processing/Analytics Tools Like Apache Flink is NOT Transactional Machine Learning**
   b) **Data Quality Checking in Data Streams**
   c) **A Fast and Simple Way To Migrate Data Streams Between Kafka Clusters: An Alternative to MirrorMaker2**
   d) **NFT (Ethereum) Price Prediction with Transactional Machine Learning, Kafka (or Redpanda) and Blockchain**
   e) **Detecting Medication Fraud at Scale with Transactional Machine Learning and Blockchain**
   f) **TML and Cybersecurity**
   g) **Contextualizing ChatGPT with Healthcare Data Streams**