

CSCE 438 - 838 Lab 6 Automating Things

Deadline by 8:29 am of October 18, 2024.

In this lab, we are going to develop tools to build a Machine Learning (ML) based automation system. You will use these to develop an end-to-end system. The instructions given below assume that you have a basic IoT Central application and a Python-based device implemented.

This lab will combine both Azure services used in labs 4 and 5 (Azure IoT Central and Azure IoT Hub).

A. Machine Learning

Machine learning is a technique of data science that helps computers learn from existing data to forecast future behaviors, outcomes, and trends. In this lab, we are going to locally train an ML model and deploy it to the cloud.

Getting the Data

We are going to use a machine learning model to make some predictions. So unlike in the previous lab, we can't use randomly generated data. We are going to use the [Rain in Australia dataset](#). This dataset contains about 10 years of daily weather observations from many locations across Australia. It also has the **RainTomorrow** column, which tells whether or not it rained the next day.

Use the above link and download the dataset from Kaggle.

Building a Predictive Model

We are going to train the model and save it as a [pickle](#) file in Python. It's a protocol to serialize Python objects and can be used to save the [skikit-learn](#).

1. Install Python Libraries

It is highly recommended that you use a virtual environment to manage dependencies.

Run the following command in terminal:

```
1 pip install pandas
2 pip install numpy
3 pip install scikit-learn
```

Depending on your environment, `pip3` may work instead of `pip` above.

2. Extract Features and Train the Model

The Python code given below processes the data through a series of operations and trains a logistic regression model using the processed data. It needs to execute from the directory where the dataset has been placed.

```
1 # Load all the libraries
2 import pandas as pd
3 import numpy as np
4
5 from sklearn.model_selection import train_test_split
6 from sklearn.preprocessing import StandardScaler
7 from sklearn.metrics import accuracy_score, classification_report
8 from sklearn.linear_model import LogisticRegression
9
10 # Load the data
```

```

11 dataset = 'weatherAUS.csv' #Give proper file location based on your system
12 rain = pd.read_csv(dataset)
13
14 # Reduce the cardinality of date by splitting it into year month and day
15 rain['Date'] = pd.to_datetime(rain['Date'])
16 rain['year'] = rain['Date'].dt.year
17 rain['month'] = rain['Date'].dt.month
18 rain['day'] = rain['Date'].dt.day
19 rain.drop('Date', axis = 1, inplace = True)
20
21 # Classify feature type
22 categorical_features = [
23     column_name
24     for column_name in rain.columns
25     if rain[column_name].dtype == 'O'
26 ]
27
28 numerical_features = [
29     column_name
30     for column_name in rain.columns
31     if rain[column_name].dtype != 'O'
32 ]
33
34 # Fill missing categorical values with the highest frequency value in the column
35 categorical_features_with_null = [
36     feature
37     for feature in categorical_features
38     if rain[feature].isnull().sum()
39 ]
40
41 for each_feature in categorical_features_with_null:
42     mode_val = rain[each_feature].mode()[0]
43     rain[each_feature].fillna(mode_val, inplace=True)
44
45 # Before treating the missing values in numerical values, treat the outliers
46 features_with_outliers = [
47     'MinTemp',
48     'MaxTemp',
49     'Rainfall',
50     'Evaporation',
51     'WindGustSpeed',
52     'WindSpeed9am',
53     'WindSpeed3pm',
54     'Humidity9am',
55     'Pressure9am',
56     'Pressure3pm',
57     'Temp9am',
58     'Temp3pm'
59 ]
60
61 for feature in features_with_outliers:
62     q1 = rain[feature].quantile(0.25)
63     q3 = rain[feature].quantile(0.75)
64     IQR = q3 - q1
65     lower_limit = q1 - (IQR * 1.5)
66     upper_limit = q3 + (IQR * 1.5)
67     rain.loc[rain[feature]<lower_limit, feature] = lower_limit
68     rain.loc[rain[feature]>upper_limit, feature] = upper_limit

```

```

69
70 # Treat missing values in numerical features
71 numerical_features_with_null = [
72     feature
73     for feature in numerical_features
74     if rain[feature].isnull().sum()
75 ]
76
77 for feature in numerical_features_with_null:
78     mean_value = rain[feature].mean()
79     rain[feature].fillna(mean_value, inplace=True)
80
81 # Encoding categorical values as integers
82 direction_encoding = {
83     'W': 0, 'WNW': 1, 'WSW': 2, 'NE': 3, 'NNW': 4,
84     'N': 5, 'NNE': 6, 'SW': 7, 'ENE': 8, 'SSE': 9,
85     'S': 10, 'NW': 11, 'SE': 12, 'ESE': 13, 'E': 14, 'SSW': 15
86 }
87
88 location_encoding = {
89     'Albury': 0,
90     'BadgerysCreek': 1,
91     'Cobar': 2,
92     'CoffsHarbour': 3,
93     'Moree': 4,
94     'Newcastle': 5,
95     'NorahHead': 6,
96     'NorfolkIsland': 7,
97     'Penrith': 8,
98     'Richmond': 9,
99     'Sydney': 10,
100    'SydneyAirport': 11,
101    'WaggaWagga': 12,
102    'Williamstown': 13,
103    'Wollongong': 14,
104    'Canberra': 15,
105    'Tuggeranong': 16,
106    'MountGinini': 17,
107    'Ballarat': 18,
108    'Bendigo': 19,
109    'Sale': 20,
110    'MelbourneAirport': 21,
111    'Melbourne': 22,
112    'Mildura': 23,
113    'Nhil': 24,
114    'Portland': 25,
115    'Watsonia': 26,
116    'Dartmoor': 27,
117    'Brisbane': 28,
118    'Cairns': 29,
119    'GoldCoast': 30,
120    'Townsville': 31,
121    'Adelaide': 32,
122    'MountGambier': 33,
123    'Nuriootpa': 34,
124    'Woomera': 35,
125    'Albany': 36,
126    'Witchcliffe': 37,

```

```

127     'PearceRAAF': 38,
128     'PerthAirport': 39,
129     'Perth': 40,
130     'SalmonGums': 41,
131     'Walpole': 42,
132     'Hobart': 43,
133     'Launceston': 44,
134     'AliceSprings': 45,
135     'Darwin': 46,
136     'Katherine': 47,
137     'Uluru': 48
138 }
139 boolean_encoding = {'No': 0, 'Yes': 1}
140
141 rain['RainToday'].replace(boolean_encoding, inplace = True)
142 rain['RainTomorrow'].replace(boolean_encoding, inplace = True)
143 rain['WindGustDir'].replace(direction_encoding, inplace = True)
144 rain['WindDir9am'].replace(direction_encoding, inplace = True)
145 rain['WindDir3pm'].replace(direction_encoding, inplace = True)
146 rain['Location'].replace(location_encoding, inplace = True)
147
148 # See the distribution of the dataset
149 print(rain['RainTomorrow'].value_counts())
150
151 # Split features and target value as X and Y
152 X = rain.drop(['RainTomorrow'], axis=1)
153 y = rain['RainTomorrow']
154
155 # Split training and test split
156 X_train, X_test, y_train, y_test = train_test_split(X, y, test_size = 0.2, random_state = 0)
157
158 # Scale input using just the training set to prevent bias
159 scaler = StandardScaler()
160 X_train = scaler.fit_transform(X_train)
161 X_test = scaler.transform(X_test)
162
163 # Train the model
164 classifier_logreg = LogisticRegression(solver='liblinear', random_state=0)
165 classifier_logreg.fit(X_train, y_train)
166
167 # Test the model accuracy
168 y_pred = classifier_logreg.predict(X_test)
169
170 print(f"Accuracy Score: {accuracy_score(y_test, y_pred)}")
171 print("Classification report", classification_report(y_test, y_pred))

```

3. Save the Model

Once the model has been trained, it can be saved using pickle library. Then, it can be loaded again when it needs to be used.

```

1 import pickle
2 with open("iot_model", "wb") as model_file:
3     pickle.dump(classifier_logreg, model_file)

```

Use the following snippet of code to load the model again:

```

1 with open("iot_model", "rb") as model_file:
2     model = pickle.load(model_file)
3 y_pred_new = classifier_logreg.predict(X_test)

```

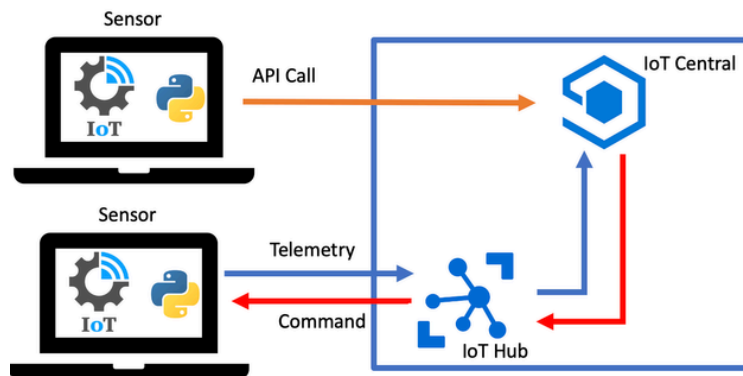
It can be useful to check your work using:

```
1 print("Output of loaded model is same as original model ?", all(y_pred == y_pred_new))
```

B. Programmatically Sending Commands to IoT Central Devices

In Lab 5, we sent IoT Central commands using the web application. IoT Central also allows using RESTful API to send commands. We are going to use this approach as it makes automating things much easier. In addition, we will use Azure portal's function as a service (FaaS) functionality. Therefore, the following sections combine functionalities of [Azure IoT Central](#), [Azure Portal](#), and local virtual IoT devices running python.

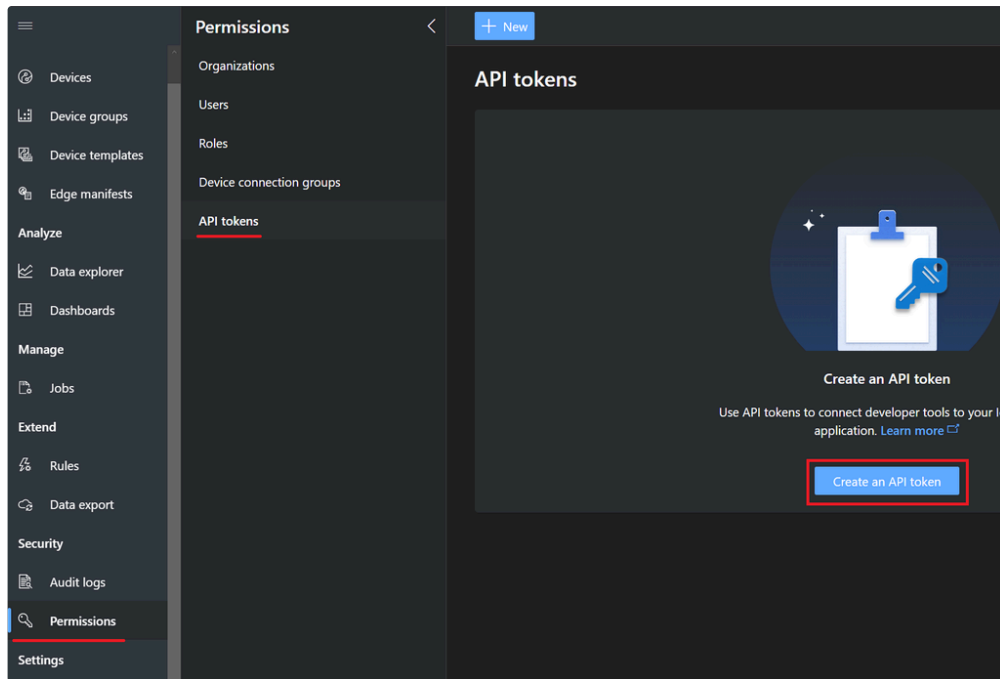
First, we will send API commands through IoT Central. Here is a diagram of what we will be doing:



If you don't have it already, install the requests library using `pip install requests`.

To send requests to IoT Central, you will need to use the API key. Go to <https://apps.azureiotcentral.com/> and do the following:

- Click 'My Apps' on the top left (third icon from the top).
- Select the IoT Central app you built in Lab 5.
- Navigate to `Permissions -> API tokens` and click `Create an API token`.



Now,

- Create a key using a token name (e.g., 'IoTazureML') and the '**App Operator**' role and then click 'Generate'.
- **Copy the 'Token' right away as you won't have access to it again.** This API key allows you to access your app in the IoT Central. Any devices within this app can be accessed.

We will then use the virtual IoT device you created in Lab 5 with the `SENSOR` (or whatever you named it) template. If necessary, create a new device using the instructions in Lab 5.

The template should have the following capabilities:

Display name	Name *	Capability type * ⓘ	Semantic type ⓘ		
Temperature	Temperature	Telemetry	Temperature	×	✓
LastTurnedOn	LastTurnedOn	Property	None	×	✓
SendData	SendData	Command		×	✓

The following snippet assumes you have IoT Central and a device. You should already have saved the `device_id` and `api_key`. For the `iotc_sub_domain`, take a look at the address field of your browser and copy the portion before `.azureiotcentral.com ...` in the address field. This is the name of your IoT Central App.

```

1 import requests
2
3 iotc_sub_domain = "your_domain"
4 device_id = "your_device_id"
5 api_key = "your_api_key"
6
7 def _command_url():
8     return f"<https://{iotc_sub_domain}.azureiotcentral.com/api/devices/{device_id}/commands/SendData?api-version=2022-05-31>"
9
10 def send_command():
11     resp = requests.post(

```

```

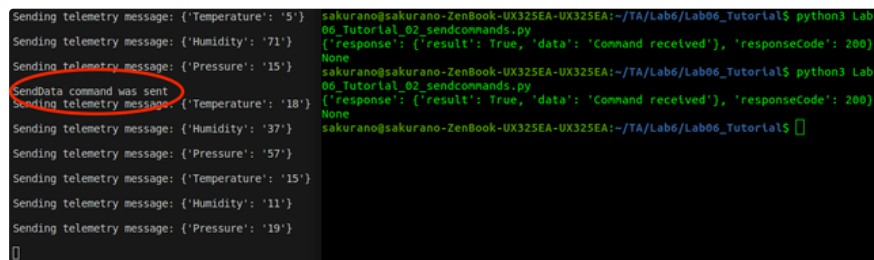
12     _command_url(),
13     json={},
14     headers={"Authorization": api_key}
15 )
16 print(resp.json())
17
18 print(send_command())

```

Please note that the sub domain MUST BE THE SAME as that is specified in `_def command_url()`: Because this API Call will look for the Command name you specified. As a default, API call is looking for the Name `SendData`. If you named the capability name as `SendData` in your template, then you do not need to modify the command url.

After making the necessary changes, you will now send commands from one virtual IoT device to another one. In a terminal window, run your Lab 5 `Lab05_IoTSample.py` code. This will start sending random temperature samples to IoT Central (and will listen for any commands)

Open another terminal window and run the above (modified) code `Lab06_Tutorial_02_sendcommands.py`. It will send commands to the former virtual IoT node.



```

Sending telemetry message: {'Temperature': '5'}
Sending telemetry message: {'Humidity': '71'}
Sending telemetry message: {'Pressure': '15'}
SendData command was sent
Sending telemetry message: {'Temperature': '18'}
Sending telemetry message: {'Humidity': '37'}
Sending telemetry message: {'Pressure': '57'}
Sending telemetry message: {'Temperature': '15'}
Sending telemetry message: {'Humidity': '11'}
Sending telemetry message: {'Pressure': '19'}
[]

sakurano@sakurano-ZenBook-UX325EA-UX325EA:~/TA/Lab6/Lab06_Tutorial$ python3 Lab06_Tutorial_02_sendcommands.py
{'response': {'result': True, 'data': 'Command received'}, 'responseCode': 200}
None
sakurano@sakurano-ZenBook-UX325EA-UX325EA:~/TA/Lab6/Lab06_Tutorial$ python3 Lab06_Tutorial_02_sendcommands.py
{'response': {'result': True, 'data': 'Command received'}, 'responseCode': 200}
None
sakurano@sakurano-ZenBook-UX325EA-UX325EA:~/TA/Lab6/Lab06_Tutorial$

```

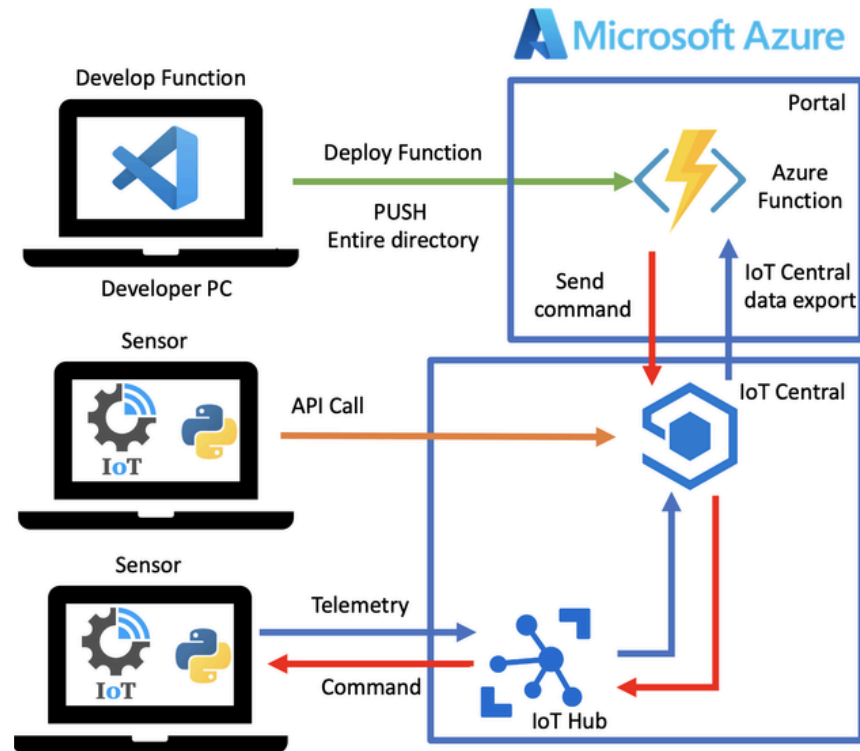
Now, you can send commands from one (virtual) IoT device to another through an API call to IoT Central!

Next, we will deploy a function in the cloud instead of running it locally.

C. Deploying Azure Functions

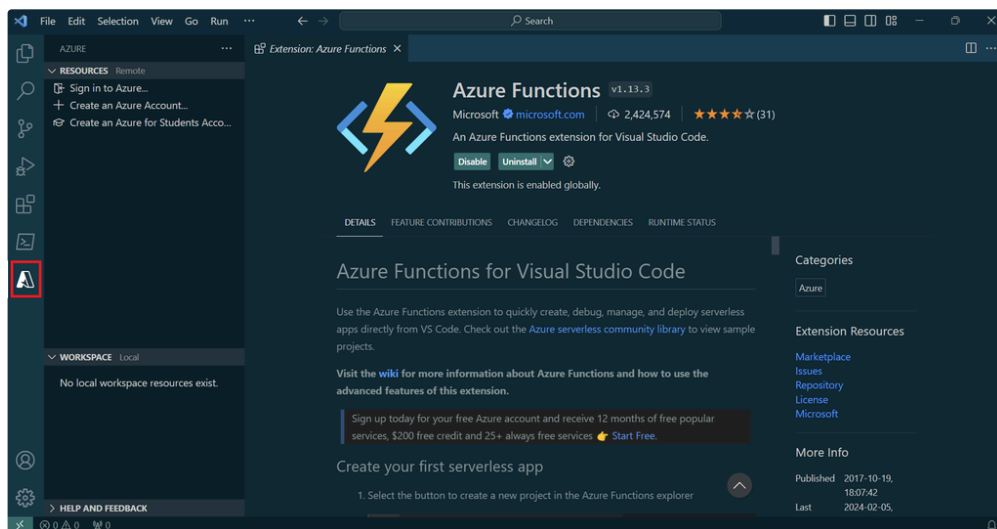
Azure Function is a function-as-a-service (FaaS) offering by Azure. It provides the runtime needed to run your application code without having to maintain the infrastructure. We will deploy a simple Python-based Azure function. We are going to use the Azure Function extension in Visual Studio Code to deploy the Azure function.

Here is a diagram of what we'll be working with:



1. If you haven't already, follow the instructions given in <https://code.visualstudio.com/docs/setup/setup-overview> to install VS Code.
2. Open VS Code and install 'Azure Functions for Visual Studio Code'. To do this, press **Ctrl or CMD + P** to open VS Code quick open and paste the following command:

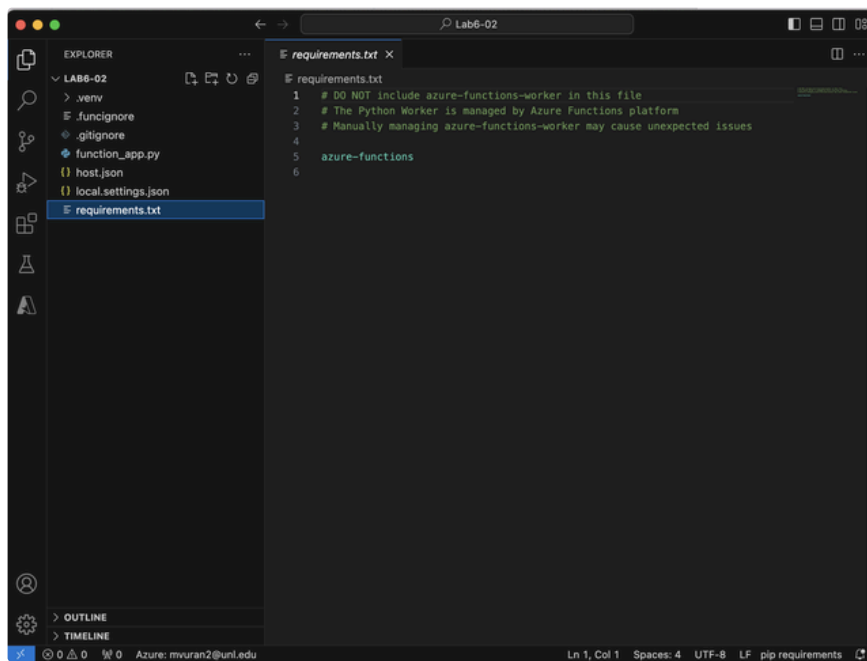
```
ext install ms-azuretools.vscode-azurefunctions
```
3. After installation is complete (may require you to reload VS Code), you will see the Azure icon on the Activity bar on the left.



- a. Click the Azure icon on the left and familiarize yourself with the interface. You will see two portions: **RESOURCES** on the top, and **WORKSPACE** at the bottom. We will go back and forth between those fields. Note that additional buttons emerge when you hover

your mouse icon on *RESOURCES* and *WORKSPACE*.

4. To get started, press 'Sign in to Azure' in the *RESOURCE* tab and complete the authentication process.
5. Now, hover over *WORKSPACE* and click the Azure Functions icon (lightning-looking icon) and then click `Create a new project`. Then, select the following options:
 - **Language:** Python
 - **Programming Model:** Latest version
 - **Interpreter:** Latest version installed in your system
 - **Template for the project:** HTTPTrigger
 - **HTTP Trigger name:** HTTPTrigger
 - **Authorization Level:** Function
6. You should have a folder structure like the one given below (you may want to click *Explorer*, the top icon on the left tab).



7. We will focus on the automatically generated `function_app.py`. Add a line to log the `req_body` `logging.info(req_body)` in the `function_app.py` file. The code should look like this:

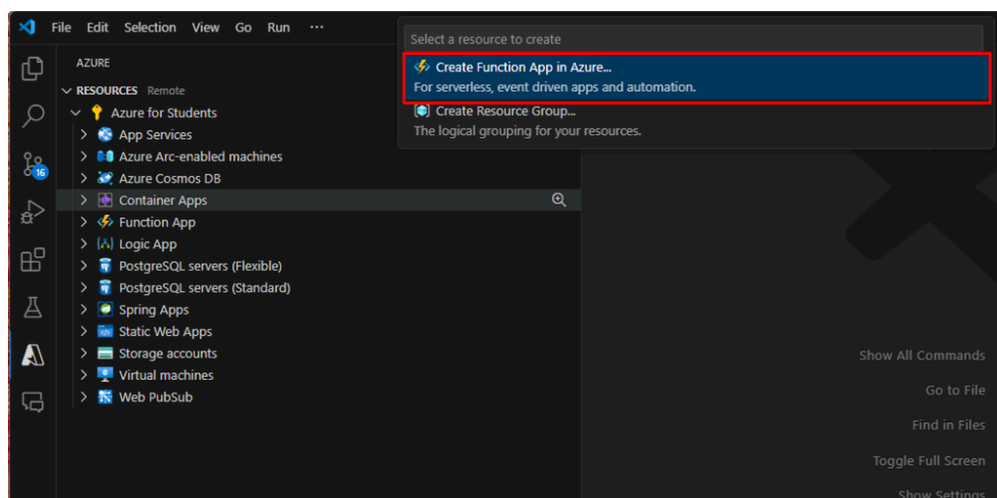
```
1 import azure.functions as func
2 import logging
3
4 app = func.FunctionApp(http_auth_level=func.AuthLevel.FUNCTION)
5
6 @app.route(route="HTTPTrigger")
7 def HTTPTrigger(req: func.HttpRequest) -> func.HttpResponse:
8     logging.info('Python HTTP trigger function processed a request.')
9
10     name = req.params.get('name')
11     if not name:
12         try:
13             req_body = req.get_json()
14             logging.info(req_body)
```

```

15     except ValueError:
16         pass
17     else:
18         name = req_body.get('name')
19
20     if name:
21         return func.HttpResponse(f"Hello, {name}. This HTTP triggered function executed successfully.")
22     else:
23         return func.HttpResponse(
24             "This HTTP triggered function executed successfully. Pass a name in the query string or in the
request body for a personalized response.",
25             status_code=200
26         )

```

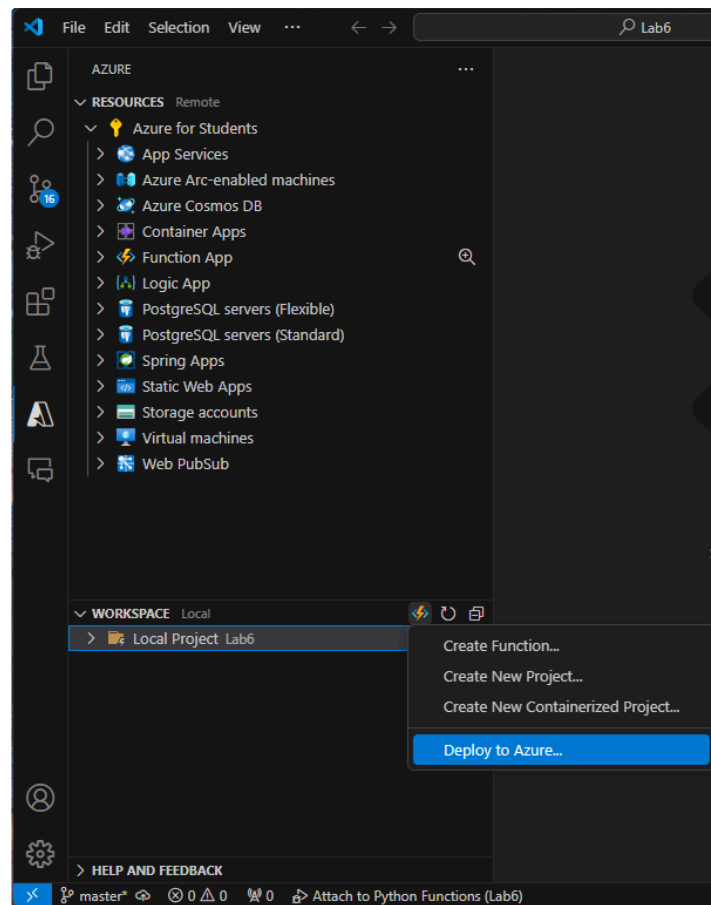
8. Now go back to the Azure tab and in the *RESOURCES* section, hover over the *RESOURCES* tab and select the '+' button, then select 'Create Function App in Azure'...



and set the following parameters:

- **Name:** anything
- **Stack:** Python 3.9
- **Location:** Central US

9. Now go to the *WORKSPACE* tab, click on the *Azure function logo*, click 'Deploy to Azure' and select the function you just created in step 7.



10 .You can visit [Azure Portal](#) and select 'Function App' from the left sidebar to view the function you just created. It should look something like this:

Home >

Function App

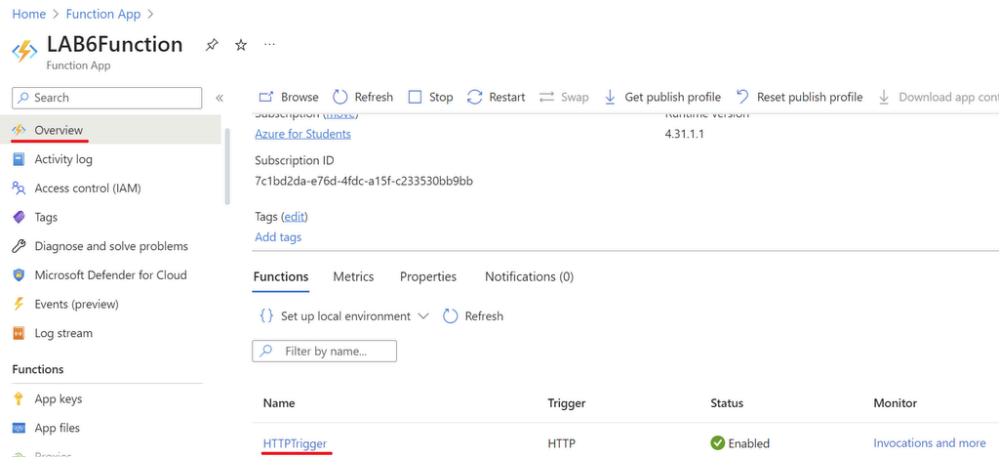
University of Nebraska-Lincoln (uofnelincoln.onmicrosoft.com)

+ Create Manage view Refresh Export to CSV Open query Assign tags Start Restart Stop Delete

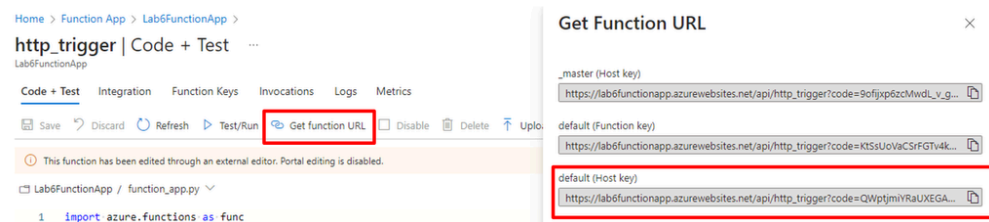
Filter for any field... Subscription equals all Resource group equals all Location equals all Add filter

Showing 1 to 1 of 1 records. No grouping List view

Name	Status	Location	Pricing Tier	App Service Plan	Subscription	App Type
LAB6Function	Running	Central US	Dynamic	ASP-LAB6Function-f5e4	Azure for Students	Function App



- Click on the function (*HTTP_Trigger*) from the above screenshot to get the **Default (Host Key) URL** with the API key. Navigate from `Functions -> HTTPTrigger -> Code+Test -> Get function URL`.

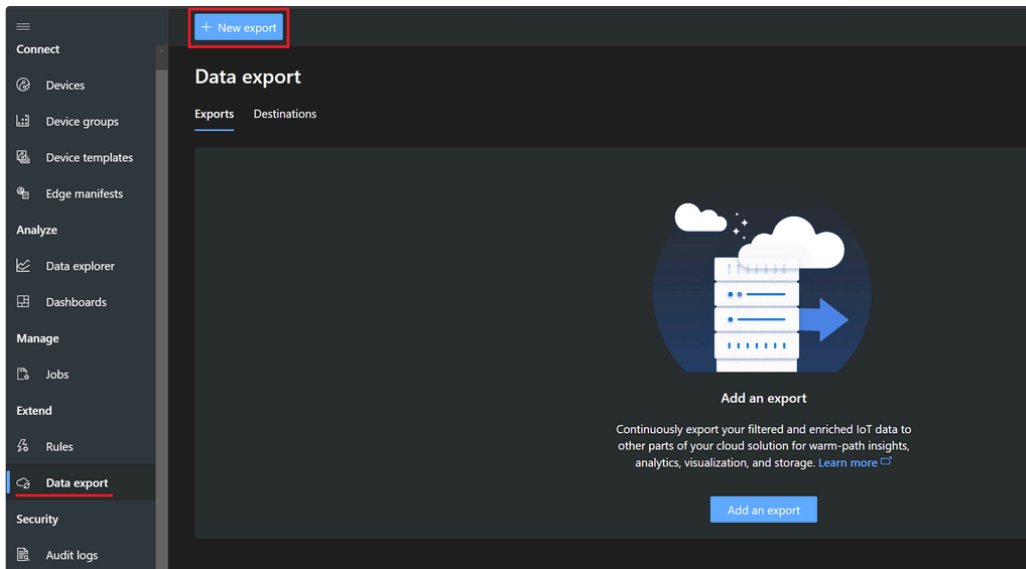


- You can monitor the logs of function execution in `Functions -> HTTPTrigger -> Monitor -> Logs`.
- If you need dependencies installed for your function to work, you can list them in the `requirements.txt` file.

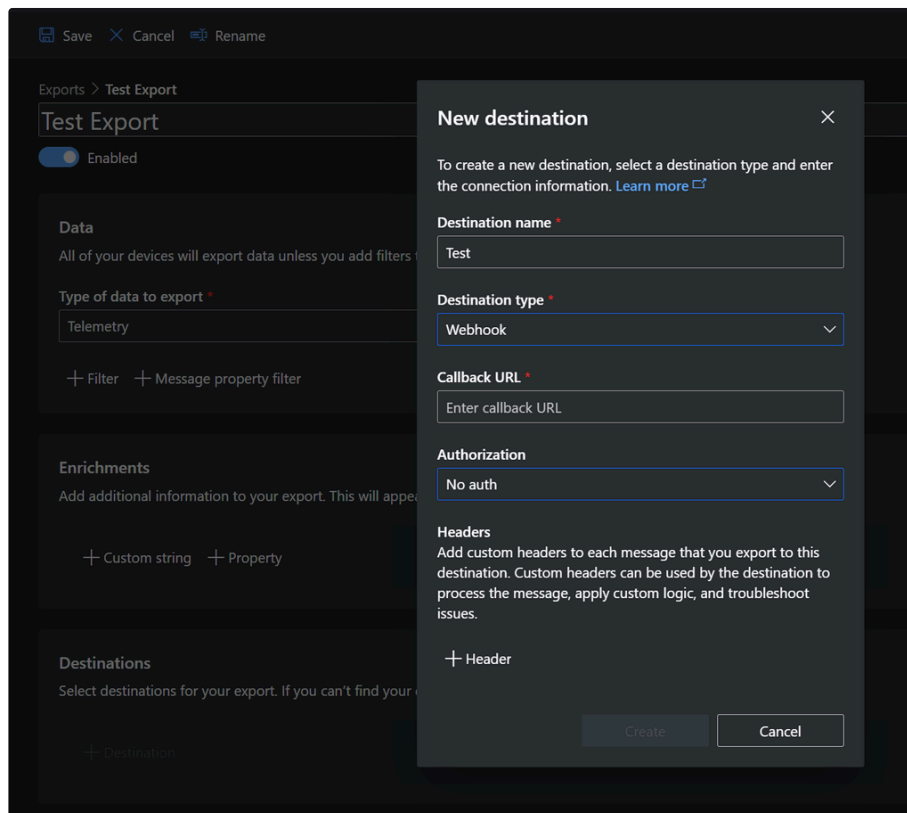
D. Exporting IoT Central to Azure Function

Now, we will link IoT Central to the Azure function we created.

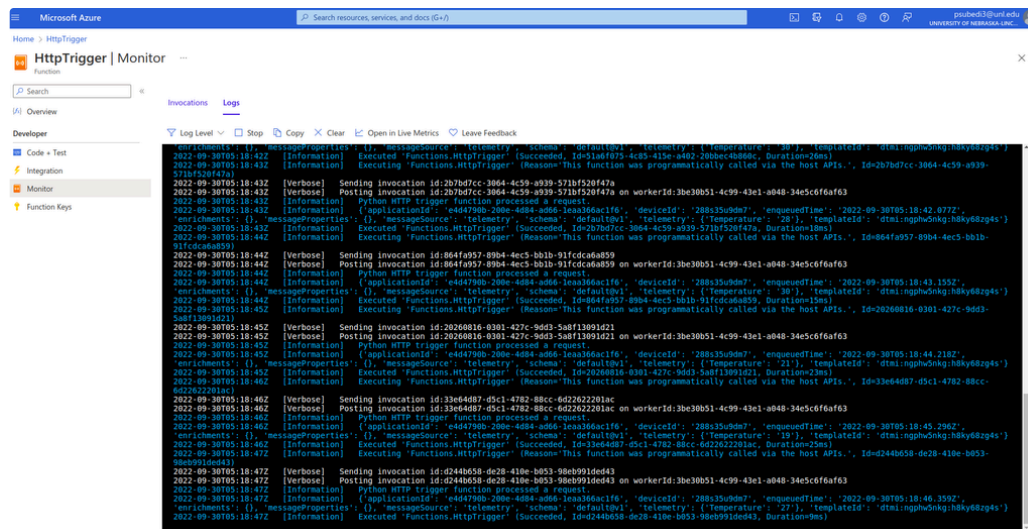
- Navigate to your custom application on [IoT Central](#) then go to the 'Data export' on the left sidebar and click '+ New Export' at the top.



2. Add any name to your export we are using "Test Export" as seen in the figure below. We will not be changing *Data* and *Enrichments*.
3. In the **Destination** section, click the highlighted '*create a new one*' text.
4. Select *Webhook* as the destination type and use the function URL from *Section C step 11* in the **callback URL** field.



5. Run your Python-based virtual IoT device from Lab 5. See the logs for the function. You should see the relevant logs in the console. It might take a while to show up after loading the screen.



Stop the app. Make sure to stop your Azure Function App to prevent exceeding your message limit. Go to your function app and click 'Stop'.

ASSIGNMENT

The task for this lab is to combine all 3 things you learned here to create an end-to-end system.

Requirements

1. Randomly split out 5% of the data from the Rain in Australia dataset before doing any processing. Use the rest of the dataset for training and testing and save the final model.
2. Create an IoT central application with a device template for devices that send data like the provided dataset. It should have a COMMAND defined.
3. Create an IoT Central device (Python based) that takes the 5% of the dataset not used for the model creation and sends data randomly to IoT Central.
4. Create another IoT Central device that only listens for a COMMAND.
5. Create an Azure function that runs on data exported from IoT Central. It should load the previously trained model and if the model predicts that it will rain tomorrow, send a command to the listening device.

Results

1. A screenshot and the Python code that fulfills each device requirement in this lab.
2. Screenshots from Azure for the completion of the requirements.

Report

- Development Process
 - Record your development process
 - **Acknowledge any resources that you found and helped you with your development (open-source projects/forum threads/books)**
 - Record the software/hardware bugs/pitfalls you had and your troubleshooting procedure.
- Results
 - Required results from the section above

Submission Instructions

1. Submit your lab on Canvas on or before the deadline.
2. Your submission should include one single PDF explaining everything that was asked in the tasks and screenshots, if any.
3. Your submission should also include all the code that you have worked on with proper documentation (Do not attach your code separately as an .ino file. Instead, copy and paste your code in the Appendix. Do not use screenshots in the Appendix.).
4. Failing to follow the instructions will make you lose points.

REFERENCES

<https://www.analyticsvidhya.com/blog/2021/06/predictive-modelling-rain-prediction-in-australia-with-python/>

 [Azure IoT Central REST API](#)

 [Azure Functions documentation](#)

<https://learn.microsoft.com/en-us/azure/azure-functions/create-first-function-vs-codepython>

Deadline by 8:29 am of October 18, 2024.