Lab: IoT Experiment Worksheet

|  |  |  |  |
| --- | --- | --- | --- |
| Name: | Gerald Oon Zhi Jian | Student ID: | A0262453H |

Submit this completed worksheet, modified Jupyter notebook and logged data (.csv) files in a single zip file on Canvas by the announced deadline. Your zip file should be named with your matriculation ID. For example, if your matriculation ID is A1234567Z, the filename will be ‘A1234567Z.zip’.

*The size of the box we allocated to the question hints at how long an answer you will need to give.*

|  |  |  |  |
| --- | --- | --- | --- |
| **Q1 – The pub-sub architecture is actually very prevalent in day-to-day activities, notably in social media and media consumption. Fill in the table below with one social media (e.g. Instagram, WeChat, X, Weibo) and one media (e.g. YouTube, Bilibili) platform as parallels to the pub-sub architecture. The magazine publisher analogy given in the summary is filled in as an example.** | | | |
|  | **Broker** | **Publish Action** | **Subscribe Action** |
| Example  **Magazine Publishing** | Magazine Publisher | Magazine writer sends magazine to publisher | Magazine reader subscribes to magazine through publisher |
| Social Media Platform  **Instagram** | **Instagram Platform** | Instagram content creator, creates content such as videos, reels, posts to the Instagram platform | Instagram audiences or followers will like and follow the content creator through the Instagram platform |
| Media Platform  **Youtube** | **YouTube Platform** | Youtube Content Creator creates and edit videos to upload to the youtube platform | The Viewers will click on the subscribe button and the notification bells icon to be alerted on new video post by the content creator |

Worksheet continues on the next page →

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Q2a – Plot the two sets of data you have collected (Graph 1: Brightness against Time; Graph 2: Temperature against Time). By visual inspection, the two sets of data appear to have some correlation. Provide a short comment to suggest a possible reason on why the two sets of data appear to be correlated.**  Graph 1    Graph 2    Comment: The two sets of data seem correlated because brightness and temperature goes linearly together, and it can be seen when it approaches peak hours such as high noon there is a spike in brightness and temperature. On the other hand, when it reach sunset hours the brightness and temperature decrease in the same pattern.  Also, fill in the following table:   |  |  | | --- | --- | | Average Sunrise Time | 04 : 23 AM | | Average Sunset Time | 09 : 47 PM | | Average Temperature in a Day | 25.03 °C | | Maximum Temperature | 31.3 °C | | Minimum Temperature | 21.30 °C | |

Worksheet continues on the next page →

|  |
| --- |
| **Q2b – Machine learning algorithms are described as learning a target function, , that best maps input variables, , to an output variable, such that .**  **Fit 3 suitable machine learning models fulfilling the following requirements:**  **Model 1: is historical values of temperature; is current temperature**  **Model 2: is historical values of brightness; is current temperature**  **Model 3: is historical values of brightness and temperature, is current temperature.**  *The algorithm used for each of the model need not be the same.*   1. *Model 1: Random Forest* 2. *Model 2: Decision Tree* 3. *Model 3: Random Forest*   **Plot curves of current temperature against time for ALL 3 models to illustrate how well each model fits the collected data. Note that we would also like to see the models’ prediction beyond the timestamp of the last collected datapoint to illustrate how well the models generalise.**  **Model 1: Random Forest – Fitted the data very nicely without much overfitting with good generalization for future predictions**    **Forecasted(Future Time stamps) – Beyond Last point Readings**  0 2022-09-17 19:51:00.698771 27.12641  1 2022-09-17 20:06:00.698771 27.13435  2 2022-09-17 20:21:00.698771 27.15062  3 2022-09-17 20:36:00.698771 27.15156  4 2022-09-17 20:51:00.698771 27.13841 |
| **Model 2: Decision Tree Algorithms**  **Model 2: X is historical values of brightness; Y is current temperature, write out the code for the program, using the decision tree algorithm, the models’ prediction beyond the timestamp of the last collected datapoint to illustrate how well the models generalise.**    Model 3: Random Forest (Regression) -  **is historical values of brightness and temperature, is current temperature.**      Future Temperature Predictions (Random Forest):  timestamp predicted\_temperature\_degC  0 2022-09-17 20:36:00.698771 26.79831  1 2022-09-17 21:36:00.698771 25.61804  2 2022-09-17 22:36:00.698771 23.37543  3 2022-09-17 23:36:00.698771 23.12085  4 2022-09-18 00:36:00.698771 22.87884  Summary Overview.  Comment: Model 3 fitted the best with high accuracy in prediction, when combining and learning from more features of the data with brightness and temperature(X), allowing more accurate predictions of temperature as shown in the above graph from the last timestamp. |