ELG 5142 Ubiquitous Sensing for Smart Cities

Assignment 3

Announcement Date: 2022-06-25

Submission Deadline: 2022-07-02

Teaching Assistant: Arda Onsu <monsu022@uottawa.ca>

In this assignment you are given a bunch of time series data that are collected from " Qbot-Qdrone Leader-follower Demo" from AVRS research studio.

In the demo, Qbot is manually controlled by user and Qdrone is supposed to follow it. We make a specific circular trajectory for the bot; however we intentionally inject some anomalies into the trajectory. For example, at some point, the Qbot will deviate from leave the trajectory and then rejoin.

Mainly, 2 different anomalies are injected into the trajectory of the QBot, and thus it supposedly reflects onto the trajectory of the QDrone as an anomaly:

- The drone will follow the bot by moving out of the trajectory temporarily. (Ex: video **case 1**)
- The bot enters a specific area that is a restricted region for the QDrone. (Ex: video **case 2**)

You will receive the Dataset from a 2-minute experiment. You will extract 4 attributes from the dataset: "follower x data", "follower y data", "leader x data" and "leader y data" ('x' and 'y' refers to coordinate). You are expected to implement and compare the performance of different machine learning algorithms and to detect those anomalies under the two demo scenarios.

You will implement 4 different unsupervised machine learning models:

1. SVM (binary)
2. KNN (unsupervised implementation)
3. PCA
4. DBSCAN

Then you need to **plot** the model results alongside with data and **compare** unsupervised models. You can use same code that you were presented in class (matplotlib.pyplot).

You are also provided the labeled version of this dataset. By using these labels, you are expected to compare the unsupervised algorithms with respect to accuracy, precision (for both anomaly and normal instances), recall (for both anomaly and normal instances) and F1 scores (for both anomaly and normal instances). You will provide a discussion on your conclusions regarding the anomaly detection performance of these algorithms for a mobile vehicular environment such as the AVRS test bed.

Deliverables and grading scheme:

1. SVM, KNN, PCA and DBSCAN implementation – Provide your implemented codes (40 pts)
2. TSNE plots of each output- provide in a report (20 pts)
3. Performance evaluation results (provide in a report) (25pts)
4. Conclusive discussions (provide in a report) (15pts)

Tips:

- You can use "**matplotlib.pyplot**" library for plotting your result. (*import matplotlib.pyplot as plt*)
- You can use "**pycaret.anomaly**" library for finding and applying anomaly detection machine learning algorithms (*from pycaret.anomaly import **). Or just by simply use "**pyod**" library (https://pyod.readthedocs.io/en/latest/) . (**helpful for first three model (KNN, PCA, SVM)**)
- You can look this web site to understand and implement pycaret library: https://towardsdatascience.com/time-series-anomaly-detection-with-pycaret-706a6e2b2427
- In order to install pycaret, open *anaconda terminal* write this command:
    - **pip install --user pycaret**
- Although the development platform is not restricted, you are recommended to use "jupyter notebook" in order to write your code and obtain the results.
    - In order to run jupyter, open *anaconda terminal* and type "**jupyter notebook**".