**Walchand College of Engineering, Sangli**

## **Machine Learning Lab (6CS372)**

**TY BTech | AY 2023-2024 | Even Sem**

**Assignment 8**

**Anomaly Detection using pyOD**

1. **Explore pyOD package**
   1. Follow documentation of pyOD and worked out examples over here <https://pyod.readthedocs.io/en/latest/example.html>

<https://www.analyticsvidhya.com/blog/2019/02/outlier-detection-python-pyod/>

* 1. You should understand the following:
     1. How data is generated (synthesized)?
     2. What is meaning of contamination?
     3. What are different detectors (algorithms) available in pyOD?
     4. How to apply single detector and evaluate its performance?
     5. How to apply multiple detectors and different methods of taking ensemble?

1. **Apply pyOD**
   1. Download dataset from here <https://github.com/GuansongPang/ADRepository-Anomaly-detection-datasets/blob/main/numerical%20data/DevNet%20datasets/creditcardfraud_normalised.tar.xz>
   2. Unzip and load in pandas.
   3. Perform necessary visualizations and pre-processing steps
   4. Apply at least 3 outlier detection algorithm from pyOD and take ensemble.
   5. Since this dataset is labelled, compare outliers given by step 3 and actual anomalies.
2. **Analysis:**
   1. Analyse the output of your code.
   2. Comment on coverage of anomalies by detectors.
   3. Can you come up with some visualizations for the anomalies?

**Deliverables:**

Jupyter Notebook or Python script containing the implementation.

Report on summarizing the findings from experimentation and analysis in notebook itself. Any additional visualizations or insights gained during the experimentation. This can go in markdown cell.

**Note**: Ensure proper documentation and comments throughout the code to make it understandable.