TVARIT

Senior Data Scientist - Assignment

Data description:

- **data.csv** file contains time series data collected from four temperature sensors installed in a machine producing metal parts.
- Each part can be identified using part_id column, and its production quality can be identified using label column. Data corresponding to each part can be extracted by grouping the data on part_id column.
- Each part can either be good (label = 0), or bad (label = 1).
- Each part takes 4 - 5 minutes to produce. The temperature values throughout this production are recorded in the four temperature columns given in the data.

Task:

1. Using above-defined temperature columns, prepare a dataset for predictive modelling. Each data point would be a window containing k consecutive parts (0, 1, 2, 3, .. kth part). The target label for each window would be the quality label of the kth part.
2. Feature Engineering: Extract the Max/Min/Std of each temperature sensor value in the window. These will be input features for our model.
3. Train a predictive model to predict the quality label of the kth wheel with the corresponding window features as input.
4. The goal is not to build the most accurate model. The goal of this assignment is to see how you approach this problem, to see your programming and storytelling/documentation skills.

Deadlines & submission process:

You are required to send your submissions via email to **shivani.kawade@tvarit.com** within 48 hours of receipt.