AI/ML Engineer (Fresher) Assignment

Time Series Anomaly Detection for IoT Sensors-SUMMARY

Name: Tharanesh P

Ph.no:9095528824

This project focuses on detecting anomalies in IoT sensor time-series data to identify potential equipment failures or maintenance needs. The implementation was done using Python in **Google Colab** with libraries such as pandas, numpy, matplotlib, scikit-learn, and TensorFlow.

Dataset and Preparation:

A synthetic time-series dataset was generated with **injected anomalies**, and optionally, the NASA Bearing dataset which created as CSV File. Data cleaning involved handling missing values and duplicates.

Featured:

Key features created include rolling mean, rolling standard deviation, and first difference to capture local trends and sudden signal changes. All numerical features were scaled using a Standard Scaler before model training.

Algorithms Used:

1. **Isolation Forest** – an unsupervised algorithm that isolates anomalies through random partitioning, which I used to know with the help of an AI tool.

2. **Autoencoder** – A deep learning model, I did my final year project in deep learning, trained to reconstruct normal data points, with high reconstruction error, were marked as anomalies.

Comparison and Results:

Both methods successfully detected the synthetic anomalies. The Isolation Forest was simpler and faster, while the Autoencoder handled non-linear patterns better. Visual comparison(charts and plots) of detected points confirmed that both approaches accurately highlighted the abnormal time segments. The result file is downloaded as a CSV file.

Conclusion

This project demonstrates an end-to-end anomaly detection pipeline covering data loading, feature creation, modelling, and visualisation. Isolation Forest is effective for quick unsupervised detection, whereas Autoencoder provides deeper insights for complex sensor data.