SUMMARY

Approach-

Firstly, Training of Deep learning model (LSTM) was done. Raw data (Taking total acceleration. And not body acceleration for better performance) was loaded on the notebook alongside numpy and Sklearn and it was then stacked and normalised to prevent any big values from lowering the accuracy of the model. Then the model and it's training was coded. However, the accuracy turned out to be 0% at first whose key cause was the Nan explosion that had occured in predicted values.Then, AI was used to find out the probable causes of explosion. The stacked raw data was checked for Nan values and replaced by 0, Range and shape of data was checked to be in the format that the LSTM model accepted. Then it was run again and the accuracy turned out to 85.78% on the provided test data. Next the Machine Learning models were trained, first from features extracted from TSFEL and then from author provided features. Starting from TSFEL approach, at first the raw data (This time using body acceleration) was loaded and the various libraries like numpy, pandas and Sklearn were imported alongside TSFEL. For faster feature extraction, GPU was required so, the features were extracted in another google collab account to use GPU since GPU runtime of current collab account ran out. The features were stored in a CSV file which was imported in the current notebook. Then the data shape and range were checked, it was put in numpy arrays (since SciKit learn models like RF, SVM and LR had to to be trained), normalised and then the models were trained on it. Accuracy of the Random Forest model was 92.30%, SVM was 95.39% and Logical Regression model was 95.62%. Next, The author extracted features were used for training. The X_train provided in the database contained author provided features and were used as input, it was then normalised and used to train the ML models and the accuracies turned out to be 92.60% for Random Forest Model, 95.22% for SVM and 95.49% for Logical Regression model.

Observations and Result-

The ML models (self and author extracted) performed better in predicting the data (avg accuracy 94.43%) as compared to the LSTM model (accuracy 85.78%). Comparing TSFEL extracted features trained ML models and author extracted feature trained ML models, RF model performed better on Author features while SVM & LR models performed better on TSFEL features (small margin). So, Conclusively, for best accuracy, LSTM models should be used when the dataset is very large and ML should be used for small or medium datasets (like one provided). Within ML models, TSFEL features outperform complex ML models, offering more flexibility as compared to author's features.