Weekly Progress Report

Submitted by:

Sandesh Pabitwar

Modern Education Society's College of
Engineering Pune

Project Title: INTP22-ML-2 Remaining Usable Life Estimation (NASA Turbine dataset)

Objectives:

1. Understand the components and working of Turbo Engine.
2. Understand the dataset and analyze dataset
3. To learn about different Python libraries and their implementation.
4. To develop model for remaining usable life estimation

Task done in this week:

1. Understood different predictive models
2. Gone through implementation of some models.
3. Started working on ARIMA model.
4. Reading research paper about ARIMA model.

Understood different predictive models.

After completing the EDA for the dataset, I started reading about different predictive models. I learned about 3 different predictive models

1. Similarity model
2. Survival model.
3. Degradation model.

When we don't know about the healthy state of the engine and if we gave data only from time of failure at that time, we use survival model.

When the failure data is not available but we know threshold safety level at that time we can use Degradation model.

When we have completed knowledge of data of different stages such as healthy state Degradation stats and failure state at that time, we can use the Similarity model.

After getting theoretical knowledge of different models, I understood these models using some examples where these models are actually implemented on different datasets.

Gone through implementation of some models.

after looking at different model's implementation understood the important states for building a model for predicting a RUL. Those are as follows

1. Data Preparation
2. model selection
3. model training
4. evaluation
5. performance tunning

Data preparation:

At this stage we select only that data which is useful for training the model. We drop the null values or some features that are not relevant. Data is balanced at this stage.

Model selection:

Based on the dataset we choose appropriate model at this stage. For example, 1000 images of different animals are given to us then here we can use classification model.

Model training:

At this stage we pass all the training data to the machine learning model to get desired output.

Evaluation:

At this stage we evaluate the results predicated by the model and actual results based on this accuracy of model is calculated.

Performance tuning:

Model is improved at this stage based on some hyperparameters. Accuracy of model is improved during this stage.

Started working on ARIMA model.

plotted the graph for

1.data of all engine for a single feature

2. data of one engine for same feature

3. Data of one engine for same feature after differencing the signal once.

After observing the graphs came to conclusions that we don't know beforehand how many times the timeseries of each sensor of each engine needs to be differenced before becoming stationary to solve this we need some functions.

Reading research paper about ARIMA model.

I have gone through the research paper titled as Time series forecasting using improved ARIMA by Mohammad Reza Keyvanpour and Soheila Mehrmolaei. This paper explains about classification of time series forecasting techniques such as estimator of short-lived and estimators of long-range. It also explains about proposed approach for time series forcasting.