**Topic: “ Predictive Maintenance Model for Electrical Machines “**

**Team No:** 9 - Naveen Rajendran

[ NUID: 001304569 ]

**Background:**

In the current world of smart devices, the role of sensors cannot be neglected. Nearly every electrical & electronic device or equipment is embedded with wide variety of industry grade sensors which provides us continuous linear or non-linear time series data. With the growing number of IoT devices & connected systems, the amount of data generated has multiplied many folds in the past few years. The sensor data can be classified into two types ( Analog & Digital), the magnitude of parameters such as temperature, pressure, vibration, humidity etc. will serve as a meaningful information inorder to take critical decision at the organization level.

**Problem Statement:**

A manufacturing companies we have machines deployed at the shop floor such as motor, pump, compressors, blowers, interstage coolers etc. These machines needed to run 24 x 7 throught the year inorder to achieve 100 % productivity in their respective businesses. The 100 % availability of production systems are highly impossible due to unexpected breakdown / damage of these critical electrical machines. Inorder to have uninterrupted industrial operations there needs to be a proper maintenance of machines installed. Maintenance practice in a factory is classified into condition-based maintenance, preventive maintenance & shutdown maintenance. These maintenance activities are performed only when the machine gets failed or when monthly / yearly based preventive schedule is planned by the engineer. Therefore, the manufacturing company faces an unexpected breakdown of production systems which directly affects the product manufacturing cycle creating losses financially to the company due to increased maintenance cost and slowdown in production of goods.

**Goal:**

As per the leading consulting firm Deloitte, “Predictive maintenance increases equipment uptime by 10 to 20% while reducing overall maintenance costs by 5 to 10% and maintenance planning time by 20 to 50%. Hence, there is a significant need to develop a predictive maintenance model where an engineer can prepare a schedule for maintenance well ahead based on the sensor data collected from the machine. In this project we are trying to resolve the unexpected machine breakdown problem using data science & machine learning techniques. For example, there is a sensor data meant for measuring temperature of the bearing on the motor. Based on the gradual increase in temperature level of the bearing we can predict that our motor is going to fail within few days.

**Dataset Description & Source Link:**

**Pump Sensor Data –** [**https://www.kaggle.com/nphantawee/pump-sensor-data/download**](https://www.kaggle.com/nphantawee/pump-sensor-data/download)

The following data is the set of sensor data collected from an electric pump. There are around **51 sensors** embedded on the machine & there is a series of float data generated from those sensors & towards extreme right there is field named status to determine the status of the machine. The values from the sensor & the corresponding status of machine ( Normal / Failure) are given in the above-mentioned dataset

**Methodology:**

Planned Approach by steps:

1. **Data Cleaning & Preprocessing** : There are multiple ‘**NaN**’ values in the dataset given, it needs to be replaced with scalar value like ‘0’ (zero) so that it would make our data processing easy
2. **Exploratory data analysis:** Planning to segregate the sensor values for which the pump is getting stopped/failed & also I will be trying to find impact score of each sensor in determining the condition of electric pump. Correlation mapping of each sensor is also important here inorder to determine the sensor impact score
3. **Machine Learning Algorithm:** Since this is a time series data with labelled entities, I am planning to use supervised machine learning techniques.

**Linear Regression :**

Statistical comparison of multiple sensor data points can predict the outcomes using time series forecasting methods. By keeping one sensor value as a dependent variable, we can estimate the outcomes of another independent variable

**K nearest neighbor:**

Using this classification algorithm, we can trace out the sensor readings falling into different clusters. Based on this we can classify the sensor readouts that are causing to fail the system.

1. **Evaluation Metrics:** Based on the results obtained from different test data , I will impute these values in Evaluation Metrics ( Confusion Matrix) inorder to deliver positive predictive value, negative predictive value & accuracy of algorithm