# Detailed Project Report

Q1. Tell me about your current project and problem statement.

This project is based on classification of wafer sensor which are used in telecom radar systems, usually the maintenance activities of the wafer sensors are done manually which requires and engages a lot of man power and also telecom towers are remotely located so it consumes a lot of time to carry out the maintenance activities. In order to automate this process and eliminate such we have build this project based on the machine learning algorithms which helps us to classify the healthy and the faulty sensors.

Q2. What was the size of the data ?

The no. of records used for training the model was around 7.5 lacs.

Q3. What was the data type ?

The data type used for model training was float16.

Q4. What was the team size for the distribution ?

The project teams consists of:-

1 Business Analyst

1 Solution Architect

2 Devops Engineer

1 Data Engineer

3 Machine Learning Engineer

1 Lead

1 Data analyst

2 UI engineer

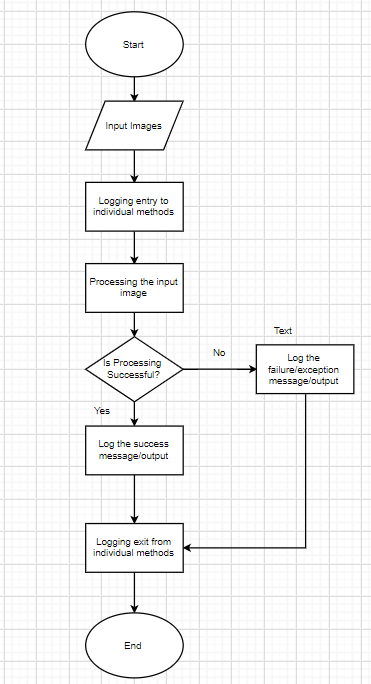
2 QA engineer

Q5. How many nodes were there in all the Dev, UAT and Prod environment ?

All the coding part and the model training part was done on local machine and after training the model we were selecting the best model which needs to be deployed in UAT & production will be send to s3 bucket, which will be further deployed in 2 UAT nodes and there are 2 nodes in production for deployment.

Q6. How were you creating and maintaining the logs ?.

The logs are maintained using the MongoDB. The logging starts with the start of the application. The start time of the application gets logged. After that there are logging for entry and exit methods for each of the methods. There are logging for the error scenarios and exception block as well.



Q7. What techniques were you using for data pre-processing for this project ?

Multiple techniques has been implemented for data preprocessing as per the master data management such as removal of columns if the column are empty, checking for the null values in the column if there are any then KNN mean imputation will be performed, removal of column with zero standard deviation because with column with zero variance will not contribute to performance of the model.

Q8.How were you maintaining the failure cases ?

Let’s say if our API end point is unable to make a prediction it will get into the exception where the email triggering set up is done if there is any kind of failure occurs it will directly send the email notifications to the concerned dept and inspection of the issue will take place.