

## **PROJECT NAME:** Predicting Life Expectancy Using Machine Learning

### **PROJECT SUMMARY:**

Life expectancy is a statistical measure of the average time a human being is expected to live, Life expectancy depends on various factors: Regional variations, Economic Circumstances, Sex Differences, Mental Illnesses, Physical Illnesses, Education, Year of their birth and other demographic factors. So I tried to provide a way to predict average life expectancy of people living in a country when various factors such as year, GDP, education, alcohol intake of people in the country, expenditure on healthcare system and some specific disease related deaths that happened in the country are given. We'll try to predict life expectancy using various Machine learning and Deep learning models.

### **PROJECT REQUIREMENTS:**

- A good dataset to be used so that our model trains without any overfitting or underfitting and also data should be balanced.
- We need to make sure that we've got the hardware and software necessary to collect, process, analyze, integrate, and store this data.
- We need to select the algorithm which gives better accuracy and less loss.

### **FUNCTIONAL REQUIREMENTS:**

- **Alert Generation:** When a machine degradation or potential asset failure is detected, this is communicated to the relevant facility stakeholders.
- **Machine Learning Methodology:** Each Predictive Asset Maintenance solution is based on a Big Data methodology. Is this a manual process or is Artificial Intelligence used to automatically select the optimal algorithm for the specific scenario?
- **Asset Visualization:** At a facility level, technicians accessing the user-interface will not be trained in Artificial Intelligence and Big Data. The key considerations when defining this requirement are the visualization of machine behavior and the ability to depict the health of machinery or the entire facility, and take specific action as a result.

## TECHNICAL REQUIREMENTS:

1. **Data itself.** Without proper datasets for training our machine learning algorithms and without incoming input once the algorithm is trained, there's no point in machine learning.
2. **Capabilities for data preparation.** We need to make sure we've got the hardware and software necessary to collect, process, analyze, integrate, and store this data.
3. **ML algorithms.** There are numerous types and kinds of machine learning algorithms and we need to make sure which algorithms suits best for the given problem.
4. **Evaluation** — What defines success? Is a 95% accurate machine learning model good enough?
- 5.

## SOFTWARE REQUIREMENTS:

- Python
- IBM Cloud
- IBM Watson

## PROJECT DELIVERABLES:

We'll try to create a platform on IBM Watson wherein a person can predict his/her life expectancy by providing data to the model.

## PROJECT TEAM:

The team includes

**Data Scientist:** To create (or select) and train algorithms, building the models for machine learning. They may create their solution under idealized assumptions about the data inputs and algorithmic outputs.

**Software Engineers:** They ensure the technology is built in a robust way so that it can be packaged and deployed into production in mission-critical systems. They might set the infrastructure, get the data pipeline in place, and ensure the data scientists have

everything they need to focus on the models.

**Data Engineer:** This role is focused on wrangling/pre-processing data to prepare it for machine learning. If an algorithm isn't working, a data engineer will try finessing the data in a different way. After a project goes into production, data engineers may work together with software engineers and data scientists to assess and optimize the data feeding back into the model.

### **PROJECT SCHEDULE:**

The estimated time required to build the model is 20 days which include data preparation, model training, model evaluation and model deployment.