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 REQUIRMENTS:

- 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:

1. At the front-end, a webpage taking the necessary inputs from the user to implement the designed model.

2. Node-Red Application for the input to gets processed and finally gives the desired output, life-expectancy.

3. Some Knowledge of Mathematics and Statistics

4. Machine Learning model based on regression

5. IBM Watson Studio and Jupyter Notebook

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?

SOFTWARE REQUIREMENTS:

- Python
- IBM Cloud
- IBM Watson
- Jupyter Notebook

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 SCHEDULE:

Task(dd/mm)	15/5-16/5	17/5-18/5	19/5-21/5	22/5-26/6	27/6-30/6	31/6-2/6
Project planning & kickoff	DONE					
Explore IBM Cloud platform		DONE				
Explore IBM Watson services			DONE			
Introduction to Watson studio				DONE		
Predicting life expectancywith Python					DONE	
Predicting life expectancy without python						DONE

FUTURE WORK AND CONCLUSION

Prognosis research using machine-learning will likely increase exponentially over the next few years, as the techniques used become more commonplace. However, prior to further translation and use in clinical settings, the analytic validity of ML algorithms needs to be established, which requires external validation and replication. Currently, this is difficult to achieve because studies using ML are varied in their nature, purpose and reporting, and are not necessarily led from clinical or health care disciplines.

This project has found deep learning and random forest ML algorithms improved accuracy of prediction of premature all-cause mortality in a large middle-aged general population, compared to standard methods. This needs to be further explored in other large databases and in other populations. The intriguing variations in ML model composition may enable new hypothesis generation for potentially significant risk factors that would otherwise not have been detected using standard approaches