Module-1 Statement of work

Name: Kartik Sojitra

**Table of Contents**

* Executive Summary
* Problem Statement
* Analytics Rationale Statement
* Data
* Data Analysis Approach
* Project Plan:
* Creating project plan outline- 24th Jan 2020
* Analyze data (i.e. Conduct EDA) -Feb 14th, 2020
* Clean dataset and Review and Chose algorithm(s)- Mar 6th, 2020
* Train model and Validate model – March 27th, 2020
* Develop finalized model and Develop final report- April 17th, 2020

Executive Summary

The project identifies patients with heart defect using machine learning algorithm, to forecast and apply the same model creating portable device to predict and diagnosis heart defects in both newborn babies and adults.

Problem Statement

“To identify and determine Children with Congenital heart defects and Adults with Heart valve defects using machine learning algorithm”

Heart valve defects can be seen at the time of birth with new born babies and also it is seen in adults as well. With this project, the algorithm will find out a prediction model to diagnose heart defects in other patients.

A Classification report metrics will be used to evaluate the algorithm to find out accurate responding model. The model will look for output variable classified as 0- Normal patients, 1- Congenital heart defect, 2- Heart Valve defect

Analytics Rationale Statement

congenital heart defects have been left undiagnosed which have results existence of heart murmur in 40-45% of Children.

Heart valve defects in adults are also undiagnosed which results heart murmur in 10% of Adults.

Early diagnosis of heart defects helps to reduce 90% of heart episodes. So, by analysing the data, the project can identify a pattern using classification model, to predict heart valve defects for other patients. A portable machine predicting heart valve defect can be developed applying the established algorithm with better accuracy.

Data

The project analyses the heart valve measurements using Echo cardiograph reports of 5009 patients at Princess Margret Hospital, in partnership with the UHN Echocardiography Lab and to establish and validate the need of pre-diagnosis of Heart defects. Having 40 independent Variables with different measurement of Heart Valves, the dependent variable will be 0- Normal patients, 1- Congenital heart defect, 2- Heart Valve defect

Data Analysis Approach

In the project, the data will be analyzed using a classification machine learning algorithm methodology, using Python and work on getting better accuracy of classification metrics.

**Project Plan**:

|  |  |  |  |
| --- | --- | --- | --- |
| **Task** | **Description** | **Start date and time** | **End Date and time** |
| **Creating Project outline** | Describing the purpose of the project and methodology to be used with timelines.  Scope   * Develop model which can help learning the pattern predicting patients with Congenital Heart defect and Heart Valve defects. * Identifying the best classification model with highest accuracy. | Jan 6, 2020 8:00 AM | Jan 24, 2020 11:59 PM |
| **Stage I** | Conduct a detail EDA to understand, Analyse the data. In this process right features and independent variables can be identified which are contributing to the dependent variable. | Feb 3, 2020 8:00 AM | Feb 14, 2020 11:59 PM |
| **Stage II** | Clean the dataset with best independent variable with key features which are contributing to the dependant variable and identify the right algorithm matching the dataset. | Feb 17, 2020 8:00 AM | Mar 6, 2020 11:59 PM |
| **Stage III** | Understanding the dataset and train the model to optimise it and validate it in the test environment | Mar 16, 2020 8:00AM | Mar27, 2020 11:59 PM |
| **Stage IV** | The final model showing accuracy with the evaluation metrics will be finalised which can help to produce final report.  The final reports show the model with highest accuracy of the classification metrics and identify the patients with Congenital Heart defect and heart valve defect and the model can be applied to develop portable devices diagnosing heart valve defect. | Apr 6, 2020 8:00 AM | Apr 17, 2020 11:59 PM |