**BIG DATA PARALLEL PROGRAMMING PROJECT REPORT**

**Shankaranarayanan BR**

**Introduction:**

In this project what we are trying to achieve is when there is higher probability of a person facing an heart stroke. In this is a typical classification problem and for performing this test we have taken the dataset which, was available for us in the following github page.

<https://github.com/surajdidwania/Healthcare-Stroke-Patients-Data/tree/master/healthcare-dataset-stroke-data>

This project was also a kaggle dataset, and it contained the following attributes. The list of attributes are given in the below table