Stroke prediction model

Introduction

With the fast evolvement of our lives and the increasing number of demands a person must handle and attain, we notice the emergence of new diseases and the spread of old ones. Stroke has become one of our century's most leading diseases. Around 1 in every 4 adults over the age of 25 will have a stroke in their lifetime worldwide. Thus, predicting a stroke case prior to its actual occurrence is a very crucial step. As this will help increasing the awareness and the commitment to following an early corrective approach in people's lives such as changing the diet plan, taking care of the mental and psychological health, avoid smoking … etc. Furthermore, enabling people from being aware of their cases and illnesses in an early stage helps in reducing the number of fatalities annually and relieve the burden on the medical sector. By the presence of Machine Learning (ML) and Deep Learning (DL) algorithms, we can build models to predict whether a person will be an ideal candidate for being diagnosed with stroke or not.

Problem statement

Predicting stroke cases using ML or DL algorithms.

Objectives

In this project I will be fulfilling the below:

- Download a large dataset 10,000+ rows with 8+ columns.
- Make some EDA
- Build an ML or DL classification model
- Evaluate the model
- Test the model

Approach

I will initially make some EDA that includes data cleaning and visualization and will be using the basic EDA libraries in Python such as pandas, matplotlib, seaborn and numpy. As well, I will build a classification model using some of the basic ML libraries in Python such as sklearn, tensorflow, or pyTorch.

References

The Stroke dataset can be found via this link: https://www.kaggle.com/lirilkumaramal/heart-stroke