Stroke Predictions using Healthcare Dataset
Kailash SenthilKumar¹, Kanthimathinathan Ramasubramanian²
Student¹, ²
Department of Electronics and Instrumentation Engineering
Chennai, India

Abstract:
Medical care has gotten one of the most significant worries on the planet. The instances of coronary illness are expanding on a quick scale among the individuals particularly among the youthful age. We can spare the lives of individuals in the event that we could distinguish the coronary illness on/before time, by getting them treated. In this matter man-made consciousness can be of an incredible assistance. Here we have gathered an informational index and afterward we have constructed a forecast model to distinguish coronary illness dependent on the different calculations that are accessible for AI.

Keywords: Python, AI, Jupyter Notebook, HTML, Stroke, Health.

I. INTRODUCTION

1. PYTHON

Python is an object-oriented, high-level programming language with integrated dynamic semantics primarily for web and app development. It is extremely attractive in the field of Rapid Application Development because it offers dynamic typing and dynamic binding options. Python is relatively simple, so it’s easy to learn since it requires a unique syntax that focuses on readability. Developers can read and translate Python code much easier than other languages. In turn, this reduces the cost of program maintenance and development because it allows teams to work collaboratively without significant language and experience barriers.

2. JUPYTER NOTEBOOK

The Jupyter Notebook is an open source web application that you can use to create and share documents that contain live code, equations, visualizations, and text. Jupyter Notebook is maintained by the people at Project Jupyter. Jupyter Notebooks are a spin-off project from the IPython project, which used to have an IPython Notebook project itself. The name, Jupyter, comes from the core supported programming languages that it supports: Julia, Python, and R. Jupyter ships with the IPython kernel, which allows you to write your programs in Python, but there are currently over 100 other kernels that you can also use.

3. HYPER TEXT MARKUP LANGUAGE (HTML)

HTML is an acronym which stands for Hyper Text Markup Language which is used for creating web pages and web applications. Let’s see what is meant by Hypertext Markup Language, and Web page.

Hyper Text: HyperText simply means "Text within Text." A text has a link within it, is a hypertext. Whenever you click on a link which brings you to a new webpage, you have clicked on a hypertext. HyperText is a way to link two or more web pages (HTML documents) with each other.

Markup language: A markup language is a computer language that is used to apply layout and formatting conventions to a text document. Markup language makes text more interactive and dynamic. It can turn text into images, tables, links, etc.

Web Page: A web page is a document which is commonly written in HTML and translated by a web browser. A web page can be identified by entering an URL. A Web page can be of the static or dynamic type. With the help of HTML only, we can create static web pages. Hence, HTML is a markup language which is used for creating attractive web pages with the help of styling, and which looks in a nice format on a web browser. An HTML document is made of many HTML tags and each HTML tag contains different content.

4. DJANGO

Django is a free and open source web application framework written in Python. A framework is nothing more than a collection of modules that make development easier. They are grouped together, and allow you to create applications or websites from an existing source, instead of from scratch. This is how websites – even simple ones designed by a single person – can still include advanced functionality like authentication support, management and admin panels, contact forms, comment boxes, file upload support, and more. In other words, if you were creating a website from scratch you would need to develop these components yourself. By using a framework instead, these components are already built, you just need to configure them properly to match your site.

II. OBJECTIVES

- The prediction of long-term outcomes in stroke patients may be useful in treatment decisions.
- Machine learning techniques are being increasingly adapted for use in the medical field because of their high accuracy.
- This study investigated the applicability of machine learning techniques to predict long-term outcomes in ic stroke patients.

III. LITERATURE REVIEW

1. Violent Python

Violent Python shows you how to move from a theoretical understanding of offensive computing concepts to a practical implementation. Instead of relying on another attacker’s tools, this book will teach you to forge your own weapons using the
Python programming language. This book demonstrates how to write Python scripts to automate large-scale network attacks, extract metadata, and investigate forensic artifacts. It also shows how to write code to intercept and analyze network traffic using Python, craft and spoof wireless frames to attack wireless and Bluetooth devices, and how to data-mine popular social media websites and evade modern anti-virus.

2. Python Passive Network Mapping:
P2NMAP is the first book to reveal a revolutionary and open source method for exposing nefarious network activity. The "Heartbleed" vulnerability has revealed significant weaknesses within enterprise environments related to the lack of a definitive mapping of network assets. In Python Passive Network Mapping, Chet Hosmer shows you how to effectively and definitively passively map networks. Active or probing methods to network mapping have traditionally been used, but they have many drawbacks - they can disrupt operations, crash systems, and - most importantly - miss critical nefarious activity. You require an accurate picture of the environments you protect and operate in order to rapidly investigate, mitigate, and then recover from these new attack vectors. This book gives you a deep understanding of new innovations to passive network mapping, while delivering open source Python-based tools that can be put into practice immediately.

3. Python Forensics
Python Forensics provides many never-before-published proven forensic modules, libraries, and solutions that can be used right out of the box. In addition, detailed instruction and documentation provided with the code samples will allow even novice Python programmers to add their own unique twists or use the models presented to build new solutions. Rapid development of new cybercrime investigation tools is an essential ingredient in virtually every case and environment. Whether you are performing post-mortem investigation, executing live triage, extracting evidence from mobile devices or cloud services, or you are collecting and processing evidence from a network, Python forensic implementations can fill in the gaps.

4. Personalized Psychiatry
Personalized Psychiatry presents the first book to explore this novel field of biological psychiatry that covers both basic science research and its translational applications. The book conceptualizes personalized psychiatry and provides state-of-the-art knowledge on biological and neuroscience method ologies, all while integrating clinical phenomenology relevant to personalized psychiatry and discussing important principles and potential models. It is essential reading for advanced students and neuroscience and psychiatry researchers who are investigating the prevention and treatment of mental disorders.

IV. STROKE PREDICTION USING HEALTH CARE DATASET
1. Data Set collection and Labelling
Initially, HTML is used to design the UI for data collection from our source, in this case Users, which is then passed to the backend of our project code as values.

We collect the data set from our source and curate them using python. In this data set we collect inputs such as age, type of work, residence type, past history of hypertension, cardiac disease, Body Mass Index (BMI) from the user.

Variables for these principle information is declared along with its type.

With the help of the principle data, we can interpret different values from the main source such as total count, mean value, etc.

Real world dataset containing more than 40,000 data points is taken.
The Machine learning model is created in python using Jupiter notebook and made to run on our Data sets.

After data processing using Machine Learning is done on the model, we start basic Data Visualisation.

Then we start our split training test in our processed data.

This Split Test data gives a more in-depth understanding for the machine learning algorithm to better predict or extrapolate the predicted value.

After completing our split train test, we move to the final stage of the process, i.e., to predict whether the user is likely to have a Stroke or not.

To predict the Stroke occurrence from the data set, we fit the logistic regression model on our dataset.
This helps us to predict whether the person/user will have a stroke or not.

V. CONCLUSION

We have successfully made a machine learning algorithm using Python, Jupyter Notebook and Django as the backend development and trained them with the help of more than 40,000 data points from real-world users, which resulted in successfully predicting the occurrence of Stroke for the user and also visually/graphically representing the same for more detailed representation and understanding.

VI. REFERENCES

[1]. Comparative case studies of open source software peer review practices
[2]. Pengembangan Pengolahan Citra Face Recognition, Face Counting dan AgeGender Detection Secara Real Time di Python
[3]. A Convolutional Neural Network Classifier for Recognition and Detection of Traffic Signs
[7]. Website: https://fmsystems.com/blog/does-bim-have-a-role-in-the-internet-of-things/
[9]. Website: http://www.proquest.com/
[10]. Website: https://www.academia.edu/material (including pictures, drawings, tables, graphs, equations)