**North-South University**



**Project Name: Multiple Disease Predictor Using Machine learning.**

**Course Code: CSE299**

**Course Title: Junior Design**

**Submitted to: Meem Tasfia Zaman (MEZ)**

**Group Member:**

|  |  |
| --- | --- |
| Mostafizur Rahman | 1912291042 |
| Sadia Binta Sarowar | 1811908642 |
| Saiful Islam | 1731604642 |

Contents

[Abstract 3](#_Toc106794925)

[Introduction 4](#_Toc106794926)

[Background/Literature review 5](#_Toc106794927)

[Aim/goal 5](#_Toc106794928)

[Application of the project in real life 5](#_Toc106794929)

[Tools needed 5](#_Toc106794929)

[Tentative schematic diagram/flow chart of the project and short description 5](#_Toc106794929)

[Projected plan 6](#_Toc106794925)

[Projected timeline 7](#_Toc106794926)

[Conclusion 7](#_Toc106794927)

[References 8](#_Toc106794928)

1. Abstract

Many academics have recently developed automated diagnostic methods based on

supervised learning models. Early identification of sickness may help reduce the number of

people who die due to these illnesses. This work created an efficient automated illness diagnostic

model using machine learning techniques. We've chosen three critical illnesses to discuss in this

paper: cancer, heart disease, and diabetes. The data is entered into a web app. The analysis is

conducted in a real-time database using a pre-trained machine learning model trained on the

same dataset and deployed in ML model. The illness detection result is eventually shown in the

web app. The computation for prediction is carried done using logistic regression.

People can predict their disease and can take the necessary step. It also can indicate the early

stage of any disease like cancer. This work appears to be promising at this stage. This isn't

limited to cancer patients. It might be used to identify various diseases that are difficult to detect

early on.

1. INTRODUCTION

During many analyses of existing healthcare systems, just one disease was evaluated at a time. The majority of articles concentrate on a single ailment. When a company wishes to analyze its patients' health records, it must use a variety of models. The existing system's technique is useful for analyzing only a single ailment. Today's mortality rate has risen as a result of a lack of precise disease identification. Even if a patient has been healed of one ailment, they may still be suffering from another. In real life, we are related in a similar circumstance.

When analyzing the disease, several current systems used only a few factors. As a result, it may be impossible to identify the diseases that will be caused as a result of the disease's side-effect. Diabetes, for example, increases the risk of heart disease, neuropathy, retinopathy, hearing loss, and dementia.

This article focused at data sets for diabetes analysis, heart disease, and Parkinson’s disease. Many other disorders may be included in the future. This study is adaptable, and it has now been expanded to include a variety of disorders. When adding a new illness analysis to this API, the developer must also include the model file for the new disease analysis.

When creating a new disease, the developer must plan for python in order to save model behavior. The developer can load a pickled file to retrieve model behavior while using this Flask API. If these methods are used to predict other diseases, then this analysis will yield the most relevant disorders.

1. BACKGROUND/LITERATURE REVIEW

Many daily duties have gotten considerably easier for us as a result of the advent of online-based applications. For the past few years, we've all been employing a collection of disease prediction models. In the beginning, a Machine Learning Model was used to detect a specific disease or one disease from a group of patients. Many of the published research focused on a specific and particular disease. When a user wants to analyze diabetes, they must use one model, and when they want to analyze heart disease, they must use another model. This is a lengthy procedure. Further, if a user has multiple diseases but the existing system can only predict one of them, there is a potential that the death rate may rise as a result of not being able to predict the other sickness in advance.

It is possible to predict more than one disease at a time using a multi-disease model. So, we are trying to build this model with the help of the existing model idea. There is no need for the user to go among multiple models in order to anticipate diseases. It will save time, and it has the potential to lower mortality rates by predicting numerous diseases at once.

1. AIM/GOAL

The objective of this article is to avoid the death rate from rising day after day by notifying patients ahead of time-based on their health conditions. This prediction model predicts numerous diseases at the same time. The disease will be projected based on the user's input. The user will be offered the option. If the user wants to predict a specific disease or if the user does not specify a disease category, the associated disease model will be launched and predicted based on the user's inputs. The advantage of a multi-disease prediction model is that it can anticipate the likelihood of multiple diseases occurring in the future, as well as minimize the mortality ratio. The cost of patient analysis can also be decreased by combining many disease models and predictions through one location. In this paper, we look into studies that use multiple supervised machine learning models for each disease recognition issue. Because evaluating the performance of a single method across multiple research settings introduces bias, resulting in imprecise results, this approach provides better comprehensiveness and precision. A few disorders in the heart, diabetes, and brain will be studied using machine learning models. Numerous methodologies will be examined by different machine learning and deep learning techniques to identify the patient's status, including logistic regression, Nave Bayes classification algorithm, SVM, Decision tree algorithm, Random Forest algorithm, and many more algorithms.

1. APPLICATION OF THE PROJECT IN REAL LIFE

The system uses web server and cloud server to predict the disease. The attribute values collected from patients are sent to a cloud server in which the computation intelligence model is placed through web server. The prediction is sent back from cloud server to patients and doctor. The web site demonstrates the predicted result for different input attributes value. Both the patient and doctor can use this application for their own purpose. First, the patients open the app and input different attributes value like name, age, sex, cheese pain type, blood pressure, or other types of symptoms etc. The input values go to the webserver and save the values there also. Inside the cloud server, the predicted model is placed and using the attributes value, the result is predicted and again acknowledge to the webserver. This result also saves in the webserver. The patients and doctor can observe the predicted result (i.e. if the patient has heart disease or not) and the risk level. The risk level is marked into percentage. Less than 30% means no risk and more than 70% means high risk

1. TOOLS NEEDED (SOFTWARE/HARDWARE) AND SHORT DESCRIPTION

Python:

Python is an interpreted, high-level, general-purpose programming language .Its language constructs and object oriented approach aim to help programmers write clear, logical code for small and large-scale projects. Python is dynamically typed, and garbage collected. It supports multiple programming paradigms, including procedural, object-oriented, and functional programming.

kaggle dataset:

Kaggle allows users to find and publish data sets, explore and build models in a web-based data-science environment, work with other data scientists and machine learning engineers, and enter competitions to solve data science challenges.

Google colab/Jupyter Notebook:

Google Colab is a web-based iPython Notebook service for interactive coding. Google Colab offers free access to Nvidia K80 GPU, and it is becoming a popular tool for quick prototyping and visualization. Colab is easy to learn and it is very similar to Jupyter Notebook, another very popular tool for interactive coding. The Jupyter Notebook App is a server-client application that allows editing and running notebook documents via a web browser. The Jupyter Notebook App can be executed on a local desktop requiring no internet access (as described in this document) or can be installed on a remote server and accessed through the internet.

Streamlit/Django:

Streamlit is an open-source Python library that makes it easy to create and share beautiful, custom web apps for machine learning and data science.

Django is a high-level Python web framework that enables rapid development of secure and maintainable websites.Django provides a secure way to manage user accounts and passwords, avoiding common mistakes like putting session information in cookies where it is vulnerable (instead cookies just contain a key, and the actual data is stored in the database) or directly storing passwords rather than a password hash.

Pickle:

Python pickle module is used for serializing and de-serializing a Python object structure. Any object in Python can be pickled so that it can be saved on disk.

VS Code:

Vs code is used to develop computer programs, as well as websites, web apps, web services and mobile apps. Visual Studio uses Microsoft software development platforms such as Windows API, Windows Forms, Windows Presentation Foundation, Windows Store and Microsoft Silverlight.

1. Tentative schematic diagram/flow chart of the project and short description.

Diagram

Description automatically generated

1. Projected plan.

|  |  |
| --- | --- |
| Sadia binte Sarowar | Web APP |
| Mostafizur Rahman | Machine learning ,Web App, UI/UX, Modeling |
| Saiful Islam | Web App and machine learning |

1. Projected timeline for the completion of your

Graphical user interface

Description automatically generated

1. Conclusion

The major goal of this paper is to use a Machine Learning algorithm to predict disease based on symptoms provided by patients. Our target is to develop a system that could accurately forecast more than one disease. This multi-disease prediction model will be used to predict several diseases at the same time. The user does not require any additional software and they don’t need to go via many websites, which saves time. The disease will be projected based on the user's input. The user will be offered the option. If the user wants to predict a specific disease or if the user doesn't specify a disease category, the associated disease model will be triggered and predicted based on the user's inputs. Early detection of diseases can extend someone’s life expectancy as well as rescuing him from financial difficulties. The advantage of a multi-disease prediction model is that it can anticipate the likelihood of multiple diseases occurring in the future, as well as minimize the mortality ratio. Applying promising technology like machine learning to the early detection of diseases would have a significant influence on society. The prognosis of heart disease in its early stages can assist in making lifestyle adjustments decisions among high-risk individuals and, as a result, decrease complications, which could represent a major breakthrough in medicine. The Each year, the number of people suffering from these diseases grows. Early detection and treatment are essential. In this case, the use of appropriate technology is beneficial. Regard for the medical community and patients can be quite valuable. In this paper, different machine learning algorithms used to measure the performance. In future, we can expand this work by adding more diseases that are trained by machine learning models and also can include the disease that involves deep learning models.

1. References.

[1] Self-Attention Equipped Graph Convolutions for Disease Prediction(Self-Attention Equipped Graph Convolutions for Disease Prediction Anees Kazi;S.Arvind krishna;Shayan Shekarforoush;Karsten Kortuem;Shadi Albarqouni;Nassir Navab 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019))

[2] Latent Class Multi-Label Classification to Identify Subclasses of Disease for Improved Prediction Awad Alsaid Alyousef;Svetlana Nihtyanova;Christopher P. Denton;Pietro Bosoni;Riccardo Bellazzi;Allan Tucker2019 IEEE 32nd International Symposium on Computer-Based Medical Systems (CBMS) Year: 2019 | Conference Paper | Publisher: IEEE

[3] Latent Class Multi-Label Classification to Identify Subclasses of Disease for Improved Prediction Awad Alsaid Alyousef;Svetlana Nihtyanova;Christopher P. Denton;Pietro Bosoni;Riccardo Bellazzi;Allan Tucker 2019 IEEE 32nd International Symposium on Computer-Based Medical Systems (CBMS) Year: 2019 | Conference Paper | Publisher: IEEE

[4] Prediction of Alzheimer's Disease Progression with Multi-Information Generative Adversarial Network Yan Zhao;Baoqiang Ma;Pengbo Jiang;Debin Zeng;Xuetong Wang;Shuyu Li IEEE Journal of Biomedical and Health Informatics Year: 2021 | Volume: 25, Issue: 3 | Journal Article | Publisher: IEEE

[5]A Research Survey on State of the art Heart Disease Prediction Systems Lakshmi Prasad Koyi;Tejaswi Borra;G. Lakshmi Vara Prasad 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS) Year: 2021 | Conference Paper | Publisher: IEEE

[6] Multi Disease Prediction System using Random Forest Algorithm in Healthcare System

R. Shanthakumari;C. Nalini;S. Vinothkumar;E.M. Roopadevi;B. Govindaraj2022 International Mobile and Embedded Technology Conference (MECON) Year: 2022 | Conference Paper | Publisher: IEEE