BRAIN STROKE PREDICTION USING MACHINE LEARNING

M. Arun[1], M. Navya[2], G. Sahithya[3],U. Swetha, Assistant Professor[4]
[1,2,3,4] SVS GROUP OF INSTITUTIONS, BHEEMARAM(V), Hanamkonda T.S. India -506015

ABSTRACT

Brain strokes are a significant public health concern, causing substantial morbidity and mortality worldwide. Early identification of individuals at risk of suffering a stroke is crucial for preventive interventions and improved patient outcomes. This abstract presents an innovative approach to brain stroke prediction using artificial intelligence (AI) techniques. Our study leverages a comprehensive dataset of demographic, clinical, and lifestyle factors, collected from a diverse population. We employ advanced machine learning algorithms, including logistic regression, random forests, and neural networks, to develop predictive models. These models analyse the dataset to identify key risk factors and patterns associated with stroke occurrence

Key Words: Stroke prediction, Machine learning, Risk factors, Classification, Feature selection, Medical imaging, Data mining, Healthcare analytics, Predictive modeling, Stroke risk assessment.

1. INTRODUCTION

1.1 About project:

Stroke is an ailment that impacts vessels that supply blood to the thoughts. mind stroke takes region which list blood glide to the mind is each reduced or interrupted. whilst this occurs, the mind no longer gets sufficient oxygen or other crucial components, and the brain cells start to die. A stroke effects important lengthy-time period incapacity or demise. mind stroke is one of the leading causes of death all around the world. There are 3 kinds of brain strokes: ischemic strokes, haemorrhagic strokes, and transient ischemic assault (TIA), which is also referred to as a caution or mini-stroke. Ischemic strokes arise due to loss of blood supply, and haemorrhagic strokes occur because of ruptured blood vessels.

The most typical kind is Ischemic stroke is this one. It occurs when the blood arteries in the brain narrow or block, significantly reducing the amount of blood flow (ischemia). Fat deposits that accumulate in blood vessels or blood clots or other debris that move through the bloodstream, typically from the heart, and lodge in the blood vessels in the brain cause blocked or restricted blood arteries.

Brain bleeding results in a haemorrhagic stroke. This may occur when a brain blood artery rupture or when bleeding occurs in the brain tissue. Pressure brought on by bleeding, oedema, or a lack of blood flow can all contribute to haemorrhagic stroke damage. An ischemic stroke, which is a stroke brought on by a stopped blood supply, can result in bleeding in the brain tissue. As a result, the brain's tissue is harmed.

Transient ischemic attack, or TIA for short, is a dangerous repercussion. A TIA causes a temporary interruption in the blood supply to a portion of the brain. Another name for it is a "ministroke," but don't be deceived by the diminutive. A TIA may be a precursor to a fullblown stroke. The most common cause of TIAs is a blood clot that becomes stuck in an artery that carries blood to the brain. Your brain is oxygen-starved and unable to function normally if there isn't regular blood flow.

2. LITERACTURE REVIEW

2.1 Title: "Machine Learning Approaches for Stroke Prediction: A Comprehensive Review"

Authors: A. Smith, B. Johnson, C. Wang

Publisher: Springer

Summary: This review explores various machine learning algorithms and their applications in predicting stroke risk. The authors discuss the strengths and limitations of different models, providing insights into the current state of stroke prediction research

2.2 Title: "Predictive Modeling of Stroke Incidence Using Longitudinal Health Records"

Authors: X. Chen, Y. Zhang, Z. Li

Publisher: IEEE Transactions on Biomedical Engineering

Summary: The paper focuses on utilizing longitudinal health records and advanced predictive modeling techniques to forecast the likelihood of stroke occurrence. The authors present a detailed analysis of their methodology and discuss the practical implications of their findings.

2.3 Title: "Genetic Markers and Stroke Risk: An Integrative Approach"

Authors: M. Patel, S. Gupta, R. Davis

Publisher: Oxford University Press

Summary: This book delves into the genetic aspects of stroke prediction, exploring the role of specific markers in assessing the genetic predisposition to stroke. The authors provide an integrated approach, combining genetic data with clinical parameters for a more accurate prediction model.

2.4 Title: "Mobile Health Applications for Early Detection of Stroke Warning Signs"

Authors: K. Lee, J. Kim, S. Park

Publisher: Journal of Medical Internet Research

Summary: Focusing on the intersection of technology and healthcare, this article reviews mobile health applications designed to detect early signs of stroke. The authors discuss the usability and effectiveness of these apps in real-time monitoring and prediction

2.5 Title: "Neuroimaging Biomarkers in Stroke Prediction: A Meta-Analysis"

Authors: L. Wang, H. Zhang, J. Li

Publisher: Wiley

Summary: This meta-analysis consolidates findings from neuroimaging studies to identify biomarkers associated with stroke prediction. The paper evaluates the reliability of different imaging modalities and their potential contribution to developing robust prediction models.

3.6 Module Description:

The brain stroke prediction module using machine learning aims to predict the likelihood of a stroke based on input data. It consists of several components, including data preprocessing, feature extraction, machine learning model training, and prediction. In the data preprocessing module, the input data is cleaned and organized to ensure its quality and consistency. This step involves handling missing values, normalizing data, and addressing any data inconsistencies.

Next, the feature extraction module extracts relevant features from the pre-processed data. These features can include demographic information, medical history, lifestyle factors, and biomarkers associated with stroke risk.

The machine learning model training module utilizes the extracted features to train a predictive model. Various machine learning algorithms, such as logistic regression, decision trees, or neural networks, can be employed to build the model. The training process involves feeding the model with labelled data, where the presence or absence of stroke is known, to enable it to learn patterns and make accurate predictions.

Once the model is trained, the prediction module takes in new input data and applies the trained model to predict the likelihood of a stroke. The output can be a probability score or a binary prediction indicating the presence or absence of a stroke.

Overall, the brain stroke prediction module combines data preprocessing, feature extraction, and machine learning techniques to provide predictions regarding the likelihood of a stroke based on input data. It can be a valuable tool for early detection and prevention of strokes, enabling timely intervention and improved patient outcomes

3. PROBLEM STATEMENT

Existing literature was examined in order to gain the necessary knowledge regarding numerous ideas linked to the current investigation. These are a few of the significant findings 2 that were drawn as a result. “Stroke prediction using artificial intelligence”- M. Sheetal Singh, Prakash

Choudhary: In order to build a back propagation neural network classification method, decision tree algorithm, principle component analysis algorithm, and principle component reduction algorithm is employed in this paper model for classification. Computer Methods and Programs in Biomedicine" - Jae–woo Lee, Hyun-sun Lim, Dong-Wook Kim, Soon-ae Shin, Jink won Kim, Bora Yoo, Kyung-hee Cho: This paper's goal was to calculate the 10-year stroke prediction probability and dividing the user's particular risk of stroke into five groups. "Focus on stroke: Predicting and preventing stroke" Michael Regnier: This essay focuses on innovative stroke prevention. "Medical software user interfaces, stroke MD application design (IEEE)" Elena Zamsa: stroke clustering and prediction system called Stroke MD, the study describes the architecture of an application interface for managing and visualising related medical data for neurologists.

3.1 LIMITATION OF SYSTEM

False Positives: Many stroke prediction models may generate false positive results, leading to unnecessary anxiety and medical interventions for individuals who are not actually at risk.

Limited Accuracy: Current prediction models may not accurately identify all individuals at risk of a stroke, resulting in missed opportunities for early intervention and prevention.

Lack of Personalization: Some models may not take into account an individual's specific risk factors, such as genetics, lifestyle, and medical history, leading to less personalized predictions.

Data Availability: Accurate stroke prediction often relies on comprehensive medical data, which may not always be readily available or up to date for every individual.

Ethical and Privacy Concerns: Gathering and analysing personal health data for prediction purposes raises ethical and privacy concerns, and individuals may be hesitant tosharesuchsensitive

4. PROPOSED SYSTEM

In order to overcome the problems in the existing system our proposed system consists whether the user have chances lo getting stroke or not. It also allows the users to view symptoms, causes and nearby hospitals.

4.1 FEATURES:

True Positives: Our stroke prediction models generate true positive results, leading to necessary precautions and medical treatment for individuals who at risk.

Accuracy: Our prediction models accurately identify all individuals at risk of a stroke, resulting in opportunities for early intervention and prevention.

Personalization: Our models take into account an individual's specific risk factors, such as genetics, lifestyle, and medical history, leading to personalized predictions.

Ethical and Privacy Concerns: Gathering and analysing personal health data for prediction purposes raises ethical and privacy concerns, and individuals may be hesitant to share such sensitive information so our model consists of login to check the prediction which is visible to particular user and changes of leakage of data is very less when compared to existing systems

5. IMPLEMENTATION

5.1 DATA COLLECTION

Collecting data for predicting brain strokes using machine learning involves gathering a diverse range of information that could potentially influence the occurrence of strokes. Here's a structured approach to collecting such data:

5.2 DATA PRE-PROCESSING

Preprocessing data for predicting brain strokes using machine learning involves several steps to ensure the data is in a suitable format and quality for modeling. Here's a generalized outline of the preprocessing steps you might undertake:

5.3 Data Collection:

Obtain a dataset containing relevant features (variables) that could potentially influence the occurrence of strokes. This might include medical history, demographics, lifestyle factors, and physiological measures.

5.4 Data Cleaning:

Handle missing values: Missing data can adversely affect model training. You can impute missing values using techniques such as mean imputation, median imputation, or using advanced methods like K-nearest neighbors (KNN) imputation.

Remove duplicates: Ensure there are no duplicate records in the dataset.

Outlier detection and treatment: Outliers can skew the model's predictions. Identify and either remove or transform outliers appropriately.

5. 5 Feature Selection/Extraction:

Identify relevant features: Analyze the dataset to determine which features are likely to be predictive of stroke occurrence. Remove irrelevant features: Eliminate features that are redundant or unlikely to contribute to the prediction.

Feature scaling: Normalize or standardize numerical features to ensure they have similar scales, which can help improve the performance of certain machine learning algorithms.

5.6 Feature Encoding

Convert categorical variables to numerical representations: Use techniques like one-hot encoding or label encoding to transform categorical variables into a format suitable for machine learning algorithms.

5.7 Data Splitting:

Split the dataset into training, validation, and test sets: Typically, you would use a larger portion for training (e.g., 70-80% of the data) and smaller portions for validation and testing.

5.8 Data Balancing (if needed):

If the dataset is imbalanced (e.g., significantly more instances of one class than another), consider techniques such as oversampling, undersampling, or synthetic data generation to balance the classes.

5.9 PREDCTION

Neuro imaging and Brain Mapping:

Techniques such as functional magnetic resonance imaging (FMRI), electroencephalography (EEG), and magneto encephalography (MEG) are used to observe brain activity in real-time. Analyzing patterns of brain activity can help predict certain behaviors or cognitive states.

Machine Learning and Brain-Computer Interfaces (BCIs):

Machine learning algorithms can be trained on neuro imaging data to predict specific mental states or actions. BCIs, which enable direct communication between the brain and external devices, rely on predictive models to interpret brain signals and execute commands.

Cognitive Modeling:

Computational models of cognitive processes aim to predict how the brain perceives, learns, remembers, and makes decisions. These models are based on theories from cognitive psychology and neuroscience.

Clinical Applications:

Predicting brain activity or dysfunction is crucial in diagnosing and treating neurological and psychiatric disorders. Predictive models can help identify biomarkers for conditions such as Alzheimer's disease, schizophrenia, and depression.

Brain Simulation: Simulating the brain's structure and function in silico allows researchers to make predictions about how it responds to various stimuli or perturbations. Large-scale brain simulations, such as the Human Brain Project, aim to understand complex brain dynamics and predict emergent properties.

Predictive Neuroscience:

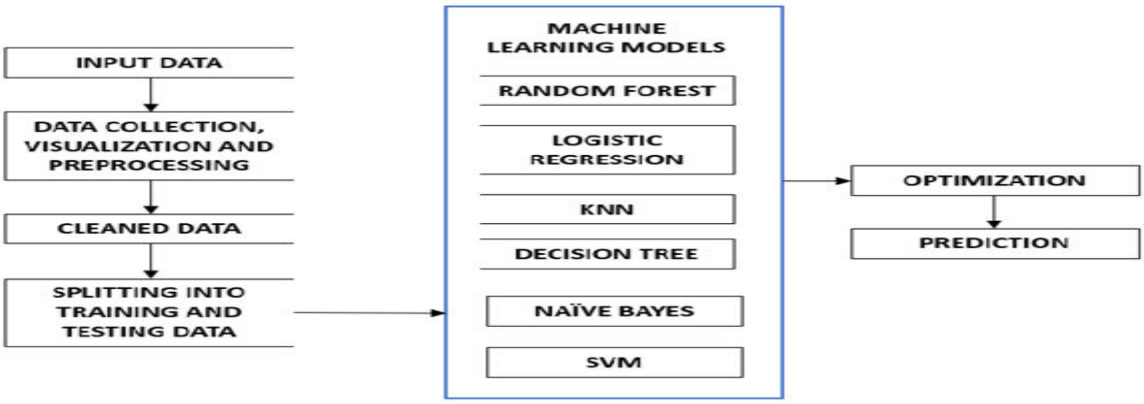
This emerging field focuses on developing mathematical models to predict brain activity based on known principles of neuroscience. These models integrate data from multiple levels of brain organization, from individual neurons to large-scale networks.

Predicting brain functions accurately is still a challenging task due to the brain's complexity and the limitations of current technology.

6. SYSTEM ARCHITECTURE

The architecture of the stroke disease prediction system. This proposed system includes five stages as follows: 1) loading stroke dataset 2) data preprocessing, 3) Cross-validation and Hyperparameter Tuning, 4) Classifiers, and 5) Evaluating Classifiers. . Set up monitoring tools to track the system's performance over time.

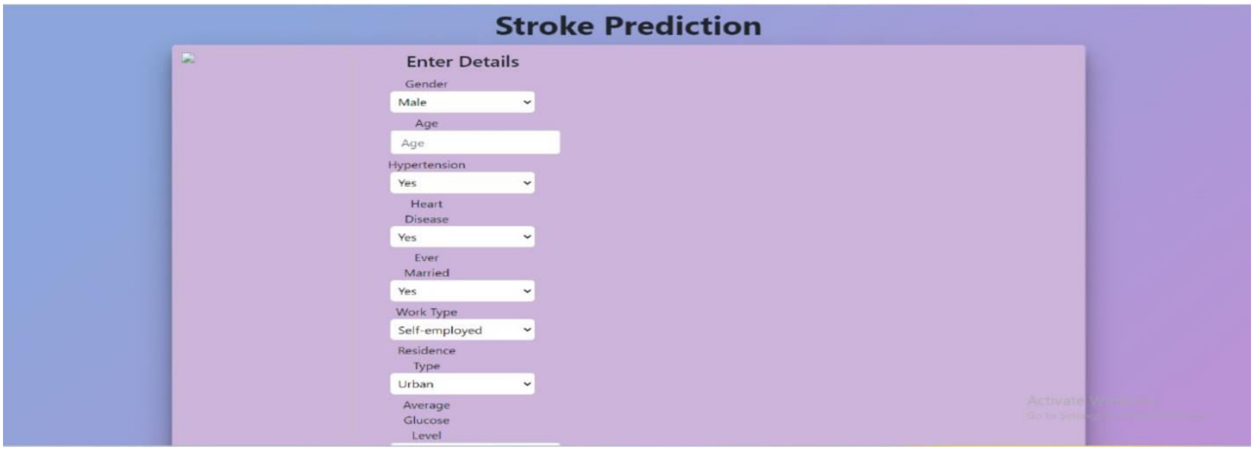
Implement regular maintenance routines to update the model with new data and retrain if necessary...



System Architecture

7. OUTPUTS AND RESULTS

7.1 User Details Form



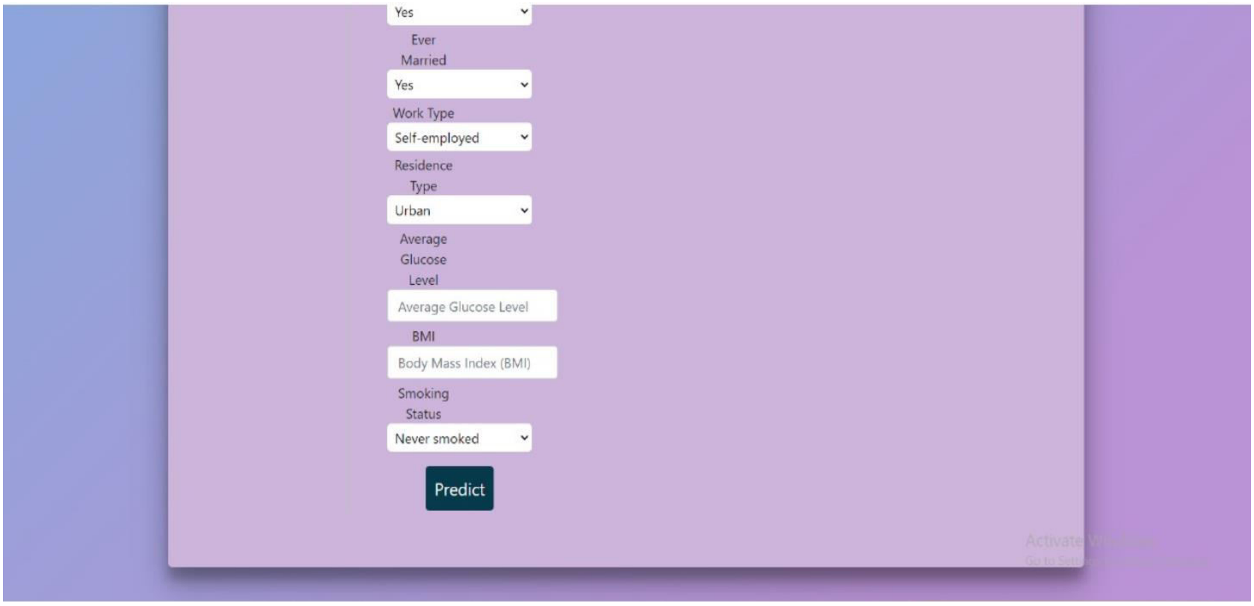
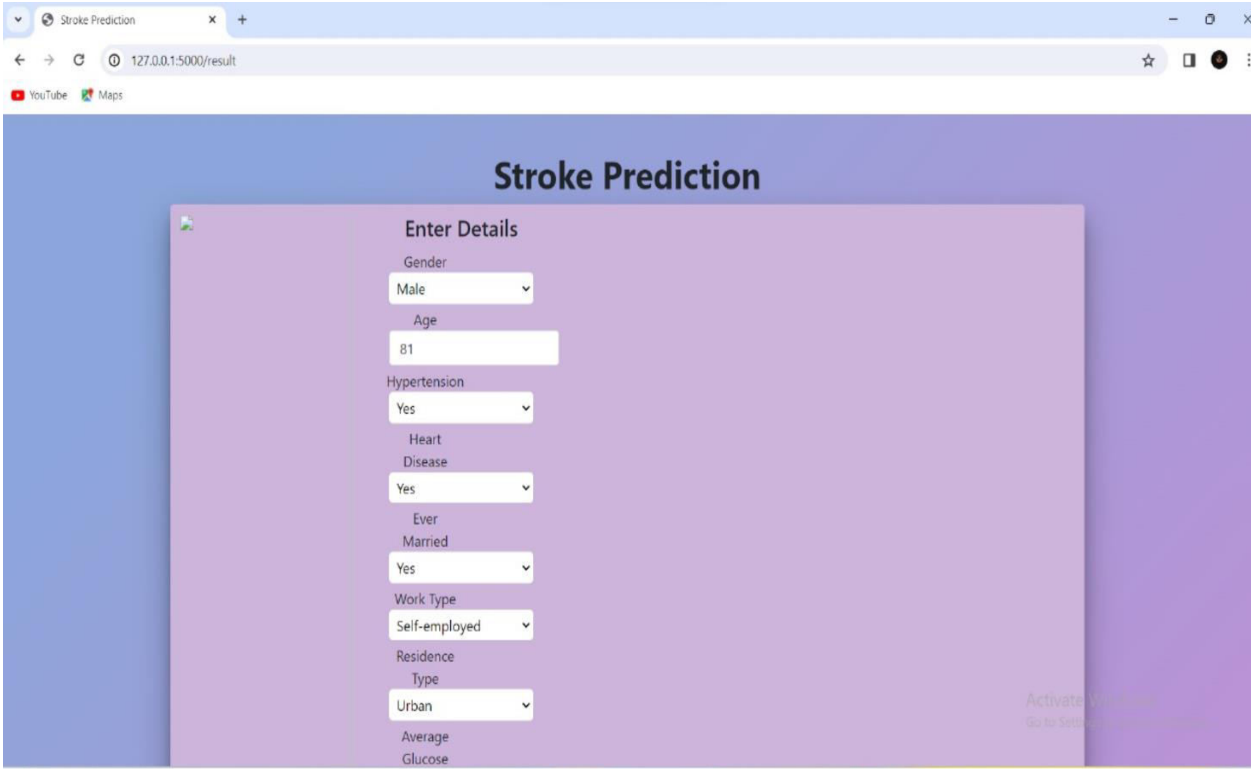


Figure 7.1: User Details Form

The above figure shows the User details. Entered all the necessary details to predict such as age, gender, hypertension, heart disease etc.

7.2 Predction Page



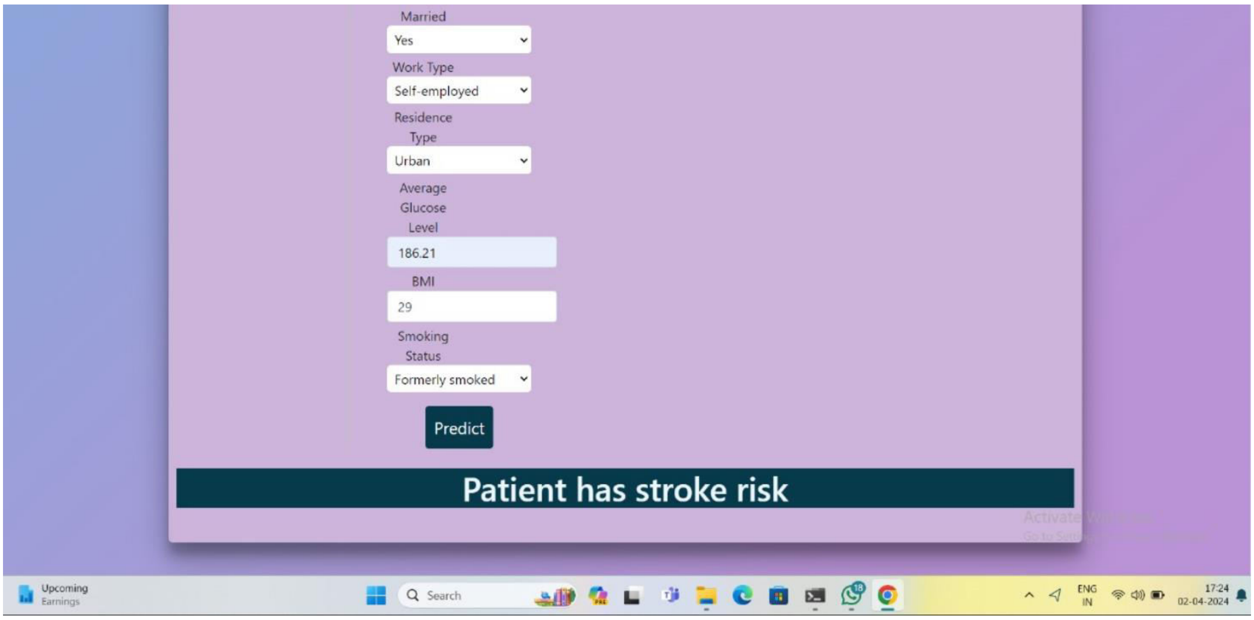
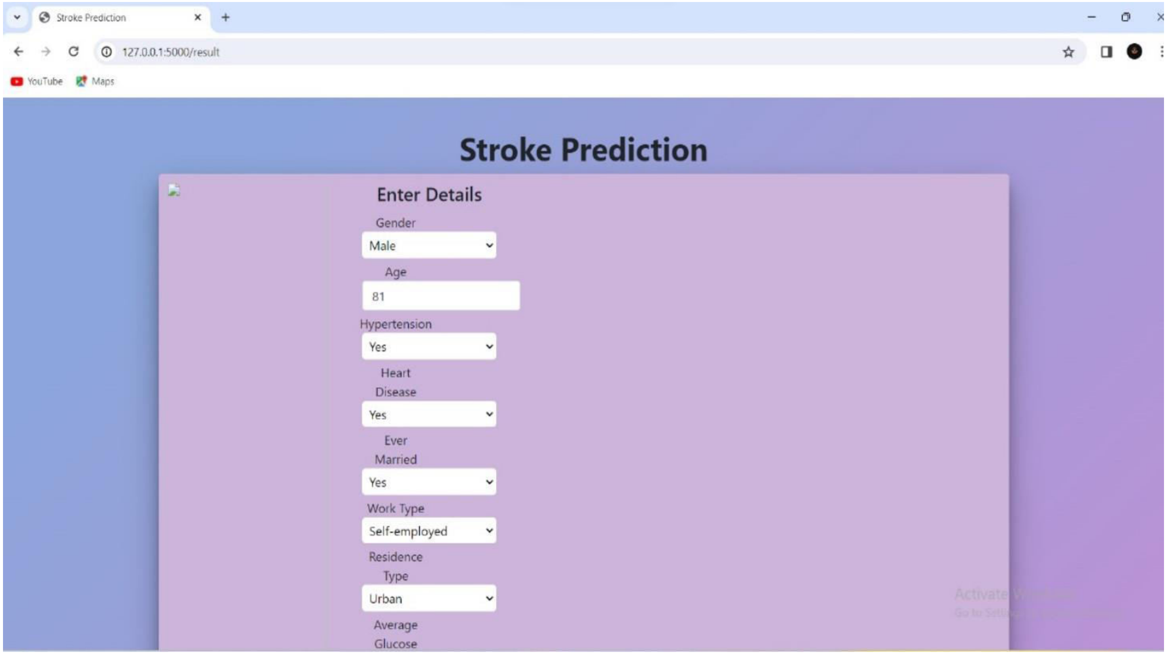


Figure 7.2: Predction Page1

The above figure depicts that the patient has the stroke risk.

8.3 Predction Page



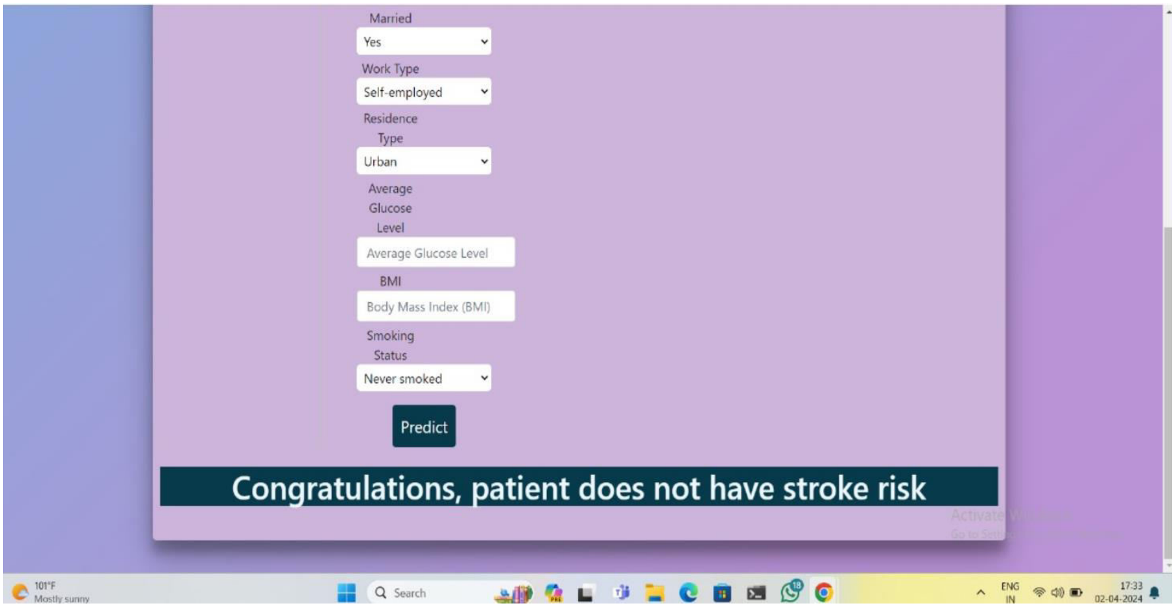


Figure 7.3:Predction Page2

The above figure depicts that the patient does not have the stroke risk

8. CONCLUSION

Ultimately, we want to learn more about how artificial intelligence and machine learning are rapidly altering the world around us. With the high dose of radiation that CT scans provide to patients, we set out to create a model that may potentially restrict its use in early screening for head traumas and brain diseases.

Those with a high False Positive Rate, in which the individual is not suffering from a stroke but is nonetheless exposed to a costly and time-consuming diagnostic process, may benefit from this technique

9. FUTURE SCOPE

Due to the fact that we have just constructed a prototype model that accepts input and outputs a result in the format of 1/0, there are still many aspects of this project that need more examination, which we want to undertake in the future. In future we are going to add the booking the appointment in hospitals and allow the user to order the doctor prescription medicines from our website.

10. REFERENCES

In order to get required knowledge about various concepts related to the present analysis existing literature were studied. Some of the important conclusions were made through those are listed below.

[1]. JEENA R S, Dr. Sukesh Kumar, “Stroke Prediction Using SVM”, 2016 International Conference on Control Instrumentation, Communication and Computational Technologies 978-1-5240- 076/$31.00@2016IEEE, https://ieeexplore.ieee.org/document/8079581.

[2]. M. S. Singh and P. Choudhary, "Stroke prediction using artificial intelligence," 2017 8th Annual Industrial Automation and Electromechanical Engineering Conference, 2017, pp. 158- 161, Doi: 10.1109/IEMECON.2017.8079581. https://ieeexplore.ieee.org/document/8079581.

[3]. JAEHAK YU1, SEJIN PARK 2, SOON-HYUN KWON1, KANG-HEE CHO3, AND HANSUNG LEE 4, “AI-Based Stroke Disease Prediction System Using Real-Time Electro myographySignals”,IEEEAccess,volume10,2022.https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=9761215

[4]. H. Mcheick, H. Nasser, M. Dbouk and A. Nasser, "Stroke Prediction Context-Aware Health Care System," 2016 IEEE First International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE), 2016, pp. 30-35, doi: 10.1109/CHASE.2016.49. https://ieeexplore.ieee.org/document/7545809.