# 1.INTRODUCTION

## 1.1 Project Scope:

The project aims to develop a stroke risk prediction system using a hybrid deep transfer learning framework. This framework will leverage pre-trained deep learning models to extract meaningful features from medical data, enhancing prediction accuracy. The system will analyse a diverse dataset of patient information, including demographics, medical history, and genetic data. By integrating both image and text-based data, it will provide a comprehensive prediction of an individual's stroke risk. The project's scope encompasses model training, data collection, feature extraction, and evaluation, with the goal of creating an accurate and interpretable tool to assist healthcare professionals in identifying

## 1.2 Project Purpose:

The purpose of the project is to develop a stroke risk prediction system using a hybrid deep transfer learning framework. This innovative approach combines deep learning techniques with transfer learning to leverage pre-trained models and medical data. By analyzing a patient's medical history, genetic information, and lifestyle factors, the system aims to accurately assess their risk of suffering a stroke. The ultimate goal is to enable early identification of individuals at high risk, facilitating timely interventions and personalized preventive measures to reduce the incidence of strokes, improve patient outcomes, and save lives.

## 1.3 Project Features:

The Stroke Risk Prediction with Hybrid Deep Transfer Learning Framework is a cutting-edge project that leverages deep learning and transfer learning techniques to predict an individual's risk of suffering a stroke. It combines the power of pre-trained neural networks with patient-specific data to enhance prediction accuracy. This innovative approach not only uses medical images and clinical data but also adapts knowledge from broader healthcare datasets to make more precise stroke risk assessments. The project aims to revolutionize stroke prevention and early intervention by providing a reliable, personalized prediction tool, ultimately saving lives and improving healthcare outcomes.

# 2.SYSTEM ANALYSIS

## 2.1 System Analysis:

System Analysis is the important phase in the system development process. The System is studied to the minute details and analyzed. The system analyst plays an important role of an interrogator and dwells deep into the working of the present system. In analysis, a detailed study of these operations performed by the system and their relationships within and outside the system is done. A key question considered here is, “what must be done to solve the problem?” The system is viewed as a whole and the inputs to the system are identified. Once analysis is completed the analyst has a firm understanding of what is to be done.

## 2.2 Problem Definition:

The problem of social media and misleading information in a democracy, from a mechanism design perspective, can be defined as follows:

In a democratic society, information plays a crucial role in shaping public opinion and decision-making. Social media platforms have become a primary source of information dissemination. However, these platforms also face challenges related to the spread of misleading or false information, often referred to as "fake news" or misinformation. This poses a significant threat to the functioning of a democracy, as it can influence public opinion, electoral outcomes, and policy decisions.

A mechanism design approach to this problem involves designing incentives and rules within social media platforms to mitigate the spread of misleading information while preserving free speech and open discourse. This may include the design of algorithms, content moderation policies, and user incentives to encourage the sharing of accurate and reliable information while discouraging the dissemination of false or misleading content. The challenge lies in striking a balance between combating misinformation and upholding principles of freedom of expression and diversity of perspectives, which are essential in a democratic society.

## 2.3 Existing System:

Proposed a feature selection algorithm combined with SVM and carried out experiments in a dataset that has 4; 988 examples with 299 occurrences of stroke. Miguel et al.used machine learning techniques to predict the functional outcome of a patient three months after the initial stroke in a dataset with 425 acute stroke patients. Lim et al. used a deep learning approach to develop SRP model using retinal images and achieved acceptable performance. But collecting retinal images is often time-consuming and expensive. From the above works, we can find that the amount of stroke data is less and the dataset is imbalanced. To solve the imbalance issue, Liu et al. proposed a DNN-based model to predict stroke on an imbalanced dataset which contains 43; 400 records of patients with 783 occurrence of stroke. However, the above methods cannot address the small data and imbalance issues of stroke data.

Intuitively, centralizing all available data from other hospitals can build a better SRP model when amounts of stroke data available in local hospitals are small.

Secure Multiparty Computation (SMC) model involves multiple party data and protects data privacy. But it demands each party knows nothing except its input and output which is difficult to achieve. Another recently proposed framework, federated learning , is also used to protect multiple party data privacy. It can exploit multiple distributed data to establish a better model. In federated learning, data in different places would not be transmitted and the model is encrypted during training. But federated learning framework cannot solve the issue of small data. Notably, the above methods have not been deployed in SRP application.

The parameter of Network Weight transfer module is an import factor of HDTL-SRP model performance. Traditionally, manual search and grid search are strategies for hyperparameter optimization in a neural network. For the same time budget, random search finds better models than grid search results by effectively searching a larger and less promising configuration space. But they find a proper hyperparameter randomly or rely on the expert's experience. Genetic Algorithm (GA) is an evolutionary search algorithm used to solve optimization. However, GA demands enough initial sample points in hyper-parameter optimization so that the optimization efficiency is usually low. Bayesian Optimization (BO) is a flexible approach in hyperparameter optimization, while BO based on Gaussian processes achieve successful implementation, as the Spearmint system.

### 2.3.1 Disadvantages Of Existing System:

- The system is not implemented HYBRID DEEP TRANSFER LEARNING FOR STROKE RISK PREDICTION.
- In an existing system, Generative Instance doesn't Transfer using External Stroke Data.

## 2.4 Proposed System:

The proposed framework can achieve a better ability in establishing SRP model. However, the parameters such as the number of transferred layer and the transferred sequence of different source domains are vital factors for model performance. Common methods such as grid and random search for parameter tuning are often inefficient due to the search space being too large . Bayesian Optimization (BO) is an approach for model-based global optimization of blackbox function and the most universally used model for BO is a Gaussian process due to its simplicity and flexibility in constructing a probabilistic model of objective function. Therefore, BO is used to find the best parameter in SRP model. The contributions of this work are as follows:

This work proposes a novel Hybrid Deep Transfer Learning-based Stroke Risk Prediction (HDTL-SRP) framework that allows simultaneous exploitation of multiple correlated sources of medical data (e.g., hypertension, diabetes, and external stroke data) to train a SRP model in a local hospital where only small and imbalanced stroke data is available.

The proposed framework is extensively compared with the state-of-the-art SRP models in synthetic scenario of stroke risk prediction. In addition, our approach is tested in a realworld dataset which contains 2; 426 stroke incidents recorded in three collaborating hospitals during 2012-2017. The empirical results show that the performance of HDTL-SRP outperform its counterparts in both synthetic and real-world scenarios.

The proposed HDTL-SRP framework can be deployed among multiple hospitals to take advantage of the distributed medical data while preserving the patients' privacy. In addition, this decentralized framework can gain efficiency when the 5G/B5G infrastructures are available among hospitals.

In this work, Bayesian Optimization is used to find the best parameters such as the number of transferred layer and the transferred sequence of different source domains for HDTLSRP model.

### 2.4.1 Advantages Of Proposed System:

- The goal of the system is to model data samples that have been used to train a ML classifier for finding an exact and high accuracy.
- The proposed system developed a Network Structure Optimization in Multiple Sources Using Bayesian Optimization with ml classifiers.

## 2.5 Feasibility Study:

The feasibility ofthe project is analyzed in this phase and business proposal is put forth with a very general plan for the project and some cost estimates. During system analysis the feasibility study of the proposed system is to be carried out. This is to ensure that the proposed system is not a burden to the company. For feasibility analysis, some understanding of the major requirements for the system is essential.

Three key considerations involved in the feasibility analysis are:

- ECONOMICAL FEASIBILITY
- TECHNICAL FEASIBILITY
- SOCIAL FEASIBILITY
- OPERATIONAL FEASIBILITY

### 2.5.1 Economical Feasibility:

This study is carried out to check the economic impact that the system will have on the organization. The amount of fund that the company can pour into the research and development of the system is limited. The expenditures must be justified. Thus the developed systemas well within the budget and this was achieved because most of the technologies used are freely available. Only the customized products had to be purchased.

### 2.5.2 Technical Feasibility:

This study is carried out to check the technical feasibility, that is, the technical requirements of the system. Any system developed must not have a high demand on the available technical resources. This will lead to high demands onthe available technical resources. This will lead to high demands being placed onthe client. The developed system must have a modest requirement, asonly minimal or null changes are required for implementing this system.

### 2.5.3 Social Feasibility:

The aspect of study is to check the level of acceptance of the system by the user. This includes the process of training the user to use the system efficiently. The user must not feel threatened by the system, instead must accept it as a necessity. The level of acceptance by the users solely depends on the methods that are employed to educate the user about the system and to make with it. His level of confidence must be raised so that he is also able to make some constructive criticism, which is welcomed, as he is the final user of the system.

### 2.5.4 Operational Feasibility:

Operational Feasibility deals with the study of prospects of the system to be developed. This system operationally eliminates all the tensions of the Admin and helps him in effectively tracking the project progress. This kind of automation will surely reduce the time and energy, which previously consumed in manual work. Based on the study, the system is proved to be operationally feasible.

## 2.6 Hardware and Software Requirements:

### 2.6.1 Hardware Requirements:

- Processor - Pentium –IV
- RAM - 4 GB (min)
- Hard Disk - 20 GB
- Key Board - Standard Windows Keyboard
- Mouse - Two or Three Button Mouse
- Monitor - SVGA

### 2.6.2 SOFTWARE REQUIREMENTS:

- Operating system : Windows 7 Ultimate
- Coding Language : Python.
- Front-End : Python.
- Back-End : Django-ORM
- Designing : Html, css, javascript
- Data Base : MySQL (WAMP Server)

# 3.ARCHITECTURE

## 3.1 Project Architecture:

Our project architecture shows the procedure followed for classification, starting from Data pre-preprocessing to Performance Analysis.

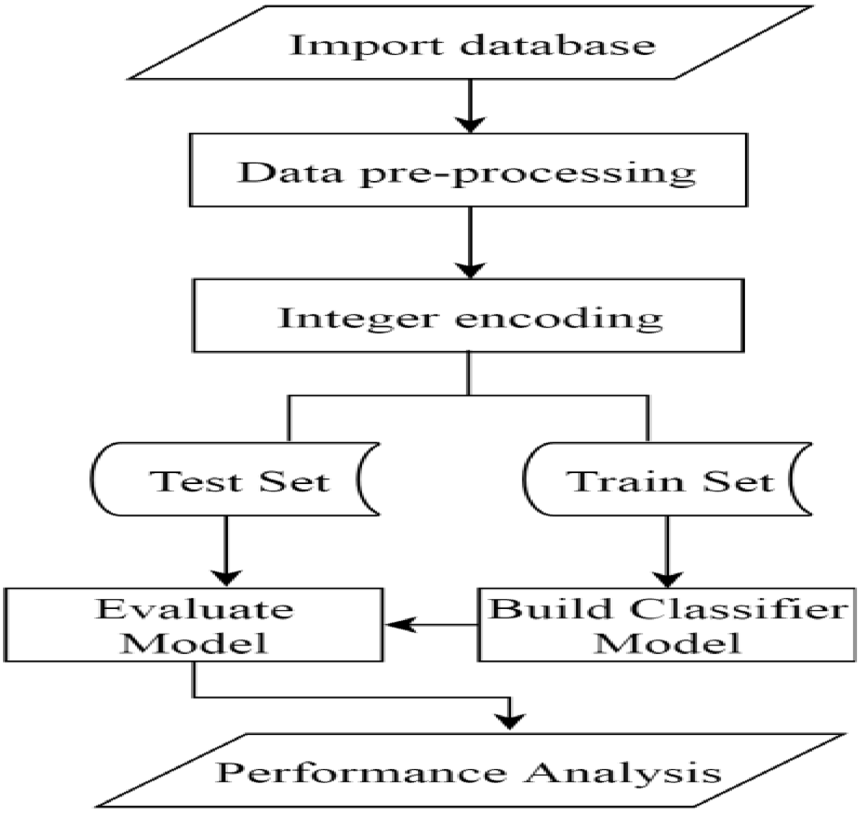


Figure 3.1: Architecture for Stroke Risk Prediction With Hybrid Deep Transfer Learning

## 3.2 Description:

The " Stroke Risk Prediction With Hybrid Deep Transfer Learning " project aims to enhance the security and confidentiality of patient data in healthcare settings. The system employs advanced encryption and authentication mechanisms to safeguard sensitive information. It incorporates real-time monitoring to detect and prevent potential privacy breaches, ensuring the forward privacy of patient data. The project also integrates decentralized storage solutions to mitigate the risk of centralized data vulnerabilities. It helps for predicting the Stroke in advance and reduce the risk of stroke causes.

## 3.3 Use-case Dia gram:

In the use case diagram, we have basically one actor who is the user in the trained model. A use case diagram is a graphical depiction of a user's possible interactions with a system. A use case diagram shows various use cases and different types of users the system has. The use cases are represented by either circles or ellipses. The actors are often shown as stick figures.

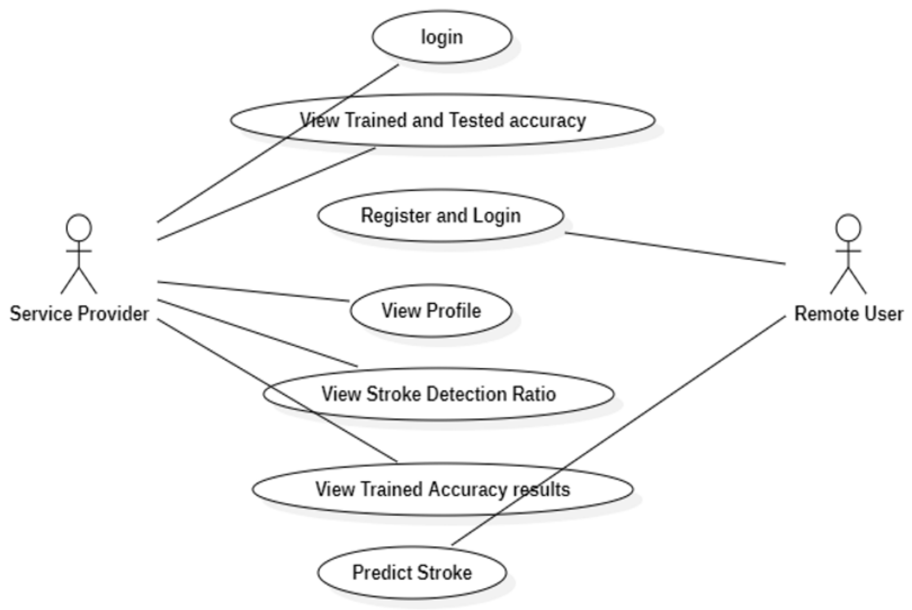


Figure 3.3: Use-case diagram for Stroke Risk Prediction With Hybrid Deep Transfer Learning

## 3.4 Class Diagram:

Class diagram is a type ofstatic structure diagramthat describesthe structure ofa system by showing the system's classes, their attributes, operations (or methods), and the relationships among objects.

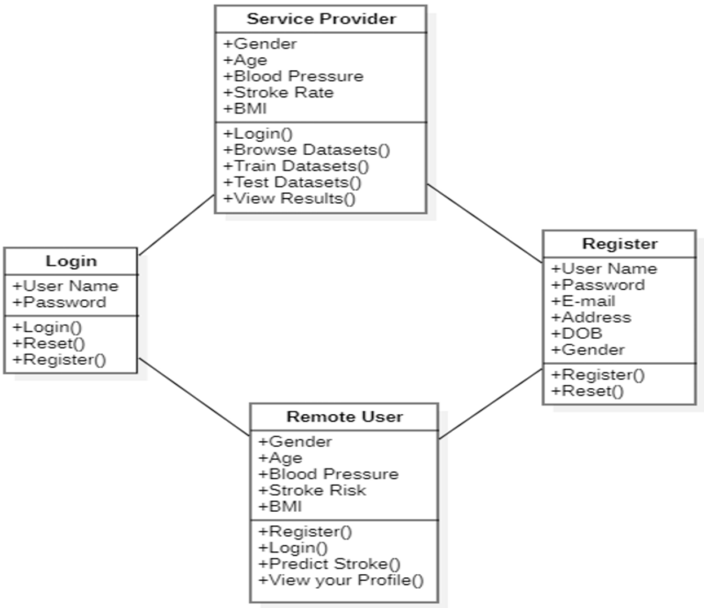


Figure 3.4: Class diagram for Stroke Risk Prediction With Hybrid Deep Transfer Learning

## 3.5 Sequence Diagram:

A sequence diagram shows object interactions arranged in time sequence. It depicts the objects involved in the scenario and the sequence of messages exchanged between the objects needed to carryout the functionality ofthe scenario. Sequence diagrams are typically associated with use case realizations in the logical view of the system under development.

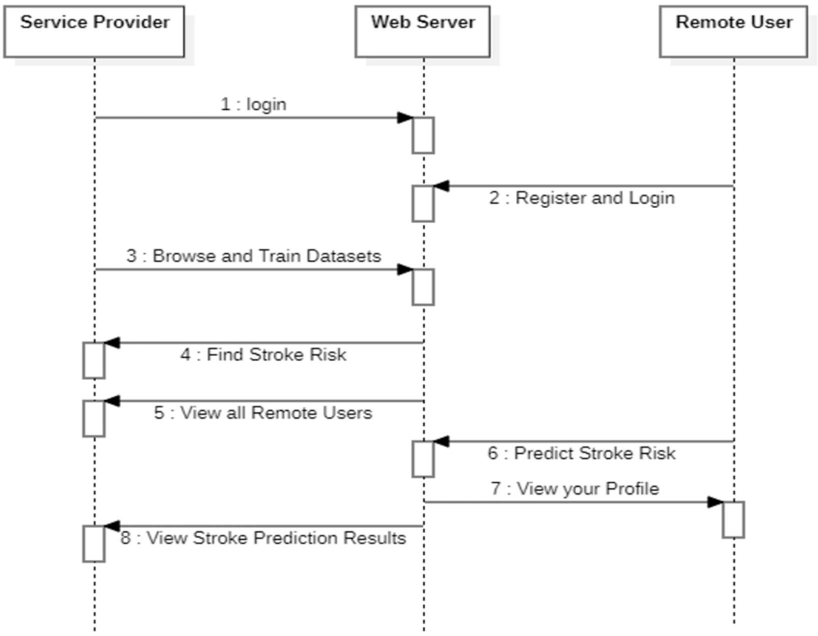


Figure 3.5: Sequence diagram for Stroke Risk Prediction With Hybrid Deep Transfer Learning

## 3.6 Activity Diagram:

Activity diagrams are graphical representations of workflows of step wise activities and actions with support for choice, iteration and concurrency. They can also include elements showing the flow of data between activities through one or more datastores.

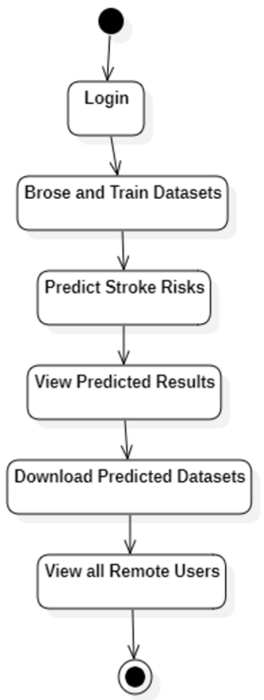


Figure 3.6: Activity diagram for Stroke Risk Prediction With Hybrid Deep Transfer Learning