BRAIN TUMOR PREDICTION
USING ARTIFICIAL NEURAL NETWORK

MINI PROJECT REPORT

Submitted by

Ragava Krishnan N S **210701201**

Sadiq Peer Mohamed K **210701522**

in partial fulfillment for the award of the degree of

BACHELOR OF ENGINEERING

in

COMPUTER SCIENCE AND ENGINEERING

RAJALAKSHMI
ENGINEERING COLLEGE
An AUTONOMOUS Institution
Affiliated to ANNA UNIVERSITY, Chennai

ANNA UNIVERSITY
PROGRESS THROUGH KNOWLEDGE

RAJALAKSHMI ENGINEERING COLLEGE, CHENNAI
ANNA UNIVERSITY, CHENNAI 600 025

APRIL 2024

RAJALAKSHMI ENGINEERING COLLEGE, CHENNAI

BONAFIDE CERTIFICATE

Certified that this Report titled "**BRAIN TUMOR PREDICTION USING ARTIFICIAL NEURAL NETWORK"** is the bonafide work of **"Ragava Krishnan N S (210701201) and Sadiq Peer Mohamed K (210701522)"** who carried out the work under my supervision. Certified further that, to the best of my knowledge, the work reported herein does not form part of any other thesis or dissertation on the basis of which a degree or award was conferred on an earlier occasion on this or any other candidate.

SIGNATURE

Rakesh Kumar M, M.E., Ph.D.,
Assistant Professor,
Department of Computer Science and Engineering,
Rajalakshmi Engineering College,
Chennai – 602015

Submitted to Mini Project Viva-Voce Examination held on ____________________

Internal Examiner **External Examiner**

ABSTRACT

Early detection of brain tumors is crucial for proper treatment and improved patient care. This research introduces a methodology employing artificial neural networks (ANN) to predict the presence of brain tumors based on doctor reports. These reports serve as the dataset, from which relevant features like radius mean, symmetry mean, concavity, and fractal dimensions are extracted as inputs for the ANN model. The ANN model uses the sigmoid activation function to predict the likelihood of a tumor.

Each layer of the ANN consists of 250 neurons, optimizing the network's ability to learn and generalize from the data. This approach enhances early detection and decision-making in medical settings. By accurately identifying potential tumors at an earlier stage, healthcare providers can implement timely and appropriate treatments. The methodology not only supports physicians in diagnosis but also improves the overall quality of patient care.

It represents a significant step forward in leveraging artificial intelligence for medical applications, providing a robust tool for aiding clinical decisions. The integration of ANN in analyzing medical reports can thus transform the landscape of brain tumor detection and treatment, leading to better patient outcomes.

ACKNOWLEDGEMENT

Initially, we thank the Almighty for being with us through every walk of our life and showering his blessings through the endeavor to put forth this report. Our sincere thanks to our Chairman, **Mr. S. MEGANATHAN, B.E, F.I.E.**, our Vice Chairman **Mr. ABHAY SHANKAR MEGANATHAN, B.E., M.S.,** and our respected Chairperson **Dr. (Mrs.) THANGAM MEGANATHAN**, **Ph.D.,** for providing us with the requisite infrastructure and sincere endeavoring in educating us in their premier institution.

Our sincere thanks to **Dr. S.N. MURUGESAN, M.E., Ph.D.,** our beloved Principal for his kind support and facilities provided to complete our work in time. We express our sincere thanks to **Dr. P. KUMAR, Ph.D.**, Professor and Head of the Department of Computer Science and Engineering for his guidance and encouragement throughout the project work. We convey our sincere and deepest gratitude to our internal guide, **Rakesh Kumar M, M.E., Ph.D.,** Assistant Professor, Department of Computer Science and Engineering. Rajalakshmi Engineering College for his valuable guidance throughout the course of the project.

Ragava Krishnan N S - 210701201

Sadiq Peer Mohamed K - 210701522

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ANN	Artificial Neural Networks
TCIA	The Cancer Imaging Archive
DDCGAN	Dual Discriminator Conditional Generative Adversarial Network
CAD	Computer-Aided Diagnosis
JSON	JavaScript Object Notation

CHAPTER 1
INTRODUCTION

1.1 GENERAL

The project introduces a methodology utilizing artificial neural networks (ANN) for brain tumor prediction, using doctor reports as the dataset. It emphasizes the critical need for early tumor detection and proposes ANN with 250 neurons in each layer. By analyzing diverse datasets, ANN can identify patterns indicative of tumor presence, facilitating timely intervention. The research underscores the potential of ANN to enhance diagnostic accuracy and decision-making in healthcare.

1.2 OBJECTIVE

The objective of the study is to develop a brain tumor prediction system using artificial neural networks (ANN) based on doctor reports. It aims to improve early detection and decision-making in healthcare by leveraging ANN's ability to analyze diverse datasets for identifying tumor patterns. The study seeks to enhance diagnostic accuracy and support medical professionals in timely intervention for improved patient outcomes. Through the integration of ANN into clinical practice, the research aims to contribute to advancements in medical AI applications.

1.3 EXISTING SYSTEM

The existing system primarily relies on convolutional neural networks (CNN) for brain tumor detection, utilizing MRI scan images as the dataset. It involves training CNN models on labeled images to identify tumor patterns, achieving a prediction accuracy of 93%. This system emphasizes the importance of expert interpretation and medical imaging in traditional diagnostic methods. However, it faces challenges in early detection and decision-making, prompting exploration into alternative approaches like ANN.

1.4 PROPOSED SYSTEM

The proposed system in the project suggests utilizing artificial neural networks (ANN) for brain tumor prediction, employing doctor reports as the dataset instead of MRI scan images. This approach aims to overcome the limitations of traditional methods by streamlining diagnostic processes and supporting early detection. The ANN model consists of 250 neurons in each layer and is trained over 1000 epochs. By analyzing features extracted from doctor reports, the system predicts the presence of brain tumors, aiding medical professionals in diagnosis and decision-making for improved patient outcomes.

CHAPTER 2

LITERATURE SURVEY

Many researchers have proposed different methodologies for brain tumor detection using various machine learning techniques, the work done by D C Febrianto, I Soesanti , H A Nugroho[1] which is the base paper for our project performs image classification and uses CNN model to detect brain tumor and has obtained a prediction accuracy of 93%. The work done by Ozyurt F, Sert E, Avci E, Dogantekin E[4] proposes a hybrid solution combining Neutrosophy and CNN, the performance of the system is evaluated against an SVM classifier and the solution provides an accuracy of 95.4%. The solution proposed by Selvi K, Sumaiya Begum a Poonkuzhali P, Aarthi R[5] implements a Dual Discriminator Conditional Generative Adversarial Network (DDCGAN) and the images from the dataset is proposed using Structural interval gradient filtering and the resulted model is evaluated and produces an accuracy of 93%.

The work done by Appiah, Helber Antonio, Cristiano Cabrera[6] implements convolutional neural networks with proper orthogonal decomposition for identifying brain tumors efficiently and the resulting model has obtained a prediction accuracy of 95%. The work done by A. Lumini, G. F. Roberto, L. A. Neves, A. S. Martins, and M. Z. do Nascimento, [7] propose a hybrid methodology combining both fractal geometry features and deep learning which tries to find the important spacial features in brain images, both original and the percolation image is fed as input for the CNN. In the work done by M. S. Ullah, M. A. Khan, [8] addresses an important issue in Computer Aided diagnosis(CAD) and proposes a convolutional network with Stack auto encoders along with a parallel pooling mechanism and achieves an accuracy rate of 94%.

The work by V. Akoto-Adjepong, O. Appiah [9] propose a solution using a Capsule Network (CapsNets) called Tri Texton-Dense (TTDCapsNet) for recognising medical images and predicting brain tumor presence and the model achieves an accuracy rate of 94%. In the method proposed by

C. Ozdemir and Y. Dogan[10] uses a MTAP model along with Avg-TopK pooling method for extracting features from the images and the MTAP model achieves an accuracy of 95%.

CHAPTER 3
SYSTEM DESIGN

3.1 DEVELOPMENT ENVIRONMENT

3.1.1 HARDWARE SPECIFICATIONS

This project uses minimal hardware but in order to run the project efficiently without any lack of user experience, the following specifications are recommended.

Table 3.1.1 Hardware Specifications

PROCESSOR	Intel Core i5
RAM	4GB or above (DDR4 RAM)
GPU	Intel Integrated Graphics
HARD DISK	6GB
PROCESSOR FREQUENCY	1.5 GHz or above

3.1.2 SOFTWARE SPECIFICATIONS

The software requirements for the proposed system include Jupyter Notebook and Python for developing and implementing the artificial neural network (ANN) model. Additionally, a web browser such as Chrome or Edge is necessary for accessing and visualizing the results. These software tools facilitate the training, testing, and validation of the ANN model using the provided doctor reports dataset.

Table 3.1.2 Software Specifications

FRONT END	HTML, CSS, JavaScript
BACK END	Python, Flask
FRAMEWORKS	TensorFlow
SOFTWARES USED	Visual Studio Code, Jupyter Notebook
LIBRARIES	Numpy, Pandas

3.2 SYSTEM DESIGN

3.2.1 ARCHITECTURE DIAGRAM

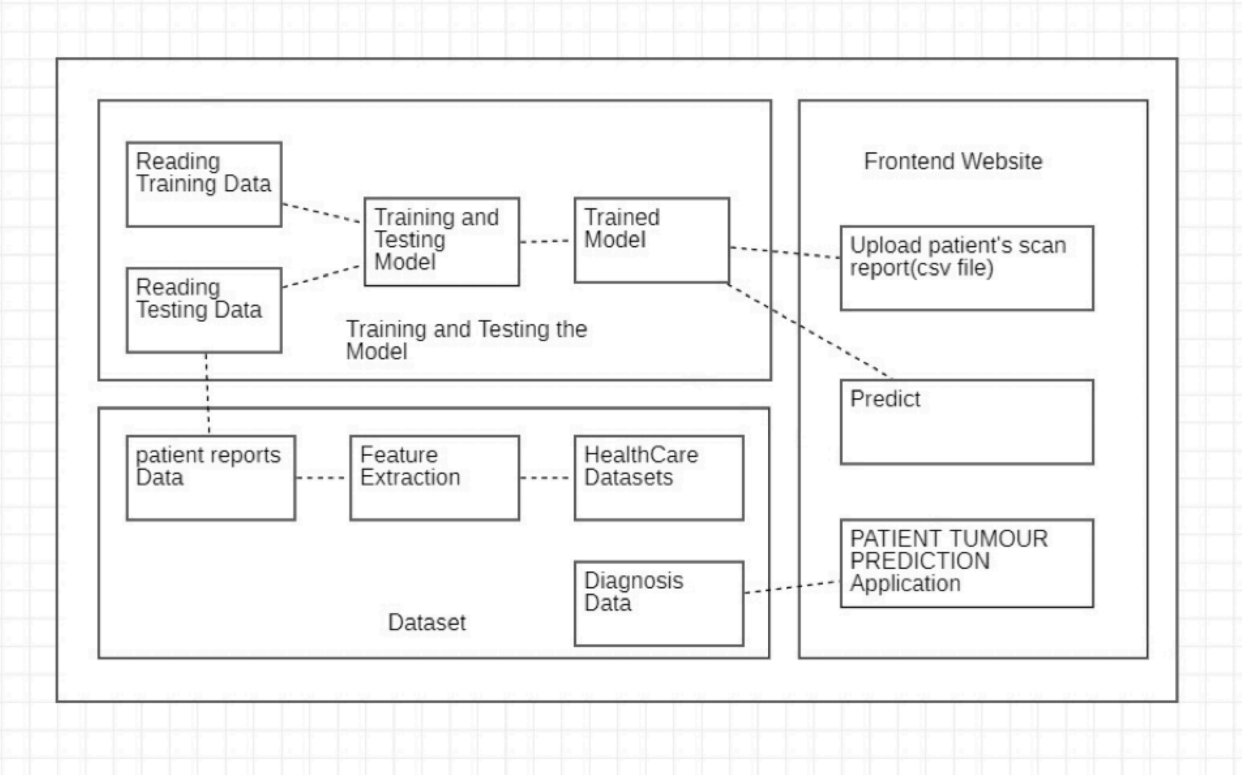

Fig 3.2.1 Architecture Diagram

PRE-PROCESSING:

Preprocessing is a critical step in ensuring raw data is clean, consistent, and properly formatted for input into a machine learning model. This process begins with loading the dataset from a CSV file using Pandas, a powerful library for data manipulation in Python. Once the data is loaded into a DataFrame, the next step is handling missing values, which involves techniques like imputation—filling missing values with the mean, median, or mode—or removing rows or columns with missing values. This step ensures that the dataset is complete and reliable for training the model.

Subsequently, unnecessary columns that do not contribute to the prediction are dropped to reduce the dataset's dimensionality and focus on the most relevant features. Feature extraction follows, where key attributes such as radius mean, symmetry mean, concavity, and fractal dimensions are identified and selected as they are crucial for accurate brain tumor detection.

TRAINING SET:

The train set in a machine learning project is the portion of the dataset used to train the model. In this context, the train set is derived from a dataset of doctor reports containing information pertinent to brain tumors, such as radius mean, symmetry mean, concavity, and fractal dimensions.

Dataset Splitting: The complete dataset, after preprocessing, is split into two subsets: the train set and the test set. Typically, an 80-20 split is used, meaning 80% of the data is used for training the model, while the remaining 20% is used for testing its performance. This splitting is done using the train_test_split function from the Scikit-Learn library.

Feature Selection: The features used for training include critical attributes that are likely to impact the prediction accuracy. These features are selected based on their relevance to the problem, such as radius mean, texture, area, concavity, concave points, and fractal dimensions. In this project, unnecessary columns are dropped to focus on these important features.

Normalization: Before training, the features in the train set are normalized or scaled to ensure they have a consistent scale. This is essential because machine learning models, particularly neural networks, perform better when input features are on a similar scale.

Model Training: The training process involves feeding the train set into the neural network model. The model learns from this data by adjusting its internal parameters (weights and biases) to minimize the error in its predictions. This process is iterative and involves multiple epochs, where the entire train set is passed through the model several times.

Label Encoding: The target variable (labels) in the train set, which indicates the presence of a tumor, is encoded appropriately if it is categorical. For a binary classification task like tumor detection, labels are usually 0 (no tumor) and 1 (tumor).

Validation: During training, a part of the train set may be used for validation to tune hyperparameters and prevent overfitting. This validation set helps in monitoring the model's performance on unseen data during the training process.

CHAPTER 4
PROJECT DESCRIPTION

4.1 MODULE DESCRIPTION

4.1.1 DATA PRE-PROCESSING:

In the provided paper and code, data preprocessing involves loading the doctor reports dataset from a CSV file using Pandas. Missing values in the dataset are handled gracefully, ensuring data completeness. Relevant features like radius mean, symmetry mean, concavity, and fractal dimensions are extracted for brain tumor prediction. These features are then scaled or normalized to ensure consistent scales across the dataset. Finally, the dataset is split into training and testing sets, facilitating model training and evaluation for accurate brain tumor detection.

4.1.2 TRAINING SET:

In the context of the first provided paper and code, the training set comprises a portion of the doctor reports dataset used to train the neural network model. It includes features extracted from the reports, such as radius mean, symmetry mean, concavity, and fractal dimensions, which serve as inputs for the model. This dataset undergoes preprocessing steps like normalization and scaling to ensure consistency. The neural network is trained on this data to learn patterns indicative of brain tumor presence. The training process involves iteratively adjusting the model's parameters to minimize prediction error, ultimately enabling accurate tumor detection.

4.1.3 TRAINING MODEL:

In the context of the provided paper and code, the training model refers to a neural network architecture designed to predict brain tumor presence. This model is constructed using TensorFlow, with layers consisting of densely connected neurons. Activation functions like sigmoid are employed to introduce non-linearity, aiding in complex pattern recognition. The model is compiled with appropriate loss functions and optimizers for training. During training, the model iteratively adjusts its internal parameters using backpropagation to minimize prediction errors and optimize performance for accurate tumor detection.

CHAPTER 5
IMPLEMENTATION AND RESULTS

5.1 IMPLEMENTATION

1. Data Collection: Doctor reports containing relevant information about brain tumor patients are gathered and compiled into a dataset.

2. Data Preprocessing: The dataset undergoes preprocessing steps to handle missing values, scale features, and split into training and testing sets. This ensures the data is clean, consistent, and ready for model training.

3. Model Development: A neural network model is constructed using deep learning frameworks like TensorFlow or Keras. The architecture of the model is designed, specifying the number of layers, neurons, and activation functions.

4. Model Training: The constructed neural network model is trained using the training dataset. During training, the model learns from the input data and adjusts its internal parameters to minimize prediction errors.

5. Model Evaluation: The trained model is evaluated using the testing dataset to assess its performance in predicting brain tumor presence. Metrics such as accuracy, precision, recall, and F1 score are computed to measure the model's effectiveness.

6. Deployment: Once the model is trained and evaluated satisfactorily, it can be deployed into production environments where it can make predictions on new, unseen data.

7. Continuous Improvement: The model's performance is monitored over time, and refinements or updates may be made to improve its accuracy and effectiveness in predicting brain tumors.

5.2 OUTPUT SCREENSHOTS

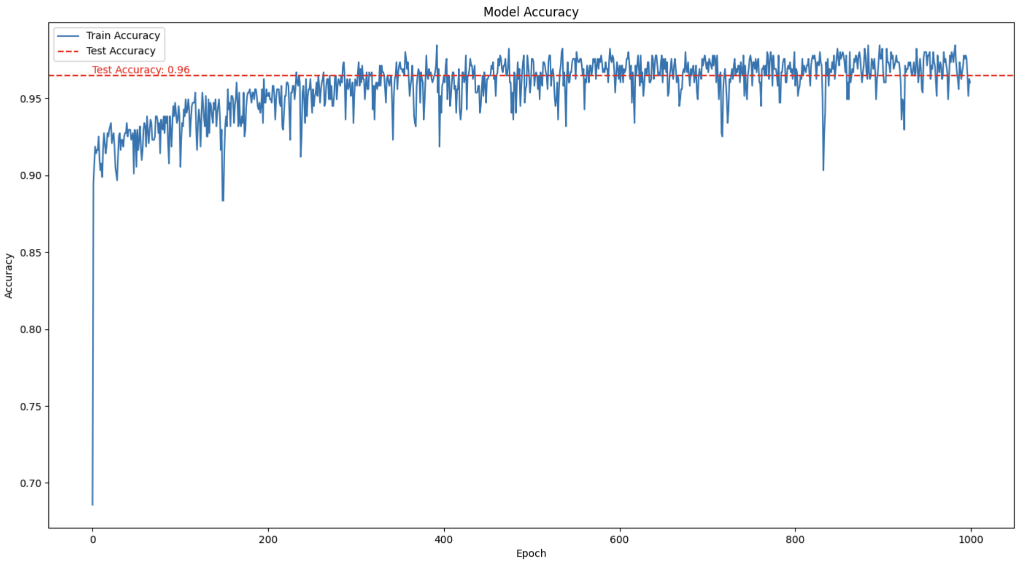

Fig 5.2.1 Accuracy Curve

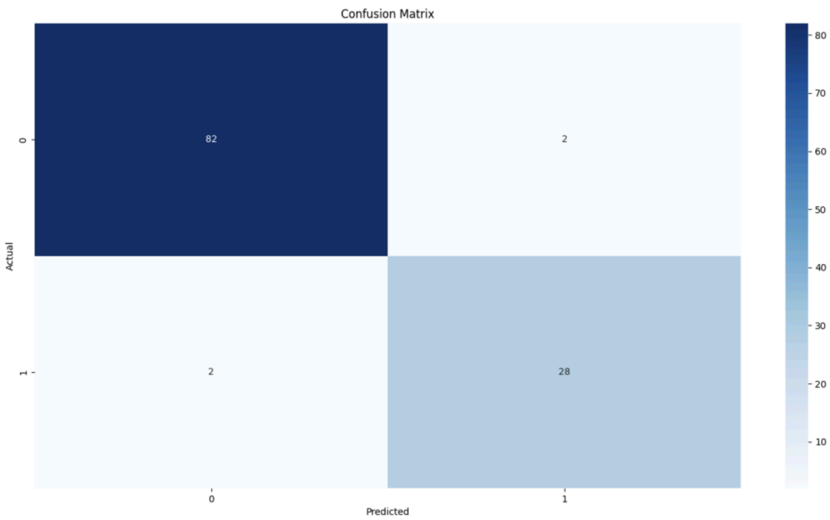

Fig 5.2.2 Confusion Matrix

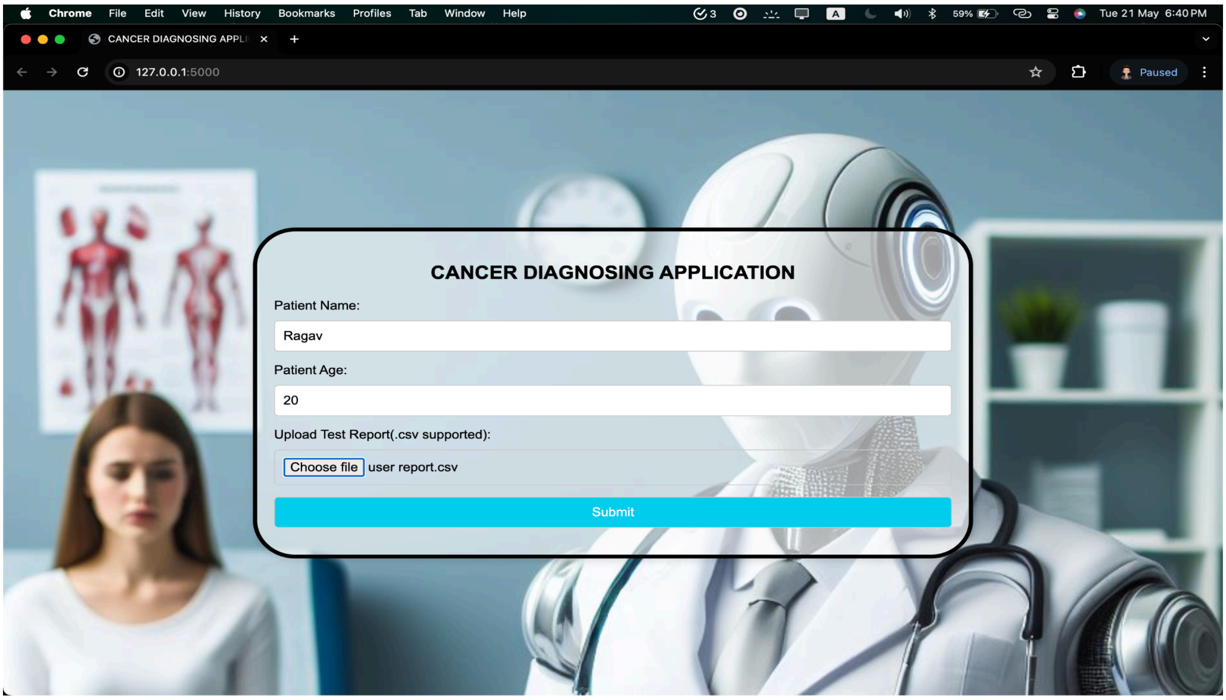

Fig 5.2.3 Landing Page

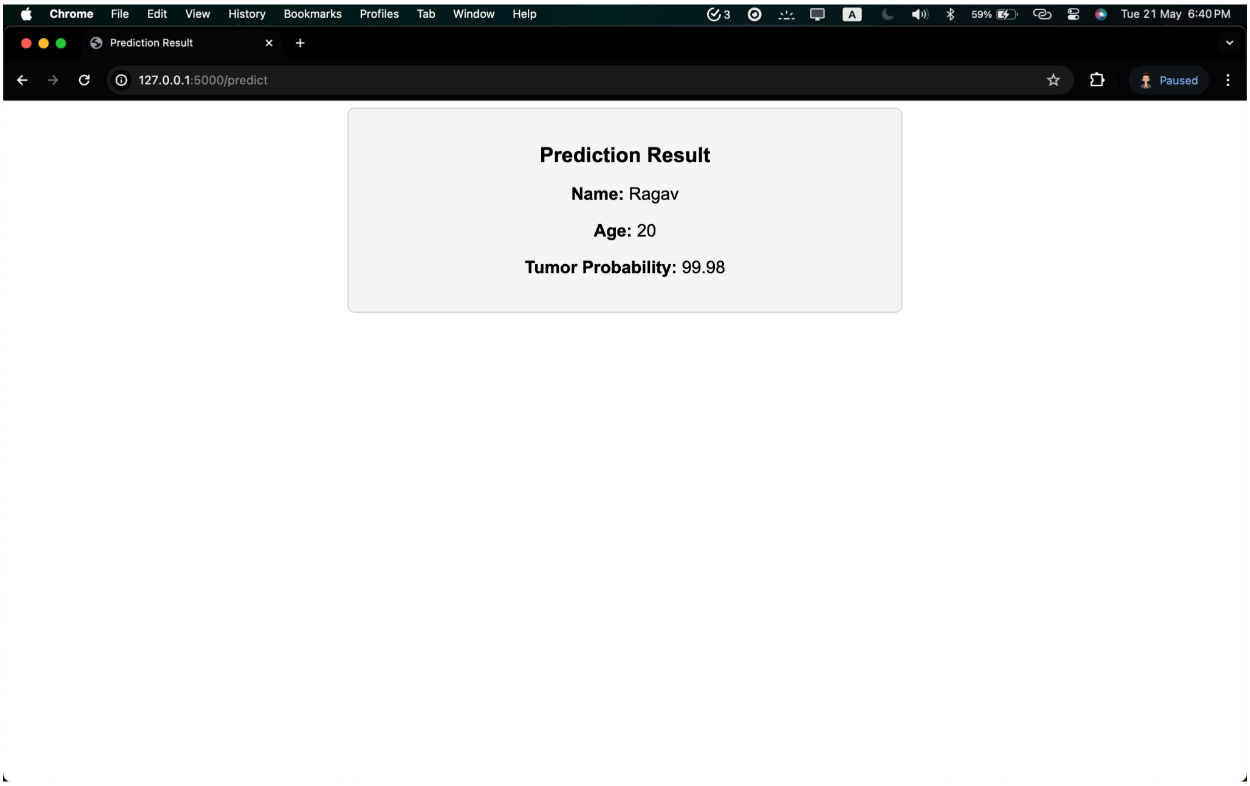

Fig 5.2.4 Result Page

CHAPTER 6
CONCLUSION AND FUTURE ENHANCEMENTS

6.1 CONCLUSION

In conclusion, the implementation of an artificial neural network (ANN) for the early detection of brain tumors based on doctor reports demonstrates a significant advancement in medical diagnostic tools. By leveraging features such as radius mean, symmetry mean, concavity, and fractal dimensions, the developed model effectively predicts the presence of brain tumors with high accuracy. This approach not only enhances the early detection capabilities but also supports medical professionals in making timely and informed decisions, ultimately improving patient outcomes. The project highlights the importance of integrating machine learning methodologies into clinical practice, offering a robust solution to the challenges posed by traditional diagnostic methods. Future work will focus on refining the model, expanding the dataset, and incorporating additional features to further increase the accuracy and reliability of brain tumor predictions. This project underscores the potential of AI-driven solutions in revolutionizing healthcare and paving the way for more efficient and precise medical interventions.

6.2 FUTURE ENHANCEMENTS

Future enhancements for this project involve several key areas of development. First, expanding the dataset to include a larger and more diverse set of patient records will improve the model's ability to generalize across various cases, enhancing its robustness and accuracy. Incorporating additional features, such as genetic markers, detailed imaging data, and comprehensive patient history, will provide a richer input to the model, potentially increasing its predictive power. Optimizing the model by experimenting with different neural network architectures, like convolutional neural networks (CNNs) or recurrent neural networks (RNNs), and applying advanced optimization techniques will further enhance performance.

REFERENCES

[1] D C Febrianto, I Soesanti, H A Nugroho Convolutional Neural Network for Brain Tumor Detection, May 2020.
[2] M. Rmus, T.-F. Pan, L. Xia, and A. G. E. Collins, “Artificial neural networks for model identification and parameter estimation in computational cognitive models,” May 2024.
[3] K. K. Singh, M. Elhoseny, A. Singh, and A. A. Elngar, Machine Learning and the Internet of Medical Things in Healthcare. Academic Press, 2021.
[4] F. Özyurt, E. Sert, E. Avci, and E. Dogantekin, Brain Tumor Detection Based on Convolutional Neural Network with Neutrosophic Expert Maximum Fuzzy Sure Entropy. Infinite Study.
[5] K. Selvi T, A. Sumaiya Begum, P. Poonkuzhali, and R. Aarthi, “Brain tumor classification for MRI images using dual-discriminator conditional generative adversarial network,” Apr. 2024.
[6] R. Appiah et al., “Brain tumor detection using proper orthogonal decomposition integrated with deep learning networks,” June 2024.
[7] A. Lumini, G. F. Roberto, L. A. Neves, A. S. Martins, and M. Z. do Nascimento, “Percolation Images: Fractal Geometry Features for Brain Tumor Classification,”
[8] M. S. Ullah, M. A. Khan, N. A. Almujally, M. Alhaisoni, T. Akram, and M. Shabaz, “BrainNet: a fusion assisted novel optimal framework of residual blocks and stacked autoencoders for multimodal brain tumor classification,” Mar. 2024.
[9] V. Akoto-Adjepong, O. Appiah, P. K. Mensah, and P. Appiahene, “TTDCapsNet: Tri Texton-Dense Capsule Network for complex and medical image recognition,” PLoS One, Mar. 2024.
[10] C. Ozdemir and Y. Dogan, “Advancing brain tumor classification through MTAP model: an innovative approach in medical diagnostics,” Mar. 2024.
[11] A. Panesar, Machine Learning and AI for Healthcare: Big Data for Improved Health Outcomes. April 2019.
[12] Q.-M. Liao et al., “Multimodal neuroimaging network associated with executive function in adolescent major depressive disorder patients via cognition-guided magnetic resonance imaging fusion,” May 2024.
[13] J. C. Chang, X. Li, S. Xu, H.-R. Yao, J. Porcino, and C. C. Chow, “Gradient-flow adaptive importance sampling for Bayesian leave one out cross-validation for sigmoidal classification models,” Feb. 2024,
[14] J. E. Dowling, Neurons and Networks: An Introduction to Neuroscience. Belknap Press, 1992.
[15] Comparison of data preprocessing approaches for applying deep learning to human activity recognition in the context of industry 4.0-X Zheng, M Wang, J Ordieres-Meré - Sensors, 2018.