**ALZHEIMER DISEASE PREDICTION**

**Project created by:**

Sree Vignesh V

Tharun Ashwin B

Krithik Raja M

1. **INTRODUCTION:**
   1. **Project Overview:**

Alzheimer's disease is a progressive neurodegenerative disorder that affects millions of people worldwide. Early diagnosis and prediction of Alzheimer's disease are crucial for timely intervention and treatment. This project aims to develop a predictive model that can help identify individuals at risk of developing Alzheimer's disease using various data sources and machine learning techniques.

* 1. **Purpose:**

Develop a predictive model to identify individuals at risk of Alzheimer's disease. Utilize a diverse dataset, including medical, genetic, and demographic information. Evaluate the model's accuracy, sensitivity, specificity, and reliability.

Implement the model for early detection and intervention in clinical settings.

1. **LITERATURE SURVEY:**
   1. **Existing Problem:**

Alzheimer's disease is a complex and heterogeneous condition, with various genetic and environmental factors contributing to its development. Predictive models often struggle to account for this heterogeneity and provide accurate risk assessments. Access to comprehensive and diverse datasets, including genetic, clinical, and lifestyle information, can be limited. This restricts the ability to train robust predictive models that consider all relevant factors. Machine learning models can inherit biases from the data they are trained on, potentially leading to unfair predictions. Ensuring that models are unbiased and generalize well to diverse populations is a critical concern. Ensuring that predictive models are validated and reproducible across different research groups and clinical settings is a significant challenge in the field. Addressing these existing problems in Alzheimer's disease prediction is essential to advance our ability to identify at-risk individuals early, ultimately improving the quality of care and potential interventions for those affected by this debilitating condition.

* 1. **Problem Statement and Definition:**

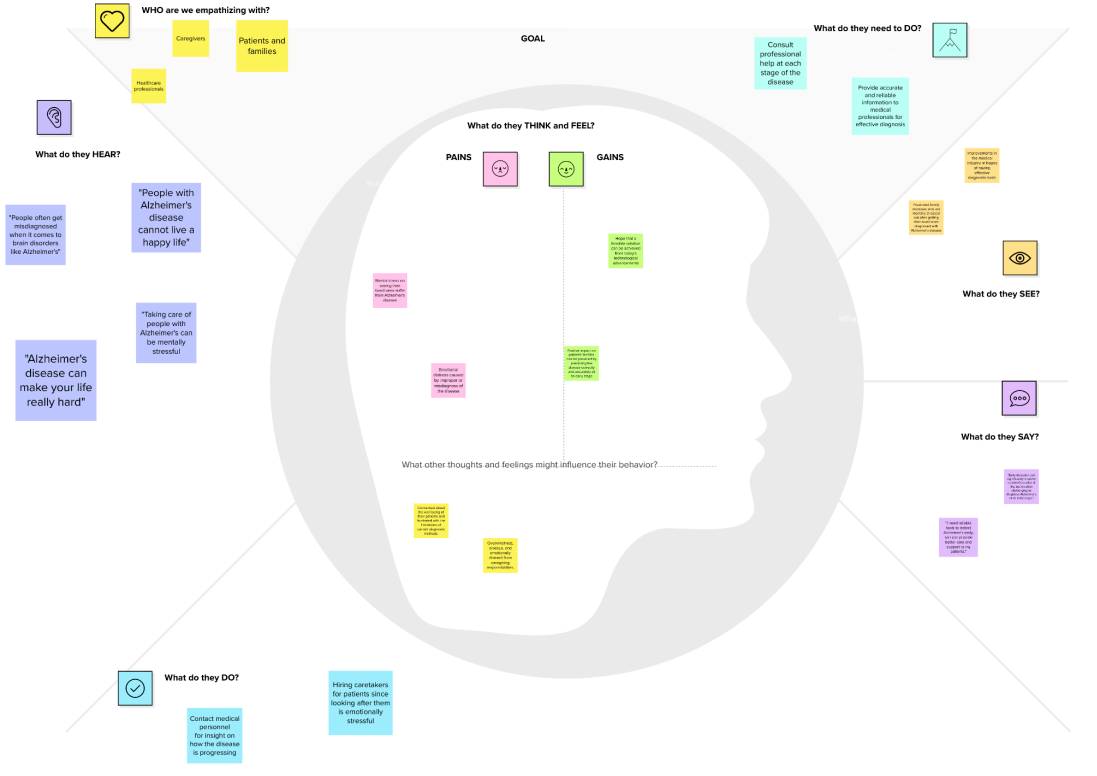
Alzheimer's disease is a prevalent and devastating neurodegenerative disorder characterized by progressive cognitive decline, memory loss, and behavioural changes. Early diagnosis and prediction of Alzheimer's disease are crucial for providing timely interventions, improving patient outcomes, and advancing research into potential therapies. The problem is to develop accurate and reliable predictive models that can identify individuals at risk of developing Alzheimer's disease, enabling early intervention and personalized care.

The problem of Alzheimer's disease prediction involves creating a predictive model that can assess an individual's risk of developing Alzheimer's disease based on various data sources, including medical records, genetic information, lifestyle factors, and demographic details. The primary objective is to develop a model that can provide a probability score or classification indicating an individual's likelihood of developing Alzheimer's disease within a specified time frame (e.g., 5 years).

1. **IDEATION & PROPOSED SOLUTION**
   1. **Empathy Map Canvas:**

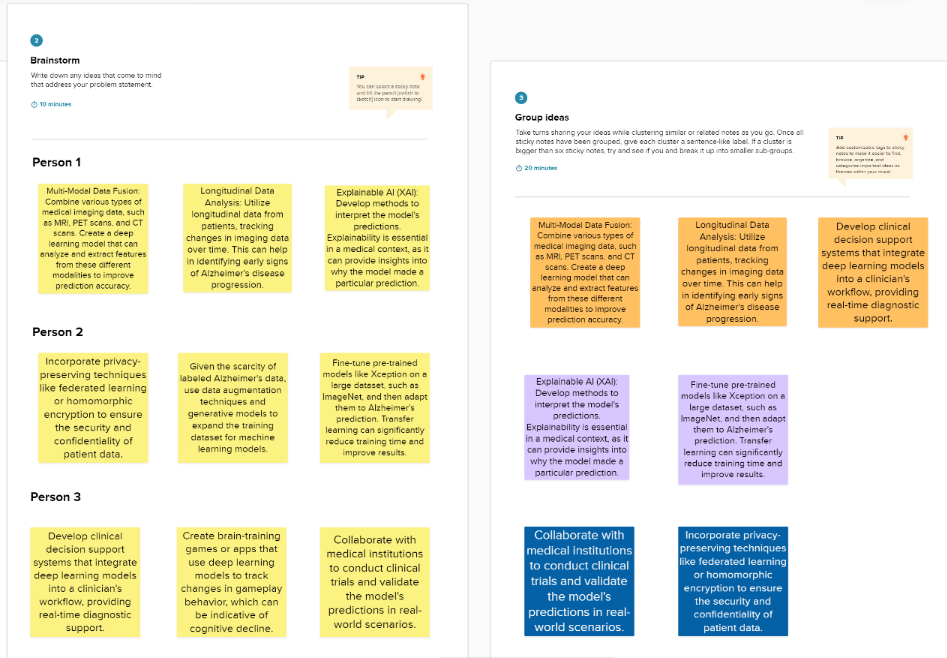
An empathy map is a simple, easy-to-digest visual that captures knowledge about a user’s behaviours and attitudes.

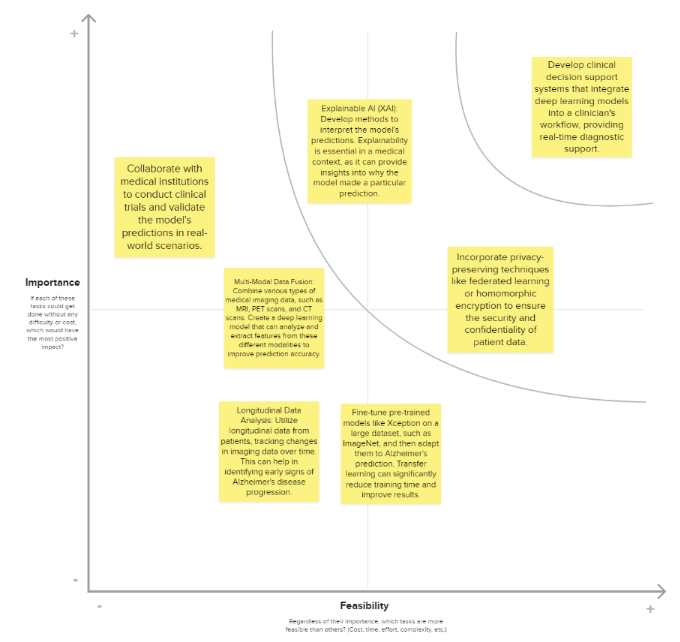
It is a useful tool to help teams better understand their users. Creating an effective solution requires understanding the true problem and the person who is experiencing it. The exercise of creating the map helps participants consider things from the user’s perspective along with his or her goals and challenges.



* 1. **Ideation & Brainstorming**

Brainstorming provides a free and open environment that encourages everyone within a team to participate in the creative thinking process that leads to problem solving. Prioritizing volume over value, out-of-the-box ideas are welcome and built upon, and all participants are encouraged to collaborate, helping each other develop a rich number of creative solutions.





**4.REQUIREMENT ANALYSIS**

* 1. **Functional Requirement:**

Functional requirements specify the features, capabilities, and behaviours that a system or application must possess to fulfil its intended purpose. In the context of Alzheimer's disease prediction, the functional requirements outline the specific functionalities and features that the predictive model and associated software or system must have.

These requirements are essential for developing a reliable and effective Alzheimer's disease prediction system. Here are some key functional requirements:

* **Data Collection:** The system should be able to collect and integrate diverse data sources, including medical records, genetic data, lifestyle information, and demographic details.
* **Data Preprocessing:** The system should clean and preprocess the data, handling missing values, outliers, and inconsistencies.
* **Model Development:** The system must use machine learning algorithms to develop a predictive model. It should support various algorithms, such as logistic regression, random forests, support vector machines, and deep learning techniques.
* **Model Evaluation:** The system should assess the model's performance using metrics such as accuracy, precision, recall, F1-score, and ROC-AUC. It must support cross-validation and hyperparameter tuning to optimize the model's performance.
  1. **Non- Functional Requirement:**

Non-functional requirements define the qualities or attributes that a system or application must possess to meet user expectations and adhere to industry standards. In the context of Alzheimer's disease prediction, non-functional requirements outline characteristics related to the performance, usability, security, and compliance of the predictive model and its associated system. Here are some key non-functional requirements:

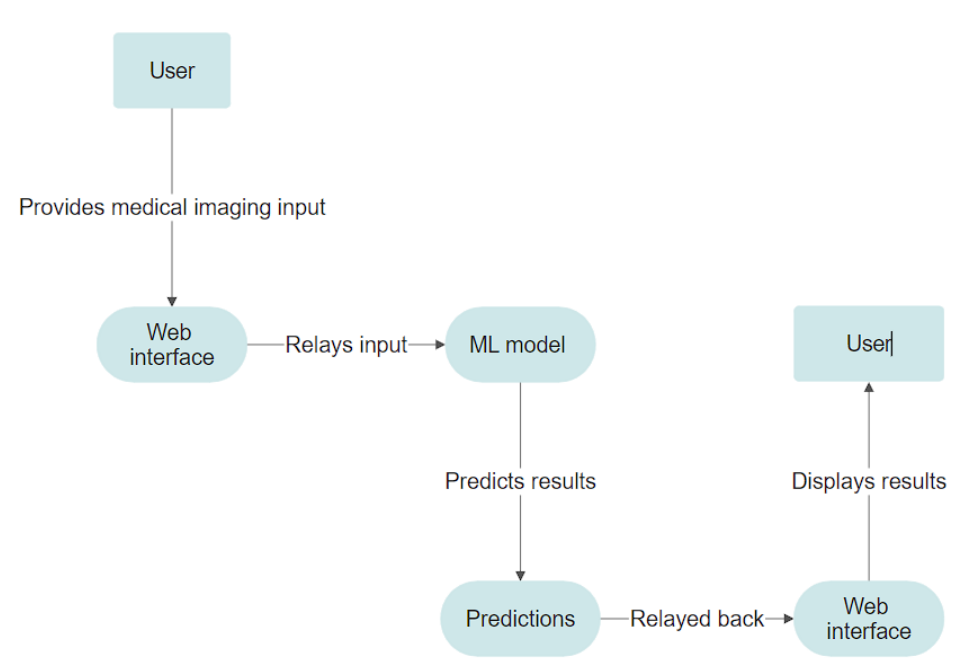
* **Performance:** Response Time: The system should provide quick responses to user queries and predictions, ensuring minimal delay in risk assessment. Scalability: The system should be able to handle a growing amount of data and users without significant performance degradation.
* **Usability:** The system's user interface should be intuitive, with clear navigation and visualizations that are easy for healthcare professionals to interpret. The system should be accessible to users with disabilities, adhering to accessibility standards.
* **Reliability:** The system should be available 24/7, especially in critical healthcare settings, with minimal downtime for maintenance. The system should continue functioning even in the presence of hardware or software failures.
* **Scalability:** The system should be designed to scale as the data and user load increase, accommodating potential future growth in healthcare data and users.

1. **PROJECT DESIGN**
   1. **Data Flow Diagrams & User Stories:**

A Data Flow Diagram (DFD) is a traditional visual representation of the information flows within a system. A neat and clear DFD can depict the right

amount of the system requirement graphically. It shows how data enters and leaves the system, what changes the information, and where data is

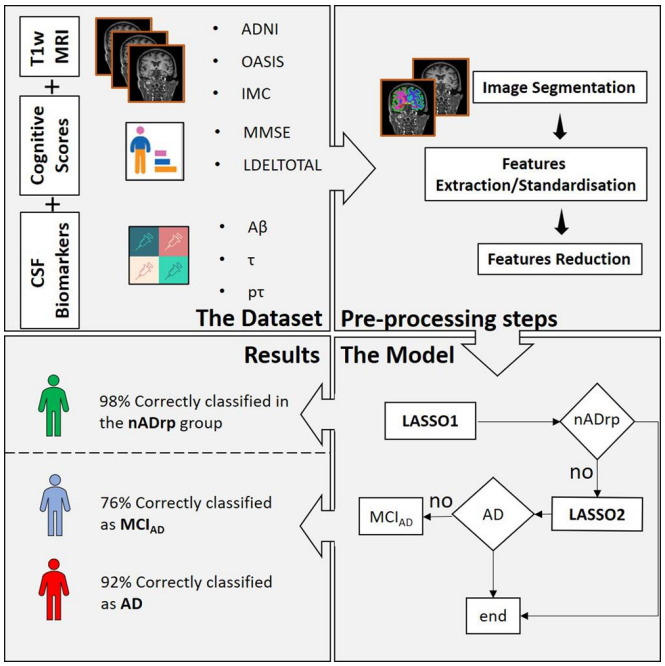
stored.



* 1. **Solution Architecture:**

The development of a machine learning solution for healthcare, especially for a complex condition like Alzheimer's disease, requires a multidisciplinary team, including data scientists, medical experts, data engineers, and regulatory specialists. Here's an outline of how you can architect a machine learning solution for Alzheimer's disease prediction:

* **Data Collection and Preprocessing:** Gather relevant data, which may include medical records, neuroimaging data, genetic information, cognitive test results, and more. Clean and preprocess the data to handle missing values, outliers, and noise.
* **Data Split:** Split the data into training, validation, and testing sets to evaluate and fine-tune the model.
* **Feature Engineering:** Identify and engineer relevant features that may be predictive of Alzheimer's disease. Feature selection and dimensionality reduction techniques might be necessary.
* **Model Architecture:** If you're using neural networks, design the architecture of your deep learning model. This may involve defining the number of layers, units, activation functions, and other hyperparameters.
* **Training:** Train the machine learning model using the training data. Experiment with different hyperparameters and optimization techniques to improve model performance.
* **Validation:** Use the validation set to fine-tune the model, avoid overfitting, and optimize hyperparameters.



1. **PROJECT PLANNING & SCHEDULING**
   1. **Technical Architecture**

**Table-1 : Components & Technologies:**

|  |  |  |  |
| --- | --- | --- | --- |
| **S.No** | **Component** | **Description** | **Technology** |
| 1. | User Interface | Website | HTML, CSS |
| 2. | Application Logic-1 | Logic applied in application | Python |
| 3. | File Storage | File storage requirements | Local Filesystem |
| 4. | Machine Learning Model | Image Classification | CNN, Xception |
| 5. | Infrastructure (Server / Cloud) | Application Deployment on Local System Local Server Configuration. | Local System. |

**Table-2: Application Characteristics:**

|  |  |  |  |
| --- | --- | --- | --- |
| **S.No** | **Characteristics** | **Description** | **Technology** |
| 1. | Open-Source Frameworks | List the open-source frameworks used | Flask, Tensor FLow |
| 2. | Security Implementations | List all the security / access controls implemented, use of firewalls etc. | NA |
| 3. | Scalable Architecture | Justify the scalability of architecture (3 – tier, Micro-services) | NA |
| 4. | Availability | Justify the availability of application (e.g. use of load balancers, distributed servers etc.) | NA |
| 5. | Performance | Design consideration for the performance of the application (number of requests per sec, use of  Cache, use of CDN’s) etc. | - |

* 1. **Sprint Planning & Estimation**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Sprint** | **Functional Requirement (Epic)** | **User Story Number** | **User Story / Task** | **Story Points** | **Priority** | **Team Members** |
| Sprint-1 | Data augmentation | USN-1 | The data has to be augmented in a suitable format for model building | 2 | High | Tharun Ashwin B |
| Sprint-1 | Data Replication | USN-2 | The data has to be replicated using SMOTE for effective prediction | 1 | High | Sree Vignesh V |
| Sprint-1 | Model Creation | USN-3 | A CNN model has to be defined and has to be tested for effectiveness in the required application | 2 | Medium | Sree Vignesh V |
| Sprint-2 | Model Training | USN-4 | The model has to be trained using a suitable split of training, testing and validation data | 2 | High | Krithik Raja M |
| Sprint-2 | Performance evaluation | USN-5 | The model’s performance has to be evaluated using performance metrics as well as a random value | 1 | Meduim | Tharun Ashwin B |
| Sprint 3 | Web application creation | USN-6 | Suitable html files has to be created to integrate the model into a web application using Flask | 2 | High | Krithik Raja M |
| Sprint 3 | Model integration | USN-7 | Suitable python script has to be written to integrate the model with the web application using flask | 2 | High | Sree Vignesh V |

* 1. **Sprint Delivery Schedule**

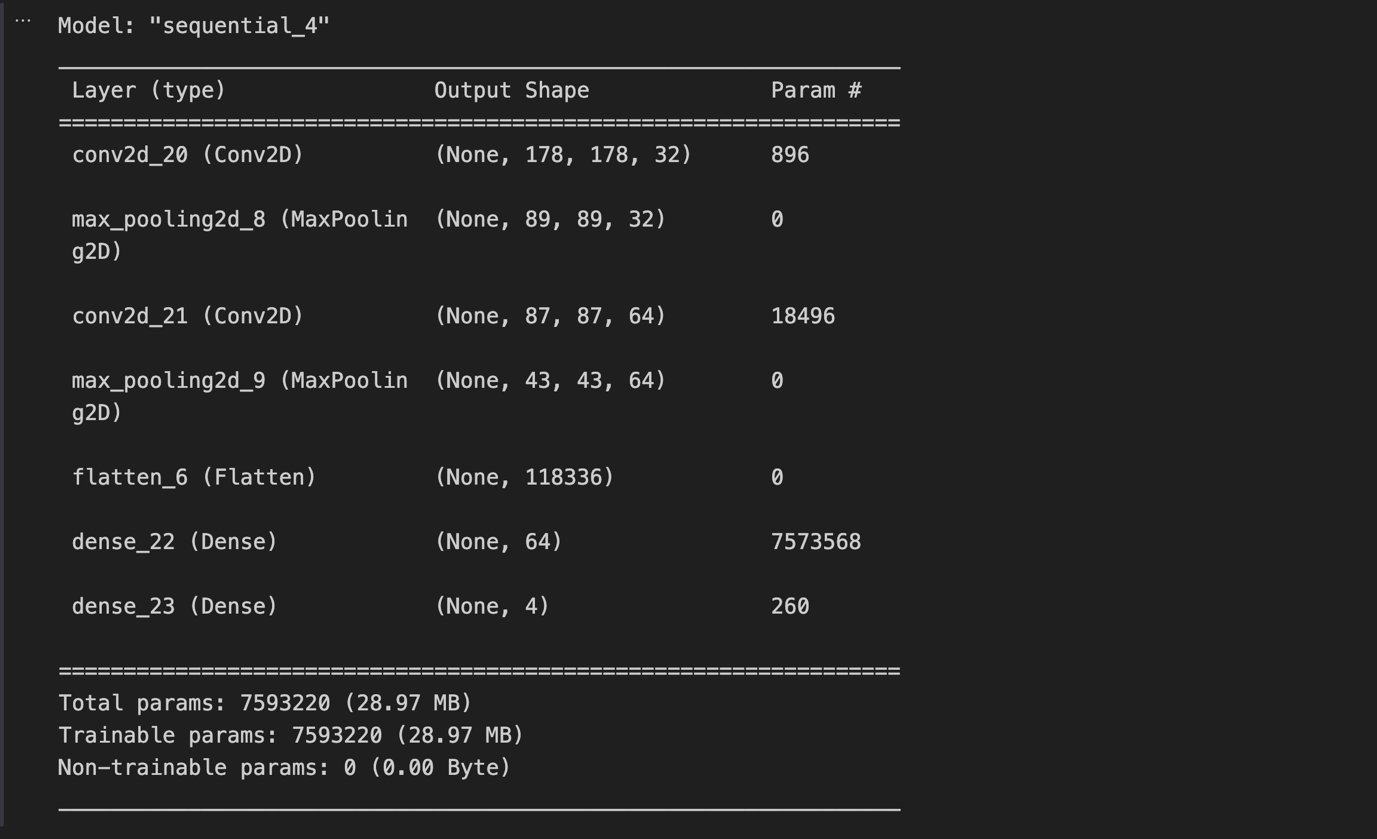
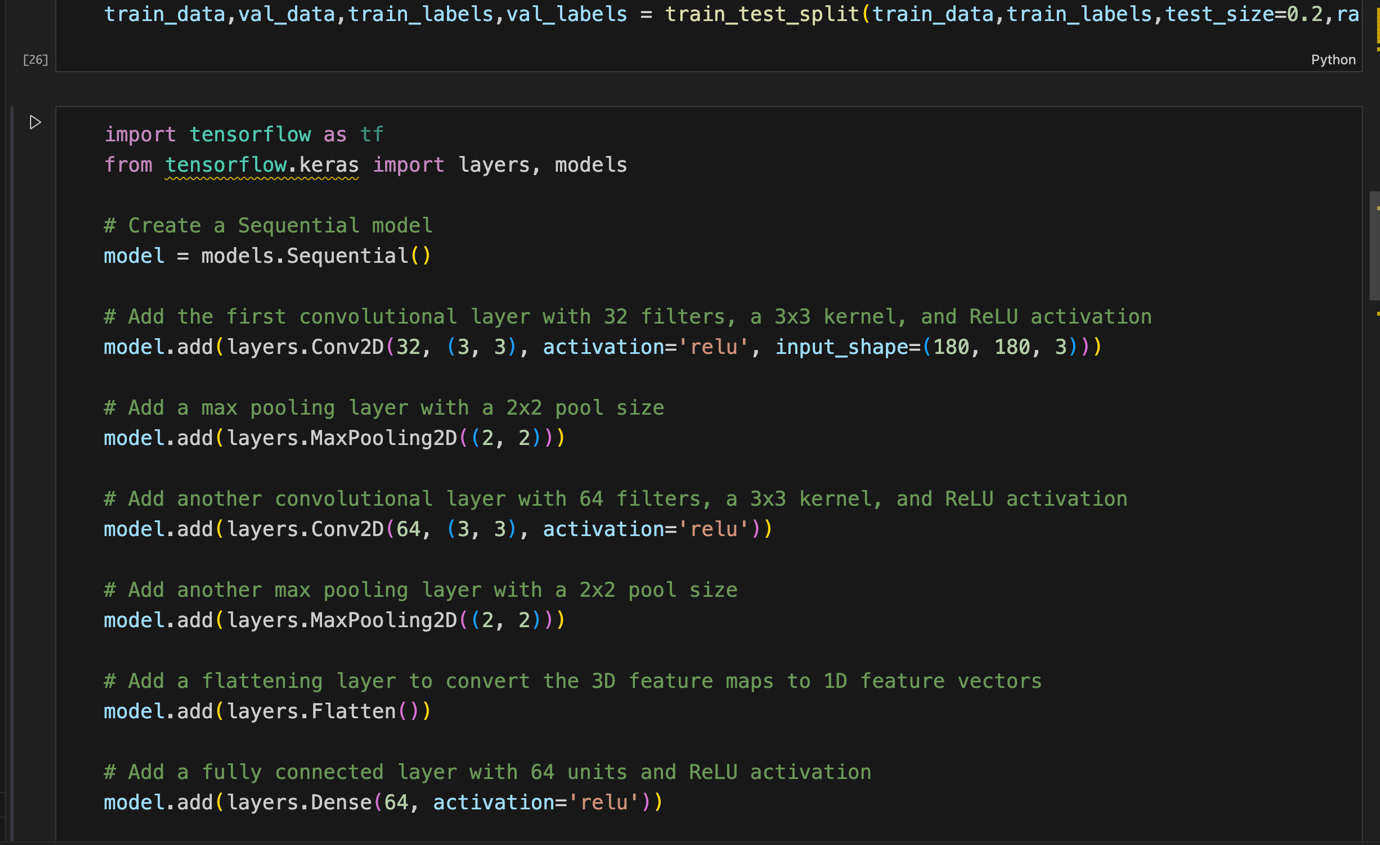
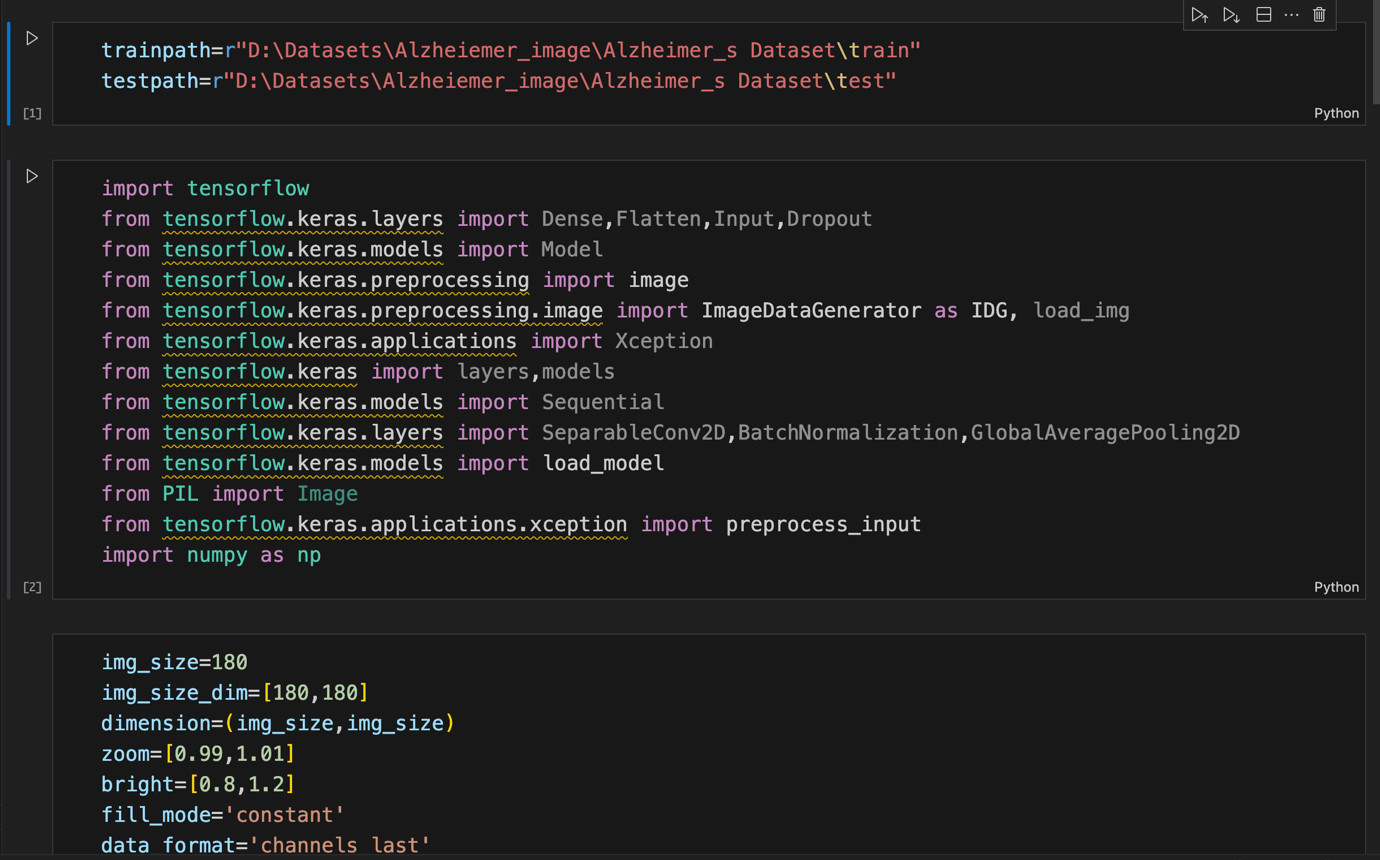
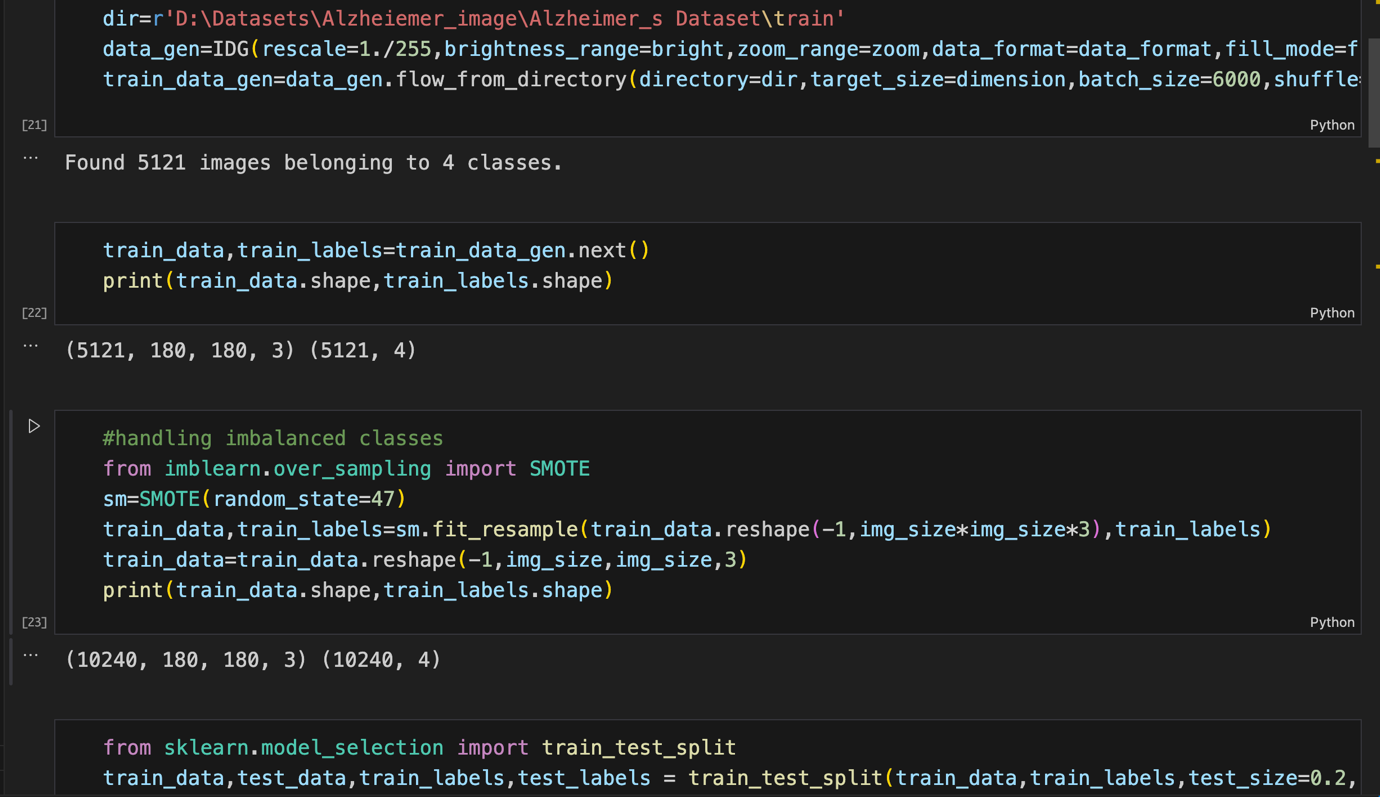
|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Sprint** | **Total Story Points** | **Duration** | **Sprint Start Date** | **Sprint End Date (Planned)** | **Story Points**  **Completed (as on Planned End Date)** | **Sprint Release Date (Actual)** |
| Sprint-1 | 6 | 6 Days | 24 Oct 2023 | 29 Oct 2023 | 6 | 30 Oct 2023 |
| Sprint-2 | 9 | 6 Days | 31 Oct 2023 | 05 Nov 2023 | 9 | 05 Nov 2023 |
| Sprint-3 | 13 | 6 Days | 07 Nov 2023 | 09 Nov 2023 | 13 | 11 Nov 2023 |

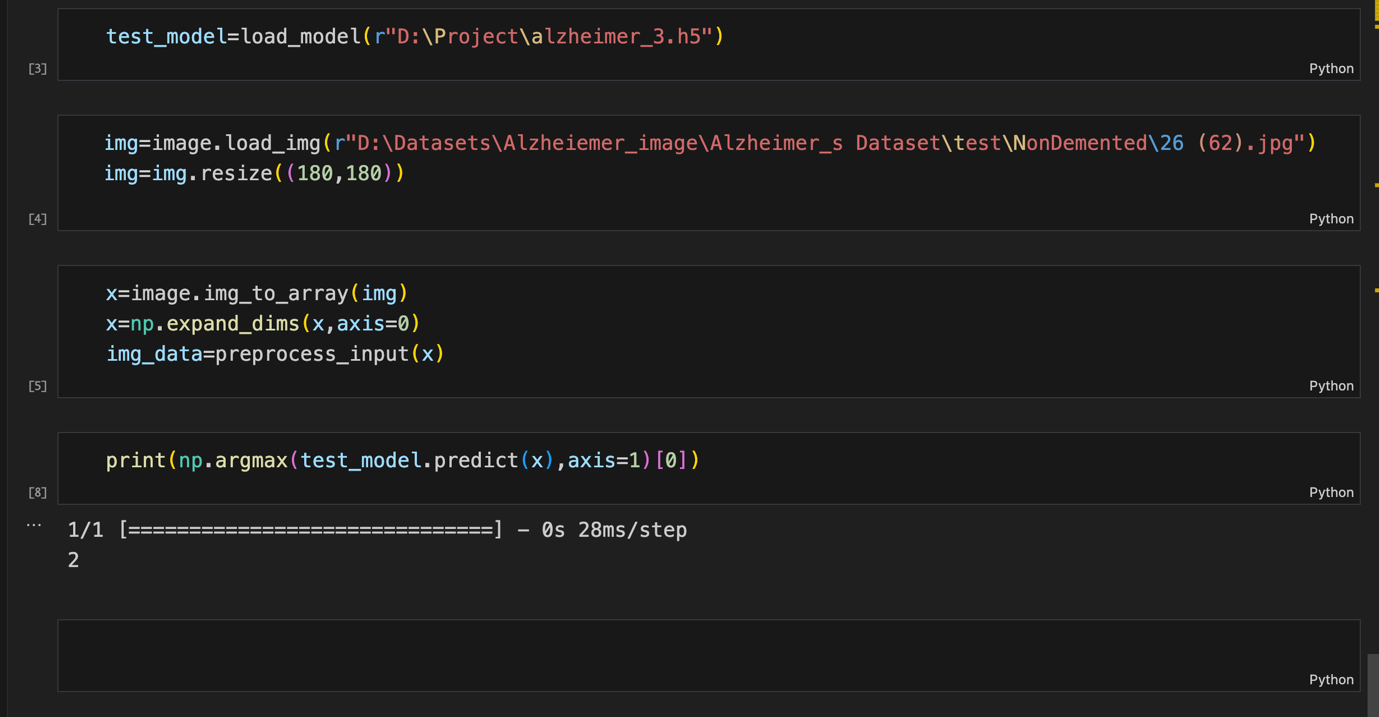
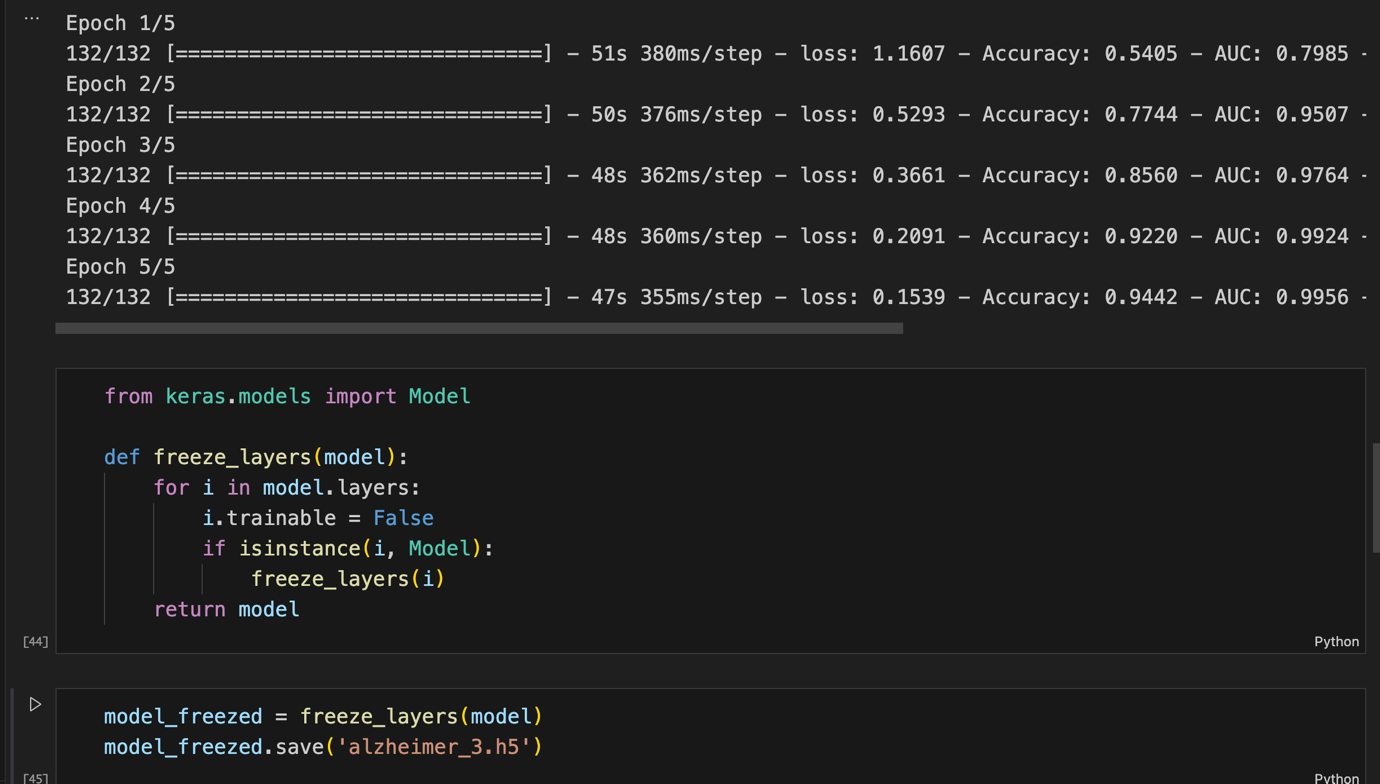
1. **CODING & SOLUTIONING**

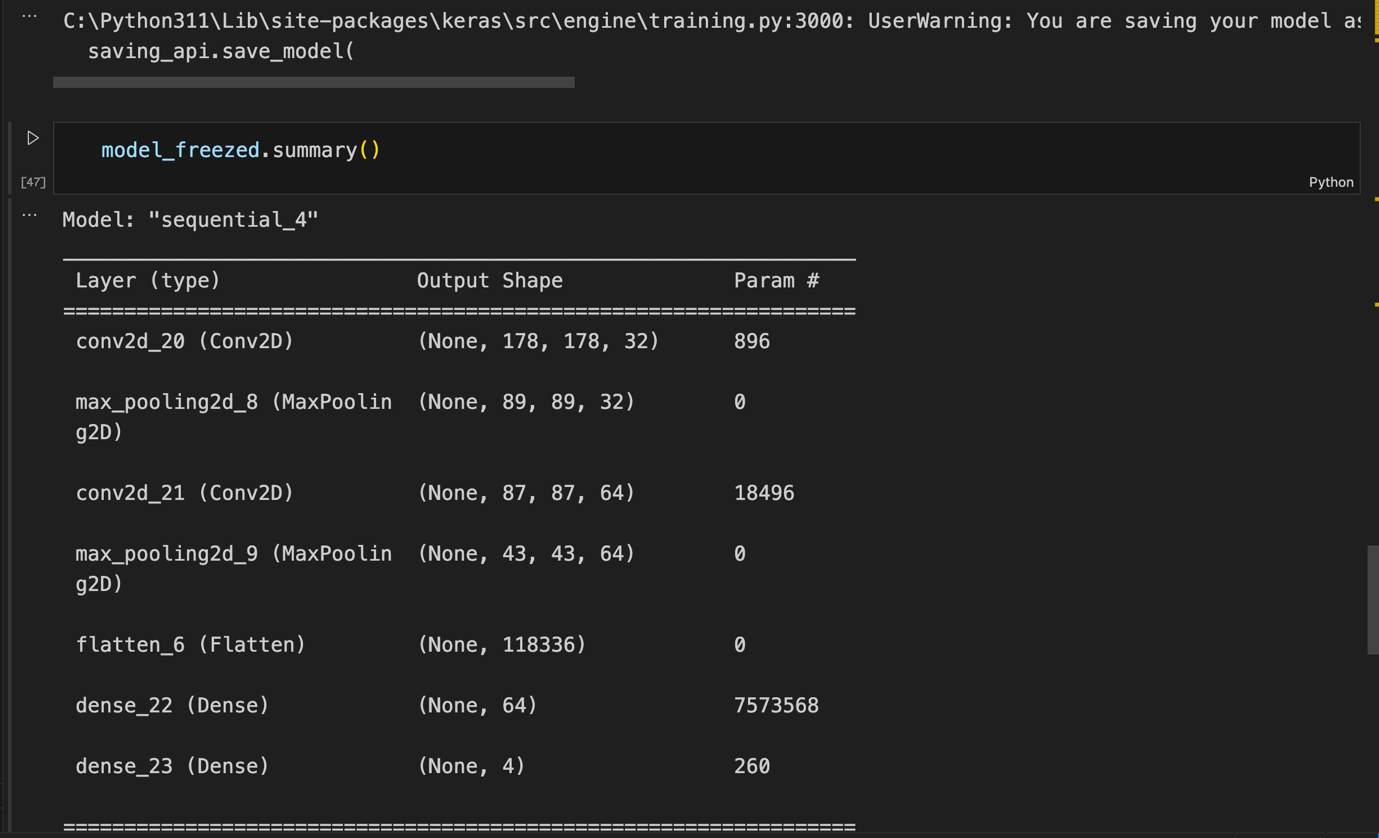
Creating a website with a CNN model for Alzheimer's disease prediction involves several steps, including data preparation, model development, and web application implementation.

Data Preparation:  
  
Gather a dataset of MRI scans labeled with Alzheimer's disease types.  
Preprocess the data, including resizing images and normalizing pixel values.

Model Development:  
  
Develop a Convolutional Neural Network (CNN) model using TensorFlow/Keras.  
Train the model using the preprocessed MRI scans data.

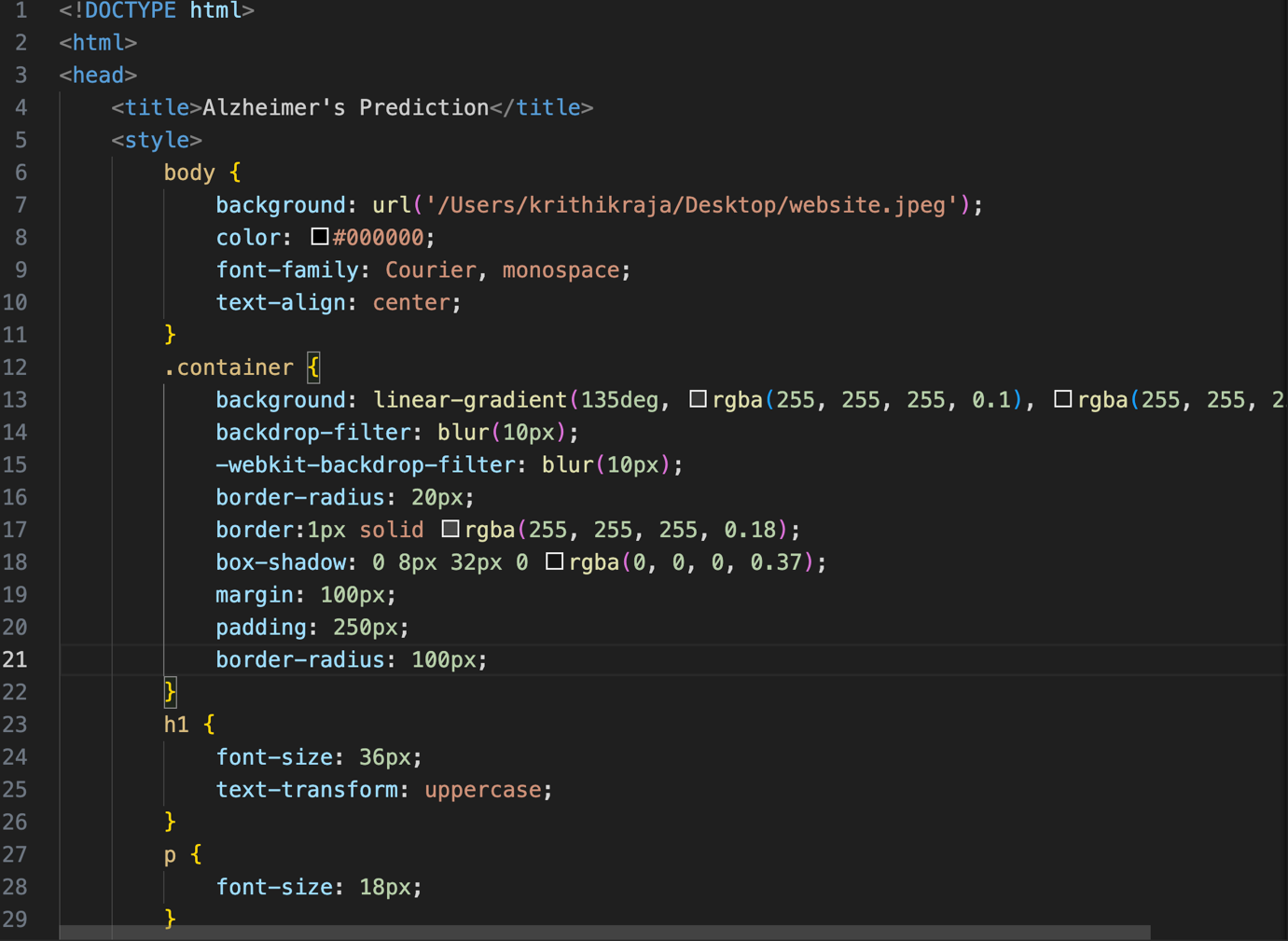


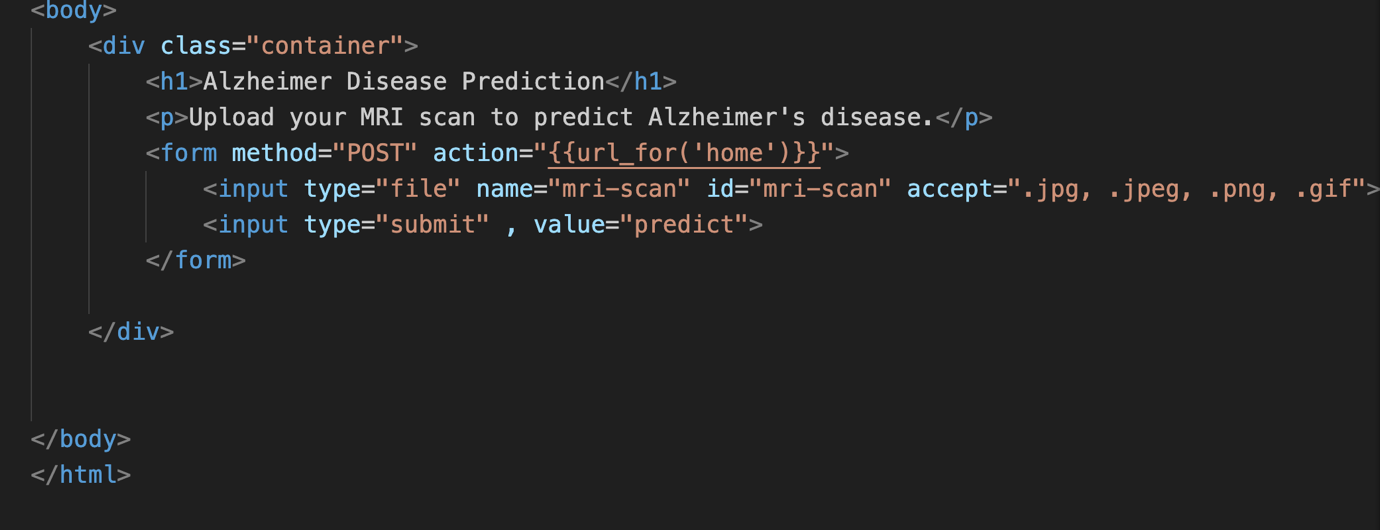




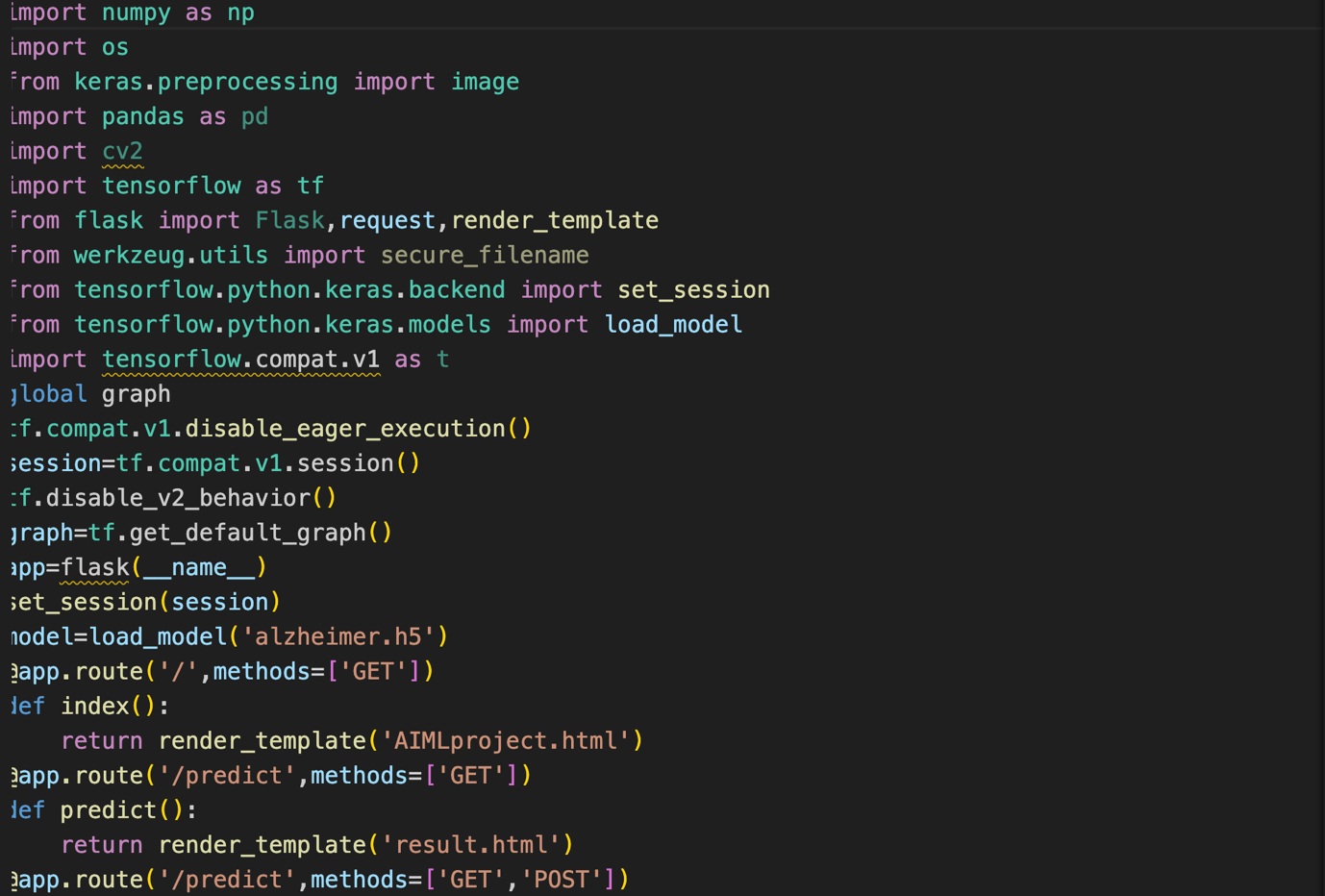
Web Application   
  
Create a Flask web application (using HTML) to deploy the trained model.  
python

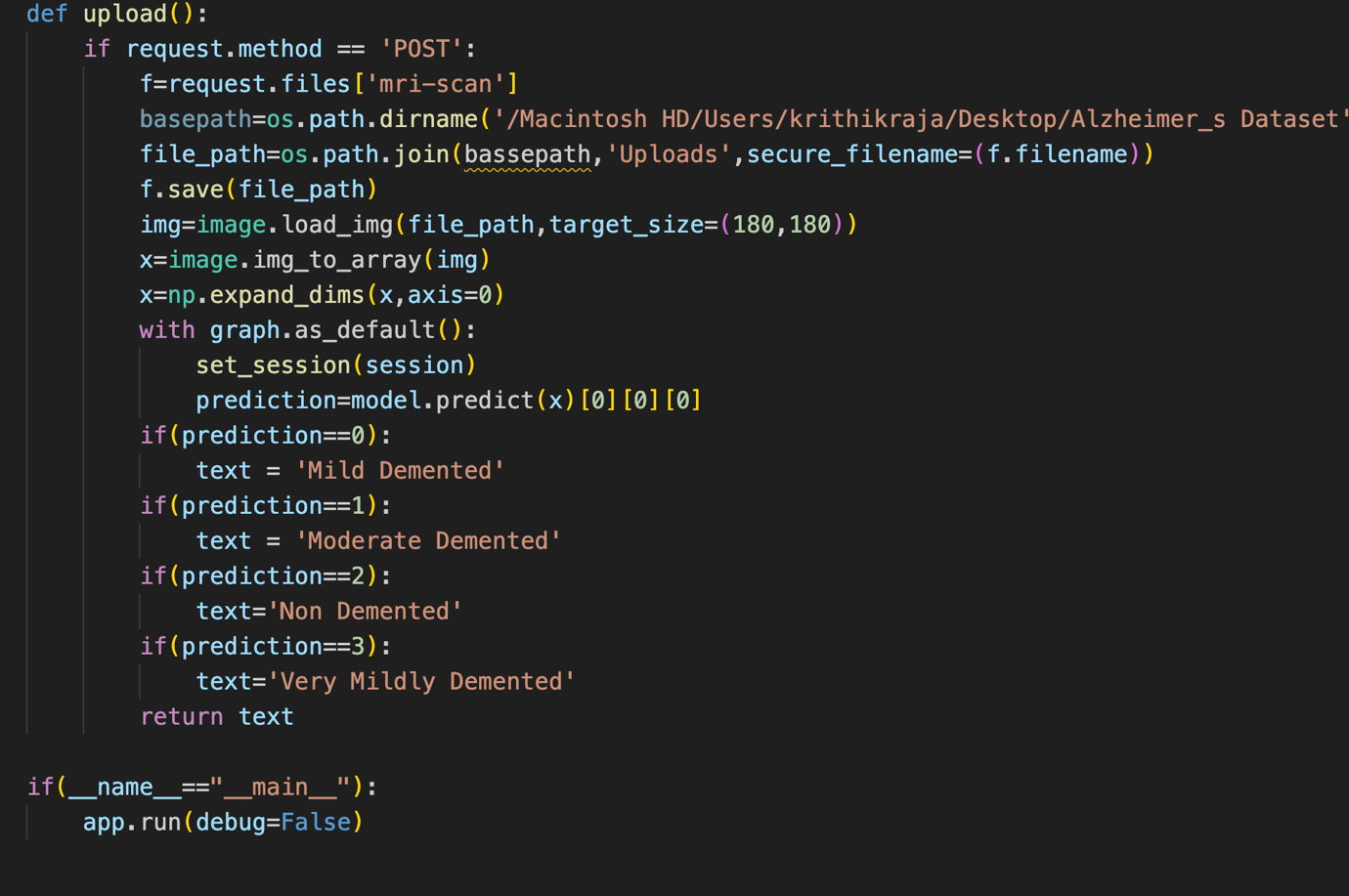
The HTML code:





The Code for the integration using Flask:





1. **PERFORMANCE TESTING**
   1. **Performance Metrics:**

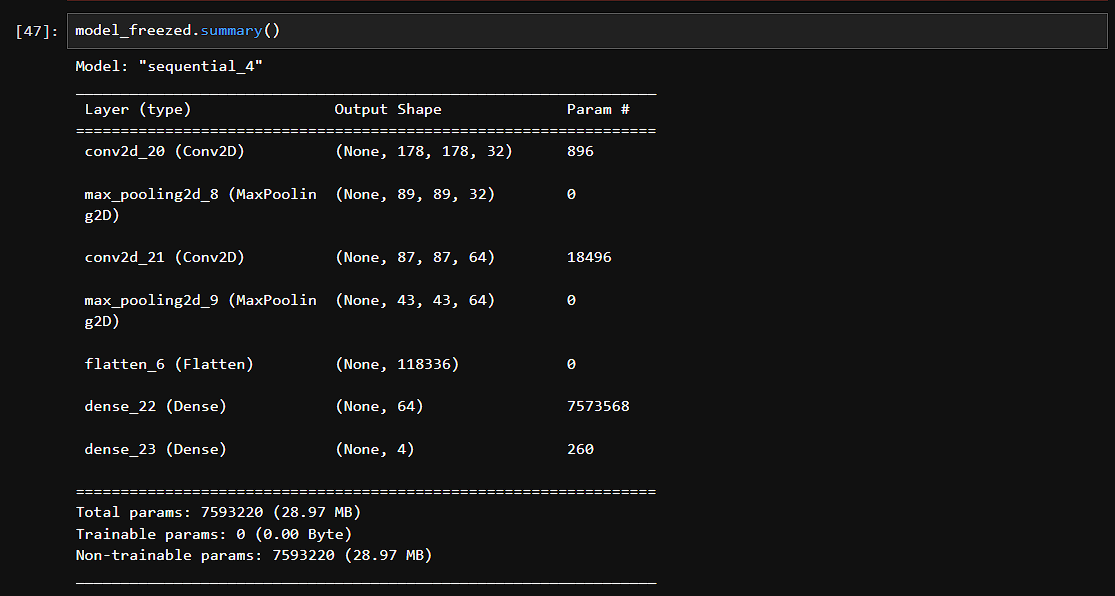
**Model Performance Testing:**

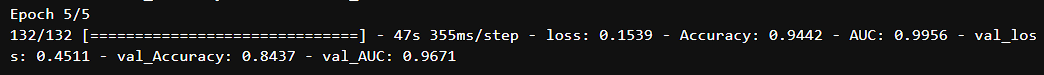
Project team shall fill the following information in model performance testing template

|  |  |  |  |
| --- | --- | --- | --- |
| **S.No.** | **Parameter** | **Values** | **Screenshot** |
|  | Model Summary | **-** |  |
|  | Accuracy | Training Accuracy - 0.9442  Validation Accuracy -0.8437 |  |

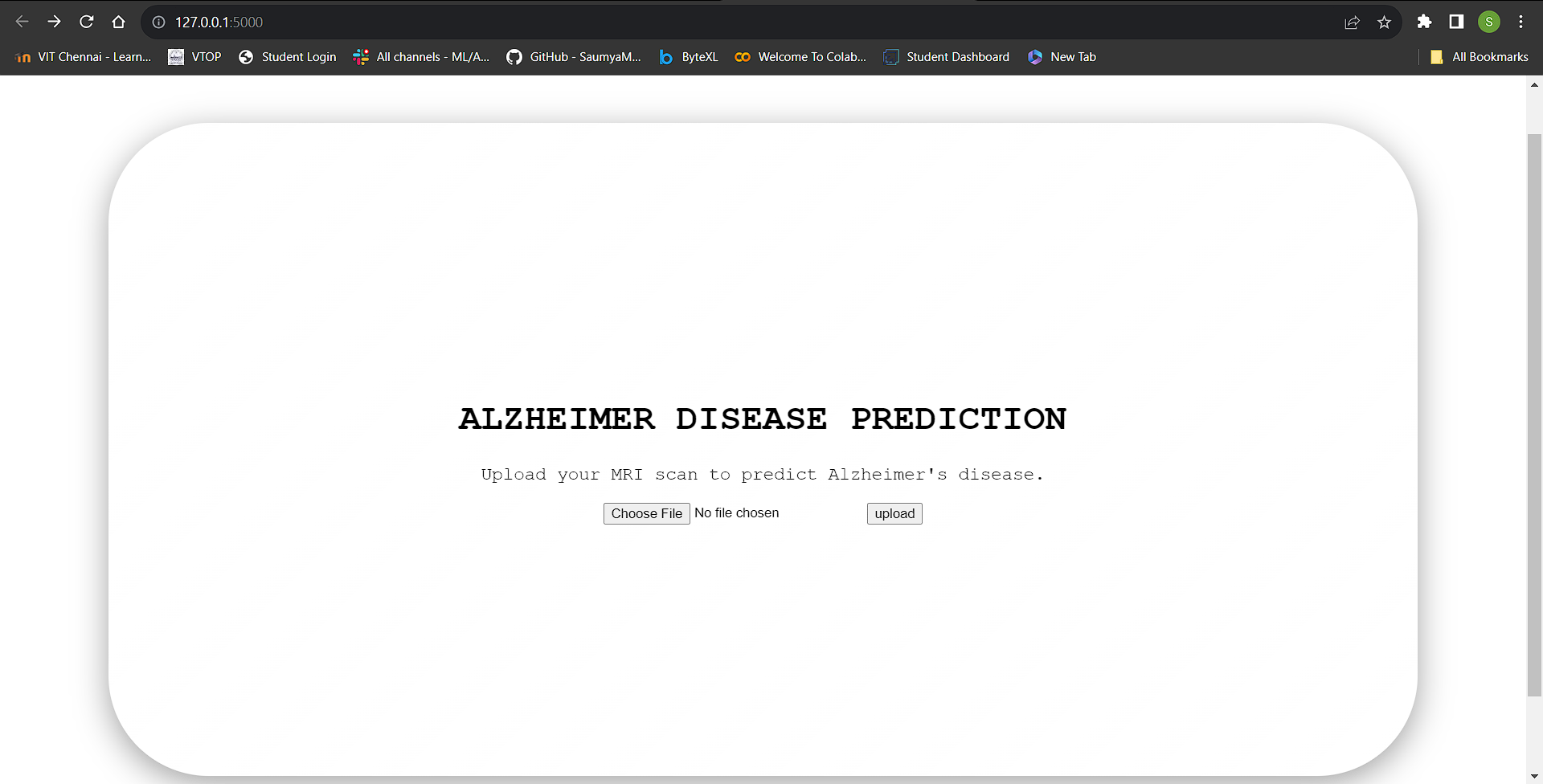
1. **Result:**
   1. **Output Screenshots:**

**Model Output and accuracy:**

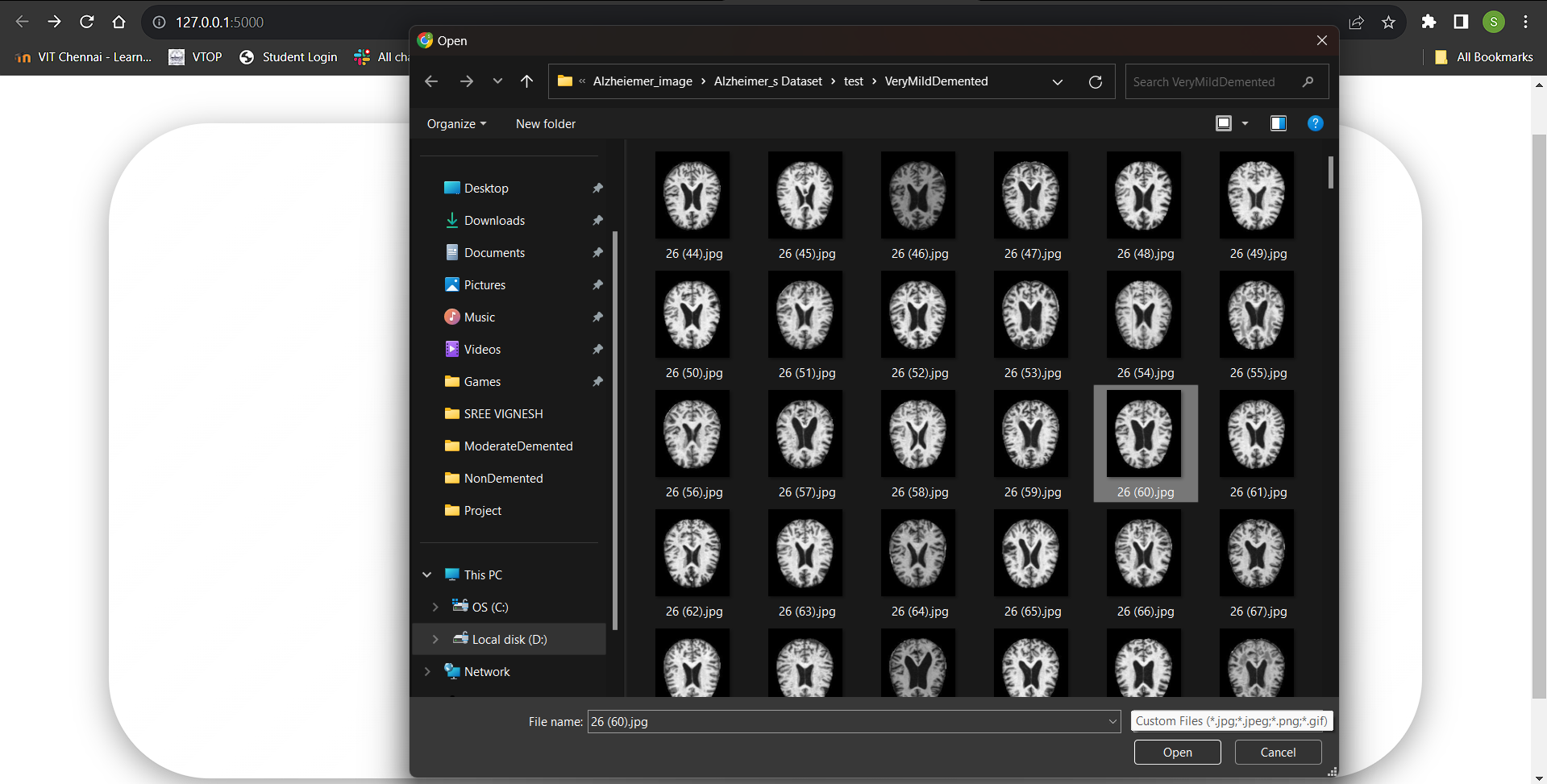




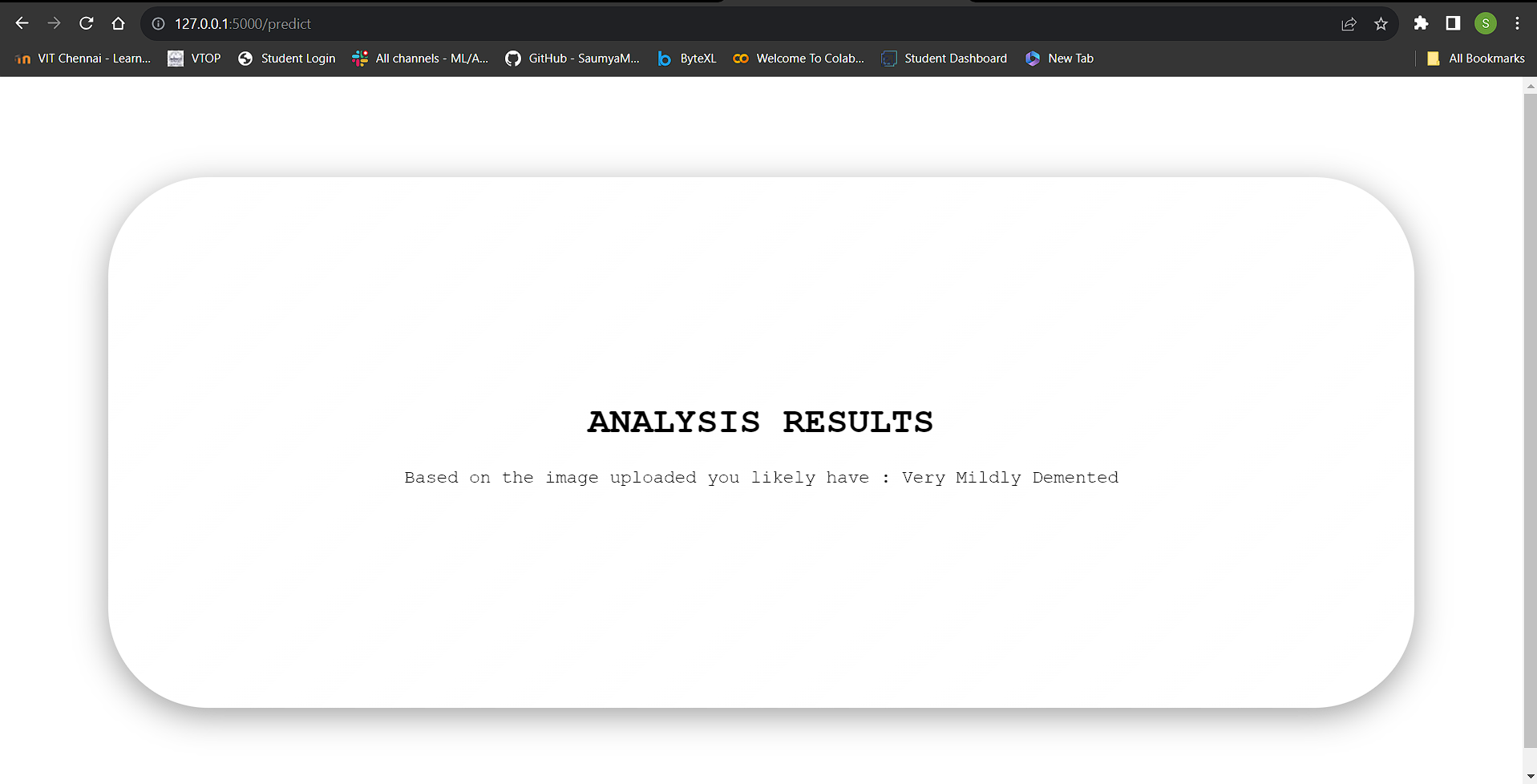
**Final website output :**



This is the home page where there is an option to upload an image to check the severity of Alzheimer’s



This is the image upload page



This is the result page where the result is displayed

1. **ADVANTAGES & DISADVANTAGES:**

Alzheimer's disease prediction models can identify individuals at risk of developing the condition before clinical symptoms become apparent. This early diagnosis enables timely intervention and treatment, potentially improving patient outcomes. Alzheimer's prediction research can contribute to a better understanding of the disease's underlying causes, risk factors, and progression, facilitating advancements in treatment and prevention strategies. Healthcare resources can be allocated more efficiently to support individuals at higher risk, ensuring that they receive the necessary care and support while reducing unnecessary interventions for lower-risk individuals. Predictive models can increase awareness among patients and their caregivers, encouraging proactive measures and lifestyle changes to mitigate risk.

Predictive models for Alzheimer's disease are not infallible and can provide false positives and false negatives. These inaccuracies can lead to unnecessary stress for some individuals and missed interventions for others. The use of predictive models must address ethical concerns related to bias, fairness, and transparency in decision-making, as well as the potential misuse of the information. Receiving a high-risk prediction for Alzheimer's disease can have a significant psychological impact on individuals and their families, potentially causing anxiety and stress. Predictive models may inadvertently contribute to the stigmatization of individuals at higher risk, affecting their social and professional lives.

1. **CONCLUSION:**

In conclusion, Alzheimer's disease prediction is a rapidly advancing field with the potential to make a profound impact on healthcare and the lives of individuals at risk. Early diagnosis and risk assessment are essential for timely intervention, personalized care, and research advancement. While predictive models offer several advantages, such as early detection and tailored treatment, they also come with challenges related to accuracy, privacy, ethics, and data quality. However, the future holds promise as technology, research, and ethical considerations evolve. Integration of multi-modal data, AI advancements, longitudinal analysis, and ethical AI will play significant roles in shaping the future of Alzheimer's disease prediction. The ultimate goal is to improve the quality of life for those affected by Alzheimer's and work toward the prevention and treatment of this devastating condition.

1. **FUTURE SCOPE:**

The field of Alzheimer's disease prediction is continuously evolving, driven by advances in technology, research, and healthcare. The future holds promising opportunities for further improving the early detection and management of Alzheimer's disease. Future research will focus on integrating diverse data sources, including genomics, proteomics, neuroimaging, and lifestyle factors. Combining these data modalities can lead to more accurate predictive models. Ongoing developments in artificial intelligence and machine learning will result in more sophisticated predictive models. Deep learning techniques and novel algorithms will enhance prediction accuracy. The integration of wearable devices and the Internet of Things (IoT) will enable real-time data collection and continuous monitoring of individuals at risk, offering more timely insights.

Future predictive models will prioritize ethical considerations, emphasizing fairness, transparency, and bias mitigation. Efforts will be made to ensure that predictions do not perpetuate disparities. Predictive models will need to gain acceptance and incorporation into clinical practice. The development of guidelines and standards for using predictive models in healthcare will be a critical focus. Governments and healthcare organizations may launch public health initiatives to raise awareness about Alzheimer's risk and the importance of early prediction, promoting preventive measures. As research on Alzheimer's disease continues, there is hope for the development of disease-modifying treatments. Predictive models will become even more valuable if effective therapies become available. Enhancements in AI explainability techniques will make predictive models more interpretable for healthcare professionals, improving their adoption.