AI-DocHelper

GUIDE :

PROF. AEJAZ KHAN

GROUP NO :

43

GROUP MEMBERS :

VIKRANT SHAH

SURBHI SINGH

ANUJA KOTHAVALE

Introduction

- For many years, many people have died due to undetected diseases. Early detection of these diseases at the micro classification stage can be useful for providing proper treatment to the patients at the early stage and could have saved a lot of lives. A lot of research is being done to detect these diseases at the earliest.

Problem Statement

- AI-DocHelper, an artificial intelligence-based system.
- The model that gives the best accuracy and minimum loss.
- The output of Final result will be a prediction.

Requirements

Hardware Requirements :

- Processor : Intel Core i3 / Pentium
- RAM : 4 GB Minimum
- Hard Disk Space : 30 GB Minimum

Requirements

Software Requirements :

- Operating System : Windows / MacOS / Linux
- Browser : Google Chrome / Mozilla Firefox
- Software Libraries : Python 3.x, Tensorflow 2.0, Annaconda, Jupyter Notebook, Numpy, Pandas, Scikit-learn, Flask/Django

Features

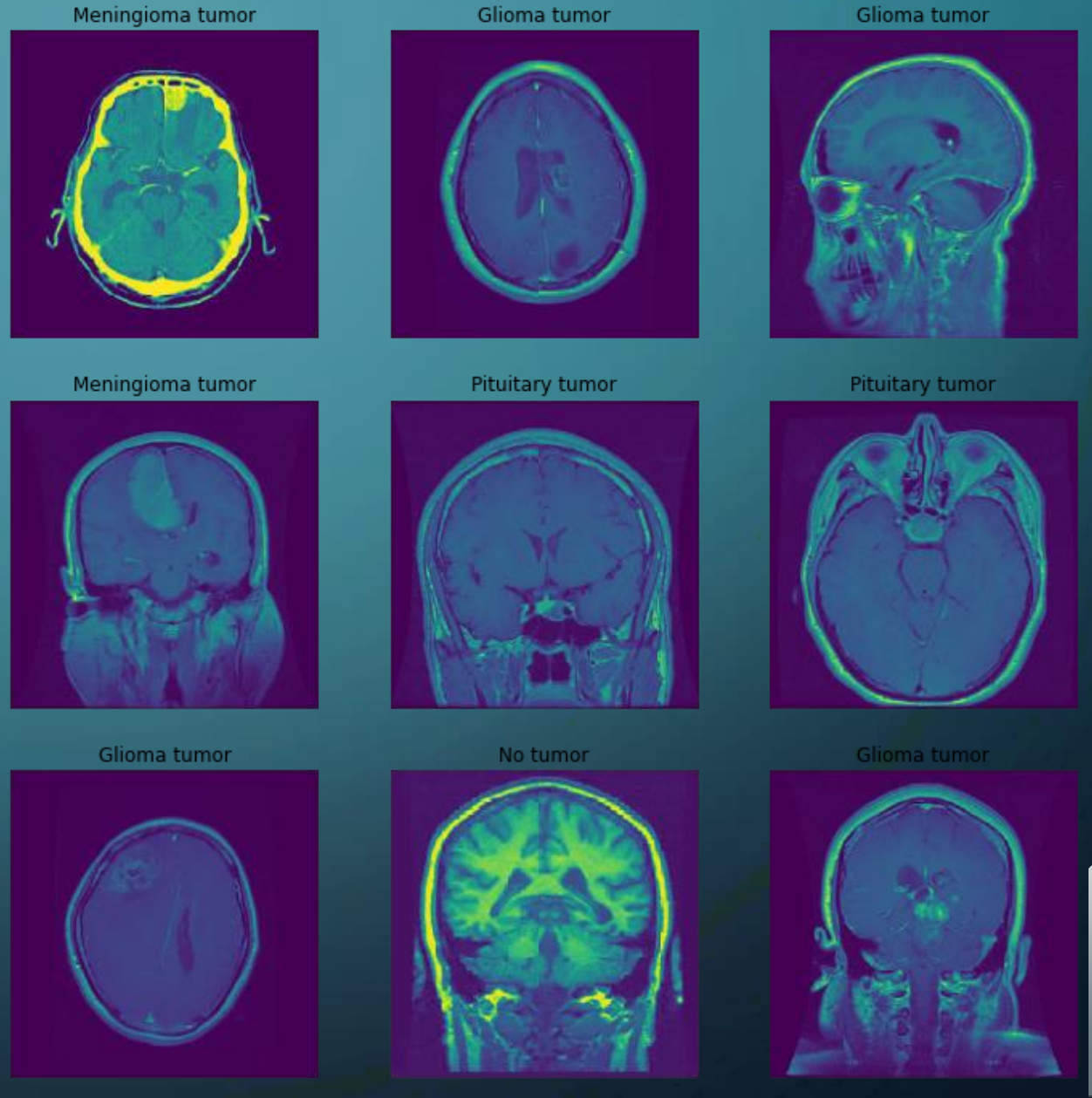
- Early detection of disease
- User friendly GUI
- Precise results
- User Data Store

Why AI-DocHelper ?

- Existing system available are able to detect only one or two diseases.
- Everything in the future will be automated.

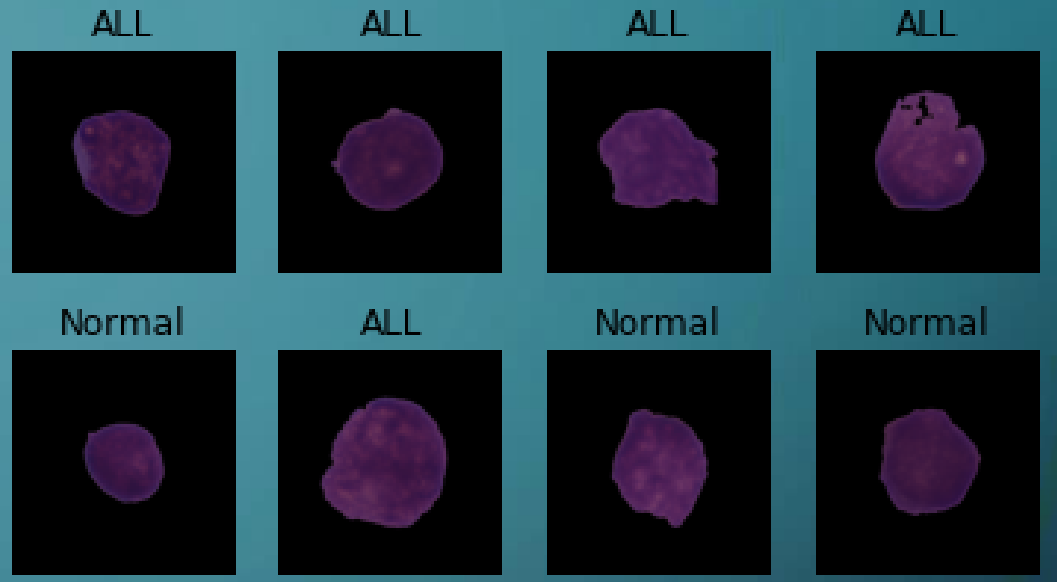
Datasets

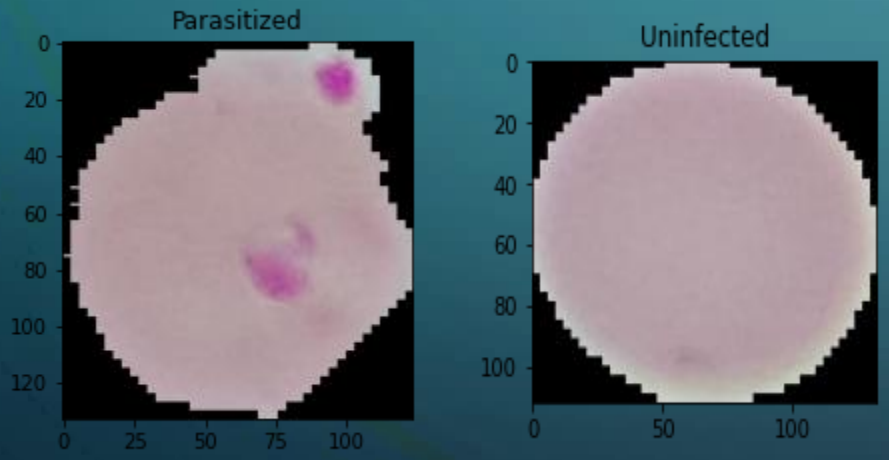
- Brain Tumor Classification (MRI) Dataset : 3624 samples and 4 classes(No Tumor, Glioma Tumor, Meningioma Tumor, Pituitary Tumor)



Datasets

- Leukemia Classification Dataset : 3527 samples and 2 classes(ALL, HEM)

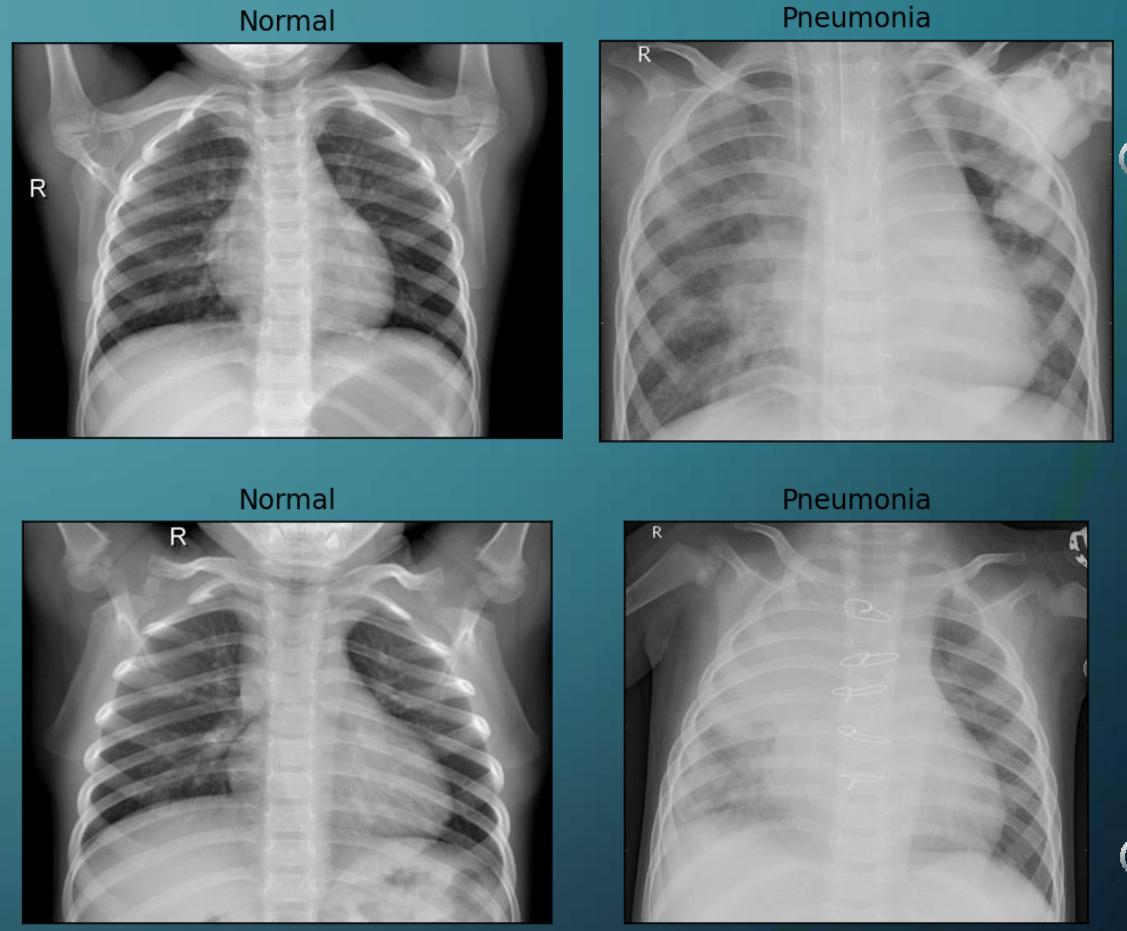




- Malaria Cell Image Dataset : 27.6k samples and 2 classes (Uninfected, Parasitized)

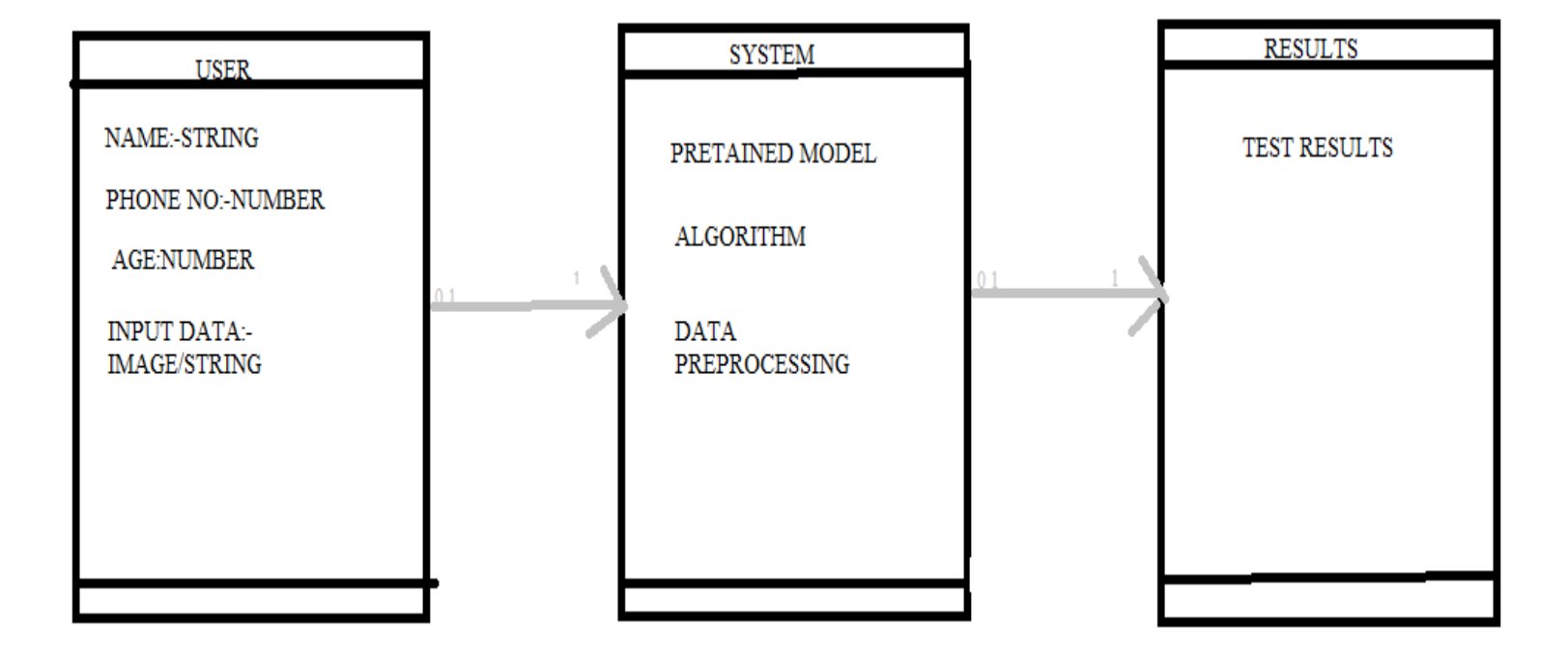
Datasets

- Pneumonia X-ray Images Dataset : 5856 samples and 2 classes (normal, opacity)

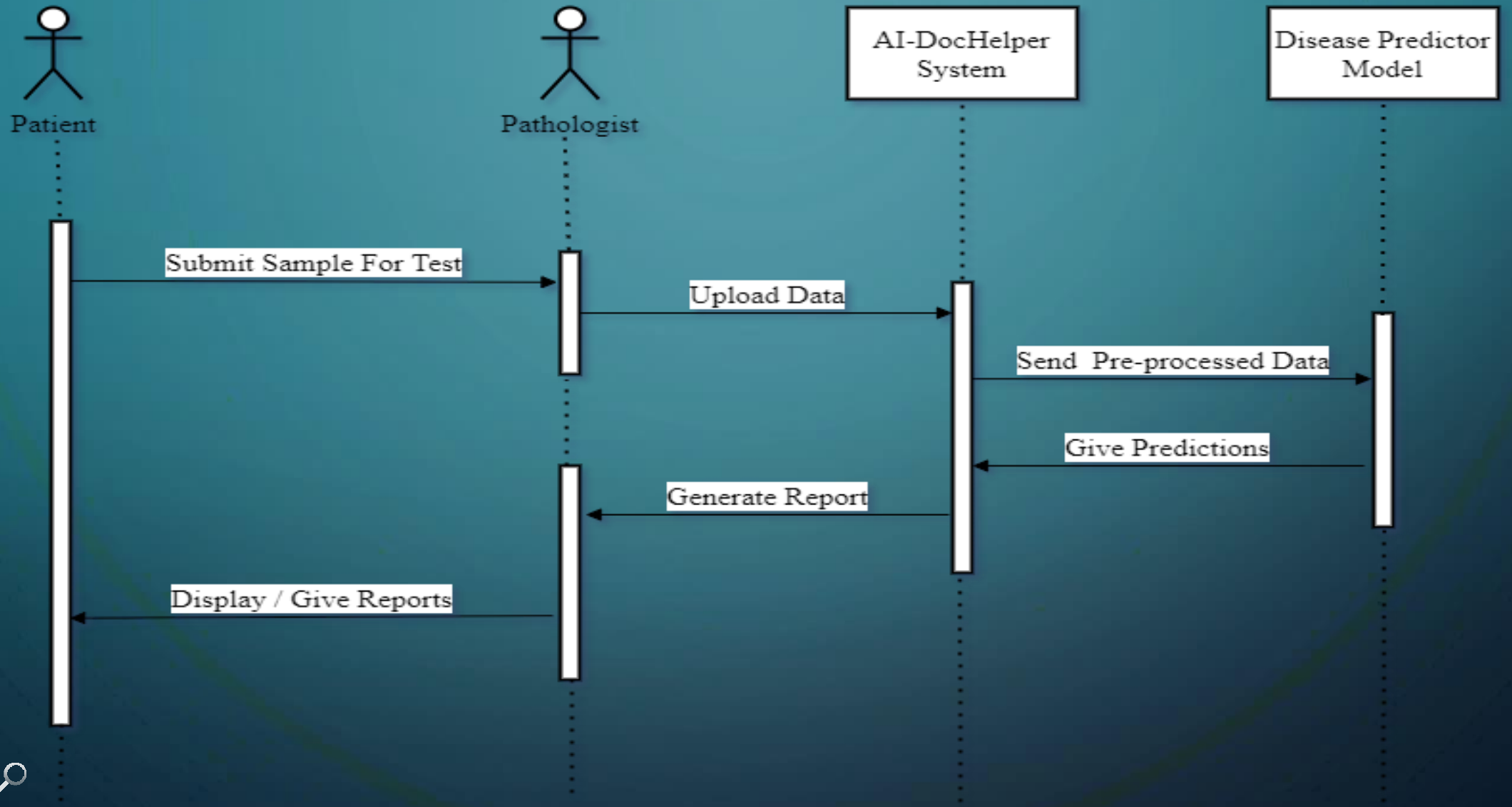


And many more...

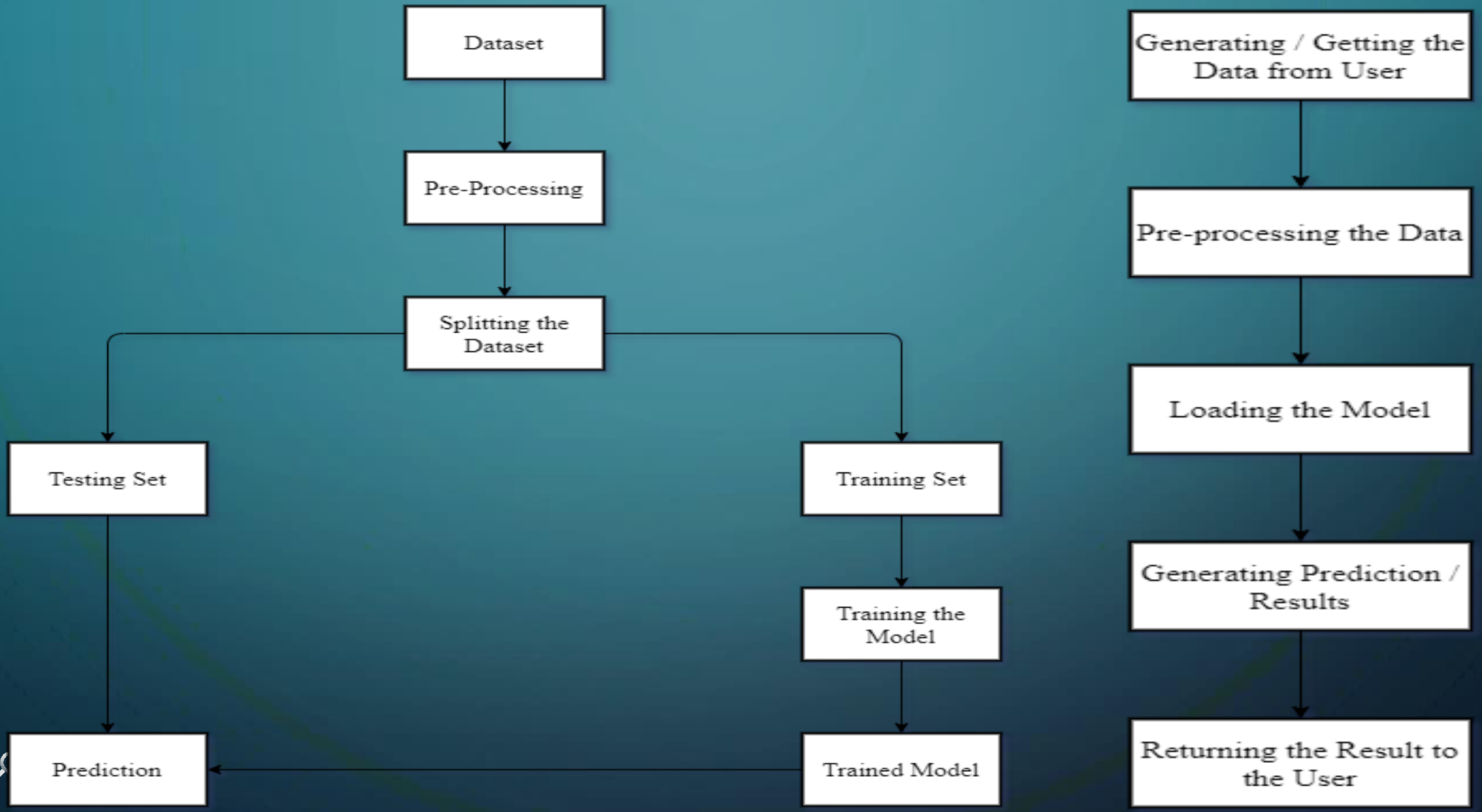
Class Diagram



Flow Of Events



Proposed System : Workflow

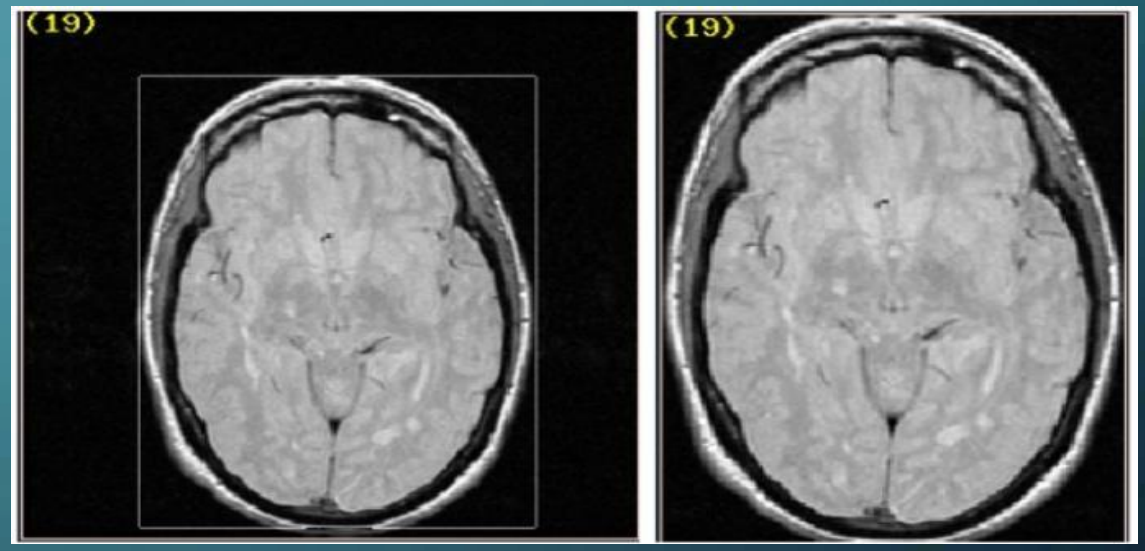


Pre-Processing The Textual Dataset

- Dropping the unnecessary fields.
- Replacing all the missing fields with a value.
- Scaling the data
- Converting categorical data to numeric data.

Pre-Processing The Image Dataset

- Dropping all the blur images.
- Region of Interest
- Filtering & Smoothing an image.
- Getting the Region of Interest / Volume of Interest.



Comparative Study For Textual Dataset

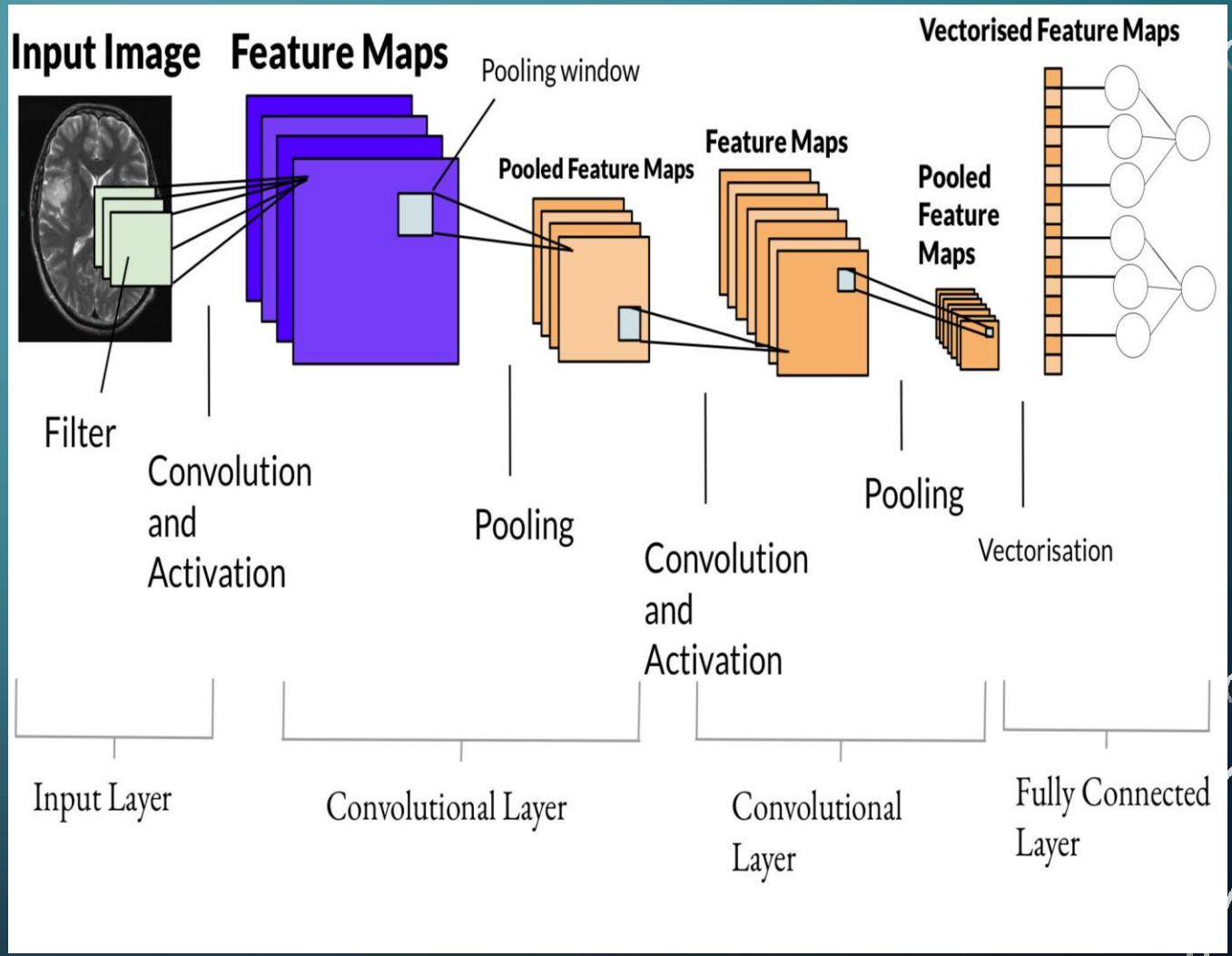
- Pima Indian Diabetes Database / Dataset
- Originally from the National Institute of Diabetes and Kidney Diseases
- Published by UCI Machine Learning
- Dataset Fields : Pregnancies, Glucose BloodPressure), SkinThickness, Insulin, BMI, DiabetesPedigreeFunction, Age, Outcome

Comparative Study For Textual Dataset

MODEL	ACCURACY
Logistic Regression	82.46%
K-Nearest Neighbour Classifier	79.87%
Support Vector Classifier	82.46%
Naïve Bayes	79.22%
Decision Tree Classifier	70.77%
Random Forest Classifier	82.46%
Artificial Neural Network	70.77%

Study For Image Dataset

- Convolutional Neural Network (CNN) with Transfer Learning will be used.



Results And Evaluation

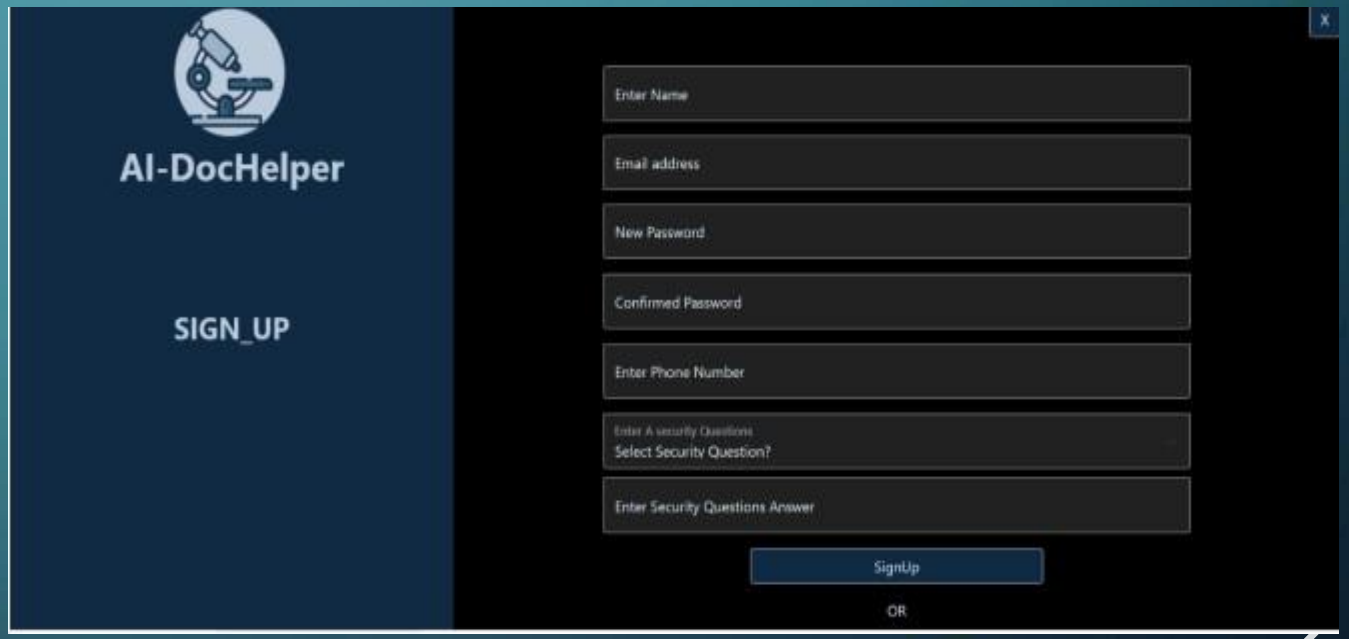
Disease Name	Training Accuracy	Training Loss	Testing Accuracy	Testing Loss
Breast Cancer	87.36%	31.24%	85.49%	38.67%
Pneumonia	95.86%	10.92%	98.25%	14.88%
Glioma Tumor	92.20%	19.64%	92.41%	19.73%
Meningioma Tumor	93.19%	18.18%	90.52%	21.80%
Diabetes	82.40%	23.10%	80.52%	25.13%
Pituitary Tumor	99.06%	3.33%	98.85%	3.03%
Melanoma (Skin) Cancer	85.12%	31.37%	83.87%	33.26%
Acute Lymphoblastic Leukemia (ALL)	84.14%	37.54%	70.43%	58.85%
Heart Disease	91.37%	25.13%	88.97%	28.83%
Chronic Kidney	89.63%	19.53%	86.63%	21.53%
Malaria	89.69%	26.80%	90.49%	24.19%
Brain Tumor	89.21%	28.49%	66.67%	33.72%
Lung Cnacer	94.41%	13.80%	95.26%	11.60%
Invasive Ductal Carcinoma (IDC)	80.57%	43.16%	78.70%	47.90%
Tuberculosis	96.69%	9.75%	95.64%	11.64%
COVID-19	86.56%	33.12%	86.90%	32.88%

Tools & Design

- Web-Based System
- Tools suitable for design.
- Advantage over other tools.

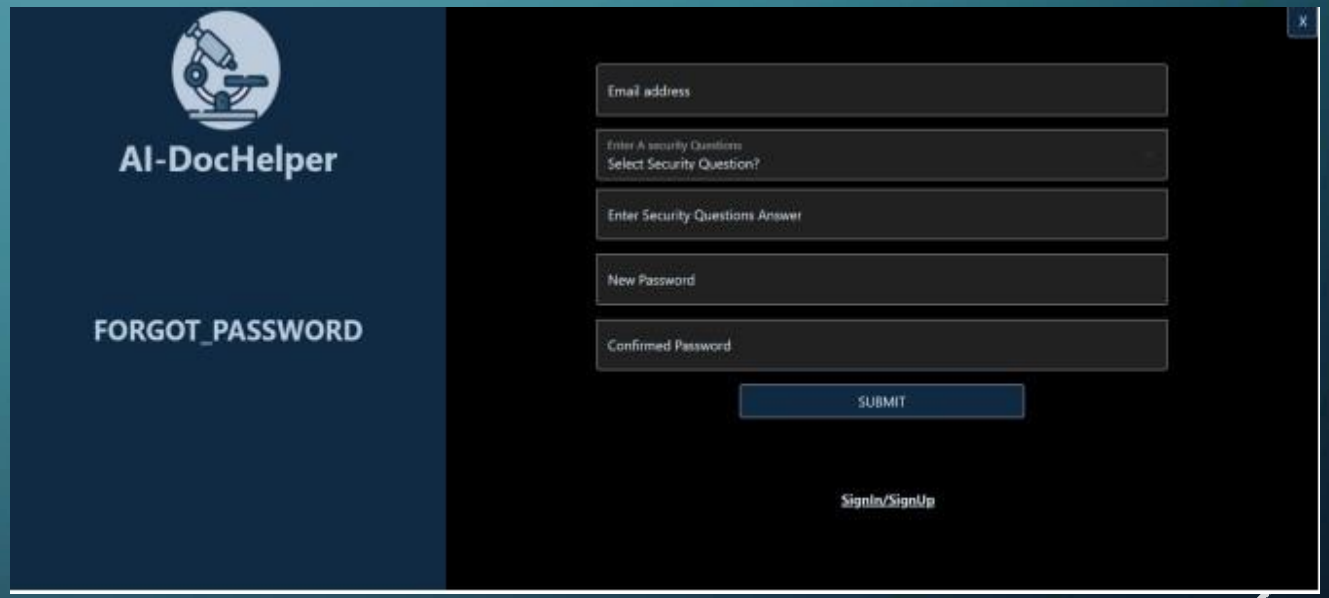
AUTHENTICATION & ACCESSBILITY

- Only registered user can access the system.
- Uses SQLAlchemy to connect to database.
- Provides security.



Updation in User Details

- What if user forgets password
- Provides easy modifications in users details.
- Commits the changes in database and allows accessing system with new details.



Conclusion

- Better than many systems available in the market as it can detect many diseases.
- Will help to save many lives with precision.
- Human errors will be reduced to a great extent.
- Course of treatment can be planned very quickly.

References

1) R. Sangeetha and K. S. Murthy, "A novel approach for detection of breast cancer at an early stage using digital image processing techniques," *2017 International Conference on Inventive Systems and Control (ICISC)*, Coimbatore, 2017, pp. 1-4, doi: 10.1109/ICISC.2017.8068625.

2) Y. Lu, J. Li, Y. Su and A. Liu, "A Review of Breast Cancer Detection in Medical Images," 2018 IEEE Visual Communications and Image Processing (VCIP), Taichung, Taiwan, 2018, pp. 1-4, doi: 10.1109/VCIP.2018.8698732.

3) S. Nayak, S. Kumar and M. Jangid, "Malaria Detection Using Multiple Deep Learning Approaches," 2019 2nd International Conference on Intelligent Communication and Computational Techniques (ICCT), Jaipur, India, 2019, pp. 292-297, doi: 10.1109/ICCT46177.2019.8969046.

4) X. Zeng, H. Chen, Y. Luo and W. Ye, "Automated Diabetic Retinopathy Detection Based on Binocular Siamese-Like Convolutional Neural Network," in IEEE Access, vol. 7, pp. 30744-30753, 2019, doi: 10.1109/ACCESS.2019.2903171.

5) T. J. Wroge, Y. Özkanca, C. Demiroglu, D. Si, D. C. Atkins and R. H. Ghomi, "Parkinson's Disease Diagnosis Using Machine Learning and Voice," 2018 IEEE Signal Processing in Medicine and Biology Symposium (SPMB), Philadelphia, PA, 2018, pp. 1-7, doi: 10.1109/SPMB.2018.8615607.

6) Y. Liu et al., "Detecting Diseases by Human-Physiological-Parameter-Based Deep Learning," in IEEE Access, vol. 7, pp. 22002-22010, 2019, doi: 10.1109/ACCESS.2019.2893877.

7) A. Shrivastava, I. Jaggi, S. Gupta and D. Gupta, "Handwritten Digit Recognition Using Machine Learning: A Review," 2019 2nd International Conference on Power Energy, Environment and Intelligent Control (PEEIC), Greater Noida, India, 2019, pp. 322-326, doi: 10.1109/PEEIC47157.2019.8976601.

Thank You