AI-based multi-disease diagnostic system

Team Member :

Vraj Patel (21BECE30244)

Harsh Patel (21BECE30191)

Vishva Patel (21BECE30242)

Charmi Rupareliya (21BECE30278)

Project Mentor:

Prof. Tejasvee Gupta

Department of Computer Engineering

Introduction

The prevalence of chronic and life-threatening diseases has been steadily rising across the globe. Conditions such as diabetes, heart disease, Parkinson’s disease, brain tumors, tuberculosis (TB), and breast cancer have emerged as significant health concerns, impacting millions of lives annually.

Challenges in Healthcare Diagnostics

High Costs

Traditional diagnostic methods are expensive, limiting accessibility for many.

Accessibility

Limited healthcare infrastructure in remote areas creates disparities in access to care.

Delayed Diagnosis

Late detection can lead to severe complications and worsened outcomes.

Limited Tools

Insufficient early-stage detection tools hinder proactive care.

Proposed Solution: AI-Powered Diagnostics

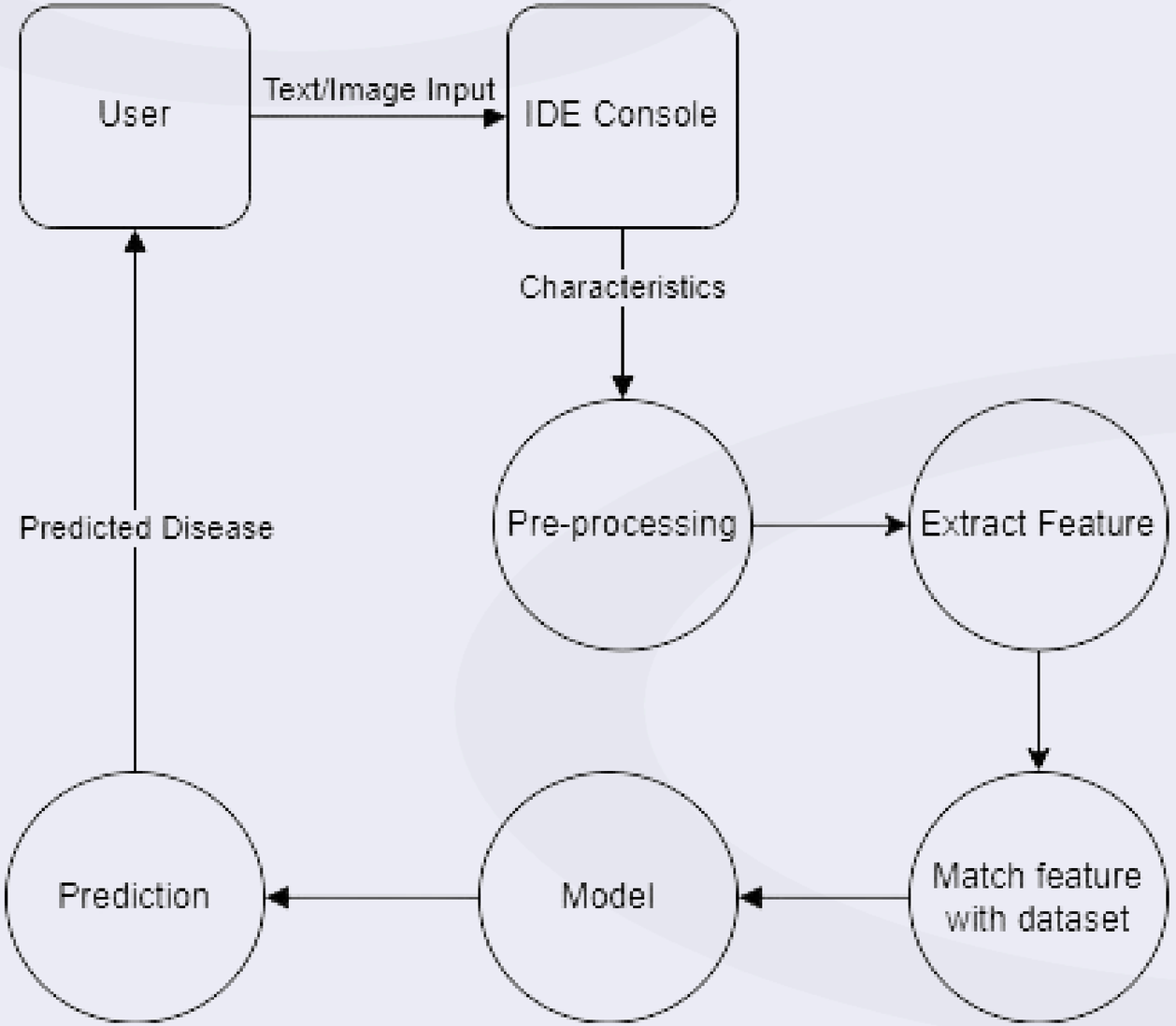
1. Provided accessible and cost-effective detection for six critical diseases, benefiting remote and underserved areas.

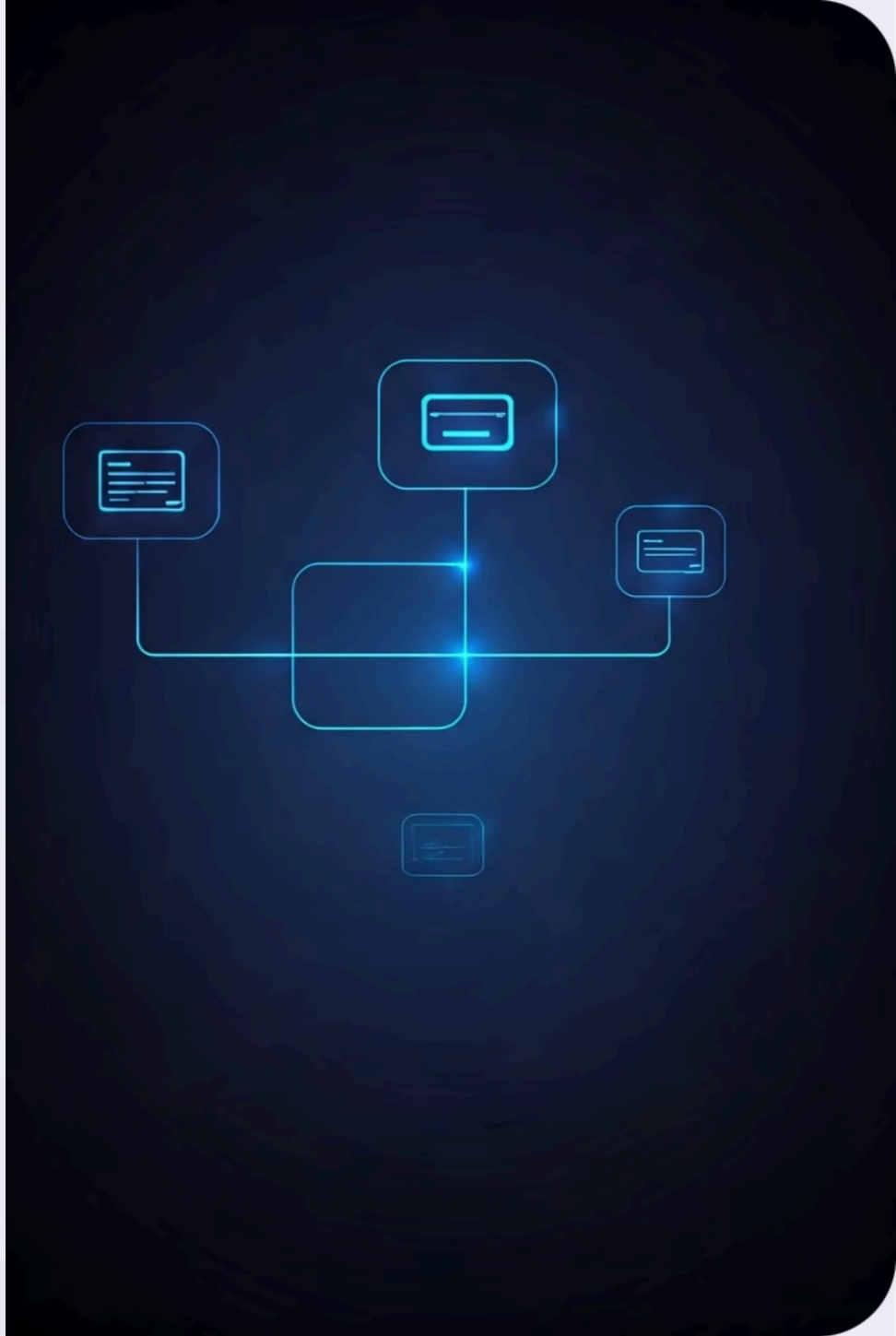
2. Enabled early diagnosis to improve treatment outcomes and reduce complications.

3. Reduced reliance on expensive diagnostic methods through AI-driven models.

4. Delivered accurate and reliable results, minimizing errors in disease detection.

Data Flow Diagram(DFD)





Diabetes Model Performance

Model Architecture	Precision	Recall	Accuracy	F1 Score
Logistic Regression	0.71	0.51	75%	0.60
Niave bayes	0.72	0.57	77%	0.63
Decision Tree	0.69	0.46	74%	0.55
Random Forest	0.67	0.5	74%	0.57
(ANN + XGBoost (Hybrid)	0.74	0.5	75%	0.55
TabNet	0.71	0.51	72%	0.50
SVM	**0.75**	**0.51**	**77%**	**0.63**

Heart Model Performance

Model Architecture	Precision	Recall	Accuracy	F1 Score
EfficientNet	0.8	0.84	80%	0.82
LSTM	0.74	0.78	73%	0.76
tabtansformer	0.78	0.90	81%	0.84
KNN	0.81	0.81	81%	0.83
AdaBoost	0.83	0.75	78%	0.79
GradientBoosting	0.77	0.72	73%	0.75
SVM	**0.81**	**0.81**	**80%**	**0.81**

parkinson’s Model Performance

Model Architecture	Precision	Recall	Accuracy	F1 Score
Random Forest	1.000	1.000	0.963	1.000
Logistic Regression	0.600	0.818	0.857	0.694
KNN	**0.667**	**0.909**	**0.887**	**0.769**
XGBoost	1.000	1.000	0.963	1.000
SVM	1.000	1.000	0.963	1.000

Breast Cancer Model Performance

Model Architecture	Precision	Recall	Accuracy	F1 Score
CNN (5 Layers)	0.81	0.81	89%	0.81
DenseNet121	**0.89**	**0.89**	**97%**	**0.89**

Tuberculosis Model Performance

Model Architecture	Precision	Recall	Accuracy	F1 Score
DenseNet121	0.91	0.44	90%	0.59
EfficientNetB0	0.72	0.57	77%	0.63
DenseNet169	**0.93**	**0.41**	**95%**	**0.56**

Brain Tumor Model Performance

Model Architecture	Precision	Recall	Accuracy	F1 Score
Yolov8	0.71	0.81	81%	0.70
CNN (8 Layers)	0.72	0.82	82%	0.73
VGG16	**0.98**	**0.91**	**99%**	**0.94**

Advantages and Challenges

Advantages

- Cost-effective compared to traditional methods.
- Early-stage detection reduces fatality risks.
- Automation minimizes human error in diagnosis.
- Scalable for future disease expansion.

Challenges

- Limited availability of labeled and high-quality datasets.
- Balancing accuracy across diseases with varying data sizes.
- Computational constraints during training on large datasets.

Project Modules: Progress and Future Work

Category	Tasks	Details
Complete Work	Data Collection	Gathered datasets for six diseases (e.g., diabetes, heart disease, Parkinson’s, etc.) from reliable sources like Kaggle and UCI.
	Preprocessing	Cleaned, normalized, and augmented the data for better model training and accuracy.
	Model Training	Trained machine learning models (SVM, CNN, KNN, etc.) for each disease on Google Colab.
	Compare Model Accuracy	Evaluated models based on accuracy, precision, and recall; identified the best fit for each disease.
	UI Development	Plan to create a user-friendly interface for seamless interaction with the diagnostic system.
Future Work	LLM Integration for Precautions	Incorporate a large language model (LLM) to provide users with disease-specific precautions and health advice based on diagnostic results.

References

- https://ieeexplore.ieee.org/document/10060903
- https://www.e3s-conferences.org/articles/e3sconf/pdf/2023/67/e3sconf_icmpc2023_01051.pdf
- https://www.irjmets.com/uploadedfiles/paper//issue_1_january_2024/48476/final/fin_irjmets1705419474.pdf
- https://www.researchgate.net/publication/381309960_MULTI_DISEASE_PREDICTION_SYSTEM_USING_MACHINE_LEARNING_