MULTI-DISEASE PREDICTION SYSTEM USING ML AND DJANGO

by

G. M. Sumanth
Roll No. 214G1A32A4

K. Nithin Kumar
Roll No. 214G1A3269

B. Pallavi
Roll No. 214G1A3270

C. Sri Varshini
Roll No. 214G1A32A3

Under the guidance of

Mrs. V. Kamakshamma M. Tech. (Ph. D.)
Assistant Professor

Department of Computer Science and Engineering (Data Science)

Srinivasa Ramanujan Institute of Technology

(Affiliated to JNTUA & Approved by AICTE) (Accredited by NAAC with 'A' Grade & Accredited by NBA (EEE, ECE & CSE)
Rotarypuram Village, B K Samudram Mandal, Ananthapuramu – 515701.

2024 - 2025

Contents

Contents

- Requirements
- UML diagrams
- Data preprocessing techniques
- References
- GitHub Dashboard

Abstract

Chronic Kidney Disease (CKD), Heart Disease, and Cervical Cancer are among the leading health challenges globally. Early detection and prevention are vital to reduce morbidity and mortality. However, traditional diagnostic processes are often resource-intensive and time-consuming, making it challenging. This project proposes a **Multi-Disease Prediction System** that leverages machine learning models to efficiently and accurately predict the likelihood of these diseases based on patient data. The Multidisease Prediction System demonstrates the potential of leveraging machine learning to address critical health challenges. The modular nature of the system allows scalability to include additional diseases in the future. The project serves as a data-driven decision-making mechanism, in medical situations.

Problem Statement

What is the problem?

- Early detection of chronic diseases like CKD, heart disease, and cervical cancer is critical but challenging due to lack of awareness.
- Late diagnosis leads to severe complications, higher costs, and mortality rates.
- Traditional methods are time consuming and require human assistance during all steps of the diagnosis.
- Conventional methods are also really costly and have a chance of giving a false positive at times.

Project Objectives

Objective No. 1:

- Build a machine learning-based system to predict the likelihood of chronic kidney disease, heart disease, and cervical cancer. Integrate multiple prediction models into a single, user-friendly interface for healthcare providers.

Objective No. 2:

- Provide awareness blogs about those diseases on the same interface about things like how to identify those diseases, what are the symptoms, how to prevent them, etc.

Literature Survey for Objective 1

No	Title	Authors	Journal Name & Year	Methodology Adapted	Key Findings	Gaps
1.	Disease Prediction using ML	Dr. N. Das, S. Gayke, N. Patel, S. Shinde	Int. Journal of Innovative Research in Science, Engg., and Technology	ML Models for Disease Prediction	Specific ML models can be developed based on the use case	No discussion of Recall Scores, No UI at all
2.	ML for Multiple Disease Prediction	K. B. Singh, A. Sharma, A. Verma, R. Maurya, Dr. Y. Perwej	Int. Journal of Scientific Research in CS, Engg. and IT (IJSRCSEIT) - 2024	A model per disease can be used	Different ML models were trained and tested against different diseases	Limited models for 3 diseases, and no UI

Literature Survey for Objective 2

No	Title	Authors	Journal Name & Year	Methodology Adapted	Key Findings	Gaps
3.	Human Disease Prediction using Machine Learning Techniques and Real-life Parameters	K. Gaurav, A. Kumar, P. Singh, A. Kumari, M. Kasar, T. Suryawanshi	International Journal of Engineering (IJE) - 2023	Machine Learning Techniques	Various ML models can be made for disease predictions	Usage of models which are too simple, no awareness blogs
4.	Multiple Disease Prediction using ML	K. Reshma, P. Niharika, J. Haneesha, K. Rajavardhan, S. Swaroop	Int. Research Journal of Modernization in Engg. Tech. And Science (IRJMETS) - 2024	An Interactive User Interface	A user interface integrating those multiple disease prediction models under one web app	Usage of old and outdated libraries, no database integration, low model performance, No articles on awareness

Proposed Work

- A proposed system for this project includes developing predictive ML models in the first step, for which Python's extensive Data Manipulation libraries like Numpy and Pandas can be used. And for making the actual ML models, Scikit-learn can be used.
- Methods to develop ML models include:
  i. Data collection (from the internet)
  ii. Data preprocessing (data cleaning, feature scaling, feature encoding)
  iii. Training and developing the ML models
  iv. Testing various ML models for each task and choosing the best performing one
- The second part of this project is the web application to interact with the models. The frontend of this application can be made using HTML, CSS, and JS (if required), and its backend can be made using Python and Django.
- Then the models can be integrated onto the web application after their development.
- Articles can be written on the web app, providing awareness on the diseases.

Existing system vs proposed system

Existing System	Proposed System
1. Average performing models	1. Better performing models
2. No discussion of recall scores	2. Models will be developed based on recall scores as the primary metric
3. No interface or outdated web interface	3. A responsive and clean interface
4. No awareness blogs of any sort	4. Blogs/Articles are to be included providing awareness on the diseases
5. No authentication for users	5. Authentication will be provided

Planning

Phase 1: Literature Survey

Phase 2: Data Collection

Phase 3: Data Preprocessing

Phase 4: Model Training

Phase 5: Model Evaluation

Phase 6: System Integration

Phase 7: Testing

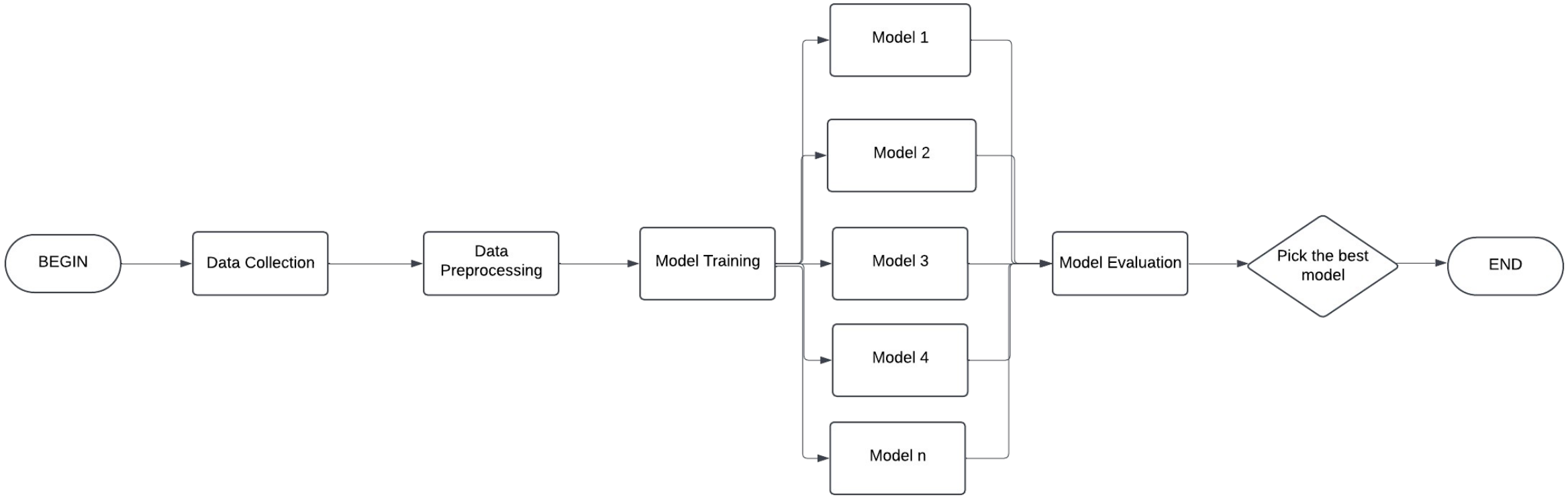
Requirements

The following are the requirements to develop this project:

1) A working OS (Windows 10 or above, Linux, or MacOS)
2) Python 3.12 or above
3) Django library (can be installed using pip)
4) Scikit-learn library (can be installed using pip
5) Data manipulation libraries: Numpy, Pandas
6) HTML5, CSS3, JS (if needed)

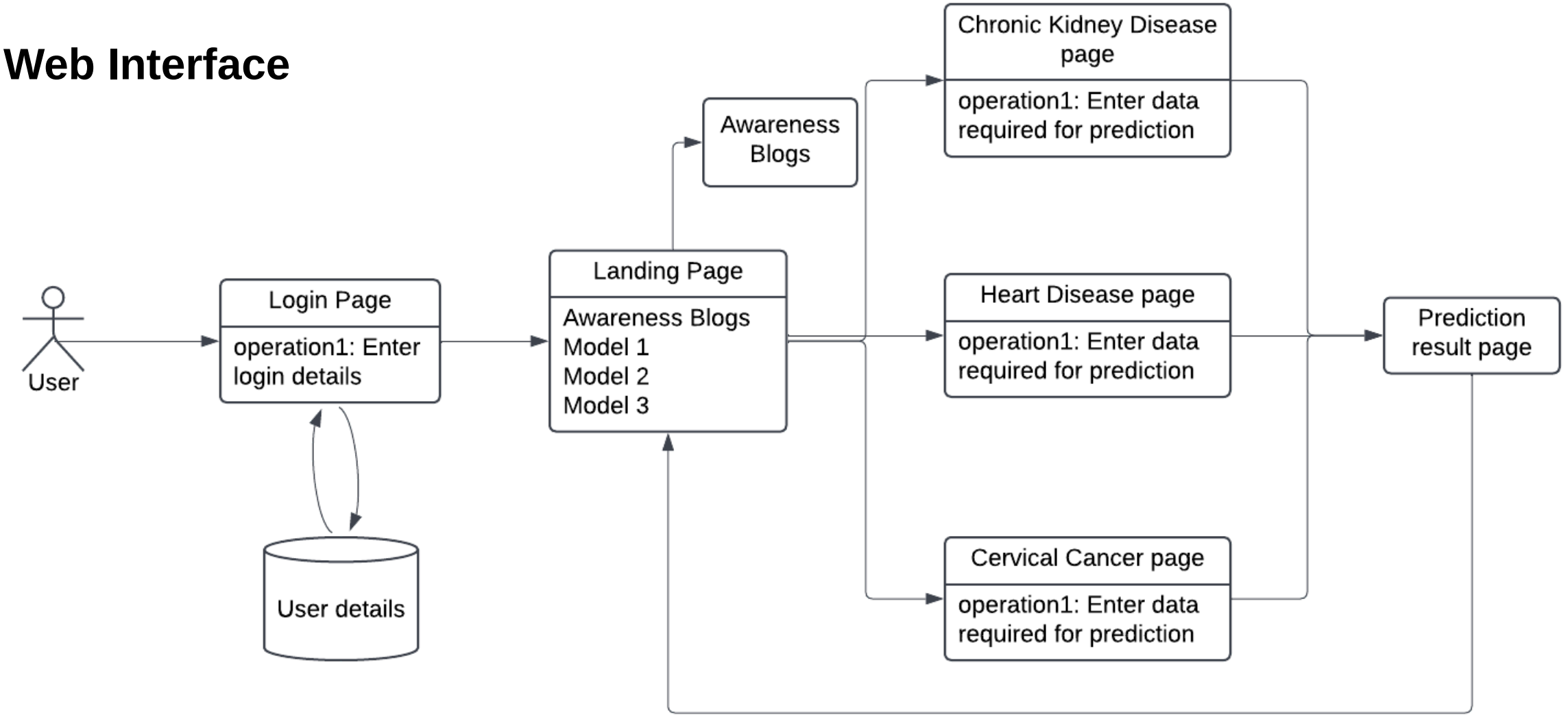
UML Diagrams

Training the models



UML Diagrams

Web Interface



Data Preprocessing Techniques

Following data preprocessing techniques are employed before training the models:

1) Handling missing values - Mean Imputation

2) Feature Encoding - Label Encoding

3) Feature Scaling - Standardization

References

[1]. K. Gaurav, A. Kumar, P. Singh, A. Kumari, M. Kasar, T. Suryawanshi, "Human Disease Prediction using Machine Learning Techniques and Real-life Parameters", International Journal of Engineering (IJE) – 2023.

[2]. K. Reshma, P. Niharika, J. Haneesha, K. Rajavardhan, S. Swaroop, “Multiple Disease Prediction using ML”, Int. Research Journal of Modernization in Engg. Tech. And Science (IRJMETS) – 2024.

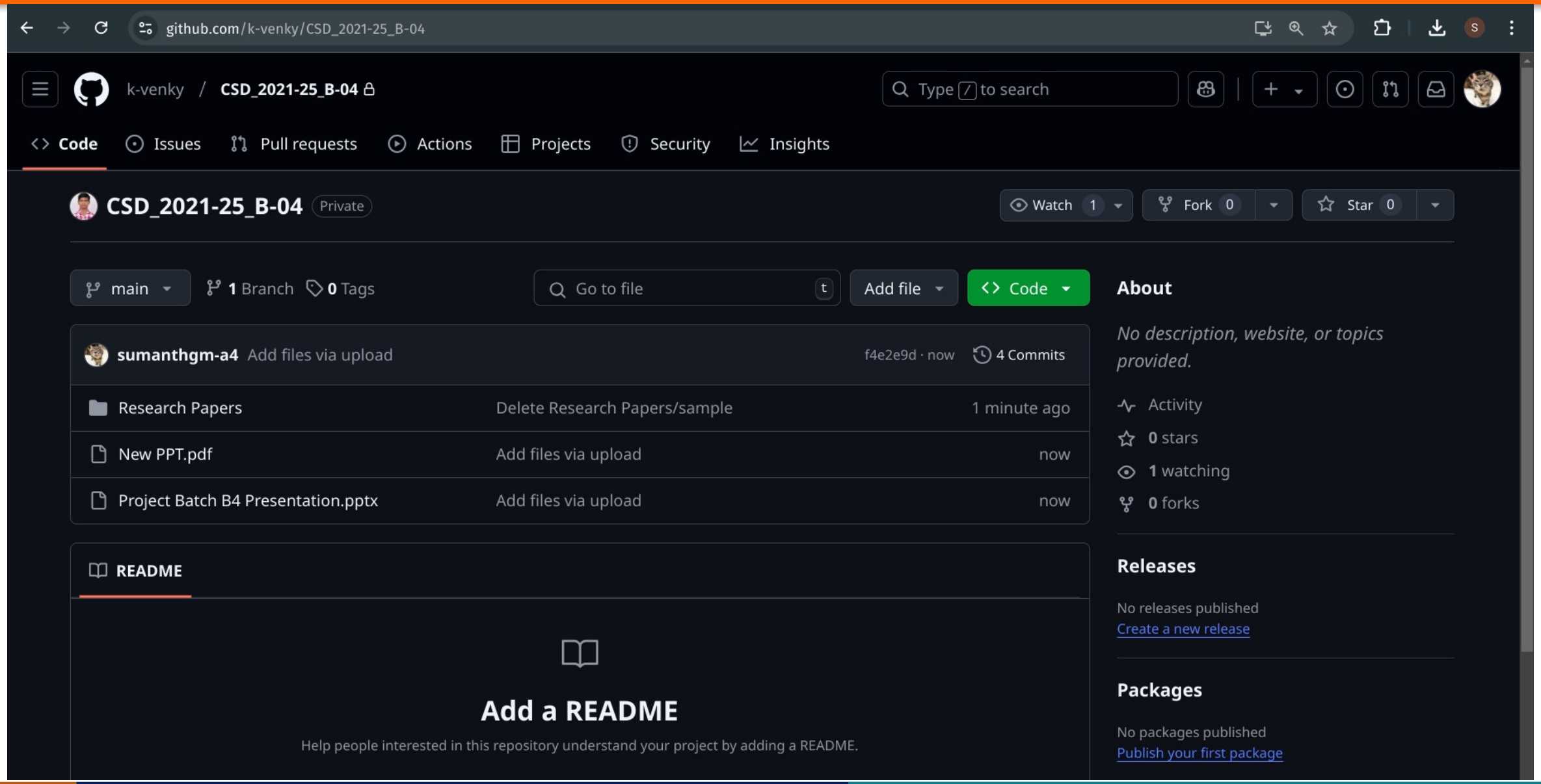
[3]. Dr. N. Das, S. Gayke, N. Patel, S. Shinde, “Disease Prediction using ML”, Int. Journal of Innovative Research in Science, Engineering and Technology (IJIRSET)

References

[4]. K. B. Singh, A. Sharma, A. Verma, R. Maurya, Dr. Y. Perwej, “ML for Multiple Disease Prediction”, Int. Journal of Scientific Research in CS, Engg. and IT (IJSRCSEIT) – 2024.

[5]. S. Yadav, V. Jaglan, Y. Singh, “ML Approaches for Disease Prediction”, Inst. Of Electrical and Electronics Engineers (IEEE)

GitHub Dashboard



Any Queries?

Thank You!!!