Major-1 Project

Synopsis Report

For

CardioSense- A cloud-based Heart Disease Detection using Explainable AI

Submitted By

|  |  |  |
| --- | --- | --- |
| Specialization | Roll Number | Name |
| CCVT(B8) | R2142201587 | Aditya Kumar Tiwari |
| CCVT(B8) | R2142201589 | Akanksha Singh |
| CCVT(B8) | R2142201405 | Harshit Ojha |
| CCVT(B8) | R2142211390 | Ajay Saini |



Cloud Software Operations Cluster

School Of Computer Science

UNIVERSITY OF PETROLEUM & ENERGY STUDIES

DEHRADUN- 248007, Uttarakhand

|  |  |
| --- | --- |
|  |  |
| Dr. Roohi Sille | Dr. Hitesh K. Sharma |
| Project Guide | Cluster Head |

**INDEX**

|  |  |  |
| --- | --- | --- |
| **S. No.** | **Topic** | **Page No.** |
| 1 | Project Title | 3 |
| 2 | Abstract | 3 |
| 3 | Introduction | 3 |
| 4 | Problem Statement | 3 |
| 5 | Literature Review | 4 |
| 6 | Objective | 4 |
| 7 | Technology Stack | 4 |
| 8 | Methodology | 5 – 7 |
| 9 | SWOT Analysis | 8 |
| 10 | Gantt Chart | 9 |
| 11 | Reference | 9 |
| 12 | GitHub Link | 10 |



**School of Computer Science**

**University of Petroleum & Energy Studies, Dehradun**

**1. Project Title**

CardioSense- A cloud-based Heart Disease Detection using Explainable AI

# **Abstract**

CardioSense aims to tackle some of the most critical global health problems related to cardiovascular diseases, still today among the leading causes of deaths worldwide. To this end, the project tries to improve diagnosis via explainable AI.

The explainable AI provides clear crystalline details of what factors drive each diagnosis. This is particularly important in a clinical setting, as the doctors must be informed behind the recommendation made by AI to be trusted and accountable. The system visualizes key diagnostic factors, connecting complex AI models to practical and clinical use. By enhancing the accuracy and explainability of heart disease diagnosis, CardioSense creates a reliable tool for clinicians to confidently use in order to make informed, timely decisions that achieve better patient outcomes.

# **Introduction**

The CVDs have remained continuously the leading cause of death worldwide, with about 17.9 million deaths annually. The diseases are supposed to be diagnosed early enough for timely intervention in order to reduce morbidity and mortality arising from heart diseases. However, the diagnosis through the identification of underlying risks is usually done using rather sophisticated tools due to the sophistication inherent in most cardiovascular conditions. In this scenario, CardioSense can be thought of as a bright solution that uses cloud computing and explainable AI in order to enable a progressive, scalable, and interpretable method of heart disease detection.

CardioSense is a cloud-based analytics platform that takes inputs from the patient's past medical history, his lifestyle, and clinical parameters to predict the possibility of heart disease. It is the aspiration of CardioSense to provide an accuracy as high as possible but also give insight into how those AI models make their decisions in a way that is as transparent as possible, using machine learning models enhanced with explainable AI techniques. That is important in instilling confidence among health professionals and patients through the provision of understandable and actionable recommendations made by AI. It will be integrated with cloud computing to let the CardioSense app provide seamless access to both health professionals and patients for immediate analysis and monitoring from any location. The integration of high-level AI technologies into practical healthcare can further empower such cooperation to support the enhancement of early detection and management capabilities of cardiac ailments, creating a better outlook for the patient.

# **4.Problem Statement**

Heart-disease is one of the leading causes of mortality among people of the world. Most of the available diagnostic methodologies lack accessibility, scalability, and transparency. The result requires a non-invasive data-based detection of heart diseases, with explainability in AI integrated into building trust for AI-driven healthcare toward efficient early diagnosis.

# **5.Literature Review**

The authors of [1] paper critically examines the current state of explainable AI, highlighting key challenges and recent advancements, while also proposing future research directions to enhance understanding and implementation in machine learning and deep learning contexts.

The authors of this research [2] offer a new view on heart disease detection problems, including such ensemble learning algorithm robustness as is impossible for any other machine learning paradigm, Bayesian optimization of hyperparameters, and SHAP-a method of interpretation of the models. The optimized model XGBoost showed the best performance with high specificity and sensitivity on both the Cleveland and Framingham datasets, which can indicate the possibility of clinical adoption.

# [3] The K-Nearest Neighbour algorithm is performed to present early prediction regarding heart diseases, showing effectiveness and efficiency in mining medical data. The selected relevant features improved accuracy with reduced runtime; hence, the study achieved an 86% accuracy rate in the prediction of heart disease, also presenting the fact that diabetes is one of the leading influencing factors that contribute to heart disease.

In the paper [4], the author proposes a deep learning model, which is a hybrid system that combines convolutional neural networks with long short-term memory for enhancing the early diagnosis of cardiovascular diseases. By the proposed model having high accuracy greater than 73%, feature engineering and explainable AI were used in order to enhance the accuracy and interpretability of the CVD prediction, hence proving its worth in being a useful tool for screening purposes within clinical practice.

# **6. Objectives**

* Early detection of Cardio Vascular diseases and integration of advanced machine learning models
* To critically explore the past performances and apply an appropriate methodology to superscribe the identified problem
* Improvement of diagnostic accuracy and clinical applicability
* To deploy the application onto a cloud platform to ensure scalability, real-time processing and global accessibility

# **7. Tech Stack**

* SQL Databases: MySQL for structured data storage
* NoSQL Databases: MongoDB
* Programming Languages: Python (with libraries like pandas, NumPy)
* Cloud Platforms: AWS: For computing (EC2), storage (S3), and database services (RDS, DynamoDB).
* Frontend: Angular for building the user interface.
* Backend: Flask (Python) for server-side logic.
* APIs: RESTful APIs
* Version Control: Git/GitHub
* SDLC Model: Agile Model
* IDE: VS Code

# **8.Methodology**

# 

Fig 1- Methodology for Heart Disease Detection using Explainable AI

* Data Acquisition and Preprocessing

Data will be collected from publicly available datasets like the Cleveland Clinic Foundation Heart Disease Dataset and other medical repositories to ensure diverse population representation. The demographic data includes age and sex, a past medical history of smoking and diabetes, and blood pressure and cholesterol as clinical variables. Missing values will be identified and handled using techniques like imputation or deletion. Numerical features will be normalized or standardized to maintain consistent scales, and categorical variables will be encoded using one-hot or label encoding. Outlier detection will be conducted to remove outliers and avoid biased results.

* Feature Extraction Feature Selection

Select the most relevant features that give a significant contribution to predicting heart disease. Perform correlation analysis or feature importance from machine learning models, or use domain-specific knowledge in order to select informative features. Create new features from the original ones, thus enabling the capture of more complex relationships. Examples include the calculation of body mass index, the interaction terms between features, and time-series analysis in case of longitudinal data.

* Model Development and Training

Select appropriate machine learning algorithms for classification problems.  
Traditional Approaches: Logistic regression, decision trees, random forests, support vector machines, neural networks. Compare models using relevant performance metrics: accuracy, precision, recall, F1-score, AUC. Optimize model parameters for the best performance. Usage of Grid search, Random search or Bayesian optimization  
Split dataset into training and testing set to check the generalization  
Train the selected model on training set by using appropriate optimization algorithms - examples include gradient descent, or stochastic gradient descent.

* Explainable AI Techniques

Generate local explanations of individual predictions by approximating the behaviour of a model with a simple interpretable model.  
Explain the prediction to each feature by using game theory notions and give global and local explanations. Model-Specific Explanations Decision Trees and Rule-Based Decision rules or paths taken in a decision tree intuitively point out how the model is generating its predictions. Neural Networks: Visualize relevant regions of the input data by using techniques such as saliency maps, class activation maps, or by visualizing attention for models that make use of attention mechanisms.

* Model Evaluation and Refining Performance Evaluation

Performance will be measured by using a number of metrics on the test set, namely: accuracy, precision, recall, F1-score, and AUC. Appropriateness: Domain-specific performance measures may be considered if relevant. Explanation Evaluation: The quality and interpretability of the generated explanations shall be verified. It shall be verified that explanations correspond to medical domain knowledge and are insightful with respect to model decision processes. If necessary, follow up on model development by adjusting feature engineering or tuning hyperparameters or trying other algorithms.  
Include explainability techniques feedback to boost model performance and interpretability.

* Deployment and Monitoring

Deploy the trained and explainable model in real-world applications like a healthcare system or mobile app. Perform proper data preprocessing and handling for new input.  
Perform continuous monitoring of model performance while at work to identify concept drift or degradation. Regularly perform model updates with newer data to maintain accuracy. Re-evaluate explainability techniques to ensure ongoing interpretability.

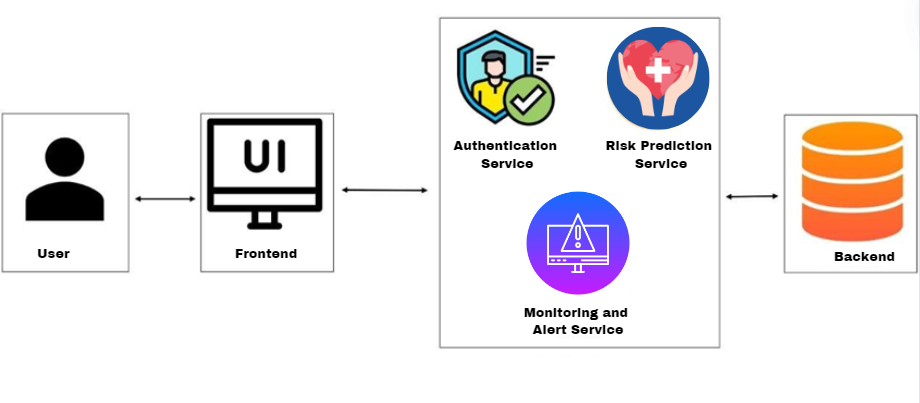


Fig 2- Monolithic Architecture for the application

* In Monolithic architecture as shown in Fig 1, all the services are part of one backend which is further connected to a single database.
* This type of architecture poses a problem when we require to scale the application as the entire backend will have to be replicated on a different system

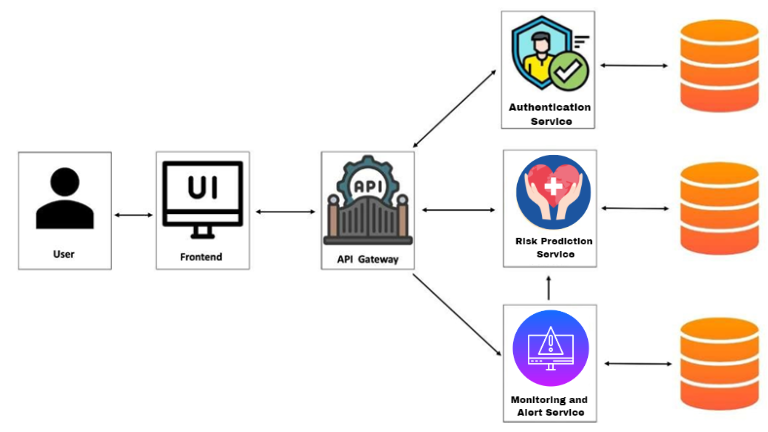


Fig 3- Microservice Architecture for the application

* In Microservice architecture as show in Fig 3, all the services have their own backend and database. This allows to scale the services individually. All the microservices can be deployed on to separate containers which can be scaled individually as per requirement.
* At the end of a sprint, a deliverable will be produced.

# **9.SWOT Analysis**

Strengths:

* Early detection by advanced AI may reduce mortality rates by heart disease through early detection with CardioSense, allowing timely interventions. Explanation: Giving insights into how decisions have been made, most AI models act like black boxes; hence, giving trust to health care professionals and patients.
* Scalability and Accessibility: Since it is cloud-based, CardioSense can easily be scaled up across multiple settings of healthcare, including remotely distributed and deprived areas, making heart disease detection more accessible.
* Non-invasive: The platform makes predictions related to the risk of heart disease by considering non-invasive data, including the history of the disease and clinical parameters, hence minimizing discomfort for the patient and the need for extensive testing.

Weaknesses:

* Data quality and availability: The quality of CardioSense predictions will naturally be dependent on the quality and comprehensiveness of the input data. Data for different healthcare providers and patient populations are going to be different.
* Integration with Existing Systems: The integration of CardioSense with existing systems and workflows at healthcare facilities could be unwieldy. This could even imply the need for customization, further increasing overall resource use.
* Technical Ability Requirement: This indeed requires some training of healthcare professionals to analyse and interpret AI model outputs and insights. This creates a learning curve that might be resisted in some cases.

Opportunities:

* Proliferation of AI in Healthcare: In a time when there is an increasing demand for AI-enabled solutions in the healthcare world, it would be an opportunity to place CardioSense at the leading position with respect to diagnosis of heart diseases.
* Expansion of Telemedicine: Once telemedicine gets expanded and remote monitoring and diagnosis becomes operational, this will surely find its place in virtual care platforms.
* Partnerships: Partnerships with large private healthcare institutions, research organizations, and technology companies will further this process of aggregating capability, data, and market reach for CardioSense.
* Regulatory Support: Interest by regulatory bodies in AI solutions for health would mean various approvals could result in more market entry and growth.

Threats:

* Data Privacy and Security: The critical health data in the cloud raises so many concern factors on data privacy and cyber security, causing more problems in terms of regulatory bodies and hesitant customers.
* Competitive: A number of companies are working on such technologies. Therefore, the AI-based health solution market shall be quite competitive, posing repercussions

# **10.Gantt Chart**



# 

Fig 4- Gantt Chart

# **11.References**

1. Angelov, P.P., Soares, E.A., Jiang, R., Arnold, N.I. and Atkinson, P.M., 2021. Explainable artificial intelligence: an analytical review. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *11*(5), p.e1424.
2. Mienye, I.D. and Jere, N., 2024. Optimized Ensemble Learning Approach with Explainable AI for Improved Heart Disease Prediction. *Information*, *15*(7), p.394.
3. Omar, S., Mohamed, N. and Elbendary, N., 2021, August. A Cardiovascular Disease Prediction Using Machine Learning Algorithms. In *The International Undergraduate Research Conference* (Vol. 5, No. 5, pp. 177-179). The Military Technical College.
4. Hossain, M.M., Ali, M.S., Ahmed, M.M., Rakib, M.R.H., Kona, M.A., Afrin, S., Islam, M.K., Ahsan, M.M., Raj, S.M.R.H. and Rahman, M.H., 2023. Cardiovascular disease identification using a hybrid CNN-LSTM model with explainable AI. *Informatics in Medicine Unlocked*, *42*, p.101370.

# **12.GitHub Link**

<https://github.com/AkankshaSingh1313/CardioSense.git>