|  |  |  |
| --- | --- | --- |
| **Tech Saksham**  Capstone Project Report  **ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING FUNDAMENTALS** |  |  |

**“HEART DISEASE PREDICTION”**

**“UNIVERSITY COLLEGE OF ENGINEERING PANRUTI”**

|  |  |
| --- | --- |
| **NM ID** | **NAME** |
| au422621105020 | MADHUSRI D |

|  |  |
| --- | --- |
|  |  |
|  | Ramar Bose |
|  | AI Master Trainer |

**ABSTRACT**

Artificial intelligence (AI) and machine learning (ML) have become pivotal tools in various fields, including healthcare. This report delves into the application of logistic regression, a fundamental ML algorithm, in predicting heart disease. As per the World Health Organization's alarming statistic revealing that four out of five cardiovascular disease (CVD) deaths result from heart attacks, the urgency to accurately identify individuals at risk is evident. Through the utilization of logistic regression, this research aims to discern the likelihood of patients being affected by CVD and subsequently predict their overall risk. This study underscores the significance of leveraging AI and ML techniques to enhance medical diagnostics and prognostics, ultimately contributing to better patient care and outcomes.

**INDEX**

|  |  |  |
| --- | --- | --- |
| **Sr. No.** | **Table of Contents** | **Page No.** |
| 1 | Chapter 1: Introduction |  |
| 2 | Chapter 2: Services and Tools Required |  |
| 3 | Chapter 3: Project Architecture |  |
| 4 | Chapter 4: Project Outcome |  |
| 5 | Conclusion |  |
| 6 | Future Scope |  |
| 7 | References |  |
| 8 | Code |  |

**CHAPTER 1**

**INTRODUCTION**

* 1. **Problem Statement**

**Cardiovascular disease (CVD) remains a leading cause of mortality globally, with heart attacks contributing significantly to these fatalities. The World Health Organization reports that four out of five CVD deaths are due to heart attacks. Identifying individuals at risk of CVD is crucial for timely intervention and prevention efforts. Thus, the challenge at hand is to develop a predictive model capable of accurately assessing the risk of heart disease in patients.**

* 1. **Proposed Solution**

**In response to this challenge, we propose utilizing machine learning techniques, specifically Logistic Regression, for heart disease prediction. Logistic Regression is a powerful algorithm widely used for binary classification tasks, making it suitable for our objective of predicting the likelihood of CVD occurrence in individuals.**

* 1. **Feature**

**Data-Driven Approach: Artificial intelligence and machine learning rely on vast amounts of data to derive insights and make predictions. In the case of heart disease prediction, relevant data such as patient demographics, medical history, and lifestyle factors can be used to train the model.**

**Complex Pattern Recognition: Machine learning algorithms excel at identifying intricate patterns within data that may not be immediately apparent to human analysts. This capability is crucial for detecting subtle indicators of heart disease risk, which may involve numerous variables and interactions.**

**Adaptability and Scalability: AI and machine learning models can adapt to evolving datasets and new information. As more data becomes available or as medical knowledge advances, the model can be updated and refined to improve its predictive accuracy. Additionally, these models can be scaled up to handle large volumes of data efficiently.**

* 1. **Advantages**

**Early Detection: By analyzing various risk factors and biomarkers, machine learning models can identify individuals who are at a heightened risk of developing heart disease, even before symptoms manifest. Early detection allows for timely interventions and preventive measures to be implemented, potentially saving lives and reducing healthcare costs.**

**Personalized Risk Assessment: Machine learning algorithms can tailor predictions to individual patients based on their unique characteristics and medical history. This personalized approach enables healthcare providers to offer targeted interventions and lifestyle recommendations, optimizing patient outcomes.**

**Continuous Monitoring and Improvement: AI-powered systems can continuously monitor patient data and update risk assessments in real-time. This dynamic feedback loop facilitates ongoing refinement of the predictive model, ensuring that it remains accurate and relevant in clinical practice.**

**CHAPTER 2**

**SERVICES AND TOOLS REQUIRED**

**To undertake this Heart Disease Prediction project, the following services and tools are essential:**

**Python Programming Language: For implementing the machine learning models and data preprocessing.**

**Scikit-learn Library: To access various machine learning algorithms, including Logistic Regression, for predictive analysis.**

**Jupyter Notebook: For interactive coding and data visualization.**

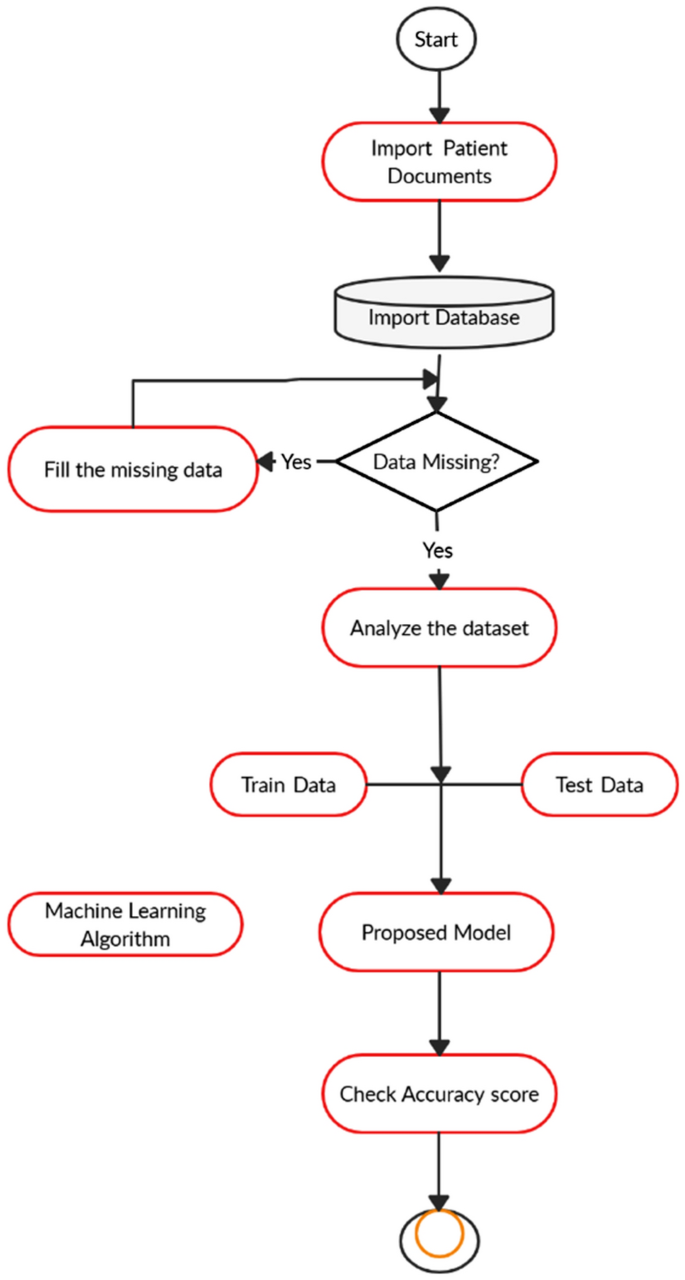
**Pandas and NumPy Libraries: For data manipulation and numerical computations.**

**Matplotlib and Seaborn Libraries: For visualizing data distributions and model performance metrics.**

**CHAPTER 3**

**PROJECT ARCHITECTURE**

**3.1 Architecture**

****

**Data Collection: Gather a comprehensive dataset containing relevant features associated with heart disease risk factors such as age, cholesterol levels, blood pressure, etc.**

**Data Preprocessing: Cleanse and preprocess the dataset by handling missing values, scaling numerical features, and encoding categorical variables.**

**Model Training: Utilize Logistic Regression algorithm to train the predictive model on the preprocessed dataset.**

**Model Evaluation: Assess the performance of the trained model using appropriate evaluation metrics such as accuracy, precision, recall, and F1-score.**

**Deployment: Deploy the trained model to make predictions on new, unseen data, enabling healthcare practitioners to identify individuals at risk of heart disease.**

|  |  |  |
| --- | --- | --- |
|  |  |  |

**CHAPTER 4**

**PROJECT OUTCOME**

**Through the implementation of Logistic Regression, the project aims to provide accurate predictions of heart disease risk for individuals based on their demographic and clinical characteristics. By leveraging AI and ML techniques, healthcare professionals can efficiently identify high-risk patients and tailor preventive interventions to mitigate the incidence of cardiovascular events.**

**CONCLUSION**

**The utilization of AI and ML in predicting heart disease risk demonstrates the potential for technology to revolutionize healthcare practices. By harnessing the power of data-driven insights, healthcare providers can proactively address cardiovascular health, ultimately leading to improved patient outcomes and reduced healthcare costs.**

**FUTURE SCOPE**

**The project lays the foundation for further research and development in the field of predictive analytics for cardiovascular diseases. Future endeavors may involve the integration of additional machine learning algorithms, feature engineering techniques, and real-time data streaming for more accurate and timely predictions. Moreover, collaborations with healthcare institutions can facilitate the deployment of predictive models in clinical settings, enabling personalized risk assessment and preventive care.**

**REFERENCES**

1. Project Github link, RamarBose , 2024

2. Project video recorded link (youtube/github), RamarBose , 2024

3. Project PPT & Report github link, RamarBose , 2024

**CODE**

**Please Provide Code through Git Hub Repo Link**

Github link :

<https://github.com/au422621105020/Heart-disease-prediction.git>

Project ppt link :

<https://docs.google.com/presentation/d/1P1FRdlD6_iq5ZqMfQ6lyFMdX2zfu5R8F/edit?usp=sharing&ouid=106213893944030120827&rtpof=true&sd=true>

Demo video link :

<https://drive.google.com/file/d/1wlP07fDWau98ofEPX0a0uPQWykEwbfek/view?usp=sharing>