Date:- 5,Apr.’2024

|  |  |
| --- | --- |
| iiitu-logo-large-final-1000x1000.jpg | **Indian Institute of Information Technology, Una [HP]**  An Institute of National Importance under MoE  **Saloh, Una (HP) - 177209**    **Website:** www.iiitu.ac.in |

**PRACTICUM REVIEW II**

**CSL306**

|  |  |  |  |
| --- | --- | --- | --- |
| **Student Name** | Keshav Agarwal | **Roll No.** | 22136 |
| **Batch No.** | B-26 | **Semester** | IV |
| **Branch** | CSE | **Supervisor** | Dr. Shivdutt Sharma |

1. **Title of the Project :-**  
   “HeartGuard” Predictive Modelling for Cardiovascular Wellness
2. **Introduction :-**

* HeartGuard focuses on the development of a robust machine learning model designed to predict heart disease risks. With a growing global concern for cardiovascular health, early detection and preventative measures are essential.
* Heart disease is a prevalent and life-threatening health condition that requires early detection for effective treatment. Machine learning models have shown promise in predicting heart disease, offering a proactive approach to healthcare.
* This machine learning (ML) model is used for predicting heart disease based on a set of relevant features. The objective is to assist healthcare professionals in early detection and proactive management of cardiovascular conditions.
* The model utilizes a dataset containing patient information and associated heart disease outcomes to train and validate its predictive capabilities.

1. **Problem Definition :-**

This project endeavours to create a machine learning model to forecast the probability of heart disease in individuals by analysing diverse health-related attributes. By leveraging data-driven insights, such as medical history, lifestyle factors, and physiological indicators, the aim is to develop a predictive tool aiding in early detection and intervention for improved cardiovascular health management.

1. **Objectives :-**

* To collect a comprehensive dataset containing relevant health features.
* To collect information regarding various features related to heart.
* To explore and analyse the dataset to understand the relationships between features and the target variable (heart disease).
* To design and implement a machine learning model for heart disease prediction.
* To evaluate the model's performance using appropriate metrics.

1. **Skillset additionally required to solve the problem :-**

* Machine Learning :-

Machine learning is a field of artificial intelligence that uses data to enable computers to learn and improve performance autonomously.

* Feature Engineering :-

Feature engineering involves selecting, transforming, and creating input variables to improve model performance and extract meaningful patterns from data.

* Data Visualization :-

Data visualization is the graphical representation of data to uncover insights, patterns, and trends for better understanding and decision-making.

* Domain Knowledge (Healthcare) :-

Domain knowledge in healthcare involves understanding medical practices, terminology, regulations, and patient needs to develop effective solutions and innovations.

* Model Evaluation :-

Model evaluation assesses the performance of machine learning models using metrics like accuracy, precision, recall, and F1-score for validation.

* Ethics and Privacy Awareness :-

Ethics and privacy awareness ensures responsible data handling, protection of individual rights, and mitigation of biases in AI applications.

1. **Timeline to achieve the skillset :-**

* Almost 16 weeks to complete the project.

1. **Block schematic :-**

Preprocessing of Data

Model Selection

Dataset

Model Training

Model Testing

Accuracy Prediction

Model Deployment

1. **Weekly milestones :-**

|  |  |
| --- | --- |
| **Week** | **Major Activities to be Completed** |
| Week 1 | Domain Exploration for ideas. |
| Week 2 | Idea Exploration. |
| Week 3 | Exploration different platforms for collecting data. |
| Week 4 | Dataset filtration. |
| Week 5 | Explore the Basics of Python and ML. |
| Week 6 | Explore the Basics of ML Algorithms. |
| Week 7 | Training of the model. |
| Week 8 | Training of the model. |
| Week 9 | Test and tune model to increase accuracy and finalize model. |
| Week 10 | Testing and debugging. |
| Week 11 | Finalize model after testing. |
| Week 12 | Deployment. |
| Week 13 | Ethical considerations. |
| Week 14 | Updates and improvements. |
| Week 15 | Updates and improvements. |
| Week 16 | Finalize all aspects of the project. |

1. **Completed Milestones :-**

* Idea exploration:-

Developing a machine learning model for predicting heart disease, aiming to enhance early detection and intervention strategies.

* Dataset Exploration:-

The dataset has been selected for the heart disease prediction model from Kaggle.

Source of the Dataset :-

<https://www.kaggle.com/datasets/rashikrahmanpritom/heart-attack-analysis-prediction-dataset>

* Basic Python and Basic Libraries Learning:-

Successfully completed learning basic Python programming and essential libraries, laying a solid foundation for further skill development.

Libraries like Matplotlib, NumPy, Pandas etc.

* Exploration of Machine Learning Models Theory :-

ML Models like KNN (K-nearest Neighbour), RandomForest & Decision Tree has been explored.

* Confusion Matrix:-

Utilized confusion matrix to evaluate the performance of classification models, aiding in understanding prediction accuracy and error distribution.

* Training and Testing of the Model:-

Conducted model training and testing, partitioning data to ensure robust evaluation and optimization of predictive performance.

* Selection of Model for Project:-

Selected the most suitable model for the project based on rigorous evaluation of performance metrics and domain-specific requirements.

1. **Milestones to be Completed :-**

* Searching better model or neural networks for the Prediction Model:-

Investigating advanced models, including neural networks, to enhance the prediction model, slated for implementation in the near future.

* Checking Ethical considerations:-

Conducting thorough examination of ethical considerations to ensure the development and deployment of the model align with societal values.

* Finalizing the Model;-

Completing the finalization process of the model, incorporating feedback and optimizations for optimal performance and reliability.

* Deployment of the Model:-

Initiating the deployment phase to make the model accessible and operational, facilitating its real-world application and impact.

1. **Expected Challenges :-**

* Imbalanced Data :- Unequal representation of positive and negative cases.
* Data Quality and Cleaning :- Noisy or inaccurate data.
* Feature Selection :- Identifying the most relevant features.
* Overfitting :- Model fitting the training data too closely.
* Ethical Considerations :- Ensuring fairness and avoiding bias.
* Continuous Model Monitoring and Updating.

1. **References :-**

* D. Zhang, Y. Chen, Y. Chen, S. Ye, W. Cai, and M. Chen, “An ECG Heartbeat Classification Method Based on Deep Convolutional Neural Network,” Journal of Healthcare Engineering, vol. 2021, 2021, doi: 10.1155/2021/7167891
* B. Deepak Kumar, S. Yellaram, S. kothamasu, S. Puchakayala, and A. Professor, “Heart Stroke Prediction using Machine Learning,” 2021. [Online]. Available: www.ijcrt.org

**Name and Signature of Student**

**Name and Signature of Supervisor**