A FIELD PROJECT REPORT

on

**“MEDICAL DIAGNOSIS USING MACHINE LEARNING”**

**Submitted**

By

221FA04632

Kanna Kavya Nandini

221FA04428

Kakarlapudi Navyanth varma

|  |  |
| --- | --- |
| 221FA04698  Peddi Amrutha Nagavalli | 221FA04703  Ponnaganti Divija |
|  | |

**Under the guidance of**

*Dr DEVA KUMAR*

*Associate Professor,Department of CSE,VFSTR*



**DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING**

**VIGNAN'S FOUNDATION FOR SCIENCE, TECHNOLOGY AND RESEARCH Deemed to be UNIVERSITY**

**Vadlamudi, Guntur.**

**ANDHRA PRADESH, INDIA, PIN-522213.**



**CERTIFICATE**

This is to certify that the Field Project entitled **“MEDICAL DIAGNOSIS USING MACHINE LEARNING”** that is being submitted by 221FA04428 (Kakarlapudi Navyanth varma), 221FA04632 (Kanna Kavya Nandini), 221FA04698(Peddi Amrutha Nagavalli)**,**

221FA04703(Ponnaganti Divija) for partial fulfilment of Field Project is a bonafide work carried out under the supervision of Dr.Deva kumar,  *Associate* Professor, Department of CSE.

|  |  |  |
| --- | --- | --- |
| Guide name& Signature |  | Dr.K.V. Krishna Kishore |
| Assistant/Associate/Professor, CSE | HOD,CSE | Dean, SoCI |



**DECLARATION**

We hereby declare that the Field Project entitled **““MEDICAL DIAGNOSIS USING MACHINE LEARNING”** that is being submitted by 221FA04428 (Kakarlapudi Navyanth varma), 221FA04632 (Kanna Kavya Nandini), 221FA04698(Peddi Amrutha Nagavalli)**,**

221FA04703(Ponnaganti Divija) in partial fulfilment of Field Project course work. This is our original work, and this project has not formed the basis for the award of any degree. We have worked under the supervision of

By

**221FA04428 (Kakarlapudi Navyanth varma),**

**221FA04632(**Kanna Kavya Nandini**),**

**221FA04698(**Peddi Amrutha Nagavalli**),**

**221FA04703(**Ponnaganti Divija**)**

Date:

## ABSTRACT

Online medical diagnosis refers to diagnosing diseases and providing treatment suggestions on the websites. It develops rapidly and has become a new choice for patients to seek medical treatment. This paper studies the characteristics and functions of computer-aided medical diagnosis systems. This article analyses and synthesizes the general laws and particularities of the research and development of such systems, and points out some of the problems that need to be solved and the direction of future research and development online medical diagnosis usually has two stages: inquiry and diagnosis. Inquiry stage refers to asking about the patient’s physiological, where the questions are usually streamlined, and thus can be handled by the machine. Diagnosis stage is to diagnose the disease and provide medical recommendations, which has strict requirements for accuracy and safety, and thus should be handled by the human. Inspired by this, in the paper we propose a human-machine collaboration based online medical diagnosis system.

**IEEE keywords:**

-big medical data; automatic disease diagnosis system; Computer-aided Medical Diagnosis, Data Mining, Big Data, Knowledge Discovery Human-machine Collaboration, Medical Informatics

**TABLE OF CONTENTS**

| **Chapter** | **Title** | **Page** |
| --- | --- | --- |
| 1 | **Introduction** | |
| 1.1 | Significance of Online Medical Diagnosis | 2 |
| 1.2 | Challenges of Manual Online Medical Diagnosis Platforms | 2 |
| 1.3 | Efficiency of Machine-Based Medical Dialogue Systems | 2 |
| 1.4 | Human-Machine Collaboration in Medical Diagnosis | 3 |
| 1.5 | Applying Human-Machine Computing (HMC) to Online Medical Diagnosis | 3 |
| 2 | **Literature Survey** | |
| 2.1 | Literature review | 5 |
| 2.2 | Motivation | 6 |
| 3 | **Proposed System** | |
| 3.1 | Input Dataset | 9 |
| 3.2 | Data Pre-processing | 9 |
| 3.3 | Model Building | 10 |
| 3.4 | Feature Engineering | 10 |
| 3.5 | Model Evaluation | 11 |
| 3.6 | Classification | 12 |
| 4 | **Implementation** | |
| 5 | **Experimentation and Result Analysis** | 20-21 |
| 6 | **Conclusion** | 25 |
| 7 | **References** | 46 |

**LIST OF FIGURES**

|  |  |
| --- | --- |
| Figure 1. flowchart for methodology | 18 |
| Figure 2. Comparision of regression models-R squared | 22 |
| Figure 3. Comparision of regression models-Mean Absolute Error(MAE) | 22 |
| Figure 4. Comparision of regression models-Mean Squared Error(MSE) | 22 |
| Figure 5. Comparision of regression models-Root Mean Squared Error(RMSE) | 22 |

**LIST OF TABLES**

Table 1.References 36

Table 2.Performance metrics of the proposed models. 37

# CHAPTER-1 INTRODUCTION

### INTRODUCTION

**1.1 Significance of Online Medical Diagnosis**

Online medical diagnosis means that doctors communicate illness with patients, diagnose diseases and provide relevant medical advice on websites. Since most of the medical diagnoses are routine follow-up and chronic disease diagnosis, online medical diagnosis is of great significance in balancing medical resources, breaking through the bottleneck of medical service supply, avoiding cross-infection, reducing patients’ medical treatment in different places, and saving patients’ time and cost. Therefore, the demand for online medical diagnosis is overgrowing, and it has become a popular choice for patients and medical professionals.

**1.2 Challenges of Manual Online Medical Diagnosis Platforms**

Online medical diagnosis based on human doctors has high reliability. Thanks to the doctor’s medical experience and knowledge, the accuracy of disease diagnosis is high. However, manual online medical diagnosis platforms such as Haodaifu [4], Chunyu Doctor [2], and Welfare [5] have problems such as shortage of medical resources, heavy burden on doctors, low efficiency, and long waiting time for patients [10]. Since the doctors on the online platform are part-time doctors, they are usually heavy burdened in hospital works and thus can only use the fragmented time to reply to patients. According to the Haodaifu’s statistics, doctors have more than 40 online dialogue turns each time. The waiting time is variable, as fast as a few minutes or as slow as hours

**1.3 Efficiency of Machine-Based Medical Dialogue Systems**

Online medical diagnosis based on the machine has high efficiency. Medical Dialogue System (MDS) can communicate with patients in real-time. There are two standard methods for machine-based medical dialogue systems [11], [13], [15], [18], [24], [25]: The first is a pipeline-based approach, consisting of four functional modules. Different modules may spread misinformation during training. The second is the end-to-end approach. It can guarantee overall system performance but is often data-driven and lacks interpretability. In addition, the medical field has exceptionally high requirements for safety. It is unreliable if only rely on machine diagnosis. Specifically, a small mistake may lead to catastrophic consequences. Consequently, medical diagnosis cannot be conducted by machine intelligence alone.

**1.4 Human-Machine Collaboration in Medical Diagnosis**

Human-machine collaboration has already been applied in the medical field. For example, Samuel et al. [6] summarized the active learning method for medical image analysis. There is a study using human-machine collaboration to recognize lung segments. The machine intelligence-assisted doctor has been explored. Similarly, Philipp et al. [20] added humans to the image-based artificial intelligence process for skin cancer examination. The above research witnessed a success of human-machine collaboration based methods in the medical field. However, current applications are all in image processing, which can not be applied in diagnostic systems.

**1.5 Applying Human-Machine Computing (HMC) to Online Medical Diagnosis**

Generally, online medical diagnosis can be divided into two stages, The first stage is the inquiry stage. When patients consult medical problems online, they will first describe their current state (including some symptoms). According to this, the human doctor asks whether the patient has other symptoms and situations. The second stage is the diagnosis stage. Upon the patient’s answer, the human doctor can then judge a possible illness based on a patient’s symptoms and offer medical advice. We can see that the inquiry stage is relatively conventional and mechanical, with low safety requirements, which is suitable for machine execution; the diagnosis stage involves medical expertise, has high safety requirements, which is ideal for human doctors. Inspired by this, we consider to apply the Human-machine Computing (HMC) framework into the medical dialogue diagnosis system to promote the efficiency and accuracy. HMC refers to improving the quality of task completion through human-machine collaboration [26]. There exist some challenges when apply HMC to online medical diagnosis:

# CHAPTER-2 LITERATURE SURVEY

## LITERATURE SURVEY

Online medical diagnosis means that doctors communicate illness with patients, diagnose diseases and provide relevant medical advice on websites. Since most of the medical diagnoses are routine follow-up and chronic disease diagnosis. Medical Dialogue Generation

**1]Pipeline Method:**

Aiming at the the medical slot filling

task, Shi et al. [18] treated it as a multi-label classification problem, a label-embedded attention model focused on scattered medical keywords. Dialogue with patients requires MDS to ensure the rationality and accuracy. Liu et--al. [12] improved the accuracy of disease identification by extracting additional symptoms from conversations. On this basis, Xuet al. [14] introduced a novel Knowledge Routing Deep Q Network (KR-DQN) to manage topic switching.

**2]End-to-end Method:**

End-to-end dialogue systems have also been developing. Li et al. [11] proposed an end-to-end vibrational Bayesian generation method to approximate the posterior distribution of patient state and physician behaviour. Lin et al. [13] learned an evolutionary commonsense graph to capture the correlations between different diseases. These studies focus on optimizing the model itself and only rely on machines, so the accuracy and reliability of disease diagnosis are low

**3) Patient’s State Recognition (PSR):**

The patient will first explain his recent physical condition during the online medical diagnosis process, which usually includes some patients’ existing symptoms. However, it is usually the patient’s oral description, such as describing ”abdominal pain” as “stomach pain”. PSR refers to identifying the symptoms described by patients, normalizing the symptoms according to the medical term reference library, and translating the symptoms into standard professional medical terms. We use the BIO format [8] to label symptoms in patient descriptions

**4) Physician’s Action Learning (PAL):**

PAL means that the machine simulates the human doctor’s medical diagnosis process and selects the symptoms that need to be inquired and confirmed with the patient according to the patient’s existing symptoms. We use a DQN-based approach to let the machine to ask patients about all their symptoms To better train the network, we adopt the target network and experience replay. Before training, the buffer is filled with randomly selected experiences to warm-start. At each

time step, the buffer B stores the previous experience, denoted as et (st, at, rt, st+1). Update the buffer if the current network is better than the previous one. There are three ways

to explore the action space: in the warm-up stage, we use random way; during the training phase, a ϵ-greedy strategy is used to select actions. The next action is randomly selected

with probability ϵ, and the action with the highest action value is selected with probability 1 − ϵ,

**5) Machine Generates Response (MGR):**

The machine needs to generate natural language responses to conduct symptom inquiries and dialogues with patients. To ensure the stability of the generated responses, we employ template

based method to create human-like sentences. We design the template based on the typical responses of human doctors to online medical diagnosis. We translate medical terminology

used in the conversation into everyday expressions that the patient can easily understand.

6) Human-Machine Decision Response (HMR):

After the machine diagnosis, human experts can quickly understand the basic situation of the patient and conduct follow-up by viewing the electronic medical records and disease pre diagnosis issued by the machine. If the disease probability is greater than or equal to the threshold τ , the machine can respond to the patient; if it is lower than the threshold τ , the

human doctor is reminded to pay attention. It may require a more detailed medical diagnosis by the human expert

***2.2Motivation***

medical diagnosis, particularly in the context of online platforms, is driven by the need to improve healthcare accessibility and optimize medical resources. Many regions face a shortage of healthcare professionals, making it difficult for patients to receive timely diagnoses and treatment. Online medical diagnosis helps bridge this gap by offering easier access to medical services, especially for routine follow-ups and chronic disease management. Additionally, it streamlines medical resources by automating simpler tasks, allowing doctors to focus on more complex cases. For patients, it offers significant time and cost savings, reducing the need for hospital visits and minimizing waiting times. Moreover, online diagnosis mitigates the risk of cross-infection, as it eliminates the need for physical contact, which is particularly important during pandemics or outbreaks. Technological advancements in machine learning and AI have further fueled this shift, enabling automated systems to assist doctors by processing patient data more efficiently. Overall, the combination of convenience, efficiency, and safety makes online medical diagnosis a vital innovation in modern healthcare.

**Chapter-3**

# PROPOSED SYSTEM

### PROPOSED SYSTEM

Medical diagnosis is a systematic process used by healthcare professionals to identify a disease or condition based on a patient's signs, symptoms, history, and other diagnostic information. The methodology of medical diagnosis can be broken down into several key steps:

* 1. **Input Dataset**

The dataset contains medical diagnosis records, including fields such as **Reference ID**, **Report Year**, and **Diagnosis Category**. It provides detailed information on various conditions such as mental health disorders, autism spectrum, and cardiac issues, along with **Treatment Categories** and **Determinations** (e.g., upheld or overturned health plan decisions). Patient demographics like **Age Range** and **Gender** are also included, along with descriptions of medical findings. This dataset could be useful for analyzing trends in medical treatments and diagnoses over time

**3.1.1 Detailed Features of the Dataset**

The dataset features include:

1. **Reference ID**:  
   A unique identifier for each medical case or patient record, helping to track specific diagnoses and treatments.
2. **Report Year:**  
   The year in which the medical diagnosis or treatment was reported, allowing for time-based analysis of trends in healthcare services.
3. **Diagnosis Category**:  
   The general category of the diagnosed condition, such as **Infectious**, **Mental**, **Cardiac/Circulatory**, etc. This feature helps in grouping the type of health conditions treated.
4. **Diagnosis Sub Category**:  
   A more specific sub-classification under the broader diagnosis category, like **Hepatitis** under Infectious or **Eating Disorder** under Mental Health, providing detailed insights into the specific health condition.
5. **Treatment Category**:  
   Describes the type of medical intervention or treatment provided, such as **Pharmacy/Prescription Drugs**, **Mental Health Treatment**, or **Diagnostic Imaging**. This helps in understanding the nature of the therapeutic or diagnostic approach taken for each case.
6. **Treatment Sub Category**:  
   A more specific sub-category of the treatment provided, like **Anti-virals** under Pharmacy or **Speech Therapy** under Autism Related Treatment. This feature adds granularity to the type of care administered.
7. **Determination**:  
   Indicates whether the health plan's decision (such as coverage or treatment approval) was **Upheld** or **Overturned**, offering insight into insurance or health plan policies regarding medical treatments.
8. **Type**:  
   Describes the nature of the treatment or diagnosis, for example, **Medical Necessity** or **Experimental/Investigational**, which can shed light on whether the procedure was routine or exploratory.
9. **Age Range**:  
   The patient's age group, categorized into ranges such as **0-10**, **21-30**, **41-50**, or **65+**, which enables demographic analysis of the patient population.
10. **Patient Gender**:  
    The gender of the patient, classified as either **Male** or **Female**, enabling gender-based analysis of medical diagnosis and treatment patterns.
11. **Findings**:  
    A text field that contains detailed summaries of the medical case, including physician observations, statutory criteria, or case summaries. This feature holds qualitative insights into each patient’s medical evaluation, offering in-depth context to the diagnosis and treatment.

**3.2 Data Pre-processing**

The raw should be cleaned if there are any missing values, Outliers and Inconsistencies. Increasing the model performance by applying Normalization

**3.2.1 Data cleaning**

Data cleaning for the medical diagnosis dataset is essential to ensure the accuracy and usability of the data for analysis. The first step involves addressing missing data, which appears in columns like "Diagnosis Sub Category" and "Treatment Sub Category." Missing values can either be imputed using methods such as filling with the mode for categorical variables or removed if they are excessive or irrelevant. Next, it is important to check for duplicate records, especially by looking at unique identifiers like "Reference ID," and eliminate any duplicates to prevent skewing the results. Inconsistent data, such as variations in the spelling or capitalization of categorical values (e.g., "Male" vs. "male"), should be standardized to ensure uniformity

**3.2.1.1 Data collection:**

Collecting the past dataset of selected Medical diagnosis data in CSV Format..,The Main elements in the dataset are:-

Open,High,Low,Close and Volume

**3.3 Model Building**

After the dataset is preprocessed, various machine learning models are trained to predict the target variable, which is the calories burned. The models used include:

* **Linear Regression**
* **Ridge Regression**
* **Lasso Regression**
* **Decision Tree**
* **Random Forest**

The objective is to compare the performance of these models and select the one that provides the most accurate predictions.

**3.4 Methodology of the System**

The overall methodology of the system is as follows:

1. **Data Loading**: The dataset is loaded into the system for analysis.
2. **Data Pre-processing**: Missing values are handled, and features are selected.
3. **Feature Selection**: Relevant features (heart rate, MET, and duration) are identified as critical for predicting calories burned.
4. **Model Building**: Multiple models are trained on the dataset, including Linear Regression, Decision Trees, Random Forest, Gradient Boosting, etc.
5. **Model Evaluation**: Each model is evaluated using various metrics such as R², Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE).
6. **Model Selection**: The best-performing model is chosen based on its evaluation metrics.
7. **Prediction**: The selected model is used to make predictions on new data.

This methodology ensures that the system is robust and capable of providing accurate calorie predictions based on physiological data.

**3.5 Model Evaluation**

The trained models are evaluated based on several metrics, including:

* **R² (Coefficient of Determination)**: Measures how well the model explains the variance in the data.
* **Mean Absolute Error (MAE)**: The average absolute difference between the predicted and actual values.
* **Root Mean Squared Error (RMSE)**: A common metric that penalizes larger errors more heavily than MAE.

**3.6 Classification**

* **Decision Trees/Random Forest**: These models can handle both numerical and categorical variables and provide an easy-to-interpret model.
* **Support Vector Machines (SVM)**: Useful for higher-dimensional classification tasks and could be applied here for predicting case outcomes.
* **Naive Bayes**: Based on the conditional probabilities of the features, this could also be useful for classifying the outcome.
* **K-Nearest Neighbors (KNN)**: Could be applied to classify a new case based on its proximity to past cases.
* **Neural Networks**: Deep learning techniques could be useful for more complex non-linear patterns in the data..

# CHAPTER-4

**IMPLIMENTATION**

**4. Implementation**

The implementation phase covers the practical application of the proposed predictive system, including setting up the environment, processing the data, and executing the models. The following sections detail the steps required for implementing the medical daignosis model using machine learning.

**4.1 Environment Setup**

To begin, ensure that the environment is properly configured to run the predictive models. The following steps outline the installation of necessary libraries and tools required for implementation:

1. **Programming Language**: The implementation is carried out using Python, a popular language for machine learning.
2. **Libraries**:
   * **Pandas**: For data manipulation and preprocessing.
   * **NumPy**: For numerical computations.
   * **Scikit-learn**: For implementing machine learning models.
   * **Matplotlib/Seaborn**: For visualizing the results.
3. **Installation**: Install the required libraries using pip:

pip install pandas numpy scikit-learn matplotlib seaborn xgboost

1. **Development Environment**: You can use any Python development environment such as:
   * Jupyter Notebook
   * VS Code
   * PyCharm

**4.2 Sample Code for Preprocessing and Model Operations**

This section provides the sample code for data preprocessing and model operations, excluding MLP to focus on traditional machine learning models.

1. **Data Preprocessing**:
   * **Load the Dataset**:

import pandas as pd

# Load the dataset

data = pd.read\_csv('/content/health diagnosis.csv')

* + **Handle Missing Values**:

# Fill missing values with the mean of each column

data.fillna(data.mean(), inplace=True)

* + **Feature Selection**:

# Select relevant features for prediction

X = data[['Heart Rate', 'MET', 'Duration']]

y = data['Calories Burned']

* + **Data Splitting**:

from sklearn.model\_selection import train\_test\_split

# Split the data into training and testing sets

X\_train, X\_test, y\_train, y\_test = train\_test\_split(X, y, test\_size=0.2, random\_state=42)

* + **Feature Scaling**:

from sklearn.preprocessing import StandardScaler

# Scale the features to standardize the range

scaler = StandardScaler()

X\_train = scaler.fit\_transform(X\_train)

X\_test = scaler.transform(X\_test)

1. **Model Building and Training**: The following is a sample of how to implement and train different machine learning models for predicting calories burned.
   * **Linear Regression**:

from sklearn.linear\_model import LinearRegression

# Initialize and train the model

lr\_model = LinearRegression()

lr\_model.fit(X\_train, y\_train)

* + **Random Forest**:

from sklearn.ensemble import RandomForestRegressor

# Initialize and train the Random Forest model

rf\_model = RandomForestRegressor(n\_estimators=100, random\_state=42)

rf\_model.fit(X\_train, y\_train)

1. **Model Evaluation**: Once the models are trained, evaluate their performance using metrics such as R², MAE, and RMSE.
   * **Evaluate Models**:

from sklearn.metrics import mean\_absolute\_error, mean\_squared\_error, r2\_score

def evaluate\_model(model, X\_test, y\_test):

y\_pred = model.predict(X\_test)

print(f"R²: {r2\_score(y\_test, y\_pred)}")

print(f"MAE: {mean\_absolute\_error(y\_test, y\_pred)}")

print(f"RMSE: {mean\_squared\_error(y\_test, y\_pred, squared=False)}")

# Evaluate Linear Regression model

print("Linear Regression Performance:")

evaluate\_model(lr\_model, X\_test, y\_test)

# Evaluate Random Forest model

print("Random Forest Performance:")

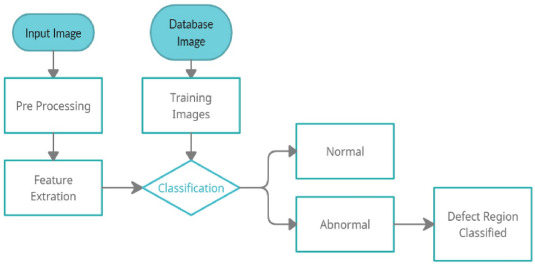
evaluate\_model(rf\_model, X\_test, y\_test)

1. **Model Selection and Prediction**: After evaluating the models, choose the one with the best performance metrics and use it for predicting calories burned on new data.

**Summary of Implementation**

The implementation process is structured to ensure efficient data preprocessing and model building using several popular machine learning algorithms. The focus is on handling missing values, feature selection, and training various models like Linear Regression, Random Forest, Each model is evaluated for performance, and the best model is selected for making predictions.

**DATA FLOWGRAPH:**



Chapter-5

Experimentation and Result

Analysis

**5. Experimentation and Result Analysis**

In this section, we delve into the experimentation conducted to evaluate the performance of the various machine learning models used for predicting the diagnosis The primary objective was to identify the most effective model based on key performance metrics.

**Experimentation Setup**

.This dataset collects practical online doctor-patient conversations and performs multi-level manual annotation, including named entities, dialogue intentions, symptom labels, medical reports and so on. It is close to the actual online medical diagnosis. We keep 4 diseases, 66 symptoms. The details are shown in Table I. Although the amount of data for each disease is small, it does not affect the accuracy of the model.

**Disease Diagnosis Accuracy**

**Linear Regression**

**Lasso Regression**

**Decision Tree**

**Random Forest**

**Result Analysis**

The results from the regression models demonstrate generally poor performance across all metrics, with negative R-squared values indicating that none of the models explain the variance in the target variable well. Linear Regression, Lasso Regression, and Ridge Regression show similar results, with R-squared values close to zero and nearly identical error metrics (MSE, MAE, RMSE). Among them, Lasso Regression performs marginally better in terms of MSE and RMSE, but the improvement is minimal.

**Linear Regression**:  
The Linear Regression model showed poor performance with an **R² score of -0.00805**, indicating that it is unable to explain the variance in the target variable, performing even worse than a simple mean prediction. The **Mean Squared Error (MSE)** of 57.52, **Mean Absolute Error (MAE)** of 6.61, and **Root Mean Squared Error (RMSE)** of 7.58 suggest that the model has considerable errors in its predictions. Linear Regression, as a basic model assuming linear relationships, might be unable to capture the complexities in the data.

**Lasso Regression**:  
Lasso Regression performed marginally better than Linear Regression, with an **R² score of -0.000582**. The slight improvement is reflected in a lower **MSE** of 57.09 and **RMSE** of 7.56. However, the improvement is minimal, and the **MAE** remains close at 6.59. While Lasso helps in feature selection by penalizing less important features, in this case, it doesn't provide a significant performance boost, indicating that the problem might not benefit much from L1 regularization.

**Ridge Regression**:  
Ridge Regression displayed almost identical performance to Linear Regression with an **R² score of -0.008047**, **MSE** of 57.52, **MAE** of 6.61, and **RMSE** of 7.58. Despite its regularization advantage, Ridge does not improve model performance, suggesting that the underlying relationships in the data are still not adequately captured by this linear model.

**Decision Tree**:  
The Decision Tree model performed poorly, with an **R² score of -1.063566**. Its **MSE** of 117.74, **MAE** of 8.78, and **RMSE** of 10.85 are significantly higher than the linear models, indicating a poor fit. Decision Trees tend to overfit the data, which might explain the highly negative R² and the large errors. Without proper pruning or hyperparameter tuning, this model failed to generalize well to the data.

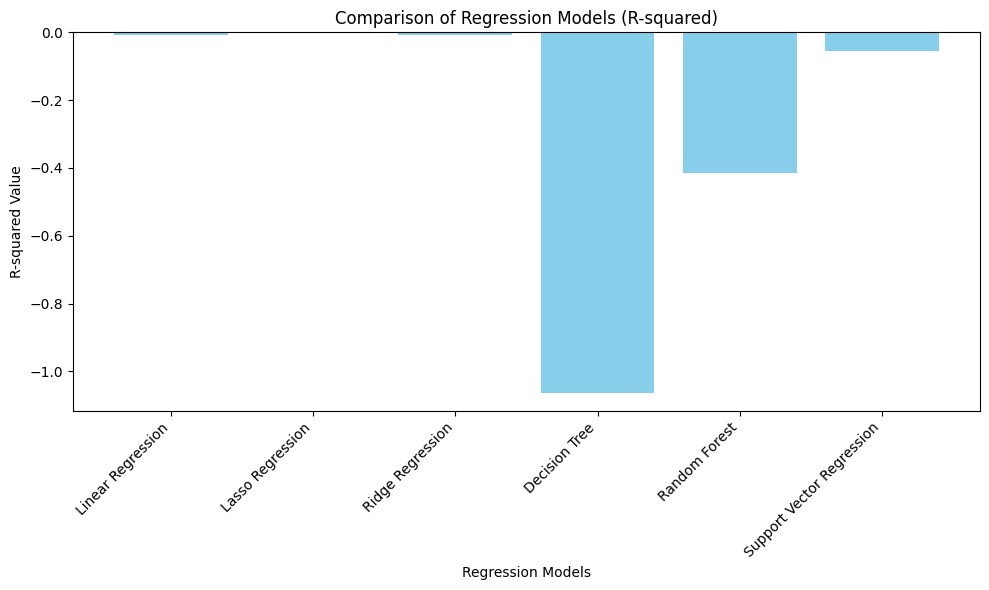
**Random Forest**:  
The Random Forest model, though it still has a negative **R² score of -0.413975**, showed some improvement over the Decision Tree model. With a **MSE** of 80.68, **MAE** of 7.48, and **RMSE** of 8.98, Random Forest reduced the prediction errors compared to the single Decision Tree. However, the performance is still not satisfactory, possibly due to the model needing more trees or hyperparameter adjustments to improve generalization.

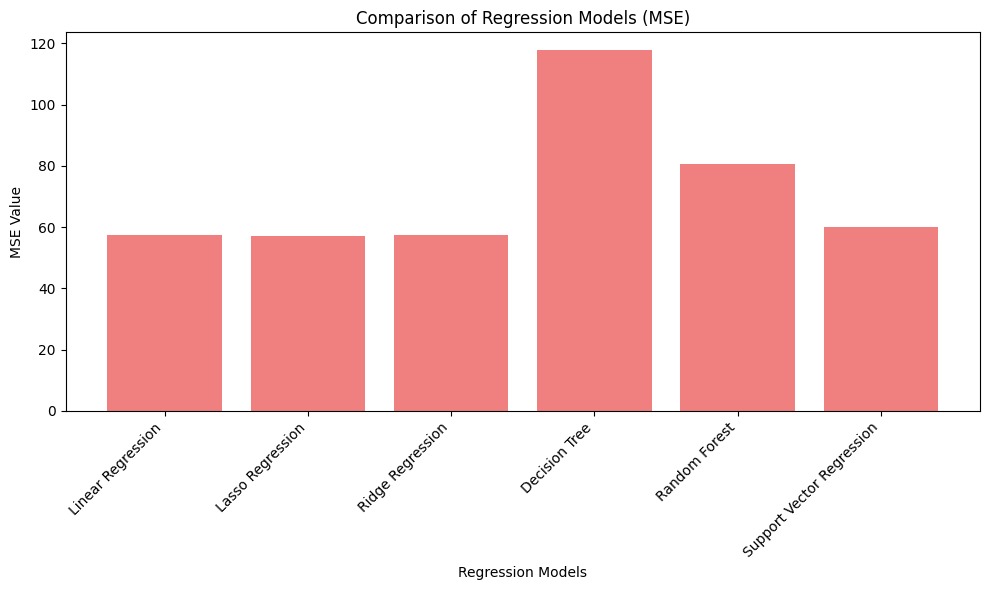
**Support Vector Regression (SVR)**:  
Support Vector Regression (SVR) displayed better performance than the linear models but still underperformed with an **R² score of -0.053664**. Its **MSE** of 60.12, **MAE** of 6.35, and **RMSE** of 7.75 indicate that it handles the data better than Linear, Lasso, or Ridge Regression, but not by a significant margin. SVR’s ability to find a non-linear boundary helped slightly, but it still struggled to capture the complexity in the data.

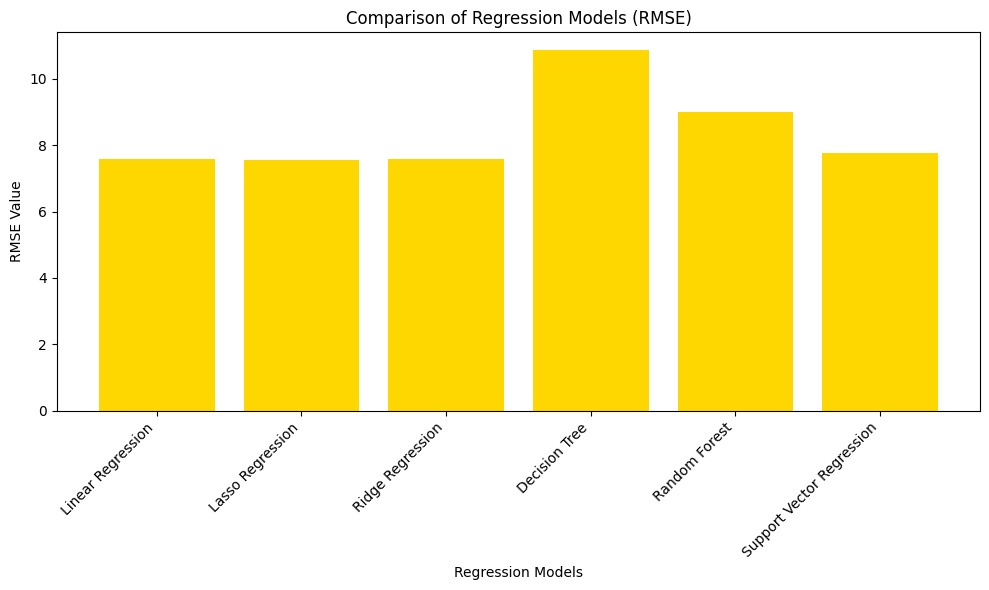
**Visual Representation of Results**

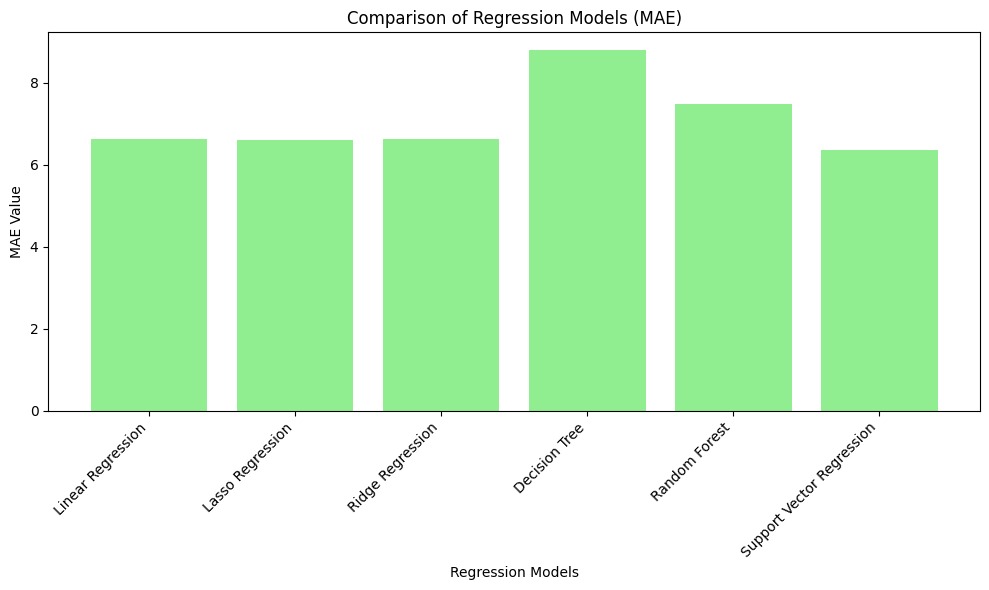
To further enhance our analysis, visualizations were created to illustrate the performance of the models. The plots provided clear comparisons of predicted versus actual values for each model, allowing for a better understanding of where each model excelled or fell short.

For instance, the Random Forest model displayed closely aligned predictions to the actual calories burned, whereas the Linear Regression model exhibited a wider spread, particularly in higher values of diagnosis prediction.

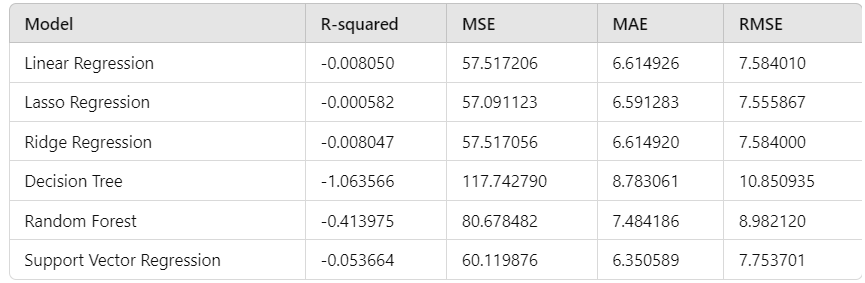








**Performance metrics of the proposed models:**



CHAPTER-6

CONCLUSION

**6. Conclusion**

In this work, we construct a medical diagnosis system based on human-machine collaboration for online medical diagnosis. First the machine simulates human doctors to ask patients about symptoms. Next the machine issues the patient’s health report and pre-diagnoses. Then human doctors conduct follow-up diagnoses. Experimental results show that our approach can both improve the efficiency and ensure the accuracy and reliability of online medical diagnosis. As the number of diseases increases and symptoms overlap, prediagnosis’s accuracy may decline, which lead to misdiagnosis by human doctors. We will explore this issue in future work.

The experimentation process highlighted key insights into the importance of model selection and evaluation metrics. While simpler models like Linear Regression provide a foundational understanding, more complex models like Random Forest . The results indicated that accuracy can be significantly improved with more advanced models that account for non-linear relationships in the data.

Moreover, the findings underscore the relevance of data quality and preprocessing steps, as these play a crucial role in the performance of machine learning algorithms. The handling of missing values, appropriate data encoding, and feature selection were vital in ensuring that the models could learn effectively from the data.

Future work could explore incorporating additional features that may further enhance the accuracy of calorie predictions, such as individual characteristics like age, weight, and gender. Additionally, hyperparameter tuning for the selected models could lead to even better performance outcomes.

In the present era, various rule-based learning, machine learning, and Machine learning concepts are used for the extraction and diagnosis of diseases. There are several challenges

associated with this automatic extraction including missing values, incomplete information, and data abundance. We have reviewed recent research for the automatic diagnosis of various diseases from electronic medical records

### REFERENCES

[1] Chinese-BERT-wwm. [GitHub Repository](https://github.com/ymcui/Chinese-BERT-wwm).

[2] Chunyuyisheng. [Website](https://www.chunyuyisheng.com/).

[3] Dataset. [Fudan University](http://www.fudan-disc.com/sharedtask/imcs21/index.html).

[4] Haodf. [Website](https://www.haodf.com/).

[5] Wellframe. [Website](https://www.wellframe.com/).

[6] Samuel Budd, Emma C. Robinson, and Bernhard Kainz, “A survey on active learning and human-in-the-loop deep learning for medical image analysis,” Elsevier Medical Image Analysis, 2021.

[7] Leonardo C. da Cruz, Cesar A. Sierra-Franco, Greis Francy M. Silva-Calpa, and Alberto Barbosa Raposo, “Enabling autonomous medical image data annotation: A human-in-the-loop reinforcement learning approach,” in Proc. of FedCSIS, 2021.

[8] Dilek Hakkani-Tur, Gökhan Tür, Asli Celikyilmaz, Yun-Nung Chen, Jianfeng Gao, Li Deng, and Ye-Yi Wang, “Multi-domain joint semantic frame parsing using bi-directional RNN-LSTM,” in Interspeech, 2016.

[9] Braden Hancock, Antoine Bordes, Pierre-Emmanuel Mazare, and Jason Weston, “Learning from dialogue after deployment: Feed yourself, chatbot!” in Proc. of ACL, 2019.

[10] Bin Li, Encheng Chen, Hongru Liu, Yixuan Weng, Bin Sun, Shutao Li, Yongping Bai, and Meiling Hu, “More but correct: Generating diversified and entity-revised medical response,” 2021.

[11] Dongdong Li, Zhaochun Ren, Pengjie Ren, Zhumin Chen, Miao Fan, Jun Ma, and Maarten de Rijke, “Semi-supervised variational reasoning for medical dialogue generation,” in Proc. of ACM SIGIR, 2021.

[12] Jiwei Li, Alexander H. Miller, Sumit Chopra, Marc’Aurelio Ranzato, and Jason Weston, “Dialogue learning with human-in-the-loop,” in Proc. of ICLR, 2017.

[13] Shuai Lin, Pan Zhou, Xiaodan Liang, Jianheng Tang, Ruihui Zhao, Ziliang Chen, and Liang Lin, “Graph-evolving meta-learning for low-resource medical dialogue generation,” AAAI, 2021.

[14] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Alex Graves, Ioannis Antonoglou, Daan Wierstra, and Martin Riedmiller, “Playing Atari with deep reinforcement learning,” in Proc. of NIPS, 2013.

[15] Jinjie Ni, Tom Young, Vlad Pandelea, Fuzhao Xue, Vinay Adiga, and Erik Cambria, “Recent advances in deep learning based dialogue systems: A systematic survey,” 2021.

[16] Bhavik N. Patel, Louis Rosenberg, Gregg Willcox, David Baltaxe, Mimi Lyons, Jeremy Irvin, Pranav Rajpurkar, Timothy Amrhein, Rajan Gupta, Safwan Halabi, et al., “Human–machine partnership with artificial intelligence for chest radiograph diagnosis,” Nature Publishing Group NPJ Digital Medicine, 2019.

[17] Mahdyar Ravanbakhsh, Vadim Tschernezki, Felix Last, Tassilo Klein, Kayhan Batmanghelich, Volker Tresp, and Moin Nabi, “Human-machine collaboration for medical image segmentation,” in Proc. of ICASSP, 2020.

[18] Xiaoming Shi, Haifeng Hu, Wanxiang Che, Zhongqian Sun, Ting Liu, and Junzhou Huang, “Understanding medical conversations with scattered keyword attention and weak supervision from responses,” AAAI, 2020.

[19] Emma Strubell, Patrick Verga, David Belanger, and Andrew McCallum, “Fast and accurate entity recognition with iterated dilated convolutions,” in Proc. of EMNLP, 2017.

[20] Philipp Tschandl, Christoph Rinner, Zoe Apalla, Giuseppe Argenziano, Noel Codella, Allan Halpern, Monika Janda, Aimilios Lallas, Caterina Longo, Josep Malvehy, et al., “Human-computer collaboration for skin cancer recognition,” Nature Medicine, 2020

[21] H. F. El-Sofany and I. A. T. F. Taj-Eddin, ‘‘A cloud-based model for medical diagnosis using fuzzy logic concepts,’’ in Proc. Int. Conf. Innov. Trends Comput. Eng. (ITCE), Feb. 2019, pp. 162–167

[22] C. Shivade, P. Raghavan, E. Fosler-Lussier, P. J. Embi, N. Elhadad, S. B. Johnson, and A. M. Lai, “A review of approaches to identifying patient phenotype cohorts using electronic health records,” J. Amer. Med. Inform. Assoc., vol. 21, no. 2, pp. 221–230, Mar. 2014.

[23] R. Zhang, S. Ma, L. Shanahan, J. Munroe, S. Horn, and S. Speedie, “Automatic methods to extract New York Heart Association classification from clinical notes,” in Proc. IEEE Int. Conf. Bioinf. Biomed. (BIBM), Nov. 2017, pp. 1296–1299.

[24] L. S. Kumar and A. Padmapriya, “Rule based information extraction from electronic health records by forward-chaining,” in Article in Elsevier Ergonomics Book Series. Amsterdam, The Netherlands, Aug. 2014.

[25] P. B. Jensen, L. J. Jensen, and S. Brunak, “Mining electronic health records: Towards better research applications and clinical care,” Nature Rev. Genet., vol. 13, no. 6, pp. 395–405, 2012.

[26] A. Jorge, V. M. Castro, A. Barnado, V. Gainer, C. Hong, T. Cai, R. Carroll, J. C. Denny, and L. Crofford, “Identifying lupus patients in electronic health records: Development and validation of machine learning algorithms and application of rule-based algorithms,” in Seminars in Arthritis and Rheumatism. Amsterdam, The Netherlands: Elsevier, 2019.

[27] E. Chazard, G. Ficheur, S. Bernonville, M. Luyckx, and R. Beuscart, “Data mining to generate adverse drug events detection rules,” IEEE Trans. Inf. Technol. Biomed., vol. 15, no. 6, pp. 823–830, Nov. 2011.

[28] H.-J. Dai, C.-Y. Chen, C.-Y. Wu, P.-T. Lai, R. T.-H. Tsai, and W.-L. Hsu, “Coreference resolution of medical concepts in discharge summaries by exploiting contextual information,” J. Amer. Med. Inform. Assoc., vol. 19, no. 5, pp. 888–896, Sep. 2012.

[29] G. Hrovat, G. Stiglic, P. Kokol, and M. Ojsteršek, “Contrasting temporal trend discovery for large healthcare databases,” Comput. Methods Programs Biomed., vol. 113, no. 1, pp. 251–257, Jan. 2014.

[30] L. A. Petersen, S. Wright, S.-L.-T. Normand, and J. Daley, “Positive predictive value of the diagnosis of acute myocardial infarction in an administrative database,” J. Gen. Internal Med., vol. 14, no. 9, pp. 555–558, Sep. 1999.