KANTIPUR ENGINEERING COLLEGE

(Affiliated to Tribhuvan University)

Dhapakhel, Lalitpur

[Subject Code: CT755]

A MAJOR PROJECT PROPOSAL ON

WEARABLE HEART RATE MONITORING DEVICE AND REAL-TIME SYSTEM PREDICTION OF HEART RATE

Submitted by:

Nitesh Poudel	[KAN075BEI03]
Pawan Bhandari	[KAN075BEI04]
Prajjwol Shrestha	[KAN075BEI05]
Ujjwal Naral	[KAN075BEI09]

A MAJOR PROJECT SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENT FOR THE DEGREE OF BACHELOR IN COMPUTER ENGINEERING

Submitted to:

Department of Computer and Electronics Engineering

May, 2022

WEARABLE HEART RATE MONITORING DEVICE AND REAL-TIME SYSTEM PREDICTION OF HEART RATE

Submitted by:

Nitesh Poudel	**[KAN075BEI03]**
Pawan Bhandari	**[KAN075BEI04]**
Prajjwol Shrestha	**[KAN075BEI05]**
Ujjwal Naral	**[KAN075BEI09]**

A MAJOR PROJECT SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENT FOR THE DEGREE OF BACHELOR IN COMPUTER ENGINEERING

Submitted to:

Department of Computer and Electronics Engineering

Kantipur Engineering College

Dhapakhel, Lalitpur

May, 2022

ABSTRACT

Low heart rate carries a risk of cardiovascular diseases that can become fatal. Monitoring the heart rate is important to determine heart's function to discover if there is any irregularity. Detection of any irregularity in early stage help doctors diagnose and treat the conditions better. Development of advanced Machine Learning algorithms has allowed health-care professionals to analyze health based data to discover risks by making accurate prediction. The proposed system consists of two phases, namely, an offline phase and an online phase. The offline phase emphasizes on developing the model with lowest root mean square error and the online phase emphasizes on predicting the heart rate in advance. Recurrent neural network (RNN), long short-term memory (LSTM), gated recurrent units (GRU), and bidirectional long short-term memory (BI-LSTM) are applied to heart rate time series to determine the best model. For the online phase, Apache Kafka and Apache Spark will be used to predict the heart rate in advance based on the best developed model. Therefore in this project we propose a heart monitoring wearable device that constantly monitors and predicts the heart rate.

Keywords – Heart rate, Machine Learning, Wearable Device

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1
INTRODUCTION

1.1 Background

Today's world revolves around minimal physical activity. However, people are aware of their health and are trying to improve it through various means. The rise of AI and ML has led to the development of various medical applications. Some of these include fitness apps that use these tools for better results.

People are choosing various kinds of wearable devices such as smart watches, smart bands, and smart clothes for tracking their activity. The features and functions of these products are increasing day by day. Some of these include monitoring fertility, temperature, and breathing rate. There are also various kinds of medical devices that can be used to monitor different aspects of a person's health, such as diabetes. Various companies such as Apple, Samsung, and Fitbit provide these devices with the necessary data to track a person's activity. Since these products are usually private, there is a need to create a machine learning model that can keep the user's data secure. This can be done through a combination of distributed learning and machine learning. For instance, in order to prevent unauthorized access to the user's data, multiple clients are involved in training a model.

Leading a healthy life has never been easier. One can track their heart rate, caloric intake and activities throughout the day using wearable tech.One of the leading causes of death in the modern world is heart-related but is much often ignored compared to other prevalent diseases such as diabetes and cancer. However, incorporating heart activity data into the lifestyle can bring about crucial health hazards before they become life-threatening.So, to accomplish these goals we first have to be able to predict the heart-rate in real-time as well as compare with the database for any anomaly in the metabolism.The prediction system gets refined as more information is fed to the algorithm.

1.2 Problem Definition

Heart diseases are one of the most prevalent and life-threatening diseases one can get but are often overlooked until it is too late. The symptoms such as shortness of breath and palpitations can be detected and accounted for by measuring and monitoring the heart-rate at various intervals.But not everyone can afford to get checkups regularly and log their information reliably to encounter heart diseases in the early stages. A person may exert fatal effort during workout or even everyday activities but if it can be predicted easily, it might prevent major accident.

1.3 Objectives

The primary objectives of this projects are as follows:

1. To develop a smart wearable capable of logging various health data
2. To find and implement the model that has lowest Root Mean Square Error value to predict heart rate
3. To use the best developed to predict heart rate in real time

1.4 Project Features

The project will be able to accomplish following:

- Accurately measure heart rate
- Predict possible heart disease
- Predict heart rate in advance

1.5 Project Application

- For medical use as critical patient monitoring system
- For personal use as personalized fitness device for chronically ill patients.

1.6 System Requirement

1.6.1 Software Requirements

The software requirements are as follows:

Operating System	Windows/Mac/Linux
Server	AWS/Azure
Hardware control	Arduino IDE
Mobile connection	Android Studio

Table 1.1: Software Requirements

1.6.2 Hardware Requirements

The hardware requirements are as follows:

Microcontroller	Arduino UNO
Band	Watch strip
Sensors	Heart rate, Tempeature, Spo2, Accelerometer and Gyro sensor
Modules	RTC module, Bluetooth module

Table 1.2: Hardware Requirements

1.7 Project Feasibility

The feasibility analysis of the system has been done from various aspects such as technical, operational and economical viewpoint.

The present technology is sufficient to meet the requirements of the system, the required algorithm exists and the device to input the data to the system is also present. The system is believed to work well when developed and installed. The requirements for the project have been accounted for and the system is built on available resources to meet the requirements. The detailed feasibility study is as follows

1.7.1 Technical Feasibility

Our project satisfies technical feasibility needs. The existing network protocols and operating system services of various operating systems would allow for feasible implementation of this application. As this service satisfies technological hardware and software capabilities of present day available personal devices, the proposed project was decided to be technically feasible.

1.7.2 Operational Feasibility

The operation of the system requires only a modern computer, the user interface will be simple and intuitive. The solution proposed for the project is operationally workable and user-friendly to end users.

1.7.3 Economic Feasibility

Economic feasibility analysis is the most commonly used method for determining the efficiency of a project. It is also known as cost analysis. It helps in identifying profit against investment expected from a project. Cost and time are the most essential factors involved in this field of study. Developed system is economically feasible. It can be developed on a simple PC which can be available at an affordable cost. System is built on open-source language, so development is free of cost. All the references and resources are freely available on the internet. So, we can say that the developed system is economically feasible.

CHAPTER 2
LITERATURE REVIEW

Heart Rate Prediction

Many studies and research have been conducted for heart rate prediction. FitRec system has been proposed by Ni et al.[1].They have extensively collected the huge amount of data from the wearables from the fitness company endomondo which included data from fitness bands and smartwatches and other wearable sensor devices. So, they have contributed in providing this fitrec dataset. It has other two major contributions. It has LSTM model captures context information in two stages. First in a specific activity like running, biking, hiking the context is captured and other one is like for different user's activity history.Also it static user embeddings are inferred from user attributes and the different workout sequences contributes to the temporal embeddings.It uses two-layer LSTM module and the attention-based encoder decoder module for the purpose. Another major contribution includes the prediction and recommendation. It estimates a user's heart rate profile for a candidate activity; and predicts and recommends suitable activities on the basis of this profiling. Also, it gives the personalized running route recommendation, by considering a variety of targets such as user preferences, goals, and environment to boost the wellness profiling performance. It uses the centralized data for its activity modelling and recommendation. So, the user privacy comes to concern as fitness data is user's personal data. The approach of wellness which accounts fitness and self-help has been tried by Tweetfit.[12].It uses the wellness related attributes and the data from the personal sensors. Also, it integrates the data from social media like twitter, Instagram, endomondo and foursquare for wellness profiling. So, this is the first to create the fusion between social media data and sensor data for wellness profiling. Bayesian inference has been used along with federated learning for heart rate prediction by Fang, Liu in 2020. It has proposed two Bayesian federated learning approaches. One is Federated learning based on Sequential Bayesian method (FD Seq Bayes) and the other one is Empirical Bayes based Hierarchical Bayesian method (FD HBayes-EB) for heart rate prediction without pooling data to the cloud for privacy preservation. This method uses the autoregression with exogenous variable (ARX) model. And the baseline for the two solution for aggregation of global model is FedAvg and HBayes-

MCMC(Markov Chain Monte Carlo). It has only monitored the ten people and the data is taken from their polar smart watches[13]. An ordinary differential equation (ODE) model for complete outdoor running exercise sessions has been proposed to predict the heart rate response. Unlike the most of the model for heart rate prediction which has been testes for indoor exercises and running activities, this model is fully focused to predict the outdoor running and exercise sessions where it measures the heart rate and does the prediction for short duration. [14] Most recent models related to predicting and controlling heart rate response to exercise has been summarized by Ludwig et al.[15] . The federated learning with clustering has been applied on the heart rate variability data by Joo Hun Yoo .[16] In this techniques unlike others which assumes data to be identical and distributed , it addresses the non-IID data which increasing accuracy in severity prediction. The Hierarchical clustering based federated learning process called personalized federated cluster models is proposed which predict the Major Depressive Disorder (MDD) severity from Heart Rate variability. So, this has approached the issue of data privacy in mental healthcare analysis and proposed the improved HRV performance for MDD in federated approach. This doesn't use the wearables data but obtained from the research at the Department of Psychiatry at Samsung Medical Center.

2.1 Related Theory

2.1.1 Recurrent Neural Network(RNN)

A recurrent neural network (RNN) is a type of artificial neural network which uses sequential data or time series data. These deep learning algorithms are commonly used for ordinal or temporal problems, such as language translation, natural language processing (nlp), speech recognition, and image captioning.Like feedforward and convolutional neural networks (CNNs), recurrent neural networks utilize training data to learn. They are distinguished by their "memory" as they take information from prior inputs to influence the current input and output. While traditional deep neural networks assume that inputs and outputs are independent of each other, the output of recurrent neural networks depend on the prior elements within the sequence.

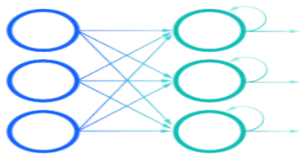

Figure 2.1: Recurrent Neural Network(RNN)

2.1.2 Long Short Term Memory(LSTM)

LSTM (Long Short-Term Memory) is a Recurrent Neural Network (RNN) which can learn long-term relationship and patterns. Unlike conventional feed-forward neural networks and RNN it can selectively remember patterns for long durations of time. A typical LSTM network is comprised of different memory blocks called cells. There are two states that are being transferred to the next cell; the cell state and the hidden state. The memory blocks are responsible for remembering things and manipulations to this memory is done through gates i.e. input gate, forget gate and output gate. The gated cell learns the priority of information over time and later such learned weights is used to rank the priority of information. On the basis of these priorities, the gate decides whether to store or forget information.

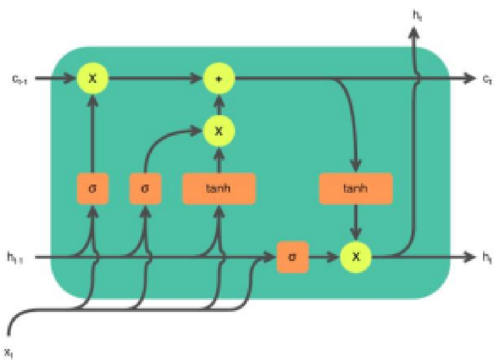

Figure 2.2: Long Short Term Memory(LSTM)

2.1.3 Gated Recurrent Units(GRU)

Gated Recurrent Units(GRU) is a variation of Recurrent Neural Networks(RNN) develop to solve the problem of Vanishing-Exploding gradients. It consists of three gates and does not maintain an Internal Cell State. The information which is stored in the Internal Cell State in an LSTM recurrent unit is incorporated into the hidden state of the Gated Recurrent Unit.The different gates of a GRU are as described below:-

- Update Gate(z):It determines how much of the past knowledge needs to be passed along into the future. It is analogous to the Output Gate in an LSTM recurrent unit.
- Reset Gate(r): It determines how much of the past knowledge to forget. It is analogous to the combination of the Input Gate and the Forget Gate in an LSTM recurrent unit.
- Current Memory Gate($\bar{h_t}$):It is often overlooked during a typical discussion on Gated Recurrent Unit Network. It is incorporated into the Reset Gate just like the Input Modulation Gate is a sub-part of the Input Gate and is used to introduce some non-linearity into the input and to also make the input Zero-mean. Another reason to make it a sub-part of the Reset gate is to reduce the effect that previous information has on the current information that is being passed into the future.

2.1.4 Bidirectional Long Short-term Memory

Bidirectional long-short term memory(bi-lstm) is the process of making any neural network have the sequence information in both directions backwards (future to past) or forward(past to future).

In bidirectional, input flows in two directions, making a bi-lstm different from the regular LSTM. With the regular LSTM, input flows in one direction, either backwards or forward. However, in bi-directional, flows in both directions to preserve the future and the past information.

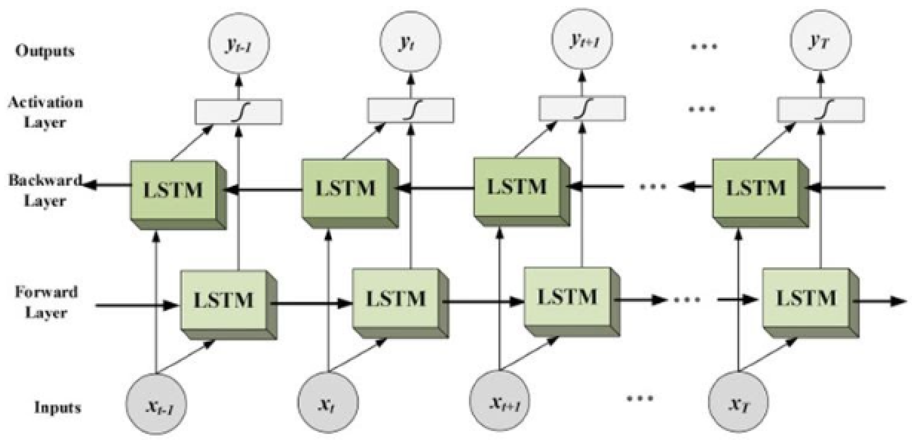

Figure 2.3: Bidirectional Long Short Term Memory(BI-LSTM)

CHAPTER 3
METHODOLOGY

Heart rate forecasting system consists of two main phases, an offline phase and an online phase.

3.1 Offline Phase

The offline phase focuses on finding the best deep learning model with the smallest Root Mean Square Error(RMSE) value.

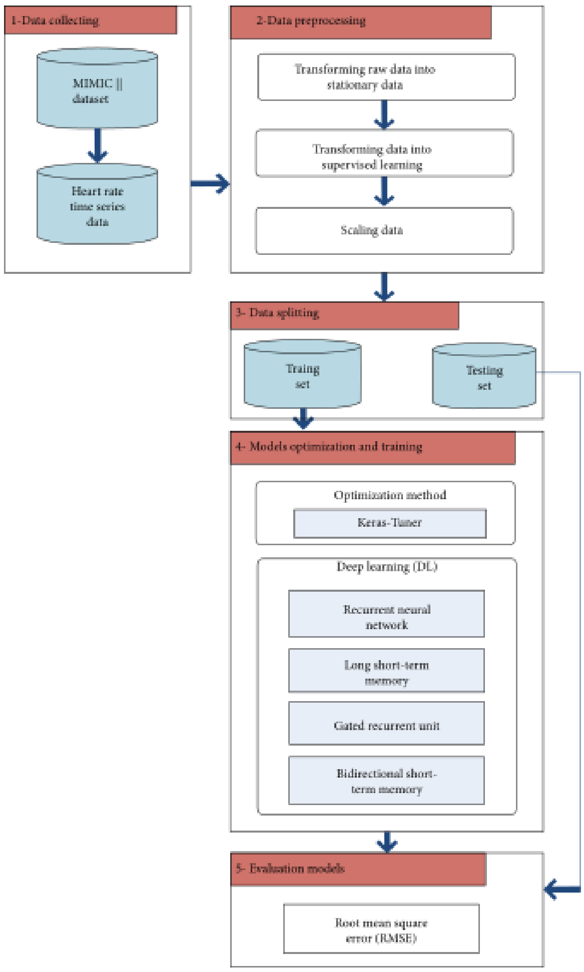

Figure 3.1: Architecture of offline phase

3.1.1 Data Collection

Along with the data collected from the wearable device, an open care dataset called Medical Information Mart for Intensive Care (MIMIC-II) will be used to train the model.

3.1.2 Data Preprocessing

In this phase, three steps will be done as follows:
Transforming raw data into stationary data: Given the nonstationarity of the Heart Rate time-series dataset, the data will be transformed into the time-series-based nonstationary dataset into a stationary dataset to be fitted in the prediction model. In particular,a transformation method will be applied called differencing. The differencing method's function computes the difference of successive operators in the sequence differencing to escape varying means.

Transforming data into supervised learning: Within the supervised learning mode, the key idea is that machine learning models learn the association function between input and output variables, denoted by (X) and (Y), respectively. According to the context of this work, which is the real-time prediction, the model will be trained based on input data (X), and then it will be tested to predict the output data (Y) in real time.

Data Scaling:As the work focuses on applying forecasting deep learning models, these models usually work on scaled data within their activation function ranges. A normalization process will be implemented to scale the original data between the rage -1 to 1.

3.1.3 Data Splitting

We split the 80 percent of the data as training data, 20 percent as testing data.

3.1.4 Model Optimization and Training

Four deep learning models are used for heart rate forecasting: RNN, LSTM, BI-LSTM, and GRU. According to 3.2 the deep learning model (i.e., neural network) consists of
(1) input sequences in terms of numbers of lags,
(2) hidden layers and then dropout layer, and
(3) output layer including dense layer that emits its output in three types of forecasting: 5 minutes, 10 minutes, or 15 minutes. The different numbers of hidden layers are applied in hidden layers, including one layer, two layers, and three layers for each model and dropout layer [19]. For the output layer, many neurons have been configured to predict three forecasting times in advance for heart rate (i.e., 5minutes, 10minutes, and 15minutes). In dense layer, rule activation function and Adam optimizer are used [20]. For loss function, mean square error (MSE) is used. For optimization models, a Keras-Tuner library [21] is used to choose the optimal value for two parameters. We configured value range for two parameters: the number of neurons and a dropout rate.

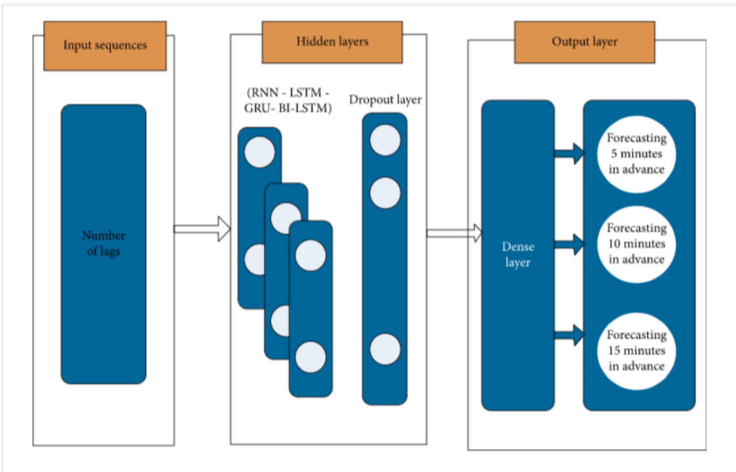

Figure 3.2: Model Optimization and Training Architecture

3.2 Online Phase

The online phase contains two steps:

1. data generation from the simulated sensor

2. online prediction.

Each step will be described as follows:

3.2.1 Data generation from simulated sensor

Apache Kafka and Apache Spark are streaming processing platforms that are used to build the online phase. In particular, Apache Kafka [22] has the guarantees of delivery and ordering data streams, which are in demanded the latency-critical applications such as predicting the heart rate in real time. The vital role of Apache Kafka is collecting the health stream data from the wearable sensors (i.e., simulated HR) and then sending it to Apache Spark. Regarding this work, a simulated sensor is developed to generate time-series-based HR dataset in JSON format. The schema of each tuple based on one-minute data consists of two parameters: HR value and timestamp. For the technical level, Kafka Producer API is used to publish the streaming data to the Kafka topic, which is consumed by Apache Spark.

3.2.2 Online prediction

Apache Spark is a micro-batch-based streaming processing platform. The core principle of Apache Spark is utilizing the memory to analyze streaming data [23]. Furthermore, one of its strengths is ingesting streaming data from different sources including medical sensor. Also, it has a set of streaming APIs, which are able to handle streaming data, and then performing set of complex window-based operations such as map, reduce, and join [24].

According to this work's context, Apache Spark streaming API is used to consume medical data (i.e., HR streams) from precreated Kafka's topic. We have utilized the windowing capability of Apache Spark. Windowing is the heart of any stream processing platform, which splits infinite data stream into chunks of finite data to execute an operation. Streaming window is defined by two parameters: window length and sliding interval. Window length defines window duration, while sliding interval specifies the interval to performing the window operation.

CHAPTER 4
EPILOGUE

4.1 Expected Output

The finalized project is expected to log various health parameters like heart rate, SpO2 and it is expected to predict the heart rate of chronically ill patient so it can be monitored to prevent exceeding a certain level.

4.2 Work Scheduled

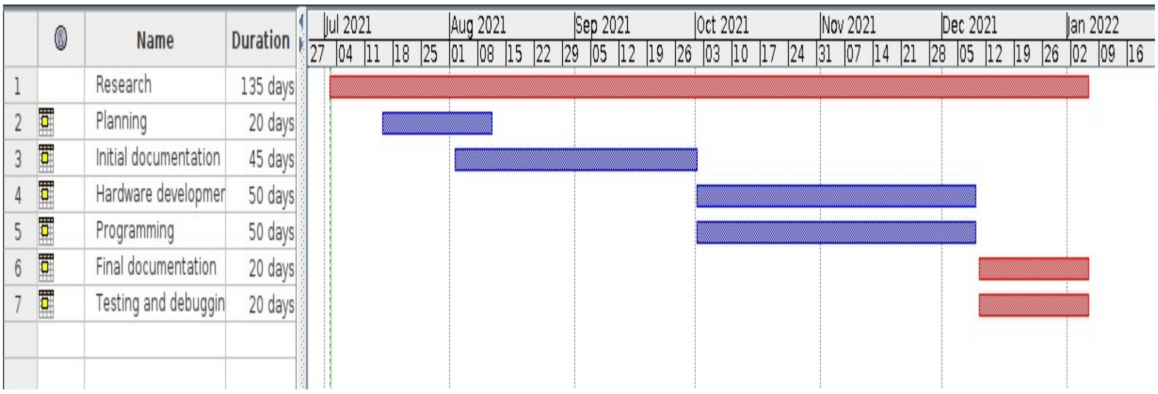

Figure 4.1: Gantt Chart

4.3 Cost Estimation

Items	No. of Item	Unit Price(Rs)	Total Price(Rs.)
Arduino UNO	1	1,000/-	1,000/-
Heart rate and Spo2	1	500/-	500/-
Temperature sensor	1	50/-	50/-
Bluetooth module	1	800/-	800/-
RTC module	1	200/-	200/-
Watch Strip	1	700/-	700/-
3D Printed Watch case	1	5,000/-	5,000/-
Total			8,250/-

Table 4.1: Cost Estimation