

Project Synopsis  
On  
**Calories Burnt Prediction Using Machine Learning**

Submitted as a part of course curriculum for

**Bachelor of Technology**  
in  
**Computer Science**



**Submitted by**

Yash Goel (2000290120195)  
Yash Puri (2000290120196)  
Vishal Verma (2000290120193)

**Under the Supervision of**

Ms. Anshula Gupta  
Assistant Professor,  
Department of Computer Science

**KIET Group of Institutions, Ghaziabad**  
**Department of Computer Science**  
**Dr. A.P.J. Abdul Kalam Technical University**  
**2022-2023**

# **TABLE OF CONTENTS**

TITLE

DECLARATION

CERTIFICATE

ACKNOWLEDGEMENT

ABSTRACT

LIST OF FIGURES

LIST OF ABBREVIATIONS

CHAPTER 1 INTRODUCTION

1.1 Introduction

1.2 Problem Statement

1.3 Objective

1.4 Scope

CHAPTER 2 LITERATURE REVIEW

CHAPTER 3 PROPOSED METHODOLOGY



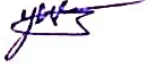
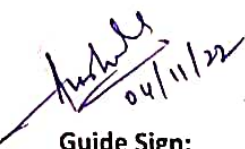
CHAPTER 4 TECHNOLOGY USED

CHAPTER 5 DIAGRAMS

CHAPTER 6 CONCLUSION

REFERENCES

## TITLE PAGE

KIET GROUP OF INSTITUTIONS
MINI PROJECT
2024 BATCH, SESSION 2022-23
APPROVAL OF TITLE BY GUIDE
Guide Name: Mns. Anshula Gupta
Project ID: PCS 23-65
Domain: Web development, Machine Learning
Title: Colonies banned prediction using ML
Date: 4/11/2022
Student Names with Signature VISHAL VERMA  YASH PURI  YASH VIJAYAL 
 04/11/22 Guide Sign:

## **DECLARATION**

We hereby declare that this submission is our work and that, to the best of our knowledge and belief, it contains no material previously published or written by another person nor material which to a substantial extent has been accepted for the award of any other degree or diploma of the university or other institute of higher learning, except where due acknowledgement has been made in the text.

**Signature of Students:**

**Yash Goel-**

**Yash Puri-**

**Vishal Verma-**

**Name:** Yash Goel, Yash Puri, Vishal Verma

**Roll No.:** 2000290120195, 2000290120196, 2000290120193

**Date:** 14-November-2022

## **CERTIFICATE**

This is to certify that Project Report entitled “**Calories Burnt Prediction Using Machine Learning**” which is submitted by **Yash Goel, Yash Puri, Vishal Verma** in partial fulfilment of the requirement for the award of degree B. Tech. in Department of Computer Science of Dr A.P.J. Abdul Kalam Technical University, Lucknow is a record of the candidates own work carried out by them under my supervision. The matter embodied in this report is original and has not been submitted for the award of any other degree.

**Date:** 14-November-2022

**Supervisor Signature**  
Ms. Anshula Gupta  
Assistant Professor,  
Department of Computer Science

## **ACKNOWLEDGEMENT**

It gives us a great sense of pleasure to present the synopsis of the B.Tech Mini Project undertaken during B.Tech. Third Year. We owe a special debt of gratitude to Ms. Anshula Gupta (Assistant Professor), Department of Computer Science, KIET Group of Institutions, Delhi- NCR, Ghaziabad, for her constant support and guidance throughout the course of our work. Her sincerity, thoroughness and perseverance have been a constant source of inspiration for us. It is only her cognizant efforts that our endeavours have seen the light of the day.

We also take the opportunity to acknowledge the contribution of Dr. Ajay Kumar Shrivastava, Head of the Department of Computer Science, KIET Group of Institutions, Delhi- NCR, Ghaziabad, for his full support and assistance during the development of the project. We also do not like to miss the opportunity to acknowledge the contribution of all the faculty members of the department for their kind assistance and cooperation during the development of our project.

Last but not the least, we acknowledge our friends for their contribution to the completion of the project.

Signature:

Yash Goel-

Yash Puri-

Vishal Verma-

Date : 14-November-2022

Name : Yash Goel, Yash Puri, Vishal Verma.

Roll No: 2000290120195, 2000290120196, 2000290120193.

## **ABSTRACT**

Machine Learning is a bunch of algorithms which allows enormous software applications to become more accurate in predicting outcomes without being explicitly programmed. The basic fundamental of machine learning is to build models and employ algorithms that can take input data and use its algorithms and analysis to predict an output while updating outputs as new data becomes available.

In this project we are trying to calculate the exact amount of calories burned during workout. For achieving this we are making a machine learning and deep learning models in which we will use certain machine learning algorithms and methods of machine learning and deep learning and will choose the best one depending on the accuracy levels. To achieve this first we will train our dataset which will include dataset for calories burned during a workout and the other dataset which we will use will be for exercise which will include various fields such as gender, duration of exercise, body temperature, height, weight etc. Then we will train our model on these dataset after training of our dataset we find out its accuracies and error or if any fixes needed in the model. When we are done with training of our model on these dataset now we will check our model with new data or we can say with testing datasets. Then we will find our model which is best working with less error values. In this prediction model we will try python libraries such as NumPY, seaborn, matplotlib, open cv and some other algorithms and regression techniques as per our requirements if necessary. These libraries and models are proposed and these may be different at the time of implementation.

## **LIST OF FIGURES**

**FIG-1.1:** Workflow Model Diagram

### **CHAPTER 1-INTRODUCTION**

Food plays a major role to provide us energy in our day to day tasks and to perform workout activities. The major portion of energy to our muscles while we workout goes from carbohydrate. Carbohydrates breaks down into simpler forms such as glucose and fructose which by using oxygen is converted into energy to supply it to our target muscles during our workout. In this project we will arrange our datasets that is our calories burned dataset and exercise dataset which includes various fields such as the user, age, gender, duration of workout, average heartbeat per minute, body temperature etc. so these parameters will be taking into consideration for input and to produce desired output. The motive of our project is about making a machine learning(or deep learning) model to calculate or predict the exact number of calories burned during workout or we can say this can be used to keep a track of amount of calories burned during workout. We are going to use various python libraries such as NumPY, pandas, SSN, matplotlib, seaborn etc. and various machine learning algorithms mainly regression algorithms such as linear regression, logistic regression, XG boost regression to train our model on dataset. All these would be used to create a prediction model also for calories burnt. Then comes the stage of testing of our model on testing dataset which contains new data and we will compare it the difference between the actual values and predicted values if the difference or error value is less then it will be the best prediction model. This will be a supervised machine learning prediction model. According to us this model will prove to be very helpful in predicting the exact amount of calories burned during any kind of physical activities.



## **PROBLEM STATEMENT**

Our problem statement in simple words would be like:

- We are going to make a machine learning model.
- This will be supervised model for prediction of calories.
- This model will predict the exact amount of calories burned during workout.
- This model will be suitable for keeping a track on calories burned.

## **OBJECTIVE**

Our objective of this project is to provide a machine learning model which will be effective for the following:-

- Takes a data about duration of exercise, gender, age, weight, height etc.
- Predicts the exact amount of calories burned during workout.
- Keeps a track on calories burned in physical activity.
- And some similar related functions like these.

## **SCOPE**

After discussion with the team and with our guide, we can consider the scope of this project as follows:

- We can create an app for calculating or predicting the exact amount of calories burned during workout.
- Health conscious people can use this model to quickly get information regarding calories they burned in their workout and to plan accordingly as per their fitness goals.

## **CHAPTER 2 -LITERATURE REVIEW**

### **2.1 Scaling Up Machine Learning: Introduction (Author- Ron Beckerman)**

In this research paper the author tries to convey that Distributed and parallel processing of very large and enormous datasets has been employed for decades in specialized, very high budget settings, such as financial applications.

The current rise in interest in scaling up in machine learning and its applications can be partially attributed to the evolution of hardware architectures and programming frameworks that make it easy to exploit the types of parallelism in many learning algorithms.

Several platforms make it convenient to implement concurrent processing of data instances or their features. This allows straight forward parallelization of many learning algorithms that view input as an unordered batch of examples and aggregate isolated computations over each of them.

Increased attention to large scale machine learning is also due to the spread of very large datasets across many modern applications.

Such datasets are often accumulated on distributed storage platforms, motivating the development of learning algorithms that can be distributed appropriately.

Finally, the proliferation of sensing devices that perform real time inference based on high dimensional, complex feature representations in learning applications.

### **2.2 Prediction of users calorie routine using CNN (Author- Sathiya T, Surya Prakash)**

By this research paper the author is trying to empower the great success of machine learning and its various algorithms in helping many people with their calorie or diet routine. As we all are aware about calories, how we can manage the calories in-take to maintain our fitness goals but we can't calculate the amount of calories that have been spent in metabolic activities during the day for the functioning of the body. By this paper the author enforced to design a machine learning intelligent system that will be convenient to help users to be more concerned about their day to day calorie intake.

To employ this model the author used neural networks to classify the food items. Author emphasis light on method of CNN which needs highly performing computational machines in order to train non linear data and it takes more time to train the data. There will be more time when we train our data using CNN but the author conveys after completion of training of dataset it can produce results very fast and more accurate in

very less time. The author tries to explain the method of CNN to train datasets and conveys that CNN consumes high computational time but they are considered to be the most suitable method or algorithm to classify images when we have huge number of classes which is a appreciable thing about this algorithm. In this paper the author trained his model on various food items using CNN but currently there are many food items which are not available but in future they can be trained using the same CNN approach for this problem and can prove to be very effective.

### **2.3 Machine learning for health care diagnostics** **(Author- K.Kalaiselvi and M.Deepika)**

By this research paper first of all the author sets up the base for health diagnostics and limitations in health care diagnostics also how can they be coped by using machine learning and its various methods and algorithms. In this research paper the author tries to explain the current limitations in health care and the benefits of introducing machine learning for diagnostics. Author conveys that systems based on machine learning can be used to find various complications before they occur at very early stage of diagnosis which will prove to be very useful in treatment if diagnosed at an early stage. This paper basically contrasts on the importance of the systems based on machine learning and its algorithms in health care for diagnostics. Very huge and enormous dataset are available in IT department of various hospitals which includes the past history of patients, disease encountered, disease diagnosed, treatment given or cure such data sets from these hospitals can be used to train our machine learning algorithms on the basis of outcome. Author says that it is comparatively easy to find patterns in such a huge dataset and predict outcomes using machine learning. Machine learning systems can be advanced and trained with more huge number of dataset to be capable of giving more accurate results as those given by a well trained clinician.

### **2.4 Neural network for the beginner** **(Author- Robin M.Snyder)**

In this research paper the author tries to explain the basics of neural networks related to machine learning and also tells some of the references where we can go and implement these neural networks.

In the starting the author tells about the neural networks that neural networks in machine learning are motivated by brain and they can be used to easily recognize patterns based on previous training. In practice, with time we can design an expert and efficient system to recognize patterns based on previous training.

We can understand more about the neural networks by taking a look into an example which is one of the most simplest example which is a fully connected three layer model that consists of input layer, the hidden layer and at last the output layer.

Training of neural networks can be difficult and can be very time consuming also but once we have trained our neural network model then they can prove to be very fast in recognizing patterns and performing other related tasks. Due to these ease and

functionalities neural networks software are becoming more available and affordable. The author concludes that if a problem has no obvious solution then various intelligent systems can be used as an approach to solve those problems and if the problem some kind of pattern recognition then neural network may act as possible solution. There can be various other possible solutions to such related problems.

## **2.5 Coronavirus Disease(COVID 19):Analysis and Prediction using machine learning techniques**

**(Author- Dimple Tiwari, Bhupesh Singh, Bharti Nagpal)**

As we all know from the history coronavirus is as deadly as it appears to be, from SARS COV in the seafood market near Wuhan in China during the end of year 2019.

And it quickly spread across the world which made a dreadful situation worldwide across the globe. Due to sudden spread and lack of specific treatment people faced major challenges.

It was observed that symptoms such as fever, throat irritation, breathing issues, but it is observed that people having diabetes or patients with blood pressure problems had more severe symptoms as compared to other people.

In some research it is also shown that covid reduced the count of white blood cells, red blood cell's which were one of the major causes of fatality rate. With the outbreak of covid 19 machine learning and data science has played a crucial role in prediction of upcoming trends of covid 19 for safety measures and implementation of various precautions that had been taken to combat or to fight the battle against covid on huge scale more effectively. With this paper the author came up to conclusion that the predicted outcomes of naive based system are very close enough to the actual confirmed cases. However other techniques of machine learning such as linear regression and support vector machine do not stand up on the marked line. And thus this can be concluded that in such cases naive bayes is more likely to be preferred in future for prediction purposes. Finally at last we can conclude that AI and ML can learn to detect an outbreak at an early stage.

## **2.6 Role of data science in managing covid 19 pandemic**

**(Author- Nikita Saxena, Priyanka Gupta, Anurag s Rathore)**

In this research paper the author is trying to convey that data science along with machine learning and its different prediction algorithms has always played an important and vital role in fighting against deadly pandemic which was popularly known as SARS COV 2.

From categorizing the covid zones based on the severity and density of infected patients to deciding the dosage and medicine for the patient on the severity of infection as a cure or treatment. Statistical analysis and mathematical model are always key towards data science field which can be further categorized into data management, data visualization and statistical machine learning with the help of its various algorithms which can further be used for organization, sorting, processing and many times for real time data analysis. The data science expert must build their model keeping in mind the limitations and

shortcomings also that are associated with it. They have to consider their models with certain reservations. The catch here proposed here by the author is that these models are acting just as an helping hands towards successful fight against covid 19 or towards the successful treatment of covid 19 infection on such a large scale.

## **2.7 Website Development Technologies: A Review**

**(Author- Pratiksha D Dutonde, Shivani S Mamidwar, Monali Sunil Korvate, Sumangla Bafna, Prof. Dhiraj D Shirbhate)**

"Web development" typically refers to the most non-design aspects of building net sites. Net development could use content management systems (CMS) to create content changes easier. There is square measure 3 styles of net developer specialization: frontend, back-end and full-stack. Traditional Technologies in Web Development: Knowing the fundamental classes of net technologies is important if you propose to figure in net development. JavaScript could be a lightweight, cross-platform, and taken scripting language. Node.js is an event-driven, non-blocking (asynchronous) I/O and it's not an artificial language. Back-end development cares with web site design, scripting, and communication with databases. Back-end code permits the communication between browsers and data from databases. Databases are necessary as a result of the permit websites and applications to handle user knowledge. There are two main sorts of databases: SQL and NoSQL. Once a business case has been developed and approved, it's time to begin building. A level-3 heading must be indented, in Italic and numbered with an Arabic numeral followed by a right parenthesis. The final check of system practicality is when the web site is ready to deploy. The developer should make sure that the positioning is responsive i.e., it seems. properly on devices of all sizes not a part of the web site ought to behave abnormally no matter the scale of the screen. HTML, CSS, and JavaScript square measure the languages used for face development. Conclusion: The Worldwide internet represents the highest technology to the perfect of a very distributed network atmosphere for polymorphic communication. As such, it should be although of as a paradigm shift aloof from earlier network protocols. Web Applications design issues the look and implementation of pc code that runs on internet servers, rather than running only on desktop computers, laptops, or mobile devices.

## **2.8 ROLE OF RESPONSIVE DESIGN IN WEB DEVELOPMENT**

**(Author- Fernando Almeida, J.Monteiro)**

Responsive design allows software developers to build a Web page that can dynamically adapt to the size of the devices. This development philosophy enables the rendering of Web pages in a fast and optimized way, ensuring a good user experience on mobile devices, tablet and desktop. In the scope of this study, we intend to explore the main advantages and limitations associated with responsive Web design. We adopted a quantitative approach based on a questionnaire filled by 181 professionals in the industry that allowed us to identify the reasons that lead software developers to the adoption of the responsive design and address the limitations felt by them. The results obtained indicate

that offering a good user experience and increasing stands out as being the most important advantages. The Web design process uses techniques to adequate structuring of information, using appropriate resources to serve on web pages, in a manner that the user can reach his goal in a direct and pleasant way. Web design distinguishes itself from other traditional forms of design. The Web is a unique channel that forces designers not to be able to control the environment around them. Elements such as colors, shapes, and layouts can be customized by the user , and there are no guarantees that all users will see the same Web page in the same way that it was designed and developed.

## **2.9 Research on Html 5 in Web Development**

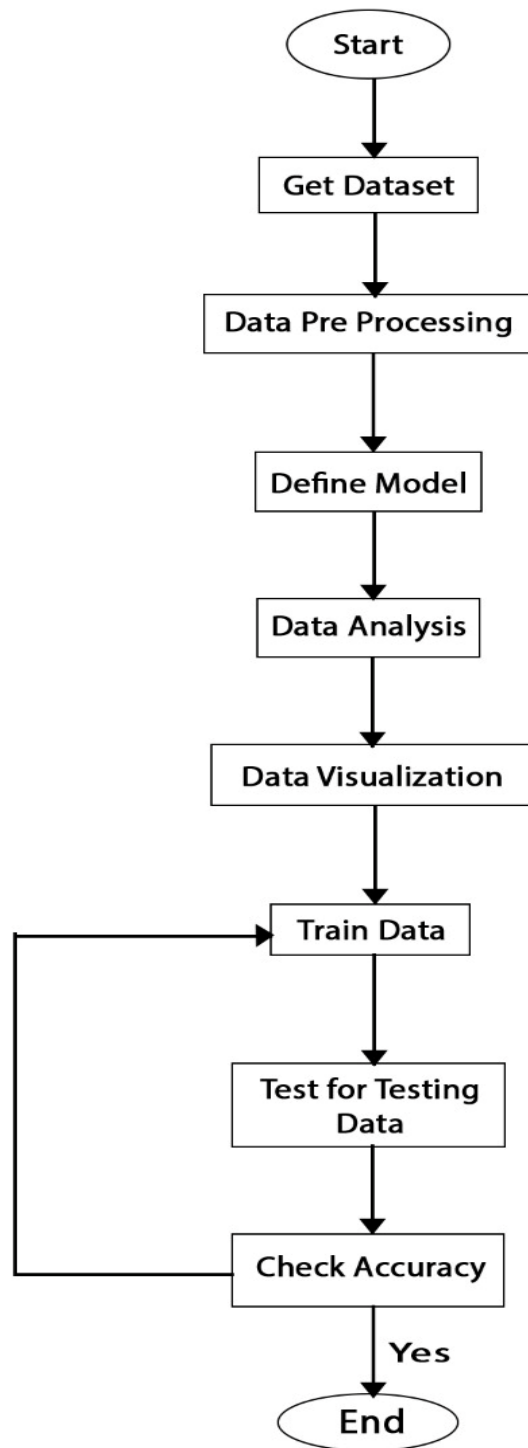
**(Author- Ch Rajesh, 2 K S V Krishna Srikanth)**

HTML5 is everywhere these days. HTML5 is the new and elegant standard for HTML that provides web users and developers enhanced functionality. The older versions of HTML, HTML 4.01, which came in 1999, and the web development have changed notably since then. HTML 4, XHTML, CSS and the HTML DOM Level 2 are now replaced with HTML5. It was brought to deliver rich content without the need for additional plug-ins and proprietary technologies. The new power of HTML5 supplies the user with everything from animation to graphics, music to movies, and can also be used to build complicated web applications and also supports cross-platform. HTML5 standard initiates the development of real-time collaborations in web browsers, which leads to less work for web developers.

HTML5 is everywhere these days. HTML5 is the new and elegant standard for HTML that provides web users and developers enhanced functionality. The older versions of HTML, HTML 4.01, which came in 1999, and the web development have changed notably since then. HTML 4, XHTML, CSS and the HTML DOM Level 2 are now replaced with HTML5. It was brought to deliver rich content without the need for additional plug-ins and proprietary technologies. The new power of HTML5 supplies the user with everything from animation to graphics, music to movies, and can also be used to build complicated web applications and also supports cross-platform. HTML5 standard initiates the development of real-time collaborations in web browsers, which leads to less work for web developers.

## CHAPTER 3 –PROPOSED METHODOLOGY

### 3.1-Flow chart:



### **3.2 Algorithm Proposed:**

- Collect our dataset
- Refine it and handle if any errors
- Visualize the dataset using distribution graphs and Sns library
- Find correlation in dataset if any
- Define the model
- Train the model using regression algorithms
- Check model on testing data
- Repeat till above steps to get more accuracy
- If not desirable redefine the model
- If desirable then use this model to predict calories on real values/ scenario
- Keep a track of steps and their results



## **CHAPTER 4-TECHNOLOGY USED**

After taking some consideration, finding some results from the web, discussing with guide and other experts we decided to use the technology which is mainly based on machine learning with the following machine learning algorithms and techniques:-

### **1) Python libraries:-**

- Numpy
- Panda
- Sns
- Seaborn
- XG boost regressor

### **2) Algorithms**

- Linear regression
- XG boost regression

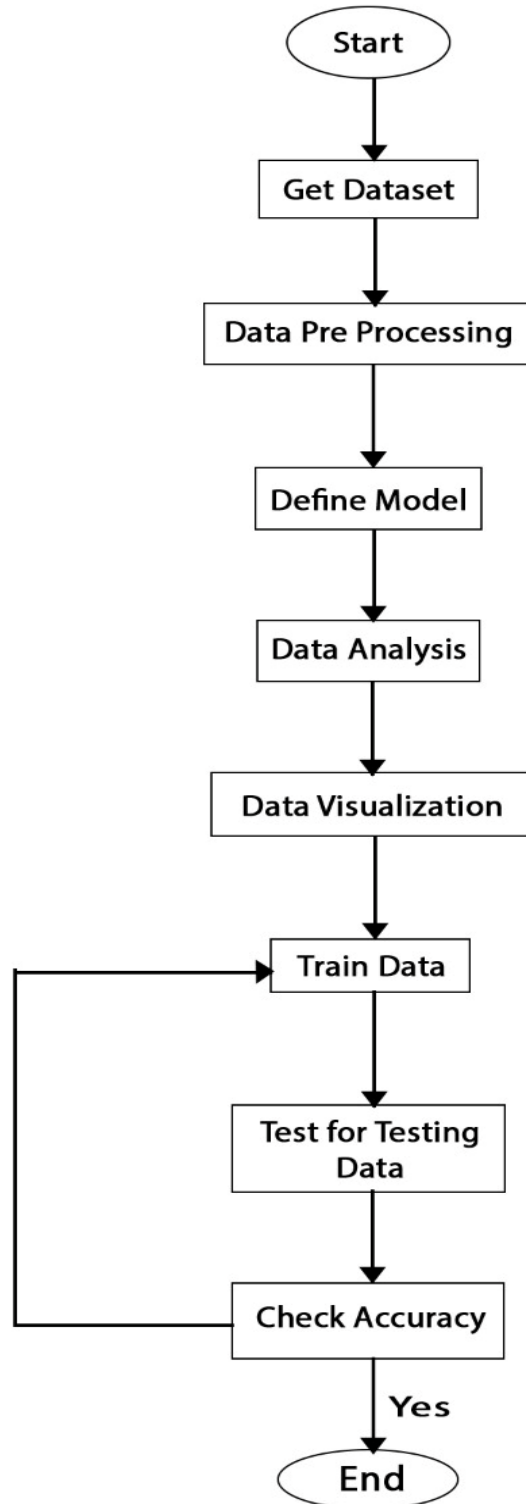
### **3) Support vector machine**

#### **Features of support vector machine:**

- It is the most famous algorithm of XG boost.
- Tianqi-Chan was the founder of XG boost.
- It is platform free.
- It is integrable with multiple systems.
- XG boost has high speed of processing.

## CHAPTER 5- DIAGRAMS

### WORK FLOW:



## **CHAPTER 6-CONCLUSION**

In conclusion of this project, we would conclude that this project will provide a great help to future as already discussed in the future scope section. This project is mainly based on machine learning algorithms and will work as a supervised machine learning calorie burned prediction model. We are using our dataset from the web, mainly from Kaggle. Our main motive of this was to calculate the exact amount of calories burned during workout or physical activity based on certain input parameters of body. This project will be of great help for those who want to keep their fitness good and maintain a healthy body. This project can also be used to provide educational content about the technologies used to concerned people.

## REFERENCES

- [1]. Smola, A., & Vishwanathan, S. V. N. (2008). Introduction to machine learning. Cambridge University, UK, 32, 34.
- [2]. Saltz, J. S., & Stanton, J. M. (2017). An introduction to data science. Sage Publications.
- [3]. Shashua, A. (2009). Introduction to machine learning: Class notes 67577. arXiv preprint arXiv:0904.3664.
- [4]. Learning, M. (1994). Neural and Statistical Classification. Editors D. Mitchie et. al, 350.
- [5]. Mitchell, T. M. (1999). Machine learning and data mining. Communications of the ACM, 42(11), 30-36.
- [6]. Downey, A. B. (2011). Think stats. "O'Reilly Media, Inc."
- [7]. Géron, A. (2019). Hands-On Machine Learning with Scikit- Learn, Keras, and TensorFlow: Concepts, Tools
- [8]. Sona p venoy and Binumon joseph calorie burned prediction linear regression.