INTERNSHIP REPORT

Submitted to

SAVEETHA INSTITUTE OF MEDICAL AND TECHNICAL SCIENCES

In partial fulfillment for the award of the degree of

BACHELOR OF ENGINEERING

IN

COMPUTER SCIENCE AND ENGINEERING

By

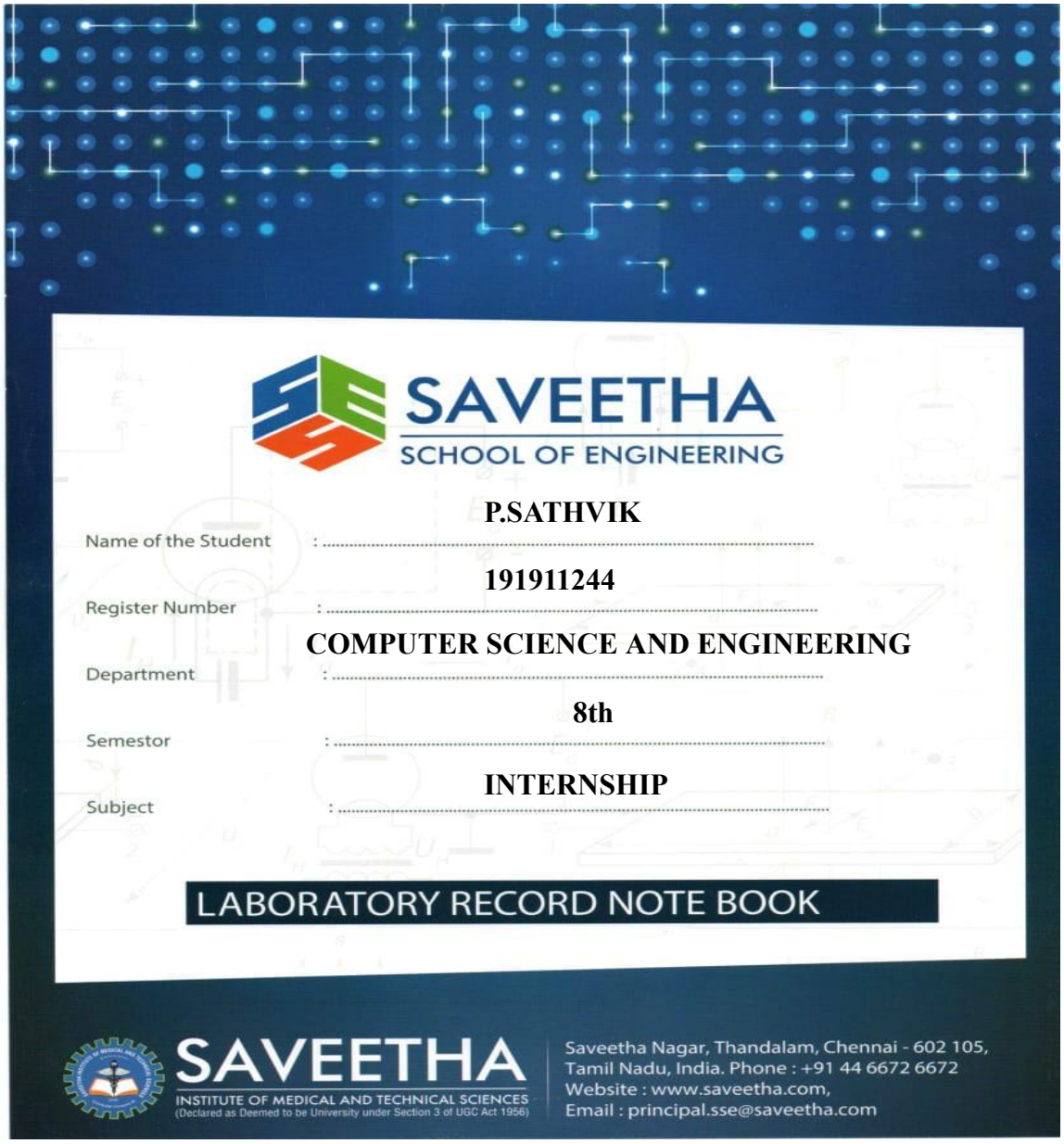
PUTTA SATHVIK
(Reg.No. 191911244)

SIMATS ENGINEERING

SAVEETHA INSTITUTE OF MEDICAL AND

TECHNICAL SCIENCES

CHENNAI - 602105



Dr.S.JOHN JUSTIN THANGARAJ

Head of the Department

Department of Networking

SIMATS Engineering

Saveetha Institute of Medical and

Technical Sciences

Chennai - 602 105

Dr.S.LOGANAYAGI

Internship Coordinator

Department of Networking

SIMATS Engineering

Saveetha Institute of Medical

and Technical Science

Chennai - 602 105

INTERNAL EXAMINER

EXTERNAL EXAMINER

DECLARATION BY THE CANDIDATE

I declare that the report entitled **"Artificial Intelligence and Machine Learning"** submitted by me for the degree of Bachelor of Engineering is the record of the internship work carrying out by me and furthermore this work will not form the basis for the award of any degree or diploma in this or any other University or other similar institution of higher learning.

PUTTA SATHVIK(191911244)

ACKNOWLEDGEMENT

This Internship work would not have been possible without the contribution of many people. It gives me immense pleasure to express my profound gratitude to our Honorable Chancellor **Dr. N. M. Veeraiyan**, Saveetha Institute of Medical and Technical Sciences, for his blessings and for being a source of inspiration. I sincerely thank our Director of Academics **Dr.Deepak Nallaswamy**, SIMATS, for his visionary thoughts and support. I am indebted to extend my gratitude to our Director **Mrs. Ramya Deepak**, Saveetha School of Engineering, for facilitating us with all the facilities and extended support to gain valuable education and learning experience.

I register my special thanks to **Dr. B. Ramesh**, Principal, Saveetha School of Engineering and **Dr.S.John Justin Thangaraj**, Head of the Department, Institute of Electronics and Communication Engineering, for the support given to me in the successful conduct of this project. I wish to express my sincere gratitude to my Internship Coordinator **Dr.S.Loganayagi**, for his inspiring guidance, personal involvement and constant encouragement during the entire course of this work.

I am grateful to Internship Coordinators, Review Panel External and Internal Members and the entire faculty of the Department of Computer Science and Engineering, for their constructive criticisms and valuable suggestions which have been a rich source to improve the quality of this work.

.

P. SATHVIK (191911244)

OFFER LETTER FOR INTERNSHIP

Dated: 10 th March 2021

To,
Putta.Sathvik

Re: Internship Offer Letter

Dear Sathvik,

We are pleased to offer you an internship opportunity as an Artificial Intelligence and Machine Learning Intern with effect from 15th March 2021. At TRIX Education, we provide you with the best training and also with good projects to work on. We strongly believe in Learn and Apply. We are confident that you would play a significant role in the overall success of the company and wish you the most enjoyable, learning packed and truly meaningful experience with TRIX Edu.

You will not be entitled for an annual salary. As part of your engagement, you shall be able to learn and apply very quickly.

Congratulations and welcome to the team!

Yours faithfully
TRIX EDU.
G Sainath Reddy.

TRIX Edu.
KPHB Hyderabad, India
trixsdu.official@gmail.com
+91 9502979770

INTERNSHIP CERTIFICATE

ABSTRACT

A modern business with a focus on artificial intelligence (AI) and machine learning (ML) technology is called TRIX.education. TRIX.education is a leader in developing original solutions to challenging issues as the world gets more and more dependent on technology. TRIX.education is committed to developing intelligent systems that maximize performance, efficiency, and productivity with a team of highly qualified individuals. By utilizing the power of AI and ML, TRIX.education is able to offer its clients innumerable insights and intelligence across a variety of industries. TRIX.education is dedicated to bringing about revolutionary change through the application of AI and ML technologies, whether it be through the creation of predictive models and different algorithms used to develop the projects in this field.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

1.1 COURSE OBJECTIVE

- Assist the student's development of the employer-valued skills such as teamwork, communications and the attention to detail.

- Expose the student to the environment and expectations of performance on the part of accountants in professional accounting practice, private/public companies or government entities.

- Enhance and/or expand the student's knowledge of a particular area(s) of accounting.

- Expose the student to professional role models or mentors who will provide the student with support in the early stages of the internship and provide an example of the behavior expected in the intern's workplace.

- Develop a solid work ethic and professional demeanor, as well as a commitment to ethical conduct and social responsibility.

- Provide training and experiential learning opportunities for the development of skills.

1.2 THE BENEFITS OF INTERNSHIP

1.2.1 Benefits to the Industry

A. Availability of ready to contribute candidates for employment.

B. Year-round source of highly motivated pre-professionals.

C. Visibility of the organization is increased on campus.

D. Quality candidate's availability for temporary or seasonal positions and projects. Freedom for industrial staff to pursue more creative projects.

1.2.2 Benefits to Students

A. An opportunity to get hired by the industry/ organization.

B. Practical experience in an organizational setting. Excellent opportunity to see how the theoretical aspects learned in classes are integrated into the practical world.

C. On-floor experience provides much more professional experience which is often worth more than classroom teaching.

D. Helps them decide if the industry and the profession is the best career option to pursue.

1.2.3. Benefits to the Institute

A. Build industrial relations.

B. Makes the placement process easier.

C. Improve institutional credibility & branding.

D. Helps in retention of the students.

1.3 COMPANY PROFILE

INTRODUCTION:

TRIX.education in Hyderabad, India is established by technology pioneers.
TRIX.education is a dynamic and innovative company that specializes in the development of artificial intelligence (AI) and machine learning (ML) technologies. With an outstanding team of experts and a wealth of experience in the field, TRIX.education is at the forefront of providing cutting-edge solutions. Another area of focus for TRIX.education is in improving decision-making. By leveraging the power of AI and ML, we are able to provide our clients with unparalleled insights and intelligence.

VISION:

Their Vision is to provide training and implement services with an edge of competitiveness in quality & price to valued customers in the ground of sustainability. They promote reputed multi brand residential, commercial and Industrial automation products. Their focus and growth will be on technological outsourcing in the field of Ai and Ml and python development.

MISSION:

Their mission is to be an ultimate solution provider with a reputation for expertise, quality, and cost effectiveness in the world of automation and to advocate sustainable environmental protection by maximizing consolidated effort through unmatchable teamwork.

- Focus on profitability and return on investments over growth.
- Maintain a lean and flexible organization.

COURSE DESCRIPTION ON AI & ML TRAINING:

Today's highly increasing competitiveness over the industry demands high quality and most consistent products with a competitive price. Automation takes a step further mechanization that uses a particular machinery mechanism aided human operators for performing a task.

CHAPTER 2

DEVELOPMENT

2.1 INTRODUCTION

Machine learning is a subfield of artificial intelligence (AI). The goal of machine learning generally is to understand the structure of data and fit that data into models that can be understood and utilized by people. The main focus of ML is to allow computer systems to learn from experience without being explicitly programmed or human intervention. Python is arguably the most popular programming language in the machine learning ecosystem. Top libraries used in data science are Pandas, NumPy, Scikit-learn, Gradio, TensorFlow, Keras, SciPy, Stats models. These libraries offer fast, flexible, and efficient methods for performing data operations. Pandas and Numpy are two popular Python libraries used for data analysis and manipulation tasks.

2.1.1 ANALYZING DATA WITH PYTHON:

- Import data sets.
- Clean and prepare data for analysis.
- Manipulate pandas Dataframe.
- Summarize data.
- Build Machine Learning models

2.2 SOFTWARE TOOLS USED FOR IMPLEMENTATION

1) PYTHON: Python is a high-level, interpreted, interactive and object-oriented scripting language. Python is designed to be highly readable. It uses English words frequently whereas other languages use punctuation, and it has fewer syntactic constructions than other languages.

- Python can be used on a server to create web applications.
- Python can be used alongside software to create workflows.

2) PANDAS: Pandas is an open-source library that is made mainly for working with relational or labeled data both easily and intuitively. It provides various data structures and operations for manipulating numerical data and time series. This library is built on top of the NumPy library. Pandas are also able to delete rows that are not relevant, or contain wrong values, like empty or NULL values. This is called cleaning the data. Pandas is mainly used for data analysis .

3) SCIKIT-LEARN: Scikit-learn is for machine learning development in python. It provides a library for the Python programming language. It helps in data mining and data analysis. It provides models and algorithms for Classification, Regression, Clustering, Dimensional reduction, Model Pre-processing.

4) OPENCV: Python is a general purpose programming language started by Guido van Rossum, which became very popular in a short time mainly because of its simplicity and code readability. It enables the programmer to express his ideas in fewer lines of code without reducing any readability. Compared to other languages like C/C++, Python is slower. But another important feature of Python is that it can be easily extended with C/C++. This feature helps us to write computationally intensive codes in C/C++ and create a Python wrapper for it so that we can use these wrappers as Python modules.

5) ANACONDA: This is a software used for python coding environments. Where different types of

python coding playgrounds are provided like Jupiter notebook and many more.

CHAPTER 3

TASKS PERFORMED

3.1 PROBLEMS/CHALLENGE

Task 1& Task 2: IMPLEMENTATION

Task 3& Task 4: RESULTS AND DISCUSSION

3.2 DURATION OF INTERNSHIP WORK

Table 1 : Work Activities of WEEK1

	DATE	DAY	NAME OF THE TOPIC/MODULE COMPLETED
1st Week	22-03-2021	Monday	Induction program
	23-03-2021	Tuesday	Installation of Tools/apps
	24-03-2021	Wednesday	Basic Understanding of Language
	25-03-2021	Thursday	Brief Introduction to Techniques

Table 2 : Work Activities of WEEK2

	DATE	**DAY**	**NAME OF THE TOPIC/MODULE COMPLETED**
2nd Week	29-03-2021	Monday	Discussion and Task assigned
	30-03-2021	Tuesday	Installation of Tools/apps
	01-04-2021	Thursday	Training in Python
	02-04-2021	Friday	Development of Project Using Python

Table 3 : Work Activities of WEEK3

	DATE	DAY	NAME OF THE TOPIC/MODULE COMPLETED
3rd Week	05-04-2021	Monday	Training of Algorithms
	06-04-2021	Tuesday	Research of Algorithms
	07-04-2021	Wednesday	Analysis of the project
	09 -04-2021	Friday	Task Submission

Table 4 : Work Activities of WEEK4

	DATE	DAY	NAME OF THE TOPIC/MODULE COMPLETED
4th Week	12-04-2021	Monday	Installing prerequisites
	14-04-2021	Wednesday	Integrating the code
	16-04-2021	Friday	Project result and discussion
	22-04-2021	Thursday	Task Submission

3.3 METHODOLOGY

Distribution of the loans is the core business part of almost every bank. The main portion of the bank's assets directly came from the profit earned from the loans distributed by the banks. The prime objective in the banking environment is to invest their assets in safe hands where it is. Today many banks/financial companies approve loans after a regress process of verification and validation but still there is no surety whether the chosen applicant is the deserving right applicant out of all applicants.

- Input Excel Dataset
- Import Libraries
- Dataset Visualization
- processing of the data
- Graphical View output
- Split Train and Test Data
- Machine learning Analysis

LOGISTIC REGRESSION: This is a classification algorithm which uses a logistic function to predict binary outcome (True/False, 0/1, Yes/No) given an independent variable. The aim of this model is to find a relationship between features and probability of particular outcome. The logistic function used is a logit function which is a log of odds in the favor of the event. Logit function develops a s shaped curve with the probability estimate similar to a step function.

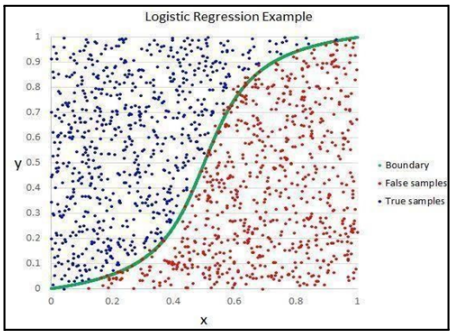
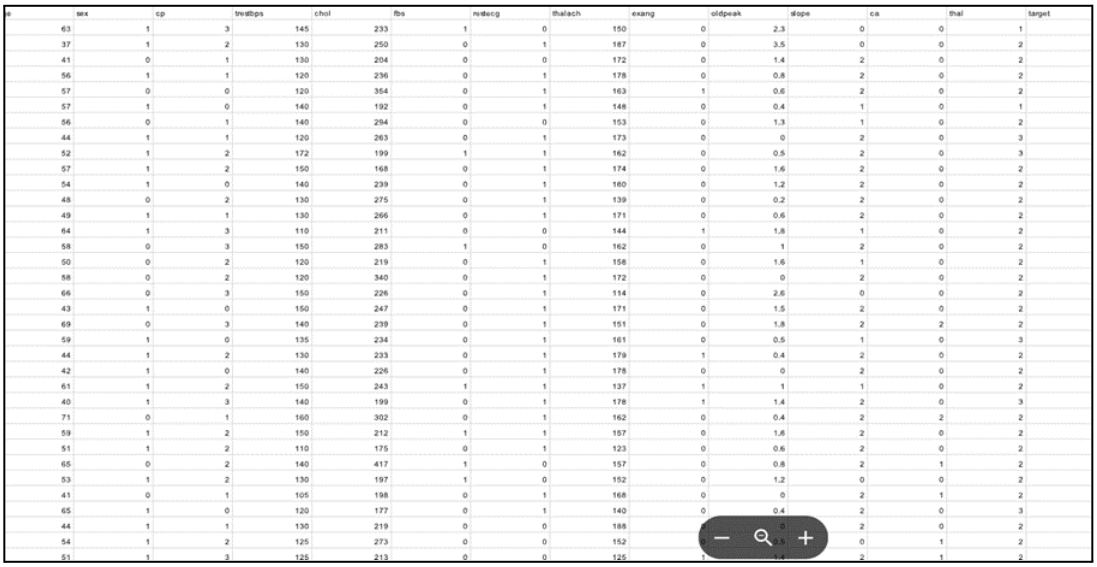


Fig:3.2.1 Logistic Regression



e	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	ca	thal	target
63	1	3	145	233	1	0	150	0	2.3	0	0	1	
37	1	2	130	250	0	1	187	0	3.5	0	0	2	
41	0	1	130	204	0	0	172	0	1.4	2	0	2	
56	1	1	120	236	0	1	178	0	0.8	2	0	2	
57	0	0	120	354	0	1	163	1	0.6	2	0	2	
57	1	0	140	192	0	1	148	0	0.4	1	0	1	
56	0	1	140	294	0	0	153	0	1.3	1	0	2	
44	1	1	120	263	0	1	173	0	0	2	0	3	
52	1	2	172	199	1	1	162	0	0.5	2	0	3	
57	1	2	150	168	0	1	174	0	1.6	2	0	2	
54	1	0	140	239	0	1	160	0	1.2	2	0	2	
48	0	2	130	275	0	1	139	0	0.2	2	0	2	
49	1	1	130	266	0	1	171	0	0.6	2	0	2	
64	1	3	110	211	0	0	144	1	1.8	1	0	2	
58	0	3	150	283	1	0	162	0	1	2	0	2	
50	0	2	120	219	0	1	158	0	1.6	1	0	2	
58	0	2	120	340	0	1	172	0	0	2	0	2	
66	0	3	150	226	0	1	114	0	2.6	0	0	2	
43	1	0	150	247	0	1	171	0	1.5	2	0	2	
69	0	3	140	239	0	1	151	0	1.8	2	2	2	
59	1	0	135	234	0	1	161	0	0.5	1	0	3	
44	1	2	130	233	0	1	179	1	0.4	2	0	2	
42	1	0	140	226	0	1	178	0	0	2	0	2	
61	1	2	150	243	1	1	137	1	1	1	0	2	
40	1	3	140	199	0	1	178	1	1.4	2	0	3	
71	0	1	160	302	0	1	162	0	0.4	2	2	2	
59	1	2	150	212	1	1	157	0	1.6	2	0	2	
51	1	2	110	175	0	1	123	0	0.6	2	0	2	
65	0	2	140	417	1	0	157	0	0.8	2	1	2	
53	1	2	130	197	1	0	152	0	1.2	0	0	2	
41	0	1	105	198	0	1	168	0	0	2	1	2	
65	1	0	120	177	0	1	140	0	0.4	2	0	3	
44	1	1	130	219	0	0	188	0	0	2	0	2	
54	1	2	125	273	0	0	152	0	0.5	0	1	2	
51	1	3	125	213	0	0	125	1	1.4	2	1	2	

Fig 3.3.5: Dataset Features

3.4 PLAN OF WORK

3.4.1 Project requirements:

1. Software conditions defined- (Anaconda)
2. Objects:
    - Different algorithms
    - Selection of algorithms

3.4.2 Plan of project

Resource allocation and work distribution to team members

All members of the team were involved equally in the following phases:

- Algorithm selection phase
- Algorithm research Phase
- Implementation Phase
- Testing Phase

Implementing the solution:

- Writing and implementing code
- Configuring database
- Build user interface
- Integration of different units to create a whole system

Testing the solution:

- Solution was checked for multiple cases with predefined correct results
- Bugs were identified and eradicated

CHAPTER 4

IMPLEMENTATION

4.1 LOGISTIC REGRESSION:

Classification techniques are an essential part of machine learning and data mining applications. Approximately 70% of problems in Data Science are classification problems. There are lots of classification problems that are available, but logistic regression is common and is a useful regression method for solving the binary classification problem.

Basically, Logistic regression measures the relationship between dependent variables and independent variables by estimating the probabilities using a logistic function.

In logistic regression, estimating the probabilities means to predict the likelihood occurrence of the event. For example, the shop owner would like to predict whether the customer who entered into the shop will buy the play station (for example) or not. There would be many features of the customer − gender, age, etc. which would be observed by the shopkeeper to predict the likelihood of occurrence, i.e., buying a play station or not. The logistic function is the sigmoid curve that is used to build the function with various parameters.

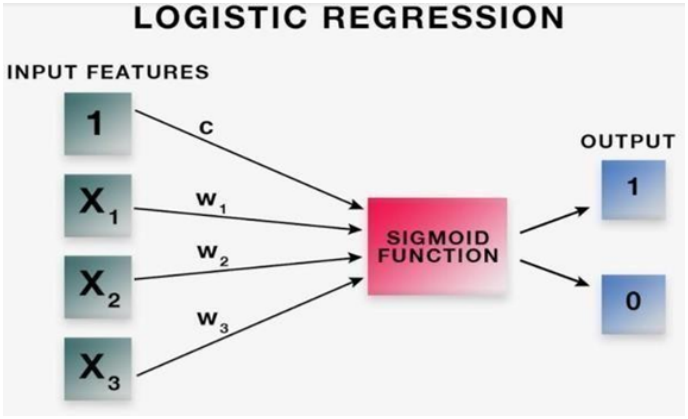


Fig:4.1 Logistic Regression

You must use Python Language in Anaconda Software,in order to load the data and get the analysis.

1) First you need to load the data which contains all the different health parameters which correspond to the person's health.

2) Once the dataset is loaded we need to process the dataset to fit into the machine learning algorithm that we are using for our project.

3) Once the data is processed we need to split our data into two parts

    i) Train Data

    ii) Test Data

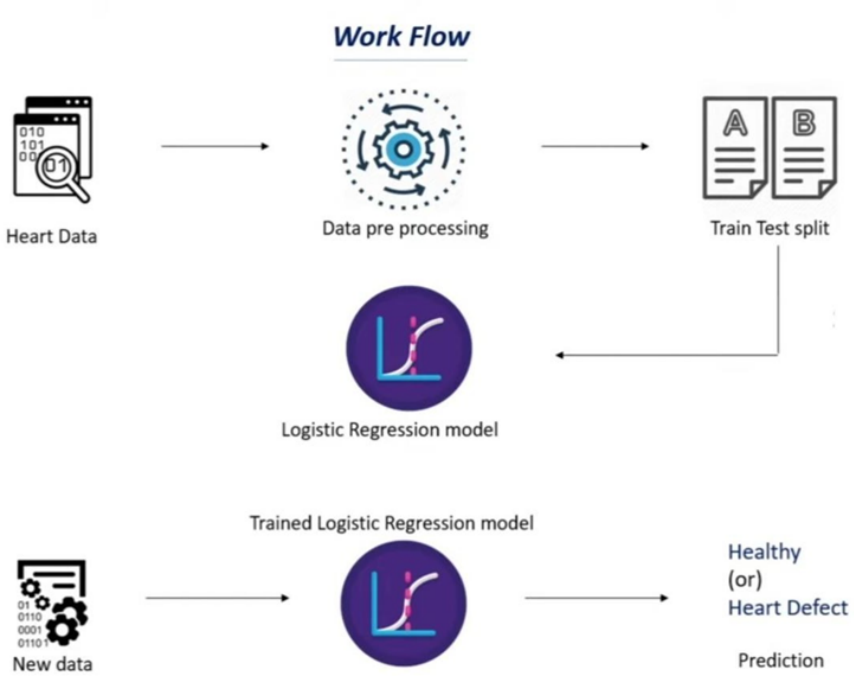
4) We will train our machine learning algorithm using trained data and we will evaluate the performance of the algorithm with test data.

5) Then we will feed our data to our machine learning model called Logistic Regression Model.As this kind of regression is binary classification as it answers in YES or NO.

6) After this we will evaluate using the test data and we will be getting the trained logistic regression model.

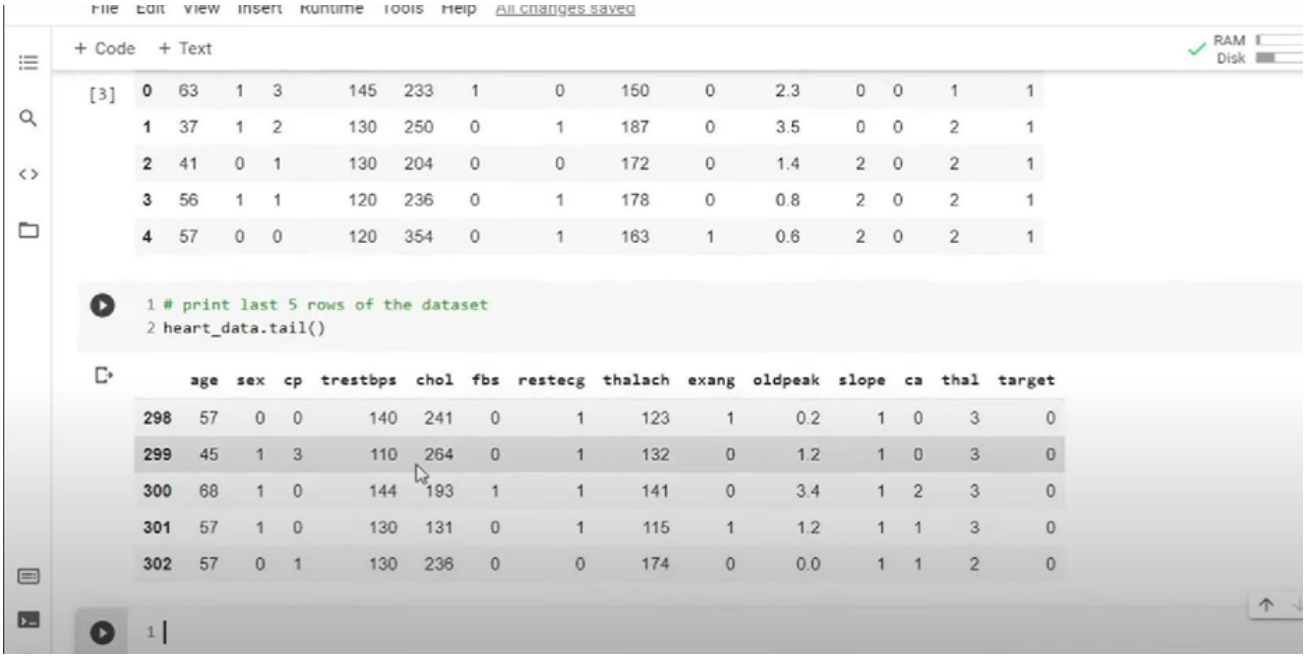
7) After this step when we add the new data about a person with a health scenario it will be predicting whether the person is having the heart disease or not.

4.2 SCREENSHOTS



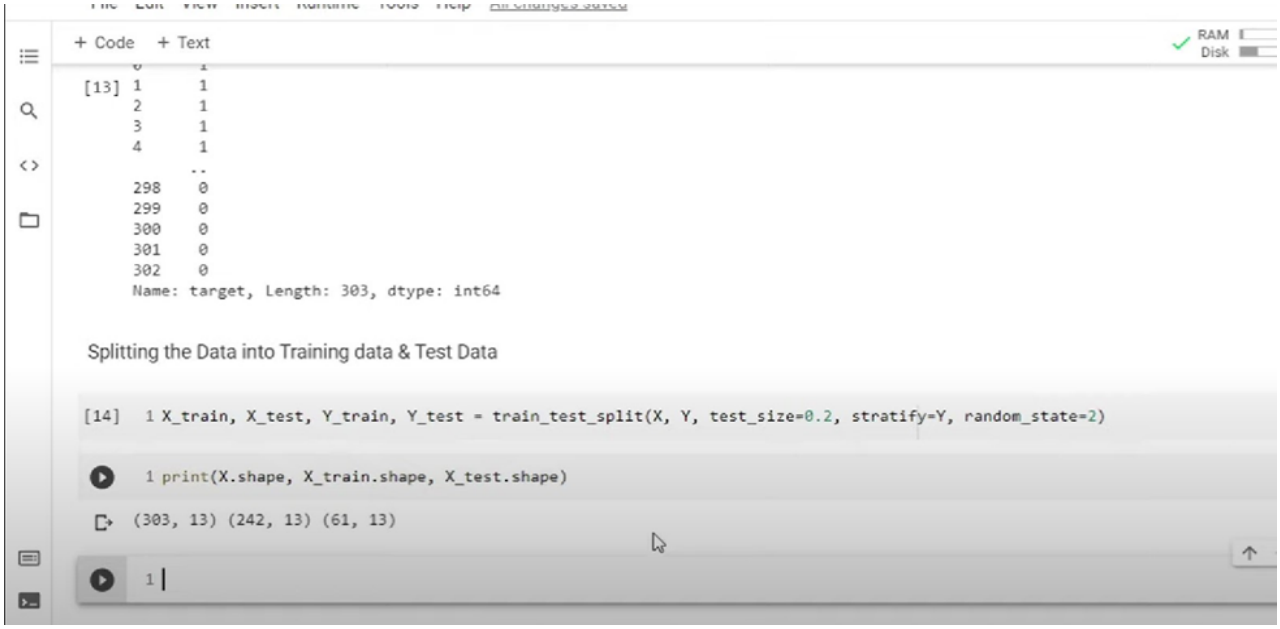
.

Loading the Data Set:

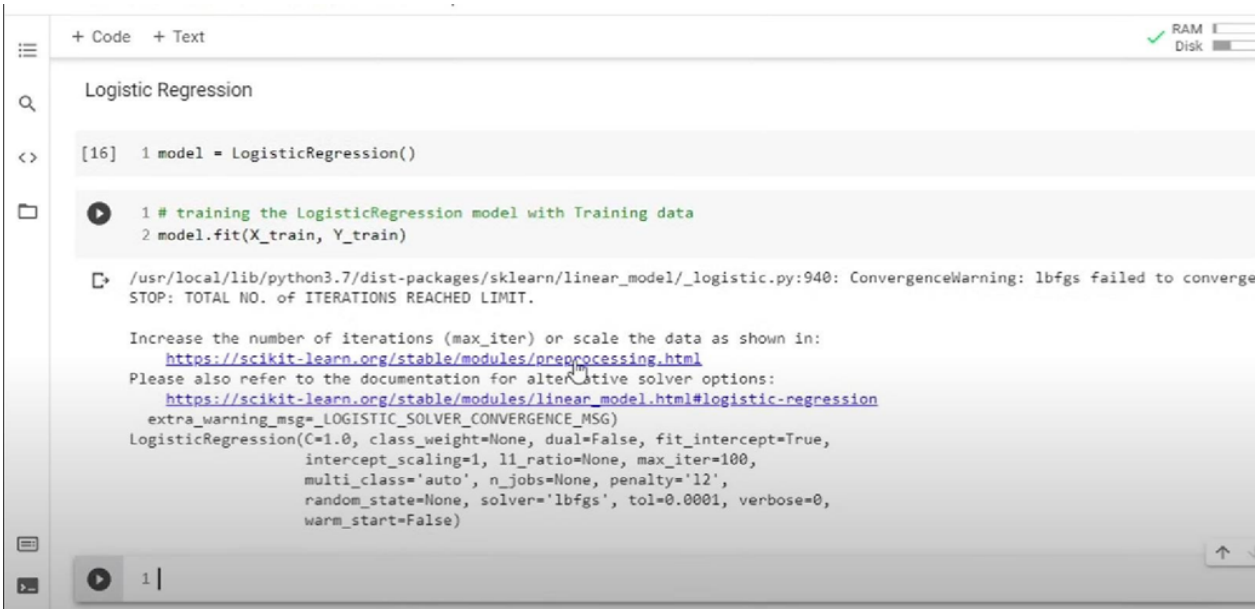


	age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	ca	thal	target
298	57	0	0	140	241	0	1	123	1	0.2	1	0	3	0
299	45	1	3	110	264	0	1	132	0	1.2	1	0	3	0
300	68	1	0	144	193	1	1	141	0	3.4	1	2	3	0
301	57	1	0	130	131	0	1	115	1	1.2	1	1	3	0
302	57	0	1	130	236	0	0	174	0	0.0	1	1	2	0

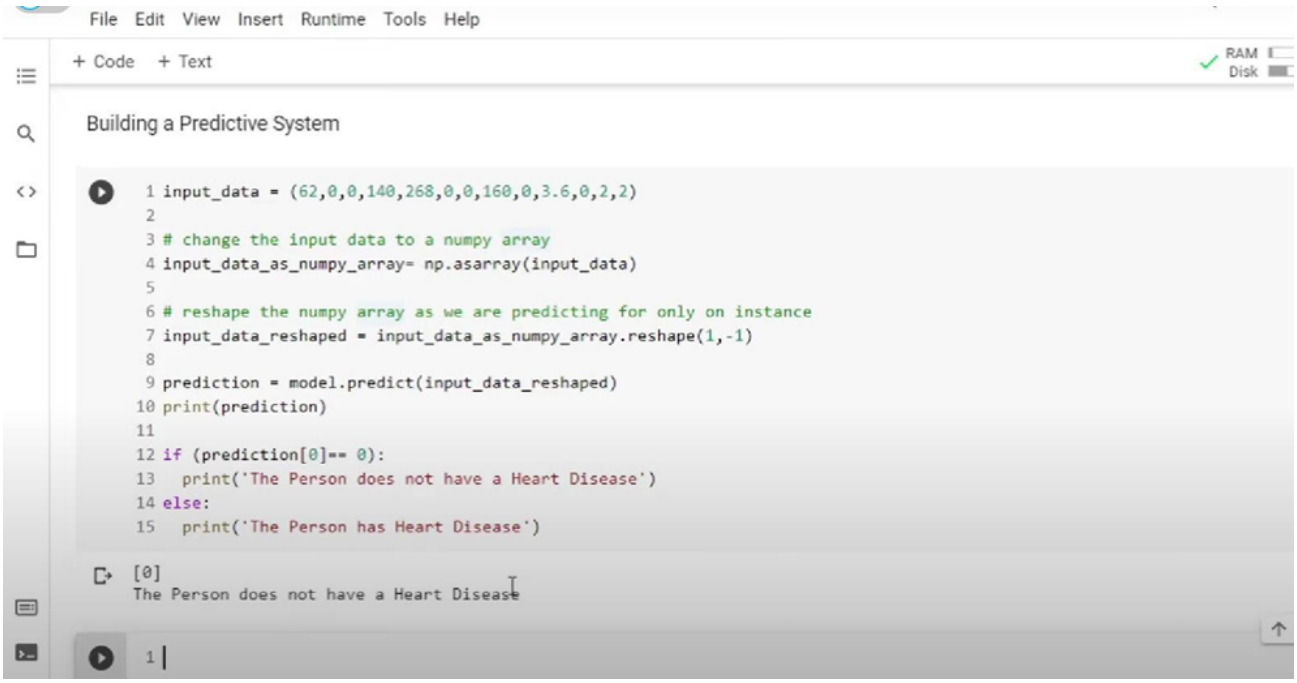
Splitting the data into Trained data and Test data



Model Training- Logistic Regression Model



Checking the Result for the given dataset.



4.3 CONCLUSION

The heart disease prediction project was a fun and fruitful internship task that entailed creating a machine learning model to estimate the chance of developing heart disease based on numerous risk factors. Data preparation, feature selection, model training, and model assessment were some of the major phases that were engaged in the project.

Through the use of machine learning techniques such as logistic regression and random forest, we were able to develop a highly accurate predictive model that could identify individuals at risk of heart disease with a high degree of accuracy. The project demonstrated the potential of machine learning to transform healthcare by enabling earlier detection and prevention of heart disease, one of the leading causes of death worldwide.

Overall, the heart disease prediction project was a challenging internship project that allowed us to develop our skills and contribute to the development of an important model. We are grateful for the opportunity to have worked on this project and thanks to the staff of TRIX.Education for this opportunity to train us in the field of Artificial Intelligence and Machine Learning as an Intern and sharing the required Knowledge.

REFERENCES

[1] https://stackoverflow.com/ (“Stack Overflow - Where Developers Learn, Share, & Build Careers”)

[2] https://www.geeksforgeeks.org/ (“GeeksforGeeks”)

[3]https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LogisticRegression.html (“sklearn.linear_model.LogisticRegression”)

[4] https://www.javatpoint.com/logistic-regression-in-machine-learning (“Website,”)

“GeeksforGeeks.” GeeksforGeeks. Accessed June 24, 2023. https://www.geeksforgeeks.org/.

“sklearn.linear_model.LogisticRegression.”Scikit-Learn. Accessed June 24, 2023. https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LogisticRegression.html.

“Stack Overflow - Where Developers Learn, Share, & Build Careers.” Stack Overflow. Accessed June 24, 2023. https://stackoverflow.com/.

“Website.” https://www.javatpoint.com/logistic-regression-in-machine-learning.