

# VISVESVARAYA TECHNOLOGICAL UNIVERSITY

JANASANGAMA, BELAGAVI – 590018



**Internship (18CSI85)**

On

**“Water Portability Detection using Machine Learning”**

*Submitted in partial fulfillment for the award of the degree of*  
**Bachelor of Engineering**

In

**INFORMATION SCIENCE AND ENGINEERING**

Submitted by:

**SATHVIK I K**

**1BI20IS083**

Internship carried out

at

**PRINSTON SMART ENGINEERS.**

**Vishwapriya Nagar, Begur, Bengaluru, Karnataka - 560068**

**Internal Guide:**

**Mrs. Pavithra N**

Assistant Professor

Dept. of ISE, BIT



**External Guide:**

**Mr. Akash V**

Project Manager

Prinston Smart Engineers



**BANGALORE INSTITUTE OF TECHNOLOGY**

**V.V. Puram, K.R. Road, Bengaluru - 560004**

**Department of Information Science & Engineering**

**2023-2024**



**BANGALORE INSTITUTE OF TECHNOLOGY**  
**V.V. Puram, K.R. Road, Bengaluru - 560004**  
**Department of Information Science & Engineering**



## **CERTIFICATE**

Certified that the internship carried out by **Mr. SATHVIK I K** bearing USN **1BI20IS083** a bonafide student of VIII Semester B.E., **BANGALORE INSTITUTE OF TECHNOLOGY** in partial fulfillment of Bachelor of Engineering in **INFORMATION SCIENCE AND ENGINEERING** of **VISVESVARAYA TECHNOLOGICAL UNIVERSITY**, Belagavi during the year 2023-2024. It is certified that all corrections / suggestions indicated for Internal Assessment have been incorporated in the report. The internship report has been approved as it satisfies the academic requirements in respect of internship work prescribed for the said Degree.

**Internal Guide:**

**Mrs. Pavithra N**

Assistant Professor

Dept. of ISE, BIT

**HOD:**

**Dr. Asha T**

Professor & Head

Dept. of ISE, BIT

Name of the Examiners:

Signature with Date:

1)

2)

## **DECLARATION**

I, **Mr. SATHVIK I K** bearing USN **1BI20IS083** student of VIII semester, Bachelor of Engineering in **INFORMATION SCIENCE AND ENGINEERING, Bangalore Institute of Technology**, Bangalore hereby declare that the internship has been carried out at **PRINSTON SMART ENGINEERS, Vishwapriya Nagar, Begur, Bengaluru**. I further declare that the matter embodied in this report has not been submitted previously to any institution or university for the award of any other Degree / Diploma Certificate.

**NAME: SATHVIK I K**

**USN: 1BI20IS083**

**Place: Bangalore**

## ACKNOWLEDGEMENT

I express my heartiest gratitude with great pleasure to **Bangalore Institute of Technology, Bangalore** that provided me an opportunity to fulfill my cherished desire to attain my goal.

I extend my thanks to our Principal **Dr. Aswath M U**, for encouraging me in all aspects to complete the internship.

I would like to thank **Dr. Asha T**, Professor & Head, Dept. of ISE, who continually helped me with her suggestions and ideas.

I immensely thank **Mr. Akash V**, Software Engineer, Prinston Smart Engineers, for giving me an opportunity to work at the project.

I immensely thank **Mrs. Pavithra N**, Associate Professor, Dept. of ISE, in supporting and guiding me in carrying out my preparations for the internship.

Finally, I thank one and all who have helped me directly or indirectly in the completion of the internship.

**SATHVIK I K**

## EXECUTIVE SUMMARY

Water Potability Detection project focused on developing predictive models to assess water potability based on various physicochemical properties. Three machine learning algorithms were employed: Logistic Regression, Support Vector Machine (SVM), and Random Forest Classifier. The dataset, comprising water quality data, underwent initial preprocessing, including feature removal for missing values and standard scaling for normalization. The dataset was split into training (85%) and testing (15%) sets.

The models were trained and evaluated using accuracy scores and confusion matrices. Logistic Regression achieved an accuracy of 0.62, while SVM and Random Forest Classifier achieved 0.63 and 0.62. Hyperparameter tuning was performed on the SVM model using grid search to optimize parameters like C, kernel type, degree, and gamma.

Overall, the models contributed valuable insights into water potability prediction. SVM, with its optimized parameters, exhibited the highest potential for accurate prediction. Future directions could involve feature engineering, experimentation with other ensemble methods like Gradient Boosting, and acquiring more diverse datasets for further model refinement.

## TABLE OF CONTENTS

CHAPTER	DESCRIPTION	PAGE No.
<b>1</b>	<b>ABOUT THE COMPANY</b>	<b>01</b>
	1.1 History of the organization	01
	1.2 Ownership	01
	1.3 Sector	01
	1.4 Customers of the company	02
	1.5 Structure of the company	02
	1.6 Services of the company	03
<b>2</b>	<b>ABOUT THE DEPARTMENT</b>	<b>05</b>
	2.1 About the Working Department	05
	2.2 Objectives of the Department	05
	2.3 Organizational procedures	06
<b>3</b>	<b>TASKS ASSIGNED AND PERFORMED</b>	<b>07</b>
	3.1 Technical Activity performed in the Company	07
	3.2 Requirement Specification	07
	3.2.1 Hardware Requirements	07
	3.2.2 Software Requirements	08
	3.3 Task assigned and performed	08
	3.3.1 Swot analysis	09
	3.4 Artificial Intelligence	09
	3.4.1 What is Artificial Intelligence?	10
	3.4.2 Advantages of Artificial Intelligence	10
	3.5 Python	11

	3.5.1 Advantages of python	11
	3.6 Basic training	13
	3.7 Machine training	14
	3.7.1 What is machine learning?	14
	3.7.2 Need for machine learning	14
	3.7.3 Machine learning algorithms	14
	3.8 Introduction of project	16
	3.8.1 Problem statement	16
	3.8.2 Implementation	17
	3.8.3 Code snippet	19
	3.8.4 Dataset	22
	3.8.5 Results	22
<b>4</b>	<b>REFLECTION NOTES</b>	<b>26</b>
	4.1 Technical Outcomes	27
	4.2 Non – Technical Outcomes	27
	4.2.1 Communication Skills	27
	4.2.2 Time Management	27
	4.2.3 Problem Solving	27
	4.2.4 Personality Development	28
	<b>REFERENCES</b>	<b>29</b>

## LIST OF FIGURES

Figure No.	Description	Page No.
Figure 1.1	Organizational Structure	02
Figure 3.1	SWOT Analysis	09
Figure 3.2	Architecture of the Project	17
Figure 3.3	Dataset	22
Figure 3.4	Data frame and Dataset information	22
Figure 3.5	Model Evaluation	22
Figure 3.6	Logistic Regression Confusion Matrix	23
Figure 3.7	SVM Confusion Matrix	23
Figure 3.8	Random Forest Confusion Matrix	24
Figure 3.9	Hyperparameter tuning for SVM	24
Figure 3.10	Test Confusion Matrix	25



## LIST OF TABLES

Table No.	Description	Page No.
Table 3.1	Hardware Requirements	07
Table 3.2	Software Requirements	08
Table 3.3	Internship Timeline	08

