

**Water Quality predicting**

Programming For Data Science

2024-2025

3rd year Engineering’s Degree in Data Science

Department of Applied Mathematics and Statistics

Institute of Technology of Cambodia

**Lecturer:** Mr. PHAUK Sokkey (Courses)

Mr. PEN Chentra (TP)

Members of Group 03 Student ID

1. NY Chantharith e20220472
2. SOEUK Bondol e20221592
3. SAMBATH Seakty e20220517
4. PHAN Wayotep e20220850
5. NGOUN Lyhorng e20221532

Water Quality predicting

Contents

1. Abstract
2. Introduction
3. Data Collection
4. Exploring Data Analysis
   1. Data Cleaning
   2. Data Visualization
   3. Data Training
5. Model Construction
6. Conclusion
7. Reference
8. Abstract

Water quality is essential for human health, agriculture, and the environment. However, monitoring and testing water quality can be expensive and time-consuming. This study explores how data analysis and machine learning can be used to predict water quality more efficiently. By using data such as pH levels, turbidity, dissolved oxygen, and other key factors, we trained a machine learning model to classify water as safe or unsafe for various uses. The results showed that machine learning can accurately predict water quality based on patterns in the data, helping to identify contamination early and improve decision-making. This approach is cost-effective and can be applied in areas with limited resources for water testing. It also supports environmental sustainability by promoting better water management practices. This work demonstrates the potential of technology to address critical challenges in water quality monitoring and protection.

1. Introduction

Water is one of the most important resources for life. Clean water is essential for drinking, farming, and keeping our environment healthy. However, water can sometimes become polluted due to chemicals, waste, or harmful substances, making it unsafe to use. Testing water quality is important, but it can be expensive and take a lot of time, especially in areas with limited resources. This project focuses on using modern technology, like data analysis and machine learning, to make water quality testing faster and easier. Machine learning is a type of computer program that learns from data to make predictions or decisions. By analyzing data such as pH, turbidity, and other water quality indicators, these programs can predict whether water is safe to use or not. The goal of this project is to create a simple, accurate, and cost-effective way to monitor water quality. This can help protect people’s health, support sustainable water use, and prevent environmental problems.