



HERALD
COLLEGE
K A T H M A N D U

UNIVERSITY PARTNER



UNIVERSITY OF
WOLVERHAMPTON

Proposal for Classification Task

Student Id : 2461859
Student Name : Ayusha Karki
Group : L5CG11
Module Leader : Siman Giri
Tutor : Robin Tuladhar
Submitted on : 27-01-2026
Word Count : 418 words

Abstract

The project deploys the classification methods to evaluate the water quality based on physicochemical water parameters. It is found on the Brisbane Water Quality Dataset that describes the way in which the key environmental indicators are used to classify water samples as safe and unsafe. The project presents a systematic machine learning project, which entails data preprocessing, exploratory data analysis, and model evaluation. The findings will be used to facilitate water quality monitoring and sustainable management of water resources.

Table of Contents

1. INTRODUCTION.....	1
2. CLASSIFICATION TASK	2
2.1 OBJECTIVES OF THIS PROPOSAL	2
3. PROPOSAL REQUIREMENTS FOR CLASSIFICATION TASK	3
3.1 RESEARCH QUESTION.....	3
3.2 DATASET DESCRIPTION	3
3.3 CONNECTION TO A SUSTAINABLE DEVELOPMENT GOAL (SDG).....	3
4. CONCLUSION	4

1. Introduction

Clean and safe water are key elements to environmental and human health. The methods of water quality monitoring by traditional methods may be time consuming and resource intensive. The machine learning classification can provide a feasible substitute because it automatically recognizes dangerous water conditions in terms of quantifiable parameters. The given project aims at categorizing water quality based on physicochemical characteristics to be able to identify possible problems with water safety much quicker and be able to promote active environmental control.

2. Classification Task

2.1 Objectives of this proposal

This proposal aims to summarize the intended final project of the course Concepts and Technologies of AI, which is based on a classification. This project is aimed at categorising environmental conditions based on real-life data in accordance with one of the United Nations Sustainable Development Goals (SDG). The project will implement the supervised classification methods to determine water quality and go through organized machine learning pipeline.

3. Proposal Requirements for Classification Task

3.1 Research Question

What are the ways through which machine learning models can be used to determine whether water quality is safe or unsafe using physicochemical water quality parameters?

3.2 Dataset Description

The data used in this classification activity is Brisbane Water Quality Dataset, which is a data retrieved through sources of environmental monitoring. The data set has about 30000 records and 20 columns.

The main independent variables are temperature, pH, dissolved oxygen, salinity, turbidity and a concentration of chlorophyll. Water Quality Status (Safe/Unsafe) is the target variable, and thus it is a binary classification problem.

The preprocessing issues to be anticipated are the management of missing values, elimination of redundant quality indicator columns, and the potential imbalance in classes of the target variable.

3.3 Connection to a Sustainable Development Goal (SDG)

This project can be aligned to SDG 6: Clean Water and Sanitation since categorizing water quality of water aids in the detection of the water quality of water that is not safe and therefore aid in proper management of water resources to achieve access to clean water and safe water.

4. Conclusion

The presented classification-based project shows how well machine learning can be used to assess the water quality conditions. The project helps in the sustainable management and protection of water and the environment by correctly identifying safe and unsafe water samples. The results of this research correspond to the global goals of sustainability and they indicate the importance of using the data in making decisions that will guarantee the availability of clean water.