

# Prediction of Different Water Quality Parameters using Regression Model

Paragkanti Chattopadhyay<sup>1</sup>, Sumit Banerjee<sup>2</sup>, Chandan Kumar Chanda<sup>3</sup>, Debasis Guha<sup>4</sup>, Sourav Bhattacharya<sup>5</sup>

<sup>1</sup>Department of Computer Science and Engineering, Dr. B. C. Roy Engineering College, Durgapur

<sup>2</sup> Department of Electrical Engineering, Dr. B. C. Roy Engineering College, Durgapur

<sup>3</sup>Department of Electrical Engineering, IEST Shibpur

<sup>4</sup>Department of Master of Computer Applications (MCA), Dr. B. C. Roy Engineering College, Durgapur

<sup>5</sup>Department of Basic Science and Humanities, Dr. B. C. Roy Engineering College, Durgapur

**Abstract:** This paper analyses water pollution data focusing on Biochemical Oxygen Demand (BOD), Chemical Oxygen Demand (COD), pH, and turbidity. Each parameter has specific acceptable ranges to ensure that water is safe and suitable for consumption. BOD, COD, pH and turbidity levels using a Comma separated value (CSV) dataset. Initially, the data is pre-processed by cleaning it to handle missing values and outliers, ensuring consistency and accuracy. The dataset was then divided into training and testing subsets. For modelling, linear regression of machine learning will be employed. The model was trained on the processed data and evaluated using metrics such as  $R^2$  score and Root Mean Square Error (RMSE) to assess pollutants label of BOD, COD, pH and turbidity of Damodar river flowing through Durgapur with some predefined threshold values based on past experience. Finally, the above analysis will be interpreted to understand the patterns and factors influencing pollution levels of BOD, COD, pH and turbidity and the results are found to be in very good agreement.

**Keywords:** Water Pollution, Biochemical Oxygen Demand (BOD), Chemical Oxygen Demand (COD), pH, turbidity, Machine Learning.

## 1. Introduction

The Damodar River, a vital water resource for communities in West Bengal, has been facing significant water pollution due to rapid industrialization and urbanization in the regions of Durgapur. This study investigates the sources, extent, and impacts of water pollution in these areas, with a focus on industrial effluents, untreated sewage, and mining activities. Water samples were collected from key sites along the river to analyze parameters such as pH, dissolved oxygen (DO), chemical oxygen demand (COD), and heavy metals. The results revealed alarming concentrations of pollutants, with some areas showing a drastic decline in water quality, making it unsafe for human consumption and aquatic life. The study also highlights the socio-economic implications of water pollution on local communities, particularly in terms of public health and livelihood and cultivation. The paper emphasizes the urgent need for comprehensive water management policies, stricter regulations on industrial discharges, and community awareness programs to mitigate the ongoing pollution crisis in the Damodar River. The Damodar River, often referred to as the "Sorrow of Bengal," has played a pivotal role in the socio-economic development of the region. Flowing through the industrial belt of West