IMPACTS OF CLIMATE CHANGE ON SURFACE WATER ANALYSIS USING MACHINE LEARNING

Project Reference No.: *47S_BE_1124*

College ***:*** *Channabasaveshwara Institute of Technology, Tumakuru*
Branch ***:*** *Department of Civil Engineering*
Guide(s) ***:*** *Mrs. Ramya H N & Mr. Prathap G P*
Student(S) ***:*** *Mr. Surag*
Mr. G N Deepu Patel
Ms. Niharika Roy H
Mr. Varun G S

Keywords:

Machine Learning, Model Selection and Training, Correlation Matrix, Interpretation and Visualization, Water Quality Monitoring, Decision Support Systems, Adaptation Strategies, Graphical Users Interface

Introduction:

Climate change poses significant challenges to the management and sustainability of surface water resources worldwide. Alterations in precipitation patterns, temperature regimes, and land use practices are leading to profound shifts in surface water dynamics, impacting water availability, quality, and distribution. Addressing these challenges requires advanced analytical tools capable of comprehensively assessing the complex interactions between climate change and surface water systems. In recent years, machine learning has emerged as a promising approach for analysing and predicting the impacts of climate change on surface water resources.

Machine learning techniques offer the capability to process large, multidimensional datasets and uncover intricate relationships between climatic variables and surface water dynamics. By leveraging algorithms such as neural networks, decision trees, and ensemble methods, machine learning enables the extraction of valuable insights from vast amounts of observational and modeling data. These insights are essential for enhancing our understanding of the complex mechanisms driving changes in surface water systems under the influence of climate change.

This introduction sets the stage for exploring the intersection of climate change, surface water analysis, and machine learning. It highlights the pressing need to develop innovative approaches for assessing and managing surface water resources in a changing climate. By integrating machine learning techniques into surface water analysis, researchers and practitioners can unlock new opportunities for informed decision-making, adaptive management, and the sustainable stewardship of water resources in the face of ongoing environmental change.

Objectives:

- To assess the influence of climate change on surface water on a study area by using machine learning.
- To examine the effects of climate change on surface water quality by using water quality parameters.

Methodology:

Location: Bugudanahalli Lake is situated in the Tumkur district, which is known for its semi-arid climate. It lies within the Devarayanadurga State Forest, surrounded by hills and lush greenery.

Ecological Importance: The lake contributes to the local ecosystem by providing habitat for various flora and fauna species. It supports diverse aquatic life and serves as a breeding ground for several migratory birds.

Water Resource: The lake is a vital water source for the surrounding villages, supporting agricultural activities and meeting the domestic water needs of the local communities.

Data Collection

Data Collection in the analysis of impact of climate change on a surface water involves gathering comprehensive and diverse datasets that capture various aspects of the water system. The collected data serves as foundation foer Machine Learning models and other analytical tools to asses changes, Identify patterns, and Predict future scenarios through OPEN DATA PLATFORMS. Here are the key considerations for data collections.

- Hydrological Data
- Water Quality Data
- Climate Data
- Land use and Land cover Data
- Geospatial Data
- Ecological Data
- Historical Data
- Citizenship Data

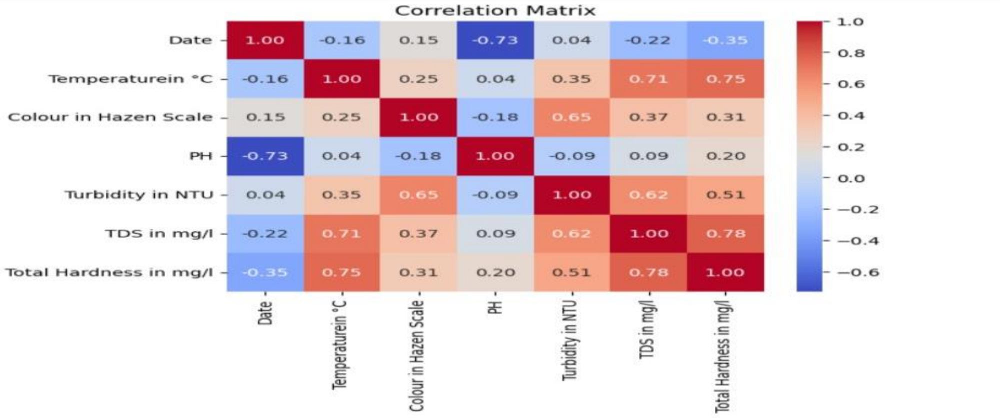
Data Analysis

Analyzing the impact of climate change on surface water using Machine Learning involves processing and interpreting the collected data to uncover patterns, common trends and relationship. The goal is to gain insights into how climate changes to influences surface water systems and to develop predictive models for future scenarios. Here is an overview on data analysis:

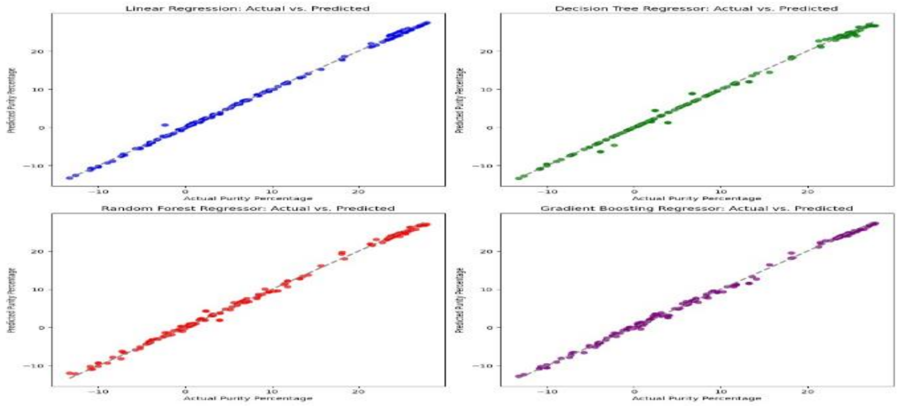
- Data preprocessing
- Exploratory data analysis
- Feature selection
- Model selection

- Training model
- Model evaluation
- Uncertainty analysis
- Graphical User Interface

The data science workflow involves several key steps: data preprocessing, exploratory data analysis (EDA), feature selection, model selection, training, and evaluation. Uncertainty analysis assesses the reliability of model predictions, while a Graphical User Interface (GUI) provides an intuitive platform for interacting with the model. These steps ensure the development of accurate and reliable models for various applications, from predicting customer behavior to analyzing financial data.



Correlation matrix between various physical parameters



Model Selection

	Date	Temperaturein °C	Colour in Hazen Scale	PH	Turbidity in NTU	TDS in mg/l	Total Hardness in mg/l	Water Edible
0	2023-01-02	28.5	10	8.180	13.8	110.8	62	1
1	2023-01-03	28	10	8.559	13.4	109.4	62	1
2	2023-01-04	28.3	10	8.661	12.1	111.8	62	1
3	2023-01-05	27.4	10	8.448	12.9	107.8	62	1
4	2023-01-06	27.4	10	8.541	13	106.9	62	1
...	...	...	...	...	...	...	...	...
248	2023-11-22	27.9	15	7.620	18.8	108.6	60	1
249	2023-11-23	27.6	15	7.480	20.5	108.8	62	0
250	2023-11-24	28.2	15	7.320	20.7	108.8	62	0
251	2023-11-25	27.5	15	7.480	20.1	108.6	60	0
252	2023-11-27	27.4	15	7.191	20.8	108.9	60	0

Water Edibility and its Percentage

Graphical User Interface

Conclusion:

Data Description

Water quality parameters such as pH, temperature, turbidity, total dissolved solids (TDS), and hardness offer vital insights into water conditions, impacting aquatic life and human use. Monitoring and predicting fluctuations in these parameters enable proactive management, aiding in pollution control, ecosystem conservation, and sustainable water resource utilization. By anticipating changes, stakeholders can make informed decisions to mitigate potential issues and ensure the long-term health of freshwater ecosystems for generations to come.

Correlation Matrix

The correlation matrix reveals a robust positive correlation (0.78) between Total Hardness and TDS, indicating their mutual influence due to shared mineral and salt content. Conversely, a weak negative correlation (-0.18) between pH and Color suggests their independence, with higher pH levels associated with lighter coloration. These findings highlight potential indicators for water purity: monitoring either Total Hardness or TDS could reflect changes in the other, while pH and color appear influenced by distinct factors. Such insights aid in understanding the interplay of water quality parameters and inform effective monitoring strategies for maintaining water purity.

Model Evaluation and Selection

The linear regression model exhibits the highest accuracy score among the evaluated models, indicating its superior fit to the data and precise prediction of water purity percentages. While deviations from ideal predictions may occur, the graph demonstrates a strong correlation between actual and predicted values, showcasing the model's effectiveness in capturing underlying data patterns. This accuracy is pivotal for informed decision-making in water quality management. The success of linear regression highlights its utility as a predictive tool, although comparative analysis with other models offers further insights into modeling approaches and their suitability for specific datasets and applications.

Purity Percentage

Limited data in the linear regression model may cause predicted water purity percentages to fall outside acceptable limits, as the model relies on available data to establish relationships. Insufficient data for certain parameter combinations can hinder accurate generalization, leading to unreliable predictions. Additionally, the model's assumption of linear relationships may not fully capture the complexity of the data, causing discrepancies between predicted and actual purity percentages.

Scope for future work:

1. **Integration with Climate Models**: Integrating machine learning models with climate models can provide more robust predictions by capturing the complex interactions between climate variables and surface water dynamics.
2. **Automation and Scalability**: Implementations may prioritize the development of automated ML pipelines that can handle large volumes of data, enabling scalable analysis across different geographical regions.
3. **Collaborative Platforms**: Building collaborative platforms that facilitate knowledge sharing and collaboration among researchers, policymakers, and stakeholders can accelerate the development and deployment of machine learning solutions for surface water analysis in the context of climate change.