**PHASE 4 PROJECT SUBMISSION**

## WATER QUALITY ANAYSIS

**TEAM MEMBERS:**

1. Aazam Khan - 2021504002
2. Sasmitha A – 2021504038
3. Swetha K – 2021504051
4. Gokul R – 2021504515

**PROJECT DEFINITION:**

The project involves analyzing water quality data to assess the suitability of water for specific purposes, such as drinking. The objective is to identify potential issues or deviations from regulatory standards and determine water potability based on various parameters. This project includes defining analysis objectives, collecting water quality data, designing relevant visualizations, and building a predictive model.

**PHASE OBJECTIVE:**

In this technology projects you will continue building your project by performing different analysis, model building and evaluation as per the project requirement. Perform different analysis and visualization using IBM Cognos.

**Dataset Link:**[**https://www.kaggle.com/datasets/adityakadiwal/=water-potability**](https://www.kaggle.com/datasets/adityakadiwal/=water-potability)

**Synopsis:**

In this phase, models such as Isolation forest, Smote , Random forest are used for water quality analysis.

**Cognos Analytics:**

**Isolation Forest:**

A screenshot of a computer screen

Description automatically generated

A screenshot of a computer code

Description automatically generated

A blue and red dots

Description automatically generated

**SMOTE:**

**RANDOM FOREST:**

**A screenshot of a computer code

Description automatically generated**

**A screenshot of a computer code

Description automatically generated**

**A screenshot of a computer

Description automatically generated**

**A screenshot of a computer

Description automatically generated**

**A graph showing different colored lines

Description automatically generated**

**A screen shot of a computer

Description automatically generated**

**A graph of blue bars with white text

Description automatically generated**

**A graph with blue bars

Description automatically generated with medium confidence**

**A screenshot of a computer program

Description automatically generatedA screenshot of a computer program

Description automatically generated**

**A screenshot of a computer program

Description automatically generated**

**A screenshot of a computer code

Description automatically generated**

**A black and white squares with white text

Description automatically generated**

**FUTURE SCOPE:**

The water quality analysis project has reached a substantial milestone by diligently addressing data cleaning, preprocessing, and employing a variety of models such as Isolation Forest, Random Forest, SMOTE, and Logistic Regression for anomaly detection and classification. To further advance this project, several promising future directions can be explored.

Firstly, improving model performance is essential. Advanced machine learning techniques like deep learning, ensemble methods, or state-of-the-art anomaly detection algorithms may be considered to enhance accuracy and generalization on complex water quality datasets. Fine-tuning hyperparameters and optimizing feature selection could also play a crucial role.

Secondly, real-time data integration can provide timely insights into water quality. Incorporating streaming data sources or sensor networks could enable continuous monitoring and immediate detection of anomalies, allowing for quicker response to potential contamination events.

Moreover, the integration of Internet of Things (IoT) devices can revolutionize data collection and analysis. Implementing sensors and smart devices for real-time data acquisition, alongside automated data processing, can enhance the project's efficiency and accuracy.

Collaboration with environmental agencies and local authorities is another avenue to consider. Sharing findings and insights with relevant stakeholders can lead to more effective water quality management strategies and quicker response to pollution incidents.

In summary, the project's future scope entails harnessing advanced machine learning techniques, embracing real-time data streams, incorporating IoT technology, and collaborating with relevant authorities. By doing so, this project can significantly contribute to the preservation and sustainable management of precious water resources, ensuring clean and safe water for communities and ecosystems.

CONCLUSION:

Thus the predictive models were built and water quality analysis was performed and different visualizations were performed using IBM cognos.