**Intelligent Water System**

**Final Project**

**Introduction to Machine Learning**

December 16th, 2022

**Predict flow And Influent Wastewater Characteristics**

Team Members – Technocrats

**Arukala Himateja,** Graduate Student, Illinois State University, [harukal@ilstu.edu](mailto:harukal@ilstu.edu)

* Expertise: Requirement Analysis and Research, Problem Solving, debugging
* Roles and Responsibilities:Team Leader, Data Processing and Cleaning, Training and testing the model, and Data Visualization.

**Yashaswini Mekala,** Graduate Student, Illinois State University, [ymekala@ilstu.edu](mailto:ymekala@ilstu.edu)

* Expertise: Data Integration and Programming
* Roles and Responsibilities:Model planning and programming

**Shivani Busireddy,** Graduate Student, Illinois State University, [sbusire@ilstu.edu](mailto:sbusire@ilstu.edu)

* Expertise: Data Mining and programming.
* Roles and Responsibilities:Interpretation of Results, Technical Report.

**Mallu Praneel Reddy,** Graduate Student, Illinois State University, [pmallu@ilstu.edu](mailto:pmallu@ilstu.edu)

* Expertise: Database management System.
* Roles and Responsibilities:Develop the model, Technical Report.

**Kancharla Deepak Reddy,** Graduate Student, Illinois State University, [dkancha@ilstu.edu](mailto:dkancha@ilstu.edu)

* Expertise: Data Mining and programming.
* Roles and Responsibilities:Model development and Validation.

**Milkuri Akhila,** Graduate Student, Illinois State University, [amilkur@ilstu.edu](mailto:amilkur@ilstu.edu)

* Expertise: Data Mining and programming.
* Roles and Responsibilities:Model development and Validation.

Problem Background Objective

**Wastewater** contains various organic components with different physical and biochemical characteristics. Water data collected from the last 20 years is analyzed and used for training the model. This data has many missing values which need to be handled. The features of this data are Total Solids, BOD5, P-TOT, SS, Org-n, PRCP\_NOOA, NH3, SO4, and TKN. Based on other features that are currently available, we evaluate the data and provide the best recommendation for forecasting each water parameter in this. We pre-processed the water dataset to replace any null values with the feature's mean value before carrying out full data analysis to ascertain the relevance of relevant features. We have used 6 different models – Linear Regressor, Logistic Regressor, Lasso Regressor, Ridge Regressor, Random Forest Regressor, and Support Vector Machine (SVM). The accuracy of the models created is calculated using Mean Absolute Percentage Error. The result has been evaluated using R2 Square Error and Accuracy. This investigation helps in comparing all the models and choosing the best model.

**INTRODUCTION**

Wastewater is any water that has been affected by human use. Wastewater is "used water from any combination of domestic, industrial, commercial or agricultural activities, surface runoff or stormwater, and any sewer inflow or sewer infiltration". Therefore, wastewater is a byproduct of domestic, industrial, commercial, or agricultural activities. The characteristics of wastewater vary depending on the source. Types of wastewater include domestic wastewater from households, municipal wastewater from communities and industrial wastewater. Wastewater can contain physical, chemical and biological pollutants. (Wastewater, n.d.)

# Bibliography

*Wastewater*. (n.d.). Retrieved 12 16, 2022, from Wikipedia: The Free Encyclopedia: http://en.wikipedia.org/wiki/Wastewater