International Journal on Science and Technology (IJSAT)
E-ISSN: 2229-7677 ● Website: www.ijsat.org ● Email: editor@ijsat.org

Under Ground Water Level Predicator Using Machine Learning

A.Vara prasad[1], M. Sai sujith Reddy[2], Sangeetha.K[3], Vadde Balaji[4], Dr. Srinivasan T R[5]

[1,2,3,4,5]Computer Science and Engineering, Presidency University Bengaluru

Abstract

Sustainable groundwater management is critical for the long-term water security of a country, particularly in those areas that depend greatly on subsurface water for agriculture, household, and industry. This project will create a predictive model of underground water level estimation based on machine learning methodologies. Using historical information like rainfall, temperature, soil moisture, water table depth, and land use patterns, the system is able to predict future groundwater levels accurately. Regression algorithms like Linear Regression, Random Forest, and Support Vector Machines (SVM)

A buzzer can be used as an alerting device to notify users or authorities if trash has to be removed. By using shaft adapters, jumper wires, and bolts, there can be a strong, compact, and extensible design, applicable for urban and smart city use.

Key Words: Water Level Predicator – Machine Learning, data set Pandas Python

1. INTRODUCTION

A buzzer can be used as an alerting device to notify users or authorities if trash has to be removed. By using shaft adapters, jumper wires, and bolts, there can be a strong, compact, and extensible design, applicable for urban and smart city use. A buzzer can be used as an alerting device to notify users or authorities if trash has to be removed. By using shaft adapters, jumper wires, and bolts, there can be a strong, compact, and extensible design, applicable for urban and smart city use.

Water level forecasting is an essential component of hydrological prediction, especially for areas that are subject to drought, groundwater over-exploitation, or ineffective management of water resources. Conventional forecasting techniques usually depend on static models or past averages, which may not possess the flexibility and precision needed under evolving environmental situations. To overcome this drawback, this project presents an "Under-Level Water Predictor" based on Bayesian Machine Learning Models (BMML).

smallholder farmers in remote areas. Without timely and accurate interventions, localized infections can escalate into widespread outbreaks, jeopardizing food security and farmer incomes.

To Bayesian Machine Learning integrates prior knowledge and revises projections as additional information comes to light, and it is especially suited for ambiguous and dynamic conditions like water table monitoring. The system designed here is meant to predict when and where water levels are likely to drop below key thresholds so that agricultural intervention, urban planning, and catastrophe prevention can be effected in a timely manner.

By using real-time sensor information and past hydrological trends, the BMML-based forecaster learns and updates iteratively, providing a strong and probabilitically valid means of prediction. It not only increases accuracy in predictions but also furnishes confidence intervals that enable decision-makers to make a better-informed decision.

Bayesian Machine Learning integrates prior knowledge and revises projections as additional information comes tolight, and it is especially suited for ambiguous and dynamic conditions like water table monitoring. The system designed here is meant to predict when and where water levels are likely to drop below key thresholds so that agricultural intervention, urban planning, and catastrophe prevention can be effected in a timely manner.

symptoms, severity, and affected plant types, which they document manually on these forms.

However, manual data collection is often susceptible to hu man error, inconsistency, and subjective interpretation. These issues can result in inaccurate records that undermine effective disease management. Moreover, correcting such inaccuracies requires additional time and expertise, often involving mul tiple stakeholders, thereby increasing the overall burden on agricultural systems.

The data correction process itself is not only resource intensive but also introduces delays in decision making delays that can be critical when prompt action is required to prevent the spread Water. The dependency on human intervention, combined with logistical and financial constraints, further complicates timely and accurate disease tracking.

Additionally, many of the current guidelines for crop dis ease identification and management are drawn from manuals, academic research, and field trials. While these resources are valuable, their effectiveness depends heavily on how well and how quickly data is collected, interpreted, and applied in the field.

In summary,groundwater is one of the most valuable natural resources and one of the prime sources of fresh water for agricultural, industrial, and domestic needs globally. In the majority of the world, especially in semi-arid and arid countries, sub-surface water is the primary source of water for local water supply systems. In spite of that, the increased demand for water, coupled with over-pumping and poor recharge, has triggered a sudden decrease in groundwater levels. Accurate prediction of ground water levels is required to maintain effective water resource management, irrigation system scheduling, and prevent water emergencies..

2. OBJECTIVE

Water is a basic resource for life, agriculture, and industry. In much of the world, though, water scarcity and varying groundwater levels threaten sustainable development gravely. Predicting precisely when water levels drop below critical thresholds — so-called "under-level water conditions" — is a must for effective water resource management, early warning systems, and environmental sustainability.

- To incorporate predictive analytics by leveraging weather patterns, soil conditions, and humidity data to forecast poten tial disease outbreaks.
- To create a predictive model with Bayesian Machine Learning (BMML) for real-time forecasting of under-level water condition
- To gather and process hydrological data, such as groundwater levels, rainfall, temperature, and soil moisture, from sensors and historical sources pertinent to the area.
- To use Bayesian inference methods for handling uncertainty and real-time updating of predictions with the availability of new data.
- To establish decisive threshold levels of waterscarcity that may invoke alerts for agricultural, municipal, or industrial consumers.
- To create an intuitive interface (dashboard ormobileapp) to display predictions, warnings, and confidence intervals for the decision-makers.
- To confirm theaccuracyand dependability of the BMML model using performance metrics like RMSE, MAE, and probability calibration

3. METHODOLOGY

The procedure for creating the Under-Level Water Predictor based on Bayesian Machine Learning Models (BMML) is organized into some of the important stages.
Identify the crucial water level limits below which situations are regarded as "under-level Normalize and standardize the data.".

a. Problem Identification and System Requirements
Define the problem:

The objective is to forecast when groundwater or surface water levels are below a predetermined critical threshold (under-levelwater for their needs in their life upto condition).

multilingual interfaces, and function offline for rural deployment.

b. Data Collection:

Groundwater Levels: From borehole sensors or government hydrology departments.. High-quality images of both healthy and diseased leaves were sourced from field visits, agricultural institutions, and open-source databases like Plant Village, with expert an notation for labeling. Additionally, IoT sensors were deployed in crop fields to monitor temperature, humidity, soil moisture, and rainfall. These sensors transmitted real-time data using wireless technologies such as Zigbee and LoRa to a centralized cloud system.

multilingual interfaces, and function offline for rural deployment.

c. Data Collection:

Groundwater Levels: From borehole sensors or government hydrology departments.. High-quality images of both healthy and diseased leaves were sourced from field visits, agricultural institutions, and open-source databases like Plant Village, with expert an notation for labeling. Additionally, IoT sensors were deployed in crop fields to monitor temperature, humidity, soil moisture, and rainfall. These sensors transmitted real-time data using wireless technologies such as Zigbee and LoRa to a centralized cloud system.

d. Data Preprocessing:

Image data underwent preprocessing steps including resizing, normalization, and noise removal using Gaussian filters. Techniques such as edge detection and segmentation were applied to extract visual features like discoloration and shape deformation. Environmental data from sensors was cleaned to remove anomalies and handle missing values through interpolation or mean imputation.
e. Data Preprocessing and Feature Engineering
Data Cleaning:

A deterministic, rule-based classification system was designed using a database of disease symptoms and environ mental triggers. This system cross-references uploaded leaf images with known symptom patterns using logical inference to identify the most probable disease affecting the crop. e. Disease Severity Assessment: To support actionable treatment, disease severity was as sessed and categorized into three levels: mild (early-stage symptoms), moderate (growth-affecting symptoms), and severe (water-threatening damage). The classification was based on lesion count and size in images and the degree of deviation from optimal environmental conditions.

e. Model Development: Bayesian Machine Learning: Depending on the diagnosed disease and its severity, the system generates specific treatment recommendations. These include chemical interventions (fungicides, bactericides), or ganic treatments (e.g., neem oil, compost teas), and preventive practices (such as crop rotation and use of resistant crop varieties). All recommendations are based on verified agricultural guidelines and expert input.

f. Prediction and Uncertainty Quantification:

An NLP module enhances accessibility by supporting multilingual voice and text inputs. It processes user queries to extract key information such as symptoms, crop names, and environmental descriptors using Named Entity Recognition (NER). The system provides responses in local languages using text-to- speech and translation services, enabling effective interaction for diverse user groups.

g. Alerting and Visualization System:

The platform integrates all modules into a cohesive sys tem accessible via a mobile application and a responsive web interface. The backend uses RESTful APIs for real time synchronization of user data, diagnostic results, and environmental metrics. The platform includes security features such as encryption, authentication, and data access logging.

h. Model Evaluation and Validation

The system's effectiveness was validated through field test ing and collaboration with agricultural experts and farmers. Diagnostic accuracy was evaluated using confusion matrices, while usability was assessed through structured interviews and user surveys. Feedback from these evaluations was used to refine the system's rule base, improve NLP performance, and enhance user interface design.

4. LITERATURE SURVEY

The Reliable prediction of water levels, especially under-level or drought conditions, is an essential issue in environmental modeling, agriculture, and resource management. Many techniques have been tried by researchers over the years, from statistical models to cutting-edge machine learning techniques. Some of the important literature relevant to the development of a Bayesian Machine Learning-based Under-Level Water Predictor is presented below

In ARIMA models have long been applied in time-series water level prediction based on their interpretability and ease of use (Box & Jenkins, 1970). But ARIMA fails to capture non-linear trends and quantify uncertainty.

Multiple Linear Regression (MLR) models were used by Singh et al. (2005) for predicting groundwater level, but they are highly sensitive to linearity and stationarity assumptions of data, which are frequently violated in environmental data.

Support Vector Machines (SVM) and Random Forests are successful in non-linear groundwater level prediction (Daliakopoulos et al., 2005). The models work with intricate relationships but do not have built-in uncertainty estimation . Artificial Neural Networks (ANNs) with a focus on Long Short-Term Memory (LSTM) networks have proved successful in aquifer and river basin water level dynamics modeling (Chitsazan et al., 2015). Classical ANNs, however, yield deterministic results and demand much tuning and big datasets.

Real-time monitoring systems with the integration of machine learning have been promising in proactive water resource management. For example, Patil et al. (2020) created an IoT-based smart irrigation system based on simple rule-based logic along with cloud data.

Nonetheless, the majority of these systems are not equipped with probabilistic forecasting, which is critical to risk-informed decision-making in uncertainty — an aspect that BMML seeks to bridge.

Most current models rely on stationarity in hydrological and climate variables, which is less and less realistic under the threat of climate change.

There is sparse deployment of models that offer uncertainty bounds, resulting in overconfident or overly conservative choices.

The application of Bayesian models to water prediction remains somewhat untapped, particularly for under-level detection compared to the prediction of flooding.

ensuring timely data delivery, which is crucial for applications like smart agriculture that require prompt decision-making.

Shekhar et al. discussed the development of intelligent infrastructure for smart agriculture, emphasizing the integra tion of food, energy, and water systems. The study advocates for the adoption of advanced technologies and data-driven strategies to enhance resource efficiency, increase , and promote sustainable farming practices.

Introduced the concept of ”Eternal- Thing,” focusing on creating a secure, aging-aware solar-energy har vester for sustainable IoT applications. The system aims to extend the operational lifespan of IoT devices through solar energy harvesting and robust security measures, contributing to the sustainability of IoT infrastructure in agriculture.

Mohanty et al. provided an extensive overview of the role of IoT in the development of smart cities, highlighting its potential to enhance various urban systems, including agriculture. The study emphasizes the importance of IoT in improving efficiency, sustainability, and responsiveness in urban environments.

Lastly, Zhou et al. presented a method for detecting Cercospora leaf spot disease in sugar beet using robust tem plate matching techniques. The approach enables accurate and automated disease detection, reducing the need for manual inspection and facilitating timely interventions to protect crop health.

5. PROPOSED WORK

While significant work has been done in water level prediction using both statistical and machine learning models, **few models address the uncertainty inherent in environmental forecasting**. Bayesian Machine Learning offers a principled way to incorporate uncertainty, prior knowledge, and real-time adaptability — making it well-suited for predicting under-level water conditions. This project aims to fill that gap by developing a **BMML-based under-level water prediction system** that is robust, adaptable, and probabilistically reliable.

The Bayesian Machine Learning Models (BMML) based Under-Level Water Predictor is an intelligent, data-driven system that can predict groundwater levels and detect impending under-level situations prior to their occurrence. The system is modular in architecture to ensure scalability, flexibility, and real-time deployment. It combines five essential modules: data acquisition, preprocessing, modeling, prediction and alerting, and visualization..

The data acquisition module captures real-time and historical information from various sources. They comprise IoT-based sensors in wells measuring groundwater, public weather APIs furnishing rainfall, temperature, and humidity data, and satellite soil moisture data. All of these data are centrally stored automatically using a database like PostgreSQL or a time-series database like InfluxDB. This provides a smooth stream of tidy and organized data for model use.

During the data preprocessing stage, raw inputs are cleaned and transformed. This involves outlier removal, filling missing values with Bayesian imputation or statistical interpolation, and normalization of numerical values for uniformity. Moreover, informative features are engineered to enhance the accuracy of the model — e.g., accumulated rainfall for the past few days, day-of-year encoding to identify seasons, and historic water level trends. The data is further divided into test and training sets according to time windows to allow temporal consistency of model assessment.

AThe Bayesian Machine Learning model forms the core of the system. Depending upon the size and complexity of data, the system implements Bayesian Linear Regression, Gaussian Process Regression (GPR), or Bayesian Neural Networks (BNNs). These models have the ability to handle uncertainty since they model probability distributions rather than single-point predictions.

This study seeks to develop a farmer-friendly, web- based plant disease identification system that allows users to upload images of affected leaves via mobile devices or basic digital platforms. The goal is to provide real-time disease predictions and actionable feedback through an interface that is intuitive, multilingual, and accessible even in low-connectivity environments. This approach democratizes access to agricultural expertise and empowers farmers to take timely preventive or corrective action.

6. SYSTEM DESIGN AND IMPLEMENTATION

After prediction, the risk assessment and alerting module checks if the likelihood of an under-level condition is above a user-specified threshold (e.g., 75%). In that case, the system automatically sends an alert through email, SMS (through Twilio),

The last layer is the user interface, which is a responsive dashboard utilizing tools such as Plotly Dash or React.js. The dashboard shows real-time sensor readings, past trends, future predictions with confidence intervals, and alert history. For convenience in rural or field settings, a mobile interface can be created using Flutter or Android Studio..

The system is organized into the following architectural layers:

a. The entire system is hosted either on a cloud (e.g., AWS, Google Cloud, Microsoft Azure) for scalability or local edge devices (e.g., Raspberry Pi) for offline and remote access..

b. Data Preprocessing Layer: Additionally, the system incorporates logging, model retraining pipelines, and a feedback loop to update the model regularly based on new observations and prediction errors, ensuring continuous improvement.

c. this BMML-based water level prediction system is designed to be a robust and adaptive solution that provides actionable, uncertainty-aware forecasts of under-level water conditions. It combines cutting-edge machine learning with real-time data streams and user-centric design to support smarter water resource management.

d. User Interface Layer: Offers mobile and web platforms for uploading images, accessing disease reports, and receiving treatment advice. It also displays disease history, treatment effectiveness, and environmental data trends.

e. The Bayesian Machine Learning Models (BMML) based Under-Level Water Predictor is designed as an end-to-end smart system which facilitates monitoring and early prediction of groundwater deficiencies. Its design consists of real-time and historic inputs, sophisticated statistical learning methods, and a human-readable interface to facilitate prompt decision-making in water resources management..

f. Continuous The system starts with an infrastructure for data acquisition that guarantees smooth and uninterrupted collection of applicable hydrological and meteorological data. This entails groundwater level data captured by digital borewell sensors, which can utilize ultrasonic or pressure transducers to sense depth.

Implementation Steps

a. Data Collection and Acquisition Images are collected by farmers and field agents using mobile phones, drones, and IoT cameras. Environmental data is continuously gathered by sensors and transmitted to the cloud.

b. Data Preprocessing Image data is enhanced through sharpening, denoising, and contrast correction. Environmental data is converted to standard formats for further analysis.

c. Rule-Based Disease Identification The system maps visual and environmental symptoms to known disease profiles using expert-defined rules. Disease severity is estimated based on threshold parameters.

d. Disease Diagnosis and Prediction The system identifies whether a crop is healthy or diseased. If diseased, it classifies the type and severity of the disease. Historical symptom pat terns and environmental data are used for disease progression forecasting.

e. Treatment Recommendations Control measures are recommended based on disease type and severity, including: Chemical treatments (e.g., fungicides, pesticides) Biological control methods Organic solutions Preventive strategies such as crop rotation and resistant varieties are also suggested.

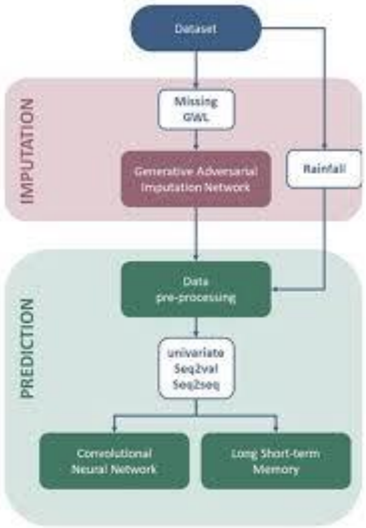
f. User Interface Development A dashboard allows farmers to upload images, receive diagnosis, and track disease and treatment history. Notifications are sent for potential disease outbreaks.

g. Cloud Storage and Deployment The system is hosted on cloud platforms to ensure secure, real-time data access. Historical records of diseases and sensor data are stored for future analysis.

h. Testing and Validation The system undergoes field testing on different crops and conditions. Accuracy is validated through comparison with agricultural expert assessments.

i. System Deployment The system is publicly deployed via web and mobile applications and integrated with local government and agricultural databases.

j. Maintenance and Upgrades Updates are made regularly based on user feedback and new research findings. Training and support resources are provided to users.



REFERENCES

1. AG. Gupta, H. P. Gupta, P. Kumari, R. Mishra, S. Saraswat, and T. Dutta, "A realtime precision agriculture monitoring system using mobile sink in wsns," in 2021 IEEE International Conference on Advanced Networks and Telecommunications Systems (ANTS). 2021, pp. 1–Eason, B. Noble, and I.
2. N. Sneddon, "On certain integrals of Lipschitz-Hankel type involving products of Bessel functions," Phil. Trans. Roy. Soc. London, vol. A247, pp. 529–551, April 1955.
3. S. P. Mohanty, U. Choppali, and E. Kougianos, "Everything you wanted to know about smart cities: The internet of things is the backbone," IEEE Consumer Electronics Magazine, Jul. 2019, vol. 5, no.
4. S. K. Ram, S. R. Sahoo, B. B. Das, K. K. Mahapatra, and S. P. Mohanty, " Eternal-Thing: A Secure Aging-Aware Solar-Energy Harvester Thing for Sustainable IoT, " IEEE Transactions on Sustainable Computing (TSUSC), vol. XX, no. YY.
5. S. Shekhar, J. Colletti, F. Munoz-Arriola, L. Ramaswamy, C. Krintz, ˜ L. R.Varshney, and D. Richardson, "Intelligent Infrastructure for Smart Agriculture:An Integrated Food, Energy and Water System," ,IEEE Access, 2020, arXiv CoRR, vol. abs/1705.01993.48
6. S. Saraswat, H. P. Gupta, T. Dutta, and S. K. Das, "Energy efficient data forwarding scheme in fog-based ubiquitous system with deadline constraints, " IEEE Transactions on Network and Service Management, 2020, vol. 17, no. 1, pp. 213–226.
7. R. Beulah and M. Punithavalli, "Prediction of sugarcane diseases using data mining techniques, " in Proc. IEEE International Conference on Advances in Computer Applications (ICACA), Oct. 2019, pp. 393– 396.
8. Peng Jiang , Yuehan Chen ,Bin Liu , Dongjian He, Chunquan Liang ,' Real-Time Detection of Apple

Leaf Diseases Using Deep Learning Approach Based on Improved Convolutional Neural Networks', IEEE Access, 06 May 2020.

9. M. Ayaz, M. Ammad-Uddin, Z. Sharif, A. Mansour, and E. M. Aggoune,"Internet-of-Things (IoT)-Based Smart Agriculture: Toward Making the Fields Talk," IEEE Access, vol. 7, pp. 129 551–129 583, 2022.
10. J. Lowenberg-DeBoer, The economics of precision agriculture, IEEE Access, Nov 2021, pp. 461–494.Bashish, D.A., Braik, M., Ahmad, S.B., 'A Framework for Detection and Classification of Plant Leaf and Stem Diseases', International Conference on Signal and Image Processing, IEEE Access, 2019, pp, 113-118