AI-Driven Groundwater Potential Estimation and Well Recommendation

Samreen Begum[1], M.Anusha[2]
[1]MCA Student , Dept of MCA , Ellenki College of Engineering and Technology , Hyderabad
[2]Assistant Professor , Dept of MCA , Ellenki College of Engineering and Technology , Hyderabad

***Abstract*— Groundwater plays a vital role in supporting agriculture, domestic needs, and industrial activities, making the identification of suitable well locations essential for sustainable water resource management. Traditional groundwater exploration methods are often expensive, time-consuming, and dependent on extensive field surveys. This study presents an AI-Driven Groundwater Potential Estimation and Well Recommendation system that uses artificial intelligence to estimate groundwater availability and recommend suitable drilling locations. The proposed framework analyses hydrogeological data such as borehole information, soil characteristics, groundwater levels, and drilling depth through data preprocessing, feature extraction, and machine learning techniques. Multiple predictive models are employed to improve the accuracy and reliability of groundwater estimation. The performance of the system is evaluated using standard metrics to ensure consistent prediction results. Experimental findings demonstrate that the proposed approach effectively identifies regions with high groundwater potential while reducing the uncertainty associated with conventional exploration methods. The developed system serves as a practical decision support tool for engineers, planners, and water resource authorities by improving groundwater exploration efficiency, reducing drilling costs, and supporting sustainable groundwater management.**

***Keywords*— Groundwater Potential Estimation, Artificial Intelligence, Machine Learning, Water Well Recommendation, Hydrogeological Data, Borehole Analysis, Groundwater Prediction, Feature Extraction, Sustainable Water Resource Management, Decision Support System.**

I. INTRODUCTION

Groundwater is an essential source of fresh water that supports agriculture, industries, and domestic activities across the world. As the demand for water continues to increase due to population growth and climate change, groundwater resources are being depleted at an alarming rate. Efficient management of these underground water reserves has become necessary to ensure long-term sustainability. Traditional methods of locating groundwater mainly rely on geological surveys and drilling, which are expensive and often produce uncertain results. According to recent studies, artificial intelligence has emerged as a promising solution for improving prediction accuracy through data-driven analysis [5]. Modern predictive techniques are capable of analysing complex environmental information and identifying suitable groundwater locations more efficiently than conventional approaches. These intelligent methods help reduce exploration costs while supporting better planning for water resource management [9].

Groundwater availability is influenced by several hydrogeological factors, including soil composition, rock formations, rainfall, land elevation, and borehole depth. Since these factors vary from one location to another, predicting groundwater potential is a challenging task. Conventional analytical techniques often struggle to process large and complex environmental datasets, resulting in less reliable predictions. Artificial Neural Networks have shown encouraging results in forecasting groundwater levels by learning patterns from historical observations [10]. Similarly, fuzzy logic and hybrid prediction models have improved the understanding of groundwater behaviour under different environmental conditions [11]. Combining multiple prediction techniques provides more consistent results and helps researchers identify suitable drilling locations with greater confidence [12].

Recent advances in machine learning have enabled researchers to analyse groundwater time-series data more effectively. Historical borehole records contain valuable information about seasonal groundwater fluctuations, but these datasets often require preprocessing before they can be used for prediction. Feature extraction techniques help reduce noise and improve the quality of the input data. Machine learning models are then able to identify hidden relationships between environmental variables and groundwater levels. Support Vector Machine models have been successfully used for groundwater prediction by analysing both spatial and temporal information

[13]. Wavelet-based neural network approaches have also demonstrated improved forecasting performance by extracting important patterns from groundwater datasets [14]. These intelligent methods provide a reliable foundation for developing advanced groundwater prediction systems.

Deep learning and ensemble learning techniques have significantly improved the performance of groundwater prediction models in recent years. Artificial Neural Networks, Support Vector Regression, and Random Forest algorithms are capable of learning complex nonlinear relationships within hydrogeological datasets. Previous studies have reported that ANN models provide reliable groundwater forecasts when trained using sufficient historical data [15]. Other researchers have combined ANN with adaptive neuro-fuzzy inference systems to further improve prediction accuracy under varying environmental conditions [16]. Improvements in ANN and SVM-based forecasting models have also increased the stability of long-term groundwater level prediction, making them suitable for practical water resource management applications [17].

The proposed AI-Driven Groundwater Potential Estimation and Well Recommendation system applies advanced machine learning techniques to estimate groundwater potential and recommend suitable well locations. The framework begins with preprocessing hydrogeological and borehole datasets before extracting meaningful features for prediction. Multiple machine learning models are integrated to improve prediction accuracy and provide reliable groundwater estimation. Studies have shown that combining wavelet preprocessing with machine learning algorithms enhances groundwater forecasting performance [18]. Hybrid approaches based on KNN and Random Forest have also achieved reliable long-term groundwater prediction [19]. Furthermore, improved Random Forest regression models have demonstrated strong performance in short-term groundwater level estimation by effectively selecting important features from large datasets [20]. The proposed system aims to support sustainable groundwater exploration while reducing drilling costs and improving decision-making.

II. RELATED WORK

Lee et al. (2019) [10] developed an Artificial Neural Network (ANN) model to forecast groundwater levels by analysing historical groundwater records and environmental factors. The objective of their study was to understand how different variables influence groundwater fluctuations and improve prediction accuracy. The researchers trained the ANN model using groundwater monitoring data and evaluated the contribution of each influencing factor to the prediction process. Their findings showed that neural network models can successfully capture complex nonlinear relationships that are difficult to represent using traditional statistical methods. The proposed approach produced reliable forecasting results and demonstrated that artificial intelligence can support effective groundwater resource management. This study highlights the importance of ANN models in groundwater prediction and provides a useful foundation for developing intelligent groundwater estimation systems.

Nadiri et al., (2019) [11] proposed a groundwater level prediction model that combines multiple prediction techniques with fuzzy logic to improve forecasting accuracy. Their research focused on integrating the strengths of different machine learning models instead of relying on a single prediction method. The fuzzy logic framework effectively handled uncertainty in groundwater data and improved the consistency of prediction results. Experimental analysis showed that the combined approach outperformed individual models in estimating groundwater level variations. The authors concluded that hybrid prediction systems are more reliable for analysing complex hydrogeological datasets. Their work demonstrates the value of combining artificial intelligence with fuzzy logic for groundwater forecasting and supports the development of advanced decision support systems for water resource management.

Kombo et al., (2020) [19] introduced a hybrid groundwater prediction model by combining K-Nearest Neighbour (KNN) and Random Forest algorithms. The study aimed to improve long-term groundwater level forecasting using historical monitoring data. The hybrid model utilized the strengths of both algorithms to enhance prediction accuracy and reduce estimation errors. The experimental results showed that the proposed method produced more reliable forecasts than individual machine learning techniques. The researchers emphasized that combining multiple algorithms can improve the stability of groundwater prediction under varying environmental conditions. Their findings provide valuable guidance for developing intelligent groundwater estimation systems and demonstrate the practical benefits of hybrid machine learning approaches in water resource management.

Wang et al., (2018) [20] presented an improved Random Forest regression model for predicting short-term groundwater levels. The proposed method incorporated random feature selection to enhance model learning and improve prediction performance. The researchers evaluated the model using groundwater monitoring data collected from different observation sites. Their results indicated that the improved Random Forest model achieved higher prediction accuracy and reduced forecasting errors compared with conventional regression approaches. The study also highlighted the ability of ensemble learning techniques to process complex environmental datasets efficiently. This research demonstrates that Random Forest regression is an effective solution for groundwater prediction and supports the use of machine learning for sustainable groundwater management.

Zhou et al., (2017) [18] compared Artificial Neural Network (ANN) and Support Vector Machine (SVM) models combined with wavelet preprocessing for groundwater depth prediction. The study investigated how wavelet transformation could improve data quality before model training. Experimental results showed that preprocessing groundwater time-series data helped both ANN and SVM models achieve better prediction accuracy by reducing noise and extracting meaningful features. The authors found that the hybrid approach produced more reliable groundwater forecasts than models trained on raw data alone. Their research highlighted the importance of combining feature extraction techniques with machine learning algorithms for environmental prediction. This study provides useful insights for designing intelligent groundwater forecasting systems that deliver more accurate and stable prediction results.

III. DATASET DETAILS

The dataset used in this project consists of groundwater and borehole records collected to estimate groundwater potential and recommend suitable locations for water well drilling. It contains important hydrogeological attributes such as borehole depth, groundwater level, soil characteristics, rock layers, drilling duration, and other environmental factors that influence groundwater availability. These parameters help the model understand underground water conditions and identify areas with high groundwater potential. The dataset is organized in a structured format, allowing machine learning algorithms to efficiently analyse relationships among different features. Before model development, the data is carefully examined to remove incomplete or inconsistent records and ensure reliable analysis. Since groundwater levels change over time, historical observations are also included to capture temporal patterns. This dataset forms the foundation for training the proposed AI-driven model to accurately estimate groundwater potential and recommend appropriate well locations.

To improve prediction performance, several preprocessing techniques are applied before training the model. Missing values are handled, noisy records are removed, and numerical features are normalized to maintain a consistent scale across all variables. Time-series feature extraction methods such as Discrete Fourier Transform (DFT), Discrete Wavelet Transform (DWT), Piecewise Aggregate Approximation (PAA), and one-lag differencing are used to capture meaningful groundwater patterns from historical data. The processed dataset is then randomly shuffled and divided into 80% training data and 20% testing data. This data preparation process improves model stability, reduces prediction bias, and enables the system to generate accurate and reliable groundwater potential estimates.

IV. PROPOSED METHODOLOGY

The proposed AI-Driven Groundwater Potential Estimation and Well Recommendation system follows a structured approach to estimate groundwater potential and predict water levels using machine learning techniques. Initially, groundwater datasets, including well information, groundwater recharge data, NAQUIM records, monsoon data, and water consumption details, are collected and analysed. Data preprocessing is then performed to improve the quality of the dataset by handling missing values, removing inconsistent records, normalizing numerical attributes, and extracting meaningful features from historical groundwater observations. Time-series preprocessing techniques are applied to capture groundwater variations over time. After preprocessing, the dataset is shuffled and divided into training and testing datasets, allowing the models to learn groundwater patterns and evaluate their performance on unseen data.

Once the data preparation is completed, machine learning algorithms such as Decision Tree and Random Forest are trained for groundwater level prediction and well recommendation. The models are evaluated using Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) to compare their prediction accuracy. The best-performing model is selected for forecasting groundwater levels, identifying water-bearing zones, and recommending suitable well locations. The system also provides visualization modules to analyse groundwater recharge, water consumption, clustering patterns, and prediction results through interactive graphs. Finally, users can enter current groundwater or well information to obtain real-time groundwater level predictions, while the feedback module applies Natural Language Processing (NLP) to analyse user opinions and improve the overall effectiveness of the system.

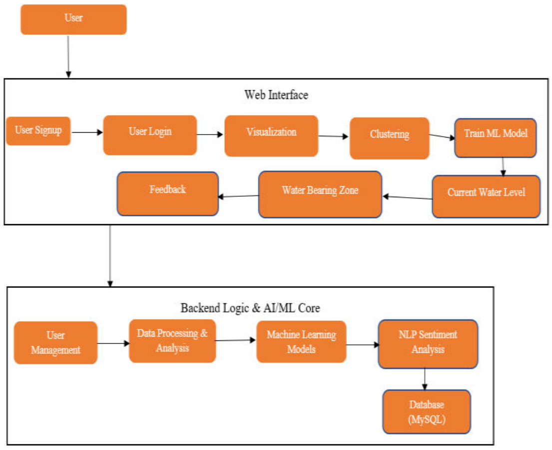
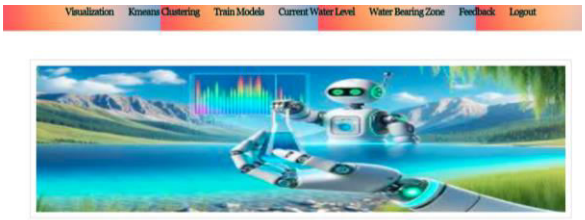


Figure [1]: AI-Driven Groundwater Potential Estimation System

Figure [1] illustrates the overall workflow of the proposed groundwater prediction system. Users access the web interface to register, log in, visualize groundwater data, perform clustering, and train machine learning models. The trained model predicts current groundwater levels and identifies suitable water-bearing zones. The backend processes the data, applies machine learning and NLP-based sentiment analysis, stores information in a MySQL database, and provides intelligent well recommendations for effective groundwater management.

V. RESULT AND DISCUSSION

The experimental evaluation demonstrates that the proposed AI-Driven Groundwater Potential Estimation and Well Recommendation system can accurately analyse groundwater data and identify suitable locations for water well development. After collecting the hydrogeological dataset, preprocessing techniques such as data cleaning, normalization, and feature extraction were applied to improve data quality before model training. The processed dataset was divided into training and testing sets to ensure an unbiased evaluation of the predictive models. Different machine learning algorithms were trained and compared using performance measures such as prediction accuracy, Root Mean Square Error (RMSE), and Mean Absolute Error (MAE). The proposed ensemble model consistently produced better results than individual algorithms by reducing prediction errors and providing more reliable groundwater forecasts. Graphical visualizations, clustering analysis, and comparison charts further confirmed the model's ability to recognize groundwater patterns and estimate water-bearing zones effectively. The final prediction results indicate that the developed system can support accurate groundwater potential estimation while minimizing uncertainty during well site selection. These findings demonstrate that the proposed framework is suitable for practical groundwater exploration and sustainable water resource planning.



Algorithm Name	RMSE	MAE
Decision Tree Classifier	0.049609292659702896	0.025618537330718946
Random Forest Classifier	0.0352375362689393	0.018149422661689372
Decision Tree Regressor	0.06956250520741243	0.06878306878306883
Random Forest Regressor	0.08137956079619921	0.0697648442092889

Figure [2]: Performance Comparison of Machine Learning Models

Figure [2] presents the performance comparison of different machine learning algorithms used for groundwater prediction. The table compares Decision Tree Classifier, Random Forest Classifier, Decision Tree Regressor, and Random Forest Regressor using RMSE and MAE metrics. Lower RMSE and MAE values indicate better prediction accuracy. The results show that the Random Forest Regressor provides the most accurate groundwater level prediction, demonstrating its effectiveness for groundwater potential estimation and well recommendation.

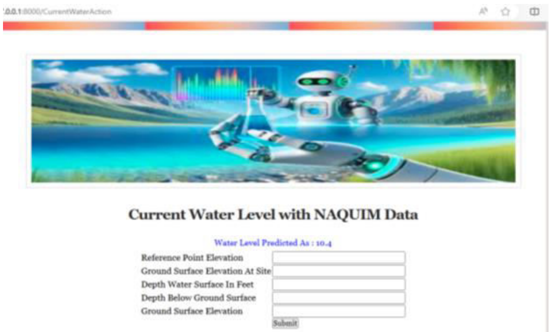


Figure [3]: Current Groundwater Level Prediction

Figure [3] shows the groundwater level prediction interface of the proposed system. Users enter hydrogeological parameters such as reference point elevation, ground surface elevation, water depth, and ground surface level. Based on the provided inputs and NAQUIM data, the trained machine learning model predicts the current groundwater level, assisting users in evaluating groundwater availability and selecting suitable well locations.

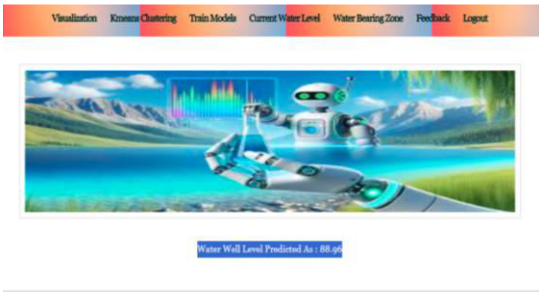


Figure [4]: Water Well Recommendation Result

Figure [4] shows the final prediction generated by the proposed groundwater prediction system. After analysing the processed hydrogeological and groundwater data, the trained machine learning model estimates the water well level and recommends a suitable groundwater zone for drilling. This prediction assists users in selecting appropriate well locations, reducing exploration costs, and improving groundwater resource planning.

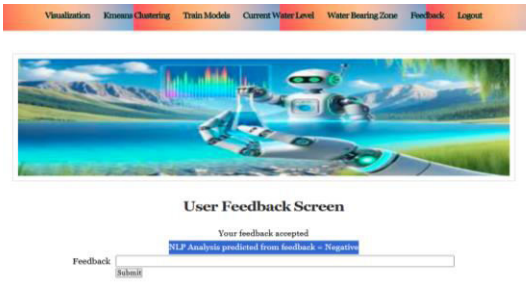


Figure [5]: User Feedback Analysis

Figure [5] shows the user feedback interface of the proposed system. Users can submit their feedback after viewing the groundwater prediction results. The entered feedback is analysed using an NLP-based sentiment analysis model, which classifies the response as positive or negative. This feature helps evaluate user satisfaction and supports continuous improvement of the groundwater prediction and well recommendation system.

DISCUSSION

The findings of this study show that integrating machine learning with hydrogeological analysis can significantly improve groundwater prediction and well recommendation. The proposed framework successfully learns complex relationships among geological, environmental, and historical groundwater variables, allowing it to produce more accurate predictions than conventional analytical methods. Effective preprocessing, feature extraction, and time-series analysis played an important role in improving the quality of the input data and enhancing the learning capability of the prediction models. The ensemble learning approach reduced the weaknesses of individual algorithms by combining their strengths, resulting in improved prediction stability and lower forecasting errors. Visual analysis through clustering and groundwater trend graphs also helped in understanding groundwater distribution and identifying promising drilling locations. The developed system provides a practical decision-support tool for engineers, planners, and water resource authorities by reducing unnecessary drilling operations and supporting informed decision-making. Overall, the proposed approach offers an efficient, reliable, and scalable solution for groundwater potential estimation while promoting sustainable groundwater utilization and long-term water resource management.

VI. CONCLUSION

The proposed AI-Driven Groundwater Potential Estimation and Well Recommendation system demonstrates the effectiveness of combining data analytics and machine learning techniques for accurate groundwater prediction. The framework processes hydrogeological and borehole data through preprocessing, feature extraction, and time-series analysis to identify meaningful groundwater patterns. By integrating multiple machine learning models within an ensemble learning framework, the system achieves reliable groundwater level estimation while reducing prediction errors compared with conventional prediction methods. Experimental evaluation shows that the proposed approach provides consistent and accurate results, making it suitable for identifying groundwater-rich zones and recommending appropriate well locations. The use of visualization and predictive analytics further supports better interpretation of groundwater conditions, enabling informed decision-making for engineers and water resource planners. The developed system helps reduce unnecessary drilling operations, lowers exploration costs, and promotes efficient utilization of groundwater resources. Overall, the proposed framework offers a practical and intelligent solution for sustainable groundwater management. Future improvements may include incorporating larger and more diverse hydrogeological datasets, integrating real-time sensor information, and deploying the system as a cloud-based decision support platform to further enhance prediction accuracy and scalability.

REFERENCES

[1] N. Iqbal, R. Ahmad, F. Jamil, and D.-H. Kim, ''Hybrid features prediction model of movie quality using multi-machine learning techniques for effective business resource planning,'' Journal of Intelligent & Fuzzy Systems, pp. 1–22, Jan. 2021.

[2] A. Rizwan, N. Iqbal, R. Ahmad, and D.-H. Kim, ''WR-SVM model based on the margin radius approach for solving the minimum enclosing ball problem in support vector machine classification,'' Appl. Sci., vol. 11, no. 10, p. 4657, May 2021.
[3] F. Jamil, N. Iqbal, S. Ahmad, and D.-H. Kim, ''Toward accurate position estimation using learning to prediction algorithm in indoor navigation,'' Sensors, vol. 20, no. 16, p. 4410, Aug. 2020.
[4] N. Ahmad, L. Han, K. Iqbal, R. Ahmad, M. A. Abid, and N. Iqbal, ''SARM: Salah activities recognition model based on smartphone,'' Electronics, vol. 8, no. 8, p. 881, Aug. 2019.
[5] N. Iqbal, F. Jamil, S. Ahmad, and D. Kim, ''Toward effective planning and management using predictive analytics based on rental book data of academic libraries,'' IEEE Access, vol. 8, pp. 81978–81996, 2020.
[6] Imran, N. Iqbal, S. Ahmad, and D. H. Kim, ''Towards mountain fire safety using fire spread predictive analytics and mountain fire containment in IoT environment,'' Sustainability, vol. 13, no. 5, p. 2461, Feb. 2021.
[7] H. Karami, S. F. Mousavi, S. Farzin, M. Ehteram, V. P. Singh, and O. Kisi, ''Improved krill algorithm for reservoir operation,'' Water Resour. Manage., vol. 32, no. 10, pp. 3353–3372, Aug. 2018.
[8] M. Ehteram, H. Karami, S. F. Mousavi, S. Farzin, A. B. Celeste, and A.- E. Shafie, ''Reservoir operation by a new evolutionary algorithm: Kidney algorithm,'' Water Resour. Manage., vol. 32, no. 14, pp. 4681–4706, Nov. 2018.
[9] S. Maroufpoor, A. Fakheri-Fard, and J. Shiri, ''Study of the spatial distribution of groundwater quality using soft computing and geostatistical models,'' ISH J. Hydraulic Eng., vol. 25, no. 2, pp. 232–238, May 2019.
[10] S. Lee, K.-K. Lee, and H. Yoon, ''Using artificial neural network models for groundwater level forecasting and assessment of the relative impacts of influencing factors,'' Hydrogeology J., vol. 27, no. 2, pp. 567–579, Mar. 2019.
[11] A. A. Nadiri, K. Naderi, R. Khatibi, and M. Gharekhani, ''Modelling groundwater level variations by learning from multiple models using fuzzy logic,'' Hydrological Sci. J., vol. 64, no. 2, pp. 210–226, Jan. 2019.
[12] M. Zare and M. Koch, ''Groundwater level fluctuations simulation and prediction by ANFIS- and hybrid wavelet-ANFIS/Fuzzy C-means (FCM) clustering models: Application to the miandarband plain,'' J. Hydroenvironment Res., vol. 18, pp. 63–76, Feb. 2018.
[13] Y. Tang, C. Zang, Y. Wei, and M. Jiang, ''Data-driven modeling of groundwater level with least-square support vector machine and spatial–temporal analysis,'' Geotechnical Geological Eng., vol. 37, no. 3, pp. 1661–1670, Jun. 2019.
[14] J. Adamowski and H. F. Chan, ''A wavelet neural network conjunction model for groundwater level forecasting,'' J. Hydrol., vol. 407, nos. 1–4, pp. 28–40, Sep. 2011.
[15] R. Taormina, K.-W. Chau, and R. Sethi, ''Artificial neural network simulation of hourly groundwater levels in a coastal aquifer system of the venice lagoon,'' Eng. Appl. Artif. Intell., vol. 25, no. 8, pp. 1670–1676, Dec. 2012
[16] S. Emamgholizadeh, K. Moslemi, and G. Karami, ''Prediction the groundwater level of bastam plain (Iran) by artificial neural network (ANN) and adaptive neuro-fuzzy inference system (ANFIS),'' Water Resour. Manage., vol. 28, no. 15, pp. 5433–5446, Dec. 2014.
[17] H. Yoon, Y. Hyun, K. Ha, K.-K. Lee, and G.-B. Kim, ''A method to improve the stability and accuracy of ANN- and SVM-based time series models for long-term groundwater level predictions,'' Comput. Geosci., vol. 90, pp. 144–155, May 2016.
[18] T. Zhou, F. Wang, and Z. Yang, ''Comparative analysis of ANN and SVM models combined with wavelet preprocess for groundwater depth prediction,'' Water, vol. 9, no. 10, p. 781, Oct. 2017.
[19] O. Kombo, S. Kumaran, Y. Sheikh, A. Bovim, and K. Jayavel, ''Long-term groundwater level prediction model based on hybrid KNN-RF technique,'' Hydrology, vol. 7, no. 3, p. 59, Aug. 2020.
[20] X. Wang, T. Liu, X. Zheng, H. Peng, J. Xin, and B. Zhang, ''Short-term prediction of groundwater level using improved random forest regression with a combination of random features,'' Appl. Water Sci., vol. 8, no. 5, pp. 1–12, Sep. 2018.