Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN：1671-5497
E-Publication：Online Open Access
Vol: 42 Issue: 02-2023
DOI 10.17605/OSF.IO/QBNEU

FRAMEWORK MODELLING FOR EFFECTIVE PRECISION AGRICULTURE USING PREDICTIVE ANALYTICS

VANISHREE K

Assistant Professor, Department of ISE, R V College of Engineering, Bengaluru, Karnataka, India.
Email; vanishreek@rvce.edu.in,

Dr. NAGARAJA G S

Professor & Associate Dean, R V College of Engineering, Bengaluru, Karnataka, India.
Email: nagarajags@rvce.edu.in

ABSTRACT

The study in this research introduces a unified framework which could assist in precision agriculture through effective predictive analytics of water quality monitoring data and also the localization of water quality monitoring sensors. It adopts analytical baseline to realize the proposed concept in the domain of wireless sensor network. The architectural design of the proposed system initially handles the redundancy of the data and performs data visualization followed by a preliminary analysis. The unified framework evaluates both classification and regression modeling to perform predictive analytics considering i) Logistic Regression Model , ii) KNN , iii) Random Forest and iv) Gaussian NB Classifier. The prime objective of this ML based solution is to accurately evaluate the reliability of the sensor data towards monitoring and assessing the water quality. This also indicates whether the sensor nodes are localized correctly or not. For this reason a set of performance metrics are considered for validation which are accuracy, precision, recall rate and F1_Score. The experimental outcome obtained from the simulation shows that the framework provides significant insight about the suitability of the reliable learning model among the four ML based approaches for predictive analytics on sensor node placement strategy. It also verifies the reliability of the water quality measurement sensor data with assessment of accuracy.

Keywords: Precision Agriculture, Wireless Sensor Networks, Artificial Intelligence, Machine Learning, Predictive Analytics, Water Quality Monitoring

1. INTRODUCTION

The idea of precision agriculture (**PA**) has evolved to bring new shifts to conventional agriculture through the advancement of various modern technologies such as the Internet of Things (IoT). However, the term PA is also often referred to as digital farming or intelligent agriculture, which uses diverse technologies in conventional farming practices to improve productivity and sustainability with lower costs and minimal human efforts [1]-[3]. Sensor nodes from wireless sensor networks (WSN) also form a basis for IoT, which, combined with the aid of artificial intelligence and diverse communication policies, build an eco-system for efficient, robust, and flexible architecture that can serve the purpose of agricultural crop monitoring and disease identification through PA [4]. WSN of on-field sensor nodes generates a large volume of data that can be used for several practices in PA. The proper understanding and interpretation of such data and the historical data corresponding to various agricultural factors leads to timely and informed decision-making and planning in PA [5]. The traditional data delivery model imposed by WSN in the physical layer get affected by various limitations of real-time monitoring. This happens

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN：1671-5497
E-Publication：Online Open Access
Vol: 42 Issue: 02-2023
DOI 10.17605/OSF.IO/QBNEU

due to the exacting and obstinate constraints of communication channels in the rural environment. There are other reasons as well, which include the improper placement of the sensor network, the election of communication protocols, channel conditions etc. which restricts the system from meeting the specific goals of monitoring cost-effectively and accurately. The following Fig. 1 shows three different views associated with the evolution of WSN deployment, where the latest deployment mode provides tremendous facility to deal with the real-time data and facilitates in PA.

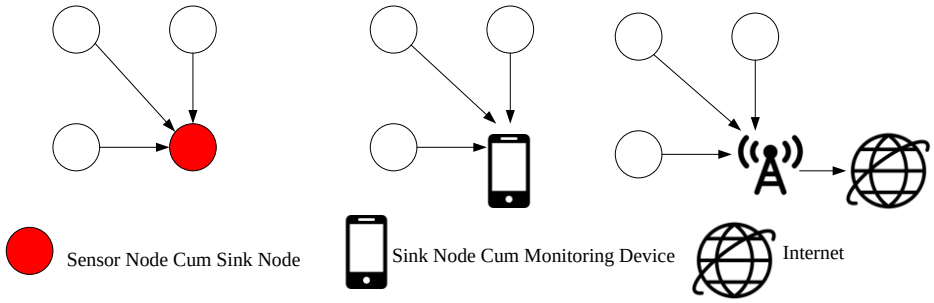


Fig: 1 Deployment Evolution of WSN in Three Different Views

The architecture deployment of WSN in the context of PA considers different scenarios as per the need of application and deployment conditions. It is observed that synchronizing the WSN with Internet communications yields better possibilities for data-driven decision-making in analytics. The prime reason is that by integrating the cloud, the predictive model-based decision support can be executed, and the outcome can be published on the user devices and distributed monitoring systems in the Edge layer of the PA eco-system [5-8]. For precision agriculture, sensor network architecture and data accuracy play an important role. A model has been designed that securely aggregates the data for transmission from sensor nodes to the base station [1]. However, it is quite essential to have an accurate measure of the data quality. The study also considers the role of water quality measurement sensors (WQMS) placement in capturing the vital environmental event data corresponding to water quality index (WQI) factors. This assists in planning and informed decision making for sustainable and productive PA. This paper uses suitable machine learning models to predict the placement of water quality measurement sensors (WQMS) based on artificial intelligence. The accurate placement of the WQMS ensures better quality for the PA.

2. Literature Review

The study in this phase of the research explores the trend of existing studies to draw the conclusion from it about the strength and limitations associated with the techniques in PA. There are various research based solution which have emphasized on assessing environment factor changes and the monitoring of various aspects such as air quality (AQ), noise pollution, emitting gas monitoring etc. have implications on the human health. Various related works also have studied how the pattern in the fluctuations of air quality index (AQI) varies with respect to different urban development conditions which have

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN：1671-5497
E-Publication：Online Open Access
Vol: 42 Issue: 02-2023
DOI 10.17605/OSF.IO/QBNEU

resulted in air pollution, water pollution due to extensive growth of industrialization and increasing vehicular traffic density among the cities. The monitoring of pollutants such as SO_2, CO_2, O_3, and NO_2 is needed to be effectively implemented. It has been observed that various machine learning (ML) based solutions have implications on correcting the time-series data which could effectively enhance the performance of the learning models. The study of (Amuthadevi et al, 2021) also introduces an LSTM based learning solution to deal with the redundant information of the time-series data which are gathered on the basis of real-time event [9]. It can be also seen that ML models are also deployed for WQ monitoring which has influenced the positive outcome in drinking water treatment which is identified in the recent study of (Li L et al, 2021) [10]. The survey work in this phase of the research basically shows how various AI based methods could assist in categorizing the contaminant and other factors of water and derive informed decision making for the drinking water plant operations. The study of (Heo S et al, 2021) [11], (Sundui et al, 2021) [12] also emphasized on the AI approaches for waste water treatment purposes from the WQ point of view. The research in the similar direction also can be observed in the studies of [13-15]. There are related studies of (Zhao et al, 2021) [17], (Chandra et al, 2021) [18] and (Van et al, 2021) [19] which have worked on the data related to noise pollution. The studies have reviewed the aspects of the problem background associated with the noise pollution and it assisted in realizing how the sound originated from daily traffic or bursting fire crackers on occasions in urban and semi-urban areas can affect both the physical and mental health of the people. The study also evaluated the propagation of the noise considering statistical correlation factor which could predict and identify the affected population and accordingly remedial steps can be taken. However, the gap in proper selection of suitable ML based approaches to correctly apply the predictive learning model restricts their implications on these data analytics which poses limitations in identifying the accurate and to be impacted zones. The prime reason is the variability in the categorization of the noises over the dataset. This accuracy is further improvised in the work of (Alvareset al, 2021) [16] considering lower-level dataset of sensor network and further effective selection of predictive learning model. However, very lesser emphasize has been found in the domain of PA where the scope of application of ML based approaches are higher. It has been also observed that despite of various research-based solutions the proper WQ monitoring is yet to achieve full-fledged solution for PA.

3. Research Problem

The role of water quality monitoring in PA is of significant concern in agriculture, and in the last few years, significant research effort has been led to develop water quality monitoring systems (WQMS) to elevate decision support systems (DSS) for PA. The idea of WQMS refers to the development of a computation model which captures vital information related to different water quality factors such as temperature, the value of dissolved oxygen (DO), pH value, conductivity, ammonia concentration, and many more. The proper statistical analysis of these factors yields proper and effective data-driven informed decision-making about the critical situations, which in return assists in enhancing the quality of PA. The analysis of existing research-based models in PA

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN：1671-5497
E-Publication：Online Open Access
Vol: 42 Issue: 02-2023
DOI 10.17605/OSF.IO/QBNEU

provides the idea that the available literature is rich and contains a variety of schemes and solutions for environmental monitoring. However, not a single research study focuses on the aspect of sensor placement concerning capturing reliable data in the case of water monitoring which could make the PA more effective. Thereby, it is quite obvious that if the WQMS node placement is not proper and not precise then it affects the process of data capturing. This scenario misleads the environmental planning and critical decision-making processes [20]. For example, it can be said that if WQMS nodes are incorrectly placed it can capture redundant and false data and the flawed data eventually affects the analysis process and the conclusion drawn from the data do not contribute to the PA. Therefore, this situation can be avoided by precise WQMS node placement strategy which not only avoids capturing of redundant data but also ensures data reliability. It has to be also noted that no traditional research-based solutions have suggested a unified framework that could have supportability of processing massive amount varied dataset corresponding to WQ as produced by the sensor nodes. The implementation of the existing systems are mostly carried out considering complex mode of execution which involves recursive operations resulting computational burden to the system. This restricts the scope of those models in mapping with the real-time use-cases of PA. It is also observed that literatures and recent publications do not provide the idea in which basis they have developed their predictive models. The literature review also outlines the implications of machine learning (ML) and deep learning (DL) models in designing predictive modeling for PA. However, it is also observed that several ML/DL based approaches have their own short of advantages and limitations. Selection of a suitable ML/DL based approach in designing the predictive model is not at all easier for different parameter settings of WQI. Thereby the problem statement of the proposed research study could be: "*It is quite challenging to design reliable and efficient ML or DL based predictive model considering empirical and evidential analysis which could assist in enhancing the decision support system for WQMS placement*".

4. SYSTEM MODEL

The system model of the proposed framework considers an effective strategy where in the initial phase of computation the system model considers water quality (WQ) data collected by the sensors during an environmental event. The prime aim of the proposed study is to design a unified framework of adaptive and reliable WQ monitoring systems through predictive analytics which enhance the quality of PA and also verifies the accurate placement of WQMS. It also influences the livable, healthy and sustainable surroundings suitable for the agricultural growth and productivity. The collected huge dataset related to WQ undergoes through data visualization and exploration phase which makes the suitable for the predictive analytics and influences insightful outcome that could assist in enhanced PA operations. The overview of the system model and its essential components are presented with the following schematic diagram which also represents the flow of execution.

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN：1671-5497
E-Publication：Online Open Access
Vol: 42 Issue: 02-2023
DOI 10.17605/OSF.IO/QBNEU

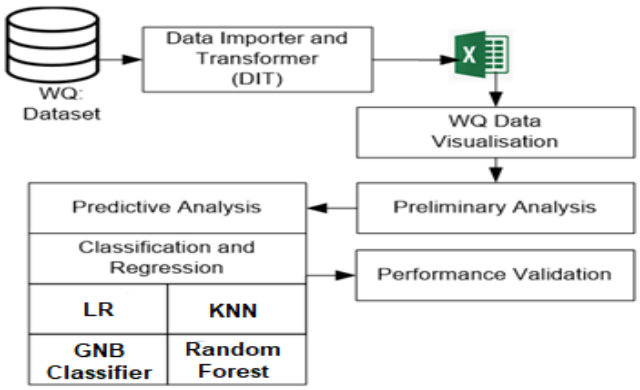


Fig: 2 Schematic Architecture of the Proposed Framework for WQ data reliability validation

4.1 Dataset Description

The King County open dataset provides open-source data for the environment related aspects that is taken for the design of the evaluation model used for the PA wherever the WQMS placement predictions are predominate. The dataset file size is 228 MB on disk as CSV file format. The samples in King county open dataset have been collected from three distinct sources such as i) streams, ii) lakes, and iii) Puget Sound. The dataset datapoint are very large to upload in the tradition data viewer. Therefore, on the load it loses many datapoints, thus it cannot be analyzed using the conventional analytics methods. The framework design in this phase considers data visualization and also performs exploratory analysis for the purpose of understanding the WQ dataset [46].

4.2 Dataset Visualization and Exploration

The model explicitly usage a data import to load the dataset from the file format to the computationally transformed structure as in the figure-1.

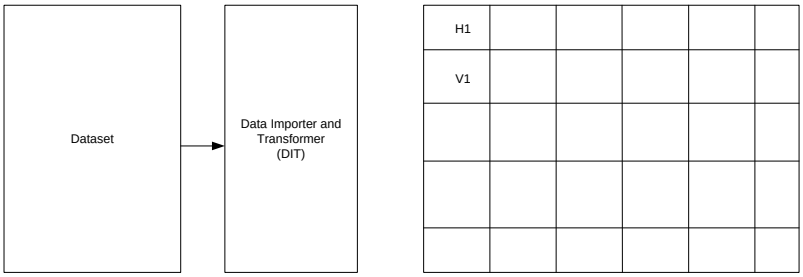


Figure-1 Dataset Transformation to Header and Value Pair

The Fig. 1 clearly shows how the dataset (d_{WQ}) is loaded into the system considering the data importer and Transformer (DIT) object module which basically computes the header (H) and value (V) pair for the respective attributes of data identifier$\{\mathrm{H},\mathrm{V}\} \in \mathrm{d}_{\mathrm{WQ}}$. It also considers the datatype (Dt) for key-value attributes. The following Table-1 exhibits the detail of the data identifier along with Dt in tabular form.

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN：1671-5497
E-Publication：Online Open Access
Vol: 42 Issue: 02-2023
DOI 10.17605/OSF.IO/QBNEU

Table-1 Data Visualization of d_{WQ} w.r.t H, V and Data-type

Sl. No	Data identifier {H}	Datatype (Dt)	Sl. No	Data identifier {H}	Datatype (Dt)
1	sample_id	Integer	**14**	Quality_id (output)	integer
2	grab_id	Integer	**15**	lab_qualifier	string
3	profile_id	Integer	**16**	mdl	decimal
4	sample_number	String	**17**	rdl	decimal
5	collect_datetime	datetime	**18**	text_value	string
6	depth_m	Decimal	**19**	sample_info	string
7	site_type	string	**20**	steward_note	string
8	area	string	**21**	replicates	integer
9	locator	string	**22**	replicate_of	integer
10	site	string	**23**	method	string
11	parameter	string	**24**	date_analyzed	date
12	value	decimal	**25**	data_source	string
13	units	string			

The further also extends the data visualization process to the preliminary analysis phase which carries out simplified operation on the data to understand its characteristic features. It also further provides better idea about the dynamics and nature associated with the dataset. This means that it indicates whether the dataset is in continuous or in discrete form. Also the preliminary operations helps to determine which variables should be considered as predictors (input) and which should be considered as response (output) towards evaluating the specific predictive model. The details of the preliminary operations are as follows from the WQ data exploration point of view.

4.3 Preliminary Analysis on d_{WQ}

The completion of the process of data visualization enables the further stage of computation for preliminary analysis of the computationally assessable structure of d_{WQ}. The following Fig. 3 shows that the depth histogram visualization of the probability distribution of the water depth measured data with respect to density. It clearly exhibits that significant amount of water lies in shallow depth measure. However, the visual interpretation of the outcome also exhibits that in some places the water depth belongs to 50 meters which is approximately 164 feet deep. It defines the density factor of the function and how the measure of probability lies within the distinct range of values and also it evaluates how for normal distribution the mean the deviation exists.

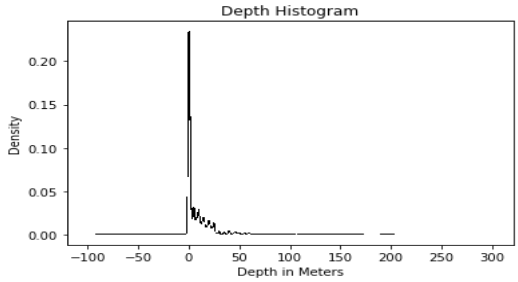
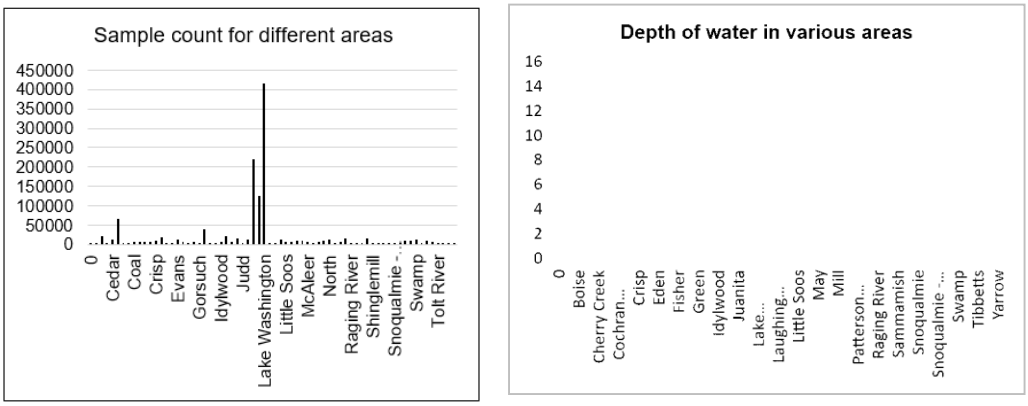


Fig. 3 Water Depth Histogram Visualization

The Following Fig. 4(a) shows the visualization of the data samples collected from different regions of the city. A proper interpretation from the visualization of the data

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN：1671-5497
E-Publication：Online Open Access
Vol: 42 Issue: 02-2023
DOI 10.17605/OSF.IO/QBNEU

shows that most of the data are taken from the lakes. It mostly ensures good WQ data collected by the sensors however, in the case of shallow water the data collected in this region might implicates low WQI which could also result in imbalance in the dataset.



(a) (b)

Fig. 4(a) Number of samples collected from different region, (b) Depth of water in various area

The analysis of Fig. 4(b) shows that depth of water in various regions. The interpretation of the dataset shows that frequency of shallow water is higher in the case creeks and other regions. On the other hand, deep water can be visualized in the case of Washington, lake Sammamish, and Lake Laughing etc. The visualization of the data also show that the most number of data is collected from the shallow waters where the number of data samples collected from the deep water is comparatively lesser in percentage.

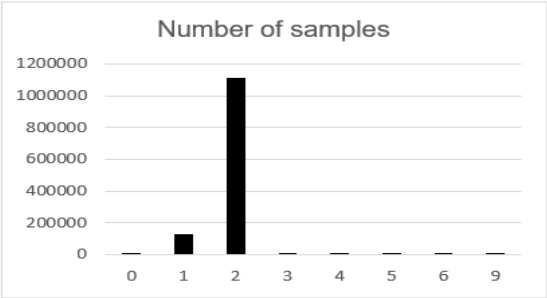


Fig: 5 Analysis of number of samples

The interpretation of the Fig. 5 shows that number of the collected samples for huge number of provisional data. The data sample range is approved through QC and also according to standard documentation. Here provisional data refers to the data where the final count might vary. The interpretation shows that high quality of data is in large number whereas limited number of poor and redundant WQ data also be in presence.

The study also defines the quality id for the collected data which is represented viz.

0 – Quality Unknown

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN：1671-5497
E-Publication：Online Open Access
Vol: 42 Issue: 02-2023
DOI 10.17605/OSF.IO/QBNEU

1 – Good Data, Passes Data Manager QC

2 – Provisional Data, Limited QC

3 – Questionable/Suspect

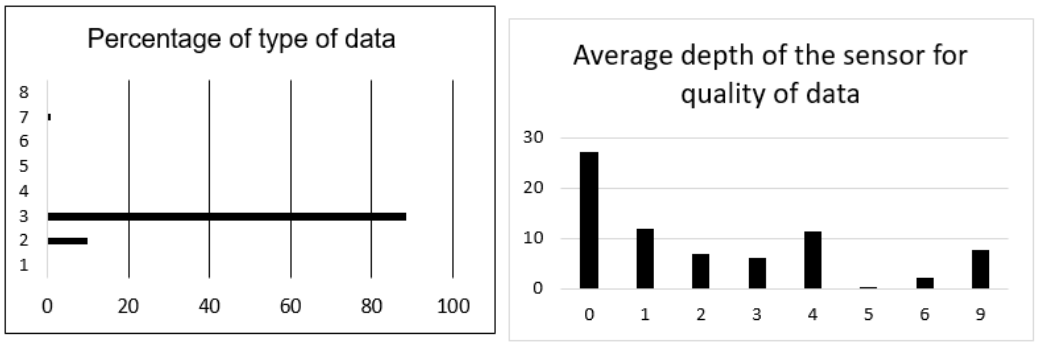
4 – Poor/Bad Data

5 – Value Changed (see Steward Note)

6 – Estimated Value

9 – Missing Value

The study also further assess the data of WQ type and evaluates the trend of the data in the following graph. The Fig. 6(a) exhibits that the quality of data value of third sample attribute poses higher degree of percentage corresponding to having poor quality and redundant attributes in comparison with the other sample type. This also indicate that the presence of third sample attribute could make the dataset imbalance and misleading. However, the study in this research work applies a regularized module to handle the inconsistency factors in the dataset and makes the data suitable for predictive analytics development strategy considering the aspects of ML.



(a) **(b)**

Fig: 6(a) Analysis of percentage of data samples, (b) Analysis of the depth measure for sensor for quality data

The Fig. 6(b) shows that average depth measure for the sensor deployed for the analysis of quality of capture WQ data. The study further performs the predictive analytics with the prepared dataset.

4.4 Predictive Analysis

During the design phase of the predictive modeling the study assess the nature of the dataset whether the dataset is continuous or discrete. This operations helps in computing the predictors (input) for the model and also it helps in determining the response variable in the form of output. This further helps in modeling the predictive framework. The design of the predictive model considers both classification and regression aspects and evaluates different ML based approaches such as Logistic Regression (LR), K-Nearest Neighbour (KNN), Random Forest Classifier (RF) and Gaussian NB Classifier (GNB) to

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN：1671-5497
E-Publication：Online Open Access
Vol: 42 Issue: 02-2023
DOI 10.17605/OSF.IO/QBNEU

carry out the predictive analysis based upon requirement. The classification techniques here evaluated for discrete dataset whereas in the similar fashion of implementation the regression models are evaluated for continuous dataset. The outcome of from the predictive analytics basically determines the reliability score of the specific model towards learning from the data. The quantified measure of statistics correspond to outcome also validates the performance of the learning model in verifying the WQMS data reliability which also helps in determining their localization factors. That means if localization of the sensors are precise then the model learning performance with respect to the reliable data will be higher with respect to accuracy. The following are the brief of the classifier and regression models used in the formulated predictive framework.

a. Logistic Regression: This refers to a baseline classification model which is having similarity with the linear regression (LR) model in its functional execution modeling. This functions with incorporating generalized linear model and also incorporates the strength factors of sigmoid as a connection function $\Sigma_c(x)$. This basis of classification model uses 1 weight factor for each variable and a bias. This study also verifies its suitability in linear data. That means it works best when the predictor and response data having linear relationship. The study have evaluated this model on the linear data and further extract the insights from the outcome as shown in the next segment 5.

b. Gaussian NB: The framework considered the potential factors of this model and utilized the concept of probabilistic classification considering the Bayes theorem on the WQ data. The model basically predict the outcome associated with the learning accuracy and also verifies the reliability of the collected WQ data. This Gaussian distribution model is best suited for probability based classification problems. The study also applied this model over the data for learning and assessed its outcome in terms of accuracy of water quality monitoring factors.

c. Random Forest Classification: This model characteristics are similar to the random forest regression schema. It differs from the random forest regression by defining a class in the leaf node instead of value. Final outcome of the classifier basically aggregates the results from the decision trees. It considers the count of classes during the aggregation process instead of averaging.

d. KNN Model: This model is a prominent classifier which maps the class corresponding to N nearest nodes and also computes the repeated count of neighbour to perform the clustering or classification of data.

5. RESULT AND DISCUSSION

The formulated predictive framework constructs a basis for implementing an efficient and automated model that could verify the reliability of the WQ data collected by the WQMS placement. This approach of validating the reliability of the WQ data not only ensures better sensor node placement strategy but also enhance the PA quality aspect from the monitoring point of view. The design and the development of the framework considers Anaconda tool where the programming language used is Python. The framework

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN：1671-5497
E-Publication：Online Open Access
Vol: 42 Issue: 02-2023
DOI 10.17605/OSF.IO/QBNEU

evaluates four different models for analyzing the WQ data which assist in further predictive analysis. The Fig. 7 further shows that comparative analysis of the outcome associated with four different ML approaches which also ensures the suitability of the implications of the ML approaches in enhancing the PA.

5.1 Performance Validation for WQ dataset

The prediction performance outcome on the WQ dataset are further applied on the classifiers KNN, Logistic Regression, Gaussian NB Classifier and Random Forest Classifier. The data samples mostly contains discrete dependent values and the models are further fed with this data to determine the reliability of the sensory data towards monitoring and assessing the WQI. The following Table 1 shows the quantified comparable outcome obtained for different performance metrices such as Accuracy, Precision, Recall and F1-score.

Table-2 shows quantified outcome obtained for the classification techniques in this context.

Algorithms	Accuracy	Precision	Recall	F1 score
KNN	0.83134	0.81252	0.8561	0.83374
Logistic Regression	0.83562	0.8524	0.82436	0.83814
Gaussian NB Classifier	0.91532	0.9142	0.91235	0.91327
Random Forest Classifier	0.9351	0.93143	0.9434	0.93737

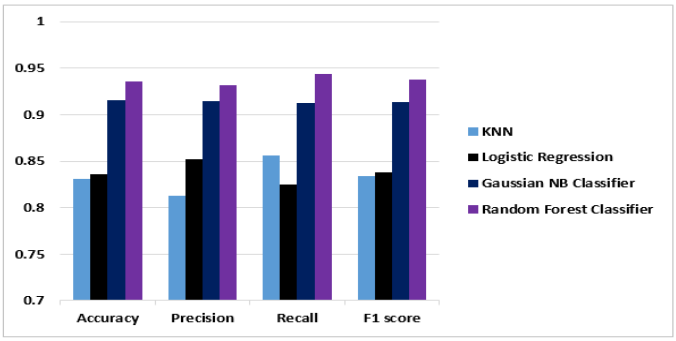


Figure: 7 Comparable outcome assessment for different models

The above Fig. 7 clearly shows that Random Forest Classifier attains significant performance in validating the reliability of the sensor data in terms of accuracy, precision. Recall and F1-score. The assessment of the metrices also shows the effectiveness of these algorithms on evaluating the training models. The overall outcome shows that although random forest outperforms other classifiers but Gaussian NB classifier also works effectively with the dataset that contains more discrete values. It has to be noted that if more discrete values are present in the dataset that indicates retention of better performance of the decision tree algorithms.

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN：1671-5497
E-Publication：Online Open Access
Vol: 42 Issue: 02-2023
DOI 10.17605/OSF.IO/QBNEU

6. CONCLUSION

The study introduces a novel framework modeling to assess the reliability of the WQMS data through predictive analytics. The study numerically designs and develops the framework in such a way where it provides a supporting baseline to carry out the implementation of four different predictive models such as KNN, Logistic Regression, Gaussian NB classier and Random Forest Classifier. The study targets to enhance the performance of WQ monitoring to a higher extent so that it could help in PA. In this regard the study not only evaluated the sensory data reliability through the predictive analytics but also the insights obtained from the learning helps in determining whether the sensor node placement strategy is effective or not. The study outcome from the classification and regression analysis viewpoint shows that Random Forest Classifier yields better accuracy, precision, Recall and F1-score that could lead to a sustainable and effective WQ monitoring environment in PA. It also provides a significant insight that RF works well with the discrete data samples as our WQ dataset also consists of more number of discrete data samples. The future work of the research further pave its path towards enhancing the performance aspects of environmental monitoring in broader spectrum through predictive analytics.

Reference

1. Vanishree K., Nagaraja G.S. (2021) Novel Secure Scheme for On-Field Sensors for Data Aggregation in Precision Agriculture. In: Silhavy R. (eds) Software Engineering and Algorithms. CSOC 2021. Lecture Notes in Networks and Systems, vol 230. Springer, Cham. https://doi.org/10.1007/978-3-030-77442-4_37
2. Sharma, A. Jain, P. Gupta, and V. Chowdary, ‘‘Machine learning applications for precision agriculture: A comprehensive review,’’ IEEE Access, vol. 9, pp. 4843–4873, 2021.
3. R. Taghizadeh-Mehrjardi, K. Nabiollahi, L. Rasoli, R. Kerry, and T. Scholten, ‘‘Land suitability assessment and agricultural production sustainability using machine learning models,’’ Agronomy, vol. 10, no. 4, p. 573, Apr. 2020, doi: 10.3390/agronomy10040573.
4. M. Keogh and M. Henry, ‘‘the implications of digital agriculture and big data for Australian agriculture,’’ Austral. Farm Inst., Sydney, NSW, Australia, Tech. Rep., 2016.
5. S. Ahmed, ‘‘Security and privacy in smart cities: Challenges and opportunities,’’ Int. J. Eng. Trends Technol., vol. 68, no. 2, pp. 1–8, Feb. 2020, doi: 10.14445/22315381/IJETT-V68I2P201.
6. K. G. Liakos, P. Busato, D. Moshou, and S. Pearson, ‘‘Machine learning in agriculture: A review,’’ Sensors, vol. 18, no. 8, p. 2674, 2018, doi: 10.3390/s18082674.
7. Borgia E. The Internet of Things vision: Key features, applications and open issues. Computer Communications. 2014 Dec 1; 54:1-31.
8. Barbosa AE, Fernandes JN, David LM. Key issues for sustainable urban stormwater management. Water research. 2012 Dec 15; 46 (20):6787-98.
9. Amuthadevi, C., Vijayan, D.S. & Ramachandran, V. Development of air quality monitoring (AQM) models using different machine learning approaches. J Ambient Intell Human Comput (2021). https://doi.org/10.1007/s12652-020-02724-2

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN：1671-5497
E-Publication：Online Open Access
Vol: 42 Issue: 02-2023
DOI 10.17605/OSF.IO/QBNEU

10. Li L, Rong S, Wang R, Yu S. Recent advances in artificial intelligence and machine learning for nonlinear relationship analysis and process control in drinking water treatment: A review. Chemical Engineering Journal. 2021 Feb 1; 405:126673.
11. Heo S, Nam K, Tariq S, Lim JY, Park J, Yoo C. A hybrid machine learning–based multi-objective supervisory control strategy of a full-scale wastewater treatment for cost-effective and sustainable operation under varying influent conditions. Journal of Cleaner Production. 2021 Apr 1; 291:125853.
12. Sundui B, Calderon OA, Abdeldayem OM, Lázaro-Gil J, Rene ER, Sambuu U. Applications of machine learning algorithms for biological wastewater treatment: Updates and perspectives. Clean Technologies and Environmental Policy. 2021 Jan 2:1-7.
13. Cai W, Long F, Wang Y, Liu H, Guo K. Enhancement of microbiome management by machine learning for biological wastewater treatment. Microbial Biotechnology. 2021 Jan; 14 (1):59-62.
14. Gencosman BC, Sanli GE. Prediction of Polycyclic Aromatic Hydrocarbons (PAHs) Removal from Wastewater Treatment Sludge Using Machine Learning Methods. Water, Air, & Soil Pollution. 2021 Mar; 232(3):1-7.
15. Yang J, Zhou A, Han L, Li Y, Xie Y. Monitoring urban black-odorous water by using hyperspectral data and machine learning. Environmental Pollution. 2021 Jan 15; 269:116166.
16. Zhao Y, Deng G, Zhang L, Di N, Jiang X, Li Z. Based investigate of beehive sound to detect air pollutants by machine learning. Ecological Informatics. 2021 Mar 1; 61:101246.
17. Chandra B, Middya AI, Roy S. Spatio-temporal prediction of noise pollution using participatory sensing. InEmerging Technologies in Data Mining and Information Security 2021 (pp. 597-607). Springer, Singapore.
18. Van Hauwermeiren W, Filipan K, Botteldooren D, De Coensel B. Opportunistic monitoring of pavements for noise labeling and mitigation with machine learning. Transportation Research Part D: Transport and Environment. 2021 Jan 1; 90:102636.
19. Alvares-Sanches T, Osborne PE, White PR. Mobile surveys and machine learning can improve urban noise mapping: Beyond A-weighted measurements of exposure. Science of the Total Environment. 2021 Jun 25; 775:145600.
20. Padwal SC, Kumar M, Balaramudu P, Jha CK. Analysis of environment changes using WSN for IOT applications. In2017 2nd International Conference for Convergence in Technology (I2CT) 2017 Apr 7 (pp. 27-32). IEEE.
21. REFERENCE FOR INTRODUCTION https://data.world/kingcounty/vwmt-pvjw#