

Ground Water Arsenic and Cancer Risk Assessment Prediction Model via Machine Learning: A step towards Modernizing Academic Research

Sobia Iftikhar¹, Sania Bhatti², Zulfiqar Ali bhatti³, Sanam Narejo⁴, Faisal Memon⁵, Mohsin Memon⁶

(Received October 03, 2020; Revised October 14, 2020; Accepted October 31, 2020)

Online First Publication

DOI: 10.33317/ssurj.232

Abstract— Ground water contamination with Arsenic (As), is one of the foremost issues in the South Asian countries, where ground water is one of the major sources of drinking water. In Asian countries, especially people of Pakistan living in rural areas are devouring ground water for drinking purpose and cleaned water is not accessible to them. This arsenic contaminated water is hazardous for human health. The persistence of this study is to understand the increasing level of arsenic in ground water, in coming years for Khairpur, Sindh Pakistan, which is also increasing the cancer rate (skin cancer, blood cancer) gradually in human body. To predict the arsenic value and cancer risk for the next five years, we have developed two models via Microsoft Azure machine learning with algorithms which includes; Support Vector Machine (SVM), Linear Regression (LR), Bayesian Linear Regression (BLR), Boosted Decision Tree (BDT). Exponential Smoothing (ETS) and Autoregressive Integrated Moving Average (ARIMA) methods are used for arsenic forecasting for the next five-years. The developed predictive model named as Arsenic Contamination and Cancer Risk Assessment Prediction (ACCRAP) model will assist for prediction of the arsenic contamination levels and the cancer rate. The obtained results demonstrated that BLR pose highest prediction accuracy of cancer rate among the four deployed machine learning algorithms.

Index Terms—Arsenic, Ground Water, Prediction Model, Health Hazards, Cancer Rate.

I. INTRODUCTION

Orthogonal Ground water is one of the major sources of drinking water in the World, specifically in Pakistan. Day by day countless changes occur in weather and environment conditions, which influences ground water and indirectly on human health. There are many elements present in ground water, like Arsenic (As), Fluoride (Fl), Zinc (Zn) etc. The undertaking of environmental changes with the passage of time generate the increasing trend of these elements which ultimately lead towards appearance of adverse types of diseases in the human body [1]. Arsenic is one amongst the toxic components reported in groundwater from over seventy countries, i.e., inflicting health hazards to concerning one hundred fifty million individuals worldwide [2]. Asia is

among one of the utmost expressively affected areas from arsenic contamination round the world. Ground water contamination due to arsenic became a vital public health issue in Sindh Pakistan [3]. The increasing level of arsenic in ground water is because of the deferral of arsenic compounds departing from Himalayas through Indus River, settling down in the course of year by geothermal, hydrological and biochemical factors becoming the part of ground water. Pakistan stood at number 80 among 122 countries relating to contaminated water quality [4]. This study predicts the elevated level of arsenic in the coming years for district Khairpur, Sindh Pakistan and the way cancer rate are raised because of arsenic changes in the body. In rural areas of Sindh, arsenic component is gradually developed in the body of the people because of consuming contaminated water and food, and its exposure comes within a variety of skin, blood and scalp cancer [5]. Thus, we have developed Arsenic Contamination and Cancer Risk Assessment Prediction (ACCRAP) model using “Microsoft Azure Machine Learning Studio” [6].

The three main contributions of this study are:

- i. **First**, to the best of knowledge this is the foremost attempt towards the forecasting of arsenic for the next five years with Exponential Smoothing (ETS) and Autoregressive Integrated Moving Average (ARIMA) methods.
- ii. **Second**, for the first-time cancer rate prediction is done with four different machine learning algorithms.
- iii. **Third**, cancer rate prediction is calculated manually and automatically then both results are compared.

The paper is consisting of six sections. **Sections II** justifies how this study modernizes academic research. **Section III** presents methodology details. **Section IV** focuses on implementation of ACCRAP Model in “Microsoft Azure Machine Learning Studio”. **Section V** discusses results. The last section, i.e., **Section VI**, discusses conclusion and future work directions.

¹Student, Department of Software Engineering, MUET, Jamshoro, Pakistan. sobiaiftikharali@gmail.com

²Associate Professor, Department of Software Engineering, MUET, Jamshoro, Pakistan. sania.bhatti@faculty.muuet.edu.pk

³Assistant Professor, Department of Chemical Engineering, MUET, Jamshoro, Pakistan. zulfiqar.bhatti@faculty.muuet.edu.pk

⁴Assistant Professor, Department of Computer Systems Engineering, MUET, Jamshoro, Pakistan. sanam.narejo@faculty.muuet.edu.pk

⁵Head of ICT Operations, Fauji Fertilizer Bin Qasim Limited Karachi, Pakistan. faisal.memon@ffbl.com

⁶Associate Professor, Department of Software Engineering, MUET, Jamshoro, Pakistan. mohsin.memon@faculty.muuet.edu.pk

II. A STEP TOWARDS MODERNIZING ACADEMIC RESEARCH

From literature study, the wide applicability of machine learning techniques has been comprehended [7-9]. This study is considered as one of the vigorous steps towards modernizing academic research because it is presenting a mapping between water research and artificial intelligence techniques. To conduct this research first of all papers related with 'As' levels in different areas of Sindh were studied [10-14]. Then health effects caused due to drinking 'As' contaminated water were studied [15]. Further applicability of machine learning techniques for 'As' prediction was analyzed [16-18]. The dataset used in the experiments of this study is collected from the published papers. During the experiments, two forecasting and four machine learning methods are deployed which make this study a preliminary step towards modernization of water research.

A. Data Collection

For this study, the dataset has been collected [19-22]. A group of researchers analyzed more than two hundred ground water samples from two sides of river Indus. From that study the arsenic levels of Khairpur district are taken for the year 2013. In another study 180 ground water samples were collected from the Khairpur Mir and Tharparkar, Sindh Pakistan [23]. They identified the arsenic in the scalp hair of different ages of people. From that study the arsenic levels of Khairpur district are taken for the year 2014. Researchers also evaluated arsenic values in groundwater of different communities in Khairpur district with the help of multivariate techniques [24]. We took arsenic values of Khairpur district for the year 2016 from this study. Authors scrutinized total arsenic, copper, iron, nickel, lead and zinc contents in two hundred forty three, hand-pump ground water samples and ninety tube-well ground water samples of Sobhoderi, Sindh, Pakistan [25]. They adapted multivariate techniques and cluster analysis for polluting elements. It was found that arsenic concentration, in most of hand-pump and tube-well samples, were higher than the World Health Organization (WHO) permissible limits. Arsenic levels of Khairpur district for the year 2017 are taken from this study.

B. Statistical Data

The datasets may contain missing quantities for different reasons. One tactic to addressing this issue is to get rid of assumptions which lack details. Yet, there is a chance of losing concentrated information with critical details. The missing quantities are due to a superior method. As such, we must construe, from the existing portion of knowledge, certain missing quantities. From the literature review we collected the data for 2013, 2014, 2016, 2017 which is in "Dataset collection" section, but for the missing data of 2015, we used mean/median for non-missing values in a column using Eq. (1). According to Eq. (1) we used previous year (2014) and one next year (2016) data points to fill up the missing values of 2015 years by using median/mean.

$$x = \frac{\sum 2014 + \sum 2016}{2} \quad (1)$$

III. METHODOLOGY

Working mechanism for this study is shown in Fig. 1. There are some important steps in order to get desired results for this research. The real time data that is 'As' levels are collected from literature for different locations of Khairpur Sindh Pakistan. The raw data is prepared in Comma-Separated Values (CSV) format which is deployable in machine learning platform. The Microsoft Machine learning studio is used for performance of various experiments with dataset file. Experiments are based on two different models. In first model, we deployed arsenic values of five years to forecast it via ETS and ARIMA methods, employing previous year's arsenic data. Exponential smoothing and ARIMA are extensively used approaches for time series estimation and work on both seasonal and non-seasonal data. For the cancer prediction we first computed cancer rate manually using Eq. (2). For the manual computation, arsenic value is used as an input parameter, so five-years forecasted arsenic value is used in this equation and results are acquired. Further, the second model for predicting, five-year cancer rate values is used with dissimilar machine learning algorithms. In the last, R-script is used for comparing both (predicted and manual) cancer rate results and outcomes are gained in the graph form.

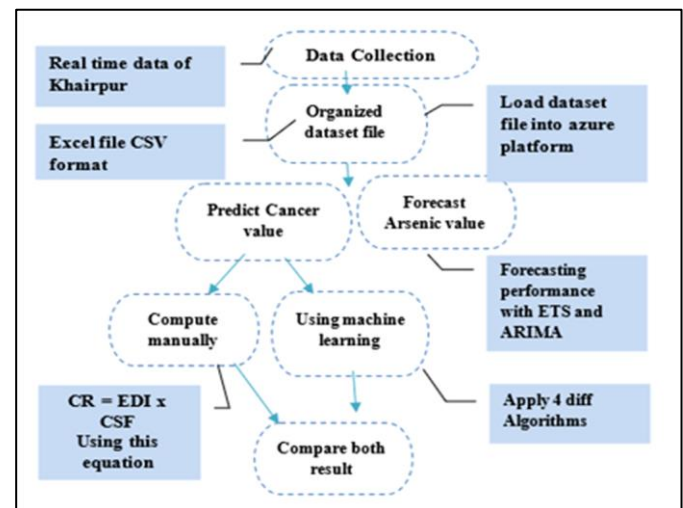


Fig. 1: Steps of Workflow

IV. IMPLEMENTATION

We used "Microsoft Azure Machine Learning Studio", for developing our predictive model. Two different models were developed, first one for arsenic prediction named as 'Arsenic Prediction Model' and second one for cancer rate prediction named as 'Cancer Prediction Model'. Fig. 2 presents 'As' prediction model. The 'As' levels of different areas of Khairpur district, are maintained in CSV dataset file. It is important to mention that 'select column in dataset' module is one commonly used feature of machine learning studio. Next two core columns 'Year and Arsenic' are selected from loaded dataset in order to apply two different forecasting methods ETS, ARIMA. These both methods are applied on the Arsenic dataset to forecast the future values by using forecast library. These both methods are scripted in R-script. The R-script is used for making an average of forecast points and further for representing the outcome into graph.

R-scripting is deployed for ETS method and its pseudocode is presented in our published work [26]. The libraries used in the R-scripting are; Forecast Zoo and Ggplot2.

A. R-Scripting

By including R-scripting in Fig. 2 model, we can accomplish a variety of modified tasks that are not available in ‘Microsoft

Azure Machine Learning Studio’ [27]. For example, ‘Create’ custom information changes. We can utilize our measurements for assessing forecasts and manufacture models utilizing calculations that are not executed as independent modules in Studio.

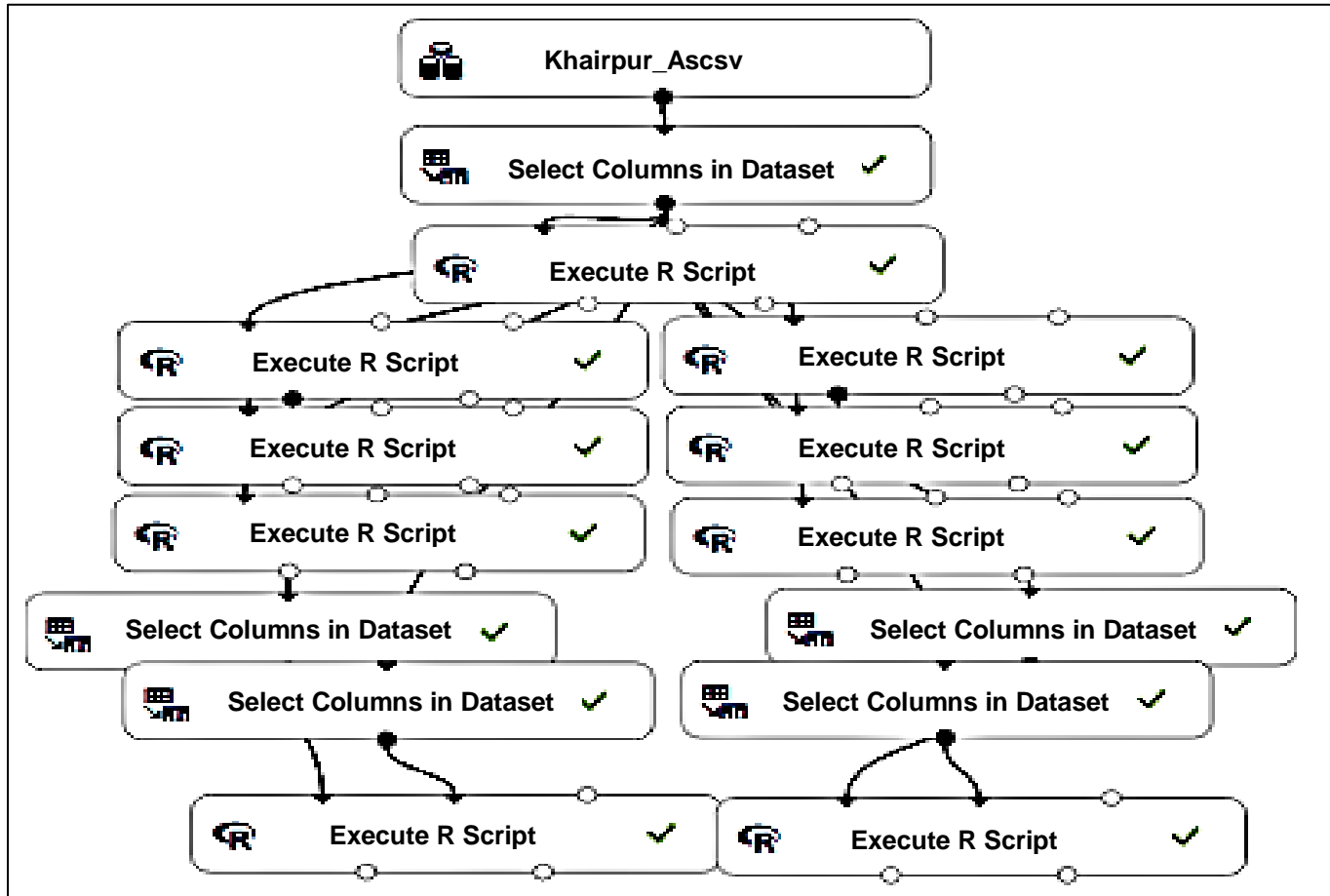


Fig. 2: Arsenic Prediction Model

The predictive model used for the cancer rate in relation to the Arsenic Prediction Model is shown in Fig. 3.

Three columns Year, P-arsenic and Cancer-value are selected with the help of “Select Column Dataset” module from the loaded datasets. Then ‘Split Data’ module is used for separating the rows of loaded data set into training and testing through specified percentages. Then BLR, BDT, LR, SVM algorithms are deployed. Among these the best algorithm is identified that gives best accuracy in Evaluate module.

The dataset ‘Cancer Rate’ for this Fig. 3 contains last five-years arsenic and cancer value, this cancer value is calculated manually by using the formulas detailed in study [2].

Evaluate module used in cancer prediction model is deployed for computing the accuracy of four machine learning algorithms. Evaluate module describes the accuracy of developed model that how much accurate model is trained for testing. Then this model is converted into predictive model using web service. We have Trained Model as an Application Programming Interface (API), for predicting new cancer

values with resultant values of arsenic from Arsenic Prediction Model.

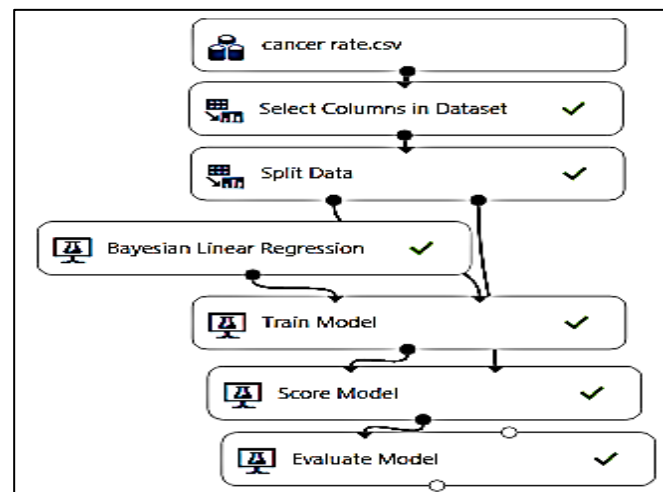


Fig. 3: Cancer Prediction Model

We used five forecasted arsenic values obtained from our Arsenic Prediction Model to give input to this Trained Model and obtained five predicted cancer values.

The Fig. 4 illustrates Trained Model in which web services are deployed intended for converting the obtained model into the predictable model. The main components of the Trained Model are web service input, data file, column selections, Cancer Prediction Model, Score Model and web service output. This web service is one of the advance features of machine learning platform; it is deployed to convert Trained Model into predicted model. This platform provides test request response environment for real time prediction based on the provided input. In this model, we deployed five forecasted ‘As’ values, which predict next five-year value of cancer with the new forecasted arsenic value at run time. This Trained Model could be used as API for commercial purpose.

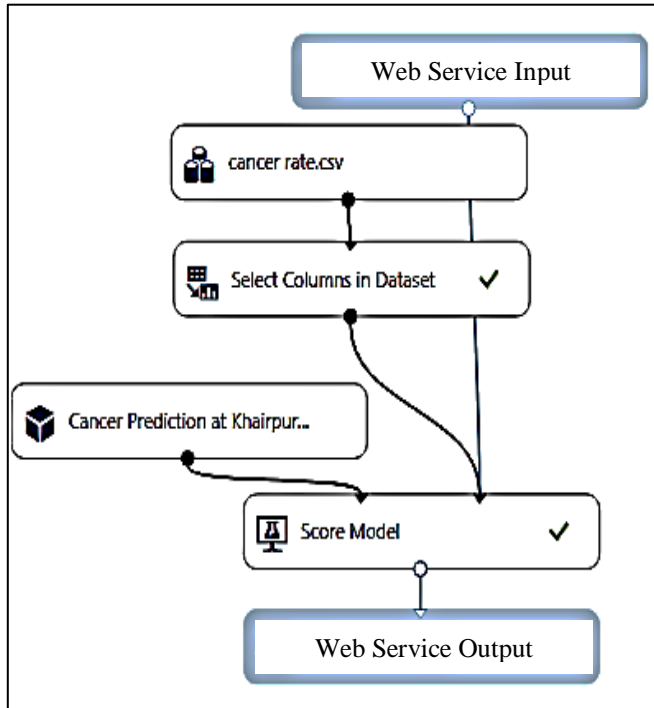


Fig. 4: Trained Model for Web Services

V. RESULTS AND DISCUSSION

First, we have organized our dataset of last five-years (2013 to 2017) of district Khairpur from studies [19-22]. District Khairpur relates to left zone of Indus river bank. During experiments, ETS and ARIMA methods are used for arsenic forecasting for the next five-years. Here BLR, BDT, LR, SVM machine learning algorithms are deployed for envision of the cancer rate. R-scripting is deployed; aimed at the middling of forecast point and imagining of the obtained results.

A. Arsenic Outcome Scrutiny

Table I reveals arsenic forecast point through columns named as “F-year”, “F-AS”, afterward applying ‘As’ predictive model. F-AS specify the average rate of forecast points altogether which is obtained via ETS and ARIMA forecasting methods. The seasonality is also generated with column named F-year through ARIMA and ETS. These both method

ETS and ARIMA are scripted in R-script. P-year (previous year) and P-As (previous arsenic value) shows previous average rate of arsenic which is collected from dataset section. With help of ARIMA and ETS trend, observed, seasonality, randomness of the arsenic results is derived. That is why these both are good for forecasting because they detect these four factors automatically with the help of previous data. The increasing trend of arsenic from 2013 to 2022 is depicted in Fig. 5. Here the curve without shade shows the previous average rate of arsenic from the year 2013 to 2017 and the curve with light blue shade shows the forecasted values of arsenic from the year 2018 to 2022. This curve is obtained after deploying R-script.

Table I: Arsenic Values According To Year (F-AS)

District	P-year	P-As	F-year	F-AS
Khairpur	2013	5.45	2018	49.03
Khairpur	2014	29.9	2019	59.85
Khairpur	2015	37.35	2020	70.63
Khairpur	2016	44.8	2021	81.49
Khairpur	2017	51.75	2022	92.31

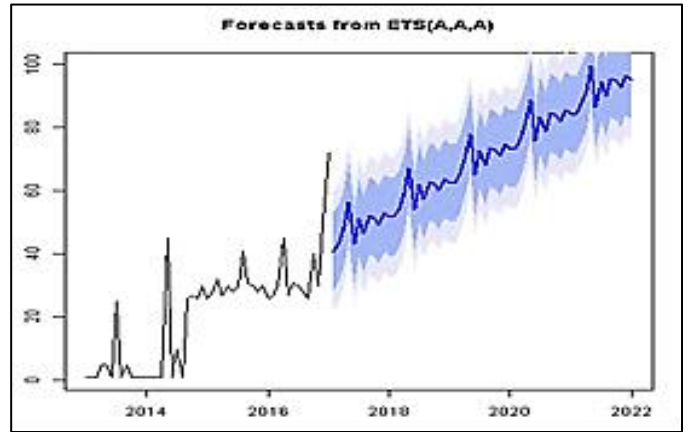


Fig. 5: Trending View of Predictive Arsenic

B. Cancer Outcome Scrutiny

We have calculated cancer rate by two methods. First, we have calculated the cancer rate manually and then we have projected subsequent five-years cancer value by means of model presented in Fig. 3. The results of both methods are presented in Table II, in which first column shows the results of manually calculated cancer-rate and second column presents predicted cancer rate values.

The cancer rate is obtained using the well-defined and extensively adapted method described in literature [19], [28], [29]. They have used Eq. (2) to calculate the cancer rate; deploying arsenic value of one year. We are forecasting five-years arsenic value via model 1 illustrated in Fig 1. The values which are obtained from that model are then deployed as an input value to the Eq. (2) and cancer rate is computed manually that totally rely on two factors, i.e., EDI and CSF. Where, EDI is the quantity of daily intake of water for human body. According to studies arsenic value is increasing in drinking water day by day that raises Cancer Slop Factor (CSF).

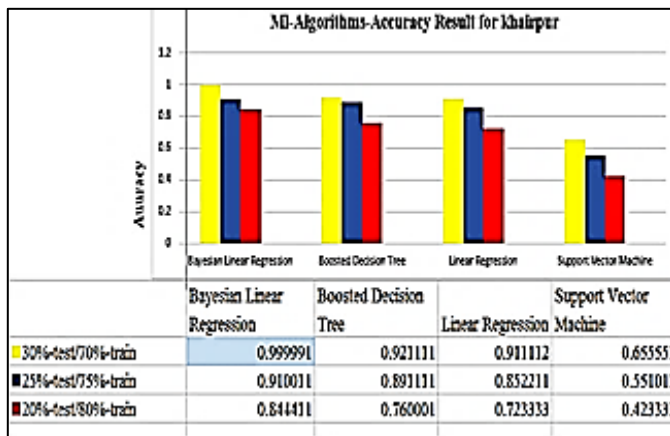
$$CR = EDI \times CSF$$

(2)

Table II: Manual and Predictive Cancer Rate

Year	As Values, Cancer Rate According to Year			
	Arsenic-Value	Manual Cancer-Rate	Predicted Cancer-Rate	Accuracy
2013	5.45	128.6588	-	-
2014	29.9	705.8529	-	-
2015	37.35	881.7259	-	-
2016	44.8	1057.599	-	-
2017	51.75	1221.668	-	-
2018	49.03	1157.489	1156.882	99.9%
2019	59.85	1412.946	1393.044	98.5%
2020	70.63	1668.403	1652.822	99.0%
2021	81.49	1923.861	1912.601	99.4%
2022	92.31	2179.318	2172.379	99.6%

After manually computation we used same forecasted arsenic values from the model demonstrated in Fig. 2, to predict the next five-year cancer rate values from the model demonstrated in Fig. 3 and the results are shown in the “Predicted” column of Table II. The accuracy column tells how much percent the predictable results are near to manual results.

**Fig. 6: Accuracy of Three Splitting Ratio by Four ML-Algorithms for Khairpur**

The cancer dataset is divided into testing and training parts. Training part of the dataset is used to train the cancer model and testing part of dataset is used for validating the Trained Model for new prediction. We split dataset into different ration and deployed cancer predictive model on three different splitting ratios as shown in Fig. 6:

- Sr-01: 30%-test and 70% (training)
- Sr-02: 25%-test and 75% (training)
- Sr-03: 20%-test and 80% (training)

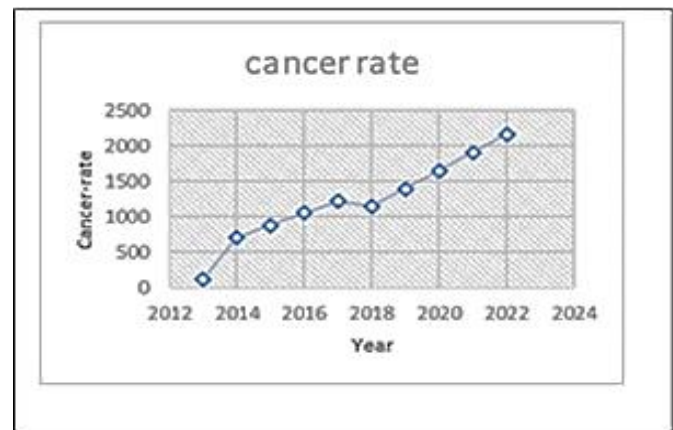
It is apparent from Fig. 6 that the 70 percent of the data deployed to train the machine learning models and remaining 30 percent of the test dataset deployed to validate the predictive model gives highest accuracy among the three splitting ratios. We used three models of regression and one model of binary classification to envision cancer rate. We found the accuracy of these four algorithms and are portrayed in Table III. It is apparent from the Table III that the accuracy of three regression algorithms is much accurate than the SVM method. The BLR algorithm portrays the highest prediction accuracy of cancer rate. It is also obvious from Fig. 6 that in

three different splitting ratios the SVM method poses lowest accuracy than the other three regression algorithms.

Table III: Accuracy Comparison of Algorithms

Algorithms	Accuracy
BLR	0.999
BDT	0.921
LR	0.911
SVM	0.65

In Fig. 7, increasing trend of cancer rate is depicted from 2013 to 2022. From the graph, it is evident that due to gradual increase in the arsenic contamination from the year 2018 to 2022 in the Khairpur district, the cancer rate is also showing the increasing trend accordingly. This graph is gained following deployment of the Trained Model shown in Fig. 4. Trained Model with the projected arsenic input gives an average of five-year cancer value at run time.

**Fig. 7: Cancer Rate with Respect to Year**

VI. CONCLUSION

Literature is evident that arsenic contamination studies have got immense attention around the world from the past decade due to its hazardous effects on the human health. This is the foremost effort to deploy ‘Microsoft Azure Machine Learning Studio’ intended for the prediction of arsenic and cancer rate. For this study; Khairpur district is carefully chosen and the arsenic values intended for the subsequent five-years and consequences on the wellbeing of the individuals in terms of cancer are predicted. Further, upturn in arsenic levels is found via the models of ‘Arsenic Prediction’ and ‘Cancer Rate’. It is manifest from the obtained results that the cancer rate is correspondingly illustrating growing trends due to increasing trend of arsenic from 2018 to 2022. Prediction and forecasting methods are implemented which makes this study different from the previous studies. One of the limitations of this work is that the web service developed in this work is currently able to predict the arsenic values for just Khairpur district.

The accuracy comparison of three regression and one binary classification algorithms revealed that the accuracy of BLR is highest among the four algorithms. One of the future directions of this work is the application of the developed ACCRAP model on the dataset of other arsenic affected regions of the Sindh, Pakistan. Another future direction could be to analyze the arsenic contamination levels in water

according to seasons because the intake of water fluctuates according to seasons.

ACKNOWLEDGMENT

This project idea achieved: Microsoft “AI for Earth Grant”. We are grateful to the Microsoft for their support and MUET for conducting this research work.

REFERENCES

- [1] Brammer, H., & Ravenscroft, P. (2009). Arsenic in groundwater: a threat to sustainable agriculture in South and South-east Asia. *Environment International*, 35(3), 647-654.
- [2] Bhatti, Z., Qureshi, K., Bhatti, I., Unar, I. N., & Khuhawar, M. Y. (2017). Determination of Arsenic and Health Risk Assessment in the Ground Water of Sindh, Pakistan.
- [3] Kori, A. H., Jakhrani, M. A., Mahesar, S. A., Shar, G. Q., Jagirani, M. S., Shar, A. R., & Sahito, O. M. (2018). Risk assessment of arsenic in ground water of Larkana city. *Geology, Ecology, and Landscapes*, 2(1), 8-14.
- [4] Azizullah, A., Khattak, M. N. K., Richter, P., & Häder, D. P. (2011). Water pollution in Pakistan and its impact on public health—a review. *Environment international*, 37(2), 479-497.
- [5] Fatmi, Z., Abbasi, I. N., Ahmed, M., Kazi, A., & Kayama, F. (2013). Burden of skin lesions of arsenicosis at higher exposure through groundwater of taluka Gambat district Khairpur, Pakistan: a cross-sectional survey. *Environmental geochemistry and health*, 35(3), 341-346.
- [6] McCaffrey, J. (2014). *Microsoft Azure : Introduction to Machine Learning Studio*. Msdn-Magazine. Retrieved from: <https://docs.microsoft.com/en-us/archive/msdn-magazine/2014/september/microsoft-azure-introduction-to-machine-learning-studio>.
- [7] Qin, S. J., & Chiang, L. H. (2019). Advances and opportunities in machine learning for process data analytics. *Computers & Chemical Engineering*, 126, 465-473.
- [8] Singh, S. K., Taylor, R. W., Rahman, M. M., & Pradhan, B. (2018). Developing robust arsenic awareness prediction models using machine learning algorithms. *Journal of environmental management*, 211, 125-137.
- [9] Barzegar, R., Moghaddam, A. A., Deo, R., Fijani, E., & Tziritis, E. (2018). Mapping groundwater contamination risk of multiple aquifers using multi-model ensemble of machine learning algorithms. *Science of the total environment*, 621, 697-712.
- [10] Uqaili, A. A., Mughal, A. H., & Maheshwari, B. K. (2012). Arsenic contamination in ground water sources of district Matiari, Sindh. *International Journal of Chemical and Environmental Engineering*, 3(4-2012).
- [11] Sanjrani, M. A., Mek, T., Sanjrani, N. D., Leghari, S. J., Moryani, H. T., & Shabnam, A. B. (2017). Current situation of aqueous arsenic contamination in Pakistan, focused on Sindh and Punjab Province, Pakistan: A review. *J Pollut Eff Cont*, 5(207), 2.
- [12] Naseem, S., & McArthur, J. M. (2018). Arsenic and other water-quality issues affecting groundwater, Indus alluvial plain, Pakistan. *Hydrological Processes*, 32(9), 1235-1253.
- [13] Majidano, S. A., Arain, G. M., Bajaj, D. R., Iqbal, P., & Khuhawar, M. Y. (2010). Assessment of groundwater quality with focus on arsenic contents and consequences. Case study of Tando Allahyar District in Sindh Province. *Intl J Chem & Environ Eng*, 1(2).
- [14] Bhatti, Z., Qureshi, K., Bhatti, I., Unar, I. N., & Khuhawar, M. Y. (2017). Determination of Arsenic and Health Risk Assessment in the Ground Water of Sindh, Pakistan.
- [15] Ali, W., Rasool, A., Junaid, M., & Zhang, H. (2019). A comprehensive review on current status, mechanism, and possible sources of arsenic contamination in groundwater: a global perspective with prominence of Pakistan scenario. *Environmental geochemistry and health*, 41(2), 737-760.
- [16] Singh, S. K., Taylor, R. W., Rahman, M. M., & Pradhan, B. (2018). Developing robust arsenic awareness prediction models using machine learning algorithms. *Journal of environmental management*, 211, 125-137.
- [17] Barzegar, R., Moghaddam, A. A., Deo, R., Fijani, E., & Tziritis, E. (2018). Mapping groundwater contamination risk of multiple aquifers using multi-model ensemble of machine learning algorithms. *Science of the total environment*, 621, 697-712.
- [18] Park, Y., Ligaray, M., Kim, Y. M., Kim, J. H., Cho, K. H., & Sthiannopkao, S. (2016). Development of enhanced groundwater arsenic prediction model using machine learning approaches in Southeast Asian countries. *Desalination and Water Treatment*, 57(26), 12227-12236.
- [19] Bhatti, Z., Qureshi, K., Bhatti, I., Unar, I. N., & Khuhawar, M. Y. (2017). Determination of Arsenic and Health Risk Assessment in the Ground Water of Sindh, Pakistan.
- [20] Jakhrani, M. A., Qasim, M., Ali, S., & Memon, S. F. (2018). Monitoring of Elemental Contamination in Groundwater Samples of Sobhodero Khairpur, Sindh, Pakistan. *International Journal of Sciences*, 7(10), 10-23.
- [21] Brahman, K. D., Kazi, T. G., Afridi, H. I., Baig, J. A., Arain, S. S., Talpur, F. N., ... & Arain, M. B. (2016). Exposure of children to arsenic in drinking water in the Tharparkar region of Sindh, Pakistan. *Science of The Total Environment*, 544, 653-660.
- [22] Qasim, M., & Ali Jakhrani, M. (2018). Physicochemical and elemental contamination assessment in groundwater samples of Khairpur Mir's, Pakistan. Human and Ecological Risk Assessment: *An International Journal*, 24(4), 870-885.
- [23] Brahman, K. D., Kazi, T. G., Afridi, H. I., Baig, J. A., Arain, S. S., Talpur, F. N., ... & Arain, M. B. (2016). Exposure of children to arsenic in drinking water in the Tharparkar region of Sindh, Pakistan. *Science of The Total Environment*, 544, 653-660.
- [24] Qasim, M., & Ali Jakhrani, M. (2018). Physicochemical and elemental contamination assessment in groundwater samples of Khairpur Mir's, Pakistan. Human and Ecological Risk Assessment: *An International Journal*, 24(4), 870-885.
- [25] Jakhrani, M. A., Qasim, M., Ali, S., & Memon, S. F. (2018). Monitoring of Elemental Contamination in Groundwater Samples of Sobhodero Khairpur, Sindh, Pakistan. *International Journal of Sciences*, 7(10), 10-23.
- [26] Bhatti, S., Memon, M. A., & Bhatti, Z. A. Groundwater Arsenic and Health Risk Prediction Model using Machine Learning for TM Khan Sindh, Pakistan.
- [27] Docs.Microsoft.com, (2020). *Execute R script*. Msdn-Magazine. Retrieved from: <https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/execute-r-script>.
- [28] Rasool, A., Xiao, T., Farooqi, A., Shafeeqe, M., Masood, S., Ali, S., ... & Nasim, W. (2016). Arsenic and heavy metal contaminations in the tube well water of Punjab, Pakistan and risk assessment: A case study. *Ecological Engineering*, 95, 90-100.
- [29] Rasool, A., Farooqi, A., Xiao, T., Masood, S., & Kamran, M. A. (2016). Elevated levels of arsenic and trace metals in drinking water of Tehsil Mailsi, Punjab, Pakistan. *Journal of Geochemical Exploration*, 169, 89-99.