Water Resources Management
https://doi.org/10.1007/s11269-022-03177-2

Predicting and Forecasting Mine Water Parameters Using a Hybrid Intelligent System

Kagiso Samuel More[1] · **Christian Wolkersdorfer**[1]

Received: 2 November 2021 / Accepted: 4 May 2022

Abstract
Water treatment plants need to stock chemicals and have enough energy as well as human resources to operate reliably. To avoid a process interruption, proper planning of these resources is imperative. Therefore, a scientifically based, practical tool to predict and forecast relevant water parameters will help plant operators to know in advance which chemicals and methods are necessary for polluted water management and treatment. This study aims to develop a system to predict and forecast mine water parameters using electrical conductivity (EC) and pH of mining influenced water from the Acid Mine Drainage treatment plant in Springs, South Africa as an example. Three machine learning algorithms, namely random forest regression, gradient boosting regression and artificial neural network (ANN) were compared to find the best learning model to be used for predictive analysis. These models were developed using historical data of the years 2016 to 2021. Input variables of the models are turbidity, total dissolved solids, SO_4 and Fe, with EC and pH being the target outputs. Results of the models have been compared with the measured data on the basis of the mean absolute error and root mean square error. The results show that random forest and gradient boosting models perform better than the ANN model, and thus these models were deployed as a web application. The Long Short-Term Memory technique was used to forecast the input parameter values for 60 days, and these values were used to get the future values for EC and pH for the same period.

Keywords Mining Influenced Water · Machine Learning · Predictive Analysis · Web Application · South Africa

✉ Christian Wolkersdorfer
Christian@Wolkersdorfer.info

Kagiso Samuel More
210042091@tut4life.ac.za

[1] SARChI Chair for Mine Water Management, Department of Environmental, Water and Earth Sciences, Tshwane University of Technology, Private Bag X680, Pretoria 0001, South Africa

Published online: 17 May 2022

1 Introduction and Background

Mining influenced water (MIW), especially acid mine or rock drainage (AMD/ARD) is a challenging problem encountered by many mining companies around the world (Verburg 2011). MIW occurs when iron-(di)sulfide minerals such as pyrite, marcasite or pyrrhotite react with water and oxygen (Blowes et al. 2014; Singer and Stumm 1970). In case no buffer minerals such as carbonates are present, these reactions will result in AMD/ARD which, depending on the geological settings, may contain potentially toxic elements (Wolkersdorfer 2008). Water treatment plants, including AMD treatment plants, need to stock chemicals, have steady and enough energy and human resources to operate reliably. To avoid a process interruption, proper planning of these resources is imperative. Therefore, a scientifically based, practical tool to predict and forecast relevant water parameters will help plant operators to know in advance which chemicals and methods to use to treat and manage polluted water.

Machine learning (ML) models are divided into three categories: supervised, unsupervised and reinforcement learning, whereas this study applied supervised learning algorithms (Online Resource 1). Supervised learning is where a model has input variables and an output value and uses an algorithm to learn the mapping function from the input to the output (Swamynathan 2017). There are several tasks within supervised learning such as classification and regression. In this study, the main aim was on accuracy of the results; thus the regression ML algorithms were explored, specifically random forest, gradient boosting and neural networks, whilst the forecasting technique used is the Long Short-Term Memory (LSTM) method.

Mine water quality can be evaluated using several parameters, e.g. electrical conductivity (EC), pH, major ions, turbidity or acidity. The amount of dissolved minerals in mine water is represented by the total dissolved solids (TDS) and EC, which is crucial for mine water conditions. Because EC can be used to calculate the TDS based on the function $TDS = f(EC)$, it can be an indicator parameter for additional mine water constituents (Hem 1985; Hubert and Wolkersdorfer 2015). Therefore, EC is one of the target outputs in the developed ML models of this study. In addition, pH measurements are important in determining metal concentrations that might be dissolved in MIW and the volume of clean water that can be discharged. Thus pH is a further target output used in this study. Treated mine water is usually discharged into receiving water courses or for industrial and commercial use. Thus, it is crucial to know the pH values of mine water entering the treatment plant beforehand so it can be increased or decreased accordingly to precipitate unwanted, potentially polluting metals.

So far, different techniques have been applied to predict the future mine water quality for optimised MIW management (e.g. Khandelwal and Singh 2005; Liu et al. 2019; Rooki et al. 2011). Modelling software such as GoldSim, MATLAB Simulink, Geochemists Workbench and PHREEQC have been previously favoured for this purpose (e.g. Nalecki and Gowan 2008; Usher et al. 2010); however, researchers are trying to optimise these predictions by including artificial intelligence (AI) technology (More et al. 2020; Vadapalli et al. 2020). This study will present ML techniques in the form of a hybrid intelligent system to predict, to our knowledge, for the first time the indicator mine water parameters EC and pH at a mine water treatment plant. For the models and forecasts, this study uses historical data of the South African Eastrand plant east of Johannesburg from the years 2016 to 2021 (Fig. 1).

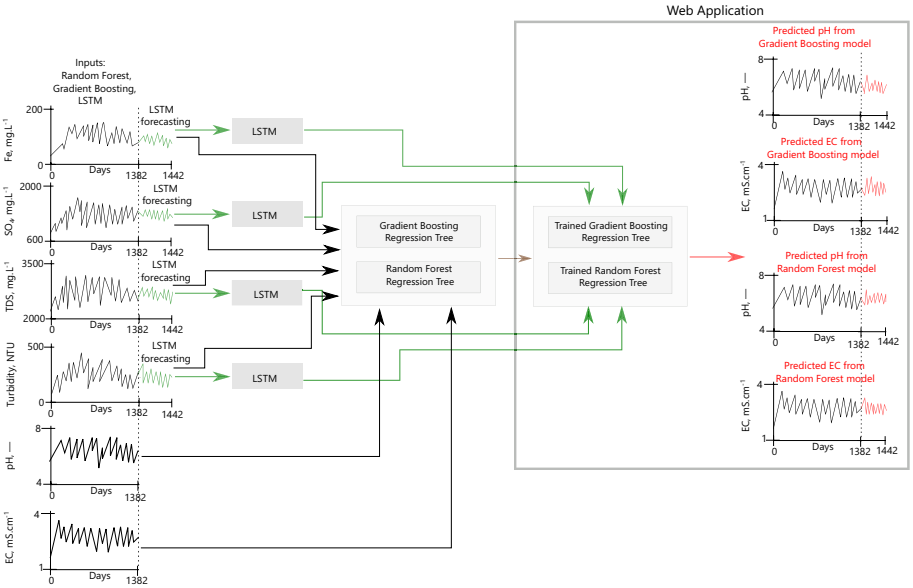

Fig. 1 Electrical conductivity and pH prediction structure showing the connection between different algorithms used in this study (modified and supplemented after Vadapalli et al. 2020)

2 Machine Learning Algorithms Used

2.1 Random Forest Regression

Random forest, proposed by Breiman in the early 2000s, is a supervised ML algorithm which uses ensemble learning to perform either classification or regression tasks (Breiman 2001) and is built on the concept of the decision trees algorithm. In this algorithm, the trees consist of decision nodes where the data are split and the leaves where the final outcomes and decisions are made (Online Resource 2; Boulesteix et al. 2012). In this study, the focus is on random forest regression, therefore regression trees were used to build a random forest model. These random forests are operated by building multiple decision trees at training time and give a mean prediction of the individual trees (e.g. Belgiu and Drăguţ 2016; Chen et al. 2020). Random forest algorithm applies a technique known as "bagging" (Mosavi et al. 2021), which is used to reduce the variance for algorithms that mainly have high variance, such as decision trees. Bagging enables random forests to make decision trees run independently and ultimately aggregates the outputs to give the final output.

2.2 Gradient Boosting Regression

Gradient boosting regression is an algorithm similar to random forest, which uses an ensemble or decision trees to predict a target label. While random forest focus on reducing variance on complex trees, gradient boosting's main aim is to decrease the bias of simple trees and make them more expressive (Johnson et al. 2017; Zhang and Haghani 2015). This technique optimises the predictive value of a model over multiple steps in the learning

process. Each iteration process of the decision tree (Online Resource 3) adjusts the values of the weights, coefficients or biases for each input variable used to build the model to predict the target value, with the main goal to reduce the loss function, i.e. the difference between the predicted and actual target values. The incremental adjustment made in each step of the model is the gradient, while boosting can be referred to as the activity of speeding up the improvement in predictive accuracy to a desired value.

2.3 Neural Networks

Artificial neural networks (ANN) are data-driven systems that work with known input data without any assumptions. These neural networks are made up of a) input layers, where data are initially presented to the model and computation is performed, b) a hidden layer, where data are processed, and c) an output layer, where the results are produced (Online Resource 4; Russell and Norvig 2002; Wolfgang 2011). Each layer in the ANN structure consists of a non-linear algebraic function which is referred to as a neuron. While the input layer neurons are connected to the neurons of the hidden layer through "channels" that are assigned numerical values known as "weight", the hidden layer neurons are associated with numerical values known as the "bias" and are added to the input sum. The resulting value is passed through a threshold function known as the activation function and the results determine whether a particular hidden layer neuron will get activated or not. Activated neurons transmit data to the neurons of the next hidden layer over the "channels", a process known as feed forward propagation (e.g. Hrnjica and Bonacci 2019; Yuan et al. 2003). Consequently, in the output layer, the neuron with the highest probability determines the output. The predicted outputs are then compared to the actual outputs, and an error can be deduced and transferred back to the network, which "weights" can be adjusted according to the derived errors, a process which is known as back propagation (e.g. Law 2000).

2.4 Long Short-Term Memory

Recurrent neural networks (RNN) are often used in time series analysis; however, they have their own limitations such as failing to process longer sequences and the vanishing gradient. The LSTM is a type of RNN that was developed to overcome the problems a basic RNN would encounter, and thus are used in this study. In general, the LSTM is a time series forecasting technique which uses the best fitting model to forecast the future observations using patterns and trends of previous and current data (Manaswi et al. 2018). When choosing a suitable model for time series prediction, it is always crucial to understand time-series data components such as a) seasonality – which will note the repeating patterns of cycles of behaviour over time, b) cyclicity – identifying repetitive changes in the time series and explain their positioning in the cycle, c) trends – which is frequently observed in a linear model and shows the decreasing and increasing behaviour of the time series, and d) anomalies – to detect observations deviating from the time series model (Chatfield 2000).

3 Mine Water Quality Data Set

The data set used in this study was generated through the monitoring and managing of MIW, with nearly daily sampling being carried out during a period of six years (2016–2021). Parameters that were measured during this period are rainfall, temperature, EC, TDS, total suspended solids

(TSS), acidity, pH, Ca, Mg, SO_4, Al, Fe and Mn and they were used in the units applied by the lab. All these parameters had some of the measurements missing, implying that not all of them have the same number of observations. Fe, turbidity and pH had the greatest number of observations, while rainfall and Mg had the very least of observations (Table 1). Therefore, the data had to be "cleaned" before it was used for training and testing ML models and for forecasting analysis.

4 Data Cleaning and Exploratory Data Analysis

4.1 Data Distribution

Graphical methods, applied in this study, show how data are distributed (Online Resource 5) and also help in visualising the spread, i.e. dispersion, variability and scatter (Online Resource 6). Additionally, dependent variables and possible relationships between the parameters were visualised using the correlation matrix. Most of the parameters display statistically highly significant correlations ($p < 0.001$) with each other (Fig. 2). Furthermore, the Kolmogorov–Smirnov and Shapiro–Wilk tests were conducted to test the normality of the data (Table 2). In most cases, the Shapiro–Wilk test works better on small sample sizes ($n < 50$), while the Kolmogorov–Smirnov test is used on larger sample sizes ($n \geq 50$). Therefore, the Kolmogorov–Smirnov test was relied on to conclude on the normality of the data. From the constructed graphs and tests conducted, it shows that not all the data were normally distributed, and all parameters have statistically significant outliers.

4.2 Outlier Visualisation, Detection and Removal

In statistical terms, an outlier is an observation that is different from the other observations (Wackerly et al. 2014), which means, it is not part of the population. This can come from random or systematic errors such as a mistake during data collection, equipment malfunctioning or just an indication of variance in data collected. Outliers can be identified using visual or statistical methods. In this study, the box plot visualisation method was used, which

Table 1 Mine water quality data supplied; n: number of measurements, $\bar{x}$: average, σ: standard deviation, min: minimum observation, max: maximum observation. pH average calculated as $-\log_{10}[\sum C_i/n]$, where C is the proton activity (www.wolkersdofer.info/pH_en); measured values and units as reported by the plant

Parameter	n	$\bar{x}$	σ	Min	Max
Rainfall, mm	59	13.7	10.6	1.0	40.0
Temperature, °C	1387	23.0	2.0	17.0	29.0
EC, mS cm^{-1}	1387	3.0	0.1	2.44	3.22
TDS, mg L^{-1}	1381	2678.0	185.0	2014	3195
pH, —	1396	6.5	0.2	5.12	7.30
Turbidity, NTU	1398	54.2	59.9	0.54	450
TSS, mg L^{-1}	1386	74.0	75.0	0.00	378
Acidity, mg L^{-1} as $CaCO_3$	898	310.0	41.0	166	442
Ca, mg L^{-1} as CaCO	598	399.0	25.0	267	560
Mg, mg L^{-1} as CaCO	591	101.0	35.0	19	389
SO_4, mg L^{-1}	1396	1474.0	204.0	658	1988
Al, mg L^{-1}	692	0.1	0.0	0.00	0.19
Fe, mg L^{-1}	1402	97.7	15.6	0.10	179.5
Mn, mg L^{-1}	1384	6.6	3.4	0.00	22.9

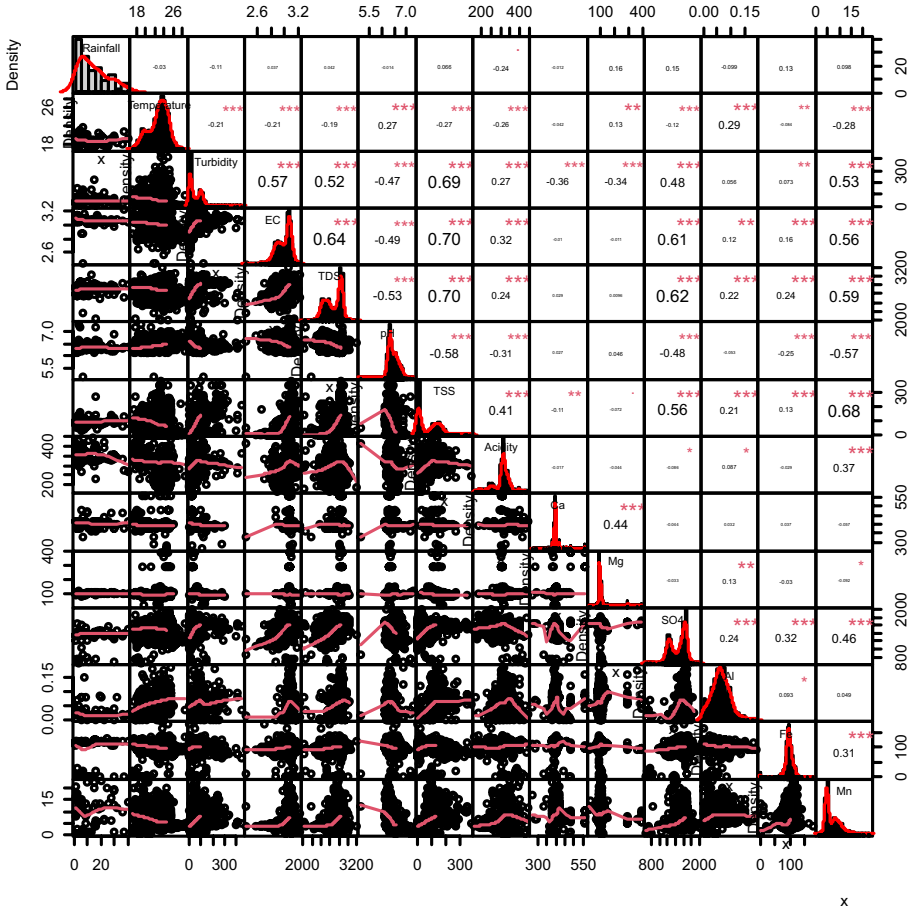

Fig. 2 Eastrand AMD treatment plant data correlation and distribution chart showing histograms on the diagonal, bivariate scatter plots with fitted line on the bottom of the diagonal, and the top of the diagonal shows values of the correlation with significance levels p shown as asterisk: $p < 0.001^{***}$, $p < 0.01^{**}$, $p < 0.05^{*}$

graphically shows the groups of numerical data using their quartiles. Outliers are plotted as points while other data are displayed within boxes (Fig. 3). Furthermore, Z-score analysis was applied as a statistical method on the data for outlier detection and removal. The further away an observation's Z-score is from zero, the more chances of it being an anomaly. Standard cut-off values in determining outliers are Z-scores of ± 3 or even further away from zero.

4.3 "Clean" Data

Data cleaning is the process of removing data that can potentially affect the performance of the models in a negative way, and also using the correct statistical methods to interpolate the missing data. During this process, various linear regression methods with single and multiple independent variables were applied to attempt to fill-in the missing values, but it was not possible due to

Table 2 Normality data test using Kolmogorov–Smirnov and Shapiro–Wilk methods; *n*: number of observations, *p*-value: probability value

		Kolmogorov–Smirnov		Shapiro–Wilk	
Parameter	*n*	Test Statistic	*p*-value	Test Statistic	*p*-value
Rainfall	59	0.143	0.004	0.912	0.000
Temperature	1387	0.089	0.000	0.970	0.000
Turbidity	1387	0.185	0.000	0.803	0.000
EC	1381	0.152	0.000	0.940	0.000
TDS	1396	0.176	0.000	0.918	0.000
pH	1398	0.111	0.000	0.947	0.000
TSS	1386	0.221	0.000	0.848	0.000
Acidity	898	0.197	0.000	0.912	0.000
Ca	598	0.341	0.000	0.469	0.000
Mg	591	0.319	0.000	0.338	0.000
SO_4	1396	0.141	0.000	0.917	0.000
Al	692	0.043	0.004	0.986	0.000
Fe	1402	0.139	0.000	0.788	0.000
Mn	1384	0.195	0.000	0.858	0.000

the number of observations for parameters not matching. Therefore, this problem was solved by a robust, fast and simple method, getting the average of the three observations before and after the missing observation. After the data cleaning process, only nine parameters showed statistically significant relationships with each other. However, within these nine parameters, acidity, TSS and

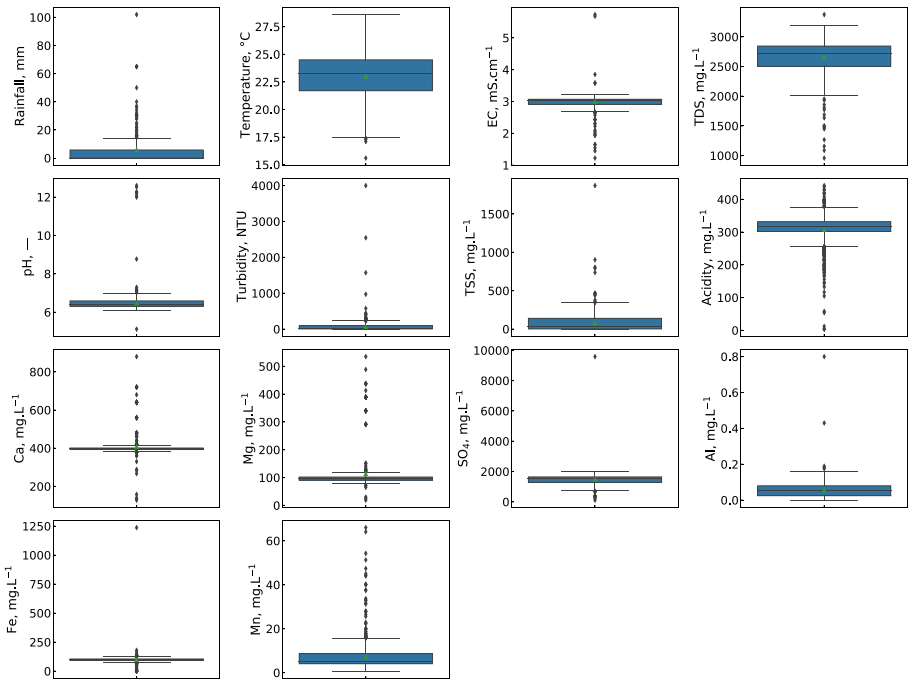

Fig. 3 Box plots with outliers for the Eastrand Acid Mine Drainage treatment plant data from 2016 to 2021

Mn were discarded. Acidity was dropped due to a lower number of observations, i.e. 898 out of 1381 when compared to other observations, while TSS and Mn contains a lot of observations with 0 mg L^{-1}, which makes them not effective for model training and testing. Therefore, parameters used for testing and training the ML models are turbidity, TDS, SO_4, Fe, EC and pH (Fig. 4).

5 Results and Discussion

5.1 Performance Results for Machine Learning Algorithms

In the random forest regression algorithm, a grid search optimisation method with 10, 50, 100, 150 and 200 decision trees was used and resulted in 10 trees giving the best results. Hyper-parameter tuning for gradient boosting regression tree algorithm resulted in 100 trees, 2 maximum tree depths and a learning rate of 0.05. The ANN structure was also

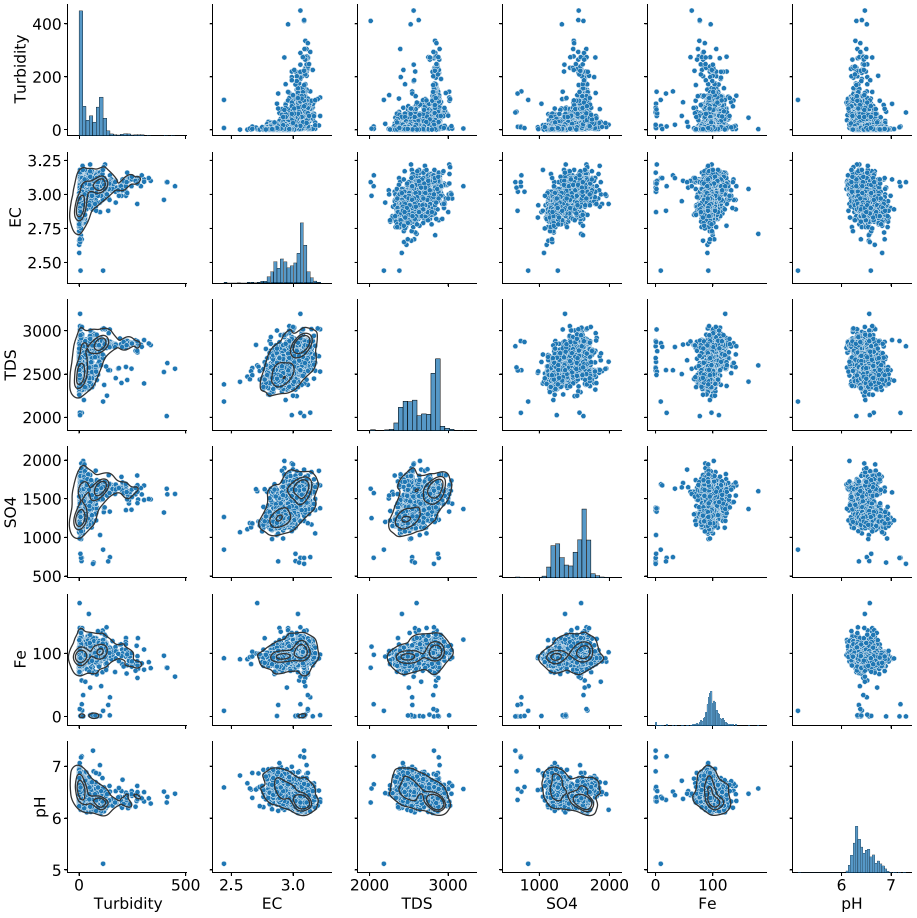

Fig. 4 Relationship between the parameters used to train the machine learning models with histograms shown on the diagonal, bivariate scatter plots with density lines on the bottom of the diagonal, and cross plots on the upper part of the diagonal

Table 3 Performance on testing and training data set for random forest, gradient boosting and neural network models

Algorithm	Testing			Training		
	MAE	RMSE	Quality	MAE	RMSE	Quality
Random Forest	0.074	0.109	Good	0.032	0.049	Good
Gradient Boosting	0.070	0.100	Good	0.069	0.100	Good
Artificial Neural Network	3.822	4.114	Bad	3.815	4.110	Bad

determined using the optimisation techniques and yielded a model with two hidden layers of six neurons and a rectified linear (ReLU) activation function for each, a sigmoid activation function for the output layer, stochastic gradient descent optimiser with learning rate and momentum of 0.001 and 0.8 respectively, batch size of 16, and 50 epochs. In all the three algorithms, the test size was set to be 20%.

The models were for predicting continuous data, and the evaluation metrics used are the mean absolute error (MAE) and root mean squared error (RMSE) for both testing and training data sets (Table 3). ANN is the only model with MAE and RMSE values greater than 1 (Online Resource 7), making it the only bad performing algorithm on the supplied data. Further hyper-parameter tuning to improve ANN algorithm performance was not conducted; therefore, this study relied on random forest and gradient boosting regression tree algorithms to perform predictive analysis.

The LSTM model was used to forecast the concentrations of the input parameters (turbidity, TDS, SO_4 and Fe) (Fig. 5). In this algorithm, the number of past days used to predict the future was set to be 100 days. A single hidden layer LSTM having 32 memory

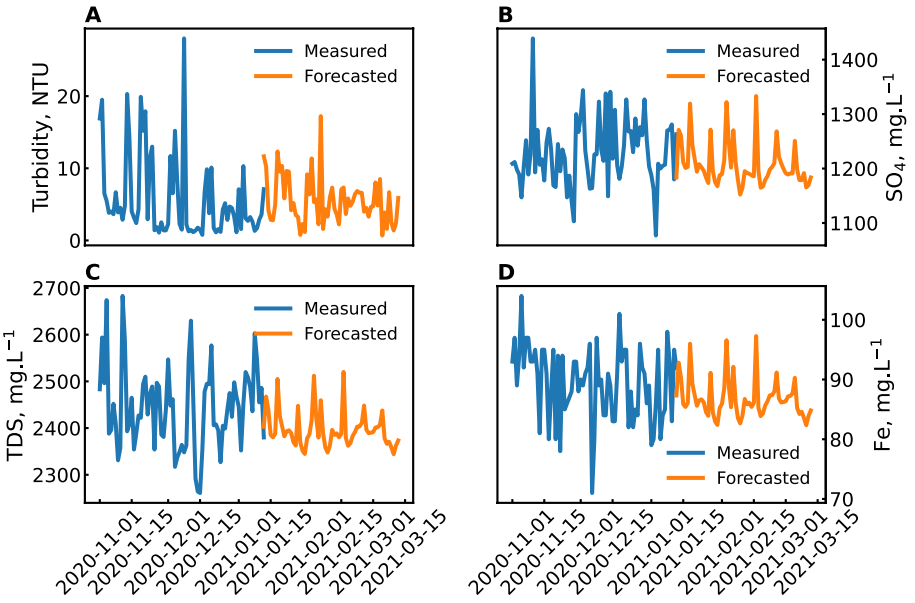

Fig. 5 Forecasted values for turbidity, SO_4, TDS and Fe for 60 days using the LSTM model. Visualisation started on 1 November 2020 due to a small forecast of 60 days and a large historical data of 1381 observations (fitting the whole 1381 observations would make it difficult to visualise the results); **A** Turbidity forecast, **B** SO_4 forecast, **C** TDS forecast, **D** Fe forecast

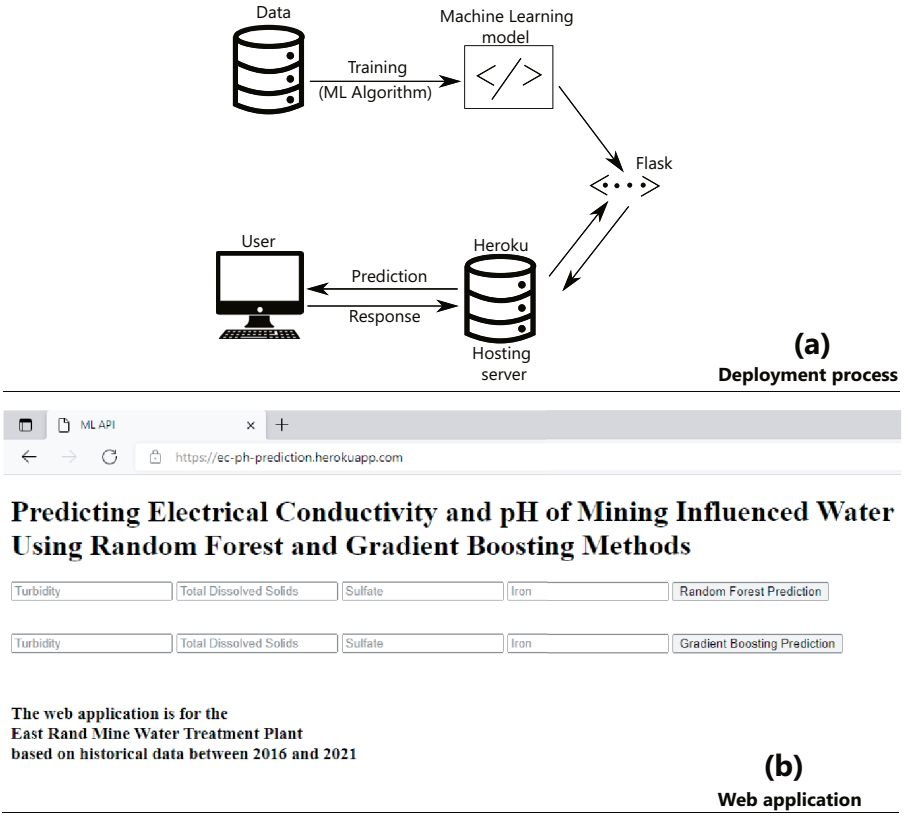

Fig. 6 Deployment process of machine learning models (**a**), and web application screenshot (**b**) used to predict EC and pH of mining influenced water for the Eastrand AMD Water treatment plant, Springs, South Africa (https://ec-ph-prediction.herokuapp.com)

units with a ReLU activation function was used. The model was fitted over 50 epochs with a batch size of 32 and validation split of 0.2. It was further compiled using the adaptive moment estimation (Adam) optimiser and mean squared error (MSE) loss, with the lower MSE value of 0.0035 implying that the model performed good (Online Resource 8).

5.2 Model Deployment and Web Application Results

The main goal of developing ML models is to solve a problem, and a ML model can only achieve that when it is deployed and used in an AI system. In this study, an application programming interface (API) was created to deploy ML models using the Flask and Heroku platforms (Fig. 6a). Flask is a web framework for Python, i.e. it provides functionality for building web applications. Heroku is a cloud platform on which applications can be deployed, managed and scaled. Steps involved in this deployment process include: a) training the model, b) creating a web application using Flask, c) committing programming code in GitHub, d) linking GitHub to the Heroku cloud platform, e) and deploying the model as

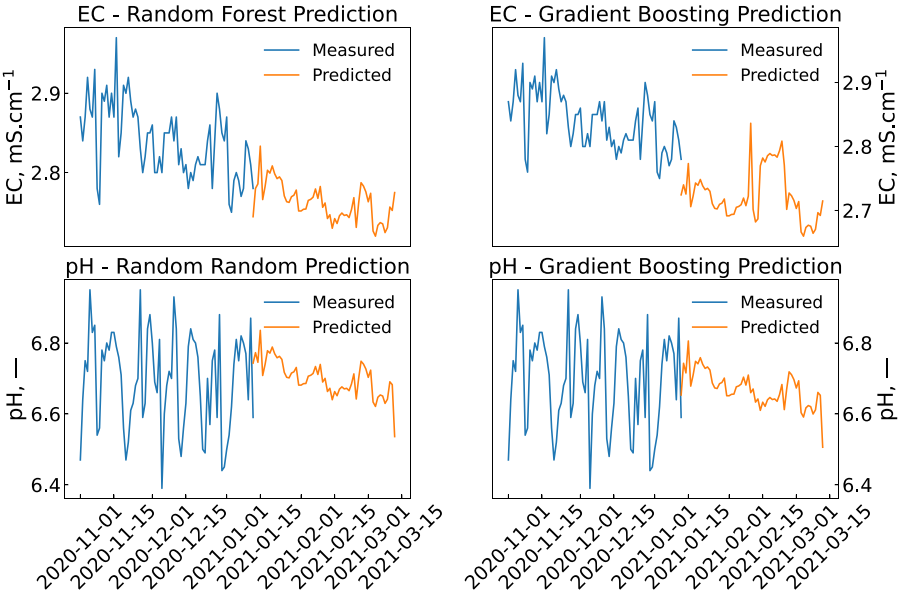

Fig. 7 EC and pH prediction visualisation using the web application

a web application (Fig. 6b). Finally, the web application was used to predict (forecast) EC and pH values for 60 days (Fig. 7).

6 Conclusions

Predicted values of EC and pH with random forest and gradient boosting algorithms were not much different from the measured ones, giving much needed confidence and reliability in the prediction of mine water chemistry. Thus, the presented approach in this study is a scientifically important contribution in knowing the future quality of mine water. In addition, the ANN model did not produce the best results; however, it can still be considered in other instances with extended data sets.

Incorporating ML models in the mines can improve the current treatment practices, give a good mine water balance, and increase the rate at which MIW is treated. Predicting the chemistry of MIW will ensure that the treatment plant operator knows in advance the quantity and type of chemicals and methods to use to treat and manage mine water. In mine water treatment, neutralising agents such as lime or quick lime are added to increase the pH for precipitating Fe and other metals. For AMD treatment plants it is therefore important to know the development of the water parameters beforehand in order to stock for chemicals or to control the mine water level in a pumped system. This research proved that computational-intelligence techniques are much more effective than traditional dynamic modelling approaches based on GoldSim, Geochemists Workbench or PHREEQC. Therefore, the proposed approach can be an efficient tool and useful alternative for forecasting and predicting mine water quality parameters. Future research will focus on forecasting additional parameters such as Fe or SO_4 from other plants.

7 Remarks

It is crucial to understand that an ML model is only as good as the input data it receives. Researchers often publish articles without accurate data collection and not having performed thorough exploratory data analysis before deciding on suitable input parameters for the outputs to be predicted using ML models. In this study, data cleaning had to be performed before deciding on the input parameters to use. Some models performed better with good accuracies, others did not, and the forecasting trend did not take the entire shape of the historical data, but the results proved that ML models are the future of mine water quality predictions and forecast. This implies that if the data were collected accurately from start to finish, without equipment malfunctioning or missing observations, model performance would have been substantially better.

A mistake commonly done is to use accuracy to evaluate regression models (e.g. Khandelwal and Singh 2005; Maier et al. 2004; Rooki et al. 2011). In ML, if the task is to predict a numerical value, then regression models can be used, and if the problem is dealing with discrete label as a result, then classification models can be applied. ML models need to be evaluated before they can be put in production. Researchers find it easy to use accuracy to evaluate ML model performance. However, accuracy only works when an observation is similar to a prediction; it focuses on whether the prediction is correct or not. In regression models, many predictions are not similar to the observations, and sometimes have low errors. Therefore, evaluation metrics recommended for regression models are RMSE and MAE, as they can indicate how spread out the prediction error is, and they can also identify the amount of errors in measurements.

The main aim to build ML models is to deploy them and make practical business decisions. In this study, a free cloud platform, Heroku, was used to show how a ML model can be put in production. Several scientific articles only show how good the algorithms can perform, and never educate the readers on the deployment of ML models. Platforms such as Microsoft Azure, Amazon Web Services (AWS) Lambda, Google Cloud, and Algorithmia are also useful tools to deploy ML models of mine water management applications.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11269-022-03177-2.

Acknowledgements The authors thank the South African National Research Foundation (NRF Grant UID 138770 and 86948) under the SARChI Chair for Mine Water Management at Tshwane University of Technology (TUT) for funding of this study, and Divan van Niekerk and the Department of Water and Sanitation, South Africa for providing us with historical mine water data. We thank the anonymous reviewers, whose comments helped to develop this study into its current version.

Author Contribution Conceptualization: K.S. More; Funding acquisition: Ch. Wolkersdorfer; Software: K.S. More; Supervision: Ch. Wolkersdorfer; Writing – original draft: K.S. More; Writing – review and editing: Ch. Wolkersdorfer.

Funding This work is funded and supported by the National Research Foundation (NRF Grant UID 138770 and 86948) South Africa under the SARChI Chair for Mine Water Management, Tshwane University of Technology (TUT).

Availability of Data and Code The data and codes that support the findings of this study are available from the corresponding author upon reasonable request.

Declarations

Ethics Approval Not applicable.

Consent to Participate All authors consent to participate.

Consent for Publication All authors consent to publish.

Competing Interests The authors declare that they have no competing interests.

References

Belgiu M, Drăguţ L (2016) Random forest in remote sensing: A review of applications and future directions. ISPRS J Photogramm Remote Sens 114:24–31. https://doi.org/10.1016/j.isprsjprs.2016.01.011

Blowes DW, Ptacek CJ, Jambor JL, Weisener CG, Paktunc D, Gould WD, Johnson DB (2014) The Geochemistry of Acid Mine Drainage. In: Holland HD, Turekian KK (eds) Treatise on Geochemistry. 2nd edn. Elsevier, Oxford, p 131–190. https://doi.org/10.1016/B978-0-08-095975-7.00905-0

Boulesteix AL, Janitza S, Kruppa J, König IR (2012) Overview of random forest methodology and practical guidance with emphasis on computational biology and bioinformatics. Wires Data Mining Knowl Discov 2(6):493–507. https://doi.org/10.1002/widm.1072

Breiman L (2001) Random Forests. Mach Learn 45(1):5–32. https://doi.org/10.1023/A:1010933404324

Chatfield C (2000) Time-series forecasting. Chapman and Hall/CRC, London Manaswi NK (2018) Deep Learning with Applications Using Python: Chatbots and Face, Object, and Speech Recognition With TensorFlow and Keras. Apress, Berkeley, CA. https://doi.org/10.1007/978-1-4842-3516-4

Chen K, Chen H, Zhou C, Huang Y, Qi X, Shen R, Liu F, Zuo M, Zou X, Wang J, Zhang Y, Chen D, Chen X, Deng Y, Ren H (2020) Comparative analysis of surface water quality prediction performance and identification of key water parameters using different machine learning models based on big data. Water Res 171:115454. https://doi.org/10.1016/j.watres.2019.115454

Hem JD (1985) Study and interpretation of the chemical characteristics of natural water. vol 2254. Department of the Interior, US Geological Survey

Hrnjica B, Bonacci O (2019) Lake level prediction using feed forward and recurrent neural networks. Water Resour Manage 33(7):2471–2484. https://doi.org/10.1007/s11269-019-02255-2

Hubert E, Wolkersdorfer C (2015) Establishing a conversion factor between electrical conductivity and total dissolved solids in South African mine waters. Water SA 41(4):490–500. https://doi.org/10.4314/wsa.v41i4.08

Johnson NE, Ianiuk O, Cazap D, Liu L, Starobin D, Dobler G, Ghandehari M (2017) Patterns of waste generation: A gradient boosting model for short-term waste prediction in New York City. Waste Manage 62:3–11. https://doi.org/10.1016/j.wasman.2017.01.037

Khandelwal M, Singh T (2005) Prediction of mine water quality by physical parameters. J Sci Ind Res 64:564–570

Law R (2000) Back-propagation learning in improving the accuracy of neural network-based tourism demand forecasting. Tour Manag 21(4):331–340. https://doi.org/10.1016/S0261-5177(99)00067-9

Liu P, Wang J, Sangaiah AK, Xie Y, Yin X (2019) Analysis and prediction of water quality using LSTM deep neural networks in IoT environment. Sustainability 11(7):2058. https://doi.org/10.3390/su11072058

Maier HR, Morgan N, Chow CWK (2004) Use of artificial neural networks for predicting optimal alum doses and treated water quality parameters. Environ Modell Softw 19(5):485–494. https://doi.org/10.1016/S1364-8152(03)00163-4

Manaswi NK, Manaswi NK, John S (2018) Deep learning with applications using python. Springer, New York

More KS, Wolkersdorfer C, Kang N, Elmaghraby AE (2020) Automated measurement systems in mine water management and mine workings — A review of potential methods. Water Resour Ind 24:1–12. https://doi.org/10.1016/j.wri.2020.100136

Mosavi A, Sajedi Hosseini F, Choubin B, Goodarzi M, Dineva AA, Rafiei Sardooi E (2021) Ensemble Boosting and Bagging Based Machine Learning Models for Groundwater Potential Prediction. Water Resour Manage 35(1):23–37. https://doi.org/10.1007/s11269-020-02704-3

Nalecki P, Gowan M (2008) Mine Water Management-Dynamic, Probabilistic Modelling Approach. In: Rapantova N (ed) 10th International Mine Water Association Congress, Karlsbad, 2008. Czech Republic, Karlsbad, pp 533–536

Rooki R, Doulati Ardejani F, Aryafar A, Bani Asadi A (2011) Prediction of heavy metals in acid mine drainage using artificial neural network from the Shur River of the Sarcheshmeh porphyry copper mine. Southeast Iran Environ Earth Sci 64(5):1303–1316. https://doi.org/10.1007/s12665-011-0948-5

Russell S, Norvig P (2002) Artificial intelligence: a modern approach. Prentice Hall, New Jersey

Singer PC, Stumm W (1970) Acidic mine drainage: the rate-determining step. Science 167(3921):1121–1123. https://doi.org/10.1126/science.167.3921.1121

Swamynathan M (2017) Mastering Machine Learning with Python in Six Steps: A Practical Implementation Guide to Predictive Data Analytics Using Python. Apress, Berkeley, CA. https://doi.org/10.1007/978-1-4842-2866-1

Usher B, Strand R, Strachotta C, Jackson J (2010) Linking fundamental geochemistry and empirical observations for water quality predictions using Goldsim. In: Wolkersdorfer C, Freund A (eds) Mine Water and Innovative Thinking—International Mine Water Association Congress, Sydney, Canada, 2010. Canada, Sydney, pp 313–317

Vadapalli VRK, Sakala E, Dube G, Coetzee H (2020) Mine Water Treatment and the Use of Artificial Intelligence in Acid Mine Drainage Prediction. In: Fosso-Kankeu E, Wolkersdorfer C, Burgess J (eds) Recovery of Byproducts from Acid Mine Drainage Treatment. Scrivener, p 23–48. https://doi.org/10.1002/9781119620204.ch2

Verburg R (2011) Mitigating Acid Rock Drainage. In: Darling P (ed) SME Mining Engineering Handbook. vol 1, 3rd edn. Society for Mining Metallurgy and Exploration, Littleton, p 1721-1732

Wackerly D, Mendenhall W, Scheaffer RL (2014) Mathematical Statistics with Applications, 7th edn. Thomson, Belmont

Wolfgang E (2011) Introduction to artificial intelligence. Undergraduate Topics in Computer Science. Springer, Cham. https://doi.org/10.1007/978-3-319-58487-4

Wolkersdorfer C (2008) Water Management at Abandoned Flooded Underground Mines – Fundamentals, Tracer Tests, Modelling, Water Treatment. 1st edn. Springer, Heidelberg. https://doi.org/10.1007/978-3-540-77331-3

Yuan H, Xiong F, Huai X (2003) A method for estimating the number of hidden neurons in feed-forward neural networks based on information entropy. Comput Electron Agr 40(1–3):57–64. https://doi.org/10.1016/S0168-1699(03)00011-5

Zhang Y, Haghani A (2015) A gradient boosting method to improve travel time prediction. Transp Res Part C Emerg Technol 58:308–324. https://doi.org/10.1016/j.trc.2015.02.019