APPLICATION OF ARTIFICIAL INTELLIGENCE IN ACID MINE DRAINAGE PREDICTION

Samiksha Sharma - Roll no 18MI31015

Kavya Venkatagiri - Roll no18MI33019

Gunti Jyothsna Sri - Roll no18MI33007

Arindam Bera - Roll no 18MI33003

Mini-Project report submitted in partial fulfillment of the requirements for the Surface Mining Technology (MIN****)***

Department of Mining Engineering

Indian Institute of Technology, Kharagpur

March 2021

TABLE OF CONTENTS **Page**

ACKNOWLEDGMENT

We would like to express our special thanks of gratitude to our project instructor **Prof Ramesh Murlidhar Bhatawdekar Sir** as well as our course instructor **Prof Khanindra Patha sir** who gave us the golden opportunity to do this wonderful project on the **APPLICATION OF ARTIFICIAL INTELLIGENCE IN ACID MINE DRAINAGE PREDICTION** , which also helped us in doing a lot of research and we came to know about so many new things And we are thankful to them.
Secondly, we would also like to thank ourselves as a team that helped us a lot in finalizing this project within the limited time frame.

ABSTRACT

This study will investigate the use of artificial neural networks(ANNs) for mine water quality prediction by implementing an AI-based model.
The coal mining in Meghalaya causes large-scale destruction and degradation of the environment. The main problems in Meghalaya are the production of Acid Mine Drainage (AMD) in nearby areas by continuous leaching of acidic waste from the coal mining sectors.
So we'll try to develop the prediction system by using parameters such as mine water (it's pH, sulphate, or heavy metals like arsenic) concentration, Rainfall, air temperature, depth to the groundwater table, and also if possible the groundwater geological aquifer parameters(like what type of rock is in between the mine surface and groundwater) will be used as inputs parameter whereas sulfate concentration as output.
A graphical user interface (GUI) will be developed by combining long short-term memory nets (LSTM) for each of the four input parameters and an ANN combining the LSTM outputs to predict future sulfate values.

CHAPTER 01- INTRODUCTION

For this mini-project report the introduction should describe the following in different sections:-

1.1 Background

Mining water-related pollution is one of the major environmental challenges in our country. Due to substantially changed the landscape in coal mining. The pollution of the groundwater is associated with acid mine drainage (AMD). Many interventions have been proposed and implemented, adapted to the volumes and quality of the AMD.

The application of artificial neural networks (ANNs) to the fields of water engineering, ecological and environmental sciences has gained momentum in the last two decades.

Chen and Mynett (2003) and Lee et al. (2003) have successfully used data-driven modeling techniques for the prediction of freshwater and seawater quality alike. Also, Najah et al. (2011) have demonstrated that ANNs can be used to predict river water quality, based on historical data.

Therefore, ANNs may be an alternative to current methods of mine-water quality prediction. ANN captures the embedded spatial and unsteady behavior inherent in the AMD-affected area. . The architecture and non-linear nature of ANNs make them more suitable than other modeling techniques. The current study intends to investigate the application of ANNs to predict the water quality with a focus on sulfate concentrations. The current project, promoting the application of historical data to predict the future may serve as a prototype for future projects.

1.2 Problem definition

AMD is recognized as one of the most serious environmental problems in the mining industry. The problem of acid mine drainage (AMD) has been present since mining activity began thousands of years ago. Mining activity has disrupted the hydrology of mining areas so badly that it is extremely difficult to predict where water would eventually re-emerge. Its causes, treatment have become the focus of several researchers. Through this project, we discuss various factors, review the techniques and models currently used to predict mine drainage quality, and at the end, we will build an AI-based model using LSTM to predict future sulphate concentration.

1.3 Project aims & objectives

This study will investigate the use of artificial neural networks(ANNs) for mine water quality prediction by implementing an AI-based model where PH of mine water, Rainfall, water table, and temperature will be input, and Sulphate concentration will be our output.

1.4 Report outline

The following list is a broad summary of development phases and associated tasks for our project. The order is general and many tasks run concurrently.

Ch 2. Acid Mine Drainage

In this chapter we will discuss some basic theory about AMD, Kind of factors which controls AMD and their effect, major source of AMD, AMD prediction tools, etc.

Ch 3. MODEL

Here we will get to know about the AI model we will build to predict sulphate concentration, like the methodology of the model, will see Accuracy and results, specifications of it, etc.

Ch 4. Conclusion and recommendation

We will conclude here about the achievements and limitations of our work, and further recommendation for this.

CHAPTER 02 - ACID MINE DRAINAGE

2.1 Acid Mine Drainage Generation

Acid mine drainage is the formation and movement of highly acidic water rich in heavy metals.Acid mine drainage (AMD) is an inevitable byproduct of the mining and mineral industry which is generally characterized by a high concentration of dissolved heavy metals, sulfate, and low pH as low as 2 and continues to be an important water pollution problem in the mining industry around the world. Although the generation of AMD occurs naturally, mining and processing of metal ores and coals can promote AMD generation through exposing sulphide minerals to both oxygen and water. AMD pollutes the receiving streams and aquifers when it is allowed to discharge without any treatment. Contaminants from mine drainage can persist for a long time after mining has stopped acidic water forms through the chemical reaction of surface water and shallow subsurface water with rocks that contain sulfur-bearing minerals, resulting in sulfuric acid. In several mining sites around the world, the presence of oxidizing sulfides in waste rock and the tailings are the cause of the most important potential environmental problem. Oxidation products can include hydrogen sulfide, partially oxidized oxyanions, such as thiosulphate and polythionates, iron sulfate in solution, elemental sulfur, various jarosite compounds, sulfuric acid, and heavy metals. Radionuclides are recognized as being of concern when uranium ore is miIn several mining sites around the world, the presence of oxidizing sulfides in waste rock and the tailings are the cause of the most important potential environmental problem. Oxidation products can include hydrogen sulfide, partially oxidized oxyanions, such as thiosulphate and polythionates, iron sulfate in solution, elemental sulfur, various jarosite compounds, sulfuric acid, and heavy metals. Radionuclides are recognized as being of concern when uranium ore is mined. Several works have reported the occurrence of acid mine drainage in the uranium mining sites. Two of the most famous mining sites to present this problem are the Rum Jungle. the presence of radioisotopes in acid mine drainage may also be of relevance in non-uranium mining sites, especially when uranium occurs in higher concentrations than the average concentration generally found in the crust. Any acid formed from pyrite oxidation will, to some extent, react with other gangue minerals within the solid sample.

The reaction to be considered :

FeS_2+ $7/O_2$ + H_2O <--> $FeSO_4$ + H_2SO_4 (equation 1)

The composition of the resulting solution will be dependent . Some examples are:

• $CaCO_3 + H_2SO_4 \rightarrow CaSO_4 + H_2O + CO_2$ (calcite dissolution)

• $CaMg(CO_3) + 2H_2SO_4 \rightarrow CaSO_4 + MgSO_4 + 2\ H_2O + 2CO_2$ (dolomite dissolution)

• $KaI[AlSi_3O_{10}](OH)_2 + H^+ + 3/2\ H_2O \rightarrow K^+ + Al_2Si_2O_5(OH)_4$ (s) (muscovite dissolution)

• $KaISi_3O_8$(s) + H^+ + $9/2\ H_2O$ -» K^+ + $2\ H_4SiO_4$ + $1/2\ Al_2Si_2O_5(OH)_4$ (s) (K-feldspar dissolution)

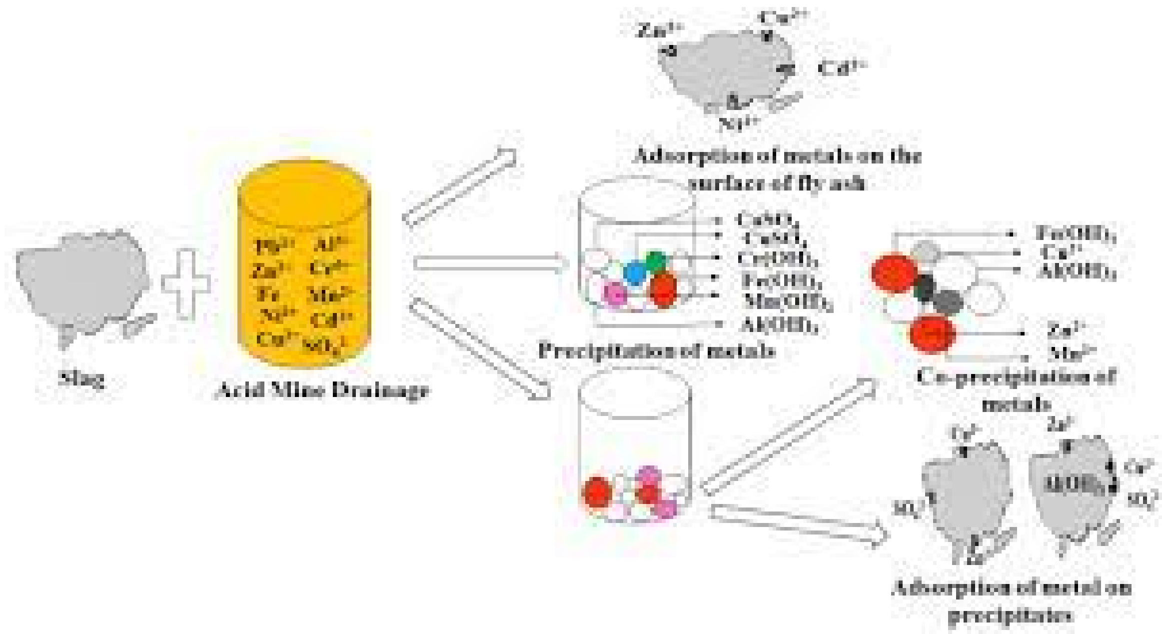


Figure 2.1

.

2.2 Factors Controlling AMD

The factors controlling AMD formation may be categorized as primary, secondary, and tertiary.

The primary factors are those directly involved in the acid production process. The primary factors include pyrite and other sulfide minerals, oxygen, water, ferric iron, and iron-oxidizing bacteria which play key roles in the acid production reactions.

Secondary factors control the consumption or alteration of the products from the acid generation reactions, The most important secondary factor is the neutralization of an acid by alkalinity released from carbonate minerals in the mine waste, such as calcite ($CaC03$) and dolomite ($CaMg(C03)2$)

Low pH-high sulfate water is the classic AMD composition.

while tertiary factors are the physical aspects of the waste materials or mine site that influence acid production, migration, and consumption.

- -The three types of factors control the AMD reactions and the quality of water emanating from the mine waste. Other physical and chemical processes may affect the quality of the drainage shortly after it leaves the acid-producing material,

This oxidized pyrite, generating AMD as well as solid weathering products in the form of acidic sulfate-bearing minerals. On the cessation of underground operations, as the underground mine workings flood, these minerals are dissolved, resulting in the first flush of acidic water. The flooding of the workings reduces the availability of oxygen, substantially reducing the oxidation of pyrite and the generation of new acidity. Younger describes the two types of acidity generated within a mine void as juvenile acidity—the acid generated by the oxidation of fresh sulfide minerals—and vestigial acidity—the acidity generated by the dissolution of secondary minerals which accumulated during active mining. Mine flooding largely excludes oxygen from the flooded workings, reducing the generation of new juvenile acidity, often leading to substantial improvements in water quality once the first flush, which liberates vestigial acidity from secondary minerals, has dissipated. Where mine workings remain air-filled after mine flooding, juvenile acidity may still be generated for many years following partial mine flooding

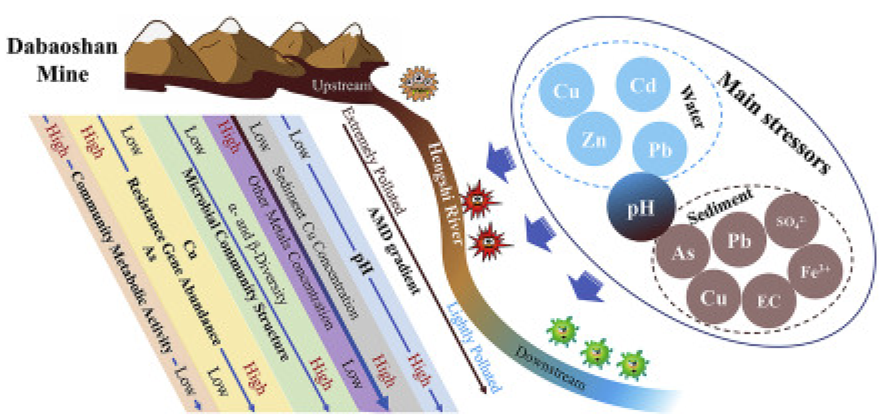


Figure 2.2

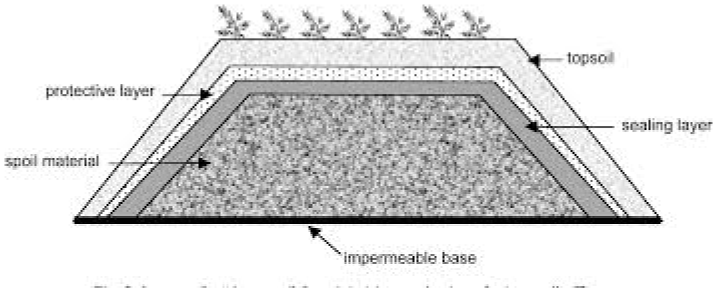
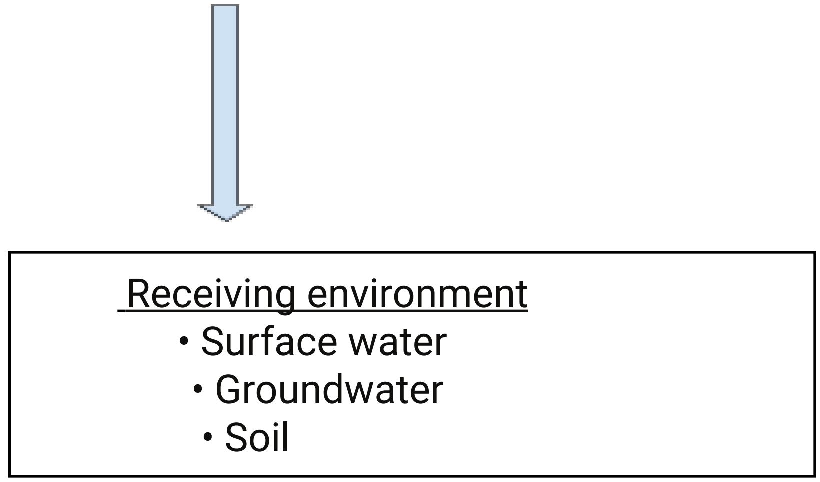


Figure 2.3

Sulfide Sources	Water Sources
• Tailings piles • Waste rock piles • Ore stockpiles • Heap leach materials • Pit walls • Underground workings • Sulfides in unmined ore Host rock	• Infiltration of rainwater • Groundwater seepage • Surface-water ingress

Mixing zones	Discharge mechanism
• Unsaturated zones in tailings/waste rock/ore piles • Open pits above pit lakes • Infiltration zones in underground mines	• Water pumped from the mine workings • Gravity-driven discharge from workings • Seepage



so, sulfide minerals are playing a major role in the formation of AMD.

2.3 Prediction of AMD

The prediction of acid mine drainage is needed to determine before mining if the quality of waters draining a mine site will exceed regulatory standards. If the oxidation processes become established, mine water contaminant levels can exceed these standards by many orders of magnitude. Thus, accurate prediction of AMD Pre-Mine Prediction of Acid Mine Drainage is required both to protect the environment and to ensure that resources are expended wisely to prevent or control AMD. From the previous discussion of the many primary, secondary, and tertiary factors that control the acid mine drainage process, the prediction of mine water quality can be an impossible task. However, only a limited number of these factors may control the acid mine drainage process in most practical situations; therefore, models used for AMD prediction may be relatively simple.

Several predictive techniques have been developed based upon simplified AMD production models and are routinely used to estimate the quality of drainage before mining. Most of the methods provide only a qualitative prediction of acid mine drainage formation. A few models have also been developed to quantitatively estimate the composition of mine drainage and these are most frequently used to predict the success of measures to control AMD formation.

Methods to predict acid mine drainage can be divided into five groups:

l. Geographical mining comparisons;

2. Paleoenvironmental and geological models;

3. Geochemical static tests;

4. Geochemical kinetic tests;
5. Mathematical models.
6. Artificial Intelligence for AMD Quality prediction

Artificial Intelligence for AMD Quality prediction

The benefits of using AI for data processing and automation can potentially supplement the human mind that has limited ability for processing data and susceptible to subjective bias, AI models can process huge amounts of data relatively faster with greater accuracy. therefore, predictive models can help reduce computational cost while increasing computational speed, reliability, and consistency in addition to reducing any human-induced

uncertainties

Different AI Techniques Used to Predict AMD Quality

Common AI tools used for predictive modeling of water quality can be classified into three broad types, namely, the knowledge-driven, the data-driven, and the hybrid type. The most common example of knowledge-driven techniques

includes fuzzy inference systems, which are used where the relationship between the input parameters and desired predicted variables are well understood.

Data-driven AI systems are computer programs that solve problems using information derived from data without explicit knowledge of the problem. Modern-day data-driven AI approaches are mainly machine learning in nature. According to the literature, the widely used data-driven tools are regression analysis, artificial neural networks (ANNs), random forest (RF), decision trees, support vector machines (SVMs), radial base functions (RBFs), and genetic algorithms.

-- the application of AI for the prediction of mine drainage is not new. Several authors used machine learning techniques such as ANN, SVM, RBF, and K-mean to model and forecast AMD quality using past physicochemical parameters of a mining area. AI tools such as machine learning are a rapid and cost-effective solution for mine water quality forecasting.

Limitations of AI Techniques in Prediction of AMD Quality

AI techniques have their limitations which include:
a) Time-consuming processing and training of the model as optimization is heuristically and repeatedly performed until a good model or learning technique is obtained.
b) Most experts find it difficult to interpret and explain decisions made by AI models, suggesting its black box nature and limiting its successful application.
For any prediction, including mine drainage quality, a thorough investigation and understanding of the fundamental physical processes are compulsory. However, it is challenging to involve domain expert knowledge directly into the data-driven approaches
Hence, hybrid approaches are developed to mitigate this problem.

Mine Water Treatment and Use of AI in AMD Prediction

Prediction of mine drainage is very complex as the conditions which produce the drainage qualities are normally nonlinear, dynamic, and change over time. In addition to this, acquiring Good quality mine drainage data is difficult and contains missing and erroneous values. approaches mentioned in the previous section or fail to learn the time-variant
parameters leading to low accuracy and poor generalization, hence specially designed algorithms that can learn time series data are required. A recurrent neural network (RNN) is used for time series data learning but is mostly not successful when learning the long-term dependencies, hence the long short-term memory (LSTM) is a specially designed RNN to overcome this problem.

CHAPTER 03 - MODEL

3.1 Limitations

Challenges with any AI model -

1. Data Collection

Data plays a key role in any use case. 60% of the work of a data scientist lies in collecting the data.

2. Variance in the dataset

Whether the dataset represents ground truth situations or not. The data we use for training should cover all the cases that occurred and that is going to occur.

3. Trust Deficit

One of the most important factors that are a cause of worry for the AI is the unknown nature of how deep learning models predict the output.

4. Challenges in obtaining good accuracy

For a deep learning model to perform at a good accuracy, it would require unprecedented finetuning, hyperparameter optimization, large dataset, and a well-defined and accurate algorithm, along with robust computing power, uninterrupted training on train data and testing on test data. That sounds a lot of work, and it's actually a hundred times more difficult than it sounds.

5. Data Scarcity

The good or bad nature of an AI system really depends on the amount of data they are trained on. Hence, the ability to gain good data is the solution to good AI systems in the future. But, in reality, the everyday data the organizations collect is poor and holds no significance of its own. Some companies are trying to innovate new methodologies and are focused on creating AI models that can give accurate results despite the scarcity of data.

6. Overfitting the data

Overfitting refers to a model that models the training data too well. **Overfitting** happens when a model learns the detail and noise in the training data to the extent that it negatively impacts the performance of the model on new data.

7. Underfitting the data

Underfitting refers to a model that **can** neither model the training data nor generalize to new data. An underfit machine learning model **is** not a suitable model and **will** be obvious as it **will** have poor performance on the training data.

Main options to reduce underfitting are:

- Feature Engineering — feeding better features to the learning algorithm.
- Removing noise from the data.
- Increasing parameters and selecting a powerful model.

3.2 Procedure

We target the first challenge to our problem - Data! To perform our machine learning or Artificial Intelligence techniques, firstly we need a dataset. We tried to collect mine site data containing information about the chemical composition and physical characteristics of mine drainage water.

As due to various company policies, confidentiality reasons as well as sensitivity of data, we couldn't obtain actual/real data of any particular mine site, in India or abroad. That's why we have used the following graphs from the research paper (IMWA 2019 "Mine Water: Technological and Ecological Challenges") to create our data table.

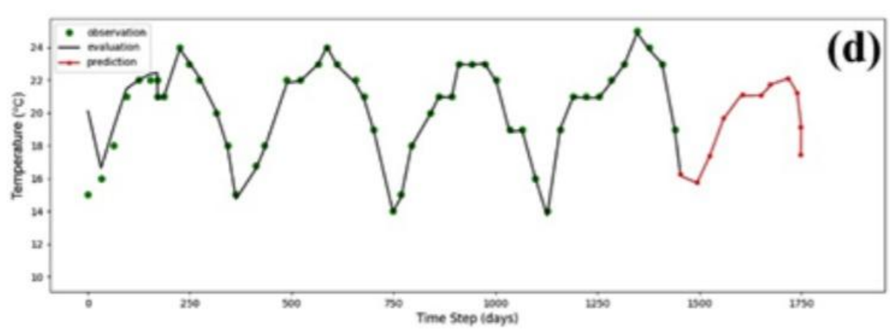
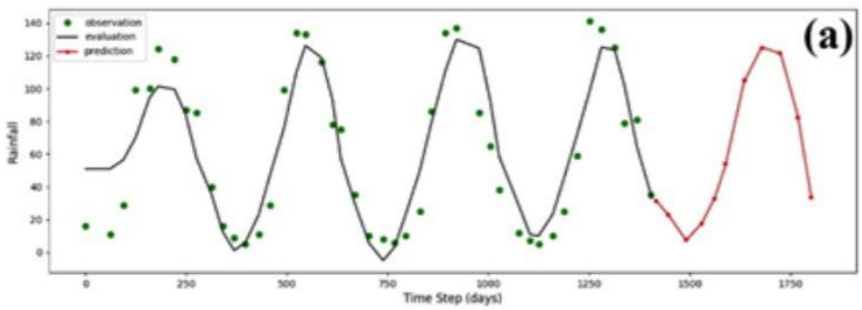
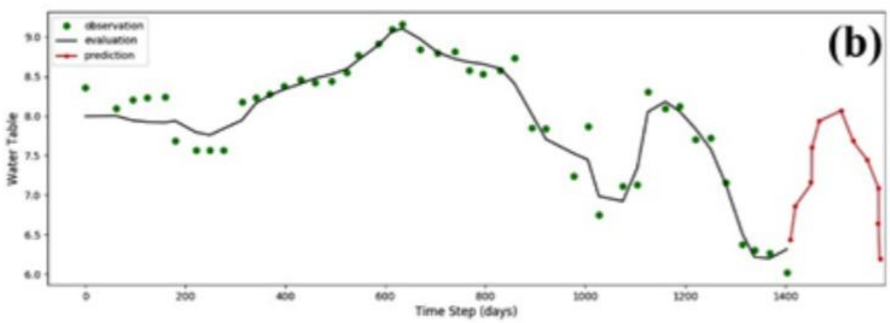
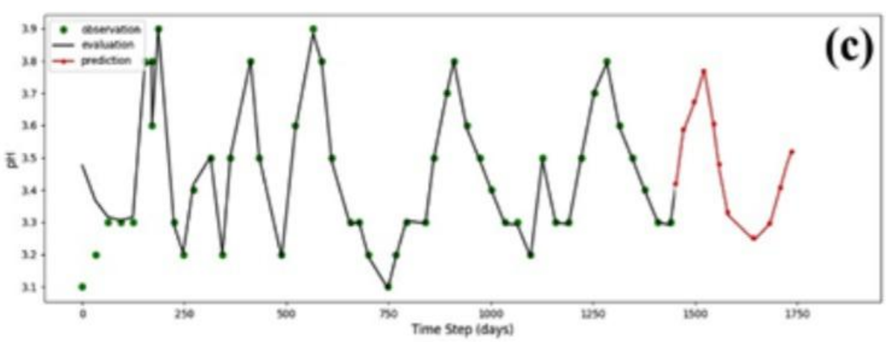


Figure 3 Testing the input predictions for a) rainfall, b) water table, c) water pH, d) air temperature and input parameters for the Witkranz discharge site.







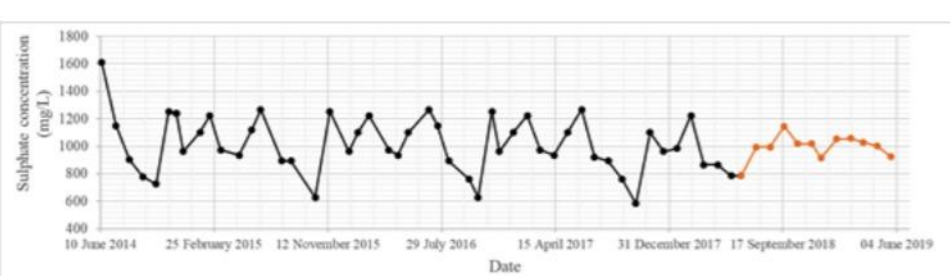
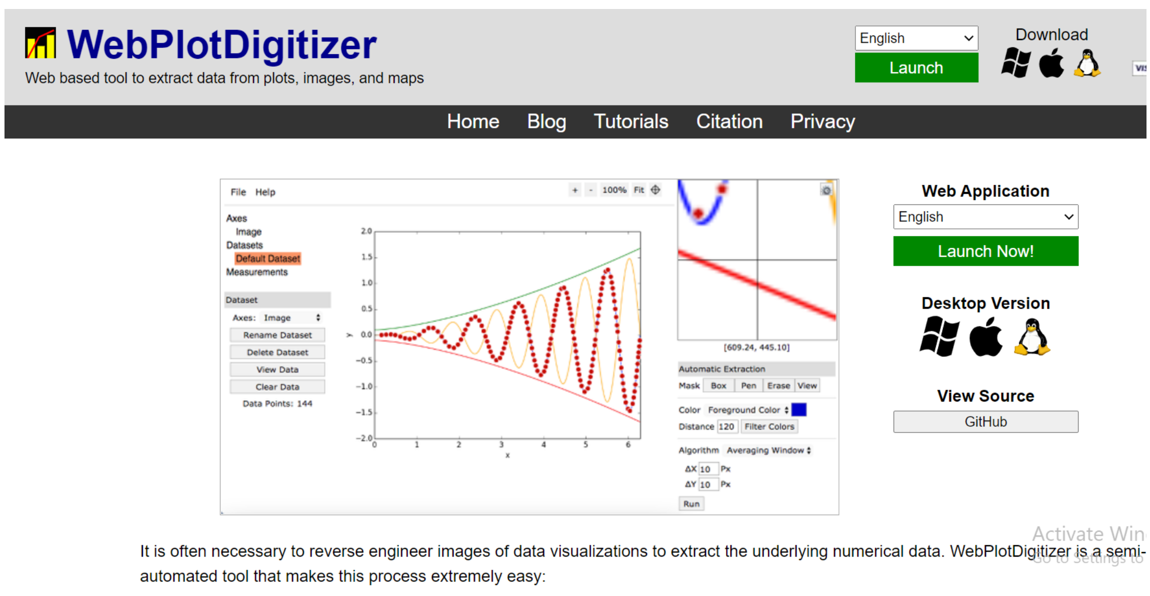


Figure 6 Prediction values for sulfate using the developed prototype.

1. Web Plot Digitizer

   This web based tool is a semi automated tool that makes the process of obtaining numeric data logs much convenient for us. It is often necessary to reverse engineer images of data visualizations to extract the underlying numerical data. Web plot digitizer does exactly this job for us. It can be used for a wide variety of charts (XY, bar, polar, ternary, maps etc.). We may opt to manually mark the points on respective graphs or plots and obtain the corresponding coordinates/values. We may also use Automatic extraction algorithms that make it even easy to extract a large number of data points. It is free to use and an open source web tool.



   This tool played a key role in our mini project to create a dataset for mine site drainage water constitution.

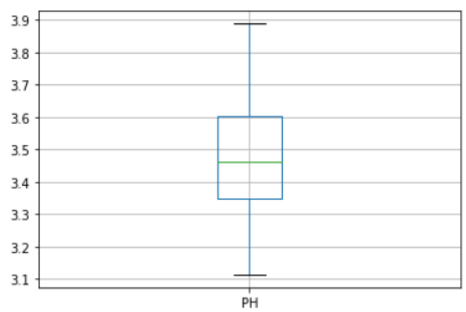
   We uploaded the above mentioned graphs obtained from the research paper to Web plot digitizer and obtained data tables for each of these graphs. Then using python data manipulation and analysis libraries - pandas and numpy we created our final merged table containing data that we shall use next to feed into our ML model. Let's name this final data table as **"Mine Site Drainage Data", abbreviated as "MSDD"** and we shall be using this same abbreviation all along in the sections to follow.

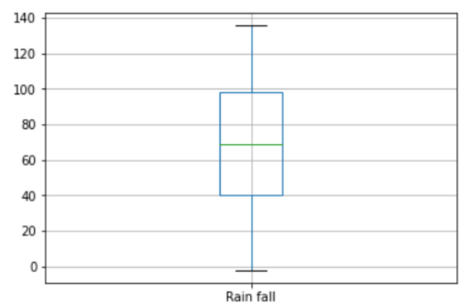
2. In MSDD, we had the following Input Parameters -

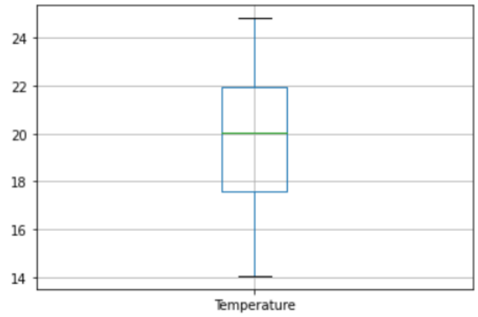
I. Days
II. pH
III. Rainfall
IV. Temperature
V. Water Table
VI. Sulphate

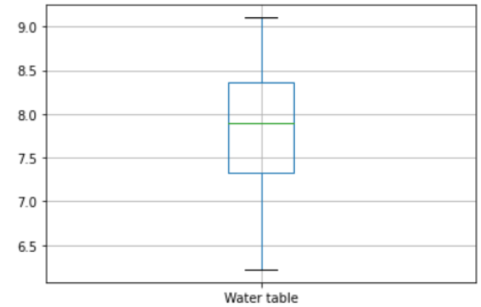
3. For each row in MSDD, the values for above data columns were available to us.

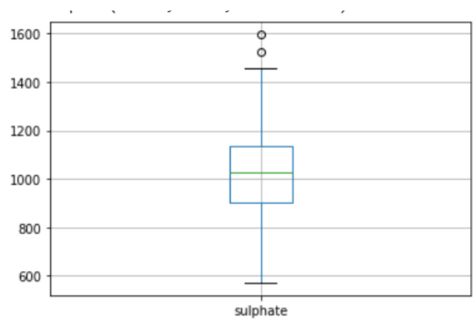
- Our dataset has 161 rows and 6 columns.
- The missing values were replaced with mode(maximum occurring value) for that particular column.
- Then we splitted the dataset as 80% training set and rest being test set.
- Using a pearson correlation heatmap, we tried to find correlation among the data columns. We performed this step to measure the strength and direction of a linear relationship between two variables.
- We did the following visualizations on the dataset -











- As we can see, pH levels show acidic levels. pH should be in 5.5 to 9 range but it is much lower than that, indicative of acidic levels in mine drainage water.
- We can add more columns as input parameters if we have its data.
- After the data visualization is over, we have run ML models on our training dataset and obtained Root Mean Squared Error value on our test dataset.
- RMSE is our evaluation metric for determining performance benchmark.
- RMSE is a frequently used measure of the differences between values predicted by a model or an estimator and the values observed.
- We compare models by the rmse obtained on the test set.
- Tried to reduce the rmse by tuning hyperparameters.

3.2 Results

Out of all the models we used, polynomial support vector machines gave us the least root mean square error.

3.3 Specifications

“Support Vector Machine” (SVM) is a supervised machine learning algorithm which can be used for both classification or regression challenges.

The important parameters having higher impact on model performance are - “kernel”, “gamma” and “C”.

Kernels can be “linear”, “rbf”,”poly”, etc. Polynomial kernel turned out to work best on our dataset.

Gamma is the kernel coefficient for ‘rbf’, ‘poly’ and ‘sigmoid’. Higher the value of gamma, will try to exactly fit the as per training data set i.e. generalization error and cause overfitting problem. We tried on gamma values, eg - 0,10,100 etc.. and found 10 working best.

C is the Penalty parameter C of the error term. It also controls the trade-off between smooth decision boundaries and classifying the training points correctly.

Chapter 4. Conclusion and recommendation

4.1 Conclusion

A system doesn’t perform well if the training set is too small, or if the data is not generalized, noisy, and corrupted with irrelevant features. We went through some of the basic challenges faced by beginners while practicing machine learning.

4.2 Recommendation for further work

Next we will work on finding more data and implementing deep learning models such as LSTM or ANN.

We would be glad to know your feedback on our work and suggestions for further improvement.

References

- Akcil, Ata, and Soner Koldas. "Acid Mine Drainage (AMD): causes, treatment and case studies." *Journal of cleaner production* 14, no. 12-13 (2006): 1139-1145.
- Simate, Geoffrey S., and Sehliselo Ndlovu. "Acid mine drainage: Challenges and opportunities." *Journal of Environmental Chemical Engineering* 2, no. 3 (2014): 1785-1803.
- Jaynes, D. B., A. S. Rogowski, and H. B. Pionke. "Acid mine drainage from reclaimed coal strip mines 1. Model description." *Water Resources Research* 20, no. 2 (1984): 233-242.
- Vadapalli, Viswanath Ravi Kumar, Emmanuel Sakala, Gloria Dube, and Henk Coetzee. "Mine Water Treatment and the Use of Artificial Intelligence in Acid Mine Drainage Prediction." *Recovery of Byproducts from Acid Mine Drainage Treatment* (2020): 23-48.
- Sakala, Emmanuel, Obed Novhe, and Viswanath Ravi Kumar Vadapalli. "Application of Artificial Intelligence (AI) to predict mine water quality, a case study in South Africa." In *Proceedings of the Mine Water Association Conference: Technological and Ecological Challenges, International Mine Water Association Annual Conference*, pp. 140-145. 2019.
- Kim, Min-Suk, Hyun-Gi Min, and Jeong-Gyu Kim. "Integrating amendment and liquid fertilizer for aided-phytostabilization and its impacts on soil microbiological properties in arsenic-contaminated soil." *Applied Sciences* 10, no. 11 (2020): 3985.
- Johnson, D. Barrie, and Kevin B. Hallberg. "Acid mine drainage remediation options: a review." *Science of the total environment* 338, no. 1-2 (2005): 3-14.
- Gazea, B., K. Adam, and A. Kontopoulos. "A review of passive systems for the treatment of acid mine drainage." *Minerals engineering* 9, no. 1 (1996): 23-42.