RESEARCH ARTICLE

International Journal of Computer Communication and Informatics

DOI: 10.34256/ijcci2313

# E-Coating Ultrafiltration System Maintenance using Machine Learning Techniques

**D. Anusruti [a,*], P.C.D. Kalaivaani [a], S. Dhanvarsha [a], R. Ashwin [a]**

[a] Department of Computer Science and Engineering, Kongu Engineering College, Erode-638060, Tamil Nadu, India.

[*] Corresponding Author: anusrutid.18cse@kongu.edu

Received: 03-03-2023, Revised: 29-04-2023, Accepted: 07-05-2023, Published: 30-05-2023

**Abstract:** Ultrafiltration process is one of the important processes in e-coating of metal parts. It is important to maintain and improve the performance of ultrafiltration membrane (UF membrane). This UF membrane should not be degraded under any situation, if this happens then it might lead to the flooding of coating fluid all over the metal parts. Hence it has to monitor properly. This will also lead to the wastage of coating fluid, wastage of materials and maintenance cost will also be high. So to avoid this, the workers have to monitor it on periodically basis. During e-coating in electrophoresis painting plant there may occur fault in filter and excess of fluid will flood in the material to be coated. The filter used in the ultrafiltration subsystem has to be monitored manually each time by the worker. Sometimes, if the worker did not monitor properly, it might lead coating to inappropriate places, it will cause wastage of material and fluid and also it takes time for clean-up and to change after fault occurs. Hence, the timely detection of faults is very important to prevent damage. In order to prevent these kinds of situations, a machine learning model is developed which predicts the flow meter readings beforehand. It uses an ensemble learning algorithm known as XGBoost which is one of the more powerful tools for prediction.

**Keywords:** Ultrafiltration, UF membrane, XGBoost, Coating, Membrane

## 1. Introduction

Electro coating or commonly referred to as E-coating is the method of applying paint to any metal surface, inside and out. Ultrafiltration (UF) has a significant role in e-coat systems and nowadays all industrial e-coat lines have a UF system. The advantage of UF system is that it improves the efficiency of electro-paint usage and also reduces its environmental impact as the amount of wastewater sent to the drain is drastically reduced.

During e-coating in an electrophoresis painting plant there may occur fault in the filter and excess of fluid will flood in the material to be coated. The filter used in the ultrafiltration subsystem has to be monitored manually each time by the worker. Sometimes, if the worker did not monitor properly, it might lead coating to inappropriate places, it will cause wastage of material and fluid and also it takes time for clean-up and to change after fault occurs. Hence, the timely detection of faults is very important to prevent damage.

Machine learning is an implementation of artificial intelligence (AI) that allows the system to learn and improve on its own without explicitly being programmed. In order to look into data and make better decisions in the future based, the learning process begins with observations or data, such as examples, or instruction. Using traditional machine learning algorithms, text is treated as a sequence of keywords; instead, a semantic analysis approach mimics the human ability to understand the meaning of a text.

### A. Xg-Boost

XG-Boost is a gradient boosting-based decision-tree-based ensemble Machine Learning technique. Artificial neural networks surpass all other algorithms or frameworks in prediction issues involving unstructured data (pictures, text, etc.). However, decision tree-based algorithms are considered best-in-class for small-to-medium structured/tabular data. The evolution of tree-based algorithms over time can be seen in the graph below. A special case of boosting where errors are minimized by gradient descent algorithm e.g. the strategy consulting firms leverage by using case interviews to weed out less qualified candidates.


**Figure1.** Features of XG-Boost

## 2. Literature Survey

Yumeng Cao *et al.* (2020), "Application of XGBoost Model on Personal Credit Evaluation" IEEE Intelligent Systems, Vol 33, pp.14-33 - This research investigates the use of the XGBoost algorithm to a big data-based credit evaluation problem [1]. We usually apply the

XGB model to the personal loan scenario using open data from the Lending Club Platform in the United States. The results of the experiment show that the XGB model outperforms the logistic regression and the other three tree-based models in terms of feature selection and classification performance.

Zhaoweijie *et al.* (2020), "Forecast Rosman Store Sales Based on XgBoost Model," Second International Conference on Economic Management, This project investigates the sales situation of Rosman's 1115 stores over the last three years, performs feature engineering in the dimensions of space and time, builds a prediction model using the XGBoost gradient promotion algorithm, and forecasts the sales volume of Rosman's 1115 stores over the next 48 days, demonstrating that the model's precision in predicting future sales of 41088 samples was 89.07 percent [2].

Li Jidong *et al.* (2018), "Dynamic Weighing Multi-Factor Stock Selection Strategy Based on XGBoost Algorithm," International Conference of Safety Produce Informatization (IICSPI), pp.566-572 – This research investigates a multi-factor dynamic weighted stock selection technique based on the XGBoost model [3]. Profit, quality, scale, growth, liquidity, valuation, momentum, and reversal are represented by seven factors. The IC coefficients of factors are predicted using the XGBoost machine learning technology. The weights of factors are continually modified based on the expected IC coefficients. The effective style factor will be given more weight, while the invalid factor will be given less.

## 3. System Description

### *3.1. Existing System*

The electro-coating ultrafiltration system consists of an e-coat tank, an UF membrane and the metal parts to be coated. This e-coat process typically involves a number of stages including cleaning stage followed by a phosphate conversion coating stage. This stage improves the metal's corrosion resistance while also providing a better foundation for the future coating. The electro painting procedure can only begin if the metal has been properly prepared. The post-rinse ultrafiltration stage is subsequently followed by the oven curing stage. The e-coat tank consists of the coating materials like resins, pigments and additives that are suspended in water and circulated in a bath.

With the electrode position method, the resin and the metal surface effectively become one. The coated metal part is removed from the bath. The next step is to remove any excess paint and this is known as cream-coat. This is carried out in the flowing post rinses of ultrafiltration. Ultra-filtrate is a by-product of the paint going through a membrane; filtering the paint and creating permeate. This is part of a closed loop system, to keep tank levels at their optimal.

Figure 2. E-Coating Ultrafiltration System

### *3.2. Dataset Description*

| Name | Original Sample Rate | Covered Time Period | Unified Sample Rate |
|---|---|---|---|
| Manual inspection records | Every 8 hours per sample | 7 years | Every 30 minutes per sample (up-sampled) |
| IIOT sensor readings | Every 10 seconds per sample | 15 days | Every 30 minutes per sample (down-sampled) |

- Over 1 lakh 10,000 samples from manual readings.
- Over 700 samples from IOT readings.

### *3.3.* Block Diagram

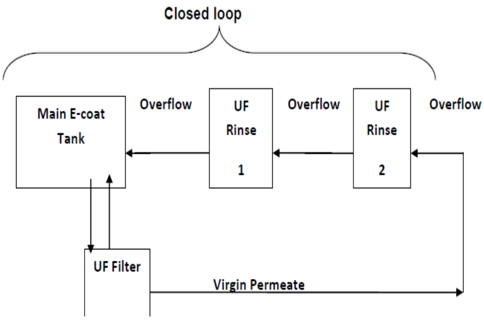


Figure 3. Block Diagram

It can be seen that there is a main e-coat tank which is the very important part of the electro-deposition process, where the metal part gets coated. A circulation system maintains the paint system uniformity by continuously mixing and agitating the e-coat bath ingredients. The rinse stage uses the permeate to remove the undeposited paint particles from metal parts coated in the e-coat bath. As the metal parts move through the rinse sections it is progressively rinsed by clean water and the undeposited paint solids are also carried back through each rinse section. This closed loop process recycles the undeposited paint particles, making it available for future use.

### 3.4. Hardware Description

### Coat Tank

Figure 4. E-Coat Tank

The paint bath in the electro-coating ultrafiltration system is known as the e-coat tank which can be seen in Figure 4. Electricity charges the materials in the e-coat tank to deposit the paint onto the metal parts which has to be coated; the charge is determined by voltage. There are two types of e-coat process tanks that are used in many industries. The first one is the transport system where the substrate load would be submerged into the tank. It is then left for a period of time while an electrical voltage is applied and then it is taken out and passed on to the next process tank. Then the other system is a inline e-coat tank. In this process the substrate will travel along a conveyor and it will submerse the product into the paint tank and removing in a continuous motion. Many industries use electro-coating for a variety of advantages because of its cost efficiency, line productivity and environmental gains. E-coating can paint high volumes of parts covering all the surfaces submerged into the tank both inside and out. The high production efficiency coupled with advanced quality results in lower unit costs, making it one of the most popular industrial processes.

### Materials Used In E-Coat Tank

- All coatings, including electro-coatings, are made from
- Polymeric resin or binder
- Pigments
- Deionized water

- Anode or cathode

The resin covers 20-30% of the tank and has qualities like corrosion resistance and ultraviolet durability. Color, gloss, and corrosion resistance are all provided by pigments. De-ionized water makes up the majority of an electro-coat bath, accounting for 80-90 percent of the total volume. Paint solids, which include resins, pigments, and a small amount of solvents, are carried by de-ionized water. The solvent ensures a smooth appearance and application of the film. Electro-coat products are either anodic or cathodic, depending on where the coating is applied. The notion of "opposite attract" is used in this process.

The mental components are positively charged in anodic electro-coating, and the negatively charged paint particles gets deposited on the metal substrates. Small amounts of metal ions migrate into the paint film during the anodic process, and these ions become trapped in the depositing paint film due to their tendency to interact with moisture, limiting the system's performance attributes. Anodic coatings are low-cost coatings that provide great colour and gloss control. The major application is for things that will be used in an interior or moderately outside environment.

The mental components are negatively charged in cathodic electro-coating, and the positively charged paint particles gets deposited on the metal substrates. This method incorporates less ions into the depositing film and improves corrosion resistance. Cathodic coatings are high-performance coatings that are resistant to corrosion.

## Uf Membrane

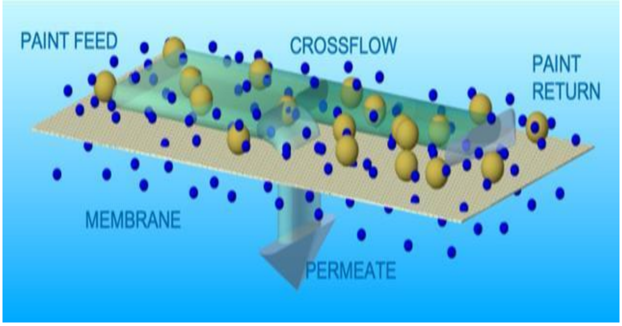


Figure 5. UF Membrane

The UF membrane which works on a cross-flow principle. The e-coat paint transfers through filter and passes across the top layer. E-coat paint is returned to the paint bath and permeate is diverted. A gel layer is formed on the polymer face as the paint flows across it. Water can pass through small pores but large components in the paint are pulled towards these pores and are unable to pass through it. At this point the large components those are not able to pass through lose their velocity and find it hard to move. In a sense they become “double parked” on the face of the polymer. If there are just a couple of double parkers, the water can pass through.

When the gel layer gets thick and viscous, the paint is restricted to pass through and permeate production is much, much lower.

## Operational Concern

Permeate is a molecular material separated from paint solids because it can pass through UF membrane's small pores. Low permeate flow is the most common complaint received from UF end users. A 7640- type UF element might have 17lpm (~4.5 gpm) of permeate when it is used for the first time. Most end users consider permeate flow below 4 lpm (~1 gpm) to be too low and it is an indication that it is time to install a new UF element.

Figure 6. Large opening in the UF membrane of the UF element.

Figure 7. Leaking UF membrane

## Low Paint Flow

Low paint flow generates less turbulence on the face of the membrane, which makes the gel layer grow. Low paint flow results in generating less active force that is needed to push the paint through the small pores of the UF membrane. The low paint flow is an important cause in many cases where the UF permeate is low or the membrane experiences short life.

## 3.5. Proposed System

### *XG-Boost Architecture*

The results of the regression problems are continuous or real values. Some commonly used regression algorithms are Linear Regression and Decision Trees. There are several metrics involved in regression like root-mean-squared error (RMSE) and mean squared-error (MSE) and Mean-Absolute-Error (MAE). These are some key members of XGBoost models, each plays an important role.

### Validating the model

The process of evaluating a trained model with a testing data set is known as model validation. The testing data set is a subset of the same data set that the training data set is derived from. Model validation is used to assess the model's accuracy and performance using historical data for which we already have actuals. Evaluation metrics considered: The suggested system's performance was assessed using the following criteria:

- SCORE
- MSE (Mean Square Error)
- RMSE (Root Mean Square Error)
- MAE (Mean Absolute Error)

### Score

Score is performed by implementing linear regression algorithm on the random sample of data. The best possible score is 1.0 and it can be negative. Score calculating based on below formula.

SCORE = (1–(u/v))

u = sum for i to N ((yi_true - yi_pred)** 2)

v = sum for i to N ((yi_true - yi_true.mean()) ** 2)

### Mean Square Error

The MSE is calculated by averaging the square of the difference between the data's original and anticipated values. It is useful when there are outliers or unexpected umbers in the dataset.

MSE (Mean Square Error) = 1 / N * sum for i to N (yi_true – yi_pred) ^2

N = Total number of samples

yi_true = ith sample's expected / actual value

yi_pred = ith sample's predicted value

## Root Mean Square Error

The standard deviation of the errors that occur when making a prediction on a dataset is known as the RMSE. This is the same as MSE, but the root of the number is taken into account when calculating the model's accuracy. The errors are squared before being averaged in RMSE. This basically means that RMSE gives greater errors more weight.

RMSE (Root Mean Square Error) = sqrt(1 / N * sum for i to N (yi_true – yi_pred)^2)

N = Total number of samples

yi_true = ith sample's expected / actual value

yi_pred = ith sample's predicted value

## Mean Absolute Error

The output of MAE is the average of this error over all samples in a dataset.

MAE (Mean Absolute Error) = 1 / N * sum for i to N abs (yi_true – yi_pred)

N = Total number of samples

yi_true = ith sample's expected / actual value

yi_pred = ith sample's predicted value

# 4. Results and Discussion

The implementation is carried out in Python language - Tensorflow. The platform chosen is Jupyter since Jupyter notebook facilitates writing and executing Python in the browser with zero configuration required. Following are the results of simulating several machine learning algorithms, such as Random Forest and Extreme Gradient Boosting.

FM1 is Flowmeter reading which is the output variable and all other Pressure sensor and temperature sensor are output variables It is visible that FM1 is not related all the input variables but proportional to some input variables and those input variables are proportional to other input variable.


Figure 9. Heap map of Manual reading dataset

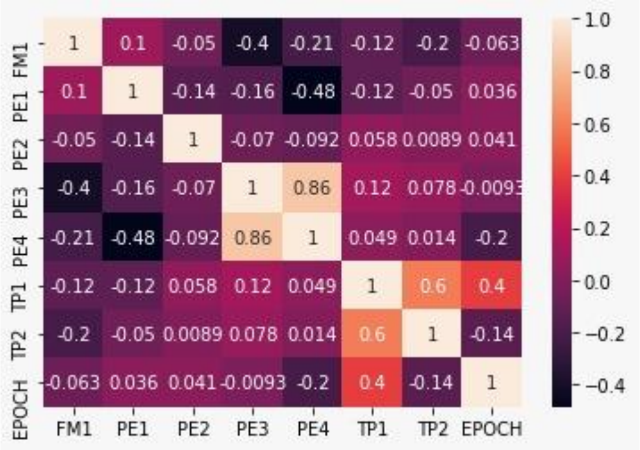


Figure 10. Heap map of IIOT dataset

Here 1 indicates positive correlation which means the two variables moved either up or down in the same direction together and -1 indicates negative correlation that means the two variables moved in opposite directions and a correlation of zero means there is no relationship between the two variables. In other words, as one variable moves one way, the other moved in another unrelated direction. The correlation in IIOT dataset more than the manual reading dataset. The correlation in IIOT dataset is more than the manual reading dataset. Hence a better model can be developed using IIOT dataset than the manual reading dataset.

It is seen that there is more relation between the variables than the Manual reading dataset.

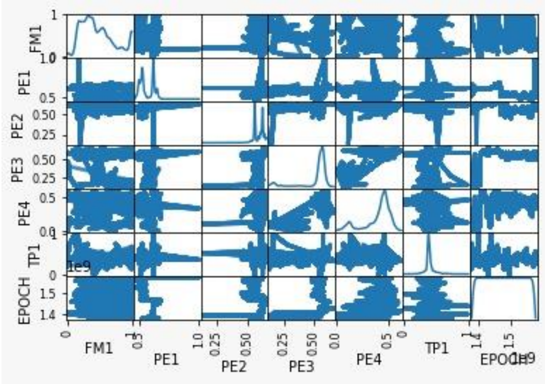


Figure 11. Scatter plot of Manual reading dataset

## Comparison of Random Forest and XGBoost

The Extreme Gradient Boost algorithm gives better accuracy in both the sensor and manual data and IIOT sensor data compared with Random forest regressor. Hence the accuracy obtained in the Extreme Gradient Boost algorithm is comparatively higher than the Random Forest regressor algorithm.

| Dataset | Accuracy | Mae | Mse | Rmse |
|---|---|---|---|---|
| Random Forest Sensor Data | 0.8815291954328363 | 0.0202983669088931 | 0.001381068940185221 | 0.037162735908235024 |
| XGBoost Sensor Data | 0.9370645104493626 | 0.01721935933857335 | 0.0007829260281190069 | 0.027980815358366648 |
| Random Forest Manual Data | 0.802555773887376 | 0.07238017800707368 | 0.013158037217939879 | 0.11470848799430615 |
| XGBoostManual Data | 0.8525591133913786 | 0.07162209195683705 | 0.009825725024421236 | 0.09912479520493969 |

The model obtained with IIOT dataset has more accuracy than the Manual reading dataset. It shows that there may be lot of errors while taking the readings manually and IIOT sensors give more accurate information than the manual readings.

## Comparision Graphs for Sensor and Manual Dataset

Figure 13 shows the comparison between the Random Forest and XGBoost algorithm accuracy and mean square value, mean absolute value and root mean square error values for the manual dataset values and the results are show that XGBoost gives the better accuracy and less error values.

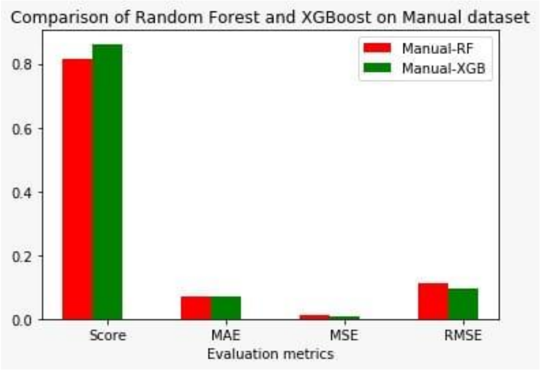


Figure 13. Comparison of Random Forest and XGBoost on Manual Dataset

Figure 14 shows the comparison between the Random Forest and XGBoost algorithm accuracy and mean square value, mean absolute value and root mean square error values for the sensor dataset values and the results are as like manual dataset results XGBoost gives the better accuracy and less error values.

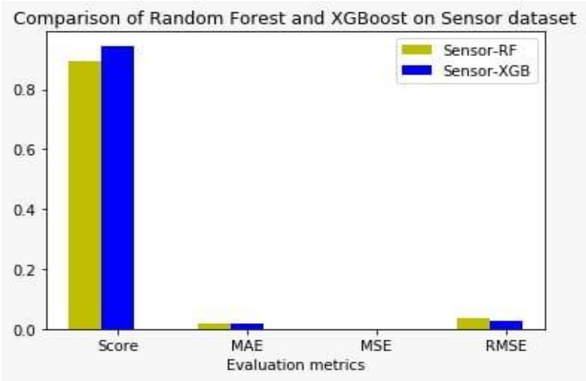


Figure 14. Comparison of Random Forest and XGBoost on Sensor Dataset

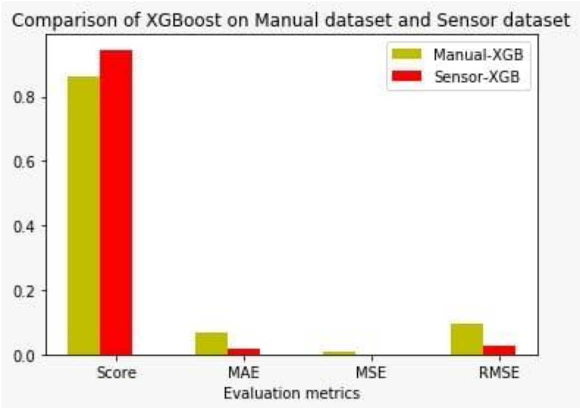


Figure 15. Comparison of Manual dataset and Sensor dataset

Figure 15 shows the Comparison of manual dataset and sensor dataset values. Which gives the results of sensor dataset values give more accuracy and less error value compared with manual dataset values. The sensors values give more accurate values.

Figure 16 shows the Running times for manual and sensor dataset values. The running time is measured in microseconds. Compared with sensor dataset values the manual dataset have more number of data values so the manual dataset takes more running time

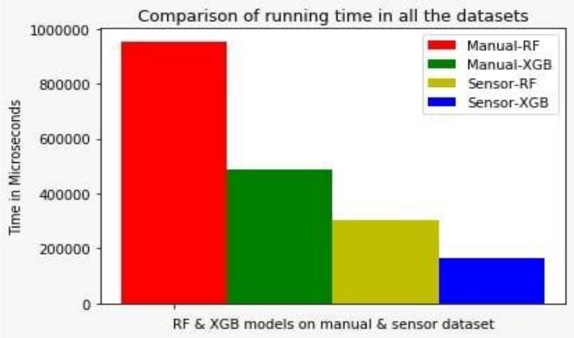


Figure16. Comparison of running time in all the datasets

## 5. Conclusion and Future Works

The machine learning model for predicting the flow meter measurements is developed using one of the best ML algorithms XGBoost. The accuracy and the overall performance of the model is high when compared to other algorithms making it ideal for prediction. Hence, by using this model, the ultrafiltration membrane of the electrophoresis painting plant can be maintained with low cost and with high efficiency. This model developed has to be tested against the challenges in real world application. Exploring the use of this model not alone in prediction but also in other fields. Explore and find new algorithms better than the existing model.

## References

[1] H. Li, Y. Cao, S. Li, J. Zhao, Y. Sun, XGBoost model and its application to personal credit evaluation, *IEEE Intelligent Systems*, 35(3), (2020) 52-61. https://doi.org/10.1109/MIS.2020.2972533

[2] Zhaoweijie, Chenliang, Hujiangmin, (2020) Forecast Rosman Store Sales Based on Xgboost Algorithm, *Second International Conference on Economic Management and Model Engineering*, 521-525. https://doi.ieeecomputersociety.org/10.1109/ICEMME51517.2020.00110

[3] Li Jidong, Zhang Ran, Dynamic Weighing Multi-Factor Stock Selection Based on XGboost Algorithm, *IEEE International Conference on Safety Produce Informatization (IICSPI)*, (2018) 868-872. https://doi.org/10.1109/IICSPI.2018.8690416

## Funding

No funding was received for conducting this study.

## Conflict of interest

The Authors have no conflicts of interest to declare that they are relevant to the content of this article.

## About The License