A Framework for Total Quality Management of Diesel Generator Fuel Consumption using Machine Learning and Internet of Things (IoT)

Ali A. Majeed Ali
Associate Professor of Production and Industrial Engineering, Faculty o Engineering, Aden University

Osama Abdulhak M. Nasher
Research Scholar at dept. of Production and Industrial Engineering, Faculty o Engineering, Aden University

Ahmed Sultan Al-Hegami
Professor of Intelligent Information Systems,
University of Sana'a, Yemen

ABSTRACT

Decision making on quantity of fuel consumption requirement is playing a very important role in industrial applications, for establishment of the production process which became more complicated and essential especially in Arab countries which have major shortage of fuel availability and price fluctuation and subsequently, Decision Making becomes very hard. Over the years, most of decisions were generated, based on personal experience which may not be effective due to many parameters such as level of experience of decision making and the state of the production system.

Developing the ability to predict fuel consumption of Diesel Generator (DG) is extremely beneficial for improvement of generator performance, reducing operation and maintenance cost and avoiding fuel misuse; however, fuel consumption is measured by the amount of fuel used during a specific time period. In this paper, we propose a framework that makes use of IIOT technology to collect data in reliable manner and construct models based on mathematical and machine learning techniques to predict the optimum time and quantity of fuel. The proposed framework is implement and experimented with real datasets. The experimental results are promising.

Keywords

Machine learning, Prediction, Fuel Consumption, Artificial Neural Network (ANN), Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM), Big Data, Fuel Consumption, Random Forests.

1. INTRODUCTION

Total Quality Management (TQM) is an approach that aims to improve quality and performance which will meet customer expectations. This can be achieved by integrating all quality-related functions and processes throughout the organization. TQM looks at the overall quality measures used by an organization including managing quality design and development, quality control and maintenance, quality improvement, and quality assurance. These quality measures plays a major role for keeping the business running [2.3,4,5,16].

Quality management systems has been coined by the industry to classify their success in latest technology advancements in Manufacturing such as Internet of Things (IoT), Cyber Physical Systems (CPS), Cloud Computing, and Big Data. Industrial Internet of Things (IIoT) currently is focusing on either managing the quality of the products via improved quality performance management, process oriented analytics, or smart manufacturing environments which are placed to make extreme gains in cost savings, improved uptime, performance and getting better benefit [6,16,18].

Machine learning has been widely used to assess the quality performance. It is a form of artificial intelligence that uses algorithms to learn from big data stored in clouds. These systems build models from data, then find patterns to make predictions. It also used to solve problems in which the relationship between input and output variables is not known or is difficult to derive [11,24]. Such predictions can be a simple as providing a recommendation to a shopper on an e-commerce website or as complex as determining if a brand of automobile should be retired, or prediction of fuel consumption by Diesel Generators [19,20,21,22].

Developing the ability to predict fuel consumption of Diesel generator (DG) is extremely beneficial for budgeting and management; however, fuel consumption is measured by the amount of fuel used during a specific time period [10]. With increasing and fluctuating fuel prices, decision makers are constantly optimizing their processes to increase fuel efficiency for Diesel generator (DG). Diesel generator (DG) sets are used by a wide variety of businesses most of the times to provide a power source [1,7,8,9,24]. It comes in varying capacities for various businesses that they serve in different sectors such as, manufacturing, offices and universities. The issue here is that unless the DG set is operated at its optimal load (usually 80% of maximum load) the amount of diesel consumed as opposed to energy produced becomes skewed [23]. Sub-optimal load can cause losses up to 75% in revenue. Also DG sets start performing sub-optimally when regular maintenance is not done. The only way this can be ascertained by the amount of energy being produced, against the diesel consumed.

In this paper, a framework is proposed which is based on Analytics &IoT to Optimize performance of Diesel generator sets. It collects data stored in cloud storage. The data in clouds are obtained automatically by means of sensors. These data are preprocessed to obtain a reduced representation of the data set that is much smaller in volume but yet produce the same (or almost the same) analytical results. Once the data is prepared, it is subjected to machine learning algorithms to construct predictive models. It makes use of the MapReduce parallel programming model for the purpose of analysis the data in parallel manner [29,30]. Finally, the predictive performances of models were evaluated using machine-learning-relevant performance measures and subsequently, the optimal predictive model is used for decision making.

2. RELATED WORKS

Several approaches have been proposed for developing machine learning algorithms to address the issue of total quality management [2,3,4,16]. [2,3] Explores the possibilities that prescriptive data analysis can provide techniques to support the maintenance decision making process. [4] Proposed integration predictive maintenance in manufacturing industry toward predictive maintenance. [16] proposes a method that allows managing the quality of product during its life cycle at design stage.

There is also a significant amount of research conducted in related fields such as fuel consumption prediction for aircraft, passenger cars, and engines using statistical analysis and machine learning methods [35,36,38]. However their studies are limited in scope and do not reach a clear conclusion about the usability of machine learning models for fuel consumption, Nor the creating an optimal fuel prediction model using IoT and clouds techniques.

Furthermore, many fuel consumption models have been proposed in literatures [4,25,26,27,28]. however, many important issues still need to be addressed to enhance the current practice of adaptive fuel consumption and to extend its application into more realistic fuel consumption of DG environments. Therefore, this study aims to provide a comprehensive framework that makes use of Cloud computing and the IoT to collect statistical data for fuel consumption, preprocessed data, perform several prediction models in a parallel manner and select the optimal predictive model to be used for decision making in the life cycle of TQM.

3. PROBLEM STATEMENT

Given a dataset D collected over the time $[t_0,t_1,t_2,…t_n]$ and stored at Clouds $[c_0,c_1,c_2,…c_n]$ respectively. At t_0, D_0 represents an empty database. At time instance t_i, an incremental dataset D_i, $i \in \{1,…,n\}$, is collected such that $D = D_1 \cup D_2 \cup … D_i$. Suppose, D is saved into processors dP_j where $1 \leq j \leq m$ and m are the number of processors. The objective is to construct models M_i using dP_j and D with highest degree of the predictive performances.

4. THE PROPOSED FRAMEWORK

In this paper, a framework is presented that efficiently builds models from Big Data to predict the optimum time and quantity of fuel consumed by Diesel generators (DG). It makes use of MapReduce [29] to deal with data in a parallel manner. The general architecture of the proposed framework is shown in Fig. 1. It comprises of 2 phases namely, Data preprocessing and Building Prediction Models. Subsequently, the predicted model can be used for decision making to improve total quality management. These phases are explained in the following subsections:

4.1 Data Preprocessing

Quality decisions must be based on quality data e.g., duplicate or missing data may cause incorrect or even misleading decisions [33,34]. The purpose of the data preprocessing stage is to minimize potential errors in the model as much as possible. Generally, a model is only as good as the data passed into it, and the data preprocessing we do ensures that the model has as accurate a dataset as possible. The Data used in this work come from sensors installed on the Diesel tanks. In this paper, a dimensionality reduction technique is used to remove unimportant attributes. Feature selection (i.e., attribute subset selection) selects a minimum set of attributes (features) that is sufficient for the prediction task. The following steps are required to preprocess the data in the proposed framework.

- **Merge the data of the two Diesel tanks**
  In this step, the data of different tanks are combined together.

- **Convert Date Time column from string to DateTime object**

- **Create columns for (Year, Month, WeekDay, Day, DayQ, Hour ..etc)**

- **Creation of new features (Refilled, Refilled Amount, Consumption)**

- **Outliers Detection**
  This stage concerns with the identification of observations which raise suspicions by differing significantly from the majority of the data. Such outlier has to be removed from a dataset before actual model buildings.

- **Feature Selection**
  It refers to the process of reducing the inputs for processing and analysis. It concerns with finding the most meaningful inputs to be used for models building.

- **Splitting Dataset to Training data and Test data**
  The dataset to be used by non-time-series models are partitioned into training set and test set which are to be used by the first tree machine learning techniques, namely, Decision tree Regressor, Random Forest Regressor and XGBoost.
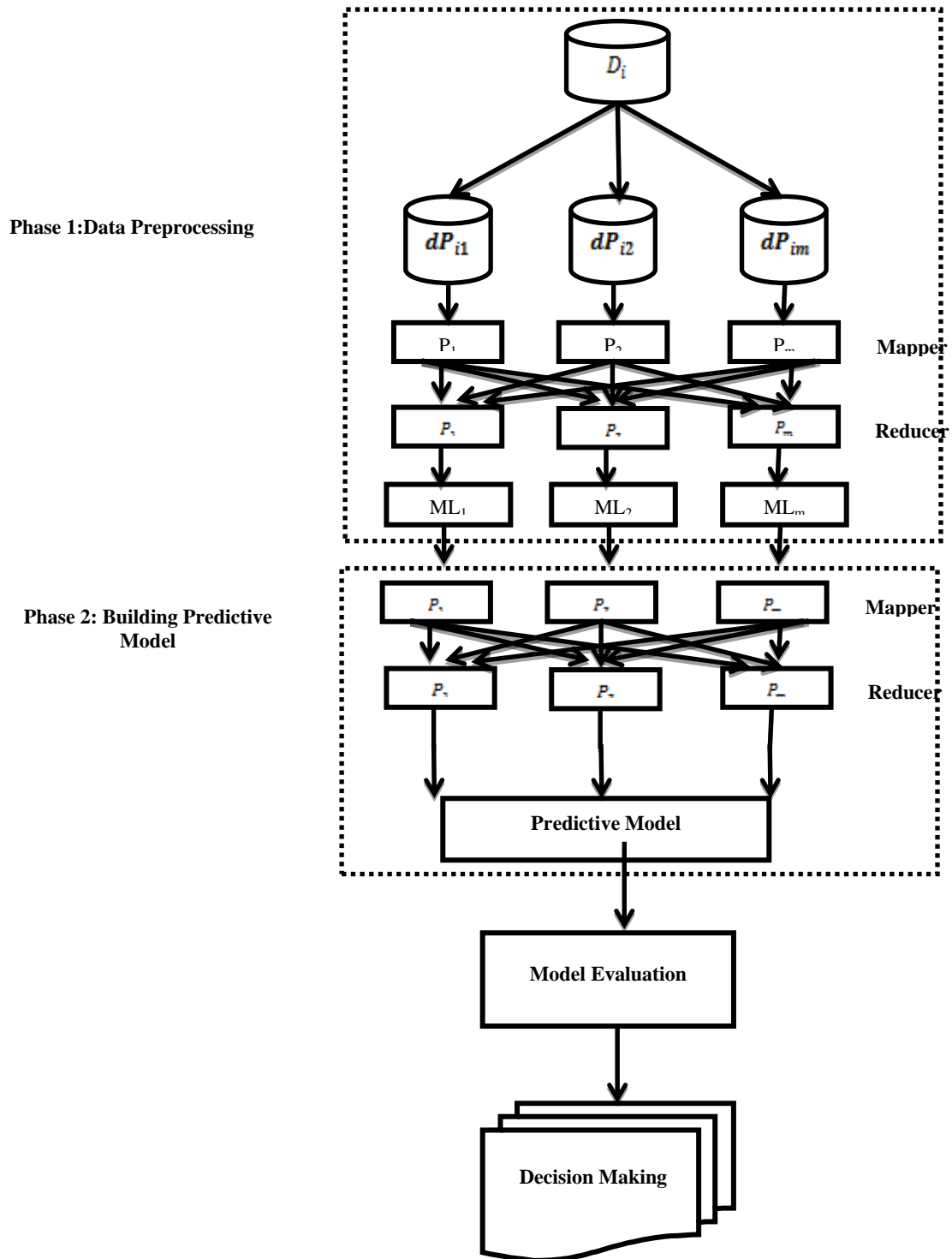
**Fig1. The General Architecture of the Framework for TQM of Diesel Generator Fuel
Consumption using Machine Learning and Internet of Things (IoT) s for Big Data**

4.2 Building Prediction Models

A framework is proposed that makes use of IIOT technology to collect data in reliable manner and construct models. It uses the MapReduce to deal with data and constructs models in a parallel manner. These models are based on mathematical and machine learning techniques in order to predict the optimum time and quantity of fuel consumption. Since there are multiple algorithms that can be used to build our models, we will compare the accuracy scores after testing and pick the most accurate model. Here 4 models under examination are presented and measure their quality to choose the best one, The following techniques are used:

4.2.1 Decision Tree Regressor:

Decision Trees are divided into Classification and Regression Trees. Regression trees are needed when the response variable

is numeric or continuous. Classification trees, as the name implies are used to separate the dataset into classes belonging to the response variable. Decision Trees (DTs) are a non-parametric supervised learning method used for classification and regression. The goal is to create a model that predicts the value of a target variable by learning simple decision rules inferred from the data features.

4.2.2 *Random Forest Regressor:*
Random Forest Regressor is one of the ensemble methods, The goal of ensemble methods is to combine the predictions of several base estimators built with a given learning algorithm in order to improve generalizability / robustness over a single estimator. In random forests, each tree in the ensemble is built from a sample drawn with replacement (i.e., a bootstrap sample) from the training set.

4.2.3 *XGBoost :*
It stands for eXtreme Gradient Boosting, which is an implementation of gradient boosted decision trees designed for speed and performance. XGBoost is one of the most popular machine learning algorithm these days. Regardless of the type of prediction task at hand; regression or classification. It is well known to provide better solutions than other machine learning algorithms. In fact, since its inception, it has become the "state-of-the-art" machine learning algorithm to deal with structured data.

4.2.4 *Long Short-Term Memory Neural Network (LSTM):*
Long Short-Term Memory networks, or LSTMs for short, is a kind of Recurrent networks, which take as their input not just the current input example they see, but also what they have perceived previously in time [31]. Here's a diagram of an early, simple recurrent net proposed by Elman [32], where the BTSXPE at the bottom of the drawing represents the input example in the current moment, and CONTEXT UNIT represents the output of the previous moment. Figure 2 shows the general architecture of Recurrent networks.
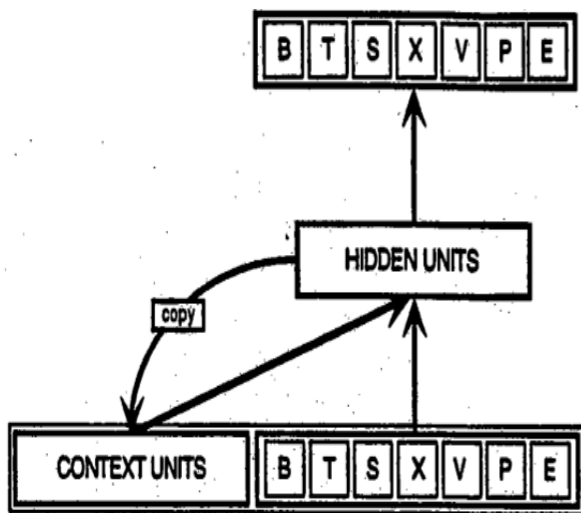

Fig 2. The SRN Network Architecture.

Each box in Figure 2, represents a pool of units and each forward arrow represents a complete set of trainable connections from each sending unit to each receiving unit in the next pool. The backward arrow, from the hidden layer to the context layer denotes a copy operation. The decision a recurrent net reached at time step t-1 affects the decision it will reach one moment later at time step t. So recurrent

networks have two sources of input, the present and the recent past, which combine to determine how they respond to new data, much as we do in life.

In this research Long Short-Term Memory network (LSTM is used. It can be applied to time series forecasting. There are many types of LSTM models that can be used for each specific type of time series forecasting problem. In our case, we use Stacked LSTM to predict the consumed amount of diesel in the next hour using consumption data of the last five hours.

4.3 Measure the Goodness of the Model
As training a model is a key step, the model generalization on future data is an important aspect that should be considered in machine learning algorithms [33]. In this section, the techniques used in evaluating the goodness of the models are explained and show how well a machine learning model generalizes to previously unseen data. Two metrics are used in the proposed frameworks to measure the quality of the trained models:

Mean Absolute Error (MAE): is a measure of difference between two continuous variables. The mean absolute error is an average of the absolute errors $|e_i|=|y_i - x_i|$, where y_i is the prediction and x_i the true value.

$$MAE = \frac{\sum_{i=1}^{n} |y_i - x_i|}{n} = \frac{\sum_{i=1}^{n} |e_i|}{n}$$

For a model to be efficient, it's Mean absolute error score MAE should be 0. As close MAE score of a model to zero, as accurate its predictions will be.

Mean Squared Error (MSE): is a measure of the quality of an estimator, It is always non-negative, and values closer to zero are better. It measures the average of the squares of the errors. that is, the average squared difference between the estimated values and the actual value.

$$MSE = \frac{1}{n}\sum_{i=1}^{n} (y_i - x_i)^2$$

5. A DETAILED EXAMPLE
For better understanding of our proposed framework, we consider a small dataset and perform data preprocessing, models building, and models testing. Subsequently, The framework selects the best model to be used for decision making.

- **Data preprocessing**

Consider the simple dataset shown in Table 1., for the sake of simplicity, we use 30 samples.

Table 1. The simple dataset

#	Date Time	Tank level	#	Date Time	Tank level
0	10/19/2019 23:00	11714.95	15	10/20/2019 14:00	11720.133 33
1	10/20/2019 0:00	11720.45	16	10/20/2019 15:00	11717.3
2	10/20/2019 1:00	11719.975	17	10/20/2019 16:00	11719.15
3	10/20/2019	11720.75	18	10/20/2019	11717.025

#			#		
	2:00			17:00	
4	10/20/2019 3:00	11719.7	19	10/20/2019 18:00	11718.1
5	10/20/2019 4:00	11718.125	20	10/20/2019 19:00	11716.475
6	10/20/2019 5:00	11718.975	21	10/20/2019 20:00	11716.525
7	10/20/2019 6:00	11718.033 33	22	10/20/2019 21:00	11716.525
8	10/20/2019 7:00	11721.325	23	10/20/2019 22:00	11717.05
9	10/20/2019 8:00	11722.4	24	10/20/2019 23:00	11715.7
10	10/20/2019 9:00	11720.85	25	10/21/2019 0:00	11646.775
11	10/20/2019 10:00	11721.35	26	10/21/2019 1:00	11402.85
12	10/20/2019 11:00	11721.35	27	10/21/2019 2:00	11403.4
13	10/20/2019 12:00	11721.75	28	10/21/2019 3:00	11402.725
14	10/20/2019 13:00	11721.85	29	10/21/2019 4:00	11403

We split Date Time column into seven columns, which are (year, month, day, weekday, weekend, dayQ, hour). Table 2 shows the first four samples of the dataset after splitting Date Time column.

Table 2. Sample of the dataset after splitting Date Time column

#	Tank level	year	month	day	weekday	weekend	hour	dayQ
0	11714.95	2019	10	19	5	1	23	4
1	11720.45	2019	10	20	6	1	0	1
2	11719.975	2019	10	20	6	1	1	1
3	11720.75	2019	10	20	6	1	2	1

After that, new features are created : (Refilled, Refilled Amount, Consumption) as shown in Table 3.

Table 3. Sample of the dataset after creating new features

#	Tank level	year	month	day	weekday	weekend	hour	dayQ	refilled	refilled_ Amount	consumption
0	11714.95	2019	10	19	5	1	23	4	0	0	0
1	11720.45	2019	10	20	6	1	0	1	1	5.5	-5.5
2	11719.975	2019	10	20	6	1	1	1	0	0	0.475
3	11720.75	2019	10	20	6	1	2	1	0	0	0

Note that the data in Table 3 contains an outlier because consumption column contains negative value, so the next step will be getting rid of outliers. Table 4 shows the data after removing the outlier.

Table 4. Sample of the dataset after getting rid of outliers

#	Tank level	year	month	day	weekday	weekend	hour	dayQ	refilled	refilled_ Amount	consumption
0	11714.95	2019	10	19	5	1	23	4	0	0	0
1	11719.975	2019	10	20	6	1	1	1	0	0	0.475
2	11720.75	2019	10	20	6	1	2	1	0	0	0
3	11719.7	2019	10	20	6	1	3	1	0	0	1.05

The next step is features selection for training the model. In this step, columns that will not be used to train the models are deleted. These columns are (year, refilled, refilled_Amount) and shown in Table 5.

Table 5. Sample of the dataset after features selection

#	Tank level	month	day	weekday	weekend	hour	dayQ	consumption
0	11714.95	10	19	5	1	23	4	0
1	11719.975	10	20	6	1	1	1	0.475
2	11720.75	10	20	6	1	2	1	0
3	11719.7	10	20	6	1	3	1	1.05

Then, consumption column is shifted one step to make it able to predict the future consumption.

Once preprocessing the data is completed, we will split the datasets to training data (20 samples), and test data (10 samples) as shown in Table 6 and Table 7 respectively.

Table 6. Training dataset

#	month	day	weekday	weekend	hour	dayQ	consumption
0	10	19	5	1	23	4	0.475
1	10	20	6	1	1	1	0
2	10	20	6	1	2	1	1.05
3	10	20	6	1	3	1	1.575
4	10	20	6	1	4	1	0
5	10	20	6	1	5	1	0.941667
6	10	20	6	1	6	2	0
7	10	20	6	1	7	2	0
8	10	20	6	1	8	2	1.55
9	10	20	6	1	9	2	0
10	10	20	6	1	10	2	0
11	10	20	6	1	11	2	0
12	10	20	6	1	12	3	0
13	10	20	6	1	13	3	1.716667
14	10	20	6	1	14	3	2.833333
15	10	20	6	1	15	3	0
16	10	20	6	1	16	3	2.125
17	10	20	6	1	17	3	0
18	10	20	6	1	18	4	1.625
19	10	20	6	1	19	4	0

Table 7. Test dataset

#	month	day	weekday	weekend	hour	dayQ	consumption
20	10	20	6	1	20	4	0
21	10	20	6	1	21	4	0
22	10	20	6	1	22	4	1.35
23	10	20	6	1	23	4	8.925
24	10	21	0	0	0	1	3.925
25	10	21	0	0	1	1	0
26	10	21	0	0	2	1	0.675
27	10	21	0	0	3	1	0
28	10	21	0	0	4	1	0
29	10	21	0	0	5	1	0

Next steps are to create three models using training data. The mathematical and machine learning techniques used are the following:

1. **Decision Tree Regressor.**

2. **Random Forest Regressor.**

3. **XGBoost.**

Once the models are constructed, the models' goodness are computed to test their performance on the test data as shown in Table 8.

Table 8. Models' Performance Measure (3 models)

#	Model Name	MAE	MSE
0	Decision Tree Regressor	6.23	113.4
1	Random Forest Regressor	5.78	107.63
2	XGBoost	6.85	127.94

The visualizing of the performance of the constructed models are shown in Figures 3,4 and 5 respectively.
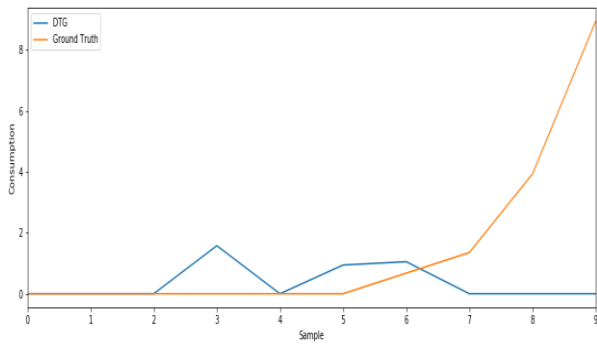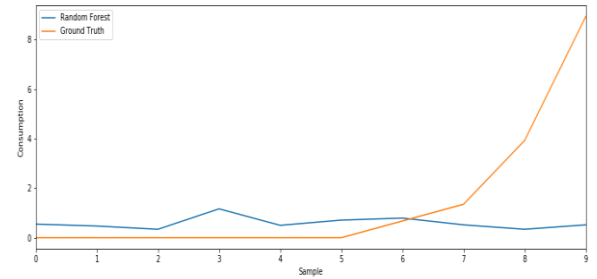
Fig 3. The Decision Tree Regressor Performance Measure
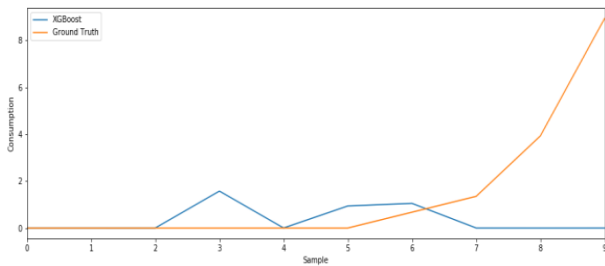

Fig 4. The Random Forest Performance Measure

Fig 5. The XGBoost Performance Measure

Notice that as show in Table 8, the performance of these models is not good enough due to the fact that their MAE and MSE are not close enough to zero. As MAE and MSE getting closer and closer to zero as the predicted accuracy getting higher and higher.

Now, we treat the data as time series data, and use a Long Short-Term Memory LSTM neural network to predict the fuel consumption. Next steps show the process of converting the previous dataset to a time series dataset.

In the time series dataset a sample will consists of two columns, first column contains the consumption data of the previous five hours, second columns holds the consumption for the next hour. Table 9 and Table 10 show the time series data after splitting the samples to training data and test data:

Table 9. Training time series data

#	X	y
0	[0.475, 0.0, 1.05, 1.575, 0.0]	0.942
1	[0.0, 1.05, 1.575, 0.0, 0.942]	0
2	[1.05, 1.575, 0.0, 0.942, 0.0]	0
3	[1.575, 0.0, 0.942, 0.0, 0.0]	1.55
4	[0.0, 0.942, 0.0, 0.0, 1.55]	0
5	[0.942, 0.0, 0.0, 1.55, 0.0]	0
6	[0.0, 0.0, 1.55, 0.0, 0.0]	0
7	[0.0, 1.55, 0.0, 0.0, 0.0]	0
8	[1.55, 0.0, 0.0, 0.0, 0.0]	1.717
9	[0.0, 0.0, 0.0, 0.0, 1.717]	2.833
10	[0.0, 0.0, 0.0, 1.717, 2.833]	0
11	[0.0, 0.0, 1.717, 2.833, 0.0]	2.125
12	[0.0, 1.717, 2.833, 0.0, 2.125]	0
13	[1.717, 2.833, 0.0, 2.125, 0.0]	1.625
14	[2.833, 0.0, 2.125, 0.0, 1.625]	0
15	[0.0, 2.125, 0.0, 1.625, 0.0]	0
16	[2.125, 0.0, 1.625, 0.0, 0.0]	0
17	[0.0, 1.625, 0.0, 0.0, 0.0]	1.35
18	[1.625, 0.0, 0.0, 0.0, 1.35]	8.925
19	[0.0, 0.0, 0.0, 1.35, 8.925]	3.925

Table 10. Test time series data

#	X	y
20	[0.0, 0.0, 1.35, 8.925, 3.925]	0
21	[0.0, 1.35, 8.925, 3.925, 0.0]	0.675
22	[1.35, 8.925, 3.925, 0.0, 0.675]	0
23	[8.925, 3.925, 0.0, 0.675, 0.0]	0
24	[3.925, 0.0, 0.675, 0.0, 0.0]	0
25	[0.0, 0.675, 0.0, 0.0, 0.0]	0.55
26	[0.675, 0.0, 0.0, 0.0, 0.55]	0
27	[0.0, 0.0, 0.0, 0.55, 0.0]	3.725
28	[0.0, 0.0, 0.55, 0.0, 3.725]	0
29	[0.0, 0.55, 0.0, 3.725, 0.0]	0

Now we can use this data to train and test the LSTM network. Figure 6 shows the performance of the LSTM mode. The performance measure of LSTM along with other Models are presented in Table 11.

Table 11. Models' Performance Measure (4 models)

#	Model Name	MAE	MSE
0	Decision Tree Regressor	25.05	6312.26
1	Random Forest Regressor	22.92	2897.86
2	XGBoost	26.67	3589.07
3	LSTM	11.55	2100.33

Note that, the values of MAE and MSE of LSTM mode are lower than other models and therefore, its predictive accuracy is the highest. Subsequently, this model will be used for prediction of fuel consumption of the organization.
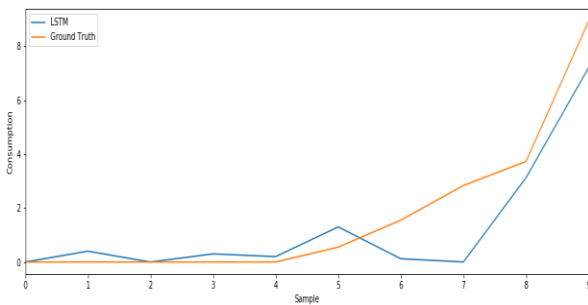

Fig 6. The LSTM Performance Measure

6. EXPERIMENTATION STUDY

The proposed approach is implemented in Python programming language and tested using real life datasets. Since, there are no other approaches available, which makes use of IIOT technology to collect data in reliable manner and constructs models based on mathematical and machine learning techniques in parallel way, to predict the optimum time and quantity of fuel consumption of Diesel Generators (DG), we could not perform any comparison against our framework. The following experiment were conducted to show the effectiveness of the framework.

The dataset belongs to Commercial Agricultural Bank (CACA BANK – HEAD OFFICE). It consists of 3527 example collected over time by reliable sensors and stored in a cloud using IoT technology.

- **Data preprocessing**
The following step are conducted in order to preprocess the collected data:

   o Date Time column is splitted into seven columns, which are (year, month, day, weekday, weekend, dayQ, hour). Table 12 shows the first four samples of the dataset after splitting Date Time column.

   o New features are created: (Refilled, Refilled Amount, Consumption)

   o Getting rid of outliers because the consumption values are negative.

   o Dimensionality reduction is applied to delete columns that will not be used to train the models, these columns are (**year, refilled, refilled_Amount).**

   o **consumption** column is shifted one step to make it able to predict the future consumption.

After processing the data, we ended up with 3480 samples. These samples are partitioned into training data (2610 samples), and test data (870 samples).

- **Models Building**
Three machine learning techniques are used to create models, train them, and test their performance on the test data. These techniques are:

   1. Decision Tree Regressor (DTG).

   2. Random Forest Regressor.

   3. XGBoost.

Once the three models are created using training dataset, these models are tested using testing dataset to measure the goodness of the models using Mean Absolute Error (MAE) and Mean Squared Error (MSE) metrics. Table 12 shows the overall quality of the models and Figures 7,8 and 9 show the visualizing of the performance of the models.

Table 12. Models' Performance Measure (3 models)

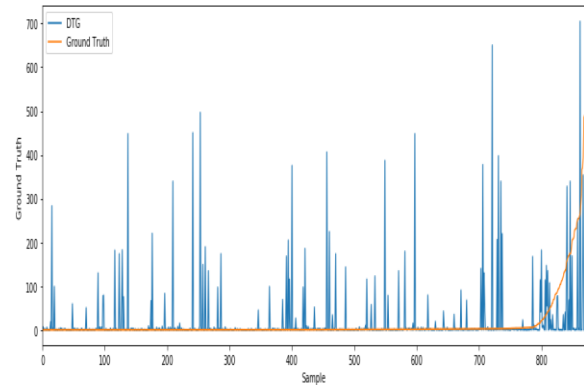| #	Model Name	MAE	MSE
0	Decision Tree Regressor	25.05	6312.26
1	Random Forest Regressor	22.92	2897.86
2	XGBoost	26.67	3589.07

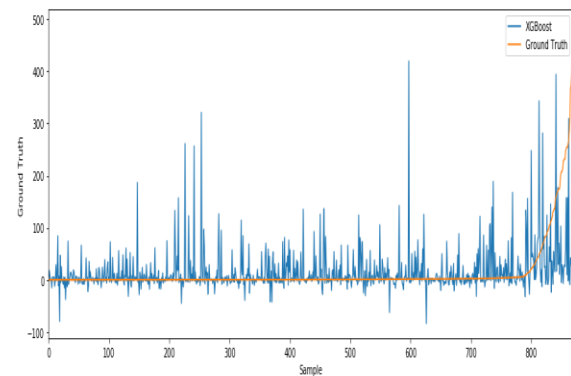**Fig 7. The Decision Tree Regressor Performance Measure**

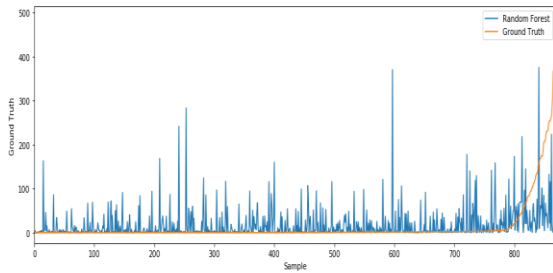**Fig 8. The Random Forest Performance Measure**

Fig 9. The XGBoost Performance Measure

It is observed from Table 12 that, the performance of these models is not good enough, due to the fact that their MAE and MSE are not close enough to zero. As MAE and MSE getting closer and closer to zero as the predicted accuracy getting higher and higher.

In order to obtain quality decisions, it is important to construct a model with a higher predicted performance. For this reason, the data is treated as time series data, and use a Long Short-Term Memory LSTM neural network to predict the fuel consumption.

In the time series dataset, the process of converting the previous dataset into a time series dataset is applied. The time series datasets consist of two columns, first column represents the consumption data of the previous five hours, and the second columns represents the consumption for the next hour.

Once the time series datasets are created, the samples are splitted into training data and test data to construct the model. Table 13 shows the overall quality or performance measure of the four models and Figure 10 shows the visualizing of the performance of the LSTM neural network model.

Table 13. Models' Performance Measure (4 models)

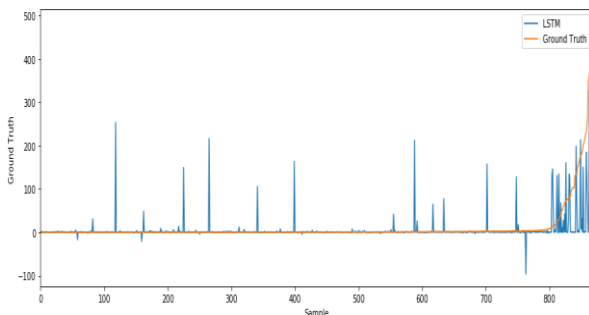| #	Model Name	MAE	MSE
0	Decision Tree Regressor	25.05	6312.26
1	Random Forest Regressor	22.92	2897.86
2	XGBoost	26.67	3589.07
3	LSTM	11.55	2100.33

Fig 10. The LSTM Performance Measure

Note that, the values of MAE and MSE of LSTM mode is lower than other models and therefore, its predictive accuracy is the highest. Subsequently, this model will be used for prediction of fuel consumption of the organization.

7. REFERENCES

[1] Begum, S., Banu, R., Ahamed, A. and Parameshachari, B. D. 2016. A comparative study on improving the performance of solar power plants through IOT and predictive data analytics. International Conference on Electrical, Electronics, Communication, Computer and Optimization Techniques (ICEECCOT), Mysuru. pp. 89-91.

[2] Klaus J. Zink. 2012. Total Quality Management as a Holistic Management Concept: The European Model for Business Excellence. Springer.

[3] Venkataraman V., 2010. Maintenance engineering and management. PHI Learning Private Limited.

[4] Efthymiou, K., Papakostas, N., Mourtzis, D. and Chryssolouris, G. 2012. On a Predictive Maintenance Platform for Production Systems. In proceedings of 45th CIRP Conference on Manufacturing Systems.

[5] Lihui, W. and Vincent Wang, Xi. 2017. Cloud-Based Cyber-Physical Systems in Manufacturing. Springer.

[6] Erdinc, K. 2020. Internet of Things (IoT): Applications for Enterprise Productivity. IGI Global.

[7] Ahn, K., Rakha, H., Trani, A., and Van Aerde, M. 2002. Estimating vehicle fuel consumption and emissions based on instantaneous speed and acceleration levels. Journal of Transportation Engineering, vol. 128, no. 2, pp. 182–190.

[8] Cartenì, A., Cantarella, G. E., and Luca, S. D. 2010. A methodology for estimating traffic fuel consumption and vehicle emissions for urban planning. In 12th World Conference for Transportation Research - WCTR, Lisboa, pp. 52–71.

[9] Siami-Irdemoosa, E., Dindarloo, S. R. 2015. Prediction of fuel consumption of mining dump trucks: A neural networks approach. Applied Energy, vol. 151, no. 1, pp. 77–84.

[10] Grisso R, Perumpral JV, Vaughan DH, Roberson GT, Pitman R , 2010. Predicting tractor diesel fuel consumption. Virginia,

[11] McQueen RJ, Garner SR, Nevill-Manning CG, Witten IH. 1995. Applying machine learning to agricultural data. Computers and Electronics in Agriculture.

[12] Federico P., Tony P. and Luis C. Neves, 2017. Application of Machine Learning for Fuel Consumption Modelling of Trucks. IEEE International Conference on Big Data, Workshop 22 - 'Applications of Big Data in the Transport Industry'.

[13] Carlos A. E. and Ruben M.-M. 2018. Machine learning techniques for quality control in high conformance manufacturing environment. Journal of Advances in Mechanical Engineering, Vol. 10(2).

[14] Máté Z. and Imre Z. 2018. Modelling fuel consumption and refuelling of autonomous vehicles|. In proceedings of MATEC Web of Conferences.

[15] Alexander S., Andy B., Brent H., Rishikesh M. B., Euzeli C. dos S. Jr., and Zina Ben M. 2019. A Machine Learning Model for Average Fuel Consumption in Heavy Vehicles. In IEEE Transactions on Vehicular

Technology.

[16] Nguyen, D.S. 2014. Total quality management in product life cycle. IEEE International Conference on Industrial Engineering and Engineering Management.

[17] Hecht, G., Josescheidt, B., Figueiredo, C.D., Moha, N., and Khomh, F. 2014. An Empirical Study of the Impact of Cloud Patterns on Quality of Service (QoS). IEEE International Conference on Cloud Computing Technology and Science. pp. 278–283.

[18] Dastjerdi, A. V. and Buyya, R. 2016. Fog computing: Helping the Internet of Things realize its potential. Computer, vol. 49, no. 8, pp. 112-116,

[19] Montgomery, D.C. 2013. Statistical Quality Control: A Modern Introduction, 7th ed.; Wiley: Hoboken, NJ, USA.

[20] Hao, L., Bian, L., Gebraeel, N., and Shi, J. 2017. Residual life prediction of multistage manufacturing processes with interaction between tool wear and product quality degradation. IEEE Trans. Autom. Sci. Eng., 14,1211–1224.

[21] Li, D.C., Chen, W.C., Liu, C.W., and Lin, Y.S. 2012. A non-linear quality improvement model using SVR for manufacturing TFT-LCDs. J. Intell. Manuf., 23, 835–844.

[22] Nada, O.A., Elmaraghy, H.A., and Elmaraghy, W.H. 2006. Quality prediction in manufacturing system design .J. Manuf. Syst., 25, 153–171.

[23] EPA fuel economy guide. Available Online at: http://www.fueleconomy.gov/feg/printGuides.shtml

[24] Zhao, Q., Chen, Q., and Wang, L. 2019. Real-Time Prediction of Fuel Consumption Based on Digital Map API. Applied sciences, 9, 1369.

[25] Ben-Chaim, M., Shmerling, E., and Kuperman, A. 2013. Analytic Modeling of Vehicle Fuel Consumption. Engergies, 6, 117–127.

[26] Post, K., Kent, J.H., Tomlin, J., and Carruthers, N. 1984. Fuel consumption and emission modeling by power demand and a comparison with other models. Transp.

Res. Part A, 18, 191–213.

[27] Frey, H.C., Rouphail, N.M., Zhai, H., Farias, T.L., and Gonçalves, G.A. 2007. Comparing real-world fuel consumption for diesel- and hydrogen-fueled transit buses and implication for emissions. Transp. Res. Part D, 12, 281–291.

[28] Weiqiang, Y., and Honggui, L. 2013. Application of grey theory to the prediction of diesel consumption of diesel generator set. Proceedings of 2013 IEEE International Conference on Grey systems and Intelligent Services (GSIS), Macao, , pp. 151-153.

[29] Dean, J., and Ghemawat, S. 2008. MapReduce: simplified data processing on large clusters. Communications of the ACM, vol. 51, pp. 107-113.

[30] Al-Hegami, A. S., and Alsaeedi, H. A. 2020. A Framework for Incremental Parallel Mining of Interesting Association Patterns for Big Data. International Journal of Computing, Volume 19, Issue 1, pp.106-117, Ukraine.

[31] Sepp, H., and Jürgen, S. *1997.* Long short-term memory". Neural Computation. 9 *(*8): 1735–1780.

[32] Elman, J. L. 1990. Finding structure in time. Cognitive Science, Volume 14, Issue 2, Pages 179-211.

[33] Müller, A. C., and Guido, S. 2016. Introduction to Machine Learning with Python. O'Reilly Media; 1st edition.

[34] Pyle, D. 1999. Data Preparation for Data Mining. Morgan Kaufmanns, San Francisco, CA, USA.

[35] Viswanathan, A. 2013. Data driven analysis of usage and driving parameters that affect fuel consumption of heavy vehicles. Master's thesis, Linköping University, Statistics, The Institute of Technology.

[36] Svärd, C. 2014. Predictive modelling of fuel consumption using machine learning techniques. Technical report, Scania CV AB.

[37] Mitchell, T. M. 1997. Machine Learning. McGraw-Hill, Inc., New York, NY, USA, 1st edition.