Application of Machine Learning in Physical Asset Management Employing Reliability

Shahzeb Anis[1], Raja Masood Larik[2*], Fatima Memon[3] and Muhammad Nadeem Iqbal[4]

[1] Grid System Maintenance and Protection Department, K-Electric Pvt Ltd,
Karachi, 75290, Pakistan (shahzebanisahmed@gmail.com)

[2*] Department of Electrical Engineering, NED University of Engineering and Technology, Karachi, 75290, Pakistan (rmlarik@neduet.edu.pk)

[3]Project Coordinator Limitless Solutions Karachi,Pakistan (memonfatima55@gmail.com)

[4] Department of Electrical Engineering Sir Syed University of Engineering and Technology

Abstract: Every power distribution utility system is equipped with a huge amount of electrical assets long-running and wide-spread including power grids, substations, power and auto transformers, transmission towers, overhead wires, insulators etc. Assets are deterioration time to time due to subsequent failures, environmental and physical impacts. This arises unexpected faults that occur within their useful-life and much sooner than their expected failure rate. This research paper shall bridge the traditional concept of asset management with application of machine learning (ML) performing reliability analysis using probabilistic & statistical techniques that will lead to modern asset management, maintenance management and reliability predictions based on ML Models. The research systematically involves recent published literature and equipment testing data in a categorical manner. The author utilizes historical failure data of power transformer of a utility to execute Weibull Analysis Reliability Index (RI) evaluation and Machine Learning based multivariate linear regression for RI Model to predict asset reliability. The findings of this research empower asset management in utility sector solving existing challenges related to Power Transformer.

Keywords: ***FMEA, Weibull Analysis, Machine Learning, Reliability***

I. INTRODUCTION

Traditionally, Asset Management was a vast concept being used in almost every public and private sector with a various number of interpretations. In domain of Electricity Utility, asset management could be defined as a systematic process of, cost effectively, operating, maintaining, upgrading of electrical assets by combining engineering practices and economic analysis with sound business practices [1].

Nowadays, regulators require these electrical power utilities to develop profitable long-term resource management strategies to reduce overall costs while maintaining system reliability. In this modern world, maintenance plays a vital role in the energy sector to improve the availability and reliability of energy resources [2]. The proper maintenance of power assets has become significant due to its high risk in operational, environmental and financial aspects.

Nowadays, intelligent or smart machines are gradually replacing and optimizing capabilities of human in different sectors. In this regard, Artificial Intelligence (AI) refers to the intelligence shown by software or machines to perform specific tasks. Furthermore, machine learning (ML) as a technology is also constantly evolving, having emerged as an advance computing technology concerning the logical patterns of algorithm, and complicated data structure designs.

This research shall bridge the traditional concept of Asset Management with modern Machine learning and Artificial Intelligence tools employing reliability with probability & statistic techniques that will lead to modern Asset Management, Maintenance Management and Failure Predictions based on Machine Learning Model.

II. PHYSICAL ASSET EMPLOYING RELIABILITY

Asset management involves making decisions to enable network companies to maximize long-term profits by providing customers with a high level of service with acceptable and manageable risk [3]. One of the major costs in “Asset Management” is the cost of maintaining system resources, for example, taking preventive measures, collectively referred to as preventive maintenance (PM) [4]. Overall reliability may be enhanced by reducing the duration/frequency of outages. PM activity can affect frequency by preventing the actual causes of failure. Therefore, PM is profitable when reliability exceeds the cost of implementing PM scale. Therefore, an architecture is needed to incorporate a system policy that links maintenance of system resources with increased system reliability. This is part of a broader concept of wealth management [5].

Every year Annual Preventive Maintenance (**APM**) carried out for almost all **78 grids** having **2 to 3** power transformer in each. Large quantity of assets indulge utility maintenance and protection engineer to record all the testing data and make a valued record for its future health and reliability predictions. In this research a **40MVA** power transformer is taken under consideration placed in one of our grid that has been installed in **1990**. Recorded data of its mechanical and electrical testing since **1992** till **2020** provides a very interesting trend to get it analyzed for its reliability and useful life assessment.

III. ARTIFICIAL INTELLIGENCE IN MAINTENANCE AND ASSETS MANAGEMENT

In the field of artificial intelligence, research has accelerated the development of smart technologies that have a major impact on people's daily lives. Engineering, science, medicine and business are undoubtedly getting smarter, thanks to their predictive ability to simplify everyday life and increase productivity. Implementing ML & AI in asset management could lead us to implement a better decision-making model for asset's Performance, its Health, the associated Risk and Consequences [7].

Artificial intelligence is used to improve, maintain and operate industrial facilities through asset management. This will provide engineers and maintenance operators with better decision-making tools [8]. Interest in state monitoring (CM) technology related to electrical equipment (especially condition monitoring technology including induction motors, generators and transformers) is increasing day by day, and it has been in power and power battery packs, windmill, solar system industry and power industry recognized. For example, artificial intelligence can be used to diagnose faults of rotating machines, it also plays an important role in diagnosing and detecting transformer faults. It can help engineers / operators to make the right decisions for any emergency intervention and choose the best one. Maintenance strategy for power transformers [9].

IV. METHODOLOGY FOR RELIABILITY ASSESSMENT

A. Weibull Analysis for Reliability Index

The Weibull distribution is implemented thoroughly on data that does not fit a straight line in a Weibull probability diagram. This type of data, especially is being followed by data points in map of probability, can indicate multiple failure modes for a set of failure times [6]. Weibull analysis provides simple graphical solution by plotted curves for better analysis. The scale at horizontal axis provides a measure of life, start/stop cycles, operating time, and equipment failure cycles. The scale at vertical axis is the event occurrence probability. The most significant part is the slope (β) that clue about the physics of failure.

The equation 4.1 shows cumulative distribution f unction (CDF) of the Weibull distribution is as f ollow,where η represent the characteristic life, or the age at which 63.2% **of units will have failed, and β represents the slope of the best-fit lineas shown in table 1.**

Mathematically,

$$F(t) = 1 - e^{-(t/\eta)^{\beta}} \tag{4.1}$$

Table 1 Shape Parameter in Weibull Distribution.

S No.	Shape Parameter	Description
1	β = 1/2	Infant Mortality • Inadequate burn-in • Misassemble • Some Quality Problem
2	β = 1.0	Random Failures • Independent of time • Maintenance error • Mixture of problems
3	β = 3.0	Early wear out • Surprise • Low cycle fatigue
4	β = 6.0	Old age wear-out (Rapid) • Bearings • Corrosions

B. METHODOLOGY OF MACHINE LEARNING MODEL

The proposed Machine Learning (ML) model is generic model for ML algorithms. In this section, there will be an overview of ML Model elements along with the availability of Transformer Failure Data compiled in this regards.

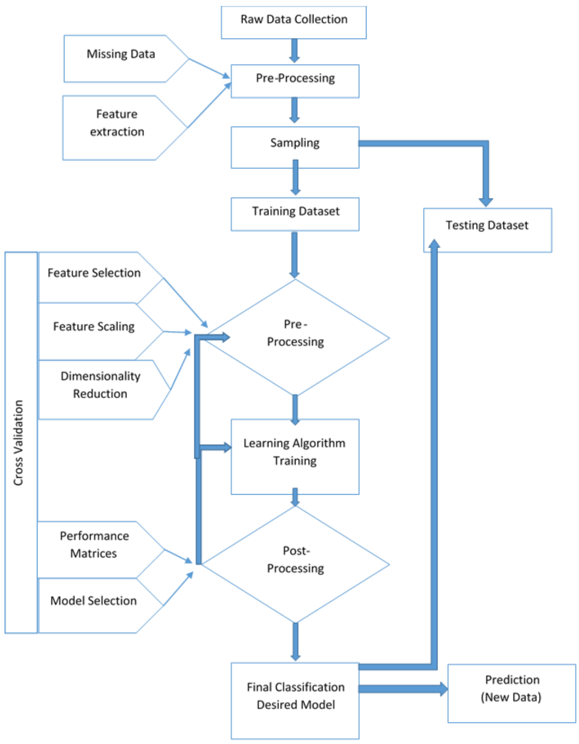

Fig. 1 Flow Chart for ML Model.

Application of Machine learning is carried out on power transformer historical data by flow chart in fig. 1. This Raw data includes Mechanical and Electrical parameter testing data. This raw data describe a brief history of Power Transformer on yearly basis from **1992 to 2019**. In order to apply machine learning model raw data shall be processed with **COLAB** (Python Programming Interface) using multivariate linear regression model.

V. RESULTS AND DISCUSSION

A. Weibull Analysis for Reliability

We consider 16 years recorded failure hour data with the cause of failure of one of the power transformer as shown in table 2. The data represent how much hour it took for a transformer to show failure and its status of fault / survive recorded from 2004 and onward.

Table 2 Failure Hour Data of Power Transformer.

B- General Failure Data Record:					
(Record before 2004 is unavailable ,78,840 hrs will be taken as initials since 1995 till 2004)					
S.NO	Recorded Year	Failure Hours	Failed/Suspended	Failure Cause	
				Details	Abbreviate
1	2004	83,512	F	Porcelain failure	PF
2	2005	89,880	F	Auxiliary control failure	ACF
3	2006	102,010	S	Out of calibration	OC
4	2007	110,770	F	Restricted oil flow	ROF
5	2008	121,330	F	Weld failure	WF
6	2009	126,312	F	Radiator clogged	RF
7	2010	133,535	F	Solid insulation failure	SIF
8	2011	145,427	F	Gasket failure	GF
9	2012	149,980	F	Bushing failure	BF
10	2013	162,770	F	Valve leak Failure	VF
11	2014	169,528	F	Over pressurization	OP
12	2015	180,298	F	Corrosion	CR
13	2016	187,470	F	Loose connection	LC
14	2017	195,675	F	Bushing CT failure	BCF
15	2018	203,026	F	Tap changer failure	TF
16	2019	213,040	F	Oil dielectric failure	ODF
17	2020	227,630	S	Oil contamination	OC

Here in table 2 each data point is accompanied by either an "**F**" or an "**S**", depending upon if the unit failed (**F**) or if the test was suspended (**S**) before the unit failed. Also note record before 2004 was unavailable, **78,840 hrs**. will be taken as initials since **1992** till **2004**.

Table 3 Minitab Results for Weibull Analysis.

	Estimate
Mean(MTTF)	156971
Standard Deviation	44959.9
Median	157886
First Quartile(Q1)	126082
Third Quartile(Q3)	188516
Interquartile Range(IQR)	62434.7

This is the most significant evaluation in table 3 for reliability estimation of power transformer. As **MTTF = 156,971** hrs. that divided by total hrs. in an year (**8760 hrs./year**) equals **17.9 years** as the expected life till its last failure.

Table 4 Minitab Results for Reliability Indices.

Time (hrs.)	Reliability	Time (hrs.)	Reliability
8760	1.00000	131400	0.69420
17520	1.00000	140160	0.62711
26280	1.00000	148920	0.55741
35040	1.00000	157680	0.48703
43800	0.99991	166440	0.41791
52560	0.99813	175200	0.35184
61320	0.99252	183960	0.29038
70080	0.98146	192720	0.23471
78840	0.96372	201480	0.18565
87600	0.93839	210240	0.14357
96360	0.90492	219000	0.10846
105120	0.86323	227760	0.07998
113880	0.81362	236520	0.05752
122640	0.75688	245280	0.04031
131400	0.69420	250000	0.03292

In above table 4 reliability indices are estimated that represents the proportion of units that would survive beyond a specific time. We can clearly observe from the results that in a span of **11.41** years **89%** assist reliability is estimated while with in an increase of time survival probabilities decreases and at almost **28 years** of life the assist existing probability descends to **1.5%.** This will allow asset engineer to take proactive decision for asset maintenance and replacement intervention and help an organization to think of their capital expenditure and operational expenditure in a better way for any individual and group of assets.

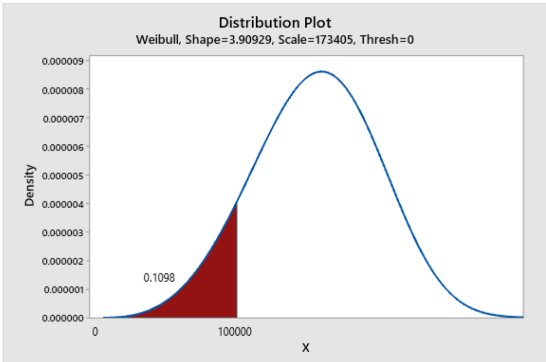

Fig. 2 Asset reliability after 10,000 hrs. of operation

The Probability Distribution plot in fig. 2 clearly indicates that at **100,000 Hrs**. that approximately become **11.41 years** only **10.98%** shall be the failure probability.

B. ML Based Multivariate Regression for Reliability

Compiled data consisting of mechanical and electrical parametric history includes below major testing result that directly or indirectly drives the asset performance, health, associated risk and consequences. A data of **23** Parameters is compiled including ***DGA (Dissolve Gas Analysis), OM (Oil Moisture), F55 (Furan levels at 55ºC), F65 (Furan levels at 65ºC), AFL (Average Furans Levels), RHVWR (Red High Voltage Winding Resistance), YHVWR (Yellow High Voltage Winding Resistance), BHVWR (Blue High Voltage Winding Resistance), AHVWR (Average High***

Voltage Winding Resistance), RLVWR (Red Low Voltage Winding Resistance), YLVWR (Yellow Low Voltage Winding Resistance), BLVWR (Blue Low Voltage Winding Resistance), ALVWR (Average Low Voltage Winding Resistance), TWR (Tertiary Winding Resistance), HLIR (HV-LV Insulation Resistance), HGIR (HV-Ground Insulation Resistance), HTIR (HV-Tertiary Insulation Resistance), LGIR (LV-Ground Insulation Resistance), LTIR (LV-Tertiary Insulation Resistance), TGIR (Tertiary-Ground Insulation Resistance), TTR (Transformer Turn Ratio), CNDFHV (Capacitance and Dissipation Factor HV), CNDFLV (Capacitance and Dissipation Factor LV) , CNDFTV (Capacitance and Dissipation Factor TV), CNDFCOM (Capacitance and Dissipation Factor COMMON), and WTH (Winding Temperature History). Along with 23 parameters 3 average parameters are also taken into count that includes Average Furans Levels (**AFL**), Average High Voltage Winding Resistance (**AHVWR**) and Average Low Voltage Winding Resistance (**ALVWR**).

Artificial Intelligence, Machine Learning and Neural Network programs requires an initial set of data. This set of data serve as a baseline foundation of training data for further utilization and analysis. We may conclude training data as a valuable resource for data analytic to build up an effective machine learning model. We train the designed model by providing them with detail, clear, and comprehensive information about the specified task. Majority Machine learning models fails due to very poor processing of data where as a successful ML model is direct impact of highly-built, properly formatted and annotated training dataset.

The most significant part of ML Model is "**Training Data**" that allows our ML model to perform predictive and forecast analysis. Major part of data is treated as training data that comprises of almost **75-80%** of the data for model-building. This data after a multi-cycle training provided algorithm a very improved accuracy. It is also a performance aspect that we need to omit the information that giving us no additional value. These are the reason why we require "**Multi-co-linearity check**" to be incorporated before training data proceedings.

We can observe higher percentages of multi-co-linearity in variables such as *DGA, AFL, CNDFHV, CNDFLV, CNDFTV and CNDFCOMM.*

```
[37] <matplotlib.axes._subplots.AxesSubplot at 0x7f1082838898>
```

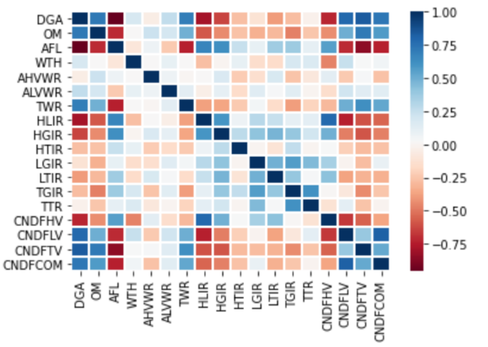

Fig. 3 Heatmap for multicolinearity check

We can observe in fig 3 the variable in diagonal are perfectly dark blue indicating the self-correlation while the dark blue position in either sides of diagonal shows how one variable is correlated with another with a great percentage. These highly correlating variable present either used of diagonal shall be taken out from regression analysis because they will ultimately disturb the results.

```
[ ] # Redinfing Trafo dataframe, adding Reiability Index in the coloumn again.
    Trafo_df_after = Trafo_df.drop([ 'DGA','AFL','CNDFHV','CNDFLV','CNDFTV','CNDFCOM'], axis = 1)


    # Defining our required inut and output varaibles
    X = Trafo_df_after.drop('RI', axis = 1)
    Y = Trafo_df_after[['RI']]

    # Splitting X an Y data into training and testing dataset
    X_train, X_test, Y_train, Y_test = train_test_split(X, Y, test_size=0.25, random_state=1)
```

Fig. 4 Test Train of ML Model in **COLAB**

Both X and Y values are provided for training data set in our machine learning model as shown in fig. 4. These variables are divided into training and testing data set. We decided to break **75%** of the data set for training and **25%** of the data set shall be utilized for testing purposes.

We have set our dependent variable **Y** as RI and independent variables $\mathbf{X = X_1 + X_2 + X_3 \ldots . X_n}$, where **n** is the no of predictors. Now we can determine the relationship between Y and X. There is also an involvement of "**estimator**" that are the coefficients of independent variables as

$\mathbf{a_0 + a_1 + a_2 \ldots a_n}$. Now we can conclude our regression function as.

$$\mathbf{Y = a_0 + a_1X_1 + a_2X_2 + a_3X_3 + \ldots a_nX_n} \qquad (5.1)$$

```
[ ] The intercept for our model is 11.78
    ------------------------------------
    The Coefficient for OM is -2.4
    The Coefficient for WTH is -1.3
    The Coefficient for AHVWR is 0.014
    The Coefficient for ALVWR is -5.4
    The Coefficient for TWR is -0.16
    The Coefficient for HLIR is 1.4e+01
    The Coefficient for HGIR is 8.1
    The Coefficient for HTIR is 1.6
    The Coefficient for LGIR is -5.4
    The Coefficient for LTIR is 4.3
    The Coefficient for TGIR is 4.3
    The Coefficient for TTR is -3.3e+01
```

Fig. 5 Cofficinents of ML based RI Model

We may fomulate the above ML model generated coefficients referring equation 5.1 with below methamatical model for our RI mentioned in fig 5.

$$f(x) = 11.78 - 2.4x_1 - 1.3x_2 + 0.014\,x_3 - 5.4x_4 - 0.16x_5 + 14x_6 + 8.1x_7 + 1.6x_8 - 5.4x_9 + 4.3x_{10} + 4.3x_{11} - 33x_{12} \quad (5.2)$$

Where

x_1 = Oil Moisture, x_2 = Winding Temperature History, x_3 = Average HV Winding Resistance, x_4 = Average LV Winding Resistance, x_5 = Tertiary Winding Resistance, x_6 = HV-LV Insulation Resistance, x_7 = HV-Gnd Insulation Resistance, x_8 =HV-Tertiary Insulation Resistance , x_9 = LV-Gnd Insulation Resistance, x_{10} =LV-Tertiary Insulation Resistance , x_{11} = Tertiary – Gnd Insulation Resistance and x_{12} = Transformer Turn Ratio.

C. Reliability Prediction

Once we divided data into X_test and X_train category, we are in position to get prediction from our model on our x_train values for year **2020**. For this using regression_model.predict () code in python interface by an iterative procedure.

Table 5 Iterative Results for RI predict

Iteration	RI Predicted
1st	2.115
2nd	59.346
3rd	30.118
4th	53.046
5th	58.225
6th	68.350

Reliability of our asset for year **2020** was found as **68.35%** while same was calculated using mathematical model using variable data for **2020** as shown in table 5.

Table 6 Testing Results for our PTR in 2020

Year	2020
OM	11.0
WTH	56.0
AHVWR	687.9
ALVWR	4.68
TWR	184.4
HLIR	18.75
HGIR	21.55
HTIR	22.5
LGIR	24.75
LTIR	32.22
TGIR	28.52
TTR	12.37

Substituting data of 2020 from table 6 in our equation 5.2.

$$f(x) = 11.78 - 2.4(11.0) - 1.3(56.0) + 0.014\,(687.9) - 5.4(4.68) - 0.16(184.4) + 14(18.75) + 8.1(21.55) + 1.6(22.5) - 5.4(24.75) + 4.3(32.22) + 4.3(28.52) - 33(12.37)$$

We get,

$$f(x) = 59.8\,\%$$

Therefore, our reliability index for 2020 is **59.8%** depending upon all those variable those are neither effected by multi-co-linearity check nor by variance inflation factor.

Table 7 Error Estimation between Results

Year	ML Predicted Reliability	Mathematically evaluated reliability	Error
2020	68.35%	59.8%	12.5%

Error estimated from output of our predicted value and mathematically estimated value is **12.5%** as shown in table 7**.**

V. CONCLUSION

Achieving results in previous section conclude comparative reliability assessment through linking traditional and ML based results. Firstly, a **16 years** recorded failure hour's data of a power transformer was utilized to perform Weibull analysis. Targeted Shape parameter is determined as 3.90 that is greater than 3 indicating the failure rate will increase with time causing wear & tear issues, lack of proper maintenance and descends in useful life of power transformer. Furthermore, surprise damages and low cyclic fatigues are also an essential failures that could be aced in near future. Estimated years of survival from Weibull analysis is found to be **17.9 years**.

Estimated survival probabilities from Weibull analysis is taken as a dependent variable for assets historical testing results. Raw data compiled for power transformer during its APM since first year of asset holding almost 18 independent variables were effectively supplied to machine learning model on COLAB interface. **75-80%** of training data by performing multivariate linear regression test on **25-20%** data predicting Reliability Index of power transformer that after iteration calculated as **68.35%** with historical results and for **2020** data we found **59.8%.**

All above software, coding aided analysis are performed to evaluate asset management requirement of power transformer which is an expensive and backbone equipment of power utility. This assessment and describe tools shall be actively utilized in evaluation of other power and auto transformer for their reliability factors estimation. Results of each step can be understood detail wise in previous.

REFERENCES

[1] Davidson, Innu. (2005). Utility asset management in the electrical power distribution sector. Proceedings of the Inaugural IEEE PES 2005 Conference and Exposition in Africa. 2005. 338 - 343. 10.1109/PESAFR.2005.1611842.

[2] A. Y. Alseiari, P. Farrel and Y. Osman, "The Impact of Artificial Intelligence Applications on the Participation of Autonomous Maintenance and Assets Management Optimisation within Power Industry: A Review," 2020 IEEE 7th International Conference on Industrial Engineering and

Applications (ICIEA), Bangkok, Thailand, 2020, pp.113-121,doi:10.1109/ICIEA49774.2020.9102017.
[3] Gavrikova, E.; Volkova, I.; Burda, Y. Strategic Aspects of Asset Management: An Overview of Current Research. Sustainability 2020, 12, 5955. https://doi.org/10.3390/su12155955
[4] Dhingra, Dr. Tarun & Velmurugan, Rama. (2015). Maintenance Strategy Selection and its Impact on Maintenance Function - A Conceptual Framework. International Journal of Operations & Production Management. 10.1108/IJOPM-01-2014-0028.
[5] Keith Bell & Simon Gill (2017) Delivering a highly distributed electricity system: Technical, regulatory and policy challenges, Energy Policy Volume 113, February 2018, Pages 765-777
[6] W. Li, "Incorporating aging failures in power system reliability evaluation," IEEE Transactions on Power Systems, vol. 17, no. 3, pp. 918-923, Aug. 2002.
[7] Schuman, C.A. and Brent, AC., 2005. Asset life cycle management: towards improving physical asset performance in the process industry. International Journal of Operations and Production Management, 25 (6), 566–579
[8] M. Žarković and Z. Stojković, “Analysis of artificial intelligence expert systems for power transformer condition monitoring and diagnostics,” Electric Power Systems Research, vol. 149, Aug. 2017, pp. 125-136.
[9] Z. He, Y. Wang, J. Li, S. Gong, and S. Grzybowski, ‘‘New mixed Weibull probability distribution model for reliability evaluation of paper-oil insulation,’’ Przegla¸d Elektrotechniczny, vol. 89, no. 1, pp. 202–206, 2013.