Grenze International Journal of Engineering and Technology, Jan Issue

Predictive Maintenance with Machine Learning

Angel Rose C J[1], Praveen A[2] and Krishna S[3]
[1-3]Department of Computer Science and Applications, Amrita School of Computing, Amrita Vishwa Vidyapeetham, Amritapuri, India
Email: angelroseecj@gmail.com, praveenyuva01@gmail.com, krishnas@am.amrita.edu

***Abstract*—Machine learning is widely used in predictive renovation to analyse current to be filled out uncover a pattern that can be used aid are anticipating failures of equipment early. A predictive model restoration extends the life of a device keeping track of its condition, reducing unplanned downtime, and repairing it when needed costs enhancing performance universal effectiveness. According to this study I will begin introduces the use of machine learning strategies for anticipating unforeseen disasters before they occur. Following that, in order to address the issue of perpetual downtime, it delves deeper into anomaly detection. It also describes the rationale for selecting these tactics. Furthermore, the methodologies in this article study having been tried using popular statistics-units, with encouraging results.**

***Index Terms*— Predictive Maintenance, Condition Based Maintenance, Machine Learning, Industry 4.0, Time Series.**

I. INTRODUCTION

Machines are now essential not just to the manufacturing business, but to all industries. Corrective maintenance or reactive remodelling is typically performed to restore machinery to suitable operating conditions following a breakdown or failure. Corrective maintenance might result in greater cost of maintenance and inconvenient downtime. On the other hand, predictive maintenance (Pd.M.) keeps the machinery in good working order on precisely forecasting when a malfunction or breakdown may occur and implementing corrective actions in the region. As a result, unplanned downtime is reduced, and the system's lifespan is extended. Some instances of Pd.M. in relation to ball bearings are: determining bearing existence (as a result of use) before it fails; determining whether a direction requires or lubrication now, raising an alert when ointment is polluted, etc. Pd.M. is commonly utilizing statistical analysis the device learning estimation algorithms abnormal behaviour, final usable time to failure (TTF) and life (RUL).

Pd.M. based on machine learning (ML) is classified into two kinds: supervised and unsupervised. Supervised instruction is largely concerned with developing prediction models are generating forecasts. Supervised instruction demands for statistics of both the input and the output, Approach labelled data should accessible. Regression and type are two examples of supervised learning algorithms. Algorithms that are mostly based on regression Take input data and convey constant output cost, For instance, the ultimate useful life (RUL) of a component is the amount of time left until the device or one of its components fails. Similar to this, category-based programmes accept input and output individually, suggesting that the success of the system or one of its components cannot be avoided. However, there isn't a classified available output or data for unsupervised learning, what's more, there is additionally no data in the statistics regarding how system errors emerge. Unsupervised learning no longer performs Forecasts, but it can be used to detect aberrant behaviour. Anomalous

Grenze ID: 01.GIJET.10.1.565_1

behaviour can be generated by employing unusual occurrences or observations. Furthermore, it provides more information regarding the intrinsic construction of the information and enables for the discovery of hidden Styles and correlations within the records. It can also be used to group records based on their similarities or differences.

II. LITERATURE REVIEW

A. THAYGO.P.CARVALHO

How much information separated from creation processes has expanded dramatically because of the expansion of detecting advancements. When handled and dissected, information can bring out significant data and information from assembling process, creation framework and gear. In enterprises, gear support is a significant key, and influences the activity season of hardware and its proficiency. {b1}Subsequently, hardware deficiencies should be recognized and addressed, keeping away from closure in the creation processes. AI (ML) techniques have been arisen as a promising device in Prescient Support (Pd.M.) applications to forestall disappointments in hardware that make up the creation lines in the manufacturing plant floor. In any case, the presentation of Pd.M. applications relies upon the proper decision of the ML strategy. This paper's goal is to give an effective writing survey of ML techniques applied to Pd.M., highlighting those that are being researched in this area and outlining the current state of the craft of ML procedures. This study is centred on two logical data sets and provides a useful background on ML techniques, their main results, challenges, and advantageous openings. It also supports future research activities in the Pd.M. field.

B. ANDEA FELICETTI

Condition monitoring and proactive maintenance of the business's electric engines and other hardware prevents severe financial losses brought on by unexpected engine failures and significantly increases system dependability.{b2}The AI architecture for prescient support presented in this paper is based on an unconventional Timberland approach. By promoting the information collection and information framework inquiry, applying the AI technique, and comparing it with the re-enactment instrument examination, the framework was tested on a real-world industrial model. Different sensors, machine PLCs, and correspondence protocols have collected information that has been made available to the Information Examination Apparatus on the Sky-blue Cloud engineering. Initial results demonstrate a proper mode of operation for the methodology on accurately predicting various machine states.

C. AMEETH KANAWADEY

Industrial Internet of Things (IIoT) refers to the application of Internet of Things (IoT) technology in the manufacturing sector, which harnesses the machine data produced by various sensors and applies various analytics to it to obtain relevant information. The date and time component that is typically included with the data that the devices collect is crucial for predictive modelling{b3}. In order to forecast potential failures and quality flaws, this research investigates the application of Autoregressive Integrated Moving Average (ARIMA) forecasting using time series data obtained from multiple sensors from a slitting machine. This would enhance the whole manufacturing process. Thus, machine learning is an essential part of the IIoT with applications in quality management and quality control, reducing maintenance costs, and enhancing overall manufacturing process.

D. JORGE BARBOSA

As of late, the fourth modern unrest has stood out around the world. A few ideas were brought into the world related to this new unrest, like prescient upkeep. This study expects to explore scholastic advances in disappointment forecast.{b4} The expectation of disappointments considers ideas as a prescient upkeep choice emotionally supportive network and a plan emotionally supportive network. We centre around structures that utilization AI and thinking for prescient upkeep in Industry 4.0. All the more explicitly, we consider the difficulties in the use of AI methods and ontologies with regards to prescient support. We direct a deliberate survey of the writing (SLR) to examine scholarly articles that were distributed online from 2015 until the start of June 2020. The screening system brought about a last populace of 38 investigations of a sum of 562 dissected.

E. ALESSANDRO BEGHI

A new classifier AI (ML) system for foresight support (Pd.M.) is presented in this paper. Given the growing demand to reduce human time and associated expenses, Pd.M. is an obvious solution for managing support difficulties{b5}. One of the challenges with PdM is determining the relationship between working costs and disappointment risk and the purported "wellbeing factors," or quantitative pointers, of the scenario with a framework associated to a given support issue. The proposed PdM approach may be applied to high-layered and controlled information problems and allows dynamical choice guidelines to be adopted for board maintenance. To achieve this, different grouping modules are prepared with various forecast skylines to provide different execution trade-offs regarding the frequency of unforeseen breaks and unutilized lifetime, and this data is then used in a working expense-based upkeep decision framework to reduce anticipated costs. Utilising a model that has been re-enacted and a benchmark semiconductor producing maintenance issue, the strategy's viability is demonstrated.

F. HARDIK.A.GOHEL

Atomic foundation structures have a big role in public safety. For the government, organisations, society, and everyday lives of citizens, atomic foundation framework capabilities and missions are essential.{b6} It is critical to plan the atomic framework for adaptability, dependability, and robustness. To achieve this, we can make use of artificial intelligence (AI), a cutting-edge technology used in a variety of industries, including autonomous vehicles, the Web of Things (IoT), and voice recognition. In this research, we suggest developing an AI calculation to execute atomic foundation maintenance in the future. The forecast will be implemented using a support vector machine and calculated strategic relapse.

III. METHODOLOGY

The prior works concentrated on various Pd.M. components and contemporary advances using ML. This research mostly concentrates on choosing ML fashions the Pd.M. and evaluating the overall performance. Conversely, however the illustration in this, article emphasises the crucial Pd.M. difficulties. Furthermore, depending on both labelled and unlabelled data, it addresses the bulk of Pd.M.-related situations.

a) Business data: Among the main problems with Pd.M. projects based on ML that they begin with statistics constructing models rather than business data. Different hobby niches have unique challenges and expectations. Enterprise targets and objectives ought to be specified first. Use scenarios that fulfil with stated objectives are observed, together with the tools and technologies needed to achieve these objectives.

b) Knowledge of statistics: After creating the subsequent business cases stage is to gather and comprehend data. Right now, two are present frequent scenarios: either the information might be/has been gathered making use of existing sensors, or new or more sensors need to be installed to collect the data needed to satisfy the specifications of the use cases. In the initial case, In the first situation, an ML model is selected to best suit the given data, whereas in the second scenario, precise data must be obtained based on a preplanner ML model. At this point, the most significant query to address is: "Can existing or likely existing facts be used to obtain the described commercial enterprise goals?" Exploratory Data Analysis (EDA) is accomplished to obtain understanding of the acquired statistics. EDA is useful for understanding the structure of the facts and determining whether further cleaning or the collection of additional records may be necessary. Hypotheses are also being created throughout this segment to successfully develop an ML-based venture for Pd.M. by combining business and records technology.

c) Records training: data processing at this segment is done to get the data ready for predictive modelling. Here, several strategies are employed, including function engineering and information cleaning. Both historical and static statistics are typically necessary to develop an ML version for Pd.M. The device's maintenance and failure history are included in old statistics. It also contains information on the incidents that led to system failure. The mechanical components of the system, the use of the equipment, working conditions, and so on are all included in static records.

d) Modelling: This division applies a variety of distinct algorithms/fashions to a large range of use cases. More modelling options are available in supervised learning than in unsupervised learning, allowing one to do predictive preservation. If a runto-failure record is available, for example, similarity version can be utilized to estimate RUL. Similarity search can uncover patterns that represent both normal working conditions and equipment failures. The paradox model, in contrast to the similarity version, can find patterns that don't fit with normal working settings. Additionally, if lifetime data that demonstrates how long it took for similar machines to fail is available, the survival model as a possibility distribution may be used. Furthermore, if there aren't any or few existence records and/or failure statistics available, a degradation version will be utilized (in view of a condition threshold that indicates failure). It is essential to discuss if the findings meet the business objectives because several models may be employed.

e) Evaluation: The best appropriate version is found using several comparison procedures to be able to successfully evaluate the presentation of the selected models. R-Squared, mean outright error (MAE), root infer square errors (RMSE), and other metrics are examples of metrics that can be used to evaluate a standard regression model. On the other hand, RUL's accuracy can show a disparity between projected and actual life. Similar techniques for evaluating a category version include confusion matrices and F1 ratings. Given that there is no labelled record accessible, evaluating an unmanaged model is less honest than evaluating one that has been supervised. However, records with the best normal behaviour or data that has already been classified by subject-matter experts allow for evaluation of the model's case F1rating, precision, recall, and accuracy; Otherwise, the domain experts must confirm the results. Reviewing, business and factual know-how levels is advised if the results still call for improvements or changes.

f) Deployment: It is crucial to continuously reveal and improve the model after deploying the excellent one to obtain better predictions.

IV. RESULTS

TABLE : EVALUATION OF REGRESSION BASED MODELS FOR RUL

Model	RMSE	R2-score	Error cycle
SVR	0.048	-4.028755379916393	-2
SVR*	0.041	-2.4280651482293845	-2
LassoLarsCV	0.050	-4.48447130580669	0
SGDRegressor*	0.017	0.393702606072853	2

TABLE : EVALUATION OF ANOMALY DETECTION MODELS

Model	Accuracy(%)	Precision(%)	Recall(%)	F1-score
One-class SVM	63.00	57.47	100.00	0.73
Elliptic Envelope	60.32	55.75	100.00	0.72
Isolation Forest	67.42	60.54	100.00	0.75
Local Outlier Factor	52.38	51.32	92.38	0.66

TABLE : EVALUATION OF CLASSIFICATION MODELS FOR FAILURE PREDICTION

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-score
Logistic Regression	46.40	48.70	90.77	0.63
K-Nearest Neighbors	84.89	89.15	80.21	0.84
LinearSVC	51.12	51.12	100.00	0.68
BFSVC	86.46	95.48	77.16	0.85
GaussianNB	86.05	94.23	77.46	0.85
Decision Tree	87.51	88.08	87.39	0.88
Random Forest Classifier	84.35	88.28	80.00	0.84

Visualization of anomaly detection results

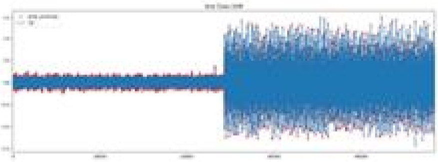

Figure 6.1 One-class support Vector Machine

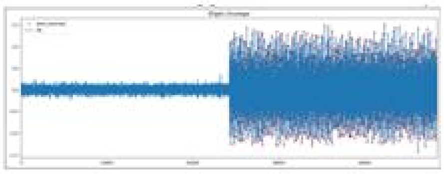

Figure 6.2 Elliptic Envelope

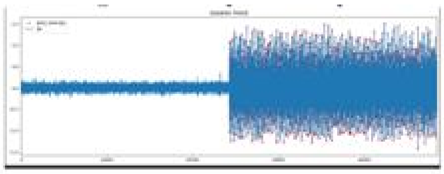

Figure 6.3. Isolation Forest (IF)

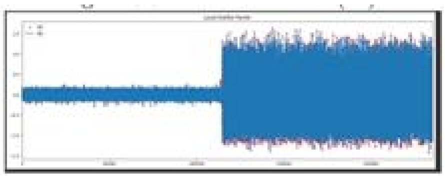

Figure 6.4. Local Outlier Factor (LOF)

Visualization of remaining usage life result

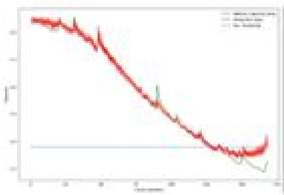

Figure 6.5. Support Vector Regression (SVR)

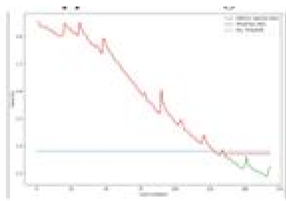

Figure 6.6. LASSO-LARS Cross validation (LassoLarsCV)

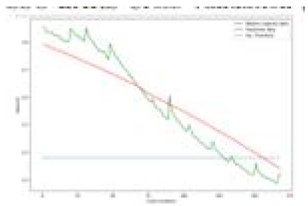

Figure 6.7. Support Vector Regression (SVR)*

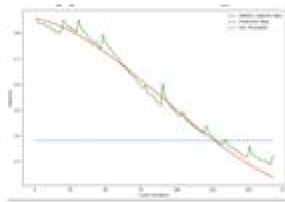

Figure 6.8. Stochastic Gradient Descent Regressor (SGDRRegressor*)

V. Discussion

A. Anomaly Detection

Utilising the CWRU bearing data-set, anomalous behaviour is estimated. In order to use unsupervised learning, the target class labels have been removed from the CWRU data set. is the same dataset used in time to failure detection. This article does not describe the data set. Additionally, in this work, anomalies are found using both novelty detection and outlier identification.While novelty detection relies on semi-supervised learning, outlier detection is based on unsupervised learning.When it comes to novelty detection, the training set only comprises normal data points whereas the testing set includes both anomalous and normal data points. Both the training set and the testing set contain both abnormal and typical data points if an outlier is discovered. Additionally, the covariance-based elliptic envelope (EE), the tree-based isolation forecast (IF), and the density-based (LOF) methods are used to identify outliers. Additionally, kernal-based one-class support vector machines (SVM) are used to identify uniqueness. The training and testing sets are divided 80/20 for EE, IF, and LOF (40/20 for one-class SVM). In contrast to one-class SVM, IF, and LOF, which are nonparametric, EE is parametric as well. All of these models may be applied to data that is multivariate or univariate.

B. Time to Failure Detection

Case Western Reserve University's (CWRU) bearing data set is utilised to forecast TTF. According {b14} to in general,30–40 % of all machine failures are caused by bearing failure.. Bearing failures can be brought on by a number of factors, including excessive loading, poor installation, inadequate lubrication, and more. When it stops production, Tens of thousands of dollars might be lost per hour due to a sudden bearing failure. Therefore, it is important to ascertain "whether a component has a high chance of failure." Both healthy and damaged bearings' vibration data are included in the CWRU data collection .The data collection contains 250 000 data points, of which 50 % are successes and 50 % are failures. The data set does not contain any missing values, but it does have 6941 duplicate entries that have been eliminated. Three features make up the data set. By anticipating the TTF, sudden bearing failure should be prevented in order to prevent unplanned downtime. In this paper, various classification models are utilised to predict TTF. Additionally, "label=0" is treated as baseline data that is normal and "label=1" is treated as failure data in the chosen data set. The data is split 80/20 into a training set and a testing set. All of the features (two in this example) are included in independent/input X, with the exception of fault class, which is specified as continuous dependent/output Y. F1-score is typically used to evaluate classification models. It is between 0 and 1. A classifier's precision and recall are correlated with their numerical average using the F1-score. While accuracy reveals the proportion of times the model was generally accurate. Table 2 also shows that, aside from logistic regression and linear SVC, the majority of the classifier models performed well in terms of F1-score.

C. Estimating Remaining Useful Life with Degradation Model

To calculate the RUL, NASA's lithium-ion battery or Li-ion battery prognostics data set is used. {b7}(Khumpromand and Yodo, 2019). These days, lithium-ion batteries are widely used in everything from electric vehicles to portable electronics. Despite being rechargeable, these batteries progressively lose capacity as a result of repeated charging and discharging (charge cycles). By anticipating the RUL of these batteries, unexpected capacity loss should be prevented in order to prevent unanticipated downtime. In this research, different degradation-based regression models are employed to forecast RUL. According to (Sahaand and Goebel, 2007){b15}, a Li-ion battery is at its end of life (EoL) when its capacity falls to 70% of its initial value.As in Figure 7, expect capacity, which is referred to as continuous dependent/output Y, may be utilised as a condition indicator to determine the number of remaining usable cycles.Following that,To compare performance, an independent/input X that only contains cycles is simulated based on the findings of the feature relevance ranking.SVR is the usage of primary regression model since the NASA data set has a high degree of dimension. SVR has the capacity to automatically regularise the features (by eliminating insignificant characteristics) and prevent over-fitting. It can tackle both linear and non-linear problems. In addition to SVR, the tree-based pipeline optimisation tool (TPOT) - AutoML4 is used in this work. The selection, comparison, and parameter adjustment of ML models are all automated using AutoML.

All the models function well in general, but SGDRegressor* did the best (Figure). The RUL prediction's SVR result is shown in Figure . The predicted line is not smooth because the data set contains a lot of features. The outcome of fitting SVR* with dimensionally reduced data is shown in Figure . Additionally, Figure shows the RUL's LassoLarsCV result. As actual and forecasted capacity values overlap, LassoLarsCV tends to over-fit.

VI. CONCLUSION

Production line failures cost the global manufacturing industry almost 50,000 USD each hour. In a similar vein, it is necessary to maintain the maximum amount of machine and structure availability in order to meet industry 4.0's needs. In order to recognise unusual behaviours and foresee equipment disasters, this article used models based on anomaly detection and prediction. The trials proved that using Pd.M. in conjunction with suggested ML techniques yields positive results. Deep learning must be researched for Pd.M., according to destiny. The supplied Pd.M. Strategies' scalability may also be investigated.

ACKNOWLEDGMENT

I offer our prayers and salutations at the Lotus Feet of the divine mother and the Chancellor of the University, Sri. Mata Amritanandamayi Devi. I would like to express our gratitude to Dr. Prema Nedungadi, Chairperson, Amrita School of Computing for providing us the opportunity to do the project. I am also grateful to Ms. Ani R, Vice Chairperson, Department of Computer Science and Applications for the support she extended throughout.I want to thank our project guide, Krishna S, for his constant support and guidance throughout the project. His valuable suggestions and feedback have been instrumental in shaping this project.

I am also grateful to Ms. Remya Rajesh and Ms.Kavitha K.R., Assistant Professors, for providing us with the necessary resources and infrastructure to carry out this project. We want to thank the Department of Computer Science and Application for allowing us to work on this project. I also would like to extend our sincere gratitude towards the reviewers of the project whose valuable comments and suggestions were instrumental in each phase of the project.Many people, especially our classmates and friends, have made valuable comments and suggestions on this proposal, who gave us an inspiration to improve our project. I thank all of them for their valuable comments.

REFERENCES

[1] {b1} Thyago P. Carvalho,Fabrízzio A. A. M. N. Soares a d, Roberto Vita c, Roberto da P. Francisco b, João P. Basto c, Symone G. S. Alcalá b, ``A systematic literature review of machine learning methods applied to predictive maintenance,'Computers and Industrial Engineering Volume 137, November 2019, 106024.

[2] {b2} Andrea Felicetti; Adriano Mancini; Emanuele Frontoni, Machine Learning approach for Predictive Maintenance in Industry 4.0, 2018 14th IEEE/ASME International Conference on Mechatronic and Embedded Systems and Applications (MESA), 18060887.

[3] {b3} Ameeth Kanawaday, Aditya Sane Machine learning for predictive maintenance of industrial machines using IoT sensor data, 2017 8th IEEE International Conference on Software Engineering and Service Science (ICSESS),17720290.

[4] {b4} Jovani Dalzochio, Rafael Kunst, Jorge Luis Victória Barbosa, Pedro Clarindo da Silva Neto, Edison Pignaton, Carla Schwengber ten Caten, Alex de Lima Teodoro da Penha.Predictive maintenance in the military domain: A systematic review of the literature,03 March 2023

[5] {b5}G. A. Susto, A. Schirru, S. Pampuri, S. McLoone and A. Beghi, "Machine Learning for Predictive Maintenance: A Multiple Classifier Approach," in IEEE Transactions on Industrial Informatics, vol. 11, no. 3, pp. 812-820, June 2015, doi: 10.1109/TII.2014.2349359.

[6] {b6} Hardik A. Gohel, Himanshu Upadhyay, Leonel Lagos, Kevin Cooper, Andrew Sanzetenea, Predictive maintenance architecture development for nuclear infrastructure using machine learning, Nuclear Engineering and Technology,Volume 52, Issue 7,2020,Pages 1436-1442,ISSN 1738-5733,

[7] {b7}Khumpromand, P. and Yodo, N. (2019). A data-driven predictive prognostic model for lithium-ion batteries based on a deep learning algorithm. Energies,12(4):660

[8] {b8}S. Unnikrishnan, S. Gokul Krishna and S. Krishna, "A Framework For Comparing Different Machine Learning Algorithm Models For Intrusion Detection In loT Environment," 2022 IEEE 3rd Global Conference for Advancement in Technology (GCAT), Bangalore, India, 2022, pp. 1-5, doi: 10.1109/GCAT55367.2022.9972026

[9] {b9}Sunil, Sharath and Sonu, P.. (2022). An Effective Approach for Detecting Acute Lymphoblastic Leukemia Using Deep Convolutional Neural Networks. 10.1007/978-981-19-0745-6-3.

[10] {b10}R. Ani, S. Krishna, N. Anju, M. S. Aslam and O. S. Deepa, "Iot based patient monitoring and diagnostic prediction tool using ensemble classifier," 2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI), Udupi, 2017, pp. 1588-1593, doi: 10.1109/ICACCI.2017.8126068.

[11] {b11}Jisha, R., Jyothindranath, A. and Kumary, L. Iot based school bus tracking and arrival time prediction. 2017 International Conference On Advances In Computing, Communications And Informatics (ICACCI). pp. 509-514 (2017)

[12] {b12}Ani, R., Krishna, S., Akhil, H. and Arun, U. An approach towards building an IoT based smart classroom. 2018 International Conference On Advances In Computing, Communications And Informatics (ICACCI). pp. 2098-2102 (2018)

[13] {b13}Raghuprasad, A., Padmanabhan, S., Babu, M. & Binu, P. Security analysis and prevention of attacks on IoT devices. 2020 International Conference On Communication And Signal Processing (ICCSP). pp. 0876-0880 (2020)

[14] {b14}Iftikhar, N.; Lin, Y. and Nordbjerg, F. (2022). Machine Learning based Predictive Maintenance in Manufacturing Industry. In Proceedings of the 3rd International Conference on Innovative Intelligent Industrial Production and Logistics - IN4PL; ISBN 978-989-758-612-5; ISSN 2184-9285, SciTePress, pages 85-93. DOI: 10.5220/0011537300003329