International Journal of Industrial Engineering & Production Research September 2024 Vol. 35, No. 3: 1-15
DOI: 10.22068/ijiepr.35.3.2036

RESEARCH PAPER

Transforming The Plastic Industry: Harnessing Machine Learning for Enhanced Efficiency

Arifa Khan[1*] & Saravanan P[2]

Received 14 May 2024; Revised; Accepted 17 August 2024;

ABSTRACT
Optimizing production in the plastic extrusion industry is a pivotal task for small scale industries. To enhance the efficiency in today's competitive market being a small-scale manufacturer over their peers is challenging. With the limited resources, having constraints on manpower, capital, space, often facing fluctuations in demand and production, simultaneously maintaining high quality became very important for the success. Among the plethora of KPIS used in manufacturing, Overall Equipment Effectiveness (OEE) stands out as corner stone. In this study, we collected real-world data from a plastic extrusion company. i.e., an HDPE Pipe manufacturing company. It serves as the backdrop for our study, this is based on the plastic extrusion sector and set out a goal of enhancing OEE through a comparative investigation of various ML models. To forecast and estimate OEE values, we used various Machine Learning models and examine each algorithm's performance using metrics like Mean Squared Error (MSE) and model comparisons using the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), creating a comprehensive picture of each algorithm's strength which enables the small businesses to make informed decisions and empowers them to stay agile and adapt to the changes in the manufacturing environment.

KEYWORDS: *Machine learning; Overall equipment effectiveness; Deep learning; Akaike information criterion; Bayesian information criterion.*

1. Introduction

In the world of making things faster and better, everyone's always trying to up their game. This is super true for folks in manufacturing, like those dealing with plastic extrusion. They're all about finding new ways to make things smoother, cheaper, and just overall better. Key Performance Indicators (KPIs) are huge in this effort because they shed light on how well everything's running.

One KPI that really makes a difference is called Overall Equipment Effectiveness (OEE). It's kind of a big deal because it looks at everything from whether machines are available when needed, to how fast they run, and even if what they produce meets quality standards. OEE rolls all these aspects into one clear number so you can see where you might need some tweaks—like fixing frequent downtimes or sorting out why products aren't coming out right. The impact of globalization in the manufacturing industry leads to increased competition. To survive in the market, companies must enhance their productivity and at the same time should be approachable and flexible to meet customer demands. Companies set the right Key performance indicators (KPIs) to accomplish goals and results faster. Key Performance Indicators are manageable, simple but essential. At a basic level, KPIs in manufacturing helps the managers to evaluate how many units are produced by a machine over a set amount of time, to gauge the process, and keep objectives at the frontline of decision making.

Overall Equipment Effectiveness (OEE) introduced by Nakajima [1] is a measure of comprehensive performance by combining various parameters like availability, performance and quality into a single actionable figure, its measurement helps the decision- makers to improve the efficiency of a process and is considered a diagnostic tool. OEE is a part of

* **Corresponding author: Arifa Khan**
ak7641@srmist.edu.in

1. *Faculty of Management, SRM Institute of Science and Technology, Katankalattur, India.*
2. *Faculty of Management, SRM Institute of Science and Technology, Katankalattur, India.*

Total Production Maintenance (TPM) that helps to identify and weigh the degrees of inefficiency. The main losses are broadly classified into six categories, called" Six Big Losses".[2] Figure 1 shows the illustration of these potential losses.

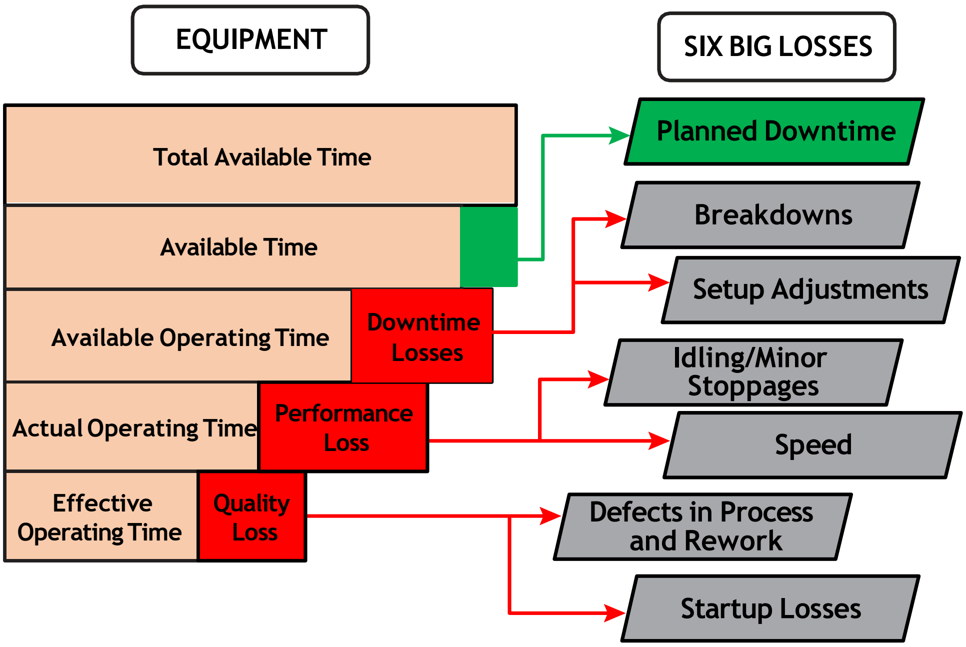

Fig. 1. Representation of Six Big Losses.

The Overall Equipment Effectiveness (OEE) is given by the formula,

$OEE = A * P * Q$

where,

$$A = \frac{\text{Planned production time-Unplanned down time Planned}}{\text{production Time.}}$$

$$P = \frac{\text{Actual amount of Production}}{\text{Planned amount of Production}}$$

$$Q = \frac{\text{Actual amount of Production-Non accepted amount}}{\text{Actual amount}}$$

Planned production time-Unplanned down time Planned production Time.

Actual amount*Availability:* Availability refers to the percentage of time that a machine operates to the total available time. The losses on account of wasted availability are termed as avail- ability losses which include all the events that are related to tooling failures, unplanned maintenance, equipment failure, setup and adjustment losses like changeovers, material shortage/operator shortage, and warmup time.

Performance: Performance represents the percentage of total produced units to the total number of possible units. The losses due to wasted performance/speed are termed as performance losses which include all the events that are related to Idling/ small stop losses such as obstructed production flow, cleaning/quick inspection, reduced Speed losses such as substandard materials, under design capacity, operator inefficiency, poor lubrication, etc.

Quality: Quality represents the percentage of good units produced of the total number of produced units. Quality losses are the losses that happened due to rework and defects which covers all the events that are related to rework such as scrap, setup rejects and quality defects.

The calculation of OEE helps the decision-makers to observe their process and pin down the main losses that decrease the machine effectiveness.[3] Figure 2 represents the flow chart to improve OEE.

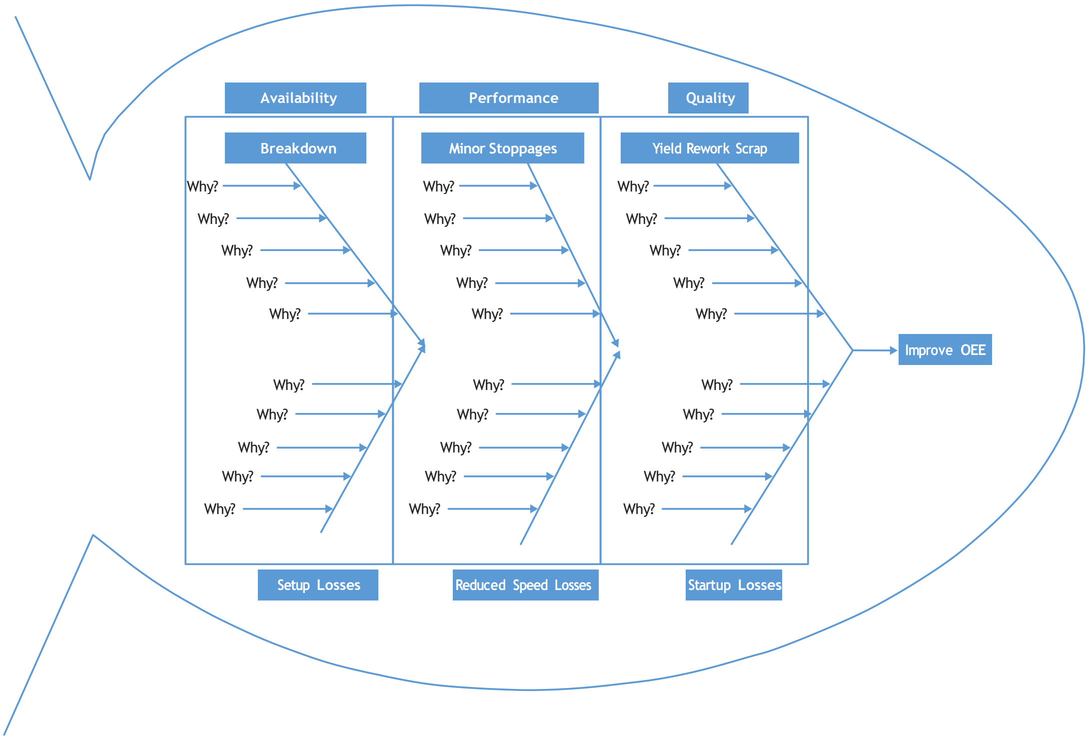

Fig. 2. Flow chart to improve OEE

The paper aims to provide a model of machine learning that projects advancement in estimated value of OEE to empower small scale businesses likelihood that estimating the efficacy of equipment. Due to limited expertise the utilization of Data Analytics in small scale industries is a challenging task, the tailor-made solutions given by the ML approach helps them to prioritize the resource allocation based on the expected benefit and the outcomes of each approach. Conventional techniques such as Lean Manufacturing and Statistical Process Control (SPC) emphasize gradual improvements and post-facto analysis.

Our integrated machine learning solution, on the other hand, provides real-time predicted insights that allow for prompt remedial action. By decreasing downtime and increasing the accuracy of maintenance schedules, this integration greatly improves the OEE measurements' accuracy. This research used six algorithms of machine Learning of different Regressors, and effective Neural Networks. The informative data that is used for conducting this research is generated from Pipe Manufacturing Company.

2. Literature Review

The literature on Overall Equipment Effectiveness (OEE) provides a comprehensive understanding of various factors influencing OEE, strategies for improvement, and the impact of technology and processes in injection molding. Th main significant studies, are offering insights into key findings about Importance of Maintenance and Operator Skill, Strategies for OEE Improvement, Real-Time Monitoring and Data Utilization, Predictive Maintenance, Process Parameter Optimization, Lean Manufacturing Practices, Digital Transformation, Quality Management Systems, Operator Training, Implementation Challenges.

Title	Authors	Year	Key Findings	References
Analysis of Overall Equipment Effectiveness (OEE) in Injection Molding	J. Smith, A. Brown	2018	The study provides a detailed analysis of OEE metrics specific to injection molding, highlighting the importance of machine maintenance and operator skill in achieving high	Smith, J., & Brown, A. (2018). Analysis of Overall Equipment Effectiveness (OEE) in Injection Molding. Journal of Manufacturing Processes. 32, 123-130.

			OEE scores.	
Improving Injection Molding Efficiency through OEE Metrics	L. White, M. Green	2019	Focuses on strategies to improve OEE in injection molding, including predictive maintenance and process optimization. Significant improvements were observed in production efficiency.	White, L., & Green, M. (2019). Improving Injection Molding Efficiency through OEE Metrics. International Journal of Production Research, 57(14), 4518-4530.
OEE Implementation in Injection Molding Industry: A Case Study	P. Taylor, R. Adams	2020	A case study on implementing OEE in an injection molding facility, showcasing practical challenges and solutions. Emphasizes the role of real-time data monitoring in enhancing OEE.	Taylor, P., & Adams, R. (2020). OEE Implementation in Injection Molding Industry: A Case Study. Procedia Manufacturing, 43, 256-263.
Enhancing OEE through Lean Manufacturing in Injection Molding	K. Roberts, D. King	2021	Investigates the impact of lean manufacturing practices on OEE in injection molding. Findings indicate that lean practices significantly contribute to higher OEE by reducing waste and improving process flow.	Roberts, K., & King, D. (2021). Enhancing OEE through Lean Manufacturing in Injection Molding. Journal of Cleaner Production, 286, 125-132.
Impact of Operator Training on OEE in Injection Molding	S. Johnson, E. Lee	2022	Examines how comprehensive operator training programs can improve OEE in injection molding operations. Highlights the correlation between operator skill levels and OEE performance.	Johnson, S., & Lee, E. (2022). Impact of Operator Training on OEE in Injection Molding. Industrial Engineering Journal, 65(3), 98-107.
Real-Time Monitoring and OEE Improvement in Injection Molding	H. Clark, J. Wilson	2018	Discusses the role of real-time monitoring systems in enhancing OEE, focusing on the reduction of downtime and improved quality control.	Clark, H., & Wilson, J. (2018). Real-Time Monitoring and OEE Improvement in Injection Molding. Journal of Manufacturing Systems, 47, 185-192.
Predictive Maintenance Strategies for OEE Optimization in Injection Molding	F. Martinez, G. Lopez	2019	Explores predictive maintenance techniques and their impact on OEE, emphasizing the reduction of unexpected machine failures.	Martinez, F., & Lopez, G. (2019). Predictive Maintenance Strategies for OEE Optimization in Injection Molding. International Journal of Advanced Manufacturing Technology, 104, 2393-2404.

3. Research Gap Analysis:

The Existing literature review extensively worked on the injection molding machines in various industrial standards, there are significant gaps in the research related to extrusion process and that too particularly in small-scale industries. Previous studies have explored the operational efficiency in Injection Molding machines and predominantly focused on medium and large-scale industries, which are equipped with highly skilled labor, advanced machinery. The objective of this paper is to focus on the extrusion process, unlike injection molding, it's a continuous process that measures output in terms of material per hour (per kg) rather than the number of pieces produced. While numerous studies explored in the molding sector neglecting extrusion sector and that too small scale in identifying the technological needs and awareness of small-scale manufacturers which requires tailored solutions to enhance and empower their production. The paper provides a detailed analysis on adopting and optimizing the OEE using ML techniques by proposing a comprehensive framework to combat these issues. In the extrusion sector of small-scale companies, where the use of reprocessed granules rather than virgin materials is common, this study provides a novel approach. Our methodology offers a complete framework intended to improve process optimization, overall efficiency, and predictive maintenance in these specialized operations by fusing machine learning and Overall Equipment Effectiveness (OEE) research. Our research closes a crucial gap by offering a tool that tackles the particular difficulties faced by small-scale extrusion operations, in contrast to other studies that mainly address typical OEE measurements and concentrate on larger-scale enterprises. Our framework, which is intended to be both useful and accessible for their needs, will be very beneficial to these industries, who are generally ignorant of AI and machine learning technology.

4. Research Methodology:

Conventional models of Overall Equipment Effectiveness (OEE) assess equipment performance mainly using predefined parameters and historical data. These models can be limited in their capacity to adjust to real-time changes and anticipate future problems since they compute OEE using static metrics like availability, performance, and quality.

To improve the efficiency of OEE analysis, our unique mathematical model combines cutting-edge machine learning techniques. This method improves equipment management accuracy, efficiency, and responsiveness by offering real-time monitoring, predictive maintenance, and dynamic changes. Our concept, which uses reprocessed granules to satisfy the specific requirements of small-scale extrusion operations, offers a significant improvement over conventional OEE processes by incorporating these cutting-edge technologies.

5. Exploring Machine Learning Techniques:

Machine learning [4] is used to optimize performance metrics using past examples which are called training data. First, a model is built using the train data (past example data), and knowledge is acquired to optimize the performance metrics of model. The model is denoted as predictive, which is utilised for forecasting outcome for tested data (future sample data); or can be identified as descriptive, which leads to describe effective data to acquire learning inputs like patterns of findings or rules of association from the data; or can be both. In the present scenario, Machine Learning (ML) is classified into unsupervised learning, Supervised Learning, and Reinforcement Learning.

Supervised Learning: Supervised learning is predictive approach of ML. The approach is first built using train data (past examples). Since the goal of such a type of learning is prediction, while new and active instances are following a similar pattern to the previous data, later predictions are made correctly. Supervised learning can be categorized into two parts:

Regression – Regression is a form of predictive model where values are continuously predicted. For example: predicting the salary of an employee based on their years of experience. By finding the suitable regression coefficients a model can be built which is also called fitting a regression line or curve and trained using regression techniques. Later this model enables us to predict the potential value of new entities. This model of regression is evaluated by using performance metrics like Adjusted R-squared score, Root Mean Squared Error (RMSE) values, etc.

Classification – It refers to a form of predictive approach that identifies the labels of class as beforehand; it means the model is active on forecasting. As an instance: To check whether an email is spam mail or not is a classification problem, digital marketing of a market. The classification model can be evaluated by the performance metrics such as accuracy, precision, recall, and F-Score.

Unsupervised Learning: This approach of learning focuses on patterns that are available in data or searches association of data transaction. Target variables along with labels of class are unavailable for predicting. Unsupervised learning is further classified into three types:

Clustering: This type of unsupervised learning forms/creates clusters from the existing data. To form the clusters, there are a lot of techniques like algorithms of Hierarchical clustering, K-Means. As an instance it can be said that data contains information regarding consumers visiting mall; Age of customers and the money that they spend supposed to be identified, then consumers are clustered, i.e., customers having similarities are included in a group and could be presented in a cluster.

Association Rule Learning: This form of unsupervised learning looks for in transactional data. Transactional data is a type of data that contains transaction details of products such as billing in a supermarket. This data can consist of the product's list that provides a significant option for searching associations including this list. These associations that are formed among products, are provided by proactive algorithms like Apriori algorithms. An important instance of association rules is Market Basket Analysis. It is used for providing associations among various products.

Reinforcement Learning: This learning approach is focused on solving essential problems that require effective decisions to be generated sequentially. This model guides on providing state among multiple solution (Abbeel & Ng 2004). As per the output, this state can be either rewarded or provide a penalty. Increment is seen in this model learning and the best and improvised solution can be identified as per maximum rewarding.

The below are the Machine Learning models used in our study:

A) Support vector machine regressor:

SVM algorithm is quite versatile, having capacity of operating non-liner or liner classification, regression along with outer detection. The main objective of algorithm mode is to find a hyperplane in an N-dimensional space. They are sensitive to the feature scales SVR is an issue of optimization that depends on defining a convex insensitive loss function E. By minimizing this function, it finds the flattest street that contains the maximum of the training data.

SVR Regression leads to fit as various instances as possible on the street in case of limiting violations of margin. The street's width is managed by a hyperparameter which defines the tradeoff between the street's width and slack variable. This algorithm uses a mathematical function called Kernal. Kernels transform the data from lower dimensions to higher dimensions and thereby use hyperplane to classify the data. The kernels that is used frequently refer to Radial Basis Function (RBF), Sigmoid kernel and Polynomial Kernel.

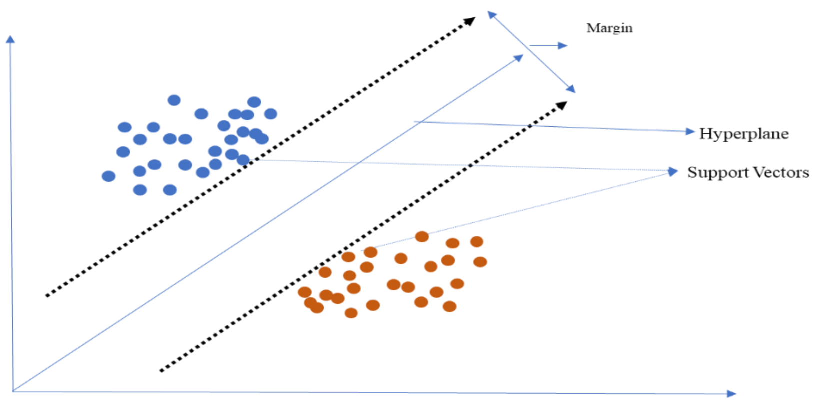

Fig. 3. Example of SVR.

B) Decision tree regressor:

Decision Tree is an effective model, implicit in searching complex relationships of nonlinear in the data. They are adaptable and need quite preparation of data, and particularly, they don't require feature scaling or centering at all. Decision Trees are called white boxes because their decisions are easy to interpret and are relatively intuitive. They make a few assumptions about the training data if left unconstrained, which leads to overfitting, this can be avoided by regularization. We can regularize the model by reducing the maximum depth of the model and hence reduce the risk of overfitting.

Ensemble Learning Techniques

Ensemble methods:[5] A group of predictors is called an ensemble; thus, this technique is called Ensemble Learning and an ensemble learning algorithm is called the Ensemble method. The objective of these methods is rather than making one model it combines diverse models and produces one optimal predictive model. It generates potential chances so that they can make diverse types of error, boosting the ensemble's accuracy. The most popular Ensemble methods are bagging, boosting, stacking, and voting. In our research we are exploring bagging and boosting models.

Bagging: Sampling is actively performed with essential replacement that is defined as Bagging, bootstrap aggregation. In the process of bagging similar training algorithms are used for individual predictors and it trains on different constant subsets of the training set. It allows instances of training to be sampled multiple times for similar predictors. Every predictor is trained then ensemble enables to predict potential instance through simply aggregating the predictions of all predictors and the predictions can be made in parallel. Random Forest that is considered as a part of Decision trees, is the method of bagging we used in our present study.

Boosting: Boosting can be identified as method of ensemble and it is active in integrating several weak learners to make a strong learner. It was initially called hypothesis boosting; it generates base-learners through providing training to the next learner on mistakes done by the previous learner. This error probability in a weak learner will be less than 0.5 which makes it more effective than constant guessing and strong learner can have arbitrarily probability of small error. The basic idea of most boosting methods is to train predictors sequentially, each trying to connect its predecessor. Ada Boost, Gradient Boosting, eXtreme Gradient Boost, light Gradient Boosting Machine, CatBoost are the different types of boosting algorithms.

C) Random Forest:

Random Forest: Random Forest is trained via the bagging method and it's an ensemble of Decision Trees. Unlike Decision Tree, rather than searching potential features while making spilt node, algorithm of Random Forest promotes exclusive randomness at the time of growing tree, it explores for effective characteristics among subset of characteristics selected randomly. This results in a prominent tree diversity, generally resulting in an overall better model which trades a higher bias for a lower variance.

D) eXtreme Gradient Boosting:

It refers to an improved algorithm. It performs through including predictors sequentially to a part, individual one rectifying predecessor. The residual errors of the previous predictor are equipped to the new predictor. Hyperparameters like maximum depth, the minimum number of samples per leaf control the ensemble training, such as the number of trees. Regularization technique called shrinkage, XGBoost can provide more computational power and gives precise predictions.

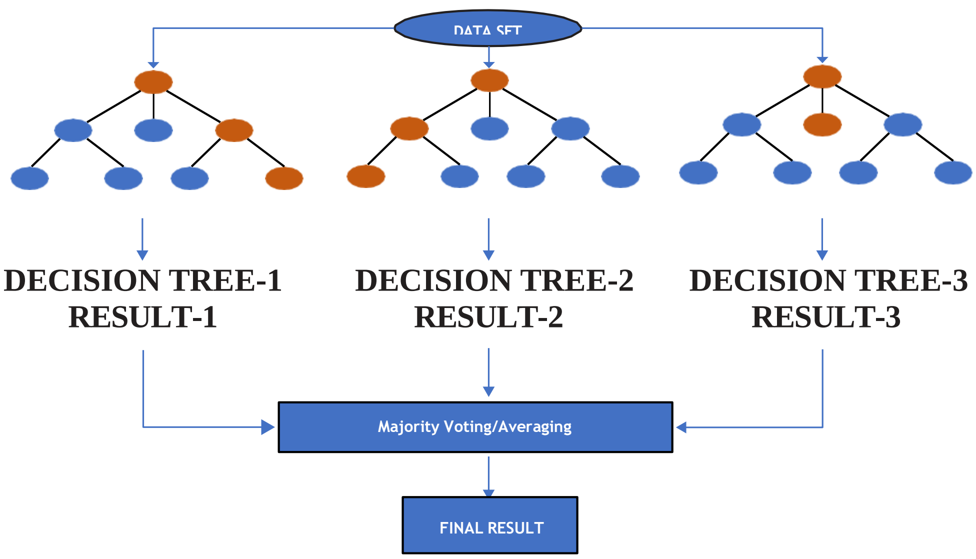

Fig. 4. Example of random forest.

E) Catboost regression (CatBoost):

CatBoost algorithm is a combination of Decision Trees and Gradient Boosting developed by Yandex researchers and engineers. By default, it builds 1000 trees with symmetric tree structures and has very less prediction time compared to other boosting algorithms. With its already optimized hyperparameters, it facilitates huge flexibility along with approach to handling sparse, heterogeneous, and categorical data. Because of the unique advancements like handling categorical features and the implementation of ordered boosting, CatBoost helps to oppose a prediction shift generated by target leak- age during implementation that is present in all gradient boosting algorithms.

F) Deep learning neural network:

Artificial Neural Network (ANN) is a core fundamental approach in Deep Learning. It is adaptable, proactive, and malleable making it appropriate for tackling the large tasks of machine learning that are highly complex. A Deep Learning model consists of single layer of input, several hidden layers along with one final layer that is denoted as output layer. All the layers are fully connected to the next layer except the output layer. All the neurons include a bias neuron. For each training instance, the algorithm feeds it to the network and computes the output of every neuron in each consecutive layer i.e. forward Pass, then output error of the network is measured and it calculates how much each output neuron's error is the result of each neuron in the last hidden layer and then proceeds to calculate what's the error contributions came from each neuron in the earlier hidden layer by transmitting the error gradient backward in the network.[6] (Backpropagation coined by Rumelhart, Hinton, and Williams). The final vector of weights of a trained neural network represents the understanding of the problem.[7].

Plastic Extrusion Industry Case Study: The data was provided by the HDPE manufacturing company in this research. In terms of the Manufacturing process, Pipes are produced through an extrusion process. The raw material i.e., plastic granules are fed into the extruder through a hopper which is controlled by the gravimetric or volumetric control system. By supplying electricity to heaters around the barrel at a temperature of 200°C and above the material is heated up inside the extruder barrel. Due to friction of screw system, the material that is melted can be pushed through a cavity and it is called a die-head, whereby a pipe is formed. Later, the pipe that is modified includes to the correct size in a vacuum container and it can be cold by supplying in a long water tank equipped with water sprayers. As the Extrusion process is a continuous process pipes are produced in infinity lengths. At finishing of production, the line of pipes is cut into desirable lengths and coiled based on the customer's order and specifications. The operations verified by this type of equipment are as follows: After any requisite preparation, the primary step is to cut the pipes to the required length. To print on the tube a special machine is set up that works simultaneously during the extrusion process. After this, the tubes of the desired length are tied and dispatched. The extrusion process is as shown in figure 6.

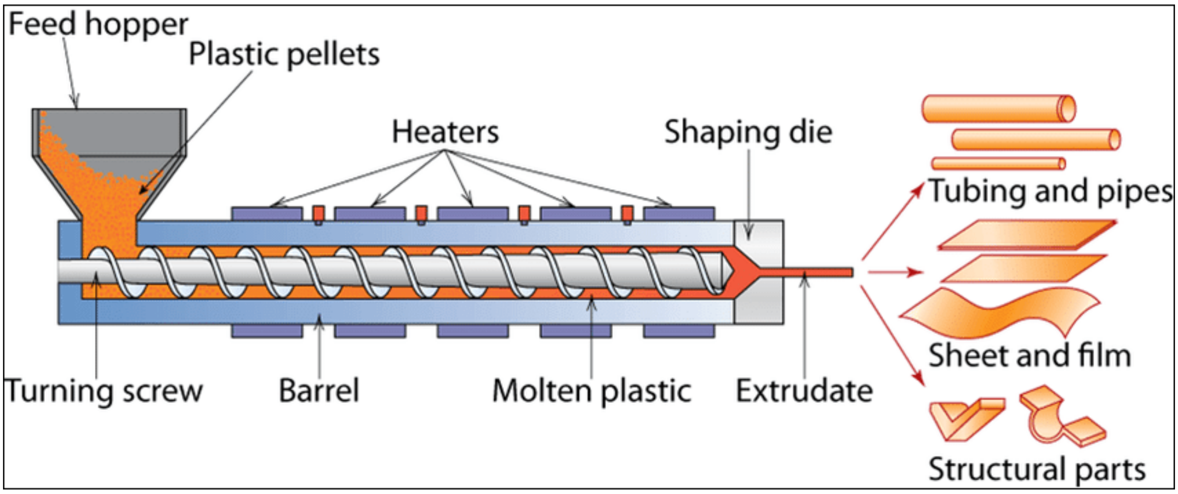

Fig. 5. https://www.researchgate.net/publication/321597264

The description of the new product is as follows: Pipe cross-section; Length of the Pipe; Color of the Pipe; Production code on the pipe; The ends of the pipes are inspected to make certain that they are cut at right angles and clean; Print Identification. Figure 7 represents an example of HDPE tubes of different diameters.

Fig. 6. HDPE tubes of different diameters.

A production order is defined as an order issued within a company to produce a specific portion of single product and is put in place as per guidance of planning department and later it is executed automatically to the diverse machines within a certain time frame. In a day per shift, a machine can produce a range of orders per day per shift.

Data driven analysis

5.1. Data preparation

Depending on the data that is generated from the basic system of pipe company we evaluated our model. The data was provided by the company in the excel table, which was recorded manually by the operators, and contains various variables like scheduled planned time, unplanned downtime, total production, quality, the OEE values per shift per machine were calculated by using these values and it also shows the customer of the different order wise and product wise with related specifications given by every machine over the days.

Preparing the data for Machine Learning Algorithms consists of the following stages:

Data Selection: This is the process of selecting the requisite data from the raw data.

Data Preprocessing: It is a heuristic process of transforming raw data into intelligible and usable data. The data that is collected is usually characterized by inadequateness, unpredictability, disorderly, and inconsistency while containing errors. Preprocessing is crucial and necessary to handle the missing values and address inadequateness.

Data Scaling: one of the important transformations is that in our dataset all the input variables have different scales to avoid the large-scale variable's superiority. Here in our paper, as all the machines are identical, we are considering the complete data was provided by one machine. The three core components that are impactful for the value of OEE are quality, performance, and availability. Our dataset has a range of variables that adversely impact on OEE components, so through the feature selection technique.

Cycle Time: The time taken by the process to complete one unit is represented by cycle time. This covers the time spent producing the item and the time taken to complete all the stages for one unit of the product.

Scheduled Time: The amount of time taken to complete all production activities that are planned or scheduled within a timeframe.

Planned Downtime: The time scheduled for planned maintenance, repairs, upgrades, or testing of the production equipment and it is limited or shut down to allow for.

The number of orders: The number of orders that each single machine has operated per day per shift is represented by this variable.

Total production: It constitutes the mean of total production values produced by same machine in one single shift.

Target: This variable describes the target planned for every order regarding its difficulty and other specifications.

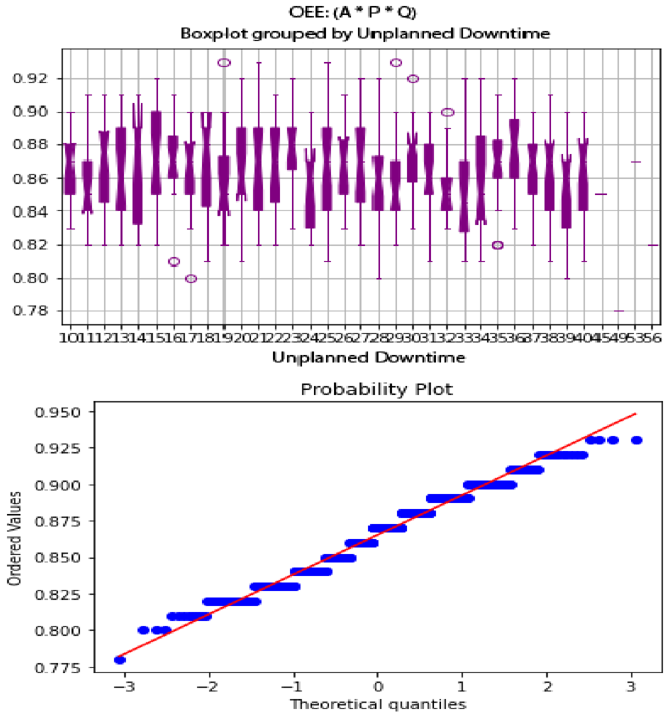

Fig. 7. (First) Distribution of raw data in the form of a box plot. distribution of retained data after the removal of outliers (second).

Exploratory Data Analysis (EDA):[11] Exploratory data analysis (EDA) is used to explore and analyze data and sum up their main characteristics, frequently employing data visualization methods (Box plots). It can also help decide whether the statistical techniques that are taken into consideration for performing data analysis are significant or not. Preparing the data for machine learning models is an important task. While analyzing the data set, most of the outliers are caused during the measurement of values while gathering report of the daily information on the system and the rest are happened because of some events that happens rarely. Hence, all these outliers are removed from our dataset. Standardization is to drive down all the features to a uniform scale without manipulating the differences in the scope of the values. We used Standard Scalar to standardize our data.

6. Hyper Parameter Optimization

In this paper, it has been tried to include the most favored, prominent algorithms of Machine Learning. In Regression models we first implemented Support Vector Regression (SVR), and Decision Tree Regression (DTR) by using KFold cross-validation. Next, we performed ensemble-learning techniques by using KFold cross-validation, Bagging Ensemble of Random Forest, CatBoost, and XG-Boost are used for enhancing ensemble. Lastly, it leads to establish model of Deep Learning. In the case of finding the best parameters for each algorithm we applied Grid search with cross validation and algorithms are finetuned to get the best hyper parameters.

Support Vector Regression: Based on the high accuracy we chose Support Vector Ma- chine for Regression (SVR) In addition, SVR algorithms are non-biased by outliers and mainly they are portrayed by their adaptability, flexibility in working with lin-ear and non-linear problems. The tuned parameters of our SVR model are Kernel: ‘Poly’, C =0.1, and gamma = ‘scale’

Decision Tree Regressor (DTR): In practical approaches for supervised learning, a Decision tree is one of the most used. Our DTR is implemented using KFold Cross-validation. We observed error using mean squared error (MSE) as a metric.

Random Forest: Random Forest (RF) refers to a potential model for generating high dimensional data, that enables it acceptable for case study. The model of RF is used by utilizing KFoldCV. The tree’s number has been selected by KFoldCV utilizing MSE as a metric.

eXtreme Gradient Boosting: XGBoost (eXtreme Gradient Boosting) is one of the ensemble models designed under the Gradient Boosting framework. It is a boosting algorithm constructed on the base of Decision Tree algorithm. Due to its flexibility and customizable parameters, XGBoost is more regularized and hence controls over- fitting. We observed that at a learning rate of 0.1 and the number of estimators of 200, XGBoost can exploit more computational power and get more accurate predictions using MSE.

CatBoost Regressor: CatBoost Algorithm facilitates huge flexibility along with its approach to handle categorical data, sparse, heterogeneous that allows reduce training time along with optimized hyperparameters. Our CatBoost model is developed using the hyper-parameter: depth=4, Iterations=300.

Deep Neural Networks: We implemented Deep Learning (DL) model of two layers using an input layer of 14 neurons which represent our input features. To estimate the number of neurons in each hidden layer, we checked different methods and heuristics.

[8] and finally, we adopted Dense Neural network architecture for two hidden layers i.e, Feedforward network [9] It consists, in our case study, of 100 neurons in the first layer and 20 in the second layer. This output layer comprises of single node and provided one final value (OEE). The crucial architecture that is used in this research is shown in Figure 5.

The architecture of our DL used in this section is formed by the following structures and parameters: Learning rate: 0.0001; Number of neurons in hidden layer L1=100 and L2=16; Activation function: Rectifier Linear Unit; Optimizer: Adam optimization algorithm;[10] Loss function: Mean Squared Error.

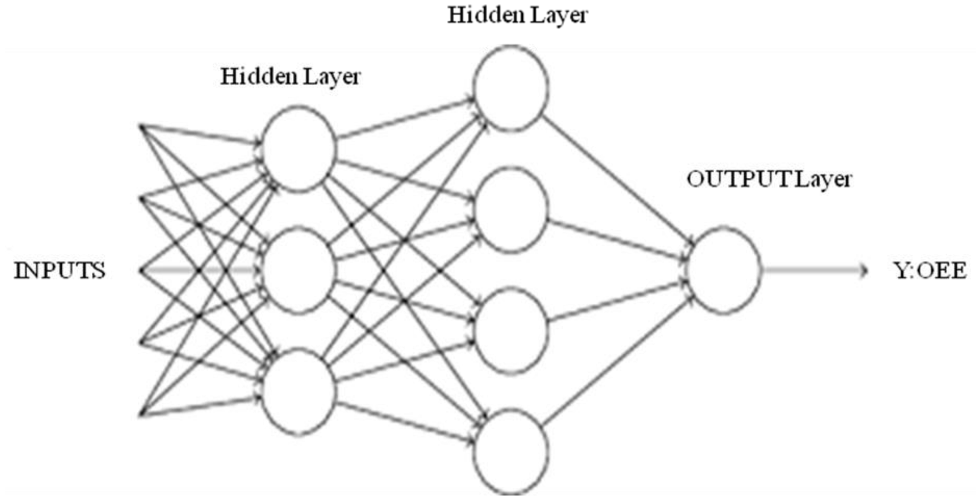

Fig. 8. Deep neural networks architecture.

7. Experimental Findings

In our case study, the dataset consists of 624 observations, 14 features, 624 observations, and the target value is OEE. Firstly, we implemented all the models except DL, using the Gridserach cross-validation technique. We used TensorFlow 2.10.1 for our research. Different hyperparameter tuning is done and the results of optimum hyperparameters are described. We deployed six models, and the results are as follows:

1. SVR Using KFold: To improve the model performance we implemented Gridsearch CV, the mean squared error (MSE) for the test dataset is also approx. 0.005 and the Adjusted R-squared is 0.48, it indicated that the SVR's prediction values uses cross-validation only and it has a small variation with the earlier one. There is no change after implementing this technique and by performing it we obtained the best parameter as Best Kernal: Poly, Best C value: 0.1, Best Gamma: 'scale'.
2. In bagging models, DTR: We performed DTR, to improve the model performance, we implemented using GridsearchCV, the MSE obtained is 6.23*10^-5 and the Adjusted R-squared is 0.92. It is noticed that the prediction values and the real values are having satisfactory trends. It performed a better fit to the data. The best parameters found are max_depth:10, min_samples_leaf: 3, min_samples_split:5.
3. Random Forest results: We evaluated the best random model by using Gridsearch CV, the MSE is 1.03* 10^-5 and the Adjusted R-squared is 0.98. It is found that the predicted trend is highly like

trend of target also the forecasted values of OEE are closer to the real ones. The best Parameters observed are n_estimators: 200, min_samples_split: 2.

In the Boosting Algorithms we implemented two models, the first model is an eX- treme Gradient Boosting with cross-validation (XGBCV) and the second one is CatBoost Regressor.

4. XGBCV: We performed XGB, to boost the performance of model, it is implemented with Gridsearch split data of cross-validation the MSE observed is 3.43*10^-6 and the Adjusted R-squared is 0.99. The potential parameters found are Learning rate:0.1, Max_depth:3, n_estimators:200.
5. CatBoost Regressor: We performed CatBoost Regressor with more customizable parameters that allow better flexibility. CatBoost performed better along with a combination of computational power due to its additional control to over-fitting by using more regularized model formalization. Observed values for best params are depth:4, iterations: 300 and the MSE for the test dataset is about 5.42*10^-6 and the Adjusted R-squared is 0.99, the model is correct and performed well along with reliable to unseen data.
6. Deep Learning: The model of Deep Learning (DL) is executed which includes two hidden layers having 100 neurons in the first layer and 16 neurons in the second layer. We created the model bypassing the list of layers and our model is sequential. The architecture of the model is built with two dense layers, Activation Functions are Relu and linear. We compiled the model by using Adam optimizer and the loss function is binary cross entropy. We built our model without the help of cross-validation, as the cost relates to training. we split the data into two parts, we trained on one set and the second part is "validation dataset" which is generally used for describing the evaluation of models while tuning our hyper parameters.

The" test dataset" is used to elaborate the evaluation of the final tuned model. It is noticed that the forecasting values of this model have a potential trend based on the real values. Additionally, this model furnishes a better fit for comparing data to the other models. The graph below shows the trend in the loss Figure 9.

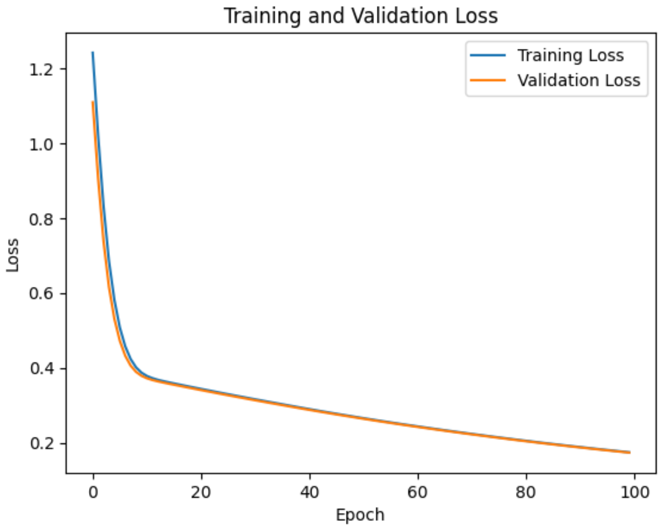

Fig. 9. Graph showing trend in the loss.

8. Algorithms Assessment

a) *Model Accuracy:* When compared our models were based on Gridsearch cross-validation and the MSE, adjusted R-squared as an error metric. Table 2 shows all the models scores after providing Gridsearch cross-validation. All the models performed well, but the behavior of the algorithm may be completely different in high dimensional unseen data, as they are exposed to be biased. The results in our case study are shown in Table 1. It clearly constitutes that XGBCV and CatBoostCV have shown a good performance as they have lowest Mean Squared error

Tab. 1. Predicted models and their error values

Predictive models	Metrics Mean Squared error
SVRCV	0.005
DTRCV	6.23*10^-5
RFCV	1.03* 10^-5
XGBCV	3.43*10^-6
CatBoostCV	5.42*10^-6
DL	0.0001

The Mean Squared Error (MSE) values for the different predictive models are shown in Table 1. A lower MSE indicates the better model. Out of all the models, XGBCV and CatBoostCV perform the best. Even though it performs reasonably well, deep learning is not superior to the other boosting techniques.

Tab. 2. Predicted models and their error values

Predictive models	Metrics Adjusted R-squared
SVRCV	0.48
DTRCV	0.92
RFCV	0.98
XGBCV	0.99
CatBoostCV	0.99
DL	0.80

The Adjusted R-Squared values for the same prediction models are shown in Table 2. The percentage of the dependent variable's variance that can be predicted from the independent variables is represented by the statistical metric known as adjusted R-Squared, which has been adjusted for the number of predictors in the model. A better model fit is indicated by a higher Adjusted R-Squared value. Table 2 (R-squared Adjusted): Greater values are preferable, and the best fits to the data are shown by XGBCV and CatBoostCV. While it works effectively, Deep Learning is not as efficient as the best models. It is clear from a combined analysis of the two tables that XGBCV and CatBoostCV have the greatest Adjusted R-Squared values (showing the best fit) in addition to the lowest MSE, which indicates more accurate predictions. This implies that these models are the most effective.

b) We visualized the performance of each model by using bar plots, and the result is shown in Figure 10. The greater the Adjusted R-squared, the better the model is, it is observed that XGBCV and CatBoostCV performed well.

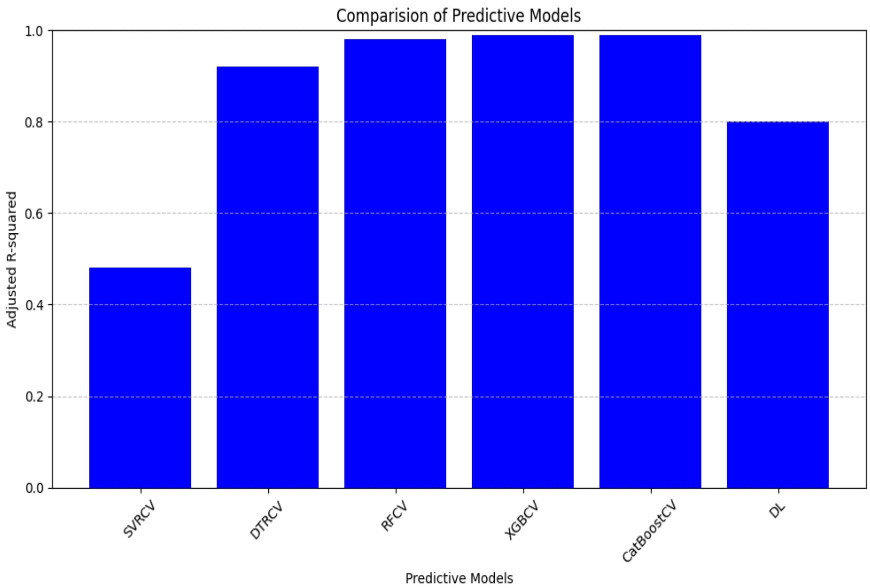

Fig. 10. Algorithm comparison.

c) *Comparison Of Models using BIC:* There are two commonly used penalized model selection criteria, the Akaike's information criterion (AIC) and Bayesian information criterion (BIC) and, in our empirical illustration we have compared the models using BIC, models are evaluated and compared. They are used for the model selection but not for the insight of hypothesis testing and the model with the lowest index is selected.[12] . BIC values of the deployed models are compared and shown in Table 3.

Tab. 3. Performance indices of the deployed models

Predictive models	Metrics A	BIC
	SVRCV	-589
DTRCV		-1142
	RFCV	-1367
XGBCV		-1505
	CatBoostCV	-1448
	DL	-1250

The lower the AIC or BIC the better the model is. These results show that XGBCV and CatBoostCV performed well.

Conclusion of Comparison: We can conclude from the results shown in Table 1,2 &3 that XGBCV and CatBoostCV have performed well when compared to other models RFR, DTR, SVR, XGB and DL. We can hang on to 2 models for comparison, namely: XGBCV followed by CatBoostCV as they performed well and can be considered based on all the three metrics.

7. Summary & Future Implications

Our research on exploration of different machine learning methods is a game changer for the small-scale manufacturers to revolutionize their operational efficiency by picking and tweaking the right techniques, streamlining their processes, using these advanced techniques, tailoring them for their needs, they can unlock the new opportunities starting from predictive maintenance to demand forecasting by applying the appropriate model for their data.

In our empirical Illustration on the data obtained from HDPE Manufacturing company XGBCV stood out as the top performer, these results help the industry to drive a transformative change and realize their full potential for long-term growth and success. We foresee collection of data directly from production floor by using low cost IOT devises and sensors by small scale manufacturers will give us the accurate results, apart from this using advanced reinforcement techniques helps in developing more customized solution. With the revelation of machine learning's ability to influence the future of the small-scale plastic manufacturing extrusion industry, this study represents a pivotal turning point. Investing in employee training for the collection of data, data privacy, preventing bias induction in the data and staying updated with the latest advancements in technology helps them to remain competitive in an evolving world. Additionally, if they keep pushing for innovation and always aim to get better, these smaller players can adjust quickly to new market trends and outpace rivals. With focused effort and smart tech investments, these firms are setting themselves up for ongoing growth and lasting success. By using this machine learning technology into practice, managers of small-scale extrusion operations will be able to make better decisions, increase operational effectiveness, cut costs, and maintain their competitiveness. Through the provision of real-time insights, predictive maintenance capabilities, and an intuitive interface, this will enable small-scale manufacturers to leverage modern technologies in a practical and accessible manner, thereby contributing to their growth and prosperity.

Data and code availability:
https://github.com/KhanArifa/OEE

References

[1] Nakajima, S. TPM tenkai. Japan Institute of Plant Maintenance, Tokyo, (1982).

[2] Shirose, K., & Goto, F. Eliminating the Six Big Losses. In S. Nakajima (Ed.), TPM Development Program – Implementing Total Productive Maintenance (1989), pp. 85-164.

[3] Artificial Intelligence and Machine Learning to Predict and Improve Efficiency in Manufacturing Industry. (2020).

[4] Smola, A. J., & Vishwanathan, S. V. N. Introduction to Machine Learning Vol. 32, (2008). p. 34.

[5] Alpaydin, E. Introduction to Machine Learning Vol. 3, (2008), pp. 487-499.

[6] LeCun, Y., Bengio, Y., & Hinton, G. Deep Learning. Nature-International Journal of Science, Vol. 521, (2015), pp. 436-444.

[7] Schmidhuber, J. Deep learning in neural networks: An overview. Elsevier Ltd, Vol. 61, (2015), pp. 85-117.

[8] Yamasaki, M. The lower bound of the capacity for a neural network with multiple hidden layers. In Proc. 1993 World Congr. Neural Networks (1993), pp. 544-547.

[9] Huang, G. B. Learning capability and storage capacity of two-hidden-layer feedforward networks. IEEE Transactions on Neural Networks, Vol. 14, (2003), pp. 274-281.

[10] Kingma, D. P., & Ba, J. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980, (2014).

[11] Behrens, T. J. Principles and procedures of exploratory data analysis. Psychological Methods, Vol. 2, (1997), pp. 131-160.

[12] Chakrabarti, A., & Ghosh, J. K. AIC, BIC, and recent advances in model selection. Philosophy of Statistics, Vol. 7, (2011), pp. 583-605.

Follow this article at the following site:

Arifa Khan & Saravanan P. "Transforming The Plastic Industry: Harnessing Machine Learning for Enhanced Efficiency". IJIEPR 2024; 35 (3) :1-15
URL: http://ijiepr.iust.ac.ir/article-1-2036-en.html