Received: February 15, 2021. Revised: April 5, 2021.

International Journal of
Intelligent Engineering & Systems

http://www.inass.org/

IoT based Machine Learning Automation Algorithm for Controlling the Industrial Loads

Bhagirathi Bai Vittal Rao[1]* **Hanume Gowda Chandrakanth[2]**

[1]*Acharya Institute of Technology, Mechatronics Engineering, India*
[2]*Sambhram Institute of Technology, Electronics and Communication, India*
* Corresponding author's Email: bhagirathistelkar@gmail.com

Abstract: Presently, industrial automation has become a popular field because of its various advantages includes higher production rates, more efficient use of materials, better product quality, improved safety, and requires minimum labours. This is achieved by utilizing Local Networking Standards (LNSs), remotely monitoring and controlling industrial devices by utilizing Raspberry Pi and Embedded Web Server (EWS) technology. This proposed research provides an idea of utilizing Internet of Things (IoT) for monitoring and controlling the automation process through Wi-Fi or wireless medium by utilizing Raspberry pi as a server system. Additionally, the prediction and error detection utilized in the machine learning process for this Improved Random Forest (IRF) method. The proposed IRF uses Out of Bag (OoB) bagging estimation technique to randomly select sub-datasets for overcoming the optimization problem. The OoB estimation is the technique used to find prediction error in IRF as every samples are not used when each tree in IRF is trained. So for all those bags unused samples estimate the prediction error for a particular bag in prediction process. IRF recursively partition the data into categories and the derived RF model and maps the results. The proposed IRF method overcomes the overfitting problem, receives the command and applies action according to it. The IRF method uses only eight devices to control and monitor more than thirty devices and this entire process is represented as IoT based Machine Learning (ML) Automation Algorithm. The experimental results show that the IRF method provides better outcomes in terms of accuracy, precision, F-measure, and recall. The RMSE values obtained for the proposed IRF model shows 0.0386 lesser error values when compared with the existing models that had achieved RMSE as 3.607 for Long Short Term Memory- Recurrent Neural Network (LSTM-RNN), Artificial Neural Network (ANN) as 1.226 and Fuzzy Gain Scheduling (FGS) of 0.4247.

Keywords: Error detection, Internet of things, Industrial automation, Machine learning Algorithm, Raspberry pi.

1. Introduction

The industrial development has a significant role in enhancing all the country's standard [1]. The application of advanced technology in manufacturing is often becoming widespread, such as Internet of Things (IoT), Wireless Sensor Networks (WSNs), Cloud Computing, Advanced Sensing Technology, Service Oriented Technology etc. Currently, the manufacturing industry is going through massive revolutionary changes and smart manufacturing technology is fascinated much attention [2]. The extensive development of industries has a significant role in enhancing the country standard and it improves economic stability, reserves foreign exchange, improves resource utilization, enhances the agriculture growth, etc. [3].

Recent advances in technology have the capacity to transform the way of interactions among humans and objects. It also shows the way to enhance the experience of products for service provider and service user [4-5]. The IoT and WSNs provides the machines and computer self-working process to monitor and control the industrial automation without human intervention. It includes many applications such as 1) Increases quality of product and productivity. 2) Reduces routine check of the product. 3) Improves the standard of safety. 4) Decreases the need of labours, time and reduces the operational cost.

But, for real-time industry automation, WSNs works on wired power supply which is not usable because it is not portable and causes tripping hazards. The WSNs meets much complexity in detection of faults, reinstallation and maintenance which results in degradation of network performance [6]. Due to the fluctuating and hardly controllable nature of renewable energy sources, the control mechanism is no longer sufficient. The industrial sector is the largest electricity consumer worldwide [1] and automation of industries has taken the advantages from IoT method to enhance its reliability, efficiency, availability and it is termed as Industry-4.0 [7,8]. However, in industries the forecasting errors have a considerable influence on Short-term load forecasting is also of great significance for dynamic state estimation and load dispatching [2].

So, IoT is most used domain in IT landscape due to its high heterogeneous integrated technology and high rise of IoT empowered applications in various area for example: Smart cities, E-health system, Smart manufacturing system and Intelligent transportation system [9, 10]. The growth of these systems is pressuring the industries to deliver cheaper and personalized products. The companies are forced to deliver smaller and much diverse batches by making the level of organization, communication and integration more complex. So, data is a most required business asset that comes polluted in many forms such as completely or incompletely lacking, includes many errors, will not be properly tracked and is unstructured [11].

To overcome the aforementioned issues, this paper highlights the designs and implementation of industrial automation system with the help of Raspberry Pi as a gateway programmed using Python language. The Raspberry pi is the sensor module used to control and monitor the different parameters of an industrial plant and energy management. In this paper, the experimental results obtained demonstrate the usefulness of IRF method in terms of accuracy. This research considered few machine-learning algorithms such as random forest, Support Vector Machine, K-means, Naïve Bayes, K- nearest neighbour and decision tree for reference according to their basic idea of research. The information on these algorithms is utilized for data collection and comparison. The novelty of the proposed method is that the IRF method algorithm overcomes the overfitting problem, receives the command and applies action according to it. The machine is responsible for the generation of random variables in terms of motors, robots, sensors, machine data, transducers, and other different parameters. Here, the Electric motors are used classified by considerations such as power source type, internal construction, application and type of motion output. Secondly the Robots has developed and integrated numerous robotic load or unload applications automated operation to determine feasibility and suitability of a robotic solution for automatic load operations. Thirdly, the machine data is obtained from machine-generated data which is the digital information created automatically by the activities and operations of networked devices for storing the load data. Fourthly, the transducers are often employed at the boundaries of automation, measurement, and control systems from where the signals are converted to and from other physical quantities which are used for load automation. The IRF method uses only eight devices to control and monitor more than thirty devices. Out of Bag (OoB) estimation is the technique used to find prediction error in IRF. Every samples are not used when each tree in IRF is trained, so for all those bags unused samples can be used to estimate the prediction error for a particular bag. The OoB error rate is obtained by averaging predicted error from each bag. This entire process is represented as IoT based Machine Learning Automation Algorithm IRF method. The scientific contribution of the proposed method is it helps business to improve the safety, saves time, quality of production will increase, reduces overall cost, and monitoring is not necessary. These benefits lead company to reach higher production, higher efficiency and more profit.

This paper is organized as follows: Section 2 provides a brief description of related works. Section 3 discussed on IRF method. Section 4 discussed results and discussion and Section 5 made the conclusion of this paper.

2. Literature survey

Runhai Jiao [7] developed Short-Term Non-Residential Load Forecasting Based on Multiple Sequences using the LSTM-RNN. The developed Non-residential load forecasting framework was developed based on the LSTM RNN having multiple sequences was utilized. However, the developed model failed to consider the external factors such as load forecasting, economic orientations which lowered the performance of the model.

Bastian Dietrich [10] developed a Machine learning based very short-term load forecasting of machine tools. The model used a blueprint to develop load forecasting models for machines in order to gain productions using the historic load profile and various machine for data processing. However, the data which was available showed optimization problems which increased energy cost as well as energy flexibility using the model.

Mani Dheeraj Mudaliar [14] developed an energy efficient monitoring system to monitor and analyse the energy usage as a pre step for the conservation of energy in industry. The developed method was based on RaspberryPi with the existing energy meters of switch gear industry. The developed energy conservation method is cost sensitive and consumes lesser power when operated by industry. In developed method data cannot be accessed when energy meters were used and graphical output was not available for the stored values.

Fotis Foukalas [15] developed a Cognitive-IoT-Platform which was utilized for the industrial fog computing application. The developed platform provides distributed intelligent that fog computing model needs to apply for the edge. The developed method was based on a predictive maintenance use case which provides services to IoT application. The developed method reduced the system cost. The developed platform was not provided the dynamic configuration of the predictive maintenance service.

Luis I. Minchala [16] developed control method for operating tank levels by combining the PID algorithm and Fuzzy Gain Scheduling (FGS) approach. The fault identification method was developed for the system using Fault Detection and Diagnosis Module (FDD) implemented with Extended Kalman Filter (EKF). The performance of the developed control method was good. The developed method was not suitable for aggressive industrial condition like dust and noise environment.

Miss Bhagyashri Tambe [17] developed the approach to assign dynamic Internet Protocol (IP) to board and tested for various dynamic IPs. The dynamic IP is used for embedded board for enabling Internet of Things (IoT). The developed method provides comfort and convenience to users in industries. The web page in the developed method shows dynamic response to sensor value which need to be enhanced by utilizing Wi-Fi module.

Shabir Ahmad [18] developed a Resource Aware (RA) method to reduce the higher period of input tasks on the basis of devices profile. The RA method allowed tasks to admit in every possible time, addressed the issue of Real Time IoT (RTIoT) systems and Industry4.0 for patient monitoring approach installed in an emergency ward. The developed resource aware method reduced the usage of power. The developed method does not consider scalability factor which is required in realistic IoT application.

3. Proposed industrial automation system

Normally, all industrial automation devices require a protocol to improve the system performances. But sometimes the industrial automated devices will not provide the accuracy in the prediction process. To overcome the above-mentioned issues, the industrial automation system is constructed by IRF method. The proposed industrial automation method is used to create a machine or system which works perfectly by training them for delivering threshold values. The values measured are used to perform the training process for the respective machines.

Fig. 1 shows the block diagram of the IRF method. The Chabot is utilized as a user interface (UI) or User Experience (UX) which is developed by utilizing Google’s dialog flow. This implements Natural Language Programming (NLP) to define and process the requests and control the signals from the user. The machine learning part of this research is developed through the K-means or random forest algorithm to detect patterns in the data. The machine is responsible for the generation of random variables, which consists of motors, robots, sensors, machine data, transducers, and other different parameters. Hadoop Distributed File System (HDFS) is used to store user data and Hadoop (H-base) database is used. The various commands are used to control the robots, motors, sensors, supply chain and monitors the automated operation. This database collects and stored over for 3months. The use of Human-Machine Interference (HMI) provides the best monitoring and controlling features of different operations. Although the entire process has operated to control the specific operations at any point in time during the process which are given to the user through the HMI. The Master Control Unit (MCU) is utilized to ensure the independent operations of each different approaches which provides the solution for failure components. Finally, this application is used in different sectors like motors, sensors, robots and supply chain management for monitoring and controlling the automated operations. The information on aforementioned algorithms are used for data collection and comparison. Some of the algorithms discussed are as follows: Random Forest, Support Vector Machine (SVM), Naive Bayes, K-Nearest Neighbour (KNN) and Decision Tree.

Primarily the programming language is utilized in making algorithms as the machine can understand the language. The brief explanation of the flow chart is given below,

Commands provided by the user is pre-processed

Received: February 15, 2021. Revised: April 5, 2021.

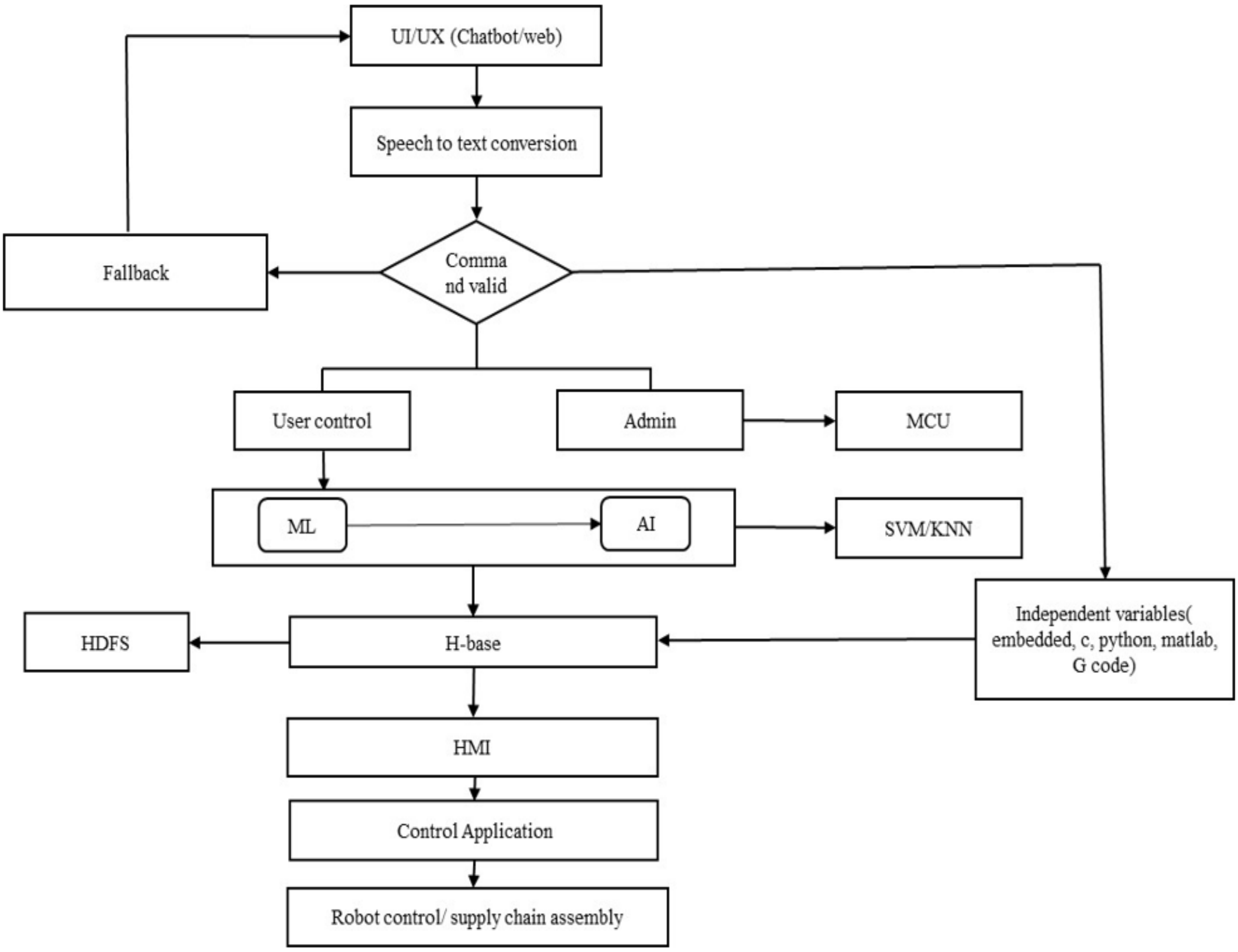


Figure. 1 Block diagram of IRF method

- If the command is utilized before, then the text to integer conversion will be easier for every word after pre-processing.
- If the command is not previously utilized, next there requires the machine to understand the command.
- The NLP is for the machine to understand the command. Every sentence is divided into words and then to the integer.
- After pattern matching, there requires training the data. If the error is identified, then it provides the detail of the trained data. Based on the error type, the trained data value gets changed.
- If the error is resolved, the value is displayed in the output section. Otherwise, it undergoes the machine learning algorithm.

In order to evaluate the classification performance of machine learning algorithms the entire classification accuracy index is taken to analyse the classification performance of the model and reduce the uncertainty effect of chosen training samples on the results. The 108 sets of the data are selected from all the data as input samples to make the prediction model.

RF is a supervised learning algorithm that is having the ability to perform classification and regression tasks by using multiple decision trees. Each and every tree present in RF is constructed on the basis of bootstrap distinct samples which is the original data. Each feature obtained based on the decision is calculated using the below Eq. (1).

$$f_{(i)} = \frac{\sum_{j\epsilon\, node\ splits\ on\ feature\ i} n_j}{\sum_{k\epsilon\, all\ nodes} n_k} \quad (1)$$

$f_{(i)}$ is the importance of each feature i
n_j is the importance of node j

The decision tree is combined forms RF and each of the object of the forest is known as tree which is classified based on decision making. For each of the class decision is performed at each object which is considered as a vote. The maximum of votes received will be selected by the forest based on the objects considered. The RF classifies the data samples or the elements into i and j using the Eq. (2).

$$f_{(i)} = \frac{\sum_{j \in all\ trees} norm\ f_{(i,j)}}{T} \tag{2}$$

$f_{(i)}$ is importance of feature i that are calculated from all trees in the Random Forest model. $norm\ f_{(i,j)}$ is the indicator function which considers random variable takes as either 1 when the event happens and 0 when not happens.

Where

$$norm\ f_{(i)} = \frac{f_{(i)}}{\sum_{j \in all\ features} f_{(j)}} \tag{3}$$

However, to improve the memory requirements and speed of the model an Improvised Random Forest is introduced in the proposed research. In the existing RF during training process the weight distributions are globally tracked and are assigned for training the sample. The weights are updated iteratively and thus the data becomes higher thereby shows difficulty during the sample classification. The proposed Improvised RF uses hamming distance and the Intrinsic Random Forest Similarity as metrics for finding neighbourhood for local performance estimates with dynamic integration. If the two instances are close together obtained based on hamming distance that shows relatively smaller to if ISM is near to the classification boundary which is calculated by using Eq. (4).

$$D_{f_i} = \sum_{i=1}^{k} |x_i - y_i| \tag{4}$$

$|x_i - y_i|$ is the distance between the two instances of samples

k is the number of trees

D_{f_i} is the hamming distance for each feature i.

The intrinsic random forest similarity demands additional space for saving information about n training instances in the leaves of the K trees.

In order to calculate the weight for model i in dynamic integration for a new instance x is calucalted by using the following Eq. (5).

$$w_i(x) = \frac{\sum_{j=1}^{k} \left(x_j \epsilon OoB_i\right) \cdot \sigma\left(x, x_j\right) \cdot mr_i\left(x_j\right)}{\sum_{j=1}^{k} I\left(x_j \epsilon OoB_i\right) \cdot \sigma\left(x, x_j\right)} \tag{5}$$

From the Eq. (5) k is the size of the neighbourhood, OoB is the set of outof-bag instances for model i, $I()$ is an indicator function, $\sigma\left(x, x_j\right)$ is a distance-based relevance coefficient and $mr_i(x_j)$ is the margin of model i on j^{th} nearest neighbour of x.

Margin is defined as usual for a classifier with crisp outputs where 1 is for a correct prediction, and –1 for a wrong one prediction using Eq. (6).

$$mr_i(X) = \begin{Bmatrix} 1 & , h_i(x) = y(x) \\ -1, & h_i(x) \neq y(x) \end{Bmatrix} \tag{6}$$

The weight represents the expected margin of model i at instance x. The weights are normalized to be nonnegative and to sum to one in order to apply them in (locally) weighted voting in dynamic integration thereby minimizes the error values.

3.1 Flow chart and control flow

This section briefly explained the flow chart and control flow for the IRF algorithm.

3.1.1. Flow chart for the IRF algorithm

The flow chart of motor and transformer is shown in Fig. 2. It is consisting of seven steps that are explained properly as follows.

Step 1: NLP Receives the commands from a user through Chatbot.

Step 2: The Commands are in the form of speech input.

Step 3: After receiving the command, NLP converts speech to text.

Step 4: when the commands are converted, the Classification of type of operation are carried out.

Step 5: For the classified commands, simulation training of a machine learning algorithm is done for further processing.

Step 6: After simulation, input values are tested with threshold values.

Step 7: When the input values are tested, Auto-correction of input values are carried out.

3.1.2. Control flow of the IRF algorithm

Fig. 3 shows the control flow of the IRF algorithm which constitutes of following blocks such as pre-processing the commands, conversion of commands

Received: February 15, 2021. Revised: April 5, 2021.

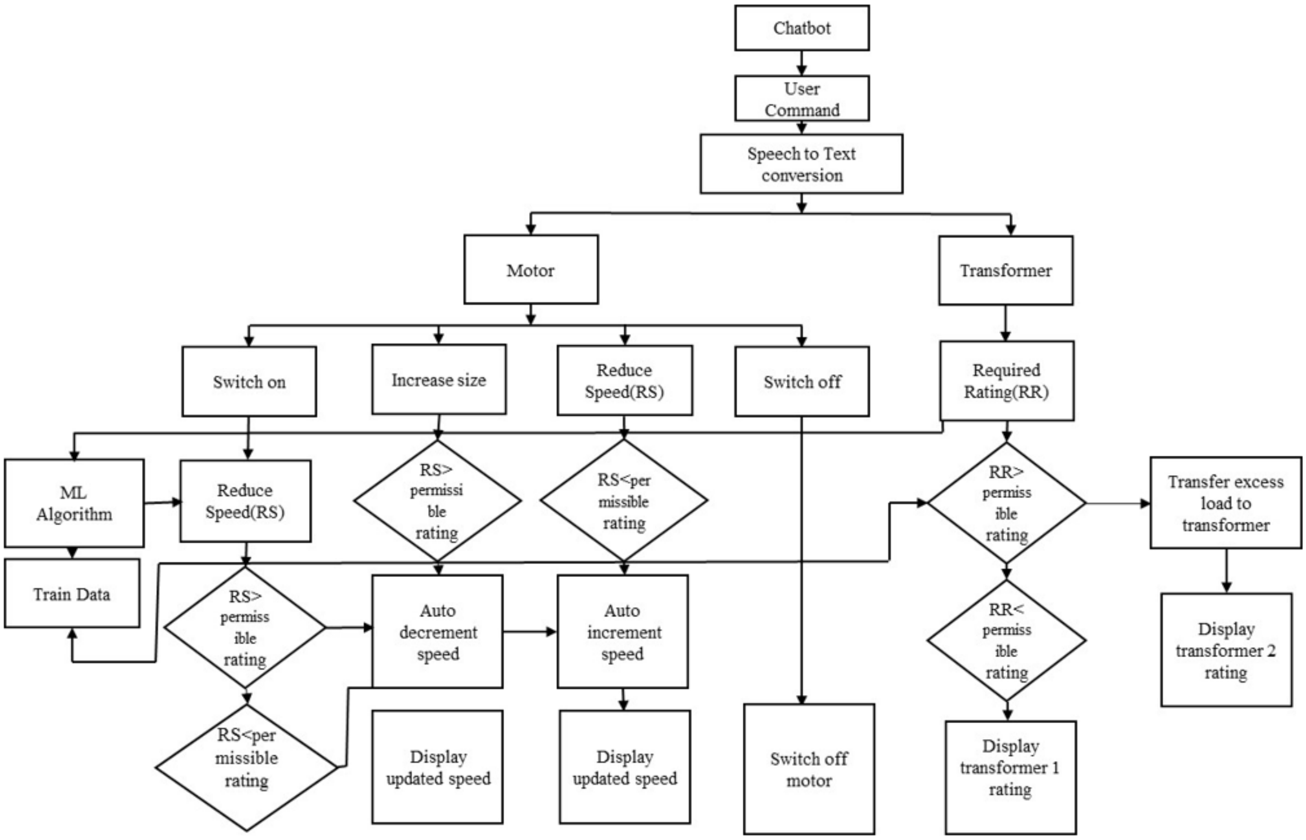


Figure. 2 Flowchart of motor and transformer

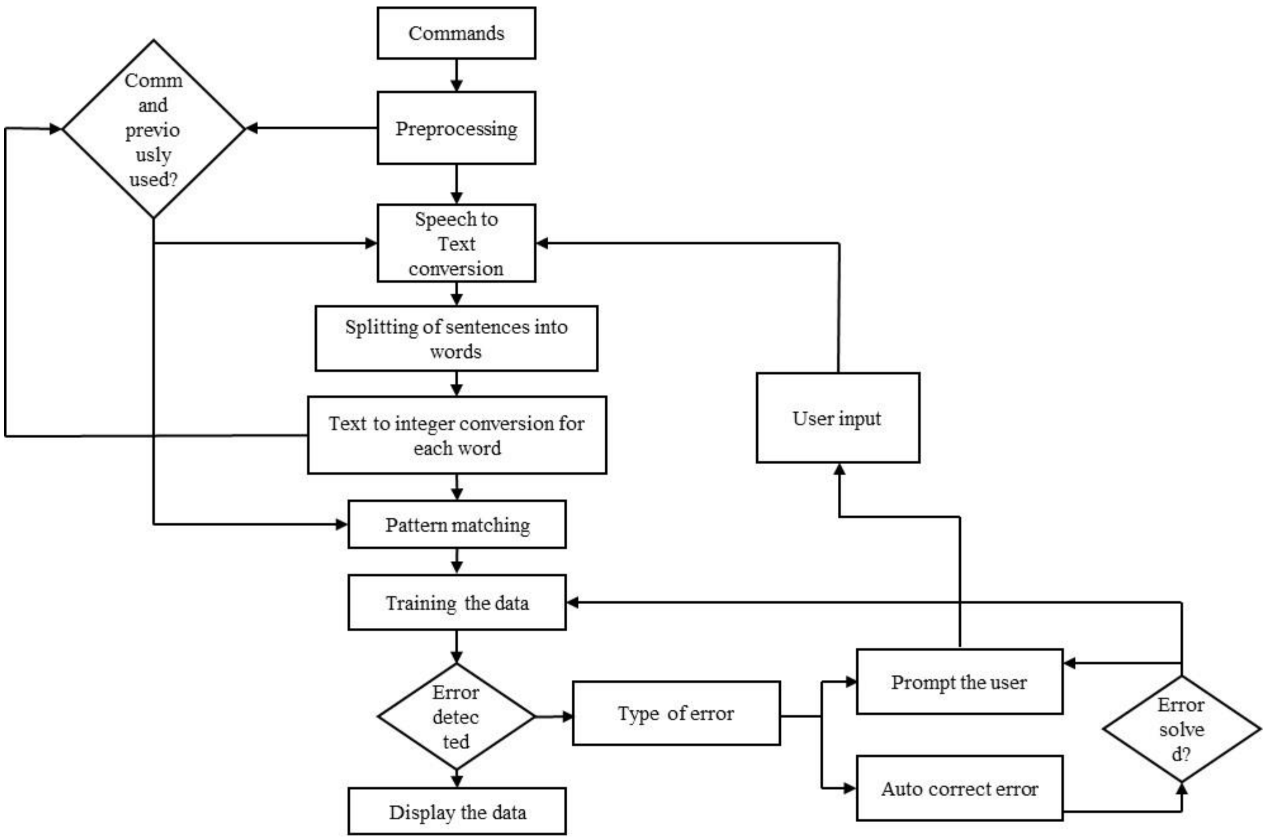


Figure. 3 Control flow of the IRF algorithm

Received: February 15, 2021. Revised: April 5, 2021.

and the conversion of speech to text. Once after conversion pattern matching is performed if in case any error is present the type of error present will be detected which are briefly explained as follows. The steps followed are shown below

Step 1: Initially, pre-processing of the commands are received. If the input command is not valid, then it suggest the user deliver a valid input.

Step 2: After receiving the command, conversion of speech to text through NLP takes place.

Step 3: After converting from speech to text, each word is assigned to an integer value. These values are further processed for pattern recognition.

Step 4: After assigning the integer value, it identifies the error and performs auto-correction.

Step 5: After auto-correction, the error correction is performed and further processing are continued.

Note:

The entire process was done through remote servers. The jupyter notebook which was used to run the python code for the machine learning algorithm was run on localhost:8080/. Hence all the voice data was converted to csv files and given as data to Jupyter notebook.

3.2 Monitoring and controlling

Raspberry pi based sensor module used for monitoring and controlling (Receiving) in the IRF method which consists of two blocks such as: controlling and monitoring.

3.2.1. Monitoring

The Monitoring IRF method consists of five sensors such as Temperature, Voltage, Current, Oil Level, Vibration and speed. The temperature sensor helps to monitor the temperature of the electrical device. The Voltmeter and current meter sensors find the Current and voltage level of the electrical devices. The sensor values updated in the web-page, which are obtained from the transmitting section based on voice commands.

3.2.2. Controlling

In the controlling block, the electrical devices are connected with the Raspberry pi. Depend upon the transmitting section command, the switching operation of the electrical devices is controlled.

4. Result and discussion

In this section, the experimental results of the proposed method are explained and the hardware design of the proposed IRF method is divided into two types. The major use of this method is to receive the commands and apply action accordingly. The initial stage Transmitting Unit (TU) consists of Raspberry Pi -3 with Raspbian Jessie Operating. To do a fair and better performance evaluation of the machine learning algorithms, this research considering simulation error. The present research work uses the real time data from small scale industries for the evaluation of results. The real-time data included in the research includes motors, robots, sensors, machine data, transducers parameters for simulation. The efficiency of these algorithms is assessed by utilizing the following terms such as Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Relative Absolute Error (RAE). The MAE is the average over the test sample of absolute differences between prediction and actual observation where all individual differences have equal weight. The RMSE is the square root of the average of the squared differences between prediction and actual observation. Operating System (OS) with Python programming includes Microphone of 3.5 mm HYPO clip MIC. Next, the proposed IRF method for controls and monitor the units, which consists of similar Raspberry Pi 3 with Raspbian Jessie OS including Python programming. The controlling and monitoring functions are processed by using the webpage update.

Accuracy: Accuracy is one of the important parameters for computing classification models. Generally, the accuracy is evaluated by the fraction of predictions. The definition of accuracy is given in Eq. (7),

$$Accuracy\ (\%) = \frac{Number\ of\ correct\ prediction}{Total\ number\ of\ prediction} \times 100 \quad (7)$$

Sensitivity:

The sensitivity parameter also called as True Positive Rate (TPR), computes the proportion of actual positives and that are appropriately identified the affected signal. The sensitivity is expressed in Eq. (8).

$$Sensitivity = \frac{TP}{TP + TN} \times 100 \quad (8)$$

False Positive Rate (FPR): This is computed as the ratio between the number of negative events which wrongly classified as the positive and total number of actual negative functions. The parameter of FPR is expressed in Eq. (9).

$$FPR = \frac{FP}{FP + TN} \times 100 \qquad (9)$$

Precision: Precision is the ratio of correctly predicted positive observation of the total predicted positive observation. The higher precision value indicates that the evaluation of the proposed method is consistent. Precision is calculated by using Eq. (10).

$$Precision = \frac{TP}{TP + FP} \times 100 \qquad (10)$$

F-measure: F-Measure is the measure of test accuracy and defined as a weighted harmonic mean of the precision and recalls the test. The F-Measure will always be nearer to the smaller value of Precision or Recall. The F-measure is mathematically expressed in Eq. (11).

$$F - mesure = \frac{2 \times recall \times precision}{Recall + Precision} \qquad (11)$$

- **Mean absolute arror**

Mean Absolute Error (MAE) is the difference between prediction and actual observation where all individual differences have equal weight. The MAE is expressed in the Eq. 12.

$$MAE = \frac{1}{n}\sum_{j=1}^{n} |y_j - \hat{y}_j| \; X(t+1) = X_p(t) - A.D \qquad (12)$$

- **Mean square error**

Mean Square Error (MSE) is the measure of the average square errors obtained when the average squared difference between the actual and the estimated value is evaluated. The MSE is expressed as shown in the Eq. (13).

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \widehat{Y_i})^2 \qquad (13)$$

- **Root-mean-square error**

The RMSE is defined as the square root of the average squared differences between the actual and prediction observations. The RMSE is expressed as shown in the Eq. (14).

Table 1. Performance evaluation of the IRF algorithm

Proposed IRF algorithm	Percentage(%)
Sensitivity	98.3
FPR	1.7
Precision	99
F-score	99
Datasets	600

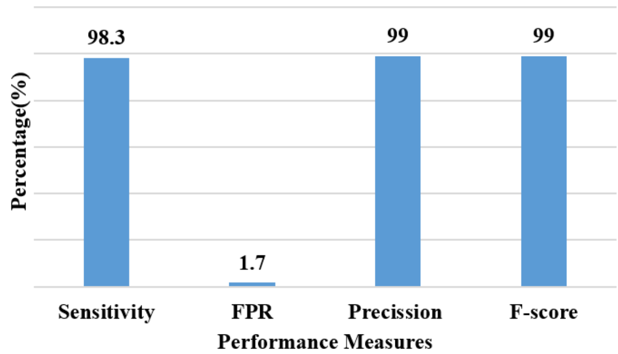


Figure. 4 Performance evaluation graph of IRF Algorithm

Table 2. Training and simulation error of the IRF algorithm

Evaluation Criteria	IRF Algorithm
MAE	0.0254
RMSE	0.0386
RAE(%)	6.22
Datasets	600

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_i - \widehat{Y_i})^2} \qquad (14)$$

where,
y_j is the actual value
$\hat{y}_j$ is the predicted value
n is the number of observations

where TP = True Positives, TN = True Negatives, FP = False Positives, and FN = False Negatives.

4.1 Quantitative analysis

The proposed IRF method was tested for the 17 different persons. The system performance is evaluated in terms of precision, recall F-score and accuracy. With the help of the web page information, the NLP receiving section operates at high efficiency. For example, if the voice command sense the temperature, then the transmitter can update the command in Drop Box File (DBF). At the same time,

Received: February 15, 2021. Revised: April 5, 2021.

the receiver section reads the text file and executes the specific monitoring and controlling process. By utilizing this IRF method, eight devices are enough to control and monitor more than thirty devices. In this scenario, this research evaluates the effectiveness of all machine learning classifiers in terms of time taken to make the model, correctly classified instances, incorrectly classified instances and classification, which is shown in Table 1. The IRF algorithm is obtained better Precision of 99 % compared to the other algorithms. The Fig. 4 is the graphical representation of performance measures obtained for the IRF algorithm. Similarly, the training and simulation error obtained for IRF algorithm in terms of error values is shown in table 2.

4.2 Comparative analysis

The Table 3 shows the comparative analysis for the proposed IRF and the exisitng LSTM-RNN and ANN model that obtained error values expressed in terms of RMSE as 3.607, 1.226. The proposed IRF showed lesser error values of 0.0386 when compared with the exisitng models. Luis I. Minchala [16] developed PID-FGS faced aggressive industrial condition like dust and noise environment resulted with RMSE of 0.4247. Similarly, Bastian Dietrich [10] utilized the data was not suitable for energy flexibility and increase in energy cost showed optimization problem in the automation obtained RMSE of 3.607 and Precision of 67 %. In order to overcome the optimization problem occurred in the exisitng methods, Runhai Jiao [7] LSTM-RNN model that failed in considering load forecasting, economic orientations factors for the automation obtained precision of 96.3 % and RMSE of 1.226. Therefore, the proposed IRF model overcomes the overfitting problem as it was based on the received command and applies action according to it. The proposed IRF achieved RMSE of 0.0386 and precision of 99 % that overcomes the problem of optimization occurred in exisitng models. The comparison of the performances for the exisitng and the proposed research is performed based on the real-time datasets used in the researches. The Fig. 5 shows the simulation error graph for the exisitng model and the proposed IRF.

4.3 Accuracy results

Accuracy is one of the important parameters for computing classification models. Generally, the accuracy is evaluated by the fraction of predictions. The definition of accuracy is given in Eq. (7). For binary classification, accuracy can also be calculated in terms of positives and negatives as Eq. (7).

Table 3. The comparative analysis for the proposed and the existing methods in terms of RMSE

Authors	Methodology	RMSE	Precision (%)
Jiao, R. [7] 2020	Time Series Features and ANN	1.226	96.3
Dietrich, B. [10] 2018	LSTM-RNN	3.607	67
Luis I. Minchala 2020 [16]	FGS-PID	0.4247	-
Proposed	**IRF**	**0.0386**	99

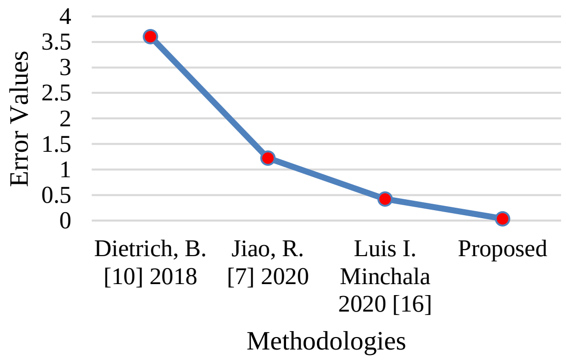


Figure. 5 Graphical representations for comparative analysis among existing and proposed method in terms of error values

Table 4 shows the accuracy of a predictive model for existing and IRF machine learning. Random forest, SVM and neural network have been achieved 87.48 %, 85.83 % and 71.66 % of prediction accuracy on 108 datasets. Accuracy of the predictive model has achieved 98.25% by proposed machine learning. The table 4 clearly shows the proposed algorithm is achieved better performance compared to the existing machine-learning algorithm. The accuracy results, idle time results, error matrix, RMS error, and actual vs predicted values are shown in Fig. 6, Fig. 7, Fig. 8, Fig. 9, and Fig. 10 respectively.

5. Conclusion

Machine learning algorithms are highly efficient for the industrial automation systems due to its accuracy. The major obstacle of the machine-learning algorithm is to solve the automation system problem with minimum error. This paper evaluated the performance of six machine-learning algorithms Naïve Bayes, SVM, SLLRM, KNN and decision tree on specific dataset. The collection of instances for various control commands are used as dataset and is given as input to different machine learning

 DOI: 10.22266/ijies2021.0831.14

Received: February 15, 2021.   Revised: April 5, 2021.

Table 4. Accuracy of the predictive model

Algorithm	**Datasets**	**Prediction Accuracy (%)**
Random forest	108	87.48
SVM	108	85.83
Neural Network	108	71.66
IRF algorithm	100	98.25

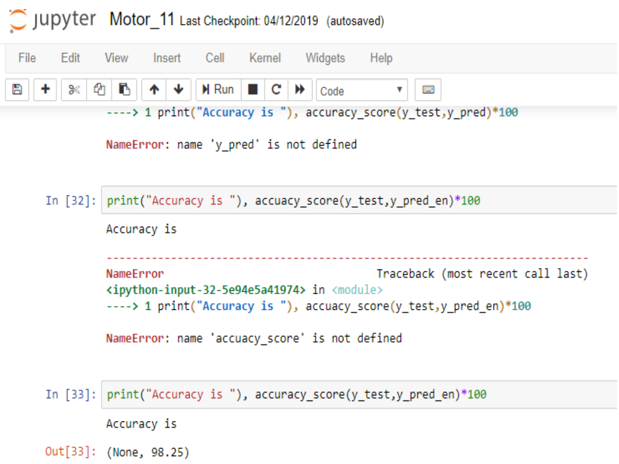


Figure. 6 Result of Accuracy

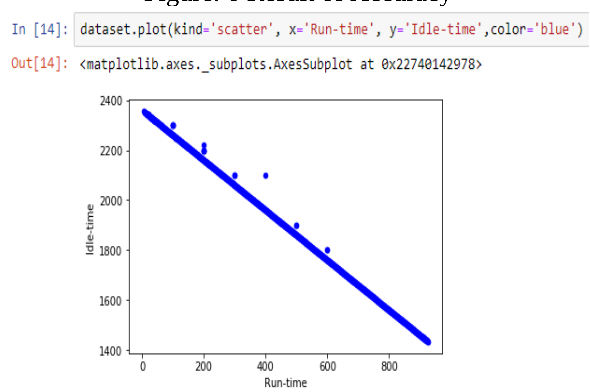


Figure. 7 Results of Idle time

algorithms. The performance comparison was made in terms of efficiency, sensitivity, FPR, precision, and F-measure, and accuracy. The proposed algorithm has been achieved better performance with accuracy 98.25 %. The RMSE values obtained for the proposed IRF model shows 0.0386 lesser error values when compared with the existing models that had achieved RMSE as 3.607 for Long Short Term Memory- Recurrent Neural Network (LSTM-RNN) and Artificial Neural Network (ANN) as 1.226. In the future, Deep Neural Network algorithm will be used for industrial automation in python language which is integrated with Raspberry Pi to improve the system accuracy.

Conflicts of Interest

The authors declare no conflict of interest.

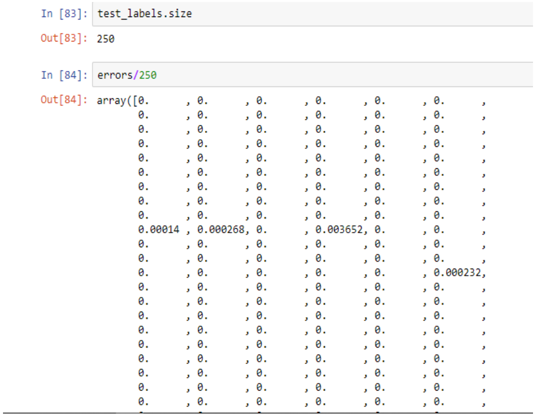


Figure. 8 Error matrix

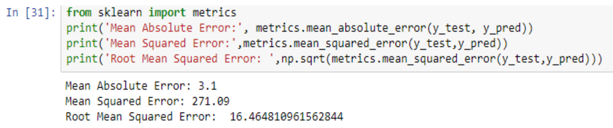
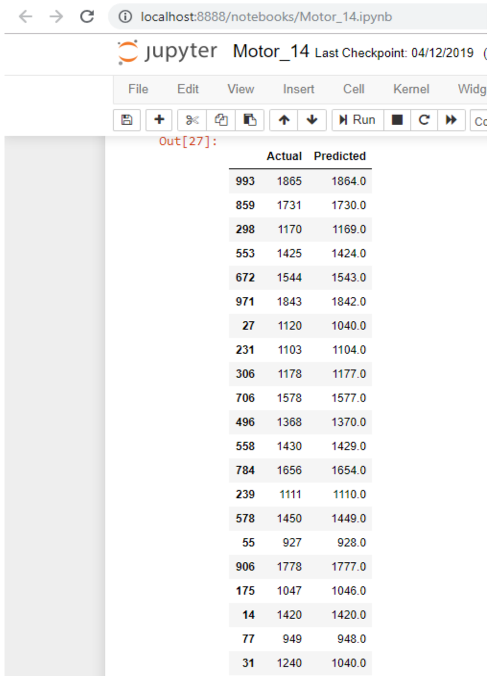


Figure. 9 RMS error



	Actual	Predicted
993	1865	1864.0
859	1731	1730.0
298	1170	1169.0
553	1425	1424.0
672	1544	1543.0
971	1843	1842.0
27	1120	1040.0
231	1103	1104.0
306	1178	1177.0
706	1578	1577.0
496	1368	1370.0
558	1430	1429.0
784	1656	1654.0
239	1111	1110.0
578	1450	1449.0
55	927	928.0
906	1778	1777.0
175	1047	1046.0
14	1420	1420.0
77	949	948.0
31	1240	1040.0

Figure. 10 Actual vs Predicted values

Author Contributions

The paper conceptualization, methodology, software, validation, formal analysis, investigation, resources, data curation, writing—original draft preparation, writing—review and editing, visualization, have been done by 1st author. The supervision and project administration, have been done by 2nd author.

References

[1] Q. Min, Y. Lu, Z. Liu, C. Su and B. Wang, "Machine learning based digital twin

Received: February 15, 2021. Revised: April 5, 2021.

framework for production optimization in petrochemical industry", *International Journal of Information Management*, Vol. 49, pp. 502-519, 2019.

[2] J. Cheng, W. Chen, F. Tao, and C.L. Lin, "Industrial IoT in 5G environment towards smart manufacturing", *Journal of Industrial Information Integration*, Vol. 10, pp. 10-19, 2018.

[3] R. R. Singh, S.M. Yash, S. C. Shubham, V. Indragandhi, V. Vijayakumar, P. Saravanan, and V. Subramaniyaswamy, "IoT embedded cloud-based intelligent power quality monitoring system for industrial drive application", *Future Generation Computer Systems*, Vol. 112, pp. 884-898, 2020.

[4] S. Ahmad, and D. Kim, "A multi-device multi-tasks management and orchestration architecture for the design of enterprise IoT applications", *Future Generation Computer Systems*, Vol. 106, pp. 482-500, 2020.

[5] K. Karunanithy and B. Velusamy, "Cluster-Tree based Energy Efficient Data Gathering Protocol for Industrial Automation using WSNs and IoT", *Journal of Industrial Information Integration*, Vol. 19, pp. 100156, 2020.

[6] K. Karunanithy and B. Velusamy, "Cluster-Tree based Energy Efficient Data Gathering Protocol for Industrial Automation using WSNs and IoT", *Journal of Industrial Information Integration*, Vol. 19, pp. 100156, 2020.

[7] R. Jiao, T. Zhang, Y. Jiang, and H. He, "Short-term non-residential load forecasting based on multiple sequences LSTM recurrent neural network", *IEEE Access*, Vol. 6, pp. 59438-59448, 2018.

[8] E. Sisinni, P. Ferrari, D.F. Carvalho, S. Rinaldi, P. Marco, A. Flammini, and A. Depari, "LoRaWAN range extender for Industrial IoT", *IEEE Transactions on Industrial Informatics*, Vol. 16, No. 8, pp. 5607-5616, 2019.

[9] P. Radanliev, D. C. De Roure, J. R. Nurse, R. M. Montalvo, and P. Burnap, "The Industrial Internet-of-Things in the Industry 4.0 supply chains of small and medium sized enterprises", *University of Oxford*, Vol. 3, No. 13, 2019.

[10] B. Dietrich, J. Walther, M. Weigold, and E. Abele, "Machine learning based very short term load forecasting of machine tools", *Applied Energy*, Vol. 276, pp. 115440, 2020.

[11] V. Casola, A. De Benedictis, M. Rak, and U. Villano, "Toward the automation of threat modeling and risk assessment in IoT systems", *Internet of Things*, Vol. 7, pp. 100056, 2019.

[12] Deshpande, P. Pitale, and S. Sanap, "Industrial automation using internet of things (IoT)", *International Journal of Advanced Research in Computer Engineering & Technology*, Vol. 5, No. 2, pp.266-269, 2016.

[13] T. Coito, M. S. Martins, J. L. Viegas, B. Firme, J. Figueiredo, S. M. Vieira, and J. M. Sousa, "A Middleware Platform for Intelligent Automation: An Industrial Prototype Implementation", *Computers in Industry*, Vol. 123, pp. 103329, 2020.

[14] M. D. Mudaliar and N. Sivakumar, "IoT based real time energy monitoring system using Raspberry Pi", *Internet of Things*, Vol. 12, pp. 100292, 2020.

[15] F. Foukalas, "Cognitive IoT platform for fog computing industrial applications", *Computers & Electrical Engineering*, Vol. 87, pp. 106770, 2020.

[16] L. I. Minchala, J. Peralta, P. Mata-Quevedo, and J. Rojas, "An Approach to Industrial Automation Based on Low-Cost Embedded Platforms and Open Software", *Applied Sciences*, Vol. 10, No. 14, pp. 4696, 2020.

[17] M. B. Tambe and M. J. Singh, "Industrial Automation using Linux based Webserver", *International Journal of Optimal Research in Science & Technology*, Vol. 3, No. 1, 2019

[18] S. Ahmad, S. Malik, I. Ullah, M. Fayaz, D. H. Park, K. Kim, and D. Kim, "An Adaptive approach based on resource-awareness towards power-efficient real-time periodic task modeling on embedded IoT devices", *Processes*, Vol. 6, No. 7, pp. 90, 2018.