iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** XI **Month of publication:** November 2024

DOI: https://doi.org/10.22214/ijraset.2024.65140

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue XI Nov 2024- Available at www.ijraset.com

Fault Diagnosis Model in Heat Exchangers Using ANN and SVM

Raj Bhamare, Athang Rajebhosale, Prajwal Pustode, Shreya Datir, Kaustubh Patil
Dept. of Chemical Engineering, Vishwakarma Institute of Technology, Pune, India

Abstract: ***Heat exchangers, often known as HEs, are essential parts of many industrial processes because they effectively transfer heat between fluids. However, these systems are susceptible to a range of issues that can compromise their performance and reliability.***
This study looks into the viability of using support vector machines (SVM) and artificial neural networks (ANN) for defect detection in heat exchangers operating in a steady state. Both deviations from normal measurement states and the physical causes of faults were analyzed.
Comparative results demonstrate that ANN and SVM methods exhibit superior sensitivity and accuracy in fault detection compared to traditional statistical and nearest neighbor classification methods, highlighting their potential for improving the dependability and effectiveness of heat exchanger operations.
Keywords: Support Vector Machine (SVM), Artificial Neural Network (ANN), Shell and Tube Heat Exchanger, Fault Diagnosis, Fouling, Fouling Factor.

I. INTRODUCTION

Heat exchangers are very important for power generation, chemical processing and HVAC systems. The performance and safety of the rest of the system must also be preserved, which mandates their efficient and reliable operation. However, a wide range of issues can affect heat exchangers—scale formation, leaks, fouling—that prohibit them from operating optimally and that can damage equipment or require expensive downtime. It is therefore important to ensure efficient fault diagnosis and rectification of these problems, for their early detection.

Traditional methods for detecting heat exchanger problems often rely on threshold-based (will things hot or cold) and manual inspection. While these methods are expensive, time-consuming and sometimes imprecise in complicated operating environments, they can be useful in some occasions. In recent years the promise of automatic, accurate and fast fault detection has caused machine learning techniques — in particular Artificial Neural Networks (ANNs) and Support Vector Machines (SVMs) — to blossom as tools for fault diagnosis.

A. Artificial Neural Networks (ANNs)

The micro-Systems operate under the same principles as a human brain is etwa artificial neural network. These consist of networked node layers (neurones) with connected weights. ANNs learn by training, a process in which the network adjusts its own weights based on inputs and outputs present in the training data set. This learning ability has made ANNs essentially useful in fault diagnosis problems possessing highly complex, nonlinear relationships. ANNs are a useful tool in the diagnosis of heat exchanger faults as they can learn to identify patterns from historical operational data that distinguish between normal operation and faulty behaviour. The trained ANN can be used to monitor all the time the heat exchanger status according to patterns learned, allowing real-time fault detection by evaluating its actual operating condition too such as those which were recognized.

B. Support Vector Machines (SVMs)

Supporting Soviet learning with vector machines applied to regression and classification issues SVMs work by determining the hyperplane that best separates data points into categories, in a higher-dimensional space. The aim is to increase the margin between classes so that it learns even better and generalizes overall.

The SVM can classify heat exchanger fault diagnosis into different operational states according to the sensor data and process parameters. Since the SVM was trained using labeled data that included normal and fault conditions, the model is able to identify different types of faults and accurately classify new observations.

©IJRASET: All Rights are Reserved

C. Combined Approach

The combination of ANNs and SVMs will combine the merits of both methods, promote an increase in accuracy for fault diagnosis. This is where ANNs are useful as it can model even complex, non-linear relationships and SVMs for good classification performance with a clear dividing line. One approach is a hybrid, with pre-processing and feature extraction by ANNs on raw data done in an initial stage before the class determination of the data using an SVM. Otherwise both models can be used together in conjunction with each other to validate the findings, making it more reliable.

II. HEAT EXCHANGER ISSUES

Sometimes the most frequent heat exchanger issues can be far more difficult to fix, which results in expensive capital costs and steadily rising operating expenses. Some of the most common heat exchanger problems that many process plants run with are as follows:

A. Fouling

The properties of the fouling process fluid, the heat transfer method, or both can typically have an impact on fouling. Biological fouling or particulate settlement can often be mitigated with low-cost process adjustments that slow down these fouling effects. Enhancing water filtration before it enters the exchanger's ability to decrease biological fouling can greatly enhance operational performance. Similarly, additional filtration processes particulates from the process streams can be eliminated upstream of the exchanger.

However, a lot of fouling types have heat transmission as a primary contributing factor. Heat transmission has an impact on the fouling rate during crystallisation, breakdown, and polymerisation. High temperatures accelerate fouling in polymerization or decomposition, so lowering temperatures can reduce the fouling rate and, consequently, the frequency of cleaning. For crystallization, higher surface temperatures can reduce fouling driven by low temperatures.

1) Different types of fouling include:
2) Crystallization
3) Decomposition
4) Polymerization and/or oxidation
5) Particles of dust, rust, or mud settling
6) Biological deposits
7) Corrosion





B. Tube Corrosion

Use of carbon steel tubes in shell and tube heat exchangers is most at danger from oxidation (corrosion) of the tubes' heat transfer surface..

The most common kind of corrosion is the result of an oxygen (O2) and iron (Fe2, Fe3) reaction. This process results in an increasing coating of iron oxide (Fe2O3) on carbon steel tubes, which eventually leads to the tubes' degradation because of a reduction in their thermal permeability. This problem is difficult to diagnose and is often not found until the tubes have corroded to the point where they no longer function as well thermally, the fluid flow has significantly decreased, or the tubes have been perforated and leak.



C. Tube Erosion

Tube erosion is the term used to describe the physical corrosion of metal caused by fluids. High levels of TDS, which can be found in saltwater containing salt, sand, and marine life, or in fluids like silica or silt, accelerate the erosion of the tubes inside the inlet tubes as well as at their leading edges. The leading edge of the inlet tubes and, if present, the U bend are often the weakest parts of the tube. But with time, all tubes are vulnerable to erosion.

U Bend Erosion



Inlet Tube End Erosion



D. Thermal Fatigue

Heat exchanger tubes can rupture and shatter due to the cumulative strains caused by large temperature differentials or frequent thermal cycling. When there is a significant temperature differential between the tubes and the shell, the tubes will bend and cause thermal fatigue.



III. LIETERATURE REVIEW

[1] Centrifugal pumps are essential in thermo fluidic systems but are prone to vibrations, potentially leading to premature failures that affect system reliability. Early fault diagnosis is crucial to prevent such breakdowns. This study uses SVM (Support Vector Machine) and ANN (Artificial Neural Network) techniques for diagnosing pump faults. Analysis of vibration signals is done in the frequency and temporal domains, with feature ranking methods like Chisquare, ReliefF, and XGBoost used to reduce dimensionality. ANN outperforms SVM in fault classification, and Chisquare and XGBoost are more effective than ReliefF for feature ranking. The findings indicate that an ANNbased approach with these ranking techniques is highly effective for centrifugal pump fault diagnosis.

[2] Chiller fault identification and diagnosis (FDD): As the building equipment with the highest energy consumption, is challenging due to their nonlinear nature and numerous parameters. In order to investigate four component level and three system level problems, this study offers a least squares support vector machine (LS-SVM) model optimised by cross validation for FDD on a 90 tonne centrifugal chiller. Eight faultindicative features were validated from an initial 64 parameters. The optimized LS-SVM model outperformed other machine learning methods in overall and individual fault detection rates, efficiency of diagnosis, rate of detection, and rate of false alarms, particularly for issues at the system level. Remarkably, the correct rates for system level faults like refrigerant leak/undercharge, refrigerant overcharge, and excessive oil were 99.59%, 99.26%, and 99.38%, respectively, with the model running 36.7% faster than a standard SVM model.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

[3] Using variable refrigerant flow (VRF) air conditioning systems in heating mode, this research presents an optimised back propagation neural network (BPNN) method for defect diagnostics. The approach involves optimizing the feature variable set using data mining techniques, starting with correlation analysis to remove redundant variables, and then improve the feature set using association rule mining. Five feature sets (FS1FS5) were created, with FS1 being the original and FS2FS5 refined through these methods. To assess the models, four fault experiments were carried out: refrigerant undercharge, overcharge, four way reversing valve failures, and outdoor unit heat exchanger fouling. Results demonstrate that correlation analysis effectively removes redundancies, and association rule mining optimizes feature sets. The BPNNFS5 model achieved the highest fault diagnosis accuracy, improving the correct rate from 88.71% to 96.40%, with hit rates for all faults exceeding 90%.

[4] Analytical PFD (Process Fault Diagnosis) methods to artificial intelligence techniques with Support Vector Machine (SVM) are widely used for fault classification, which has appeared as an exceptional method involving the capability of penetrating into unseen data. Multi-label support vector machine (MLSVM) is another technique, which overcomes the limitations of mono label artificial neural network (MLANN) techniques but it deals with a large amount of data and computation-intensive. While ML-SVM has its own advantages over tradition SVM, it generally has more worse classification performance compared with other typical classifier. AbstractThis work provides a novel solution with various regularization parameters via ML-SVM to enhance fault classification in one of the important process called Dew Point Process.

This new approach has better classification performance than traditional MLSVM and performs comparably to MLANN, but still requires (that) fewer data points, easier data collection …in addition to maintaining memory-based dual form learning of a classification problem in the non-linear feature due to kernel trick..

[5] This study proposes improved prediction methods using supervised ML algorithms to evaluate the effect of transverse baffles and air injection on a shell and tube heat exchanger's thermohydraulic performance. By injecting air at various flow rates, the goal is to optimize performance without relying on complex mathematical models or costly experiments. The outlet temperatures and pressure drop were predicted using four algorithms: knearest neighbours (kNN), support vector machine (SVM), random vector functional link (RVFL), and social media optimisation (SMO).

The techniques were applied on laboratory data. The findings demonstrate that RVFL provides better prediction accuracy by effectively capturing the nonlinear interactions between process responses and operating conditions. Compared to SMO, SVM, and kNN, RVFL's mean relative error (0.016167) and root mean square error (0.719167) were much lower, indicating its improved capacity to forecast thermohydraulic performance.

IV. METHODOLOGY

A. Data Loading

In this step, the dataset containing various features related to heat exchanger operations, along with a target variable indicating the presence of leakage, is loaded into the environment. This dataset is necessary to train machine learning algorithms that anticipate heat exchanger failures.

B. Data Preprocessing

Cleaning and converting the raw data into a format appropriate for analysis and modelling is known as data preparation. In order to guarantee that the dataset is correct and full, this involves addressing missing values.. Additionally, converting categorical variables to numeric form allows machine learning algorithms to process them effectively.

C. Feature Engineering

In order to improve the performance of machine learning models, feature engineering entails adding new features or changing ones that already exist. In this context, additional target variables for different types of faults (such as Fouling Fault, Corrosion Fault, and Blockage) are generated based on specific conditions in the dataset. This step enriches the dataset with more relevant information for fault detection.

D. Train-Test Split and Feature Scaling

The dataset is split into testing and training sets so that machine learning models can be appropriately evaluated. This separation ensures that the models are evaluated on unseen data, helping to prevent overfitting. In order to ensure that all of the input features have a similar scale and to keep any one feature from controlling the learning process, feature scaling is used.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

E. Machine Learning Models

Multiple ML algorithms are employed to predict different types of faults in heat exchangers:

1) Logistic Regression

A statistical technique called logistic regression is applied to binary classification problems (e.g., yes/no, true/false) in which the target variable has two possible outcomes. Despite its name, logistic regression is a linear model, not a regression model. The logistic (sigmoid) function, which transfers any real-valued number to the range [0, 1], is used to describe the likelihood that a given input belongs to a specific class. Logistic regression estimates the parameters (coefficients) of the linear equation using optimization techniques like gradient descent. Because of its efficiency and interpretability, it is frequently used for binary classification problems where there is a roughly linear connection between the target variable and the characteristics.

2) Random Forest Classifier

During training, numerous decision trees are built using the Random Forest ensemble learning technique, which produces the mean (regression) or mode (classification) prediction of each individual tree. A random selection of training data and characteristics are used to train each decision tree in the random forest. This helps to decorrelate the trees, minimising overfitting and enhancing generalisation. Every tree in the forest classifies the input data on its own before adding together all of the trees' predictions to arrive at the final forecast. Random forests are renowned for their scalability, resilience, and capacity to manage complicated feature interactions in high-dimensional data. They are especially successful at jobs involving classification with unbalanced datasets and noisy features.

3) Support Vector Machine (SVM)

SVM (Support Vector Machine) is a powerful supervised learning algorithm used for both classification and regression tasks. SVM seeks to identify the ideal hyperplane for classifying data points by dividing them with the greatest margin possible. The margin—the distance between the hyperplane and the closest data points, or support vectors, from each class—is maximised while selecting the hyperplane. SVM works well for problems involving a large number of features since it is efficient in high-dimensional domains. It can handle a variety of kernel functions (such as linear, polynomial, and radial basis functions) to handle data that is linearly separable or non-linearly separable. SVMs are renowned for their strong resistance to overfitting and good generalisation skills, particularly when working with small to medium-sized datasets.

4) Convolutional Neural Networks (CNN)

A family of deep learning models called CNNs (Convolutional Neural Networks)1 are intended to interpret structured grid data, such pictures or sequences. They are made up of several layers, such as fully connected, pooling, and convolutional layers. CNNs employ convolutional filters, or kernels, across the input data to extract features at various levels of abstraction, taking advantage of the geographic locality and hierarchical structure of the data. By lowering the dimensionality of the feature maps, pooling layers aid in the management of overfitting and computational complexity. When it comes to tasks like object identification, image segmentation, and image classification, CNNs are extremely efficient because they can automatically build hierarchical representations of features from raw data. They are widely employed in many different applications, including as natural language processing, autonomous vehicles, and medical imaging, and they have completely transformed the field of computer vision.

F. Model Evaluation

The machine learning models are assessed using a variety of metrics, including ROC curves, confusion matrices, and classification reports, once they have been trained. These measurements, which include accuracy, precision, recall, and area under the ROC curve (AUC), shed light on how well the models work. Assessing models through several measures guarantees a thorough comprehension of their efficacy in identifying faults.

1) Confusion Matrix

A table that lists a classification model's results given a set of test data for which the true values are known is called a confusion matrix. It compares the predicted classes with the actual classes in a tabular format. The matrix consists of four elements:

- True Positive (TP): The quantity of positive occurrences that were accurately predicted.
- The quantity of correctly anticipated negative cases is known as True Negative (TN).
- False Positive (FP): The quantity of cases that, although being Type I errors, are anticipated as positive.
- False Negative (FN): The quantity of cases (Type II error) that are expected to be negative but turn out to be positive.

The model's performance, including accuracy, precision, recall, and specificity, may be understood from the confusion matrix. It is particularly useful for assessing binary classification models and identifying areas where the model may be misclassifying certain instances.

2) Classification Report

A tabular summary of the different assessment metrics for a classification model is called a classification report. It contains metrics for every class in the dataset, including support, recall, F1 score, and accuracy. These measures offer a thorough grasp of the model's performance in several classes, particularly when datasets are unbalanced and one class may outperform the others. The metrics in a classification report are defined as follows:

- Precision: The model's ability to prevent false positives is indicated by the ratio of true positive predictions to all positive predictions.
- Recall (sensitivity): The model's capacity to detect every positive case is indicated by the ratio of true positive predictions to the total number of actual positive instances.
- F1 Score: the harmonic mean of recall and precision, which offers a fair assessment of the model's effectiveness.
- Support: The quantity of real instances of every class in the dataset.

The classification report aids in comprehending the model's advantages and disadvantages for every class, guiding further improvements or adjustments to the model.

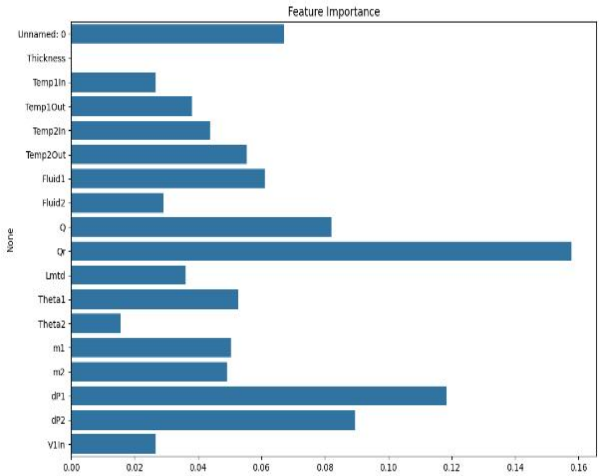
3) ROC Curve (Receiver Operating Characteristic Curve):

A binary classification model's performance at various threshold values is shown graphically via the ROC curve. At different threshold levels, it shows the true positive rate (TPR) against the false positive rate (FPR). The y-axis represents the true positive rate (sensitivity or recall), and the x-axis represents the false positive rate (1 - specificity). A distinct threshold setting is shown by each point on the ROC curve, which shows how sensitivity and specificity are traded off. The model's performance is summarised using the area under the ROC curve (AUC-ROC) for all potential threshold settings. A higher AUC-ROC value indicates stronger discrimination ability, with a value of 1 reflecting a perfect classifier and 0.5 suggesting a random classifier. ROC curves are helpful in assessing binary classification models and comparing how well various models perform. They shed light on the model's capacity to distinguish between positive and negative instances at different threshold settings, which aids in choosing the best threshold for making decisions.

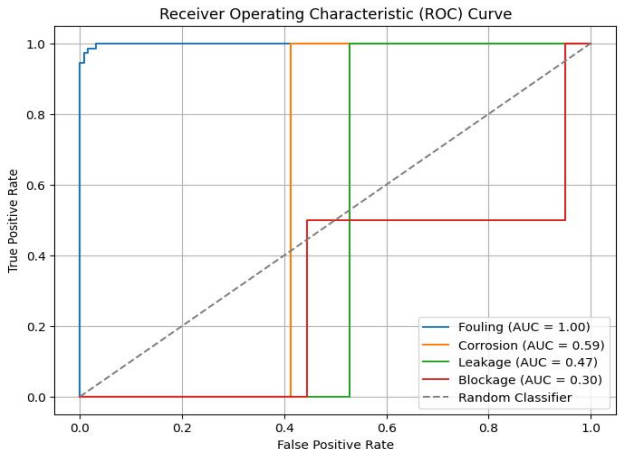
G. Predictions on New Data

The models are used to predict outcomes on fresh data samples after they have been trained and assessed. This stage shows how the models can be used practically in real-world situations to detect heat exchanger problems in a timely manner.

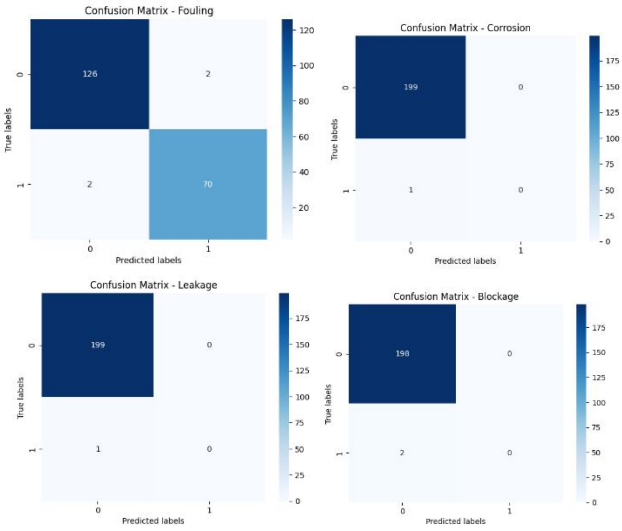
V. RESULT AND DISCUSSION



Overall Feature Extraction



ROC Curve



VI. DISCUSSION

The method is called alkaline electrolysis and it uses electricity to split water into hydrogen and oxygen. This process is performed using an electrolytic cell, which contains anode(+ve) and cathode(-ve) on the opposite ends of electrolyte. In general, KOH or NaOH solutions are used as the electrolyte. When an electric current is introduced to the cell, water molecules break and hydroxide (OH-) and hydrogen (H+) ions are created. Once over at the cathode, the hydrogen ions turn into H2 (hydrogen gas). The hydroxide ions are attracted to the anode and become oxidized to give oxygen gas (O2). Alkaline electrolysis is a relatively efficient process, with an efficiency of up to 70-80%. However, a variety of variables, including as the operating temperature, the current density, and the electrolyte's purity, might have an impact on the process' efficiency.

VII. CONCLUSION

The work carried out in this paper consists of the development and validation of fault diagnosis models for heat exchangers using both ANN and SVM. Extensive datasets controlled the training of models, comprising a multiple feature set covering heat exchanger operations, which underwent prediction and fault types classification. Our results indicate that ANN and SVM are both good options for identifying fault occurrences but each has strengths in particular areas.

The ANN model showed the best capability to describe complex, nonlinear relationships within the data, which may explain why it was so resilient in detecting subtle fault modes. On the other hand, the SVM model gave a very good classification power with clear decision boundaries that can be effectively used to differentiate between fault types [29] By integrating these models, we can exploit their complementations for an improved fault diagnosis. Future work will focus on Real-time Monitoring Systems where integration of these models in further enhancing diagnostic accuracy and reliability needs to be enhanced with hybrid techniques.

REFERENCES

[1] Nagendra Singh Ranawat, Pavan Kumar Kankar and Ankur Miglani, "Fault Diagnosis in Centrifugal Pump using Support Vector Machine and Artificial Neural Network" System Dynamics Lab, Indian Institute of Technology, Indore, Madhya Pradesh, India. DOI : 10.36909/jer.EMSME.13881.

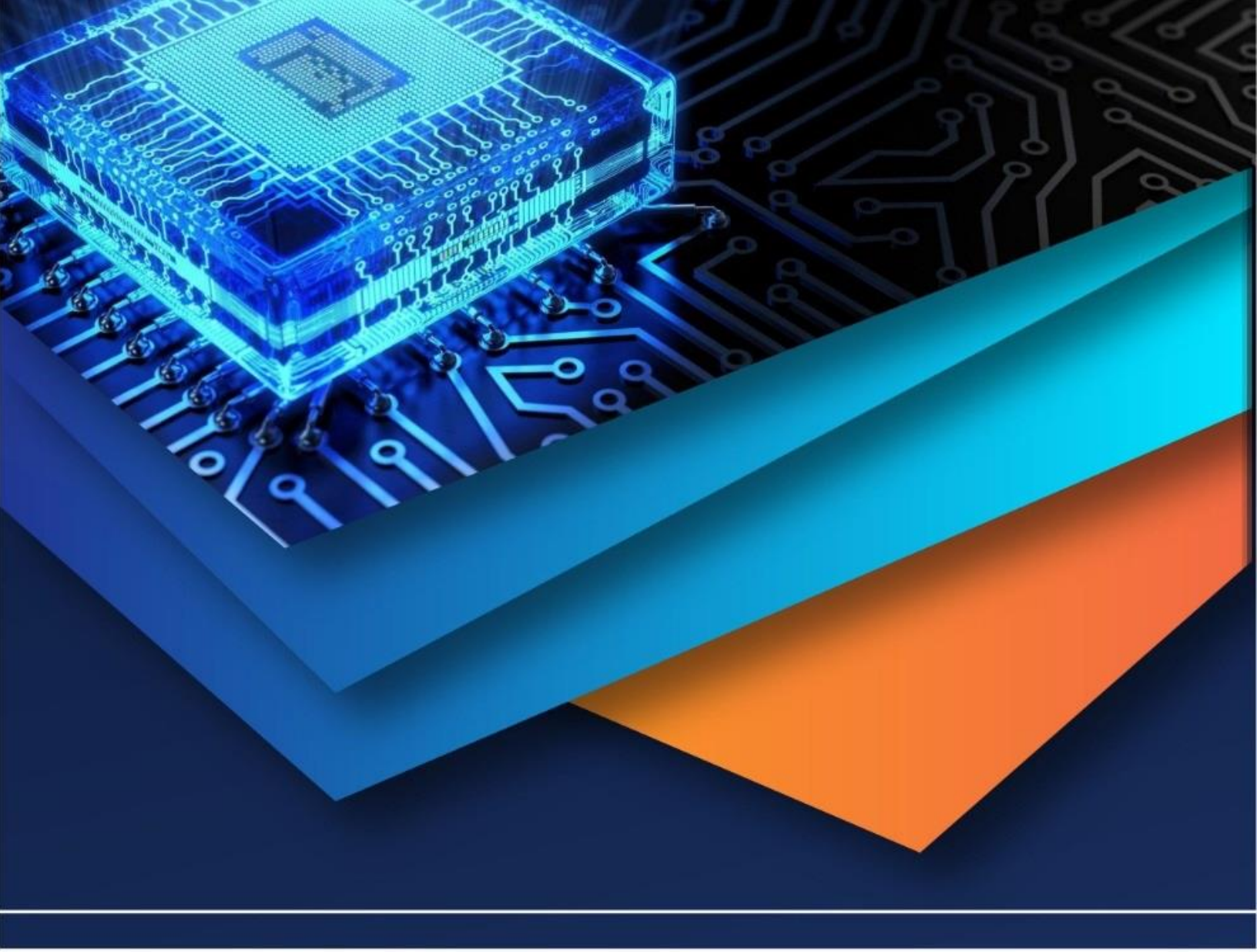
[2] Han, H., Cui, X., Fan, Y., & Qing, H. (2019). Least squares support vector machine (LSSVM)based chiller fault diagnosis using fault indicative features. Applied Thermal Engineering, 154, 540–547. doi:10.1016/j.applthermaleng.2019.

[3] Guo, Y., Li, G., Chen, H., Wang, J., Guo, M., Sun, S., & Hu, W. (2017). Optimized neural networkbased fault diagnosis strategy for VRF system in heating mode using data mining. Applied Thermal Engineering, 125, 1402–1413.

[4] Pooyan, N., Shahbazian, M., salahshoor, K., & Hadian, M. (2015). Simultaneous Fault Diagnosis using multi class support vector machine in a Dew Point process. Journal of Natural Gas Science and Engineering, 23, 373–379. doi:10.1016/j.jngse.2015.01.043.

[5] Emad M.S. ElSaid, Mohamed Abd Elaziz, Ammar H. Elsheikh, Machine learning algorithms for improving the prediction of air injection effect on the thermohydraulic performance of shell and tube heat exchanger, Applied Thermal Engineering, Volume 185, 2021.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)