Energy Reports 13 (2025) 3469–3488

Contents lists available at ScienceDirect

Energy Reports

journal homepage: www.elsevier.com/locate/egyr

Research paper

Enhancing fault detection and classification in distribution transformers using non-contact magnetic measurements: A comparative study of tree models and neural networks

Sufiyan Rao [a], Syed Ali Abbas Kazmi [a,*], Muhammad Zubair Iftikhar [b], Thamer A.H. Alghamdi [c,d], Mohammed Alenezi [c,*]

[a] US -Pakistan Center for Advanced Studies in Energy (USPCAS-E), National University of Sciences and Technology (NUST), H-12 Campus, Islamabad 44000, Pakistan
[b] Department of Electrical and Electronic Engineering, The University of Melbourne, VIC 3010, Australia
[c] Wolfson Centre for Magnetics, School of Engineering, Cardiff University, Cardiff CF24 3AA, UK
[d] Electrical Engineering Department, Faculty of Engineering, Al-Baha University, Al-Baha 65779, Saudi Arabia

ARTICLE INFO

Keywords:
Artificial neural network
Magnetic flux density
Random Forest

ABSTRACT

Ensuring the reliability of an electric power supply network (ESPN) requires accurate and rapid fault detection in distribution transformers. This paper presents a Finite Element Analysis (FEA) approach for non-contact magnetic measurements to capture magnetic flux density (MFD) values during both short-circuit (SC) and open-circuit (OC) fault conditions. These MFD measurements are then classified using various machine learning algorithms, including decision tree (DT), gradient boosting (GB), random forest (RF), and artificial neural network (ANN). The RF model achieves the highest accuracies, with 99.74 % for SC faults and 93.02 % for OC faults, outperforming all other models. The ANN model shows accuracies of 98.71 % and 92.38 %, while the DT model achieves 92.85 % for SC faults and 88.75 % for OC faults, and the GB model records 95.63 % for SC faults and 90.55 % for OC faults. Additionally, the DT model demonstrates fast prediction times of 0.0028 s for 7203 SC samples and 0.0019 s for 4802 OC samples. The novelty of this research lies in the use of FEA-based non-contact magnetic measurements to collect MFD values, which enhances safety and fault detection accuracy compared to traditional voltage and current signals. Unlike previous studies focused on overhead line protection, this method provides equipment-specific protection for transformers. Furthermore, integrating MFD data with machine learning models significantly improves fault classification speed and accuracy, providing a significant advancement in transformer fault detection.

1. Introduction

An EPSN consists of three different interconnecting sections: generation, transmission, and distribution. These three interconnected sections of EPSN constitute a considerable portion of a state's economy. Thus, all the power system's components must operate at their maximum efficiency and be protected from faults. EPSN is more susceptible to faults due to its complex structure. These faults may lead to issues like damaging electrical devices, service disruption, and system instability which ultimately reduces system reliability (Singh et al., 2015; Thomas et al., 2023).

Nowadays, uninterrupted power supply to the end customers is the first and foremost plan of distribution utilities. Unfortunately, more than 80 % of the interruptions in the power supply occurred due to symmetrical and unsymmetrical shunt faults in the distribution line (Gonen, 2015). These distribution lines' faults come into view for several reasons, including poor weather conditions, tree branches' growth that interfere with power lines, resulting in the breaking of electric lines, and electric equipment failure (Bompard et al., 2013). Open-circuit (OC) faults in distribution networks can also result from physical damage to equipment, loose connections, environmental conditions, or corrosion. Unlike transmission lines, distribution networks faced challenges in fault diagnosis because of limited measurement options, dynamic topology, and numerous branches and tap loads (Choi et al., 2004).

* Corresponding authors.
E-mail addresses: alirao427@gmail.com (S. Rao), saakazmi@uspcase.nust.edu.pk (S.A.A. Kazmi), muhammadzubairiftikhar@gmail.com (M.Z. Iftikhar), Alghamdit1@cardiff.ac.uk (T.A.H. Alghamdi), AleneziM1@cardiff.ac.uk (M. Alenezi).

https://doi.org/10.1016/j.egyr.2025.03.011
Received 1 November 2024; Received in revised form 3 March 2025; Accepted 10 March 2025
Available online 18 March 2025

In the distribution structure, a significant part of the capital cost is spent on the distribution transformers (Dawood et al. 2023). It is a common electrical device used in the power system that undertakes the critical tasks of changing AC voltage. In (Ehsanifar, 2021), the paper discusses inter-turn faults in transformers because of short circuit currents that can cause severe damage to the transformer if left undetected, using finite element analysis (FEA). When short-circuit faults occur, the transformer can undergo a strong electromagnetic force that can damage or deform its winding or insulation, leading to the transformer outage from the EPSN (Li et al., 2021). In (Jin, 2024), it is also discussed that SC faults lead to transformer winding failure. OC faults, on the other hand, result in high voltage that causes insulation breakdown, as shown in (48th North American Power Symposium, 2016). The intensity of OC faults, as discussed in (Alpsalaz and Mamiş, 2024), shows that insulation failure leads to winding arcs, which may result in SC faults, ultimately causing extreme damage to the transformer. Additionally, (2016 IEEE Power and Energy, 2016) highlights how OC faults can severely damage transformers, resulting in fires and posing serious safety risks. So, quick and automated fault detection and diagnosis are particularly important.

In practical situations, a relay and circuit breaker handle fault-related operations. However, the activation of the relay is a slow process, and this task can be speeded up using Machine Learning (ML). In the domain of EPSN, researchers have pointed their attention toward detecting and classifying faults using ML. The ML models were used in (Balan et al., 2023) and (Alenezi et al., 2024) for fault recognition in distribution transformers, relying on current and voltage signals to generate trip signals for relays and circuit breakers. In (He et al., 2023), the CNN model was utilized for transformer fault diagnosis using a dataset based on current signal samples. However, it did not consider all SC fault scenarios or analyze the model's effectiveness in comparison to other ML algorithms. In (Venkata et al., 2022), a support vector machine (SVM) model was trained with 25,168 samples to classify symmetrical and unsymmetrical shunt faults on transmission lines. The RF classifier was used in (Fonseca et al., 2022) to classify transmission faults, and it outperformed the neural network in terms of time and accuracy, achieving 91.96 % compared to the ANN, which got 89.59 %. In (Zheng et al., 2022), an RF classifier was trained on 4200 samples collected during several faults in the industrial process, achieving an accuracy of 93.14 %, which outperformed other algorithms used in this process, such as K-nearest neighbors (KNN), SVM, and ANN. Similarly, in (El Mrabet et al., 2022), the RF model outclassed other state-of-the-art models, including SVM, DT, KNN, and ANN, in fault detection and processing time. In (Ogar et al., 2022), a transmission line fault model was simulated in MATLAB, and the supervised ML model, such as the CatBoost boosting classifier, was used for fault classification, resulting in an accuracy of 99.54 %. The discrete wavelet transform (DWT) technique was applied in (Gangwar and Shaik, Jul. 2023) to sample current signals. After that, the KNN method was adopted for precise fault detection and classification (FDC) in distribution networks. In (Moloi et al., 2022), the current signal was processed by using wavelet packet decomposition (WPD). Then, this processed dataset was applied to the SVM algorithm, which achieved an accuracy of 99.5 % for classifying ten types of faults.

For FDC, neural networks such as ANN in (Usman et al., 2020), a multi-layer perceptron (MLP) based on the Bayesian approach in (Ferreira et al., 2020), long short-term memory (LSTM) in (Rafique et al., 2021) and a capsule network with sparse filtering (CNSF) in (Fahim et al., 2021) were used for classifying transmission line faults. An extreme learning machine (ELM) model was utilized in (Goni, 2023) for the FDC in the transmission line by considering a total of ten types of faults and gaining an average accuracy of 99.18 %. The voltage and current samples dataset, based on the discrete Fourier transform (DFT), was used as input to the ANN algorithm for classifying ten types of faults in (Gutierrez-Rojas et al., 2022). In (Uddin, 2022), the voltage signal was sampled by applying DWT to provide accurate data features, and in (Mamuya et al., 2020), a similar strategy was utilized to obtain crucial features from three-phase current signals to train ANN, MLP, and ELM neural networks for FDC in distribution networks. In (Yuan and Jiao, 2023), a hybrid convolutional neural network (CNN) and LSTM model were used to detect faulty feeders for just single-phase-to-ground faults on the distribution side.

Both voltage and current signals in (Balan et al., 2023; Alenezi et al., 2024; He et al., 2023; Venkata et al., 2022; Fonseca et al., 2022; Zheng et al., 2022; El Mrabet et al., 2022; Ogar et al., 2022; Gangwar and Shaik, 2023; Moloi et al., 2022; Usman et al., 2020; Ferreira et al., 2020; Rafique et al., 2021; Fahim et al., 2021; Goni, 2023; Gutierrez-Rojas et al., 2022; Uddin, 2022; Mamuya et al., 2020; Yuan and Jiao, 2023) are utilized for FDC in the transmission and distribution network. For the attainment of voltage and current signals, a current transformer (CT) and a potential transformer (PT) are used at the electric power substation. During symmetrical and unsymmetrical faults, the operating capabilities of CT may be limited because of its core saturation. The physical interaction between the instrument transformer and the high-voltage line disregards safety protocols and regulations. To address these challenges and to increase the accuracy and precision of fault detection, non-contact magnetic field sensing offers a hopeful solution (Khadse et al., 2021). Furthermore, techniques like Fourier Transform or Wavelet Analysis are often needed to process voltage and current signals, as seen in (Balan et al., 2023; Alenezi et al., 2024; He et al., 2023; Venkata et al., 2022; Fonseca et al., 2022; Zheng et al., 2022; El Mrabet et al., 2022; Ogar et al., 2022; Gangwar and Shaik, 2023; Moloi et al., 2022; Usman et al., 2020; Ferreira et al., 2020; Rafique et al., 2021; Fahim et al., 2021; Goni, 2023; Gutierrez-Rojas et al., 2022; Uddin, 2022; Mamuya et al., 2020; Yuan and Jiao, 2023). In contrast, magnetic field measurements can provide direct knowledge regarding fault location and type with minimal signal processing. Magnetic fields can also pass through insulating materials, enhancing the ability of magnetic field sensors to detect anomalies effectively. In (Chen, 2022), contactless magnetic sensors are utilized for current measurements instead of traditional current transformers, which provide accurate data for load monitoring and load imbalances but don't consider how magnetic sensors would behave in faulty conditions. In (Kwasi Anane et al., 2021), fault analysis was conducted on the transmission line using MATLAB software, and magnetic sensors were used to differentiate only between fault and non-fault conditions. In (De Oliveira Neto et al., 2021), magnetic sensors were used to consider different types of symmetrical and unsymmetrical faults, but they didn't apply any ML model to classify these faults. In (Malik et al., 2020), contactless magnetic sensors were employed for unbalanced load detection, and the ANN algorithm was trained to differentiate between balanced and unbalanced loading. However, this study did not consider multiple FDC case scenarios. In (Khadse et al., 2021) a neural network was utilized to locate and classify faults on transmission lines by acquiring data from magnetic sensors mounted on the pole towers. However, this approach did not account for open-circuit (OC) fault conditions or the behavior of MFD across various machine learning algorithms.

2. Contribution of the paper

There is limited research on fault identification in complex and congested secondary distribution networks. Contemporary fault detection methodologies mainly focus on transmission and primary distribution networks. Moreover, previous studies didn't consider the impact of symmetrical and unsymmetrical faults on expensive equipment, such as transformers in the secondary distribution network. Most existing techniques rely on voltage and current signals, for which CT and PT are used. However, this method doesn't provide reliable, accurate data and a safer environment for the workers. Furthermore, there is a need for intelligent methods, such as machine learning, to manage and classify SC and OC fault MFD data into their specific fault types. The contribution of this article includes the following:

Table 1
Designed parameters of the 3D transformer.

S.no	Designed Parameters	Values
1	Power Rating	630 kVA
2	Primary Voltage / Secondary Voltage	11 kV / 0.4 kV
3	Primary Current / Secondary Current	33.06 A / 909.35 A
4	Frequency	50 Hz
5	Number of turns in HV winding/ LV winding	1375 / 50
6	Phase Configuration	3-Phase (Delta / Star)
7	Core Material / Winding Material / Tank Material	AK Steel / Copper / Steel 1008

- FEA software ANSYS Maxwell was used for the 3-Dimensional (3D) design and modeling of the 630 kVA distribution transformer.
- Twin Builder, a software from ANSYS, was integrated with ANSYS Maxwell through co-simulation to provide the power supply system to the distribution transformer.
- Positioning of contactless magnetic measurements at locations around the main tank where the strength of the leakage MFD is most prominent and easily detectable to perform fault analysis.
- Simulating SC and OC faults on the secondary side of the distribution line and fetching the MFD measurements through contactless magnetic measurements.
- Collected MFD measurement data was provided to various ML algorithms for performance evaluation related to fault classification tasks.

The rest of the article is organized as follows. Section 3 presents a review of Maxwell's equations, 3D modeling of the distribution transformer, SC and OC fault case scenarios, and ML algorithms. Section 4 describes the comparative analysis of the results obtained with magnetic measurements and ML algorithms. Section 5 concludes this article.

3. Methodologies

3.1. 3D distribution transformer modelling and simulation

The working principle of a transformer is based upon Faraday's law, which states that if a conductor is positioned in a changing magnetic field, it results in an induced voltage in the conductor, which causes the current to flow. In a transformer, the electrical energy is transferred from its input winding to its output winding through a magnetic field. In an ideal transformer, when time-varying primary voltage is applied to the primary winding, it causes the current in the primary coil, which generates core flux, and this flux induces a secondary voltage across the secondary coil. However, practically not all flux links both coils, and it leaks through the surrounding air around the transformer. This type of flux is called leakage flux, and the MFD of leakage flux is utilized in this study to monitor the transformer during SC and OC faults.

3.1.1. Mathematical model

For complex engineering problems, the Finite Element Method (FEM) is used as a numerical method to figure out intricate engineering tasks (ÖZÜPAK, 2021). It divides the complex geometry into a small number of areas which functions as an efficient tool to solve MFD-related problems. In this study, FEA for 3D transformer design is carried out through the ANSYS Maxwell software. In the case of transformers, the MFD is determined by the currents flowing through the windings, the geometry of the core, and the magnetic properties of the materials. The relationship between the current and the magnetic field in transformers is governed by Ampere's Law. The general mathematical expression for Ampere's Law in integral and differential form is expressed in Eq.(1) and Eq.(2) as:

$$\oint \vec{B}.d\vec{l} = \mu_0 I_{enc} \tag{1}$$

$$\nabla \times \vec{B} = \mu_0 \vec{J} \tag{2}$$

In Eq.(1) and Eq.(2), $\vec{B}$ is the magnetic field density, I_{enc} is the enclosed current and $\vec{J}$ shows the current density per unit area. In FEA, Eq.(1) and Eq.(2) are discretized as shown in Eq.(3), and the geometry is divided into smaller elements to iteratively compute the magnetic field at each point based on these equations.

$$\vec{B}_{normal} = \int \mu.\vec{J}_{normal} dV \tag{3}$$

During faults, the current density in transformer winding regions increases severely. The resulting MFD at a faulted area can be expressed in Eq.(4) as:

$$\vec{B}_{fault} = \int \mu.\vec{J}_{fault} dV \tag{4}$$

where $\vec{J}_{fault}$ is calculated according to the Eq.(5):

$$\vec{J}_{fault} = \frac{I_{fault}}{A} \tag{5}$$

If the transformer faces a short circuit, the increase in $\vec{J}_{fault}$ leads to spikes in $\vec{B}_{fault}$, especially in areas near the fault. These Eq.(3), Eq.(4) and Eq.(5) are implemented in the FEA-based transformer model to solve the simulations and provide MFD measurements during faulty current conditions.

3.1.2. ANSYS maxwell

ANSYS Maxwell is a high-performance user-interactive software that employs FEA to solve 3D electric and magnetic transient problems. Transient solver provides a better understanding of magnetic field behavior over time and how its behavior impacts the electrical circuits connected to them. The magnetic field formulation for transient problems used by ANSYS Maxwell is based on Maxwell's equations as stated above. These complicated mathematical problems are described by differential equations, and FEM is applied to solve these equations numerically. This technique breaks down complicated geometry into small parts known as finite elements. The basic unit of a finite element in Maxwell 3D is a tetrahedron and these units are combined to form a finite element mesh. The targeted field in each finite element is estimated by using a basic second-order quadratic polynomial Eq.(6).

$$H_x(x,y,z) = a_0 + a_1x + a_2y + a_3z + a_4xy + a_5yz + a_6xz + a_7x^2 + a_8y^2 + a_9z^2 \tag{6}$$

Eq.(6) shows how the magnetic field component (H_x) is estimated at a point (x,y,z) by using these coefficients (a_0, a_9). To create this polynomial, field quantities are calculated at 10 points in a 3D simulation. After the equilateral tetrahedra mesh is specified, the finite elements are arranged in a sparse matrix expressed in Eq.(7).

$$[S][H] = [\quad J] \tag{7}$$

This sparse matrix is shown in Eq.(7) is then solved by applying the standard technique such as Sparse Gaussian Elimination.

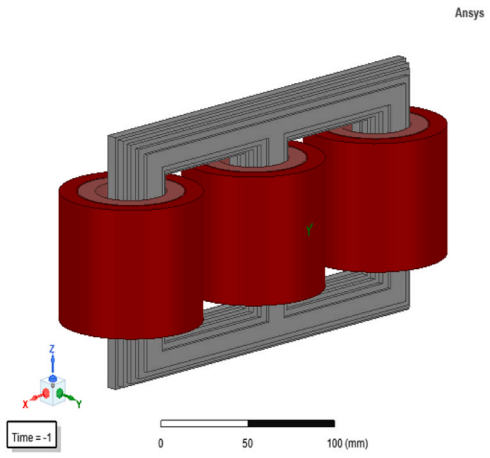

Fig. 1. 3D-designed model of core and windings (inner view).

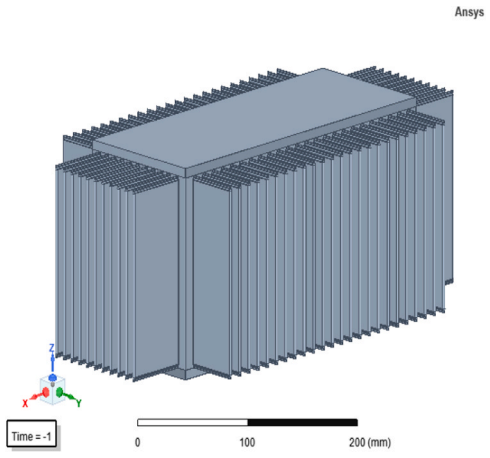

Fig. 2. Transformer steel tank (outer view).

3.1.3. *Distribution transformer model design*

The real design parameters of the utility distribution transformer are utilized for designing the distribution transformer in ANSYS Maxwell. Integrating these real parameters into the FEA transformer model guarantees that the MFD simulation results during fault analysis closely align with the real behavior of the physical transformer. The designed parameters of the 3D distribution transformer are displayed in Table 1.

The distribution transformer model is designed with a core-type structure using AK steel for its efficient magnetic properties, and the core consists of 9 stages of stacked laminated steel. Copper is selected for winding due to its high conductivity and low resistive losses. These copper windings are connected in a delta-star configuration. The transformer consists of three sets of copper windings, each placed on separate legs of the core. The HV coil, used as the primary side, is placed above the LV coil, used as the secondary side. This configuration results in three terminals for both the primary and secondary sides. The 3D-designed model of the core and windings is shown in Fig. 1. The steel tank of the distribution transformer is designed based on the real parameters as displayed in Fig. 2.

In ANSYS Maxwell, a feature exists to simplify a complex 3D model by dividing it into half if symmetry is present within it. This feature offers advantages, primarily a reduction in computational time and memory usage during simulation. Due to the symmetry of the transformer model across the opposite faces, a half model of the transformer as shown in Fig. 3, was taken for FEA. This half-model strategy reduces the complexity of the transformer, and it also doesn't impact the MFD measurement captured by magnetic measurement points.

The precise solution of the MFD measurement depends upon the mesh regions. The MFD measurement result becomes more precise by creating smaller and finer mesh elements as shown in Table 2, but it increases the simulation time. So, it's better to maintain the balance

Table 2
3D transformer model mesh element numbers.

S.no	Mesh Objects	Number of tetrahedra elements
1	High Voltage Coil	2204
2	Low Voltage Coil	1963
3	Core	9633
4	Transformer Steel Tank	12373
5	Region Around Transformer	27459

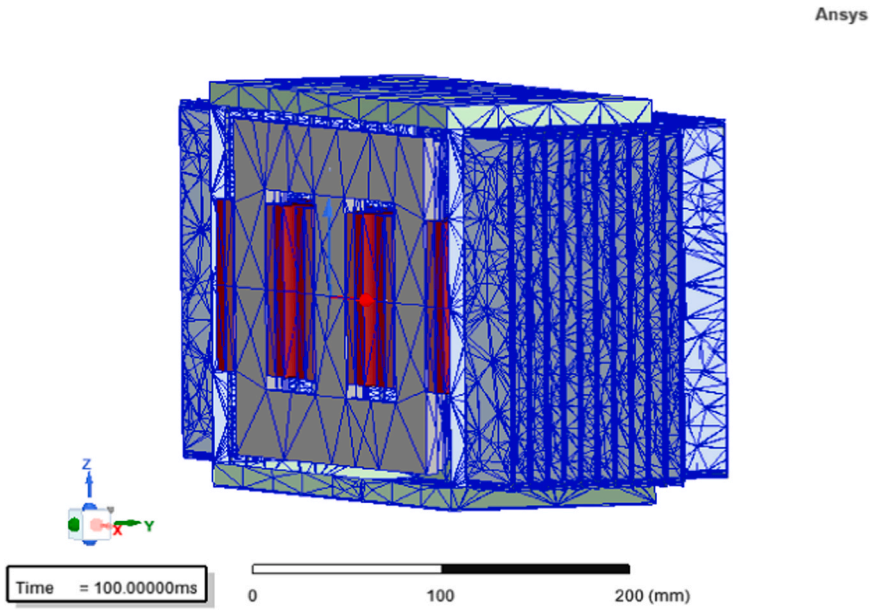

Fig. 3. 3D model splitting and creation of mesh elements for MFD analysis.

Table 3
Transient solver setup parameters.

S.no	Parameters	Value
1	Start time	0 s
2	Stop time	3 s
3	Step size	1 ms
4	Sample frequency	1 kHz
5	Total collected samples per case scenario	3001

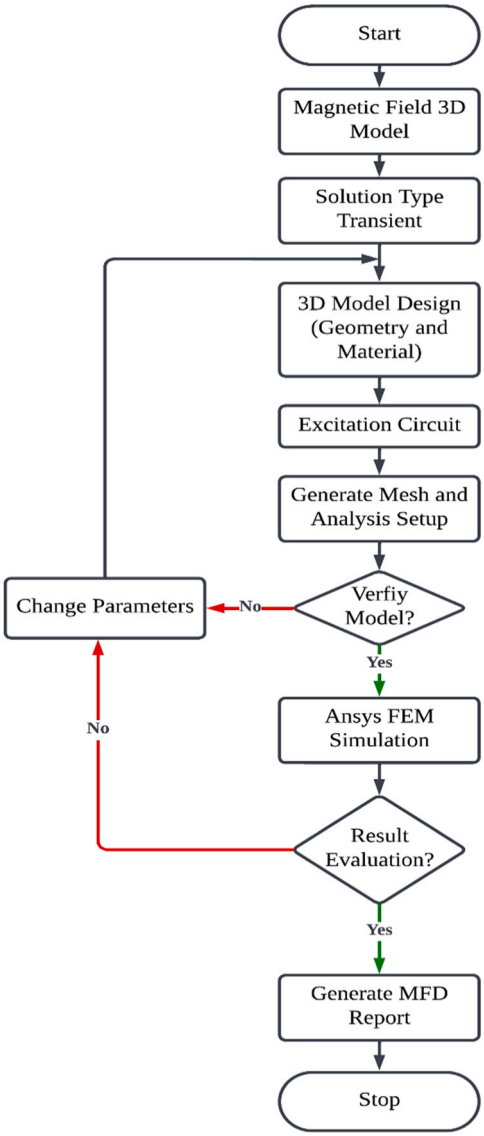

Fig. 4. 3D Transformer design and magnetic analysis process using ANSYS Maxwell Simulation.

between mesh elements and simulation time. The mesh elements for this analysis are shown in Fig. 3. The meshing method "Length-based inside selection" is used to create a mesh within a 3D transformer model. The tetrahedrons are used as mesh or finite elements that are dimensioned based on length considerations.

The MFD changes its value rapidly over time and it becomes difficult to accurately predict the MFD measurement due to its dynamic behavior. To counter this issue, the magnetic transient solution type is utilized. In this study, the distribution transformer was simulated for 3 seconds for each case, including normal conditions, 11 types of SC faults, and 7 types of OC faults. This resulted in a total simulation time of 60 seconds covering all fault scenarios and the normal condition. The transient analysis solver parameters are illustrated in Table 3, and the complete design procedure of the 3D transformer model is depicted using the flowchart in Fig. 4.

3.1.4. Placement of magnetic field points

In this study, magnetic points are placed near the center of the transformer on each face side to detect MFD around the distribution transformer. Ansys Maxwell provides a magnetic field point strategy to measure the magnetic field at various positions. These points interact with the magnetic field, recording the behavior of MFD over time. Five magnetic points are positioned at the center of five different faces of the transformer (left, right, top, bottom, and back) where the magnetic field strength is stronger, as shown in Fig. 5, to measure the MFD measurements. The central position on each side is selected because the MFD is typically intense near the core and windings, which are centrally placed within the transformer, and the intensity decreases with distance from the windings. The data collected from these field points is then exported in CSV format. This report presents data on the density of magnetic flux passing through a specified point area in the magnetic field.

3.1.5. Model simulation

To perform SC and OC fault analysis, Twin Builder software is combined with ANSYS Maxwell software through co-simulation. The Twin Builder provides an external excitation circuit that is applied to the windings of the distribution transformer. The external excitation circuit, shown in Fig. 6, is connected in a delta-star connection. In Fig. 6, the choice of a thyristor rectifier with a parallel R-L load simulates conditions where both resistive and inductive components are present in the load. In this parallel configuration, the voltage across the inductor and resistor is the same, ensuring proper voltage across the load. Furthermore, the current division between the resistive and inductive branches allows for efficient transformer operation by minimizing reactive power flow and improving the power factor. Moreover, the line reactance was excluded to simplify the circuit, as it primarily plays a significant role in long-distance transmission scenarios. In Fig. 7, the magnetic field vectors are shown, changing over time in response to the input provided to the distribution transformer. During various fault scenarios, the MFD changes rapidly, and these changes are effectively recorded by the non-contact magnetic points.

3.2. SC faults on 3-phase system

Most faults in a power system result in a SC condition. When such a situation occurs, a heavy current, known as the short-circuit current, passes through the electrical equipment, leading to considerable equipment damage and resulting in disruptions of services for consumers. These SC faults are classified as symmetrical and unsymmetrical faults.

An unsymmetrical fault, also called an unbalanced fault, is a fault that results in unbalanced fault currents with unequal phase displacement among the line currents. It interrupts the balanced nature of a three-phase system. These unsymmetrical faults include the Single Line to Ground (L-G), Line to Line (L-L), and Double Line to Ground (L-L-G). L-G fault occurs when one of the lines becomes shorted to the ground. L-L fault occurs when two lines are shorted together without including ground, and L-L-G fault occurs when two lines experience a short circuit with the ground. For these unsymmetrical faults, there are three possible fault case scenarios for each fault type.

A symmetrical fault, also called a balanced fault, is a fault that generates balanced fault currents with 120° phase displacement among the line currents. It disturbs all three lines of a three-phase system equally. These symmetrical faults include the Three Line (L-L-L) and Three Line to Ground (L-L-L-G). L-L-L fault occurs when all three lines come into direct contact without involving the ground, and L-L-L-G fault occurs when all three lines become shorted with the ground. For these

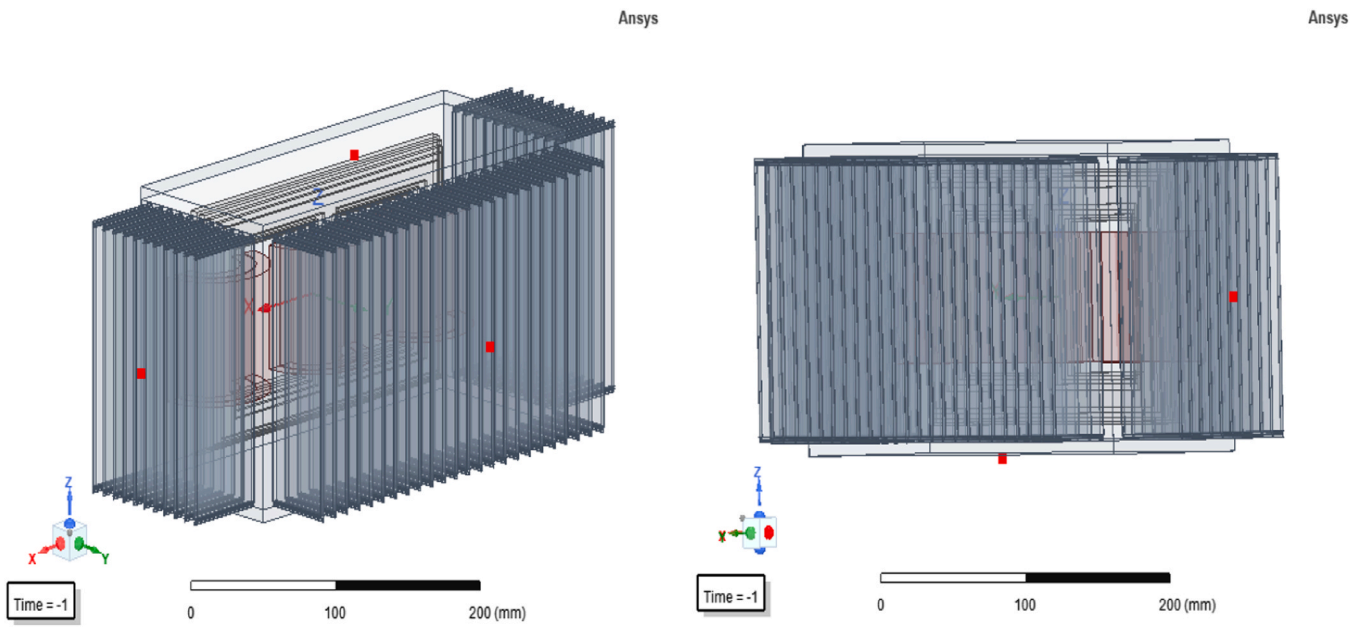

Fig. 5. Magnetic points placement at positions (left, right, top, bottom and back) to collect MFD data.

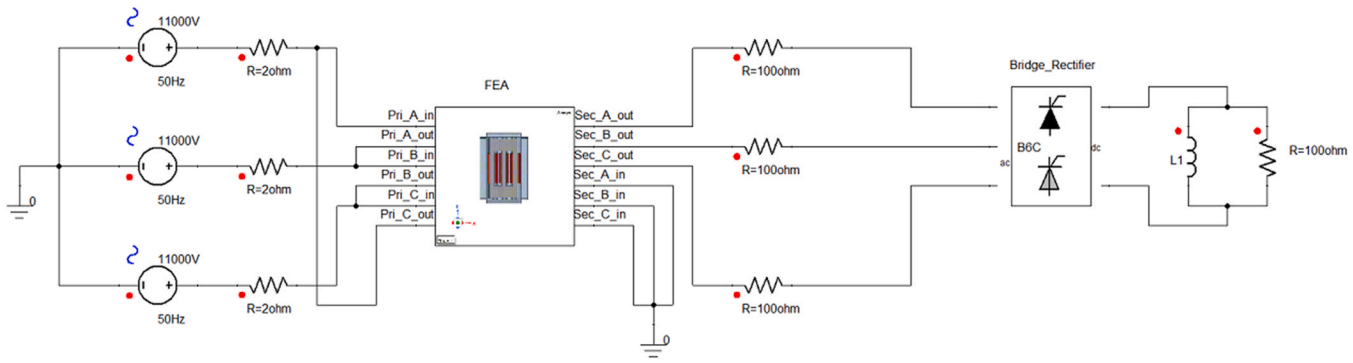

Fig. 6. Delta-star connected external circuit in Twin builder.

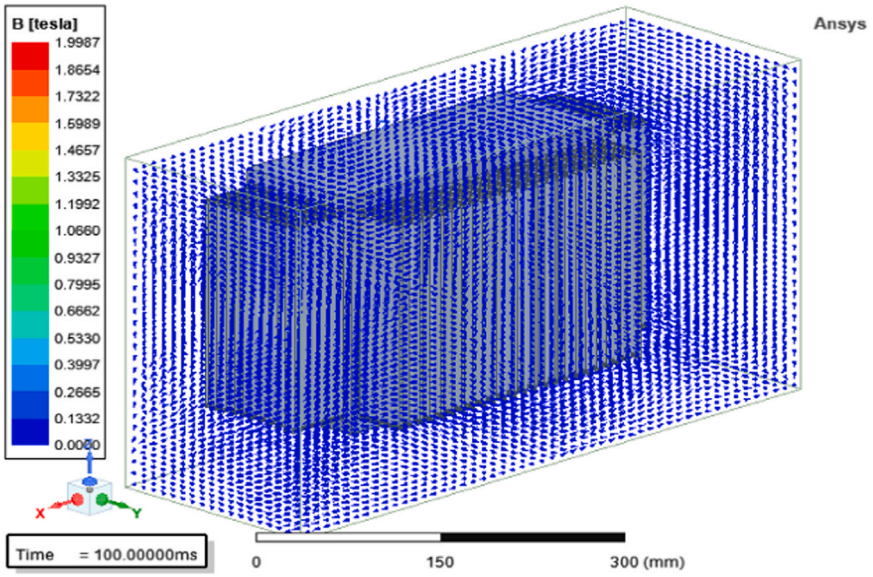

Fig. 7. Vector representation of MFD around the 3D transformer steel tank.

Table 4
Short circuit fault case scenarios with severity extent (Gururajapathy et al., 2017).

S.no	Fault Case	Fault Type	Fault Occurrence	Classification	Severity extent
1	AG	L-G	70 %	Unsymmetrical	Moderate
2	BG				
3	CG				
4	AB	L-L	15 %		
5	BC				
6	AC				
7	ABG	L-L-G	10 %		
8	BCG				
9	ACG				
10	ABC	L-L-L	5 %	Symmetrical	Extreme
11	ABCG	L-L-L-G			

Table 5
Training and testing sample datasets for SC case scenario.

Case Scenario	Label	Training Sample Counts	Testing Sample Counts
Normal	0	2398	603
AG	1	2421	580
BG	2	2403	598
CG	3	2391	610
AB	4	2416	585
BC	5	2390	611
AC	6	2371	630
ABG	7	2406	595
BCG	8	2400	601
ACG	9	2389	612
ABC	10	2425	576
ABCG	11	2399	602
Total Sample Counts		28,809	7203

Table 6
Training and testing sample datasets for OC case scenario.

Case Scenario	Label	Training Sample Counts	Testing Sample Counts
Normal	0	2427	574
A_open	1	2420	581
B_open	2	2391	610
C_open	3	2378	623
AB_open	4	2388	613
BC_open	5	2413	588
AC_open	6	2399	602
ABC_open	7	2390	611
Total Sample Counts		19,206	4802

symmetrical faults, there is just one fault case scenario for each fault type. In this study, the no-fault (NF) case and all 11 fault case scenarios are considered to examine their impacts on the MFD measurements around the transformer steel tank. The percentage of fault occurrence chances and severity of each fault type can be seen in Table 4.

3.3. Train and test data of SC MFD measurements

For NF and SC case scenarios, the MFD measurements through contactless magnetic points located at five different positions around the steel tank provided the sample dataset. The resulting data was exported from ANSYS Maxwell software as CSV files. These files were then imported to the coding platform to apply ML algorithms for automated FDC. The entire dataset, comprising all case scenarios, was divided into training and testing sections. The training and testing datasets are different because of the random sampling technique used in the ML algorithm, as presented in Table 5. This technique avoids human bias towards any specific fault scenario. By randomly selecting the data, the algorithm ensures that the model does not overfit to any particular fault type. The 80 % training dataset is used to teach the model to learn

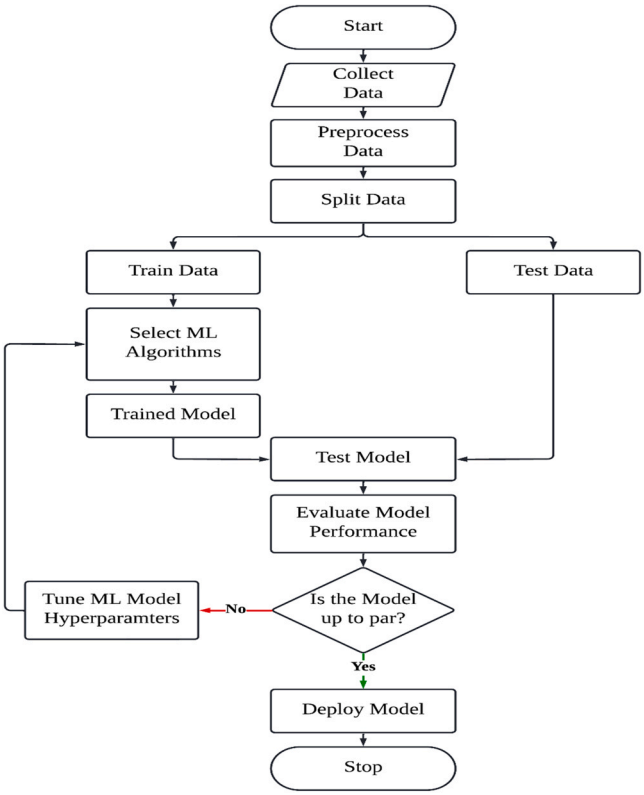

Fig. 8. Workflow of ML process utilized for the four algorithms in this study.

dataset behavior, while the 20 % testing dataset is reserved for evaluating its performance on unseen data to ensure how efficiently it learns the dataset's behavior. During training, the model learns the intricate MFD values and relationships within MFD data for each case scenario. The test dataset was used as a benchmark to assess the trained model's performance in accurately recognizing the unseen MFD measurements.

3.4. OC faults on 3-phase system

Open-circuit (OC) faults are common electrical issues that occur when there is an interruption in an electrical circuit, which impedes the flow of current. These faults can disrupt the normal operation of a three-phase system and can arise from various causes. In a three-phase system, OC faults can be categorized into three types: open one-phase, open two-phase, and open three-phase faults. An open one-phase fault occurs when one of the three phases becomes disconnected. An open two-phase fault happens when two of the three phases are interrupted. In an open three-phase fault, all three phases are disconnected, preventing any current flow.

3.4.1. Train and test data of OC MFD measurements

The entire dataset, comprising all OC case scenarios, was divided into training and testing sections. Here, 80 % of the MFD data was employed for training and 20 % for testing purposes as presented in Table 6.

3.5. ML approach for fault detection and classification

The standard procedure for any ML process is illustrated in Fig. 8, which begins with the data collection relevant to the task and preprocessing of data to enhance the data quality. Once the data is prepared, it is split into training and testing sections. The data allocated for training is then provided to the selected ML algorithm, permitting it to learn complex MFD measurements and relationships within data.

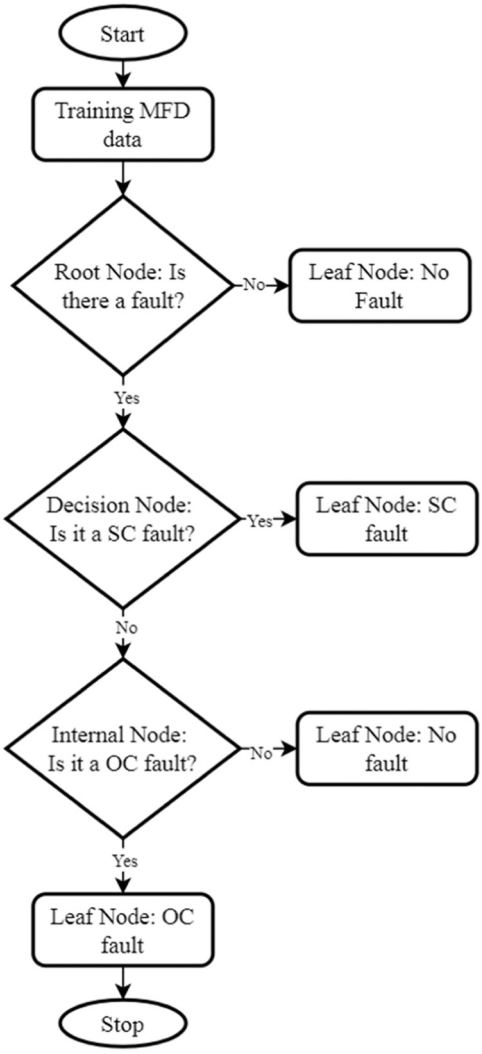

Fig. 9. Basic flowchart of the DT algorithm (Vidal et al., Oct. 2024).

Afterward, the algorithm's performance is evaluated using separate testing data. If the ML algorithm satisfies the predetermined performance criteria, it can be implemented for the task and it signifies the ending of the process. Otherwise, the selected ML algorithm's hyperparameters are tuned until the targeted level of performance is achieved.

3.6. ML algorithms

In this section, a concise mathematical overview of the four algorithms implemented in this study is demonstrated. To improve the model's performance based on the algorithm types, feature selection techniques like StandardScaler are used in tree models and ANN models to ensure all features contribute equally to the model. Furthermore, hyperparameter optimization is applied, such as adjusting learning rates, the number of layers, batch size in neural networks, and the number of trees and their maximum depth in tree-based models. These adjustments improve model accuracy, ensure faster convergence, and reduce training time. However, excessive tuning results in overfitting, causing the model to perform poorly on unseen datasets. Beyond a certain point, further tuning provides minimal improvements, making additional changes inefficient. Therefore, a balanced approach should be used. These four algorithms are chosen for their proven effectiveness in handling complex datasets and capturing non-linear relationships. The tree-based models offer clear views by breaking down the decision process into steps and demonstrating robustness against overfitting,

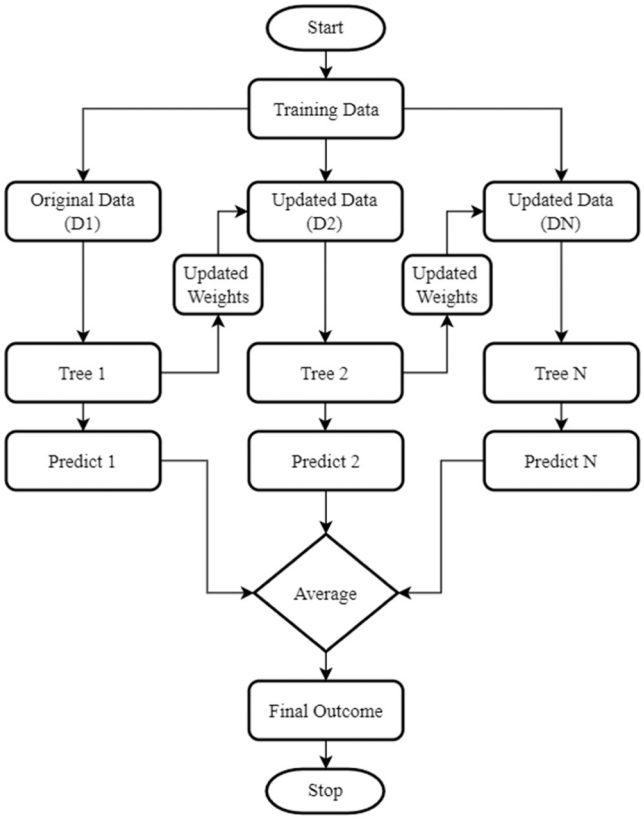

Fig. 10. Basic flowchart of the GB algorithm (Nhat-Duc and Van-Duc, 2023).

while ANN provides powerful capabilities for detecting complex MFD values in dataset. This combination enables the effective evaluation of NF, SC, and OC fault detection and classification. The ML algorithms with their mathematical expressions are described below:

3.6.1. Decision tree algorithm

In this research, a DT Classifier, a supervised learning algorithm, was utilized to perform FDC. The algorithm works by iteratively splitting the data into subsets based on the feature values, constructing a tree-like model of decisions as shown in Fig. 9. The classifier was trained and tested using a labeled dataset containing 20 scenarios, collected from monitoring data of MFD around the transformer. A train-test split was performed, with 80 % of the data used for training and 20 % for testing, ensuring that the model could effectively extend its predictions to new, unseen data. The Decision Tree model was trained on this dataset, with the training time denoted as t_{train} and the test time denoted as t_{test}, represented in the Eqs.(8)and (9).

$$t_{train} = t_{end\_train} - t_{start\_train} \tag{8}$$

$$t_{test} = t_{end\_test} - t_{start\_test} \tag{9}$$

These time measurements are important for evaluating the efficiency of the classifier in fault detection scenarios, where quick decision-making is necessary. Additionally, the Decision Tree utilizes metrics like entropy and information gain to decide the best splits, assuring that the tree grows optimally by selecting the most informative features at each step. This process limits uncertainty and improves classification accuracy.

$$Entropy(S) = -\sum_{i=1}^{c} p_i \quad \log_2(p_i) \tag{10}$$

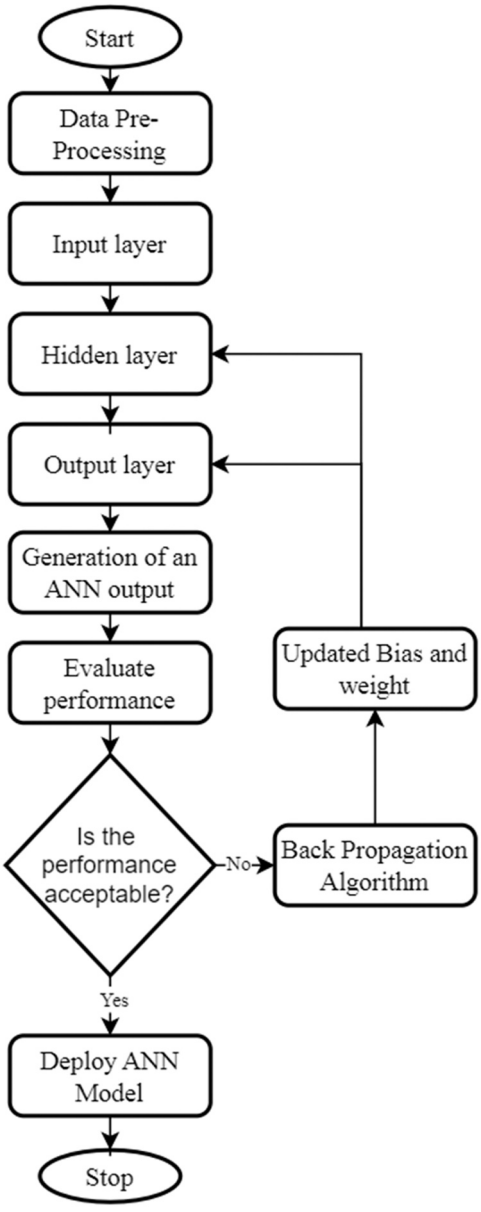

Fig. 11. Basic flowchart of the ANN algorithm (Hussain et al., 2023).

Table 7
List of parameters designed for the ANN algorithm.

S.no	Model Parameters		Value
1	Architecture		Feed-forward MLP
2	Number of Neurons	Input layer	5
		1st Hidden layer	24
		2nd Hidden layer	16
		Output layer	12 for SC, 8 for OC
3	Total Weights		748
4	Batch Size		64
5	No. of epochs		100
6	Optimizer		Adam
7	Activation function hidden layers		Tanh
8	Activation function output layer		SoftMax
9	Data normalization		Feature Scaling
10	Loss Function		Sparse Categorical Cross-Entropy

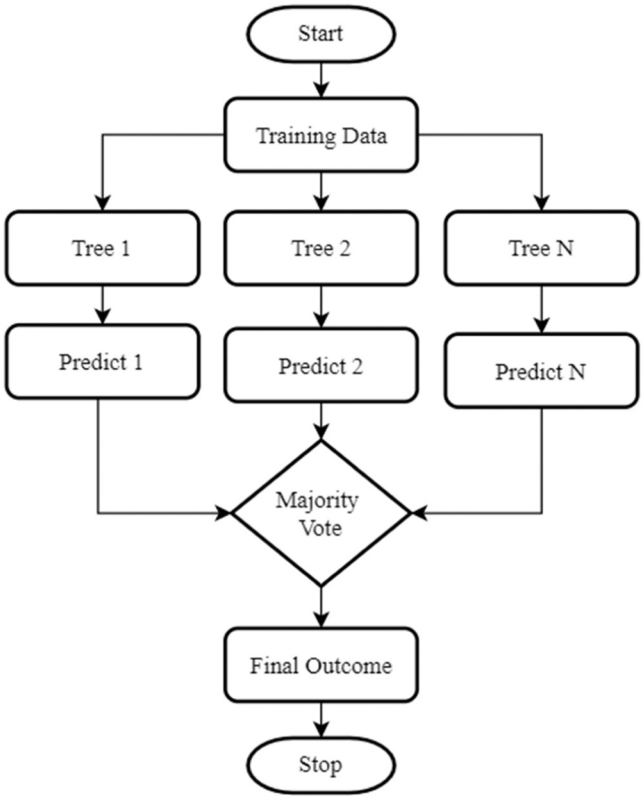

Fig. 12. Basic flowchart of the RF algorithm (Abbas, 2024).

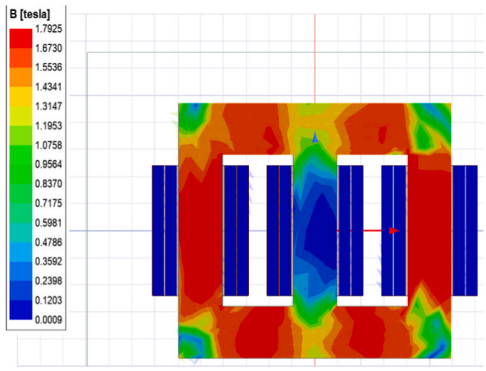

Fig. 13. Core saturation under fault conditions.

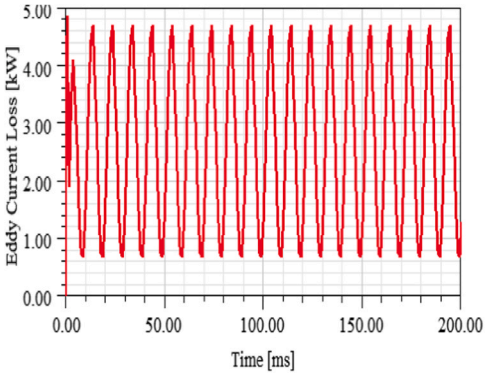

Fig. 14. Eddy current loss under saturated core.

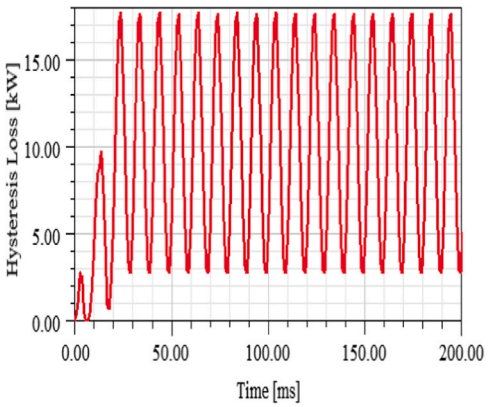

Fig. 15. Hysteresis loss under saturated core.

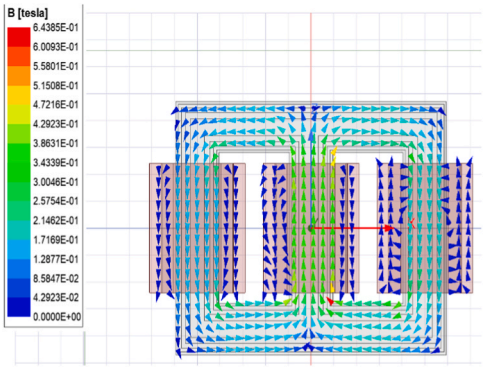

Fig. 16. Magnetic field patterns under normal case.

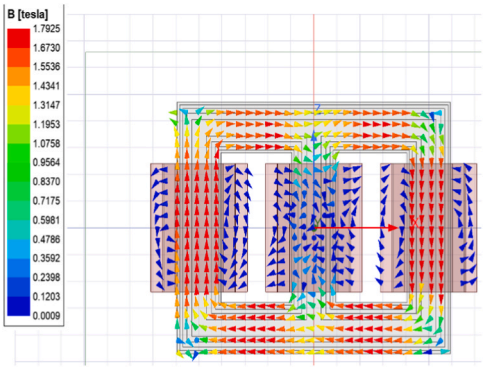

Fig. 17. Magnetic field patterns under fault case.

$$Information\ Gain(S,A) = Entropy(S) - \sum_{vinA} \frac{|S_v|}{|S|} \times Entropy(S_v) \tag{11}$$

In Eqs.(10), and (11), the term S represents the dataset, c represents the class labels and p_i represents the proportion of samples in class i relative to the total number of samples in the dataset.

3.6.2. Gradient boosting algorithm

In this study, a GB Classifier was utilized, which uses an ensemble technique that builds models sequentially, where each new model attempts to correct the errors made by the previous ones as presented in Fig. 10. This approach not only increases predictive performance but also provides robustness against overfitting, particularly when dealing with complex MFD datasets. The process begins with the initialization of the model, where a base learner is constructed. In this case, the model

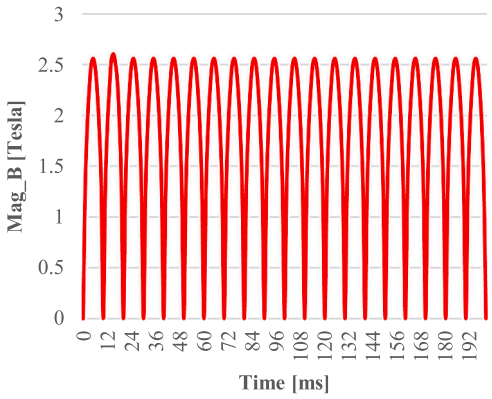

Fig. 18. MFD measurement at the left-side magnetic point under no-fault scenarios.

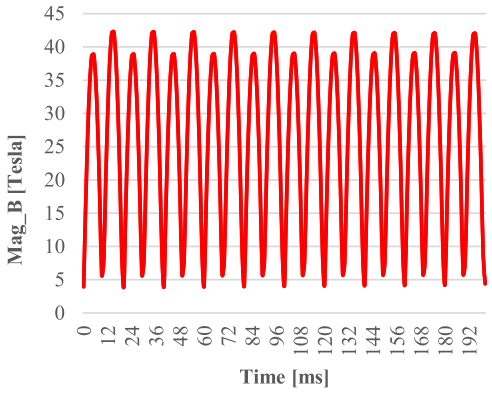

Fig. 19. MFD measurement at the left-side magnetic point under ABCG SC fault scenario.

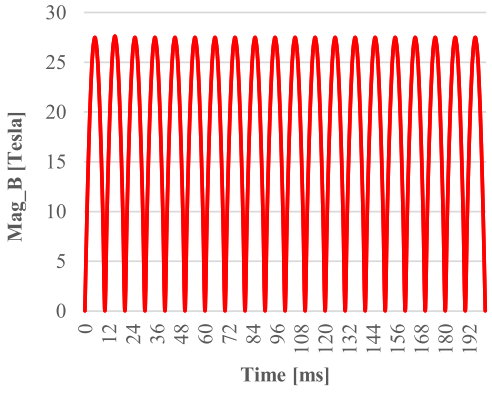

Fig. 20. MFD measurement at the left-side magnetic point under ABC OC fault scenario.

was defined with 50 estimators, meaning that 50 weak learners were unified to form a strong predictive model. The optimization of a loss function in this study is represented in the Eq. (12), where L is the overall loss function, y is the true output, $\hat{y}$ is the predicted output, and l is a loss function that calculates the error for observation.

$$L(y,\hat{y}) = \sum_{i=1}^{n} l(y_i,\hat{y}_i) \tag{12}$$

Once the training process was completed, the model's predictions were evaluated using various performance metrics, including the

Table 8
Performance evaluation metrics with formula description.

S.no	Metric	Formula
1	False Positive Rates (FPR)	$\frac{FP}{FP+TN}$
2	Recall or True Positive Rates (TPR)	$\frac{TP}{TP+FN}$
3	Precision	$\frac{TP}{TP+FP}$
4	F-Measure	$\frac{2\times Precision\times Recall}{Precision+Recall}$
5	Accuracy	$\frac{TP+TN}{TP+FN+TN+FP}$

Table 9
Classification report of the DT algorithm for SC case.

S.no	Case Scenario	Precision	Recall	F1-score	Count	Accuracy (%)
1	NF	0.93	0.89	0.91	603	-
2	AG	0.78	0.78	0.78	580	-
3	BG	1.00	1.00	1.00	598	-
4	CG	0.99	0.99	0.99	610	-
5	AB	0.96	0.94	0.95	585	-
6	BC	0.92	0.92	0.92	611	-
7	AC	0.88	0.91	0.89	630	-
8	ABG	0.99	0.98	0.98	595	-
9	BCG	0.98	0.99	0.99	601	-
10	ACG	0.99	0.98	0.99	612	-
11	ABC	0.92	0.91	0.91	576	-
12	ABCG	0.81	0.85	0.83	602	-
13	-	**Average**			**Total**	-
14	-	0.9291	0.9283	0.9283	7203	92.85

classification report and confusion matrix.

3.6.3. Artificial neural network

In this study, a feed-forward ANN model is utilized on MFD measurements for FDC in distribution transformers. The architecture is designed with multiple layers to handle intricate relationships within the dataset as displayed in Fig. 11. The network comprises of an input layer with 5 neurons linked to the features of the dataset, two hidden layers with 24 and 16 neurons, respectively, and an output layer with 12 neurons for SC faults and 8 neurons for OC faults, representing a total of 20 distinct fault categories including normal scenarios. The mathematical operations for each layer are performed using weights and biases, and the overall network is optimized to minimize error through backpropagation. The activation function used for both hidden layers is the tanh function, mathematically represented in the Eq.(13):

$$\tanh(x) = \frac{e^{x}-e^{-x}}{e^{x}+e^{-x}} \tag{13}$$

For the output layer, the SoftMax function is applied, converting raw prediction scores into probabilities for each fault class where z_i represents the input to the neuron and n is the total number of output neurons. The SoftMax function is shown in Eq.(14):

$$SoftMax(z_i) = \frac{e^{z_i}}{\sum_{j=1}^{n} e^{z_j}} \tag{14}$$

The ANN model was compiled using the Adam optimizer, an adaptive learning rate optimization algorithm that enhances both convergence speed and model performance. The optimizer is mathematically represented in the Eqs.(15), (16), and (17), where β_1 and β_2 are decay rates, α is the learning rate, and θ_t shows the model parameters.

$$m_t = \beta_1 m_{t-1} + (1-\beta_1) g_t \tag{15}$$

$$v_t = \beta_2 v_{t-1} + (1-\beta_2) g^2{}_t \tag{16}$$

Table 10
Classification report of the DT algorithm for OC case.

S.no	Case Scenario	Precision	Recall	F1-score	Count	Accuracy (%)
1	NF	1.00	0.90	0.95	574	-
2	A_open	0.75	0.90	0.82	581	-
3	B_open	1.00	0.90	0.95	610	-
4	C_open	1.00	0.70	0.82	623	-
5	AB_open	1.00	0.90	0.94	613	-
6	BC_open	1.00	0.82	0.90	588	-
7	AC_open	0.66	1.00	0.80	602	-
8	ABC_open	0.91	1.00	0.95	611	-
9	-	**Average**			**Total**	-
10	-	0.915	0.89	0.891	4802	88.75

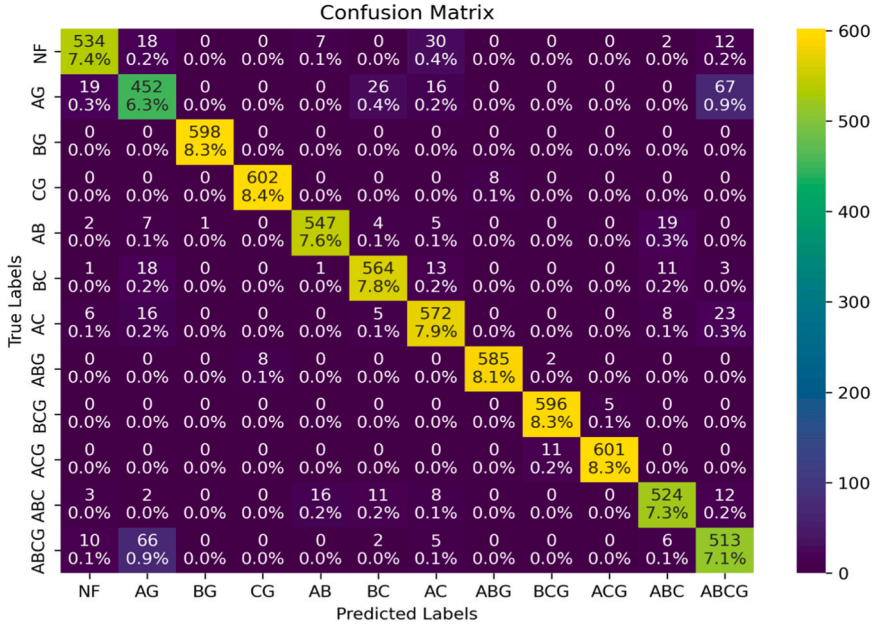

Fig. 21. CM of DT algorithm for SC case scenario.

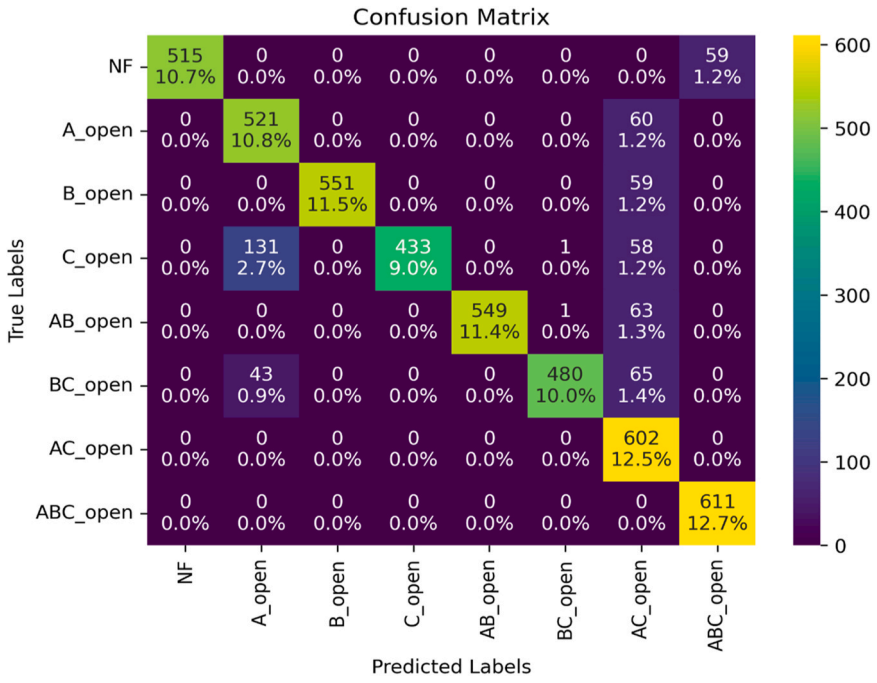

Fig. 22. CM of DT algorithm for OC case scenario.

Table 11
Classification report of the GB algorithm for each case scenario.

S.no	Case Scenario	Precision	Recall	F1-score	Count	Accuracy (%)
1	NF	0.93	0.93	0.93	603	-
2	AG	0.86	0.92	0.89	580	-
3	BG	1.00	0.99	1.00	598	-
4	CG	0.99	1.00	0.99	610	-
5	AB	0.97	0.99	0.98	585	-
6	BC	0.94	0.94	0.94	611	-
7	AC	0.94	0.93	0.94	630	-
8	ABG	1.00	0.98	0.99	595	-
9	BCG	1.00	0.99	1.00	601	-
10	ACG	0.99	1.00	0.99	612	-
11	ABC	0.97	0.90	0.93	576	-
12	ABCG	0.89	0.90	0.89	602	-
13	-	**Average**			**Total**	-
14	-	0.9566	0.9558	0.9558	7203	95.63

$$\theta_t = \theta_{t-1} - \alpha \frac{m_t}{\sqrt{v_t} + \epsilon} \tag{17}$$

$$L = \sum_{i=1}^{n} y_i \log(\hat{y}_i) \tag{18}$$

The model was trained over 100 epochs with a batch size of 64. The Sparse Categorical Cross-Entropy loss function was used shown in the Eq.(18), which is suitable for multi-class classification tasks. The specifications of this MLP model are displayed in Table 7.

3.6.4. Random forest algorithm

In this research, the RF Classifier, a robust ensemble learning technique, was utilized to address the classification challenges related to MFD datasets. This classifier functions by creating multiple decision trees during the training phase and integrating their outputs. The final prediction resulted from the majority vote of the individual trees, which improves predictive accuracy as shown in Fig. 12. The number of decision trees, denoted as $n_{estimators}$, was set to 50 in this study, indicating that 50 individual trees would be used to form a comprehensive model. The classification output can be mathematically expressed as follows:

$$\hat{y} = mode\{h_1(x), \quad h_2(x), \quad \ldots, \quad h_n(x)\} \tag{19}$$

In Eq.(19) $\hat{y}$ is the predicted label, $h_i(x)$ shows the output of the i^{th} decision tree for input x, and n is the total number of trees in the forest. After training, the model's performance was evaluated using evaluation metrics.

4. Results and discussions

4.1. Impact of SC faults on core in distribution transformers

SC faults are linked to core saturation in transformers because they cause sudden, high-magnitude currents that force the core to operate beyond its normal operating flux. When an SC fault occurs, the transformer experiences a rapid surge of current, similar to inrush current but with even higher magnitudes. These high-magnitude currents force the core flux to rise severely, often beyond its operating limits, leading to deep core saturation, as shown in Fig. 13, which increases core losses, generates harmonics, and causes overheating, thereby degrading the transformer's performance and lifespan. This saturation of the core increases core losses such as eddy current and hysteresis losses shown in Fig. 14, Fig. 15, as higher MFD causes excessive circulating currents in the core and more energy dissipation, reducing efficiency and increasing heat.

4.2. MFD pattern in normal and fault conditions

During both normal and fault conditions, complex and continuously changing magnetic patterns are captured through non-contact magnetic sensors in the form of MFD measurements. Under normal operating conditions, these magnetic patterns signify that the transformer core is operating efficiently with minimal losses as shown in Fig. 16. However, during fault scenarios, the patterns indicate a significant increase in core losses due to excessive heating as shown in Fig. 17. In visual representations, blue magnetic pattern vectors indicate minimal heating, while red magnetic pattern vectors show severe heating. If fault conditions remain, the transformer's efficiency deteriorates rapidly. To prevent

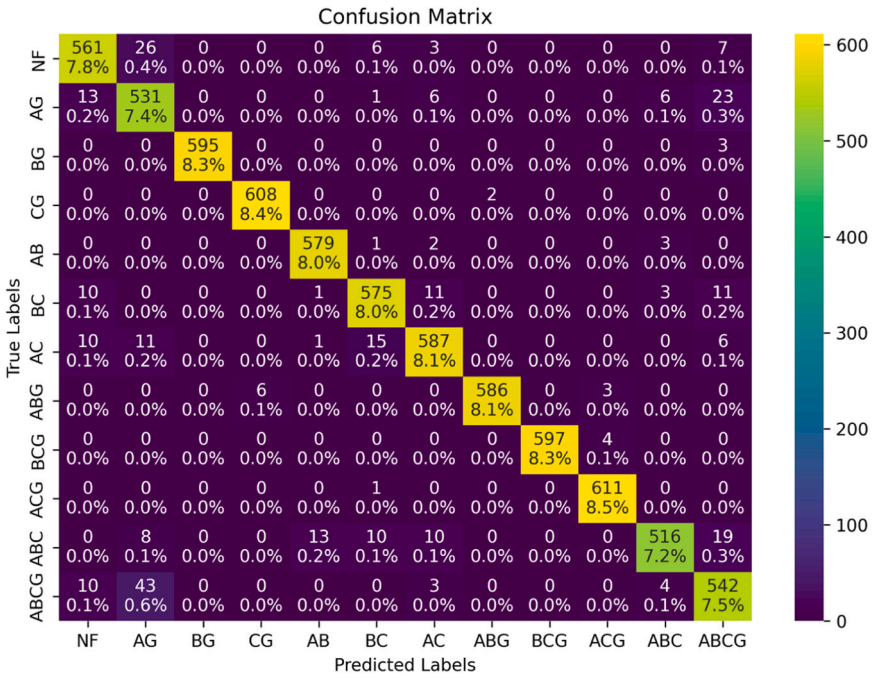

Fig. 23. CM of GB algorithm for SC case scenario.

Table 12
Classification report of the GB algorithm for OC case scenario.

S.no	Case Scenario	Precision	Recall	F1-score	Count	Accuracy (%)
1	NF	1.00	1.00	1.00	574	-
2	A_open	0.79	0.83	0.81	581	-
3	B_open	0.79	0.62	0.70	610	-
4	C_open	1.00	0.91	0.95	623	-
5	AB_open	0.84	0.90	0.87	613	-
6	BC_open	1.00	1.00	1.00	588	-
7	AC_open	1.00	1.00	1.00	602	-
8	ABC_open	0.83	1.00	0.91	611	-
9	-	**Average**			**Total**	-
10	-	0.906	0.907	0.905	4802	90.55

such degradation, an effective protection mechanism is necessary. In this study, non-contact magnetic sensors rapidly capture these magnetic patterns in the form of MFD values and provide them to ML algorithms, which rapidly classify different fault scenarios.

4.3. MFD measurement in normal and fault conditions

From the Eq.(4), it is evident that the MFD changes in response to fault current variations in the transformer. This section presents a comparative analysis of the left-side non-contact magnetic measurements, shown in Fig. 18, Fig. 19, and Fig. 20 to illustrate how MFD varies with changing fault currents. A similar behavior is observed in the other four sides of the non-contact magnetic measurement points, with slight variations in magnetic and current values. For simplicity, the comparison includes the first 200 MFD measurement samples for no-fault, ABCG SC, and ABC OC fault scenarios. The MFD graph is sinusoidal due to the AC current flowing through the windings of the transformer, which causes the magnetic field to change its magnitude over time.

In Fig. 18, the MFD behavior is normal during no-fault conditions. However, under the ABCG SC fault condition in Fig. 19, it rises to a significantly higher value, while under the ABC OC fault condition in Fig. 20, it peaks at a comparatively lower value than in the SC fault scenario. These results indicate that, under SC conditions, the MFD can reach higher values, causing more severe damage to the transformer compared to an OC fault. To mitigate such risks, a fast protection mechanism is important to protect the transformer from high MFD peaks by swiftly sending a trip signal to the circuit breaker. ML application offers an effective, automated solution for FDC, using non-contact magnetic MFD measurements to immediately detect faulty conditions and protect the transformer from damage.

4.4. Performance evaluation metrics

In the domain of ML, performance evaluation metrics perform a key function in figuring out the efficiency of the algorithms. The algorithms previously mentioned, such as DT, GB, ANN, and RF, are initially trained with the provided training data. After this, these algorithms are assessed on the test data, and the effectiveness of these algorithms is examined through performance evaluation metrics, as illustrated in Table 8. From Table 8, it's evident that true positive (TP), true negative (TN), false positive (FP), and false negative (FN) terms played an important role in building the performance evaluation report for algorithms. TP represents actual positive samples correctly predicted by the model as positive, and TN indicates where actual negative samples are rightly predicted as negative. FP represents actual negative samples incorrectly predicted as positive, and FN indicates where positive samples are mistakenly classified as negatives.

Precision, Recall, and F1 score use these positive and negative samples to measure a model's accuracy when making predictions. Precision focuses on the accuracy of positive sample predictions, while recall emphasizes the model's capability to capture actual positive samples. Accuracy is the important metric which is the ratio of correctly predicted samples to the total number of samples. For a better understanding of the error a model commits and the distribution of these errors across different classes, the confusion matrix (CM) has emerged as an important tool. It is a $n \times n$ matrix also called an error matrix where n represents the number of classes. The x-axis of this CM shows the predicted labels whereas the y-axis represents the actual labels. The diagonal of this CM indicates the correctly predicted samples and the off-diagonal represents the misclassified predicted samples. This helps in evaluating a model's performance and locating areas for improvement in classification problems. The last factor considered for assessing the performance of the above-discussed algorithms is the processing time

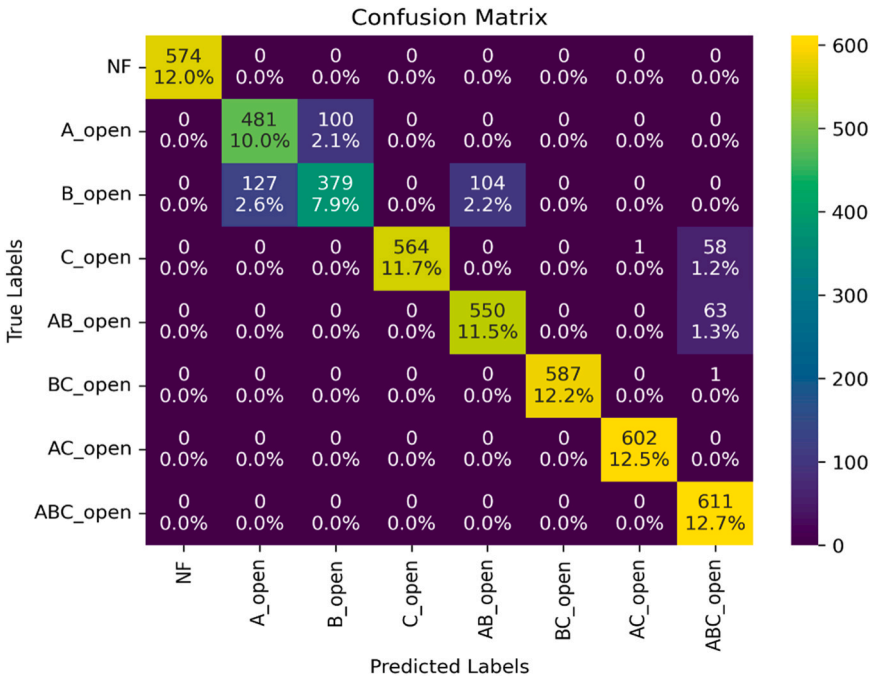

Fig. 24. CM of GB algorithm for OC case scenario.

Table 13
Classification report of the ANN algorithm for each case scenario.

S.no	Case Scenario	Precision	Recall	F1-score	Count	Accuracy (%)
1	NF	0.97	0.96	0.97	603	-
2	AG	0.96	0.99	0.97	580	-
3	BG	1.00	1.00	1.00	598	-
4	CG	1.00	1.00	1.00	610	-
5	AB	1.00	0.98	0.99	585	-
6	BC	0.97	0.97	0.97	611	-
7	AC	1.00	0.98	0.99	630	-
8	ABG	1.00	1.00	1.00	595	-
9	BCG	1.00	1.00	1.00	601	-
10	ACG	1.00	1.00	1.00	612	-
11	ABC	0.98	0.99	0.99	576	-
12	ABCG	0.97	0.97	0.97	602	-
13	-	**Average**			**Total**	-
14	-	0.9875	0.9866	0.9875	7203	98.71

taken during training and testing while executing all algorithms on the same platform. After that, these times are compared to draw comparisons among the algorithms. The acquired results for each algorithm are shown below:

4.4.1. Evaluation of DT model performance for SC case

After implementing the DT model on the MFD dataset, the achieved results are summarized in Table 9 which includes performance evaluation metrics for each case scenario. The classification report highlights that the DT algorithm experiences more challenges when predicting cases such as NF, AG, AC, and ABCG. In the case of AG, the precision shows that this model's predictions were true positive 78 % of the time, while the remaining 22 % were false positives. For ABCG, the recall indicates that the model correctly identified 85 % of the actual positive samples and spotted false negatives 15 % of the time. Moreover, the algorithm achieved an accuracy score of 92.85 % which signifies the overall performance of the DT algorithm is below average.

The CM presented in Fig. 21, clearly indicates that the DT algorithm struggles while differentiating the AG case from the ABCG case as it predicted 67 times the AG samples as ABCG and 66 times the ABCG samples predicted as AG samples. In total, the AG was misclassified 128 times and the ABCG was misclassified 89 times which indicates poor performance. Similarly, 30 NF case samples are incorrectly predicted as AC samples by the DT algorithm.

4.4.2. Evaluation of DT model performance for OC case

The classification report in Table 10 highlights that the DT algorithm experiences more challenges when predicting cases such as A_open, C_open, AC_open, and ABC_open. In the case of A_open, the precision shows that this model's predictions were true positive 75 % of the time, while the remaining 25 % were false positives. For BC_open, the recall indicates that the model correctly identified 82 % of the actual positive samples and spotted false negatives 18 % of the time. Moreover, the algorithm achieved an accuracy score of 88.75 %.

The CM presented in Fig. 22, clearly indicates that the DT algorithm struggles while differentiating the A_open case from the AC_open case as it predicted 60 times the A_open samples as AC_open. Similarly, 65 BC_open case samples are incorrectly predicted as AC_open samples by the DT algorithm.

4.4.3. Evaluation of GB model performance for SC case

The classification report displayed in Table 11, clearly indicates that the GB model faced challenges when predicting these cases such as NF, AG, and ABCG but it performs much better than the DT algorithm. Significantly, the algorithm obtained high accuracy in predicting cases like BG, BCG, and ACG. In the case of BG, the precision highlights that the model's predictions were 100 % true positive, while in the AG case, the model's precision predictions were only 86 % true positive. For BG, the recall indicates that the algorithm identified only 1 % of false negatives, whereas, for AG, the false negatives were 7 % of the time. Furthermore, the accuracy score of the GB algorithm is 96.63 %, suppressing the accuracy of the DT algorithm by 3.78 %.

The GB model's CM is displayed in Fig. 23. The CM indicates that the ABCG case was mixed up with the AG case because the classifier predicted the ABCG samples as AG samples 43 times and the AG samples as ABCG samples 23 times. Notably, the ACG case was predicted 611 times accurately, with just 1 sample predicted as BC among the total 612 MFD test samples allocated for this case. Moreover, a comparison between the GB and DT algorithms displays that the AG case faced only 49

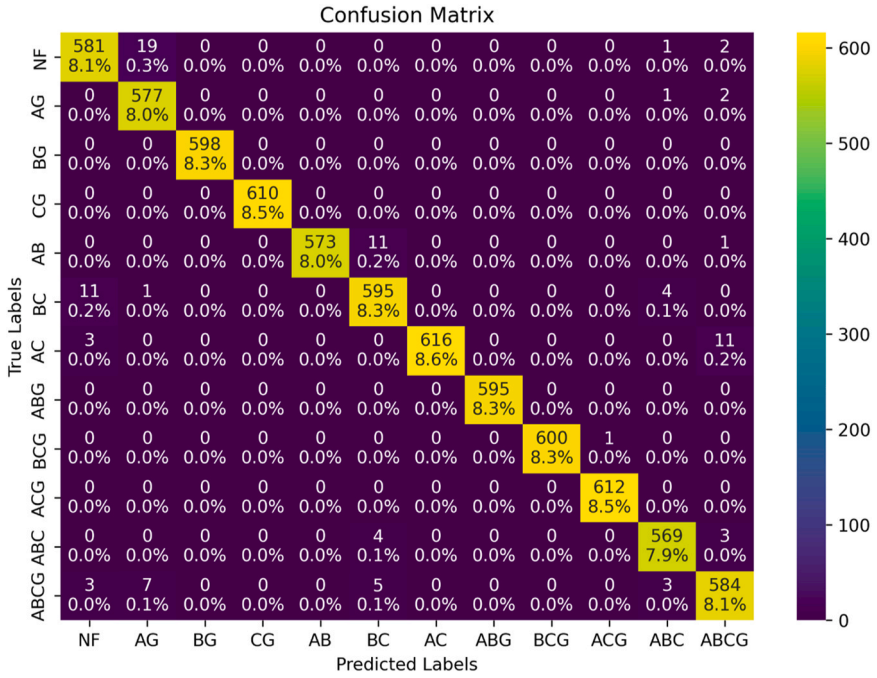

Fig. 25. CM of ANN algorithm for SC case scenario.

Table 14
Classification report of the ANN algorithm for OC case scenario.

S.no	Case Scenario	Precision	Recall	F1-score	Count	Accuracy (%)
1	NF	1.00	1.00	1.00	574	-
2	A_open	0.61	1.00	0.76	581	-
3	B_open	1.00	0.90	0.95	610	-
4	C_open	1.00	0.91	0.95	623	-
5	AB_open	1.00	0.90	0.95	613	-
6	BC_open	1.00	0.89	0.94	588	-
7	AC_open	1.00	0.91	0.95	602	-
8	ABC_open	1.00	0.89	0.94	611	-
9	-	**Average**			**Total**	-
10	-	0.951	0.925	0.93	4802	92.38

misclassifications with GB as compared to the 128 misclassifications witnessed in the DT algorithm.

4.4.4. Evaluation of GB model performance for OC case

The classification report displayed in Table 12, clearly indicates that the GB model faced challenges when predicting these cases such as A_open, B_open, and AB_open. In the case of BC_open, precision highlights that the model's predictions were 100 % true positive, while in the A_open case, the model's precision predictions were only 79 % true positive. Furthermore, the accuracy score of the GB algorithm is 90.55 %, suppressing the accuracy of the DT algorithm by 1.8 %.

The GB model's CM is displayed in Fig. 24. The CM indicates that the A_open case was mixed up with the B_open case because the classifier predicted the A_open samples as B_open samples 100 times. Notably, the BC_open case was predicted 587 times accurately, with just 1 sample predicted as ABC_open among the total 588 MFD test samples allocated for this case.

4.4.5. Evaluation of ANN model performance

The performance evaluation report of the ANN is presented in Table 13 where it is evident that this classifier performs efficiently in predicting all NF and SC fault case scenarios. The precision, recall, and F1 scores for specific case scenarios like BG, CG, ABG, BCG, and ACG present an outstanding accuracy of 100 % in predicting true positive samples. The previous algorithms faced difficulty in case scenarios such as NF, AG, AC, and ABCG but the ANN algorithm also exhibits high efficiency in these cases. The average of all metric measurements is almost equal to 0.987 demonstrating the algorithm's capability to predict maximum true positive rates and minimum false positives and false negatives. The algorithm's accuracy score is 98.71 % which also endorses the robust performance of the ANN.

The CM is shown in Fig. 25. The CM indicates that ANN predicted case scenarios like BG, CG, ABG, and ACG without any misclassification. However, the NF case scenario was mixed up with AG, as NF samples were incorrectly predicted as AG 19 times. In the complete NF case, accurate predictions occurred 581 times, with 22 test samples misclassified out of a total of 603 MFD test samples. By considering all case scenarios together, the analysis reveals that 7110 MFD test samples were predicted accurately from a complete pool of 7203 MFD test samples.

4.4.6. Evaluation of ANN model performance for OC case

The performance evaluation report of the ANN is presented in Table 14, where it is evident that it performed efficiently in NF, B_open, C_open and AC_open case scenarios. The precision score for specific case scenarios like NF, B_open, C_open, BC_open, AC_open and ABC_open present an outstanding accuracy of 100 % in predicting true positive samples. The algorithm's accuracy score is 92.38 % which also highlights the robust performance of the ANN.

The CM is shown in Fig. 26. The CM indicates that ANN predicted case scenarios like NF and A_open without any misclassification. However, the A_open case scenario was mixed up with all the remaining case scenarios. By considering all case scenarios together, the analysis reveals that 4501 MFD test samples were predicted accurately from a complete pool of 4802 MFD test samples.

4.4.7. Evaluation of RF model performance

The results achieved by the RF model are displayed in Table 15, consisting of performance evaluation metrics. The classification report highlights that the RF algorithm performed efficiently for all case scenarios in predicting true positive samples. For NF and SC fault case scenarios, except for AG, the precision indicates that the algorithm predicted 100 % true positive samples and 0 % false positives. The recall metric displays that the algorithm predicted 1 % false negatives for three

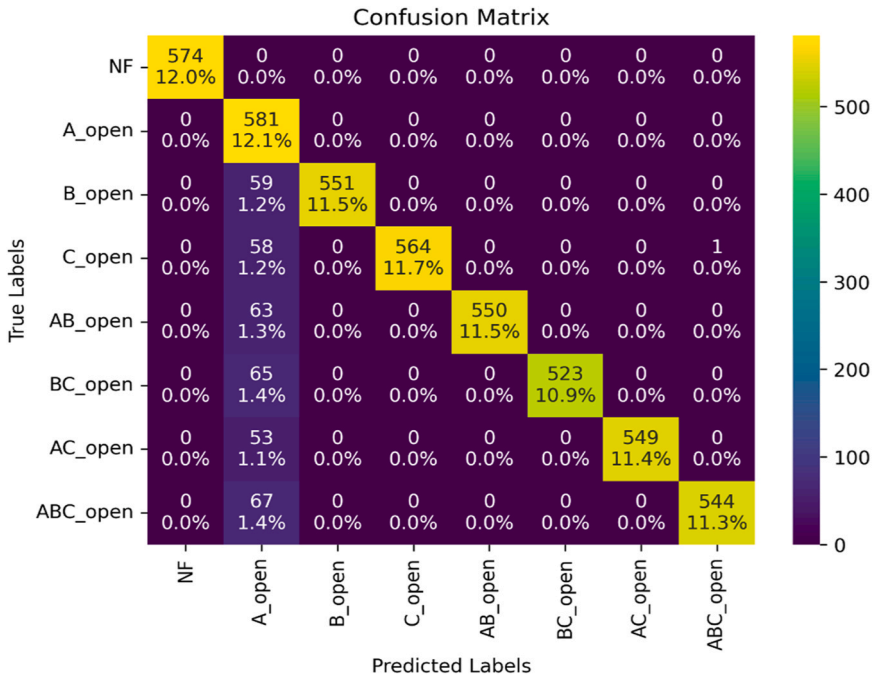

Fig. 26. CM of ANN algorithm for OC case scenario.

Table 15
Classification report of the RF algorithm for each case scenario.

S.no	Case Scenario	Precision	Recall	F1-score	Count	Accuracy (%)
1	NF	1.00	1.00	1.00	603	-
2	AG	0.98	0.99	0.99	580	-
3	BG	1.00	1.00	1.00	598	-
4	CG	1.00	1.00	1.00	610	-
5	AB	1.00	1.00	1.00	585	-
6	BC	1.00	1.00	1.00	611	-
7	AC	1.00	1.00	1.00	630	-
8	ABG	1.00	1.00	1.00	595	-
9	BCG	1.00	1.00	1.00	601	-
10	ACG	1.00	1.00	1.00	612	-
11	ABC	1.00	0.99	1.00	576	-
12	ABCG	1.00	0.99	0.99	602	-
13	-	**Average**			**Total**	-
14	-	0.9983	0.9975	0.9983	7203	99.74

cases AG, ABC, and ABCG, while it showed 0 % for the remaining fault cases. The average of these metrics is nearly equal to 100 % which indicates the enhanced performance of the RF algorithm. Furthermore, the algorithm's accuracy score is 99.74 % which outperformed all the preceding algorithms like DT, GB, and ANN.

The CM is presented in Fig. 27. The detailed view of the CM explains the efficient performance of the RF algorithm in accurately classifying the MFD test samples for each case scenario. Unlike previous algorithms, the RF algorithm handled misclassification challenges efficiently in cases like NF, AG, AC, and ABCG. Notably, the highest prediction errors occurred in ABCG, with only 5 out of the total 602 MFD test samples being misclassified. In case scenarios like BG, CG, AB, ABG, and ACG, the RF algorithm predicted all the MFD test samples accurately without any error. By considering all the case scenarios together, the analysis uncovers that out of a total of 7203 MFD test samples, the algorithm accurately predicted 7184 MFD test samples. It means only 19 samples were misclassified by the algorithm, which outclassed all the previously implemented algorithms in this study.

4.4.8. *Evaluation of RF model performance for OC case*

The classification report in Table 16 highlights that the RF algorithm performed efficiently for all case scenarios in predicting true positive samples. For NF, AB_open, AC_open, and ABC_open, the precision, recall, and F1-score are 100 % which means the RF model predicted 100 % true positive samples in these cases. Furthermore, the algorithm's accuracy score is 93.02 % which outperformed all the preceding algorithms like DT, GB, and ANN.

The detailed view of the CM in Fig. 28 explains the efficient performance of the RF algorithm in accurately classifying the MFD test samples for each case scenario. By considering all the case scenarios together, the analysis uncovers that out of a total of 4802 MFD test samples, the algorithm accurately predicted 4467 MFD test samples. It means only 335 samples were misclassified by the algorithm.

4.5. *Comparative analysis*

For comparative analysis, the accuracy results obtained from all the above-applied algorithms are displayed in Fig. 29 and Fig. 30. From this analysis, it's evident that the RF classifier outclassed all the other algorithms, mainly due to its use of Bootstrap Aggregating (bagging), an ensemble learning technique where multiple decision trees are combined to create a robust and highly accurate model. This technique ensures that the model is less prone to overfitting compared to individual decision trees. With a single decision tree, there is a higher chance of overfitting and misclassification, especially when handling complex data like MFD samples. However, using multiple trees in RF allows each tree to learn different aspects of the data, effectively minimizing overfitting. In our study, 50 trees were used, improving RF's ability to combine predictions from multiple trees. As a result, the RF classifier effectively handled the MFD data samples and learned intricate relationships within the data. Following this, the accuracy score of the ANN is slightly lower than the RF, but it performed well in detecting and classifying the faults. The ANN algorithm misclassified a few case scenarios by mixing up their MFD data samples. The GB algorithm is known for its robustness that sequentially creates weak learners, but its overfitting nature crafted complications for the GB classifier in understanding complex MFD data samples. Consequently, this led to low accuracy for the FDC task. Moreover, the least performer is the DT algorithm in

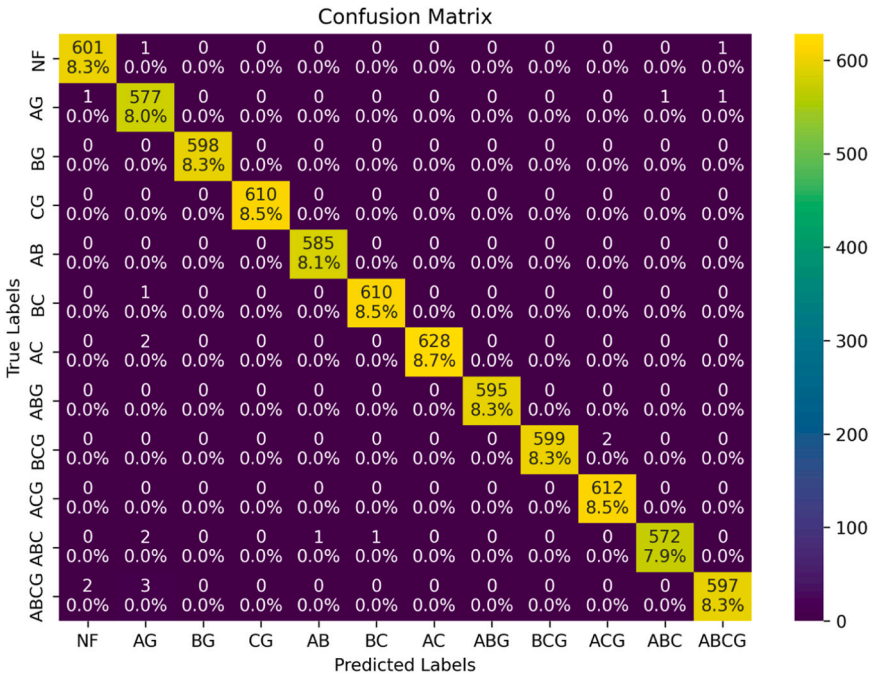

Fig. 27. CM of RF algorithm for SC case scenario.

Table 16
Classification report of the RF algorithm for OC case scenario.

S.no	Case Scenario	Precision	Recall	F1-score	Count	Accuracy (%)
1	NF	1.00	1.00	1.00	574	-
2	A_open	0.72	1.00	0.84	581	-
3	B_open	1.00	0.64	0.78	610	-
4	C_open	1.00	0.82	0.90	623	-
5	AB_open	1.00	1.00	1.00	613	-
6	BC_open	0.84	1.00	0.91	588	-
7	AC_open	1.00	1.00	1.00	602	-
8	ABC_open	1.00	1.00	1.00	611	-
9	-	**Average**			**Total**	-
10	-	0.945	0.932	0.928	4802	93.02

classifying the fault case scenarios accurately.

Apart from the accuracy score, the processing time of each ML algorithm provides intensive knowledge about the model's performance. The training time and testing time of DT, GB, ANN, and RF are presented in Table 17. The DT and RF algorithms have the shortest training and testing times when learning the MFD data and predicting fault case scenarios. The simplicity of the DT, which consists of a single tree structure, leads to faster processing time. In contrast, the RF algorithm simultaneously builds multiple decision trees and aggregates their predictions, making it slightly slower than DT but faster than ANN and GB. The GB algorithm takes more time to learn the MFD data due to its sequential nature, but it predicts unseen case scenarios swiftly. The training time of ANN is higher compared to the others due to its backpropagation process, which adjusts parameters to minimize error. However, it takes less time to predict fault case scenarios as it doesn't involve the backpropagation process.

Based on this analysis, we can conclude that a balance should be taken into consideration between accuracy score and processing time when selecting the appropriate ML model for a specific task. The RF algorithm stands out as the first choice for fault detection and classification tasks, having achieved a higher accuracy score of 99.74 % and 93.02 % for both fault cases while maintaining a balanced training and testing time. As discussed above, a major portion of the country's economy is allocated to the EPSN. Hence, it's important to protect the power equipment which can be achieved if the classifiers accurately predict the fault nature. For this reason, the ANN algorithm emerges as the second choice with an accuracy score of 98.71 % and 92.38 % for both fault cases, mainly due to its ability to handle complicated real-world scenarios efficiently, a factor that both GB and DT models lack.

5. Conclusion

This paper aims to enhance the safety and reliability of the EPSN by introducing an automatic FDC approach using the ML algorithms for the secondary distribution system, focusing on 18 different fault types. This research uses non-contact magnetic measurements to monitor the MFD behavior around the distribution transformer. An FEA-based electromagnetic analysis of a 630 KVA was conducted, where five magnetic points were placed to gather data on the MFD behavior around the steel tank of the secondary distribution transformer during SC and OC fault conditions. A dataset consisting of 28,809 MFD SC samples and 19,206 MFD OC samples is used as input for the training of four ML algorithms to assess the performance of each algorithm in handling these complex MFD measurements. The RF algorithm showed the best performance among these four algorithms with 99.74 % and 93.02 % accuracy in correctly detecting and classifying fault scenarios followed by ANN at 98.71 % and 92.38 %, GB at 95.63 % and 90.55 %, and DT at 92.85 % and 88.75 %. This research highlights the advantages of using MFD measurements collected through non-contact magnetic points, which increase safety and fault detection accuracy compared to traditional CT and PT. The core saturation of CT also limits its accuracy, whereas the non-contact nature of these magnetic measurements eliminates the need for physical contact with potentially hazardous equipment, thus reducing risks. Additionally, the ML models efficiently processed these intricate MFD measurements, providing an effective solution for FDC to protect the transformer during fault times. Future work will focus on addressing noise interference and transient effects, which were not considered in this study.

Authors role and contributions

Regarding roles, Sufiyan Rao as the first is responsible for formal analysis, investigation, methodology, software, visualization, and

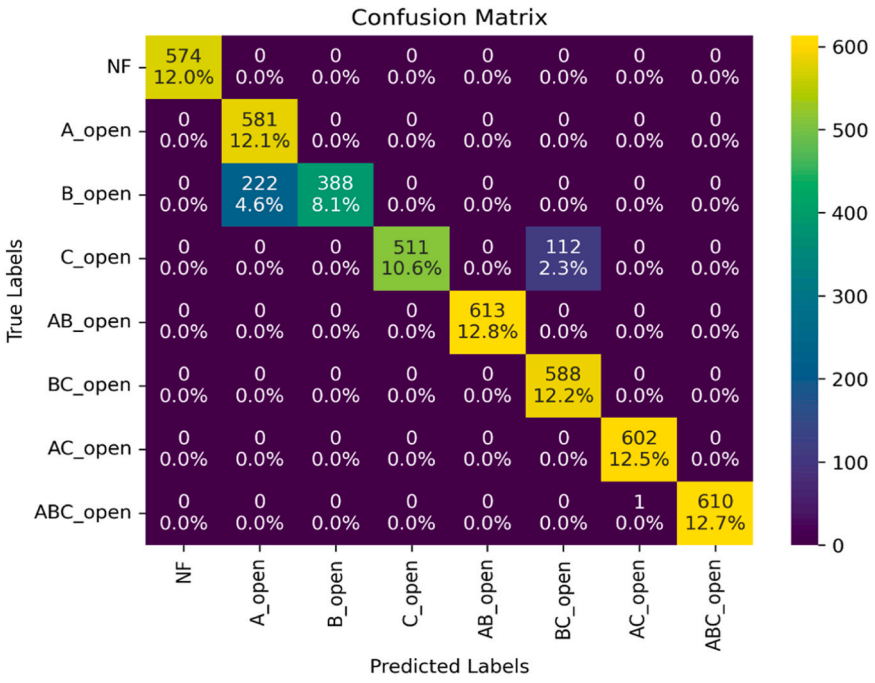

Fig. 28. CM of RF algorithm for OC case scenario.

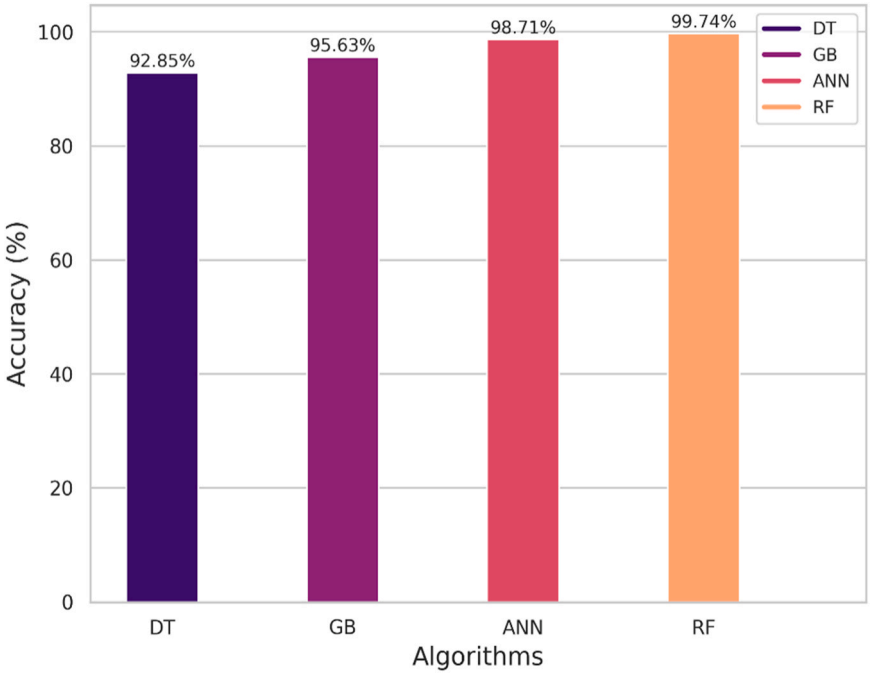

Fig. 29. Comparing accuracy scores of DT, GB, ANN, and RF for SC cases.

writing – original draft while Syed Ali Abbas Kazmi (second and corresponding author) provide support in conceptualization, investigation, supervision, writing – review & editing. Muhammad Zubair Iftikhar (third author) is responsible for data formal analysis, funding acquisition, methodology, validation, writing - editing and reviewing the draft of paper. Dr. THAMER A. H. ALGHAMDI (fourth author) provides funding acquisition, investigation, project administration, and software. Mohammed Alenezi(fifth and corresponding Author) have done funding acquisition, project administration, validation, writing – review & editing. All the authors have added considerable contribution towards paper completion and numerical checks.

CRediT authorship contribution statement

Rao Sufiyan: Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis. **Alghamdi Thamer A. H.:** Software, Project administration, Investigation, Funding acquisition. **Alenezi Mohammed:** Writing – review & editing, Validation, Project administration, Funding acquisition. **Kazmi Syed Ali Abbas:** Writing – review & editing, Supervision, Investigation, Conceptualization. **Iftikhar Muhammad Zubair:** Writing – review & editing, Validation, Methodology, Funding acquisition, Formal analysis.

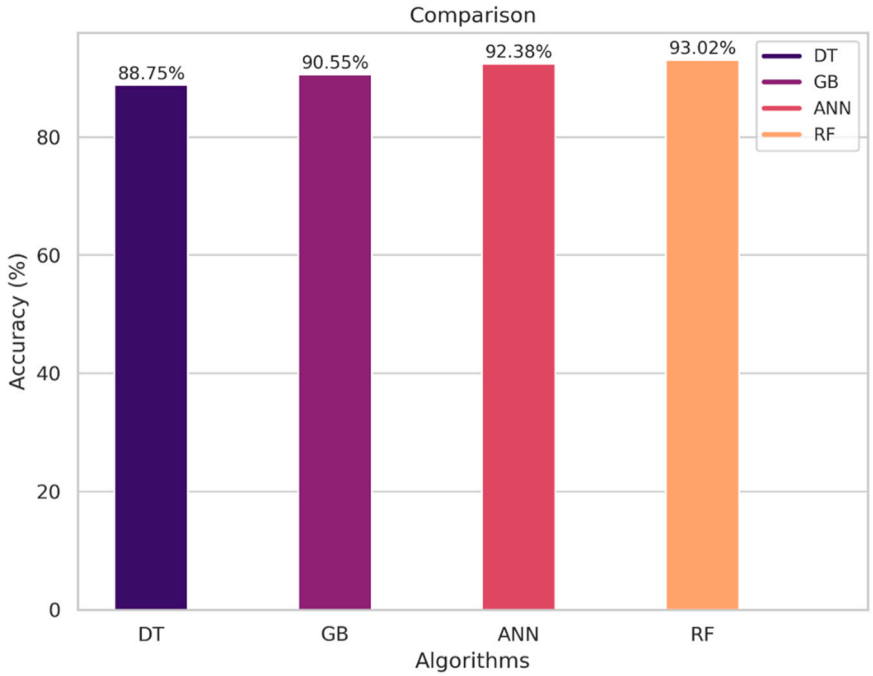

Fig. 30. Comparing accuracy scores of DT, GB, ANN, and RF for OC cases.

Table 17
Train and Test time for DT, GB, ANN, and RF algorithms.

S.no		Algorithm	Train Time (s)	Test Time (s)
1	SC case	Decision Tree	0.194	0.0028
2	scenarios	Gradient Boosting	78.41	0.1267
3		Artificial Neural Network	115.41	0.4875
4		Random Forest	3.363	0.0759
5	OC case	Decision Tree	0.036	0.0019
6	scenarios	Gradient Boosting	2.296	0.0509
7		Artificial Neural Network	41.48	0.2656
8		Random Forest	1.250	0.0413

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors do not have permission to share data.

References

2016 IEEE Power and Energy Society General Meeting (PESGM): date, 17-21 July 2016. IEEE, 2016.
48th North American Power Symposium: Denver, Colorado, September 18-20, 2016: proceedings. IEEE, 2016.
Abbas, J., et al., 2024. Classification and comprehension of software requirements using ensemble learning. Comput., Mater. Contin. 80 (2), 2839–2855. https://doi.org/10.32604/cmc.2024.052218.
Alenezi, M., Anayi, F., Packianather, M., Shouran, M., Dec. 2024. Enhancing transformer protection: a machine learning framework for early fault detection. Sustainability 16 (23). https://doi.org/10.3390/su162310759.
Alpsalaz, F., Mamiş, M.S., Oct. 2024. Detection of arc faults in transformer windings via transient signal analysis. Appl. Sci. 14 (20). https://doi.org/10.3390/app14209335.
Balan, A., Srujan, T.L., Manitha, P.V., Deepa, K., 2023. Detection and analysis of faults in transformer using machine learning, IDCIoT 2023 - International Conference on Intelligent Data Communication Technologies and Internet of Things, Proceedings, Institute of Electrical and Electronics Engineers Inc.47748210.1109/IDCIoT56793.2023.10052786.
Bompard, E., Huang, T., Wu, Y., Cremenescu, M., Sep. 2013. Classification and trend analysis of threats origins to the security of power systems. Int. J. Electr. Power Energy Syst. 50 (1), 50–64. https://doi.org/10.1016/J.IJEPES.2013.02.008.
Chen, K.L., Jul. 2022. Intelligent contactless current measurement for overhead transmission lines. IEEE Trans. Smart Grid 13 (4), 3028–3037. https://doi.org/10.1109/TSG.2022.3151863.
Choi, M.-S., Lee, S.-J., Lee, D.-S., Jin, B.-G., Jan. 2004. A new fault location algorithm using direct circuit analysis for distribution systems. IEEE Trans. Power Deliv. 19 (1), 35–41. https://doi.org/10.1109/TPWRD.2003.820433.
Dawood, K., Sönmez, O., Kömürgöz, G., Apr. 2023. Optimization and analysis of tapping position on leakage reactance of a two winding transformer. Alex. Eng. J. 69, 363–369. https://doi.org/10.1016/j.aej.2023.02.006.
De Oliveira Neto, J.A., Sartori, C.A.F., Junior, G.M., 2021. Fault location in overhead transmission lines based on magnetic signatures and on the extended Kalman Filter. IEEE Access 9, 15259–15270. https://doi.org/10.1109/ACCESS.2021.3050211.
Ehsanifar, A., et al., Sep. 2021. Transformer inter-turn winding fault detection based on no-load active power loss and reactive power. Int. J. Electr. Power Energy Syst. 130. https://doi.org/10.1016/j.ijepes.2021.107034.
El Mrabet, Z., Sugunaraj, N., Ranganathan, P., Abhyankar, S., Jan. 2022. Random forest regressor-based approach for detecting fault location and duration in power systems. Sensors 22 (2). https://doi.org/10.3390/s22020458.
Fahim, S.R., Sarker, S.K., Muyeen, S.M., Das, S.K., Kamwa, I., Dec. 2021. A deep learning based intelligent approach in detection and classification of transmission line faults. Int. J. Electr. Power Energy Syst. 133. https://doi.org/10.1016/j.ijepes.2021.107102.
Ferreira, V.H., Zanghi, R., Fortes, M.Z., Gomes, S., Alves da Silva, A.P., Aug. 2020. Probabilistic transmission line fault diagnosis using autonomous neural models. Electr. Power Syst. Res. 185. https://doi.org/10.1016/j.epsr.2020.106360.
Fonseca, G.A., Ferreira, D.D., Costa, F.B., Almeida, A.R., Apr. 2022. Fault classification in transmission lines using random forest and notch filter. J. Control Autom. Electr. Syst. 33 (2), 598–609. https://doi.org/10.1007/s40313-021-00844-4.
Gangwar, A.K., Shaik, A.G., Jul. 2023. k-Nearest neighbour based approach for the protection of distribution network with renewable energy integration. Electr. Power Syst. Res. 220. https://doi.org/10.1016/j.epsr.2023.109301.
Gonen, T., 2015. Electric Power Distribution Engineering. CRC Press. https://doi.org/10.1201/b16455.
Goni, M.O.F., et al., Mar. 2023. Fast and accurate fault detection and classification in transmission lines using extreme learning machine. e-Prime - Adv. Electr. Eng., Electron. Energy 3. https://doi.org/10.1016/j.prime.2023.100107.
Gururajapathy, S.S., Mokhlis, H., Illias, H.A., 2017. Fault location and detection techniques in power distribution systems with distributed generation: a review. Elsevier Ltd. https://doi.org/10.1016/j.rser.2017.03.021.
Gutierrez-Rojas, D., Christou, I.T., Dantas, D., Narayanan, A., Nardelli, P.H.J., Yang, Y., Mar. 2022. Performance evaluation of machine learning for fault selection in power transmission lines. Knowl. Inf. Syst. 64 (3), 859–883. https://doi.org/10.1007/s10115-022-01657-w.
He, A., Jiao, Z., Li, Z., Liang, Y., Oct. 2023. Discrimination between internal faults and inrush currents in power transformers based on the discriminative-feature-focused CNN. Electr. Power Syst. Res. 223. https://doi.org/10.1016/j.epsr.2023.109701.

Hussain, M.T., Sarwar, A., Tariq, M., Urooj, S., BaQais, A., Hossain, M.A., Jul. 2023. An evaluation of ANN algorithm performance for MPPT energy harvesting in solar PV systems. Sustainability 15 (14). https://doi.org/10.3390/su151411144.

Jin, M., et al., Apr. 2024. A path of power transformers failure under multiple short-circuit impacts. Electr. Power Syst. Res. 229. https://doi.org/10.1016/j.epsr.2024.110216.

Khadse, C., Patharkar, A.A., Chaudhari, B.S., 2021. Electromagnetic field and artificial intelligence based fault detection and classification system for the transmission lines in smart grid. Energy Sources, Part A: Recovery Util. Environ. Eff. https://doi.org/10.1080/15567036.2021.1948637.

Kwasi Anane, P.O., Huang, Q., Bamisile, O., Ayimbire, P.N., Dec. 2021. Fault location in overhead transmission line: a novel non-contact measurement approach for traveling wave-based scheme. Int. J. Electr. Power Energy Syst. 133. https://doi.org/10.1016/j.ijepes.2021.107233.

Li, Y., Xu, Q., Lu, Y., Nov. 2021. Electromagnetic force analysis of a power transformer under the short-circuit condition. IEEE Trans. Appl. Supercond. 31 (8). https://doi.org/10.1109/TASC.2021.3107799.

Malik, S.Z., Habib Khawaja, A., Janjua, A.K., Kazim, M., Oct. 2020. A contactless method for unbalanced loading detection in power distribution lines by magnetic measurements. IEEE Trans. Instrum. Meas. 69 (10), 7472–7483. https://doi.org/10.1109/TIM.2020.2983339.

Mamuya, Y.D., Lee, Y.Der, Shen, J.W., Shafiullah, M., Kuo, C.C., Jul. 2020. Application of machine learning for fault classification and location in a radial distribution grid. Appl. Sci. 10 (14). https://doi.org/10.3390/app10144965.

Moloi, K., Ndlela, N.W., Davidson, I.E., Dec. 2022. Fault classification and localization scheme for power distribution network. Appl. Sci. 12 (23). https://doi.org/10.3390/app122311903.

Nhat-Duc, H., Van-Duc, T., Apr. 2023. Comparison of histogram-based gradient boosting classification machine, random Forest, and deep convolutional neural network for pavement raveling severity classification. Autom. Constr. 148, 104767. https://doi.org/10.1016/j.autcon.2023.104767.

Ogar, V.N., Hussain, S., Gamage, K.A.A., Jul. 2022. Transmission line fault classification of multi-dataset using CatBoost classifier. Signals 3 (3), 468–482. https://doi.org/10.3390/signals3030027.

ÖZÜPAK, Y., Jan. 2021. Performing structural design and modeling of transformers using ANSYS-Maxwell. Brill. Eng. 2 (2), 38–42. https://doi.org/10.36937/ben.2021.002.005.

Rafique, F., Fu, L., Mai, R., Oct. 2021. End to end machine learning for fault detection and classification in power transmission lines. Electr. Power Syst. Res. 199. https://doi.org/10.1016/j.epsr.2021.107430.

Singh, M.R., Chopra, T., Scholar, M.Tech, 2015. Fault classification in electric power transmission lines using support vector machine. Int. J. Innov. Res. Sci. Technol. | 1. ⟨www.ijirst.org⟩ ([Online]. Available).

Thomas, J.B., Chaudhari, S.G., Shihabudheen, K.V., Verma, N.K., 2023. CNN-based transformer model for fault detection in power system networks. IEEE Trans. Instrum. Meas. 72. https://doi.org/10.1109/TIM.2023.3238059.

Uddin, M.S., et al., Nov. 2022. On the protection of power system: transmission line fault analysis based on an optimal machine learning approach. Energy Rep. 8, 10168–10182. https://doi.org/10.1016/j.egyr.2022.07.163.

Usman, M.U., Ospina, J., Faruque, Md.O., Jan. 2020. Fault classification and location identification in a smart DN using ANN and AMI with real-time data. J. Eng. 2020 (1), 19–28. https://doi.org/10.1049/joe.2019.0896.

Venkata, P., Pandya, V., Vala, K., Sant, A.V., Dec. 2022. Support vector machine for fast fault detection and classification in modern power systems using quarter cycle data. Energy Rep. 8, 92–98. https://doi.org/10.1016/j.egyr.2022.10.279.

Vidal, J., Jha, S., Liang, Z., Delgado, E., Deneke, B.S., Shasha, D., Oct. 2024. Dynamic decision trees. Knowledge 4 (4), 506–542. https://doi.org/10.3390/knowledge4040027.

Yuan, J., Jiao, Z., May 2023. Faulty feeder detection for single phase-to-ground faults in distribution networks based on patch-to-patch CNN and feeder-to-feeder LSTM. Int. J. Electr. Power Energy Syst. 147. https://doi.org/10.1016/j.ijepes.2022.108909.

Zheng, J., Liu, Y., Ge, Z., Jun. 2022. Dynamic ensemble selection based improved random forests for fault classification in industrial processes. IFAC J. Syst. Control 20, 100189. https://doi.org/10.1016/j.ifacsc.2022.100189.