technologies

MDPI

*Article*

# Classification and Prediction of Chip Diameter in High-Power Semiconductor Devices Through Electrical Parameters Using Machine Learning

**Fawad Ahmad [1,2], Luis Vaccaro [1,*], Armel Asongu Nkembi [3], Mario Marchesoni [1] and Federico Portesine [2]**

[1] Department of Naval, Electrical, Electronic and Telecommunications Engineering (DITEN), University of Genova, 16145 Genova, Italy; fawad.ahmad@edu.unige.it (F.A.); marchesoni@unige.it (M.M.)
[2] Poseico Power Electronics S.p.A., 16153 Genova, Italy; portesine.federico@poseico.com
[3] Department of Engineering and Architecture, University of Parma, 43121 Parma, Italy; armelasongu.nkembi@unipr.it
* Correspondence: luis.vaccaro@unige.it

**Abstract**

The applications of machine learning (ML) are rapidly expanding across various fields to reduce their complexity and improve efficiency. In power electronics, where design tasks require complex analytical computations and accurate predictions, ML techniques are becoming increasingly important for reliable device design and robust manufacturing. With the growing demand of power density of high-power semiconductor devices, such as diodes and thyristors, the electrical parameters critically influence the physical dimensions and geometry of the chip. In this article, a comprehensive survey of high-power thyristors is conducted, analyzing the influence of chip diameter and thickness on both electrical and thermal performance. Moreover, a dedicated dataset is developed by extracting electrical parameters from the leading semiconductor manufacturer's datasheet of multiple models. Furthermore, multiple machine learning algorithms, including Artificial Neural Networks (ANNs), Support Vector Machines (SVMs), and Ensemble methods, are implemented and compared. The developed models provide manufacturers with efficient predictive tools to determine optimal chip dimensions for specific power ratings, thereby supporting efficient and reliable device design.

**Keywords:** machine learning in semiconductors; chip diameter prediction; high power semiconductors; electrical parameters analysis; data-driven semiconductor design

check for updates

Academic Editor: Jianxiong Zhu

Received: 7 November 2025
Revised: 30 November 2025
Accepted: 2 December 2025
Published: 4 December 2025

**Citation:** Ahmad, F.; Vaccaro, L.; Nkembi, A.A.; Marchesoni, M.; Portesine, F. Classification and Prediction of Chip Diameter in High-Power Semiconductor Devices Through Electrical Parameters Using Machine Learning. *Technologies* **2025**, *13*, 567. https://doi.org/10.3390/technologies13120567

**Copyright:** © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

## 1. Introduction

In recent years, the integration of machine learning and Artificial Intelligence (AI) has significantly expanded across various domains; however, their application in power electronics is still evolving. Before the advent of these technologies, semiconductor manufacturers focused on developing devices with reduced size and higher power densities. These advancements have optimized the design and operation of power devices such as power converters [1] and motor drives [2], enabling them to handle higher currents and power levels. Despite this progress, efficient power dissipation remains a challenge, as it is crucial to maintain safe operation within the rated junction temperature limits to prevent device failure. In high-power semiconductor devices, a decrease in diameter may increase resistance, leading to reduced current flow and overheating. Conversely, a larger diameter improves heat dissipation by reducing thermal resistance through distribution over a

larger surface area, thereby lowering the risk of failure. Manufacturers like ABB, Infineon, and Poseico Power Electronics have developed high-power semiconductor devices such as press-pack diodes and thyristors with enhanced current and voltage capabilities. ABB, for instance, progressed from a 78 mm thyristor (4310 A, 1800 V, 125 °C) to larger 143 mm variants for traction, energy, and industrial applications, offering ratings of 8500 V/5499 A and 7200 V/6220 A at 110 °C. Similarly, Infineon designed a 5200 V/3400 A thyristor for crowbar and pulsed power applications with full blocking capability at 50/60 Hz over a wide temperature range. Moreover, Poseico Electronics developed multiple devices with diameters up to 101 mm, voltage ratings from 800 V to 6000 V, and average currents from 3344 A to 8099 A, operating at 120 °C for industrial and traction applications. Increasing chip diameter generally reduces electrical and thermal stresses, enabling higher power and current capabilities; this can result in non-uniform current distribution and localized hotspot formation. Hence, the selection of an optimal diameter remains a critical preliminary design consideration in high-power semiconductor devices [3].

Manufacturers typically rely on conventional analytical formulations to determine these parameters on the basis of material properties, thermal limits, and electrical performance requirements. However, such formula-based approaches may not fully capture complex inter-dependencies among multiple design variables such as on-state voltage drop, thermal impedance, and transient current handling capability, particularly when operating conditions vary between applications. In this context, data-driven prediction and classification methods offer a complementary pathway, enabling designers to explore the optimal diameter within the feasible operational envelope and tailor device specifications more precisely to target application requirements. Machine learning techniques primarily rely on data-driven models trained on specific scenarios, eliminating the need to manually formulate complex analytical expressions and compute design parameters. These models possess self-learning capabilities, achieving high accuracy even with larger datasets while maintaining computational efficiency. Machine learning can be broadly categorized into supervised and unsupervised learning. Supervised learning is used to develop models in which predictions are made based on both input and output data, enabling accurate forecasting. It can be further divided into classification and regression; classification assigns data to predefined categories, while regression predicts numerical values based on previous data patterns. These models are widely employed in power electronics and semiconductor applications, transforming processes from design and manufacturing to testing and optimization [4,5]. Recent studies have demonstrated the effectiveness of machine learning techniques in modeling [6], prediction [5], and fault diagnosis of IGBTs and power electronic devices [7]. In [8], junction temperature and device degradation were estimated by predicting the on-state voltage using K-NN, SVM, and decision tree (DT) models. ML approaches have also been applied to predict the lifetime of IGBT modules, achieving higher accuracy than conventional lifetime models [7]. Data-driven convolutional neural network (CNN) methods have enabled reliable classification of IGBT failure modes under varying operating conditions, highlighting their potential for condition monitoring and predictive maintenance [9]. Furthermore, [10] proposed a hybrid approach combining compact device models with ANNs to predict IGBT output characteristics, obtaining Hefner static model parameters such as $V_{Gate}$ and $V_{Anode}$ from structural parameters. Comparative evaluations of machine learning algorithms, including Random Forest, backpropagation neural networks, and decision trees, have also been conducted for fault diagnosis of power electronic devices in converter systems, demonstrating their applicability across different operational scenarios [11]. These advances demonstrate the versatility of ML across semiconductor manufacturing, device modeling, and predictive analysis. Despite significant progress, most existing studies emphasize reliability assessment, failure mode classifica-

tion, and prediction of electrical parameters. However, modeling frameworks capable of predicting and classifying fundamental design parameters, such as chip diameter in high-power devices (e.g., diodes and thyristors), remain largely unexplored. Moreover, the existing literature shows that datasets available for training and evaluating machine learning model performance are limited for various applications in the semiconductor manufacturing industry and power electronics. In particular, this includes datasets related to chip diameter, which can significantly optimize design and cost. The development of such domain-specific datasets would provide deeper insights into model behavior under varying diameters and electrical parameters, thereby assisting manufacturers in achieving optimized device designs and optimal chip diameter for specific applications. Additionally, Shapley analysis is incorporated to highlight the importance of crucial parameters and to identify the impact of each parameter on model outputs. This can reduce design complexity by guiding parameter tuning towards optimal diameter configurations. Consequently, all influential parameters can be prioritized according to design requirements during device development. This work may enable manufacturers and researchers to optimize parameter selection for specific application requirements. The applicability of the developed dataset from the manufacturer datasheet extends well beyond the scope of this study, providing a foundational resource for machine learning models to predict and classify a variety of critical parameters, such as thermal resistance and average current, in high-power semiconductor devices. Furthermore, the proposed methodology can be utilized to design or evaluate the performance of other types of power electronics components, such as IGBTs, diodes, etc. The main contributions of this paper are as follows:

1. A comprehensive market survey of high-power thyristors is conducted, analyzing the role of chip diameter and current capacity in device performance.
2. A dedicated dataset is developed by extracting electrical parameters from datasheets of multiple leading semiconductor manufacturers.
3. The influence of chip diameter and thickness on electrical and thermal performance is investigated.
4. Multiple machine learning algorithms are implemented and compared, including ANN, SVM, and Ensemble methods.

The rest of paper is organized as follows. Section 2 presents a market survey of high-power devices, discussing available thyristor and diode variants. In Section 3, the influence of material geometry on current distribution is analyzed. Section 4 describes the proposed methodology, with the corresponding results presented in Section 5. Finally, Section 6 concludes the study by summarizing the key findings and their implications.

## 2. Market Analysis of Power Devices: Electrical Characteristics with Respect to Diameter

The manufacturing of high-power semiconductor devices represents an advanced and evolving field, driving innovation across multiple industries by enabling efficient solutions for a wide range of applications. Key device types include diodes, transistors (such as MOSFETs and IGBTs), and thyristors. Diodes are two-terminal devices that permit current flow in only one direction, offering simple and robust functionality. In contrast, transistors and thyristors, featuring three or more terminals, enable signal amplification and switching capabilities, making them suitable for more complex power management tasks. Each of these devices operates based on distinct physical mechanisms, offering specific advantages and trade-offs. However, their primary objective is to maximize energy efficiency in high-power electronic systems. The demand for efficient conversion and growing industrialization has driven the dominance of power semiconductors, according to Market.us; the global power semiconductor market is expected to increase to USD 90.5 billion by 2033, up from USD

54.0 billion in 2023 [12], with Asia–Pacific countries holding the largest share, as depicted in Figure 1. Another report projects a 9.1% increase during 2025 and 2034, and highlights that the impact of AI has contributed to the semiconductor market growth, reaching USD 321.1 billion by 2034 [13]. The largest share of total power devices use belongs to applications such as Industrial (drives, converters, renewable) [14], Automotive (including EV traction applications) [15], IT, Telecom & Consumer Electronics [16], and Others (military, aerospace, other) as shown in Figure 2.

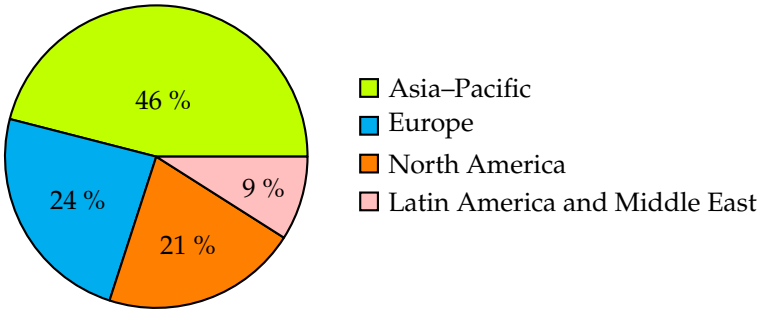



**Figure 1.** Global market share: Power semiconductors.

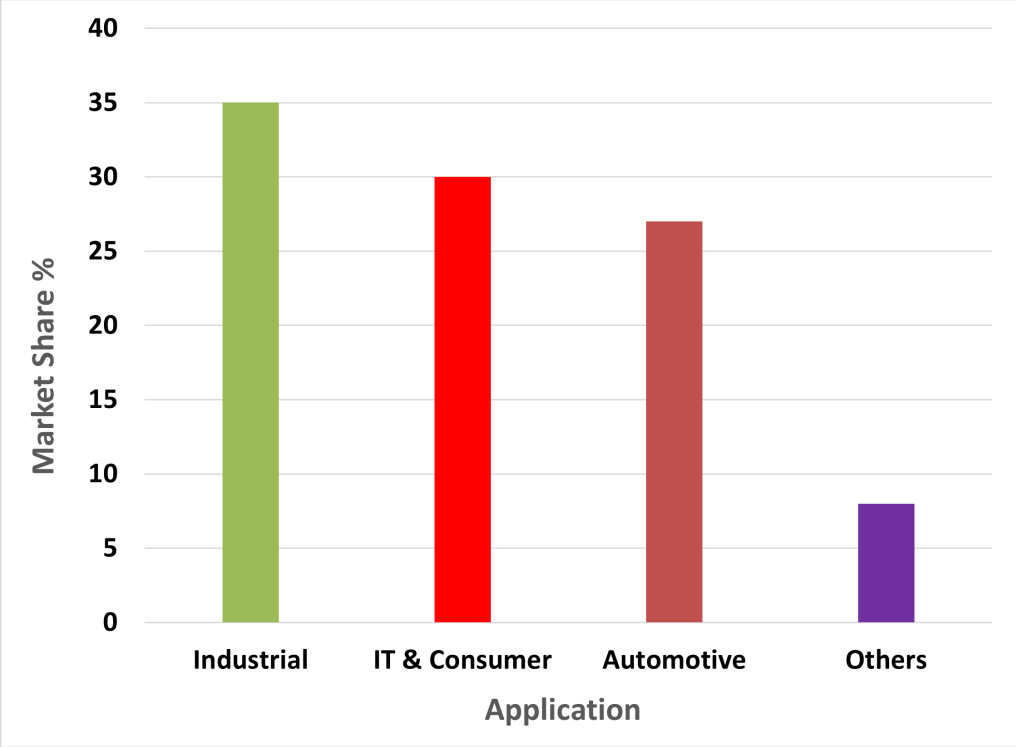



**Figure 2.** Estimated market share of high-power devices by application.

Recent advancements in power semiconductor device design have seen several manufacturers develop devices capable of supporting significantly higher current ratings as evident from Table 1. For instance, ABB has achieved an average current rating of up to 7612 A, Infineon up to 6110 A, Dynex up to 8750 A, and Poseico Power Electronics S.p.A. up to 8100 A. While these increases in current handling are notable, it is important to highlight that manufacturers have also optimized key parameters, particularly $V_{RRM}$ (peak reverse repetitive voltage) and chip diameter to facilitate these improvements. The diameter of the chip exhibits a directly proportional relationship with current capacity. In contrast, increasing $V_{RRM}$ tends to reduce current capability, as observed in Figure 3. The exponential curves in Figure 3 further confirm the linear dependence of current on chip diameter, while

also highlighting the decreasing trend of $V_{RRM}$ with increasing average current. For example, ABB offers devices with current ratings ranging from 5000 A to 6880 A over a $V_{RRM}$ span of 1800 V to 8500 V. The device 5STP 42L1800, featuring a 78 mm chip diameter and $V_{RRM}$ of 1800 V, achieves a current rating of 5356 A. Comparatively, multiple ABB devices with a larger 100 mm chip diameter but different $V_{RRM}$ values (1800 V, 2800 V, and 4200 V) demonstrate a clear trend: the 5STP 50Q1800 ($V_{RRM}$ = 1800 V) attains the highest current of 7612 A. ABB has further increased chip diameters to 110 mm and 143 mm. Devices with a 110 mm diameter use $V_{RRM}$ values of 5200 V and 6500 V, while those with 143 mm chips use 7200 V and 8500 V. Despite the increase in chip size, devices with lower $V_{RRM}$ values consistently achieve higher current ratings, such as those with 5200 V and 7200 V for 110 mm and 143 mm diameters, respectively. Furthermore, devices with 143 mm diameters demonstrate lower thermal resistance, contributing to improved current handling performance. Infineon Technologies has introduced devices with chip diameters ranging from 86 mm to 115 mm. The highest current rating of 6110 A is attained by the device T4771N22TOF PR, featuring a 100 mm chip diameter and $V_{RRM}$ of 2200 V. In contrast, a device with a larger 115 mm diameter and a higher $V_{RRM}$ of 5200 V (T4021N52TOH) achieves a lower average current rating of 5460 A. Several other Infineon devices with 100 mm chips and $V_{RRM}$ of 5200 V also demonstrate reduced current capabilities, reinforcing the inverse correlation between $V_{RRM}$ and current. Dynex Semiconductors has developed the DCR5240H52, which features a 115 mm chip and 5200 V $V_{RRM}$, achieving 5685 A. Notably, Dynex has also used 115 mm diameters with varied $V_{RRM}$ values of 4200 V and 6500 V for the DCR6140H42 and DCR4420H65, yielding current ratings of 6485 A and 4774 A, respectively. The highest current capability among Dynex devices is 8750 A, attained by the DCR7610H28, which uses a 110 mm chip and a relatively moderate $V_{RRM}$ of 2800 V. This again emphasizes the importance of balancing $V_{RRM}$ and chip size for maximizing current. Similarly, Poseico Power Electronics S.p.A. has developed several devices with a consistent chip diameter of approximately 101 mm, paired with varying $V_{RRM}$ levels: 800 V, 2900 V, 3200 V, 3400 V, and 5200 V. The device AT908, with the lowest $V_{RRM}$ of 800 V, achieves the highest current of 8100, while the AT975, with a $V_{RRM}$ of 5200 V, records the lowest current value of 4624 A among the group. Across all manufacturers, maximum current ratings are predominantly achieved with chip diameters around 100 mm and relatively lower $V_{RRM}$ values. While increasing chip size generally enhances current capacity, higher $V_{RRM}$ levels introduce trade-offs due to increased electric field stresses and associated thermal limitations. As silicon-based technologies approach their physical and thermal performance limits, particularly in terms of power density, manufacturers are increasingly investigating alternative materials. Among these, wide-bandgap semiconductors such as silicon carbide (SiC) and gallium nitride (GaN) are gaining significant attention. These materials offer superior electrical properties, including higher breakdown voltage, faster switching speed, and improved thermal conductivity, making them ideal candidates for next-generation high-power applications.

**Table 1.** Comparative summary of high-current thyristor devices available in market.

| Manufacturer | Device Code | Diameter | Current | $V_{RRM}$ |
|---|---|---|---|---|
| ABB | 5STP 50Q1800 | 100 | 7612 | 1800 |
| | 5STP 45Q2800 | 100 | 6879 | 2800 |
| | 5STP 38Q4200 | 100 | 5326 | 4200 |
| | 5STP 52U5200 | 110 | 6232 | 5200 |
| | 5STP 48Y7200 | 143 | 6220 | 7200 |
| | 5STP 45Y8500 | 143 | 5494 | 8500 |
| | 5STP 42L1800 | 78 | 5356 | 1800 |

**Table 1.** *Cont.*

| Manufacturer | Device Code | Diameter | Current | $V_{RRM}$ |
|---|---|---|---|---|
| Dynex | DCR7610H28 | 100 | 8743 | 2800 |
| | DCR6140H42 | 115 | 6485 | 4200 |
| | DCR5240H52 | 115 | 5685 | 5200 |
| | DCR4420H65 | 115 | 4774 | 6500 |
| | DCR4940V18 | 73 | 5572 | 1800 |
| | DCR4060V22 | 73 | 4562 | 2200 |
| Infineon | T4771N22TOF PR | 100 | 6110 | 2200 |
| | T4021N52TOH | 115 | 5460 | 5200 |
| | T3441N52TOH | 100 | 4410 | 5200 |
| | T3801N36TOF VT | 100 | 5370 | 3600 |
| | T3160N18TOF VT | 75 | 4620 | 1800 |
| | T2810N22TOF VT | 75 | 4070 | 2200 |
| Poseico | AT908 | 101 | 8099 | 800 |
| | AT906 | 101 | 7114 | 800 |
| | AT960 | 101 | 6057 | 3200 |
| | AT975 | 101 | 4624 | 5200 |
| | AT706 | 80 | 4947 | 800 |
| | AT970 | 101 | 4780 | 3400 |

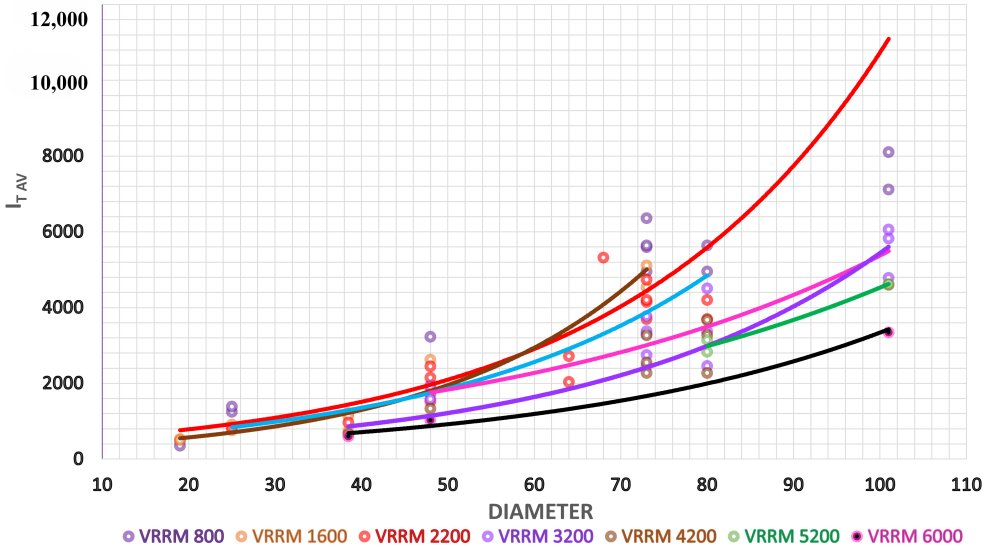



**Figure 3.** Impact of diameter on electrical characteristics average current and $V_{RRM}$.

## 3. Influence of Material Physical Dimension on Device Performance

The geometry and dimensions of materials significantly influence thermal management, structural integrity, and electrical properties in semiconductor devices, thereby affecting thermal resistance and, in turn, power dissipation. In particular, thickness and diameter play critical roles in defining the threshold voltage and slope resistance. The effects of these parameters were simulated following our previous work [17] for various materials under different operating conditions, as detailed below, providing design guidance for optimizing diameters in high-power semiconductor devices.

### *3.1. Effect of Change in Material Thickness on Device Performance*

This section presents the influence of material thickness on thermal impedance, for silicon, copper, and molybdenum. Simulations capture both transient and steady-state behavior, exhibiting the direct relationship between thickness and thermal resistance as shown in

Figure 4a–c. Both copper and molybdenum exhibit negligible variation in transient response at shorter times. However, at a steady state, copper shows greater deviation with thickness due to its higher thermal conductivity, while molybdenum demonstrates comparatively smaller differences. In contrast, silicon shows nonlinear deviations at short time scales as shown in Figure 4d: for pulses below 10 ms, thicker materials exhibit lower transient thermal impedance when the chip diameter exceeds approx 0.8 mm, as illustrated in Figure 5. An increase in thickness leads to higher thermal resistance, which limits power-handling capability and causes elevated temperatures under high current and voltage conditions. This compromises reliability, accelerates device degradation and failure, and requires enhanced cooling strategies.

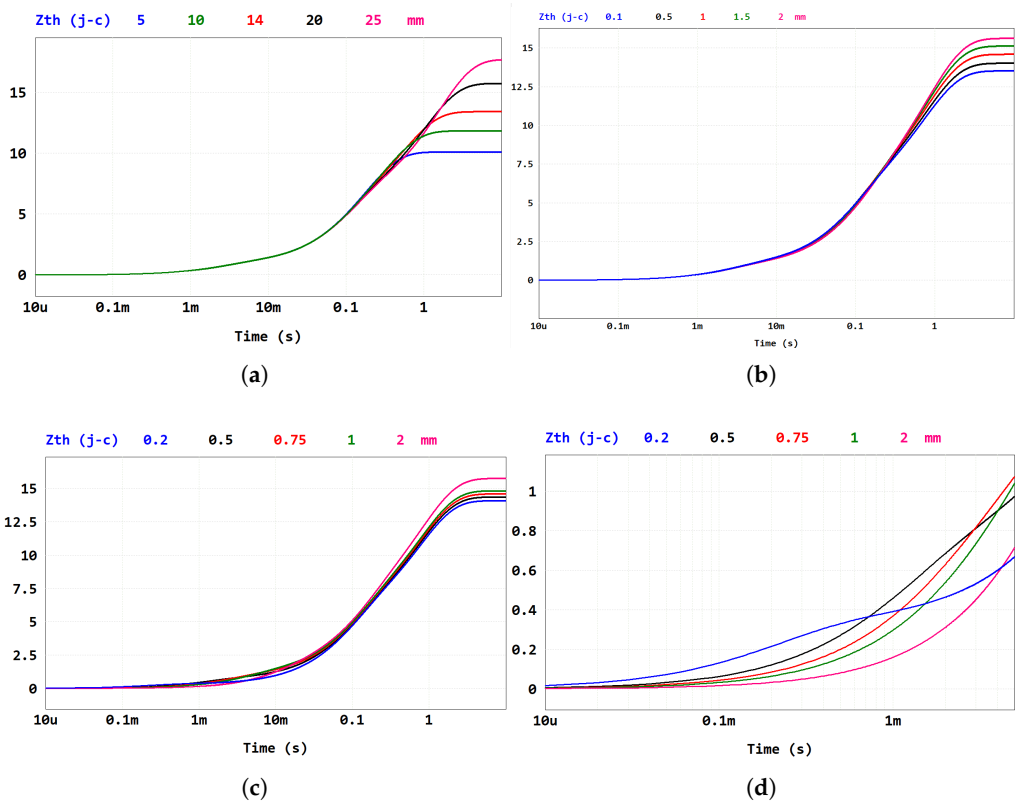



**Figure 4.** Impact of change in thickness on thermal resistance in different materials (**a**) Copper. (**b**) Molybdenum. (**c**) Silicon. (**d**) Silicon at 5 ms.

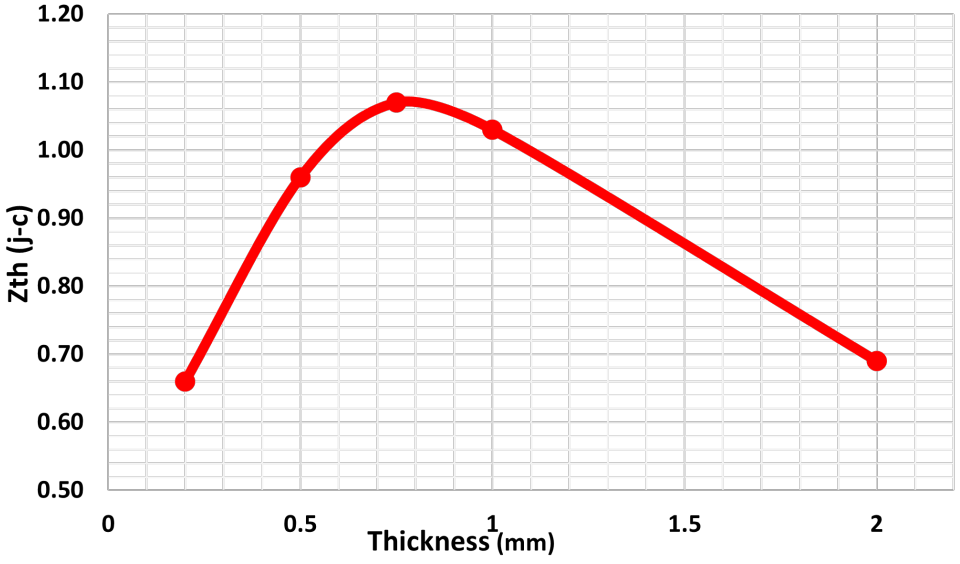



**Figure 5.** Behavior of device performance due to change in thickness at 5 ms.

### *3.2. Impact of Change in Material Diameter on Device Performance*

The device is simulated to investigate the effect of diameter variation on performance, while keeping the thickness constant. The results indicate that increasing the diameter reduces thermal resistance as shown in Figure 6, as the heat is distributed over a larger surface area. Consequently, devices with larger diameters can operate at higher power ratings; however, larger diameters also result in higher capacitance, which may reduce switching speed. Furthermore, uneven current distribution at the edges of larger chips can lead to localized hotspots and larger temperature gradients across the die during high-power operation, potentially causing performance mismatches within the same device.

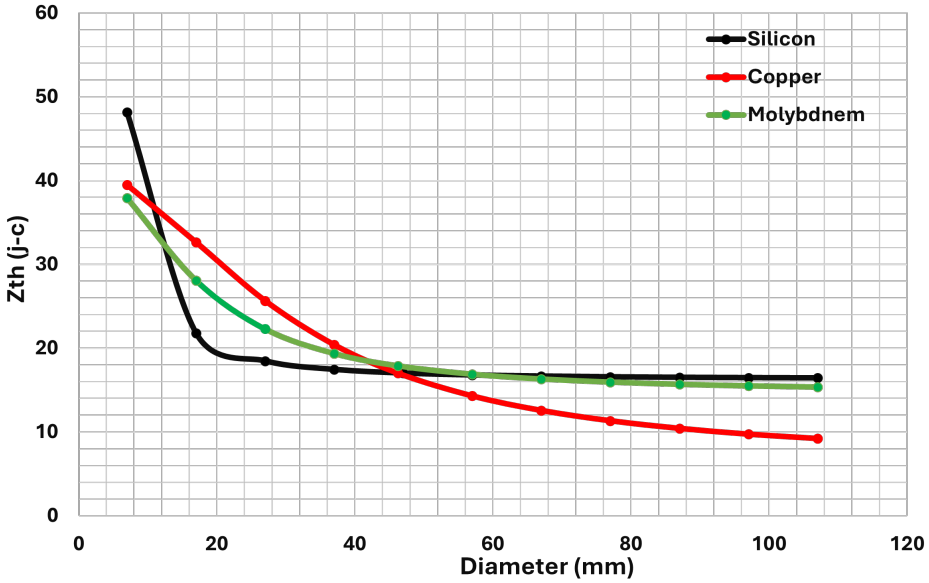



**Figure 6.** Impact of change in diameter on device thermal impedance.

### *3.3. Performance Evaluation of Geometric Variations*

As discussed in the previous sections, it is evident from simulation results that variations in the diameter and thickness of the material significantly influence the device characteristics. In particular, changes in these parameters affect the thermal resistance. The performance of the device is evaluated in terms of current capability as shown in Figure 7. The results clearly demonstrate the influence of geometric parameters on device performance. A reduction in thickness decreases the thermal resistance, thereby enabling higher current capability, and vice versa, as summarized in Tables 2 and 3 for silicon and molybdenum, respectively. Conversely, an increase in diameter enhances the current capability. While most parameters affecting device characteristics are reported in manufacturer datasheet, a comprehensive analysis of these factors is essential for optimized device design. Other parameters, such as device diameter, $V_{th}$ and rT, etc., are kept constant; however, all parameters are interdependent. These findings emphasize the critical role of material geometry in governing overall device behavior.

**Table 2.** Change in silicon current.

| Thickness (mm) | $R_{th}$ | Current |
|---|---|---|
| 0.2 | 14 | 4856 |
| 0.5 | 14.3 | 4793 |
| 0.75 | 14.5 | 4752 |
| 1 | 14.8 | 4692 |
| 2 | 15.7 | 4523 |

**Table 3.** Molybdenum current variation.

| Thickness (mm) | $R_{th}$ | Current |
|---|---|---|
| 0.1 | 13.5 | 4966 |
| 0.5 | 14 | 4856 |
| 1 | 14.5 | 4752 |
| 1.5 | 15 | 4653 |
| 2 | 15.6 | 4541 |

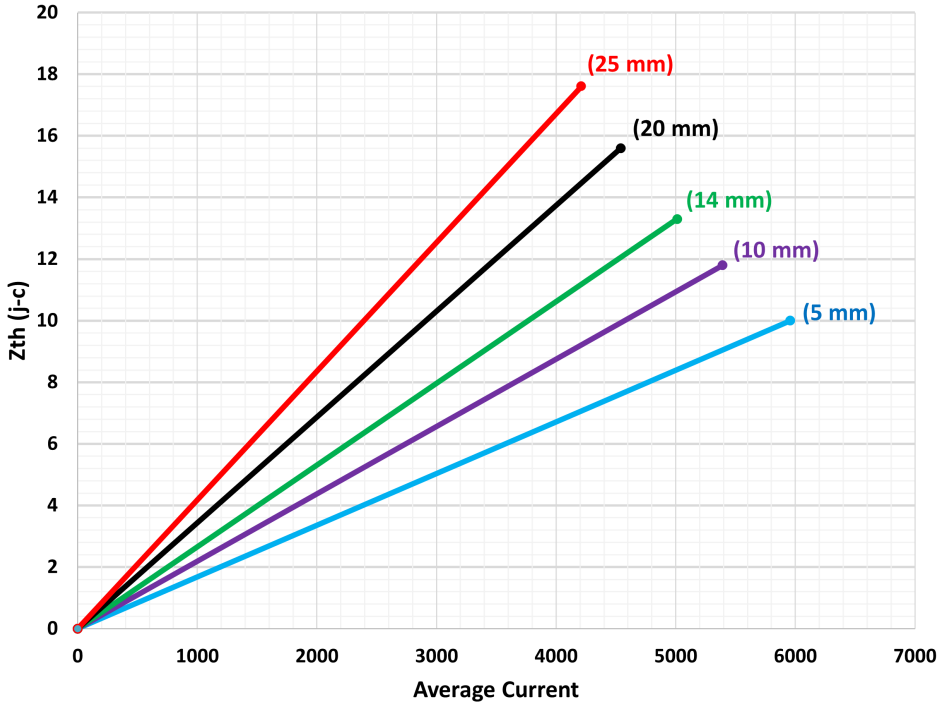



**Figure 7.** Impact of change in material geometry on device average current in copper.

## 4. Methodology

### *4.1. Data Acquisition*

The dataset used in this study is derived from the technical datasheets of leading thyristor manufacturers, namely ABB, Infineon, Dynex, and Posieco Power Electronics. Six key electrical parameters such as peak reverse repetitive voltage ($V_{RRM}$), average current ($I_{AV}$), threshold voltage ($V_{TH}$), thermal resistance ($R_{th}$), slope resistance (rT), and device diameter were systematically extracted and standardized to ensure cross-manufacturer comparability. The dataset comprises 250 thyristor devices, with diameters ranging from 19 mm to 115 mm, $V_{RRM}$ ratings from 800 V to 8500 V, and average current ratings up to 15,000 A. For subsequent analysis, diameters were categorized into defined classes as shown in Table 4, enabling a structured representation of characteristics.

**Table 4.** Diameter classes and corresponding included diameters.

| Diameter Class | Included Diameters |
|---|---|
| 25 | 19, 25, 31 |
| 38.5 | 35, 38, 41 |
| 47 | 44, 47, 51 |
| 63 | 57, 63 |
| 76 | 73, 78, 80 |
| 100 | 100, 101 |
| 112 | 110, 115 |

This detailed dataset provides a wide range of industrial thyristors with various electrical characteristics, thereby providing a robust foundation for machine learning models aimed at performance prediction, classification, and design optimization. Table 5 shows the six features of the dataset which have been considered for the proposed model to classify the diameter of the chip. All features have numerical values assigned. All these features impact the device performance and help in determining the size of device using the ML system.

**Table 5.** Sample of dataset features used in the proposed model.

| $I_{AV}$ | $R_{th}$ | $V_{RRM}$ | $V_{th}$ | $rT$ | Diameter (mm) |
|---|---|---|---|---|---|
| 1609.56 | 56.59 | 811.23 | 0.7219 | 0.2523 | 25 |
| 2143.54 | 38.68 | 995.51 | 0.7129 | 0.1259 | 38.5 |
| 3370.69 | 18.11 | 2513.90 | 0.7786 | 0.1607 | 48 |
| 4718.06 | 9.16 | 1713.85 | 0.9849 | 0.0692 | 63 |
| 5458.89 | 9.16 | 1852.87 | 0.7451 | 0.1033 | 76 |
| 7157.89 | 6.98 | 1826.56 | 0.7812 | 0.0630 | 100 |
| 9621.60 | 4.41 | 4799.44 | 0.7994 | 0.0711 | 112 |

### *4.2. Data Augmentation*

Although a variety of thyristor devices are available in the market with different specifications, still the availability of datasheets of well known manufacturers are limited, especially for each diameter which reduces the size of the dataset. In order to balance the classes, data augmentation was performed in Matlab to increase the samples of classes in the dataset [18–20].

Augmentation was performed by generating synthetic samples through controlled random perturbations applied within manufacturer-specified tolerance ranges. All parameters were varied by injecting Gaussian noise at a fractional level of ±5% relative to their nominal values, simulating measurement uncertainties and natural parameter variations. This noise-based approach preserves the underlying statistical distribution and interparameter relationships inherent to each device class. Mathematically, for each original feature vector $x$, multiple augmented samples $x'$ were created as follows:

$$\mathbf{x}' = \mathbf{x} + \boldsymbol{\epsilon} \odot \mathbf{x}$$

where $\boldsymbol{\epsilon}$ is a Gaussian noise vector with zero mean and a standard deviation corresponding to the specified noise level, and $\odot$ denotes element-wise multiplication. To ensure physical plausibility, augmented parameter values were constrained within manufacturer-specified limits; for instance, $V_{RRM}$ was bounded between 800 V and 8500 V, and thermal resistance $R_{th}$ was restricted to positive, realistic values. Augmentation was applied separately for each device class to maintain class-specific characteristics, generating five synthetic samples per original instance. This process increased the dataset size significantly and improved class balance, particularly for underrepresented diameter categories. A controlled augmentation of 1:5 was applied to maintain balance in all data classes throughout the dataset by normalizing each class relative to its original sample size, ensuring that the effective representation after augmentation remained proportionally consistent across all classes. To verify augmentation quality, statistical comparisons demonstrated substantial overlap between the distributions of original and augmented data. Furthermore, principal component analysis (PCA) revealed that synthetic samples projected onto the first principal component (PC1) lie entirely within the range of the original data distribution, as illustrated in Figure 8. Moreover, the PCA scores showed variation in the original and augmented datasets were very similar, with variance decreasing by about 2.6% on PC1 and 2.5% on PC2. This highlights that the augmentation results in higher sample density while keeping the original data distribution the same. Subse-

quently, these results also confirm that the augmentation process did not introduce unrealistic parameter combinations or distort the original data manifold.

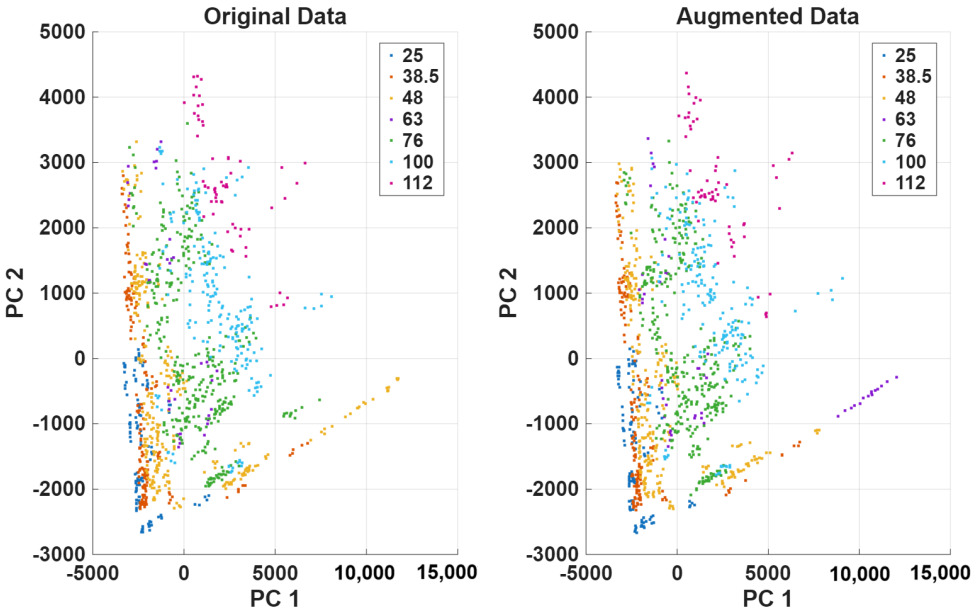



**Figure 8.** Validation of original data and augmented data.

### *4.3. Feature Optimization*

To identify the most relevant features contributing to the classification and prediction performance of the machine learning models, the Minimum Redundancy Maximum Relevance (MRMR) algorithm was employed [21–24]. MRMR selects features that maximize relevance with the target variable while minimizing redundancy among features, thus ensuring an optimal subset that balances information content and independence. Applying MRMR to the dataset revealed that parameters such as $R_{th}$, $I_{AV}$, and $V_{RRM}$ contribute most significantly to accuracy as shown in Figure 9, consistent with their critical roles in the performance of thyristor. Other parameters showed lower importance scores, but were retained to preserve structural information. Based on these results, models trained with the top-ranked features demonstrated improved efficiency and comparable or better accuracy, validating the effectiveness of the feature optimization process in refining the dataset for machine learning applications.

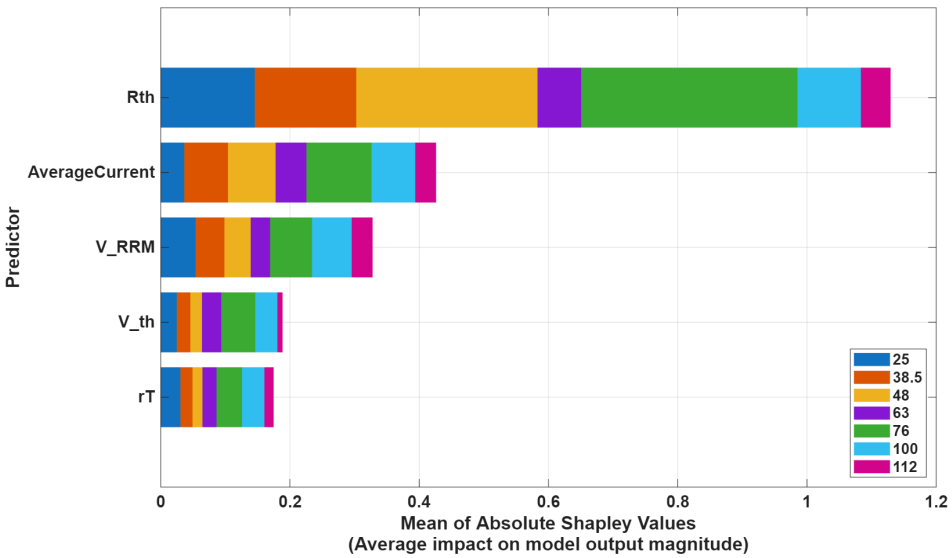



**Figure 9.** Contribution of feature on output for diameter classification.

### *4.4. Implementation of Machine Learning Models*

To ensure robust classification and prediction of chip diameter, three machine learning models were implemented as shown in Figure 10 and compared including Support Vector Machine, Artificial Neural Network, and Ensemble Learning. These algorithms were selected due to their complementary strengths in handling complex, nonlinear relationships, diverse dataset distributions, and high-dimensional feature spaces.The architecture of implemented models is listed in Table 6. All models were trained and evaluated using identical dataset partitions, with 90% of the data allocated for training and 10% for testing. Performance was measured using accuracy, F1-score, precision, and recall, averaged over 7-fold cross-validation. Hyperparameters were tuned via Bayesian optimization for each model to ensure optimal performance.

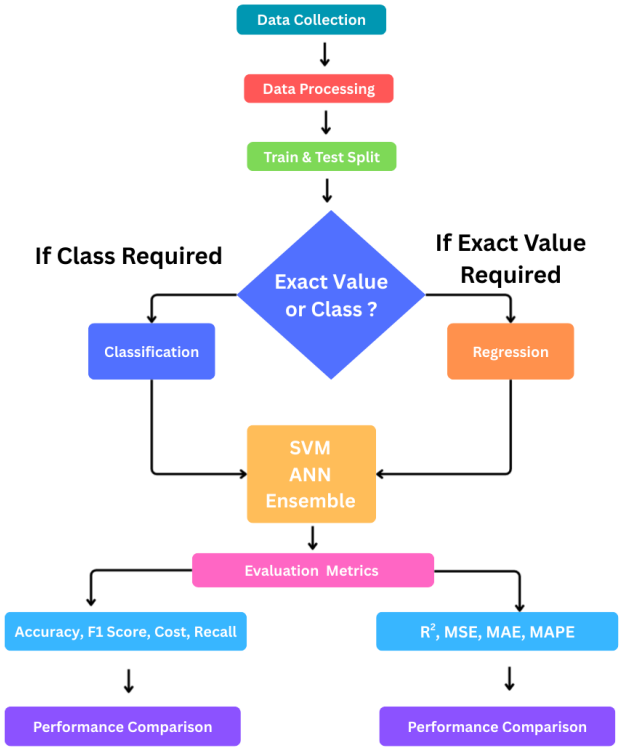



**Figure 10.** Methodology for diameter classification and prediction.

**Table 6.** Architecture of machine learning models.

| Model | Architecture |
|---|---|
| ANN | Hidden layers: 1<br>Number of neurons in each layer: 263<br>Activation function: Tanh<br>Optimizer: Bayesian optimization<br>Learning rate: $9.2024 \times 10^{-09}$ |
| Ensemble | Method: Bagging<br>Splits: 1023<br>Learners: 13 (base learner: decision tree) |
| SVM | Kernel: Cubic<br>Regularization parameter: $C = 1$<br>Kernel scale: 25.5 |

#### 4.4.1. Support Vector Machine

Support Vector Machine is a widely used machine learning algorithm, particularly effective for both classification and regression tasks across diverse application domains. Its strength lies in the ability to model both linear and nonlinear relationships within data by constructing an optimal decision boundary, defined by maximizing the margin between the hyperplane and the closest data points, known as support vectors. These support vectors are critical in determining the classification boundary, as they alone influence the separation of data into distinct categories. In contrast, SVM regression constructs a predictive function by fitting the model within a specified tolerance margin known as epsilon, ensuring that deviations smaller than this margin do not contribute to the prediction error. During the implementation, the classification model employs a cubic kernel function with a box constraint of 25.5 and a one-vs-one multiclass coding scheme, while the regression model uses a quadratic kernel with a box constraint of 3.7, an epsilon of 2.09, and hyperparameters optimized via Bayesian optimization; the different kernels were chosen to provide flexible nonlinear separation for classification and smooth function fitting for regression. The algorithm automatically determined the optimal kernel value during training through iterative optimization within the range of 0.001–1000.

#### 4.4.2. Artificial Neural Network

The ANN model was implemented in MATLAB® 2025a classification learner framework using a feedforward architecture with single hidden layer layer. Generally, ANN models with multiple hidden layers can be mathematically expressed as follows:

$$y = f\left(\sum_{i=1}^{n} w_i x_i + b\right),$$

where $x_i$ are the inputs, $w_i$ the corresponding weights indicating the importance of each feature, $b$ the bias term, adjusting the output independently of the input values and $f(\cdot)$ the activation function. the values of $w_i$ and $b$ are adjusted to reduce errors between its predictions and actual data. In this study, Bayesian optimization was employed to tune the hyperparameter and determine efficient ANN architectures. For classification, the optimal configuration consists of a single hidden layer with 263 neurons, utilizing a tanh activation function and trained for 1000 iterations to achieve stable convergence. For regression, the optimized architecture comprises two hidden layers with 296 and 3 neurons, respectively, employing a ReLU activation function for effective nonlinear approximation. These architectures were selected to balance model complexity with generalization capability, ensuring reliable performance across both tasks.

#### 4.4.3. Ensemble Learning

Ensemble methods are machine learning techniques that combine multiple learners trained on the same dataset to improve prediction accuracy and robustness. Two commonly used strategies are bagging and boosting. In bagging, multiple models are trained on random subsets of the dataset, and predictions are aggregated by averaging (for regression) or voting (for classification). Boosting, in contrast, trains a sequence of weak learners iteratively, increasing the weights of misclassified samples to reduce errors in subsequent rounds, with the final prediction obtained by combining weighted outputs. In this study, bagging, a decision tree ensemble classifier was employed for classification, where the maximum number of splits is 1023, and the ensemble consisted of 13 learners. For regression, a bagging ensemble was used, with decision trees having a minimum leaf size of 2 and considering 4 variables at each split with 13 number of learners; predictions were averaged across all trees to reduce variance. Hyperparameters for both classification

and regression ensembles were optimized using Bayesian optimization to achieve robust architecture and improved generalization performance. Bagging was employed for both classification and prediction to improve the stability and accuracy of the model. Both techniques are fundamental to ensemble learning, aiming to enhance the overall predictive performance and robustness of machine learning algorithms [25].

## 5. Results and Discussion

### *5.1. Comparison of Classification Accuracies*

This study determined that machine learning algorithims can be used to accurately classify the diameter of high-power thyristors. Both ANN and SVM classifiers achieved an accuracy of 95%, while Ensemble resulted in an accuracy of 96.7%. The performance of all algorithms are satisfactory in this study and achieved an accuracy of over 95% as listed in Table 7. Among the evaluated models, the ANN achieves the best overall performance, delivering 95.0% testing accuracy along with stable F1-score and recall, i.e., 96.1% across validation, demonstrating robustness against overfitting. In comparison, the SVM model attains a similar accuracy of 95% but shows lower F1 score and recall of 96% and 96.1%, respectively, while the Ensemble reaches the highest accuracy of 96.7% yet suffers from degraded recall 94.8% and significantly higher computational cost. Furthermore, ANN exhibits the lowest validation cost of 42, compared to 43 for SVM and 56 for Ensemble, reflecting greater efficiency and reduced sensitivity to feature tuning and class-boundary effects. While the Ensemble model offers competitive accuracy, its higher class-specific error rates and lack of end-to-end feature learning result in reduced generalization. The ANN's architecture effectively captures nonlinear feature interactions, enabling reliable deployment in high-stakes, real-time applications, whereas the SVM's superior accuracy is offset by its computational overhead as shown in Table 8. Overall, the ANN achieves the best trade-off between accuracy, efficiency, and robustness, establishing state-of-the-art performance.

**Table 7.** Model performance on validation and test datasets.

| Parameter | Artificial Neural Network | Ensemble | Support Vector Machine |
|---|---|---|---|
| **F1 score** | | | |
| Test | 95.1 | 96.6 | 95.1 |
| Validation | 96.1 | 94.8 | 96.0 |
| **Recall score** | | | |
| Test | 95.0 | 96.7 | 95.0 |
| Validation | 96.1 | 94.8 | 96.1 |
| **Accuracy** | | | |
| Test | 95.0 | 96.7 | 95.0 |
| Validation | 96.1 | 94.9 | 96.1 |
| **Precision** | | | |
| Test | 96.4 | 96.8 | 95.6 |
| Validation | 96.1 | 94.9 | 96.0 |

**Table 8.** Computational Performance Metrics for Classification and Regression-based Prediction.

| Model | Classification | | | Regression-Based Prediction | | |
|---|---|---|---|---|---|---|
| | Training Time (s) | Model Size (kB) | Prediction Speed (obs/s) | Training Time (s) | Model Size (kB) | Prediction Speed (obs/s) |
| ANN | 700.12 | 29 | 23,000 | 220.4 | 23 | 100,000 |
| Ensemble | 95.12 | 313 | 12,000 | 38.2 | 60 | 90,000 |
| SVM | 893.66 | 28 | 6000 | 107.1 | 37 | 130,000 |

5.1.1. Confusion Matrix of All Features

Confusion matrix analysis shows that most models achieve class-wise accuracies above 90% except few classes such as 100, 63, and 38.5. For class 38.5, both ANN and SVM reach 88.9% accuracy as shown in Figures 11 and 12, with misclassifications distributed in class 25.

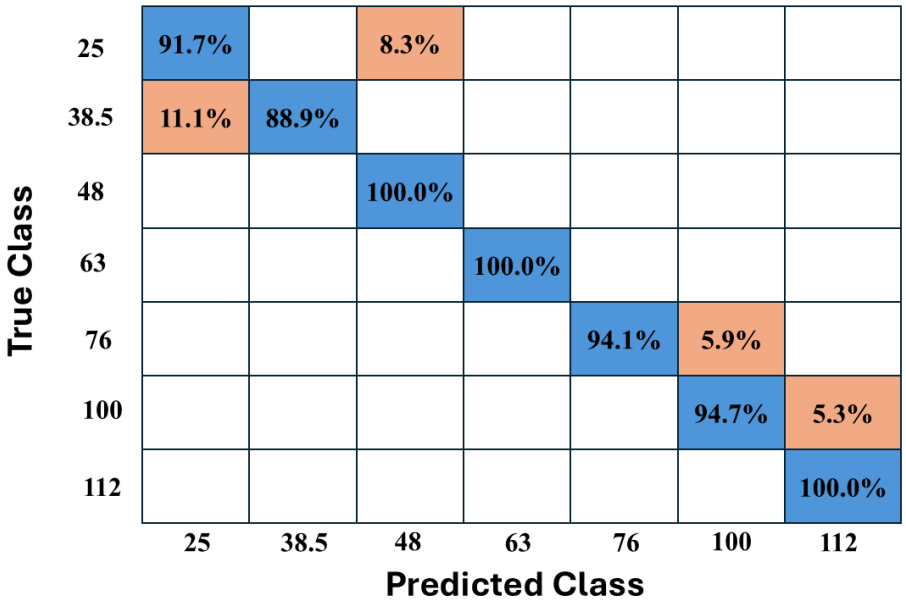



**Figure 11.** Confusion matrix of test accuracy for classification of diameter by ANN.

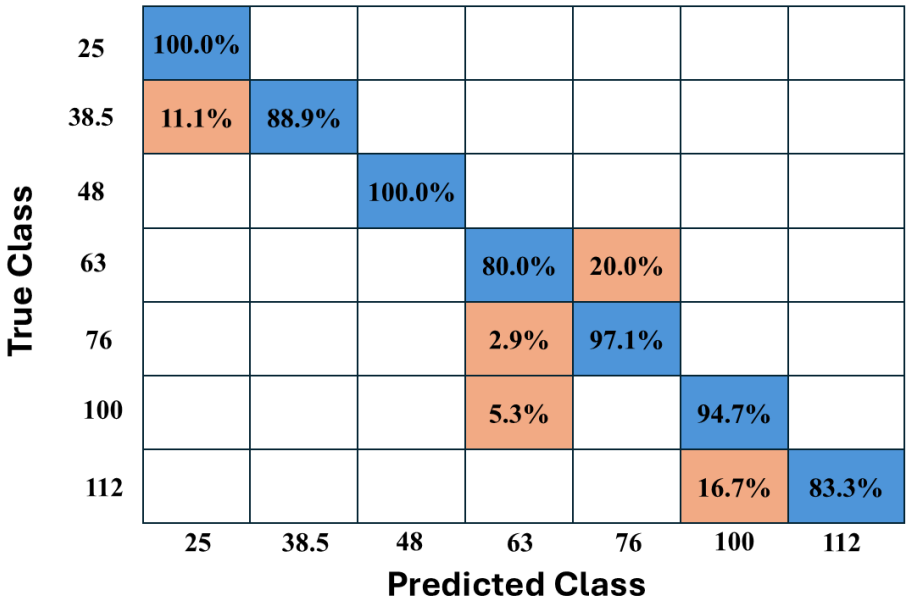



**Figure 12.** Confusion matrix of test accuracy for the diameter classification for SVM.

For class 63, both the Ensemble Learner and SVM achieved an accuracy of 80%, with the Ensemble model misclassifying samples as class 100 as shown in Figure 13, while the SVM misclassified them as class 76. In contrast, the ANN exhibited superior performance, attaining 100% accuracy for class 63 and misclassifying only 5.3% of samples in class 100, compared to 10.5% by the Ensemble model, reflecting its stronger ability to define clear decision boundaries and avoid local minima. Additionally, the SVM accuracy also decreases for class 112 to 83.3%, misclassifying some samples as class 100. This reduction is likely due to the SVM's reliance on maximizing class separation, which is

challenging for small datasets, making it more prone to overfitting. In contrast, ANNs and Ensemble methods are more robust on limited data, as they can generalize through multi-layered architectures or by aggregating multiple diverse models to reduce individual errors. Overall, apart from these exceptions, all models maintain high accuracies above 95%, confirming their robustness for diameter classification.

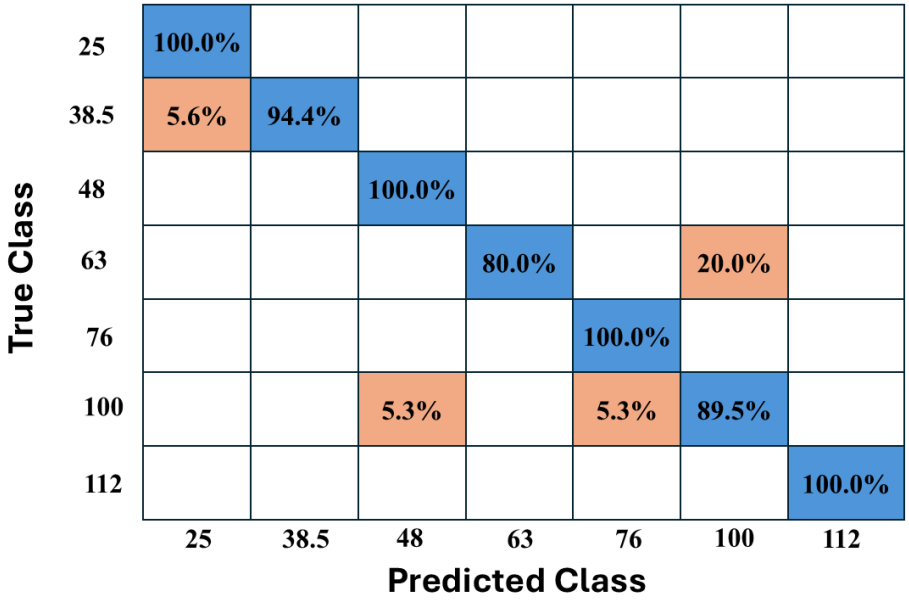



**Figure 13.** Confusion matrix of test accuracy for ensemble.

5.1.2. Confusion Matrix of Optimized Features

The models were evaluated using the optimized feature parameters described in Section 4.3. Results confirm that these parameters enable accurate diameter classification with overall performance approaching 90% accuracy. Specifically, the ANN and SVM models achieved 94.2% accuracy, while the Ensemble reached 90.9% as depicted in Table 9.

**Table 9.** Model performance on validation and test datasets with optimized parameters.

| Parameter | Artificial Neural Network | Ensemble | Support Vector Machine |
|---|---|---|---|
| **F1 Score** | | | |
| Test | 94.3 | 91.1 | 94.3 |
| Validation | 94.9 | 94.2 | 94.7 |
| **Recall Score** | | | |
| Test | 94.2 | 90.9 | 94.2 |
| Validation | 94.9 | 94.2 | 94.7 |
| **Accuracy** | | | |
| Test | 94.2 | 90.9 | 94.2 |
| Validation | 94.9 | 94.2 | 94.7 |
| **Precision** | | | |
| Test | 94.8 | 91.4 | 94.7 |
| Validation | 94.9 | 94.3 | 94.8 |

Although the SVM achieves comparable accuracy to the ANN, the ANN attains this performance at a substantially lower computational cost. Specifically, the ANN demonstrates superior efficiency with a prediction speed of approximately 48,000 observations per second (obs/s), compared to 18,000 obs/s for the SVM, establishing it as the most effective and computationally efficient model among those evaluated. Nonetheless, the full

feature models deliver a higher accuracy of 95%; the models achieve approximately 94.2% accuracy with the optimized feature parameters. This reduction in accuracy is not very high and indicates that the optimized features pose sufficient classification capability, while significantly reducing complexity, making it acceptable. The confusion matrix of ANN, as shown in Figure 14 reveal that the performance of optimized features are reduced for class 63, 76, and 112 while improved for class 100. Similarly for the ensemble, class 76 and 112 are misclassified as illustrated in Figure 15 and improved for class 38.5. Moreover, the confusion matrix of the SVM, shown in Figure 16, indicates improvement in classification of class 38.5, 48, and 76, while the other classes indicate a close match with the full feature classification.

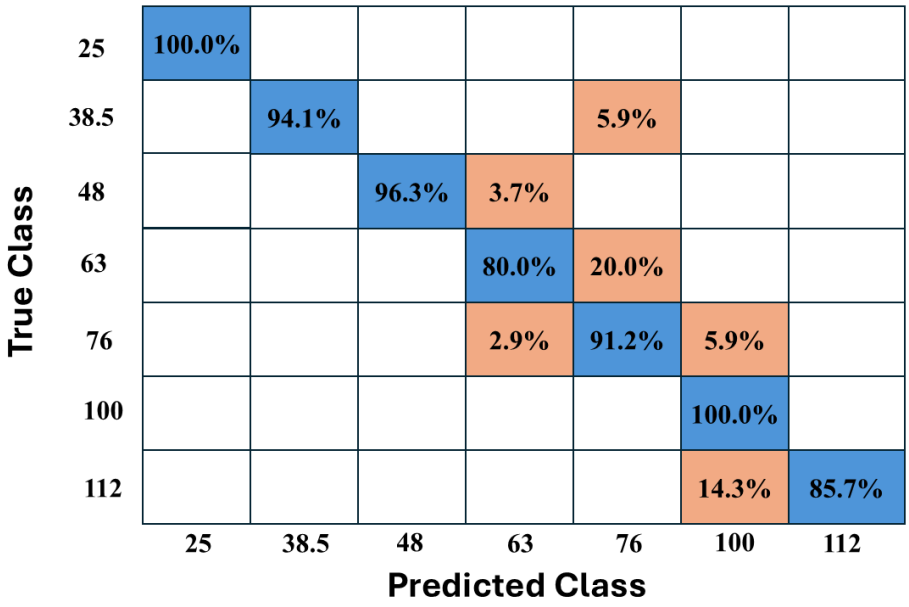



**Figure 14.** Confusion matrix of test accuracy for the diameter classification with feature optimized ANN.

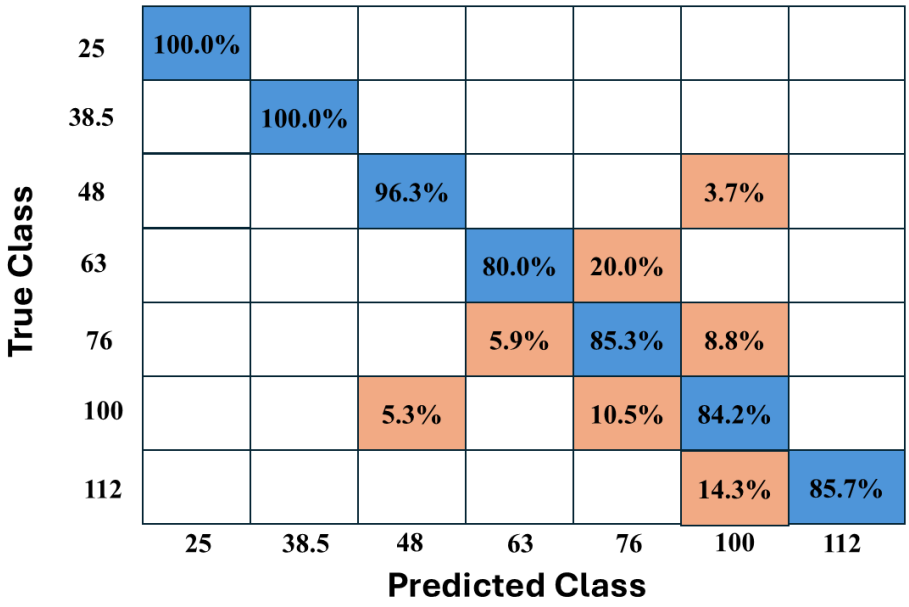



**Figure 15.** Confusion matrix of test accuracy for the diameter classification with optimized features ensemble.

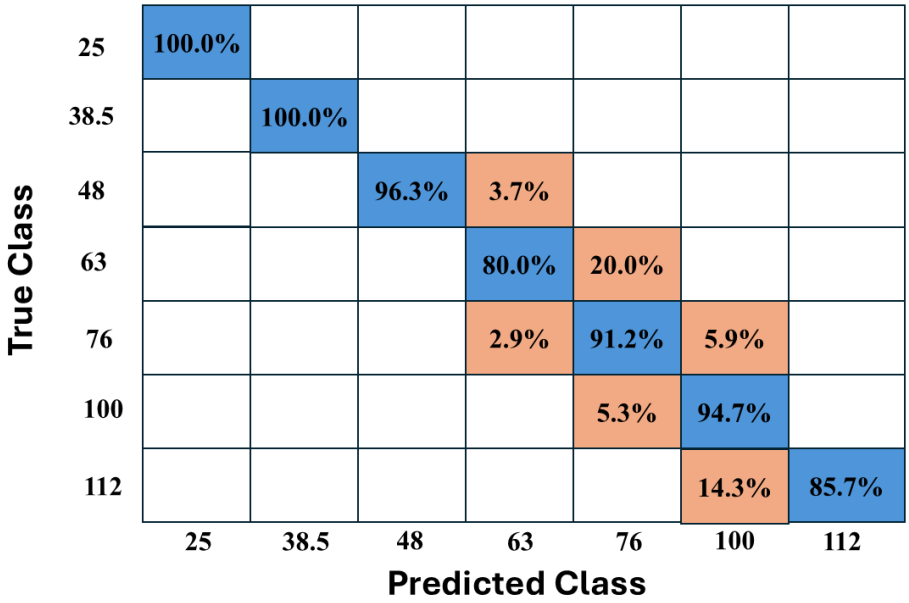



**Figure 16.** Confusion matrix of test accuracy for the diameter classification with optimized features SVM.

### *5.2. Prediction*

Regression analysis is a fundamental approach for predicting continuous process variables, enabling precise forecasting of critical parameters. In this study, the predictive performance of the three optimized ML models, ANN, Ensemble, and SVM, were evaluated for diameter prediction. Table 10 summarizes the performance metrics, highlighting that the Ensemble model achieved the highest accuracy with a minimum prediction error of 0.03. SVM yielded comparable accuracy but with a higher average error of 0.07 (2.3× that of Ensemble). In contrast, ANN demonstrated the worst performance, with $R^2$ = 0.76 and an average error of 0.24 (8× higher than Ensemble). The ensemble and SVM models achieved $R^2$ values of 0.97 and 0.93, respectively, which are considered acceptable for applications in the manufacturing industry. The deviations between actual and predicted diameters for Ensemble, SVM, and ANN are illustrated in Figure 17, respectively. The ANN model exhibits the largest discrepancies, in contrast to the closer alignment achieved by Ensemble and SVM. This reduced performance is attributed to the limited dataset size, as ANN typically requires larger datasets to effectively learn complex patterns, resulting here in overfitting and degraded generalization. These findings establish the Ensemble model as the most efficient and reliable predictor as detailed in Table 8, with predicted diameters closely matching the actual values. This capability enables manufacturers to develop devices precisely tailored to the requirements of their specific applications.

**Table 10.** Comparative performance of optimized regression models for average current forecasting.

| Model | $R^2$ | MSE | MAE | MAPE (%) |
|---|---|---|---|---|
| Optimized Ensemble | 0.97 | 20.4 | 1.7 | 3.4 |
| Support Vector Machine | 0.93 | 47.7 | 4.8 | 8.8 |
| Artificial Neural Network | 0.76 | 161.09 | 9.78 | 18.1 |

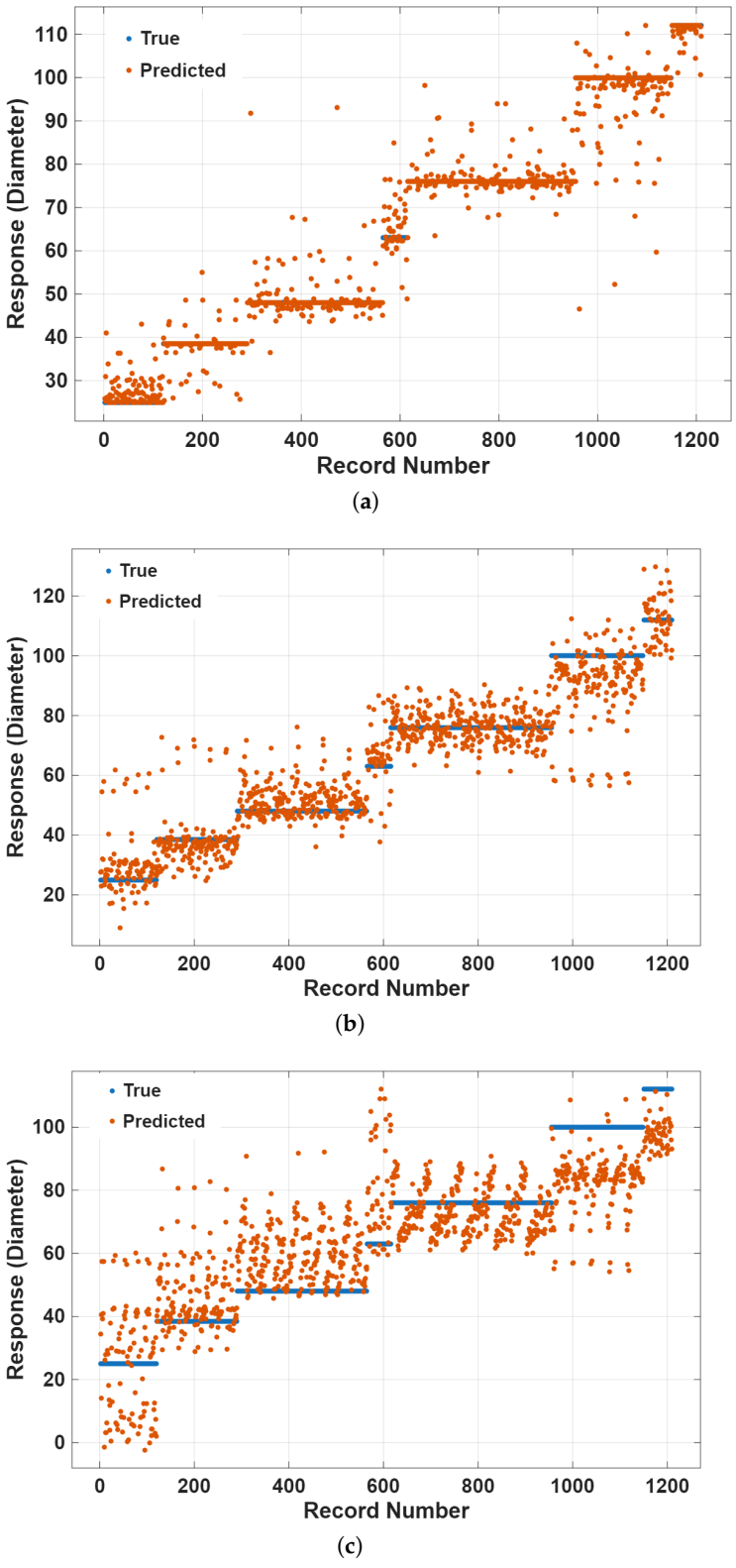



**Figure 17.** Results of accuracy of predicted diameter (**a**) Ensemble model. (**b**) SVM model. (**c**) ANN model.

## 6. Conclusions and Future Work

In this work, a comprehensive market survey was conducted to explore high-current thyristors, and a dataset was generated from leading manufacturers through the extraction of various electrical parameters. To increase the dataset size, data augmentation techniques were applied, and the device diameters were categorized into seven classes. Classification

and prediction of diameter based on these parameters were performed using three machine learning algorithms: ANN, Ensemble, and SVM. Although the SVM achieved slightly higher accuracy, its computational cost was significantly greater, making the ANN the most efficient and reliable choice for diameter classification. However, in the regression-based prediction tasks, both the ensemble and SVM models demonstrated superior predictive performance compared to the ANN, indicating their enhanced capability in capturing the underlying nonlinear relationships within the dataset. Future work will focus on extending this methodology to predict and classify current ratings, as well as implementing other advanced AI models and algorithms to further enhance performance and generalization.

**Author Contributions:** Conceptualization: L.V., M.M. and F.A.; Methodology: F.A.; Modeling: F.A. and A.A.N.; Software: F.A.; Formal Analysis: F.P.; Writing—Original Draft Preparation: F.A.; Writing—Review and Editing: L.V. and M.M.; Supervision: L.V. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was supported by Programma Regionale Fondo Sociale Europeo+ 2021–2027, Priorità 2—Istruzione e Formazione—ESO 4.6 (OS-f), Operazione 'Scienze e tecnologie per l'ingegneria elettrica e i sistemi complessi per la mobilità', carried out in collaboration with Poseico Power Electronics S.p.A., Genoa, Italy. The funding does not cover the cost of this publication.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Data are available and may be provided by the authors upon request.

**Conflicts of Interest:** Author Federico Portesine was employed by the company Poseico Power Electronics S.p.A. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## References

1. Carbone, L.; Marchesoni, M.; Passalacqua, M.; Vaccaro, L.; Formentini, A. H2-LMI-Based High Performance Control for Matrix Converter. *IEEE Trans. Ind. Electron.* **2024**, *71*, 10029–10038. [CrossRef]
2. Benevieri, A.; Marchesoni, M.; Passalacqua, M.; Pozzobon, P.; Vaccaro, L. Synchronous DTC for torque sub-harmonic reduction in low switching frequency induction motor drives. *Control Eng. Pract.* **2025**, *154*, 106133. [CrossRef]
3. Wang, G.; Wen, H.; Liu, W.; Li, F. Physical Structure, Characteristics, and Applications of Monolithic Bidirectional Switches: A Comprehensive Review. *IEEE Trans. Power Electron.* **2025**, *40*, 9187–9199. [CrossRef]
4. Chen, Y.L.; Sacchi, S.; Dey, B.; Blanco, V.; Halder, S.; Leray, P.; Gendt, S.D. Exploring Machine Learning for Semiconductor Process Optimization: A Systematic Review. *IEEE Trans. Artif. Intell.* **2024**, *5*, 5969–5989. [CrossRef]
5. Ghoshhajra, R.; Biswas, K.; Sarkar, A. A Review on Machine Learning Approaches for Predicting the Effect of Device Parameters on Performance of Nanoscale MOSFETs. In Proceedings of the 2021 Devices for Integrated Circuit (DevIC), Kalyani, India, 19–20 May 2021; pp. 489–493. [CrossRef]
6. Wang, Z.; Li, L.; Leon, R.C.C.; Yang, J.; Shi, J.; van der Laan, T.; Usman, M. Improving Semiconductor Device Modeling for Electronic Design Automation by Machine Learning Techniques. *IEEE Trans. Electron Devices* **2024**, *71*, 263–271. [CrossRef]
7. Zhao, S.; An, T.; Qin, F. Research on Machine Learning-Assisted Numerical Simulation Methods for IGBT Module Bonded Interface Fatigue Damage. *IEEE J. Emerg. Sel. Top. Power Electron.* **2025**, *13*, 3259–3272. [CrossRef]
8. Thekemuriyil, T.; Rohner, J.D.; Minamisawa, R.A. Machine learning-based prediction of on-state voltage for real-time health monitoring of IGBT. *Power Electron. Devices Compon.* **2023**, *6*, 100049. [CrossRef]
9. Yang, X.; Zhang, Y.; Wu, X.; Liu, G. Failure Mode Classification of IGBT Modules Under Power Cycling Tests Based on Data-Driven Machine Learning Framework. *IEEE Trans. Power Electron.* **2023**, *38*, 16130–16141. [CrossRef]
10. Yao, Q.; Chen, J.; Dai, Y.; Yao, J.; Zhang, J.; Zhang, M.; Liu, J.; Tang, W.; Zhang, J.; Zhang, B.; et al. A Novel Prediction Technology of Output Characteristics for IGBT Based on Compact Model and Artificial Neural Networks. *IEEE Trans. Electron Devices* **2023**, *70*, 4885–4891. [CrossRef]
11. Liu, L.; Zhang, G.; Zhao, C.; Lin, C.; Liu, J.; Zhang, Y. Study on Fault Diagnosis of Power Electronic Devices in Power Conversion System Based on Machine Learning. In *Proceedings of the 1st Electrical Artificial Intelligence Conference, Volume 2: EAIC 2024, 6–8 December, Nanjing, China*; Qu, R., Song, Z., Ding, Z., Mu, G., Xiong, R., Han, L., Eds.; Springer: Singapore, 2025; pp. 52–61.

12. Market.us. Semiconductor Materials Market Report. Available online: https://market.us/report/semiconductor-materials-market/ (accessed on 5 August 2025).
13. Gujar, S.; Ugale, S. Semiconductor Foundry Market Size, Share, Forecast 2025–2034. February 2025. Available online: https://www.gminsights.com/industry-analysis/semiconductor-foundry-market (accessed on 15 August 2025).
14. SNS Insider pvt Ltd. *Power Semiconductor Market Size to Hit USD 75.13 Billion by 2032*; GlobeNewswire: Austin, TX, USA, 4 February 2025.
15. PW Consulting Information & Electronics Research Center. Power Semiconductor Chips Market. Website, 2025. "Power Semiconductor Chips Market" Report. Available online: https://pmarketresearch.com/it/power-semiconductor-chips-market/ (accessed on 7 August 2025).
16. Mordor Intelligence. Power Semiconductor Market—Share, Companies & Industry Size. 2025. Available online: https://www.mordorintelligence.com/industry-reports/semiconductor-industry-landscape (accessed on 5 August 2025).
17. Ahmad, F.; Portesine, F.; Ashero, M.; Marchesoni, M.; Vaccaro, L. Transient Thermal Impedance Characterization of New High Power Press Pack Diodes. In Proceedings of the 2024 International Symposium on Power Electronics, Electrical Drives, Automation and Motion (SPEEDAM), Napoli, Italy, 19–21 June 2024; pp. 81–86. [CrossRef]
18. Wang, Z.; Wang, P.; Liu, K.; Wang, P.; Fu, Y.; Lu, C.T.; Aggarwal, C.C.; Pei, J.; Zhou, Y. A Comprehensive Survey on Data Augmentation. *IEEE Trans. Knowl. Data Eng.* 2025, *early access*. [CrossRef]
19. Jung, J.; Apsari, D.P.; Lee, D.C. Robust Open-Switch Fault Diagnosis of Three-Level NPC Inverters Based on Data Augmentation With White Noise Injection. *IEEE Trans. Power Electron.* **2025**, *40*, 3553–3565. [CrossRef]
20. Ge, J.; Yao, Z.; Wu, M.; Almeida, J.H.S., Jr.; Jin, Y.; Sun, D. Tackling data scarcity in machine learning-based CFRP drilling performance prediction through a Broad Learning System with Virtual Sample Generation (BLS-VSG). *Compos. Part B Eng.* **2025**, *305*, 112701. [CrossRef]
21. Kavitha, K.; Kumar, R.A.; Mol C, M. A Maximum Relevance Minimum Redundancy and Random Forest based feature selection and classification of gene expression data. In Proceedings of the 2024 5th International Conference for Emerging Technology (INCET), Belgaum, India, 24–26 May 2024; IEEE: New York, NY, USA, 2024; pp. 1–5.
22. Wang, Y.; Li, X.; Ruiz, R. Feature Selection With Maximal Relevance and Minimal Supervised Redundancy. *IEEE Trans. Cybern.* **2023**, *53*, 707–717. [CrossRef] [PubMed]
23. Bezerra, F.E.; Oliveira Neto, G.C.d.; Cervi, G.M.; Francesconi Mazetto, R.; Faria, A.M.d.; Vido, M.; Lima, G.A.; Araújo, S.A.d.; Sampaio, M.; Amorim, M. Impacts of feature selection on predicting machine failures by machine learning algorithms. *Appl. Sci.* **2024**, *14*, 3337. [CrossRef]
24. Wu, M.; Yao, Z.; Verbeke, M.; Karsmakers, P.; Gorissen, B.; Reynaerts, D. Data-driven models with physical interpretability for real-time cavity profile prediction in electrochemical machining processes. *Eng. Appl. Artif. Intell.* **2025**, *160*, 111807. [CrossRef]
25. Choi, S.; Hur, J. An ensemble learner-based bagging model using past output data for photovoltaic forecasting. *Energies* **2020**, *13*, 1438. [CrossRef]


**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.