

# ZERO CROSSING POINT DETECTION IN A DISTORTED SINUSOIDAL SIGNAL USING DECISION TREE CLASSIFIER

Venkataramana VEERAMSETTY<sup>1</sup> , Pravallika JADHAV<sup>2</sup> , Eslavath RAMESH<sup>2</sup> ,  
Srividya SRINIVASULA<sup>2</sup> , Surender Reddy SALKUTI<sup>3</sup> 

<sup>1</sup>Center for AI and Deep Learning, SR University, 506371 Warangal, India

<sup>2</sup>Department of Electrical and Electronics Engineering, School of Engineering,  
SR University, 506371 Warangal, India

<sup>3</sup>Department of Railroad and Electrical Engineering, Woosong University,  
171 Dongdaeyeon, Dong-gu, 34606 Daejeon, Republic of Korea

dr.vvr.research@gmail.com, pravallikajadhav03@gmail.com, ramesheslavath303@gmail.com,  
srividyasrinivasula17@gmail.com, surender@wsu.ac.kr

DOI: 10.15598/aeec.v20i4.4562

Article history: Received May 22, 2022; Revised Oct 26, 2022; Accepted Nov 01, 2022; Published Dec 31, 2022.  
This is an open access article under the BY-CC license.

**Abstract.** Zero-crossing point detection in a sinusoidal signal is essential in the case of various power systems and power electronics applications like power system protection and power converters controller design. In this paper, 96 data sets are created from a distorted sinusoidal signal based on MATLAB simulation. Distorted sinusoidal signals are generated in MATLAB with various noise and harmonic levels. In this paper, a decision tree classifier is used to predict the zero crossing point in a distorted signal based on input features like slope, intercept, correlation and Root Mean Square Error (RMSE). Decision tree classifier model is trained and tested in the Google Colab environment. As per simulation results, it is observed that decision tree classifier is able to predict the zero-crossing points in a distorted signal with maximum accuracy of 98.3 % for noise signals and 100 % for harmonic distorted signals.

## Keywords

*Decision tree, distorted sinusoidal signal, harmonics, noise, zero-crossing point.*

## 1. Introduction

In many electrical domains like industrial electronics, grid synchronization, power quality and power system

protection etc., accurate Zero-Crossing Point (ZCP) detection is critical. To provide a proper firing angle to trigger switching devices in power converters, a ZCP detection is required as the firing angle is measured as a delay angle from ZCP. Similarly, the circuit breaker at ZCP should be open upon the occurrence of a fault for fast arc extinction. Practical line voltages are distorted in general, and they usually include a lot of harmonics and noise, which can cause synchronisation issues. ZCP detection is an easy task in case of a pure sinusoidal signal, it can be identified using simple comparator circuits. However, the ZCP detection in a distorted sinusoidal signal using comparator circuits is not appropriate as it contains multiple false ZCPs. So there is a need to build an accurate mechanism to detect ZCPs in a distorted sinusoidal signal.

Many researchers are working on ZCP detection problem and had provided various solutions. A Deep Neural Network (DNN) model is developed in [1] to predict the ZCPs in the distorted signal. Distorted signal simulated in MATLAB with noise levels 10 % to 50 %, and with Total Harmonic Distortion (THD) levels 10 % to 50 %. Data samples are extracted from these signals with a window size of 15. A phase-delay free method is proposed in [2] to detect the ZCPs of back electromotive force in spindle motors. The impact of asymmetric machine parameters and resistance tolerance of back emf measurement circuit on ZCP detection based sensor-less control of high speed brush less DC motor is studied in [3]. In this study,

the authors did not analyze the impact of asymmetric mutual inductance on ZCP detection.

A digital zero-crossing detector circuit is used for phase synchronization and frequency tracking to control the grid-tie power converter for an efficient energy conversion system in [4]. Zero-crossing point detection based methodology is proposed in [5] to estimate the synchronization between the signals. This technique observes the synchronization between signals by detecting the phase change within the half cycle. This methodology works well in the frequency range of 50 Hz to 52 Hz only. Power quality is analyzed based on measurements like RMS value, frequency and harmonics. For accurate measurements, proper zero-crossing detection is required. In [6], a comparative analysis between digital filters for ZCP detection in power quality measurement in presence of 3<sup>rd</sup> and 5<sup>th</sup> harmonics and noise is presented.

ZCP detection using a digital pulse-frequency modulator based on Field Programmable Gate Array (FPGA) is implemented in [7] to identify the zero current and zero voltage transition. This methodology is implemented to change the resonant pulse width in a quasi-resonant pulsed converter under different load conditions. Analog ZCP detection based on digital zero-crossing detection algorithm with signal reconstruction and least square fitting technologies is used in [8] for high precise time difference measurement in ultrasonic flow measurement unit. ZCP detection in line voltage based on a multistage filter, least square line fitting model and extrapolation of the ZCP is implemented in [9]. This methodology is implemented only on a 50 Hz sinusoidal signal but caes that this can be applicable to a signal upto 60 Hz. ZCP detection in inductor current for high current switched mode DC-DC converters is presented in [10]. In this paper, voltage polarity detector based on the transistor memory cell and auto zero-comparator is used for ZCP detection.

Identification of Safety operation area represented by back emf ZCP in a high speed Brushless DC (BLDC) motor in terms of free wheeling angle is implemented in [11]. The effect of Pulse Width Modulation (PWM) techniques mechanism on free wheeling angle is investigated. For given motor parameters, torque and speed area of BLDC motor are identified. A robust ZCP detection mechanism is developed in [12] using support vector machine. In this study, authors considered noise level up to 20 % and THD level 50 %. ZCP detection using voltage sensors, voltage shifter and micro controller is discussed in [13]. Machine learning is a powerful approach to find the solution for various problems in electrical engineering like load forecasting [14], [15] and [16] and health care [17] etc., Most of the researchers are also using machine learning based approach to detect the ZCP in distorted signals. In this

paper, also machine model called a decision tree classifier model is used for ZCP detection. ZCP detection in a distorted sinusoidal signal with a smaller range of noise and THD using decision tree model is discussed in [18]. ZCP detection in a distorted sinusoidal signal with a larger range of noise and THD using a logistic regression model is discussed in [19].

A digital frequency measuring approach is presented in [20] to solve the obstacle that the standard dual mixer time difference measurement method's single ZCP identification is sensitive to noise. Sinusoidal beat technology, multi-channel synchronous acquisition technology, and digital frequency measurement technology are all used in [21] for sensorless control of a BLDC motor based on ZCP detection in the Back Electromotive Force (BEMF). Estimation line back electromotive force via a sensor-less control technique is proposed in [22]. The phase relationship between the optimal commutation points of the brush-less direct current motor and the ZCPs of the line back-electromotive force is used to establish a commutation rule for different rotor locations in this study. Most of the researchers are also using machine learning based approach to detect the ZCP in distorted signals.

Main contributions of this paper are as follows:

- ZCP detection in a wide range of distorted signals by considering noise levels from 10 % to 60 %, THD level from 10 % to 60 %.
- Decision tree classifier which is a machine learning model is used for the first time for ZCP detection.
- New data consists of 96 datasets that are developed to work on ZCP detection problem and are available in <https://data.mendeley.com/datasets/d2hs6zt8gw/1>.
- Performance of the decision tree classifier model to detect ZCP in distorted signal with various window sizes is observed.

The remaining part of this paper is organized as: Sec. 2. describes dataset and machine learning model, Sec. 3. presents results discussion and Sec. 4. demonstrates the conclusions of the paper.

## 2. Methodology

This section consists of the architecture and training process of decision tree classifier used for the ZCP detection. The information about the complete data that is used to train and test the decision tree model is considered from [19]. The features like slope, correlation, intercept and Root Mean Square Error (RMSE) are extracted from various distorted signal as mentioned in [12].

### 2.1. Decision Tree Classifier

Decision Tree Classifier has tree based architecture which is used for both binary and categorical classification problems [23]. It follows a supervised learning [24] approach and is used to solve the classification problems like discussed in [25], [26], [27], [28] and also to solve regression problems like [29]. Any one of the two cost functions called entropy and Gini index can be used to construct the tree based on data samples. The mathematical representation of the entropy cost function is shown in Eq. (1) and Eq. (2), and for Gini index is shown in Eq. (3) and Eq. (4). Entropy is a measure of a random variable's uncertainty; it characterises the impurity of any arbitrary collection of samples. The greater the entropy, the greater the information content. The Gini Index is a statistical measure of how frequently a randomly selected piece is erroneously recognised. This suggests that a characteristic with a lower Gini index should be chosen. The performance of the decision tree classifier is measured in terms of accuracy [30], [31] and [32] as shown in Eq. (5). The complete work for ZCP detection in the form of block diagram is shown in Fig. 1. The main strengths of decision tree methods are as follows:

- Decision trees can provide rules that are easy to grasp.
- Decision trees perform classification with little processing.
- Decision trees can handle both continuous and discrete data.
- Decision trees clearly show which fields are most relevant for prediction or categorization.

$$Entropy (label) = - \sum_{i \in labels} -p(i.value) \log_2 p(i.value). \tag{1}$$

$$Entropy (label, feature) = \sum_{j \in featurecategories} -probability(j) Entropy(j). \tag{2}$$

$$Gini (feature.value) = 1 - \sum_{feature.value \in class} [p(feature.value)]^2. \tag{3}$$

$$Gini (class, feature) = \sum_{feature.value \in class} [p(feature.value)] Gini (feature.value). \tag{4}$$

$$Accuracy = \frac{TZCP + TNZCP}{TZCP + TNZCP + FZCP + FNZCP}. \tag{5}$$

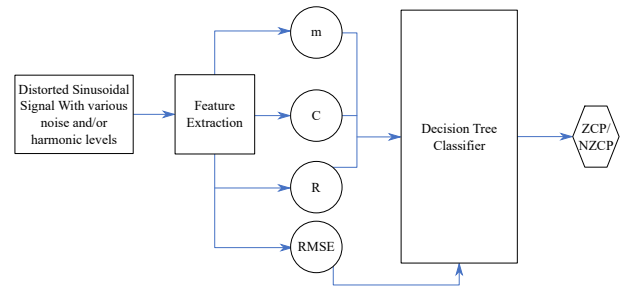


Fig. 1: Block diagram representation of proposed work.

## 3. Result Analysis

Decision tree classifier is trained with all 96 datasets which are created with various levels of noise, THD and window size in Google Colab. The performance of the model is observed in terms of accuracy.

### 3.1. Decision Tree Classifier Model Performance on Distorted Signal with Noise

The Decision Tree (DT) model is trained and tested on datasets which are created with a distorted sinusoidal signal with various noise levels from 10 % to 60 %. To get a better DT model, this model is trained by tuning the hyperparameters like depth of the tree and cost functions i.e. entropy and Gini index. Testing and training accuracy for the DT model on these 28 datasets for various combinations of several trees and the cost functions are presented in Tab. 1. The combination of depth of the tree and cost function that gave better testing accuracy is considered as optimal DT model for each dataset and is highlighted in Tab. 1.

The performance of the DT model on ZCP detection in distorted signals with various noise levels is observed in terms of testing and training accuracy with respect to various window sizes i.e. 5, 10, 15 and 20 are presented in Fig. 2. If testing accuracy of the DT model for different window sizes is equal then the model that gives better training accuracy considered as an optimal model. To avoid a DT model specific to a particular noise level, A new DT model is trained and tested on a distorted signal with all noise levels 10 % to 60 % using datasets "ZCP-Noise-25", "ZCP-Noise-26", "ZCP-Noise-27" and "ZCP-Noise-28" with different window sizes i.e. 5, 10, 15 and 20. The impact of window size on the performance of the DT model in terms of testing accuracy is presented in Fig. 3. From Fig. 3, it is observed that the DT model at performs well during training and testing with a window size of 20.

**Tab. 1:** Testing and training accuracy of DT model to detect ZCP in distorted signal with different noise levels.

Dataset	Entropy			Gini			Dataset	Entropy			Gini			
	Depth	Accuracy		Testing	Accuracy			Depth	Testing	Accuracy		Testing	Accuracy	
		Testing	Training		Testing	Training				Testing	Training		Testing	Training
ZCP-NOISE-01	12	<b>98.3</b>	<b>100</b>	97.3	100	7	ZCP-NOISE-15	10	<b>97.3</b>	<b>100</b>	95.6	100	9	
	9	97.3	99.7	97.6	99.7	6		9	97.3	99.8	94.3	98.2	6	
	6	97.3	98.9	97	99.1	4		6	97.3	97.9	96	96.8	4	
	3	97.6	97.8	96.3	98.1	2		3	93.6	94.9	95.3	93.5	2	
ZCP-NOISE-02	9	94	100	<b>94.6</b>	<b>100</b>	<b>11</b>	ZCP-NOISE-16	9	96	100	<b>96.3</b>	<b>100</b>	<b>9</b>	
	6	93.6	97.9	94.6	98.4	6		8	96	99.7	96	98.8	6	
	4	93.3	97.1	94.6	97.1	4		6	96	98.9	95.6	97.5	4	
	2	93.6	95.2	94.3	95.8	2		3	95	96.4	95.6	95.6	2	
ZCP-NOISE-03	9	98	100	95.6	96.8	9	ZCP-NOISE-17	12	96.3	100	96.6	100	9	
	6	97.6	99.5	<b>98.3</b>	<b>99.5</b>	<b>6</b>		9	96.3	99.5	96	99.5	6	
	4	97.6	98.5	98	98.5	4		6	96	98.9	96.6	98.8	4	
	2	95.6	96.8	95.3	96.8	2		<b>3</b>	<b>97.6</b>	<b>98.1</b>	95.6	97.5	2	
ZCP-NOISE-04	12	96.3	100	<b>96.6</b>	<b>100</b>	<b>11</b>	ZCP-NOISE-18	12	94.6	100	95	100	8	
	9	96.3	99.5	95	97.9	6		9	94.6	99.4	94	99.4	6	
	6	96.3	97.9	95	95.9	4		6	93.6	96.9	<b>95.6</b>	<b>97.7</b>	<b>4</b>	
	3	94.6	94.9	95	93.5	2		3	94.6	95.9	94.6	94.6	2	
ZCP-NOISE-05	9	96	100	<b>96</b>	<b>100</b>	<b>8</b>	ZCP-NOISE-19	9	94.3	100	<b>96</b>	<b>100</b>	<b>9</b>	
	6	95.6	99.1	96	99.2	6		8	94.6	99.8	95	99.7	6	
	4	96	98.7	95.6	98.7	4		6	95	99.7	94.3	98.7	4	
	2	95.6	97.5	95.6	97.5	2		3	94.3	97.2	95.3	96.2	2	
ZCP-NOISE-06	10	93	94.5	93.3	100	9	ZCP-NOISE-20	8	94.6	100	95.6	100	9	
	<b>9</b>	<b>96</b>	<b>99.8</b>	94	99.2	6		6	95.3	98.7	<b>96.3</b>	<b>99.4</b>	<b>6</b>	
	6	93.6	98.2	95.6	96.9	4		4	93.3	96.2	95.3	97.4	4	
	3	93.6	93.1	93	94.5	2		3	92.6	95.6	95	95.5	2	
ZCP-NOISE-07	11	96.6	100	97	97.9	8	ZCP-NOISE-21	6	96	100	96	100	6	
	<b>9</b>	<b>97.3</b>	<b>99.7</b>	96.6	99.5	6		<b>5</b>	<b>96.6</b>	<b>99.4</b>	96.6	99.5	4	
	6	96.6	98.9	97	98.5	4		4	96.6	98.8	96.6	98.8	3	
	3	97	97.9	96.3	97.5	2		3	96.6	98.1	96.6	98.1	2	
ZCP-NOISE-08	12	95.3	100	<b>97.3</b>	<b>100</b>	<b>10</b>	ZCP-NOISE-22	<b>13</b>	<b>93.3</b>	<b>100</b>	91.6	100	10	
	9	96	99.4	97	98.9	6		9	91.6	98.1	92.6	98.1	6	
	6	94.6	98.2	96.3	97.8	4		6	92.3	96.9	92.6	96.5	4	
	3	95.3	95.9	95.6	95.6	2		3	91.3	93.6	91.6	95.2	2	
ZCP-NOISE-09	6	95.3	95.9	97.6	100	6	ZCP-NOISE-23	<b>13</b>	<b>93.3</b>	<b>100</b>	91.6	95.2	10	
	4	98	98.5	97.6	99.8	5		8	91.6	97.7	91.6	99.5	6	
	<b>3</b>	<b>98.3</b>	<b>98.4</b>	97.6	99.4	4		6	92.3	96.9	92.6	96.5	4	
	2	97.6	98.1	97.6	98.1	2		3	91.3	93.6	91.6	95.2	2	
ZCP-NOISE-10	16	92.3	100	<b>94.3</b>	<b>100</b>	<b>11</b>	ZCP-NOISE-24	<b>8</b>	<b>95.3</b>	<b>100</b>	95	100	10	
	9	92.6	98.1	94	98.5	6		7	95.3	99.5	95	99.2	6	
	6	94	96.8	93	96.9	4		6	95	98.4	95	97.8	4	
	3	91.3	93.9	91.3	95.2	2		3	95	95.1	93.6	96.7	2	
ZCP-NOISE-11	8	97.6	100	97	100	8	ZCP-NOISE-25	27	94.9	100	95.7	100	18	
	7	96.6	99.8	96.6	99.8	6		14	95	98.3	95.5	98.3	10	
	6	96.6	99.8	97	98.8	4		6	96.1	96.8	96.1	97.1	6	
	3	98	97.9	<b>98</b>	<b>97.9</b>	<b>2</b>		3	96.3	96.1	<b>96.4</b>	<b>96</b>	<b>2</b>	
ZCP-NOISE-12	9	95.6	100	95.3	100	9	ZCP-NOISE-26	23	89.1	100	89.5	100	21	
	<b>8</b>	<b>96</b>	<b>99.7</b>	94.6	98.8	6		15	89.6	98.8	90.5	98	13	
	6	95.6	98.9	95.6	97.5	4		6	90.8	93	<b>90.8</b>	<b>96.4</b>	<b>10</b>	
	3	94.6	96.4	95.3	95.6	2		3	88.3	91	90	94.8	8	
ZCP-NOISE-13	<b>10</b>	<b>96.3</b>	<b>100</b>	96.3	100	11	ZCP-NOISE-27	24	92	100	92.2	100	23	
	9	96.3	99.8	96	99.1	6		20	92.3	99.6	91.8	99.4	16	
	6	96	99.1	96	98.7	4		<b>14</b>	<b>92.7</b>	<b>98.3</b>	91.7	95.9	9	
	3	96	98.5	95.6	97.5	2		8	92	94.6	90	91	2	
ZCP-NOISE-14	15	94	100	95	100	9	ZCP-NOISE-28	19	94	100	<b>94.7</b>	<b>100</b>	<b>19</b>	
	9	95.3	98.4	<b>96</b>	<b>98.9</b>	<b>6</b>		10	93.5	96.4	94.5	97.5	10	
	6	93.3	97.5	95.6	96.7	4		6	91.8	93.2	93.8	96.3	8	
	3	93.6	93.5	96.6	94.8	2		3	89.1	88.4	87.4	87.1	2	

The performance of the DT model on ZCP detection in distorted signals with various noise levels is observed in terms of testing and training accuracy with respect to various window sizes i.e. 5, 10, 15 and 20 are presented in Fig. 2. If testing accuracy of the DT model for different window sizes is equal then the model that gives better training accuracy considered as an optimal model. To avoid a DT model specific to a particular noise level, A new DT model is trained and tested on a distorted signal with all noise levels 10 % to 60 % using datasets "ZCP-Noise-25", "ZCP-Noise-26", "ZCP-Noise-27" and "ZCP-Noise-28" with different window sizes i.e. 5, 10, 15 and 20. The impact of window size on the performance of the DT model in terms of testing accuracy is presented in Fig. 3. From Fig. 3, it is observed that the DT model

at performs well during training and testing with a window size of 20.

Finally, the parameters of optimal DT models like cost function and depth of the tree, the performance of the optimal models in terms of training and testing accuracy with respect to each noise level are presented in Tab. 2. From Tab. 2, DT model has the maximum testing accuracy of 98.3 % with a noise level of 10 % and 30 % but if all noise signals are added then it has the minimum accuracy of 96.4 %. The information about correctly and incorrectly predicted ZCPs are presented in Tab. 3. From Tab. 3, it is observed that the DT model can detect ZCP with good accuracy for noise levels 10 % to 60 % but when all signals are combined then it has more false ZCP detection problems. The confusion matrix which is built with

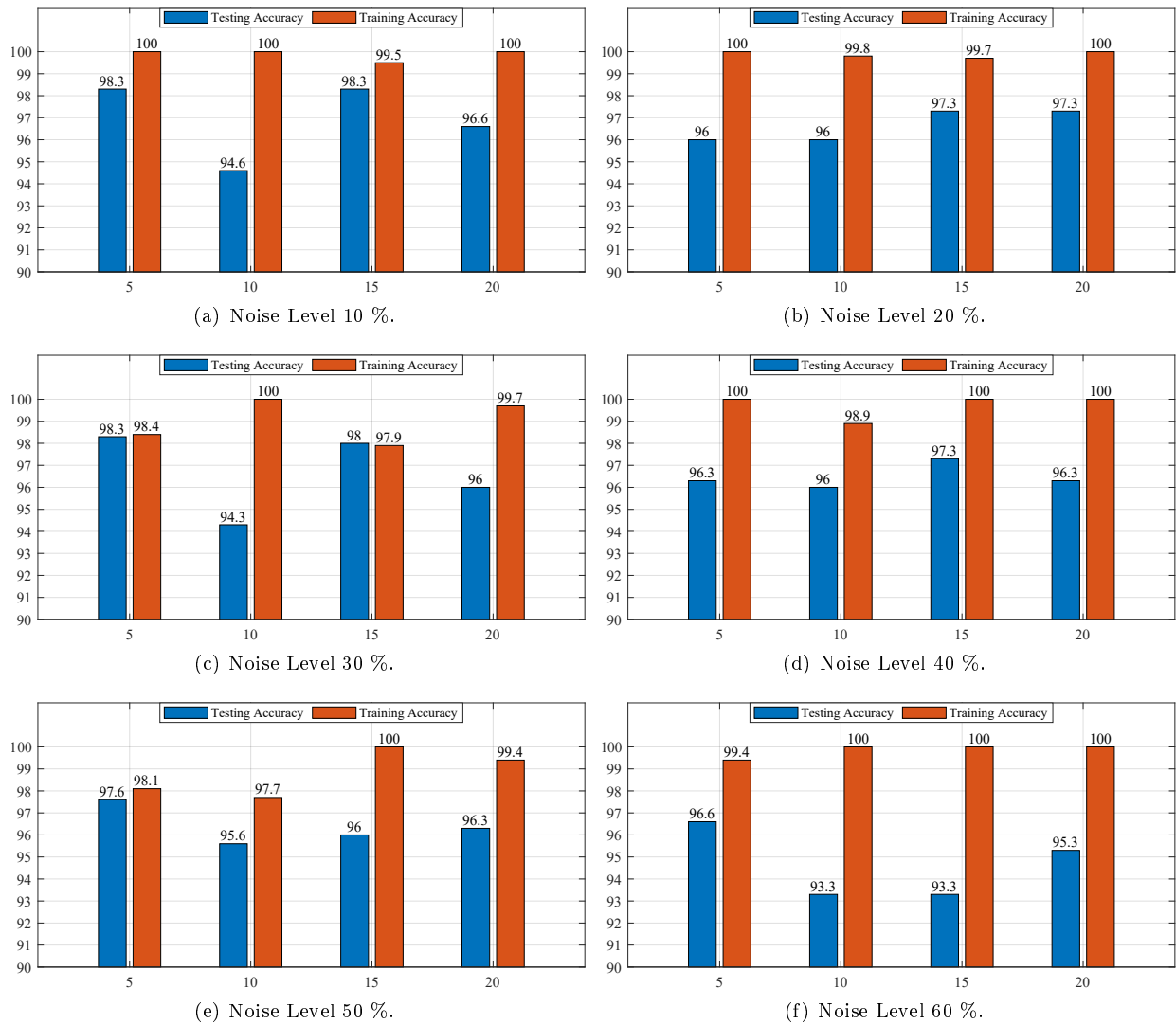


Fig. 2: Impact of window size on testing accuracy of DT model on distorted signal with various noise levels.

Tab. 2: Optimal DT model parameters for various noise levels.

Noise Level	Dataset	Accuracy		Depth	Cost function	
		Testing	Training		Entropy	Gini
10 %	ZCP-NOISE-01	98.3	100	12	Yes	No
20 %	ZCP-NOISE-08	97.3	100	10	No	Yes
30 %	ZCP-NOISE-09	98.3	98.4	3	Yes	No
40 %	ZCP-NOISE-15	97.3	100	10	Yes	No
50 %	ZCP-NOISE-17	97.6	98.1	3	Yes	No
60 %	ZCP-NOISE-21	96.6	99.4	4	Yes	No
10-60 %	ZCP-NOISE-25	96.4	96	2	No	Yes

Tab. 3: Training and testing accuracy details of optimal DT models for various noise signals.

Dataset	Testing Data				Training Data			
	NZCP		ZCP		NZCP		ZCP	
	TRUE	FALSE	TRUE	FALSE	TRUE	FALSE	TRUE	FALSE
ZCP-NOISE-01	289	4	6	1	682	0	15	0
ZCP-NOISE-08	245	5	46	4	566	0	131	0
ZCP-NOISE-09	284	2	11	3	662	2	24	9
ZCP-NOISE-15	269	4	23	4	633	0	64	0
ZCP-NOISE-17	281	6	12	1	663	1	21	12
ZCP-NOISE-21	279	8	11	2	663	1	30	3
ZCP-NOISE-25	1726	1	6	62	3995	5	29	159

testing data extracted from all distorted noisy signals is presented in Tab. 4.



**Fig. 3:** Impact of window size on testing accuracy of DT model on distorted signal with all noise levels.

**Tab. 4:** Confusion Matrix with testing data for various noise signals.

Confusion Matrix		Predicted	
		ZCP	NZCP
Actual	ZCP	115	30
	NZCP	77	3373

### 3.2. The Decision Tree Classifier Model Performance on Distorted Signal with THD

The DT model is trained and tested on datasets which are created with a distorted sinusoidal signal with various THD levels from 10 % to 60 %. To get the better the DT model, it is trained by tuning the hyperparameters like the depth of the tree and the cost functions i.e. entropy and Gini index. Testing and training accuracy for DT model on these 28 datasets for various combinations of several trees and the cost functions is presented in Tab. 5. The combination of the depth of the tree and cost function that gave better testing accuracy is considered as optimal DT model for each dataset and is highlighted in Tab. 5.

The performance of the DT model on ZCP detection in distorted signals with various THD levels is observed in terms of testing and training accuracy using various window sizes i.e. 5, 10, 15 and 20 are presented in Fig. 4. If testing accuracy of the DT model for different window sizes is equal then model which gives better training accuracy is considered as the optimal model. To avoid a DT model specific to a particular THD level, a new DT model was trained and tested on a distorted signal with all THD levels 10 % to 60 % using datasets "ZCP-THD-25", "ZCP-THD-26", "ZCP-THD-27" and "ZCP-THD-28" with different window sizes i.e. 5, 10, 15 and 20. The impact of window size on performance of the DT model in terms of testing accuracy is presented in Fig. 5. From Fig. 5, it is observed that the DT model

performs well during training and testing with a window size of 10.

Finally, the parameters of an optimal DT models like the cost function and the depth of the tree, the performance of the optimal models in terms of training and testing accuracy with respect to each THD level are presented in Tab. 6. From Tab. 6, the DT model has maximum testing accuracy of 100 % with noise level 10 % and 60 % but with all the noise signals, it has minimum accuracy of 96.4 %.

The information about correctly and incorrectly predicted ZCPs are presented in Tab. 7. From Tab. 7, it is observed that DT model has more false ZCP detection problem with combined signals. The impact of THD level on testing accuracy of the DT model is shown in Fig. 6, from the Fig. 6 it is observed that the performance of the DT model is slightly decreasing for distorted signals with all combined THD levels. Confusion matrix which is build with testing data that extracted from all distorted signals due to harmonics is presented in Tab. 8.

### 3.3. Decision Tree Classifier Performance on Distorted Signal with Harmonics and Noise

The DT model is trained and tested on datasets which are created with a distorted sinusoidal signal with various THD and noise levels. To get the better the DT model, it is trained by tuning the hyperparameters like the depth of the tree and the cost functions i.e. entropy and Gini index. Testing and training accuracy for DT model on these 40 datasets for various combinations of several trees and the cost functions are presented in Tab. 9. The combination of the depth of the tree and cost function that gave better testing accuracy is considered as optimal DT model for each dataset and is highlighted in Tab. 9.

The performance of the DT model on ZCP detection in distorted signals with various THD and noise level combination is observed in terms of testing and training accuracy using various window sizes i.e. 5, 10, 15 and 20 are presented in Fig. 7. If testing accuracy of the DT model for different window sizes is equal then model that gives better training accuracy is considered as the optimal model. To avoid a DT model specific to a particular THD and noise level combination, a new DT model was trained and tested on a distorted signal with all THD and noise level combinations using datasets "ZCP-NTHD-37", "ZCP-NTHD-38", "ZCP-NTHD-39" and "ZCP-NTHD-40" with different window sizes i.e. 5, 10, 15 and 20.

**Tab. 5:** Testing and training accuracy of DT model to detect ZCP in distorted signal with different THD levels.

Dataset	Entropy			Gini			Dataset	Entropy			Gini		
	Depth	Accuracy		Accuracy		Depth		Depth	Accuracy		Accuracy		Depth
		Testing	Training	Testing	Training				Testing	Training	Testing	Training	
ZCP-THD-01	2	<b>100</b>	<b>100</b>	100	100	2	ZCP-THD-15	5	<b>100</b>	<b>100</b>	100	100	5
	1	97.67	97.84	97.67	97.84	1		3	98	97.99	98	97.99	3
ZCP-THD-02	7	98.33	100	<b>100</b>	<b>100</b>	7	ZCP-THD-16	6	<b>100</b>	<b>100</b>	100	100	6
	5	98	99.71	98	99.57	5		3	96	95.98	96	95.98	3
	3	97.66	98.56	97.67	98.56	3		2	91	90.67	91	90.67	2
	1	95.67	97.27	95.67	97.27	1		1	87.33	84.36	91	90.67	1
ZCP-THD-03	2	<b>100</b>	<b>100</b>	100	100	2	ZCP-THD-17	3	<b>100</b>	<b>100</b>	100	100	3
	1	97.66	97.85	97.66	97.85	1		2	100	99.85	100	99.85	2
ZCP-THD-04	3	<b>100</b>	<b>100</b>	100	100	2	ZCP-THD-18	8	<b>100</b>	<b>100</b>	100	100	8
	2	97.66	97.85	97.67	97.85	1		6	96.67	97.13	96.67	97.13	6
	1	95.33	95.26					4	93.33	94.41	93.33	94.41	4
ZCP-THD-05	3	99.67	100	99.67	100	3	ZCP-THD-19	2	<b>100</b>	<b>100</b>	100	100	2
	2	<b>100</b>	<b>99.85</b>	100	99.85	2		1	91.67	89.67	91.67	89.67	1
ZCP-THD-06	2	<b>100</b>	<b>100</b>	100	100	2	ZCP-THD-20	8	<b>98</b>	<b>100</b>	98	100	8
	1	95.33	95.26	97.66	97.84	1		6	98	99.14	98.33	99.14	6
ZCP-THD-07	2	<b>100</b>	<b>100</b>	100	100	2	ZCP-THD-21	7	<b>100</b>	<b>100</b>	100	100	7
	1	95	95.41	95	95.41	1		5	98	97.7	98.33	97.84	5
ZCP-THD-08	7	97	100	97	100	7	ZCP-THD-22	8	<b>100</b>	<b>100</b>	100	100	8
	5	<b>97.33</b>	<b>99.71</b>	97.33	99.71	5		6	97.33	98.42	97.33	98.42	6
	3	96.33	98.13	96.33	98.13	3		4	90.67	92.4	90.67	92.4	4
	1	93	94.84	93	94.84	1		2	87	85.93	91.67	92.25	2
ZCP-THD-09	2	<b>100</b>	<b>100</b>	100	100	2	ZCP-THD-23	3	<b>100</b>	<b>100</b>	100	100	3
	1	95	95.41	95	95.41	1		2	100	99.85	100	99.85	2
ZCP-THD-10	2	<b>100</b>	<b>100</b>	100	100	2	ZCP-THD-24	10	<b>100</b>	<b>100</b>	100	100	9
	1	95	95.41	95	95.41	1		8	96.33	96.41	96	96.27	7
ZCP-THD-11	3	<b>100</b>	<b>100</b>	100	100	3	ZCP-THD-25	9	99.55	100	99.44	100	9
	2	100	99.85	100	99.85	2		7	99.27	99.95	<b>99.61</b>	<b>99.78</b>	7
	1	95	95.41	95	95.41	1		5	98.9	99.23	98.1	98.1	5
ZCP-THD-12	3	100	100	<b>100</b>	<b>100</b>	5	ZCP-THD-26	11	99.49	100	99.33	100	12
	2	95	95.41	97.67	98.13	3		9	<b>99.66</b>	<b>99.95</b>	99.27	99.88	9
	1	91.67	89.67	97.67	98.13	2		6	99.11	99.4	97.43	98.49	6
ZCP-THD-13	5	99.67	100	<b>99.67</b>	<b>100</b>	6	ZCP-THD-27	17	<b>99.44</b>	<b>100</b>	99.16	100	16
	3	99.33	99.13	99.33	99.13	3		12	98.38	99.31	98.66	99.54	12
	2	99.33	99.13	99.33	99.13	2		8	96.49	96.63	96.49	96.96	8
	1	93.33	92.53	93.33	92.53	1		4	91.69	92.12	91.69	92.12	4
ZCP-THD-14	6	<b>98.33</b>	<b>100</b>	98.33	100	6	ZCP-THD-28	22	99.27	100	<b>99.49</b>	<b>100</b>	13
	3	97.33	98.27	97.66	98.85	3		16	98.38	99.16	97.04	97.87	8
	2	97.66	98.13	97.33	98.27	2		9	94.31	94.26	93.7	93.69	5
	1	91.67	91.82	91.33	91.96	1		2	81.89	81.2	83.45	82.92	2

**Tab. 6:** Optimal DT model parameters for various THD levels.

Noise Level	Dataset	Accuracy		Depth	Cost function	
		Testing	Training		Entropy	Gini
10 %	ZCP-THD-01	<b>100</b>	<b>100</b>	2	Yes	No
20 %	ZCP-THD-06	<b>100</b>	<b>100</b>	2	Yes	No
30 %	ZCP-THD-09	<b>100</b>	<b>100</b>	2	Yes	No
40 %	ZCP-THD-15	<b>100</b>	<b>100</b>	5	Yes	No
50 %	ZCP-THD-17	<b>100</b>	<b>100</b>	3	Yes	No
60 %	ZCP-THD-21	<b>100</b>	<b>100</b>	7	Yes	No
10-60 %	ZCP-THD-26	<b>99.66</b>	<b>99.95</b>	9	Yes	No

**Tab. 7:** Training and testing accuracy details of optimal DT models for various THD signals.

Dataset	Testing Data				Training Data			
	NZCP		ZCP		NZCP		ZCP	
	TRUE	FALSE	TRUE	FALSE	TRUE	FALSE	TRUE	FALSE
10 %	286	0	14	0	664	0	33	0
20 %	286	0	14	0	664	0	33	0
30 %	275	0	25	0	625	0	72	0
40 %	262	0	38	0	588	0	109	0
50 %	262	0	38	0	588	0	109	0
60 %	248	0	52	0	552	0	145	0
10-60 %	1635	0	154	6	3765	0	420	2

Finally, the parameters of an optimal DT models like the cost function and the depth of the tree, the performance of the optimal models in terms of training and testing accuracy with respect to each THD and noise level combination are presented in Tab. 10.

From Tab. 10, the DT model has minimum testing accuracy of 99.4 % with all noise and THD and noise level combinations due to highly diversified data.

The information regarding correctly and incorrectly predicted ZCPs are presented in Tab. 11.

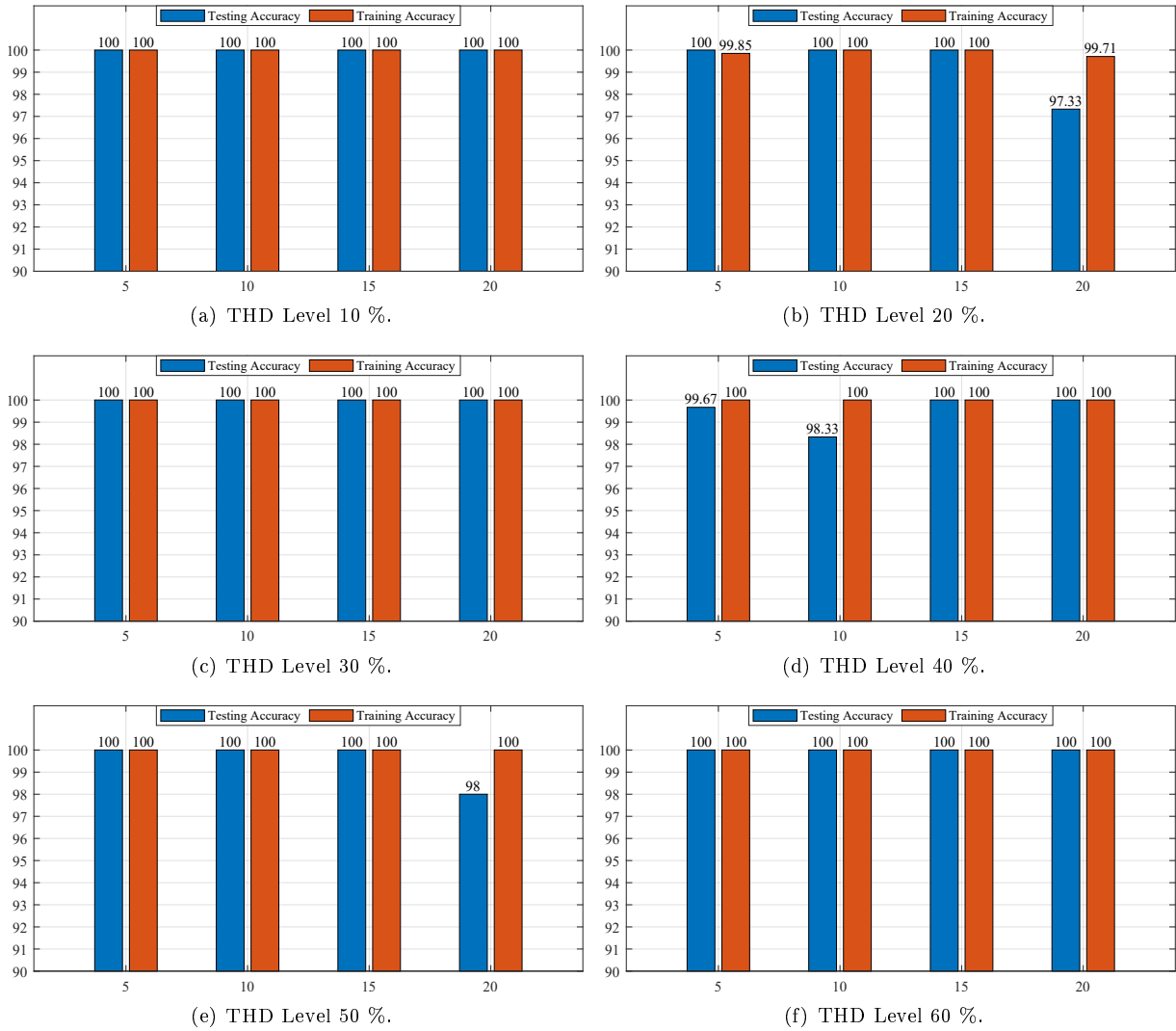


Fig. 4: Impact of window size on testing accuracy of DT model on distorted signal with various THD levels.

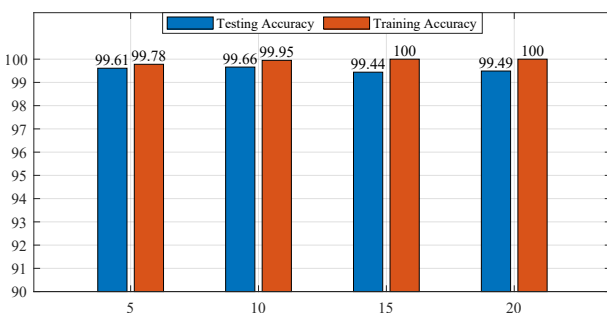


Fig. 5: Impact of window size on testing accuracy of DT model on distorted signal with all THD levels.

Tab. 8: Confusion Matrix with testing data for all THD signals.

Confusion Matrix		Predicted	
		ZCP	NZCP
Actual	ZCP	181	0
	NZCP	0	3254

From Tab. 11, it is observed that the DT model has more false ZCP detection problem with consideration of all THD and noise level combinations. The confusion matrix which is built with testing data that extracted from all distorted signals due to both noise and harmonics is presented in Tab. 12.

### 3.4. Comparative Analysis

The comparison of the performance of each model i.e., logistic regression and decision tree classifier in terms of testing accuracy for detection of ZCP on various signals is presented in this section.

#### 1) Comparative Analysis on Noise Signals

Comparison between machine learning models i.e. logistic regression and decision tree classifier based on the performance of ZCP detection on various



**Tab. 9:** Testing and training accuracy of DT model to detect ZCP in distorted signal with different THD and noise level combinations.

Dataset	Entropy			Gini			Dataset	Entropy			Gini		
	Depth	Accuracy		Accuracy		Depth		Depth	Accuracy		Accuracy		Depth
		Testing	Training	Testing	Training				Testing	Training	Testing	Training	
ZCP-NTHD-01	<b>2</b>	<b>99.66</b>	<b>100</b>	99.66	100	3	ZCP-NTHD-22	<b>3</b>	<b>100</b>	<b>100</b>	100	100	3
	1	96.66	96.7	99	98.99	2		2	100	99.85	100	99.85	2
ZCP-NTHD-02	<b>4</b>	<b>99.66</b>	<b>100</b>	99	100	3	ZCP-NTHD-23	1	95	95.4	95	95.4	1
	3	99	99.85	97.66	98.99	2		<b>3</b>	<b>100</b>	<b>100</b>	100	100	3
ZCP-NTHD-03	2	96.66	98.42	97	98.56	1	2	100	99.85	100	99.85	2	
	6	98.33	100	98	100	5	1	93.33	92.53	93.33	92.53	1	
	4	<b>98.66</b>	<b>99.71</b>	97.33	99.13	3	20	97.33	100	<b>98</b>	<b>100</b>	<b>19</b>	
	3	97.33	99.13	97	98.56	2	15	94	95.98	92	93.68	13	
ZCP-NTHD-04	2	96.66	98.42	97	98.56	1	10	89.33	89.23	88.33	86.94	7	
	6	<b>98.66</b>	<b>100</b>	97.66	100	9	5	86	85.22	83	81.2	1	
	4	98	98.7	97.33	99.28	6	<b>19</b>	<b>98</b>	<b>100</b>	97.66	100	15	
ZCP-NTHD-05	3	99.33	99.33	98.33	98.27	3	13	98	98.42	97	98.42	10	
	2	96.66	94.83	95.66	95.55	2	7	96.33	96.84	96.33	96.98	5	
ZCP-NTHD-06	<b>2</b>	<b>99.66</b>	<b>100</b>	99.66	100	3	1	95.33	95.26	95.33	95.4	2	
	1	94	95.83	99.66	99.85	2	<b>3</b>	<b>100</b>	<b>100</b>	100	100	3	
ZCP-NTHD-07	<b>4</b>	<b>99.33</b>	<b>100</b>	99.33	100	3	2	100	99.85	100	99.85	2	
	3	99.33	99.825	99.33	99.42	2	1	95	95.4	95	95.4	1	
ZCP-NTHD-08	2	99.66	99.56	98.66	98.13	1	<b>2</b>	<b>100</b>	<b>100</b>	100	100	2	
	<b>3</b>	<b>99.66</b>	<b>100</b>	99.66	100	3	1	93.33	92.53	93.33	92.53	1	
ZCP-NTHD-09	2	99.66	99.85	99.66	99.85	2	<b>2</b>	<b>99.33</b>	<b>100</b>	99.33	100	2	
	1	93.66	99.53	93.66	92.53	1	1	91	89.67	91	89.67	1	
ZCP-NTHD-10	6	97.66	100	97	100	6	5	99.66	100	<b>100</b>	<b>100</b>	<b>4</b>	
	4	97.66	99.85	97.66	99.85	4	3	99.33	99.42	100	99.71	3	
	<b>3</b>	<b>98</b>	<b>95.56</b>	98	99.56	3	2	99.33	99.28	100	99.28	2	
	2	98	99.42	98	99.56	2	1	94.33	97.13	98	98.56	1	
ZCP-NTHD-11	<b>4</b>	<b>99.66</b>	<b>100</b>	99.66	100	4	<b>3</b>	<b>100</b>	<b>100</b>	100	100	3	
	3	99.33	99.56	99	99.56	3	2	100	99.85	100	99.85	2	
	2	99.33	99.42	99.33	99.42	2	1	95	95.4	95	95.4	1	
ZCP-NTHD-12	1	98.33	97.77	98.33	97.77	1	7	97	100	97	100	7	
	<b>3</b>	<b>100</b>	<b>100</b>	100	100	3	5	97.66	99.85	96	99.85	5	
ZCP-NTHD-13	2	100	99.85	100	99.85	2	3	97	98.99	97.66	99.13	3	
	1	95	95.4	95	95.4	1	1	97.33	97.13	<b>98.66</b>	<b>97.41</b>	<b>1</b>	
ZCP-NTHD-14	<b>3</b>	<b>99.66</b>	<b>100</b>	99.66	100	3	6	97.33	100	97	100	7	
	2	99.66	99.85	99.66	99.85	2	4	97.33	99.13	<b>97.66</b>	<b>99.71</b>	<b>5</b>	
ZCP-NTHD-15	1	93.66	92.53	93.66	92.53	1	3	97.33	98.7	97	98.42	3	
	<b>2</b>	<b>100</b>	<b>100</b>	100	100	2	2	97.33	97.27	97.33	99.13	1	
ZCP-NTHD-16	1	91.66	89.69	91.66	89.69	1	<b>6</b>	<b>100</b>	<b>100</b>	99.66	100	6	
	<b>4</b>	<b>100</b>	<b>100</b>	100	100	4	4	99.33	99.28	99.66	99.71	4	
	3	99.33	99.56	99.56	99.33	3	3	98.66	99.13	99.66	99.71	3	
ZCP-NTHD-17	2	99.33	99.42	99.33	99.42	2	2	98.66	99.13	98	97.27	2	
	1	98.33	97.7	98.33	97.7	1	4	99	100	99	100	4	
ZCP-NTHD-18	<b>3</b>	<b>100</b>	<b>100</b>	100	100	3	<b>3</b>	<b>99.33</b>	<b>99.85</b>	99.33	99.85	3	
	2	100	99.85	100	99.85	2	2	99.33	99.13	99.33	99.13	2	
ZCP-NTHD-19	1	95	95.4	95	95.4	1	1	99.33	99.7	99.33	98.7	1	
	<b>3</b>	<b>99.66</b>	<b>100</b>	99.66	100	3	<b>7</b>	<b>98.66</b>	<b>100</b>	98.66	100	6	
ZCP-NTHD-20	2	99.66	99.85	99.66	99.85	2	5	98.66	99.28	98.33	99.13	4	
	1	93.66	92.53	93.66	92.53	1	3	98.33	97.84	98.66	98.7	3	
ZCP-NTHD-21	<b>3</b>	<b>100</b>	<b>100</b>	100	100	3	1	97.33	97.27	98	97.41	2	
	2	100	99.85	100	99.85	2	<b>9</b>	<b>96</b>	<b>100</b>	94.33	100	10	
ZCP-NTHD-22	1	91.66	89.67	91.66	89.67	1	6	94.33	98.99	93.66	99.56	7	
	<b>4</b>	<b>100</b>	<b>100</b>	100	100	4	3	94.33	96.98	94.66	97.27	4	
	3	99.33	99.56	99.33	99.56	3	2	89	88.23	83.33	85.07	1	
	2	99.33	99.42	99.33	99.42	2	<b>24</b>	<b>99.4</b>	<b>100</b>	99.21	100	24	
ZCP-NTHD-23	1	98.33	97.7	98.33	97.7	1	18	99.21	99.84	99.03	99.84	18	
	<b>4</b>	<b>99.66</b>	<b>100</b>	99.66	100	4	12	99.21	99.64	99.07	99.63	12	
ZCP-NTHD-24	3	99.33	99.85	99.33	99.85	3	6	97.99	97.99	98.06	98.04	6	
	2	100	99.71	100	99.71	2	<b>10</b>	<b>99.44</b>	<b>100</b>	99.4	100	11	
ZCP-NTHD-25	1	95	95.26	95	95.26	1	7	99.33	99.77	99.07	99.61	7	
	<b>8</b>	<b>98</b>	<b>100</b>	98	100	8	4	98.4	98.29	98.1	97.93	4	
ZCP-NTHD-26	6	98	99.85	98	99.85	6	1	93.68	93.28	93.68	93.56	1	
	4	98	99.71	98	99.71	4	<b>13</b>	<b>98.84</b>	<b>100</b>	<b>98.88</b>	100	12	
ZCP-NTHD-27	2	98	99.28	98	99.28	2	9	98.84	99.72	98.95	99.93	9	
	6	97.66	100	97.66	100	6	6	98.51	98.93	98.73	99.23	6	
ZCP-NTHD-28	4	97.66	99.85	97.66	99.85	4	3	95.02	94.85	95.02	94.85	3	
	<b>3</b>	<b>98</b>	<b>99.42</b>	98	99.42	3	22	98.17	100	<b>98.32</b>	<b>100</b>	<b>21</b>	
ZCP-NTHD-29	2	98	99.42	98	99.42	2	17	97.91	99.53	97.91	99.52	17	
	<b>4</b>	<b>100</b>	<b>100</b>	100	100	4	12	97.43	98.67	97.58	98.59	12	
ZCP-NTHD-30	3	100	99.85	100	99.85	3	7	96.43	96.78	96.84	96.95	7	

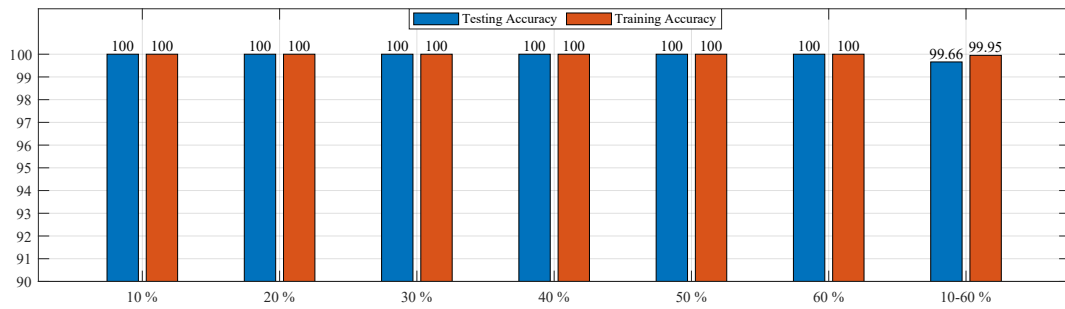


Fig. 6: Impact of THD level on testing accuracy of DT model.

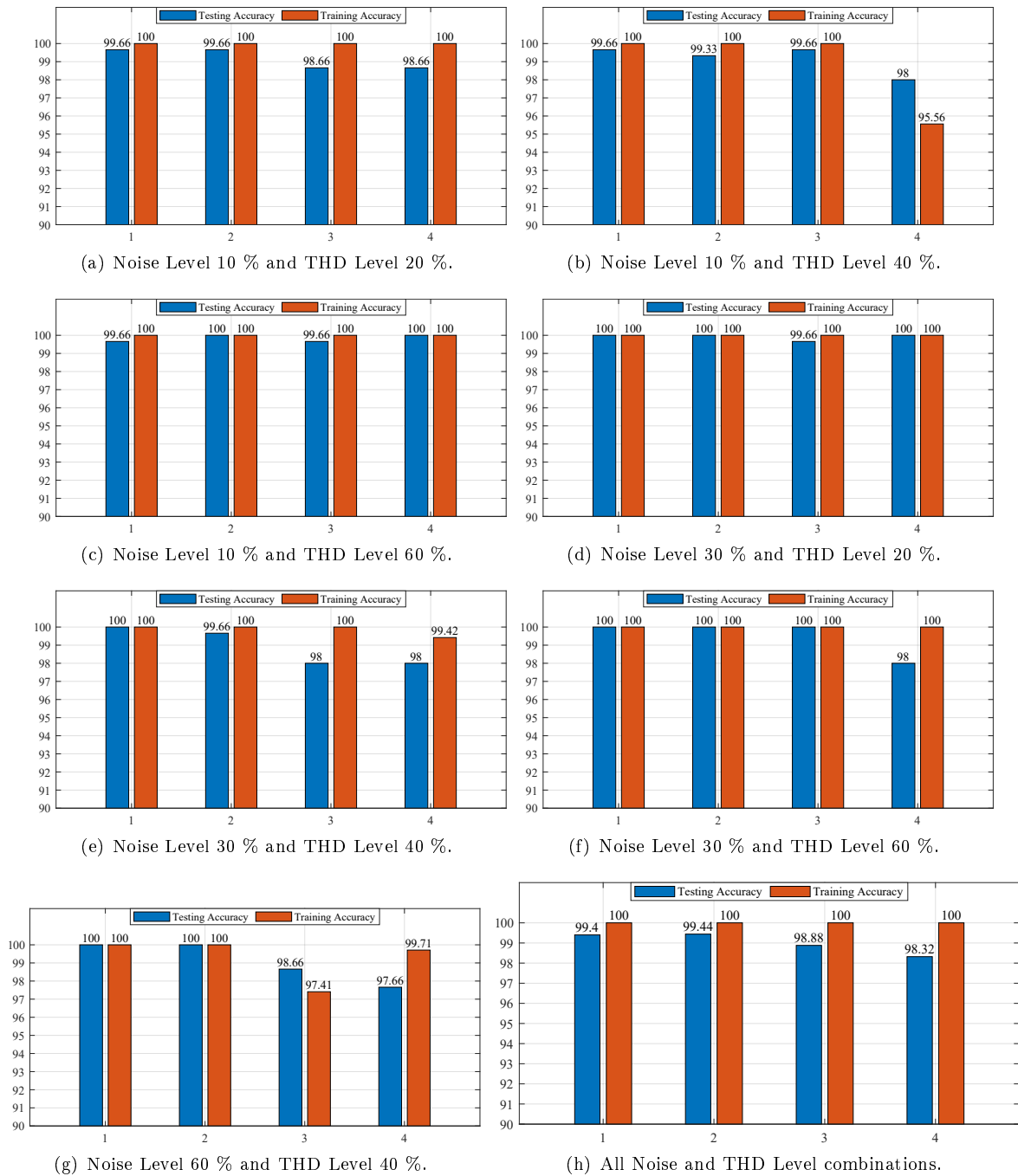


Fig. 7: Impact of window size on testing accuracy of DT model on distorted signal with various THD and noise level combinations.

Tab. 10: Optimal DT model parameters for various THD and noise level combinations.

Noise Level	THD	Dataset	Accuracy		Depth	Cost function	
			Testing	Training		Entropy	Gini
10 %	20 %	ZCP-NTHD-02	<b>99.66</b>	<b>100</b>	4	Yes	No
10 %	40 %	ZCP-NTHD-07	<b>99.66</b>	<b>100</b>	3	Yes	No
10 %	60 %	ZCP-NTHD-10	<b>100</b>	<b>100</b>	3	Yes	No
30 %	20 %	ZCP-NTHD-13	<b>100</b>	<b>100</b>	3	Yes	No
30 %	40 %	ZCP-NTHD-17	<b>100</b>	<b>100</b>	4	Yes	No
30 %	60 %	ZCP-NTHD-21	<b>100</b>	<b>100</b>	4	Yes	No
60 %	20 %	ZCP-NTHD-26	<b>100</b>	<b>100</b>	3	Yes	No
60 %	40 %	ZCP-NTHD-29	<b>100</b>	<b>100</b>	4	NO	YES
60 %	60 %	ZCP-NTHD-33	<b>100</b>	<b>100</b>	6	Yes	No
10 %, 30 %, 60 %	20 %, 40 %, 60 %	ZCP-NTHD-37	<b>99.4</b>	<b>100</b>	24	Yes	No

Tab. 11: True and false zcp detection information of optimal DT models for various THD and noise level combinations.

Dataset	Testing Data				Training Data			
	NZCP		ZCP		NZCP		ZCP	
	TRUE	FALSE	TRUE	FALSE	TRUE	FALSE	TRUE	FALSE
ZCP-NTHD-02	275	2	22	1	626	0	68	0
ZCP-NTHD-07	262	1	37	0	588	0	109	0
ZCP-NTHD-10	275	0	25	0	625	0	72	0
ZCP-NTHD-13	286	2	12	0	664	3	30	0
ZCP-NTHD-17	288	0	12	0	667	0	30	0
ZCP-NTHD-21	288	0	12	0	664	0	33	0
ZCP-NTHD-26	275	0	25	0	625	0	72	0
ZCP-NTHD-29	285	1	14	0	664	0	33	0
ZCP-NTHD-33	286	0	14	0	664	0	33	0
ZCP-NTHD-37	2570	0	103	11	5987	0	294	0

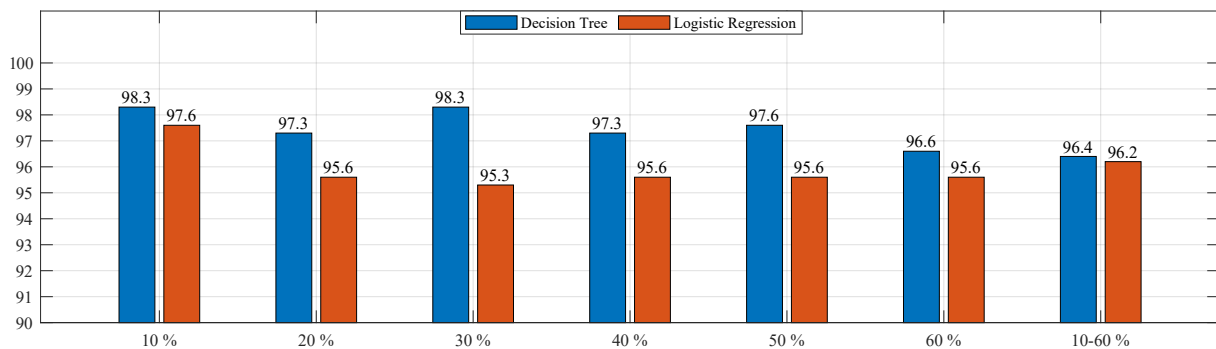


Fig. 8: Comparison of performance of machine learning models on noisy signals.

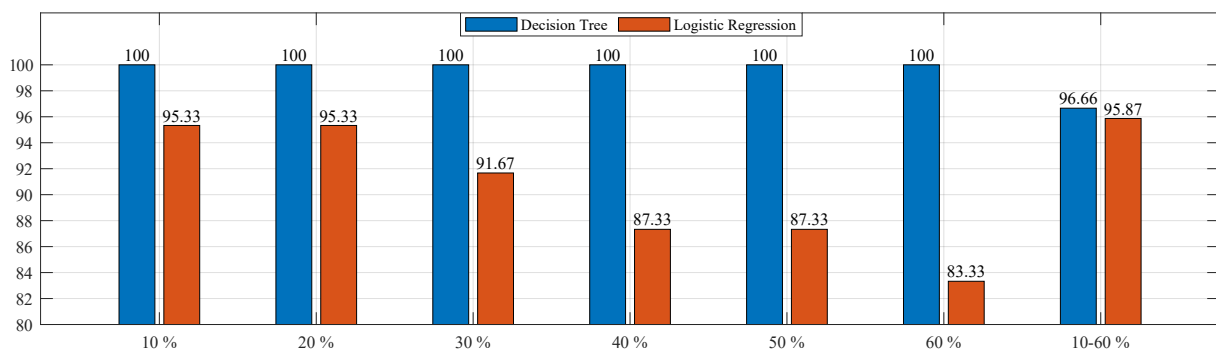


Fig. 9: Comparison of performance of machine learning models on harmonic signals.

noise signals in terms of testing accuracy is presented in Fig. 8. From Fig. 8, it is observed that the decision tree classifier model can detect ZCPs on a noisy signal with good accuracy than the logistic regression model.

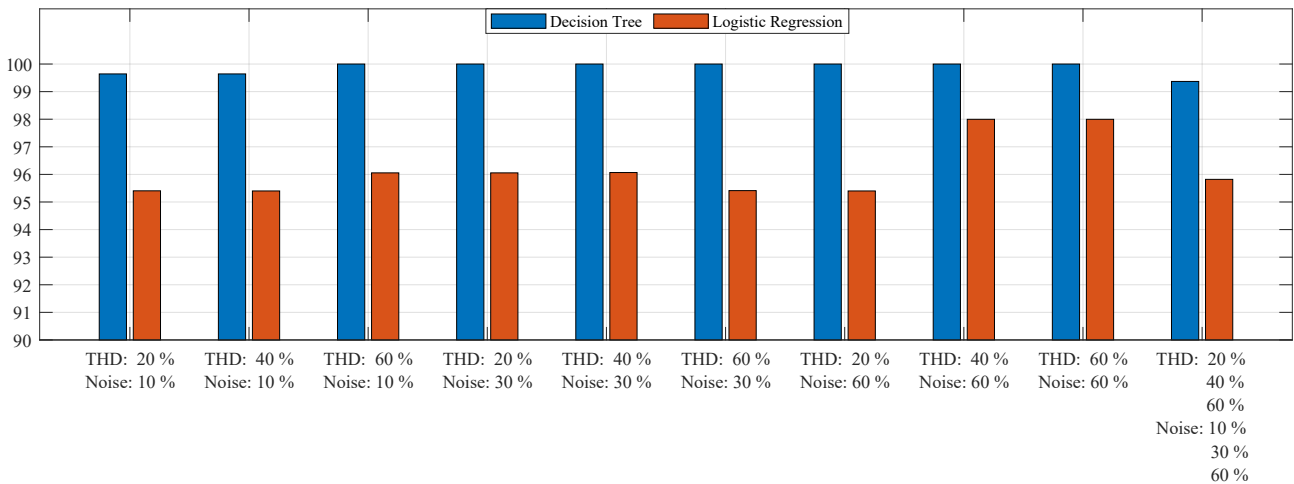


Fig. 10: Comparison of performance of machine learning models on combined harmonic and noise signals.

Tab. 12: Confusion Matrix with testing data for all combined noise and THD distorted signals.

Confusion Matrix		Predicted	
		ZCP	NZCP
Actual	ZCP	173	1
	NZCP	6	5090

### 2) Comparative Analysis on Harmonic Signals

Comparison between machine learning models i.e., the logistic regression and the decision tree classifier based on the performance of ZCP detection on various harmonic signals in terms of testing accuracy is presented in Fig. 9. From Fig. 9, it is observed that the decision tree classifier model can detect ZCPs on harmonic signals with better accuracy than the logistic regression models.

### 3) Comparative Analysis on Harmonics and Noise Combined Signals

Comparison between machine learning models i.e., the logistic regression and the decision tree classifier based on the performance of ZCP detection on various harmonics and noise combined signals in terms of testing accuracy is presented in Fig. 10. From Fig. 10, it is observed that the decision tree classifier model can detect ZCPs on combined noise and harmonic signals with better accuracy than the logistic regression model.

## 4. Conclusions

Accurate ZCP detection in a distorted signal is a complex task and essential to operate the power system network without power quality issues, protect the system against faults and for efficient power electronics

converter controller design. In this study, the performance of the decision tree classifier model on accurate ZCP detection in a distorted sinusoidal signal is discussed.

The distorted sinusoidal signal is generated in MATLAB with various noise and THD levels and from each signal features like slope, intercept, correlation and RMSE are extracted in order to prepare the data for training and testing the decision tree classifier. From the observations based on simulation results, decision tree classifier model can predict ZCP points with good accuracy and is performing better than the logistic regression model.

ZCP detection problem is further extended by incorporating the voltage swell in the sinusoidal signal also by considering other machine learning models like random forest, deep learning sequence models like recurrent neural network, long-short term memory and gated recurrent unit.

## Author Contributions

V.V., P.J., E.R. and S.S. developed the theoretical formalism, performed the analytic calculations, Data Collection and performed the numerical simulations. V.V. and S.R.S. contributed to the analysis of the results and to the proofreading of this manuscript. All the authors provided critical feedback, helped to shape the research and conduct the analysis and thus contributed to the final version of the manuscript.

## References

[1] VEERAMSETTY, V., B. R. EDUDODLA and S. R. SALKUTI. Zero-Crossing Point

- Detection of Sinusoidal Signal in Presence of Noise and Harmonics Using Deep Neural Networks. *Algorithms*. 2021, vol. 14, iss. 11, pp. 1–19. ISSN 1999-4893. DOI: 10.3390/a14110329.
- [2] JIANG, Q., C. BI and R. HUANG. A new phase-delay-free method to detect back EMF zero-crossing points for sensorless control of spindle motors. *IEEE Transactions on Magnetics*. 2005, vol. 41, iss. 7, pp. 2287–2294. ISSN 1941-0069. DOI: 10.1109/TMAG.2005.851841.
- [3] YANG, L., Z. Q. ZHU, B. SHUANG and H. BIN. Adaptive Threshold Correction Strategy for Sensorless High-Speed Brushless DC Drives Considering Zero-Crossing-Point Deviation. *IEEE Transactions on Industrial Electronics*. 2020, vol. 67, iss. 7, pp. 5246–5257. ISSN 1557-9948. DOI: 10.1109/TIE.2019.2931501.
- [4] MAITI, A., P. SYAM and K. MUKHERJEE. Alternate computation of the unit vectors synthesis towards synchronization of current-controlled grid-tie converter for renewable power system: An embedded outlook. *Engineering Science and Technology, an International Journal*. 2022, vol. 28, iss. 1, pp. 1–9. ISSN 2215-0986. DOI: 10.1016/j.jestch.2021.06.003.
- [5] SAXENA, H., A. SINGH and J. N. RAI. Design and Testing of Frequency Adaptive Zero-Crossing Detector as a Synchronizing Technique. In: *2020 IEEE 9th Power India International Conference (PIICON)*. Sonapat: IEEE, 2020, pp. 1–6. ISBN 978-1-72816-664-3. DOI: 10.1109/PIICON49524.2020.9113009.
- [6] RODRIGUES, N. M., F. M. JANEIRO and P. M. RAMOS. Digital filter performance for zero crossing detection in power quality embedded measurement systems. In: *2018 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*. Houston: IEEE, 2018, pp. 1–6. ISBN 978-1-5386-2222-3. DOI: 10.1109/I2MTC.2018.8409701.
- [7] YERSHOV, R. D. FPGA-Based Pulse-Frequency Modulator with Adaptive Zero-Crossing Detection for Quasi-Resonant Pulsed Converters. In: *2018 IEEE 38th International Conference on Electronics and Nanotechnology (ELNANO)*. Kyiv: IEEE, 2018, pp. 721–726. ISBN 978-1-5386-6383-7. DOI: 10.1109/ELNANO.2018.8477499.
- [8] CHEN, Z. and Z. LI. Robust Precise Time Difference Estimation Based on Digital Zero-Crossing Detection Algorithm. *IEEE Transactions on Instrumentation and Measurement*. 2016, vol. 65, iss. 8, pp. 1739–1748. ISSN 1557-9662. DOI: 10.1109/TIM.2016.2562318.
- [9] PATIL, T. and S. GHORAI. Robust zero-crossing detection of distorted line voltage using line fitting. In: *2016 International Conference on Electrical, Electronics, Communication, Computer and Optimization Techniques (ICEECOT)*. Mysuru: IEEE, 2016, pp. 92–96. ISBN 978-1-5090-4697-3. DOI: 10.1109/ICEECOT.2016.7955192.
- [10] MICHAL, V. Inductor Current Zero-Crossing Detector and CCM/DCM Boundary Detector for Integrated High-Current Switched-Mode DC–DC Converters. *IEEE Transactions on Power Electronics*. 2014, vol. 29, iss. 10, pp. 5384–5391. ISSN 1941-0107. DOI: 10.1109/TPEL.2013.2292600.
- [11] YANG, L., Z. Q. ZHU, H. BIN, Z. ZHANG and L. GONG. Safety Operation Area of Zero-Crossing Detection-Based Sensorless High-Speed BLDC Motor Drives. *IEEE Transactions on Industry Applications*. 2020, vol. 56, iss. 6, pp. 6456–6466. ISSN 1939-9367. DOI: 10.1109/TIA.2020.3012594.
- [12] GHOSH, M., C. KOLEY and N. K. ROY. Robust support vector machine-based zero-crossing detector for different power system applications. *IET Science, Measurement & Technology*. 2019, vol. 13, iss. 1, pp. 83–89. ISSN 1751-8830. DOI: 10.1049/iet-smt.2018.5025.
- [13] VOROBYOV, M. and K. VITOLS. Low-Cost Voltage Zero-Crossing Detector for AC-Grid Applications. *Electrical, Control and Communication Engineering*. 2014, vol. 6, iss. 1, pp. 32–37. ISSN 2255-9159. DOI: 10.2478/ecce-2014-0015.
- [14] VEERAMSETTY, V., A. MOHNOT, G. SINGAL and S. R. SALKUTI. Short Term Active Power Load Prediction on A 33/11 kV Substation Using Regression Models. *Energies*. 2021, vol. 14, iss. 11, pp. 1–21. ISSN 1996-1073. DOI: 10.3390/en14112981.
- [15] VEERAMSETTY, V., D. R. CHANDRA and S. R. SALKUTI. Short-term electric power load forecasting using factor analysis and long short-term memory for smart cities. *International Journal of Circuit Theory and Applications*. 2021, vol. 49, iss. 6, pp. 1678–1703. ISSN 1097-007X. DOI: 10.1002/cta.2928.
- [16] VEERAMSETTY, V., K. R. REDDY, M. SANTHOSH, A. MOHNOT and G. SINGAL. Short-term electric power load forecasting using random forest and gated recurrent unit. *Electrical Engineering*. 2022, vol. 104, iss. 1, pp. 307–329. ISSN 1432-0487. DOI: 10.1007/s00202-021-01376-5.

- [17] VEERAMSETTY, V., G. SINGAL and T. BADAL. Coinnet: platform independent application to recognize Indian currency notes using deep learning techniques. *Multimedia Tools and Applications*. 2020, vol. 79, iss. 31, pp. 22569–22594. ISSN 1573-7721. DOI: 10.1007/s11042-020-09031-0.
- [18] EDUDODLA, B. R., V. S. VELDANDI and V. VEERAMSETTY. Detection of zero crossing points for a sinusoidal signal under the presence of harmonics and noise using decision trees. In: *International Conference on Research in Sciences, Engineering & Technology*. Warangal: 2022, pp. 1-21. DOI: 10.1063/5.0081953.
- [19] VEERAMSETTY, V., S. SRINIVASULA and S. R. SALKUTI. Non-Zero Crossing Point Detection in a Distorted Sinusoidal Signal Using Logistic Regression Model. *Computers*. 2022, vol. 11, iss. 6, pp. 1–19. ISSN 2073-431X. DOI: 10.3390/computers11060094.
- [20] XIAO, B., Y. LIU, X. LI, Z. DENG and Y. XUE. High-Resolution Multi-Channel Frequency Standard Comparator Using Digital Frequency Measurement. *Sensors*. 2021, vol. 21, iss. 16, pp. 1–13. ISSN 1424-8220. DOI: 10.3390/s21165626.
- [21] HO, T.-Y., C.-K. HUYNH, T.-H. LIN and S.-W. YANG. The Design and Implementation of a Sensorless Power Tool Based on a Microcontroller. *Electronics*. 2020, vol. 9, iss. 6, pp. 1–22. ISSN 2079-9292. DOI: 10.3390/electronics9060921.
- [22] TONG, C., M. WANG, B. ZHAO, Z. YIN and P. ZHENG. A Novel Sensorless Control Strategy for Brushless Direct Current Motor Based on the Estimation of Line Back Electro-Motive Force. *Energies*. 2017, vol. 10, iss. 9, pp. 1–20. ISSN 1996-1073. DOI: 10.3390/en10091384.
- [23] NEVLUD, P., M. BURES, L. KAPICAK and J. ZDRALEK. Anomaly-based Network Intrusion Detection Methods. *Advances in Electrical and Electronic Engineering*. 2013, vol. 11, iss. 6, pp. 468–474. ISSN 1804-3119. DOI: 10.15598/aeec.v11i6.877.
- [24] OUACHTOUK, I., S. EL HANI and K. DAHL. Intelligent Bearing Fault Diagnosis Method Based on HNR Envelope and Classification Using Supervised Machine Learning Algorithms. *Advances in Electrical and Electronic Engineering*. 2021, vol. 19, iss. 4, pp. 282–294. ISSN 1804-3119. DOI: 10.15598/aeec.v19i4.4183.
- [25] HAI, N. T., N. T. NGUYEN, M. H. NGUYEN and S. LIVATINO. Wavelet-Based Kernel Construction for Heart Disease Classification. *Advances in Electrical and Electronic Engineering*. 2019, vol. 17, iss. 3, pp. 306–319. ISSN 1804-3119. DOI: 10.15598/aeec.v17i3.3270.
- [26] MISAK, S., M. KRATKY and L. PROKOP. A Novel Method for Detection and Classification of Covered Conductor Faults. *Advances in Electrical and Electronic Engineering*. 2016, vol. 14, iss. 5, pp. 481–489. ISSN 1804-3119. DOI: 10.15598/aeec.v14i5.1733.
- [27] KHALEGHI, A., M. O. SADEGH, M. GHAZIZADEH-AHSAEE and A. M. RABORI. Transient Fault Area Location and Fault Classification for Distribution Systems Based on Wavelet Transform and Adaptive Neuro-Fuzzy Inference System (ANFIS). *Advances in Electrical and Electronic Engineering*. 2018, vol. 16, iss. 2, pp. 155–166. ISSN 1804-3119. DOI: 10.15598/aeec.v16i2.2563.
- [28] LOJEK, S. and A. OTCENASOVA. Classification of Voltage Dips in Distribution Networks. *Advances in Electrical and Electronic Engineering*. 2011, vol. 4, iss. 4, pp. 241–244. ISSN 1804-3119.
- [29] TRIWIYANTO, T., O. WAHYUNGGORO, H. A. NUGROHO and H. HERIANTO. An Investigation Into Time Domain Features of Surface Electromyography to Estimate the Elbow Joint Angle. *Advances in Electrical and Electronic Engineering*. 2017, vol. 15, iss. 3, pp. 448–458. ISSN 1804-3119. DOI: 10.15598/aeec.v15i3.2177.
- [30] SWEENEY, C., E. ENNIS, M. MULVENNA, R. BOND and S. O'NEILL. How Machine Learning Classification Accuracy Changes in a Happiness Dataset with Different Demographic Groups. *Computers*. 2022, vol. 11, iss. 5, pp. 1–15. ISSN 2073-431X. DOI: 10.3390/computers11050083.
- [31] EL MARHRAOUI, Y., H. AMROUN, M. BOUKALLEL, M. ANASTASSOVA, S. LAMY, S. BOUILLAND and M. AMMI. Foot-to-Ground Phases Detection: A Comparison of Data Representation Formatting Methods with Respect to Adaption of Deep Learning Architectures. *Computers*. 2022, vol. 11, iss. 5, pp. 1–26. ISSN 2073-431X. DOI: 10.3390/computers11050058.
- [32] KALAFI, E. Y., A. JODEIRI, S. K. SETAREHDAN, N. W. LIN, K. RAHMAT, N. A. TAIB, M. D. GANGGAYAH and S. K. DHILLON. Classification of Breast Cancer Lesions in Ultrasound

Images by Using Attention Layer and Loss Ensemble in Deep Convolutional Neural Networks. *Diagnostics*. 2021, vol. 11, iss. 10, pp. 1–11. ISSN 2075-4418. DOI: 10.3390/diagnostics11101859.

System Award (PPSA), India in 2013. He is a member of IEEE and IEEE Power and Energy Society.

## About Authors

**Venkataramana VEERAMSETTY** was born in Srikakulam, India, in 1985. He received his B.Tech. from Jawaharlal Nehru Technological University (JNTU), Hyderabad in 2006, and M.E. from university college of engineering, Osmania University in 2011, and Ph.D. from National Institute of Technology Waranagl in 2018. His research interests include power system deregulation and application of artificial intelligence in electrical engineering.

**Pravallika JADHAV** was born in Warangal, India, in 2001. She received her B.Tech. from SR University, Warangal in 2022. Her research interests include application of artificial intelligence in electrical engineering.

**Eslavath RAMESH** was born in Warangal, India, in 2001. He received her B.Tech. from SR University, Warangal in 2022. His research interests include application of artificial intelligence in electrical engineering.

**Srividya SRINIVASULA** was born in Warangal, India, in 2001. She received her B.Tech. from SR University, Warangal in 2022. Her research interests include application of artificial intelligence in electrical engineering.

**Surender Reddy SALKUTI** (corresponding author) received the Ph.D. degree in electrical engineering from Indian Institute of Technology, New Delhi, India, in 2013. He was a postdoctoral researcher at Howard University, Washington, D.C., USA, from 2013 to 2014. He is currently working as an associate professor in the Department of Railroad and Electrical Engineering, Woosong University, Daejeon, Republic of Korea. He published 120+ international journal papers, and 20+ international conference papers. His current research interests include power system restructuring issues, ancillary service pricing, real and reactive power pricing, congestion management, and market clearing, including renewable energy sources, demand response, smart grid development with integration of wind and solar photovoltaic energy sources, battery storage and electric vehicles, artificial intelligence applications in power systems, and power system analysis and optimization. He received Distinguished Researcher Award from Woosong University Educational Foundation, Republic of Korea in 2016, and POSOCO Power

## Appendix

### Appendix A Conversion of Continuous Features into Categorical Features

This section demonstrates step by step procedure that is used to convert continuous features into categorical features.

- Read data that is shown Tab. 13.

**Tab. 13:** Few data samples from the dataset that prepared from noisy distorted signal.

$m$	$c$	$R$	$E$	Class
-448	4.4	0.96	0.027	0
-313	3.1	0.98	0.010	0
-331	3.3	0.99	0.009	0
-299	3.0	0.99	0.009	1
-269	2.7	1.00	0.009	1
-301	3.0	1.00	0.003	1
-293	2.9	1.00	0.004	1
-319	3.2	1.00	0.003	1
-303	3.0	0.99	0.005	0
-324	3.2	0.99	0.007	0

- Identify the unique threshold values for each input feature i.e.  $m$  (Slope),  $c$  (Intercept),  $R$  (Correlation) and  $E$  ( $RMSE$ ) as shown below.
  - Unique threshold values for  $m$ :  $\geq -448, \geq -313, \geq -331, \geq -269, \geq -299, \geq -301, \geq -293, \geq -319, \geq -303, \geq -324$ .
  - Unique threshold values for  $c$ :  $\geq 4.4, \geq 3.0, \geq 2.7, \geq 3.1, \geq 3.2, \geq 2.9$ .
  - Unique threshold values for  $R$ :  $\geq 0.96, \geq 0.98, \geq 0.99, \geq 1$ .
  - Unique threshold values for  $E$ :  $\geq 0.027, \geq 0.01, \geq 0.009, \geq 0.004, \geq 0.003, \geq 0.005, \geq 0.007$ .
- Now identify the best threshold value for each feature to convert continuous data into categorical based on Information Gain (IG) as shown below.

**1.1. Identification of Best Threshold for  $m$**

**1) Calculate IG for Threshold  $\geq -448$  Using Tab. 14**

**Tab. 14:** Sub Table:  $m$  Vs. class with threshold  $\geq -448$ .

$m$	Class	$m$	Class
$\geq -448$	0	$< -448$	1
$< -448$	0	$< -448$	1
$< -448$	0	$< -448$	1
$< -448$	1	$< -448$	0
$< -448$	1	$< -448$	0

- Calculate probability for class label.

Parameter	1	0
Count	5	5
Probability	$P(1) = \frac{5}{10}$	$P(0) = \frac{5}{10}$

- Calculate entropy for class label.

$$E(class) = E(1, 0) = -P(1) \log_2 P(1) - P(0) \log_2 P(0). \tag{6}$$

$$E(class) = E(1, 0) = -\frac{5}{10} \log_2 \frac{5}{10} - \frac{5}{10} \log_2 \frac{5}{10} = 1. \tag{7}$$

- Calculate probability with respect to feature  $m$  threshold  $\geq -448$ .

Threshold	1	0	Probability
$\geq -448$	5	5	$\frac{10}{10} = 1$
$< -448$	0	0	0

- Calculate entropy with respect to feature  $m$  threshold  $\geq -448$ .

$$E(m \geq -448) = E(5, 5) = -\frac{5}{10} \log_2 \frac{5}{10} - \frac{5}{10} \log_2 \frac{5}{10} = 1. \tag{8}$$

$$E(m < -448) = E(0, 0) = 0. \tag{9}$$

- Calculate entropy of feature with respect to class.

$$E(class, m) = P(m \geq -448) E(m \geq -448) + P(m < -448) E(m < -448) = (1)(1) + (0)(0) = 1. \tag{10}$$

- Calculate information gain of feature  $m \geq -448$ .

$$IG(m \geq -448) = E(class) - E(class, m) = 1 - 1 = 0. \tag{11}$$

**2) Calculate IG for Threshold  $\geq -313$  Using Tab. 15**

**Tab. 15:** Sub Table:  $m$  Vs. class with threshold  $\geq -313$ .

$m$	Class	$m$	Class
$< -313$	0	$\geq -313$	1
$\geq -313$	0	$\geq -313$	1
$< -313$	0	$< -313$	1
$\geq -313$	1	$\geq -313$	0
$\geq -313$	1	$< -313$	0

- Calculate probability with respect to feature  $m$  threshold  $\geq -313$ .

Threshold	1	0	Probability
$\geq -313$	4	2	$\frac{6}{10} = 0.6$
$< -313$	1	3	$\frac{4}{10} = 0.4$

- Calculate entropy with respect to feature  $m$  threshold  $\geq -313$ .

$$E(m \geq -313) = E(4, 2) = -\frac{4}{6} \log_2 \frac{4}{6} - \frac{2}{6} \log_2 \frac{2}{6} = 0.723. \tag{12}$$

$$E(m < -313) = E(1, 3) = -\frac{1}{4} \log_2 \frac{1}{4} - \frac{3}{4} \log_2 \frac{3}{4} = 0.811. \tag{13}$$

- Calculate entropy of feature with respect to class.

$$E(class, m) = P(m \geq -313) E(m \geq -313) + P(m < -313) E(m < -313) = (0.6)(0.723) + (0.4)(0.811) = 0.7582. \tag{14}$$

- Calculate information gain of feature  $m \geq -313$ .

$$IG(m \geq -313) = E(class) - E(class, m) = 1 - 0.7582 = 0.2418. \tag{15}$$

**3) Calculate IG for Threshold  $\geq -331$  Using Tab. 16**

**Tab. 16:** Sub Table:  $m$  Vs. class with threshold  $\geq -331$ .

$m$	Class	$m$	Class
$< -331$	0	$\geq -331$	1
$\geq -331$	0	$\geq -331$	1
$\geq -331$	0	$\geq -331$	1
$\geq -331$	1	$\geq -331$	0
$\geq -331$	1	$\geq -331$	0



**Tab. 17:** Probability for each split:  $m \geq -313$  and  $m < -313$ .

Threshold	1	0	Probability
$\geq -331$	5	4	$\frac{9}{10}=0.9$
$< -331$	0	1	$\frac{1}{10}=0.1$

- Calculate probability with respect to feature  $m$  threshold  $\geq -331$  and presented in Tab. 17.

$$E(m \geq -331) = E(5, 4) = -\frac{5}{9} \log_2 \frac{5}{9} - \frac{4}{9} \log_2 \frac{4}{9} = 0.991. \tag{16}$$

$$E(m < -331) = E(0, 1) = -\frac{0}{1} \log_2 \frac{0}{1} - \frac{1}{1} \log_2 \frac{1}{1} = 0. \tag{17}$$

- Calculate entropy of feature with respect to class.

$$E(class, m) = P(m \geq -331) E(m \geq -331) + P(m < -331) E(m < -331) = (0.9)(0.991) + (0.1)(0) = 0.8919. \tag{18}$$

- Calculate information gain of feature  $m \geq -331$ .

$$IG(m \geq -331) = E(class) - E(class, m) = 1 - 0.8919 = 0.1081. \tag{19}$$

**4) Calculate IG for Threshold  $m \geq -269$  Using Tab. 18**

**Tab. 18:** Sub Table:  $m$  Vs. class with threshold  $\geq -269$ .

$m$	Class	$m$	Class
$< -269$	0	$< -269$	1
$< -269$	0	$< -269$	1
$< -269$	0	$< -269$	1
$< -269$	1	$< -269$	0
$\geq -269$	1	$< -269$	0

- Calculate probability with respect to feature  $m$  threshold  $\geq -269$  and presented in Tab. 19.

**Tab. 19:** Probability for each split:  $m \geq -269$  and  $m < -269$ .

Threshold	1	0	Probability
$\geq -331$	1	0	$\frac{1}{10}=0.1$
$< -331$	4	5	$\frac{9}{10}=0.9$

- Calculate entropy with respect to feature  $m$  threshold  $\geq -269$ .

$$E(m \geq -269) = E(1, 0) = -\frac{1}{1} \log_2 \frac{1}{1} - \frac{0}{1} \log_2 \frac{0}{1} = 0. \tag{20}$$

$$E(m < -269) = E(4, 5) = -\frac{4}{9} \log_2 \frac{4}{9} - \frac{5}{9} \log_2 \frac{5}{9} = 0.991. \tag{21}$$

- Calculate entropy of feature with respect to class.

$$E(class, m) = P(m \geq -269) E(m \geq -269) + P(m < -269) E(m < -269) = (0.1)(0) + (0.9)(0.991) = 0.8919. \tag{22}$$

- Calculate information gain of feature  $m \geq -269$ .

$$IG(m \geq -269) = E(class) - E(class, m) = 1 - 0.8919 = 0.1081. \tag{23}$$

**5) Calculate IG for Threshold  $m \geq -299$  Using Tab. 20**

**Tab. 20:** Sub Table:  $m$  Vs. class with threshold  $\geq -299$ .

$m$	Class	$m$	Class
$< -299$	0	$< -299$	1
$< -299$	0	$\geq -299$	1
$< -299$	0	$< -299$	1
$\geq -299$	1	$< -299$	0
$\geq -299$	1	$< -299$	0

- Calculate probability with respect to feature  $m$  threshold  $\geq -299$  and presented in Tab. 21.

**Tab. 21:** Probability for each split:  $m \geq -299$  and  $m < -299$ .

Threshold	1	0	Probability
$\geq -299$	3	0	$\frac{3}{10}=0.3$
$< -299$	2	5	$\frac{7}{10}=0.7$

- Calculate entropy with respect to feature  $m$  threshold  $\geq -299$ .

$$E(m \geq -299) = E(3, 0) = -\frac{3}{3} \log_2 \frac{3}{3} - \frac{0}{3} \log_2 \frac{0}{3} = 0. \tag{24}$$

$$E(m < -299) = E(2, 5) = -\frac{2}{7} \log_2 \frac{2}{7} - \frac{5}{7} \log_2 \frac{5}{7} = 0.8631. \tag{25}$$

- Calculate entropy of feature with respect to class.

$$E(class, m) = P(m \geq -299) E(m \geq -299) + P(m < -299) E(m < -299) = (0.3)(0) + (0.7)(0.8631) = 0.6042. \tag{26}$$

- Calculate information gain of feature  $m \geq -299$ .

$$IG(m \geq -299) = E(class) - E(class, m) = 1 - 0.6042 = 0.3958. \tag{27}$$

**6) Calculate IG for Threshold  $m \geq -301$  Using Tab. 22**

**Tab. 22:** Sub Table:  $m$  Vs. class with threshold  $\geq -301$ .

$m$	Class	$m$	Class
$< -301$	0	$\geq -301$	1
$< -301$	0	$\geq -301$	1
$< -301$	0	$< -301$	1
$\geq -301$	1	$< -301$	0
$\geq -301$	1	$< -301$	0

- Calculate probability with respect to feature  $m$  threshold  $\geq -301$  and presented in Tab. 23.

**Tab. 23:** Probability for each split:  $m \geq -301$  and  $m < -301$ .

Threshold	1	0	Probability
$\geq -301$	4	1	$\frac{5}{10}=0.5$
$< -301$	1	4	$\frac{5}{10}=0.5$

- Calculate entropy with respect to feature  $m$  threshold  $\geq -301$ .

$$E(m \geq -301) = E(4, 1) = -\frac{4}{5} \log_2 \frac{4}{5} - \frac{1}{5} \log_2 \frac{1}{5} = 0.7212. \tag{28}$$

$$E(m < -301) = E(1, 4) = -\frac{1}{5} \log_2 \frac{1}{5} - \frac{4}{5} \log_2 \frac{4}{5} = 0.7212. \tag{29}$$

- Calculate entropy of feature with respect to class.

$$E(class, m) = P(m \geq -301) E(m \geq -301) + P(m < -301) E(m < -301) = (0.5)(0.7212) + (0.5)(0.7212) = 0.7212. \tag{30}$$

- Calculate information gain of feature  $m \geq -301$ .

$$IG(m \geq -301) = E(class) - E(class, m) = 1 - 0.7212 = 0.2788. \tag{31}$$

**7) Calculate IG for Threshold  $m \geq -293$  Using Tab. 24**

**Tab. 24:** Sub Table:  $m$  Vs. class with threshold  $\geq -293$ .

$m$	Class	$m$	Class
$< -293$	0	$< -293$	1
$< -293$	0	$\geq -293$	1
$< -293$	0	$< -293$	1
$< -293$	1	$< -293$	0
$\geq -293$	1	$< -293$	0

- Calculate probability with respect to feature  $m$  threshold  $\geq -293$  and presented in Tab. 25.

**Tab. 25:** Probability for each split:  $m \geq -293$  and  $m < -293$ .

Threshold	1	0	Probability
$\geq -293$	2	0	$\frac{2}{10}=0.2$
$< -293$	3	5	$\frac{8}{10}=0.8$

- Calculate entropy with respect to feature  $m$  threshold  $\geq -293$ .

$$E(m \geq -293) = E(2, 0) = -\frac{2}{2} \log_2 \frac{2}{2} - \frac{0}{2} \log_2 \frac{0}{2} = 0. \tag{32}$$

$$E(m < -293) = E(3, 5) = -\frac{3}{8} \log_2 \frac{3}{8} - \frac{5}{8} \log_2 \frac{5}{8} = 0.9544. \tag{33}$$

- Calculate entropy of feature with respect to class.

$$E(class, m) = P(m \geq -293) E(m \geq -293) + P(m < -293) E(m < -293) = (0.2)(0) + (0.8)(0.9544) = 0.7635. \tag{34}$$

- Calculate information gain of feature  $m \geq -293$ .

$$IG(m \geq -293) = E(class) - E(class, m) = 1 - 0.7635 = 0.2364. \tag{35}$$

**8) Calculate IG for Threshold  $m \geq -319$  Using Tab. 26**

- Calculate probability with respect to feature  $m$  threshold  $\geq -319$  and presented in Tab. 27.

- Calculate entropy with respect to feature  $m$  threshold  $\geq -319$ .

**Tab. 26:** Sub Table:  $m$  Vs. class with threshold  $\geq -319$ .

$m$	Class	$m$	Class
$< -319$	0	$\geq -319$	1
$\geq -319$	0	$\geq -319$	1
$< -319$	0	$\geq -319$	1
$\geq -319$	1	$\geq -319$	0
$\geq -319$	1	$< -319$	0

**Tab. 27:** Probability for each split:  $m \geq -319$  and  $m < -319$ .

Threshold	1	0	Probability
$\geq -319$	5	2	$\frac{7}{10}=0.7$
$< -319$	0	3	$\frac{3}{10}=0.3$

$$E(m \geq -319) = E(5, 2) = -\frac{5}{7} \log_2 \frac{5}{7} - \frac{2}{7} \log_2 \frac{2}{7} = 0.8631. \tag{36}$$

$$E(m < -319) = E(0, 3) = -\frac{0}{3} \log_2 \frac{0}{3} - \frac{3}{3} \log_2 \frac{3}{3} = 0. \tag{37}$$

- Calculate entropy of feature with respect to class.

$$E(class, m) = P(m \geq -319) E(m \geq -319) + P(m < -319) E(m < -319) = (0.7)(0.8631) + (0.3)(0) = 0.6042. \tag{38}$$

- Calculate information gain of feature  $m \geq -319$ .

$$IG(m \geq -319) = E(class) - E(class, m) = 1 - 0.6042 = 0.3958. \tag{39}$$

**9) Calculate IG for Threshold  $m \geq -303$  Using Tab. 28**

**Tab. 28:** Sub Table:  $m$  Vs. class with threshold  $\geq -303$ .

$m$	Class	$m$	Class
$< -303$	0	$\geq -303$	1
$< -303$	0	$\geq -303$	1
$< -303$	0	$< -303$	1
$\geq -303$	1	$\geq -303$	0
$\geq -303$	1	$< -303$	0

- Calculate probability with respect to feature  $m$  threshold  $\geq -303$  and presented in Tab. 29.
- Calculate entropy with respect to feature  $m$  threshold  $\geq -303$ .

$$E(m \geq -303) = E(4, 1) = -\frac{4}{5} \log_2 \frac{4}{5} - \frac{1}{5} \log_2 \frac{1}{5} = 0.7212. \tag{40}$$

**Tab. 29:** Probability for each split:  $m \geq -303$  and  $m < -303$ .

Threshold	1	0	Probability
$\geq -303$	4	1	$\frac{5}{10}=0.5$
$< -303$	1	4	$\frac{5}{10}=0.5$

$$E(m < -303) = E(1, 4) = -\frac{1}{5} \log_2 \frac{1}{5} - \frac{4}{5} \log_2 \frac{4}{5} = 0.7212. \tag{41}$$

- Calculate entropy of feature with respect to class.

$$E(class, m) = P(m \geq -303) E(m \geq -303) + P(m < -303) E(m < -303) = (0.5)(0.7212) + (0.5)(0.7212) = 0.7212. \tag{42}$$

- Calculate information gain of feature  $m \geq -303$ .

$$IG(m \geq -303) = E(class) - E(class, m) = 1 - 0.7212 = 0.2788. \tag{43}$$

**10) Calculate IG for Threshold  $m \geq -324$  Using Tab. 30**

**Tab. 30:** Sub Table:  $m$  Vs. class with threshold  $\geq -324$ .

$m$	Class	$m$	Class
$< -324$	0	$\geq -324$	1
$\geq -324$	0	$\geq -324$	1
$< -324$	0	$\geq -324$	1
$\geq -324$	1	$\geq -324$	0
$\geq -324$	1	$< -324$	0

- Calculate probability with respect to feature  $m$  threshold  $\geq -324$  and presented in Tab. 31.

**Tab. 31:** Probability for each split:  $m \geq -324$  and  $m < -324$ .

Threshold	1	0	Probability
$\geq -324$	5	3	$\frac{8}{10}=0.8$
$< -324$	0	2	$\frac{2}{10}=0.2$

- Calculate entropy with respect to feature  $m$  threshold  $\geq -324$ .

$$E(m \geq -324) = E(5, 3) = -\frac{5}{8} \log_2 \frac{5}{8} - \frac{3}{8} \log_2 \frac{3}{8} = 0.9544. \tag{44}$$

$$E(m < -324) = E(0, 2) = -\frac{0}{2} \log_2 \frac{0}{2} - \frac{2}{2} \log_2 \frac{2}{2} = 0. \tag{45}$$

- Calculate entropy of feature with respect to class.

$$E(class, m) = P(m \geq -324) E(m \geq -324) + P(m < -324) E(m < -324) = (0.8)(0.9544) + (0.2)(0) = 0.7635. \tag{46}$$

- Calculate information gain of feature  $m \geq -324$ .

$$IG(m \geq -324) = E(class) - E(class, m) = 1 - 0.7635 = 0.2365. \tag{47}$$

- Information Gain of each split of feature  $m$  is presented in Tab. 32. From the Tab. 32, it is observed that split  $\geq -299$  and  $\geq -319$  have highest information gain value i.e 0.3958. Hence,  $\geq -299$  is considered as a best split.

Tab. 32: Information Gain of each split of feature  $m$ .

Split	IG	Split	IG
$\geq -448$	0	$\geq -301$	0.2788
$\geq -313$	0.2418	$\geq -293$	0.2364
$\geq -331$	0.1081	$\geq -319$	<b>0.3958</b>
$\geq -269$	0.1081	$\geq -303$	0.2788
$\geq -299$	<b>0.3958</b>	$\geq -324$	0.2365

## 1.2. Identification of Best Threshold for $c$ i.e. Intercept Feature

### 1) Calculate IG for Threshold $\geq 2.9$ Using Tab. 33

Tab. 33: Sub Table:  $c$  Vs. class with threshold  $\geq 2.9$ .

$c$	Class	$c$	Class
$\geq 2.9$	0	$\geq 2.9$	1
$\geq 2.9$	0	$\geq 2.9$	1
$\geq 2.9$	0	$\geq 2.9$	1
$\geq 2.9$	1	$\geq 2.9$	0
$< 2.9$	1	$\geq 2.9$	0

- Calculate probability with respect to feature  $c$  threshold  $\geq 2.9$  as shown in Tab. 34.

Tab. 34: Probability for each split:  $c \geq 2.9$  and  $c < 2.9$ .

Threshold	1	0	Probability
$\geq 2.9$	4	5	$\frac{9}{10}=0.9$
$< 2.9$	1	0	$\frac{1}{10}=0.1$

- Calculate entropy with respect to feature  $c$  threshold  $\geq 2.9$ .

$$E(c \geq 2.9) = E(4, 5) = -\frac{4}{9} \log_2 \frac{4}{9} - \frac{5}{9} \log_2 \frac{5}{9} = 0.991. \tag{48}$$

$$E(c < 2.9) = E(1, 0) = -\frac{1}{1} \log_2 \frac{1}{1} - \frac{0}{1} \log_2 \frac{0}{1} = 0. \tag{49}$$

- Calculate entropy of feature with respect to class.

$$E(class, c) = P(c \geq 2.9) E(c \geq 2.9) + P(c < 2.9) E(c < 2.9) = (0.9)(0.991) + (0.1)(0) = 0.8919. \tag{50}$$

- Calculate information gain of feature  $c \geq 2.9$ .

$$IG(c \geq 2.9) = E(class) - E(class, c) = 1 - 0.8919 = 0.1081. \tag{51}$$

### 2) Calculate IG for Threshold $\geq 3.2$ Using Tab. 35

Tab. 35: Sub Table:  $c$  Vs. class with threshold  $\geq 3.2$ .

$c$	Class	$c$	Class
$\geq 3.2$	0	$< 3.2$	1
$< 3.2$	0	$< 3.2$	1
$\geq 3.2$	0	$\geq 3.2$	1
$< 3.2$	1	$< 3.2$	0
$< 3.2$	1	$\geq 3.2$	0

- Calculate probability with respect to feature  $c$  threshold  $\geq 3.2$  as shown in Tab. 36.

Tab. 36: Probability for each split:  $c \geq 3.2$  and  $c < 3.2$ .

Threshold	1	0	Probability
$\geq 3.2$	1	3	$\frac{4}{10}=0.4$
$< 3.2$	4	2	$\frac{6}{10}=0.6$

- Calculate entropy with respect to feature  $c$  threshold  $\geq 3.2$ .

$$E(c \geq 3.2) = E(1, 3) = -\frac{1}{4} \log_2 \frac{1}{4} - \frac{3}{4} \log_2 \frac{3}{4} = 0.811. \tag{52}$$

$$E(c < 3.2) = E(4, 2) = -\frac{4}{6} \log_2 \frac{4}{6} - \frac{2}{6} \log_2 \frac{2}{6} = 0.9183. \tag{53}$$

- Calculate entropy of feature with respect to class.

$$\begin{aligned}
 E(class, c) &= P(c \geq 3.2) E(c \geq 3.2) + \\
 &+ P(c < 3.2) E(c < 3.2) = \tag{54} \\
 &= (0.4)(0.811) + (0.6)(0.9183) = 0.8753.
 \end{aligned}$$

- Calculate information gain of feature  $c \geq 3.2$ .

$$\begin{aligned}
 IG(c \geq 3.2) &= E(class) - E(class, c) = \tag{55} \\
 &= 1 - 0.8753 = 0.1247.
 \end{aligned}$$

**3) Calculate IG for Threshold  $\geq 3.1$  Using Tab. 37**

**Tab. 37:** Sub Table:  $c$  Vs. class with threshold  $\geq 3.1$ .

$c$	Class	$c$	Class
$\geq 3.1$	0	$< 3.1$	1
$\geq 3.1$	0	$< 3.1$	1
$\geq 3.1$	0	$\geq 3.1$	1
$< 3.1$	1	$< 3.1$	0
$< 3.1$	1	$\geq 3.1$	0

- Calculate probability with respect to feature  $c$  threshold  $\geq 3.1$  as shown in Tab. 38.

**Tab. 38:** Probability for each split:  $c \geq 3.1$  and  $c < 3.1$ .

Threshold	1	0	Probability
$\geq 3.1$	1	4	$\frac{5}{10}=0.5$
$< 3.1$	4	1	$\frac{5}{10}=0.5$

- Calculate entropy with respect to feature  $c$  threshold  $\geq 3.1$ .

$$\begin{aligned}
 E(c \geq 3.1) &= E(1, 4) = \tag{56} \\
 &= -\frac{4}{5} \log_2 \frac{4}{5} - \frac{1}{5} \log_2 \frac{1}{5} = 0.7212.
 \end{aligned}$$

$$\begin{aligned}
 E(c < 3.1) &= E(4, 1) = \tag{57} \\
 &= -\frac{1}{5} \log_2 \frac{1}{5} - \frac{4}{5} \log_2 \frac{4}{5} = 0.7212.
 \end{aligned}$$

- Calculate entropy of feature with respect to class.

$$\begin{aligned}
 E(class, c) &= P(c \geq 3.1) E(c \geq 3.1) + \\
 &+ P(c < 3.1) E(c < 3.1) = \tag{58} \\
 &= (0.5)(0.7212) + (0.5)(0.7212) = 0.7212.
 \end{aligned}$$

- Calculate information gain of feature  $c \geq 3.1$ .

$$\begin{aligned}
 IG(c \geq 3.1) &= E(class) - E(class, c) = \tag{59} \\
 &= 1 - 0.7212 = 0.2788.
 \end{aligned}$$

**Tab. 39:** Sub Table:  $c$  Vs. class with threshold  $\geq 3.3$ .

$c$	Class	$c$	Class
$\geq 3.3$	0	$< 3.3$	1
$< 3.3$	0	$< 3.3$	1
$\geq 3.3$	0	$< 3.3$	1
$< 3.3$	1	$< 3.3$	0
$< 3.3$	1	$< 3.3$	0

**4) Calculate IG for Threshold  $\geq 3.3$  Using Tab. 39**

- Calculate probability with respect to feature  $c$  threshold  $\geq 3.3$  as shown in Tab. 40.

**Tab. 40:** Probability for each split:  $c \geq 3.3$  and  $c < 3.3$ .

Threshold	1	0	Probability
$\geq 3.3$	5	2	$\frac{7}{10}=0.7$
$< 3.3$	0	3	$\frac{3}{10}=0.3$

- Calculate entropy with respect to feature  $c$  threshold  $\geq 3.3$ .

$$\begin{aligned}
 E(c \geq 3.3) &= E(5, 2) = \tag{60} \\
 &= -\frac{5}{7} \log_2 \frac{5}{7} - \frac{2}{7} \log_2 \frac{2}{7} = 0.8631.
 \end{aligned}$$

$$\begin{aligned}
 E(c < 3.3) &= E(0, 3) = \tag{61} \\
 &= -\frac{0}{3} \log_2 \frac{0}{3} - \frac{3}{3} \log_2 \frac{3}{3} = 0.
 \end{aligned}$$

- Calculate entropy of feature with respect to class.

$$\begin{aligned}
 E(class, c) &= P(c \geq 3.3) E(c \geq 3.3) + \\
 &+ P(c < 3.3) E(c < 3.3) = \tag{62} \\
 &= (0.7)(0.8631) + (0.3)(0) = 0.6041.
 \end{aligned}$$

- Calculate information gain of feature  $c \geq 3.3$ .

$$\begin{aligned}
 IG(c \geq 3.3) &= E(class) - E(class, c) = \tag{63} \\
 &= 1 - 0.6041 = 0.3959.
 \end{aligned}$$

**5) Calculate IG for hreshold  $\geq 3.0$  Using Tab. 41**

**Tab. 41:** Sub Table:  $c$  Vs. class with threshold  $\geq 3.0$ .

$c$	Class	$c$	Class
$\geq 3.0$	0	$\geq 3.0$	1
$\geq 3.0$	0	$< 3.0$	1
$\geq 3.0$	0	$\geq 3.0$	1
$\geq 3.0$	1	$\geq 3.0$	0
$< 3.0$	1	$\geq 3.0$	0

**Tab. 42:** Probability for each split:  $c \geq 3.0$  and  $c < 3.0$ .

Threshold	1	0	Probability
$\geq 3.0$	3	5	$\frac{8}{10}=0.8$
$< 3.0$	2	0	$\frac{2}{10}=0.2$

- Calculate probability with respect to feature  $c$  threshold  $\geq 3.0$  as shown in Tab. 42.
- Calculate entropy with respect to feature  $c$  threshold  $\geq 3.0$ .

$$E(c \geq 3.0) = E(3, 5) = -\frac{3}{8} \log_2 \frac{3}{8} - \frac{5}{8} \log_2 \frac{5}{8} = 0.9546. \tag{64}$$

$$E(c < 3.0) = E(2, 0) = -\frac{2}{2} \log_2 \frac{2}{2} - \frac{0}{2} \log_2 \frac{0}{2} = 0. \tag{65}$$

- Calculate entropy of feature with respect to class.

$$E(class, c) = P(c \geq 3.0) E(c \geq 3.0) + P(c < 3.0) E(c < 3.0) = (0.8)(0.9546) + (0.2)(0) = 0.7636. \tag{66}$$

- Calculate information gain of feature  $c \geq 3.0$ .

$$IG(c \geq 3.0) = E(class) - E(class, c) = 1 - 0.7636 = 0.2364. \tag{67}$$

**6) Calculate IG for Threshold  $\geq 2.7$  Using Tab. 43**

**Tab. 43:** Sub Table:  $c$  Vs. class with threshold  $\geq 2.7$ .

$c$	Class	$c$	Class
$\geq 2.7$	0	$\geq 2.7$	1
$\geq 2.7$	0	$\geq 2.7$	1
$\geq 2.7$	0	$\geq 2.7$	1
$\geq 2.7$	1	$\geq 2.7$	0
$\geq 2.7$	1	$\geq 2.7$	0

- Calculate probability with respect to feature  $c$  threshold  $\geq 2.7$  as shown in Tab. 44.

**Tab. 44:** Probability for each split:  $c \geq 2.7$  and  $c < 2.7$ .

Threshold	1	0	Probability
$\geq 2.7$	5	5	$\frac{10}{10}=1$
$< 2.7$	0	0	$\frac{0}{10}=0$

- Calculate entropy with respect to feature  $c$  threshold  $\geq 2.7$ .

$$E(c \geq 2.7) = E(5, 5) = -\frac{5}{10} \log_2 \frac{5}{10} - \frac{5}{10} \log_2 \frac{5}{10} = 1. \tag{68}$$

$$E(c < 3.0) = E(0, 0) = 0. \tag{69}$$

- Calculate entropy of feature with respect to class.

$$E(class, c) = P(c \geq 2.7) E(c \geq 2.7) + P(c < 2.7) E(c < 2.7) = (1)(1) + (0)(0) = 1. \tag{70}$$

- Calculate information gain of feature  $c \geq 2.7$ .

$$IG(c \geq 2.7) = E(class) - E(class, c) = 1 - 1 = 0. \tag{71}$$

**7) Calculate IG for Threshold  $\geq 4.4$  Using Tab. 45**

**Tab. 45:** Sub Table:  $c$  Vs. class with threshold  $\geq 4.4$ .

$c$	Class	$c$	Class
$\geq 4.4$	0	$< 4.4$	1
$< 4.4$	0	$< 4.4$	1
$< 4.4$	0	$< 4.4$	1
$< 4.4$	1	$< 4.4$	0
$< 4.4$	1	$< 4.4$	0

- Calculate probability with respect to feature  $c$  threshold  $\geq 4.4$  as shown in Tab. 46.

**Tab. 46:** Probability for each split:  $c \geq 4.4$  and  $c < 4.4$ .

Threshold	1	0	Probability
$\geq 4.4$	0	1	$\frac{1}{10}=0.1$
$< 4.4$	5	4	$\frac{9}{10}=0.9$

- Calculate entropy with respect to feature  $c$  threshold  $\geq 4.4$ .

$$E(c \geq 4.4) = E(0, 1) = -\frac{0}{1} \log_2 \frac{0}{1} - \frac{1}{1} \log_2 \frac{1}{1} = 0. \tag{72}$$

$$E(c < 4.4) = E(5, 4) = -\frac{5}{9} \log_2 \frac{5}{9} - \frac{4}{9} \log_2 \frac{4}{9} = 0.9911. \tag{73}$$

- Calculate entropy of feature with respect to class.

$$\begin{aligned}
 E(class, c) &= P(c \geq 4.4) E(c \geq 4.4) + \\
 &+ P(c < 4.4) E(c < 4.4) = \tag{74} \\
 &= (0.1)(0) + (0.9)(0.9911) = 0.8919.
 \end{aligned}$$

- Calculate information gain of feature  $c \geq 4.4$ .

$$\begin{aligned}
 IG(c \geq 4.4) &= E(class) - E(class, c) = \tag{75} \\
 &= 1 - 0.8919 = 0.1081.
 \end{aligned}$$

- Information Gain of each split of feature  $c$  is presented in Tab. 47. From the Tab. 47, it is observed that split  $\geq 3.3$  gives highest information gain value i.e 0.3959. Hence,  $\geq 3.3$  is considered as a best split.

Tab. 47: Information Gain of each split of feature  $c$ .

Split	IG	Split	IG
$\geq 4.4$	0.1081	$\geq 2.9$	0.1081
$\geq 3.0$	0.2364	$\geq 3.3$	<b>0.3959</b>
$\geq 2.7$	0	$\geq 3.1$	0.2788
$\geq 3.2$	0.1247		

### 1.3. Identification of Best Threshold for $R$ i.e. Correlation Feature

#### 1) Calculate IG for Threshold $\geq 0.96$ Using Tab. 48

Tab. 48: Sub Table:  $R$  Vs. class with threshold  $\geq 0.96$ .

$R$	Class	$R$	Class
$\geq 0.96$	0	$\geq 0.96$	1
$\geq 0.96$	0	$\geq 0.96$	1
$\geq 0.96$	0	$\geq 0.96$	1
$\geq 0.96$	1	$\geq 0.96$	0
$\geq 0.96$	1	$\geq 0.96$	0

- Calculate probability with respect to feature  $R$  threshold  $\geq 0.96$  as shown in Tab. 49.

Tab. 49: Probability for each split:  $R \geq 0.96$  and  $R < 0.96$ .

Threshold	1	0	Probability
$\geq 0.96$	5	5	$\frac{10}{10}=1$
$< 0.96$	0	0	$\frac{0}{10}=0$

- Calculate entropy with respect to feature  $R$  threshold  $\geq 0.96$ .

$$\begin{aligned}
 E(R \geq 0.96) &= E(5, 5) = \tag{76} \\
 &= -\frac{5}{10} \log_2 \frac{5}{10} - \frac{5}{10} \log_2 \frac{5}{10} = 1.
 \end{aligned}$$

$$E(R < 0.96) = E(0, 0) = 0. \tag{77}$$

- Calculate entropy of feature with respect to class.

$$\begin{aligned}
 E(class, R) &= P(R \geq 0.96) E(R \geq 0.96) + \\
 &+ P(R < 0.96) E(R < 0.96) = \tag{78} \\
 &= (1)(1) + (0)(0) = 1.
 \end{aligned}$$

- Calculate information gain of feature  $R \geq 0.96$ .

$$\begin{aligned}
 IG(R \geq 0.96) &= E(class) - E(class, R) \tag{79} \\
 &= 1 - 1 = 0.
 \end{aligned}$$

#### 2) Calculate IG for Threshold $\geq 0.98$ Using Tab. 50

Tab. 50: Sub Table:  $R$  Vs. class with threshold  $\geq 0.98$ .

$R$	Class	$R$	Class
$< 0.98$	0	$\geq 0.98$	1
$\geq 0.98$	0	$\geq 0.98$	1
$\geq 0.98$	0	$\geq 0.98$	1
$\geq 0.98$	1	$\geq 0.98$	0
$\geq 0.98$	1	$\geq 0.98$	0

- Calculate probability with respect to feature  $R$  threshold  $\geq 0.98$  as shown in Tab. 51.

Tab. 51: Probability for each split:  $R \geq 0.98$  and  $R < 0.98$ .

Threshold	1	0	Probability
$\geq 0.98$	5	4	$\frac{9}{10}=0.9$
$< 0.98$	0	1	$\frac{1}{10}=0.1$

- Calculate entropy with respect to feature  $R$  threshold  $\geq 0.98$ .

$$\begin{aligned}
 E(R \geq 0.98) &= E(5, 4) = \tag{80} \\
 &= -\frac{5}{9} \log_2 \frac{5}{9} - \frac{4}{9} \log_2 \frac{4}{9} = 0.9911.
 \end{aligned}$$

$$\begin{aligned}
 E(R < 0.98) &= E(0, 1) = \tag{81} \\
 &= -\frac{0}{1} \log_2 \frac{0}{1} - \frac{1}{1} \log_2 \frac{1}{1} = 0.
 \end{aligned}$$

- Calculate entropy of feature with respect to class.

$$\begin{aligned}
 E(class, R) &= P(R \geq 0.98) E(R \geq 0.98) + \\
 &+ P(R < 0.98) E(R < 0.98) = \tag{82} \\
 &= (0.9)(0.9911) + (0.1)(0) = 0.892.
 \end{aligned}$$

- Calculate information gain of feature  $R \geq 0.98$ .

$$\begin{aligned}
 IG(R \geq 0.98) &= E(class) - E(class, R) = \tag{83} \\
 &= 1 - 0.892 = 0.108.
 \end{aligned}$$

3) Calculate IG for Threshold  $\geq 0.99$  Using Tab. 52

Tab. 52: Sub Table:  $R$  Vs. class with threshold  $\geq 0.99$ .

$R$	Class	$R$	Class
$< 0.99$	0	$\geq 0.99$	1
$< 0.99$	0	$\geq 0.99$	1
$\geq 0.99$	0	$\geq 0.99$	1
$\geq 0.99$	1	$\geq 0.99$	0
$\geq 0.99$	1	$\geq 0.99$	0

- Calculate probability with respect to feature  $R$  threshold  $\geq 0.99$  as shown in Tab. 53.

Tab. 53: Probability for each split:  $R \geq 0.99$  and  $R < 0.99$ .

Threshold	1	0	Probability
$\geq 0.99$	5	3	$\frac{8}{10}=0.8$
$< 0.99$	0	2	$\frac{2}{10}=0.2$

- Calculate entropy with respect to feature  $R$  threshold  $\geq 0.99$ .

$$E(R \geq 0.99) = E(5, 3) = -\frac{5}{8} \log_2 \frac{5}{8} - \frac{3}{8} \log_2 \frac{3}{8} = 0.9544. \tag{84}$$

$$E(R < 0.99) = E(0, 2) = -\frac{0}{2} \log_2 \frac{0}{2} - \frac{2}{2} \log_2 \frac{2}{2} = 0. \tag{85}$$

- Calculate entropy of feature with respect to class.

$$E(class, R) = P(R \geq 0.99) E(R \geq 0.99) + P(R < 0.99) E(R < 0.99) = (0.8)(0.9544) + (0.2)(0) = 0.7635. \tag{86}$$

- Calculate information gain of feature  $R \geq 0.98$ .

$$IG(R \geq 0.99) = E(class) - E(class, R) = 1 - 0.7635 = 0.2365. \tag{87}$$

4) Calculate IG for Threshold  $\geq 0.99$  Using Tab. 54

Tab. 54: Sub Table:  $R$  Vs. class with threshold  $\geq 0.1$ .

$R$	Class	$R$	Class
$< 0.1$	0	$\geq 0.1$	1
$< 0.1$	0	$\geq 0.1$	1
$< 0.1$	0	$\geq 0.1$	1
$< 0.1$	1	$< 0.1$	0
$\geq 0.1$	1	$< 0.1$	0

Tab. 55: Probability for each split:  $R \geq 0.1$  and  $R < 0.1$ .

Threshold	1	0	Probability
$\geq 0.1$	4	0	$\frac{4}{10}=0.4$
$< 0.1$	1	5	$\frac{6}{10}=0.6$

- Calculate probability with respect to feature  $R$  threshold  $\geq 0.1$  as shown in Tab. 55.
- Calculate entropy with respect to feature  $R$  threshold  $\geq 0.1$ .

$$E(R \geq 0.1) = E(4, 0) = -\frac{4}{4} \log_2 \frac{4}{4} - \frac{0}{4} \log_2 \frac{0}{4} = 1. \tag{88}$$

$$E(R < 0.1) = E(1, 5) = -\frac{1}{6} \log_2 \frac{1}{6} - \frac{5}{6} \log_2 \frac{5}{6} = 0.65. \tag{89}$$

- Calculate entropy of feature with respect to class.

$$E(class, R) = P(R \geq 0.1) E(R \geq 0.1) + P(R < 0.1) E(R < 0.1) = (0.4)(1) + (0.6)(0.65) = 0.79. \tag{90}$$

- Calculate information gain of feature  $R \geq 0.1$ .

$$IG(R \geq 0.1) = E(class) - E(class, R) = 1 - 0.79 = 0.21. \tag{91}$$

- Information Gain of each split of feature  $R$  is presented in Tab. 56. From the Tab. 56, it is observed that split  $\geq 0.99$  gives highest information gain value i.e 0.2365. Hence,  $\geq 0.99$  is considered as a best split.

Tab. 56: Information Gain of each split of feature  $R$ .

Split	IG	Split	IG
$\geq 0.96$	0	$\geq 0.99$	<b>0.2365</b>
$\geq 0.98$	0.108	$\geq 0.1$	0.21

1.4. Identification of Best Threshold for RMSE i.e. Root Mean Square Error Feature

1) Calculate IG for Threshold  $\geq 0.027$  Using Tab. 57

- Calculate probability with respect to feature  $RMSE$  threshold  $\geq 0.027$  as shown in Tab. 58.



**Tab. 57:** Sub Table: *RMSE* Vs. class with threshold  $\geq 0.027$ .

<i>RMSE</i>	Class	<i>RMSE</i>	Class
$\geq 0.027$	0	$< 0.027$	1
$< 0.027$	0	$< 0.027$	1
$< 0.027$	0	$< 0.027$	1
$< 0.027$	1	$< 0.027$	0
$< 0.027$	1	$< 0.027$	0

**Tab. 58:** Probability for each spilt: *RMSE*  $\geq 0.027$  and *RMSE*  $< 0.027$ .

Threshold	1	0	Probability
$\geq 0.027$	0	1	$\frac{1}{10}=0.1$
$< 0.027$	5	4	$\frac{9}{10}=0.9$

- Calculate entropy with respect to feature *RMSE* threshold  $\geq 0.027$ .

$$E(RMSE \geq 0.027) = E(0, 1) = -\frac{0}{1} \log_2 \frac{0}{1} - \frac{1}{1} \log_2 \frac{1}{1} = 0. \tag{92}$$

$$E(RMSE < 0.027) = E(5, 4) = -\frac{5}{9} \log_2 \frac{5}{9} - \frac{4}{9} \log_2 \frac{4}{9} = 0.9911. \tag{93}$$

- Calculate entropy of feature with respect to class.

$$E(class, RMSE) = P(RMSE \geq 0.027) E(RMSE \geq 0.027) + P(RMSE < 0.027) E(RMSE < 0.027) = (0.1)(0) + (0.9)(0.9911) = 0.8919. \tag{94}$$

- Calculate information gain of feature *RMSE*  $\geq 0.027$ .

$$IG(RMSE \geq 0.027) = E(class) - E(class, RMSE) = 1 - 0.8919 = 0.1081. \tag{95}$$

**2) Calculate IG for Threshold  $\geq 0.01$  Using Tab. 59**

**Tab. 59:** Sub Table: *RMSE* Vs. class with threshold  $\geq 0.01$ .

<i>RMSE</i>	Class	<i>RMSE</i>	Class
$\geq 0.01$	0	$< 0.01$	1
$\geq 0.01$	0	$< 0.01$	1
$< 0.01$	0	$< 0.01$	1
$< 0.01$	1	$< 0.01$	0
$< 0.01$	1	$< 0.01$	0

- Calculate probability with respect to feature *RMSE* threshold  $\geq 0.01$  as shown in Tab. 60.

**Tab. 60:** Probability for each spilt: *RMSE*  $\geq 0.01$  and *RMSE*  $< 0.01$ .

Threshold	1	0	Probability
$\geq 0.01$	2	3	$\frac{5}{10}=0.5$
$< 0.01$	3	2	$\frac{5}{10}=0.5$

- Calculate entropy with respect to feature *RMSE* threshold  $\geq 0.01$ .

$$E(RMSE \geq 0.01) = E(2, 3) = -\frac{2}{5} \log_2 \frac{2}{5} - \frac{3}{5} \log_2 \frac{3}{5} = 0.971. \tag{96}$$

$$E(RMSE < 0.01) = E(3, 2) = -\frac{3}{5} \log_2 \frac{3}{5} - \frac{2}{5} \log_2 \frac{2}{5} = 0.971. \tag{97}$$

- Calculate entropy of feature with respect to class.

$$E(class, RMSE) = P(RMSE \geq 0.01) E(RMSE \geq 0.01) + P(RMSE < 0.01) E(RMSE < 0.01) = (0.5)(0.971) + (0.5)(0.971) = 0.971. \tag{98}$$

- Calculate information gain of feature *RMSE*  $\geq 0.01$ .

$$IG(RMSE \geq 0.01) = E(class) - E(class, RMSE) = 1 - 0.971 = 0.029. \tag{99}$$

**3) Calculate IG for Threshold  $\geq 0.003$  Using Tab. 61**

**Tab. 61:** Sub Table: *RMSE* Vs. class with threshold  $\geq 0.003$ .

<i>RMSE</i>	Class	<i>RMSE</i>	Class
$\geq 0.003$	0	$\geq 0.003$	1
$\geq 0.003$	0	$\geq 0.003$	1
$\geq 0.003$	0	$\geq 0.003$	1
$\geq 0.003$	1	$\geq 0.003$	0
$\geq 0.003$	1	$\geq 0.003$	0

- Calculate probability with respect to feature *RMSE* threshold  $\geq 0.003$  as shown in Tab. 62.

**Tab. 62:** Probability for each spilt: *RMSE*  $\geq 0.003$  and *RMSE*  $< 0.003$ .

Threshold	1	0	Probability
$\geq 0.003$	5	5	$\frac{10}{10}=1$
$< 0.003$	0	0	$\frac{0}{10}=0$

- Calculate entropy with respect to feature *RMSE* threshold  $\geq 0.003$ .

$$E(RMSE \geq 0.003) = E(5, 5) = -\frac{5}{10} \log_2 \frac{5}{10} - \frac{5}{10} \log_2 \frac{5}{10} = 1. \tag{100}$$

$$E(RMSE < 0.003) = E(0, 0) = 0. \tag{101}$$

- Calculate entropy of feature with respect to class.

$$E(class, RMSE) = P(RMSE \geq 0.003) E(RMSE \geq 0.003) + P(RMSE < 0.003) E(RMSE < 0.003) = (1)(1) + (0)(0) = 1. \tag{102}$$

- Calculate information gain of feature *RMSE*  $\geq 0.003$ .

$$IG(RMSE \geq 0.003) = E(class) - E(class, RMSE) = 1 - 1 = 0. \tag{103}$$

**4) Calculate IG for Threshold  $\geq 0.004$  Using Tab. 63**

**Tab. 63:** Sub Table: *RMSE* Vs. class with threshold  $\geq 0.004$ .

<i>RMSE</i>	Class	<i>RMSE</i>	Class
$\geq 0.004$	0	$< 0.004$	1
$\geq 0.004$	0	$\geq 0.004$	1
$\geq 0.004$	0	$< 0.004$	1
$\geq 0.004$	1	$\geq 0.004$	0
$\geq 0.004$	1	$\geq 0.004$	0

- Calculate probability with respect to feature *RMSE* threshold  $\geq 0.004$  as shown in Tab. 64.

**Tab. 64:** Probability for each spilt: *RMSE*  $\geq 0.004$  and *RMSE*  $< 0.004$ .

Threshold	1	0	Probability
$\geq 0.004$	3	5	$\frac{8}{10}=0.8$
$< 0.004$	2	0	$\frac{2}{10}=0.2$

- Calculate entropy with respect to feature *RMSE* threshold  $\geq 0.004$ .

$$E(RMSE \geq 0.004) = E(3, 5) = -\frac{3}{8} \log_2 \frac{3}{8} - \frac{5}{8} \log_2 \frac{5}{8} = 0.9554. \tag{104}$$

$$E(RMSE < 0.004) = E(2, 0) = -\frac{2}{2} \log_2 \frac{2}{2} - \frac{0}{2} \log_2 \frac{0}{2} = 0. \tag{105}$$

- Calculate entropy of feature with respect to class.

$$E(class, RMSE) = P(RMSE \geq 0.004) E(RMSE \geq 0.004) + P(RMSE < 0.004) E(RMSE < 0.004) = (0.8)(0.9554) + (0.2)(0) = 0.7643. \tag{106}$$

- Calculate information gain of feature *RMSE*  $\geq 0.004$ .

$$IG(RMSE \geq 0.004) = E(class) - E(class, RMSE) = 1 - 0.7643 = 0.2357. \tag{107}$$

**5) Calculate IG for Threshold  $\geq 0.005$  Using Tab. 65**

**Tab. 65:** Sub Table: *RMSE* Vs. class with threshold  $\geq 0.005$ .

<i>RMSE</i>	Class	<i>RMSE</i>	Class
$\geq 0.005$	0	$< 0.005$	1
$\geq 0.005$	0	$< 0.005$	1
$\geq 0.005$	0	$< 0.005$	1
$\geq 0.005$	1	$\geq 0.005$	0
$\geq 0.005$	1	$\geq 0.005$	0

- Calculate probability with respect to feature *RMSE* threshold  $\geq 0.005$  as shown in Tab. 66.

**Tab. 66:** Probability for each spilt: *RMSE*  $\geq 0.005$  and *RMSE*  $< 0.005$ .

Threshold	1	0	Probability
$\geq 0.005$	2	5	$\frac{7}{10}=0.7$
$< 0.005$	3	0	$\frac{3}{10}=0.3$

- Calculate entropy with respect to feature *RMSE* threshold  $\geq 0.005$ .

$$E(RMSE \geq 0.005) = E(2, 5) = -\frac{2}{7} \log_2 \frac{2}{7} - \frac{5}{7} \log_2 \frac{5}{7} = 0.862. \tag{108}$$

$$E(RMSE < 0.005) = E(3, 0) = -\frac{3}{3} \log_2 \frac{3}{3} - \frac{0}{3} \log_2 \frac{0}{3} = 0. \tag{109}$$

- Calculate entropy of feature with respect to class.

$$E(class, RMSE) = P(RMSE \geq 0.005) E(RMSE \geq 0.005) + P(RMSE < 0.005) E(RMSE < 0.005) = (0.7)(0.862) + (0.3)(0) = 0.6034. \tag{110}$$

- Calculate information gain of feature  $RMSE \geq 0.005$ .

$$\begin{aligned}
 IG(RMSE \geq 0.005) &= \\
 &= E(class) - E(class, RMSE) = \quad (111) \\
 &= 1 - 0.6034 = 0.3966.
 \end{aligned}$$

**6) Calculate IG for Threshold  $\geq 0.007$  Using Tab. 67**

**Tab. 67:** Sub Table:  $RMSE$  Vs. class with threshold  $\geq 0.007$ .

$RMSE$	Class	$RMSE$	Class
$\geq 0.007$	0	$< 0.007$	1
$\geq 0.007$	0	$< 0.007$	1
$\geq 0.007$	0	$< 0.007$	1
$\geq 0.007$	1	$< 0.007$	0
$\geq 0.007$	1	$\geq 0.007$	0

- Calculate probability with respect to feature  $RMSE$  threshold  $\geq 0.007$  as shown in Tab. 68.

**Tab. 68:** Probability for each split:  $RMSE \geq 0.007$  and  $RMSE < 0.007$ .

Threshold	1	0	Probability
$\geq 0.007$	2	4	$\frac{6}{10}=0.6$
$< 0.007$	3	1	$\frac{4}{10}=0.4$

- Calculate entropy with respect to feature  $RMSE$  threshold  $\geq 0.007$ .

$$\begin{aligned}
 E(RMSE \geq 0.007) &= E(2, 4) = \\
 &= -\frac{2}{6} \log_2 \frac{2}{6} - \frac{4}{6} \log_2 \frac{4}{6} = 0.918. \quad (112)
 \end{aligned}$$

$$\begin{aligned}
 E(RMSE < 0.007) &= E(3, 1) = \\
 &= -\frac{3}{4} \log_2 \frac{3}{4} - \frac{1}{4} \log_2 \frac{1}{4} = 0.8112. \quad (113)
 \end{aligned}$$

- Calculate entropy of feature with respect to class.

$$\begin{aligned}
 E(class, RMSE) &= \\
 &= P(RMSE \geq 0.007) E(RMSE \geq 0.007) + \\
 &+ P(RMSE < 0.007) E(RMSE < 0.007) = \\
 &= (0.6)(0.918) + (0.4)(0.8112) = 0.8753. \quad (114)
 \end{aligned}$$

- Calculate information gain of feature  $RMSE \geq 0.007$ .

$$\begin{aligned}
 IG(RMSE \geq 0.007) &= E(class) - \\
 &+ E(class, RMSE) = 1 - 0.8753 = 0.1247. \quad (115)
 \end{aligned}$$

- Information Gain of each split of feature  $RMSE$  is presented in Tab. 69. From the Tab. 69, it is observed that split  $\geq 0.005$  gives highest information gain value i.e 0.3966. Hence,  $\geq 0.005$  is considered as a best split.

**Tab. 69:** Information Gain of each split of feature  $RMSE$ .

Split	$IG$	Split	$IG$
$\geq 0.007$	0.1247	$\geq 0.004$	0.2357
$\geq 0.005$	<b>0.3966</b>	$\geq 0.003$	0
$\geq 0.01$	0.029	$\geq 0.027$	0.1081

Based on identification of best split for each feature, the continuous data shown in Tab. 13 is converted into categorical data as presented in Tab. 70, and the data shown in Tab. 70 is used to build decision tree model.

**Tab. 70:** Categorical dataset that is used to build decision tree.

$m$	$c$	$R$	$RMSE$	Class
$< -299$	$\geq 3.3$	$< 0.99$	$\geq 0.005$	0
$< -299$	$< 3.3$	$< 0.99$	$\geq 0.005$	0
$< -299$	$\geq 3.3$	$\geq 0.99$	$\geq 0.005$	0
$\geq -299$	$< 3.3$	$\geq 0.99$	$\geq 0.005$	1
$\geq -299$	$< 3.3$	$\geq 0.99$	$\geq 0.005$	1
$< -299$	$< 3.3$	$\geq 0.99$	$< 0.005$	1
$\geq -299$	$< 3.3$	$\geq 0.99$	$< 0.005$	1
$< -299$	$< 3.3$	$\geq 0.99$	$< 0.005$	1
$< -299$	$< 3.3$	$\geq 0.99$	$\geq 0.005$	0
$< -299$	$< 3.3$	$\geq 0.99$	$\geq 0.005$	0

## Appendix B Decision Tree Formulation using Information Gain (IG)

In this section, decision tree formulation based on data available in Tab. 70 using information gain is explained.

### 2.1. Identification of Root Node

#### 1) Calculate IG for Feature $m$

- Calculate probability with respect to feature  $m$  as shown in Tab. 71.

**Tab. 71:** Probability for the feature  $m$ .

Threshold	1	0	Probability
$\geq -299$	3	0	$\frac{3}{10}=0.3$
$< -299$	2	5	$\frac{7}{10}=0.7$

- Calculate entropy with respect to feature  $m$ .

$$\begin{aligned}
 E(m \geq -299) &= E(3, 0) = \\
 &= -\frac{3}{3} \log_2 \frac{3}{3} - \frac{0}{3} \log_2 \frac{0}{3} = 0.918.
 \end{aligned}
 \tag{116}$$

$$\begin{aligned}
 E(m < -299) &= E(2, 5) = \\
 &= -\frac{2}{7} \log_2 \frac{2}{7} - \frac{5}{7} \log_2 \frac{5}{7} = 0.8631.
 \end{aligned}
 \tag{117}$$

- Calculate entropy of feature with respect to class.

$$\begin{aligned}
 E(class, m) &= P(m \geq -299) E(Rm \geq -299) + \\
 &+ P(m < -299) E(m < -299) = \\
 &= (0.3)(0) + (0.7)(0.8631) = 0.6041.
 \end{aligned}
 \tag{118}$$

- Calculate information gain of feature  $m$ .

$$\begin{aligned}
 IG(m) &= E(class) - E(class, m) = \\
 &= 1 - 0.6041 = 0.3959.
 \end{aligned}
 \tag{119}$$

### 2) Calculate IG for Feature $c$

- Calculate probability with respect to feature  $c$  as shown in Tab. 72.

Tab. 72: Probability for the feature  $c$ .

Threshold	1	0	Probability
$\geq 3.3$	0	2	$\frac{2}{10}=0.2$
$< 3.3$	5	3	$\frac{8}{10}=0.8$

- Calculate entropy with respect to feature  $c$ .

$$\begin{aligned}
 E(c \geq 3.3) &= E(0, 2) = \\
 &= -\frac{0}{2} \log_2 \frac{0}{2} - \frac{2}{2} \log_2 \frac{2}{2} = 0.
 \end{aligned}
 \tag{120}$$

$$\begin{aligned}
 E(c < 3.3) &= E(5, 3) = \\
 &= -\frac{5}{8} \log_2 \frac{5}{8} - \frac{3}{8} \log_2 \frac{3}{8} = 0.9544.
 \end{aligned}
 \tag{121}$$

- Calculate entropy of feature with respect to class.

$$\begin{aligned}
 E(class, c) &= P(c \geq 3.3) E(c \geq 3.3) + \\
 &+ P(c < 3.3) E(c < 3.3) = \\
 &= (0.2)(0) + (0.8)(0.9544) = 0.7635.
 \end{aligned}
 \tag{122}$$

- Calculate information gain of feature  $c$ .

$$\begin{aligned}
 IG(c) &= E(class) - E(class, c) = \\
 &= 1 - 0.7635 = 0.2365.
 \end{aligned}
 \tag{123}$$

### 3) Calculate IG for Feature $R$

- Calculate probability with respect to feature  $R$  as shown in Tab. 73.

Tab. 73: Probability for the feature  $R$ .

Threshold	1	0	Probability
$\geq 0.99$	5	3	$\frac{8}{10}=0.8$
$< 0.99$	0	2	$\frac{2}{10}=0.2$

- Calculate entropy with respect to feature  $R$ .

$$\begin{aligned}
 E(R \geq 0.99) &= E(5, 3) = \\
 &= -\frac{5}{8} \log_2 \frac{5}{8} - \frac{3}{8} \log_2 \frac{3}{8} = 0.9544.
 \end{aligned}
 \tag{124}$$

$$\begin{aligned}
 E(R < 0.99) &= E(0, 2) = \\
 &= -\frac{0}{2} \log_2 \frac{0}{2} - \frac{2}{2} \log_2 \frac{2}{2} = 0.
 \end{aligned}
 \tag{125}$$

- Calculate entropy of feature with respect to class.

$$\begin{aligned}
 E(class, R) &= P(R \geq 0.99) E(R \geq 0.99) + \\
 &+ P(R < 0.99) E(R < 0.99) = \\
 &= (0.8)(0.9544) + (0.2)(0) = 0.7635.
 \end{aligned}
 \tag{126}$$

- Calculate information gain of feature  $R$ .

$$\begin{aligned}
 IG(R) &= E(class) - E(class, R) = \\
 &= 1 - 0.7635 = 0.2365.
 \end{aligned}
 \tag{127}$$

### 4) Calculate IG for Feature $RMSE$

- Calculate probability with respect to feature  $RMSE$  as shown in Tab. 74.

Tab. 74: Probability for the feature  $RMSE$ .

Threshold	1	0	Probability
$\geq 0.005$	2	5	$\frac{7}{10}=0.7$
$< 0.005$	3	0	$\frac{3}{10}=0.3$

- Calculate entropy with respect to feature  $RMSE$ .

$$\begin{aligned}
 E(RMSE \geq 0.005) &= E(2, 5) = \\
 &= -\frac{2}{7} \log_2 \frac{2}{7} - \frac{5}{7} \log_2 \frac{5}{7} = 0.8631.
 \end{aligned}
 \tag{128}$$

$$\begin{aligned}
 E(RMSE < 0.005) &= E(3, 0) = \\
 &= -\frac{3}{3} \log_2 \frac{3}{3} - \frac{0}{3} \log_2 \frac{0}{3} = 0.
 \end{aligned}
 \tag{129}$$

- Calculate entropy of feature with respect to class.

$$\begin{aligned}
 E(\text{class}, RMSE) &= \\
 &= P(RMSE \geq 0.005) E(RMSE \geq 0.005) + \\
 &+ P(RMSE < 0.005) E(RMSE < 0.005) = \\
 &= (0.7)(0.8631) + (0.3)(0) = 0.6041.
 \end{aligned}
 \tag{130}$$

- Calculate information gain of feature *RMSE*.

$$\begin{aligned}
 IG(RMSE) &= E(\text{class}) - \\
 + E(\text{class}, RMSE) &= 1 - 0.6041 = 0.3959.
 \end{aligned}
 \tag{131}$$

- Information Gain of each feature is presented in Tab. 75. From the Tab. 75, it is observed that both *m* and *RMSE* have highest information gain value i.e 0.3959. Hence, *m* is considered as root node.

Tab. 75: Information Gain of each feature.

Split	IG	Split	IG
<i>m</i>	0.3959	<i>c</i>	0.2365
<i>R</i>	0.2365	<i>RMSE</i>	0.3959

After the identification of best feature, the data shown in Tab. 70 is split into two sub tables based on categorical values in feature *m* and these sub tables shown in Tab. 76 and Tab. 77. From Tab. 77, the branch  $\geq -299$  for the root node *m* has a leaf node with value '1'.

Tab. 76: Categorical dataset with *m* < -299.

<i>c</i>	<i>R</i>	<i>RMSE</i>	Class
$\geq 3.3$	$< 0.99$	$\geq 0.005$	0
$< 3.3$	$< 0.99$	$\geq 0.005$	0
$\geq 3.3$	$\geq 0.99$	$\geq 0.005$	0
$< 3.3$	$\geq 0.99$	$< 0.005$	1
$< 3.3$	$\geq 0.99$	$< 0.005$	1
$< 3.3$	$\geq 0.99$	$\geq 0.005$	0
$< 3.3$	$\geq 0.99$	$\geq 0.005$	0

Tab. 77: Categorical dataset with *m*  $\geq -299$ .

<i>c</i>	<i>R</i>	<i>RMSE</i>	Class
$< 3.3$	$\geq 0.99$	$\geq 0.005$	1
$< 3.3$	$\geq 0.99$	$\geq 0.005$	1
$< 3.3$	$\geq 0.99$	$< 0.005$	1

## 2.2. Identification of Decision Node under Branch *m* < -299

### 1) Calculate IG for Feature *c*

- Calculate probability with respect to feature *c* based on data in Tab. 76 and presented in Tab. 78.

Tab. 78: Probability for the feature *c*.

Threshold	1	0	Probability
$\geq 3.3$	0	2	$\frac{2}{7}=0.286$
$< 3.3$	2	3	$\frac{5}{7}=0.714$

- Calculate entropy with respect to feature *c*.

$$\begin{aligned}
 E(c \geq 3.3) &= E(0, 2) = \\
 &= -\frac{0}{2} \log_2 \frac{0}{2} - \frac{2}{2} \log_2 \frac{2}{2} = 0.
 \end{aligned}
 \tag{132}$$

$$\begin{aligned}
 E(c < 3.3) &= E(2, 3) = \\
 &= -\frac{2}{5} \log_2 \frac{2}{5} - \frac{3}{5} \log_2 \frac{3}{5} = 0.971.
 \end{aligned}
 \tag{133}$$

- Calculate entropy of feature with respect to class.

$$\begin{aligned}
 E(\text{class}, c) &= P(c \geq 3.3) E(c \geq 3.3) + \\
 &+ P(c < 3.3) E(c < 3.3) = \\
 &= (0.286)(0) + (0.714)(0.971) = 0.6932.
 \end{aligned}
 \tag{134}$$

- Calculate information gain of feature *c*.

$$\begin{aligned}
 IG(c) &= E(\text{class}) - E(\text{class}, c) = \\
 &= 1 - 0.6932 = 0.3068.
 \end{aligned}
 \tag{135}$$

### 2) Calculate IG for Feature *R*

- Calculate probability with respect to feature *R* based on data in Tab. 76 and presented in Tab. 79.

Tab. 79: Probability for the feature *R*.

Threshold	1	0	Probability
$\geq 0.99$	2	3	$\frac{5}{7}=0.714$
$< 0.99$	0	2	$\frac{2}{7}=0.286$

- Calculate entropy with respect to feature *R*.

$$\begin{aligned}
 E(R \geq 0.99) &= E(2, 3) = \\
 &= -\frac{3}{5} \log_2 \frac{3}{5} - \frac{2}{5} \log_2 \frac{2}{5} = 0.971.
 \end{aligned}
 \tag{136}$$

$$\begin{aligned}
 E(R < 0.99) &= E(0, 2) = \\
 &= -\frac{0}{2} \log_2 \frac{0}{2} - \frac{0}{2} \log_2 \frac{0}{2} = 0.
 \end{aligned}
 \tag{137}$$

- Calculate entropy of feature with respect to class.

$$\begin{aligned}
 E(\text{class}, R) &= P(R \geq 0.99) E(R \geq 0.99) + \\
 &+ P(R < 0.99) E(R < 0.99) = \\
 &= (0.714)(0.971) + (0.286)(0) = 0.6932.
 \end{aligned}
 \tag{138}$$

- Calculate information gain of feature  $R$ .

$$IG(R) = E(class) - E(class, R) = 1 - 0.6932 = 0.3068. \tag{139}$$

**3) Calculate IG for Feature RMSE**

- Calculate probability with respect to feature  $RMSE$  based on data in Tab. 76 and presented in Tab. 80.

**Tab. 80:** Probability for the feature  $RMSE$ .

Threshold	1	0	Probability
$\geq 0.005$	0	5	$\frac{5}{7}=0.714$
$< 0.005$	2	0	$\frac{2}{7}=0.286$

- Calculate entropy with respect to feature  $RMSE$ .

$$E(RMSE \geq 0.005) = E(0, 5) = -\frac{0}{5} \log_2 \frac{0}{5} - \frac{5}{5} \log_2 \frac{5}{5} = 0. \tag{140}$$

$$E(RMSE < 0.005) = E(2, 0) = -\frac{2}{2} \log_2 \frac{2}{2} - \frac{0}{2} \log_2 \frac{0}{2} = 0. \tag{141}$$

- Calculate entropy of feature with respect to class.

$$E(class, RMSE) = P(RMSE \geq 0.005) E(RMSE \geq 0.005) + P(RMSE < 0.005) E(RMSE < 0.005) = (0.714)(0) + (0.286)(0) = 0. \tag{142}$$

- Calculate information gain of feature  $RMSE$ .

$$IG(RMSE) = E(class) - E(class, RMSE) = 1 - 0 = 1. \tag{143}$$

- Information Gain of each feature is presented in Tab. 81. From the Tab. 81, it is observed that  $RMSE$  has highest information gain value i.e 1. Hence,  $RMSE$  is considered as decision node under branch  $m < -299$ .

**Tab. 81:** Information Gain of each feature.

Feature	$IG$	Feature	$IG$	Feature	$IG$
$c$	0.3068	$R$	0.3068	$RMSE$	1

After the identification of the decision node, the data shown in Tab. 76 is split into two sub tables based

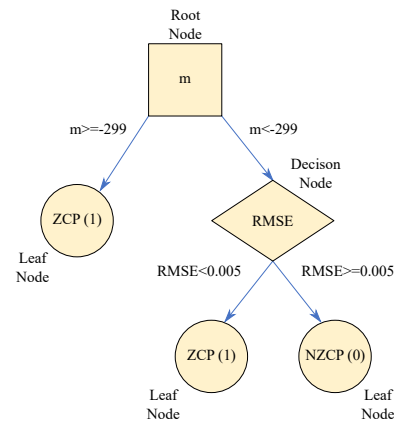
on categorical values in feature  $RMSE$  and these sub tables shown in Tab. 82 and Tab. 83. From Tab. 82, the branch  $\geq 0.005$  for the decision node  $RMSE$  has a leaf node with value '0'. Similarly, From Tab. 83, the branch  $< 0.005$  for the decision node  $RMSE$  has a leaf node with value '1'. Complete decision tree model is shown in Fig. 11. From the Fig. 11, it is observed that tree end with leaf nodes and it will not grow further.

**Tab. 82:** Categorical dataset with  $RMSE \geq 0.005$ .

$c$	$R$	Class
$\geq 3.3$	$< 0.99$	0
$< 3.3$	$< 0.99$	0
$\geq 3.3$	$\geq 0.99$	0
$< 3.3$	$\geq 0.99$	0
$< 3.3$	$\geq 0.99$	0

**Tab. 83:** Categorical dataset with  $RMSE < 0.005$ .

$c$	$R$	Class
$< 3.3$	$\geq 0.99$	1
$< 3.3$	$\geq 0.99$	1



**Fig. 11:** Decision tree model with information gain.

**4) Accuracy of Decision Tree Model**

In this section, the performance of the decision tree model shown in Fig. 11 is observed using data shown in Tab. 13 based on accuracy. The predicted class label for the data shown in Tab. 13 is presented in Tab. 84.

- Formation of confusion matrix.

Confusion matrix is shown in Tab. 85. Accuracy of the developed decision tree model is calculated using

**Tab. 84:** Predicted label for few data samples from the dataset that prepared from noisy distorted signal.

<i>m</i>	<i>c</i>	<i>R</i>	<i>RMSE</i>	Class	Pred
-448	4.4	0.96	0.027	0	0
-313	3.1	0.98	0.010	0	0
-331	3.3	0.99	0.009	0	0
-299	3.0	0.99	0.009	1	1
-269	2.7	1.00	0.009	1	1
-301	3.0	1.00	0.003	1	1
-293	2.9	1.00	0.004	1	1
-319	3.2	1.00	0.003	1	1
-303	3.0	0.99	0.005	0	0
-324	3.2	0.99	0.007	0	0

below equation and it is equal to 100 %.

$$\begin{aligned}
 \text{Accuracy} &= \frac{TNZCP + TZCP}{TNZCP + TZCP + FNZCP + FZCP} \\
 &= \frac{5 + 5}{5 + 5 + 0 + 0} = 100 \%.
 \end{aligned}
 \tag{144}$$

**Tab. 85:** Confusion Matrix for the decision tree model with information gain.

Confusion Matrix		Acutal Label	
		0	1
Predicted label	0	TNZCP:5	FNZCP:0
	1	FZCP:0	TZCP:5

## Appendix C Decision Tree Formulation using Gini Index (GI)

In this section, decision tree formulation based on data available in Tab. 70 using Gini Index (GI) is explained.

### 3.1. Identification of Root Node based on GI

#### 1) Calculate GI for Feature *m*

- Calculate GI with respect to feature *m* as shown in Tab. 86.

**Tab. 86:** GI for the feature *m*.

Thres.	1	0	No. of samples	P(1)	P(0)	P( <i>m</i> )	<i>GI</i>
$\geq -299$	3	0	3	$\frac{3}{3}$	$\frac{0}{3}$	$\frac{3}{10}$	0
$< -299$	2	5	7	$\frac{2}{7}$	$\frac{5}{7}$	$\frac{7}{10}$	0.408

- Calculate GI with respect to feature *m*.

$$\begin{aligned}
 GI(m \geq -299) &= GI(3, 0) = \\
 &= 1 - \left[ \left(\frac{3}{3}\right)^2 + \left(\frac{0}{3}\right)^2 \right] = 0.
 \end{aligned}
 \tag{145}$$

$$\begin{aligned}
 GI(m < -299) &= GI(2, 5) = \\
 &= 1 - \left[ \left(\frac{2}{7}\right)^2 + \left(\frac{5}{7}\right)^2 \right] = 0.408.
 \end{aligned}
 \tag{146}$$

$$\begin{aligned}
 GI(\text{class}, m) &= P(m \geq -299) GI(m \geq -299) + \\
 &+ P(m < -299) GI(m < -299) = \\
 &= \frac{3}{10} \cdot 0 + \frac{7}{10} \cdot 0.408 = 0.2856.
 \end{aligned}
 \tag{147}$$

#### 2) Calculate GI for Feature *c*

- Calculate GI with respect to feature *c* as shown in Tab. 87.

**Tab. 87:** GI for the feature *c*.

Thres.	1	0	No. of samples	P(1)	P(0)	P( <i>m</i> )	<i>GI</i>
$\geq 3.3$	0	2	2	$\frac{0}{2}$	$\frac{2}{2}$	$\frac{2}{10}$	0
$< 3.3$	5	3	8	$\frac{5}{8}$	$\frac{3}{8}$	$\frac{8}{10}$	0.468

- Calculate GI with respect to feature *c*.

$$\begin{aligned}
 GI(c \geq 3.3) &= GI(0, 2) = \\
 &= 1 - \left[ \left(\frac{0}{2}\right)^2 + \left(\frac{2}{2}\right)^2 \right] = 0.
 \end{aligned}
 \tag{148}$$

$$\begin{aligned}
 GI(c < 3.3) &= GI(5, 3) = \\
 &= 1 - \left[ \left(\frac{5}{8}\right)^2 + \left(\frac{3}{8}\right)^2 \right] = 0.468.
 \end{aligned}
 \tag{149}$$

$$\begin{aligned}
 GI(\text{class}, c) &= P(c \geq 3.3) GI(c \geq 3.3) + \\
 &+ P(c < 3.3) GI(c < 3.3) = \\
 &= \frac{2}{10} \cdot 0 + \frac{8}{10} \cdot 0.468 = 0.374.
 \end{aligned}
 \tag{150}$$

#### 3) Calculate GI for Feature *R*

- Calculate GI with respect to feature *R* as shown in Tab. 88.

- Calculate GI with respect to feature *R*.

**Tab. 88:** GI for the feature *R*.

Thres.	1	0	No. of samples	P(1)	P(0)	P( <i>m</i> )	GI
≥ 0.99	5	3	8	$\frac{5}{8}$	$\frac{3}{8}$	$\frac{8}{10}$	0.468
< 0.99	0	2	2	$\frac{0}{2}$	$\frac{2}{2}$	$\frac{2}{10}$	0

$$GI(R \geq 0.99) = GI(5, 3) = 1 - \left[ \left(\frac{5}{8}\right)^2 + \left(\frac{3}{8}\right)^2 \right] = 0.468. \tag{151}$$

$$GI(R < 0.99) = GI(0, 2) = 1 - \left[ \left(\frac{0}{2}\right)^2 + \left(\frac{2}{2}\right)^2 \right] = 0. \tag{152}$$

$$GI(class, R) = P(R \geq 0.99)GI(R \geq 0.99) + P(R < 0.99)GI(R < 0.99) = \frac{8}{10} \cdot 0.468 + \frac{2}{10} \cdot 0 = 0.374. \tag{153}$$

**4) Calculate GI for Feature *RMSE***

- Calculate GI with respect to feature *RMSE* as shown in Tab. 89.

**Tab. 89:** GI for the feature *RMSE*.

Thres.	1	0	No. of samples	P(1)	P(0)	P( <i>m</i> )	GI
≥ 0.005	2	5	7	$\frac{2}{7}$	$\frac{5}{7}$	$\frac{7}{10}$	0.408
< 0.005	3	0	3	$\frac{3}{3}$	$\frac{0}{3}$	$\frac{3}{10}$	0

- Calculate GI with respect to feature *RMSE*.

$$GI(RMSE \geq 0.005) = GI(2, 5) = 1 - \left[ \left(\frac{2}{7}\right)^2 + \left(\frac{5}{7}\right)^2 \right] = 0.408. \tag{154}$$

$$GI(RMSE < 0.005) = GI(3, 0) = 1 - \left[ \left(\frac{3}{3}\right)^2 + \left(\frac{0}{3}\right)^2 \right] = 0. \tag{155}$$

$$GI(class, RMSE) = P(RMSE \geq 0.005)GI(RMSE \geq 0.005) + P(RMSE < 0.005)GI(RMSE < 0.005) = \frac{7}{10} \cdot 0.408 + \frac{3}{10} \cdot 0 = 0.2856. \tag{156}$$

- GI of each feature is presented in Tab. 90. From the Tab. 90, it is observed that both *m* and *RMSE* have lowest GI value i.e 0.2856. Hence, *RMSE* is considered as root node.

**Tab. 90:** GI value of each feature.

Feature	GI	Feature	GI
<i>m</i>	0.2856	<i>c</i>	0.374
<i>R</i>	0.374	<i>RMSE</i>	<b>0.2856</b>

After the identification of the best feature, the data shown in Tab. 70 is split into two sub tables based on categorical values in feature *RMSE* and these sub tables shown in Tab. 91 and Tab. 92. From Tab. 91, the branch < 0.005 for the root node *RMSE* has a leaf node with value '1'.

**Tab. 91:** Categorical dataset with *RMSE* < 0.005.

<i>m</i>	<i>C</i>	<i>R</i>	Class
< -299	< 3.3	≥ 0.99	1
≥ -299	< 3.3	≥ 0.99	1
< -299	< 3.3	≥ 0.99	1

**Tab. 92:** Categorical dataset with *RMSE* ≥ 0.005.

<i>m</i>	<i>C</i>	<i>R</i>	Label
< -299	≥ 3.3	< 0.99	0
< -299	< 3.3	< 0.99	0
< -299	≥ 3.3	≥ 0.99	0
≥ -299	< 3.3	≥ 0.99	1
≥ -299	< 3.3	≥ 0.99	1
< -299	< 3.3	≥ 0.99	0
< -299	< 3.3	≥ 0.99	0

**3.2. Identification of Decision Node under Branch *RMSE* ≥ 0.005**

**1) Calculate GI for Feature *m***

- Calculate GI with respect to feature *m* as shown in Tab. 93 using data shown in Tab. 92.

**Tab. 93:** GI for the feature *m* with respect to *RMSE* ≥ 0.005.

Thres.	1	0	No. of samples	P(1)	P(0)	P( <i>m</i> )	GI
≥ -299	2	0	2	$\frac{2}{2}$	$\frac{0}{2}$	$\frac{2}{7}$	0
< -299	0	5	5	$\frac{0}{5}$	$\frac{5}{5}$	$\frac{5}{7}$	0

- Calculate GI with respect to feature *m*.

$$GI(m \geq -299) = GI(2, 0) = 1 - \left[ \left(\frac{2}{2}\right)^2 + \left(\frac{0}{2}\right)^2 \right] = 0. \tag{157}$$



$$\begin{aligned}
 GI(m < -299) &= GI(0, 5) = \\
 &= 1 - \left[ \left(\frac{0}{5}\right)^2 + \left(\frac{5}{5}\right)^2 \right] = 0. \tag{158}
 \end{aligned}$$

$$\begin{aligned}
 GI(class, m) &= P(m \geq -299) GI(m \geq -299) + \\
 &+ P(m < -299) GI(m < -299) = \\
 &= \frac{2}{7} \cdot 0 + \frac{5}{7} \cdot 0 = 0. \tag{159}
 \end{aligned}$$

**2) Calculate GI for Feature c**

- Calculate GI with respect to feature *c* as shown in Tab. 94 using data shown in Tab. 92.

**Tab. 94:** GI for the feature *c* with respected to  $RMSE \geq 0.005$ .

Thres.	1	0	No. of samples	P(1)	P(0)	P( <i>m</i> )	GI
$\geq 3.3$	0	2	2	$\frac{0}{2}$	$\frac{2}{2}$	$\frac{2}{7}$	0
$< 3.3$	2	3	5	$\frac{2}{5}$	$\frac{3}{5}$	$\frac{5}{7}$	0.48

- Calculate GI with respect to feature *c*.

$$\begin{aligned}
 GI(c \geq 3.3) &= GI(0, 2) = \\
 &= 1 - \left[ \left(\frac{0}{2}\right)^2 + \left(\frac{2}{2}\right)^2 \right] = 0. \tag{160}
 \end{aligned}$$

$$\begin{aligned}
 GI(c < 3.3) &= GI(2, 3) = \\
 &= 1 - \left[ \left(\frac{2}{5}\right)^2 + \left(\frac{3}{5}\right)^2 \right] = 0.48. \tag{161}
 \end{aligned}$$

$$\begin{aligned}
 GI(class, c) &= P(c \geq 3.3) GI(c \geq 3.3) + \\
 &+ P(c < 3.3) GI(c < 3.3) = \\
 &= \frac{2}{7} \cdot 0 + \frac{5}{7} \cdot 0.48 = 0.342. \tag{162}
 \end{aligned}$$

**3) Calculate GI for Feature R**

- Calculate GI with respect to feature *c* as shown in Tab. 95 using data shown in Tab. 92.

**Tab. 95:** GI for the feature *R* with respected to  $RMSE \geq 0.005$ .

Thres.	1	0	No. of samples	P(1)	P(0)	P( <i>m</i> )	GI
$\geq 0.99$	2	3	5	$\frac{2}{5}$	$\frac{3}{5}$	$\frac{5}{7}$	0.48
$< 0.99$	0	2	2	$\frac{0}{2}$	$\frac{2}{2}$	$\frac{2}{7}$	0

- Calculate GI with respect to feature *R*.

$$\begin{aligned}
 GI(R \geq 0.99) &= GI(2, 3) = \\
 &= 1 - \left[ \left(\frac{2}{5}\right)^2 + \left(\frac{3}{5}\right)^2 \right] = 0.48. \tag{163}
 \end{aligned}$$

$$\begin{aligned}
 GI(R < 0.99) &= GI(0, 2) = \\
 &= 1 - \left[ \left(\frac{0}{2}\right)^2 + \left(\frac{2}{2}\right)^2 \right] = 0. \tag{164}
 \end{aligned}$$

$$\begin{aligned}
 GI(class, R) &= P(R \geq 0.99) GI(R \geq 0.99) + \\
 &+ P(R < 0.99) GI(R < 0.99) = \\
 &= \frac{5}{7} \cdot 0.48 + \frac{2}{7} \cdot 0 = 0.342. \tag{165}
 \end{aligned}$$

- GI of each feature is presented in Tab. 96. From the Tab. 96, it is observed that *m* has lowest GI value i.e 0. Hence, *m* is considered as decision node.

**Tab. 96:** GI value of each feature with respected to  $RMSE \geq 0.005$ .

Feature	GI	Feature	GI	Feature	GI
<i>m</i>	0	<i>c</i>	0.342	<i>R</i>	0.342

After the identification of the decision node, the data shown in Tab. 92 is split into two sub tables based on categorical values in feature *m* and these sub tables shown in Tab. 97 and Tab. 98. From Tab. 97, the branch  $\geq -299$  for the decision node *m* has a leaf node with value '1'. Similarly, from Tab. 98, the branch  $< -299$  for the decision node *m* has a leaf node with value '0'. The final decision tree based on GI value is shown in Fig. 12.

**Tab. 97:** Categorical dataset with  $m \geq -299$ .

<i>C</i>	<i>R</i>	Label
$< 3.3$	$\geq 0.99$	1
$< 3.3$	$\geq 0.99$	1

**Tab. 98:** Categorical dataset with  $m < -299$ .

<i>C</i>	<i>R</i>	Label
$\geq 3.3$	$< 0.99$	0
$< 3.3$	$< 0.99$	0
$\geq 3.3$	$\geq 0.99$	0
$< 3.3$	$\geq 0.99$	0
$< 3.3$	$\geq 0.99$	0

**4) Accuracy of Decision Tree Model**

In this section, the performance of the decision tree model shown in Fig. 12 is observed using data shown in Tab. 13 based on accuracy. The predicted class label for the data shown in Tab. 13 is presented in Tab. 99.

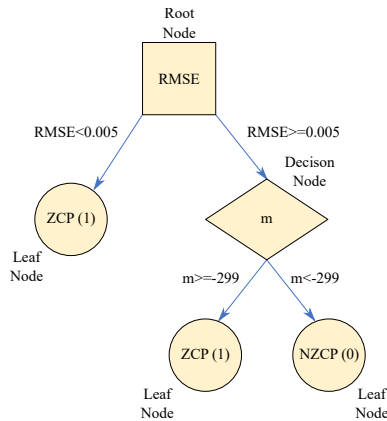


Fig. 12: Decision tree model with Gini Index.

$$\begin{aligned}
 Accuracy &= \frac{TNZCP + TZCP}{TNZCP + TZCP + FNZCP + FZCP} = \\
 &= \frac{5 + 5}{5 + 5 + 0 + 0} = 100 \%.
 \end{aligned}
 \tag{166}$$

Tab. 99: Predicted label for the dataset that prepared from noisy distorted signal using decision tree with GI.

<i>m</i>	<i>c</i>	<i>R</i>	<i>RMSE</i>	Class	Pred
-448	4.4	0.96	0.027	0	0
-313	3.1	0.98	0.010	0	0
-331	3.3	0.99	0.009	0	0
-299	3.0	0.99	0.009	1	1
-269	2.7	1.00	0.009	1	1
-301	3.0	1.00	0.003	1	1
-293	2.9	1.00	0.004	1	1
-319	3.2	1.00	0.003	1	1
-303	3.0	0.99	0.005	0	0
-324	3.2	0.99	0.007	0	0

- Formation of confusion matrix.

Confusion matrix is shown in Tab. 100. Accuracy of the developed decision tree model based on GI is calculated using below equation and it is equal to 100 %.

Tab. 100: Confusion Matrix for decision tree model based on GI.

Confusion Matrix		Acutal Label	
		0	1
Predicted label	0	TNZCP:5	FNZCP:0
	1	FZCP:0	TZCP:5