PERFORMANCE ANALYSIS OF HIGHER-ORDER STATISTICAL FEATURES IN CLASSIFICATION OF SOME MODULATION TYPES

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
ATILIM UNIVERSITY

BY

REMZİYE BÜŞRA TEZEL

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
ELECTRICAL AND ELECTRONICS ENGINEERING

JANUARY 2020

Approval of the Graduate School of Natural and Applied Sciences, Atilim University.

Prof. Dr. Ali KARA
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of **Master of Science in Electrical and Electronics Engineering Department, Atilim University**.

Assoc.Prof.Dr. Efe ESELLER
Head of Department

This is to certify that we have read the thesis PERFORMANCE ANALYSIS OF HIGHER-ORDER STATISTICAL FEATURES IN CLASSIFICATION OF SOME MODULATION TYPES submitted by REMZİYE BÜŞRA TEZEL and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Science.

Prof. Dr. Ali KARA
Supervisor

Examining Committee Members:

Assoc.Prof.Dr Enver ÇAVUŞ

Asst.Prof.Dr. Beytullah YILDIZ

Asst.Prof.Dr. Özgür ERGÜL

Asst.Prof.Dr. Hakan TORA

Prof.Dr. Ali KARA

Date: January 22, 2020

I declare and guarantee that all data, knowledge and information in this document has been obtained, processed and presented in accordance with academic rules and ethical conduct. Based on these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name : REMZİYE BÜŞRA TEZEL

Signature :

ABSTRACT

PERFORMANCE ANALYSIS OF HIGHER-ORDER STATISTICAL FEATURES IN CLASSIFICATION OF SOME MODULATION TYPES

Tezel, Remziye Büşra

M.S., Department of Electrical and Electronics Engineering

Supervisor : Prof. Dr. Ali KARA

January 2020, 53 pages

Modulation Classification algorithms are used to determine the modulation type of signal obtained at the receiver and to use the appropriate demodulator. There are 2 types as Feature-based(FB) and Likelihood-based(LB). In this thesis, FB method is used, which is less complex in structure. Algorithm has been developed to classify the signals that were modulated by 12 Analog and Digital Modulation types. Statistical features, Higher-order Moments(HOMs) and Higher-order Cumulants(HOCs) were used as features. Signals, which are recorded as over-the-air adding synthetic simulated channel effects, were classified with Linear, Quadratic, and Cubic Support Vector Machine(SVM). The classification performance of the signals examined at SNR from 0 dB to 20 dB were presented. As a result, the classification performance was found to be stable between 10 dB and 20 dB and is approximately 73%. The highest value of performance was observed in Quadratic SVM as 75.5% at 12dB. In this thesis, the limits of the developed modulation classification algorithm successfully presented with the features and SVM structure of 12 modulation types.

Keywords: Modulation Classification, Feature Extraction, Support Vector Machine, Analog Modulations, Digital Modulations

ÖZ

BAZI MODÜLASYON TÜRLERİNİN SINIFLANDIRILMASINDA YÜKSEK MERTEBEDEN İSTATİSTİKSEL ÖZELLİKLERIN PERFORMANS ANALİZİ

Tezel, Remziye Büşra

Yüksek Lisans, Elektrik ve Elektronik Mühendisliği

Tez Yöneticisi : Prof. Dr. Ali KARA

Ocak 22, 2020, 53 sayfa

Modülasyon Sınıflandırma algoritmaları, alıcıda elde edilen sinyalin modülasyon tipini belirlemek ve uygun demodulator seçimi için kullanılır. Özellik tabanlı ve Olabilirlik tabanlı olmak üzere 2 tür vardır. Bu tezde yapı olarak daha az karmaşık olan FB yöntemi kullanılmıştır. 12 Analog ve Dijital Modülasyon tipli sinyalleri sınıflandırmak için algoritma geliştirilmiştir. İstatistiksel özellikler, Yüksek Dereceli Momentler ve Yüksek Dereceli Kümülantlar kullanılmıştır. Havadan kaydedilen ve sentetik simüle kanal etkileri eklenen sinyaller Lineer, Kuadratik ve Kübik Destek Vektör Makinesi (DVM) ile sınıflandırıldı. SNR'de 0 dB ile 20 dB arasında incelenen sinyallerin sınıflandırma performansı sunulmuştur. Performansın 10 dB ve 20 dB arasında kararlı olduğu ve yaklaşık %73, en yüksek performansın ise Karesel SVM'de 12dB'de % 75.5 olduğu gözlenmiştir. Bu tezde, geliştirilen algoritmasının sınırları, 12 modülasyon tipinin özellikleri ve SVM yapısı ile başarılı bir şekilde sunulmuştur.

Anahtar Kelimeler: Modülasyon Sınıflandırma, Öznitelik Çıkarımı, Destek Vektör Makinesi, Analog Modülasyonlar, Dijital Modülasyonlar

To my family, my fiance and everyone who touches my life.

ACKNOWLEDGMENTS

I would like to thank my advisor Prof. Dr. Ali KARA for his patience and support, which enabled me to successfully complete the master's degree, for the great improvement in my career and ideas.

I will thank the rest of my thesis committee: Assoc.Prof. Dr Enver ÇAVUŞ, Asst. Prof. Dr. Beytullah YILDIZ, Asst. Prof. Dr. Özgür ERGÜL, Asst. Prof. Dr. Hakan TORA, for presenting their ideas for improving the work with the questions and comments they asked. I also thank Asst. Prof. Dr. Hakan TORA for contributing his ideas and help in my thesis study.

In addition, my friend that we have accomplished a lot in 2 years when our paths crossed at Atılım University, my colleague, my seatmate, and my confidant. I would like to thank her for her contribution to my work and for her motivation.

My fiance, we grew up together and survived every challenge. Thank you for being with me in every obstacle that I encountered in 8 years.

My family, I would like to thank you for your efforts, patience, support, and motivation to bring me to where I am. I am grateful to you for your contributions to my life.

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

LIST OF SYMBOLS

μ	:	Mean of Complex Signal
σ^2	:	Variance of Complex Signal
γ	:	Skewness of Complex Signal
μ_{42}	:	Kurtosis of Complex Signal
p	:	Order of Moment
q	:	Power of Signal's Conjugate
M_{pq}	:	Moment of Complex Signal
C_{pq}	:	Cumulant of Complex Signal
$a(n)$	:	Instantaneous Amplitude of Complex Signal
$f(n)$	:	Instantaneous Frequency of Complex Signal
$\varphi(n)$	:	Instantaneous Phase of Complex Signal
α_i	:	Support Vectors
$K(.,.)$	:	Kernel Function

CHAPTER 1

INTRODUCTION

Communication systems are the efficient and safe transmission of information from one place to another. Developing communication systems to meet the needs of people to communicate has an important place in today's technology. A general communication system consists of 3 main parts: Transmitter, Transmission Channel and Receiver [2]. This system is shown in Figure 1.1.

Figure 1.1: General Communication System

The transmitter allows the information signal to be transmitted on the system. The transmission channel provides the environment for the signal sent by the transmitter to reach the receiver. The receiver receives the signal from the channel to obtain the information signal. There are 2 types of transmission channels in the communication systems: wired and wireless. In wired communication, transmission channels such as twisted pairs, coaxial cables, and fiber-optic cables are used for data transmission, and air or space is used in wireless communication. Wireless communication utilizes infrared, radio and microwave signals.

There are two kinds of wireless transmission channels in the communication systems, namely Additive White Gaussian Noise (AWGN) and Multi- fading, which are called wireless transmission environment. Additive white Gaussian noise, which is added to the signals to mimic the wireless environment in nature, is a channel noise that is widely mentioned in the literature. However, due to many factors in nature, the AWGN channel is insufficient to represent nature. For this reason, A multi-fading channel is presented for a more complex structure for signals. There are effects such as reflection and refraction instead of noise due to objects in the wireless environment. Because of this complex structure, applications show lower performance when working in multi-fading channels. However, in order to use the system efficiently in real life, studies are conducted on multi-fading channels and the problems that may be in practice are tried to be minimized.

Multi-fading channels consist of Rician and Rayleigh channels. The difference between these channels can be explained by Line of Sight (LOS). The LOS is the absence of any obstructions between the receiver and transmitter. If LOS is provided, this channel is called the Rician channel. In other cases, LOS cannot be provided. This channel is called Rayleigh. In real life, there are many obstacles in the wireless communication environment and are subject to effects such as signal reflection and refraction. Therefore, the use of the Rayleigh channel expresses more real effects.

Due to problems such as the wireless environment conditions and antenna size, it cannot be performed properly in the transmission of low frequency signals. Therefore, high frequency signals are used to transmit low frequency signals in wireless environment. The high-frequency signal used is called a carrier signal, and the transmission of a low-frequency information signal by means of a carrier signal is called modulation. There are 2 types of modulation, Analog Modulation and Digital Modulation [3].

An algorithm for detecting the modulation type when a signal in which modulation type is unknown reaches the receiver has an important place in the technology of communication systems. The algorithms developed for this purpose are called modulation classification algorithms. The lack of prior knowledge of the received signal in communication systems creates a difficult situation for the Modulation method. This method allows analyzing the modulated signal reaching the receiver with the appropriate demodulator. It has seen in literature that Modulation Classification methods are used generally in the military, civil and commercial fields [4]. The location of the modulation classification algorithm in a communication system is shown in Figure 1.2.

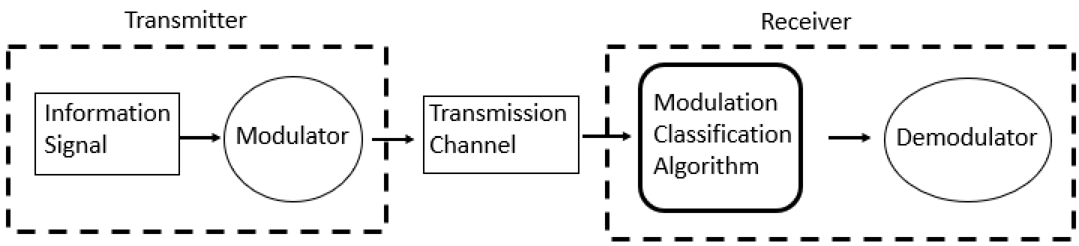

Figure 1.2: Modulation Classification Algorithm in Communication System

Modulation classification methods consist of two approaches: Likelihood-based (LB) and Feature-based (FB). The first method, Likelihood-based approach uses the likelihood function and problems are identified by multiple hypothesis testing [5]. In the Likelihood-based approaches, Average Likelihood Ratio Test (ALRT), Generalized Likelihood Ratio Test (GLRT), Hybrid Likelihood (HLRT) [5] and Quasi-Hybrid Likelihood Ratio Test (QHLRT) [6] are available in the literature. This approach does not provide good performance in the presence of phase and frequency offset [7]. The Second method, Feature-based method can be defined as extracting useful information from an incoming signal that does not have information about it. The feature extraction can be performed in the frequency domain or time domain. In the literature, Statistical extracted from instantaneous amplitude, phase and frequency, spectrum symmetry, wavelet transform, Fourier transform, High order Moments (HOMs) and Cumulants (HOCs), Higher-order Cyclic Cumulants, Very High-order Cumulants, Very Higher-order Statistics (VHOS) are used as features. The performance of these features may vary depending on factors such as signal characteristics, modulation types, channel models. Using the right feature is an important factor.

LB shows the optimal classifier property while FB shows the suboptimal property. Since the complexity of the FB method is lower than LB, it can be said that it is preferred more in practical applications. Also, the LB method is susceptible to nonideal situations and model mismatch. Considering the advantages and disadvantages of FB and LB methods, it can be said that the FB method is more preferred in literature [8].

The extracted features go through a classification process to separate the signals by modulation type. In general, there are many classification algorithms like Artificial Neural Networks (ANN), Support Vector Machines (SVM), decision-tree in the literature.

In [9, 10, 11], intantaneous amplitude, phase and frequency information of the signal were used as feature. Known signal processing techniques were used to extract this characteristic information from the signal and the necessary pre-processing was completed before the information was given to the classification algorithm. It was seen that decision tree or neural network [12] algorithms were used in classification step of the articles used by these features.

In [13], wavelet transform was used as a method to extract features from signals. It was preferred due to low computational complexity. The multi-layer neural network which has increased in popularity recently was used as a classifier [14]. In [15], the phase probability density function of the received signal was acquired using Tikhonov function and the suboptimal classifier was improved. In [16], the MPSK signal was represented using the Fourier series expansion unlike the paper was used Tikhonov function. The asymptotic optimal classification algorithm is examined using the theoretical approach. In both, the signals were simulated in the AWGN channel.

In [17], An AMC was developed using the HOS as a solution of estimation of the blind channel and pattern recognition in the Multipath channel. In [18] and [19], it was used the 4th order cumulant for feature extraction step without the channel estimation coefficient. An important difference between the two articles is that [19] uses MIMO systems. In [20]. In addition to the 4th order cumulant, the 6th cumulant was added and the results were compared. The multipath channel is selected as the channel. Similarly, the features were extracted with 6th cumulant [21]. Signals have been classified with SCAEs, which is a learning method (Stacked convolutional auto-encoders). In [22], the classification performance was improved by using the algorithm generated by creating a relationship between 4th order cumulants. In [23], instantaneous values and higher-order statistics such as cumulants and moments were used to generate features. In the classification of signals, different neural networks such as MLP neural networks, radial basis function neural networks, probabilistic neural networks, Multi-class SVM based classifiers were used. In [24], second, fourth and sixth order cumulants were used. Hierarchical polynomial structure showed a high performance in contrast to conventional polynomial classifer. In [25], the performance of the Modulation classification algorithm under the conditions of SUMC (Sigle user modulation classification) and MUMC (Multi user modulation classification) were examined using fourth and sixth order cumulants. In [26], the fourth cumulants were used and it is aimed to achieve higher performance at low SNR using the hierarchical scheme. In [27], it has used the correlation function, cyclic cumulants (CC), and cumulants for the feature extraction step. The extracted features were examined to classify using the tree-based algorithm. As the types of antenna, it was used SISO, SIMO, MISO, MIMO in frequency selective fading channel. Also, in terms of the algorithm which is simulated in the computer environment, this article is important. In [28], the statistical features are extracted from the received signals have classified with deep neural network method (DNN). Because DNN provides an advantage for a complex structure a fading channel. The support vector machine method which is one of machine learning methods has been used for classification in [29]. The importance of frequency offset and fading problems have been emphasized. To solve this problem, the new algorithm has been improved. In [30], it was used the correlation function for the received signal. A blind modulation classification algorithm was performed using SISO, SIMO, MISO, MIMO with frequency selective channel.

In this thesis, a modulation classification algorithm has been developed in order to determine which modulation type the signal is modulated. Modulation types used in the developed algorithm include Mary-QAM (Quadrature Amplitude Modulation), i.e. 16-QAM, 32-QAM, 64-QAM, 128-QAM, and 256-QAM, GMSK (Gaussian Minimum Shift Keying) and OQPSK (Offset Quadrature Phase-shift Keying) as digital modulation and AM-DSB-WC (Double Sideband with Carrier), AM-DSB-SC (Double Sideband Suppressed Carrier), AM-SSB-WC (Single Sideband with Carrier), AM-SSB-SC (Single Sideband Suppressed Carrier) and FM (Frequency Modulation) as analog modulations. Firstly, features that contain the characteristics of the modulated signals were extracted from the signals. Afterward, the signals were separated according to the modulation type by certain classification methods. Success rates were examined. Then the results were presented.

According to these purposes, the thesis was organized as follows; In the first chapter, a general infrastructure of communication systems, modulation classification, and analysis of the information obtained in the literature review were presented. In the second chapter, the topics were divided into 4 main topics. The characteristics of the data used in the study were defined under the subtitle ”Description of Data”. The structure of the signals in this data was examined under the subtitle ”Signal Model”. The structure of Modulation types with equations was examined under the subtitle ”Modulation Types”. The characteristic information extracted from the signals to distinguish with which modulation type the signals are modulated is explained together with the equations under the subtitle ”Feature Extraction”. Then, The classification of signals according to the modulation type and used classification algorithms were described under the subtitle ”Modulation classification”. In the third chapter, the details of the study were presented and the performance of the modulation classification algorithm according to different SNR values and different modulation types were examined.

CHAPTER 2

METHOD

This chapter details the modulation classification algorithm created and includes 4 subtitles. Under the title of ”Description of Data”, the content of the data is explained with details included the modulation types, channel effect, SNR values, and lengths of the signals. Under the title ”Signal Model”, the information obtained from the important parts of the signal is given in order to analyze the signals in the I and Q domains. Under the title of ”Modulation Types,”, The structure of the modulations which are used in the algorithm were presented. These include Mary-QAM (Quadrature Amplitude Modulation), i.e. 16-QAM, 32-QAM, 64-QAM, 128-QAM, and 256-QAM, GMSK (Gaussian Minimum Shift Keying) and OQPSK (Offset Quadrature Phase-shift Keying) as digital modulation and AM-DSB-WC (Double Sideband with Carrier), AM-DSB-SC (Double Sideband Suppressed Carrier), AM-SSB-WC (Single Sideband with Carrier), AM-SSB-SC (Single Sideband Suppressed Carrier) and FM (Frequency Modulation) as analog modulations. Under the title of ”Feature extraction”, it is described that the details of how to create a feature space by using the important information obtained from the signals in a smaller and summarized format. It is used that Statistical features, extracted from instantaneous amplitude, phase and frequency, Higher-order Moments (HOMs) and Higher-order Cumulants (HOCs) as a feature. Under the title of ”Modulation Classification”, the modulation classification algorithms which are used to analyze an unknown signal and to place it in the correct class were presented. The details of the support vector machine algorithm used in this thesis, to examine at the correct classification performance of the signals are given.

2.1 Description of Data

The dataset includes recorded signals as over-the-air. Synthetic simulated channel effects are added to these signals. In addition to these channel distortions, using Software Defined Radio (SDR) was used to generate and transmit the modulated signal and then to receive back the signals in the lab which indoor wireless channel on the 900MHz ISM band. Data is available on the RadioML website [31]. The signals are generated in 2 layers as I and Q domains. The number of samples of the signals is 1024 and in the next step feature extraction, the features will be extracted with one information per 1024 samples. Tha data was used in [32]. Signals generated by synthetic simulated channel effects were recorded with an SNR of 0 dB to 20 dB. There are 45056 signals for each modulation type. For only one dB value, there are 4096 signals. A simple concept for one modulation is given in Figure 2.1.

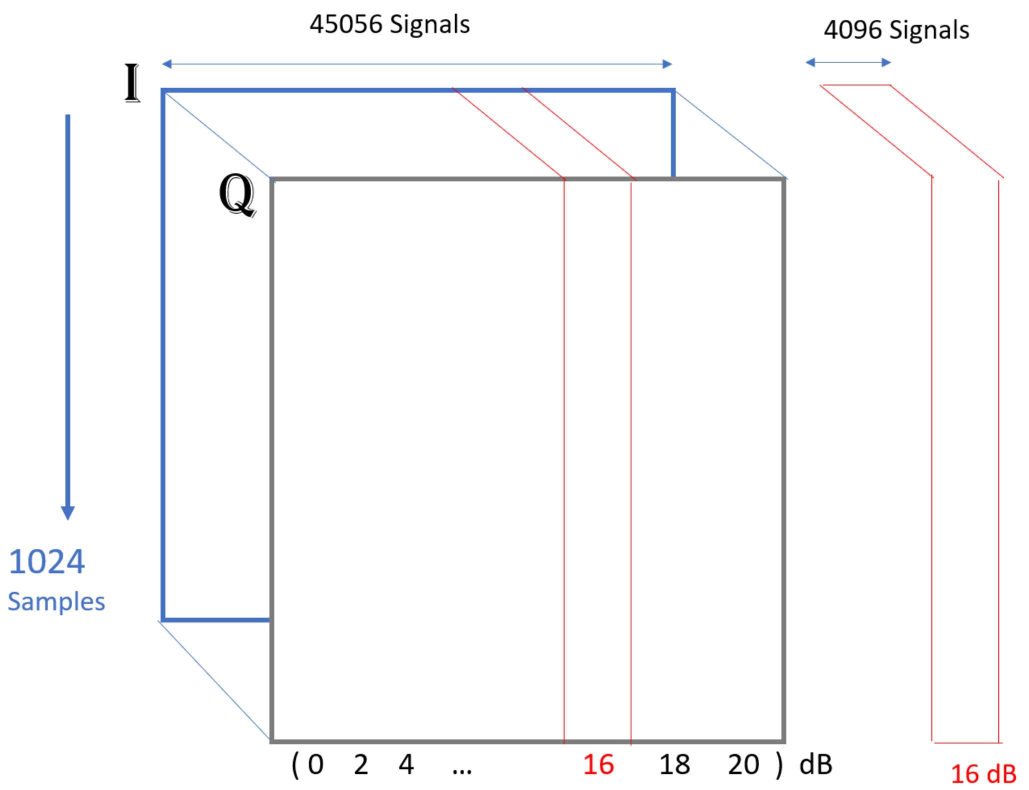

Figure 2.1: A Simple Concept of Data for One Modulation

Signals modulated by analog and digital modulation types. Data includes Higher-order Modulation types such as 128-QAM and 256-QAM. Modulation types in this data set are given in Figure 2.2 .

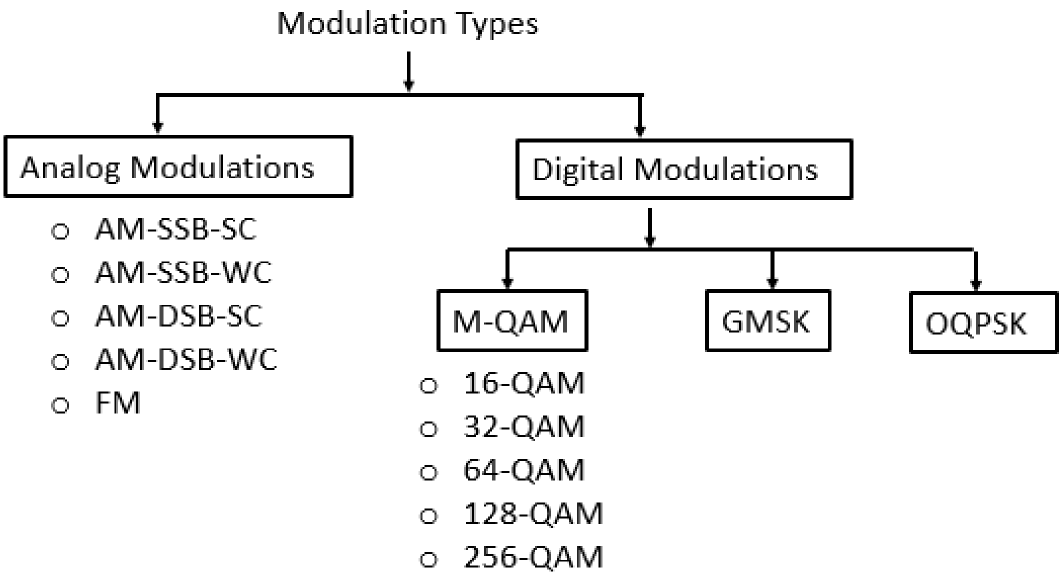

Figure 2.2: The Used Modulations For Algorithm

2.2 Signal Model

A complex signal which expresses the output of the Hibert Filter for communication systems is expressed as in Equation 2.1. $I(n)$ represents the real part of the complex signal, while $Q(n)$ represents the imaginary part of the complex signal. $I(n)$ and $Q(n)$ are the instantaneous in-phase and quadrature-phase components, respectively.

$$s(n) = I(n) + jQ(n) \tag{2.1}$$

The instantaneous amplitude $a(n)$, phase $\varphi(n)$ and frequency $f(n)$ informations obtained from the complex signal is used to extract feature. Initially, amplitude and phase signal characteristics have been used to obtain information from signals, and frequency information has recently been utilized. $a(n)$, $\varphi(n)$ and $f(n)$ the characteristics are given in Equation 2.2, 2.3 and 2.4 , respectively.

$$a(n) = \sqrt{I^2(n) + Q^2(n)} \tag{2.2}$$

$$\phi(n) = tan^{-1}[\frac{Q(n)}{I(n)}] \tag{2.3}$$

$$f(n) = \frac{1}{2\pi} \frac{\phi(n) - \phi(n-1)}{\Delta n} \tag{2.4}$$

The instantaneous amplitude, frequency and phase values obtained are used for feature extraction after being centered to avoid the system bias. The centered operation of these values is performed by subtracting the mean value. Centered the instantaneous amplitude $a_c(n)$, frequency $f_c(n)$ and phase $\varphi_{cnl}(n)$ values are given in Equation 2.5, 2.6 and 2.8 respectively.

$$a_c(n) = a(n) - \mu_a \tag{2.5}$$

$$f_c(n) = f(n) - \mu_f \tag{2.6}$$

Unlike the amplitude and frequency characteristic, the instantaneous phase $\phi_{nl}(n)$ characteristic is subtracted from the linear component before the mean subtraction. Linear component subtraction operation for phase characteristic is given 2.7.

$$\phi_{nl}(n) = \phi(n) - 2\pi\mu_f(n)\Delta_t \tag{2.7}$$

$$\phi_{cnl}(n) = \phi_{nl}(n) - \mu_{\phi_{nl}} \tag{2.8}$$

where n is the sample number of the signal.

Preliminary operations were presented in order to use the instantaneous amplitude, frequency and phase values of the signals efficiently in the feature extraction section.

2.3 Modulation Types

The transmission of low-frequency signals by means of high-frequency signals is called modulation. There are two types of analog and digital modulation in communication systems. Nowadays, digital modulation is used more than analog modulation. This is because they are more resistant to noise and are more suitable for coding and cryptography. Thanks to repeaters, they are more successful in transmitting data over longer distances. However, although it is thought that there is a transition to digital modulation in practice, the use of analog modulation cannot be ignored [3].

2.3.1 Analog Modulation Types

The change of the carrier signal characteristics according to the information signal is called analog modulation. If the amplitude of the carrier signal varies with the information signal, this type of modulation is called AM. If the frequency of the carrier signal varies with the information signal, this type of modulation is called FM. If the phase of the carrier signal varies with the information signal, this type of modulation is called Phase Modulation (PM) The main types of analog modulation are given in Figure 2.3.

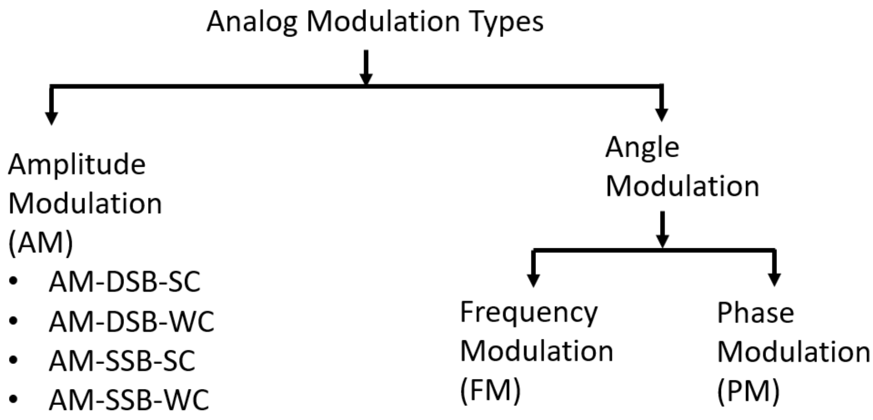

Figure 2.3: Analog Modulation Types

2.3.1.1 Amplitude Modulation (AM)

If the amplitude of the carrier signal $c(t)$ varies with the information signal $m(t)$, this type of modulation is called AM. It is a linear modulation. AM includes Single-Sideband Amplitude Modulation (AM-SSB) and Double-Sideband Amplitude Modulation (AM-DSB). In amplitude modulation with no high bandwidth, more signals are transmitted in a narrow space. At the same time, because the bandwidth is narrow, it can be described as less complex and cost-effective. However, in wireless communication, it is susceptible to noise and disturbing effects in transmission. The carrier signal with amplitude value A_c and frequency value f_c is given in Equation 2.9

$$c(t) = A_c cos(2\pi f_c t) \tag{2.9}$$

AM signal waveform acquires with the multiplication of the carrier signal and message signal. AM signal waveform $s(t)$, which the amplitude of the carrier signal was changed according to the information signal and the frequency was fixed, is given Equation 2.10

$$s(t) = A_c[1 + k_a m(t)]cos(2\pi f_c t) \tag{2.10}$$

where k_a is constant which called the amplitude sensitivity. The waveform in the frequency domain which acquired using AM waveform trigonometric transformations is given in Equation 2.11 is obtained. The $fm - fc$ and $fm + fc$ frequencies represent the sidebands. If it is present in two sidebands, it is called AM-DSB modulation. Otherwise, if one of the sidebands is suppressed, it is called AM-SSB.

$$S(f) = \frac{A_c}{2}A_c[\delta(f - f_c) + \delta(f + f_c)] + \frac{k_a A_c}{2}A_c[\delta(f - f_c) + \delta(f + f_c)] \tag{2.11}$$

In AM modulation, there are two types of modulation, the carrier of which is pressed or not pressed. The carrier of Double Sideband - Suppressed Carrier Modulation (AM-DSB-SC) and Single Sideband - Suppressed Carrier Modulation (AM-SSB-SC) are suppressed. The unsuppressed version is called Double Sideband - with Carrier (AM-DSB-WC) and Single Sideband - with Carrier (AM-SSB-WC) Modulations. The characteristics of the information signal to be sent in the communication environment are available in the sidebands of the modulated signal. Therefore, modulation types with suppressed carrier information are more efficient to use.

2.3.1.2 Frequency Modulation (FM)

If the frequency of the carrier signal varies with the information signal, modulation is called FM. FM has been developed because AM is susceptible to noise and disturbing effects. It is the nonlinear modulation. The structure of the FM signal is more complex than the AM signal. The equation for FM signal is given in Figure 2.12

$$s(t) = A_c = cos[2\pi f_c t + \beta sin(2\pi f_m t)] \tag{2.12}$$

where β is modulation index. The frequency of the carrier signal is called the central frequency and the sidebands are formed by changes in frequency. The number of sidebands depends on the modulation index. FM signal in frequency domain is given in Equation 2.13.

$$S(f) = \frac{A_c}{2} \sum_{n=-\infty}^{\infty} J_n(\beta)[\delta(f - f_c - nf_m) + \delta(f + f_c + nf_m)] \tag{2.13}$$

where $J_n(\beta)$ is Bessel function. FM needs higher bandwidth due to the sidebands it has. This analysis is done with the Bessel function, since the calculation of these sidebands leads to complexity.

2.3.2 Digital Modulation Types

Digital modulation techniques are used to make digital signals suitable for communication channels. It is used more because of its advantages such as noise resistance, allowing for more data transmission and information security. The modulation is the changing of the amplitude, frequency, and phase of the carrier signal depending on the information signal. The main types of digital modulation are given in Figure 2.4

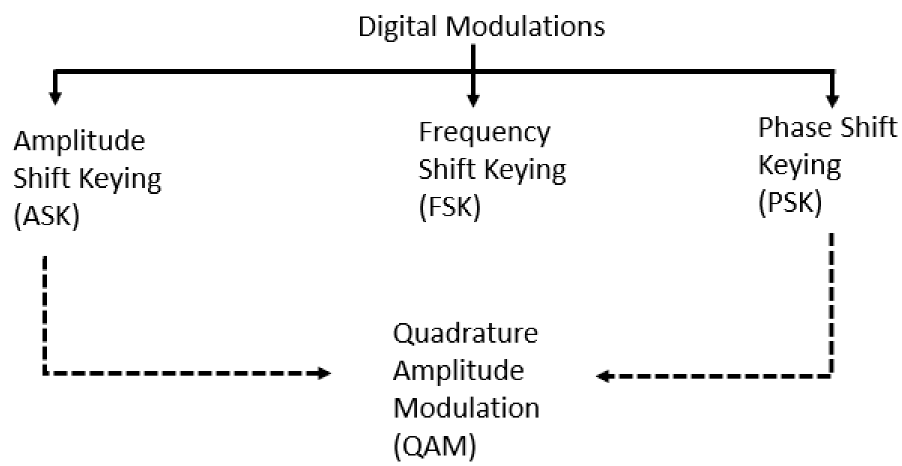

Figure 2.4: Digital Modulation Types

The amplitude shift of a sinusoidal carrier signal relative to the information signal is called Amplitude Shift Keying (ASK). The frequency shift of a sinusoidal carrier signal relative to the information signal is called Frequency Shift Keying (FSK). The phase switching of a sinusoidal carrier signal relative to the information signal is called Phase Shift Keying (PSK). If both the amplitude and the phase of the carrier signal are switched together, it is called Quadrature Amplitude Modulation (QAM) [1].

2.3.2.1 Gaussian Minimum Shift Keying (GMSK)

Unlike MSK modulation that produces a half sinusoidal pulse, a narrower main lobe and fewer side lobes are obtained using the Gaussian pulse shape in the GMSK. This provides spectral efficiency and bandwidth advantages. GMSK, which is a form of continuous phase modulation type, is a member of the MSK family. MSK can be defined as a special case of FSK modulation. Unlike standard FSK, the disadvantage of the wide sidebands that occurs in FSK has been tried to be reduced in MSK and GMSK. The use of wireless communication is quite high because MSK is resistant to noise and disturbing effects and has an intense power spectrum [33].

In order to reduce the sidebands occurring in the MSK, the signal is passed through a Gaussian filter $G(f)$ which is an ideal filter to obtain GMSK. GMSK has a wide range of second and third-generation cellular communication systems thanks to its narrow bandwidth advantage. Although GMSK is passed through a nonlinear filter, it does not suffer deterioration. In addition, the signal is more resistant to noise because the information is not stored in amplitude. These are the other advantages of GMSK modulation.

A more intense power spectrum causes the symbols to mix. This mixing is called the Inter Symbol Interference (ISI). In this modulation type, ISI is allowed to obtain a more intense power spectrum. For GMSK modulation [34], transmitter structure is given in Figure 2.5

Figure 2.5: Trasmitter Structure for GMSK

The output of the Gaussian Filter is given in Equation 2.15.

$$\sum_{k=-\infty}^{\infty} a_k p_c(t - kT) \tag{2.14}$$

where a_k are the binary information symbols and $p_c(t)$ is the rectangular pulse response of the Gaussian filter. And, for the output of the modulator, GMSK signal expression is given in Equation 2.15.

$$s(t) = cos\left(2\pi f_c t + \frac{\pi\beta}{T} \int_{-\infty}^{t} \left(\sum_{k=-\infty}^{\infty} b_k p(\mu - kT) \right) d\mu \right) \tag{2.15}$$

where β is modulation index and equals to 0.5, and $p(t)$ is the truncated pulse.

2.3.2.2 Offset Quadrature Phase-shift Keying (OQPSK)

OQPSK modulation, which is the same structure as the QPSK modulation structure, has taken place in the literature. In the QPSK, the pulses in the I and Q domain simultaneously change, while in OQPSK, any of the pulses are delayed, resulting in a new modulation type. This difference may observe for QPSK in Figure 2.6 and for OQPSK in Figure 2.7. The main purpose of this process is to prevent zero transitions. That is, phase changes in both the I and Q domains do not occur. The phase change in QPSK is at most 180 degrees, whereas in OQPSK it is 90 degrees.

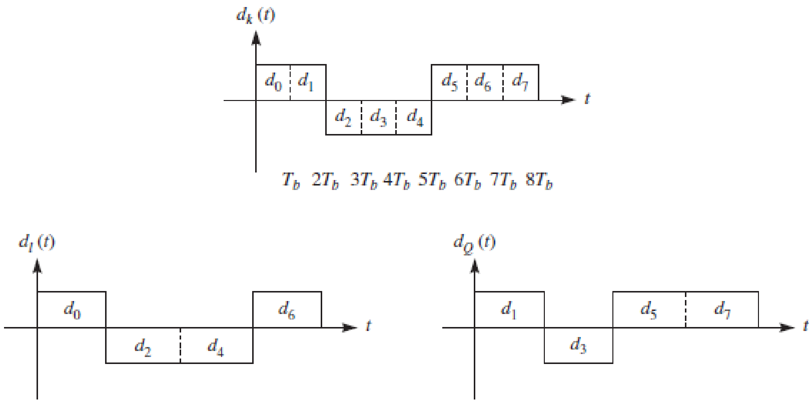

Figure 2.6: The in-phase and quadrature components for QPSK [1]

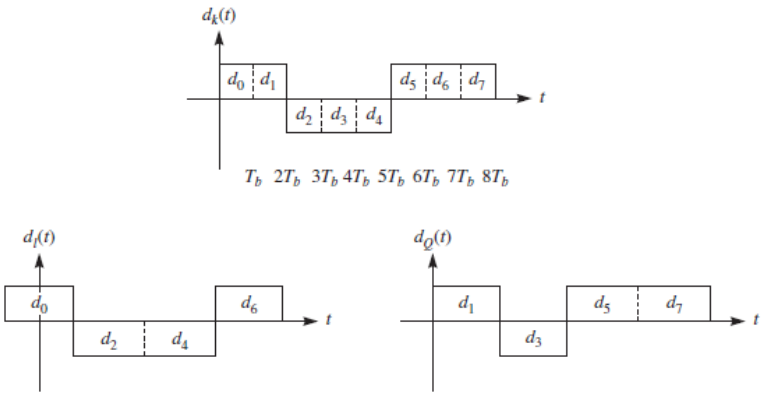

Figure 2.7: The in-phase and quadrature components for OQPSK [1]

Sudden phase shifts are observed 2 times more in QPSK modulation. It uses 4 different values of the information to be transmitted with phase information. OQPSK signal equation is given in Equation 2.16

$$s(t) = A\left[\left(\sum_{n=-\infty}^{\infty} I_{2n} g(t-2nT)\right) cos 2\pi f_c t + \left(\sum_{n=-\infty}^{\infty} I_{2n+1} g(t-2nT-T))\right) sin 2\pi f_c t\right] \tag{2.16}$$

2.3.2.3 M-ary Quadrature Amplitude Modulation (M-QAM)

It is produced to provide wide bandwidth in wireless communication systems. The amplitude and phase value of the two carrier signals which have 90 degrees phase difference at the same frequency is taken as output. One of the signals is called I and the other is called Q. Mathematically, one of the signals is expressed with the sine and the other with the cosine wave. Combined with the modulated signals, the transmission is realized. One of the advantages is that QAM transmits more data than standard amplitude and phase switching. The transmitted M-ary QAM signal is defined in Equation 2.18

$$s_k(t) = \sqrt{\frac{2E_0}{T}}a_k cos(2\pi f_c t) - \sqrt{\frac{2E_0}{T}}b_k sin(2\pi f_c t) \qquad (2.17)$$

where E_0 is the energy of the signal with the lowest amplitude, a_k and b_k refer to amplitude values. M in M-QAM expression refers to the number of conditions a digital signal carries according to its characteristics. M refers to the number of conditions a digital signal carries according to its characteristics. It also refers to the situation that occurs when more than one bit of data is sent in a transmission. This has the advantage of transmitting more data to the receiver in one transmission. The number of bits of data to be sent under certain conditions is defined in Equation 2.18

$$N = \log_2 M \qquad (2.18)$$

where N is the number of bits and M is the number of conditions. Mary-QAM creates different modulation types in order to transmit different numbers of symbols in data transmission. The constellation diagrams of 16-QAM, 32-QAM, 64-QAM, 128-QAM and 256-QAM , which are widely used in literature, are shown in Figure 2.8, 2.9, 2.10, 2.11, and 2.12, respectively.

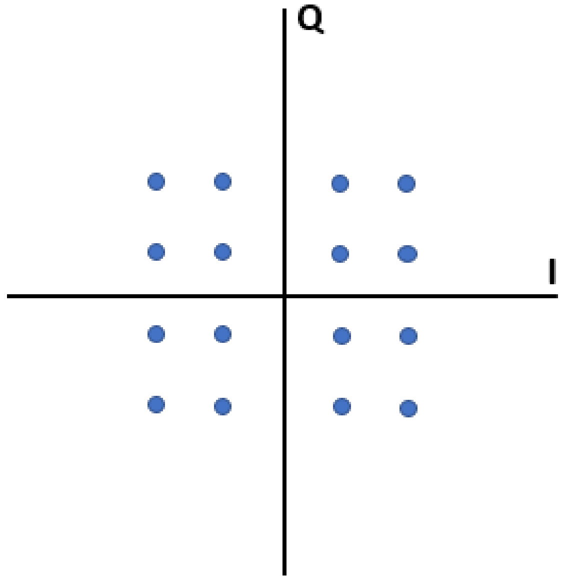

Figure 2.8: 16-QAM Constellation Diagram

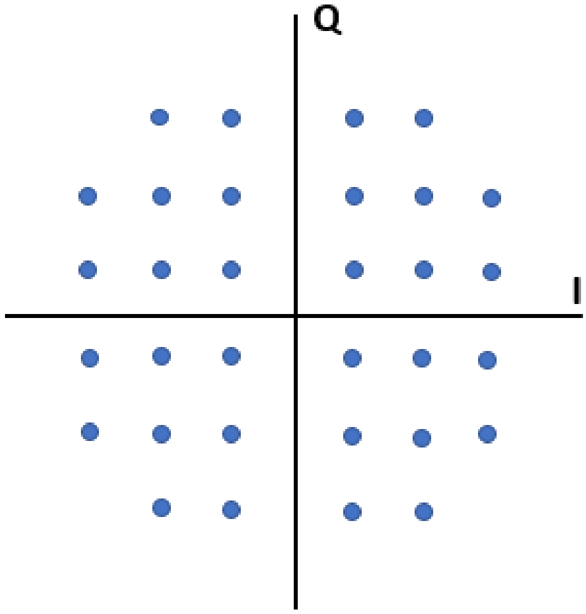

Figure 2.9: 32-QAM Constellation Diagram

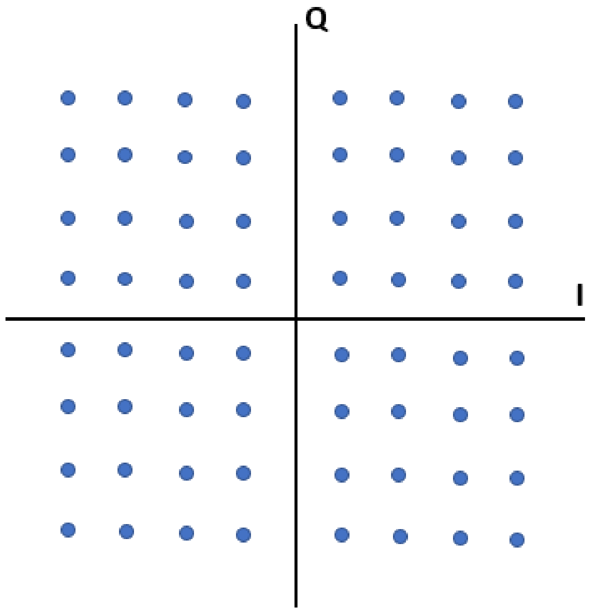

Figure 2.10: 64-QAM Constellation Diagram

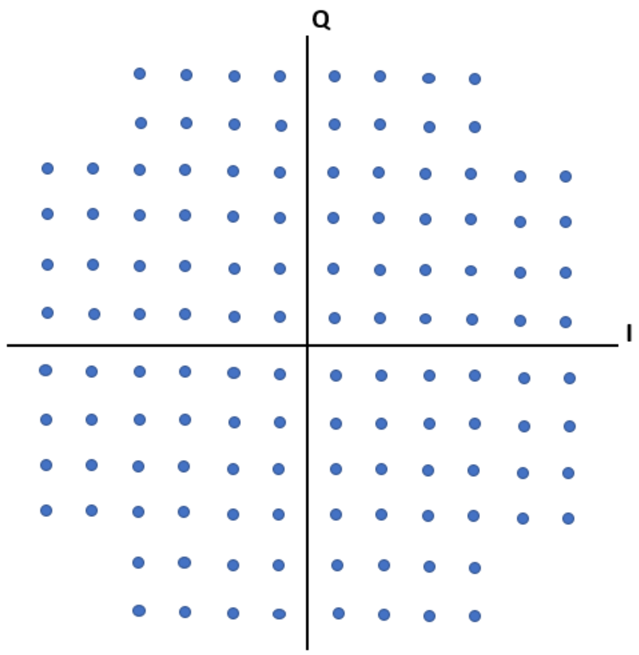

Figure 2.11: 128-QAM Constellation Diagram

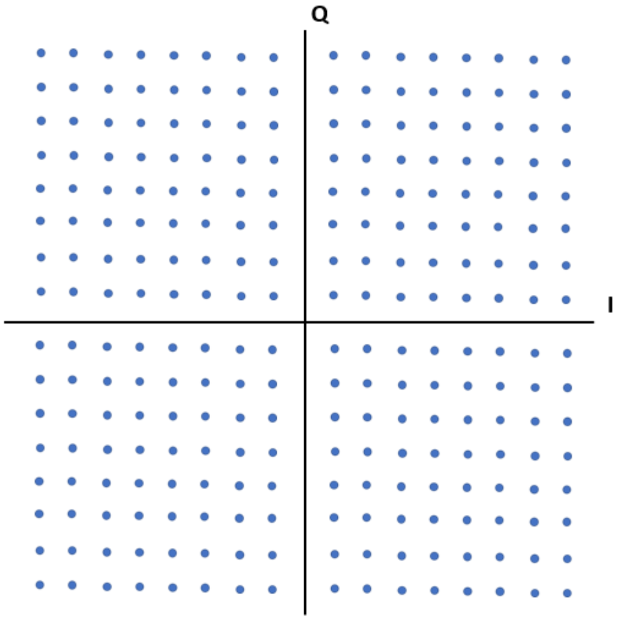

Figure 2.12: 256-QAM Constellation Diagram

2.4 Feature Extraction

The purpose of feature extraction, which is a process in modulation classification, is to reduce the dimension of the information in the existing data set and create a smaller set of data that can summarize the information in a precise way. In this way, feature space dimension and computational complexity reduce thanks to features derived using signal characteristics, and the information is thus easier to manage and process. A feature vector was created with the information extracted from the signals and prepared for the modulation classification process. Statistical features extracted from instantaneous amplitude, phase and frequency, Higher-order Moments (HOMs) and Higher-order Cumulants (HOCs) were used in this study as a feature.

2.4.1 Statistical Features

The mean μ, variance σ^2, kurtosis μ_{42}, skewness γ are used as Statistical features. These features are extracted using the amplitude frequency and phased characteristics of the signal. The first moment of a distribution is mean, the second moment is variance, the third is skewness, and the fourth one is kurtosis. Unlike mean and variance in Equation 2.19, 2.20, kurtosis and skewness are used as normalized features with standard deviation in modulation classification applications.

$$\mu_x = \frac{1}{N_x}\sum_{k=1}^{N_x} x(k) \tag{2.19}$$

$$\sigma_x^2 = \frac{1}{N_x}\sum_{k=1}^{N_x}[x(k) - \overline{x}]^2 \tag{2.20}$$

Kurtosis and skewness statistics are less common than mean and variance, but they are important criteria when interpreting a data distribution, although the mean and variance are common in signal processing applications. Kurtosis and skewness feature help in executing ideas about data distribution. Kurtosis relates to the weight of a data distribution tail. Skewness examines whether data distribution is symmetry. Although these two concepts are thought to have a similar relationship, they should not be confused one another. Skewness can take negative and positive values because it is the 3rd moment, and kurtosis is the 4th moment, it only takes positive values. The formulas of kurtosis and skewness are given Equation 2.21, 2.22, respectively.

$$\gamma_x = \frac{1}{\sigma_x^3 N_x}\sum_{k=1}^{N_x}[x(k) - \overline{x}]^3 \tag{2.21}$$

$$\mu_{42} = \frac{1}{\sigma_x^4 N_x}\sum_{k=1}^{N_x}[x(k) - \overline{x}]^4 \tag{2.22}$$

where $\overline{x}$ is the mean of $x(k)$. These features can be found for intantaneous amplitude, frequency and phase, depending on the use case. So, $x(k)$ can be $a_c(n)$, y $f_c(n)$ or $\varphi_{cnl}(n)$ value. HOSs features are given in Table 2.4

Table 2.1: Statistical Features

μ_a	Mean of the instantanous centered amplitude of the complex signal
μ_f	Mean of the instantanous centered frequency of the complex signal
μ_p	Mean of the instantanous centered nonlinear phase of the complex signal
σ_a^2	Variance of the instantanous centered amplitude of the complex signal
σ_f^2	Variance of the instantanous centered frequency of the complex signal
σ_p^2	Variance of the instantanous centered nonlinear phase of the complex signal
γ_a	Skewness of the instantanous centered amplitude of the complex signal
γ_f	Skewness of the instantanous centered frequency of the complex signal
γ_p	Skewness of the instantanous centered nonlinear phase of the complex signal
μ_{42a}	Kurtosis of the instantanous centered amplitude of the complex signal
μ_{42f}	Kurtosis of the instantanous centered frequency of the complex signal
μ_{42p}	Kurtosis of the instantanous centered nonlinear phase of the complex signal

2.4.2 Higher-order Moments (HOMs)

HOMs are preferred in signal processing and modulation classification applications for feature extraction process because they are less affected by noise which created by multi-fading and AWGN channel. Moments are the generalized state of the expected value which is changed according to the power of the moment in Equation 2.23.

$$M_{pq} = E[s^{p-q}(s^*)^q] \tag{2.23}$$

where p is the order of the moment and q is the power of the signal's conjugate. In the same way, the general formula for HOMs is given in Equation 2.24.

$$M_{pq} = \frac{1}{N}\sum_{n=1}^{N} s[n]^{p-q} s[n]^{*q} \tag{2.24}$$

where N is the number of samples and $s[n]$ is complex signal and $s[n] = s[1], s[2], \ldots, s[N]$. The HOMs were used both as a feature and to obtain cumulants. The used HOMs are given in Table 2.2.

Table 2.2: Higher-order Moments

2nd order moments	M_{20}, M_{21}
4th order moments	M_{40}, M_{41}, M_{42}, M_{43}
6th order moments	M_{60}, M_{61}, M_{62}, M_{63}
8th order moments	M_{80}, M_{81}, M_{82}, M_{83}, M_{84}

2.4.3 Higher-order Cumulants (HOCs)

HOCs resist to the effect of the multi fading channel or AWGN. Therefore, the use of HOCs has become more important in literature. HOCs are the statistical characteristic of random variables such as HOMs. The used HOCs are given in Table 2.3

Table 2.3: Higher-order Moments

2nd order cumulants	C_{20}, C_{21}
4th order cumulants	C_{40}, C_{41}, C_{42}
6th order cumulants	C_{60}, C_{61}, C_{62}, C_{63}
8th order cumulants	C_{80}, C_{81}, C_{82}, C_{83}, C_{84}

HOCs are produced using HOMs. HOCs are given in Table 2.5 with calculations [35].

Table 2.4: Higher order cumulants (HOCs) expression - 1

$C_{s,2,0}$	$E_{s,2,0} = E[x^2.(\overline{x})^0]$
$C_{s,2,1}$	$E_{s,2,1} = E[x^1.(\overline{x})^1]$
$C_{s,4,0}$	$E_{s,4,0} - 3.(E_{s,2,0})^2$
$C_{s,4,1}$	$E_{s,4,1} - 3.E_{s,2,0}.E_{s,2,1}$
$C_{s,4,2}$	$E_{s,4,2} - (E_{s,2,0})^2 - 2(E_{s,2,1})^2$
$C_{s,6,0}$	$E_{s,6,0} - 15E_{s,2,0}E_{s,4,0} + 30(E_{s,2,0})^3$
$C_{s,6,1}$	$E_{s,6,1} - 10E_{s,2,0}E_{s,4,1} - 5E_{s,2,1}E_{s,4,0} + 30(E_{s,2,0})^2E_{s,2,1}$
$C_{s,6,2}$	$E_{s,6,2} - E_{s,2,0}E_{s,4,0} - 8E_{s,2,1}E_{s,4,1} - 6E_{s,2,0}E_{s,4,2} + 6(E_{s,2,0})^3 + 24(E_{s,2,1})^2E_{s,2,0}$
$C_{s,6,3}$	$E_{s,6,3} - 6E_{s,2,0}E_{s,4,1} - 9E_{s,2,1}E_{s,4,2} + 18(E_{s,2,0})^2E_{s,2,1} + 12(E_{s,2,1})^3$
$C_{s,8,0}$	$E_{s,8,0} - 35(E_{s,4,0})^2 - 630(E_{s,2,0})^4 + 420(E_{s,2,0})^2(E_{s,4,0})$
$C_{s,8,1}$	$E_{s,8,1} - 35E_{s,4,0}E_{s,4,1} - 630(E_{s,2,0})^3E_{s,2,1} + 210E_{s,4,0}E_{s,2,0}E_{s,2,1} + 210E_{s,2,1}E_{s,4,1}$
$C_{s,8,2}$	$E_{s,8,2} - 15E_{s,4,0}E_{s,4,2} - 20(E_{s,4,1})^2 + 30E_{s,4,0}(E_{s,2,0})^2 + 60E_{s,4,0}(E_{s,2,1})^2 + 240E_{s,4,1}E_{s,2,1}E_{s,2,0} + 90E_{s,4,2}(E_{s,2,0})^2 - 90(E_{s,2,0})^4 - 540(E_{s,2,0})^2(E_{s,2,1})^2$
$C_{s,8,3}$	$E_{s,8,3} - 5E_{s,4,0}E_{s,4,1} - 30E_{s,4,1}E_{s,4,2} + 90E_{s,4,1}(E_{s,2,0})^2 + 120E_{s,4,1}(E_{s,2,1})^2 + 180E_{s,4,2}E_{s,2,1}E_{s,2,0} + 30E_{s,4,0}E_{s,2,0}E_{s,2,1} - 270(E_{s,2,0})^3E_{s,2,1} - 360(E_{s,2,1})^3E_{s,2,0}$
$C_{s,8,4}$	$E_{s,8,4} - (E_{s,4,0})^2 - 18(E_{s,4,2})^2 - 16(E_{s,4,1})^2 - 54(E_{s,2,0})^4 - 144(E_{s,2,1})^4 - 432(E_{s,2,0})^2(E_{s,2,1})^2 + 12E_{s,4,0}(E_{s,2,0})^2 + 96E_{s,4,1}E_{s,2,1}E_{s,2,0} + 144E_{s,4,2}(E_{s,2,1})^2 + 72(E_{s,4,2})(E_{s,2,0})^2 + 96E_{s,4,1}E_{s,2,0}E_{s,2,1}$

Table 2.5: Higher order cumulants (HOCs) expression - 2

C_{20}	M_{20}
C_{21}	M_{21}
C_{40}	$M_{40} - 3M_{20}^2$
C_{41}	$M_{41} - 3M_{20}M_{21}$
C_{42}	$M_{42} - M_{20}^2 - 2M_{21}^2$
C_{60}	$M_{60} - 15M_{20}M_{40} + 30M_{20}^3$
C_{61}	$M_{61} - 10M_{20}M_{41} - 5M_{21}M_{40} + 30M_{20}^2M_{21}$
C_{62}	$M_{62} - M_{20}M_{40} - 8M_{21}M_{41} - 6M_{20}M_{42} + 6M_{20}^3 + 24M_{21}^2M_{20}$
C_{63}	$M_{63} - 6M_{20}M_{21} - 9M_{21}M_{42} + 18M_{20}^2M_{21} + 12M_{21}^3$
C_{80}	$M_{80} - 35M_{40}^2 - 630M_{20}^2 + 420M_{20}^2M_{40}$
C_{81}	$M_{81} - 35M_{40}M_{41} - 630M_{20}^3M_{21} + 210M_{40}M_{20}M_{21} + 210M_{20}M_{41}$
C_{82}	$M_{82} - 15M_{40}M_{42} - 20M_{41}^2 + 30M_{40}M_{20}^2 + 60M_{40}M_{21}^2 + 240M_{41}M_{21}M_{20} + 90M_{42}M_{20}^2 - 90M_{20}^4 + 540M_{20}^2M_{21}^2$
C_{83}	$M_{83} - 5M_{40}M_{41} - 30M_{41}M_{42} + 90M_{41}M_{42} + 120M41M_{21}^2 + 180M_{42}M_{21}M_{20} + 30M_{40}M_{20}M_{21} - 270M_{20}^3M_{21} - 360M_{21}^3M_{20}$
C_{84}	$M_{84} + M_{40}^2 - 18M_{42}^2 - 16M_{41}^2 - 54M_{20}^4 - 144M_{21}^4 - 432M_{20}^2M_{21}^2 + 12M_{40}M_{20}^2 + 966M_{41}M_{21}M_{20} + 144M_{42}M_{21}^2 + 72M_{42}M_{20}^2 + 72M_{42}M_{21}^2 + 96M_{41}M_{20}M_{41}$

When the power of cumulants is increased to $\frac{2}{p}$, a reduction in processing time occurs when used in machine learning algorithms. This reduction provides a significant advantage in practice. For example, normalization for C_{62} is given in Equation 2.25.

$$\widehat{C_{62}} = C_{62}^{\frac{1}{3}} \tag{2.25}$$

Magnitude values of the complex moment and cumulant features are created to use effectively and to protect against phase-shifting in the classification algorithms.

2.5 Modulation Classification

To distinguish between modulated signals, the characteristic information of each modulation class must be known. Because an appropriate demodulator must be used in the system to regain the information signal at the receiver. If the modulation type of the signal is not detected correctly and an incorrect signal is processed, the demodulator may be damaged. For these reasons, the correct functioning of modulation classification algorithms is extremely important. For the classification algorithms, it is easy for some modulation types to distinguish signals according to modulation types and difficult for others. For example; The higher-order of the modulation type, the more difficult it is to distinguish. 64-QAM and 128-QAM modulations are more difficult to separate, while BPSK and QPSK modulations are easier to separate. The other important thing about correct working is the channel effect. It has a significant impact on the success rate of modulation classification algorithms. Separating signals becomes more difficult, especially when there is a complex channel effect. It is seen in the literature that the AWGN channel is used more in classification algorithms. However, algorithms are also being developed for multipath channels.

It is seen in the literature that machine learning algorithms and deep learning are widely used in modulation classification. Machine learning algorithms can be summarized under two headings as supervised and unsupervised learning [36]. In the Supervised Method, the machine is trained with specific examples and labels for the data reserved for the train. In other words, it knows and learns the signal is modulated with which modulation type. It then separates the signals in the test data according to their class during the test process. Supervised algorithms include Support vector machine, Neural network, Design tree. A system model for supervised learning is given in Figure 2.13.

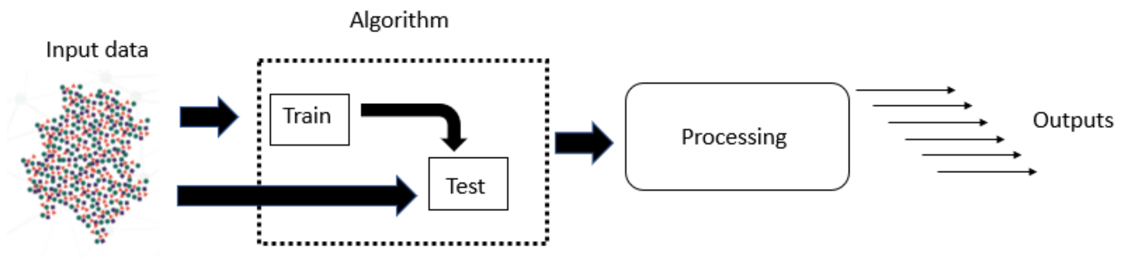

Figure 2.13: A System Model for Supervised Learning

In the Unsupervised method, the machine estimates the results without having been trained with data that have a label. Unsupervised algorithms include Cluster algorithms, K-means, Hierarchical clustering. A system model for supervised learning is given in Figure 2.14

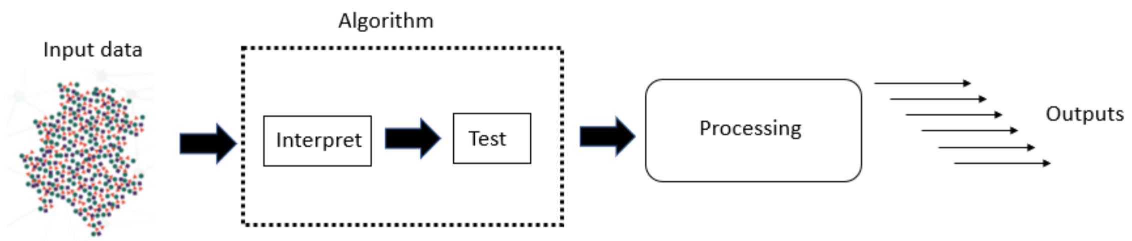

Figure 2.14: A System Model for Unupervised Learning

Supervised learning is more preferred in applications. Unsupervised learning is often used as a preprocessing of supervised algorithms. Studies on unsupervised algorithms are increasing in the literature.

One of the Supervised learning, the neural network algorithm consists of 3 stages. The first one is Feature Extraction to determine a characteristic structure based on the modulation type of the signals obtained in the receiver. The second one is a training process for the algorithm to learn the correct structures. Finally, the test process to distinguish the signals according to the modulation type. Neural network algorithms aim to achieve a higher classification performance by using multiple neural layers in the modulation classification process. In addition, these layers minimize the time spent in the whole modulation classification process. [37]

One of the supervised machine learning algorithms, the decision tree algorithm breaks down data into small clusters with specific decision levels. Step by step structures provides a system for easier classification. The decision tree provides ease of visualization and interpretation. However, it can create complex structures and a problem of overfitting, which is critical to machine learning algorithms, may arise [38].

The other supervised algorithm, deep learning algorithms, which are widely used in the literature, such as machine learning algorithms, have shown rapid growth with developing technology. In cases where machine learning algorithms have problems and where a versatile system is needed, deep learning algorithms are an alternative. One of the most important advantages is that the algorithm interprets, learns and gives results without any feature extraction or pre-processing.

2.5.1 Support Vector Machine (SVM)

The SVM, which is one of the supervised algorithm, introduced in 1992 by Boser, Guyon, and Vapnik [39]. ANN algorithms are widely used in the literature, but train constraints negatively affect the performance of the algorithm. SVM provides an alternative solution to these restrictions [40]. When it was first presented, it provided a solution for the 2-class structure. But, since classes in real life are more than 2, today SVM algorithm has fulfilled regression and multi-class tasks. Since SVM is resistant to multi-class high data, it can perform poorly on data with fewer classes. The SVM algorithm is called a non-parametric model because the number of parameters is not constant when creating a model. There are two types as linear and nonlinear. While the lines that Linear SVM uses for classification in a multidimensional plane are straight, in nonlinear SVM, the lines can change shape to separate objects in the multidimensional plane. The operating logic of the linear SVM algorithm is simply shown in Figure 2.17

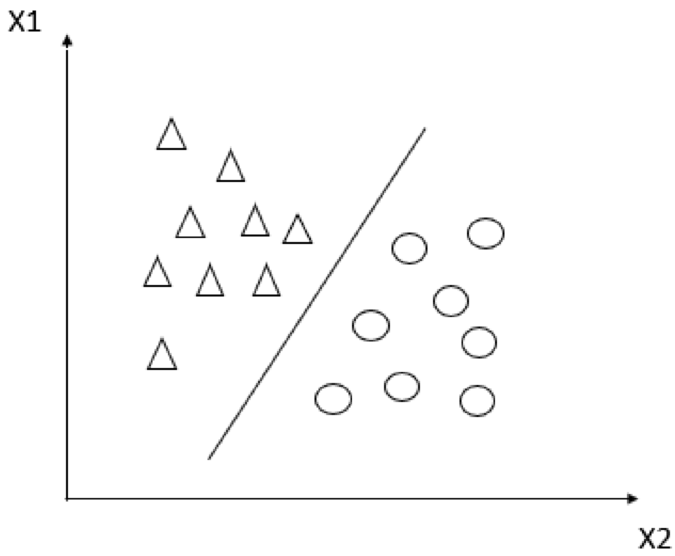

Figure 2.15: The problem of Linear Support Vector Machine Algorithm

A problem arises when some of the data remain in the other class set when the data is linearly separated. In order to solve this problem, in addition to a linear separation, another line is added to the model, which varies according to the data. This form of decomposition is called nonlinear SVM and is shown in Figure 2.18.

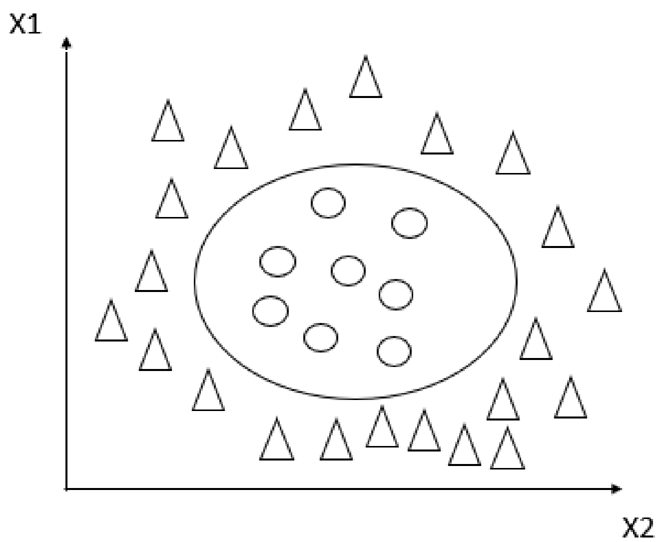

Figure 2.16: The Problem of Nonlinear Support Vector Machine Algorithm

The Decision Tree algorithm decreases performance when the number of samples decreases in feature space and the number of classes is high. SVM is preferred in such cases. Among these, the popularity of Support Vector Machine algorithms which is a machine learning algorithm has increased recently due to its high performance and low computational complexity. Solving the classification problem into an optimization problem is one of the most important advantages of SVMs. Structure of the SVM includes the ideas of ”large margin” and ”mapping data into a higher-dimensional space,” and the kernel functions in the SVM. Simply, SVM aims to draw two lines on the existing data and to make the distance between the lines as large as possible. SVM uses the hyperplane when classifying in feature space. It refers to the largest space between data points of different classes. Among all available hyperplanes, SVM aims to make an accurate classification by choosing the hyperplane at maximum distance [41].

Unlike traditional algorithms, machine learning algorithms reduce train errors and the complexity of the classification process. The SVM classification algorithm is expressed as in Equation 2.26.

$$f(x) = \sum_{i=1}^{N} \alpha_i K(x, xi) + b \tag{2.26}$$

where x is feature model, N is the number of support vectors, K(.,.) is the kernel function, b is constant, $\alpha_i \neq 0$ represent the support vectors [42].

Kernel functions are used in pattern analysis and recognition subjects. It is used to compute the classifier. It performs data with dot product function. The use of the Kernel function in SVM algorithms is widely seen in the literature. There are many types of the Kernel function like Linear, Nonlinear, Polynomial, Gaussian RBF (Radial Basis Function), Sigmoid Kernel. Polynomial and Gaussian RBF are widely used in literature [43]. The Polynomial kernel is a non-stationary kernel and is defined in Equation 2.27

$$K(x, x_i) = (xx_i + c)^n \tag{2.27}$$

where c is constant and n is the power of the kernel function. SVM algorithms with 3 kernel functions are used in the thesis as linear, quadratic and cubic [44]. Linear SVM represents 1. degree polynomial kernel function in Figure 2.15. Quadratic SVM represents 2. degree polynomial kernel function in Figure 2.16and Cubic SVM represents 3. degree polynomial kernel function in Figure 2.19.

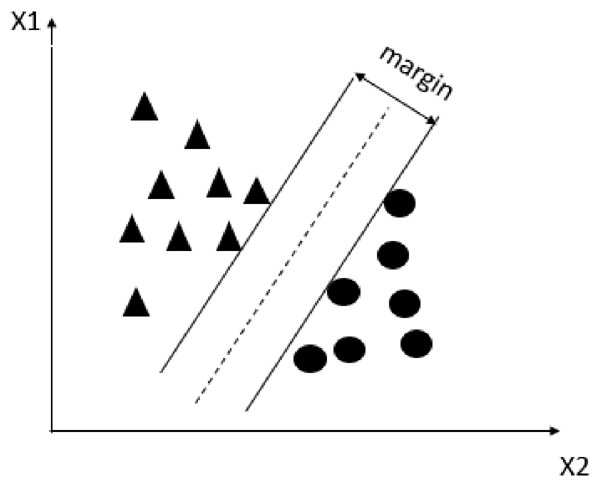

Figure 2.17: Basic Concept of Linear Support Vector Machine Algorithm

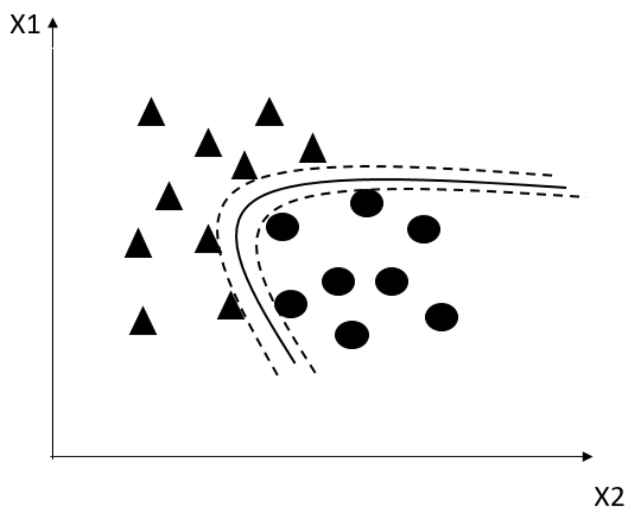

Figure 2.18: Basic Concept of Quadratic Support Vector Machine Algorithm

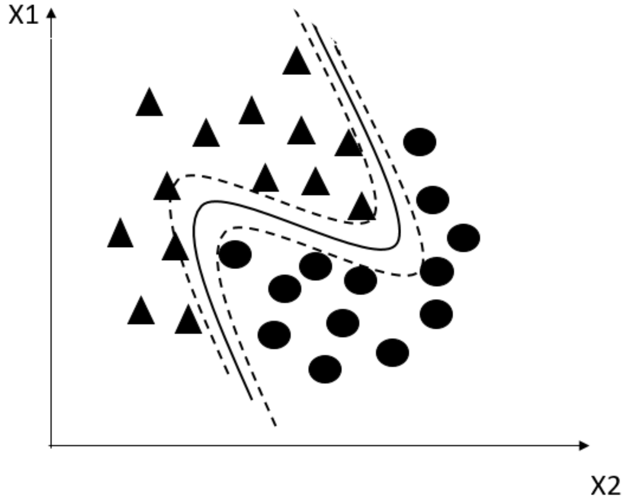

Figure 2.19: Basic Concept of Cubic Support Vector Machine Algorithm

The support vectors are closest to the separator hyperplane and form the boundary. The Hyperplane size is directly related to the number of features discussed in the classification algorithm. If the number of features is 2, the hyperplane is a line in Figure 2.20, if the number of features is 3, it is a two-dimensional plane in Figure 2.21 , and as the number of features increases, the hyperplane becomes complex, difficult to select with the eye.

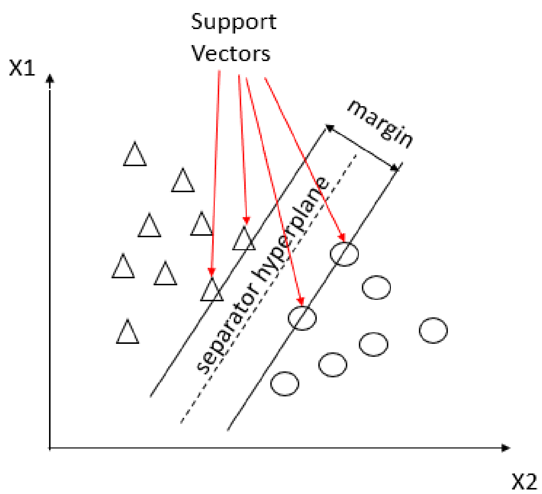

Figure 2.20: Support Vectors and Line Hyperplane

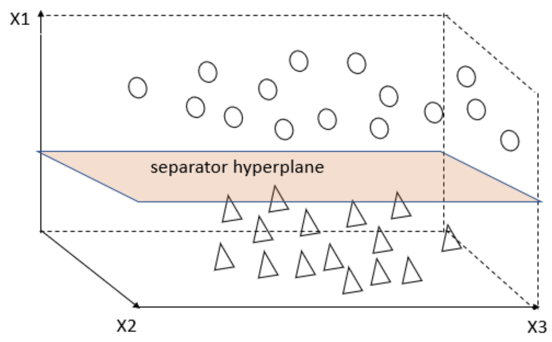

Figure 2.21: Two-dimensional Hyerplane

CHAPTER 3

RESULTS

This chapter includes the results of the performance of the Modulation Classification Algorithm developed with the feature extraction process and the Support Vector Machine Algorithm (SVM). The results of the Confusion Matrix were presented separately at different SNR values to examine. The used Linear, Quadratic and Cubic SVM algorithms at different SNR values were presented in the graph which includes performance percent too to examine together.

Modulation classification algorithms mostly used in military fields are widely available in the literature. The modulation classification algorithm can be defined as the analysis of a signal, about which no information is available, transmitted from a transmitter to the receiver. Because the modulation type of a signal processed in the Demodulator needs to be known. SVM, which is a Machine Learning algorithm is used in this classification algorithm. Feature extraction, which is pretreatment of SVM algorithm, provides information for the separation of modulated signals. The 3-dimensional structure used before feature extraction step and includes SNR values and signals were created for 12 modulations is given in Figure 3.1 .

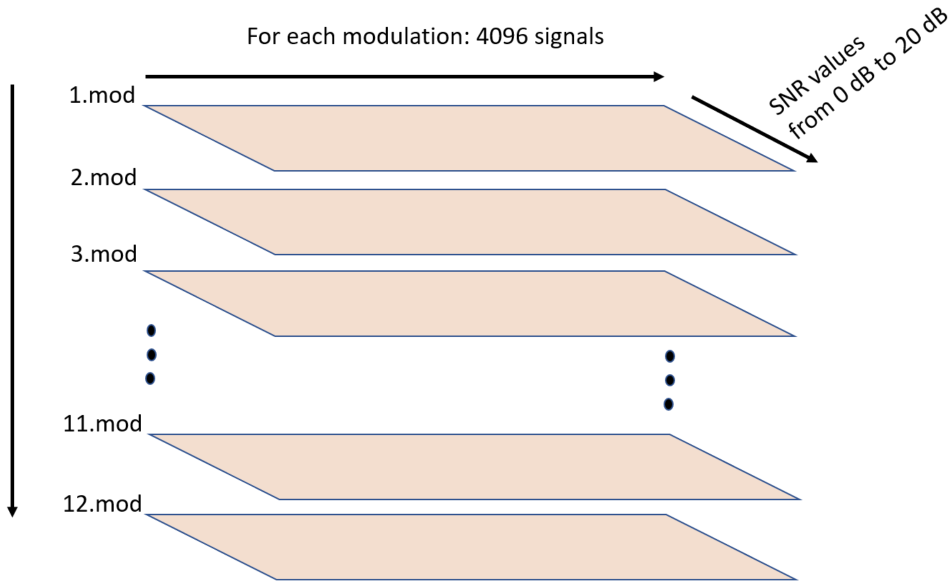

Figure 3.1: The General Form for Data Analysis

12 modulation types were used for the modulation classification algorithm. These include 16-QAM, 32-QAM, 64-QAM, 128-QAM, 256-QAM, GMSK and OQPSK as digital modulation, also AM-DSB-SC, AM-DSB-WC, AM-SSB-SC, AM-SSB-WC and FM as analog modulation. In the developed algorithm, 41 features were produced to classify the signals according to their modulations. These features include Statistical features, extracted from instantaneous amplitude, phase and frequency, Higher-order Moments (HOMs) and Higher-order Cumulants (HOCs). Each of all features is calculated from each 1024 samples of the signal. After extracting the features, Matlab's Classification Learner Application is used to separate the signals into modulation types. The Classification learner app provided the opportunity to see and use the performance of classification algorithms.

A classification process consists of two stages: training and testing. It is desirable to establish a learning mechanism by explicitly giving the training data to the machine together with the correct classroom answer. It is asked from the machine, which is trained with this data, to analyze the new test data and reach the correct outputs. 80% of the data is reserved for train step and 20% for test step to be used in the machine learning algorithm. In the other words, 3277 of the 4096 signals found for each modulation in total were used in the training process, while 819 signals were used in the test process to measure the performance of classification algorithms. To avoid overfitting, 5-fold cross-validation was applied to the data. Machine learning algorithms are first given a train data set to learn the structure of modulated signals. The train data form with features and signals is given in Figure 3.2.

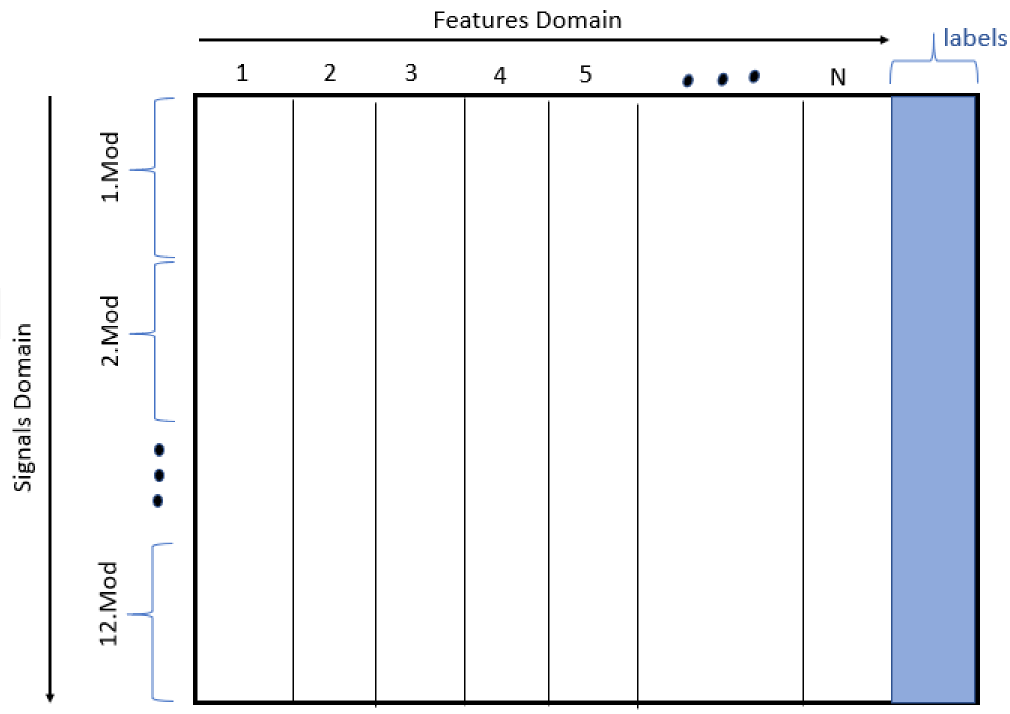

Figure 3.2: The Data Form for Train Process

Labels in the train data set represent the true answers for modulations. The test data set doesn't include the labels. It just includes features and modulated signals information. It is wanted that the machine learning algorithm finds the correct modulation types utilizing the information created in the training process. At the end of the test, the confusion matrix is used to observe the performance rate of the modulations.

Data has been trained and tested with 22 different classifiers in Matlab's classification toolbox. Among them, all known to be successful in pattern recognition and linear classification, Linear SVM, Quadratic SVM and Cubic SVM are outperform in Figure 3.3, 3.4 and 3.4, respectively.

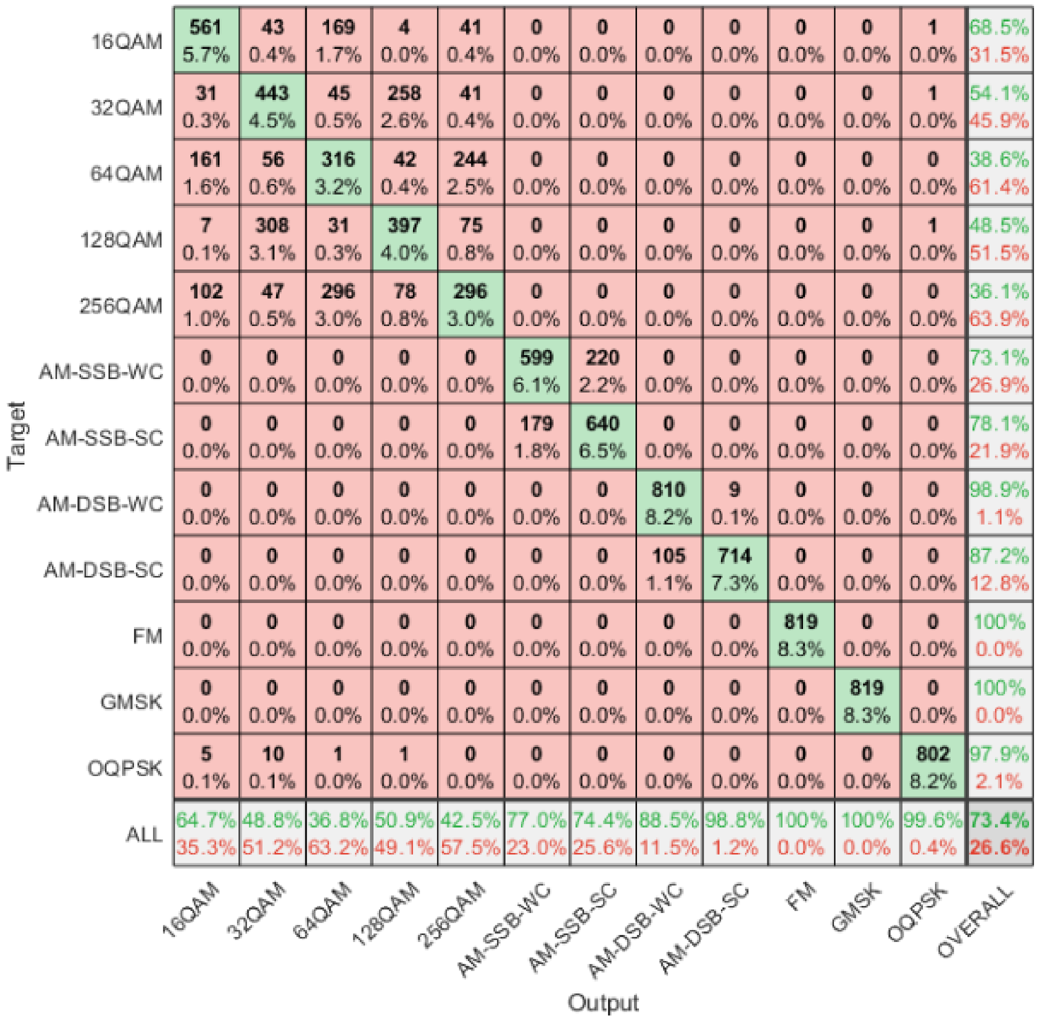

Figure 3.3: The Confusion Matrix of Linear SVM at 10 dB

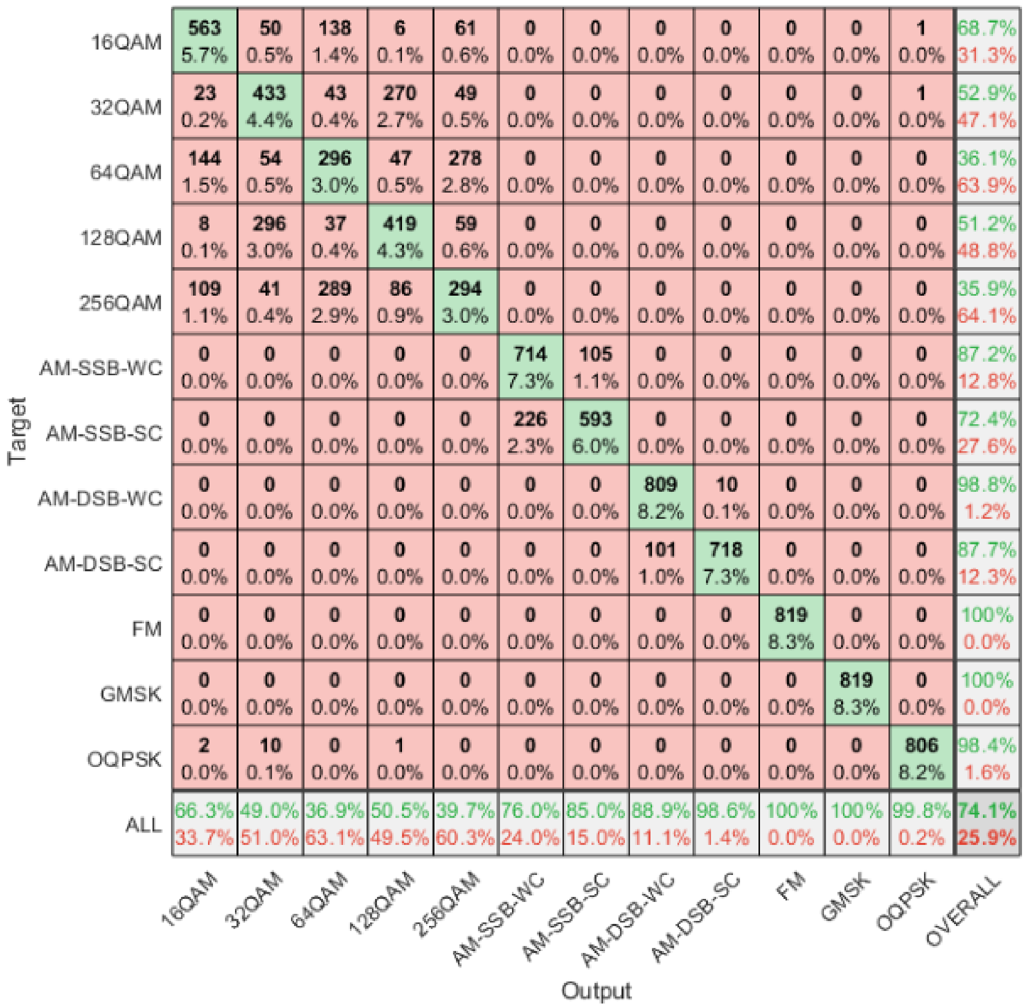

Figure 3.4: The Confusion Matrix of Quadratic SVM at 10 dB

Target \ Output	16QAM	32QAM	64QAM	128QAM	256QAM	AM-SSB-WC	AM-SSB-SC	AM-DSB-WC	AM-DSB-SC	FM	GMSK	OQPSK	OVERALL
16QAM	527 5.4%	40 0.4%	153 1.6%	12 0.1%	86 0.9%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	1 0.0%	64.3% 35.7%
32QAM	24 0.2%	424 4.3%	57 0.6%	263 2.7%	49 0.5%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	2 0.0%	51.8% 48.2%
64QAM	138 1.4%	53 0.5%	300 3.1%	44 0.4%	284 2.9%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	36.6% 63.4%
128QAM	13 0.1%	312 3.2%	54 0.5%	372 3.8%	67 0.7%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	1 0.0%	45.4% 54.6%
256QAM	103 1.0%	63 0.6%	270 2.7%	82 0.8%	301 3.1%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	36.8% 63.2%
AM-SSB-WC	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	688 7.0%	131 1.3%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	84.0% 16.0%
AM-SSB-SC	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	205 2.1%	614 6.2%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	75.0% 25.0%
AM-DSB-WC	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	799 8.1%	20 0.2%	0 0.0%	0 0.0%	0 0.0%	97.6% 2.4%
AM-DSB-SC	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	102 1.0%	717 7.3%	0 0.0%	0 0.0%	0 0.0%	87.5% 12.5%
FM	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	819 8.3%	0 0.0%	0 0.0%	100% 0.0%
GMSK	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	819 8.3%	0 0.0%	100% 0.0%
OQPSK	1 0.0%	6 0.1%	0 0.0%	2 0.0%	2 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	808 8.2%	98.7% 1.3%
ALL	65.4% 34.6%	47.2% 52.8%	36.0% 64.0%	48.0% 52.0%	38.1% 61.9%	77.0% 23.0%	82.4% 17.6%	88.7% 11.3%	97.3% 2.7%	100% 0.0%	100% 0.0%	99.5% 0.5%	73.1% 26.9%

Figure 3.5: The Confusion Matrix of Cubic SVM at 10 dB

It is observed that the performance result of these 3 algorithms is close to each other, when the performance results of the modulation classification algorithm of Linear, Quadratic and Cubic SVM algorithms were examined. At 10dB, Linear SVM performance is 73.4%, Quadratic SVM performance is 74.1% and Cubic SVM performance is 73.1%. Since the performance results of the 3 algorithms are close, the modulation classification performance results at different SNR values are given only for one algorithm. That is the Quadratic SVM.

Considering Table 3.1, performance success can be seen for each modulation type for Linear SVM, Quadratic SVM and Cubic SVM. Apart from SSB modulation type, it can be said that it does not make a significant difference for 3 different SVM algorithms.

Table 3.1: Comparing Linear SVM, Quadratic SVM and Cubic SVM For Each Modulation Types

	Linear SVM	Quadratic SVM	Cubic SVM
16-QAM	68.5	68.7	64.3
32-QAM	54.1	52.9	51.8
64-QAM	38.6	36.1	36.6
128-QAM	48.5	51.2	45.4
256-QAM	36.1	35.9	36.8
SSB-WC	73.1	87.2	84
SSB-SC	78.1	72.4	75
DSB-WC	98.9	98.8	97.6
DSB-SC	87.2	87.7	87.5
FM	100	100	100
GMSK	100	100	100
OQPSK	97.9	98.4	98,7

Classification performance of 12 modulations are examined at 0 dB in Figure 3.6, 6 dB in Figure 3.7, 12 dB in Figure 3.8 and 18 dB in Figure 3.9. Thanks to the confusion matrix, modulations can be analyzed one by one. And, it can be examined which modulations mixed with each other.The result of the classification performance is 44.5% at 0 dB, it is 69.9% at 6 dB, it is 75.5 at 12 dB, it is 73.6% at 18 dB.

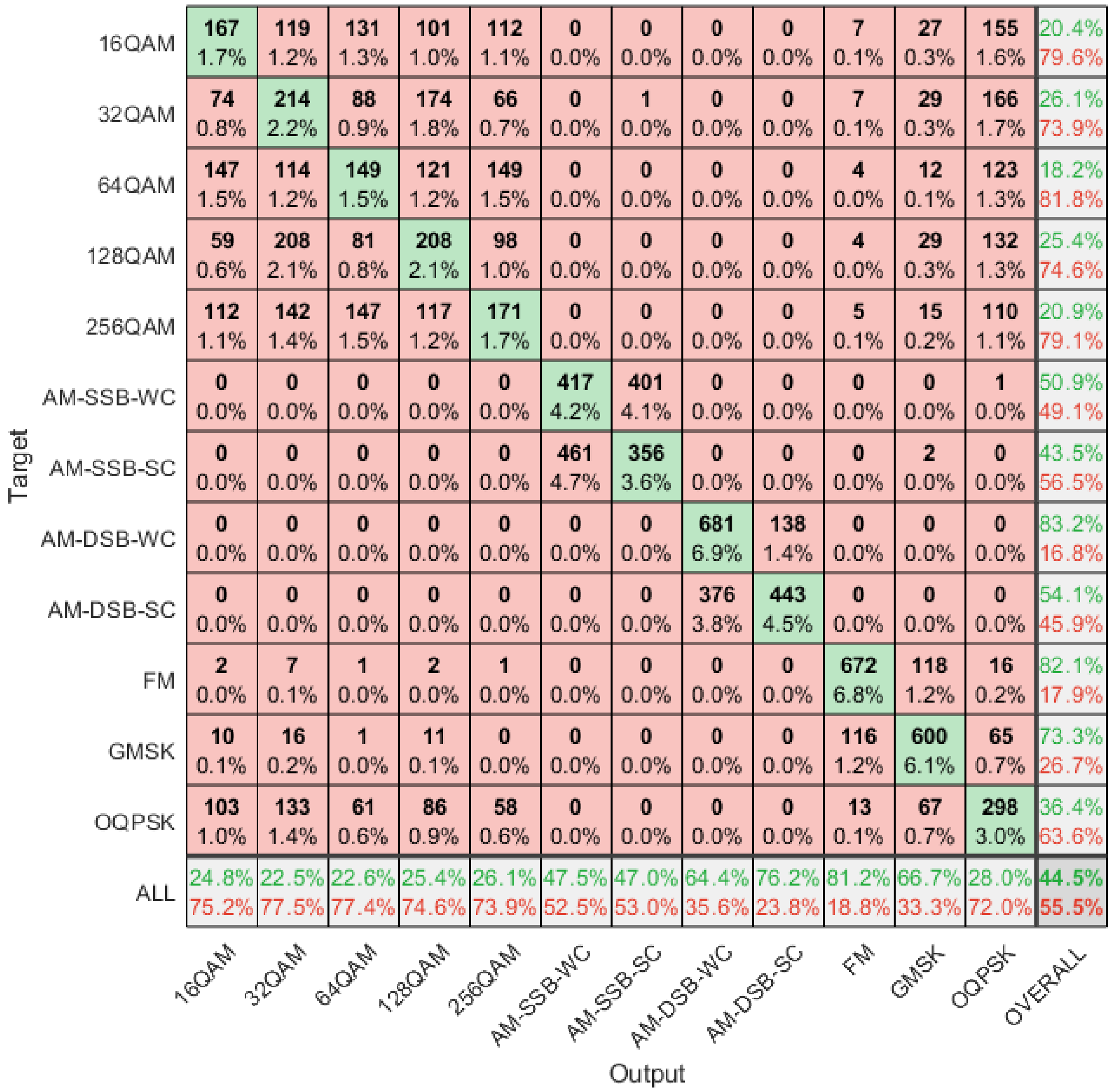

Figure 3.6: The Confusion Matrix of Quadratic SVM at 0 dB

Target \ Output	16QAM	32QAM	64QAM	128QAM	256QAM	AM-SSB-WC	AM-SSB-SC	AM-DSB-WC	AM-DSB-SC	FM	GMSK	OQPSK	OVERALL
16QAM	**540** 5.5%	**44** 0.4%	**143** 1.5%	**10** 0.1%	**81** 0.8%	**0** 0.0%	**0** 0.0%	**0** 0.0%	**0** 0.0%	**0** 0.0%	**0** 0.0%	**1** 0.0%	65.9% 34.1%
32QAM	**37** 0.4%	**428** 4.4%	**40** 0.4%	**268** 2.7%	**42** 0.4%	**0** 0.0%	**0** 0.0%	**0** 0.0%	**0** 0.0%	**0** 0.0%	**0** 0.0%	**4** 0.0%	52.3% 47.7%
64QAM	**187** 1.9%	**63** 0.6%	**237** 2.4%	**68** 0.7%	**264** 2.7%	**0** 0.0%	**0** 0.0%	**0** 0.0%	**0** 0.0%	**0** 0.0%	**0** 0.0%	**0** 0.0%	28.9% 71.1%
128QAM	**20** 0.2%	**309** 3.1%	**60** 0.6%	**362** 3.7%	**66** 0.7%	**0** 0.0%	**0** 0.0%	**0** 0.0%	**0** 0.0%	**0** 0.0%	**0** 0.0%	**2** 0.0%	44.2% 55.8%
256QAM	**141** 1.4%	**52** 0.5%	**238** 2.4%	**87** 0.9%	**301** 3.1%	**0** 0.0%	**0** 0.0%	**0** 0.0%	**0** 0.0%	**0** 0.0%	**0** 0.0%	**0** 0.0%	36.8% 63.2%
AM-SSB-WC	**0** 0.0%	**0** 0.0%	**0** 0.0%	**0** 0.0%	**0** 0.0%	**553** 5.6%	**266** 2.7%	**0** 0.0%	**0** 0.0%	**0** 0.0%	**0** 0.0%	**0** 0.0%	67.5% 32.5%
AM-SSB-SC	**0** 0.0%	**0** 0.0%	**0** 0.0%	**0** 0.0%	**0** 0.0%	**248** 2.5%	**571** 5.8%	**0** 0.0%	**0** 0.0%	**0** 0.0%	**0** 0.0%	**0** 0.0%	69.7% 30.3%
AM-DSB-WC	**0** 0.0%	**0** 0.0%	**0** 0.0%	**0** 0.0%	**0** 0.0%	**0** 0.0%	**0** 0.0%	**773** 7.9%	**46** 0.5%	**0** 0.0%	**0** 0.0%	**0** 0.0%	94.4% 5.6%
AM-DSB-SC	**0** 0.0%	**0** 0.0%	**0** 0.0%	**0** 0.0%	**0** 0.0%	**0** 0.0%	**0** 0.0%	**150** 1.5%	**669** 6.8%	**0** 0.0%	**0** 0.0%	**0** 0.0%	81.7% 18.3%
FM	**0** 0.0%	**0** 0.0%	**0** 0.0%	**0** 0.0%	**0** 0.0%	**0** 0.0%	**0** 0.0%	**0** 0.0%	**0** 0.0%	**819** 8.3%	**0** 0.0%	**0** 0.0%	100% 0.0%
GMSK	**0** 0.0%	**0** 0.0%	**0** 0.0%	**0** 0.0%	**0** 0.0%	**0** 0.0%	**0** 0.0%	**0** 0.0%	**0** 0.0%	**0** 0.0%	**819** 8.3%	**0** 0.0%	100% 0.0%
OQPSK	**6** 0.1%	**12** 0.1%	**0** 0.0%	**2** 0.0%	**0** 0.0%	**0** 0.0%	**0** 0.0%	**0** 0.0%	**0** 0.0%	**0** 0.0%	**0** 0.0%	**799** 8.1%	97.6% 2.4%
ALL	58.0% 42.0%	47.1% 52.9%	33.0% 67.0%	45.4% 54.6%	39.9% 60.1%	69.0% 31.0%	68.2% 31.8%	83.7% 16.3%	93.6% 6.4%	100% 0.0%	100% 0.0%	99.1% 0.9%	**69.9% 30.1%**

Figure 3.7: The Confusion Matrix of Quadratic SVM at 6 dB

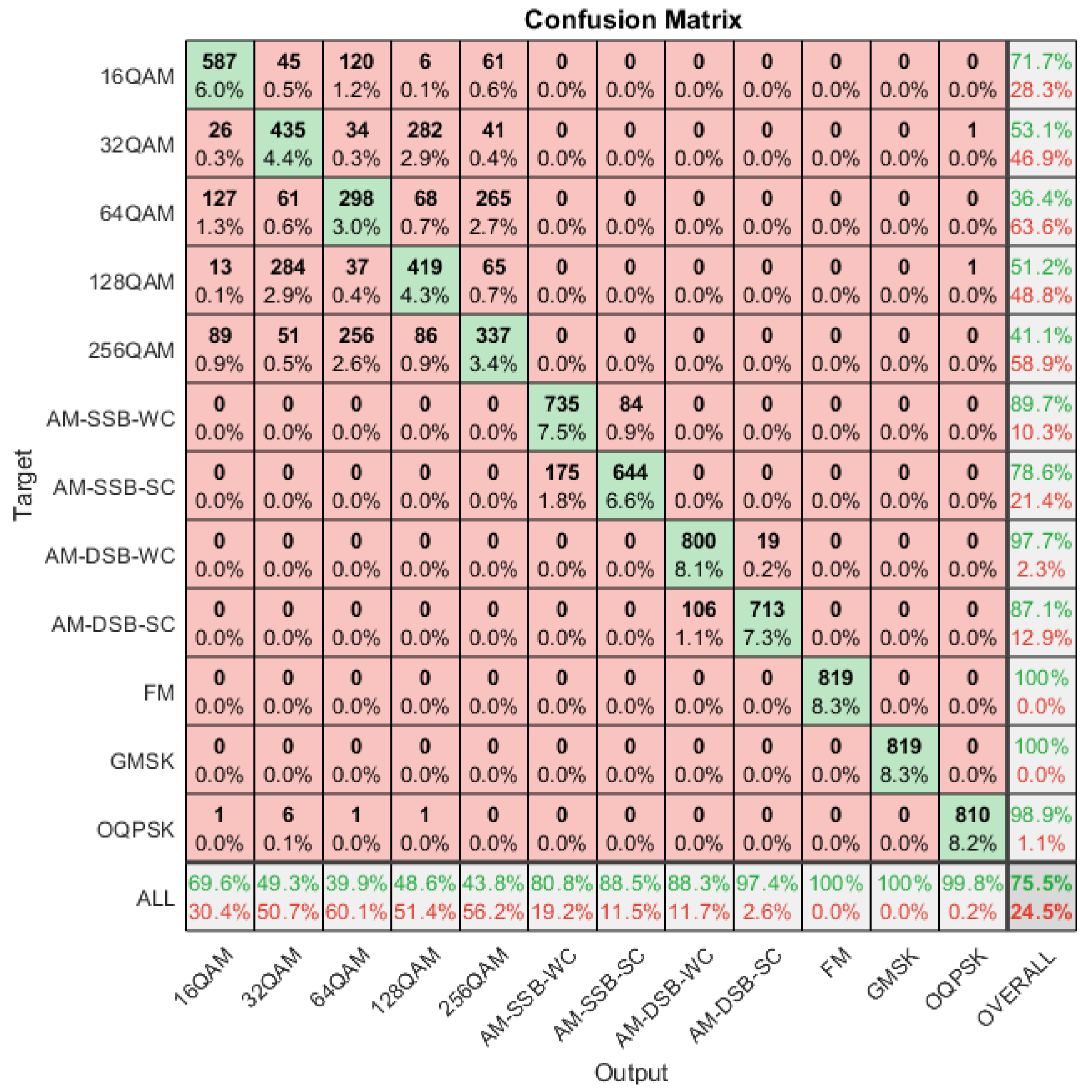

Figure 3.8: The Confusion Matrix of Quadratic SVM at 12 dB

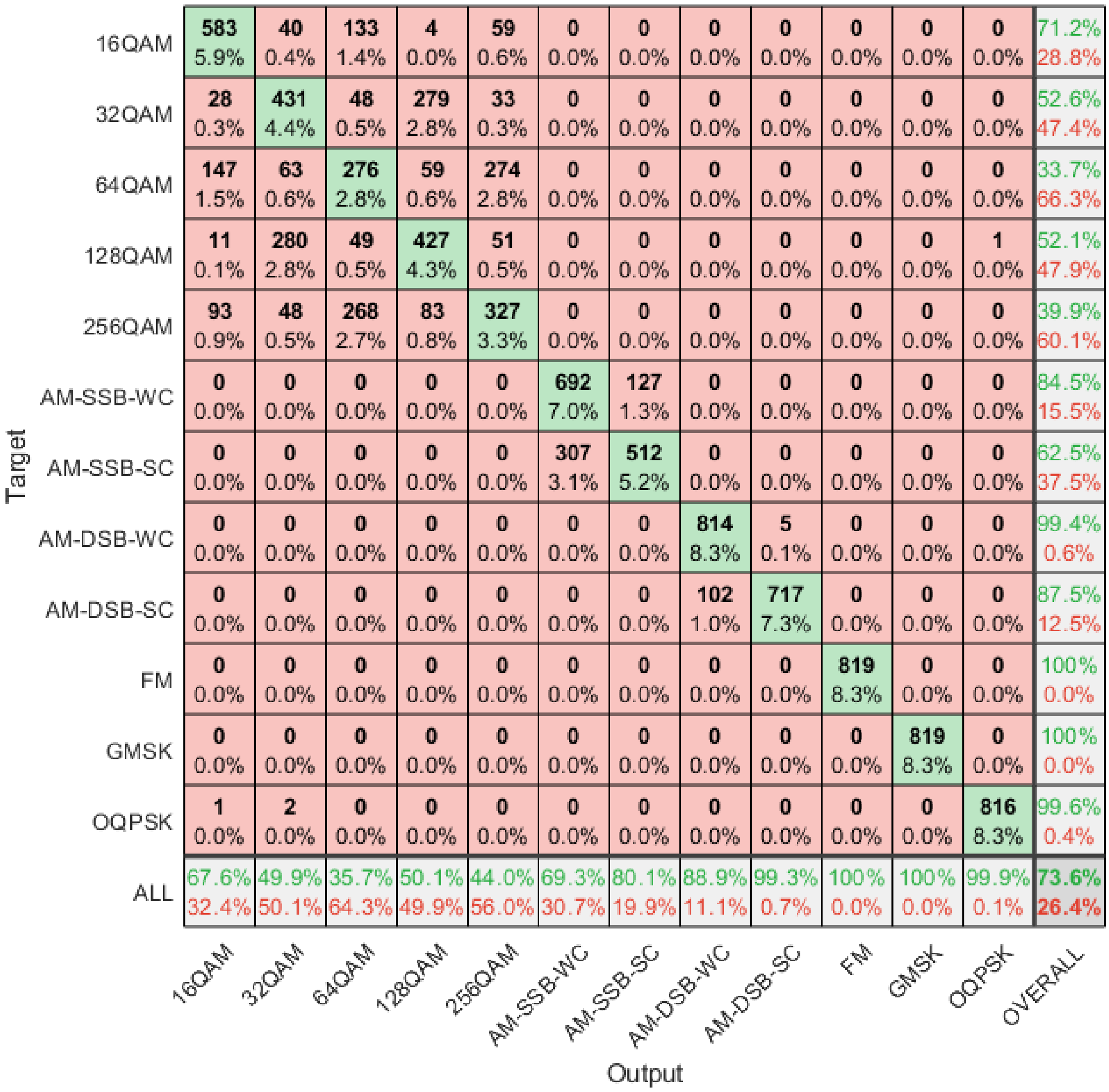

Figure 3.9: The Confusion Matrix of Quadratic SVM at 18 dB

The results are presented in Table 3.2 for a more convenient observation of performance at different SNR values for each modulation. When the confusion matrix was examined for each modulation, 8 modulations could be separated with a performance of more than 70%. For others, it was observed that 32-QAM and 128-QAM were classified with a performance of approximately 50%, while 64-QAM and 256-QAM were classified with a percentage less than 50%. FM and GMSK modulations showed approximately 100% classification performance starting from 2dB. DSB-WC and OQPSK have been classified with an average of 90% success starting from 4dB. While DSB-SC has a performance of more than 80% since 6dB, SSB-SC and SSB-

WC have achieved 70% success from 10dB. For M-ary QAM, classification performance remained stable from 6dB.

Table 3.2: Classification Performance at Different dB Values For each Modulation Type

	0 dB	2 dB	4 dB	6 dB	10 dB	14 dB	18 dB
16-QAM	20.4	38.6	57.1	65.9	68.7	71.1	71.2
32-QAM	26.1	38.5	47.3	52.3	52.9	52.3	52.6
64-QAM	18.2	17.8	27.1	28.9	36.1	30.5	33.7
128-QAM	25.4	37.5	43.6	44.2	51.2	49	52.1
256-QAM	20.9	30.3	35.4	36.8	35.9	37.6	39.9
AM-SSB-WC	50.9	49.3	55.2	67.5	87.2	91.9	84.5
AM-SSB-SC	43.5	56.3	59.1	69.7	72.4	76.9	62.5
AM-DSB-WC	83.2	82.9	88.6	94.4	98.8	99.3	99.4
AM-DSB-SC	54.1	61.4	74.4	81.7	87.7	87.8	87.5
FM	82.1	100	100	100	100	100	100
GMSK	73.3	99	99.9	100	100	100	100
OQPSK	36.4	74.2	92.4	97.6	98.4	99.1	99.6

The results of the classification performance at all SNR values for Linear SVM, Quadratic SVM, and Cubic SVM are given in Figure 3.10. It is seen that performance is stable from 10 dB to 20dB and observed the performance accuracy equals to approximately 73%. When it is examined the performance at lower SNR, It is seen that the performance of the algorithm is much lower. For example, the performance is approximately 43% at 0dB and the performance is approximately 64% at 4dB. The Quadratic SVM performs slightly better than the others at a rate of 1%.

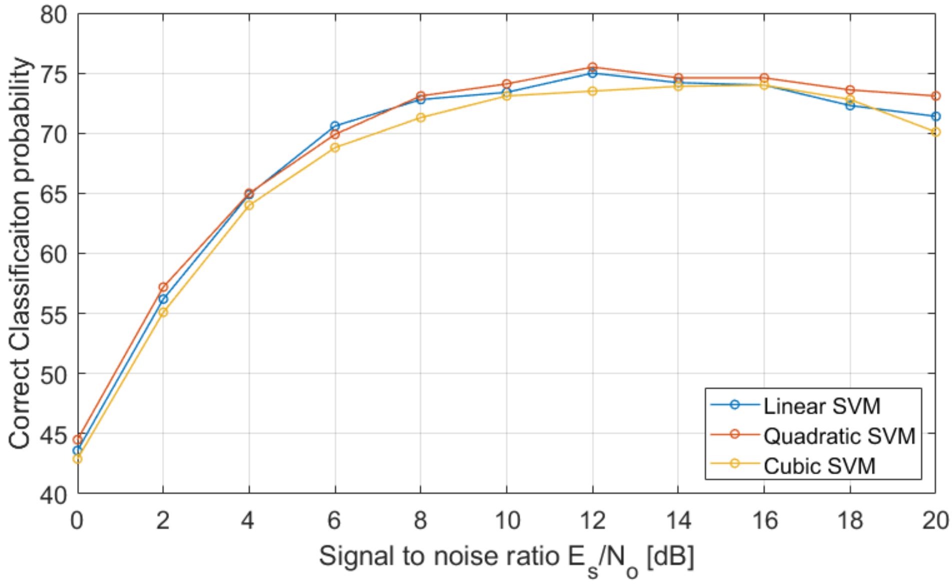

Figure 3.10: Correct Classification Probability at all SNR values

CHAPTER 4

CONCLUSION

The aim of this thesis is to classify the modulated signals with 12 different modulation types according to the modulation type with SVM which is a machine learning algorithm with the help of features extracted from the signals. For this purpose, Statistical features, extracted from instantaneous amplitude, phase and frequency, Higher-order Moments (HOMs) and Higher-order Cumulants (HOCs) were used in order to provide information from signals at different SNR values. Linear SVM, Quadratic SVM, Cubic SVM algorithms which are trained using these features and then used for the testing phase were presented. The results of the performance of these algorithms in the classification of signals from 0 dB to 20 dB were presented. As a result of these analyzes, it was observed that the performances of these 3 algorithms were very close to each other, but with a small difference, Quadratic SVM performance was superior.

In general, when confusion matrices at different dB values are examined, while for Mary-QAM (Quadrature Amplitude Modulation), i.e. 16-QAM, 32-QAM, 64-QAM, 128-QAM, and 256-QAM, false prediction rates are high, GMSK (Gaussian Minimum Shift Keying) and OQPSK (Offset Quadrature Phase-shift Keying) are separated from other modulations more easily. In analog modulations, the algorithm has difficulty in differentiating the modulation types in case of the presence and absence of carriers of the same modulation type. That is, the Algorithm is forced to separate AM-DSB-WC (Double Sideband with Carrier) and AM-DSB-SC (Double Sideband Suppressed Carrier) modulations and to separate AM-SSB-WC (Single Sideband with Carrier) and AM-SSB-SC (Single Sideband Suppressed Carrier). Also, FM (Frequency Modulation) was separated easily.

It can be observed that higher-order modulations are more difficult to classify than others. In Mary-QAM types, classification performance for 256-QAM is 36.8% and classification performance for 128-QAM is 44.2%, while classification performance for 16-QAM is 65.9% and classification performance for 32-QAM is 52.3%. Furthermore, based on the formula $2^n = M$, it can be observed that n exhibits more confusing than double and uniqueness. For example, 32-QAM and 128-QAM are confused, and the other modulation types which are confused are 64-QAM and 256-QAM.

As a result, when the results of the classification performances are examined at all SNR values, the stable result which is observed from 10 dB to 20 dB is approximately 73%. The most difference in classification algorithm performance is observed at between 0 dB and 12 dB and it is approximately 31%. It performs well in analog modulations and in digital modulations such as GMSK, OQPSK and low-order M-ary QAM. Algorithm performance decreased in high order digital modulation types.

The developed modulation classification algorithm can be extended to include more modulation than the existing modulation types. It may be important to increase the number of digital modulations, especially as technology tends to use digital modulation types more. In addition, an experimental environment can be prepared in a laboratory environment and a unique dataset can be prepared by transmitting the modulated signals from the transmitter to the receiver. In this way, all the information about the content of the generated dataset will be obtained and, thus possible errors are prevented. By using multi-fading channel effects instead of synthetic channel effects, the algorithm can be adapted more easily to real life. In the classification process, Deep learning applications, which have gained popularity in recent years, can be adapted to algorithm. In the light of these recommendations, the success of the algorithm developed for modulation classification can be increased and its scope can be improved.

REFERENCES

[1] J. G. Proakis and M. Salehi, *Digital communications*, vol. 4. McGraw-hill New York, 2001.

[2] H. Simon, *Communication systems*. Springer Nature, 1999.

[3] S. Haykin, *An introduction to analog and digital communication*. John Wiley, 1994.

[4] Z. Zhu and A. K. Nandi, *Automatic modulation classification: principles, algorithms and applications*. John Wiley & Sons, 2015.

[5] P. Panagiotou, A. Anastasopoulos, and A. Polydoros, "Likelihood ratio tests for modulation classification," in *MILCOM 2000 Proceedings. 21st Century Military Communications. Architectures and Technologies for Information Superiority (Cat. No. 00CH37155)*, vol. 2, pp. 670–674, IEEE, 2000.

[6] O. A. Dobre and F. Hameed, "Likelihood-based algorithms for linear digital modulation classification in fading channels," in *2006 Canadian Conference on Electrical and Computer Engineering*, pp. 1347–1350, IEEE, 2006.

[7] Z. Zhu and A. K. Nandi, "Blind digital modulation classification using minimum distance centroid estimator and non-parametric likelihood function," *IEEE Transactions on Wireless Communications*, vol. 13, no. 8, pp. 4483–4494, 2014.

[8] O. A. Dobre, A. Abdi, Y. Bar-Ness, and W. Su, "Survey of automatic modulation classification techniques: classical approaches and new trends," *IET communications*, vol. 1, no. 2, pp. 137–156, 2007.

[9] E. E. Azzouz and A. K. Nandi, "Automatic identification of digital modulation types," *Signal Processing*, vol. 47, no. 1, pp. 55–69, 1995.

[10] E. Azzouz and A. Nandi, "Procedure for automatic recognition of analogue and digital modulations," *IEE Proceedings-communications*, vol. 143, no. 5, pp. 259–266, 1996.

[11] A. Nandi and E. E. Azzouz, "Automatic analogue modulation recognition," *Signal processing*, vol. 46, no. 2, pp. 211–222, 1995.

[12] A. K. Nandi and E. E. Azzouz, "Algorithms for automatic modulation recognition of communication signals," *IEEE Transactions on communications*, vol. 46, no. 4, pp. 431–436, 1998.

[13] T. Cai, C. Wang, G. Cui, and W. Wang, "Constellation-wavelet transform automatic modulation identifier for m-ary qam signals," in *2015 IEEE 26th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, pp. 212–216, IEEE, 2015.

[14] M. Walenczykowska and A. Kawalec, “Type of modulation identification using wavelet transform and neural network,” *Bulletin of the Polish Academy of Sciences Technical Sciences*, vol. 64, no. 1, pp. 257–261, 2016.

[15] Y. Yang and S. S. Soliman, “A suboptimal algorithm for modulation classification,” *IEEE transactions on aerospace and electronic systems*, vol. 33, no. 1, pp. 38–45, 1997.

[16] Y. Yang and C.-H. Liu, “An asymptotic optimal algorithm for modulation classification,” *IEEE communications letters*, vol. 2, no. 5, pp. 117–119, 1998.

[17] H.-C. Wu, M. Saquib, and Z. Yun, “Novel automatic modulation classification using cumulant features for communications via multipath channels,” *IEEE Transactions on Wireless Communications*, vol. 7, no. 8, pp. 3098–3105, 2008.

[18] P. Liu and P.-L. Shui, “A new cumulant estimator in multipath fading channels for digital modulation classification,” *IET Communications*, vol. 8, no. 16, pp. 2814–2824, 2014.

[19] M. S. Muhlhaus, M. Oner, O. A. Dobre, and F. K. Jondral, “A low complexity modulation classification algorithm for mimo systems,” *IEEE Communications Letters*, vol. 17, no. 10, pp. 1881–1884, 2013.

[20] V. D. Orlic and M. L. Dukic, “Automatic modulation classification algorithm using higher-order cumulants under real-world channel conditions,” *IEEE Communications Letters*, vol. 13, no. 12, pp. 917–919, 2009.

[21] Z. Zhang, Z. Hua, and Y. Liu, “Modulation classification in multipath fading channels using sixth-order cumulants and stacked convolutional auto-encoders,” *IET communications*, vol. 11, no. 6, pp. 910–915, 2017.

[22] D.-C. Chang and P.-K. Shih, “Cumulants-based modulation classification technique in multipath fading channels,” *IET Communications*, vol. 9, no. 6, pp. 828–835, 2015.

[23] A. Ebrahimzadeh and R. Ghazalian, “Blind digital modulation classification in software radio using the optimized classifier and feature subset selection,” *Engineering Applications of Artificial Intelligence*, vol. 24, no. 1, pp. 50–59, 2011.

[24] A. Abdelmutalab, K. Assaleh, and M. El-Tarhuni, “Automatic modulation classification based on high order cumulants and hierarchical polynomial classifiers,” *Physical Communication*, vol. 21, pp. 10–18, 2016.

[25] M. Zaerin and B. Seyfe, “Multiuser modulation classification based on cumulants in additive white gaussian noise channel,” *IET Signal Processing*, vol. 6, no. 9, pp. 815–823, 2012.

[26] A. Swami and B. M. Sadler, “Hierarchical digital modulation classification using cumulants,” *IEEE Transactions on communications*, vol. 48, no. 3, pp. 416–429, 2000.

[27] R. Gupta, S. Majhi, and O. A. Dobre, “Design and implementation of a tree-based blind modulation classification algorithm for multiple-antenna systems,” *IEEE Transactions on Instrumentation and Measurement*, 2018.

[28] J. Lee, J. Kim, B. Kim, D. Yoon, and J. Choi, “Robust automatic modulation classification technique for fading channels via deep neural network,” *Entropy*, vol. 19, no. 9, p. 454, 2017.

[29] S.-J. Kim and D. Yoon, “Automatic modulation classification in practical wireless channels,” in *2016 International Conference on Information and Communication Technology Convergence (ICTC)*, pp. 915–917, IEEE, 2016.

[30] M. Marey and O. A. Dobre, “Blind modulation classification algorithm for single and multiple-antenna systems over frequency-selective channels,” *IEEE Signal Processing Letters*, vol. 21, no. 9, pp. 1098–1102, 2014.

[31] “Deepsig.” `https://www.deepsig.io/datasets`. Accessed: 12.12.2019.

[32] T. J. O’Shea, T. Roy, and T. C. Clancy, “Over-the-air deep learning based radio signal classification,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 1, pp. 168–179, 2018.

[33] D. K. Sharma, A. Mishra, and R. Saxena, “Analog & digital modulation techniques: An overview,” *International Journal of Computing Science and Communication Technologies*, vol. 3, no. 1, p. 2007, 2010.

[34] K. Kuchi and V. K. Prabhu, “Power spectral density of gmsk modulation using matrix methods,” in *MILCOM 1999. IEEE Military Communications. Conference Proceedings (Cat. No. 99CH36341)*, vol. 1, pp. 45–50, IEEE, 1999.

[35] N. P. Geisinger, “Classification of digital modulation schemes using linear and nonlinear classifiers,” tech. rep., NAVAL POSTGRADUATE SCHOOL MONTEREY CA, 2010.

[36] J. Jagannath, N. Polosky, A. Jagannath, F. Restuccia, and T. Melodia, “Machine learning for wireless communications in the internet of things: a comprehensive survey,” *Ad Hoc Networks*, p. 101913, 2019.

[37] S. Wei, S. Zou, F. Liao, W. Lang, and W. Wu, “Automatic modulation recognition using neural architecture search,” in *2019 International Conference on High Performance Big Data and Intelligent Systems (HPBD&IS)*, pp. 151–156, IEEE, 2019.

[38] J. Tsiligaridis, “Classification with neural network and svm via decision tree algorithm,” in *AIP Conference Proceedings*, vol. 1982, p. 020040, AIP Publishing, 2018.

[39] B. E. Boser, I. M. Guyon, and V. N. Vapnik, “A training algorithm for optimal margin classifiers,” in *Proceedings of the fifth annual workshop on Computational learning theory*, pp. 144–152, ACM, 1992.

[40] A. Hazza, M. Shoaib, S. A. Alshebeili, and A. Fahad, “An overview of feature-based methods for digital modulation classification,” in *2013 1st International Conference on Communications, Signal Processing, and their Applications (ICCSPA)*, pp. 1–6, IEEE, 2013.

[41] Y. Ma and G. Guo, *Support vector machines applications*. Springer, 2014.

[42] S.-W. Lee and A. Verri, *Pattern Recognition with Support Vector Machines: First International Workshop, SVM 2002, Niagara Falls, Canada, August 10, 2002. Proceedings*, vol. 2388. Springer, 2003.

[43] J. H. Min and Y.-C. Lee, “Bankruptcy prediction using support vector machine with optimal choice of kernel function parameters,” *Expert systems with applications*, vol. 28, no. 4, pp. 603–614, 2005.

[44] I. Mejía-Guevara and Á. Kuri-Morales, “Mp-polynomial kernel for training support vector machines,” in *Iberoamerican Congress on Pattern Recognition*, pp. 584–593, Springer, 2007.