



REPUBLIC OF TÜRKİYE

ALTINBAŞ UNIVERSITY

Institute of Graduate Studies

Electrical and Computer Engineering

**NEW MEDICAL IMAGE RECOGNITION SYSTEM
FOR COVID-19 DETECTION USING
CONVOLUTIONAL NEURAL NETWORK**

Othman Hussein Alwan TIKREETI

Master's Thesis

Supervisor

Asst. Prof. Dr. Ayça Kurnaz TÜRKBEN

Istanbul, 2022

NEW MEDICAL IMAGE RECOGNITION SYSTEM FOR COVID-19 DETECTION USING CONVOLUTIONAL NEURAL NETWORK

Othman Hussein Alwan TIKREETI

Electrical and Computer Engineering

Master's Thesis

ALTINBAŞ UNIVERSITY

2022

The thesis titled NEW MEDICAL IMAGE RECOGNITION SYSTEM FOR COVID-19 DETECTION USING CONVOLUTIONAL NEURAL NETWORK prepared by OTHMAN HUSSEIN ALWAN TIKREETI and submitted on 15/12/2022 has been **accepted unanimously** for the degree of Master of Science in Electrical and Computer Engineering.

Asst. Prof. Dr. Ayça Kurnaz TÜRKBEN

the Supervisor

Thesis Defense Committee Members:

Asst. Prof. Dr. Ayça Kurnaz
TÜRKBEN

Department of Software
Engineering,

Altınbaş University

Asst. Prof. Dr. Abdullahi Abdu
IBRAHIM

Department of Computer
Engineering,

Altınbaş University

Asst. Prof. Dr. Serdar KARGIN

Department of Biomedical
Engineering,

İstanbul Arel University

I hereby declare that this thesis meets all format and submission requirements of a Master's thesis.

Submission date of the thesis to Institute of Graduate Studies: ____/____/____

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Othman Hussain Alwan TIKREETI

Signature

DEDICATION

This thesis is dedicated to All, I love Specially My dearest parents My wife and daughter My brother For their endless encouragement, patience, and support and for being a great source of motivation and inspiration All my friends And to my homeland, Iraq.



PREFACE

I would like to thank my supervisor Asst. Prof. Dr. Ayça Kurnaz TRKBEN for his help and assistance throughout my work and to my family.



ABSTRACT

A NEW MEDICAL IMAGE RECOGNITION SYSTEM FOR COVID-19 DETECTION USING CONVOLUTIONAL NEURAL NETWORK FOR

Tikreeti, Othman Hussain Alwan

MSc., Electrical and Computer Engineering, Altınbaş University,

Supervisor: Asst. Prof. Dr. Ayça Kurnaz TÜRK BEN

Date: 12/2022

Pages: 66

In this study new method based convolutional neural network presented for covid-19 detection. The proposed method applied to detect covid-19 in three different datasets. The proposed method show remarkable results when compared with other traditional techniques. The proposed method presented more than 99% accuracy, precision and sensitivity. This mean that the combination of the model are very effective to predicate optimum results. The presented model combined CNN, which is new technique, used in feature extraction problems and extracted high level features from input x-ray images. The CNN trained using ALO which is new optimization algorithm and applied in various optimization problems. The using ALO lead to obtain best weight and basis to present best accuracy and detection rate. On the other hand, the genetic algorithm applied for select best features from features that are extracted by CNN. The genetic algorithm role is to minimize the size of CNN output features to decrease the execution time and increase the performance of the classifiers that applied in this study.

Keywords: Covid-19, Genetic Algorithm, Ant Colony Optimiztion Algorithm , KNN, CNN.

TABLE OF CONTENTS

	<u>Pages</u>
ABSTRACT	vii
LIST OF FIGURES.....	x
ABBREVIATIONS.....	xii
1. INTRODUCTION.....	1
1.1 DEFINITION OF SARS-COV-2.....	3
1.2 CONTRIBUTIONS.....	7
2. LITREATURE REVIEW.....	8
2.1 MACHINE LEARNING.....	8
2.1.1 Univariate Statistical Models	9
2.1.2 Simple Regression Analysis.....	11
2.1.3 Single Discriminant Analysis.....	12
2.1.4 Markov Chain Analysis.....	13
2.1.5 Multivariate Discriminant	15
2.1.6 Multivariate Regression Model.....	17
2.1.7 Logistic Regression Model.....	18
2.1.8 Logit and Probit Models.....	21
2.1.9 Artificial Neural Network	22
2.1.10 Support Vector Machines Model (SVM).....	25
2.1.11 Decision Tree Model	26
2.1.12 Fuzzy Logic.....	28
2.1.13 K-Nearest Neighbor Classifier	30
2.1.14 Random Forest Method.....	32
3. MATERIAL AND METHODS.....	35
3.1 DEEP LEARNING	35

3.1.1	Autoencoders.....	37
3.1.2	Generative Adversarial Networks (GAN).....	38
3.1.3	Recurrent Neural Networks (RNN).....	40
3.1.4	Long Short Term Memory (LSTM).....	42
3.1.5	Gated Recurrent Unit-GRU.....	44
3.1.6	Convolutional Neural Networks (CNN)	45
3.2	PROPOSED CONVOLUTIONAL NEURAL NETWORK BASED METHOD	49
4.	EXPERIMENTAL RESULTS AND DISCUSSIONS	51
5.	CONCLUSION.....	54
	REFERENCES	55

LIST OF FIGURES

	<u>Pages</u>
Figure 2.1: Univariate Statistical Models	11
Figure 2.2: Simple Regression	12
Figure 2.3: Discriminant Analysis.....	14
Figure 2.4: Markov Chain Analysis	15
Figure 2.5: Markov Chain Analysis	16
Figure 2.6: Multivariate Regression	18
Figure 2.7: Logistic Regression.....	20
Figure 2.8: Logit and Probit Regression.....	22
Figure 2.9: Artificial Neural Network	24
Figure 2.10: SVM.....	25
Figure 2.11: DT	28
Figure 2.12: Fuzzy Logic	30
Figure 2.13: KNN.....	32
Figure 2.14: RF.....	34
Figure 3.1: Autoencoders	38
Figure 3.2: GAN.....	40
Figure 3.3: RNN	41
Figure 3.4: LSTM.....	43

Figure 3.5: LSTM Step 2.....	44
Figure 3.6: Gate Recurrent Unit	45
Figure 3.7: CNN	46
Figure 3.8: Convolution Layer	47
Figure 3.9: Pooling Layer	48
Figure 3.10: FC.....	49
Figure 3.11: Proposed Method	50
Figure 4.1: Evaluation Parameters of Binary Dataset	51
Figure 4.2: Evaluation Parameters of Three Classes Dataset.....	52
Figure 4.3: Evaluation Parameters of Four Classes Dataset.....	52
Figure 4.4: Execution Time with Three Datasets	53

ABBREVIATIONS

SVM	:	Support Vector Machine
NN	:	Neural Network
RBF	:	Radial Basis Function
ESD	:	Energy Spectral Density
IoTs	:	Internet of Things
ACO	:	Ant Colony Optimization
GAN	:	Generative Adversarial Network
CNN	:	Convolutional Neural Network
LSTM	:	Long Short Term Memory
RNN	:	Recurrent Neural Network

1. INTRODUCTION

Coronaviruses are beta coronaviruses in the family Coronaviridae. The agent detected in a throat swab culture taken from a patient after the sudden onset of pneumonia cases was reported in Wuhan, China in the last month of 2019, was named a new coronavirus (2019-nCoV) by the Chinese Center for Disease Control and Protection (CDC), and then SARS. - Named as CoV-2 by the World Health Organization (WHO) (1). Due to the high contagiousness of the virus, the epidemic was declared as a pandemic by WHO on March 11, 2020, as a result of its rapid intercontinental spread [1].

The disease shows a clinical course that progresses from an upper respiratory tract infection-like picture to severe lower respiratory tract infection and acute respiratory distress syndrome (ARDS). COVID-19 patients usually present with non-specific symptoms (cough, malaise, fever, muscle and joint pain). The reason for hospitalization of patients who need inpatient treatment is mostly shortness of breath due to pulmonary involvement of COVID19. In the follow-up and treatment follow-up of the patients, renal involvement, cardiovascular and central nervous system involvement, and other system involvements are also accompanied in addition to lung involvement. Early diagnosis and timely isolation of infected patients is very important in controlling the pandemic. However, due to the cases that were asymptomatic and progressed with mild symptoms, it became difficult to control the epidemic. Real-time Reverse-Transcriptase Polymerase Chain Reaction (rRT-PCR) test and antigen test are used in the diagnosis of the disease. The Coronavirus (Cov) family is a group of viruses with enveloped, single-stranded and positive-sense RNA belonging to the class of Nidoviruses, which infect humans and animals. They are spherical in shape and have diameters of 80-120 nm (13). Due to the rod-like extensions on the surface of these viruses, they are named coronavirus (crown virus), which is called crown in Latin [2].

There are 4 types according to their genetic structure. These; Alpha-coronavirus, Beta-coronavirus, Gamma-coronavirus and Deltacoronaviruses (15). While alpha-coronavirus and beta-coronaviruses infect mammals, delta-coronavirus and gamma-coronaviruses are known to infect birds. The coronavirus (CoV) is structurally composed of proteins: Envelope (E) protein,

Transmembrane (M) protein, Nucleocapsid (N) and Spike (S) protein. The envelope (E) protein is the smallest component in the coronavirus structure and is a hydrophobic transmembrane protein. It has an important place in viral pathogenesis. While it is synthesized in large amounts in the replication of infected cells, only a small amount is included in the structure of the virion envelope (20). The transmembrane (M) protein basically ensures the formation of new viral particles in the cell membrane. It is also responsible for removing some proteins of the host cell from the viral envelope during fusion. It is the most abundant component in the structure of coronavirus [3].

When epitope mapping was performed, it was observed that some amino acids of the M protein and the patient's serum affect each other, and it is thought that this situation stimulates the immune response of the host. The N protein, which is the same in the entire coronavirus family, is the main protein of the nucleocapsid. It takes part in the cellular response of the host to viral infection and in the replication of viral RNA. In some studies, it has been determined that the N protein plays an important role in the formation of viral RNA and in the transport of the genome to the virion. In addition, antibodies against the N protein have been detected in some of the patients infected with coronavirus, and it is thought that the N protein can be used in clinical diagnosis. Spike (S) spike protein is a glycoprotein protein with two subunits as receptor-binding subunit S1 and membrane fusion subunit S2, which gives the typical crown structure to the coronavirus with spikes, and both subgroups induce the release of neutralizing antibodies.

It also plays a role in the fusion of viral membranes and attachment of the membrane to cells. The virus can infect the cell by binding to the target receptors on the host cell with the S protein. Different types of coronavirus can use different regions as receptors. For example, MERS-CoV that uses dipeptidyl peptidase-4 (DPP-4) as a receptor to bind to the surface, while those that use angiotensin-converting enzyme-2 (ACE-2) are HCoV-NL63 (Human Coronavirus NL63) and SARSCoV.

1.1 DEFINITION OF SARS-COV-2

The first coronavirus was isolated in 1937 as an infectious bronchiolitis agent in sick chickens, and in 1965, the coronavirus was first detected in humans by culturing a swab taken from the nasal 5 cavities of a person with a cold complaint. There are seven types of coronaviruses that cause infection in humans. These are HCoV-HKU1, HCoV-229E, HCoV-NL63, HCoV-OC43 with low disease-causing effect and mild respiratory symptoms, and severe and mortal respiratory diseases such as MERS-CoV, SARS-CoV, SARS-CoV-2. are subtypes [5].

In the city of Wuhan, China, on 12 December 2019, due to the increase in pneumonias, which are thought to be related to the seafood and live animal market, the cause of which is unknown, it was shown that this pneumonia agent is a new subspecies from the coronavirus family in the first week of January 2020, with the genome separation. It was found to be similar to SARS-CoV and to be a beta coronavirus and was named SARS-CoV-2. It has been determined that the structure of the receptor-binding gene region of SARS-CoV-2 uses ACE2 to enter the cell. Similarly, SARS-CoV uses ACE2, and SARS-CoV-2 has a higher affinity for ACE2. As a result of MERS (Middle East Respiratory Syndrome) and SARS (Severe Acute Respiratory Syndrome) research and finally SARS-CoV-2 research, it has been revealed that these three coronaviruses, which can cause disease in humans, use the same receptor, ACE2. ACE2 belongs to the dipeptidyl-carboxydipeptidase converting enzyme (ACE) family.

It is involved in many physiological functions such as inflammatory response, cell proliferation, hypertrophy, fluid balance and blood pressure. ACE2 is secreted in some tissues and organs. In a study, ACE2 protein level was detected in human tissues and organs including brain, lung, skin, bone marrow, lymph nodes, thymus, stomach, liver, small intestine, spleen, kidney, colon. As a result, ACE2 is released in large amounts in the small intestine and lungs; It was found to be highly expressed in smooth muscle cells and endothelial cells of all other organs. In line with these data, it was determined that SARS-CoV-2 is likely to spread in the circulatory system by blood flow and that SARS-CoV-2 not only affects the respiratory system, but also affects the circulatory system, central nervous system, digestive system and urogenital system and is a potential threat. shows. SARS-CoV-2 infection is characterized by a large amount of

inflammatory response and rapid release of cytokines such as interleukin-1 (IL-1), interleukin-6 (IL-6), interferon- γ (IFN- γ) and tumor necrotizing factor- α (TNF α). triggers the release. Therefore, patients are susceptible to multiple organ failure (MODS) and acute respiratory distress syndrome (ARDS). Cytokine cascades and low lymphocytes play an important role in the progression and severity of the disease.

Recent studies have confirmed the presence of high inflammatory cytokines and low lymphocytes in most of the SARS-CoV-2 cases (36) (37). When the cytokine cascade is triggered for some reason, the prognosis of the disease worsens if the situation is not acted upon immediately. Therefore, cytokine is a serious risk factor for mortality in critical SARS-CoV-2 cases (38). SARS-CoV-2 was first isolated in respiratory secretions; later, RNA of the virus was demonstrated in nasopharyngeal swab sample, saliva, serum, urine and feces. The virus is transmitted mainly through droplets. It can also be transmitted through contact on contaminated surfaces. 7 Coronaviruses are generally unstable to the external environment.

The duration of infectivity can vary with many factors such as the texture of the contaminated surface, moisture content and temperature of the environment. An estimated 3000 droplets are formed in a single cough, and the coronavirus is thought to remain in airborne droplets in active infectious form for about 3 hours. Transmission from person to person occurs by direct droplets at a distance of less than two meters or indirectly through contaminated surfaces. In addition, it has been shown that coronaviruses remain viable and infectious on metal and plastic surfaces for an estimated 72 hours.

The coronavirus has also been found in blood and stool samples, and it has been stated that infectivity in this way is not very important and can be neglected by WHO. In the analyzes performed, no virus was detected in the semen and testicular samples, therefore it was predicted that sexual transmission would not occur. Asymptomatic and pre-symptomatic people are considered contagious because they have viruses in their respiratory tract. Transmission occurs from pre-symptomatic individuals approximately 1-3 days before symptoms appear [6].

In a study conducted in China on this subject, the transmission rate of non-symptomatic individuals was found to be 12.6%. It was observed that the viral load increased rapidly during

the onset of symptoms, and then suddenly decreased on the 7th-14th days following it. It has been observed that the viral load is at least 60 times higher in cases with a severe course of the disease. It has also been found that viral excretion continues in stool samples and nasopharyngeal swab, rarely after the 14th day in more severe cases (40).

In a multicenter retrospective cohort study, it was shown that there is a mean viral clearance of 20 days and a maximum viral clearance of 37 days. In a study on the detection time of viral RNA, viral RNA was not detected in the nasopharyngeal swab sample taken 10 days after the onset of the complaints in 85% (18 out of 21) patients with a mild course of COVID-19. viral RNA detection time was found to be prolonged. According to the COVID-19-related isolation and quarantine practice updated on 14 January 2022 by the General Directorate of Public Health of the Ministry of Health in our country, 9 SARS-CoV-2 PCR tests are positive in asymptomatic or mild cases (those with a stay at home or hospital stay shorter than 24 hours).) isolation period regardless of vaccination status; Persons without symptoms are placed in isolation for 7 days from the first PCR test result, and for those with symptoms, from the onset of symptoms for 7 days.

At the end of the 7th day, the isolation of individuals who have no symptoms or have mild symptoms but have not taken antipyretic in the last 24 hours is terminated without PCR testing. These people should still pay attention to the mask and distance for 10 days. In those with high risk (close) contact with COVID-19, the isolation period is planned in two groups according to the vaccination status and whether or not to have the disease. In the first group; There are people who have had the disease in the last 3 months and have been given a reminder dose of the COVID-19 vaccine.

If these people are in close contact, they are not quarantined. They can continue their daily lives by following the mask, personal hygiene and distance rules very carefully for 10 days. If they have symptoms, they should give PCR testing on the day of symptom development. In the second group, quarantine is applied for 7 days in people who have not had the disease in the last 3 months, have not received a reminder dose of the COVID-19 vaccine, or have not been vaccinated. In the absence of symptoms, quarantine is removed without testing on the 7th day.

However, distance, mask and hygiene rules should be observed for at least 10 days. When the symptom develops, the COVID-19 PCR test is given on the day the symptom develops. Diagnosis is made using clinical, laboratory and radiological features.

First of all, the possibility of COVID-19 infection should be considered in patients with newly emerging fever and / or shortness of breath, cough. It should also be considered in patients with severe lower respiratory tract disease of unknown cause, diarrhea, myalgia, malaise, smell or taste disorders. Currently, the gold standard diagnostic method for detecting viral RNA for COVID-19 is reverse transcriptase polymerase chain reaction (RT-PCR). method. 10 The Ministry of Health of the Republic of Turkey has published a COVID-19 guideline that is constantly updated with new information.

In this guideline, the diagnosis of COVID-19 is made by nasal, oropharyngeal swab sample or nasopharyngeal washing from the upper respiratory tract and by sampling with bronchoscopy or tracheal aspirate from the lower respiratory tract.

Although the RT-PCR technique is the gold standard diagnostic method, its sensitivity is between 50-62%. For this reason, many studies have been conducted that have determined that a negative PCR test will not rule out COVID-19 disease. In a study conducted in China, a total of 1014 patients were recruited, and 601 patients had positive PCR tests and covid pneumonia findings on thorax CT, with a thoracic CT sensitivity of 97%. PCR test was negative in 413 patients, and typical findings were found in thorax CT in 308. In addition, in 15 patients with negative PCR test and thoracic CT involvement, the subsequent PCR test was found to be positive [7].

Another diagnostic method used in the diagnosis of COVID-19 is ELISA (enzyme-linked immunosorbent assay). With this method, immunoglobulins developed against SARS-CoV-2 are detected. Although there is no clear information yet, it has been observed that IgM antibodies are detected in the first five days of infection, it can still be high in the 2nd and 3rd weeks of the disease, and IgG antibodies are detected approximately 14 days after the onset of symptoms.

1.2 CONTRIBUTIONS

This study present several contributions as listed below:

- i. The first contribution is combining CNN with ALO and GA which is new model and presented remarkable results.
- ii. The proposed method presented very low execution time when compared with traditional techniques.



2. LITREATURE REVIEW

2.1 MACHINE LEARNING

Machine learning is a set of algorithms or techniques that increase the ability of systems to perform tasks and solve problems without human intervention. In other words, they learn to solve problems on their own using these systems and methods. Machine learning is used in many applications, from computer vision and robotics to autonomous driving and medical surgery.

Machine learning was born out of the question: Can computers learn to perform certain tasks, order food and do more than we know? Can computers surprise us? Instead of programmers manually creating data processing rules, can computers present data and learn those rules automatically? These questions opened the door to new programming paradigms. In classical programming, an AI signal model creates rules for human (software) input, data, and responses processed according to those rules. In machine learning, a person enters data and 64 expected responses come from the data, and the rules appear.

The diagram below presents a new programming notation for machine learning. Machine learning systems are trained without explicit programming. Various task models are provided that create statistical structures that allow the system to eventually find rules that automate the task. For example, if you want to automate the task of tagging vacation photos, you can feed the machine learning system many examples of photos that people have already tagged, and the system will learn statistical rules that match photos with specific tags.

Here we need to determine the distance between the current output of the algorithm and the expected output. The metrics are used as feedback signals to tune the performance of the algorithm. This adjustment step is called training. A machine learning model is the process of transforming input data into meaningful "learned" output data without using known input and output examples. Therefore, the key to machine learning and deep learning lies in transforming data in useful ways. That is, learn a useful representation of the input, a representation that approximates the expected output.

What is the concept of representation? Basically, it's a different way of looking at data. Tools for representing or encoding data. For example, a color image can be encoded in RGB (Red-Green-Blue) or HSV (Hue Saturation Value) format. These are two different representations of the same data. Some tasks that may be difficult to complete may be easier to complete others. For example, the task of "selecting all red pixels in the image" is simpler in RG format, while "making the image less saturated" is simpler in HSV format. Machine learning models are all about finding suitable representations for input data.

It is the process that makes data more suitable for the task at hand, such as the task of classification. After giving general information with machine learning, information will be given below about the k nearest neighbor classifier and random forest methods, which are among the machine learning techniques used for analysis in the method part of the thesis, which are more preferred in the prediction of financial failure [8].

2.1.1 Univariate Statistical Models

Univariate statistical models are statistical models used to predict financial failure. In these models, financial ratios are considered individually to measure the financial performance of the enterprises and they are the models that contribute to the financial information providers about the success or failure of the enterprises by making an assessment of the financial situation of the enterprise. The main purpose of these models is to evaluate the relationship of this variable with a single independent variable and reach a conclusion by accepting that the variable in question is a single independent variable and all units are considered constant [9].

In other words, this method is a method used to make predictions about the future of the enterprise by evaluating the success and failure of the enterprise with the help of financial ratios one by one. In this model, successful and unsuccessful businesses are determined according to their maximum profitability, and it is aimed to determine the ratios that differ the most between the financial ratios in the successful and unsuccessful businesses. Since the weights of the financial ratios used in determining the failure cannot be fully revealed here, the results cannot reveal a clear idea about the future of the enterprise.

The univariate analysis model is based on a number of restrictive assumptions. In particular, univariate analysis assumes a linear relationship between rates, independent variables, and institutional failure, the dependent variable. However, this assumption may not hold in practice, thus making any result from univariate analysis questionable. For example, a company is classified as failing at a certain rate and not failing by a different rate. This limitation is called the “inconsistency problem”.

The criticism leveled at univariate analysis is that a single ratio alone cannot cover the nonlinear and multidimensional situation of a company. Also, in a univariate analysis, the optimal intersection point is selected retrospectively on a "trial and error" basis. This means that the cut-off point is sample-specific and thus the true preliminary prediction accuracy of the univariate model is less than that obtained. Moreover, univariate analysis fails to take into account the relative cost of Type I and Type II errors.

Statistical techniques used in univariate statistical models are simple regression analysis, single discriminant analysis and markov chain analysis. Criticisms of the one-dimensional model are;

- Since only one variable is used in this model, accurate analysis cannot be made and false information is presented to financial information providers.

The effect of the financial ratio used in this model may be low, but in another model, the effect of this financial ratio may be higher. - In this model, it is impossible to evaluate the entire enterprise analyzed. Since these methods have lost their importance in the literature and left their place to multivariate statistical methods, they are less mentioned in this section. In addition, studies on univariate statistical models will be mentioned in the literature section.

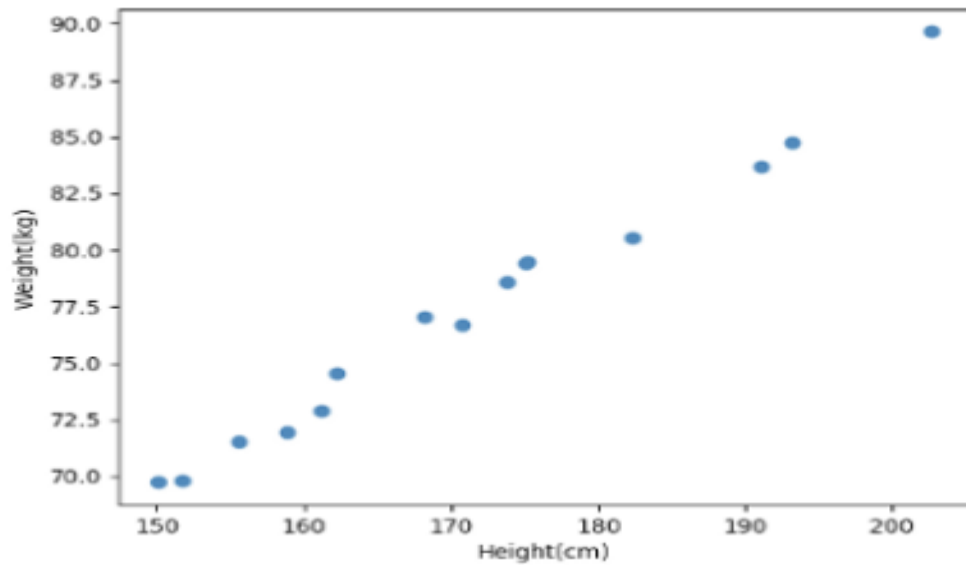


Figure 2.1: Univariate Statistical Models.

2.1.2 Simple Regression Analysis

It is a statistical analysis used to estimate the relationship between a normally distributed dependent and an independent variable. This analysis is generally an equation represented by the function $y=a+bx$. While the “a” shown in this equation represents a constant term showing the value of y when $x=0$; If “b” is a unit in x. It expresses the regression coefficient showing the average change caused by the change in y.

In other words, only one independent variable is used in this analysis. And these variables are in a cause and effect relationship with each other. If we explain with an example; The cheap price of a product (independent variable -x) may affect the sales of that product (dependent variable -y). Or, the quality of this product (independent variable -x) may positively affect its promotion in the market (dependent variable -y).

As can be understood from this equation, simple regression analysis explains the relationship between two variables. In the simple regression analysis, the observation values belonging to the variables included in the equation are taken into consideration and the model to be established in the study to be calculated is estimated.

It is important in the application of this method that the distribution between the variables to be considered in the model is in the normal distribution. The reason why it is not used very often today is due to the fact that the distribution of the financial ratios considered in the failure estimation is not normal. However, since the financial ratios used in financial failure predictions do not show a normal distribution, simple regression analysis is not used very often in today's financial failure predictions [10].



Figure 2.2: Simple Regression.

2.1.3 Single Discriminant Analysis

In general, discriminant analysis is to determine more than one combination of the variables to be used and to create one or more functions that measure the levels of differences between groups.

In other words, discriminant analysis is expressed as univariate or multivariate analysis in terms of the number of variables they use. If a single variable is used in the analysis, it is defined as single discriminant analysis, and if more than one variable is used, multivariate analysis is defined as discriminant analysis. Since the results will not fully reflect the reality in a failure analysis using single discriminant analysis, the use of this analysis will affect healthy results [11].

Since it is understood that the failure estimation to be made by taking one financial ratio into the analysis does not give healthy results, it has lost its prevalence.

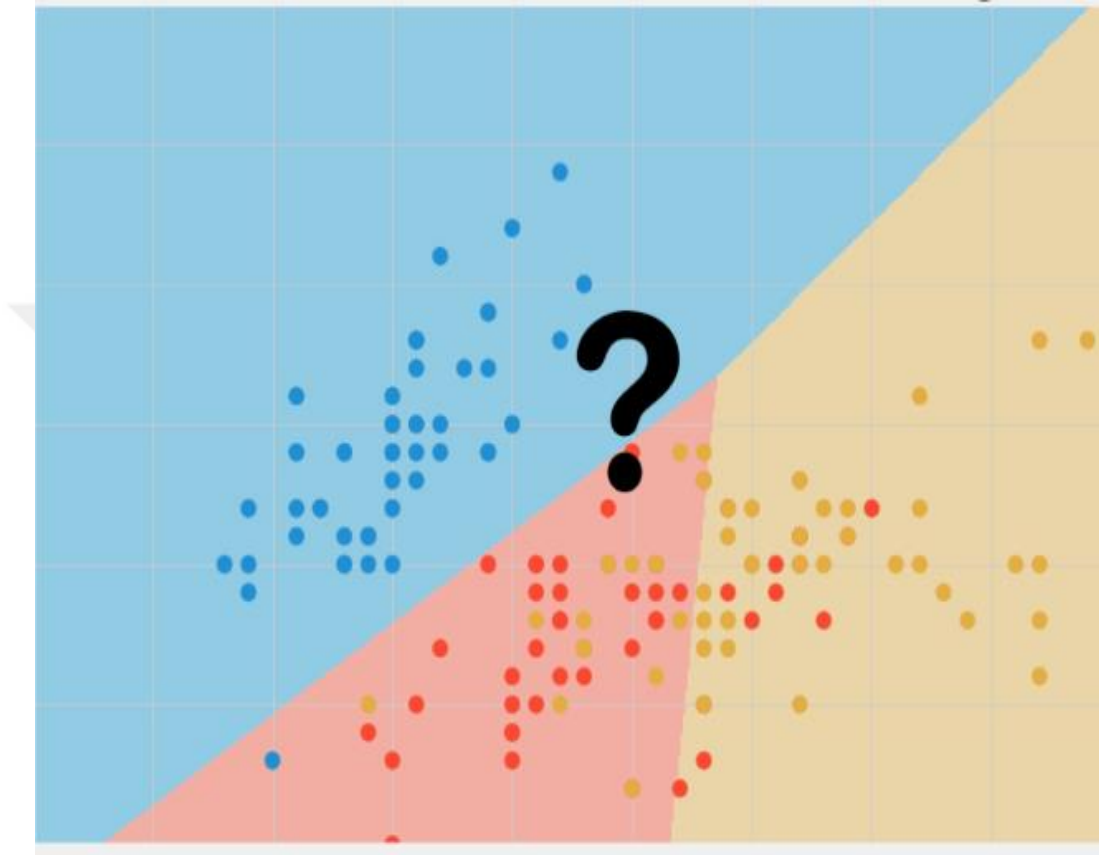


Figure 2.3: Discriminant Analysis.

2.1.4 Markov Chain Analysis

Markov chain analysis is an analysis process that provides information on the conditional probability of its future behavior by examining the past and present state of an event. The basis of this analysis was revealed by Markov's work, also known as the Brownian method, at the beginning of the 20th century.

It is known that Markov analysis is used a lot in predicting future events. As it is known, most of the events in our lives change over time. In analysis, if the system explains more than one value at the same time, this system is called dynamic system. When the transition probabilities depend only on the current state of the system and not on the previous states, the process is

called a "Markov process". It is a fundamental aspect of the Markov model and is often called the Markov hypothesis. Because of the Markov hypothesis, the Markov model states that it is not necessary to know the past state of the system to predict the future. Markov processes and extensions of Markov chains, called Markov decision processes (MDP), are processes that allow designers to apply tasks to a system and interact with objects in the system.

Applying a task to the system can be thought of as a modification of the transition matrix for a certain number of time steps. PDM is used to study the effects on decision making when the outcome is partly random and partly under the control of the decision maker. Another reason for the widespread use of Markov chains is that systems created by evaluating variable and probabilistic systems can lead to changes between different states over time. Another characteristic of this method is the ability to predict the immediate state of the created system and the situations that will arise in the long run. This method is not widely used today, but since it is a simulation model, it is called an assessment tool.

Multivariate statistical models The univariate statistical models above lost their reliability due to low success rates or showed different results in the same study and were relegated to multivariate statistical models. Using this model, we analyzed the financial structure of the company in detail. After Altman (1968), most subsequent researchers have argued that a model for predicting percent failure was too simple to reflect the complexity of financial failure.

As a result, many have begun to use multivariate statistical models to derive estimation formulas involving combinations of ratios. In univariate models, the use of a single financial ratio remains low in explaining the results, while the explanatory power of the results is higher in models created using more than one ratio.

Moreover, considering the periods when univariate models started to be used, adequate analysis could not be done in studies with large data sets, while the accuracy of the analyzes to be made increases with the development of the analysis techniques used today, no matter how many variables are used. The use of these models has become widespread because the results obtained from multivariate models, which emerged due to the criticized aspects of univariate models, are more reliable and stable [12].

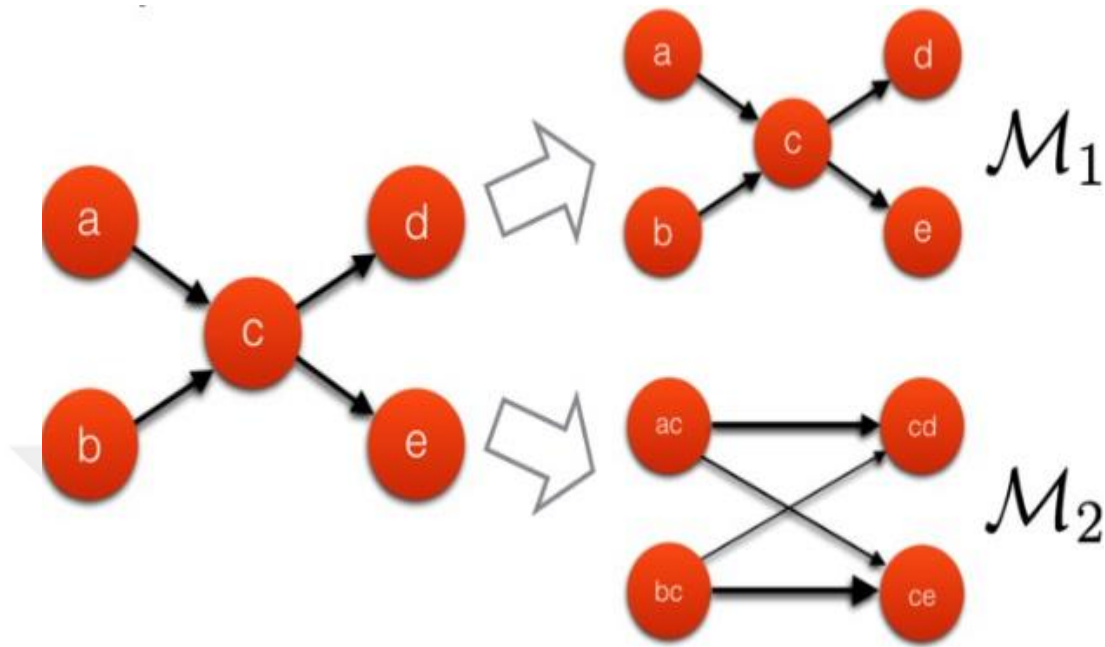


Figure 2.4: Markov Chain Analysis.

2.1.5 Multivariate Discriminant

Altman analysis and z-score determination analysis are commonly used to determine the contribution of each selected variable in a group by finding a linear connection of "k" measurement variables. The minimum number of linear connections used is 1. You can also use a linear or auxiliary function to assign or predict a new function to one of the predefined sets. You can also use this model to determine the contribution of multiple variables or to investigate the relationship between multiple variables and specific cases. After discriminant analysis, a mathematical model can be built. This model can be used to estimate group membership. Multivariate analysis consists of several methods used to analyze the "k" values of data for many variables measured in each experimental unit [13].

Researchers are often interested in determining the contribution of individual variables to splits between different groups. This analysis is a multivariate statistical technique used to classify observations based on two or more independent variables. This method determines which group the annotation belongs to. Multivariate discriminant analysis classifies observations into a given group and creates a linear set of explanatory variables called a discriminant function or

classification function, assuming that the probability of a correct classification is highest. It is also designed to maximize the sum of squares/fist between groups. It is designed to compute the first distribution function for optimal group separation. The second class function is computed to better separate the groups and has nothing to do with the first. Therefore, a calculation is performed to find the maximum number of class functions [14].

Multivariate discriminant analysis (MDA) remains the most widely used method for solving classification problems.

It uses discriminant analysis to determine the best financial performance that distinguishes failing companies from healthy ones, and annually examines the predictive performance of various models. The MDA model consists of a linear combination of scores. The discriminator function aims to better distinguish between failed and healthy companies. The application of MDA is fundamental in the assumption that the multivariate explanatory variables have a normal distribution and a covariance matrix of equal variance between the normal and rejection classes. However, it is consistently claimed that most BFP studies based on MDA (a classic method of statistical forecasting often used in corporate bankruptcy) have not been tested to see if the actual data matches the assumptions.

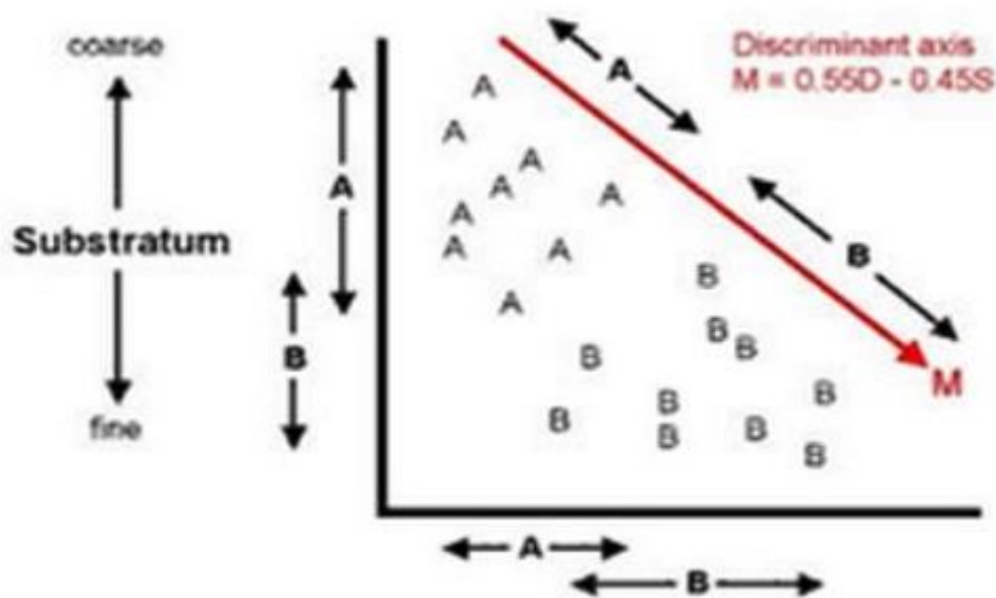


Figure 2.5: Markov Chain Analysis.

2.1.6 Multivariate Regression Model

Regression analysis is at the heart of applied statistics. A regression object is a continuous response based on one or more predictors. Univariate evaluation. If the model contains only one predictor variable, then it is a simple linear regression. When there is more than one predictor, it is called multiple linear regression. Multiple linear regression attempts to formulate a combined prediction of multiple predictors for a single response variable. Multiple linear regression differs from simple linear regression in that the model contains multiple predictors. Multiple regression evaluates the contribution of one predictor variable while emphasizing the influence of another predictor variable.

Linear regression is an extension of Pearson's correlation. Correlation determines whether two variables are correlated, and regression determines whether one variable can predict the outcome of another. Since we are more interested in predicting some variables than others, we can specify which variables are dependent (outcome) and independent (predictive) variables. Unlike correlation analysis, regression variables must be classified as predictors or outcome variables. Simple linear regression uses a single predictor variable to predict or predict an outcome variable. The standard equation for simple linear regression is $Y_i = (b_0 + b_1X_i) + e_i$. The optimal straight line is determined by the point closest to the slope of the straight line or the regression coefficient ($b = 1$) of the estimator. This line intersects the vertical Y axis (b_0). Y_i is the expected outcome variable, X_i is the participant's score for the predictor variable, and e_i is the difference between the optimal error or residual line and the participant's actual observed score. Multiple linear regression is called "multivariate" linear regression. It can be confused with the term "multivariate", which is used only when there is more than one outcome variable [15].

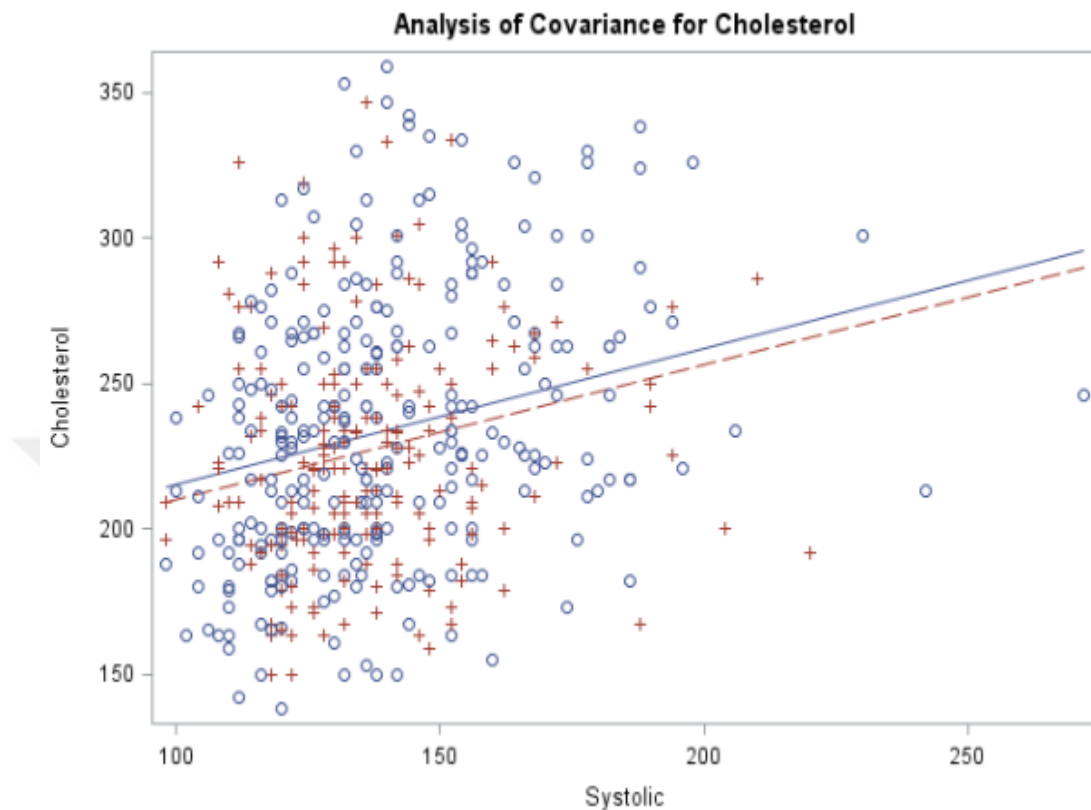


Figure 2.6: Multivariate Regression.

2.1.7 Logistic Regression Model

It is suitable for models dealing with decision problems and is often used in statistical analysis found in the economic and financial literature. Regression models serve two purposes. (1) Outcome variables can be evaluated against new expectations. (2) The settings of each predictor automatically map that variable to the outcome variable, automatically testing the effects of other predictors. then. Logistic regression models allow you to investigate the effects of many independent variables. Dependent variable Y . This variable Y takes only 2 values and is divided by 2. These values are encoded as 1s and 0s. A value of 1 in variable Y means that the target event has occurred. Otherwise, when an unwanted event occurs, this variable is set to '0'. The regression process can help you identify the most important factors to ignore and how they affect each other. Logistic regression uses logistic functions with outputs in the range (0, 1) for simplicity.

A simple least squares linear regression equation produces a sigmoid curve if the model violates the model constraint. You need to create a curve model that represents your data set. A simple logistic regression model can easily be extended to three or more independent variables. In other words, the more variables you have, the more difficult it is to get multiple observations for all variables at all levels.

Therefore, most logistic regression is performed on multiple independent variables using the maximum likelihood method. In logistic regression, the dependent variable is a categorical variable, but the independent variable provides a convenient way to estimate the economic burden. Logistic regression is often used for two reasons. Second, it is not necessary to satisfy statistical assumptions such as linear regression.

When the MDA hypothesis is formulated, logistic regression and MDA are coupled. The record format depends on the global logistics function. Therefore, a company may belong to the pre-financial analysis group. Logistic regression is useful when estimates are based on values of independent variables. Logistic regression is a nonlinear regression defined by a binary dependent variable. Use logistic regression to determine the OR of each independent variable used in the model.

This regression allows you to set up a multivariate regression between the dependent and independent variables. The result of this regression analysis determines the probability of two responses (0 and 1) or (yes or no) depending on the definition of the two response values. Logistic regression predicts a specific outcome. Group B membership based on binary, continuous, discrete, or combinations thereof. Response variables are usually categorical. So it is classified as 'None/Available' or 'Fail/Pass'. 43 The logistic regression (LOG) model systematically estimates the likelihood of fraud based on observed tissue characteristics, using two extreme kurtosis probabilities as dependent variables.

There are known structural relationships (additive and linear) between the observed characteristics of the firm and the underlying probabilities. Model parameters are usually estimated using the maximum likelihood method. LOG greatly simplifies interpretation of results by reducing some of the assumptions required for discriminant analysis (covariance

matrix and multivariate normality) and providing different hypothetical probabilities to individual borrowers. However, many restrictions apply. There are non-linear or complex interactions between observations and standards, which makes it difficult to maximize large data sets with high sensitivity to outliers or missing data (Moscatelli, 2020: 3). Logistic regression is a technique designed to overcome the shortcomings of multiple ANOVA, but it has some drawbacks. First, he said, there are studies that do not take into account the cost of type I and type II error rates when determining the optimal threshold for log models. However, since the cost of misclassification is subjective, most researchers implicitly assume the same cost of misclassification as the overall error rate drops. The problem of multicollinearity (when there is a high correlation between specific independent variables) in a logistic regression (LA) model.

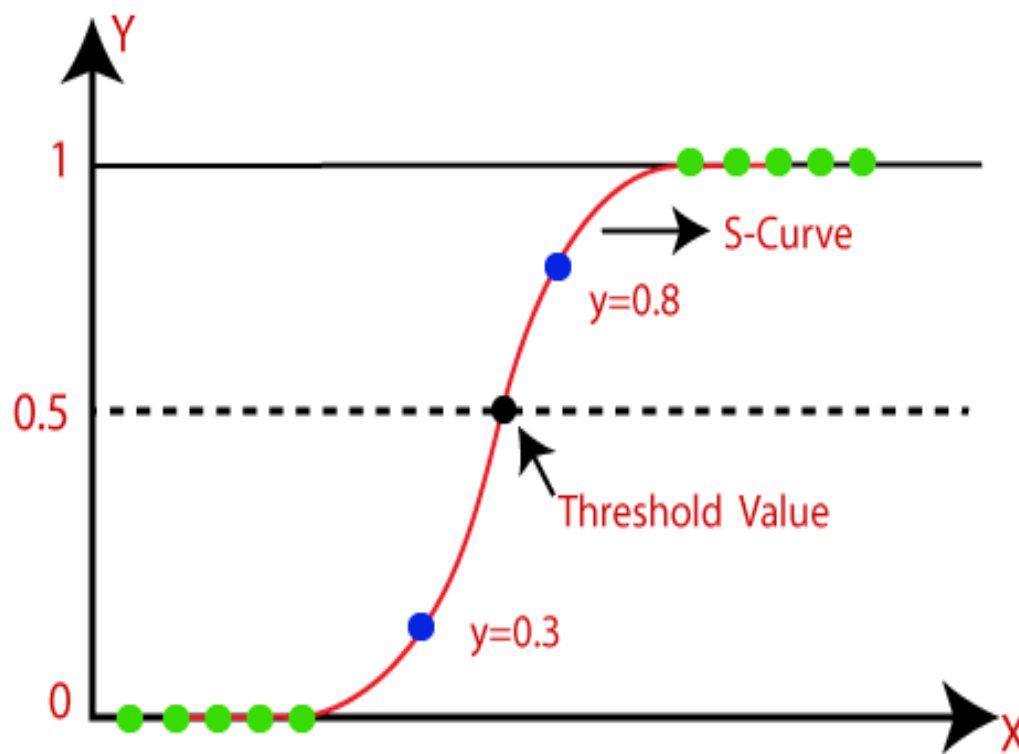


Figure 2.7: Logistic Regression.

2.1.8 Logit and Probit Models

Logistic regression has been the standard method of mathematical statistics for decades. It is useful when the dependent variable is not continuous, binary, or alternating. If the explanatory variables are continuous, then classical linear regression is used here. This means that only two values are allowed. The variable y (the independent variable), which must be written as a ratio, takes discrete values. 1 if it works, 0 if it fails. Therefore, instead of "classical" regression, you should use a modified regression analysis called logarithm and probability models.

The logarithmic model is characterized by estimating the probability of an event occurring or not occurring. Thus, the calculated probability is 1 or 0. To create this condition, we need to perform a logarithmic transformation in logistic regression. This logit transformation is based on the "happy hope relationship". This transformation allows for an ideal relationship between the dependent vector variable y and the independent variable x . The probability y is close to 0 when the values of the independent variable are very small, and the probability y is close to 1 when the values of the independent variable are large.

Logistic regression uses variables defined by categories. There are different classifications in the literature. These classification variables are nominal (classifier), rank (ranking), rank (discrete) and modulus (proportional). When estimating models with dependent binary variables, nonlinear models such as logarithm and probability are used to overcome performance problems and common specifications of least squares methods.

These models perform two functions. First, the effect of changing the independent variable depends on the initial value of the dependent variable (i.e. the probability that the dependent variable will take a certain value). Secondly, these models are "interactive" in all their various forms. The effect of changing the independent variable on the dependent variable depends on the values of all other independent variables. These characteristics suggest that attention should be paid when examining systematic differences in stimulus sensitivity between responders (hereafter, response heterogeneity) or interaction effects between model variables [18].

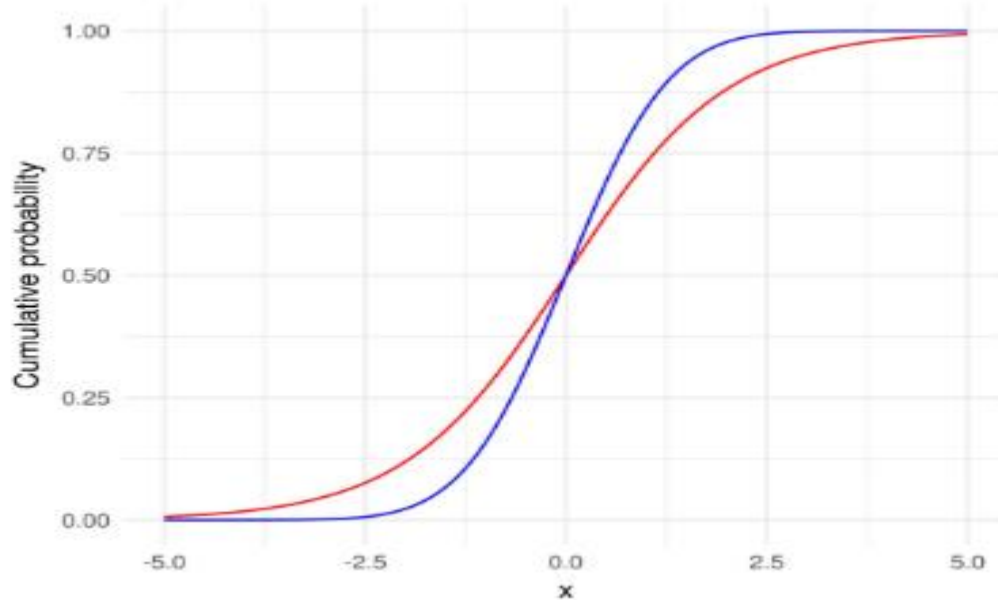


Figure 2.8: Logit and Probit Regression.

2.1.9 Artificial Neural Network

Artificial neural network models The creation of intelligent systems that can simulate human behavior has received worldwide attention in recent years. Therefore, it is not surprising that technologies such as neural networks are showing great interest in pattern recognition. General biological neurons are described to understand the concepts used in artificial neuron models. This basic concept describes a simple neural model that includes the basic elements of a neural network: input layers, hidden layers, output layers, connections between layers, types of neural networks, transfer functions, learning algorithms. ANN is a parallel computing model consisting of highly adaptive processors.

These networks are complex parallel implementations of static or dynamic nonlinear systems. Neural networks are designed to model the regulatory principles of the central nervous system, and the computational power of biologically inspired ANNs enables them to perform cognitive and sensory tasks more easily and successfully than traditional sequential processors. An important feature of these networks is their ability to adapt to "code learning by example" when solving problems. With this feature, these computational models are very interesting for

applications where training data can be easily obtained before the problem to be solved is fully understood. In general, neural networks can be thought of as "black box" devices that accept inputs and produce outputs. Their uses are nearly unlimited, but they can be grouped into a few simple categories: data classification, modeling/prediction, and systems administration. Advances in computing, learning, and artificial intelligence include the development of powerful computing skills, closed-access programming of inaccessible systems, and an attempt to learn and simulate the functions of the human brain.

Artificial Neural Networks are computer structures based on biological simulations of the central nervous system. Networks consist of many interconnected processing elements (neurons) that work in parallel to solve a particular problem. The main advantages of neural networks are their nonparametric techniques, their ability to process incomplete or complex data, and their ability to approximate complex nonlinear maps and continuous functions. A network usually consists of three layers: an input layer, a hidden layer, and an output layer. The leading edge represents the weight and direction of information flow in a network.

ANNs are often used to evaluate economic loads due to their non-linear and non-parametric adaptive learning properties. The first attempt to predict errors using neural networks was made in 1990. The national neural network receives information from the central nervous system. It consists of several interconnected levels. Each level consists of several interconnected nodes. There is a common form of artificial neural network known as multi-layered perceptual, where each node and layer consists of feedback. External information, consisting of key financial numbers, is entered into the input and problem-solving layer or, if there are no problems, generated in the output layer. Several layers are hidden between the input and output layers. Each node in the network can be thought of as a small computer system that accepts and converts inputs and provides outputs.

The input layer requires some financial cost in the network. In the hidden layer, the activation values for the input nodes are weighted and aggregated during the learning process. The second becomes the activation value of the node through its transfer function. This is the input to the output layer node. Finally, a baseline is created that compares the work to the desired value

(mean problem or not). Some ANNs add direct connections between the input and output layers for more flexibility. The foundation stage is a very important step in the use of artificial neural networks. There are some limitations such as univariate analysis, multivariate discriminant analysis, logistic regression analysis, logarithms and probability analysis, multilinear steps, probability distributions and nonlinear relationships. If the data used or the model developed does not agree with the above assumptions, the results may be questionable. Therefore, many researchers believe that neural networks can be a useful alternative methodology if the above limitations can be completely ignored. According to some studies, neural networks outperform multivariate analyzes due to their relatively low error costs. Moreover, we found that neural networks are more effective in explaining the complex relationships between variables than logistic regression analysis [19].

We concluded that neural networks should be used for more complex dependent variables, considering only the accuracy of the predictions. After describing how artificial neural networks work, it is clear that how this system works and the paths between networks have not yet been fully described. Therefore, the network cannot be managed after an ANN is released. However, studies clearly show that the results obtained when analyzing in this way are superior to other methods.

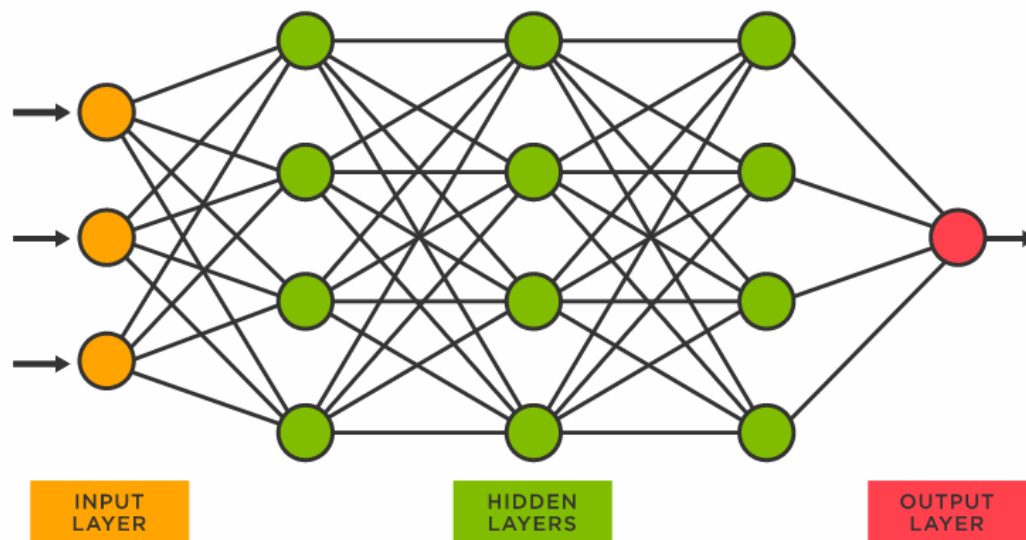


Figure 2.9: Artificial Neural Network.

2.1.10 Support Vector Machines Model (SVM)

Support (SVM) is a leadership-based method with good generalization. It is designed to use empirical risk to design SVMs as mechanistically risky. The Support Machine (SVM) is a powerful Bell technology developed in 1995 by Vapnik and Labs. It is a basic static learning method between SVM and local connection. Reduce the risks associated with venture capital. This greatly increases the complexity of the model (i.e., rich SVM capabilities, constrained solution of quadratic linear problems, and integral optimization). For Current Neural Networks and Review: VM SVM SVM prediction includes functional regression models, sequence structures, contexts, and models. This role is one of great teachers and students.

A double trick in aging education. H is usually an appropriate value for the class. DVM is a non-parameter user. It must be able to determine the relational information or relationship between input and input. This is also important. For smaller virtual machines, we recommend using the extension plugin. There are also innovations in using these core functions: processes that are difficult to scale but can easily handle large amounts of data. SVM is also suitable for outliers or large amounts of data. This added value is very important. In many ways, these are really beautiful and visual reviews [20].

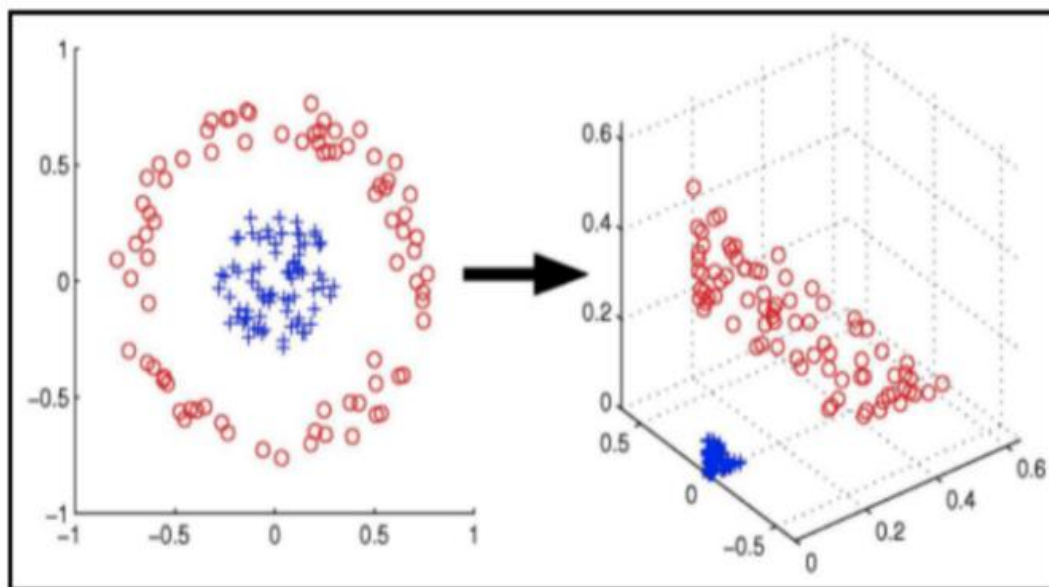


Figure 2.10: SVM.

2.1.11 Decision Tree Model

Decision trees, one of the data mining techniques, have emerged as a method frequently used in recent years in predicting financial failure. Decision trees are a popular way to provide classification and scoring capabilities. Focus on your data and use inductive reasoning to create a spatial tree model. Decision trees are a common technique that has been used successfully in real world problems. Decision trees can be modeled from databases that contain numerical and categorical data. The main modeling algorithms for decision tree analysis include interactive discrimination models 3 (ID3) and C5.0, classification and regression trees (CART), and automatic interactive chi-square detector models (CHAID).

The ID3 algorithm creates a simple tree structure by choosing the best way to find information as the first function of the branch decision tree. However, data collection is used as criteria for selecting features for segmentation, and the ID3 algorithm has the disadvantage that the feature focuses on large values. Therefore, if there is only one data value left in the subtree after partitioning the database, the data gain will be large and the partition size will be small. The C4.5 ID3 algorithm integrates class rules and arithmetic methods. Instead, we use probabilities as a feature-separating metric to mitigate the shortcomings of the ID3 hash node. The C5.0 algorithm is a commercial version of C4.5 and enhances the C4.5 rule generation process.

It is particularly suitable for managing large databases. The C5.0 algorithm is faster and more memory efficient than C4.5. The tree classifier divides the sample space into multiple subgroups and a CART approximation. CART and CHAID provide a set of rules that can be applied to unclassified databases to predict records with specific results. CART divides the database into bivariate partitions and CHAID partitions the database into multidimensional partitions using the chi-square test. CART generally requires less data to be configured than CHAID. The CHAID algorithm is based on the Chi-Square test and is generated by dividing a portion of the region containing all records into several child nodes.

Combine the pairs of allowed expected classes for each node to determine the best distribution in which there is no significant difference between the pairs of target variables. This CHIAD algorithm handles interactions between independent variables that can be obtained directly from

tree traversals. Although CHAID is not an ideal method for achieving optimal clip size, it can help researchers correct clip size discrepancies to find the optimal clip size. Of course, for large databases, the CHAID algorithm determines which hash variable comes first. Decision trees are one of the most widely used methods for providing classification and prediction capabilities. Focus on your data and use inductive reasoning to create a spatial tree model. Decision trees are a common technique that has been used successfully in real world problems. A decision tree can be modeled as a database containing numerical and categorical data.

The main modeling algorithms for decision tree analysis include ID3, C5.0, CRAT (Classification and Regression Tree) and CHAID (Automatic Interactive Chi-square Detector) models. The following diagram shows the structure of the decision tree in a simplified form. Here the values of the analyzed objects are displayed, grouped by arrows from the root to the leaves of the decision tree. Popular decision tree learning algorithms use iterative subdivisions and winning strategies to build decision trees.

A child node is created for each subsection and the subsections are linked to the subsections. This process is then repeated for each child node data until the termination criteria are met. Decision tree or classification tree is a non-parametric technique for data mining. A decision tree is created through an iterative process that separates information as it moves up and down. Build a business failure prediction tree that ranks each company as either successful or failed to predict financial failure. Moreover, each unordered node represents a partitioning rule [21].

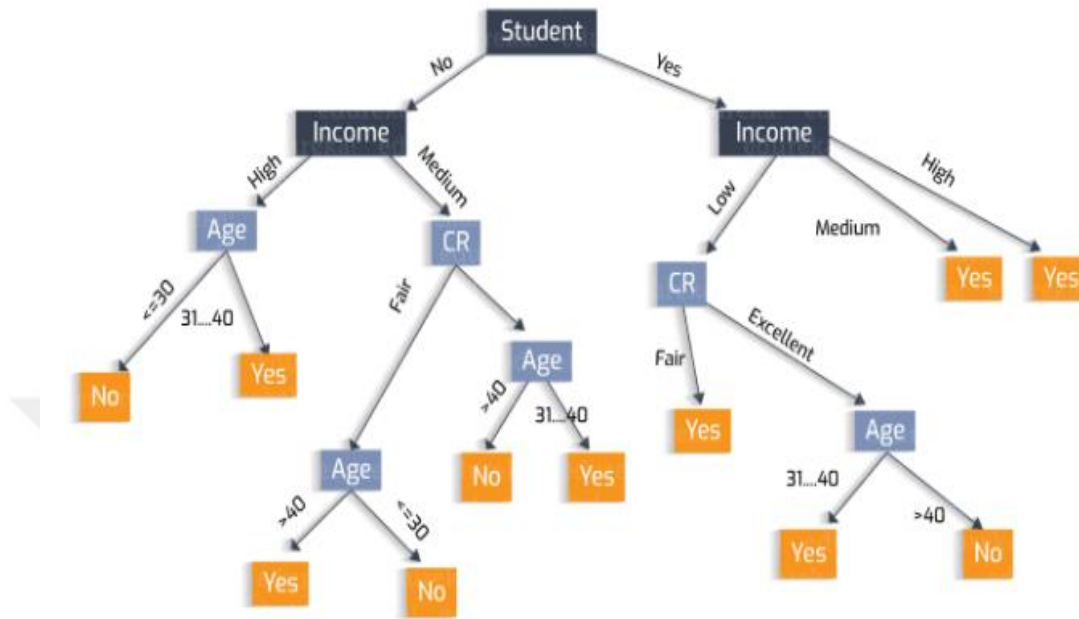


Figure 2.11: DT.

2.1.12 Fuzzy Logic

Fuzzy logic models are used to model erroneous thinking, such as logical reasoning about uncertain and complex processes. Fuzzy group theory is similar to human reasoning using raw data to make decisions. An important contribution of fuzzy control theory is that it can solve many practical problems that conventional control methods cannot adequately solve. Fuzzy models are based on rule-based system descriptions and use conditional rules to express relationships between variables. A set of 60 complex files was used to further define the variables and test the performance of the model.

A fuzzy set is defined as an organic function (denoted by the symbol μ) following the spatial element under study during the period $[0, 1]$. The specified methods use status values or I/O data pairs. The system obtains initial conditions and computes the error between the actual system output and the desired output. This error is used to determine the membership function value and reuse the next input. Equilibrium is achieved when the radial work does not change significantly. The fuzzy model thus obtained is the model used for complex systems [22]. In

short, fuzzy logic is a framework of knowledge that is able to represent the way people understand the world, think about problems and situations, and formalize qualitative and ambiguous concepts. Incorporating the fuzzy logic of evaluation and decision making is interesting. It is a reliable method that attracts managers, experts and analysts. Linguistic characteristics such as “low,” “high,” “good,” “very good,” “positive,” “interesting,” “effective,” and “appropriate” can be formulated with ambiguous logic. You can use multiple properties of the same variable and display them on the same graph. The main steps in the design of the fuzzy system:

- i. Selection of fuzzy system type; A modular system consisting of several interconnected fuzzy modules is chosen.
- ii. Define input and output variables and define linguistic properties (fuzzy values) for each variable using membership functions (unobvious inputs and outputs).
- iii. Definition of fuzzy set of rules.
- iv. Fuzzy The inference method is chosen.
- v. The fuzzy output is converted to a "clear" value.
- vi. The fuzzy system is tested: Modification of membership functions and fuzzy rules if necessary, adjustment of the fuzzy system, verification of results by sensitivity analysis.

When creating fuzzy expert systems, it is very important to use fuzzy inputs to create fuzzy rule blocks. A common method is to rely on information obtained from interviews with experts on the subject or generally accepted information. In recent years, the number of scientific studies with fuzzy logic in the economic and financial sector has increased significantly. There are studies showing that fuzzy logic can be used for safe and profitable applications in business, finance and industry.

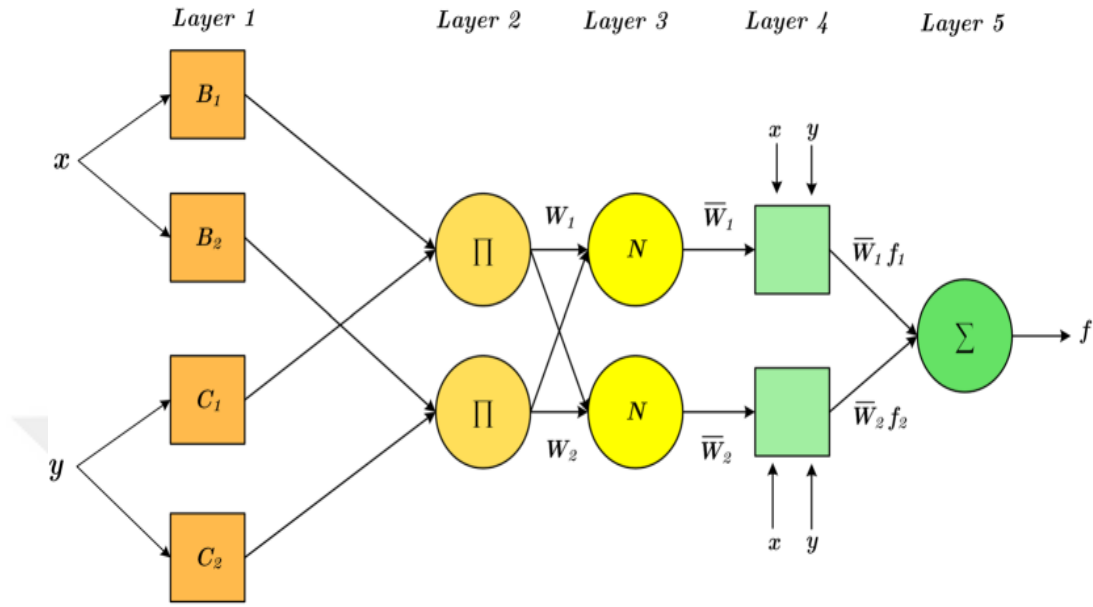


Figure 2.12: Fuzzy Logic.

2.1.13 K-Nearest Neighbor Classifier

K-Nearest Neighbor Classifier KNN, which is a non-parametric and lazy learning method, has been examined for regression and classification. The model structure is non-parametric because it is based on a dataset and makes no assumptions about the distribution of the data. This feature is useful because realistic datasets do not necessarily match theoretical assumptions. A lazy algorithm means that all training datasets are used in the testing phase. This guide is simple and intuitive. Therefore, it is widely used for classification due to its short computation time and easy interpretation. ANNs are used in a variety of industries, including finance, healthcare, and handwriting recognition. Credit scores are used by financial institutions or banking institutions to evaluate a customer's credit information and determine whether the customer is eligible for a loan. The structure of the National Institute of Statistics Research is as follows: Estimate the signal by calculating the distance and finding the nearest neighbor [23].

In this method, K is the method's nearest neighbor and control variable. Each dataset contains its own optimal amount of K . If K is small, the noise effect will be large, and if K is large, the computational cost will be large. Studies have shown that small K values are flexible, with large

diffusion and low distortion, while large K values are inelastic, with low diffusion and high distortion. The advantage of this method is that the ANN training phase is much faster than other classifiers and it is quite intuitive and easy as it does not require recruitment. ANNs help solve multilevel and regression problems. The weakness of this method is that it requires a large amount of memory to store all the training datasets, so the next step of checking K queries is expensive and slows down processing. ANNs are sensitive to unusual sizes and values because they select neighbors based on distance. Large functions take precedence over smaller functions. It also doesn't handle large amounts of data, so scaling is required to improve performance. Also, this method does not solve the missing value problem.

The last neighbor entered for the first time can be considered as a modelless classification method. This is because the new condition classification is entirely dependent on the data points in the training set. The training set is denoted by N . $\{(x_{t-1}, R_t)\}_{N_t=1}$, where x_{t-1} is the feature value vector and R_t is the polynomial response given by the equation. To rank the new XN data points, we need to find the k data points closest to the new data points in the training data based on distance measurement. Euclidean distance is the most used alternative. These k data points are called new neighbors XN . The final ranking is based on the majority votes of the last k neighbors. This process is repeated for each data point in the test set. This shows how the final K classification works on a small three-class synthetic dataset. The left side of Figure 7 shows the nearest neighbors of the two new data points as a function of Euclidean distance. Suppose the number of new neighbors is 1 or 3. If k is 3 in the two common places named x , then the solid gray circle represents the nearest neighbor and the dotted circle represents the neighbor.

The right side of Figure 2.13 shows the decision limits for this dead data when k is equal. Because the transition was so simple, the nearest K -score was used as the study citation method. The only optimization parameter of this method is the closest value to the number of k neighbors. Larger k values increase the resolution limit and decrease detail. Although this method is simple, Naver's method has recently been shown to be effective in many financial applications, such as forecasting currency rates or stock movements. Because features in a dataset can have different values, each feature is usually measured with a mean of 0 and a variance of 1.

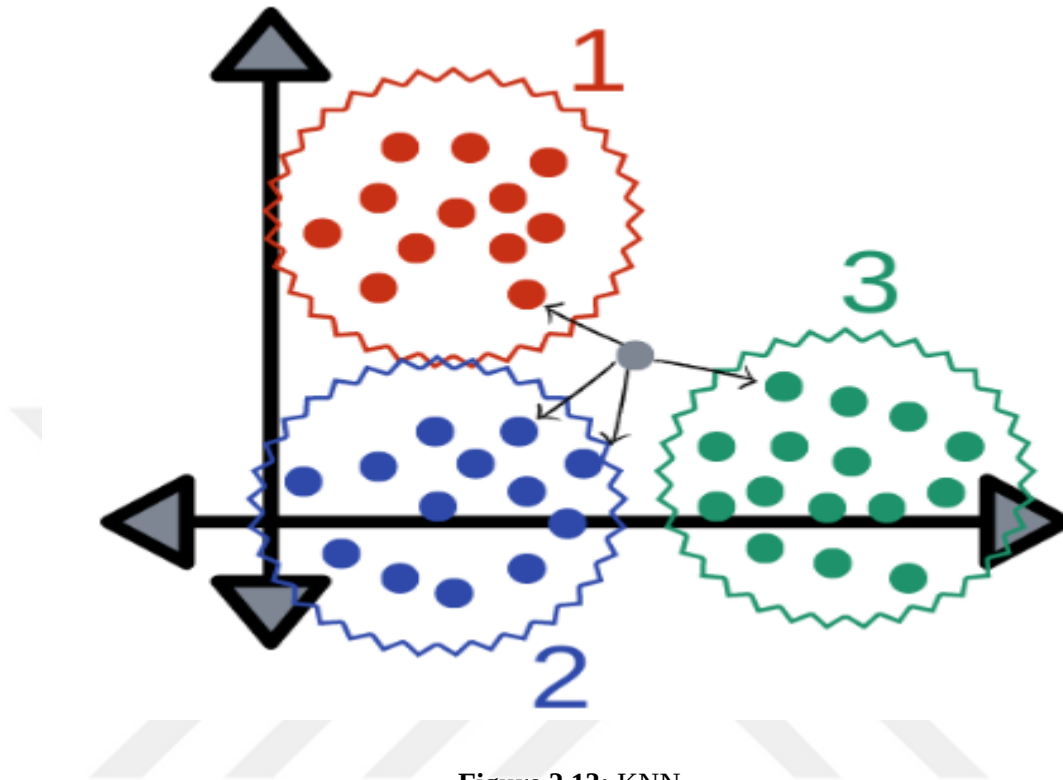


Figure 2.13: KNN.

2.1.14 Random Forest Method

It is not necessary to reduce the size of the function because the amount of data is large, the data is dimensional, it contains many similar examples, and the default settings can better handle the random use of the collection. Also, the data sample is unstable, which allows the random forests to better adjust the error level. Random Forest is the most widely used algorithm in machine learning. That is, it creates multiple decision trees from a set of training data and performs classification and regression using a tool called population. Each decision tree represents a prediction of the class, the method collects votes from these decision trees, and the class that gets the most votes is the last class. During training, each reference model is generated independently by training several random subsamples of the data. Samples are obtained through a process called packaging or storage.

This means that a given state can be used multiple times in the same decision tree. Training each decision tree with a different subsample results in low forest variance but a high degree of bias,

even if each tree has high variance for a given training data set. The test measures the average of the predictions for each decision tree to get the overall prediction. This process is called boxing. The main reason why this method performs better than decision trees is that many independent decision trees can protect each other from individual errors to generate cumulative predictions, thus suboptimal results in terms of computational accuracy. , can reduce processing and prediction problems.

This procedure has other advantages. You can work efficiently with large data sets and even predict important variables in classification. Random forests are superior to linear models in that they allow nonlinear relationships between objects and functions. The downside to this method is that the decision tree is a building block, so we can't use the scarcity functions. Also, the data sample is unstable, which allows the random forests to better adjust the error level. Random Forest is the most widely used algorithm in machine learning.

That is, it creates multiple decision trees from a set of training data and performs classification and regression using a tool called population. Each decision tree represents a prediction of the class, the method collects votes from these decision trees, and the class that gets the most votes is the last class. During training, each reference model is generated independently by training several random subsamples of the data. Samples are obtained through a process called packaging or storage. This means that a given state can be used multiple times in the same decision tree. Training each decision tree with a different subsample results in low forest variance but a high degree of bias, even if each tree has high variance for a given training data set. Therefore, we need a preprocessing process that follows the model and feeds the system. To implement a random set, the number of trees to be created must be specified, the minimum number of leaf observations allowed in each node, and the number of variables that must be included in each subsection. In general, the use of this method requires relatively little modification of these parameters to obtain acceptable estimates. The most common parameter is the number of candidate variants per polymorphism [25].

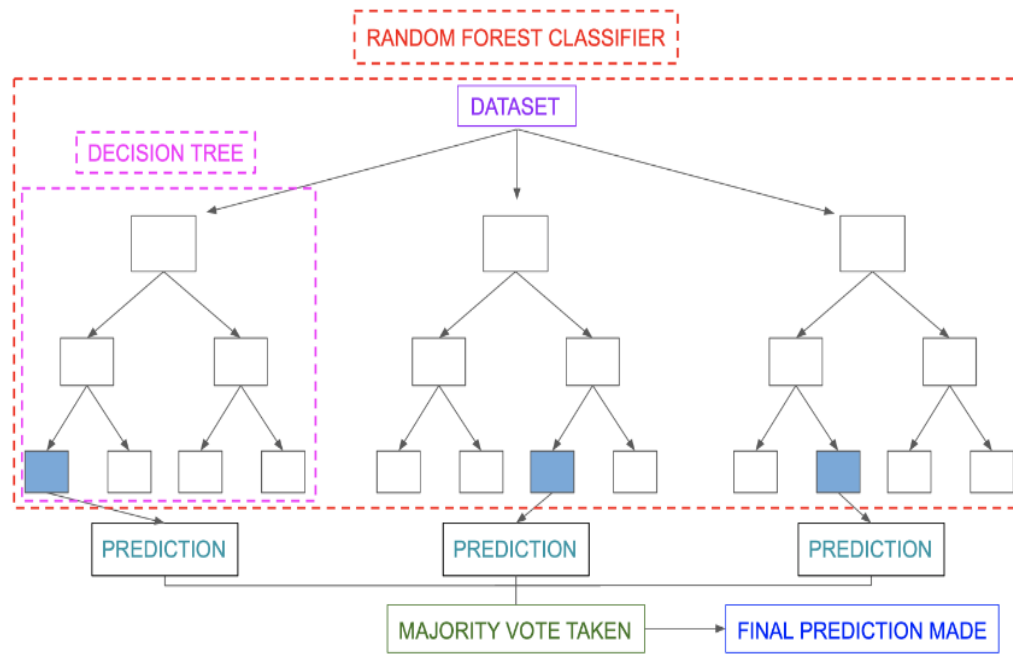


Figure 2.14: RF.

3. MATERIAL AND METHODS

3.1 DEEP LEARNING

Deep learning, also known as hierarchical learning or deep structured learning, is part of a large family of machine learning techniques that rely on learning from representations of data rather than models. Traditional methods of activity. You can learn how to submit data with the help of a classroom teacher approach or an unsupervised teacher.

Deep learning models mimic the information processing and communication patterns observed in the biological nervous system. For example, B. Neurocoding attempts to explain and explain the relationship between different stimuli and their associated neural responses in the brain. Deep learning is a field of research that deals with artificial neural networks and similar machine learning algorithms with one or more hidden layers.

That is, at least one ANN is used, and in many algorithms the computer searches for new data from existing data. Deep learning can be done with a teacher, with a class teacher or without a teacher. Artificial deep neural networks have also shown good results using an advanced training method. Deep generative models describe the spatial distribution of all possible data (eg, face images), so that data points can often be derived from random variables representing hidden regions. Deep generative models use deep neural networks to determine the distribution of data and produce impressive imaging results.

Two types of deep generator models are used to represent missing data. Variable Autoencoder (VAE) and Generative Enemy Network (GAN), RNN, LSTM and CNN [26]. Deep learning, a machine learning technique, uses multi-level non-linear objects to determine the properties of objects fed to the computer and transforms them into the desired properties. Subsequent levels take the output of the previous level as input. Algorithms defined in deep learning can be modified with or without tracking. It is an architecture based on researching the functional levels of AI on more specific data. A step-by-step tutorial is provided to identify critical and non-critical data based on data attributes. Views written according to this criterion have different levels of representation corresponding to different levels of abstraction. Therefore, this method

is mainly based on a representative study of the data. Machine learning differs in that it can multitask at different levels, and the entire system can be built independently rather than manually. In other words, deep learning is a special type of machine learning that provides more power and flexibility in learning to represent the world as a nested hierarchy of concepts and uses a more abstract representation that describes each concept rather than a simple concept. . A short summary. Deep learning is a special field of machine learning. It is defined as a new way of looking at learning representations from data, with an emphasis on learning successive levels of representations from more meaningful representations. The depth of deep learning does not mean that this approach can lead to deeper understanding. Rather, it is a set of settings at the level of attention. The number of layers that make up the data model is called the depth of the model. Other proper names of this direction are hierarchical prospective studies and hierarchical prospective studies [27].

Modern deep learning often involves dozens or hundreds of successive layer representations that are automatically learned when exposed to training data. On the other hand, other machine learning approaches focus on representing training data in one or two layers. For this reason, it is also called surface learning. In deep learning, these multi-layered representations are (in most cases) learned using models called neural networks. Artificial neural networks consist of physical layers superimposed on each other. Although the term "neural networks" refers to neuroscience, some of its basic concepts. Given the benefits of deep learning, the more training data available, the less economical this method is. Deep learning models have evolved over time as the computing infrastructure (both hardware and software) for deep learning advances.

Deep learning has been shown to solve complex applications with increasing accuracy over time. Deep learning, which has recently gained popularity, is widely used in the fields of video recognition, speech recognition and natural language processing, but there does not seem to be a lot of study on its application in finance, security, and business administration. A typical application is the use of cyclic neural networks (RNNs) to predict stock price fluctuations, as they are suitable for time series analysis. Another important deep learning model is the CNN (Convolutional Neural Network) which predicts literature errors.

We found that this model outperformed common traditional models using more traditional machine learning methods. Additionally, limited research shows that deep learning outperforms text-based bankruptcy predictions when it comes to correlating specific textual data (news, stock reports, etc.) with conventional numerical data (indexes data providers, etc.). performance and better accuracy. it's clear. Therefore, machine learning methods provide new insights and inspire research in the field of error prediction.

3.1.1 Autoencoders

A variety of machine cipher is a type of neural network used to learn how to efficiently encode data without supervision. The purpose of the autoencoder is to teach the network to ignore the "noise" in the signal. In general, you are learning a low dimensional representation of the data set. As an adaptive approach to the missing data problem, VAE adds random predicates to the model. It uses a unique probabilistic neural model based on a variable autoencoder, in which the observed data is placed in a random subset. Recorded values can be numeric or categorical. The model was trained using Bayesian random variance. Experimental prediction, mapping and presentation tasks on synthetic data show the effectiveness of the proposed approach and the diversity of the samples produced [28, 29].

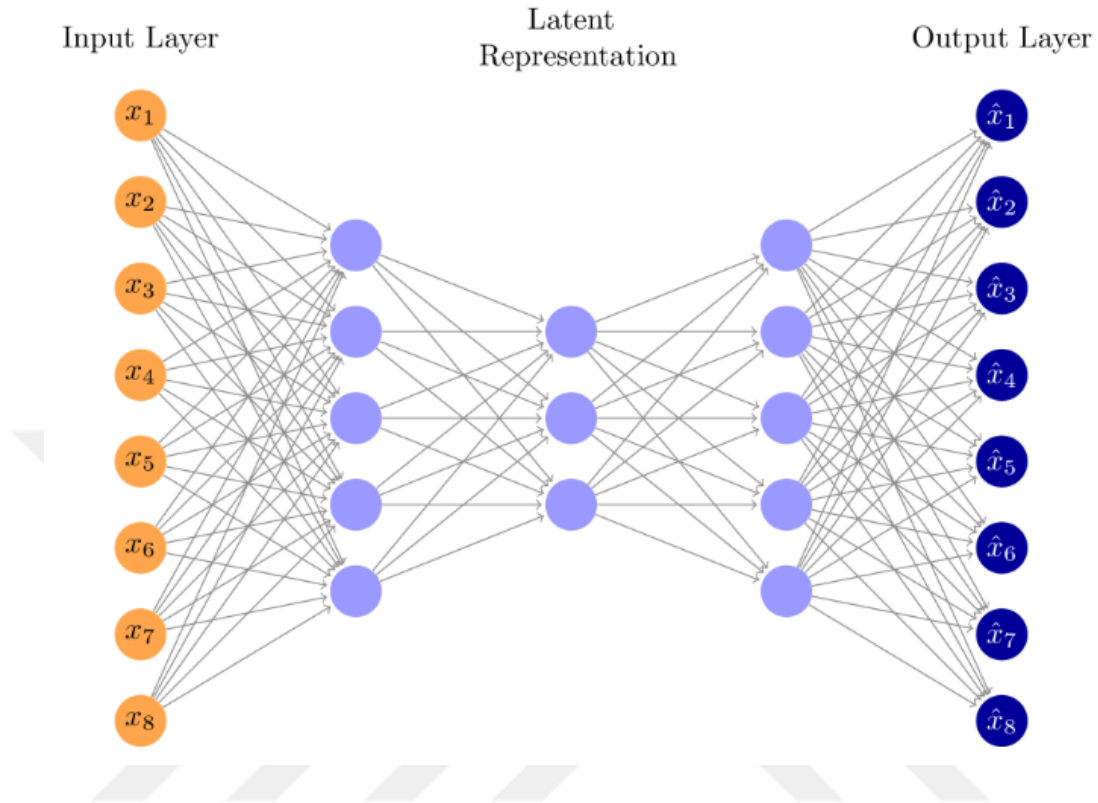


Figure 3.1: Autoencoders.

3.1.2 Generative Adversarial Networks (GAN)

Generative Contention Networks is machine learning algorithm is responsible for recognizing patterns in the available data, and this understanding is based on the problems of classification (assigning samples to the correct categories) and regression (for example, predicting numerical values based on multiple inputs). [30].

However, if these algorithms are required to generate new data, a computer cannot do so. Machine learning algorithms can beat chess masters, predict stock price movements, and determine whether a credit card transaction is fraudulent. On the other hand, the chat options are very limited as the app is used for secondary tasks like Amazon Alexa or Apple Siri. Whether it's having a fun conversation or creating a unique product, the simplest and most basic human technology can digitally transform even the most sophisticated supercomputers. It revolves around the invention of Generative Adversarial Network (GAN). [31].

GAN is a technology that allows computers to generate real data using two separate neural networks instead of one. GANs were not the first computer programs used to generate data, but the power and versatility of GANs made the technology stand out. GAN is a class of machine learning techniques that consists of two models trained simultaneously. One part (the generator) can generate fake data, and the other part (the discriminator) can distinguish fake data from real data samples. GAN transforms ordinary graffiti into 57-inch photos or videos of zebras or horses in action to create realistic artwork. It was a great success. This is crazy. For systems without synthetic features, this can be done without training data. The word constructor refers to the overall purpose of the module. Create new data. The data a GAN learns to generate depends on the selection of data loaded into the model.

This data set can consist of a music stream, photos or video data. The term advertising refers to the competitive dynamics between the two models that make up the GAN framework. products and distributors. The goal of the generator is to create a sample that does not differ from the actual data in the training sample. The objective of the separator is to separate the fake model generated by the manufacturer from the real model in the training data set. These two networks always try to avoid each other. The more data your product generates, the better it should be able to distinguish between real and fake credentials. Finally, the term "network" refers to neural networks. This is a class of machine learning models commonly used to refer to generators and constraints.

Depending on the complexity of the GAN application, it can range from simple neural networks to convolutional neural networks or more complex options. Therefore, the purpose of the constructor is to create samples that reflect the properties of the data set used to train the model, so that the generated samples are indistinguishable from the training data. Unlike constructors, defined object recognition is defined as a pattern. Object recognition algorithms can distinguish content by recognizing patterns in images [32].

Instead of recognizing a pattern, the generator basically learns to create a pattern from scratch, and the input starts as a vector of random numbers. Manufacturers learn from your feedback. The identifier determines whether a given sample is genuine (from the training database) or fake

(by the manufacturer). Thus, if the delimiter classifies invalid data as true, the producer knows which data the delimiter has classified as true. On the other hand, if the reviewer deliberately rejects the image the product creates, the manufacturer will receive feedback that can be improved. At the same time, the space expands [33].

Like other classifiers, it learns how far its predictions are from the correct label (true or false). The gap between real and fake data is widening, and both systems continue to evolve as manufacturers get better at generating real data.

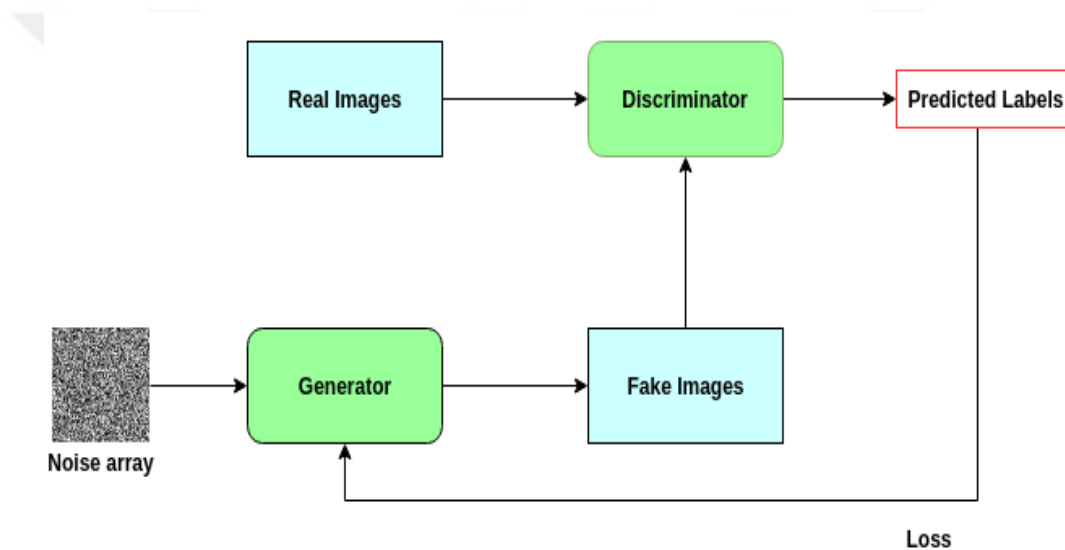


Figure 3.2: GAN.

3.1.3 Recurrent Neural Networks (RNN)

Recurrent neural networks (RNN) is a kind of product artificial neural network that can use information transmitted continuously in its network structure. A circular neural network is a neural network model designed specifically for sequential information processing [35].

Educational neural networks specialize in information processing in areas such as data processing while circular neural networks specialize in sequential processing of sets of values. Iterative neural networks are useful for learning sequential data, using the recorded data to evaluate the following inputs. In this sense, circular neural networks differ from feeder neural systems from the point of view of information processing. In a forward neural network,

information from the input layer passes through the hidden layers and reaches the output layer without crossing the neuron again [36].

On the other hand, circular neural networks consider current as well as past inputs when processing information over a decision-making cycle. This structure is shown in Figure 3.3.

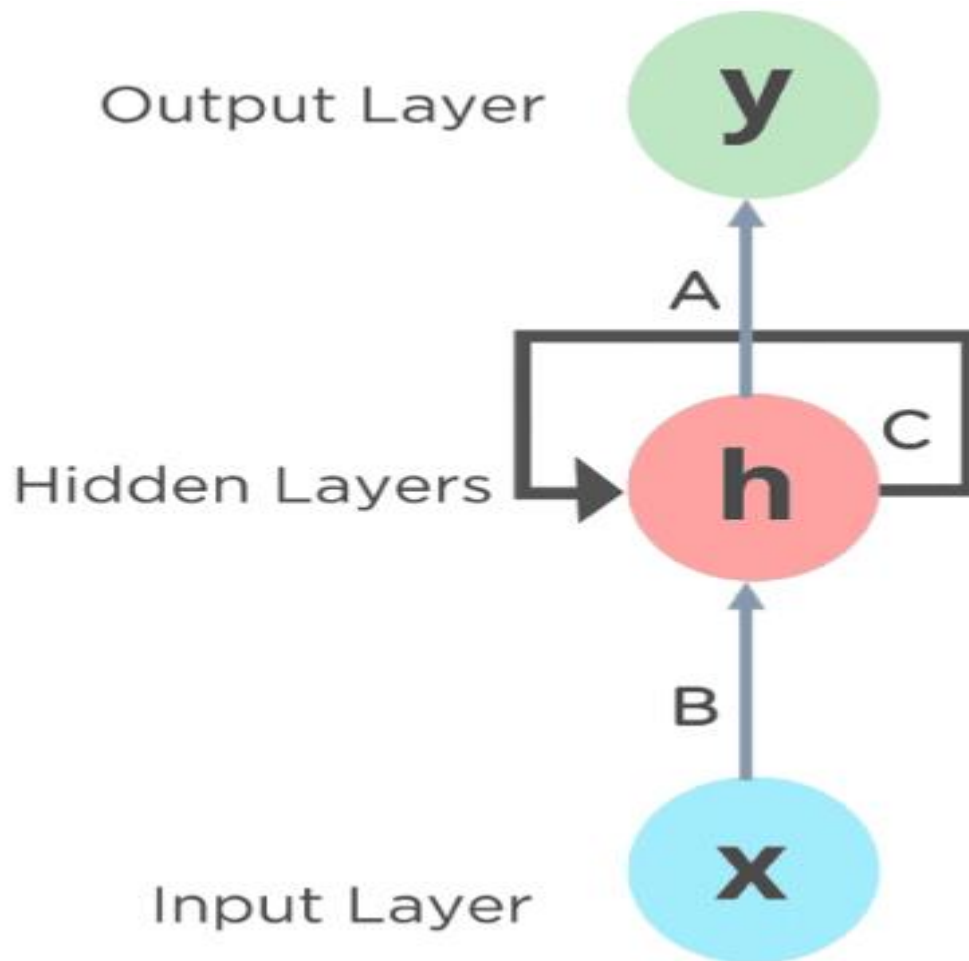


Figure 3.3: RNN.

While feedforward neural networks are trained using the back propagation algorithm, the loop shown in Figure 3.3 in recurrent neural networks creates an iterative process, which is called back propagation in time. Back propagation in time is the training algorithm used by iterative neural networks. Backpropagation in time is based on processing the data series in order by connecting the data in each time step to the data in the previous time steps. Recurrent neural

networks take the output of each hidden layer and output layer and feed itself back through neurons called delay nodes. Delay units are units that allow information to pass from one step to the next. Time series data contains patterns ordered by time, and the information necessary for data generation is contained in these patterns [37].

Through the lag units, these patterns in the time series exhibit persistence, creating a short-term memory that allows the recurrent neural network to learn and generalize the inputs. It does this by storing the activation values and output values in the hidden layer of the previous time step and transferring these values back to the network in the next time step.

3.1.4 Long Short Term Memory (LSTM)

In RNN, with the increase of time delays, very deep feed-forward neural networks emerge and this situation creates the problem of vanishing gradient [38].

Long-short-term memory model (LSTM) and gated repetitive unit model (GRU), developed to solve the problem of gradient extinction in recurrent neural networks, are neural network models designed to determine optimal time delays by giving cells the ability to determine when they will forget certain information. In recurrent neural networks, the gradient of the weights becomes too small or too large as the time step gets longer. If the gradient is too big, it will explode; if the gradient is too small, it will spread out and disappear, reflecting changes in the parameters being studied. Another form of circular neural networks, the LSTM and GRU models, have been developed to address the leakage gradient problem faced by circular neural networks.

On the other hand, LSTMs (long and short memory models) are a special type of repetitive neural networks with memory cells that are used to store information for a long time. In this sense, LSTM is a network architecture designed to more accurately model long-term dependencies on time series. Standard iterative neural networks are limited in that they model long-term dependencies, whereas LSTM architectures can store previous iterations needed for cells holding and retain old values in the presence of forgotten ports to efficiently capture

dependencies. long-term. Don't let the cells know. forget your value. The schematic view of the LSTM model is shown in Figure 3.4.

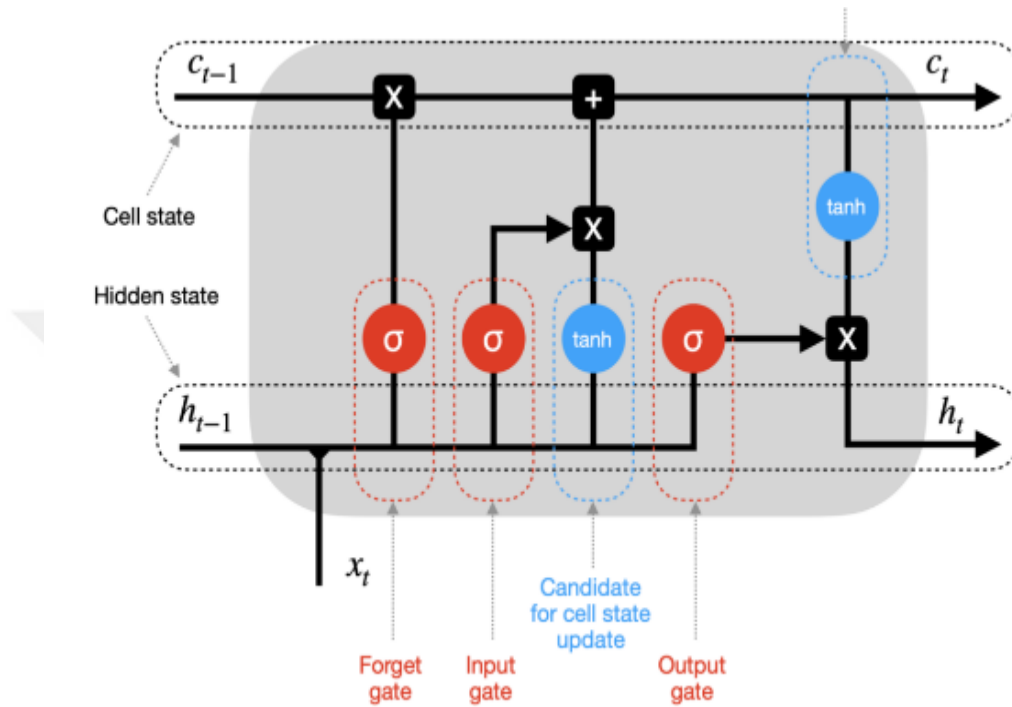


Figure 3.4: LSTM.

The LSTM layer consists of a series of blocks, called memory blocks, that are repeatedly linked together. Each of the memory blocks in Figure 3.4 includes a few or other repetitively attached recollection units and three units (input port, exit port, and forget port) that deliver continuous script, read out, and reset procedures. Network interacts with cells through gates.

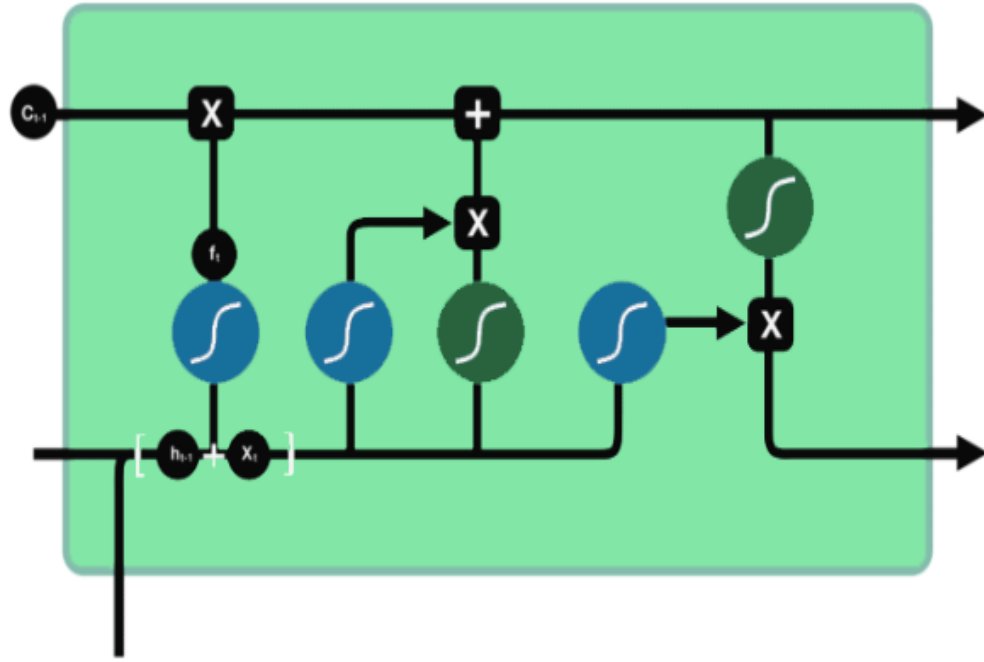


Figure 3.5: LSTM Step 2.

3.1.5 Gated Recurrent Unit-GRU

A closed recurrent unit (CRU) is a type of recurrent neural network designed to solve the gradient delay and gradient burst problems encountered in conventional recurrent neural networks by identifying the long-term dependence of data. Like the LSTM model, the GRU model is designed to collect data from previous periods. LSTM model uses three ports: input port, output port and output port, while GRU model uses two ports as reset port and update port. A retrieval gate is used to determine the amount of previous information the model should forget or retain, while an update gate determines the amount of information to be carried over to future steps. The network architecture of the GRU model consists of repeating blocks of blocks with ports that manage memory updates and emissions. GRUs with fewer parameters than LSTMs allow for faster learning [39].

In a GRU block, data flow occurs in the block's hidden layer, not in individual memory cells. A recovery gate that can send and receive data from a previous time step rejects data if the data from the previous time step is not current. Update ports help identify long-term dependencies

by tracking how much old storage is lost and how much new storage is added. When the content value is set for further use, the update gate is closed and the memory content is moved to another step. In this case, the content of the last memory consisted mainly of the content of the new memory and the content of the old memory [40].

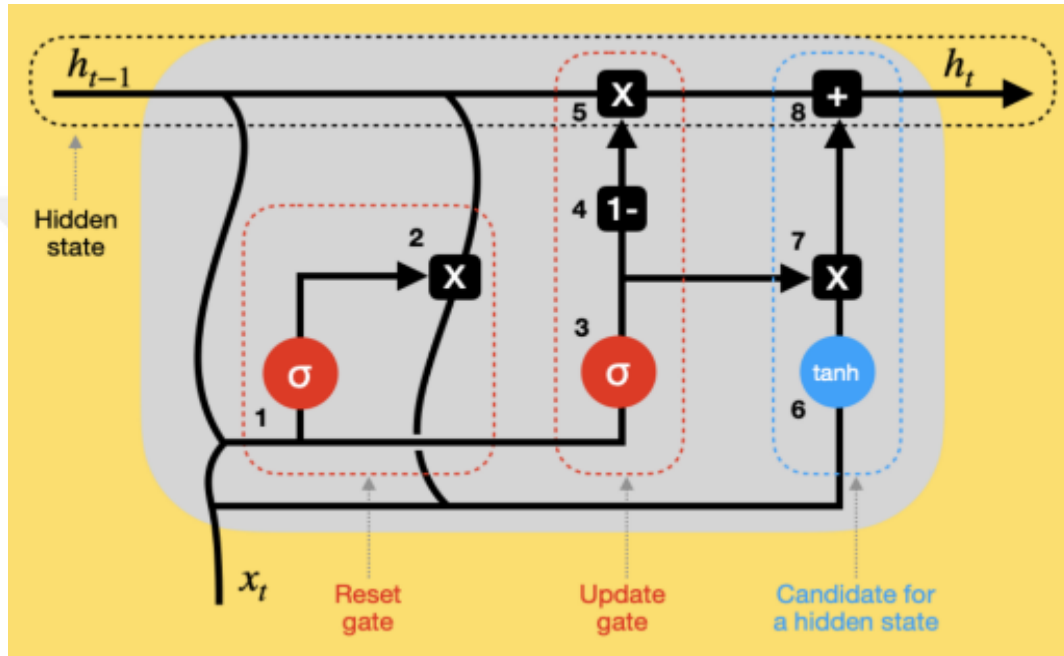


Figure 3.6: Gated Recurrent Unit.

3.1.6 Convolutional Neural Networks (CNN)

CNNs (Convolutional Neural Networks) are neural networks that use a special linear mathematical process, convolution, instead of the usual multiplication of matrices in at least one layer to process 1D time series data and 2D image data containing pixels. Learning neural networks are similar to general neural networks in that they are made up of neurons that improve during the learning process [41].

Similar to neural networks, convolutional neural networks work with the input of each neuron. Here we can compare the expected result with the actual result of the last level to obtain the loss function. In composite neural networks, the same technique is used in artificial neural networks. However, unlike neural networks, convolutional neural networks use the pattern recognition

field to recognize inputs in the first stage. Neural network training is a deep learning model designed to learn how to transform low-level features in network data, such as images, into high-level models consisting of three levels: composition, sum, and composition. Fully connected. [42].

In neural networks, the convolution and cluster layers extract functionality and the fusion layer presents the extracted functionality as final output. A trained neural network uses many building blocks because it is a network architecture designed to obtain attributes from input data via backpropagation. learn. Feature extraction consists of abstract representations from the network to extract low-level functionality from the input. These extracted features are implemented using a classifier developed from scratch [43].

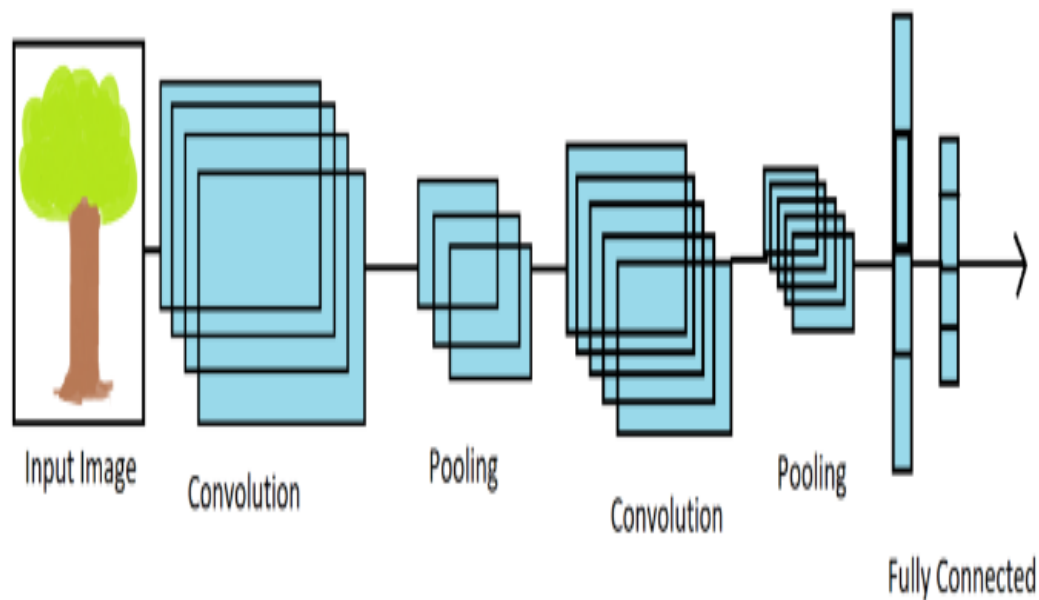


Figure 3.7: CNN.

With the ability of neural networks to extract input features and perform classifications based on these features, image classification, object recognition, face recognition, voice recognition, etc. can be used. It is based on numerous studies in the specialized literature. This represents a wide range of applications. The first effective layer in the CNN is shown in Figure 3.8.

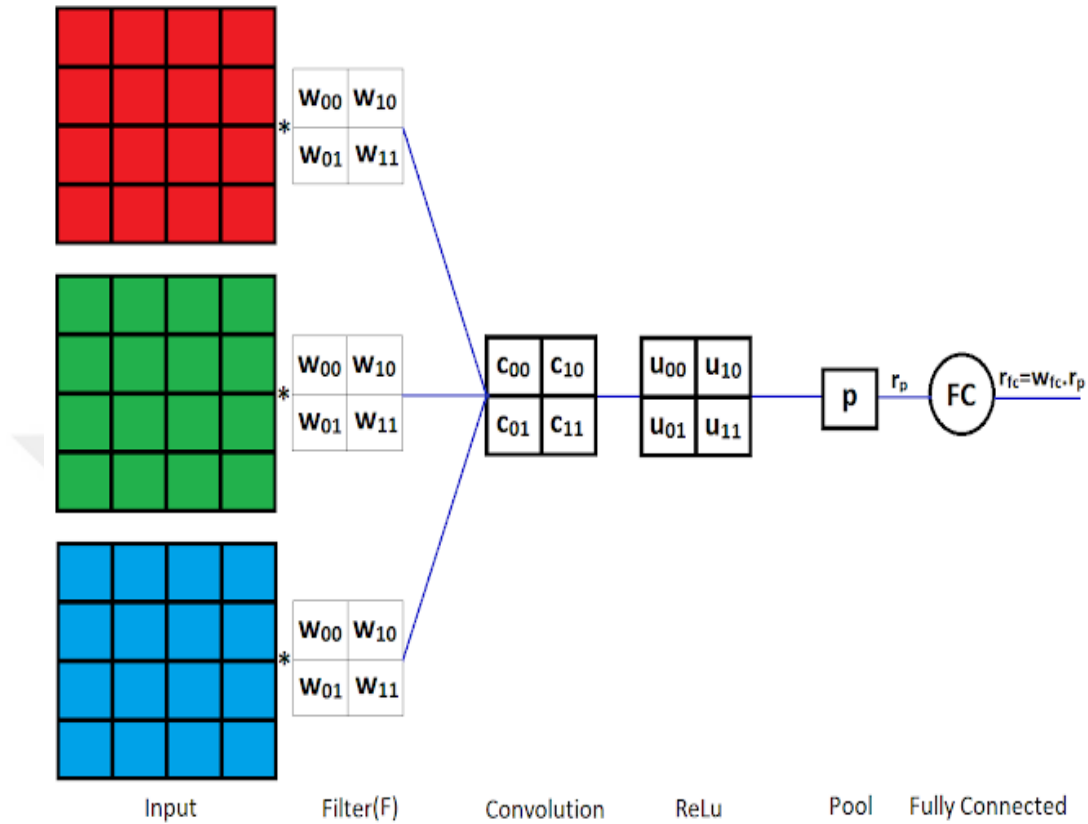


Figure 3.8: Convolution Layer.

In particular, it is easier to design than traditional neural networks of similar size due to fewer connections and fewer parameters. The general model of a composite neural network consists of compound stacks, group couches, and fully connected couches [44].

In neural network models, the main part of the network, the convolutional layers, extracts the attributes from the input data. The subtraction layer consists of multiple feature maps, and each neuron belonging to the same feature map is used to extract a different feature from the previous layer. Convolution is a special type of linear function used to extract features, where the input takes a series of numbers called a tensor and a sequence of numbers called a kernel. Since each kernel element is calculated for each position of the tensor, we repeat this process to generate several feature maps that represent the different properties of the input sensor. This process called, compression, reduces the size of the intake sensor. The main feature of the push is the distribution of the weight with the core.

By replacing it with a functional map, we can get a more complete picture of the function. The extraction results are reported by activation functions and nonlinear activation functions. Convolutional neural network architectures use pooling layers together with input layers to gradually reduce the size of the input data, making greater use of the number of parameters and the computational complexity of the convolutional input assembly model. Full sofa as shown. 3.9.

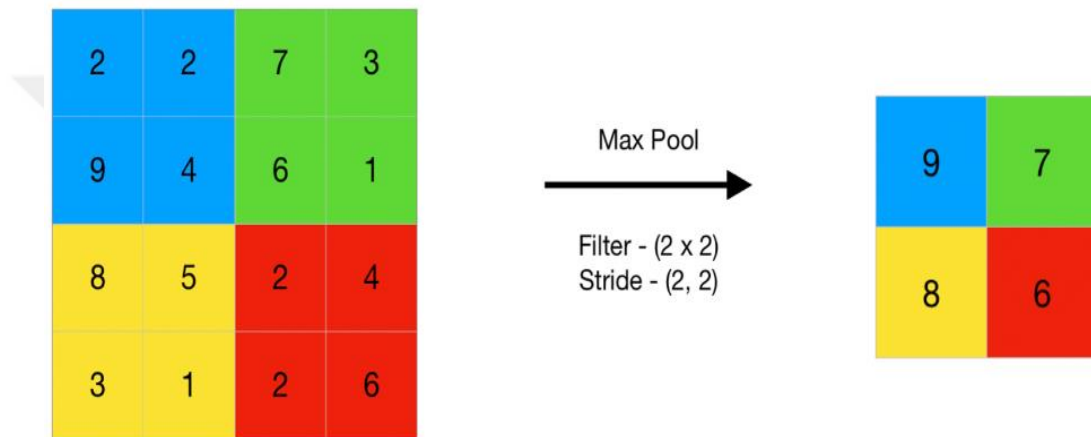


Figure 3.9: Pooling Layer.

Output loss for a fully bonded layer. The pool level controls the size of each activation card in the input data. A classification system in the architecture of the nervous system consists of one or more fully connected layers. A layer in a classification network receives information from the previous layer, connects to each neuron of the current layer, and transmits information from there to the last connected layer. The last fully connected layer transmits this information to the source layer. Features extracted from convolution layers and selected into pool layers are transformed into a 1D numerical matrix (plane) associated with the fully connected layers and mapped to the final mesh result of the fully connected layers. After the final output loss function is calculated, the error propagation process begins, and the optimization algorithm updates the kernel and weights according to this loss function [45].

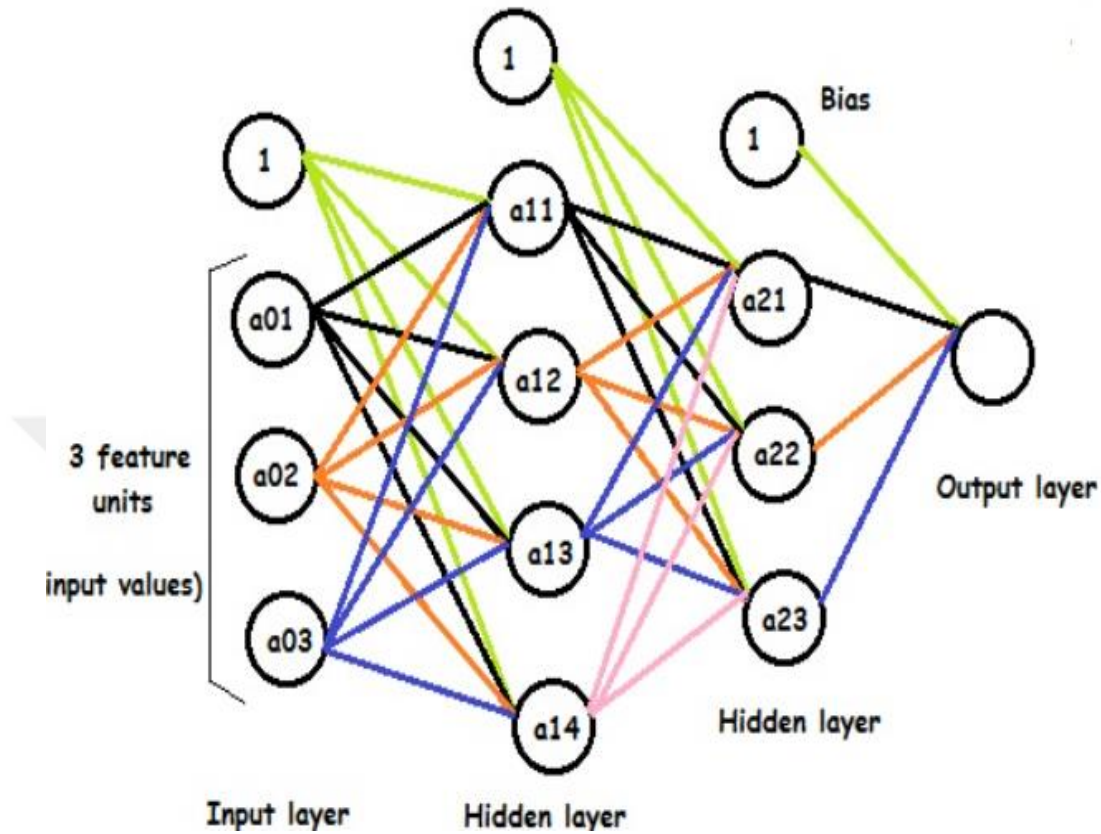


Figure 3.10: FC.

3.2 PROPOSED CONVOLUTIONAL NEURAL NETWORK BASED METHOD

In this study new method based convolutional neural network presented for covid-19 detection. The proposed method combined convolutional neural network with optimization algorithm called Ant lion optimization algorithm and particle swarm optimization algorithm. The aim of applying ACO is for training CNN to obtain best classification rate this is the first stage of the proposed method. In the second stage, the features extracted using convolutional neural network which convolutional neural network have ability to extract high level features from input images. This mean the features that are extracted by the convolutional neural network represented the x-ray image very well. In the 3rd stage, the genetic algorithm applied for feature selection to reduce the size of input features. The genetic algorithm is very power and robust optimization algorithm applied in various optimization problems and presented remarkable results.

On the other hand, several classifier applied to classify the extracted features by the CNN and optimization algorithms genetic algorithm and ant Conny optimization algorithm.

The flowchart of the proposed method presented in the Figure.

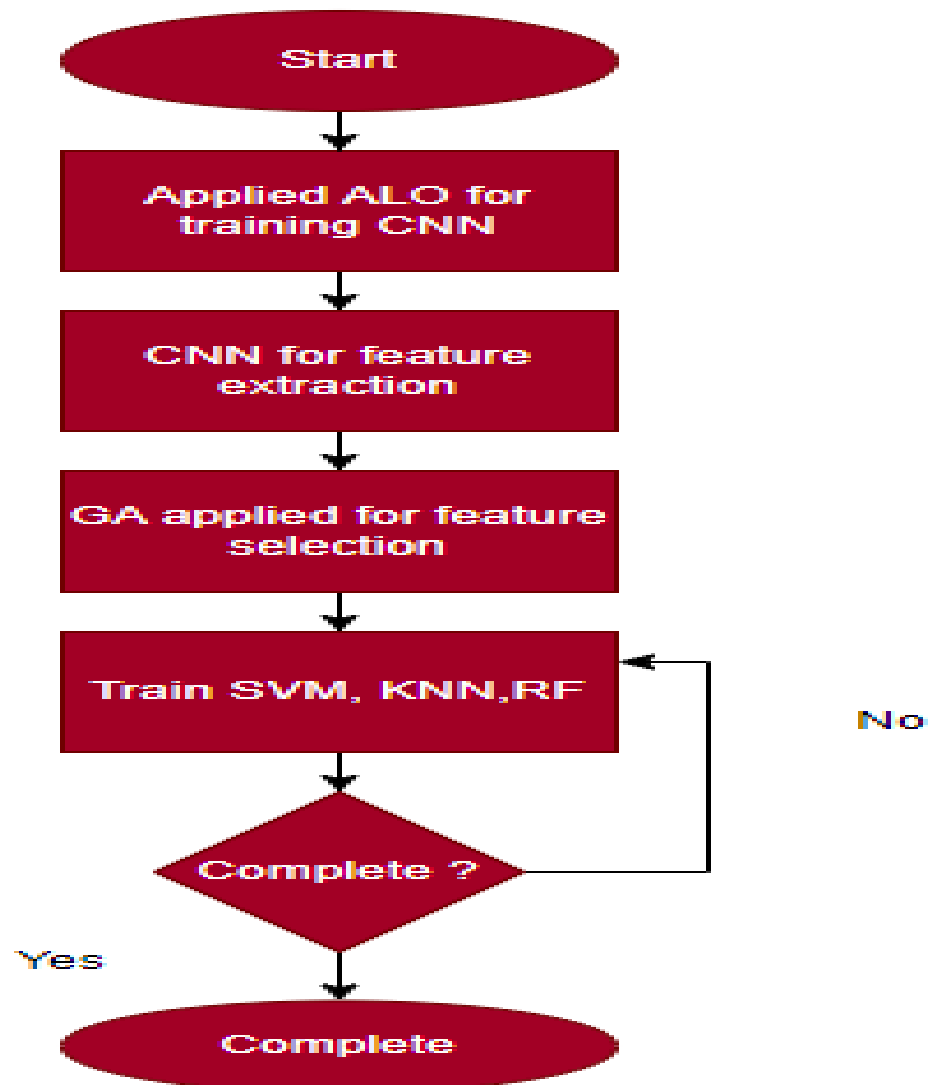


Figure 3.11: Proposed Method.

4. EXPERIMENTAL RESULTS AND DISCUSSIONS

In this section, the proposed method applied for three datasets to evaluate the results and after that compare the execution time. Three parameters sensitivity, precision and accuracy are calculated to evaluate the performance of the model. The proposed method show that the CNN based ALO and GA-RF show that its performance best of other techniques that are executed in this study.

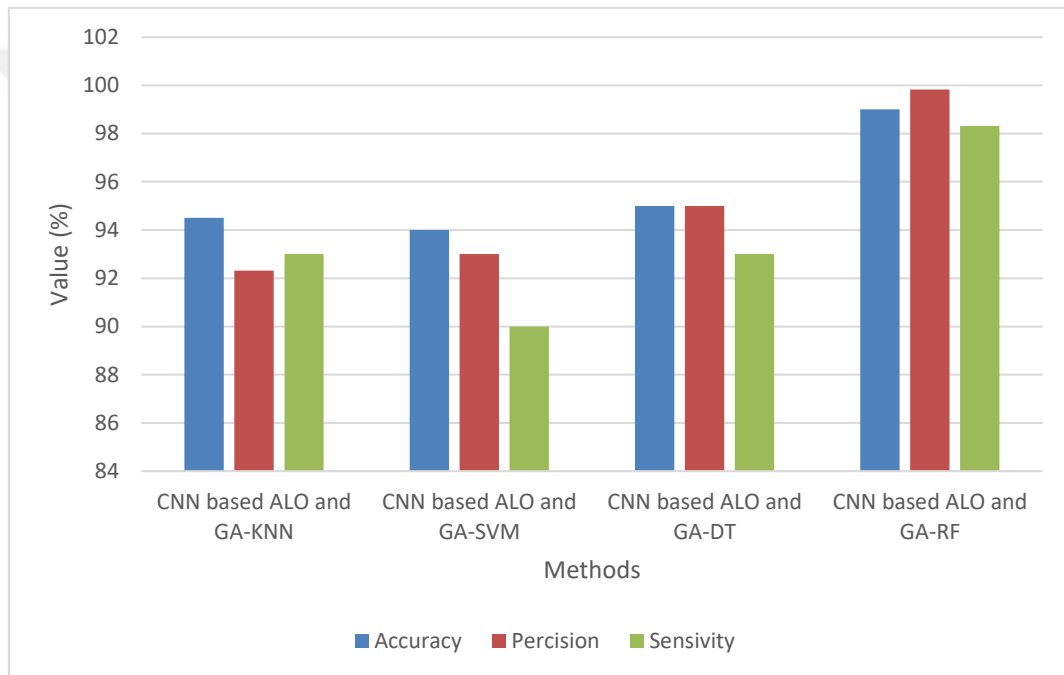


Figure 4.1: Evaluation Parameters of Binary Dataset.

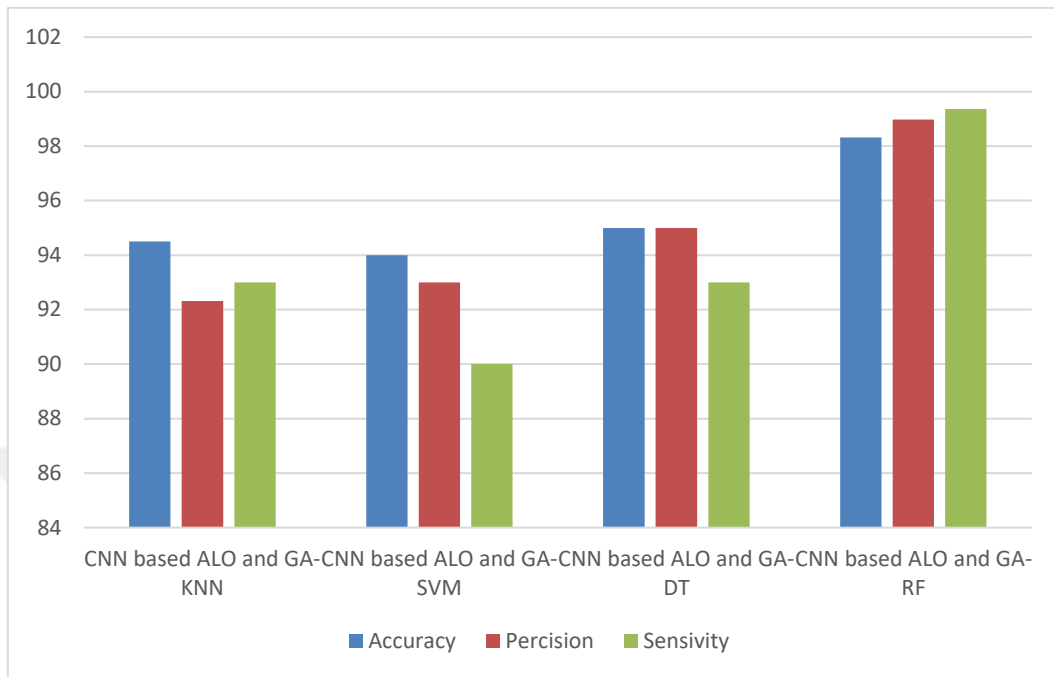


Figure 4.2: Evaluation Parameters of Three Classes Dataset.

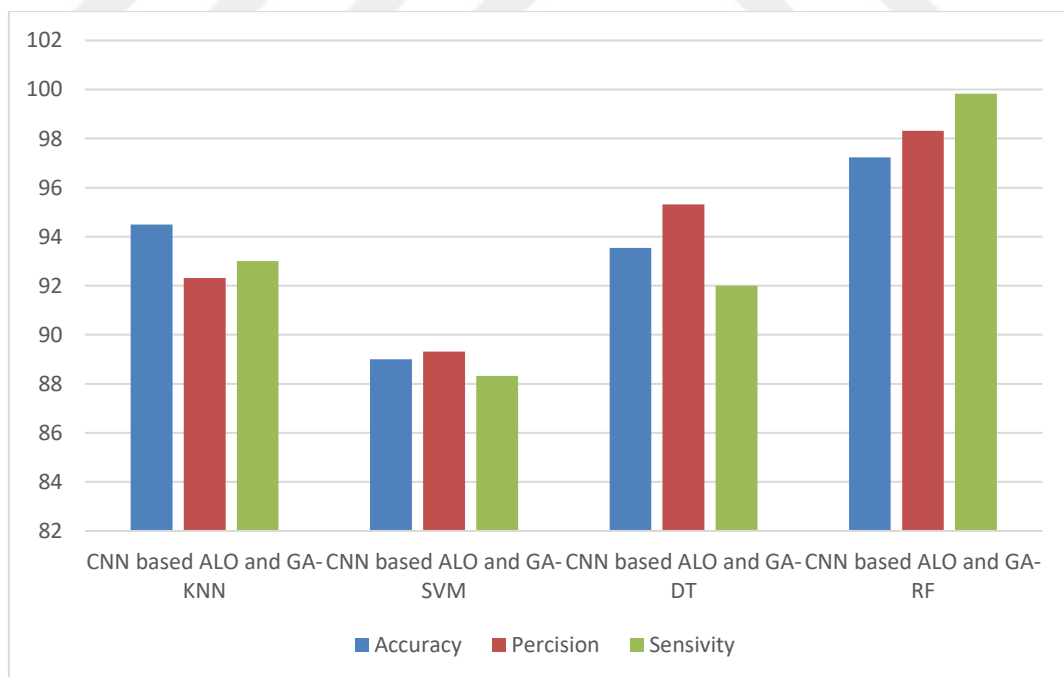


Figure 4.3: Evaluation Parameters of Four Classes Dataset.

The other evaluation metric is execution time which is play great role in selecting machine learning methods. This mean we must calculate the execution time performance of the methods with several datasets as shown in the Figure 4.4.

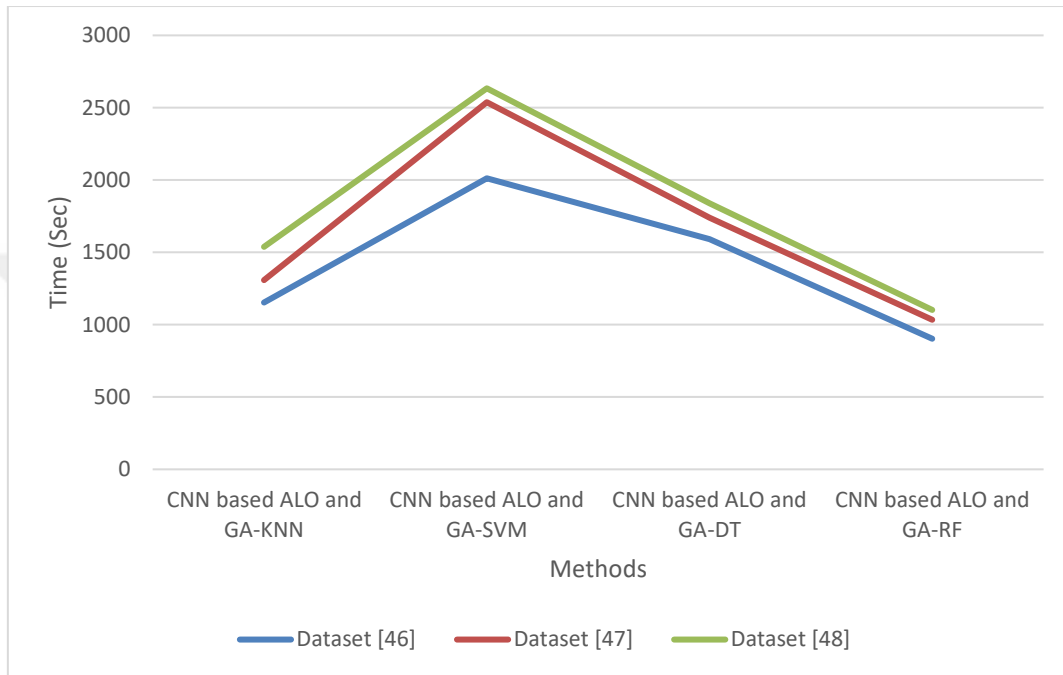


Figure 4.4: Execution Time with Three Datasets.

5. CONCLUSION

COVID-19 spread rapidly all over the world and caused serious problems for the people it infected. Detection of COVID-19 in the early stages, and timely and accurate intervention will both reduce the spread of the epidemic and minimize the health problems that may occur. In recent years, machine learning has been used in many fields. Machine learning also has the potential to be used in the early diagnosis of diseases.

In this study new method based convolutional neural network presented for covid-19 detection. The proposed method applied to detect covid-19 in three different datasets. The proposed method show remarkable results when compared with other traditional techniques. The proposed method presented more than 99% accuracy, precision and sensitivity. This mean that the combination of the model are very effective to predicate optimum results. The presented model combined CNN which is new technique used in feature extraction problems and extracted high level features from input x-ray images. The CNN trained using ALO which is new optimization algorithm and applied in various optimization problems. The using ALO lead to obtain best weight and basis to present best accuracy and detection rate.

On the other hand, the genetic algorithm applied for select best features from features that are extracted by CNN. The genetic algorithm role is to minimize the size of CNN output features to decrease the execution time and increase the performance of the classifiers that applied in this study.

Then, several classifiers applied to classify the features that are selected by the combination of the CNN based genetic algorithm and ALO. The RF presented best classification rate then other classifiers such as SVM, KNN, and DT.

In the future works, the researchers can applied other deep learning techniques to detect abnormality in the z-ray images. Furthermore, other optimization algorithms can applied to for training the CNN and other deep learning models. Moreover, the optimization algorithms also can applied for features selection stages.

REFERENCES

- [1] Lee S, Myung H. Receding horizon particle swarm optimisation-based formation control with collision avoidance for non-holonomic mobile robots. 2015;9,2075–2083.
- [2] Su K, Wang Y. Robot path planning based on random coding particle swarm optimization. 2015;6(4):58–64.
- [3] Abukhalil T, Patil M, Patel S, Sobh T. Coordinating a heterogeneous robot swarm using robot utility-based task assignment (RUTA). In 2016 IEEE 14th International Workshop on Advanced Motion Control (AMC). IEEE, 2016;(57-62).
- [4] Ferrús RM, Somonte MD. Design in robotics based in the voice of the customer of household robots. *Rob. Auton. Syst.*, 2016;79,99–107.
- [5] Diprose J, MacDonald B, Hosking J, Plimmer B. Designing an API at an appropriate abstraction level for programming social robot applications. *J. Vis. Lang. Comput.*, 2017;39,22–40.
- [6] Das B, Couceiro MS, Vargas PA. MRoCS : A new multi-robot communication system based on passive action recognition, *Rob. Auton. Syst.*, 2016;82,46–60.
- [7] R.A. Saeed, Diego Reforgiato Recupero, Paolo Remagnino, A Boundary Node Method for path planning of mobile robots, *Robotics and Autonomous Systems*, Volume 123, 2020, 103320, ISSN 0921-8890, <https://doi.org/10.1016/j.robot.2019.103320>.
- [8] Meng Zhao, Hui Lu, Siyi Yang, Yinan Guo, Fengjuan Guo, A fast robot path planning algorithm based on bidirectional associative learning, *Computers & Industrial Engineering*, Volume 155, 2021, 107173, ISSN 0360-8352, <https://doi.org/10.1016/j.cie.2021.107173>.
- [9] Abhishek Kumar Kashyap, Dayal R. Parhi, Manoj Kumar Muni, Krishna Kant Pandey, A hybrid technique for path planning of humanoid robot NAO in static and dynamic terrains, *Applied Soft Computing*, Volume 96, 2020, 106581, ISSN 1568-4946, <https://doi.org/10.1016/j.asoc.2020.106581>.

- [10] Feiyi Xu, Haolun Li, Chi-Man Pun, Haidong Hu, Yujie Li, Yurong Song, Hao Gao, A new global best guided artificial bee colony algorithm with application in robot path planning, *Applied Soft Computing*, Volume 88, 2020, 106037, ISSN 1568-4946, <https://doi.org/10.1016/j.asoc.2019.106037>.
- [11] Baoye Song, Zidong Wang, Lei Zou, An improved PSO algorithm for smooth path planning of mobile robots using continuous high-degree Bezier curve, *Applied Soft Computing*, Volume 100, 2021, 106960, ISSN 1568-4946, <https://doi.org/10.1016/j.asoc.2020.106960>.
- [12] Mohammad Samadi Gharajeh, Hossein B. Jond, An intelligent approach for autonomous mobile robots path planning based on adaptive neuro-fuzzy inference system, *Ain Shams Engineering Journal*, 2021, ISSN 2090-4479, <https://doi.org/10.1016/j.asej.2021.05.005>.
- [13] Hamid Taghavifar, Subhash Rakheja, Giulio Reina, A novel optimal path-planning and following algorithm for wheeled robots on deformable terrains, *Journal of Terramechanics*, 2020, ISSN 0022-4898, <https://doi.org/10.1016/j.jterra.2020.12.001>.
- [14] Anirudh Krishna Lakshmanan, Rajesh Elara Mohan, Balakrishnan Ramalingam, Anh Vu Le, Prabakar Veerajagadeshwar, Kamlesh Tiwari, Muhammad Ilyas, Complete coverage path planning using reinforcement learning for Tetromino based cleaning and maintenance robot, *Automation in Construction*, Volume 112, 2020, 103078, ISSN 0926-5805, <https://doi.org/10.1016/j.autcon.2020.103078>.
- [15] Yang Chen, Mengqing Chen, Zhihuan Chen, Lei Cheng, Yanhua Yang, Hui Li, Delivery path planning of heterogeneous robot system under road network constraints, *Computers & Electrical Engineering*, Volume 92, 2021, 107197, ISSN 0045-7906, <https://doi.org/10.1016/j.compeleceng.2021.107197>.
- [16] Ritam Sarkar, Debaditya Barman, Nirmalya Chowdhury, Domain knowledge based genetic algorithms for mobile robot path planning having single and multiple targets, *Journal of King Saud University - Computer and Information Sciences*, 2020, ISSN 1319-1578, <https://doi.org/10.1016/j.jksuci.2020.10.010>.

- [17] Navid Kayhani, Hosein Taghaddos, Ali Mousaei, Saeed Behzadipour, Ulrich Hermann, Heavy mobile crane lift path planning in congested modular industrial plants using a robotics approach, *Automation in Construction*, Volume 122, 2021, 103508, ISSN 0926-5805, <https://doi.org/10.1016/j.autcon.2020.103508>.
- [18] Divya Agarwal, Pushpendra S. Bharti, Implementing modified swarm intelligence algorithm based on Slime moulds for path planning and obstacle avoidance problem in mobile robots, *Applied Soft Computing*, Volume 107, 2021, 107372, ISSN 1568-4946, <https://doi.org/10.1016/j.asoc.2021.107372>.
- [19] Jiankun Wang, Baopu Li, Max Q.-H. Meng, Kinematic Constrained Bi-directional RRT with Efficient Branch Pruning for robot path planning, *Expert Systems with Applications*, Volume 170, 2021, 114541, ISSN 0957-4174, <https://doi.org/10.1016/j.eswa.2020.114541>.
- [20] Milad Nazarahari, Esmaeel Khanmirza, Samira Doostie, Multi-objective multi-robot path planning in continuous environment using an enhanced genetic algorithm, *Expert Systems with Applications*, Volume 115, 2019, Pages 106-120, ISSN 0957-4174, <https://doi.org/10.1016/j.eswa.2018.08.008>.
- [21] Fatin H. Ajeil, Ibraheem Kasim Ibraheem, Mouayad A. Sahib, Amjad J. Humaidi, Multi-objective path planning of an autonomous mobile robot using hybrid PSO-MFB optimization algorithm, *Applied Soft Computing*, Volume 89, 2020, 106076, ISSN 1568-4946, <https://doi.org/10.1016/j.asoc.2020.106076>.
- [22] Shasha Tian, Yuanxiang Li, Yilin Kang, Jiening Xia, Multi-robot path planning in wireless sensor networks based on jump mechanism PSO and safety gap obstacle avoidance, *Future Generation Computer Systems*, Volume 118, 2021, Pages 37-47, ISSN 0167-739X, <https://doi.org/10.1016/j.future.2020.12.012>.
- [23] Abderraouf Maoudj, Abdelfetah Hentout, Optimal path planning approach based on Q-learning algorithm for mobile robots, *Applied Soft Computing*, Volume 97, Part A, 2020, 106796, ISSN 1568-4946, <https://doi.org/10.1016/j.asoc.2020.106796>.

- [24] Lin Zhang, Yingjie Zhang, Yangfan Li, Path planning for indoor Mobile robot based on deep learning, Optik, Volume 219, 2020, 165096, ISSN 0030-4026, <https://doi.org/10.1016/j.ijleo.2020.165096>.
- [25] Ben Beklisi Kwame Ayawli, Albert Yaw Appiah, Isaac Kofi Nti, Frimpong Kyeremeh, Esinam Irene Ayawli, Path planning for mobile robots using Morphological Dilation Voronoi Diagram Roadmap algorithm, Scientific African, Volume 12, 2021, e00745, ISSN 2468-2276, <https://doi.org/10.1016/j.sciaf.2021.e00745>.
- [26] Lars Larsen, Jonghwa Kim, Path planning of cooperating industrial robots using evolutionary algorithms, Robotics and Computer-Integrated Manufacturing, Volume 67, 2021, 102053, ISSN 0736-5845, <https://doi.org/10.1016/j.rcim.2020.102053>.
- [27] Changwei Miao, Guangzhu Chen, Chengliang Yan, Yuanyuan Wu, Path planning optimization of indoor mobile robot based on adaptive ant colony algorithm, Computers & Industrial Engineering, Volume 156, 2021, 107230, ISSN 0360-8352, <https://doi.org/10.1016/j.cie.2021.107230>.
- [28] Desheng Lyu, Ziwei Chen, Zesu Cai, Songhao Piao, Robot path planning by leveraging the graph-encoded Floyd algorithm, Future Generation Computer Systems, Volume 122, 2021, Pages 204-208, ISSN 0167-739X, <https://doi.org/10.1016/j.future.2021.03.007>.
- [29] Gong D, Zhang J, Zhang Y. Multi-objective particle swarm optimization for robot path planning in environment with danger sources. 2011;6(8):1554–1561.
- [30] Patle B. K., Parhi D., Jagadeesh A., Sahu O. P. Real time navigation approach for mobile robot. 2017;12(2):135–142.
- [31] Tang Q., Eberhard P. Relative observation for multi-robot collaborative localisation based on multi-source signals. 2014;26(4):571–591.
- [32] Ahmadzadeh S., Ghanavati M., Branch A., Branch M. Navigation of mobiler robot using the PSO particle swarm. 2012;2(1):32–38.
- [33] Islam, M. R., Tajmiruzzaman, M., Muftee, M. M. H., & Hossain, M. S. Autonomous robot

- path planning using particle swarm optimization in dynamic environment with mobile obstacles & multiple target. In International conference on mechanical, industrial and energy engineering, 2014;1-6.
- [34] Deepak, B. Parhi, D.R. 2012. Pso based path planner of an autonomous mobile robot. Cent. Eur. J. Comp. Sci. 2(2):152-168.
 - [35] Zhang, Y. Gong, D. Zhang, J. 2013. Robot path planning in uncertain environment using multi-objective particle swarm optimization. Neuro computing 103: 172– 185.
 - [36] Tang, Y. Li, Q. Wang, L. Zhang, C. Yin, Y. 2010. An improved pso for path planning of mobile robots and its parameters discussion, Intelligent Control and Information Processing, Dalian, China.
 - [37] Wang, L. Liu, Y. Deng, H. Xu, Y. 2006. Obstacle-avoidance path planning for soccer robots using particle swarm optimization, Robotics and Biomimetics, Kunming, China.
 - [38] Saska, M. Macas, M. Preucil, L. Lhotska, L. 2006. Robot path planning using particle swarm optimization of ferguson splines. Emerging Technologies and Factory Automation.
 - [39] Nasrollahy, A. Z. Javadi, H. H. S. 2009. Using particle swarm optimization for robot path planning in dynamic environments with moving obstacles and target. Sim European Symposium on Computer Modeling and Simulation.
 - [40] Solea, R. Cernega, D. 2016. Online path planner for mobile robots using particle swarm optimization, ICSTCC, Sinaia, Romania.
 - [41] Z. Batık, G. Atalı, D. Karayel, and S. S. Özkan, “Using heuristic distance functions in path planning of mobile robots,” Electronics World, pp. 34–36, Oct. 2018.
 - [42] Z. Batık, D. Karayel, S. S. Özkan, and G. Atalı, “A Comparative Study on Heuristic Distance Functions in Path Planning of Mobile Robots,” presented at the ICOME2017, İstanbul, 2017.
 - [41] Z. Batık, G. Atalı, D. Karayel, and S. S. Özkan, “Using heuristic distance functions in path planning of mobile robots,” Electronics World, pp. 34–36, Oct. 2018.

- [42] Gao, Z. Zeng, X. Wang, J. Liu, J. 2008. FPGA implementation of adaptive IIR filters with particle swarm optimization algorithm. *Communication Systems*.
- [43] Gigras, Y. Choudhary, K. Gupta, K. Vandana. 2015. A hybrid ACO-PSO technique for path planning. *INDIACom*, 1616-1621.
- [44] Sung-Il Kim, Ji-Young Lee, Young-Kyoun Kim, Jung-Pyo Hong, Y. Hur and Yeon-Hwan Jung, "Optimization for reduction of torque ripple in interior permanent magnet motor by using the Taguchi method," in *IEEE Transactions on Magnetics*, vol. 41, no. 5, pp. 1796-1799, May 2005, doi: 10.1109/TMAG.2005.846478.
- [45] M. Sarshenas and Z. H. Firouzeh, "A new wideband closed-form Green's function for microstrip structures using hybrid Taguchi-gradient optimization method," *7th International Symposium on Telecommunications (IST'2014)*, 2014, pp. 230-233, doi: 10.1109/ISTEL.2014.7000703.
- [46] D. Narin and T. Ö. Onur, "Investigation of the Effect of Edge Detection Algorithms in the Detection of Covid-19 Patients with Convolutional Neural Network-Based Features on Chest X-Ray Images," *2021 29th Signal Processing and Communications Applications Conference (SIU)*, 2021, pp. 1-4, doi: 10.1109/SIU53274.2021.9477882.
- [47] M. Wu, C. Heng, H. Zhu and H. Cai, "COVID-19 detection based on Computer Vision and Big Data," *2022 7th International Conference on Intelligent Computing and Signal Processing (ICSP)*, 2022, pp. 474-477, doi: 10.1109/ICSP54964.2022.9778429.
- [48] R. Qu, Y. Yang and Y. Wang, "COVID-19 Detection Using CT Image Based On YOLOv5 Network, " *2021 3rd International Academic Exchange Conference on Science and Technology Innovation (IAECST)*, 2021, pp. 622-625, doi: 10.1109/IAECST54258.2021.9695714.