X-RAY ANALYSIS PREDICTION OF BGA COMPONENTS IN PCBA PRODUCTION WITH NEURAL NETWORKS

A Thesis

by

Bayram Yurdakurban

Submitted to the
Graduate School of Sciences and Engineering
In Partial Fulfillment of the Requirements for
the Degree of

Master of Science

in the
Department of Computer Science

Özyeğin University
January 2021

X-RAY ANALYSIS PREDICTION OF BGA COMPONENTS IN PCBA PRODUCTION WITH NEURAL NETWORKS

Approved by:

Assistant Professor Furkan Kıraç, Advisor
Department of Computer Science
Özyeğin University

Assistant Professor Reyhan Aydoğan
Department of Computer Science
Özyeğin University

Associate Professor Arzucan Özgür
Department of Computer Engineering
Boğaziçi University

Date Approved: 14 January 2021

To mother and sister,

ABSTRACT

Printed circuit boards are the most important part of all electronic devices used today, and the production of these boards consists of many critical processes. Assembly production lines include different inspection machines such as Solder Paste Inspection (SPI), Automatic Optical Inspection (AOI), X-Ray Inspection Device to detect defects and misplaced components on the circuit boards. SPI and AOI machines determine these problems by checking that different measurement values remain between certain threshold values, but these machines cannot provide a perfect detection mechanism since threshold values are determined by a maintenance technician by trial and error. Therefore, the circuit boards are also controlled by X-Ray Inspection machines, thus the solder paste that cannot be inspected visually by human can be easily inspected. This project aims to create a neural networks model that uses SPI and AOI measurement parameters as input features and predicts potentially defective boards before the boards go through X-Ray inspection. This model allows suspicious circuit boards to be tested in X-Ray Inspection machine, instead of testing of some circuit boards randomly. Thus, the number of X-ray devices required to test the boards is reduced, the circuit boards to be selected for sampling are selected from the boards that are likely to be defective rather than randomly selected. By virtue of this model, suspicious solders and components will be marked on the boards that will be sent to X-Ray test station as suspicious and the operator will be prevented from missing the defects that may occur in these areas. It is foreseen that this system will also identify contact without connection defects, also known as head-in-pillow defects, which cannot be detected in SPI, AOI and Functional Verification Test (FVT) processes, thus the quality of the boards produced is likely to increase. The data will be collected for

17MB170R4 model circuit boards, which are the main board of televisions, produced in one of the production lines at Vestel Electronics Factory, and the model will be applied for four BGA elements on this board. In this project, it is aimed to work on BGA elements because the solder areas of these elements cannot be inspected visually after the component is placed on the board. Due to the model, it is possible to detect most of the defective boards by applying X-Ray testing to only 1% of all circuit boards produced.

ÖZETÇE

Baskılı devre kartları günümüzde kullanılan tüm elektronik cihazların en önemli bileşenidir ve bu kartların üretimi pek çok kritik süreçten oluşmaktadır. Montaj üretim hatları devre kartları üzerindeki kusurları ve yanlış yerleştirilmiş elemanları tespit edebilmek için Lehim Pastası Muayenesi (SPI), Otomatik Optik Muayene (AOI), X-Ray Muayene Cihazı gibi farklı muayene cihazları içermektedir. SPI ve AOI cihazları bu problemleri farklı ölçüm değerlerinin belirli eşik değerler arasında kaldığını kontrol ederek tespit eder ancak bu cihazlar mükemmel bir tespit mekanizması sağlayamaz çünkü eşik değerler bir bakım teknisyeni tarafından deneme yanılma yöntemi ile belirlenir. Bu yüzden devre kartları aynı zamanda X-Ray muayene cihazları tarafından kontrol edilir, bu sayede insan tarafından görülmesi mümkün olmayan lehim pastaları kolayca denetlenebilir. Bu proje SPI ve AOI ölçüm parametrelerini girdi özellikleri olarak kullanan ve devre kartları X-Ray muayenesine girmeden kusurlu olabilecek kartları önceden tahmin eden bir sinir ağı modeli oluşturmayı amaçlamaktadır. Bu model rastlantısal olarak bazı devre kartlarına X-Ray testi yapmak yerine gerçekten hata vermesi olası şüpheli devre kartların X-Ray testinin yapılmasını sağlar. Böylece kartların testlerinin yapılması için gerekli X-ray cihaz sayısı azalır, örnekleme için seçilecek devre kartları rastlantısal seçilmek yerine arızalı olması olası kartlardan seçilir. Bu model sayesinde X-Ray test istasyonuna şüpheli olarak gönderilecek kartlarda şüpheli lehimler ve elemanlar işaretlenerek operatörün de bu bölgelerde oluşabilecek kusurları gözden kaçırması önlenecektir. Bu sistemin aynı zamanda SPI, AOI ve Fonksiyon Doğrulama Testi (FVT) süreçlerinde tespit edilemeyen, yastık içinde baş hatası olarak da bilinen bağlantı olmadan temas kusurla-

rını da yakalayacağı öngörülmektedir, bu sayede üretilen kartların kalitesi de artacaktır. Veriler Vestel Elektronik Fabrikası'ndaki üretim hatlarından birinde üretilen, televizyonların ana kartı olan 17MB170R4 model devre kartları için toplanacaktır, bu kart üzerindeki elemanlardan dört adet BGA elemanı için bu model uygulanacaktır. Bu projede BGA elemanları üzerinde çalışılması hedeflenmiştir çünkü bu elemanların lehim bölgeleri eleman karta yerleştirildikten sonra görülemez. Model sayesinde üretilen tüm devre kartlarının sadece %1 ine X-Ray testi uygulayarak arızalı kartların büyük çoğunluğunu tespit etmek mümkündür.

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to Asst Prof Dr Furkan Kıraç, my advisor, for patiently answering my endless questions at every stage and for always supporting me. I owe a debt of gratitude to my dear friend Ezgi Özçoban, who revised my mistakes in each section, did not spare her support and always provided motivation and support for me to complete this project. Without her, this thesis would not be possible. I would like to thank my family, whose support I always feel by my side, who always trusts me in every decision I make and in my difficult times. I would like to thank Vestel Electronics employees for their efforts, who showed all the support for the commissioning of this project and made every effort to collect data from production at the most important stage. Finally, I would like to give my deepest respect and longing to Asst Prof Dr Pınar Boyraz Baykaş, my undergraduate advisor, who enlightened us, her students, on both scientific and engineering ethics, who loved for her job, and deeply saddened us with passing away too early. I promise that I will never leave the science and research in the way she taught me throughout my career.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

1.1 Overview of PCB Layout and Its Components

Printed Circuit Boards (PCBs) are generally green, non-conductive, epoxy resin boards and they contain copper conductive tracks, pads and other connections, thus providing electrical connection of the parts on them. PCBs contain electrical parts only on one surface, as well as different connections and parts on the other side. PCBs can be single-layered or many-layered that are isolated from each other and connected by conductive pads called vias.

The electrical part assembled by soldering to the conductive pads on the PCB is called component. Components are assembled in two ways based on their mechanical structure: Trough Hole Technology (THT) and Surface Mounting Technology (SMT). Components assembled by THT method have legs in the form of conductive wires and these legs are soldered after being inserted into the holes on the PCB. Components assembled by SMT method are soldered directly to the surface without passing through the holes in the PCB. Components that are assembled with SMT method are smaller than those suitable for THT method and it is possible to place more components on PCB. In addition, the production processes of SMT are faster and cost effective. In THT method, the components are relatively larger, they can only be mounted on the top surface of the PCB and they are costly compared to SMT components. The most important advantage of the THT method is its high reliability. Since the legs belonging to the component are mechanically inserted into the holes on the PCB surface, they are more suitable for adverse conditions such as mechanical impact and extreme temperatures. Therefore, they are especially preferred in the circuits used in defense industry and aviation industry.

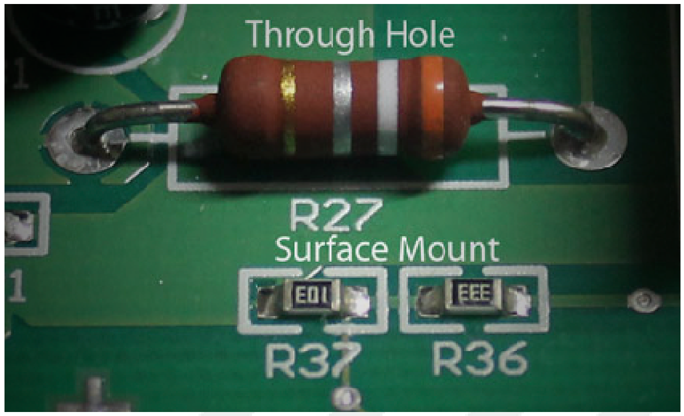

Figure 1: Resistors that are through-hole mounted and surface mounted. [1]

Nowadays, components suitable for SMT method are mostly used due to their low cost, less required space than THT components and their ability to be mounted on lower, upper and intermediate layers. Components that can be assembled by SMT method are called Surface Mounting Device (SMD). There are many different SMD components such as Metal Electrode Leadless Face (MELF), Small Outline Transistor (SOT) and Integrated Circuit (IC). Among these components, IC components are the most common types. IC components are small components that can contain many different circuits on a semiconductor material, and there are different IC components such as Small Outline Integrated Circuit (SOIC), Thin Small Outline Package (TSOP), Ball Grid Array (BGA). Unlike other IC components, BGA components are components whose solder points are in the form of a square matrix on the bottom surface of the component, not on the edges of the component. BGA components have solid solders at their soldering points. After solder is applied to the solder folds on the PCB, BGA components are placed, and the fusing process ensures that the solders on the component and the PCB are fused. The fact that the solder points are under the surface enables the PCB surface to be used more effectively, its electrical and thermal performance is better, and thinner boards can be produced with a very low component thickness, so they are frequently used in this respect [21]. The only disadvantage of this component is that the solder pads cannot be seen visually because they are under the component and the rework process is difficult. BGA type components will be examined in this project since they are used on almost all of boards today and it is very difficult to figure out if there is a problem for the solder paste on the pads.

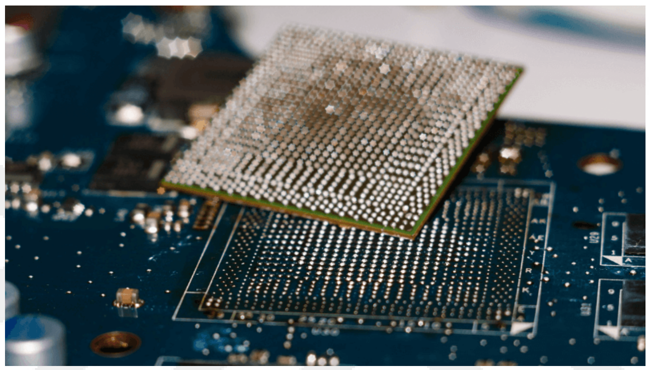

Figure 2: A BGA component and solder pads on both PCB and the component. [2]

After all the components are placed and soldered on the PCBs and they are ready to be used, the boards are called Printed Circuit Board Assembly (PCBA). PCBAs are named as PCBs in the beginning phase of the production. After the solder paste is applied on the solder folds, the components are placed, and the boards are heated up in oven to fuse the solder. Then the boards leave the production line as PCBAs. Process phases that are applied on PCBs in the production will be examined in Section 1.2. in detail.

PCBAs are produced as a whole board, named as a panel, combined with same boards which have the same design and components for the ease of production. Each PCBA with the same features in a panel is called multiplication. A panel can be composed of one multiplication as well as two or more of them. In Vestel PCBA factory, each PCBA board has a serial number and the serial number of each PCBA is different from each other. Each panel which includes many same PCBAs has its own unique number, this number is called master code. PCBAs exist as panels throughout placing and soldering of components. At the end of the production line, an operator divides the panels into PCBAs by using laser cutting machine. Therefore, while the tests on the production line such as Solder Paste Inspection (SPI), Pre Automatic Optic Inspection (Pre-AOI), Post Automatic Optic Inspection (Post-AOI) are performed with master code on a panel based, the tests which are performed after separating PCBAs such as X-Ray Inspection, Functional Verification Test (FVT), In-Circuit Test (ICT) are matched by serial numbers.

Each component on PCBAs has a different name so that they can be distinguished from each other during tests. In Vestel PCBA factory, components are named using one or two alphanumeric characters and adjacent numerical characters. For the conductive legs of each component to be soldered on the PCB, there are corresponding conductive copper folds. Generally, the conductive folds on the PCB are called pads, and the legs corresponding to these pads on the component side are called pins. However, after the components are soldered and joined to the PCB, pad and pin names can also be used to address to the same point. In this thesis, both the pad and pin names will be used frequently in the same sense to describe these conductor folds on the PCB.

1.2 PCB Assembly Process in Production Lines

As stated in the previous section, the boards are PCBs without any components at the very beginning of the production line, after all production processes are completed, they are packaged and shipped as PCBAs. In this section, the processes which are applied in the transition from PCB to PCBA will be examined in detail. As PCBs move along a production line, they pass through a series of machines arranged in series and they are processed and controlled by each different machine. The first machine that PCB is processed by is Screen Printer (SP) machine. In this station, a

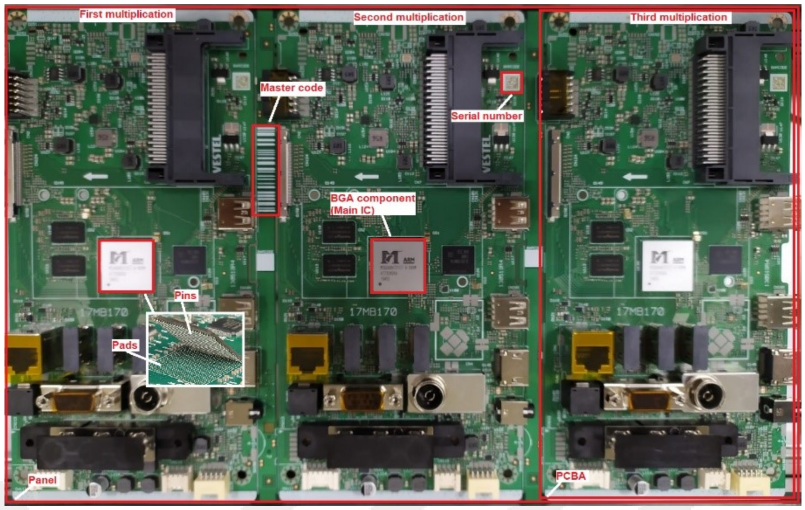

Figure 3: A panel which is named as 17MB170R3 and produced in Vestel.

plate called stencil is placed on the PCB, on which only the regions corresponding to the solder pads on the PCB are empty and all other parts are solid. There is clay-like solder paste on one side of the stencil before the process. Solder paste is a material that contains a mixture of solder and flux. A squeegee pass over the stencil and rubs the solder paste over the stencil, ensuring that the solder paste is equally distributed over the gaps on the stencil. Next, the stencil is removed from PCB, so that the solder paste is remained only on the solder pads. It is possible to adjust different parameters on the machine to ensure that the amounts and shapes of the solders are homogeneous. Adjustable machine parameters such as the speed of squeegee, the amount of pressure the squeegee applies on the stencil, the angle of the squeegee, the speed of separating the stencil from PCB directly affect the quality of the solder applied to the solder pads. When a new type of PCBs is being produced for the first time in a production line, a technician adjusts and saves these parameters until high-quality solders are obtained on PCBs. Screen printing process is very critical because if there is any problem in the amount or shape of the solders that SP machine applies to the solder pads, when the components are placed on the PCB, the solders on the pads may dissipate and problems caused by soldering such as short circuit, missing solder, cold soldering may occur. For these reasons, right after SP machine, PCBs are controlled by Solder Paste Inspection (SPI) machine to inspect applied solders.

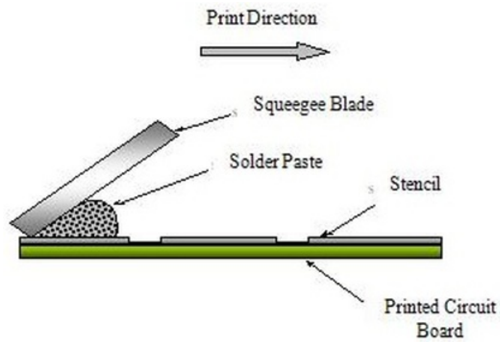

Figure 4: The process of rubbing solder paste over a PCB in SP machine. [3]

SPI machine is the first machine in the production line that the control process is applied. In this machine, optical control of the solder paste on the pads is performed. The machine has more than one angled camera and lighting. Angled cameras capture the image of solder paste on the PCB from different angles and the machine calculates and records measurement values such as volume, area, height, offset for each pin by image processing techniques. It is possible to set lower and upper limit values for each measurement, so if any deviation is detected in any of the controlled pins as a result of the measurement, the machine generates failure result for that pin and warns the operator. The main technical features of SPI machine are the number of cameras and lights, and the speed of test completion. On relatively low-cost machines, the number of lights is limited, which may sometimes cause shadows in the images captured by the machine and measurement results are less reliable. Test completion speed is another feature that determines how many seconds the control process last with units of sec/cm^2, regardless of PCB size. High speed machines are preferred in production lines with high daily production quantities.

The measurement results of SPI machine are one of the most important control processes in the production line because if there is a defect in the solder paste that is caused by SP machine, this defect must be caught in SPI machine. Otherwise, the PCB will continue to move along the line and all the components will be placed and soldered. Thereafter, if a problem is observed in the tests performed after the production, the problematic component is heated, separated from the board, the solder is cleaned and a new component is placed by applying solder again, this process is called rework. Rework operations are costly and risky because interference of a technician is required. However, in the case of a defect in the solder paste is caught directly in SPI machine, the PCB is removed from the line and transferred to Stencil Room, all solder paste is washed and cleaned, and the clean PCB without solder is moved to the beginning of the production line. This process is a very low cost and more reliable process than rework process.

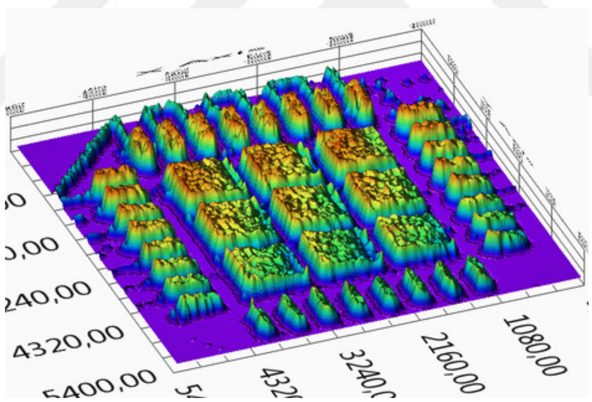

Figure 5: 3D image of the solder paste on a PCB obtained by SPI machine. [4]

After the solder paste on the pads is controlled by SPI machine, if the machine has decided that all solders are valid or if the solders with some measurement values very close to the threshold values are visually checked by the operator and no problem is observed, the PCB goes inside Pick & Place (P&P) machine to make the components placed on them. In P&P machine, the PCB is fixed mechanically first, and the cameras inside the machine check for the reference marks on the panel called fiducial marks to make sure that the PCB is positioned correctly. P&P machine has feeding mechanisms called feeder where the components to be placed on PCB are stored. These mechanisms can be in the form of reel / tape, tube or tray. An operator continuously feeds new components to the machine through the feeder. After the board is positioned correctly, a robotic arm operating in Cartesian coordinates (x, y and z axes that are perpendicular to each other) picks the components from the feeder mechanism via vacuum pads and places them in their positions on the PCB according to a previously created software on the machine. There is also a camera on the arm, so it is visually checked whether all the components taken are not physically damaged and that all legs are in good condition and straight.

As previously described, a PCBA production line is a line where multiple machines are lined up one after another along a single vertical line. In order to provide the production line to work continuously without stopping, the processing time of each operation performed in each machine must be very close to each other. Otherwise, the boards start to accumulate in the stations which are behind the station that the processing time is too long, and the production runs intermittently. Since the longest process time in PCBA production belongs to the process of placing the components on the PCB, all the components are not placed at once, but are partly placed at two or more stations. For this reason, there are more than one P&P machine in the lines and the component placement process is completed after PCB is processed in all these machines one after another.

After SMD components that are suitable for SMT type mounting are placed on the PCB via P&P machines, the boards go to Pre-Reflow Automated Optic Inspection (Pre-AOI) machine where the second control process is applied. The cameras inside this machine check the soldering of the components, the correctness of their polarity, and that there are no missing or short circuit components. Pre-AOI machine also records the measurement values of all components such as position, offset, angle and probability of failure. When there is a problem with measurement values of the

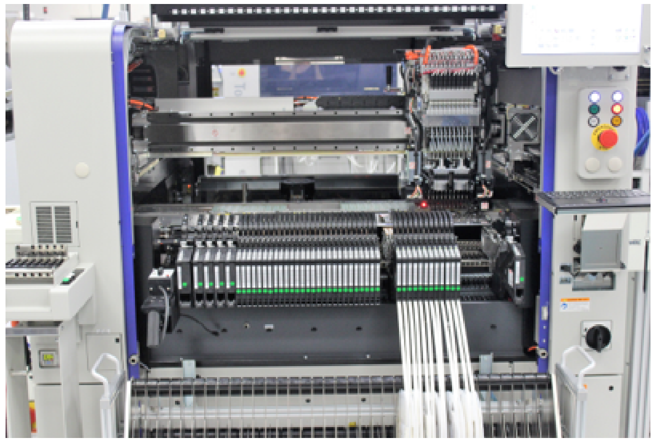

Figure 6: A pick and place machine and its reel-typed feeder mechanism. [5]

components, the machine warns the operator. If the operator detects a problem with any of the components, he / she transfers the board to Stencil Room in order to make the board cleaned from all the components and solder paste and then restarted the whole processes again from the beginning. Pre-AOI control machine is one of the most important stations because it is the last checkpoint before the board goes to the oven. If a defect in the components cannot be detected here and the board goes to the oven, then a rework process must be done to repair this component, and this process requires human interference, so it requires a lot of cost and may involve errors.

The boards that pass all the controls in Pre-AOI machine successfully goes to Pin in Hole Reflow (PIHR) machine, which is the last station before the oven. In this machine, the components which are suitable for THM method (Through Hole Device - THD) are placed by passing their legs through the holes on the PCB as in P&P machine. The solder paste on the solder pads with holes is applied in SP machine. When PIHR machine places the components that are suitable for THM, the solder paste is pushed down through the hole. In this way, THD components are heated up in oven together with SMD components without the need for wave soldering the bottom of PCB in a separate station.

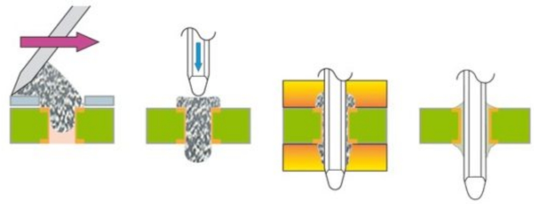

Figure 7: Assembly of THD components by PIHR method. [6]

By placing THD components as well, the PCBs, of which the solder paste is applied and the placement of all components has been completed, are heated up in Reflow Oven in order to fuse the solder on the pads and pins, and cooled down to permanently mount the components to the PCB. The length of Reflow Oven depends on the production capacity of the line and the melting temperature of the solder paste used in SP machine. In addition, while determining all parameters of this machine, the following features of the board should be considered: Solder paste type, PCB material, PCB thickness, number of layers, copper density on the PCB, the number of SMD components and the type of SMD materials [10]. PCBs in Reflow Oven go through four different stages, and these stages are as follows: pre-heating, soaking, reflow and cooling. In pre-heating stage, the boards are heated to reach the soaking temperature. Heating rate is very important at this stage and other stages also, because if heating occurs too fast, components may crack, or small solder balls may form around the solder paste. In soaking phase, the aim is to make sure that melting temperature is reached enough in every part of the board before entering the reflow phase. In reflow phase, the ambient temperature is adjusted to a temperature higher than the melting temperature of the solder, and all solders on the board become fluid. Ambient temperature and exposure time of the board are very important in this stage. If the board is exposed to too much heat, some of the components may deteriorate due to excessive heat, if exposed to too little heat, the solders may not melt properly, and problems called cold solder may occur. In reflow phase, nitrogen is applied on the boards to clean intensive fluxes. The last stage is cooling stage, at this stage, the board is cooled down in a controlled manner. If cooling occurs too fast, bends and distortions may occur in PCB and components.

When the PCB comes out of the oven, all solders are fused, and the components are permanently mounted on the PCB. However, it is possible for the components or solders to be damaged in Reflow Oven machine, as the PCB is heated to high temperatures and cooled again. For this reason, the boards are also controlled by Post-Reflow Automated Optic Inspection (Post-AOI) machine, which is the last station that the control process is applied and the end of the production line. The controls made by this machine are similar to those made on Pre-AOI machine. Post-AOI machine checks whether there is slippage, distortion, deflection or deficiency in the components after the heat treatment applied in Reflow Oven and records the measurement values such as position, misalignment, angle and probability. Unlike the other control points, since all components are mounted permanently after the oven, in case of a problem in the components or solders, they cannot be directed to Stencil Room to be sent to the beginning of the line after cleaning process. If the defective component is a simple and easy-to-install component, the repairman standing next to the production line heats up the component and solders a new component. If the defective component is complex and contains too many solder pads, the board is transferred to Rework Room for the replacement of this component, where a more knowledgeable and skilled technician performs the replacement of the component by using more professional equipment.

PCBAs that pass all controls on Post-AOI machine leave the production line. In a separate laser cutting station, an operator divides panels which contain many multiplexed boards into single PCBAs. Afterwards, the separated boards are directed to the test line to ensure that they are working correctly and meet desired performance

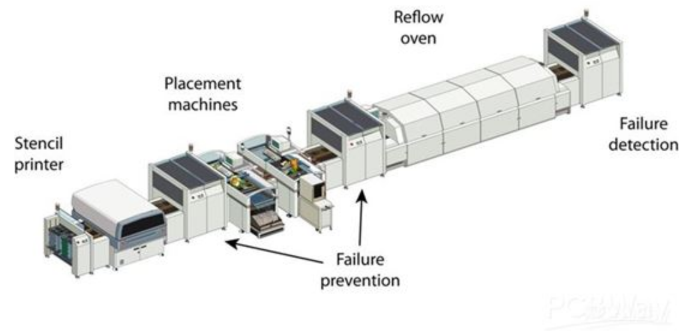

Figure 8: A general layout scheme of a PCBA production line. [7]

criteria before shipping to the customer. In test line, operators place the boards in a mechanism called gauge or jig and energize the boards. They install the necessary software and firmware into the processor of the boards via a desktop computer and run the test software to make sure that all the technical features are working correctly. This process is called Functional Verification Test (FVT). The boards then pass through ICT station. In this station, PCB is fixed with clamps first and a mechanism consisting of many conductive needles called bed of nails is contacted on the test points on the PCB. The electrical properties of the boards such as short circuit, resistance, and capacitance are examined by applying different voltages and currents to the circuit. If the PCBA is not successful in any of these tests, the board is directed to Rework Room, after the analysis is performed to find out relevant errors, component replacement is made by the technician if necessary. If the boards pass all these tests successfully, they are placed in boxes which have Electrostatic discharge (ESD) feature by a mechanism and then shipping operations are provided.

Some of the boards that passes the controls of Post-AOI machine successfully are randomly sampled and examined on X-Ray Inspection machine before the test line. The main purpose of X-Ray control is to find errors that cannot be detected during the controls performed along the production line, and to determine the stations related to these errors and to make improvement studies. Additionally, since the solder points of some components are under the component, which is similar to BGA type components, it is not possible to control these solders on AOI machines or visually by the operators along the production line. These components are especially examined in the boards selected by random sampling method, and in case of any problem, relevant stations are reconsidered.

It is not possible to include X-Ray Inspection machine in the production line because the control process time of these machines is quite long, and it is not suitable for the capacity of the production line. Moreover, these machines are very expensive, so X-ray testing of all produced boards is not preferred due to both time and cost excess. With the work carried out in this project, a model will be trained with measurement values taken from SPI, Pre-AOI and Post-AOI machines, and it will be provided that the prediction of X-Ray result of each pin for BGA components is correct. In this way, the boards with very low error probability according to the model will not be tested and only suspicious boards will be tested and a more logical sampling method without the risk of error will be achieved.

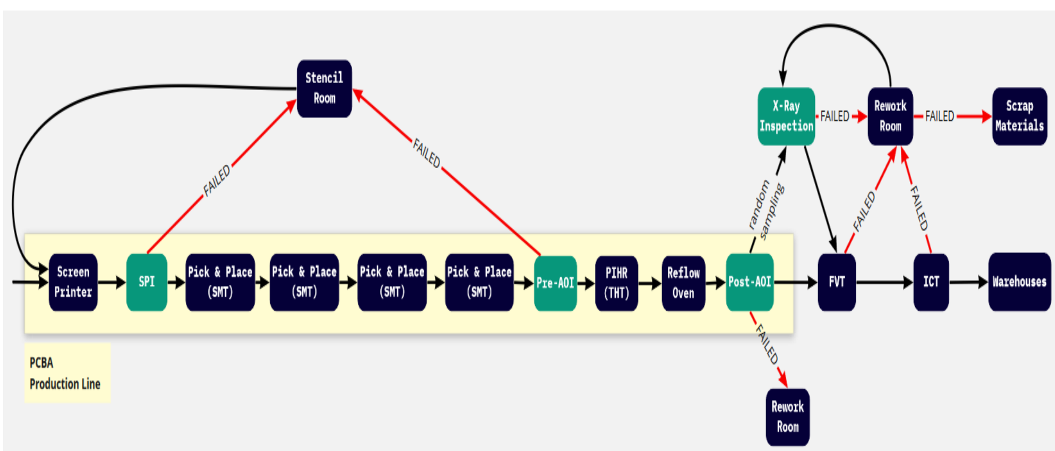

Figure 9: A diagram that shows whole processes of PCBA production in Vestel.

1.3 Main Challenges in PCB Assembly Process

PCBA production consists of processes that involve many risks in each step. It is possible to categorize these risks into material-based risks, production-based risks, human intervention-based risks and inspection-based risks. This section will be examined under the headings of these risks related to PCBA.

Generally, PCBA production is entire of assembly processes. Solder paste is applied on the PCB and all electrical components are mounted on the PCB. Therefore, the PCB, the solder paste and all the components are supplied from different manufacturers. These materials are tested and examined by a department called Quality Assurance by sampling randomly. In case of an error, the batch of materials containing that material is completely rejected and the manufacturer is warned. However, since it is not possible to control all incoming materials due to high costs, it is of great importance to identify defective materials and components during the production process.

Figure 10: Solder printing process in SP machine. [8]

As explained in detail in the previous section, the production line consists of many machines on which many different processes are applied on the PCB. These machines are automated machines, but their operating characteristics are determined by the parameters which are set by a technician. In SP machine, while the solder paste is applied on the PCB, the stencil which has holes only in the areas where correspond to the pads is placed and the paste is allowed to flow through the holes to the pads. If there is an obstruct in one of the holes on the stencil, the paste will not flow through this hole and the pad will remain without solder. The angle of the squeegee and the pressure that is applied by the squeegee in order to rub the solder paste over the stencil are very important. If the technician sets the pressure value lower than required value, the solder paste cannot penetrate to the pads sufficiently through the holes on the stencil and lack of solder defect occurs. After the solder paste is applied on the stencil, the stencil is detached from the PCB and only the solder paste remains on the solder pads. If the stencil is detached too quickly during this separation process, the solders on the pads are dispersed and the shape of the solder is distorted. The defects that can be caused by solder paste are summarized in Figure 11. Slumped print defect is usually caused by melting of solder paste when the PCB is processed at temperatures above the recommended value for the solder paste. Scavenged print defect occurs when the pressure that is applied by the squeegee to the stencil is too high. Due to the excessive pressure, the liquid in the middle of the solder is distributed towards the edges and a hole occurs in the middle. Bridging defect is an error caused by the stencil not being properly placed on the board or the stencil not being clean. This defect is very dangerous because two solder pads are conducting with each other and a short circuit occurs when the board is energized. Peaking defect, also called dog ear, is caused by the rapid separation of the stencil from the PCB.

In P&P machine where the components are placed on the PCB, it is possible to have some risks if maintenance of the machine is not performed periodically. If the vacuum nozzle on the robotic arm carrying the components and placing them on the PCB is not cleaned properly or is not replaced with new nozzle due to wear and corrosion over time, the component may fall off the arm, or the component may slip

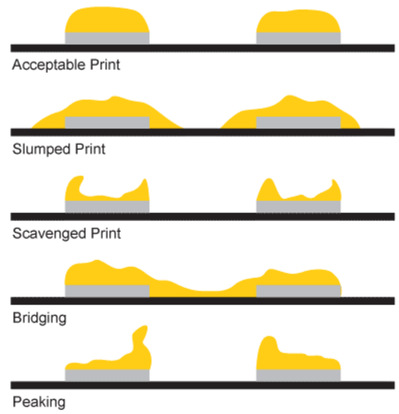

Figure 11: Possible defects that can be seen for the solder paste on the pads. [9]

during the transportation process, which results an offset as it is placed on the PCB. In addition, if the nozzle which is composed of right material is not used, it may damage the components while taking and cause the PCBA to fail in future tests. One of the processes with the highest probability of error is undoubtedly Reflow Oven machine. Since the PCB is first heated and then cooled with a predefined heat profile, many different types of defects may occur if the PCB is exposed to undesired temperature change. If the PCB is subjected to extremely rapid temperature changes during pre-heating phase in Reflow Oven machine, the components may crack, or small solder balls may form around the solder paste. If the heat in soaking phase is applied for too long to ensure that the temperature is balanced throughout the board, the solder paste may fatigue, and bubbles may occur on the solder. In reflow phase, the PCB is heated above the melting temperature of the solder paste, if the PCB is exposed to more heat than it should, the components on it may deteriorate or the material characteristics of the solder may change and become more brittle. This results in the lack of permanent adhesion and the component can fall off easily.

A common defect which is called tombstone, is seen as the component levitating to one side with more adhesion in one of the pads contacted by the component after the solder is heated and cooled. It is known that this error is usually caused by the

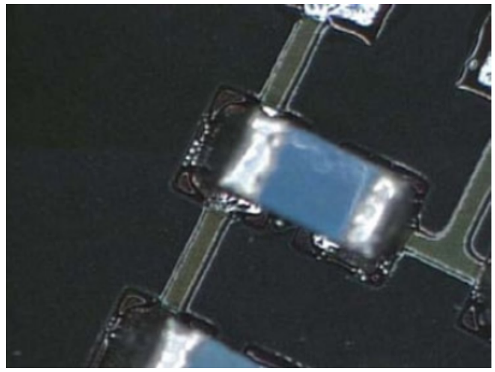

Figure 12: Solder ball issues caused by excessive heating in Reflow Oven. [10]

wrong heat profile applied to the PCB which results the solder on one pad melts and the other does not melt, or the wideness on one pad being more than the other. Perhaps one of the worst defects that can occur in Reflow Oven machine is called "head-in-pillow" defect. This defect usually happens in BGA-like components and is seen as the solder on the pins on the bottom of the component and the solder paste on the pads of the PCB contacting each other but not fully fused. This kind of defect cannot be caught in Pre-AOI and Post-AOI inspection, as well as in tests such as FVT and ICT, because the solders on the problematic pad and pin contact each other even though they are not fully fused. After the PCBA is shipped to the customer and used for a while, the unfused solder is dispersed again due to the mechanical and thermal stress on the component and the PCBA will fail. Although there are many different reasons for this defect, one of the most important reasons is that the quality of the solder paste is poor, and the heat profile applied in Reflow Oven is not correct [22].

All processes in PCBA production are carried out by automated machines without any human intervention. However, there are open-top conveyors between these machines, the boards are transported from one machine to another by means of these conveyors. Since the components are not permanently mounted on the PCB before Reflow Oven machine, during the movement of the boards, the wrong intervention by the operator may cause the component to slip and the solder paste to deteriorate.

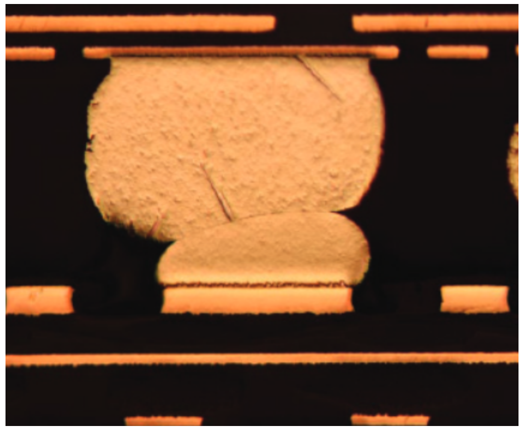

Figure 13: Head-in-pillow defect that can be seen on the solder pads of BGA typed components. [11]

Furthermore, if SPI, Pre-AOI and Post-AOI control machines calculate that some measurement values do not exceed the limit values, but they are too close, they give a warning instead of evaluating the board as defective. The final decision is expected by the operator in this situation. The operator visually checks the component or solder paste and decides whether there is a problem. Although this decision is made by authorized and knowledgeable operators, it does not contain absolute accuracy, the operator may not notice any defects and these boards may fail in the next control station or test line.

Inspection-based risks are caused by incorrect threshold values set on the control machines. In SPI, Pre-AOI and Post-AOI control machines, different measurements are made for the solder paste and components on the PCB, and each measurement has lower and upper limit values. If the measurement values for all components in a PCB are between these two limit values, the control machine does not give a warning and the PCB passes through the machine and proceeds to the next station. If any of the measurement values are close to the upper or lower limit, the machine warns the operator. If a value is outside of these limits, that board leaves the production line as defective. In other words, it is determined by the machine whether the card is passed or failed according to the measurement values and this decision directly depends on the lower and upper limit values. Limit values of control machines are set by authorized technicians, these technicians determine the limit values with their experience and knowledge. However, although the limit values are set to suitable values as much as possible, it is not possible to set the most optimum limit values in practice. If the limit values are set in a very narrow range, the machine defines many boards as defective and unnecessary interventions cause the production line to slow down. If the limit values are set in a very wide range, since the measurement values of the boards that may fail will remain within the limit range, they pass to the next station without any warning. In this case, the components that are actually failed cannot be caught by the control machine.

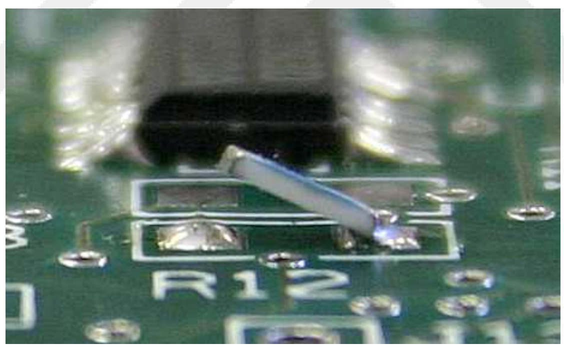

Figure 14: Tombstone defect that can happen after the reflow process. [12]

Most of the failures that can occur during the production described above are detected by SPI, Pre-AOI and Post-AOI control machines, but there are some defects that these machines cannot identify. These defects are those that cannot be detected visually by human eyes or cameras when looking at the top surface of the board and can only be detected by X-ray beams. The defects may occur in the components which stay in the intermediate layers of the PCB, or some defects can be detected inside of the components. The defects can also be observed in the solder pads that are between the bottom surface of the BGA type components and the top surface of the PCB. In this project, it is aimed to predict the results of X-Ray inspection with the measurement values of SPI, Pre-AOI and Post-AOI machines, for a specific type, BGA type components. By identifying the boards that may be suspicious to pass X-Ray analysis, it will be ensured that the boards that may contain defects are transferred to X-Ray Inspection without being missed. In this way, instead of performing X-Ray testing by random sampling, X-Ray test will be performed for the boards with suspected defects and a more logical sampling method will be provided. By testing fewer boards, the cost of operation will be reduced, the maintenance cost will be reduced due to the less frequent use of existing machines, and the cost will be reduced by decreasing the number of X-Ray devices that are needed to be purchased to inspect the boards in all production lines by X-Ray Inspection. Solder defects such as head-in-pillows, which cannot be detected in any test on the production line, will be detected using measurement values in the control machines and the quality of the PCBAs obtained as the final products will be increased since the possibility of missing the defect is removed.

CHAPTER II

LITERATURE REVIEW

The difficulties during the processes in PCB and PCBA production and the defects that occur in the products caused by these difficulties have led to a number of studies in this field. S. S. Zakaria et al (2020), who examined the studies related with detecting defects in PCB production in the literature, stated that the studies conducted in this area are limited and early detection of defects is still not fully discovered [23]. One of the most important factors of the limitation of these studies is the lack of structure to collect data from all machines in the production line in PCBA factories. Foster (2019) stated that Flex, one of the biggest companies in this sector, has 1500 PCBA production lines all over the world, does not have a sufficient structure to record the data and images taken from the machines [24].

In light of research about PCBA production, it is realized that the most important reason of the defects in PCBA production is caused by SP machines. Parameters such as the pressure value applied by the squeegee on the stencil, separation speed of the squeegee, quality of the solder paste affect the quality of the solder paste directly and the quality of the produced PCBAs indirectly. In a study, it was observed that 52-71% of the defects observed in the assembly of components suitable for the SMT method were directly caused by the application of solder paste [13]. Consequently, the research projects in this area are generally focused on examining SPI data or examining the solder paste on the PCB with computer vision algorithms in order to detect the defects in solder paste.

It is possible to categorize the studies on the automatic detection of defects in PCBA production under two subcategories with studies made on PCB images using

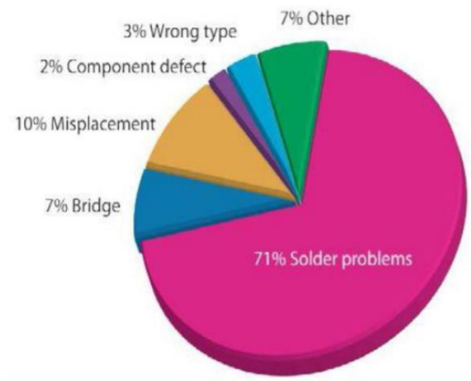

Figure 15: Defect distribution of SMT components in PCBA production. [13]

computer vision algorithms and studies made on the measurement data obtained from inspection machines. In this section, these studies will be examined under these subcategories and their advantages and disadvantages will be discussed. To briefly summarize the research in the literature, the studies to predict defects in PCBA production have been limited due to the minority of the data. Since the data collected to train deep layered neural networks is not sufficient, these models have been underfitted and cannot not show sufficient performance. In addition, in studies that use measurement values, only the values obtained from SPI machine were generally used as inputs and while these values were labeled, the results of ICT or AOI stations at the end of the production line were considered. Since the boards are performed by many different processes such as component placement and reflowing solder between SPI and AOI stations, the recall value remains low due to the fact that the labels do not only depend on the values obtained from SPI machine. In this project, by means of previous work to collect as much as possible data from the machines in the production lines, the dataset is created with the measurement data of SPI, Pre-AOI and Post-AOI control machines. Furthermore, in the studies related with detecting defects, the results of the boards such as defective or not defective were used to label all pin values of a board instead of using more precise pin-based results. In this project, the measurement values are labeled at the X-Ray Inspection station with defective or not defective results on pin-based level. In other words, data collected from the machines are labeled more precisely.

2.1 Research Related with Computer Vision

Two approaches are generally observed in the studies for defect detection with computer vision. These studies can be focused on detecting defective solder paste or component by using two or three-dimensional images obtained from existing SPI, AOI or X-Ray Inspection control machines on the production line, as well as a different imaging system can be set up and defect detection analysis can be done by different cameras. Using images by AOI inspection machine, Song et al (2018) developed a model to classify complex defects in ten categories from the direction and location of the component, the quantity and quality of solder paste, and both the component and the solder paste. They divided the component images they obtained from AOI machine into regions called feature extract regions. Thereafter, they obtained average intensity of blue and red color values for each region in the images obtained with blue, green and red lighting. Initially, they divided the images into 3, 9 and 15 regions and classified them using Support Vector Machine (SVM) and Multi Layered Perceptron (MLP) models. In this study, they found that using SVM model by dividing images into 9 regions gives better results. Furthermore, instead of manually dividing the images into a certain number of regions, they found the optimum number of regions using Genetic Algorithm method. As a result of this study, they determined the optimum number of feature extract regions as 47 regions and achieved 96% accuracy with SVM method. These studies were applied only to the resistor and capacity components, which have only two solder pads, and the scope of work is limited because the system can only be applied to the images taken with RGB lighting [25].

Acciani et al (2005) developed a system similar to AOI control machine with basic cameras and created a model that classifies the solder in the legs of the components on PCBs into five classes from less solder to excessive solder. They developed an algorithm that captures the board at first, then the relevant component and the solders of the component in PCB images obtained from the camera, and thus determines Region of Interest (ROI) in the PCB image. Afterwards, they obtained Geometrical and Wavelet transform features for the solder areas they acquired and classified them with Linear Vector Quantization (LVQ), MLP, K-Nearest Neighbor (KNN) models. The models were trained on number of 640 PCB images. As a result, MLP model trained using GW features achieved 98.8% success for recognition rate (RR), which is the ratio of correct predictions to all predictions. One of the greatest achievements of this study is that it does not require a special mechanism that provides precise positioning of PCBs and a system that includes expensive advanced cameras for classification of solder pads on PCBs [26].

Wu et al (2005) developed a system that classifies the solder paste on PCBs into five categories by using the cameras of SP machine. In the system they developed, the PCB is aligned first and ROIs are obtained. Later, they obtained 2D and 3D features over ROIs with Otsu algorithm and they trained an artificial neural network model using these features [27].

2.2 Research Related with Measurement Data

Automatic Mistake Reduction (AMR) system was developed by Yi-Ming Chang et al (2018) that uses measurement values of SPI machine to reduce the number of boards with false failures on AOI control machine located after Reflow Oven machine and to increase the efficiency of the production. They generally formed this system in online and offline stages. In the offline phase, they preprocessed the volume, area, height and offset values obtained from SPI control machine and observed the distribution in the data. In the online phase, they trained a five-layer deep learning model that includes an input, an output, and three hidden layers with number of 15000 training data. They labeled the data that they used while training this model with fail and false fail results of AOI machine. As a result of the study, they obtained 98.1% Root Mean Square Error (RMSE) values in training data and 92.5% in validation data with the model [28].

LaCasse et al (2020) developed a model that uses SPI measurement values to predict ICT results of BGA typed components. Because the defects of BGA typed components are difficult to detect and the rework process is costly, they aimed to predict possible defects in ICT with the system they proposed, and to clean solder paste from PCBs and reapply again without placing the components in SPI station. Hence, they developed two models, the first model is pin-based, and the second model is component-based. In the first model, they extracted 16 features with the measurement values they obtained from SPI control machine and used the results of the ICT station as the output of the model. In the second model, they extracted features for each component with TSFRESH which is used as a feature extraction tool in time series data and used Decision Tree model. When training the models, they used two separate datasets containing 930955 and 806321 records and compared the results. As a result, it was observed that the first model achieved 38% precision and the second model achieved 83.3% precision and the recall values remained lower than the precision [14]. One of the biggest disadvantages in the study is that the board is exposed to very different processes between SPI and ICT stations. Since there are many different stations such as P&P, Reflow Oven between these two stations, not all of the results in ICT reflect SPI measurement values. In addition, the measurement values obtained from SPI station are pin-based level and the results at ICT station are component-based level. This caused the pin-based model to give a lower performance result. Another drawback in the pin-based model is that the model does not contain the neighborhood information of the pins. As a result of SPI, the measurement values of solder pastes may be within the desired range, but if measurement values of two adjacent pins are close to the limit values, this may cause a problem in the downstream processes. In the component-based model, since the measurement values of all pins are used to train the model at once, unlike the pin-based model, this model can learn the relationship between the values of neighboring pins [14].

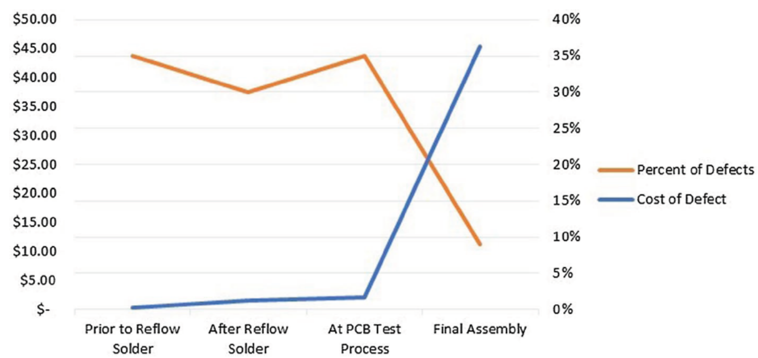

Figure 16: Cost of unrecognized defects and percentage of defects for the stations in PCBA production. [14]

Foster (2019) developed a model that predicts X-Ray Inspection results of boards using the measurement and parameter values in SP and SPI machines. The purpose of this study is to intervene PCBs in SPI station in order to remove the solder paste that may cause problems in X-Ray Inspection, and to eliminate the defect at the very beginning of the line. Because the cost of loss increases exponentially at each station, unless a defect is not caught immediately. In this way, even if a defect is not identified at SPI control station and instead of being detected in X-Ray station, PCBs will be cleaned and reprocessed as a result of the output of the model, and high-cost errors will be eliminated [24]. For this purpose, panel-based and component-based models were developed. In the panel-based model, required features were acquired by applying Principle Component Analysis (PCA) on the properties obtained from the data in SP and SPI machines, and then developed Random Forest, Convolutional Neural

Network and Boosted Tree models. While training the models, dataset is created with the measurement data of number of 1689 panels, and for labeling this data, X-Ray Inspection results of the panels were used. As a result, it was understood that panel-based Boosted Tree model gave the best result with 0.74 AUC. In the component-based model, the min, max and average values of all pins for a measurement value of a component were used, since the number of solder pads of each component and the number of SPI measurement values are different. It has been observed that the component-based model performs less than the panel-based model, but Random Forest model succeeds when the model is used for only a certain type of component such as IC components [24].

The most important advantage of this project from the studies in the literature is that the model that is developed for defect detection has been trained not only with measurement values from SPI control machine, but also from SPI, Pre-AOI and Post-AOI control machines. Thus, the possible problems that may occur in every processes of production can be learned, and a more accurate defect prediction mechanism can be achieved. Besides, the datasets in the previous studies could not be labeled precisely due to the limited data collection system, which allowed the data were labeled only on panel-based or component-based level. In this project, as a result of preparatory work done for X-Ray Inspection machine, defects are labeled on pin-based level and matched with measurement values of other control machines. Consequently, the data is labeled more precisely, and it is aimed to train a model with higher precision.

CHAPTER III

METHODOLOGY AND RESULTS

3.1 A Brief of Main Problem and Target Achievements

While manufacturing PCBAs along the production line, the PCBAs go through many different critical processes and these processes are controlled by inspection machines at the control points. As explained in detail in Section 1.3, it is not possible to adjust the measurement threshold values in control machines to the most optimum value in order to ensure a perfect control mechanism. Along with that, the measurement results depend on the decision of the operators at some points. Since all possible errors cannot be detected by the machines, the produced boards are sampled and inspected in detail in the X-Ray inspection device. However, the inspected boards are selected by random sampling, it is possible to miss out PCBAs being not selected and containing subtle defects which are not determined by the control machines. That may results dispatching the defective boards without being detected. In this project, it is aimed to build a neural network model will be trained with the data obtained by merging measurement values from the control machines and the results in X-Ray inspection device. In this way, the PCBAs having high probability of failure in X-Ray inspection according to the measurement values obtained from SPI, Pre AOI and Post AOI machines will be detected by the model and it will be provided that these boards are inspected on X-Ray device. Compared to randomly sampling method, the boards will be sampled in a more rational way, therefore defective boards that may be dispatched without inspection will be detected and the defects of end products will be reduced to zero.

X-Ray inspection devices are more expensive than other control machines due to their technology and measurement capabilities, and high maintenance costs. In addition, X-Ray inspection requires more time than other inspections, so these devices do not take part in the production line and they are used as a separate station in order not to affect the capacity of the production lines negatively. Factors such as lack of sufficient number of X-Ray machines due to the excessive cost of X-Ray machines, increase in maintenance cost due to a need of changing the parts of the machine more frequently in case of intensive use, cause the produced boards to be tested in the X-Ray device with less counts. With the aimed model, according to the measurement values taken at the control points on the production line, the pins that are likely to be caught as defective in the X-Ray device will be determined and only the boards with these pins will be tested in X-Ray devices. In this way, the boards without any problems will be removed, the need for X-Ray machines will be reduced and the maintenance cost of the existing machines will be reduced by enabling fewer boards to be tested. Approximately 5% of the cards produced at the Vestel Electronics Factory, where the project is implemented, are randomly selected and inspected by X-Ray devices. In addition, if the boards are shipped to another factory and any defect is detected afterwards, all of the boards in the same shipment are inspected in X-Ray devices once more. With the proposed model, it is aimed to reduce the number of sampling in production and to avoid any errors in the shipped boards.

The defects that may occur due to soldering of basic components such as resistors and capacitors on PCBA are relatively simple and can be easily detected at the control stations in the production line. However, the solder defects of the components such as BGA components cannot be caught in these stations because the solder pads are located under the component, but they can only be examined with X-Ray devices. Due to this fact, the proposed model will be applied for BGA components that cannot be fully detected in the production line and need to be inspected in X-Ray device.

It is aimed to identify all possible BGA components failures and ensure them to be checked in X-Ray device.

X-Ray control devices can either be a manual type which inspections are performed visually by an operator, or automated type that defective components are identified automatically. In the devices that requires manual inspection by a human being, the operator examines all components one by one and reports the relevant component in case of a problem. Although the inspection process is carried out by knowledgeable and experienced operators, a perfect inspection is not possible, minor flaws that the operator does not notice can be transferred to the next station without reporting. With the model to be obtained in this study, a system will be developed to identify the pins and components that are likely to contain failures, and to notify the operator to carefully examine these components. Thus, the problems that may occur in these components will be reported and intervened without being unnoticed.

3.2 Collecting Data and Building Dataset

The necessary data for training proposed machine learning models will be collected from the machines in the production line. It will be implemented on a production line which is numbered as "Line-1913" in Vestel Electronics PCBA Factory being located in Manisa, Turkey. 17MB170R4 model panels produced on this line, known as the motherboard of TVs, will be considered. There are four BGA components on 17MB170R4 model, and these components are Main Integrated Circuit (IC), two Random Access Memory (RAM) ICs and one NAND Flash Memory IC components. Each component on the circuit has a position name. Position names and pin counts of these components are given in Table 1. Each PCBA has these four components, 17MB170R4 model has three multiplications in a panel, hence a panel has 12 components and 3309 pins in total.

Table 1: Components on 17MB170R4 model.

Component Name	Position Name	Number of Pins
RAM IC-1	U110	96
RAM IC-2	U113	96
NAND Flash Memory IC	U128	153
Main IC	U139	758
Total		1103

The data to be used in this project will be obtained by combining measurement values from SPI, Pre AOI and Post AOI control machines and results from FVT stations and X-Ray inspection device. In SPI machine, measurements of different properties such as area, volume and height are provided for the solder paste on all pins on the board. The machine assigns numbers to the pins of each component starting from 1 and increasing by one and records measurement values for each pin in ".xml" file format. In Pre AOI and Post AOI machines, measurements such as offset and angle for each component are obtained on component based and the measurement values for each component are recorded with the position name in ".xml" file format. In this project, it is aimed to create a dataset which is pin based, each record belongs to one pin, in order to encourage the model to learn better and to obtain precise results. For this reason, measurement results of the components obtained from Pre AOI and Post AOI machines will be replicated to the records of the pins of each component. The X-Ray device used in the production line is a manual type device, the operator should check the components visually with the help of X-Ray beams and detect defective pins and components. X-Ray device can name the pins of each component on the board by predefined filters. The operator examines BGA components on 17MB170R4 board, if she does not see any defects, she records the board as "Passed", which means that there is no defect. If the operator observes a defect in the pins in any of these components, she reports the board's serial number as well as the position name and the wrong pin names as "Failed". All pins that the operator does not specify are recorded as passed, and the pins whose names are specified are recorded as failed to the system on pin basis.

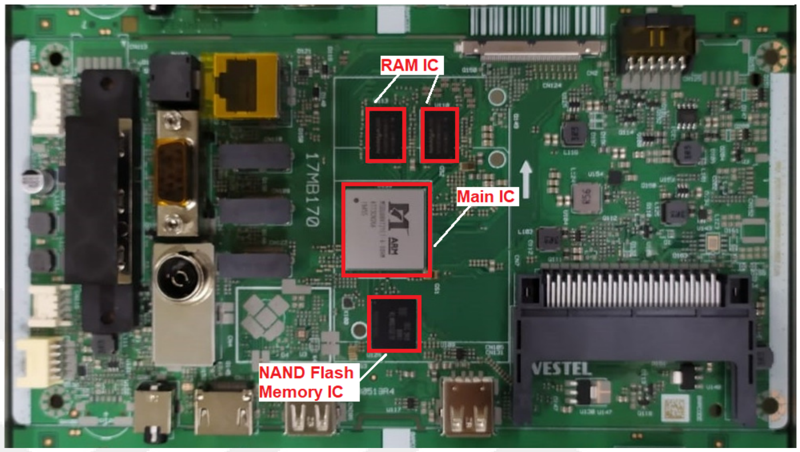

Figure 17: Four BGA type components on 17MB170R4 model.

A software management system that collects data from the machines in production exists in the factory. In this system, information such as the number of products, the model ids of produced products, measurement results are collected from all machines suitable for data acquisition. This data is stored in SQL Server database. Thanks to this system, the current number of products that have been produced can be tracked from the dashboard interfaces, line-stops can be automatically recorded, and relevant staff can be automatically notified by e-mail as long as the problem occurring line-stop exists. In addition, all measurements and tests performed in production can be followed and the efficiency of production can be observed. In this system, the data is collected also from SPI, Pre AOI and Post AOI machines, but these data only includes the information that the test results of the panels as a whole are passed or failed, they do not include the measurement values for each pin and component. Within the scope of this project, it is planned to establish a separate database in order to record measurement values and to combine them with X-Ray test results. The main software runs on the server and it collects data from the machines and saves it in the database, and this software has been provided to transfer the necessary information to the new established database for the project purpose.

The main purpose of this database, which was established to create the necessary dataset, is to record SPI, Pre AOI, Post AOI measurement values on pin basis, label them with the results from FVT tests and X-Ray tests and obtain pin-based data. For this purpose, a SQL Server based database containing relational tables has been established. SPI, Pre AOI and Post AOI measurement results of the panels are available in MLRecord table, the detailed measurement values of these results for each pin and component are recorded in MLResult table. There is a unique ID number for each component and pins measured, this information is available in Pin and Component table. Measurement values are transferred to MLResult table using ComponentID and PinID numbers in these tables. X-Ray results and FVT results of the boards are acquired from different databases and recorded in XRayResult and FVTResult table. If a panel completes all its tests, measurement values and test results are combined using serial numbers and master numbers in SerialNumber table and transferred to the targeted record table which is PinResult table.

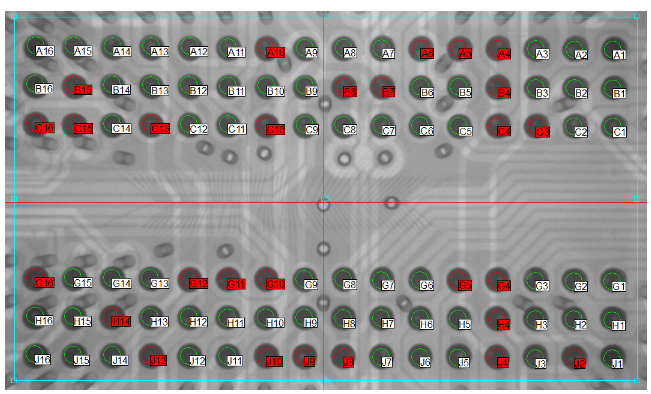

Figure 18: The image of the pins of NAND Flash IC component on X-Ray device.

All required measurement values and test results are recorded in the tables, but a software is needed to combine this information in the tables and transfer them to PinResult table. For this purpose, server software has been developed in C#

language which runs on one of the server computers in the factory. The purpose of the software is to periodically examine all the tables in the database and to transfer the measurement results of the boards with all completed tests results to PinResult table. When a board starts its production cycle and SPI is performed in order to check solder quality after the solder paste is applied in SP machine, its first record is transferred to MLRecord and MLResult tables. Since this measurement is made on a panel specific, its results are moved to the database with its MasterNo number which is specific to each panel. SerialNo numbers which are specific to each PCBA are obtained from SAP system according to MasterNo number of the panel and recorded in SerialNumber table. Afterwards, the board goes through other processes, and measurement values of the components which are obtained in Pre-AOI and Post-AOI machines are recorded in MLRecord and MLResult table.

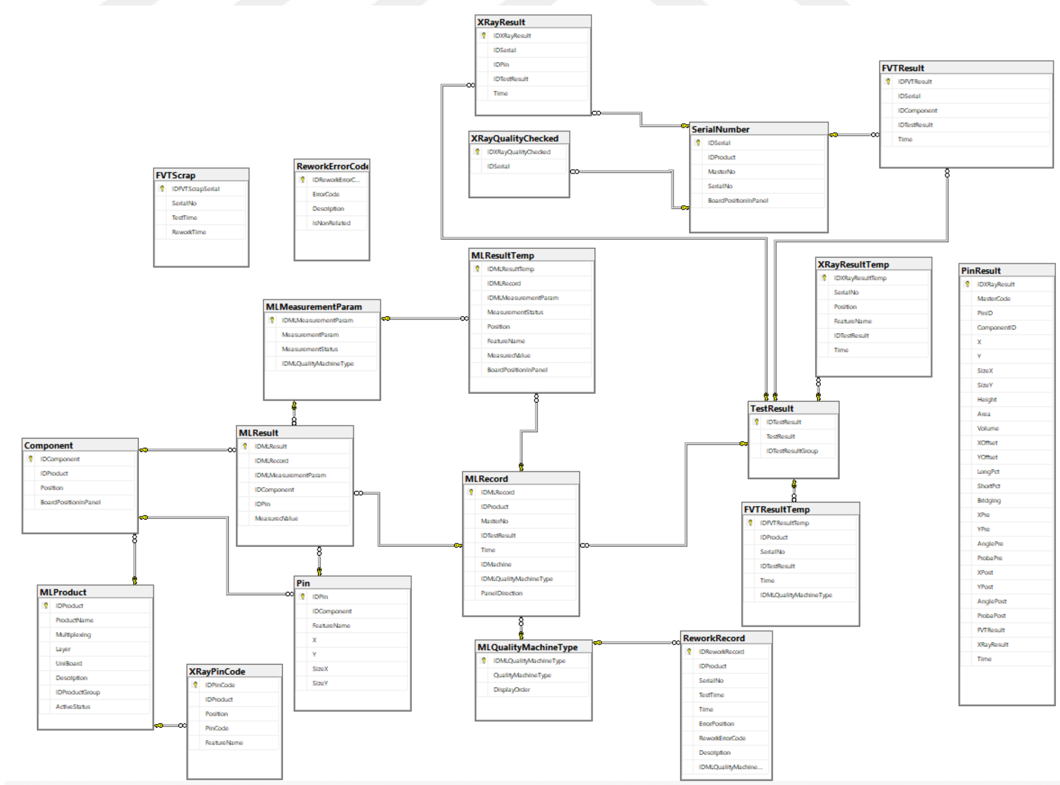

Figure 19: Diagram that shows relations between tables in the database.

The boards produced in Line-1913 are randomly sampled and sent to X-Ray inspection station. If a board is not inspected in X-Ray device, it is expected that X-Ray test result will be performed for three days after the first record is entered in the database. If the board is not sampled during this period and does not take any X-Ray testing, all records related with its MasterNo and SerialNo are deleted from all the tables by the server software. If the board is inspected in X-Ray device, these records from a different database are first transferred to XRayResultTemp and then to XRayResult tables. X-Ray test results are reported with the pin names as the machine names the pins, and SPI measurement results are reported with the numbering system of the SPI machine. In XRayPinCode table, there are records that match the naming of the X-Ray machine and the numbering of the SPI machine, so X-Ray test results are translated according to the numbering in SPI machine and saved with PinID numbers. The board is transferred to FVT station after X-Ray testing. In this station, a software is installed on the board and functional tests are performed. If a problem occurs during FVT, the board is sent to Rework Room for rework process, then FVT process is applied again after the rework process. After successfully completing FVT operations for the board, it passes to the other stations and the production process is completed. FVT results of the board are transferred to FVTResultTemp table. If the server software detects a new SerialNo number in this table, it queries the rework records of this PCBA from ReworkRecord table. If the PCBA does not have any rework records belonging to BGA components, FVT results of all BGA components are transferred to FVTResult table with passed label. If a component has a rework record, FVT result of this component is recorded in the same table as failed label. Server software periodically checks all tables. If SPI, Pre AOI, Post AOI, X-Ray and FVT test results of a panel exist, the measurement results of this panel are combined by the software and transferred to PinResult table. After merging process, all relevant records belonging to this panel are deleted from the tables.

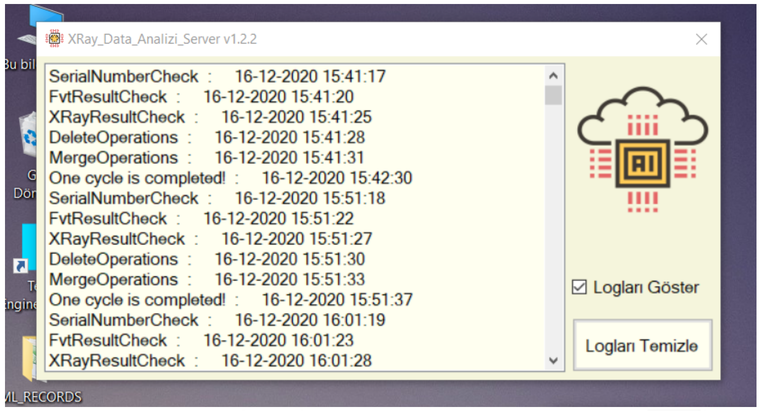

Figure 20: Server software that manages X-Ray database operations.

PinResult table contains measurement values and test results of BGA pins. In addition, there is MasterNo, PinID, ComponentID information that includes the panel, pin and component addresses of these pins in the table. While the server software transfers the results to this table, it also saves these records in the ".csv" file format on the computer it is working on. Each csv file consists of 100000 lines and the software converts the data into multiple csv files.

MasterCode	PinID	ComponentID	X	Y	SizeX	SizeY	Height	Area	Volume	XOffset	YOffset	LongPct	ShortPct	Bridging	XPre	YPre	AnglePre	ProbaPre	XPost	YPost	AnglePost	ProbaPost	FVTResult	XRayResult	Time
2367856800001547	1	1	287.01	99.60	0.40	0.40	0.12	0.11	0.01	0.01	-0.04	9.40	3.12	0.00	-0.03	-0.02	-0.08	0.00	-0.07	-0.06	0.00	0.04	1	1	7.10.2020 13:47:30
2367856800001547	2	1	287.01	100.40	0.40	0.40	0.13	0.11	0.02	0.01	-0.02	6.06	3.61	0.00	-0.03	-0.02	-0.08	0.00	-0.07	-0.06	0.00	0.04	1	1	7.10.2020 13:47:30
2367856800001547	3	1	287.01	101.20	0.40	0.40	0.14	0.11	0.02	0.01	-0.03	6.33	3.19	0.00	-0.03	-0.02	-0.08	0.00	-0.07	-0.06	0.00	0.04	1	1	7.10.2020 13:47:30
2367856800001547	4	1	287.01	104.40	0.40	0.40	0.14	0.11	0.02	0.01	-0.04	9.41	1.70	0.00	-0.03	-0.02	-0.08	0.00	-0.07	-0.06	0.00	0.04	1	1	7.10.2020 13:47:30
2367856800001547	5	1	287.01	105.20	0.40	0.40	0.13	0.11	0.01	0.01	-0.03	6.86	3.37	0.00	-0.03	-0.02	-0.08	0.00	-0.07	-0.06	0.00	0.04	1	1	7.10.2020 13:47:30

Figure 21: Sample records from PinResult Table.

There are 25 columns or features in the dataset obtained. From these columns, 20 columns contain the measurement values taken from SPI, Pre-AOI and Post-AOI machines. FVTResult and XRayResult columns contain passed / failed result information of the measurement values. 3086861 lines of data were collected within the scope of the project. Of these data, 2831 rows of data X-Ray result failed, the remaining data are the data of the pins that passed X-Ray result.

3.3 Preprocessing and Sampling Data

As explained in the previous section, data was collected from the machines in the production for the training proposed machine learning model and 31 csv files containing 100000 lines were obtained. When these files were analyzed by combining all of them, it was deemed appropriate to remove some columns. Since the collected data was obtained in about two months, no defective components were observed in FVT stations during this period. Because FVT results of all records are labeled as passed and it does not contain meaningful information, this column has been omitted from the dataset. Although there are models developed using timestamp information in the literature [24], since a model with time series will not be used in this project, the column containing record time has also been removed from the dataset. Due to some problems that may occur during data collection, dataset may contain duplicate rows which have the same values for all features. Therefore, the cleaning process has been made so that only the last record will remain from the rows containing the same values. In X-Ray test results in the database, pins with the passed label are recorded as 1, and the pins with the failed label are recorded as 3. In order to create reasonable labels in the dataset for model training, failed labeled records containing a minority of samples were changed to 1 in the sense of defective, records labeled passed with a majority of samples were changed to 0 in the meaning of not defective. The resulting features and variable types are listed in Table 2.

In the models to be discussed, XRayResult feature will be used as output and other features will be used as input. When the data is classified according to XRayResult feature, since there is a enormous difference between the numbers of the defective and non-defective samples, the train-test splitting process was made in a way that

Table 2: Features and their data types in the dataset.

Feature Name	Data Type	Feature Name	Data Type
MasterCode	string	LongPct	float64
PinID	string	ShortPct	float64
ComponentID	string	Bridging	float64
X	float64	XPre	float64
Y	float64	YPre	float64
SizeX	float64	AnglePre	float64
SizeY	float64	ProbaPre	float64
Height	float64	XPost	float64
Area	float64	YPost	float64
Volume	float64	AnglePost	float64
XOffset	float64	ProbaPost	float64
YOffset	float64	XRayResult	float64

the ratio of the samples from minority class to the samples from majority class is equal for both training and test datasets. After 80% train-test separation, datasets with class ratio of 0.00094 were obtained, with 2015 defective, 2141216 non-defective pins in the training data, 504 defective pin samples, and 535304 non-defective pin samples in the test data.

After the data was split as train and test data, cleaning operations were applied as the second step in the training dataset. All of the preprocessing steps that will be explained hereafter in this section are applied only to training data. Since operations such as standardization and dimensionality reduction are the operations that change the characteristics of the data, the same transformations applied in the training dataset are applied directly to the test data in order to obtain logical results, and no statistical analysis study has been conducted on the test data to prevent the model from cheating.

In the training dataset, the samples having missing values were cleaned in the preprocessing stages at the beginning. Since the rows with too many missing values do not contain meaningful information, samples containing more than four empty values

	X	Y	SizeX	SizeY	Height	Area	Volume	XOffset	YOffset	LongPct	ShortPct	Bridging	XPre	YPre	AnglePre	ProbaPre	XPost	YPost	AnglePost	ProbaPost	XRayResult
count	214...	214...	214...	214...	190...	190...	190...	190...	190...	190...	190...	190...	189...	189...	189...	189...	189...	189...	189...	189...	214...
mean	161.24	101.53	0.35	0.35	0.13	0.08	0.01	0.01	-0.01	4.03	4.43	0.00	0.00	-0.01	0.15	0.08	0.03	0.03	0.01	0.04	0.00
std	90.39	6.17	0.04	0.04	0.02	0.02	0.00	0.02	0.02	3.17	3.87	0.00	0.04	0.03	0.15	0.10	0.04	0.05	0.05	0.06	0.03
min	23.39	86.10	0.27	0.27	0.00	0.00	0.00	-0.30	-0.33	0.00	0.00	0.00	-0.18	-0.14	-0.21	0.00	-0.16	-0.16	-0.18	-0.16	0.00
25%	58.48	97.50	0.35	0.35	0.12	0.07	0.01	-0.00	-0.02	1.59	1.53	0.00	-0.02	-0.03	0.00	0.01	0.01	0.01	0.00	0.01	0.00
50%	160.99	101.50	0.35	0.35	0.13	0.08	0.01	0.01	-0.01	3.37	3.39	0.00	0.00	-0.01	0.12	0.04	0.04	0.04	0.03	0.04	0.00
75%	262.83	105.92	0.35	0.35	0.14	0.09	0.01	0.02	0.00	5.77	6.36	0.00	0.02	0.01	0.31	0.11	0.05	0.06	0.05	0.05	0.00
max	299.01	112.42	0.40	0.40	0.49	0.50	0.21	0.32	0.29	82.16	79.23	0.51	0.23	0.14	0.91	0.65	0.16	0.19	0.16	0.55	1.00

Figure 22: Statistical properties of the features in the training data.

have been removed from the training dataset. After this process, the imputation process has been applied for the values that are still missing. With this process, the most common values for each feature, which are statistically called mode, were used in order to fill the missing values. Afterwards, the standardization process was applied for each column, and the transformation process was performed such that the mean of feature values is 0 and the standard deviation is 1.

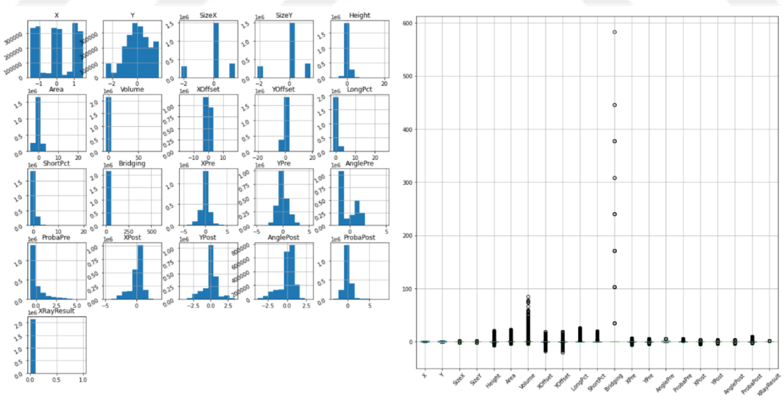

Figure 23: Histogram and boxplot graphs of the training dataset after standardization.

After the standardization process, it is observed that some features have values far from the distribution. Therefore, additional preprocessing step was tried on the dataset to obtain more uniform distribution. For each feature, the values are sorted in ascending and first and third quartiles are obtained. With the difference between these two values, the interquartile range (IQR), which is the statistical feature of the data was calculated. Samples with values 1.5 IQR less than the first quartile value or 1.5 IQR more than the third quartile value were removed from the dataset. After this process, 1258603 non-defective and 1253 defective pin samples were obtained.

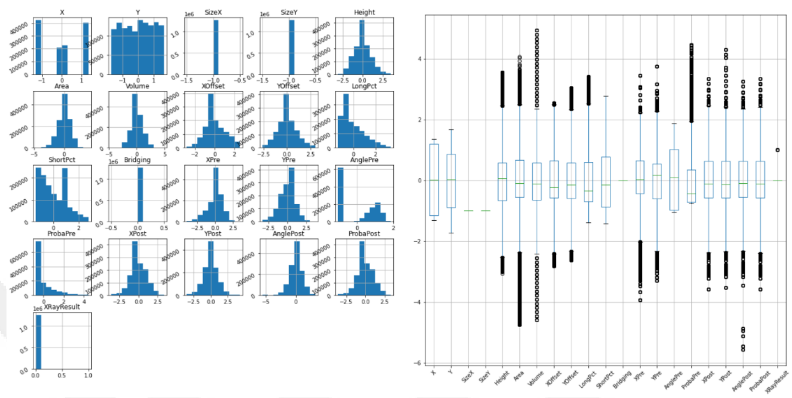

Figure 24: Histogram and boxplot graphs of the training dataset obtained after removing excessive values.

With the help of this process, more evenly distributed dataset was obtained. However, although the validation performance was very high when training a neural networks model with this dataset, very low values were obtained in test performance. The reason for this problem is that the information belonging to a significant part of the data is lost by discarding very high values. The model was not successful in samples with these values in the test data. Therefore, model training processes were applied with the dataset obtained without applying this process.

The dataset contains samples from defective and not-defective classes with highly imbalanced ratio, which may cause the model to overfit in the training process of the model. There are different resampling methods for imbalanced datasets in order to acquire more balanced classes. The process of reducing the data of majority class by sampling is called under-sampling, and the process of reproducing the data of minority class with different methods is called over-sampling. Both methods have advantages and disadvantages. The biggest disadvantage of the over-sampling method is that the model may overfit in the training process with the balanced dataset and it performs

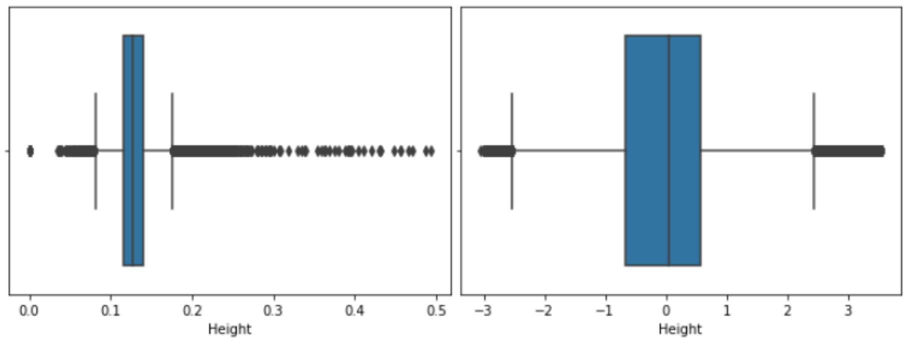

Figure 25: Boxplot graphics of the Height feature before and after throwing excess values.

poorly in imbalanced test data. The biggest disadvantage of the under-sampling method is that the model does not learn the data characteristics as a result of the loss of important information in the data, and therefore it shows poor performance in the test data.

In this project, Random Under Sampling, Tomek-Links and Edited Nearest Neighbor under-sampling methods are used to balance the data. In Tomek-Links method, majority samples that are close to the minority samples are calculated, and the pairs of minority and majority samples which are called Tomek-Links are identified in the dataset. If the aim is to reduce the majority class, the majority class samples are deleted from these pairs, if the goal is to clear the data, the samples belonging to both classes are removed from the pairs. In Edited Nearest Neighbor (ENN) method, classes of neighboring samples are checked for each sample. If a sample belongs to the majority class and the closest neighbors belong to the minority class, this sample from the majority class is removed. If a sample belongs to the minority class and the closest neighbors belonging to the majority class, the neighboring samples are removed. With these methods, it is aimed to make the model perform better by creating a more reliable and precise dataset by reducing the examples of the majority class very close to the minority class [29].

In order to observe the effect of under-sampling methods in the resampled dataset,

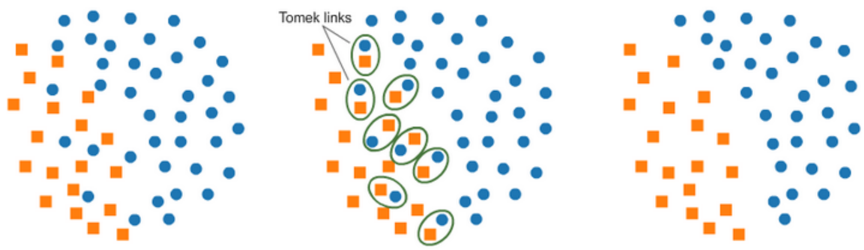

Figure 26: The cleaning process of the data using Tomek-Links method. [15]

it is aimed to reduce the dimensions of the dataset to two in order to visualize it. Dimensionality reduction process is the process applied in machine learning projects to reduce the data having too many dimensions into less dimensions without losing too much information. One of the methods with carrying this purpose is the method called Principal Component Analysis (PCA). This method will be discussed in detail later in this section and will be used as a data preprocessing method. Training data was reduced from 20 dimensions to 2 dimensions by PCA method for visualization purposes.

Figure 27: Scatter graph obtained by reducing the training dataset in two dimensions.

Under-sampling methods were applied to the majority class in order to make the dataset more balanced and increase the model performance. First of all, randomly selected samples from the majority class were removed from the dataset, which is called Random Under Sampling method. Then, Tomek-Links was found in the dataset and samples belonging to the majority class were deleted from these pairs. ENN under-sampling method was applied to clean samples from the majority class, which are very close to the minority class samples. In this step, while determining the number of neighbors, trial and error method was applied and 120 neighbors were selected as the most suitable number of neighbors. Afterwards, Tomek-Links method was applied once again to clean up the remaining majority classes. Tomek-Links and ENN under sampling methods are highly computational expensive methods, each sample on the dataset must be compared with other samples and the distance calculated. Since the dataset used is relatively large and contains too many dimensions, it is not possible to apply these methods directly. For this reason, Random Under Sampling method was applied so the ratio of minority to majority class was increased to 0.002 in order to reduce the samples in the dataset. After all under-sampling methods applied, 2141216 majority samples were reduced to 792984 samples.

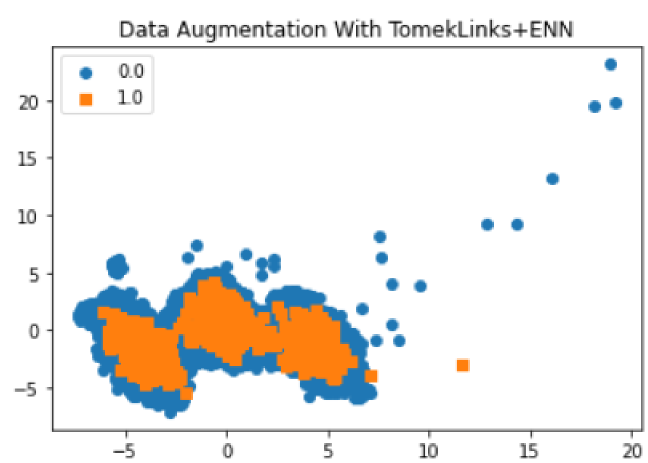

Figure 28: Scatter graph of the training dataset after applying resampling methods.

Having too many dimensions in the dataset used for training in machine learning models may cause some problems. When the dataset has too many dimensions, the space that samples composed of becomes sparse and the number of samples required for the model to learn the data increases exponentially, this problem is also called Curse of Dimensionality in statistics [30]. There are methods to eliminate some features or obtain new enhanced features in order to reduce the number of input features, i.e. the size of the data. The most common technique among these methods is PCA method. The aim of the PCA method is to obtain fewer dimensions which are independent of each other with high variance. For this purpose, normalization process is performed in all dimensions first. Then the covariance matrix is calculated by Formula 1 and the eigen vectors and eigen values of this matrix are calculated [31]. Eigen vectors are ordered in descending order according to their eigen values, and the projection of the data on these vectors is calculated by taking the number of k largest eigen vectors and reduced to the data into k dimensions.

$$Cov(X,Y) = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y}) \tag{1}$$

PCA method was applied to the training data which the missing values and duplicate rows were omitted first and then the standardization process was applied. By taking the first 13 vectors from eigen vectors, 94.74% of the data variance was covered. Three different datasets were obtained as a result of data preprocessing methods. In the first dataset, only imputation and standardization processes were applied. The second dataset was obtained by applying under sampling methods to the samples in the first dataset. The third dataset was constructed by reducing the dimensionality in the first dataset from 20 to 13 using the PCA method. When these datasets are used in the training of a neural network model, it is observed that the best test performance was obtained in the first dataset. Therefore, the models to be discussed in the next sections were trained on the first dataset.

MasterCode	PinID	ComponentID	pca_0	pca_1	pca_2	pca_3	pca_4	pca_5	pca_6	pca_7	pca_8	pca_9	pca_10	pca_11	pca_12	XRayResult
2366459000000016	2461.00	11.00	0.25	0.89	-0.05	-0.70	-0.44	-0.69	-0.11	0.08	-0.05	1.86	-0.65	-0.21	-0.50	0.00
2359695200002466	208.00	3.00	-0.10	1.13	-1.83	-0.54	-0.71	0.22	-0.66	1.84	0.23	-0.29	0.22	0.19	-0.80	0.00
2363880300002308	528.00	4.00	0.54	0.50	0.10	-0.30	-1.03	-0.52	0.02	0.17	-0.05	-0.34	1.20	0.76	0.28	0.00
2363859900002272	2451.00	11.00	-5.12	-1.74	-0.68	0.32	-0.97	-0.68	0.57	0.25	0.34	-0.21	-0.44	-0.91	1.76	0.00
2362583200001366	2740.00	12.00	-1.75	-0.47	0.17	-0.92	0.14	-0.68	0.56	-0.34	0.41	0.44	-0.47	1.32	0.23	0.00

Figure 29: Sample rows from the third dataset which is obtained by PCA method.

3.4 Machine Learning Models

Due to predicted test results in X-Ray inspection device by analyzing measurement values from SPI, Pre-AOI and Post-AOI control machines, machine learning models will be examined in this section. As explained in the previous section, all models will only be trained on the first dataset which is created by imputation and standardization preprocessing steps.

The most important characteristic of the dataset being constructed with measurement values collected from the machines is imbalanced class distribution. In imbalanced datasets, unlike other datasets, additional attention is required at every stage of training and testing to ensure that the model is not overfitted or not trained with insufficient data. One of the most important factors when training a model in imbalanced datasets is the key metric. If inappropriate key metric is selected to evaluate the model, the model may seem as successful, but its actual performance may be too poor. The formula for accuracy term which is commonly used in model trainings is given in Formula 2. For instance, if the accuracy is selected as the performance criterion for the first dataset and the model outputs non-defective class to all samples without any estimation, the accuracy is calculated as 99.9%, which shows that the accuracy is not a correct metric for this case.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{2}$$

There are many metrics to measure the actual performance of the model in imbalanced datasets, the most basic metrics are precision and recall metrics. The precision is a measure of how accurate the samples are found to be positive; the recall is a measure of how well the model can detect samples that are actually positive. It is aimed to identify all defective examples by the model in this study. If the model outputs defective label for non-defective samples, these samples are examined by the operator and they are reported as non-defective by the operator and the boards are sent for the next operations. However, if the model classifies defective samples as non-defective, these boards are transferred to the next stations without being detected. Thus, the desired behavior of model is having performance with high recall value. It is not sufficient if just recall value is high, because if the model outputs all samples as defective without any learning, 100% recall is obtained, but it cannot be said to be a successful learning. For this purpose, key metrics which recall is prior metric, but precision is also important should be preferred. The metrics that meet these conditions are F2-score and Precision-Recall Area Under Curve (PR-AUC) metrics. Precision, recall and F2-score can be calculated with Formula 3, 4 and 5 [32]. While the models are being trained, PR-AUC metric will be considered as the main metric.

$$Precision = \frac{TP}{TP + FP} \tag{3}$$

$$Recall = \frac{TP}{TP + FN} \tag{4}$$

$$F2 - Score = \frac{5.Precision.Recall}{4.Precision + Recall} \tag{5}$$

Four type machine learning models will be examined, these models are Random Forest, Multi-Layered Perceptron (MLP), Decision Tree and Neural Networks (NN). The main purpose of this study is to examine the performance of neural network models when predicting X-Ray results and to develop an appropriate model. For this reason, model training will be applied with few parameters in Random Forest, MLP and Decision Tree models due to benchmark with developed models. For NN model, different model parameters and the effect of these parameters will be examined. While evaluating model performances, the training data will be divided into five folds with K = 5 by K-Fold Cross Validation method, and the training on four parts and the validation on one part will be applied. Afterwards, PR AUC metrics on the test data will be compared with the model parameters that performs best learning.

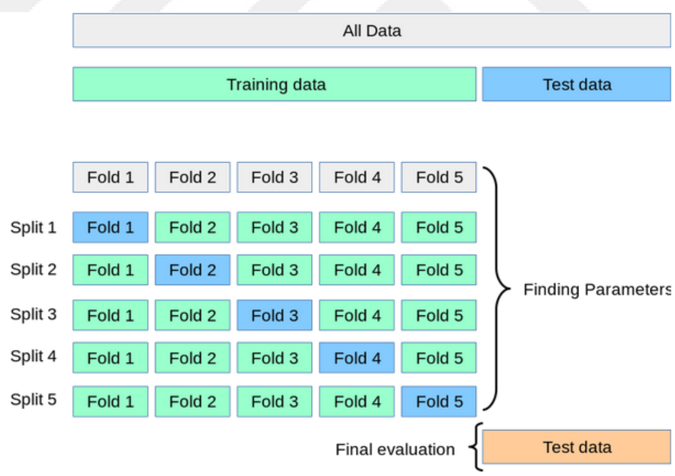

Figure 30: The evaluation of model performance by K-Fold Cross Validation. [16]

3.4.1 An Introduction to Neural Networks

Neural Network (NN) model is a mathematical model that imitates the working principle of the neurons in the human brain. The first neuron, or namely perceptron, developed by Frank Rosenblatt (1958) was used for a pattern recognition task [33]. It is known that the signal received from the dendrites of a neuron are transmitted to the next neuron if it exceeds a certain threshold value. Signal transmission between neurons occurs through spaces called synapses, and these spaces allow different signals to be transmitted to the cell by weighting them. Neurons work according to the all or nothing law, i.e. if the incoming signals exceed the threshold value, they transmit a signal to the next neuron, otherwise they do not allow the signal to pass.

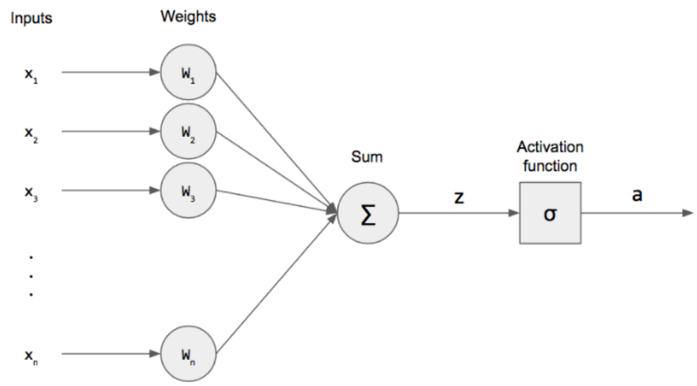

Figure 31: A diagram that shows mathematical model of a perceptron. [17]

According to this working principle, it is possible to create mathematical model of a perceptron as in Figure 31. In this model, input values indicated by x_i are multiplied by the weight coefficients indicated by w_i, and z variable is obtained by summing them with the parameter called bias. This part of the model is similar to the transmission of signals from the synapse to the neuron. The summation is passed through the activation function, so if z value is above a certain threshold value, an output arises. A signal obtained at the output of the neuron can be calculated by the following formulas [17]:

$$z = \sum_{i=1}^{n} W_i x_i + b = \sum_{i=0}^{n} W_i x_i = W^T x \tag{6}$$

$$\sigma(q) = \begin{cases} 1 & q \geq 0 \\ 0 & q < 0 \end{cases} \tag{7}$$

$$a = \sigma(W^T x) \tag{8}$$

Sigma is a step function as the activation function in these formulas. Although perceptron model is successful in classifying data that can be separated linearly, it cannot classify more complex data where the decision boundary is not linear. Nonlinear activation functions are preferred to classify such datasets. Besides, multi-layered neural network models with multiple layers are preferred in order to achieve better results with more adjustable parameters. Frequently used nonlinear activation functions are Sigmoid, Tanh, Rectified Linear Unit (ReLU) and Leaky ReLU functions. The plots of these functions can be seen in Figure 32.

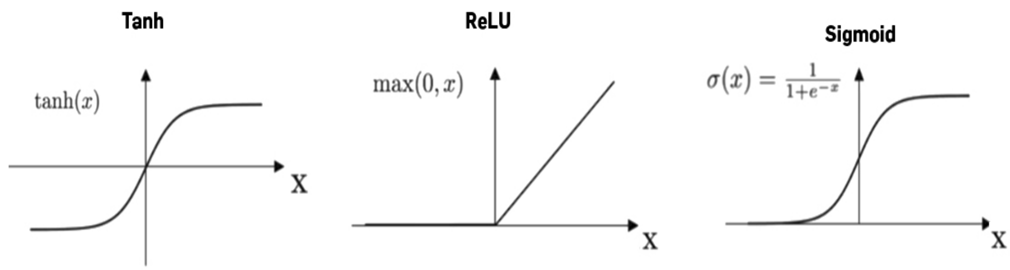

Figure 32: The plots of frequently used activation functions. [18]

In neural networks in which multiple layers with many neurons exist, the first layer is the input layer, the last layer is the output layer, and the intermediate layers are called the hidden layers. In these models, classification of a sample of the model is performed with the following steps. Each feature of the data represents a neuron in the input layer, the values of these features are multiplied by the weight coefficients and summed with the bias. The resulting total is passed through the activation function and the output is obtained. After performing these operations for all neurons in the first hidden layer, operations are performed for the other layers in such a way that the input layer is the output layer of the previous neuron and the output is acquired in the output layer. This process is called forward propagation. The difference between the expected value of the model in response to an input and the actual value of the model output is called an error. Different cost functions are available to calculate this error.

A model is desired to output with the least error, so it is aimed to minimize the cost function. For this reason, the gradients of the cost function are calculated from the last layer to the first layer and the weight coefficients are updated by multiplying the gradients with a coefficient. This process is called back propagation. The structure of the multi layered neural network model is shown in Figure 34.

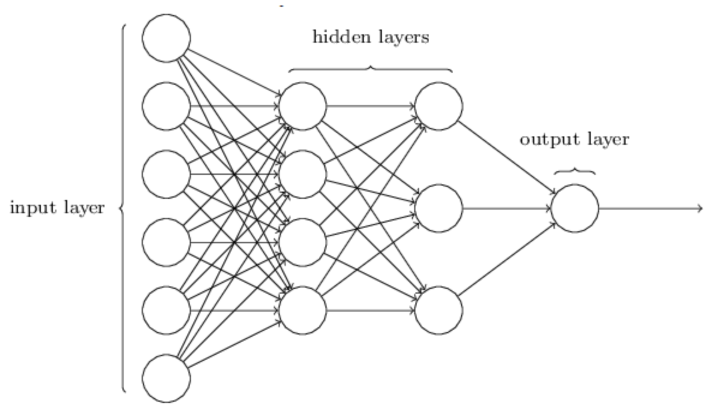

Figure 33: A scheme that explains the architecture of multi layered neural networks. [19]

Supervised machine learning models are generally used for classification and regression purposes. While the aim of regression models is to estimate the value of a continuous variable, the classification models are intended to estimate the value of the variable with discrete values or to determine the class of a sample. Different cost functions are used according to the main purpose of the model. While Mean Squared Error (MSE) is generally preferred in regression models, Log Loss functions are preferred in classification models. In this project, Binary Cross Entropy Loss (BCE Loss) function has been chosen as the cost function. This function is calculated by the formula below [34]:

$$H_p(q) = -\frac{1}{N}\sum_{i=1}^{N} y_i . \log(p(y_i)) + (1 - y_i) . \log(1 - p(y_i)) \tag{9}$$

In this formula, y_i is the label or expected value and $p(y_i)$ is the predicted probability as actual value for the output of the model. After the cost function is chosen, an optimizer that updates the model parameters is determined due to minimize the cost function during training process. The commonly preferred optimizer is Stochastic Gradient Descent (SGD). In this method, the cost function is calculated for each sample in the data and the necessary gradients are calculated to minimize the cost, then the model parameters are updated by multiplying with learning rate. Since this method updates the model parameters for only one sample at a time, it causes a noisy training process, which means it can be stuck at the local minimum point in the error function. As a solution, instead of training with a single sample one by one, samples are given to the model in groups called batches and the parameters are updated in this way. This method is known as Mini-batch Gradient Descent. One disadvantage of this method is that the same learning rate is used for all model parameters. Adaptive Moment Estimation (Adam) method allows using a separate learning rate for each parameter [20]. It also performs an update process by using the gradient values of the parameters in the previous iterations, thus the error value reaches to its global minimum without being stuck in the flattened areas in the error function. Although both SGD and Adam optimizer are examined in this project, Adam optimizer is mostly preferred because of giving better performance.

3.4.2 Pin Based Models

With this type of model, it is aimed to accurately predict X-Ray analysis results with pin-based measurement values. For this purpose, MasterCode, PinID and ComponentID features that include unique id numbers and do not contain any numerical meaning are removed from the dataset. In this model, 20 measurement features which are gathered from SPI, Pre-AOI and Post-AOI machines are chosen as input features and XRayResult feature which includes the results of the X-Ray test is chosen as an

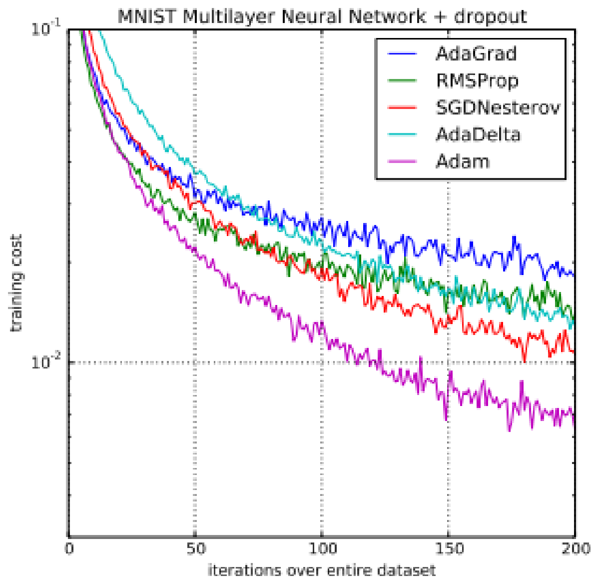

Figure 34: The cost curves of different optimizers on MNIST dataset. [20]

output. It is aimed to use Random Forest, MLP, Decision Tree and NN algorithms for this model.

	X	Y	SizeX	SizeY	Height	Area	Volume	XOffset	YOffset	LongPct	ShortPct	Bridging	XPre	YPre	AnglePre	ProbaPre	XPost	YPost	AnglePost	ProbaPost	XRayResult
0	-1.48	0.32	-2.14	-2.14	-0.15	-2.09	-2.03	0.82	1.88	1.55	0.98	-0.00	0.39	0.48	-0.88	-0.72	1.25	0.94	1.27	0.76	0.00
1	0.95	0.32	-2.14	-2.14	1.07	-2.12	-1.83	0.72	0.25	-1.00	0.81	-0.00	-2.20	0.53	-0.88	2.68	-0.44	-0.45	-0.04	-0.34	0.00
2	1.25	-1.29	0.07	0.07	0.89	-0.39	0.14	0.24	-1.51	1.55	-0.28	-0.00	0.36	0.73	0.50	1.23	0.74	2.78	-2.50	-0.51	0.00
3	-1.48	0.08	-2.14	-2.14	1.78	-2.36	-1.83	0.88	1.23	0.32	1.07	-0.00	-1.62	3.78	-0.88	-0.56	-0.22	-0.28	0.13	-0.20	0.00
4	-1.17	1.13	0.07	0.07	-0.50	0.14	0.18	-1.32	1.31	0.07	-0.22	-0.00	1.84	-0.37	1.33	0.48	0.78	0.55	0.90	0.45	0.00

Figure 35: Sample records from the dataset for pin based models.

Random Forest algorithm is an ensemble algorithm and it contains more than one tree models. Each tree gives an output by evaluating the input features and all models are provided to give a single output by evaluating the outputs of the trees. In this study, Random Forest models are trained by using different values for the number of estimators, number of maximum depth and splitting criteria, the other model parameters are left in their default values. Different values are chosen for these three parameters and the performances of all parameter groups were obtained with Grid Search algorithm. The performed parameter values are shown in Table 3. At the end of this experiment, it is observed that the separation criterion "Entropy" performed better than "Gini" criterion. Increasing the number of estimators also increased the performance of the model. Although raising the maximum depth also increases the model success, the model performance did not increase in the models with a depth of more than 50. The parameters which performed the best performance are shown in Table 3. With these parameters, 0.9184 PR AUC for average was achieved during training, and 0.9230 PR AUC was obtained in the test data.

Table 3: Grid Search parameters for pin based models.

Random Forest			
	Estimator Size	**Maximum Depth**	**Splitting Criterion**
Parameter Options	10, 30, 50	10, 50, 100	Gini, Entropy
Best Parameters	50	50	Entropy
MLP			
	Hidden Layers	**Activation Function**	**Learning Rate Type**
Parameter Options	[10,20,5], [20,30,10], [30,60,20]	Tanh, ReLU	Constant, Adaptive
Best Parameters	[20,30,10]	ReLU	Adaptive
Decision Tree			
	Maximum Depth	**Splitter**	**Splitting Criterion**
Parameter Options	10, 20, 40	Best, Random	Gini, Entropy
Best Parameters	40	Best	Entropy

MLP algorithm is a machine learning algorithm that contains multiple hidden layers which contain multiple neurons and the neurons in these layers are fully connected to each other. Each layer output is passed through a nonlinear function called the activation function, thus complex models are obtained, and complex data characteristics can be understood. In order to determine the best parameters and achieve the highest success in MLP model, different values for number of neurons, type of learning rate and activation functions are selected and K = 5 Cross Validation is applied for all parameter groups. The performed values for these parameters are given in Table 3. Increasing the number of neurons in the hidden layers positively affected the model performance. ReLU and Tanh functions showed similar results for PR AUC score. The parameters with the best score are shown in Table 3. With these parameters, PR AUC value of 0.9094 was obtained in the test data.

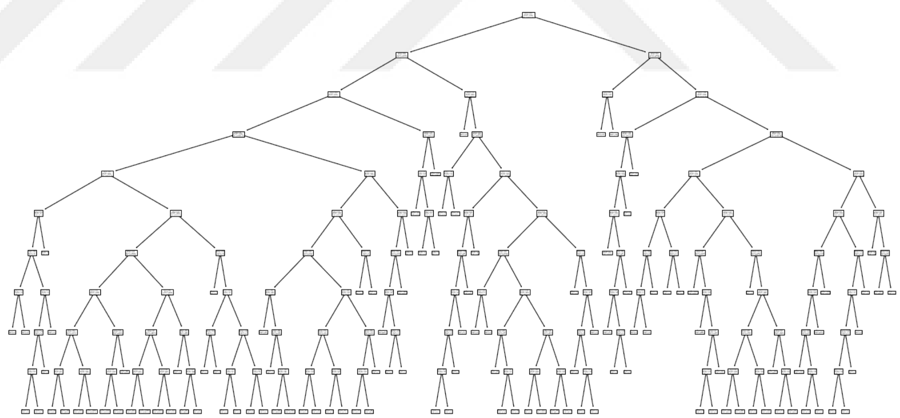

Figure 36: Visual scheme of Pin Based-v2 Decision Tree Model.

Decision Tree algorithm is a model with a tree structure consisting of multiple branches and leaves. It is an algorithm that determines rules in a way that minimizes the entropy of the data in each cluster of each branch and separates all samples in the data. Its biggest advantage over other models is its observability. After the model training, the tree structure can be expressed visually, and the rules derived by the model can be observed. The disadvantage of this model is that if there are too many branches and leaves, the model may become complicated and the test performance may decrease. Another parameter which affects the model performance during training is the approach used during splitting called splitter. If this parameter is chosen as "best", the feature which provides the best splitting is chosen. If "random" is selected, splitting is done with a feature chosen randomly from the best features. In this project, the maximum depth, splitter and splitting criterion parameters for the pin-based model are used, other parameters are left as default. Parameter values used in training are given in Table 3. As a result, it is seen that "Entropy" criterion performed better than "Gini" criterion. Using the "best" approach in the splitting approach gave better results. Although different values of the maximum depth number give close results, the best result is obtained with max depth as 10. The best parameter groups are explained in Table 3, the model with these parameters achieved 0.8899 PR AUC in the test data.

Table 4: Test performances of pin based models.

Model	PR AUC	F2-Score	Recall
Random Forest	0.9230	0.8906	0.8690
MLP	0.9094	0.8740	0.8512
Decision Tree	0.8899	0.8840	0.8651

Neural Networks model is a model that includes many different structures such as Recurrent Neural Network (RNN), Convolutional Neural Network (CNN), including MLP model. In this study, a feed forward neural network model will be developed similar to MLP, and different experiments will be performed to increase the performance of the model. In order to compare the performance of previous models, the training process in Random Forest, MLP and Decision Tree algorithms has been carried out by using built-in functions and default parameters from "sklearn" library. In NN model, "pytorch" library, which allows the development of custom modules for more manipulation and adjustment to the algorithm, was preferred.

NN model is a feed forward neural network model that generally consists of more than one layer, contains many neurons in each layer, and allows neurons to output with certain threshold values by the activation function at each layer output. In the training process, samples in the dataset are given as input to the model and an output is provided. Then, with a predefined loss function, the error between the expected output of the model and the actual output is calculated. In order to minimize this error, the necessary model parameters are calculated by going from the output layer to the input layer and the model parameters are updated. If there are too many neurons or layers in the model, that is, when the complexity of the model increases, the validation performance of the model increases during training, but the test performance decreases at the end of the training, this is known as overfitting. There are different regularization methods to prevent this problem, one of these methods is dropout method. In dropout method, randomly selected neurons at some ratio determined for each layer are inactivated during the training process and prevented from giving output. In this way, training process takes place with a smaller number of neurons in each iteration and overfitting of the model is prevented. During the validation and test steps, the dropout is removed, and all neurons are allowed to output. Another factor affecting the model performance is the initial values of the weights of the neurons. If the value range of the initial values of the weight coefficients is chosen too narrow, the model weights start training with very close values and low performance is observed. If the initial range of the weights is chosen too wide, a problem called vanishing gradient can be seen in the neurons that receive randomly large values and causes an unsuccessful training due to the increase in the model output. Weight coefficients by default in Pytorch library,

$$[\frac{1}{-\sqrt{L_{in}}}, \frac{1}{\sqrt{L_{in}}}] \tag{10}$$

It takes value in its range [35]. In this formula, L_{in} is number of input neurons in the hidden layer. Xavier et al (2010) suggested the following initial range that allows the model to converge better [36]:

$$[\frac{-\sqrt{6}}{\sqrt{L_{in}+L_{out}}}, \frac{\sqrt{6}}{\sqrt{L_{in}+L_{out}}}] \tag{11}$$

In which L_{in} is the number of input neurons and L_{out} being the number of output neurons in the hidden layer. In the developed module, the number of hidden layers and neurons, activation functions of the model, dropout feature are provided to be adjustable. Default value, Xavier initial value and Kaiming initial values were chosen as the initial values of the weight coefficients in the module.

The performance of NN model depends not only on the model parameters but also on the hyperparameters which affect the training of the model. For instance, if the learning rate is set too large for updating the weights during the training, an unstable learning occurs, the model converges to a local minimum or not at all. If it is adjusted as too low, unsuccessful learning occurs and the training is completed without learning much from training set. Learning rate is kept as adjustable in the developed training module. In addition, training with a high learning rate in first epochs and lower learning rates in following epochs is allowed. Early stopping feature has been added because the performance decreases when the number of epochs is too high and the model repeats over the training set too much.

Many different issues to be considered during training on imbalanced dataset. To illustrate, for a successful performance measurement, it should be ensured that the samples belonging to the minority class in K-Fold Cross Validation method are equally distributed among all folds. If the dataset used is too large and the training set will be divided into mini batches and training process is applied with more than one iteration, the distribution of minority samples in batches can be problematic. In the experiments performed, low performance was observed for the training performed with the method in which the minority samples were not evenly distributed, but randomly distributed. For this reason, a sampler module has been developed and implemented to determine the distribution of batches in each epoch. In this module, the number of batches and the number of minority samples required in each batch are determined. Another parameter of the module is the priority of the majority class. If this parameter is set to True for the majority class to have priority, the iteration continues until all the majority samples in the dataset are passed over, while the minority class is used completely, the same minority samples are reused again. If this parameter set to False, the iterations end when all minority samples are used. In the studies conducted, it was observed that the models trained by selecting the majority class with priority showed better performance. In this way, minority samples were randomly re-selected and a procedure similar to the Random Over Sampling method is applied.

Four different model parameters and hyper parameters were chosen for Pin Based NN model. The most successful model performs 0.8711 for PR AUC metric in the test dataset. Parameter values of these models are included in Table 5.

Table 5: Model parameters for Pin Based NN Model.

	Models			
	Model-1	Model-2	Model-3	Model-4
Hidden Layers	[32,64,16,8]	[32,64,16,8]	[32,64,16,8]	[32,64,16,8]
Activation Function	ReLU	ReLU	ReLU	ReLU
Initialization	Kaiming	Kaiming	Kaiming	Kaiming
Dropout	False	False	False	False
Drop Ratio	0	0	0	0
Optimizer	Adam	Adam	Adam	Adam
Batch Size	1024	1024	512	2048
Minority Size	128	128	128	256
Learning Rate	0.001	0.01	0.01	0.01
Weight Decay	0	0.001	0.001	0
Early Stopping	True	True	True	True
Test PR AUC Score	**0.8557**	**0.7874**	**0.7453**	**0.8711**

The most successful model for PR AUC score among pin based models is Random Forest model and best PR AUC score is 0.9230. However, recall value for this model and other pin based models remained very poor compared to NN models. Since this study focuses on determining all possible defects in the dataset and obtaining high recall value, the best recall value is obtained as 0.9544 with four-layered NN model which is named as Model-4. The Loss - Accuracy graph and confusion matrix obtained in the training of this model are shown in Figure 37 and Table 6. As seen in this model, increasing the batch size and the number of minority samples in the batch increased the model performance.

Figure 37: Lost-Accuracy plot of the training of Model-4.

Table 6: Confusion matrix for test performance of Model-4.

	Prediction	
Real	**Negative**	**Positive**
Negative	531219	4085
Positive	23	481

3.4.3 Pin Based-v2 Models

Unlike the previous model, ComponentID feature has been added to the input features in this type of model. ComponentID is a property that indicates which component a sample pin belongs to. For 17MB170R4 panel, there are 12 different components in the panel, with four BGA components in each PCBA, so ComponentID feature consists of integer values between 1 and 12.

When determining input features in machine learning models, features containing reasonable numerical features can be directly given as input to the model. However, features that do not represent a meaningful value and contain discrete values are called categorical features and require preprocessing before being input to the model. Two preprocessing methods are applied to such categorical features, named as Label encoding and One-hot encoding methods. Label encoding method is applied to categorical features that have meaning such as more-less or good-bad, which contain a certain amount of semantics. Since ComponentID feature does not contain such a meaning, 12 features have been transformed with One-hot encoding method, with one feature as 1 in each sample and the other features being 0.

XPost	YPost	AnglePost	ComponentID_0	ComponentID_1	ComponentID_2	ComponentID_3	ComponentID_4	ComponentID_5	ComponentID_6	ComponentID_7	ComponentID_8	ComponentID_9	ComponentID_10	ComponentID_11
1.25	0.94	1.27	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00
-0.44	-0.45	-0.04	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.74	2.78	-2.50	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
-0.22	-0.28	0.13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00
0.78	0.55	0.90	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00

Figure 38: Sample records from the dataset for pin based-v2 models.

Similar to the first model, Random Forest, MLP, Decision Tree and NN algorithms were used in this model. Grid Search algorithm is applied for each model by taking the parameter values used in pin based models and these values can be seen in Table

3. In Random Forest model, the model consisting of 50 estimators showed the best performance, this model achieved average 0.9194 PR AUC score during training and 0.9236 PR AUC in test data. In MLP model, the model with three hidden layers of 30, 60 and 20 neurons showed the best success with 0.9031 PR AUC in test data. Adam optimizer is used as optimizer in all MLP models. In Decision Tree model, the model with a maximum depth of 10 performed the best success. A structure with 10 depth and 113 leaves is obtained in this model and 0.8835 PR AUC is achieved. The parameters of Random Forest, MLP and Decision Tree models which perform the best success are given in Table 7.

Table 7: Best parameters for pin based-v2 models.

Random Forest		
Estimator Size	**Maximum Depth**	**Splitting Criterion**
50	50	Entropy
MLP		
Hidden Layers	**Activation Function**	**Learning Rate Type**
[30,60,20]	Tanh	Constant
Decision Tree		
Maximum Depth	**Splitter**	**Splitting Criterion**
10	Best	Entropy

Table 8: Test performances of pin based-v2 models.

Model	**PR AUC**	**F2-Score**	**Recall**
Random Forest	0.9236	0.8910	0.8690
MLP	0.9031	0.8635	0.8413
Decision Tree	0.8835	0.8847	0.8651

In Pin Based-v2 NN model, since the number of input features increased compared to the first model, the models are chosen so that there would be more neurons in the first hidden layer. In this model, training is performed with seven different models and hyper parameters, and the model parameters and performances can be seen in Table 10. As is seen, adding dropout to the model increases the performance. When comparing Model-1 and Model-2 models, adding dropout increased the model performance by 4%. Comparing Model-2 and Model-4, it has been observed that ReLU activation function is more successful in this dataset than Rectified ReLU function. From Model-5 and Model-6, it is seen that increasing the batch size and learning rate at the same time increased the model performance by 12%.

The model with the highest PR AUC score among Pin Based-v2 models is Random Forest model with 0.9236 PR AUC but this model performed insufficiently with 0.8690 recall score. However, four-layered NN model named Model-2 achieved the highest success with a value of 0.9025 PR AUC and 0.9583 recall score. This model has proven that dropout is an important regularization method that improves model performance. The loss-accuracy graph and confusion matrix of the model predictions are as follows.

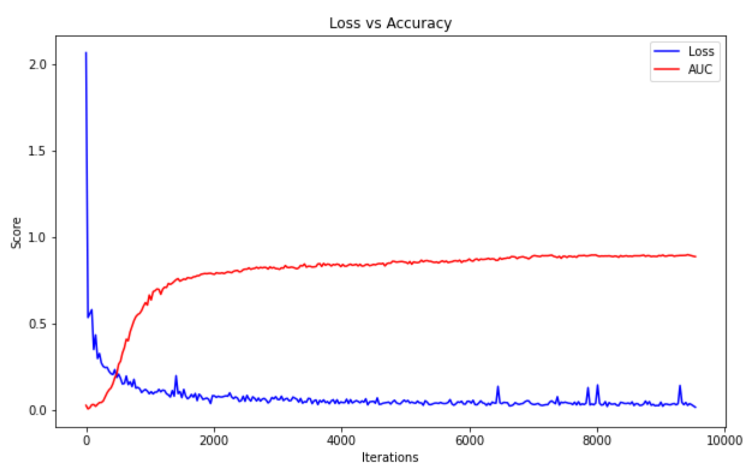

Figure 39: Lost-Accuracy plot of the training of Model-2.

Table 9: Confusion matrix for test performance of Model-2.

	Prediction	
Real	**Negative**	**Positive**
Negative	531532	3772
Positive	25	479

Table 10: Model parameters for Pin Based-v2 NN Model.

	Models						
	Model-1	Model-2	Model-3	Model-4	Model-5	Model-6	Model-7
Hidden Layers	[64,32, 32,8]	[64,32, 32,8]	[64,32, 32,8]	[64,32, 32,8]	[32,64, 16,8]	[32,64, 16,8]	[128,64, 32]
Activation Function	ReLU	ReLU	Leaky ReLU	Leaky ReLU	ReLU	ReLU	ReLU
Initialization	Kaiming	Kaiming	Kaiming	Kaiming	Kaiming	Kaiming	Default
Dropout	False	True	False	True	False	False	False
Drop Ratio	0	0.2	0	0.25	0	0	0
Optimizer	Adam	Adam	Adam	Adam	Adam	Adam	Adam
Batch Size	1024	1024	1024	1024	1024	2048	1024
Minority Size	128	128	128	128	128	256	128
Learning Rate	0.001	0.001	0.001	0.001	0.001	0.005	0.005
Weight Decay	0	0	0	0	0	0	0
Early Stopping	True	True	True	True	True	True	True
Test PR AUC Score	0.8622	0.9025	0.8291	0.8641	0.7777	0.8963	0.8808

3.4.4 Component Based Models

The main purpose of this model is to evaluate the pins belonging to the same component together and to predict the result of X-Ray on component basis, assuming that a problem that may occur in soldering component will be seen in all pins. Thus, all pin records belonging to a component are grouped and virtual features are examined. In the dataset, ComponentID property shows that each pin belongs to which component. However, since there are different pin measurements of the same ComponentID and different panels, all pin samples are grouped by MasterCode and ComponentID columns. Afterwards, min, max, mean and standard deviation values of pin measurements for each feature were found for each group and a total of 80 virtual features were obtained.

(XPost, min)	(XPost, max)	(XPost, mean)	(XPost, std)	(YPost, min)	(YPost, max)	(YPost, mean)	(YPost, std)	(AnglePost, min)	(AnglePost, max)	(AnglePost, mean)
-1.27	-1.27	-1.27	0.00	-1.14	-1.14	-1.14	0.00	-0.68	-0.68	-0.68
-1.99	-1.99	-1.99	0.00	-1.73	-1.73	-1.73	0.00	-1.24	-1.24	-1.24
-0.01	-0.01	-0.01	0.00	-0.10	-0.10	-0.10	0.00	0.29	0.29	0.29
0.12	0.12	0.12	0.00	0.01	0.01	0.01	0.00	0.40	0.40	0.40
-1.50	-1.50	-1.50	0.00	-1.33	-1.33	-1.33	0.00	-0.86	-0.86	-0.86

Figure 40: Sample records from the dataset for component based models.

Similar to the previous models, Grid Search algorithm was applied with the model parameters in Table 3, and the parameter values different from the best parameters obtained in pin based-v2 model showed the best results in component based models, these parameters are included in Table 11. In Random Forest model, the model with 30 estimators performed the highest score in validation with 0.6878 PR AUC, but it succeeded 0.4920 PR AUC in test data. The most successful three-layered MLP model achieved 0.4652 PR AUC in test performance. In Decision Tree model, the model with depth of 40 and 83 leaves and "Gini index" as the splitting criterion obtained 0.3126 PR AUC.

Since this model has 80 input features and has much more input dimensions than other models, models with more layers and numbers of neurons are preferred. In models trained with four different parameters, it was observed that the high number of layers did not improve the model performance much. Apart from that, when the

Table 11: Best parameters for component based models.

Random Forest		
Estimator Size	**Maximum Depth**	**Splitting Criterion**
30	50	Entropy
MLP		
Hidden Layers	**Activation Function**	**Learning Rate Type**
[10,20,5]	Tanh	Constant
Decision Tree		
Maximum Depth	**Splitter**	**Splitting Criterion**
40	Random	Gini

Table 12: Test performances of component based models.

Model	PR AUC	F2-Score	Recall
Random Forest	0.4920	0.4144	0.3846
MLP	0.4652	0.4891	0.4615
Decision Tree	0.3126	0.3298	0.4872

models are compared, it has been observed that better performance is obtained in models with more minority samples in batches.

Five-layered NN model with 0.5486 PR AUC score showed the best performance among Component Based models. All Component Based models showed lower performance than Pin-Based and Pin Based-v2 models. One of the significant reasons for this situation is that there are not too many data on the basis of components in the collected data. After grouping by MasterCode and ComponentID features, 9634 non-defective and 78 defective samples were obtained in the training dataset, 9654 non-defective and 39 defective samples in the test dataset. Since the dataset contains very few data, the models could not learn sufficiently. In addition, pin-based X-Ray test results were used while labeling samples during data collection. In component

Table 13: Model parameters for Component Based NN Model

	Models			
	Model-1	Model-2	Model-3	Model-4
Hidden Layers	[128,64,32,16,8]	[128,64,32,16,8]	[128,64,32]	[32,16,8]
Activation Function	ReLU	ReLU	ReLU	ReLU
Initialization	Kaiming	Kaiming	Kaiming	Kaiming
Dropout	True	True	False	False
Drop Ratio	0.4	0.5	0	0
Optimizer	Adam	Adam	Adam	Adam
Batch Size	64	64	64	64
Minority Size	16	8	8	16
Learning Rate	0.001	0.001	0.0001	0.005
Weight Decay	0	0	0	0.001
Early Stopping	True	True	True	True
Test PR AUC Score	**0.5487**	**0.3693**	**0.4325**	**0.4508**

based approach, while labeling is performed after the grouping process, if there is any defective pin record of a component, that component is deemed to be defective. In this case, components with a small number of defective pins and having all other pins are normal but are labeled as defective cause the model not to learn the data accurately, since even the min, max, mean and std values are close to non-defective samples.

Table 14: Confusion matrix for test performance of Model-1.

	Prediction	
Real	**Negative**	**Positive**
Negative	9614	40
Positive	13	26

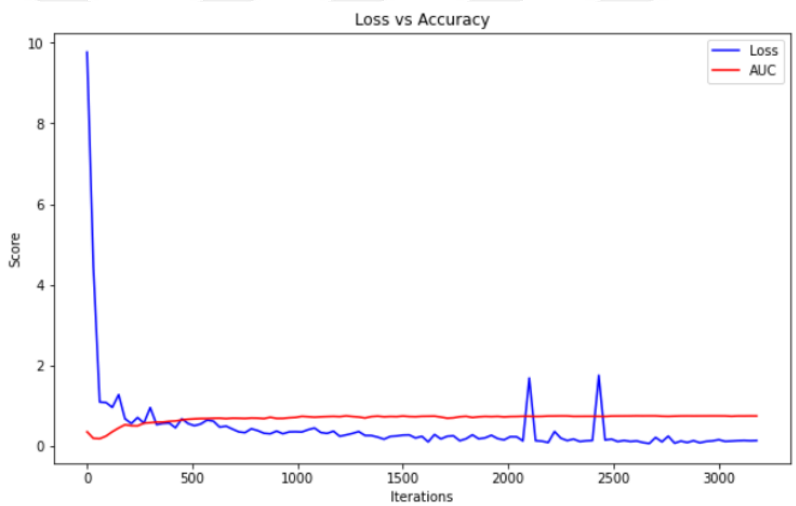

Figure 41: Lost-Accuracy plot of the training of Model-1.

CHAPTER IV

CONCLUSION

4.1 Discussion

Machine learning models were constructed using three different models and four different algorithms and trained each model in Chapter 3 for the prediction of X-Ray test results with the measurement values taken from SPI, Pre-AOI and Post-AOI machines. It has been observed in the studies that the most successful model for both PR AUC and recall scores is four-layered Pin Based-v2 NN model. This model achieved 0.9025 PR AUC, 0.95 Recall and 0.38 F2-Score in the test data. Between pin based models, Random Forest model achieved the highest PR AUC score with 0.9230 value but its recall performance remained poor with 0.8690 score. On the other hand, four-layered Pin Based NN model also achieved high success with 0.8711 PR AUC score and 0.9544 recall score. The highest performance was obtained in component based models with 0.5487 PR AUC and 0.7179 recall value in five layered NN model.

Table 15: Model parameters of the best performing model without Grid search.

Pin Based-v2 NN Model			
Hidden Layers	[64,32,32,8]	**Batch Size**	1024
Activation Function	ReLU	**Minority Size**	128
Initialization	Kaiming	**Learning Rate**	0.001
Dropout	True	**Weight Decay**	0
Drop Ratio	0.2	**Early Stopping**	True
Optimizer	Adam	**PR AUC Score**	0.9025

The reason that the performances of the component based models are lower than the other two models is the data obtained after grouping operation was insufficient for model training since there were not enough measurements from different components in the data. In addition, in this model, the component with defects in any of the pins was labeled as defective while creating the dataset, even though the measurement values of all other pins were close to non-defective pin measurements, which caused the dataset to have a noise structure and affected the performance.

```
begin
first_k_thresh ← 0.75
all_k_thresh ← 0.80
best_score ← 0
for parameter in parameters do
   val_list ← empty list
   for k = 1 to 5 do
      train(model)
      score ← validation(model)
      val_list.insert(score)
      avg ← mean(val_list)
      if (k = 1 and avg < first_k_thresh) or
         ( k > 1 and (avg < all_k_thresh or avg < best_score))
         break
      end if
   end for
   if avg > best_score
      best_score ← avg
   end if
end for
end
```

Figure 42: Pseudo code that explains Grid search algorithm.

It has been observed that the best successful model is Pin Based-v2 NN Model, but this model may not have the optimum parameters. In order to find the optimum parameters of the model, a Two-Stage Grid Search algorithm has been applied. At the first stage, all parameter values that had good performance in previous trials were taken. Average PR AUC values were obtained for each value of the parameters with K = 5 K-Fold Cross Validation method. In order for Grid search to give faster results, training was stopped before five iterations were completed for the models with low performance in the first iterations and it was passed to the next parameter group. A total of 1152 models were trained for different parameter values and the average scores were recorded. At the second stage, the first 50 parameter groups with the

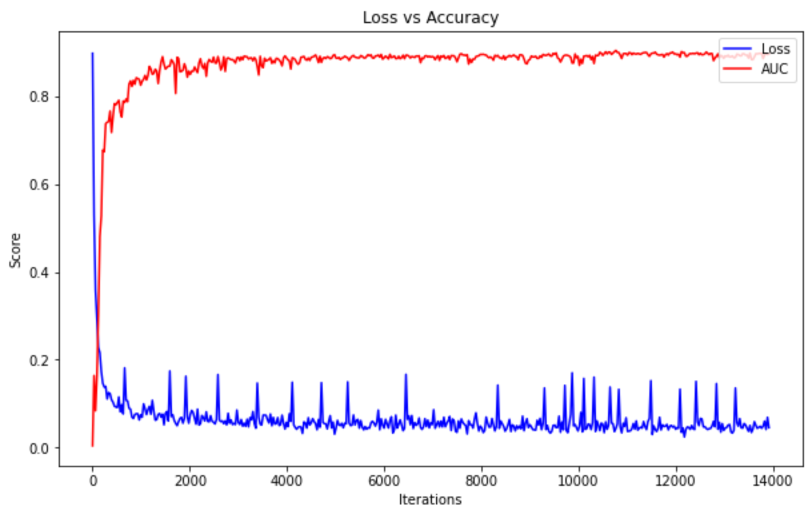

Figure 43: Lost-Accuracy plot for the model with best parameters.

highest value of PR AUC above 0.85 in the first phase were trained to complete all iterations for K = 5 and the parameters with the highest performance was chosen. As the last step in the model with the most optimum parameter values, training process with five epochs was applied by passing over all training data. The obtained model has achieved 90.62% success for PR AUC metric.

Table 16: Parameters values chosen for Grid search.

	Parameter Options	**Best Parameter**
Hidden Layer	[128, 64, 32], [32,64,16,8], [32,64,32,16,8]	[128,64,32]
Drop Ratio	0.0, 0.2, 0.4	0.2
Initialization	"default", "he"	"default"
Optimizer	Adam, SGD	Adam
Batch Size	2048, 1024	1024
Minority Size	128, 256	128
Learning Rate	0.001, 0.005, 0.01, 0.03	0.001
Weight Decay	0.0, 0.001	0.0

Pin Based NN model was examined to understand the effect of batch size, number of minority samples in batches and epoch number on the model performance. In this experiment, the number of batch and number of minority samples were kept low and the number of epochs was kept high in the first training. In addition, in the 16th epoch, the learning rate was decreased by 0.1 and an adaptive learning rate was applied. In the second training, the number of batch size and minority samples was chosen as high and epoch number was chosen as low for the same model. Test performance of the model was 0.8114 PR AUC in the first training, and 0.8037 PR AUC in the second training. As a result of this study, it was seen that keeping the epoch number high by decreasing the batch number and minority samples increases the model performance.

Table 17: Training parameters to see the effect of batch size and minority size on the model.

Model Parameters	Hidden Layers: [32,64,16,8], Activation: ReLU, Initialization: Kaiming, Dropout: False
Hyper Parameters for Training-1	Batch Size:64, Minority Size:16, Epoch:20, LR: 0.003, Weight Decay:0.001, LR Decay Epoch:16, LR Decay Ratio:0.1
Hyper Parameters for Training-2	Batch Size:1024, Minority Size:128, Epoch:5, LR: 0.003, Weight Decay:0.001

Another experiment is examined in the pin based-v2 model in order to observe the effect of "majority priority" parameter in developed batch sampler module over training sessions. The task of batch sampler module is to provide an equal number of minority samples in each of the batches applied to the model during training. Majority priority parameter in this module provides specifying how majority instances will be used during the training of the model. If this parameter is set to True, it is aimed to use all majority samples during an epoch, iteration continues until all majority

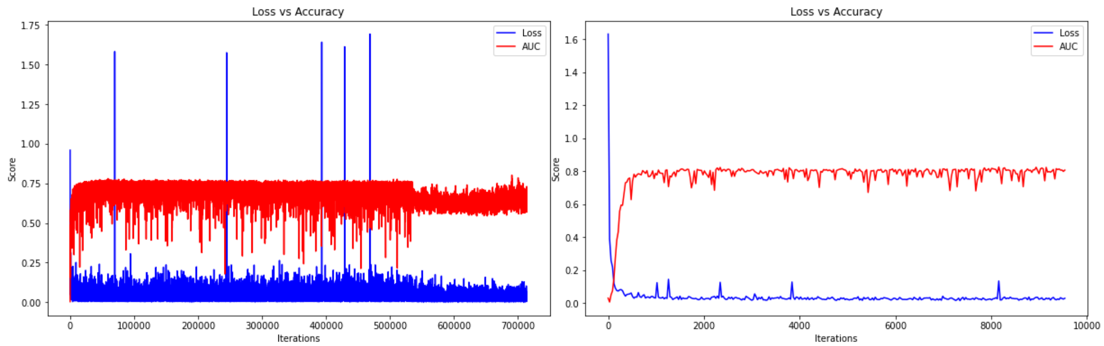

Figure 44: Lost-Accuracy plots for Training-1 and Training-2.

samples are used. If all minority samples are trained during the epoch, the minority samples are repeated. Since the same minority samples are used multiple times for an epoch, it has the same effect as applying over-sampling method for minority samples. If this parameter is set to False, it is aimed to use all minority samples during an epoch, iteration continues until all minority samples are used. Since some of majority samples are chosen randomly for an epoch, this process has the same effect as applying under sampling method for majority samples. In the experiment, the model parameters that performs the best success in the pin based-v2 model were chosen. These parameters is explained in Table 10. Three different training sequences were created for this model. The majority priority parameter is set to True in the first sequence and False in the second and third sequences. Since the number of minority samples is very low in the dataset and an epoch is completed when all minority samples are used for False value of majority priority, the number of iterations in epochs in the second and third sequences is less than in the first sequence. The number of epochs was set differently in three training sequences, so it is ensured that number of iteration for each sequence remains the same. Training parameters are explained in Table 18.

As a result of the experiment, 0.9035 PR AUC in Training-3, 0.9017 PR AUC in Training-4, and 0.8506 PR AUC in Traning-5 were obtained in the test dataset. It was observed that similar results were obtained when similar iterations were provided

Table 18: Training parameters in the experiment to measure the effect of majority priority parameter on training.

	Majority Priority	Batch Size	Minority Size	Epoch Count
Training-3	True	1024	128	5
Training-4	False	1024	128	800
Training-5	False	1024	16	100

in Training-3 and Training-4, but lower performance was observed in Training-5. In Training-5 sequence, unlike Training-4, batches with less minority size was used, which involved to train with more iterations and to use more majority samples, but the performance was not sufficient. In addition, it was observed in previous experiments that the low number of minority samples in batches negatively affected the model performance. Although Training-5 sequence remained lower than the other models in PR AUC and Recall values, it achieved a better result for F2-Score. The loss-accuracy curves and confusion matrices for the training are can be observed in Figure 45, 46 and 47 and Table 19 .

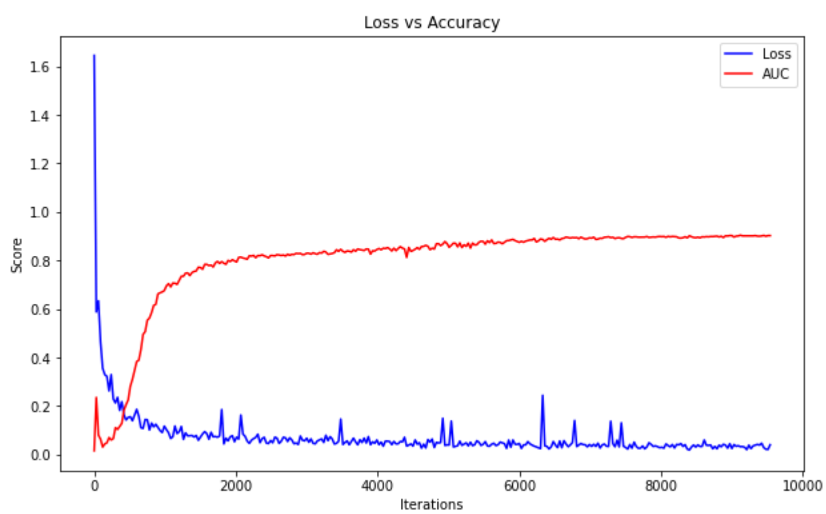

Figure 45: Lost-Accuracy plot of the training of Training-3.

Figure 46: Lost-Accuracy plot of the training of Training-4.

Table 19: Confusion matrix for test performance of Training-3, Training-4 and Training-5.

	Training-3		**Training-4**		**Training-5**	
Real	**Negative**	**Positive**	**Negative**	**Positive**	**Negative**	**Positive**
Negative	531724	3580	531695	3609	535238	66
Positive	16	488	16	488	58	446

While creating a dataset from the data obtained from the control machines in production, three different datasets were created, and the training process of the models was performed with the first dataset. The other two datasets are obtained by applying PCA analysis and resampling to the first dataset. In the second dataset, 13 eigen vectors were obtained from 20 features by PCA analysis and the input dimensions were reduced to 13 dimensions. In the third dataset, Random Under Sampling, Tomek-Links and ENN methods were applied and records belonging to the majority class were sampled and reduced from 2141216 samples to 920152. Training with both datasets had low performance. The reason for this issue is that some information in the data was lost in PCA analysis and under-sampling method, the model was not successful in predicting different samples in the test data.

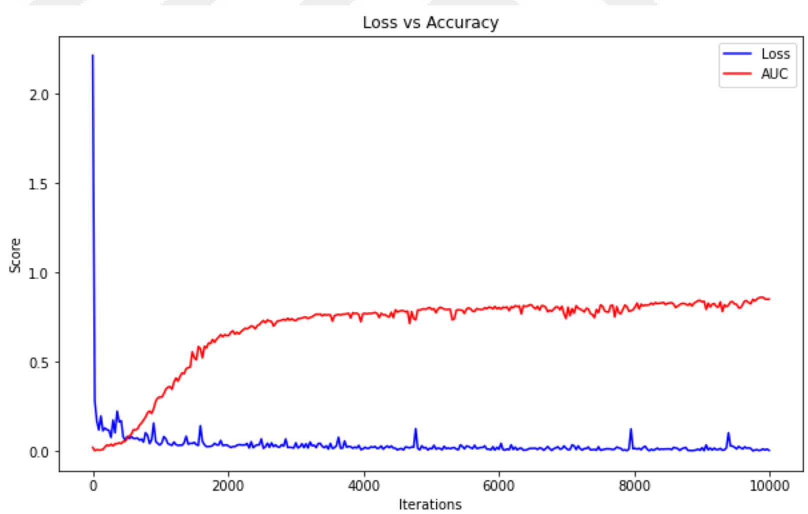

Figure 47: Lost-Accuracy plot of the training of Training-5.

Comparing the training with the dataset obtained by applying under sampling to majority samples with Tomek-Links and ENN methods and the training with random under sampling by adjusting the majority priority parameter to False, the training with the dataset that is applied by Tomek-Links and ENN method under sampling was less successful. Although majority instances are sampled for an epoch in the training with random under sampling method, the samples not included in training for an epoch may be included in the training randomly in the next epoch. However, the majority samples excluded by Tomek-Links and ENN methods are not included in the training anymore. This may cause some information to be lost in the training dataset and the model to be underfitted.

In Vestel Electronics Factory, X-Ray tests are not processed on the production line, but the boards are chosen by random sampling tested in a separate station. Random sampling of the boards has two main problems. First, if a defect in solder paste or component cannot be detected by the control machines on the production line and is not chosen by random sampling, this board can complete its production cycle and can be shipped without any defect detection. Second, because a method does not exist to choose the boards that need to be analyzed by X-Ray inspection device, overmuch boards are tested with random sampling. This causes excessive use of X-Ray inspection devices which have high initial and maintenance costs. By virtue of proposed model, the targeted success was achieved by reducing the number of sampling by 4% in Line-1913.

4.2 Future Work

In PCBA production, many critical operations are applied to the boards. These processes are controlled by the quality control machines in the production line and the boards containing defective components are detected and thrown away during production. Perfect defect detection mechanism does not exist in the quality control of the produced boards, because the measurement results of the machines performing the control depend on the threshold values set on the machine. It is not possible to adjust these threshold values to the optimum values in practical. In addition, in measurement results which are very close to the threshold value, the decision is made by the operator, which affects the reliability and stability of the process. In order to eliminate these problems and to detect errors that cannot be caught in the production line, the boards selected by random sampling are analyzed with X-Ray inspection device. Random sampling may cause missing defective cards that are not identified in the production line and not selected for sampling. In this project, it was aimed to develop a machine learning model that predicts X-Ray test results of the pins with the measurement values taken from SPI, Pre-AOI, Post-AOI machines. Thanks to this model, the components containing the pins labeled by the model as defective and the boards with these components are determined. Instead of randomly sampled boards, the boards that the model determines as defective are tested in X-Ray inspection device. Thus, with a more reasonable sampling, all faulty samples are detected on the X-Ray device without missing. In addition, in manual type X-Ray devices, the components detected by the model as defective are reported to the operator and the operator is prevented from overlooking these defects. With fewer defective boards detected than random sampling, the use of X-Ray machines and maintenance costs are reduced due to less operating times, and the number of X-Ray machines required to detect defective boards on all production lines is decreased.

A machine learning model has been developed to detect defects that may occur in BGA components on PCBAs known as the main board of televisions at the Vestel Electronics Factory. Server software has been developed in order to collect measurement values from SPI, Pre-AOI and Post-AOI machines and to label them with the test results on the X-Ray detection device, and approximately 3 million data have been collected for two months. Different models have been developed with three different models and four different machine learning algorithms. As a result of the studies, PR AUC score of 90.62% and recall score of 97.42% was obtained in the test data with the three-layered Pin Based-v2 NN model. With the commissioning of the model in production, the number of X-Ray sampling for Line-1913 was reduced by 4%.

Because the amount of data collected from the machines is not sufficient, low success was achieved in component based models. If data is collected in a longer period of time, it is expected that the model performance will increase in all models. It is aimed to use the results in X-Ray tests and FVT tests while labeling the measurement values collected from the control machines. Thus, if a component labeled as non-defective in X-Ray test fails FVT test, this component and its pins would be marked as defective. However, no defects were detected in the FVT test during data collection, so this feature was not used during training. In future studies, a model including this feature will give more accurate results. Furthermore, data was collected from only three control machines in the production line in this project. If a model is developed by taking data also from SP, P&P and Reflow Oven machines, it will be possible to obtain a more precise and better model in future works.

Since it is not possible to determine a possible flaw in the solder pads under BGA components in the production line and these components have a higher risk of being undetected, models that predict the X-Ray analysis result of BGA components have been developed in this project. However, among the studies, the most successful model was achieved with a pin-based model that is independent of the component. It is possible to analyze the results of this model in future studies by applying to different components on the boards such as resistors and capacitors. Moreover, an offline training system has been established in this project, it has been planned to train and update model parameters with accumulated data with certain periods. In the long run, it is aimed to develop a dynamic model in which model parameters are continuously updated with online training.

Bibliography

[1] *PCB basic knowledge.* https://tinyurl.com/y9b6nplp (Accessed: 15.08.2020).

[2] *BGA Rework Station with Easy Operation and High Performance.* https://www.seamarkzm.com/products/bga-rework-station (Accessed: 15.08.2020).

[3] *Solder Paste Printing Process.* https://www.surfacemountprocess.com/solder-paste-printing-process.html (Accessed: 17.08.2020).

[4] *VersaPrint 3D-SPI.* https://www.kurtzersa.com/electronics-production-equipment/festival-of-innovations/versaprint-3d-spi.html (Accessed: 17.08.2020).

[5] *SMT.* https://www.newonics.com/SurfaceMountTechnology-SMT.html (Accessed: 18.08.2020).

[6] *Pin-In-Paste Technology, The Cheapest Way To Solder THT Components.* https://myproto.eu/news/pin-in-paste-technology-the-cheapest-way-to-solder-tht-components/ (Accessed: 20.08.2020).

[7] *The PCB Assembly Process.* https://www.pcbway.com/blog/PCB_Assembly/The_PCB_Assembly_Process.html (Accessed: 20.08.2020).

[8] *Gigabyte Factory Tour 2011.* https://bit-tech.net/reviews/gigabyte-factory-tour/1/ (Accessed: 17.08.2020).

[9] *Automated Inspection Process.* https://www.venturatech.com/quality/inspection/ (Accessed: 21.08.2020).

[10] *Reflow Soldering Process.* https://www.surfacemountprocess.com/reflow-soldering-process.html (Accessed: 20.08.2020).

[11] M. Scalzo, "Addressing the challenge of head-in-pillow defects in electronics assembly," *IPC APEX EXPO Proceedings.*

[12] S. Sattel, *What Is PCB Tombstoning?* https://www.autodesk.com/products/eagle/blog/what-is-pcb-tombstoning/ (Accessed: 22.08.2020).

[13] H. Biemans, "5d solder paste inspection-merits beyond 3d technology," *Global SMT & Packaging*, vol. 24, no. 3, pp. 8–13, 2011.

[14] P. M.LaCasse, W. Otieno, and F. P. Maturana, "Predicting contact-without-connection defects on printed circuit boards employing ball grid array package types: a data analytics case study in the smart manufacturing environment," *SN Applied Sciences*, vol. 2, Jan 2020.

[15] *Resampling strategies for imbalanced datasets*. https://www.kaggle.com/rafjaa/resampling-strategies-for-imbalanced-datasets#t1 (Accessed: 26.08.2020).

[16] *Cross-validation: evaluating estimator performance*. https://scikit-learn.org/stable/modules/cross_validation.html (Accessed: 06.09.2020).

[17] M. Deshpande, *Perceptrons: The First Neural Networks*. https://pythonmachinelearning.pro/perceptrons-the-first-neural-networks/ (Accessed: 10.11.2020).

[18] *Activation Function*. https://docs.paperspace.com/machine-learning/wiki/activation-function (Accessed: 11.11.2020).

[19] *Chapter 1: Using Neural Nets to Recognize Handwritten Digits*. http://neuralnetworksanddeeplearning.com/chap1.html (Accessed: 11.11.2020).

[20] G. Singh, *From SGD to Adam*. https://medium.com/mdr-inc/from-sgd-to-adam-c9fce513c4bb (Accessed: 12.11.2020).

[21] *Ball Grid Array, BGA*. https://www.electronics-notes.com/articles/electronic_components/surface-mount-technology-smd-smt/ball-grid-array-bga.php (Accessed: 13.08.2020).

[22] K. Seelig, "Head-in-pillow bga defects," tech. rep., AIM, Cranston, Rhode Island, USA.

[23] S. S. Z. et al, "Automated detection of printed circuit boards (pcb) defects by using machine learning in electronic manufacturing: Current approaches," *IOP Conf. Ser.: Mater. Sci. Eng.*, vol. 767, Dec 2020.

[24] A. W. Foster, "Predicting solder defects in printed circuit board assembly (pcba) process," Master's thesis, Jun 2019.

[25] J. D. Song, Y. G. Kim, and T. H. Park, "Smt defect classification by feature extraction region optimization and machine learning," *The International Journal of Advanced Manufacturing Technology*, vol. 101, 2018.

[26] G. Acciani *et al.*, "Application of neural networks in optical inspection and classification of solder joints in surface mount technology," *IEEE Transactions On Industrial Informatics*, vol. 2, Aug 2006.

[27] *A Real-time Machine Vision System for Solder Paste Inspection*, International Conference on Advanced Intelligent Mechatronics, Jul 2008.

[28] Y. M. Chang *et al.*, *Classification of Solder Joints via Automatic Mistake Reduction System for Improvement of AOI Inspection.*

[29] G. E. A. P. A. Batista, R. C. Prati, and M. C. Monard, *A Study of the Behavior of Several Methods for Balancing Machine Learning Training Data*. Sao Carlos - SP, Brazil.

[30] *Curse of dimensionality*. `https://en.wikipedia.org/wiki/Curse_of_dimensionality` (Accessed: 28.08.2020).

[31] *What is Covariance?* `https://corporatefinanceinstitute.com/resources/knowledge/finance/covariance/` (Accessed: 01.09.2020).

[32] T. Wood, *F-Score*. `https://deepai.org/machine-learning-glossary-and-terms/f-score` (Accessed: 05.09.2020).

[33] F. Rosenblatt, "The perceptron: A probabilistic model for information storage and organization in the brain," *Pyschological Review*, vol. 65, no. 6, 1958.

[34] D. Godoy, *Understanding Binary Cross-Entropy / Log Loss: A visual Explanation*. `https://towardsdatascience.com/understanding-binary-cross-entropy-log-loss-a-visual-explanation-a3ac6025181a` (Accessed: 12.11.2020).

[35] J. Santarcangelo, *Neural Network initialization Weights*. `https://www.coursera.org/learn/deep-neural-networks-with-pytorch/lecture/IUPoB/8-3-neural-network-initialization-weights` (Accessed: 20.08.2020).

[36] X. Glorot and Y. Bengio, *Understanding the difficulty of training deep feedforward neural networks*. Montreal, Quebec, Canada.

VITA

EDUCATION

Bachelor of Engineering in Control and Automation Engineering at Electrical and Electronics Faculty, Istanbul Technical University, June 2015.
Thesis Title: *Implementing Improvement of Push Recovery Algorithm for Humanoid Robot Nao*

Bachelor of Engineering in Mechanical Engineering at Mechanical Engineering Faculty, Istanbul Technical University, August 2014.
Thesis Title: *Adaptive Solar Panel System Design*

ACADEMIC AWARDS

Student Graduation Certificate – Istanbul Technical University (within 2nd percentile 5 year prior from the graduation)