DOKUZ EYLÜL UNIVERSITY

ENGINEERING FACULTY

DEPARTMENT OF COMPUTER ENGINEERING

QUALITY CONTROL DECISION SUPPORT SYSTEM

by

Zeynep KAYA

Nilay YÜCEL

Advisor

Prof. Dr. Recep Alp KUT

June, 2022

İZMİR

QUALITY CONTROL
DECISION SUPPORT SYSTEM

A Thesis Submitted to the
Dokuz Eylül University, Department of Computer Engineering
In Partial Fulfillment of the Requirements for the Degree of B.Sc.

by
Zeynep KAYA
Nilay YÜCEL

Advisor
Prof. Dr. Recep Alp KUT

June, 2022
İZMİR

SENIOR PROJECT EXAMINATION RESULT FORM

We have read the thesis entitled **"QUALITY CONTROL DECISION SUPPORT SYSTEM"** completed by **Nilay YÜCEL** and **Zeynep KAYA** under advisor of **Prof. Dr. Alp KUT** and we certify that in our opinion it is fully adequate, in scope and in quality, as a thesis for the degree of B.Sc.

Prof. Dr. Alp KUT

Advisor

Committee Member

Committee Member

Prof. Dr. Yalçın ÇEBİ

Chair

Department of Computer Engineering

ACKNOWLEDGEMENTS

We would like to thank Prof. Dr. Alp Kut, our consultant for his support during these studies, and Totomak Makina ve Yedek Parça San.Tic. A.Ş., who provided the necessary data for the conduct of our studies. In addition, we would like to thank our friends for their support during this difficult process and our families who have always been with us throughout our education life.

Zeynep KAYA
Nilay YÜCEL

QUALITY CONTROL DECISION SUPPORT SYSTEM

ABSTRACT

Quality has been in our lives as an important concept in every field from past to present. It is also of great importance in the manufacturing sector. From this point of view, it would not be difficult to estimate the place of quality control in this sector. In the production sector, which accelerates as mechanization increases, it has become a very difficult situation to monitor and control quality.

“If the role of the machine in production is increasing, why not in quality control?”. With this question, machines were needed to support human power and decision. With these systems, which have started to take their place in this sector as decision support systems, it is aimed to enable people to make decisions like humans by teaching their behaviors/decisions to a machine.

Within the framework of this purpose, the methods, models and results used in researches and studies in this field were examined. It has been observed that very successful results have been achieved with a correct working method. At the end of the preliminary studies, the real quality control measurement values taken from a factory producing machine parts and the production decisions made by people according to these values were decided to be processed. However, it has been observed that this data, which is not collected for the purpose of teaching the machine, will complicate the work.

As an additional and preliminary study, it was decided to collect the reports containing the quality control measurement values in a system belonging to us and to make the human decision through this system. In this way, the report content has been converted into a suitable form for machine learning modelling. At the same time, the data for the human decision, namely the target column of the study, were collected from a single source. A data set ready for modeling was obtained by applying the methods take as a result of the research on the collected data.

Algorithms with high success were determined in the studies conducted in this area. The modeling process was carried out by selecting the best parameters.

When the results were examined, it was observed that very high success values were obtained. These results greatly increase the confidence in the decision of the decision support system.

KALİTE KONTROL KARAR DESTEK SİSTEMİ

ÖZET

Kalite geçmişten günümüze her alanda önemli bir kavram olarak hayatımızdadır. Üretim sektöründe de büyük bir önem taşır. Buradan yola çıkılacak olursa kalite kontrolün de bu sektördeki yerini tahmin etmek zor olmaz. Makineleşme arttıkça hızlanan üretim sektöründe kalitenin takibi ve denetlenmesi oldukça güç bir durum haline gelmiştir.

"Üretimde makinenin rolü arttıysa kalite kontrolde neden artmasın?" sorusuyla birlikte insan gücüne ve kararına destek amacıyla yine makinelere ihtiyaç duyulmuştur. Karar destek sistemleri olarak bu sektörde yerini almaya başlayan bu sistemlerin, insanların davranışlarını/kararlarını bir makineye öğreterek insan gibi karar vermesini sağlamak amaçlanmaktadır.

Bu amaç çerçevesinde, bu alanda yapılan araştırmalar ve çalışmalar incelenerek kullanılan yöntemler, modeller ve çalışmaların sonuçları incelenmiştir. Doğru bir çalışma yöntemiyle oldukça başarılı sonuçlara ulaşıldığı gözlemlenmiştir. Ön çalışmalar sonunda, makine parçaları üreten bir fabrikadan alınan gerçek kalite kontrol ölçüm değerleri ve bu değerlere göre insan tarafından verilen üretim kararları alınarak işlenmesine karar verilmiştir. Fakat makineye öğretme amacıyla toplanmayan bu verilerin işi zorlaştıracağı gözlemlenmiştir.

Ek ve ön bir çalışma olarak kalite kontrol ölçüm değerlerini içeren raporların bize ait bir sistemde toplanmasına ve insan kararının bu sistem üzerinden verilmesine karar verilmiştir. Bu sayede rapor içeriği makine öğrenmesi modellemesi için uygun forma dönüştürülmüş oldu. Aynı zamanda verilen insan kararı, yani çalışmanın hedef sütunu için veriler de tek elden toplanmış oldu. Toplanan veriler üzerinde, araştırma sonucu ulaşılan yöntemler, uygulanarak modellemeye hazır bir veri seti elde edildi. Bu alanda yapılan çalışmalarda yüksek başarılar elde eden algoritmalar belirlendi. En iyi parametlerin seçilmesiyle modelleme işlemi gerçekleştirildi.

Sonuçlar incelendiğinde oldukça yüksek başarı değerlerinin elde edildiği gözlemlenmiştir. Bu sonuçlar karar destek sisteminin vereceği karara olan güveni oldukça arttırmaktadır.

CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE

INTRODUCTION

1.1. Background Information

Quality has been defined in many ways from the past. When all definitions are examined, in summary, quality is a criterion that shows how much of the features that should be found in a product or service are met. According to known records, the history of the concept of quality dates back to BC. It dates back to 2150 years. Article 229 of the Law of Hammurabi, published by Hammurabi, the sixth king of the First Babylonian Dynasty, states, "If a builder builds a house for a man and the house is not strong enough and collapses on the owner, causing his death, that builder's head will be cut off." The foundation of the concept of quality with its most primitive definition was laid here and it has become an important part of our lives today.

It is possible to examine quality in many sub-branches such as design, conformity, use, distribution, and relationship. In order to examine the quality, the definition of quality control comes out. Quality control is a system that measures the quality of a product or service and produces results by comparing it with the required standard.

Quality control aims to provide the best quality production/service at the most affordable cost. Quality control has great importance for both production/service providers and customers. At first, the implementation of these systems seems to be costly, but it is not important compared to the losses caused by poor quality. It helps to prevent big financial losses if the production/service providing organizations pass the work done through a correct quality control process. In fact, in the long term, it is seen that it reduces the production/service cost considerably.

In addition to the financial benefits, with the increase in the quality of the product/service offered to the consumer, a significant increase in customer satisfaction is also achieved. With the satisfaction of the consumer, there is an increase in the prestige to rival companies. Accordingly, increasing the motivation of the manager, strengthens the relations with the employee.

With the concept of 'learning machines' that emerged in the 1970s, software-based solutions began to emerge for the problems experienced. The increase in these solutions was followed by methods such as analytical modeling, analyzing data and evaluating relationships. With the processing of the obtained data in machines that learn with these methods, meaningful results have begun to be obtained. Thus, these machines, which transform raw data into meaningful information, are used to support decision makers. These systems, which help the decision-making process to achieve more efficient and precise results in the quality control process, have started to become widespread today.

1.2. Problem Definition

Nowadays, quality control systems are generally used by large companies engaged in mass production to detect products that do not comply with the standard. In case of faulty production of a product, that production batch should be checked and production should be stopped in order not to produce more faulty products. The cause of the damage should be determined, the damage should be corrected, and then production should be resumed as soon as possible in accordance with the standard.

Machine learning applications have started to be used in all areas of our lives with the improved technology. However, machine learning applications have just started to be used in quality control systems. All these quality control, production shutdown, damage assessment and repair works are carried out manually with company employees. This causes some deficiencies and problems arising from limited human power. Due to human physiological nature, he cannot examine meaningless large data pools in a short time and draw meaningful conclusions. At the same time, due to its psychological nature, it is affected by external factors very quickly and may not be able to make quality decisions. The aim is to control the quality anyway. At this point, a contradiction arises, such as controlling the quality of human decisions. Even if humans are advanced beings, machine learning applications should be integrated into the functioning of these companies so that human-induced errors can be reduced to zero, quality control can be carried out in the fastest / safest way, companies can survive in today's competitive sector without financial loss and can improve themselves.

1.3. Motivation/Related Works

Proving the positive effects of decision support systems and increasing the benefits of these systems as a result of the project to be realized is the main source of motivation. According to the results of a previous survey on the benefits of decision support systems in relation to this project, the general average of the benefits of decision support systems is 4.38 out of 7. It is important to be able to increase the average of these items, which are directly related to the project to be realized, such as "Provides evidence in support of a decision or confirms existing assumptions", "Reduces decision-making costs", "Improves satisfaction with decision outcomes" in Figure 1.1 (Holsapple et al., 2005). This survey not only provides motivation for the project to be carried out in the most qualified way, but also has a great impact on the goal of adding a value to decision support systems.

Important sectoral objectives are to reduce the financial losses of companies that use this system, to increase their efficiency and income in production, to bring companies to a leading company position in a competitive environment and to increase their appreciation in the sector, when the project to be realized is completed. As a result of all these, quality control systems can increase the criteria of the concept of quality and also customer satisfaction can be maximized. When viewed from a great perspective, the spread of quality control systems to all sectors plays a role in increasing the welfare level of customers, that is, indirectly, humanity. This is a source of motivation, which is a long-term but idealistic approach.

Degree to which enterprise system enables organizations to achieve decision-support benefits

Decision-support benefit	Mean[a]
Enhances decision makers' ability to process knowledge	4.84
Improves the reliability of decision processes or outcomes	4.73
Provides evidence in support of a decision or confirms existing assumptions	4.65
Improves or sustains organizational competitiveness	4.65
Shortens the time associated with making decisions	4.59
Enhances decision makers' ability to tackle large-scale complex problems	4.59
Reduces decision-making costs	4.51
Improves coordination of tasks performed by participants jointly making a decision	4.47
Improves coordination of tasks performed by an individual decision maker	4.39
Enhances communication among participants involved in jointly making a decision	4.35
Improves coordination of tasks performed by participants making inter-related decisions	4.33
Enhances communication among participants involved in making inter-related decisions	4.27
Enhances communication among decision-making participants across organizational boundaries	4.25
Improves coordination of tasks performed by decision-making participants across organizational boundaries	4.16
Enables decentralization and employee empowerment in decision making	4.15
Encourages exploration or discovery on the part of decision makers	4.14
Stimulates new approaches to thinking about a problem or a decision context	4.14
Improves satisfaction with decision processes	4.03
Improves satisfaction with decision outcomes	3.98

[a] On a 7-point scale with end-points labeled "to a great extent" and "not at all" and a middle point labeled "moderately".

Figure 1.1 Degree to which enterprise system enables organizations to achieve decision-support benefits.

1.4. Goal/Contribution

The project aims to perform the quality control process in the production sector in the fastest and most reliable way. Today, decision support systems use machines that have developed learning skills with developing technology and methods. These machines produce meaningful and useful results.

In production, there are many factors that occur due to human or malfunction and cause stopped or will stop production. It is very easy to detect these situations as early as possible or to predict them before they happen, with correctly trained models.

This project will provide many benefits to the organization that will use the decision support system. Considering a mass production organization, anticipating potential problems will help to avoid losses in terms of both time and cost. The fact that the problems that have occurred and their causes will be determined in the fastest way will minimize time, cost and workforce losses.

As a result, the use of decision support systems in quality control processes aims to raise the concepts of time, cost and labor, which are the most important elements of production, to a level that will benefit the organization, employee and customer. The realization of this purpose depends on the most appropriate processing of the data and the extraction of the most accurate model. At the same time, it is aimed to contribute to the decision support systems that will be developed in the future, thanks to the efficient models to be created.

1.5. Project Scope

Within the scope of the project, the reports of the institutions/companies that make mass production as a result of the quality control systems will be interpreted. At the same time, a machine learning application will be made that will minimize the erroneous or incomplete decisions made by the employees. Within the scope of the data to be examined, the evaluation of the erroneous measurements resulting from the quality control will be made and as a result of this evaluation, it will be decided whether the production will be stopped or not. In this way, the financial loss that will occur due to unnecessarily stopping the production in the factories or not stopping the production, when necessary, will be prevented.

The products produced, how the quality control of these products is done, the maintenance to be done after the production is stopped, the conditions under which the production will be continued, technical measurements will not be discussed within the scope of the project. According to the result of the model trained within the scope of the project, a machine learning application will be made about when the production will be stopped or not. In this way, both the number of faulty production and the factors causing faulty production will be minimized, and it will be possible to intervene immediately when the production is faulty.

1.6. Methodology/Tools/Libraries

While developing machine learning applications, the first and most important step is to make sense of the data. In the project to be worked on, statistical values will be obtained by examining the data set consisting of real data. These data should be examined in detail, and incorrect and missing values should be detected. Later, visualizations will be added, and the data pre-processing stage will be completed using appropriate algorithms in the data set to be studied. Then, the most efficient models will be created by using Supervised/Unsupervised learning types according to the type of this processed data. In the last step, the accuracy of the model will be tested, and parameter improvements will be made if necessary, by looking at the validation methods and accuracy metrics. When you are sure that the model is ready, the deploy operation will take place.

During the project, it will be worked on Jupyter/Apachi Spark/Spyder using Python programming language. It is planned to create a model by making use of the libraries of Python that will help to work on machine learning, such as Sklearn, Seaborn, Numpy, Pandas, Mathplotlib.

CHAPTER TWO

LITERATURE REVIEW

Studies in the field of decision support systems were reviewed. As a result of the researches, it has been observed that the Linear Regression method generally achieves the best results in similar works. However, in some studies, it is observed that different algorithms achieve much more successful results than regression.

2.1. Related Works

2.1.1. Acceptance-Rejection Decision Support System in the Entry Quality Control Process

In this study, the quality control and acceptance process of the yarns belonging to a company producing fabrics has been dealt with. There are 4 different parameters taken into consideration in yarn quality. In addition to these parameters, yarn acceptance processes are carried out by taking into account the opinions of authorized personnel. In this study, it is aimed to develop a system (Eroğlu ,2019), in order to accelerate the acceptance process and to support the authorized personnel in the decision stage. In the study, in which 1231 data are analyzed, instances that give a result of 1 indicate that the thread is accepted, and those that receive a result of -1 indicate that it is rejected.

When similar studies were examined, it was seen that these studies were generally conducted using regression. This study examined whether there is a more effective algorithm by trying different algorithms. Two different methods were used in the study. The first method is classification using a hybrid genetic algorithm. In the study that the k-NN algorithm is used as a classifier, parameter selections were made with an experimental design. The accuracy of the classifications was calculated using the 10-cross validation method. When the analyzes were completed, the population number, crossover rate and mutation rate values that gave the most successful results were determined. In the second method, classification study was carried out with

Hybrid RBF. This structure consists of three layers. In the hidden layer, which is the second layer of the structure, the K computation unit was determined and it was decided which samples were the central points with the K-means algorithm. In addition to these two methods, the performances of Naive Bayes and IBk classification algorithms were obtained using Weka software.

When the average accuracy results obtained were examined, it was observed that the results of Hybrid RBF and Hybrid Genetic Algorithm were very close to each other. Also, according to the accuracy rate in the inspired method, Regression, these two algorithms performed significantly better. As a result, the performance of the hybrid genetic algorithm was examined by examining the best chromosome, and it was decided to work on the acceptance decision support system with this algorithm. (Eroğlu , 2019)

2.1.2. Using Intelligent Control Systems to Predict Textile Yarn Quality

Yarn strength is the main component in yarn quality and is very important for spinning quality. For this reason, estimating the yarn strength is of great importance. Research has been done by examining the relationship between fiber and yarn. A great deal of success has been achieved as a result of these studies. Many mathematical models have been used to make sense of and predict the relationship between yarn and fiber parameters. Linear regression is a widely used approach for yarn strength estimation. Is aimed to use new approaches to predict new smart technologies in this study (Nurwaha et al.,2012).

In this study, to examine the importance of fiber properties in yarn strength, The Multilayer Perceptron Neural Network (MLPNN), Support Vector Machines (SVM), the Radial Basis Function Network (RBFN), the General Regression Neural Network (GRNN), the Group Method of Data Handling, Polynomial Neural Network (GMDHPNN), and Gene Expression Programming (GEP) models were used. These models are often referred to as intelligent technical models. Statistical calculations, Mean Squared Error, Mean Absolute Error and Mean Absolute Percentage Error, were performed for each model to measure the prediction performance. Except for

GEP and GRNN, the 10-cross validation method was used as the validation method. The leave one out method was used in GRNN. In GEP, on the other hand, accuracy/fitness is based on good modeling.

As a result of the comparison of the analyzes, the model that gave the lowest statistical error results was determined as the Support Vector Machine. This result is followed by Linear Regression and GEP models. The highest error values were observed in the GRNN model. As a result, it was decided that the GRNN model is not suitable to be used in this data. (Nurwaha et al.,2012)

2.1.3. A Hybrid Decision Support System For Automating Decision Making In The Event Of Defects In The Era Of Zero Defect Manufacturing

An article written on a hybrid Decision Support System to automate decision-making in the event of production defect in the era of Zero Defect Manufacturing has broadly similar points to the project to be realized. In the event of a faulty production in the mass production machines within the scope of the project to be realized, a decision must be made to stop the faulty production machine. This article, on the other hand, contains the details of the project on determining whether the manufactured products are faulty.

It is not possible to produce all products completely accurately in mass production factories. Historical data is used to detect faulty products or to determine whether the machines producing these products produce faulty products. With the developing world, production is increasing exponantially today. At the same time, the control of these productions with the developing technology contributes to fast and reliable production. While both the world and technology are developing so much, zero defect manufacturing has become a very important issue (Figure 2.1). Decision support systems (DSS) provide great support to companies in the era of Zero Defect Manufacturing, such as reducing production costs by looking at historical data, predicting possible errors, predicting the errors that have occurred as soon as possible. It minimizes the errors caused by the human factor, which is one of the most important factors of faulty production.

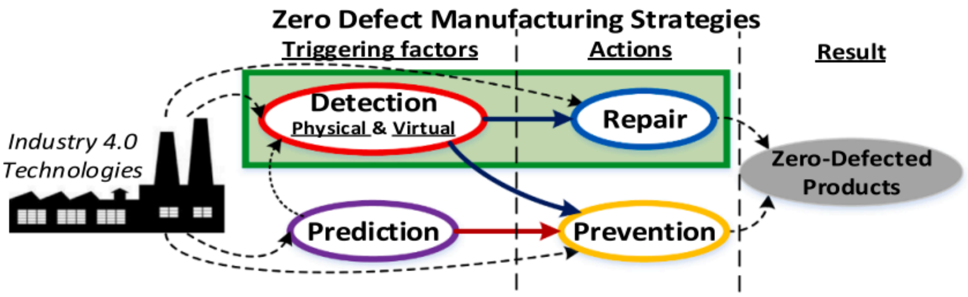

Figure 2.1 Zero defect manufacturing framework & strategies (Psarommatis, 2021)

In this article, five different types of DSS are mentioned, namely Model Driven DSSs, Data driven DSSs, Communication driven DSSs, Document Driven DSSs, Knowledge-driven DSSs, and Relative DSSs. Data driven DSSs, which works by making use of the company's historical data, is seen as the type of DSS to be used within the scope of the project to be realized. At the same time, the Document Driven DSSs type, which can quickly get data from the documents for the project to be realized, is a method that can be used for processing the documents after quality control.

As a result of the research mentioned in this article, 39% of companies spend an average of two hours a day to find solutions to problems. In this direction, it is mentioned that modeling data using ontologies provides great help in reaching relevant information. The aim of the project mentioned in the article is to automate the decision-making process involved in the production process. Relational and Triplestore databases, a blackboard architectural pattern, an Apache Cassandra to process the data, and MySQl database to save this data were used. The information was transformed into information using the Triplestore database and the project turned into a rule-based system using SPARQL queries. Real-life industrial situation is used to check the accuracy of this system.

Within the scope of the project mentioned in this article, 82.20% error correction was achieved without using DSS. This is 10.83% more than using DSS. However, this is due to the repair of all faulty parts, regardless of all the dynamics of the system in the manual decision-making process. When the project is examined with an overview, DSS provides superiority compared to manual decision making. Thanks to DSS, the production cost per unit decreased by 11.80%. The use of linear interpolation produced workable, accurate and high quality process plans. At the same time, the cost of raw materials was reduced by 15.49% thanks to DSS, resulting in earlier completion of certain orders. Consequently, the results show that the DSS generated decisions of higher quality than the manual method. **(**Psarommatis, 2021**)**

CHAPTER THREE

REQUIREMENTS / REQUIREMENT ENGINEERING

3.1. Functional Requirements

The data to be used in the decision support system should be collected accurately and completely. This data will be collected with a created Web application. Until the product goes to the quality control stage, there should be no data loss in between. Models should be trained with a smooth, logical and complete data set.

While creating a decision support system, different models should be trained and grid search should be applied when necessary. The highest accurancy should be obtained as a result of the trained models.

The recommendation obtained as a result of the decision support system should consist of the most appropriate and accurate results. The long-term quality control errors and expenses of the factory should be minimized.

Since the system will be used by more than one user, different user types should be integrated into the system. Software should be developed for different types of users who bring the products to quality control, take the products quality control and decide to stop production.

Errors originating from the system or the user should be taken into account when the system is running. These errors should be minimized. When errors occur, the system should continue to work without losing data.

3.2. Non-Functional Requirements

Since data storage has a cost, database should use with the most efficient . In this way, the annual cost required to use the system will be reduced.

A fast and secure system is developed. Training models and machine learning algorithms are time-consuming processes. The most important thing here is that the user should be able to use a fast and secure software after the deploy.

Decision support systems are systems that are just beginning to be used today. For this reason, it will take time for users who have been using the same manual system for years to get used to the system. In order to solve this problem, software should be developed with a user frindly design.

CHAPTER FOUR

DESIGN

4.1. Architectural View

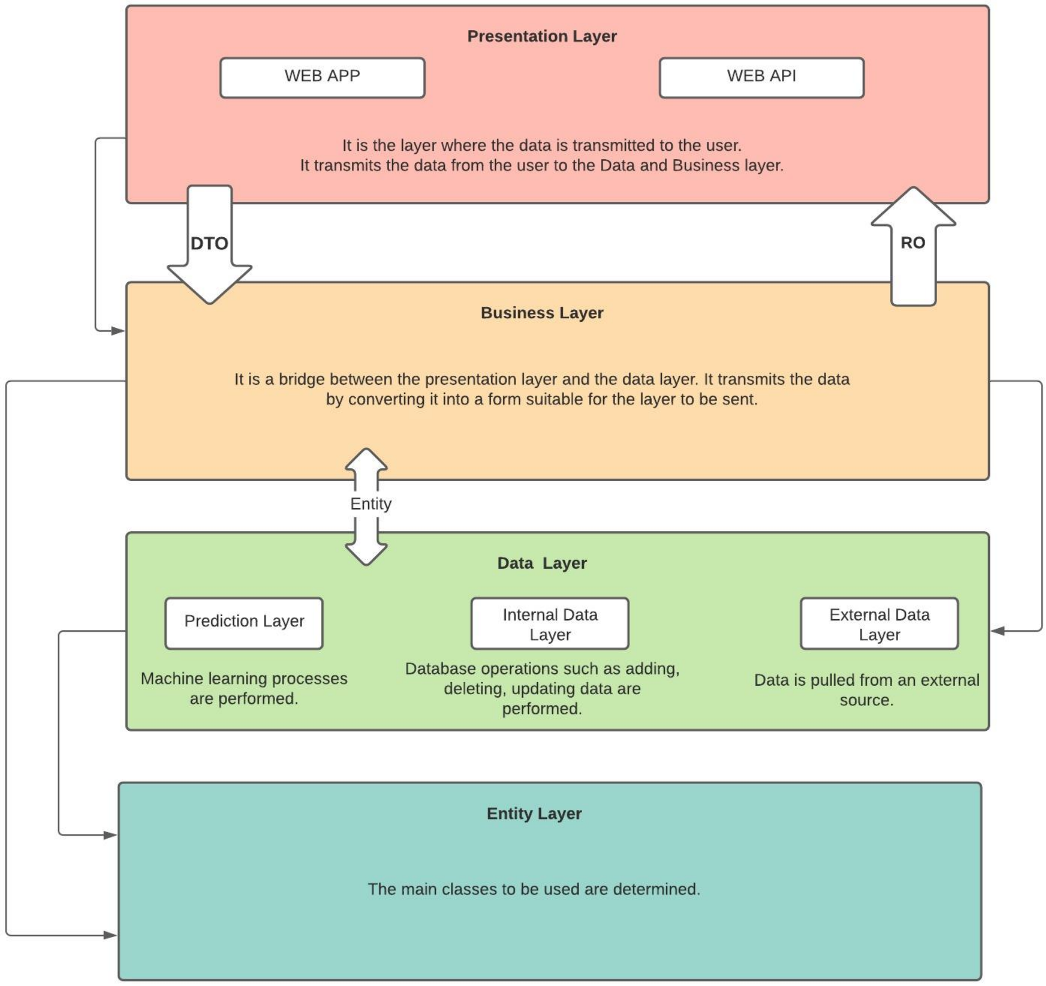

Figure 4.1 Architectural View

4.2. Database Design/ER Diagram

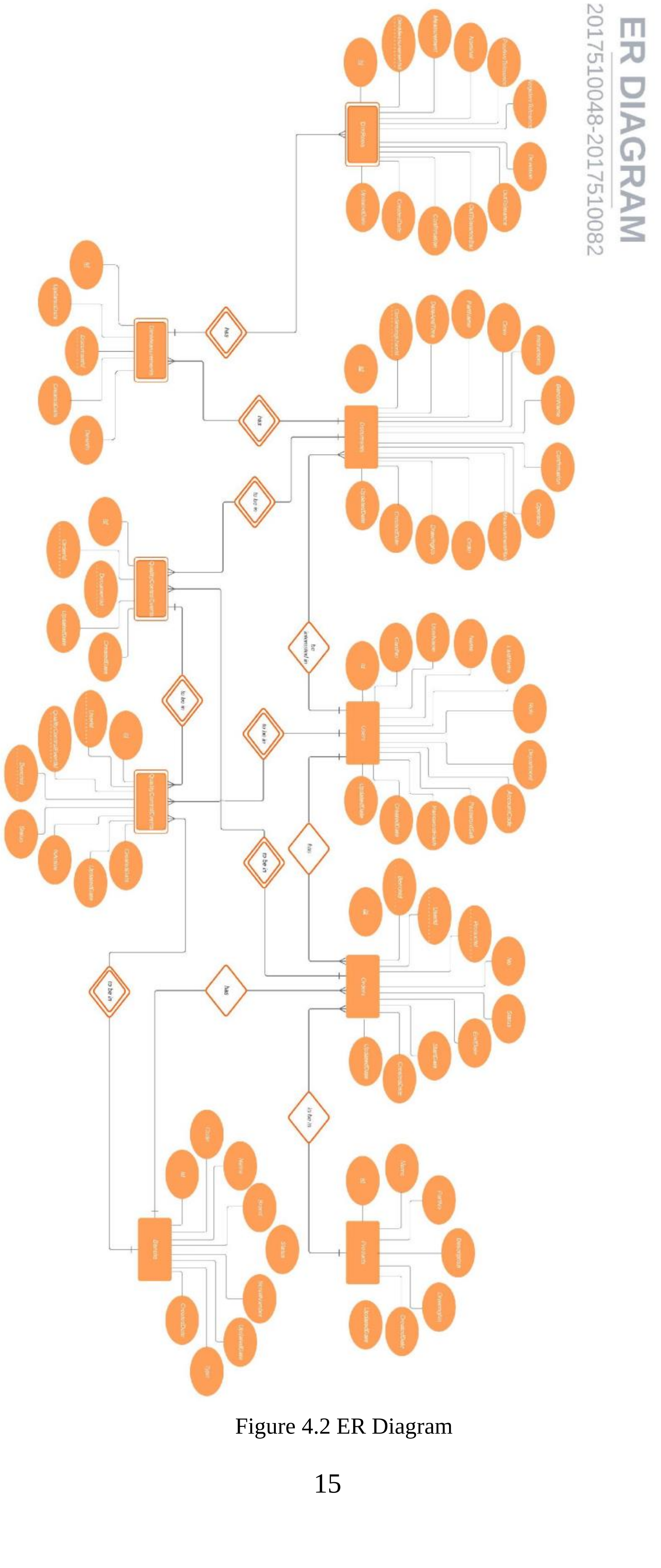

Figure 4.2 ER Diagram

4.3. UI Design

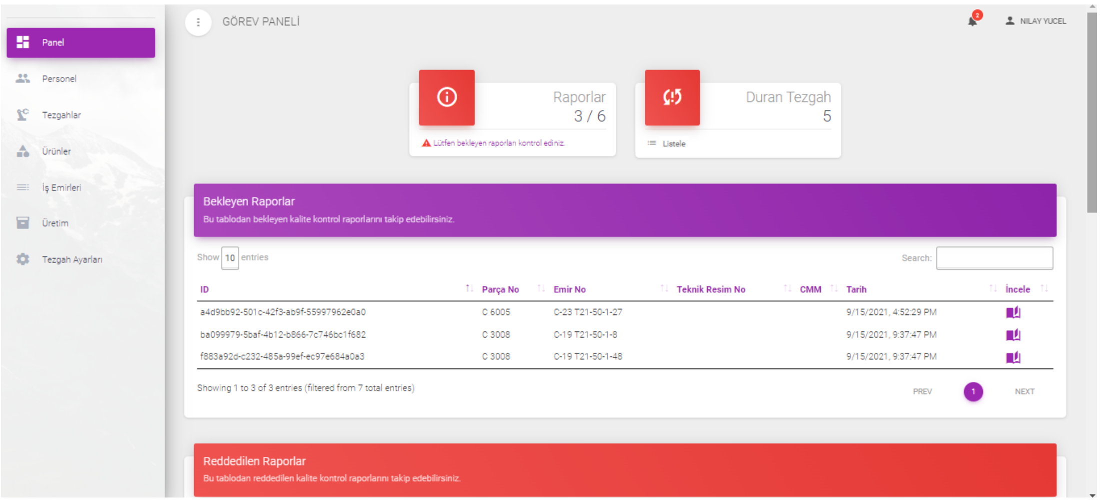

Figure 4.3 Homepage Screen

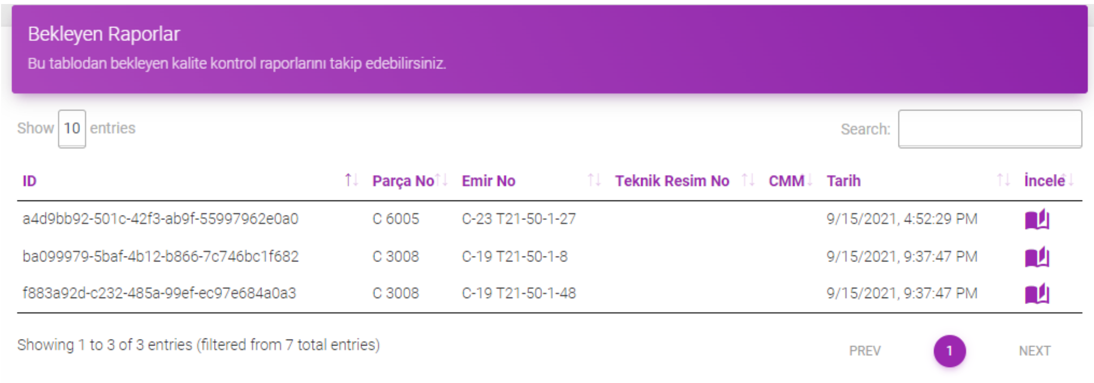

Figure 4.4 Waiting Report Table

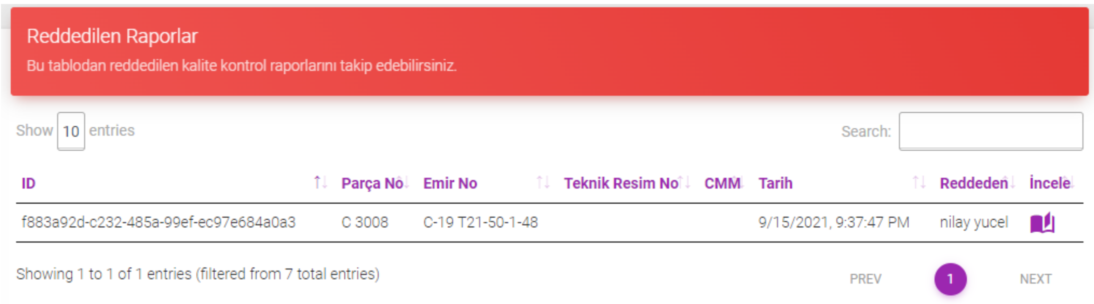

Figure 4.5 Rejection Report Table

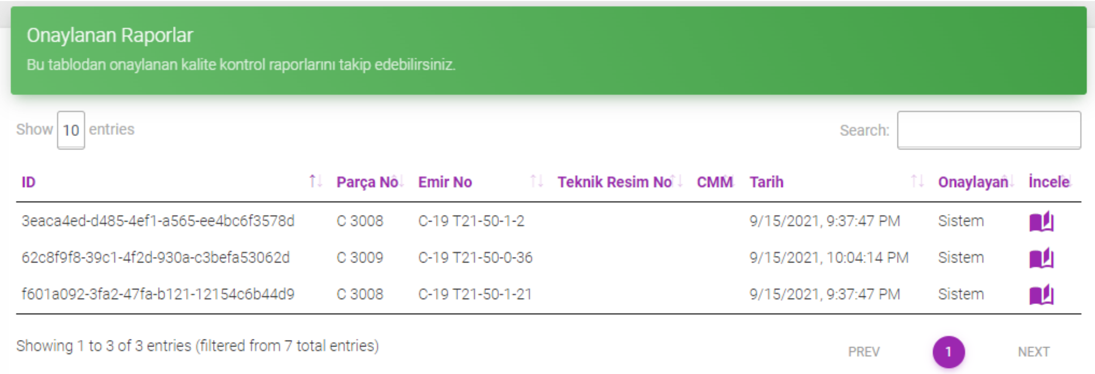

Figure 4.6 Confirmation Report Table

Figure 4.7 Manual Report Table

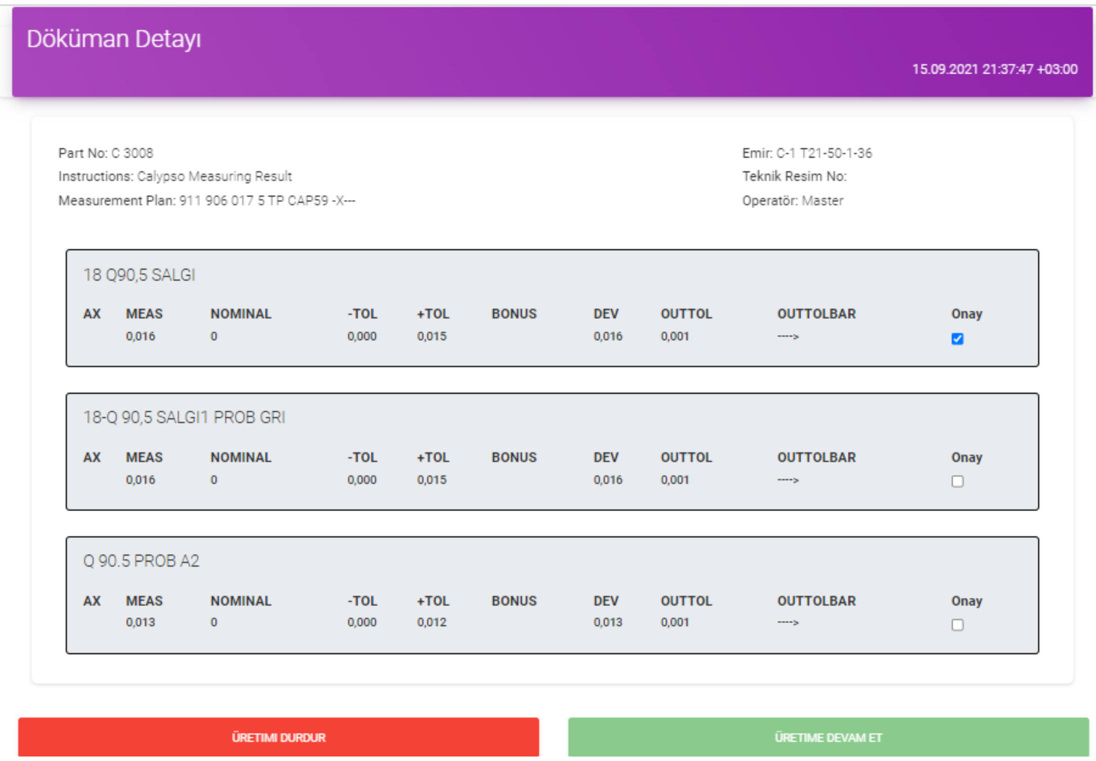

Figure 4.8 Confirmation/Rejection Table

4.4. Use Cases

Table 4.1 Use Case - Conversion of The Document

Use Case Name	**Conversion of the document containing the quality control measurements into a form that will be submitted to the production manager.**
Actors	Indoor Worker Quality Control Staff Quality Control Device Parser Application
Pre-Condition	The product must be brought to the quality control area by the internal field worker and the product must be registered as it leaves the inner field.
Flow of Events	The products, which are exited from the inner area, are taken into the quality control area by the quality control employee in order. The product is given to the quality control device. The results obtained from the device are output in excel format. Excel files containing quality control results are converted into models with the parser application.
Post-Condition	The models are saved in the database to be used in the application to be prepared for the production stop decision.
Special Requirements	Measurement files coming out of quality control machines must be in predetermined formats. If a new measuring device is brought, the file format of that machine should be added to the Parser application and a new model design should be made.

Table 4.2 Use Case - Confirmation/Rejection Decision

Use Case Name	**For each measurement of the product, the manager decides whether to stop production with the approval/rejection decision.**
Actors	Production Manager Management Panel
Pre-Condition	The production manager must be logged into the administration panel screen.
Flow of Events	The production manager monitors the products entering the quality control from the timeline screen on the management panel. When the reports that receive the quality control completion information are listed as "Pending Reports", these reports are examined. Each product has a confirm button for all dimensional measurements. If manager thinks these measurements will not cause problems in production, he approves the measurements. In cases where there are measurements that he does not approve in the report, he can stop the machine that produces this product by clicking the stop production button.
Post-Condition	Production should continue or be stopped according to the decision of the production manager.
Special Requirements	In order to access the management panel, the production manager must be logged in. Other employees cannot access this panel. The decision to stop/continue production is only possible with instantaneous reports from the machine. Reports added to the system manually can be reviewed, but they cannot affect the production decision.

4.5. Sequence Diagram

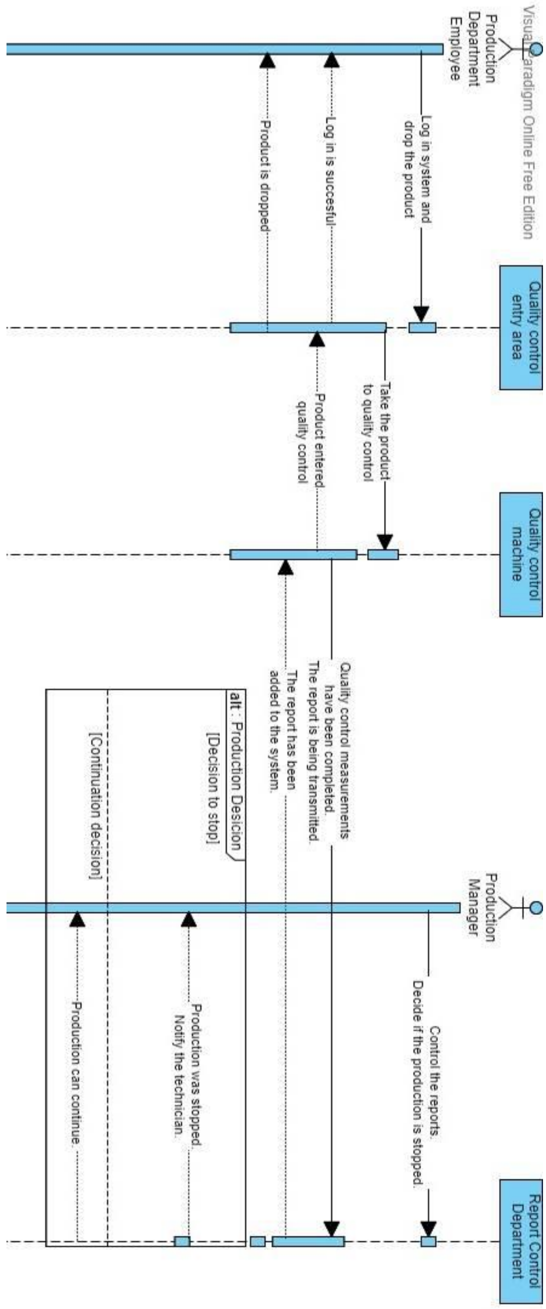

Figure 4.9 Sequence Diagram – Confirmation/Rejection Decision

4.6. Activity Diagram

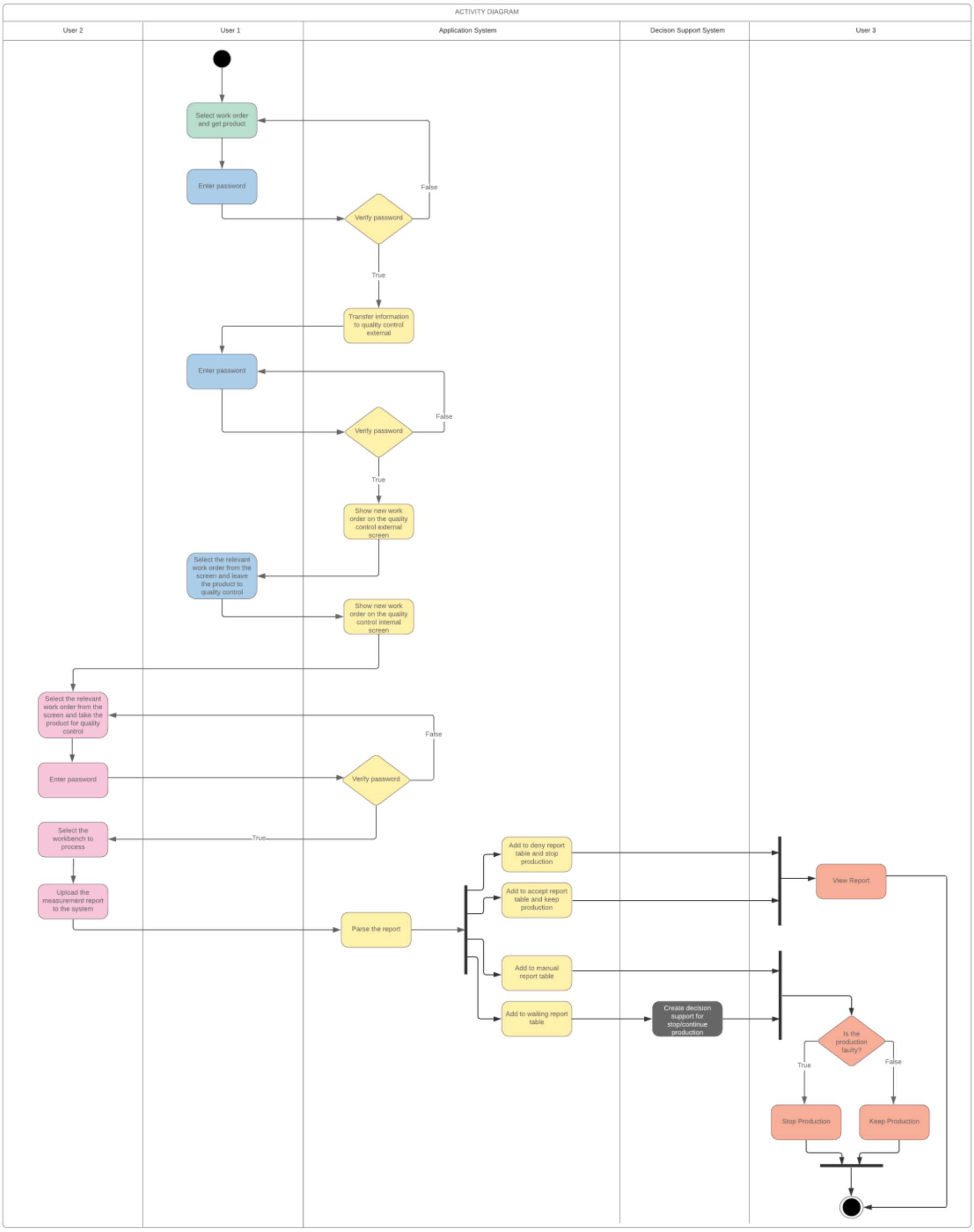

Figure 4.10 Activity Diagram

4.7. Deployment Diagram

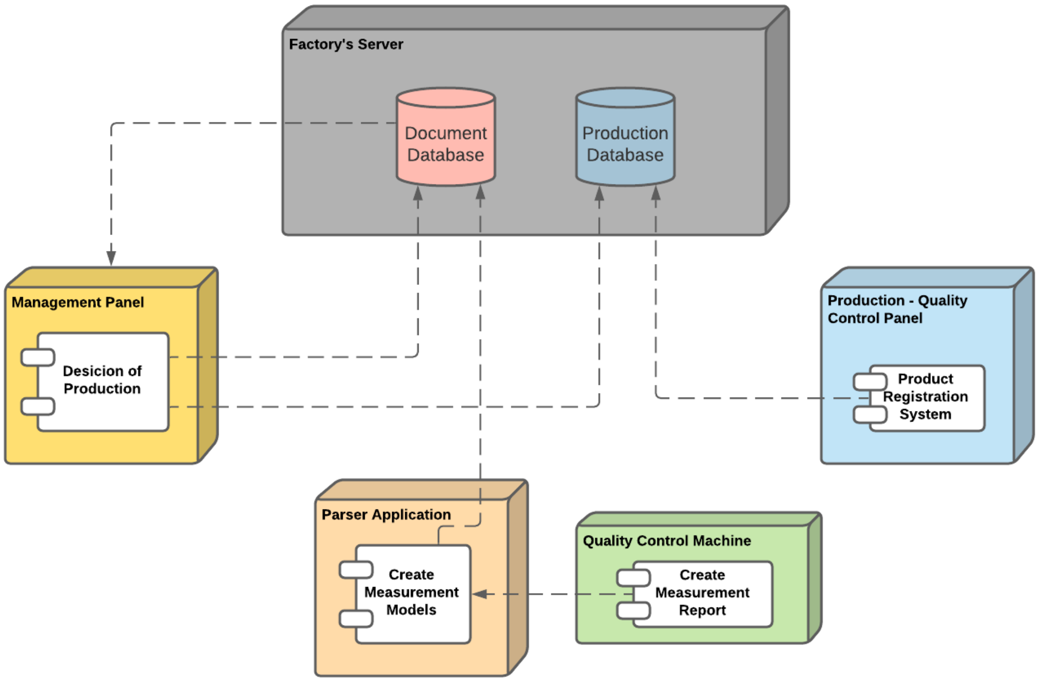

Figure 4.11 Deployment Diagram

CHAPTER FIVE

IMPLEMENTATION

5.1. Web Application

Within the scope of the project, it was decided to develop a web application in order to receive the data to be used for the decision support system. First of all, the requirements were determined. Afterwards, the coding process was started according to the designs made.

In this process, a web application was started to be developed, which was designed to be used in a factory producing machine parts, followed the quality control processes of the manufactured parts and ensured that the production continued efficiently. In the project, Visual Studio was used to develop applications, Postman to test the Back-end APIs, pgAdmin4 applications for database management, as well as .NET framework working with C# language. ASP.NET has been preferred due to its performance, cross platform, modular infrastructure, dependency injection, asynchronous operations, and easy maintenance.

It was decided to keep the reports from the quality control process in a folder so that they can be integrated into the project. A service has been written to listen to this folder and retrieve the information in the incoming files. This service, which was created using the FileSystemWatcher class, listened to the folder specified asynchronously in the back as long as the application was running and sent the quality control documents to the parsera.

A parse model was created to make the documents received with the Listener service available in API and APP sections. In this model, the information to be extracted from the files and the changes to be made in the parsing processes according to the file format were determined. Socket programming was used to be in constant communication with the file listening service. Then, the determined models were created by parsing and the data were saved in the database.

After creating the models of the project with Parser, it was necessary to create an application programming interface (API) that provides the connections in order for the project to be usable in the application. Then, the data transfer objects (DTOS) required for the models created in the parser to be used in the project were created. Functions that may be required in the application were written using data transfer objects. In order to import data into these functions, it was necessary to establish a connection with the database and at the same time send the correct queries to the database. After the appropriate queries were written for each object, the testing phase was started. GET, POST, PUT, UPDATE, PATCH and DELETE queries were tested for each object created with Postman, which is used for back-end API tests.

After the API operations were completed, the APP part of the project was started. First of all, the pages of the application and the functions and models to be used on these pages were determined. Models drawn from DTOs were processed in the Controller and sent to the View, where the pages that the user can see will be created. After creating the layout of these pages in View, .css and .js files were designed for the necessary designs and controls. Layered architecture was used in all these processes. The web application was completed after the page designs were made.

5.2. Decision Support System

The trial version was made ready for data collection and testing, after the web application was completed. For the trial version, the necessary installations were made by going to the factory. The web program has been actively used for about 3 months. After this process, the data required for the decision support system were collected in the database.

5.2.1. Data Description

A total of 3 different tables will be used for the decision support system. The first of these tables is the document table. In this table, the general features of the report that came out of the quality control are kept. Considering the possibility of generating different report types, 24 different attributes have been defined. A total of 2905 reports were added to the database. (Figure 5.1)

```
Data columns (total 25 columns):
 #   Column            Non-Null Count  Dtype
---  ------            --------------  -----
 0   Id                2905 non-null   category
 1   DateAndTime       2905 non-null   category
 2   documentType      2905 non-null   category
 3   PartName          2905 non-null   category
 4   RevisionNo        0 non-null      float64
 5   SerialNo          0 non-null      float64
 6   StatsCount        0 non-null      float64
 7   ReportNo          0 non-null      float64
 8   OperationNo       2905 non-null   object
 9   ResRevNo          2905 non-null   object
 10  Cmm               0 non-null      float64
 11  Instructions      2905 non-null   category
 12  BenchName         2905 non-null   category
 13  Confirmation      2905 non-null   category
 14  ConfirmingUserId  0 non-null      float64
 15  Operator          2905 non-null   category
 16  Measurer          0 non-null      float64
 17  Shift             0 non-null      float64
 18  StandNo           0 non-null      float64
 19  MaterialCode      0 non-null      float64
 20  MeasurementPlan   2905 non-null   category
 21  Order             2905 non-null   category
 22  DrawingNo         2905 non-null   category
 23  CreatedDate       2905 non-null   category
 24  UpdatedDate       2905 non-null   category
dtypes: category(13), float64(10), object(2)
```

Figure 5.1 Document Table

There can be more than one measurement in the Document table. Before these measurements are evaluated, the document is marked as faulty or error-free by the software according to the tolerance range. In the tables, this variable is kept under the name of confirmation. If the confirmation value is 1, it means that the document has been approved by the software. If the confirmation value is 2, it means that the document has been rejected for re-examination by the supervisors working at the factory. If the Confirmation value is 4, the document has not been rejected or approved by the software. These reports are shown directly to supervisors for review. (Figure 5.2)

```
[25]: doc["Confirmation"].value_counts().plot.barh()

[25]: <AxesSubplot:>
```

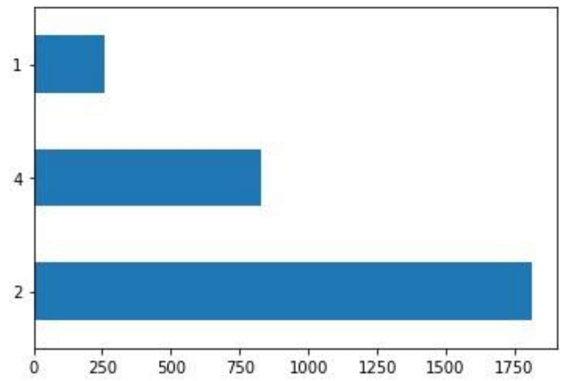

Figure 5.2 Document Table Confirmation Attribute

The information of each measurement in the Document table is stored in the DimMeasurement table (Figure 5.3). This table connects the values obtained from each measurement and the Document table.

```
[39]: dimInfo.shape

[39]: (7259, 5)

[36]: dimInfo.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 7259 entries, 0 to 7258
Data columns (total 5 columns):
 #   Column       Non-Null Count  Dtype
---  ------       --------------  -----
 0   Id           7259 non-null   object
 1   DocumentId   7259 non-null   object
 2   DimInfo      7259 non-null   object
 3   CreatedDate  7259 non-null   object
 4   UpdatedDate  7259 non-null   object
dtypes: object(5)
memory usage: 283.7+ KB
```

Figure 5.3 DimMeasurement Table

There are different measurements included in each measurement. Incorrect ones from these measurements are stored in the DimRow table (Figure 5.4). Each measurement is evaluated separately. This evaluation is made by the supervisors working in the factory. The attributes to be recorded in the database were determined by taking into account the different measurement variables that may come from different quality control devices.

```
Data columns (total 43 columns):
 #   Column               Non-Null Count  Dtype
---  ------               --------------  -----
 0   Id                   7259 non-null   object
 1   DimMeasurementId     7259 non-null   object
 2   Axis                 0 non-null      float64
 3   Measurement          7259 non-null   float64
 4   Nominal              7259 non-null   float64
 5   PositiveTolerance    7259 non-null   float64
 6   NegativeTolerance    7259 non-null   float64
 7   Bonus                0 non-null      float64
 8   Deviation            7259 non-null   float64
 9   OutTolerance         7259 non-null   float64
 10  OutToleranceBar      7259 non-null   object
 11  PlanId               0 non-null      float64
 12  PartNo               0 non-null      float64
 13  PlanProductId        0 non-null      float64
 14  Type                 0 non-null      float64
 15  IdSymbol             0 non-null      float64
 16  WarningLimitCf       0 non-null      float64
 17  FeatureId            0 non-null      float64
 18  FeatureSigma         0 non-null      float64
 19  Comment              0 non-null      float64
 20  Link                 0 non-null      float64
 21  LinkMode             0 non-null      float64
 22  Mmc                  0 non-null      float64
 23  UserUpperTol         0 non-null      float64
 24  UserLowerTol         0 non-null      float64
 25  Fftphi               0 non-null      float64
 26  FftphiUnit           0 non-null      float64
 27  ZoneRoundnessAngle   0 non-null      float64
 28  GroupName            0 non-null      float64
 29  GroupName2           0 non-null      float64
 30  DatumAid             0 non-null      float64
 31  DatumBid             0 non-null      float64
 32  DatumCid             0 non-null      float64
 33  NatUpperTolId        0 non-null      float64
 34  NatLowerTolId        0 non-null      float64
 35  DecimalPlaces        0 non-null      float64
 36  FeaturePosX          0 non-null      float64
 37  FeaturePosY          0 non-null      float64
 38  FeaturePosZ          0 non-null      float64
 39  Unit                 0 non-null      float64
 40  Confirmation         7259 non-null   category
 41  CreatedDate          7259 non-null   object
 42  UpdatedDate          7259 non-null   object
dtypes: category(1), float64(37), object(5)
```

Figure 5.4 DimRow Table

The approval and rejection status of each measurement in the DimRow table is kept in the confirmation variable (Figure 5.5). A confirmation value of 1 indicates that the error in the relevant measurement is insignificant and production can continue. If the confirmation value is 2, it means that the relevant measurement is rejected by the supervisor. If any of these measurements is rejected, production is stopped on the relevant bench.

```
[67]: rows["Confirmation"].value_counts().plot.barh()

[67]: <AxesSubplot:>
```

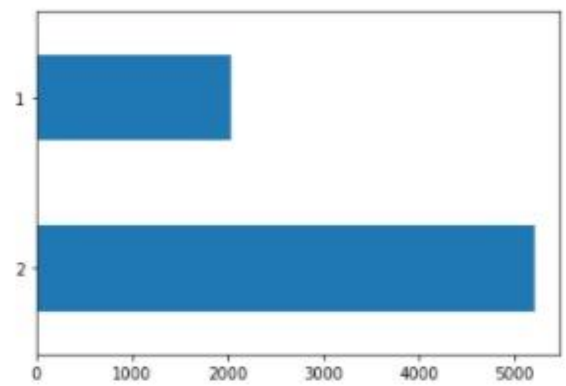

Figure 5.5 DimRow Table Confirmation Attribute

It was decided to combine these three tables in order to examine the associations of the information contained in the data and to carry out preprocessing on them. First of all, the unnecessary/empty columns in the tables were cleaned and then the merge operation was performed.

Empty, unique, repetitive data in the tables were removed and matched over ids. The new data set (Figure 5.6), consisting of a total of 7259 rows and 17 columns, was brought ready for necessary examinations and preprocessing.

```
Data columns (total 17 columns):
 #   Column                Non-Null Count  Dtype
---  ------                --------------  -----
 0   DocumentId            7259 non-null   object
 1   DateAndTime           7259 non-null   category
 2   PartName              7259 non-null   category
 3   DocumentConfirmation  7259 non-null   category
 4   MeasurementPlan       7259 non-null   category
 5   Order                 7259 non-null   category
 6   DrawingNo             7259 non-null   category
 7   DimId                 7259 non-null   object
 8   DimInfo               7259 non-null   category
 9   RowId                 7259 non-null   object
 10  Measurement           7259 non-null   float64
 11  Nominal               7259 non-null   float64
 12  PositiveTolerance     7259 non-null   float64
 13  NegativeTolerance     7259 non-null   float64
 14  Deviation             7259 non-null   float64
 15  OutTolerance          7259 non-null   float64
 16  RowConfirmation       7259 non-null   category
dtypes: category(8), float64(6), object(3)
```

Figure 5.6 Merged Table / Data Set

5.2.2. Data Preprocessing

5.2.2.1. Dimensionality Reduction

The current data set was examined and it was decided to remove the retained ids to consolidate the tables. In addition, time information has been removed since time series work cannot be done.

As a result of the dimensionality reduction, a total of 12 columns remained in the data set. These columns will be examined in detail and what can be done will be observed.

5.2.2.2. *Derived Data*

It has been observed that some columns contain two different information together. Since a wide variety of categorical results occur when kept together, it was thought that evaluating them separately would give better generalization results.

The PartName column (Figure 5.7) contains two different values. The letters represent the shift and the number represents the team number. In addition, lines containing extra characters cause an unnecessary increase in the number of value_counts, since they are user-entered data. This column, where these two properties are together, contains a total of 96 different values. By breaking it down, two different features named “Shift” and “TeamNo” will be added to the table.

```
df.PartName.value_counts()

C 6001       467
B 5004       450
B 5001       402
B 5005       335
B 5002       311
            ...
B  5003        1
B - 5005       1
A 1008         1
A 1005         1
B - 5007       1
Name: PartName, Length: 96,
```

```
[56]: tdf.Shift.value_counts()

[56]: B    3194
      A    2058
      C    2007
      Name: Shift, dtype: int64

[57]: tdf.TeamNo.value_counts()

[57]: 6001    471
      5004    461
      5001    404
      5005    355
      5002    318
             ...
      4016      2
      5018      2
      5019      2
      1008      1
      1005      1
      Name: TeamNo, Length: 66, dtype: int64
```

Figure 5.7 Derived Features I - PartName column

The Order column (Figure 5.8) also contains more than one feature. A portion of these values (like C133) represents the machine to be worked on. The section starting with T21/T22 represents the year, week, which arm of the machine will work and which part will be produced. The time values and the number of pieces in this section are unnecessary information for the study. For this reason, keeping the information on which arm of the machine will be produced will be sufficient for the study.

```
df.Order.value_counts()

C 133 T21-46-1-9 TKR        57
C 133 T21-46-1-9            56
C 133 T21-46-1-29           55
C 112 T22-4-3-117 ONAY 1    46
C 112 T22-4-4-117 ONAY 1    44
                            ..
C 76 T21-42-2-1              1
C 76 T21-42-1-62             1
C 57 T22-7-2-116             1
C 76 T21-42-1-29             1
C 76 T21-42-1-72             1
Name: Order, Length: 1785, dtype: int64
```

```
[60]: tdf.BenchNo.value_counts().head(5)

[60]: C133    1696
      C76     1408
      C57     1025
      C128    1020
      C123     931
      Name: BenchNo, dtype: int64
```

```
[55]: tdf.PartofMachine.value_counts()

[55]: 2    2124
      1    2039
      3    1600
      4    1496
      Name: PartofMachine, dtype: int64
```

Figure 5.8 Derived Features II - Order Columns

5.2.2.3. *Data Hiding*

When the distribution of categorical data is examined, the excess of some values suppresses the other values.

MeasurementPlan values are hidden for data privacy. As it can be seen in Figure 5.9 and Figure 5.10, the data with a small number were combined in the OTHER category without changing the distribution.

```
tdf.MeasurementPlan.value_counts()

    5891
     979
     246
      36
      36
      29
      24
       9
       7
       2
```

```
[50]: tdf.MeasurementPlan.value_counts()

[50]: Measurement Plan X    5891
      Measurement Plan Z     979
      Measurement Plan Y     246
      OTHERS                 143
      Name: MeasurementPlan, dtype: int64
```

Figure 5.9 MeasurementPlan value_counts()

```
[51]: tdf.DrawingNo.value_counts()

[51]: 911 9060 175         5920
      412 3520 165         1015
      912 5180 145          246
      411 1540 115           60
      K3659390 OPR REV1       9
      912 5100 105            7
      001 0521 00             2
      Name: DrawingNo, dtype: int64
```

```
[54]: tdf.DrawingNo.value_counts()

[54]: 911 9060 175    5920
      412 3520 165    1015
      912 5180 145     246
      OTHERS            78
      Name: DrawingNo, dtype: int64
```

Figure 5.10 DrawingNo value_counts()

At the same time, the DimInfo column contains 266 different values. Its distribution is also quite uneven. In order to eliminate this irregularity, values with similar properties were combined. At the same time, those with a low distribution were combined with the "OTHER" category. As a result of this process, the number of DimInfo values decreased to 58.

5.2.2.4. Outlier Detection

Before performing outlier analysis, outliers were observed with the help of boxplots (Figure 5.11).

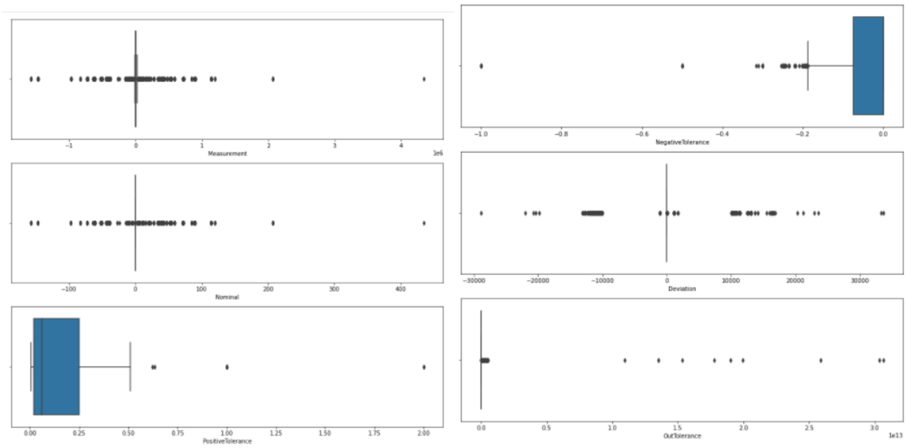

Figure 5.11 Box Plots

After examining the boxplots, outlier analysis was performed with the IQR method. Then, outlier values are filled with the KNN algorithm. In this case, the change in statistical values (Figure 5.12 , Figure 5.13) is quite evident.

	count	mean	std	min	25%	50%	75%	max
Measurement	7259.0	8.118432e+03	3.764781e+05	-1.575447e+06	0.015527	0.094012	10322.539000	4.342325e+06
Nominal	7259.0	8.804946e-01	3.998015e+01	-1.574000e+02	0.000000	0.000000	0.000000	4.350000e+02
PositiveTolerance	7259.0	1.401914e-01	1.623393e-01	5.000000e-03	0.018000	0.060000	0.250000	2.000000e+00
NegativeTolerance	7259.0	-6.629957e-02	1.442132e-01	-1.000000e+00	-0.075000	0.000000	0.000000	0.000000e+00
Deviation	7259.0	-5.521102e+00	1.819369e+03	-2.887719e+04	-0.018209	0.039901	0.271569	3.366888e+04
OutTolerance	7259.0	2.932739e+10	7.720649e+11	1.023000e-04	0.003402	0.014115	0.052612	3.066888e+13

Figure 5.12 Statistical Description Before KNN

	count	mean	std	min	25%	50%	75%	max
Measurement	7259.0	290.854461	4504.066942	-14110.008000	0.013261	0.060001	0.416924	25382.745000
Nominal	7259.0	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
PositiveTolerance	7259.0	0.127873	0.121517	0.005000	0.018000	0.060000	0.236577	0.508794
NegativeTolerance	7259.0	-0.032456	0.049875	-0.187074	-0.050000	0.000000	0.000000	0.000000
Deviation	7259.0	0.103480	0.202893	-0.440533	0.010718	0.050848	0.266998	0.706065
OutTolerance	7259.0	0.029419	0.036503	0.000102	0.003181	0.012831	0.043166	0.126254

Figure 5.13 Statistical Description After KNN

5.2.2.5. Normalization

At the end of all operations, it was ensured that there was no contradictory, empty and unnecessary data. In order to equalize the weights of the numeric values, normalization was performed with the min-max standardization method and all values were arranged to be in the range of 0-1.

It became undefined during the normalization process because the Nominal column contains too many 0 and near 0 values (Figure 5.14). For this reason, it was decided to remove this column.

	count	mean	std	min	25%	50%	75%	max
Measurement	7259.0	0.364646	0.114048	0.0	0.357281	0.357282	0.357291	1.0
Nominal	0.0	NaN	NaN	NaN	NaN	NaN	NaN	NaN
PositiveTolerance	7259.0	0.243895	0.241203	0.0	0.025804	0.109172	0.459666	1.0
NegativeTolerance	7259.0	0.826507	0.266604	0.0	0.732726	1.000000	1.000000	1.0
Deviation	7259.0	0.474458	0.176952	0.0	0.393557	0.428555	0.617070	1.0
OutTolerance	7259.0	0.232391	0.289359	0.0	0.024405	0.100900	0.341366	1.0

Figure 5.14 Statistical Description

5.2.2.6. Transformation

At the end of the preprocessing processes, the data set consisting of 7123 rows and 13 columns is almost ready for modeling. In order to work more easily with algorithms, it was decided that all features should be of the same type. For this reason, one hot encoding process was applied for all the features of the object and category type, except the target column of the data set. As a result of the operation, categorical variables are converted to int type.

5.2.2.7. Over Sampling

As a result of the target column scatterplot examinations, it was observed that the rejected cases were 2.5 times more than the accepted ones. In this case, it was decided to apply an over sampling or under sampling process so that the weight does not shift to a single target.

As an oversampling technique, SMOTE process was applied and synthetic data was produced so that the target value distribution was equal (Figure 5.15).

```
Before SMOTE : Counter({'Rejected': 4107, 'Confirmed': 1591})
After SMOTE : Counter({'Rejected': 4107, 'Confirmed': 4107})
```

Figure 5.15 Over Sampling

5.2.3. Algorithms for Modelling

Machine learning models are used to draw meaningful conclusions from the data collected while developing decision support systems. After the preprocessing is done, the data is brought into a format that the models can understand and machine learning models can be developed on it. Within the scope of this project, a total of 6 different algorithms have been used, including k-Nearest Neighbors, Support Vector Machine, Naive Bayes, Random Forest, Gradient Boosting Machine and XGBoost algorithms.

5.2.3.1. k-Nearest Neighbor

It is often used to solve Classification problems. In the basic logic, as shown in the Figure 5.16, the k value (tuple) closest to the value to be estimated is checked. Estimation is made according to the classes to which the k closest values belong. To find the closest value, distance metrics such as Euclidean and Manhattan Jacard are used.

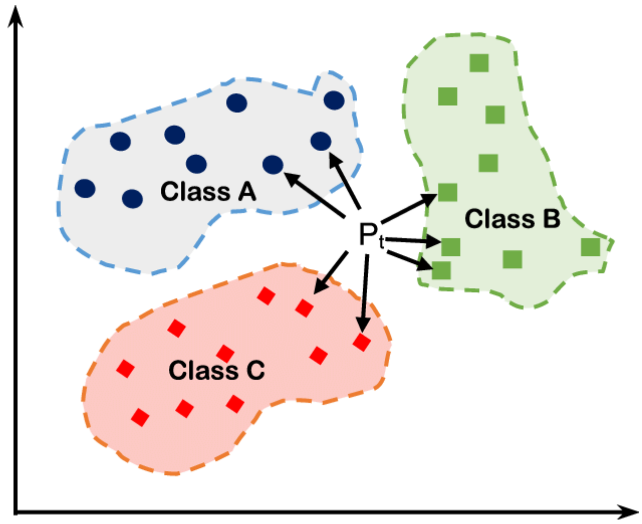

Figure 5.16 k-Nearest Neighbor (Atallah, 2019)

5.2.3.2. *Support Vector Classifier*

It is often used to solve Classification problems. In its basic logic, the data set is divided into classes with the help of a hyperplane as shown in the Figure 5.17. In addition to this hyperplane, support vectors defined at a certain distance are also used. Whichever of the dividing regions the value to be estimated enters, the class label of the value becomes the label of that region.

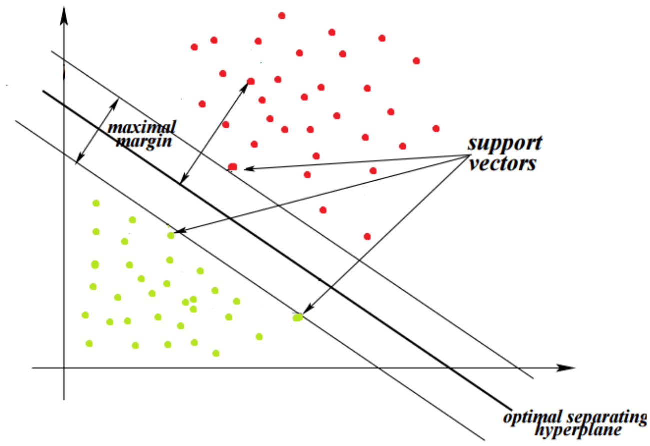

Figure 5.17 Support Vector Machine (Kumar, 2021)

5.2.3.3. *Naïve Bayes*

It is used to solve classification problems. In the basic logic, the probability of the variables to be estimated and the classes to which they belong is calculated separately. Then, the probabilities of the classes to which the value to be estimated may belong are found by the formula shown in the Figure 5.18, and the class with the highest probability value is defined as the class of the value to be estimated.

Probability of B occurring given evidence A has already occurred

Probability of A occurring

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$

Probability of A occurring given evidence B has already occurred

Probability of B occurring

Figure 5.18 Naïve Bayes (Akdağlı,2021)

5.2.3.4. Random Forest

It is used to solve both classification problems and regression problems. In its basic logic, there is a tree structure. As shown in the Figure 5.19, different subsets are created within the data set and a separate tree is created for each separated subset. At the same time, it has created a solution to the overfitting problem that occurs in tree-based progressive models. The class that emerges the most among the results is determined as the class of the value to be estimated.

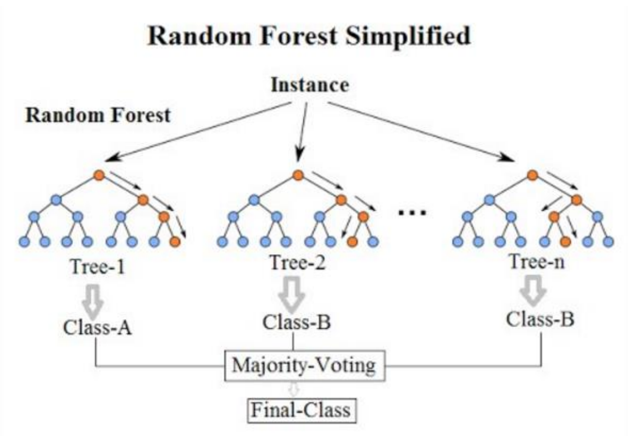

Figure 5.19 Random Forest (Wikipedia, 2021)

5.2.3.5. *Gradient Boosting Machine*

It is a type of boosting algorithm. As shown in the Figure 5.20, the general working principle is based on trying to correctly estimate the incorrectly predicted values each time with different iterations.

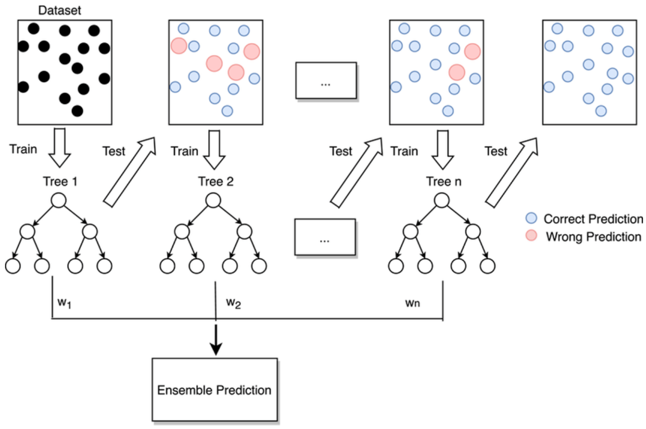

Figure 5.20 GBM (Zhang, 2021)

5.2.3.6. *Light GBM*

The working logic is the same as the Gradient Boosting Machine. The improved version of the GBM algorithm is the XGBoost algorithm.

CHAPTER SIX

TEST/EXPERIMENTS

Modeling algorithms mentioned in Chapter 5.2.3 were run and models were obtained. While modeling, the reliability of the accuracy was also observed by using grid search and some measurement matrices.

6.1. Grid Search

The algorithms mentioned in Chapter 5.2.3 work with more than one parameter. These parameters can have different values. According to the data set used, the most suitable ones from these parameters should be selected and the model is trained. In this way, the best accuracy can be achieved.

However, it is impossible to predict from the beginning which values of the parameters will give the best results for which data set. For this reason, grid search is used. Grid search looks at which parameters give better results for that data set and shows us the best parameter values to use Table 6.1 shows the best parameter values obtained as a result of grid search.

Table 6.1 Parameters Table

Algorithms	**Best Parameters**	
KNN	*n_neighbors*: 1	
Random Forest	*bootstap*: True *max_depth*: 10 *min_sample_*leaf:1 *n_estimaors*: 673	*criterion*: entropy *max_features*: auto *min_samples_split*: 2
SVC	*C*: 10 *gamma:* 1	*degree*: 1 *kernel*: poly
GBM	*learning_rate*: 0.01 *n_estimators*: 500	*max_depth*: 5 *subsample*: 0.5

XGBoost	*n_estimaors 1000* *max_depth*: 7 *min_samples_split*: 2	*subsample:0.5* *learning_rate*: 0.01
Naive Bayes	*var_smoothing*: 0.0001	

6.2. Comparison

As a result of the researches and studies, classifiers were made with 6 machine learning algorithms. As seen in Table 6.2, accuracy and measurement values were recorded before and after Grid Search.

In line with the studies, the XGBoost algorithm has achieved high results in both accuracy and measurement values.

In line with the studies, better parameters and different algorithms will be tried and studies will be carried out to increase the accuracy value, which is currently high.

Table 6.2 Comparison Table

Algorithm	Before Grid Search				After Grid Search			
	Accuracy	Precision	Recall	F1	Accuracy	Precision	Recall	F1
XGBoost	92%	92%	92%	92%	92%	93%	92%	92%
SVC	88%	89%	88%	88%	92%	92%	%92	92%
GBM	91%	92%	91%	91%	91%	92%	91%	91%
RF	91%	92%	92%	92%	90%	92%	91%	91%
KNN	80%	82%	80%	80%	85%	86%	86%	86%
Naïve Bayes	70%	%78	%70	%68	78%	82%	78%	78%

CHAPTER SEVEN

CONCLUSION

Customer satisfaction, whose importance is increasing day by day, has also carried the concept of quality to higher levels. The production sector is one of the areas that have to progress based on this concept. Measuring and evaluating quality is a very important step. A mistake made in this step can lead to huge financial losses. Quality control systems have been created to prevent these losses. Although systems consist of machines, a human decision/opinion is needed in the last step. However, since human decisions can also vary, machine models are developed that can support humans and make more precise and accurate decisions at this stage.

The project aims to obtain successful models developed with machine learning algorithms to support the decision made during the quality control phase. In this direction, the process of processing and modeling real data taken from a factory was carried out. First of all, it was decided to process the data in a web application so that it can be obtained more reliably and regularly. These processed data are recorded together with the decisions made by the human being. With this system, which was used for about 3 months, quality control reports of 2905 parts were obtained. Separate tolerance ranges and separate decisions are available for each size of these parts. The decision made for a total of 7259 different dimensions was recorded.

This obtained data set also includes the target value (production decision). For this reason, it was decided to make it suitable for classification studies. The data set was run with 6 different algorithms after appropriate preprocessing. Different results have been obtained, ranging from 70% to 92%. Tunning process was performed to measure the reliability of these results. Among the given parameters, the best ones were selected and new models were created. Recent results range from 78% to 92%.

When the results were examined, it was determined that the model that reached both the highest and the most reliable results was created with the XGBoost algorithm. This model can be added to the created web application and used to assist human decision.

REFERENCES

Akdağlı, E. (2021) Makine Öğrenmesinde Naive Bayes Yöntemi. Medium. https://ece-akdagli.medium.com/makine-öğrenmesinde-naive-bayes-yöntemi-98d28cfe46de

Atallah, D. M., Badawy, M., & El-Sayed, A. (2019). Intelligent feature selection with modified K-nearest neighbor for kidney transplantation prediction. SN Applied Sciences, 1(10), 1-17.

EROĞLU D. Y. (2019). Giriş Kalite Kontrol Sürecinde Kabul-Ret Kararı Destek Sistemi. Gazi Üniversitesi Fen Bilimleri Dergisi Part C: Tasarım ve Teknoloji, 7(3), 576 - 590. Doi: 10.29109/gujsc.549890

Holsapple, C. W., & Sena, M. P. (2005). ERP plans and decision-support benefits. Decision Support Systems, 38(4), 575-590.

Kumar, N. (2021) Introduction to Support Vector Machines (SVMs). MarkTechPost. https://www.marktechpost.com/2021/03/25/introduction-to-support-vector-machines-svms/

Nurwaha, D., & Wang, X. H. (2012). Using intelligent control systems to predict textile yarn quality. Fibres & Textiles in Eastern Europe, 20(1), 23-27.

Psarommatis, F., & Kiritsis, D. (2021). A hybrid Decision Support System for automating decision making in the event of defects in the era of Zero Defect Manufacturing. Journal of Industrial Information Integration, 100263.

Zhang, T., Lin, W., Vogelmann, A. M., Zhang, M., Xie, S., Qin, Y., & Golaz, J. C. (2021). Improving convection trigger functions in deep convective parameterization schemes using machine learning. Journal of Advances in Modeling Earth Systems, 13(5), e2020MS002365.

Wikipedia. (2022, May) Random forest.
https://en.wikipedia.org/wiki/Random_forest