Development of Machine Learning based User Friendly Control Charts

Chetan Paliwal, Sidharth Goyal, Pradhuman Jain, Dr Suman Kant, Dr Mandeep Dhanda

Department of Production and Industrial Engineering, Punjab Engineering College, Chandigarh, India

Abstract

This research paper consists of the development of a user-friendly software application that makes control charts and uses machine learning for trend analysis and detects the error. Control charts are a popular tool in statistical process control, used to monitor and analyze process performance. The proposed work uses machine learning algorithms to analyze data and automatically identify patterns and trends, making it easier for users to make accurate and effective control charts. This paper presents the methodology used to develop the application, including the selection of machine learning algorithms and the training and testing of the model. The results of the study demonstrate the effectiveness of the user friendly interface in making control charts that accurately identify process variations and trends. The paper concludes with a discussion of the potential applications of technology in industry and the potential for future research in this area.

Keywords – Machine Learning, Control Charts, Quality Control, User Interface

1. Introduction

Importance of Quality

Quality refers to meeting or surpassing customer expectations, regulatory requirements, and industry standards in terms of products, services, or processes. It can be defined differently based on the context and industry, but generally refers to the ability to fulfill the intended purpose or function.

In manufacturing, quality means meeting design specifications, reliable performance, and absence of defects. In services, quality means delivering customer needs on time and without errors. Quality can be measured through inspections, testing, customer feedback, or statistical analysis. Quality management systems like ISO 9001 establish processes to ensure quality throughout production or service delivery.

The manufacturing company applies several techniques for quality control (QC) to increase the quality of the progression by decreasing its variability [1].

Quality is vital in industries due to:

1. Customer satisfaction: High-quality products build customer loyalty and satisfaction, leading to repeat business and referrals.

2. Brand reputation: Providing quality products builds a positive brand reputation, attracting more customers.

3. Cost savings: Quality management systems reduce waste, rework, and defects, resulting in cost savings and increased profitability.

4. Regulatory compliance: Meeting quality regulations and standards avoids legal and financial penalties.

5. Employee satisfaction: Quality work environments with tools and equipment enhance employee productivity and motivation.

In conclusion, quality is crucial in industries as it enhances customer satisfaction, brand reputation, cost savings, regulatory compliance, and employee satisfaction. Prioritizing quality improves the bottom line and fosters a sustainable business.

Quality Control Tools

Quality control tools are techniques and methods used in industries to ensure products and services meet customer expectations, regulatory requirements, and industry standards. These tools identify and address defects, errors, and other quality issues. Common quality control tools include flowcharts, checklists, Pareto charts, scatter diagrams, statistical process control (SPC), and control charts. Flowcharts identify process issues, checklists ensure all steps are completed, Pareto charts prioritize significant issues, scatter diagrams reveal data patterns, SPC monitors and controls processes, and control charts track process performance over time.

By utilizing these tools, companies can improve product or service quality by addressing potential issues and taking corrective actions. The techniques for quality control can be categorized into the basic, intermediate, and advanced level, although there is no unanimity amongst researchers about it. For example, Xie et al. [2] deliberate the DoE as an intermediary level technique while Antony et al. [3] classified that technique as advanced.

An essential knowledge of the control charts is to analyze the hypothesis that there are few common reasons of alternative versus variability, that there are exceptional causes by continuously observing the process. The manufacturing company could avoid defect items to be administered in the subsequent stage and to take instant corrective action while the process exists to be out of control [4]

Control Charts

Control charts are a statistical tool used in quality control to monitor and control a process over time. Control charts are used to identify when a process is performing within acceptable limits and when it is out of control. Control charts are useful in quality control because they help to identify potential issues before they result in defects or errors.

The basic idea behind control charts is to plot data points on a chart over time and analyze the pattern of the data points. Control charts typically have a centerline, which represents the mean or average value of the data, and upper and lower control limits, which represent the range of variation that is acceptable for the process.

There are different types of control charts used in quality control, depending on the type of data being analyzed. Some of the most common control charts include:

1. X-bar and R Chart: This chart is used for variables data, such as measurements or weights. The X-bar chart shows the average value of the data over time, while the R chart shows the range of variation in the data.

2. Individual and Moving Range (I-MR) Chart: This chart is used for variables data, such as measurements or weights, when the sample size is small. The I chart shows the individual data points over time, while the MR chart shows the range of variation in the data.

3. P Chart: This chart is used for attribute data, such as the number of defects or errors. The P chart shows the proportion of defective items over time.

4. C Chart: This chart is used for attribute data, such as the number of defects or errors, when the sample size is constant. The C chart shows the count of defects or errors over time.

There are control charts for special uses in literature which can be listed as cumulative sum control charts, attributes control charts, modified and acceptance control charts, group control charts, chi-square control chart, difference control charts, control charts for contrasts, run sum and zone control charts, adaptive control charts, residual control charts, control charts for six-sigma processes, acceptance control charts, T2control charts, exponentially weight means square control charts, multivariate EWMA control charts, one-sided EWMA control charts, moving centerline EWMA control charts, and one-sided CUSUM control chart [5]

In order a process to achieve the intended result, the causes of the mentioned process must be kept under control. To this end, control charts are used [6]

Machine Learning

Machine learning is a branch of AI that involves creating algorithms and models capable of learning and improving from experience without explicit programming. These algorithms use statistics to uncover patterns in vast amounts of data, enabling predictions and actions.

There are three main types of machine learning: supervised learning, unsupervised learning, and reinforcement learning. Supervised learning uses labeled data to train models, while unsupervised learning discovers patterns in unlabeled data. Supervised learning is fairly common in classification problems because the goal is often to get the computer to learn a classification system that we have created [7].

Unsupervised learning refers to algorithms to identify patterns in data sets containing data points that are neither classified nor labelled. The algorithms are thus allowed to classify, label and group the data points within the data sets without having any external guidance in performing that task. The users do not need to supervise the model [8].

Reinforcement learning trains models to make decisions based on environmental feedback. Reinforcement learning is learning through interaction with an environment by taking different actions and experiencing many failures and successes while trying to maximize the received rewards. The agent is not told which action to take. Reinforcement learning is similar to natural learning processes where a teacher or a supervisor is not available and learning process evolves with trial and error, different from supervised learning, in which an agent needs to be told what the correct action is for every position it encounters [9]

Machine learning finds application in diverse fields, including natural language processing, image recognition, recommendation systems, and predictive modeling.

Machine Learning application in control charts

Machine learning plays a growing role in quality control, enhancing traditional control charts used in statistical process control. Control charts monitor and control processes over time, identifying when they are within acceptable limits or out of control. Machine learning improves the accuracy and speed of detecting out-of-control situations. Algorithms uncover patterns that may be unnoticed by humans, enabling early issue identification. Predicting the next data point anticipates process deviations, reducing detection time and preventing waste.

By the ease of use of controls charts and the wide application of ML, this combination is increasingly researched and applied. This is because many types of problems that are arising during the implementation of control charts in nowadays complex processes can be effectively solved with the help of ML approaches (see for example Kang et al. [10] and Qiu and Xie [11]).

Automation of pattern and trend identification through machine learning lightens human analysts' workload, allowing them to focus on interpretation and decision-making. Machine learning also enhances control chart accuracy by identifying relevant data. Algorithms analyze the data to determine vital variables for the chart, eliminating insignificant ones.

One of the main contributions of applying ML techniques in designing control charts is that the modern (production, insurance, healthcare, and etc) processes generate huge data sets with a large degree of diversity by means of modern measurement systems like sensors. In such situations, the traditional statistical monitoring methods fail to handle the monitoring procedure of such processes while ML techniques are able to provide impressive results (Weese et al. [12]).

In conclusion, machine learning enhances traditional control charts in quality control by improving accuracy and speed, automating pattern identification, and enhancing chart precision. Implementing machine learning in control charts boosts efficiency, effectiveness, and cost reduction in quality control processes.

User Interface

A user interface (UI) is the means by which a user interacts with a computer system, including hardware and software. The quality of the user interface can have a significant impact on the user experience, and ultimately on the success of a product or system. Here are a few ways in which a good user interface can help:

1. Ease of use: A well-designed UI reduces the learning curve, enhances efficiency, and prevents errors, making it easier for users to interact with the system.

2. Efficiency: A good UI streamlines tasks and provides easy access to information, increasing productivity by minimizing time and effort required to complete tasks.

3. User satisfaction: An intuitive and efficient UI boosts user satisfaction, leading to higher adoption rates, increased usage, and positive recommendations.

4. Accessibility: A UI can be designed to accommodate diverse users, including those with disabilities. Features like adjustable font sizes and support for assistive technologies ensure inclusivity.

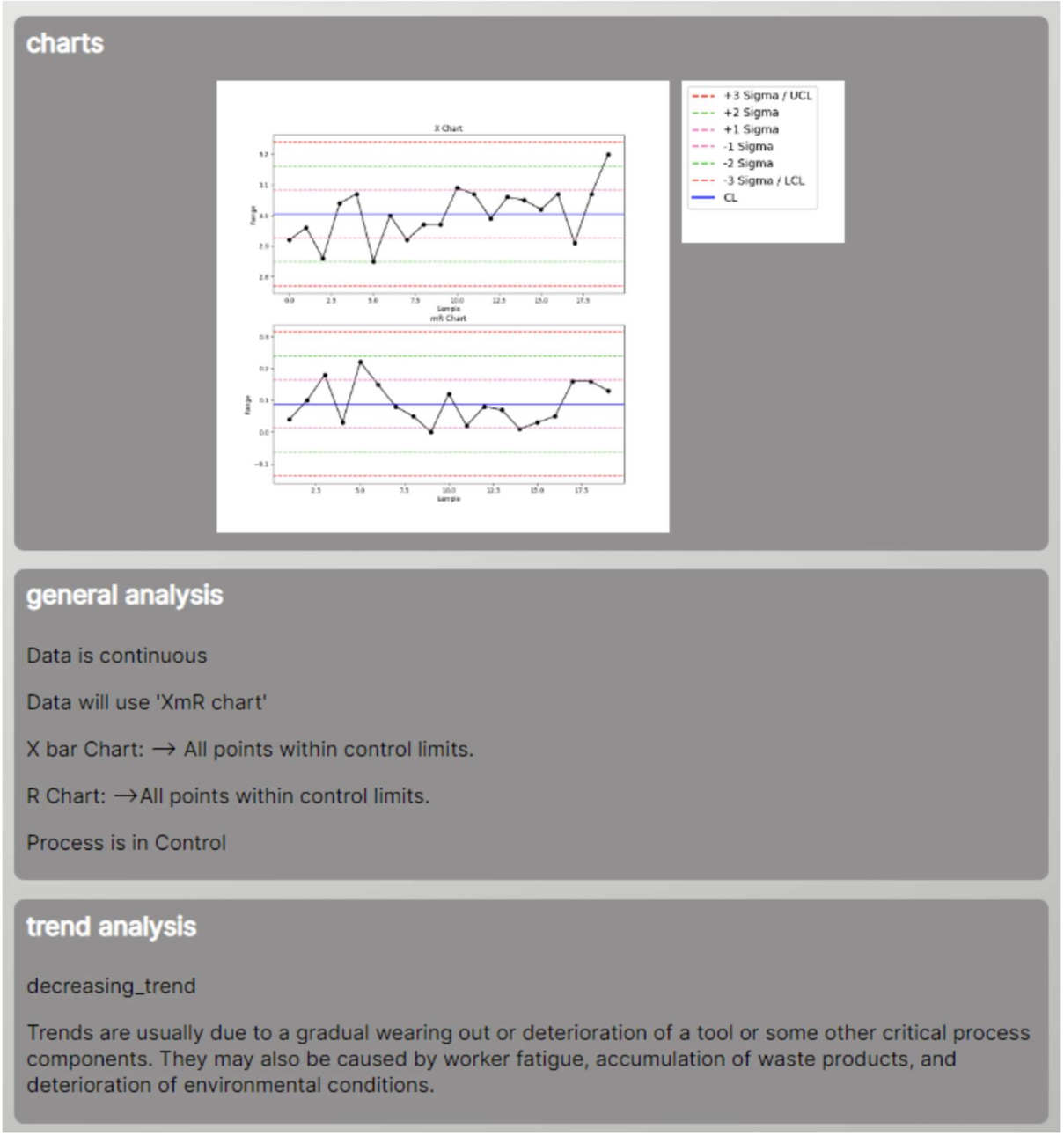
5. Branding: A well-designed UI reinforces brand identity through consistent use of logos, colors, and typography, creating a cohesive and recognizable user experience.

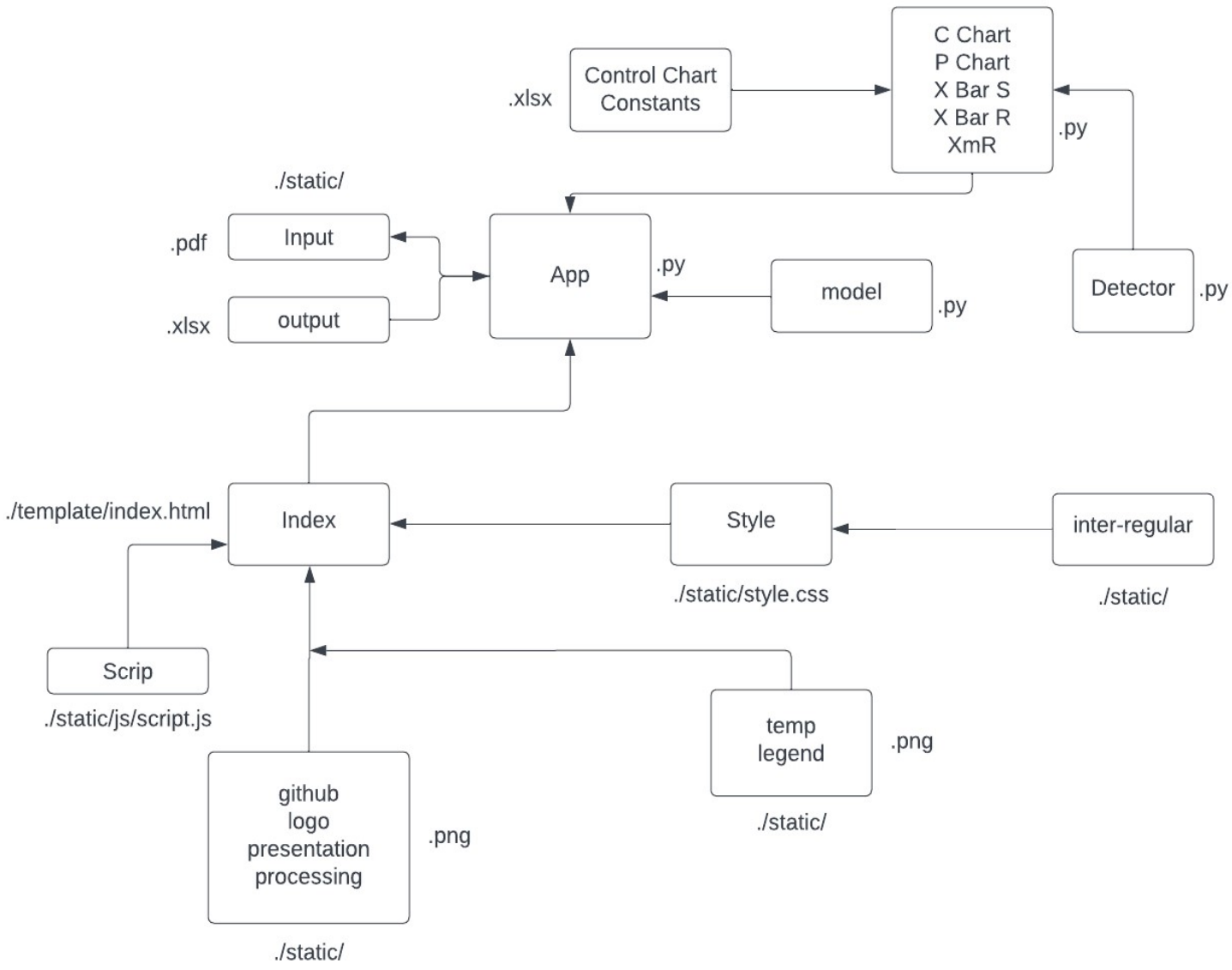
To date, substantial efforts have been dedicated to the development and application of control charts in conjunction with machine learning. However, limited attention has been given to the imperative aspect of enhancing user-friendliness, allowing individuals with minimal expertise to seamlessly incorporate them into their everyday activities.

2. Methodology

Design for the Application:

:





Website Structure

Parts of the Application

1) Frontend of the website which is the user interface.
2) The website's backend includes a server and an Application Programming Interface (API) responsible for transmitting user-uploaded data to the backend for processing.
3) Machine Learning algorithms which process the data and create the control charts.

Tools & Technologies:

1. **Python:**

Python is a high-level, interpreted, interactive and object-oriented scripting language. Python is designed to be highly readable. It uses English words frequently whereas other languages use punctuation, and it has fewer syntactic constructions than other languages. Created by Guido van Rossum and first released in 1991, Python's design philosophy emphasizes code readability with its notable use of significant whitespace. Its language constructs and object- oriented approach aim to help programmers write clear, logical code for small and large-scale projects. Python is dynamically typed and garbage- collected. It supports multiple programming paradigms, including procedural, object-oriented, and functional programming. Python is often described as a "batteries included" language due to its comprehensive standard library.

2. **HTML**

HTML stands for HyperText Markup Language and it is the standard markup language used to create and structure content for the web. HTML is used to define the structure and content of a web page, including headings, paragraphs, images, links, and other elements. It uses a series of tags and attributes to define the content, which are interpreted by web browsers to display the page.

3. **CSS**

CSS stands for Cascading Style Sheets and it is a language used to style and format HTML documents. CSS allows web developers to separate the content and structure of a web page from its presentation and layout, which makes it easier to create and maintain consistent designs across multiple web pages.With CSS, developers can define various styles for HTML elements such as fonts, colors, sizes, and positioning, among others

4. **Flask**

Flask is a lightweight web application framework for Python. It is designed to be simple and flexible, allowing developers to quickly build web applications using Python programming language. Flask is based on the WSGI (Web Server Gateway Interface) specification and can be used with any web server that supports the WSGI interface.

Flask provides a number of features to make web development easier, such as template rendering, URL routing, and request handling. It also allows developers to easily integrate third-party libraries and plugins, making it a popular choice for web development.

5. Javascript

JavaScript is a high-level programming language that is commonly used for creating interactive and dynamic web pages. JavaScript allows web developers to add functionality to their websites, such as form validation, pop-up windows, and animations, among others.
JavaScript is a client-side language, which means that it runs in the user's web browser, rather than on the server-side. This allows for faster and more responsive web applications as users can interact with the website without the need for page reloads.

6. Pandas

Pandas is an open-source library that is built on top of NumPy library. It is a Python package that offers various data structures and operations for manipulating numerical data and time series. It is mainly popular for importing and analyzing data much easier. Pandas is fast and it has high-performance & productivity for users. After the pandas have been installed into the system, you need to import the library.

7. Numpy

NumPy is a general-purpose array-processing package. It provides a high-performance multidimensional array object, and tools for working with these arrays. It is the fundamental package for scientific computing with Python. Besides its obvious scientific uses, NumPy can also be used as an efficient multi-dimensional container of generic data. Arbitrary data-types can be defined using NumPy which allows NumPy to seamlessly and speedily integrate with a wide variety of databases.

Features Included:

1. Data Upload button (formats allowed: xls, pdf)
2. Control Chart built by AI based on the data given.
3. Simple User Interface which is easy to understand.
4. A HOW TO USE section for the user.
5. Process Control Detection

6. Trend Analysis of the data.
7. Some general insights about the data.

Application Flow:

1. Data Uploaded by user through the user interface.
2. Data goes to the backend.
3. Data is processed by the decision tree algorithm.
4. The algorithm recommends an appropriate control chart.
5. With the help of python libraries like PANDAS and NUMPY, a control chart is created.
6. Also, machine learning algorithms give analysis of trend and process control analysis.

Machine Learning Algorithms:

Support Vector Machines

Support vector machines (SVMs) are a type of supervised machine learning algorithm that have been successfully applied to a wide range of classification and regression problems. SVMs have been used in various applications in manufacturing and quality control, including the development of control charts.

Control charts are typically constructed using statistical methods, but machine learning techniques such as SVMs have been proposed as an alternative method for constructing control charts.

One of the advantages of SVMs is their ability to handle non-linear relationships between variables, which may be useful for identifying complex patterns in the data. SVMs can also handle datasets with high dimensionality, making them useful for processing large amounts of data.

In the context of control chart analysis, SVMs have been used to identify abnormal patterns in the data that may indicate an out-of-control process. SVMs can be used to classify new data points as either in control or out of control based on patterns in the historical data. SVMs can also be used to estimate the control limits for the control chart, which can be useful when the underlying distribution of the data is not well-defined.

SVMs have been shown to be effective in identifying complex patterns in the data and have been used successfully in a variety of applications.

Decision Tree

Decision trees are a type of supervised machine learning algorithm that have been used in a wide range of applications, including quality control and process improvement. One application of decision trees in quality control is the selection of appropriate control charts for a given dataset.

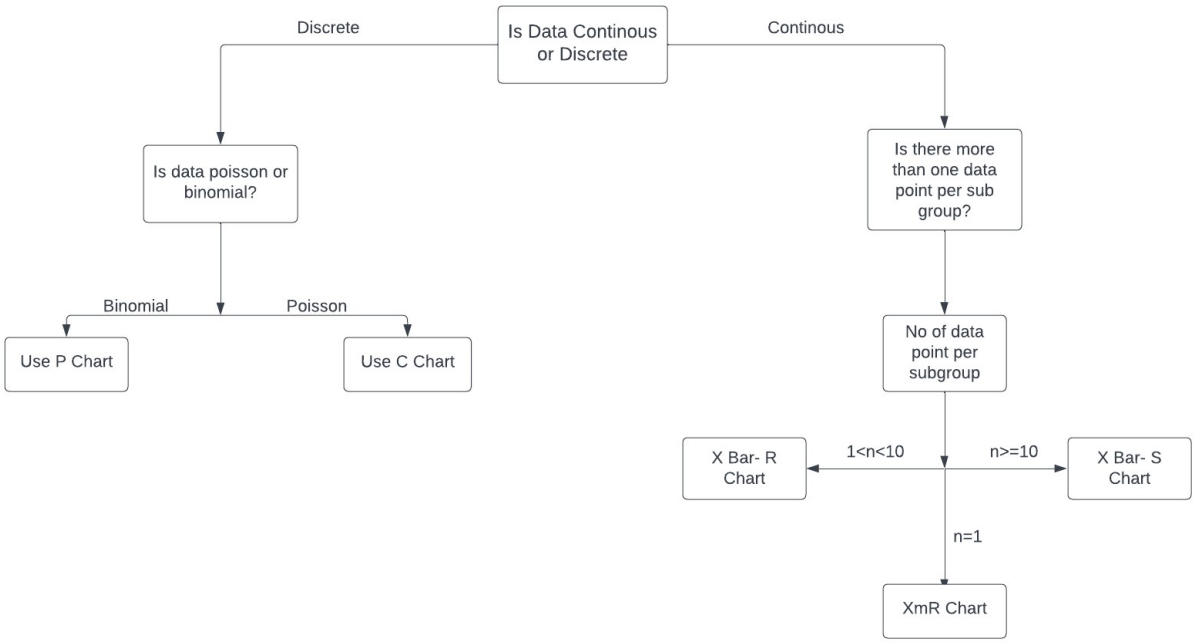
The choice of control chart to be used depends on the nature of the data being monitored, as well as the objectives of the process being analyzed. Selecting an appropriate control chart can be a complex task, requiring an understanding of statistical theory and the characteristics of the data being analyzed.

Decision trees are a useful tool for selecting appropriate control charts because they can handle both categorical and continuous data, and can identify important variables and interactions between variables. Decision trees can be used to analyze data and identify patterns, and can be used to predict which control chart is appropriate for a given dataset.

In the context of control chart selection, decision trees can be used to classify data points into different categories based on the characteristics of the data.

Decision trees can also be used to identify interactions between variables that may affect the choice of control chart. For example, a decision tree could be used to identify whether a given dataset contains a trend or a shift, and then select an appropriate control chart based on this information.

By using decision trees to analyze data and identify patterns, practitioners can make informed decisions about the choice of control chart, improving the accuracy and efficiency of quality control processes.



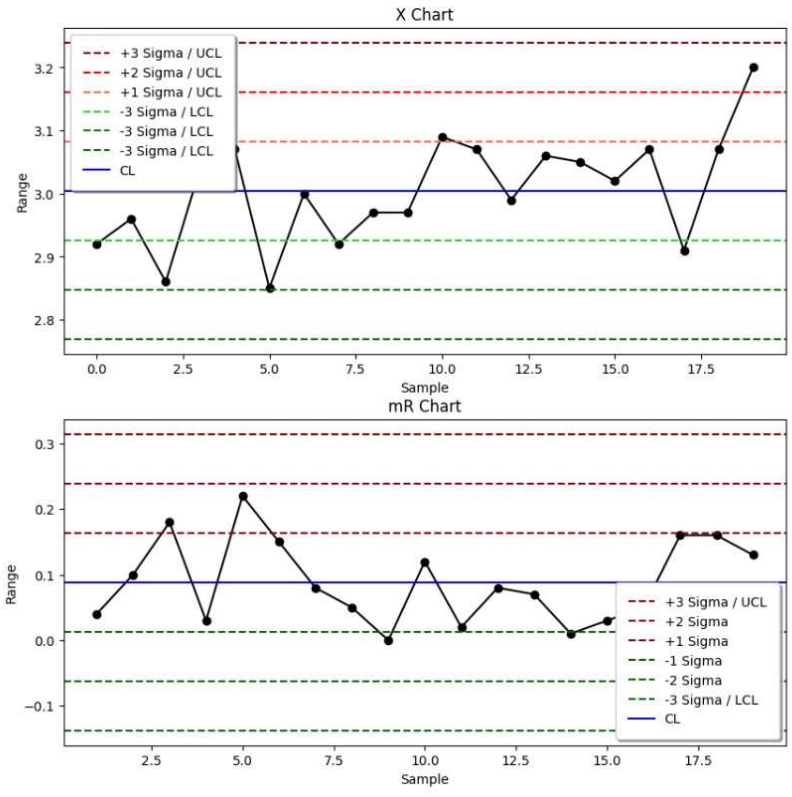
3. Results And Discussion

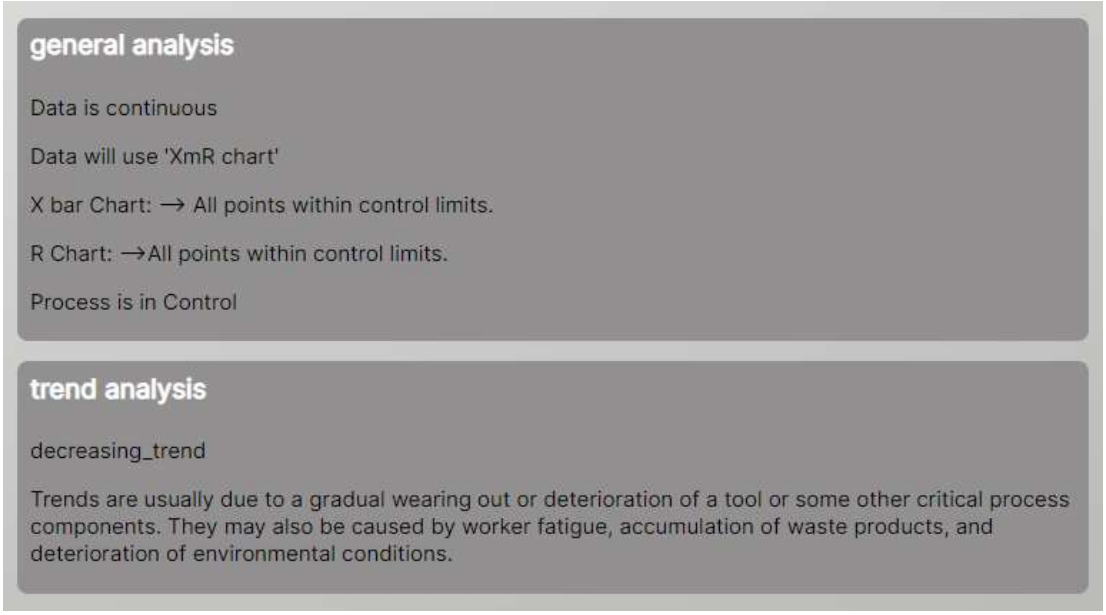
A web-based application was developed that allows users to upload data and generate control charts based on their input, using machine learning techniques to automatically detect out-of-control points. The application was built using Python and its Flask web framework and incorporates machine learning libraries such as PANDAS.

Users are presented with a simple interface that allows them to upload data in CSV format, choose a type of control chart, and customize chart parameters such as chart title, x and y axis labels, chart width and height, and other options. The application calculates the relevant control chart statistics based on the uploaded data using statistical methods and libraries such as NumPy in Python.

In addition to the standard control chart calculations, the application also uses machine learning techniques to automatically detect out-of-control points in the data. This has been done using the help of machine learning algorithms like support vector machines (SVMs). Overall, the application was found to be a useful tool for generating control charts based on user-uploaded data, incorporating machine learning to automatically detect out-of-control points. Future work could include improving the accuracy and efficiency of the machine learning approach, adding more machine learning algorithms and incorporating additional types of control charts.

The use of machine learning in control charts has the potential to improve the accuracy and efficiency of control chart analysis, providing a useful tool for practitioners in the field of statistical process control.





4. Conclusion

In this paper, we presented a web-based application that allows users to upload data and generate control charts based on their input, incorporating machine learning techniques to automatically detect out-of-control points.

Future work could include improving the efficiency and robustness of the machine learning approach and incorporating additional types of control charts. In addition to SVMs, other machine learning techniques such as neural networks and decision trees have also been proposed for constructing control charts.

Overall, this research demonstrates the potential of machine learning in the field of statistical process control, providing a useful tool for practitioners in a variety of industries and applications, including manufacturing, quality control, and data analysis. The integration of machine learning techniques in control chart analysis has the potential to improve the accuracy and efficiency of control chart analysis, providing a valuable contribution to the field.

5. References

1 International Standards Organization (ISO). ISO 15189:2007. Medical Laboratories – Requirements for Quality and Competence. 2nd ed. 2007

2 Xie M, Lu XS, Goh TN, Chan LY. A quality monitoring and decision-making scheme for automated production processes. International Journal of Quality and Reliability Management. 1999;16(2):148-157

3 Antony J, Kaye M, Frangou A. A strategic methodology to the use of advanced statistical quality improvement techniques. The TQM Magazine. 1998;10(3):169-176

4 Hairulliza MJ, Hazura M, Erna BN. Aplikasi Carta Kawalan Dalam Industri Produk Minuman. Seminar Kebangsaan Komputeran Industril Comp. 2005:43-47

5 Montgomery DC. Introduction to Statistical Quality Control. 4th ed. New York: John Wiley & Sons, Inc.; 2001. 796 pp

6 Atilla A. Exercises on ISO 9001 Quality Insurance System (ISO 9001) Kalite Güvence Sisteminin Örnekleri. İstanbul: Çağlayan Publications (Kitabevi); 1996. 412 pp

7 Taiwo, O. A. (2010). Types of Machine Learning Algorithms, New Advances in Machine Learning, Yagang Zhang (Ed.), ISBN: 978-953-307-034-6, InTech, University of Portsmouth United Kingdom. Pp 3 – 31

8 Barlow, H. B., "Unsupervised learning," Neural computation, Vol.1, No.3, pp.295-311, 1989.

9 S. Russell and P. Norvig, Artificial Intelligence: A Modern Approach, 3rd ed. Pearson, 2009.

10 Z. Kang, C. Catal, and B. Tekinerdogan. Machine learning applications in production lines: A systematic literature review. Computers & Industrial Engineering, 149:106773, 2020.

11 P. Qiu and X. Xie. Transparent sequential learning for statistical process control of serially correlated data. Technometrics, (just-accepted): 1–29, 2021.

12 M. Weese, W. Martinez, F.M. Megahed, and L.A. Jones-Farmer. Statistical learning methods applied to process monitoring: An overview and perspective. Journal of Quality Technology, 48(1):4–24, 2016.

13 https://github.com/DevilsSnare/MLforControlCharts