1. **Introduction**

## The Fourth Industrial Revolution and Artificial Intelligence

Artificial intelligence plays a crucial role in the fourth industrial revolution. (Li et al. 2017; Zhong et al. 2017).The focus of value generation, artifactual paradigms, enabling technologies, applied methodologies, and business strategies of design are continually enriched. This is a fact of the matter. (Horváth, I. 2021). With the acquisition of more appropriate knowledge and information and more applications, a person or system can become smarter. (Rindermann and Ceci, 2009).

* 1. Challenges and Opportunities in Smart Manufacturing with IoT and Computational Intelligence

Smart manufacturing is being touted as the next indus- trial revolution (M. Bryner 2012).  
The objective of the smart manufacturing movement is to establish production operations that incorporate information. Operations undergo a transformation from being reactive to proactive, taking action before responding, shifting from compliance to performance, transitioning from tactical to strategic, and expanding from local to global. Tata Motors' smart factory implements the Nano model and possesses the capability to predict bottlenecks and malfunctions through the utilisation of automation technologies such as sensors, microprocessors, and motor controllers. Additionally, it possesses the capability to procure components from suppliers in real time (M. Bryner 2012). It is widely believed by numerous researchers that the implementation of a computational intelligence-assisted design framework is imperative for the development of smart systems. The operational and behavioural self-adaptation requires a dedicated system intelligence (Ashby, 1947). In other words, the ability of the system to expand its knowledge base and improve its reasoning mechanisms when necessary will be a fundamental measure of the intelligence of systems. In a simplified manner, the level of intelligence exhibited by intellectualised engineering systems can be assessed based on their capacity to address a variety of complex real-world application problems. Additionally, their cognitive abilities can be evaluated based on their capacity to expand their knowledge base and problem-solving mechanisms. In the domain of system engineering, there exists a direct correlation between the inherent self-adaptive capabilities and potentialities (Sabatucci et al., 2018).

* 1. Industrial Internet of Things (IIoT) and Advanced Analytics for Process Surveillance

Intelligent products and services are now widely available thanks to growth of the Internet of Things (Atzori et al., 2010). The volume of data generated from Internet of Things systems is a significant challenge to state-of-the-art big data solutions. Generating insights from the extensive amount of data requires the development of efficient data engineering systems and algorithms tailored to the specific type of sensor data being analysed (Maurya, 2016). In their study, Bures et al. (2020) examined the manifestation of smart systems, which are characterised by a diverse and interconnected landscape comprising different applications of the Internet of Things (IoT), Cyber-Physical Systems (CPSs), and smart sensing systems. In addition, the observers have witnessed a conventional implementation of a smart system application. This implementation consists of various autonomous components that inherently collaborate with each other. These components encompass hardware units that operate on specific networks, as well as software components that are associated with them. The achievement of smartness in this application is realised through the combined capabilities of sensing and operation, which are carried out both autonomously and collaboratively. Components proactively sense the environment and provide their knowledge to other components in order to enable them to make intelligent and informed decisions. A typical conceptualization of a smart system is that it changes its reasoning strategy and activates problem solving agents accordingly. It also learns new models to process a changing and growing set of (sensor) input data or knowledge base contents. The computing mechanisms are preprogrammed to perform this task, although the necessary adaptation and computational resources are determined during runtime. Systems adapt themselves within an anticipated envelope of changes. The concept of smartness does not imply that all decisions are made solely by the mechanisms of reasoning, but rather that humans are involved in the operational loops of the system (Schirner et al., 2013).   
In the realm of Industrial Internet of Things (IIoT), manufacturing processes and product lifecycles are under constant surveillance to detect any irregularities or failures (Xu et al., 2014)  
Utilising advanced analytics is crucial in extracting valuable insights and business value from the vast amount of data collected through IoT. Nevertheless, there are certain difficulties associated with IoT data analytics. The datasets collected using sensors are more complex and challenging to interpret compared to the structured datasets collected through human interaction on Internet websites (Maurya, 2016) In the realm of Industrial Internet of Things (IIoT), manufacturing processes and product lifecycles are under constant surveillance to detect any irregularities or failures (Xu et al., 2014)  
We consider data, information, and knowledge as separate and distinct tiers. Data are commonly perceived as a compilation of factual elements, whereas information encompasses the significance derived from interconnected data. Knowledge, on the other hand, represents the capacity to solve problems by integrating and abstracting information. Lastly, intellect denotes the aptitude to effectively apply acquired knowledge across diverse and ever-changing circumstances. Sensors collect signals from various sources and subsequently transform them into data. Information structures serve the purpose of capturing and encoding relationships among data, thereby revealing their inherent meaning. Advanced reasoning mechanisms, such as artificial neural networks, have the capability to uncover hidden relationships within vast data streams and transform them into patterns of knowledge. The field of context management involves acquiring and utilising meta-knowledge to effectively apply problem-solving knowledge (Horváth, 2020a). The collection of data serves as the initial and fundamental step in the realm of industrial big data, assuming a crucial role in facilitating data-driven maintenance practises within the industrial sector. The establishment of a fine-grained data flow for subsequent statistical analysis and machine learning is challenging due to the integration of devices from various equipment suppliers, the diversity of communication protocols, varying degrees of device openness, and differing levels of intelligence among the devices (J. Wan, 2017). As the integration of information technologies continues to expand across various sectors of industry, there has been a growing emphasis on the incorporation of increasingly sophisticated algorithms in the realm of active preventive maintenance. However, in the industrial setting, the need for durability and immediate processing becomes much more important. For example, the utilisation of deep learning has shown exceptional effectiveness in the fields of image identification and natural language processing. Empirical research has confirmed that deep learning algorithms may effectively utilise stored data to generate remarkable outcomes in various domains. However, caution must be exercised when considering the suitability of deep learning in the context of active preventive maintenance. Algorithms that may not possess a high level of sophistication, but when coupled with expert knowledge, have the potential to yield enhanced performance (J. Wan, 2017).

## Methods in Data Science for Real-Time Monitoring in Manufacturing

Manufacturing companies are exploring ways to adapt to industry 4.0 due to the growing need for advancements that can enhance their product yield. (Lee et al., 2014)  
Quality control has long been a crucial component of production and an essential component. More advanced sensor technologies, such Internet of Things (IoT) and Radio Frequency Identification (A. Zaslavsky, et al., 2012), allow data to be collected at every level of the production process. Companies can collect crucial process data from different stages of their production line during the manufacturing process. By harnessing copious amounts of accessible data, firms can utilise advanced data analytics to effectively analyse and elucidate uncertainties. This enables them to make more perceptive and well-informed judgements regarding future product development. (Carbery et al., 2018)  
AI techniques have been utilised to enhance the early identification or prediction of faults in production lines within the field of intelligent manufacturing (IM) (Kusiak 2017; Tao et al. 2018).  
Manufacturing processes exhibit significant intricacy as they typically involve multiple sequential steps and necessitate various sorts of resources, including financial, human, and technological resources. The recent technological breakthroughs have provided several options to extract valuable information from manufacturing processes. This information can be used to optimise production and identify any factors that may result in wasteful allocation of resources. Analysing the data that describes the fabrication processes is highly tough because the products undergo numerous intricate operations. As a result, the data may exhibit characteristics such as missing information and a large number of features.( Moldovan et al, 2019)  
Utilising standard statistical approaches to extract information from data that describes manufacturing processes is frequently insufficient for identifying procedures that may result in errors in the final outcomes. .( Moldovan et al, 2019)  
  
Recent research efforts have delved into machine learning techniques for simulating manufacturing domains, focusing on predictive maintenance and defect detection (G. Susto, et al., 2015; T. Wuest et al., 2014; B. Kroll, et al., 2015; G. Wang, et al., 2017). The utilization of advanced machine learning and deep learning techniques for forecasting product quality has demonstrated high value, offering more precise predictions. However, challenges, such as selecting the most suitable prediction model and addressing missing information, need to be addressed, particularly in light of the intricate nature of manufacturing data (Moldovan et al., 2019).  
Given the importance of staying competitive and increasing productivity, it is clear that incorporating data science methods for real-time monitoring of manufacturing processes is a logical progression. An notable accomplishment occurred in 2012 when Intel managed to reduce its manufacturing expenses by $3 million by employing predictive analytics to prioritise inspections of their silicon chips (Ronen and Burns 2013). As further examples, Raytheon, a defence manufacturing corporation, introduced the manufacturing execution system (MES) at its missile facility in Huntsville, Alabama (J. Hagerty, 2013). This system gathers and examines data from the factory shop floor, enabling precise determination of the optimal number of screw rotations required for a flawless critical component.  
Following this trend, Bosch competed in a Kaggle competition (Kaggle, 2016) and released its dataset, which consisted of anonymized records of measurements and tests made for each component along the assembly line. The company then challenged the Kaggle community to predict product part failures, which ultimately enabled Bosch to provide quality products to the end user at lower costs.

## General Project Aim

In this context, in this study highlights the importance of using advanced data analytics and machine learning techniques to understand, detect and solve problems encountered in Bosch's production line. The primary goal of the study is to identify the root causes of failures occurring in Bosch production lines and to increase production efficiency to finding solutions for these reasons.  
This study is important in the following key areas:

1. **Fast and Accurate Detection of Problems:** Using data analytics and machine learning models, problems on production lines can be quickly detected. This allows businesses to take a proactive rather than reactive approach.
2. **Determination of Root Causes:** The study addresses the root of the problems by identifying the root causes of failures in production lines. In this way, similar problems can be prevented from recurring.
3. **Data Based Decision Making:** Advanced data analytics and machine learning techniques enable businesses to make data-driven decisions. This optimizes business processes and increases efficiencies.

In conclusion, this study, using production data shared by Bosch on Kaggle, highlights the importance of data analytics and machine learning applications in industrial production. Successful adoption of these technologies enables Bosch to more effectively manage problems in production processes and increase production efficiency. Therefore, the study is expected to be an important reference source for companies seeking data-driven solutions in industrial production.

## Objectives

This study includes two main objectives set to understand and solve problems in Bosch's production line:

**RQ1: What are the primary causes of manufacturing failures on Bosch's production line, and how can these root causes be identified and addressed?**

The primary objective is to identify and examine malfunctions that have transpired in Bosch's manufacturing facilities and comprehend their underlying origins. To effectively address these factors, it is essential to get a significant outcome through the analysis. Within this framework, the objective is to ascertain the underlying causes of issues in the production line and devise solution plans through the utilisation of analyses and learning algorithms. This goal aims to focus on key problems to prevent efficiency losses in production processes and improve product quality.

**RQ2: How can advanced data analytics and machine learning techniques be leveraged to predict and mitigate manufacturing failures in real-time on the Bosch production line?**

The second goal is to use advanced data analytics and machine learning techniques to predict and prevent production errors in real time on Bosch's production line. Thanks to the models to be used, potential errors in the production process will be predicted in advance, paving the way for instant interventions to these errors and optimizing production processes. This goal aims to ensure that production lines operate more reliably and efficiently by strengthening predictive maintenance strategies.  
The purposes are to achieve these two objectives in order to offer more efficient and strategic solutions to the issues in Bosch's manufacturing processes, resulting in cost reduction and enhanced customer satisfaction. Furthermore, this study seeks to offer a thorough comprehension of the potent utilisation of data analytics and machine learning in industrial production.

## DATA EXPLORATION

# Background and Related Work ? altbaşlıklar ne olmalı ? ?