2023

Production planning in the context of Industry 4.0 with focus on efficient job allocation & workers' real-time status

Abdullah Mohammed Albassam
Western New England University

Production Planning in the Context of Industry 4.0 with Focus on Efficient Job Allocation & Workers' Real-Time Status

by

Abdullah Mohammed Albassam

A dissertation submitted to the Faculty of
Western New England University
in partial fulfillment of the requirements for the
Degree of Doctor of Philosophy
in Engineering Management

Springfield, MA

08/15/2023

Approved by

Dr. Seyed A. Niknam, Chair, Associated Professor of IEEM	5/22/23 Date
Dr. Christian Salmon, Member, Chair and Assoc. Professor IEEM	5/4/2023 Date
Dr. Joseph Ekong, Member, Assistant Professor of IEEM	5/4/2023 Date
Dr. Vedang Chauhan, Member, Assistant Professor of EE	5/4/2023 Date
Dr. Christian Salmon, IEEM Department Chair	5/4/2023 Date
Dean, College of Engineering	5/30/2023 Date

Abstract

Industry 4.0 (I4.0) has emerged a distinct impact on industrial workforce and created demand for diverse set of workforce skills and domain knowledge. Accordingly, I4.0 production systems are in need for developing and utilizing an appropriate workforce planning that considers workers with different type of skills to cope with the production requirements and keep up an efficient production. The I4.0 philosophy advocates the usage of advanced wearable technologies. Such wearable devices are able to monitor workers' status and record vital signs and physiological data. It is well known in literature that workers performance in production systems is linked to their job satisfaction level as well as psychological well-being. There is much active research in the area of advanced physiology measurement technologies and incorporating the workers' health data into industrial applications in real time. In essence, it is expected that smart wearable health devices provide the ability to boost job satisfaction, reduce human errors, and affect performance by helping managers for more efficient task matching and scheduling.

This research is focused on developing job assignment models in the context of I4.0 and has considered both workers' physiological status and the skills required to achieve the production goals. The ultimate goal of the proposed models is to maximize productivity by matching operations tasks to workers with different required skills and various skill levels. This study also considers workers' performance indicator which is predicted by machine learning models using workers' physiology measurement. The assignment model could provide promising results in moving toward real-time application of workers' physiological status in order to better assign production tasks and maximize production value.

Acknowledgment

I would like to express my deepest appreciation to my advisor Dr. Seyed Niknam who made this work possible. The completion of my dissertation would not have been possible without his support and nurturing. His vision, knowledge, and encouragement have been invaluable throughout all stages of my dissertation. I would also like to acknowledge my warmest thanks to my committee members, Dr. Christian Salmon, Dr. Joseph Ekong, and Dr. Vedang Chauhan, for their guidance, comments, and suggestions throughout my research journey.

This work would not have been possible without the financial support of the Saudi Cultural Mission in the US. I would like to take the opportunity to extend my thanks to this institution.

Last but not least, special thanks go to my parents, my wife, my daughter Reem, my little boy Yousef, and my brothers and sisters for the extreme support. They mean a lot to me, and they are the most important people in my life. I dedicate this dissertation to them.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

Abbreviation	Definition
AdaBoost	Adaptive Boosting
AI	Artificial Intelligence
ALWABP	Assembly Line Worker Assignment and Balancing Problem
ANS	Autonomic Nervous System
ATMT	Advanced Trail Making Test Task
BVP	Blood Volume Pulse
CANTAB	Cambridge Neuropsychological Test Automatic Battery
CASI	Cognitive Abilities Screening Instrument
CIPD	Chartered Institute of Personnel and Development
COP	Constraint Optimization Problem
CPS	Cyber- Physical Systems
CS	Centralized Solution
DS	Decentralized Solution
DSC	Digit Symbol Coding
DT	Decision Trees
ECG	Electrocardiograph
EDA	Electro Dermal Activity
EWB	Equal Width Binning
GBoost	Gradient Boosting
HF	power in High Frequency range
HR	Heart Rate
HRV	Heart Rate Variability
ICT	Information and Communication Technologies
IIOT	Industrial Internet of Things
IoT	Internet Of Things
IQR	Interquartile Range
IT	Information Technology
IT-OT	Information Technology - Operational Technology
KNN	k-Nearest Neighbor
LASSO	Least Absolute Shrinkage and Selection Operator
LDA	Linear Discriminant Analysis
LF	Power in Low Frequency range
LF-HF	A ratio of Low to High Frequency power
LSVM	Linear Support Vector Machine
LTA-FIT	Learning, Training, Assistance– Formats, Issues, Tools
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MAS	Multiagent System
ME	Mean Error
ML	Mean Latency
ML	Machine Learning

MOMA	Multiskilling Optimization Model for Allocation
MOT	Motor Screening
MRBCRS	Multiple Rule Based Constructive Randomized Search
MRT	Mean correct Reaction Time on go trials
MSDs	Musculoskeletal Disorders
MSE	Mean Squared Error
NRMSE	Normalized Root Mean Squared Error
PC	Personal Computer
PCA	Principal Component Analysis
PH	Probability of Hit
pNN50	Percentage of successive Normal-Normal intervals that differ by more than 50 ms
PPG	Photoplethysmography
PSS	Proportion of Successful Stops
RBF-SVM	Radial Basis Function kernel- Support Vector Machine
RF	Random Forest
RMSE	Root Mean Squared Error
RMSSD	Root Mean Square of Successive Differences
RR Intervals	R-R Intervals is the time between two successive heartbeats measured in milliseconds
RULA	Rapid Upper Limb Assessment
RVP	Rapid Visual information Processing
SDNN	Standard Deviation of Normal-Normal intervals
SME	Small and Medium size Enterprises
SSRT	Stop Signal Reaction Time
SST	Stop Signal Task
ST	Skin Temperature
ST	Strategy
SVM	Support Vector Machine
SWM	Spatial Working Memory
VAS	Visual Analogue Scale

1. Introduction

1.1 Background

In recent years, certain technologies such as cyber-physical systems, variety types of sensors, Internet of Things (IoT), and smart networks have evolved and started to influence different areas in both service and manufacturing sectors. These developments are referred to as the fourth industrial revolution, also known as "Industry 4.0" [1].

As these technologies contribute to more automation for simple manufacturing processes, one can expect an increase for highly complex workspaces. In addition, in near future, structures for communication and collaboration will be modified and therefore, processes will become more interconnected. This will lead to a transformation in working models which is defined as how work is planned, organized, and managed. Consequently, the technical, organizational, and social aspects of work activities will interweave. I4.0 will not only affect technology and production, but our way of working will be one of the most affected dimensions [1].

One important aspect of I4.0 is empowering customers with the so-called individualized products in small batches at reasonable costs by enabling an adaptive production process and optimizing value chains and value-added networks [2]. Hence, the manufacturing tasks are changing to accommodate the individual customized products and the associated demands which lead to a need for adaptive employees who are able to perform different types of tasks. Moreover, I4.0 will alter the nature of work processes and that would result in conversion of employees' roles and required skills [3]. This transformation will alter job profiles, necessitating the acquisition of a new and diverse set of worker skills which create the need for multi-skills workers [1], [4].

A production manager in the era of I4.0 is in need for developing and utilizing an appropriate workforce planning that considers workers with different type of skills to cope with

I4.0 requirements and achieve efficient production. Chartered Institute of Personnel and Development (CIPD) defines workforce planning as "*A core process of human resource management that is shaped by the organizational strategy and ensures the right number of people with the right skills, in the right place at the right time to deliver short-and long-term organization objectives*" [5]. Therefore, the question here is what these skills are and how to assign workers to jobs in a way that achieves efficient results in production performance.

Workers' performance has always been an important factor in manufacturing processes. I4.0 introduces advanced technologies that aim to improve both workers and production performance. A branch of I4.0 technologies, known as wearable devices, has focused on collecting health data and analyzing workers individual performance. Some of these wearable devices provide additional strength to workers helping them to perform tough jobs like lifting heavy loads. Other devices monitor workers' status and measure vital and physiological markers. Certain researchers have started investigating how advanced physiology measurement technologies can be incorporated into an industrial application to increase worker's performance and wellbeing at assembly stations via assessing worker wellbeing in real time [6]. In essence, smart devices have the ability to boost job satisfaction, reduce confusion and errors, and affect performance by providing visible clues to managers for matching task to workers [7]. Moreover, these smart devices are useful not just for personal well-being, but also for improving overall operation [8]. It should be also noted that performance is correlated with psychological wellbeing and health [9][10]. High performance has been linked to a high level of job satisfaction as well as a high level of psychological well-being [11]. Individuals with diverse knowledge and capabilities in their work settings are more likely to experience work-related stress when the work demand is not matched with their capacities[12], [13]. In this regard, boredom and under-incentive have been identified to influence operator performance [14].

Workers' assignment models in I4.0 have to consider both workers skills and the skills required to achieve potential improvement in production and profitability. Moreover, I4.0 provides means that measures workers' physiological status in real-time. Since the worker's status proven to affect his/her performance, engaging this information in job assignment models is promising in order to increase model effectiveness toward improving production performance. Accordingly, this research is focusing on building job assignment models considered the above-mentioned aspects of I4.0.

1.2 Research Objectives

The objective of this research is to develop an efficient job allocation model in the context of industry 4.0 that considers identifying the right skills needed for I4.0 and matching workers who possess these skills with jobs in order to achieve optimum performance level. The research objectives are listed below:

- Reviewing and selecting the current skill sets and models in the published literature to define the skills needed for the jobs related to I4.0 systems.
- Developing job assignment models that can assign workers to jobs based on their skill levels and the skills required by job.
- Advancing the job assignment models by engaging a significant real-time factor that affect workers performance.
- Developing a machine learning model to predict the performance factor value which is used in the assignment model.

Chapter 2 provides a review of the published literature with respect to (i) required skills for Industry 4.0, (ii) job assignment models that are suitable for 4.0, (iii) models that consider multi-skills worker and job rotation, and (iv) real time monitoring in workplace using wearable devices. Chapter 3 presents the details on skills and competencies selection model, and presents the developed job assignment models. Chapter 4 provides and discusses the results of the job assignment models. Chapter 5 provides the conclusions and presents the future directions of this research.

2. Literature review

2.1 Industry 4.0 definitions

Schumacher et al.[15] defined Industry 4.0 as recent technological advances where physical objects, human actors, intelligent machines, production lines and processes across organizational boundaries are integrated by the internet and its related technologies (e.g. sensors, embedded systems) which is called smart network in order to form an intelligent and agile value chain.

Wang et al. [16] described the main idea of Industry 4.0 as the use and implementation of emerging technologies to integrate business and engineering processes in order to get flexible, efficient, high quality, and low-cost production operations. Tay et al. [17] defined the Industry 4.0 as a combination of new technologies and ideas that will transform the existing value chain toward a connected value chain by smart systems and intelligent technologies that is able to be self-organizing and provides dynamic control within the organization.

2.2 Workforce Planning in Industry 4.0

2.2.1 Skills Requirement

Shaw et al. [18] reviewed the automation effects on employee’s skill development and workspaces complexity. The authors provided recommendations to ensure the retention of jobs and workforce in an industry 4.0 environment. The authors concluded that the implementation of smart production systems will result in automation for simple and routine processes in which human involvement will be decreased. On the other hand, some processes become more complex and probably need high-skilled employees. Therefore, they recommended organizations to have a qualification strategy for employees to cope with this dynamic working environment.

Balalle et al. [19] discussed what skills employees need to develop and how to adapt to I4.0 workplace. They believed physical and manual skills will decline in new workplace, while technological, social and emotional skills are in demand. Regarding the adaptation of new skills, the strategy of continuous learning process is recommended to be an essential organization culture in order to improve employee skills according to work requirements.

A Competency Model for "Industry 4.0" employees developed by Prifti et al. [1] focusing on competencies needed in Industry 4.0 environment. They used SHL Universal Competency Framework which is "*a generic foundation for building competency models"* along with focus group to develop their model. The model is represented in a table which displays number of competencies clustered in eight groups (leading, supporting, interacting, analyzing, creating, organizing, adapting, and enterprising), followed by 20 competency dimension that divide these groups to 112 component competencies mapped to three disciplines; information systems, information technology, and engineering.

Rehe et al. [20] emphasized that the integration of new technologies will have radical impact on various tasks and job profiles for all staff members as these technologies alter the requirements across the whole value chain and lead to new processes. The authors presented a qualification model named LTA-FIT (Learning, Training, Assistance– Formats, Issues, Tools). The model focusses on learning and training existing employees in three levels for small and medium size enterprises (SME) in Germany. Each level has its own format, issues, and tools according to the organization requirements.

Schinner et al. [21] introduced a competence classification overview that matches Industry 4.0 requirements. Their classification includes technical, methodological, social, and self-management competencies. The technical competencies defined as specialized competences applicable in specific areas such as machine operation. The methodological competencies are defined as methods with a clear functional focus such as analytical ability. The social competencies can be defined as interactional competencies such as communication skills. The self-management competencies are defined as competences relevant to self-organization such as willingness to learn, reliability, and openness to change.

Baena et al. [22] argued that Learning Factories were considered to be effective for developing theoretical and practical knowledge in a real production environment. The Learning Factories is a hands-on prototyping and design laboratory that provides students with modern tools for prototyping, manufacturing, and training in a safe environment. They concluded that Learning Factories may contribute to leverage the way towards new manufacturing trends such as industry 4.0.

Linda and Mattias [23] comprised a literature review on recent studies that analyze the implications of Industry 4.0 and cyber physical systems on labor and work organization. Their research revealed that Industry 4.0 would lead to a significant decline in low-skill jobs and would result in an increase in high-skill activities and a rising complexity in many job profiles. They also anticipated a significant growing of continuous learning, training, and education for workforce to adapt to future qualification required by Industry 4.0 technologies.

Hecklau et al. [24] presented competence model and showed an approach of how companies can make use of the model to meet arising challenges in Industry 4.0. The model identifies main challenges in Industry 4.0 by conducting an extensive literature review. Then, based on a further analysis of those challenges, a list of essential core competencies for employees was derived. By visualizing the identified core competencies, a user can detect competence gaps at the first sight. The visualization in [24] is based on the concept of radar charts, which are used to display multivariate data in a two-dimensional chart. Competencies are clustered around their categories and the red areas symbolize the minimum required competence level for each competence as shown in Figure 1.

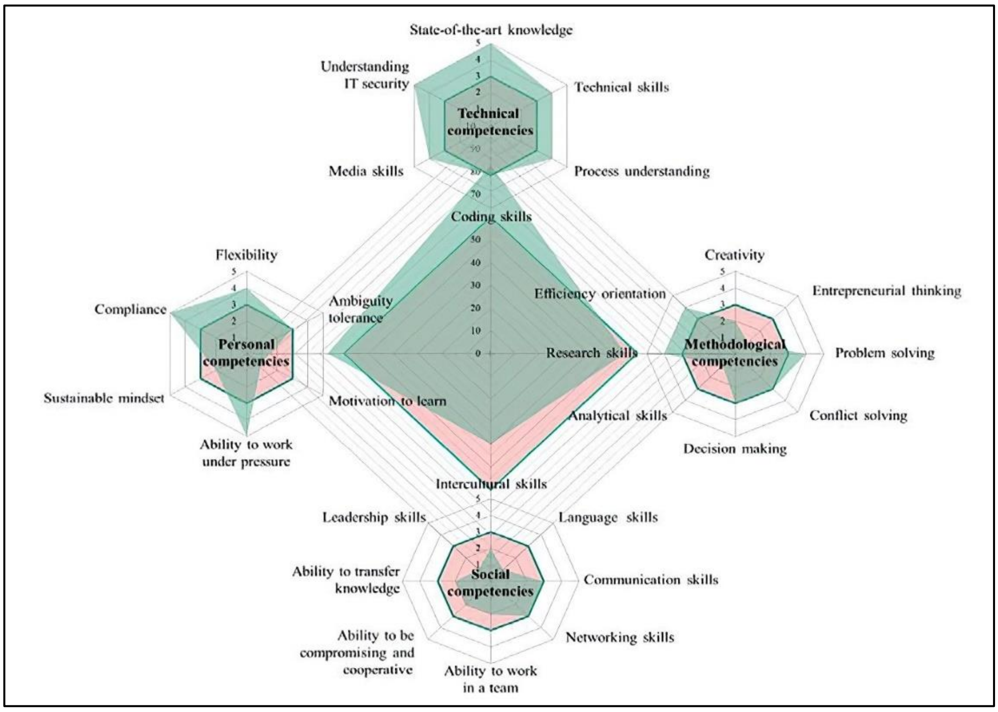

Figure 1. Visualized competence model [24]

Fareri et al. [25] studied the effects of Industry 4.0 on business value chain in terms of competencies and job profiles. The authors applied Porter value-chain model. In [26], the authors included five primary functions and four supporting functions in addition to the six professional archetypes 4.0 proposed by Fantoni et al. [27] to visualize which of these professional archetypes are distributed across different business functions. This was done by applying automated text mining to analyze the literature in order to associate the emerging job profiles resulting from Industry 4.0 for the primary and supporting business functions [27]. The results are presented in Table 1.

Kazancoglu and Ozen [28] proposed structural competency model for workforce 4.0 in five stages. The modeling was initiated by identifying the needs for an organization to be transformed to an Industry 4.0 organization. The second stage is about identifying the needs from workforce 4.0 which represents the understanding of job profile changes. The third stage deals with defining criteria for workforce 4.0 such as ability of dealing with complexity and problem solving. After that, the core stage of the model begins where Fuzzy DEMATEL is used as the multi-criteria decision technique. The advantage of using Fuzzy DEMATEL is the ability to examine criteria's prominence and causal relations. The result of the Fuzzy DEMATEL represents the fifth stage which can be used to support human resource departments with the selected criteria in recruitment processes.

Table 1. Professional archetypes distribution across business functions

Professional archetypes	Business functions (Primary and supporting)
Data Architect	All
IT Architect	Logistics and IT
Geek	Business Management and Facilities
Investigator	Facilities and Quality control
Perfectionist	Facilities, Quality control, and Accounting
Prophet	IT and production
Strategist	Marketing, Business Management, and R&D

Dworschak et al. [29] presented competence needs for Cyber- Physical Systems (CPS) in manufacturing. The competence needs are drawn from technology forecasting and early identification of skill needs by using a model of industry maturity based upon technological, work organization, product and service innovation [30]. The required skills and knowledge include technical, social and collaboration, deep operational and business informational, IT and engineering knowledge.

Hartmann and Bovenschulte [31] proposed a methodology to analyze the skill needs for Industry 4.0 based on European Technology Platform on Smart Systems Integration roadmap [32]. The methodology considered two aspects of skill requirements. First, various models of industrial organization will result in different skills needs. Second, for different subsets of the technologies, there will be different skills needs that depend also on workforce segments in different sectors. The second aspect can be viewed as a matrix of different needed skills in different workforce segments (e.g., managers, engineers, labor, etc.) and in different sectors (e.g., production, sales, finance, etc.) depending on different subsets of technologies.

Adolph et al. [33] derived competencies features for future production based on small scale literature review for production challenges and megatrends such as globalization, shortening of product life cycles, new technologies, scarce resources, knowledge society, and demographic change. They concluded that competencies should have positive impacts on flexibility, changeability, resource efficiency, and process efficiency.

Mittelmann [34] used European Political Strategy Centre framework (see Figure 2) as the starting point to derive competencies needed for work 4.0. The author expanded and restructured the framework into categories of competencies and supported the selection of the competencies by performing small scale literature review. The author presented a model which consists of three main categories: intrapersonal, interpersonal, and a category related to Information and Communication Technologies (ICT) along with their corresponding competencies as shown in Table 2.

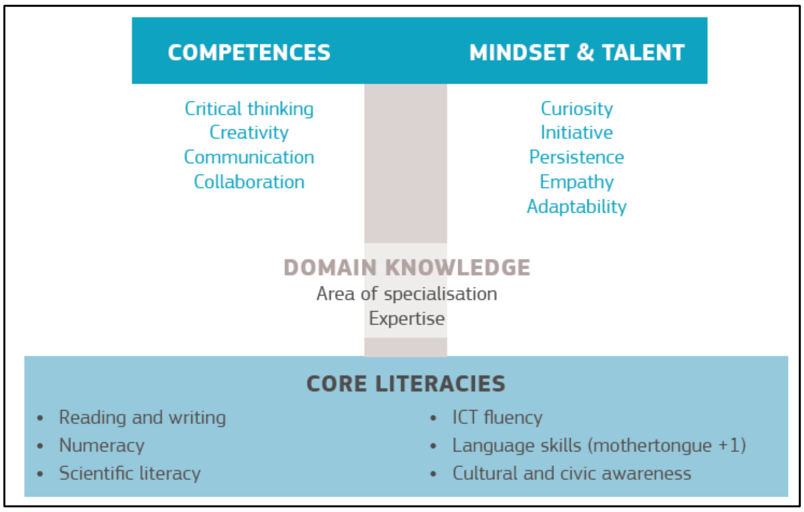

Figure 2. T-shaped skills framework [35]

Galaske et al. [2] presented a toolbox for workforce management 4.0 that can be used to assess the organization's readiness for Industry 4.0 by identifying the current situation of the organization and determining the next steps to reach Industry 4.0 vision. A toolbox named WM4.0 was developed based on the VDMA (German Engineering Federation) Guideline Industry 4.0 and Generic Procedure Model for Small and Medium-sized Enterprises [36][37].

Table 2. Categories and correspondent competencies[34]

Intrapersonal competencies	
Critical thinking	*Using good judgement and common sense as well as logic and reasoning to identify the strengths and weaknesses of alternative solutions, conclusions, or approaches to problems.*
Sense-making	*Determining the deeper meaning or significance of what is being expressed visually or in written or spoken texts*
Novel and adaptive thinking	*Routinely thinking across boundaries and coming up with responses and solutions beyond that which is rote or rule-based*
Transdisciplinary	*Understanding concepts across multiple disciplines and crossing many disciplinary boundaries to create holistic solutions*
Self-direction	*Guiding and organizing oneself, steering and controlling one's learning, and maximizing cognitive functioning with respect to well-being*
Interpersonal competencies	
Communication	*Active listening, conveying information comprehensibly, having difficult conversations with ease to avoid resp. resolve conflicts*
(Virtual) collaboration	*Working productively, driving engagement, and demonstrating presence as a member of a (virtual) team*
Social intelligence	*Connecting to others in a deep and direct way, sensing and stimulating reactions and desired interactions*
Intercultural competency	*Operating effortlessly in different cultural settings*
ICT-related competencies	
ICT fluency	*Using computers, communication technologies and applications to access, manage, integrate, evaluate, and create information in order to take part in a knowledge society*
Computational thinking	*Identifying general principles and patterns in data, processes, or problems, effectively explaining the purpose and meaning of problems and their potential computational solutions*
Social media literacy	*Critically assessing and developing content that uses social media forms, and leveraging these media for persuasive communication*
Information security awareness	*Realizing the consequences of revealing personal information on the web, and taking appropriate actions to protect personal information from misuse and unwanted dissemination*

Four categories shaped the WM4.0 toolbox: hard skills, soft skills, usability and operability, and work environment. Each category has three application fields depending on a specific category such as IT (Information Technology) and business process knowledge for hard skills category and social competences for soft skills category. In each application, there are five development phases towards Industry 4.0 vision where the highest level of the development phase represents the vision of Industry 4.0.

Mourtzis [38] derived competencies required for Industry 4.0 based on literature review of related technologies and categorized them into four groups: technical, methodological, social, and personal. The author associated these groups of competencies and their level of importance (low, intermediate, and major) to the employees' roles in the enterprise, i.e. technical workforce, production engineer, and executives. For example, the technical competences group is at low importance level for executives. The technical competences group includes equipment operation, process understanding, and knowledge interpretation. Methodological competences group includes creativity, decision making, and problem solving. Social competences group includes communication, knowledge transfer, and leadership. Personal competences group includes responsibility, motivation, and flexibility.

Kusmin et al. [39] proposed a visualization diagram for competency management platform that can be used as a communication channel among the stakeholders interested in preparing and upskilling the workforce for industry 4.0. Stakeholders involved in the competency management process includes individuals, employers, governments, policy makers, education, and training providers. In this respect, businesses would be able to assess their employees' competencies and identify the gap based on their needs and market trends. Furthermore, academic and training organizations would be able to visualize future competency and estimate the relevant trends. In addition, governments can direct their initiatives toward needed competencies. Likewise, a case study was conducted in an ICT firm where stakeholder interactions across competency management have been studied. Here, the concentration was on the workers to assess their own competencies as well as their team members based on the model presented in [24] (which was discussed earlier). The firm could identify the gap and concluded that the inclusion of workers would result in achieving innovation in the workplace designing process as recommended by [40]. The authors support the idea where the responsibility for upskilling and "lifelong learning" seems to be moving more and more from educational institutions and companies to individual workers.

Lupicka and Grzybowska [41] tried to answer the question of what competencies required for industry 4.0 managers. Based on literature review, the authors identified three categories of core managerial competencies each composed of number of skills and ability. Then a survey was conducted amongst selected practitioners' experts, researchers, and students to rank these skills and abilities. The three categories: technical, managerial, and social and their dependent skills and abilities ranking are shown in Figure 3, Figure 4, and Figure 5 respectively.

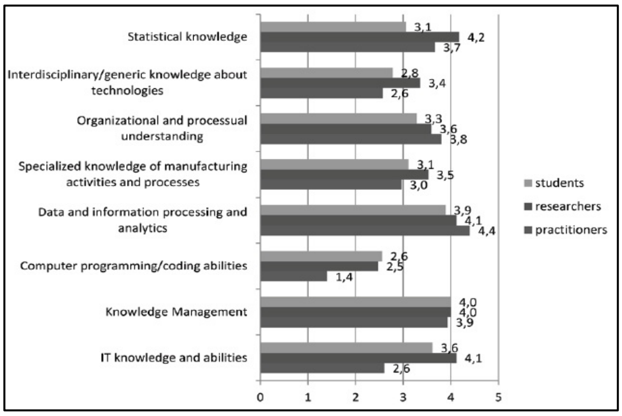

Figure 3. Technical competencies [41]

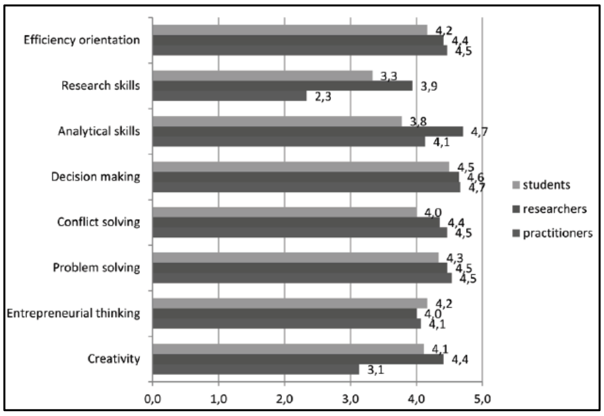

Figure 4. Managerial competencies [41]

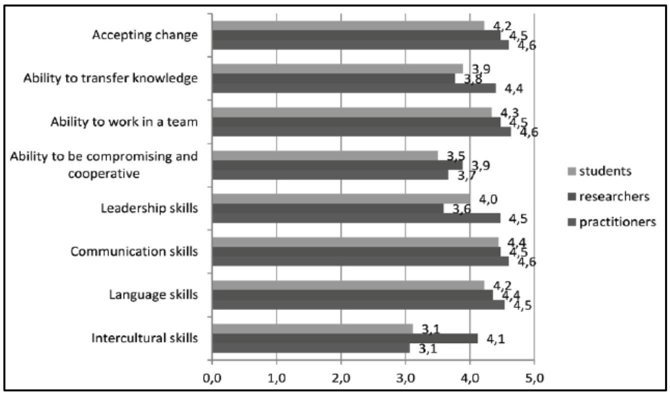

Figure 5. Social competencies [41]

Acerbi et al. [42] developed an assessment tool for skills 4.0 based on maturity models. The tool evaluates soft and hard skills, and knowledge about ICT for managers and operators working in manufacturing companies. It is a self-assessment tool with five levels: proficient, competent, practiced, aware, and basic. Proficient indicates that a worker is completely able to manage the emerging technologies, willing to improve his/her capabilities, and being able to use tablets, PC (personal computer), and the installed software. Basic indicates that a worker is not aware of the majority of new technologies, not willing to improve his/her capabilities, and not familiar with the use of tablets or PC. The assessment is achieved by conducting a questionnaire tailored to each job profile on three dimensions: soft skills, hard skills, and knowledge about ICT. The authors then applied the tool to a real case assessing ten different job profiles such as production manager, maintenance supervisor, warehouse operator, production operator, and data scientist. The results indicate that, all the job profiles are just about the third level in soft skills compare to knowledge about ICT where they are around level five except the warehouse and production operators being around the third level.

Pinzone et al. [43] developed a list of technical skills related to five organizational areas within Industry 4.0: operations management, supply chain management, product-service innovation management, data science management and IT-OT (operational technology) integration management. The identification of the organizational areas and their related technical skills was based on literature review. Technical skills classifications for the manufacturing field were analyzed based on classifications provided by national and international bodies [44][45]. Technical skills were used as input for focus groups consist of recognized manufacturing experts to enhance and specify the initial list of skills. Finally, interviews with technology providers and manufacturing companies' stakeholders were performed to check and refine the technical skills according to the five organizational areas.

Bensghir et al. [46] discussed the necessary qualities for future workers adapted from [47] where the qualities categorized in two main groups; personal and technical. In each group, there are three levels of qualities. The first level includes qualities that workers must have such as IT knowledge which belongs to the technical group. The second level includes qualities that workers should have such as mindset for continuous improvement and lifelong learning and categorized as personal. The third level includes qualities that workers could have such as specialized knowledge about technologies which belong to the in the technical group. The authors predicted that industries are going to need individuals who have adequate level of education and experience in social science due to interdisciplinary and multicultural workplaces. Consequently, the authors stated that advanced education institutions are able to provide individuals with these qualities. It was also emphasized that the main labor's tasks are going to be related to supervising and regulating the automated complex processes and the applications executed by machines [48].

Bermúdez and Juárez [49] conducted a study to identify the required competencies of operational management personnel for 10 automotive part suppliers through a literature review. They ranked these competencies in terms of importance, opportunity, and strength by performing qualitative study which also includes 15 operation managers. In addition, the authors supported the idea by Lorentz [50] which points out that "soft" skills will be more important for employees due to the need for high flexibility to adapt new roles and work environments and becoming familiar with constant interdisciplinary learning to increase communication and trust.

The authors in [50] highlighted that companies in the manufacturing sector will face major challenges in transition to Industry 4.0 due to the excessive use of robots and software. This demands new skills obtained by workers such as high qualification in information technologies, analytics, research and development. Therefore, the workers are able to complete tasks associated with data-based quality control, robot-assisted production, automotive vehicles for logistics use, intelligent supply network, predictive maintenance, and self-organized production.

Bongomin et al. [51] identified required skills for Industry 4.0 through a broad literature review in the field of engineering, production, and management. The identified skills in their study are divided into technical and personal skills. The technical skills are subdivided into theory, hardware, software, and algorithm skills. The authors also supported the idea that there will be transition in job demand from lower-skilled to highly-skilled jobs [52]. This study emphasized that there is still a need for human skills for tasks that artificial intelligence (AI) is unable to perform [53]. Therefore, to protect most jobs, individuals and organizations have to invest in improving the necessary skills via higher education and advanced training [54]. In this regard, higher education institutes specially the universities have to prepare their graduates for the transforming workforce landscape [55][56].

Maisiri et al.[57] explored the required skills for Industry 4.0 in the engineering profession by performing a systematic literature review and then categorized the result in two main categories; technical and non-technical (soft skills). Each category divided into subcategories and then skills set as shown in Table 3. Among these skills sets, the top three soft skills required in industry 4.0 employees as stated by [58] are creativity, emotional intelligence, and proactive thinking where they grant the employee to adapt the incremental changes in Industry 4.0 environment. Another aspect brought by [59] and endorsed by [57] is the fact that collaboration of humans and robots will be inescapable to improve productivity in industry 4.0. In addition, Intelligent machines cannot apply common-sense reasoning or show empathy where it needed to increase productivity and that can be performed by humans [60].

Table 3. Skills required for I4.0 [57]

Skills Category	***Skills sub-category***	***Skills set***
Technical skills	*Technological skills*	*Designing skills that incorporate virtualizing, simulating, interoperability, modularizing, decentralizing capabilities.*
		Fault and error recovery skills
		Application and use of technological skills
		Process digitalization and understanding
		Ability to work with the internet of Things, autonomous robots, 3D printing, and other advanced technologies. Interaction with modern interface
	Programming skills	*Computational skills*
		Simulation skills
		Coding
		Computer and software programming skills Software Development
	Digital skills	*Data analytics / data processing*
		IT/data/cyber security
		Cloud computing skills
		IT knowledge and abilities
		Artificial intelligence skills
		Digital content creation skills
Non-technical skills/soft skills	*Thinking skills*	*Creativity, innovation, practical ingenuity*
		Critical and logical thinking
		Flexibility
		Complex problem solving, troubleshooting
		Analytical thinking skills
		Technical and literate communication
		Collaboration (including machine-human)
		Interdisciplinary skills
	Social skills	*Teamwork*
		Perspective-taking
		Professional ethics
		Understanding of diversity
		Self-awareness, self-organization
		Interpersonal skills
	Personal skills	*Social responsibility and accountability*
		Lifelong learning skills
		Leadership skills/people management
		Emotional intelligence
		Negotiation skills
		Entrepreneurship
		Adaptability

Based on the reviewed literature, it is clear that most of the researchers agree on dividing the skills needed for industry 4.0 into three main categories: technical, managerial/methodological, and social. Some research studies divided the social category into interpersonal and intrapersonal, while others consider personal as subgroup of social. Two of these studies identified the skills needed for Industry 4.0 along with the level of each skill.

There are two important research outcomes from [24] and [41] are used in this research. The modeling in this research will utilize the chart presented in the former to select which skills and skill level are needed for each job. The latter focused on what competencies required for industry 4.0 managers.

2.2.2 Job Assigning

Tsarouchi et al. [61] proposed a method for tasks assigning to a human and robot in separate workstation. This method consists of a set of decisions based on an evaluation of three criteria: (i) resource suitability to ensure that a resource is suitable to execute the task, (ii) resource availability, and (iii) operation time where a task will be assigned to resources with minimum operation time. A resource could be a human or a robot. The decision on suitability criteria is binary. The authors did not provide details about what considered in this criterion nor the evaluation method.

Fiasche et al. [62] proposed a model to maximize the match between worker capability and job requirements where a worker who has suitable capabilities for the job will be selected. Besides, the model considers the worker's preferences to perform a job. The model assumes that that each worker is described by a certain level of physical, sensorial, and cognitive capacities and a defined level of proficiency associated with each knowledge and skill. Moreover, each worker has defined preferences for the jobs he/she would like to perform. Similarly, each job requires a certain level of physical, sensorial, and cognitive capacities, and relevant knowledge and skills. However, the authors did not discuss how a job requires physical, sensorial, and cognitive capacities nor the type of scaling used to determine these skill levels.

Campbell [63] developed a two-stage stochastic model for scheduling and allocating cross-trained workers (i.e. those who trained to work for more than one department or task) in a multi-department with random demands. The model consists of two stages; (1) dealing with scheduling and (2) dealing with job assigning. Dividing the model into two stages is based on a framework for workforce planning and scheduling decisions developed in [64] and structured as planning, scheduling, and allocation. The proposed model formulation is based on a model developed in [65] but it was extended to multiple departments and time period. The first stage covers scheduling over a time horizon where the model considers the cost of workers, and the optimization objective is to minimize the cost while maximizing the scheduling of required workers over a period of time. The second stage covers job assigning to allocating available workers in each department to accommodate demands where the model objective is to maximize allocation of workers considering worker daily productivity. The productivity of workers is given as a value between 0 to 1, but it is not clear how that value is determined and based on what factors.

Digiesi et al. [66] proposed an optimization model to balance the workload when adopting job rotation to reduce the ergonomic risk mainly musculoskeletal disorders (MSDs) risk and to meet required production rate. The model focuses on repetitive assembly tasks and use the RULA (Rapid Upper Limb Assessment) method as risk evaluation tool where the total risk level of each worker when perform different job during planning period must be under a certain level. For the production part, the model uses a production factor for each worker that depends on worker skills and age, but the authors did not provide details about the evaluation methods of that factor.

Norman et al. [67] proposed a worker assignment model that maximize profit that based on three components: productivity, quality costs and training costs. The model considers workers' skill with different skill levels. The authors divided the skills into two categories; technical and human skills. There is a given weight for each category in the proposed model. Each job has a necessary skill level of certain skills where workers must have at least this skill level value to perform a job. Productivity and quality components in the objective function are associated with a worker's skill level required to perform a job. A worker who has the highest level of skills required by a job, his/her productivity and quality level will get a value of 1.0, otherwise it will be less than 1. The authors did not discuss the determination of the productivity and quality level, and it is assumed as given values.

Akyol and Baykasoğlu [68] proposed a solution for the assembly line worker assignment and balancing problem (ALWABP). This type of problem occurs when the execution of every task is varied depending on workers. Therefore, the primary objective function to model this type of problem is to minimize the cycle time. The authors used a model for this problem proposed by [69]. They solved the model by using multiple rule based constructive randomized search (MRBCRS) algorithm. The algorithm considers 39 task priority rules and 4 worker priority rules used to sequence tasks and select workers. The rules in the approach are obtained from [70] and [71]. In their problem-solving approach, the suitability of a worker to perform a task is based on processing time. The worker processing time of a job is assumed to be known and it is not determined by the model.

Weckenborg et al. [72] developed a mixed-integer programming model for balancing and scheduling of assembly lines with collaborative robots. In [72], a task can be performed in three ways: by human only, robot only, or both in collaborative way. The authors endorsed the idea that humans and robots are complementary to each other and not competing each other. This idea was also indicated by [73] and [74] where human has characteristics like adaptability, flexibility, decision making skills, and creativity while robot has attributes like strength, endurance, speed, and accuracy. In their model, the authors focused on part scheduling considering task flow where predecessor task finished before starting the next one. In essence, that part of the model is an extension of formulation that has been proposed in [75] and [76]. The model objective is to minimize the cycle time. The main assumption here is that the task can be performed by either human, robot, or collaboratively. It is to be noted that no information was mentioned about the task nature and if its required certain ability such as strength where robot become the best choice. Instead, the main factor to assign the task is the execution time. Therefore, the task will be assigned to whoever can perform it in the shortest time either by human, robot, or collaboratively and the execution time assumed to be known.

Gomar et al. [77] developed a linear programming model to optimize the allocation of multiskilled workers to certain jobs and tasks in construction projects. Based on a study conducted by [78], the authors assumed that the success of multiskilling relies on the manager's ability to assign workers to appropriate tasks. Their model was named Multiskilling Optimization Model for Allocation (MOMA) and focused on minimizing the total number of workers, switching, hires and fires. However, the parameters value used for switching, hiring and firing in the objective function are not clear. Moreover, the model simulation results show that a worker can be assigned for a job for certain time because he/she obtained a particular skill. It seems the model assumed a job will require only one skill and not a combination of skills. Moreover, it considers only whether the worker has the skill or not. This model has not considered the level of skill obtained by a worker.

Lian et al. [79] proposed a model to solve a multi-skilled worker assignment problem in the context of Seru production systems. This refers to a type of cellular manufacturing with the purpose of deconstructing long multi-product conveyor assembly line in order to improve flexibility in order to consider differences in workers' skill sets and proficiency levels. The authors insisted on production flexibility and advocated other authors' opinion who argue that the due to the changing nature of products in industry 4.0, production systems need to obtain a high-level flexibility and agility [80]. In their literature review, the authors concluded that several studies considering workers differences from one another in the skill set, but their skills level are defined as binary parameters where a worker either perform the task as specialist or cannot perform it at all. However, in their model, skill level is ignored too, i.e. they considered one skill per task. They solved the model by letting the product to require more than a task/skill. In terms of proficiency level, the authors measure workers' skill level by comparing the time he needs to perform a task compared to the standard time. In essence, measuring the time each worker needs to perform a task may not be a good representation of their skills.

Savino et al. [81] addressed the problem of disruption occurrence in workforce assignment and scheduling in U-shaped assembly lines where the number of operators is less than the number of workstations. The authors argued that the workforce assignment usually addressed with centralized approaches, which fail when dynamic events occur, forcing to re-schedule the workforce several times. Therefore, the proposed approach consists of a two-step's procedure. The first step involves a centralized scheduling process based on a constraint optimization problem (COP) for initial operator scheduling, which was introduced by [82]. This step aims to maximize throughput considering different workstations with different capacity to perform jobs, number of jobs in different buffers, different shifts, and number of operators whom skills level is neglected and considered in a same level of proficiency. Second, a decentralized algorithm performed by a multiagent system (MAS) to manage workers in case of unforeseen events mainly workstation failure or operator missing addressed with centralized approaches. The MAS can play a key role for managing disruptions because agents (Operators and Workstations) are able to interact, cooperate, and negotiate tasks dynamically by acknowledging their current status.

The problem is approached in two main steps. First the COP will initiate scheduling of workers to all the workstations as a centralized solution (CS). Second, a decentralized solution (DS) is performed with MAS to support the CS in case of unforeseen events. Therefore, in the case of machine failure, if more than one machine is available, the operator is assigned to the machine with the longest waiting time. If there is no such waiting machine, the operator will be assigned to the next available machine. The authors could validate their approach by comparing the throughput when applying only the CS and also with the support of DS. They conclude that applying both CS and DS increases the throughput.

Ayough et al.[83] proposed a job assignment and rotation model considering human characteristics. In particular, workers boredom was considered. The authors in [84] advocate the opinion that integrating human factors into job assignments is essential where physical/cognitive human characteristics and behaviors are being considered. The model rotates workers during a given planning horizon with an objective to minimize total assignment boredom and cost. Their model was built based on a model proposed by [85] that minimizes the workforce cost and the level of boredom due to assigning the same jobs to the worker during the planning horizon. However, the authors investigate further about boredom and defined it as function of two jobs similarity (‘one if identical, zero if totally different’) and if performed by a worker in two consecutive periods. Assigning cost was not defined in detail, instead a predetermined cost value is provided to each worker for each job they undertake.

Mossa et. al [86] developed mixed integer nonlinear programing model to find optimal job rotation assignment for manual tasks with a high frequency of repetition that maximizing the production rate, reducing and balancing human workloads and ergonomic risk. According to [87], job rotation is the most common labor flexibility method in the case of repetitive assembly tasks aiming to increase productivity and smooth workload and to deal with the related ergonomic risk among employees. During a shift of eight hours there will be five breaks and workers will be rotated after each break time. The model considered workers' performance based on skills level and measured it based on task completion time. It is evident that lower skill level will result in longer task completion time. The skills levels are assumed as three levels; high (1.00), medium (1.15), and low (1.25). It was given as direct input to the model, i.e. not based on calculations. The authors applied the model to an industrial case study and concluded that the model has an ability to identify optimal job rotation assignment achieving productivity and ergonomic risk goals.

Moussavi et al .[88] proposed an assignment model with an objective to minimize the total production time by assigning workers with different skill levels to workstations based on the worker's task executing time. The task executing time depends on the workers' efficiency. The authors used the procedure proposed in [89] to calculate worker's efficiency. The efficiency based on utility function is used to calculate accumulative value of four contributing factors: skills, height, age, and experience level of workers. In the utility function, there are three intervals value for each factor; forbidden "zero" where the level of worker is considerably below or higher than the acceptable range required by the job; excellent "one" where the level is within the range; and acceptable "below one" where the level of worker is slightly below or higher than the upper level. In this model, the authors applied skills factor only and used the worker's efficiency value to determine the executing time for each task. The model also considered workers' availability and the possibility of job rotation. However, this model deals only with one type of skills and neglects the variety of skills where in reality different tasks require different set of skills. Moreover, the rotation procedure was a relatively simple approach which is based on executing the model after reducing the worker availability in order to get different assignment combinations.

Bentefouet and Nembhard [90] studied the impact of workers variability in performance on production output and characterize the optimal switching time between workers in worksharing systems. The authors claimed that the influence of human behavior on production systems has been underestimated while the most attention focused on technological aspects. The authors developed their study based on mathematical analogy of two workers with two and three tasks. Their results suggested that when workers switching and cross-training policies are implemented, it is feasible to achieve both higher flexibility and higher production output where a calculus rule could be used to determine the optimal switching time.

Xin et al. [91] proposed an assignment model considering cooperative multi-skilled workers with objective to minimize human cost and to balance workload of both assembly stations and processes. The authors believe that the multi-skilled worker is a key element in assembly production system [92]. They also mentioned that in order to improve the productivity and quality of that system, there should be cooperation between multi-skilled workers [93] where workers can help each other, and the workload should be distributed. Since workers have variation in their experience, knowledge, and training, there will be difference in workers' skills [94] and that should be considered in the assignment procedure.

The authors' core idea was to divide workers into 3 types: elementary, intermediate, or senior-skilled worker based on their work experience in order to provide a quantitative relationship between the cycle time and multi-skilled workers. Then, the model will decide which two workers need to be combined to perform cooperatively one of the station's processes that will result in reducing cycle time and balancing the load for that station. Therefore, elementary-skilled worker must cooperate with at least one intermediate-skilled or senior-skilled worker to learn and avoid making mistakes. However, the authors only considered experience as the main factor that differentiates workers' productivity, and they ignore other personal factors like cognitive ability and learning rate.

Liu et al. [95] developed a mixed integer linear programming assignment model with an objective of minimizing both makespan and total flow time. Makespan is defined as the completion time of a machine with the longest execution duration. Flowtime is the sum of all jobs' finalization times. The model considers multiskilled worker with different skills level in a hybrid flow shop. The authors claim that multiskilled worker assignment problem has been studied in the literature without considering different skills level. Moreover, since workers have different skills and skills level, job processing time will differ according to workers performance. In their model, the authors assume that each job has a standard processing time on a specific machine, but the actual processing time of that job is based on worker efficiency which is determined by skills level. They divided workers' efficiency into five levels. However, the authors did not consider individual skill level nor the importance of that skill for the production system. In other words, they ignored the diversity of skills acquired by each worker and assumed that all workers have the same skill set but differ in skill levels.

Wu et al. [96] developed a mathematical model to assign workers with different skills level to various jobs in divisional and rotating Seru production system in order to maximize throughput and balance the workload. In a divisional Seru, jobs are divided on stations and assigned to several multi-skilled workers, so a worker performs one or several jobs. In rotating Seru, each worker has to perform all the jobs starting from the first station to the last one and rotate again. According to [97], worker skills and the required jobs are the two factors that influence the throughput of Seru production system. The authors supported [98] the impression that multiskilled workers in Seru production systems should possess the capability of handling both technical and managerial jobs.

According to [99], workers with higher comprehension abilities should be selected for complex Seru. However, the authors in their model considered workers' skills level variation to not only maximizing throughput but also to balance the workload. The authors assume that workers' skills levels are different for different jobs. Hence, the jobs processing time will vary according to the workers' skills level. However, they did not explain how a worker's skills level could vary according to the job. Moreover, the job processing time assumed to be a function of the worker's skill level ratio, yet the calculation of this ratio was not explained.

Wu et al. [100] developed a mathematical programing model with an objective to minimize training cost and maximize the balance of workload by assigning workers to tasks based on workers' skills level and available working time in one shift. The authors emphasized on ideas presented in [101] where skill types and workers' skill levels should be considered in worker assignment. The model could assign a worker who has higher skills level to perform the task compared to a standard skill level in order to minimize the difference between those levels which result in lower training cost. The training cost was calculated based on that difference. The worker's available time is considered to balance the workload. Even though the authors are advocate of considering different types of skills in assignment model, in their model they did not consider that. Instead, they define the skill as worker proficiency of performing a task.

Samouei and Ashayeri [102] presented an assignment model for assembly line considering three types of operators: robots, workers, and workers with assistant robot. The model objective is to minimize cost and cycle time. The task processing time and execution cost are based on the operator's skills level. Tasks could be performed by robot or a human worker with various skill levels (low, medium, or high) alone or with assistant robot. The authors did not explain how the value of the task processing time and cost is determined. They assumed those as given and known values.

Based on previous studies, it is clear that most of the published research dealt with worker skills either as only one type or different skills but with only one level. Some research studies emphasized on the importance of considering different skills with different levels but in their assignment model. In essence, those studies ignored differences in skill type and consider the skill level as fixed number (or given input) by managers. A summary of the reviewed models is presented in Table 4.

It should be also noted that none of the previous studies consider real-time data that affect workers performance. A worker has a certain level of different skills and he/she is expected to perform in certain performance levels. Nevertheless, in reality, the real time factors affect worker performance. Physical and mental health is an example of the real-time factor. All the previous studies have not considered real-time physical and mental health when assigning workers to perform jobs. Therefore, real-time data should be used to assign workers to jobs where the production rate is affected by the performance of those workers.

In addition, suitability of a job to worker should be considered. In other words, a worker may perform well in one type of job compared to another one even though both jobs required same skills level. Thus, there should be a link between worker performance level and his/her current real-time data. The question now is which real-time factors that affect worker's performance should be considered and can be measured during performance. One of the technologies that Industry 4.0 has provided is industrial internet of things (IIOT). In particular, personal wearable devices could provide real-time data about worker's status. In the next section, studies related to the use of these device and their impact on performance improvement are reviewed.

Table 4. A summary of reviewed assignment models

Ref.	Objective	Skills included*	Different Skills	Skill level	Production Measures*	Comments
[61]	Min tasks execution time	×	×	×	×	By assign task to the available faster (worker/robot)
[62]	maximizes worker-job fit	✓	✓	✓	×	By reduce difference between chosen worker qualification lev and the required by job

[63]	to minimize the worker cost & maximize allocation of workers considering worker productivity	×	×	×	Cost and productivity	productivity of worker is given as a value between 0 to 1, no data about how that value is determined
[66]	balance the workload	✓	×	×	Meet demand	adopting job rotation to reduce the ergonomic risk mainly musculoskeletal disorders MSDs risk and to meet required production rate. production factor for each worker that depend on worker skills and age, (given value)
[67]	maximize profit that based on: productivity, quality costs and training costs.	✓	✓	✓	Productivity and cost	worker who has the highest level for skills required by a job, his/her productivity and quality level will get a value of 1.0, otherwise it will be something less than 1.0. Given value not calculated
[68]	to minimize the cycle time	×	×	×	cycle time	The worker processing time of a job assumed to be known not determined by the model.
[72]	to minimize the cycle time where a task can be performed by human only, robot only, or both in collaborative way.	×	×	×	cycle time	The operator processing time of a job assumed to be known not determined by the model.
[77]	minimizing the number of workers, switching, hires and fires	✓	✓	×	Meet demand	the model assumed a job will requires only one skill not a combination of skills
[79]	minimizes the deviations of the average workload per serus (Station) and per worker	×	×	×	Meet demand Cell/station capacity	Worker proficiency level measure by comparing the time he needs to perform a task with the stander time of that task

[81]	maximizing throughput	×	×	×	Cell/station capacity	By proposing Reassignment approach in case of machine failure or worker missing using multi agent system where both machine and worker acknowledge their availability.
[83]	To minimize workers' boredom level due to assigning the same jobs to the worker during the planning horizon	×	×	×	×	By rotates workers during a given planning horizon. Boredom is defined based on similarity of two jobs and if performed by a worker in two consecutive periods.
[86]	maximizing the production rate by to finding optimal job rotation assignment for multiple periods.	✓	×	✓	production rate Cell/station capacity	workers performance based on skills level and measured based on task completion time
[88]	to minimize the total production time	✓	×	✓	cycle time	task executing time depend worker's efficiency which is based on the workers' skill level.
[91]	to minimize human cost considering cooperative multi-skilled workers where workers can help each other, and workload is distributed	✓	×	✓	cycle time cost	the model will decide which two workers need to be combined to perform task cooperatively where workers are divided into 3 types based on their work experience
[95]	minimizing both makespan and the total flow time	✓	×	✓	total flow time	processing time of the job is based on the worker's efficiency which is determined by his skills level (5 given levels)
[96]	to maximize throughput and balance the workload in divisional and rotating seru production system	✓	×	✓	throughput	workers' skills levels are different for different jobs, so the jobs processing time will vary according to the workers' skills level. (given)

[100]	minimize training cost and maximize the balance of workload by assigning workers to tasks based on the worker's skills level and available working time in one shift	✓	×	✓	×	The model will assign worker who has higher skills level to perform the task compared to a stander skill level in order to minimize the difference which result in lower training cost as the training cost calculated based on that difference
[102]	to minimize cost and cycle time. considering three types of operators: robots, workers, and workers with assistant reboot	✓	×	✓	cycle time cost	The task processing time and execution cost are based on the operators type and skills level in case of an operator is a human worker (3 levels), assumed as given value.

*Skills: skills/qualification/capabilities.

**Production measures: productivity, stations capacity, cost, revenue, cycle time

2.3 Real Time Monitoring and the Use of Wearable Devices

Mattsson et al. [6] investigated empirical ways through case studies, laboratory tests, workshop, and literature review to realize how advanced physiology measurement technologies can be incorporated into an industrial application to increase worker's performance and wellbeing at assembly stations. The study focused on how to assess worker wellbeing in real-time. The study revealed that smart devices have the ability to boost job satisfaction, reduce complexity and errors, and affect performance by providing visible cues to managers to match tasks to workers [7]. Moreover, these smart devices are useful not just for personal well-being, but also for improving operator performance [8]. It has been shown that performance is correlated with psychological wellbeing and health [9][10]. High performance has been linked to a high level of job satisfaction as well as a high level of psychological well-being [11]. Individuals with diverse knowledge and capabilities in their work settings are more likely to experience work-related stress when the work demand is not matched with their own capacities[12][13]. For example, boredom and under-incentive are considered as factors that influence operator performance [14].

The literature suggests that Electro Dermal Activity (EDA), Blood Volume Pulse (BVP), and Heart Rate Variability (HRV) can be used as useful physiological markers for determining operator wellbeing in an industrial setting. These physiological measurements can be used to assess operator wellbeing in real-time quickly and reliably. According to Can et al.[103], researchers have begun to use industrial smart wearable devices for detecting individuals' stress from physiological signals. The most used physiological signals in the literature are HRV and EDA[104].

According to [105], operator's well-being at work can be assessed by smart wearables (i.e. body-sensors for assessing changes in the operator's cognitive states based on skin conductance, blood-pressure, heart rate, breathing, and/or temperature measurements or by assessing eye-movement). Romero et al. [106] believed that cognitive workload can be managed by addressing the cognitive tasks assigned to the operator in order to maintain an ideal 'stress level'. This could be accomplished by associating new variations to an operator's routine whose vital data indicate she or he is bored with the current routines and there is a need for new task assignment.

Khakurel et al.[107] conducted a systematic literature study in order to present a heuristic overview of contemporary wearable technologies and evaluate their potential. Their findings demonstrate that wearable technologies have the potential to boost employee productivity, improve physical well-being, and prevent work-related injuries. These findings were also confirmed in [108]. The authors concluded that the most significant advantages of wearable devices in a workplace are related to employee's health, workplace safety and improving work performance.

Battinin et al. [109] discussed the use of Industry 4.0 technologies mainly wearable devices to design assembly processes considering workers physical health and production efficiency. The main idea in this research was to collect energy expenditure data while workers perform tasks and then using the collected data in the assembly system design phase. This could help to identify which tasks are assigned to each workstation and how to perform workstation sequencing in the assembly system. It was mentioned that heart rate (HR) can be used to evaluate people fatigue in terms of energy expenditure [110]. In this way, smartwatches were used as wearable devices to monitor workers' health and collect data about the physical efforts via counting the heartbeat per minute. The above-mentioned works proved that Industry 4.0 technologies such as wearable devices can be used as guidance in work-cell design phase to achieve ergonomics aspects and human centric concept.

2.3.1 Using Heart Rate Variability to Assess Stress, Fatigue, and Mental and Physical Health.

Jebelli et al. [111] developed a framework for measuring workers' stress in non-invasive way by evaluating changes in workers' physiological signals collected from a wearable biosensor (a wristband-type). As workers encounter varying levels of stress during work, the framework collects patterns of physiological data and learns these patterns using a supervised-learning algorithm. Measured signals were HRV, HR, EDA and skin temperature (ST). The authors rely on previous studies which indicated that higher stress might lead to lower HRV [112][113][114]. To examine the performance of the proposed framework, the authors collected physiological signals of 10 construction workers while working under different conditions and measured their cortisol levels. Cortisol is strongly linked to an individual's stress level and known as a stress hormone[115][116]. The authors could validate the proposed framework's ability to measure employees' stress using support vector machine (SVM) algorithm with an accuracy of 84.48 % in predicting stress levels between low and high, and 73.28 % in predicting stress levels between low, medium, and high.

Gancitano et al. [117] assessed the changes in the HRV parameters of 112 personnel of male special force while engaging in several duty tasks by using wearable devices They processed the HRV data via a multivariate linear regression analysis. The comparison between different task groups showed that activities with a high demand for concentration, precision, and acute stress - as is the case with paratroopers and dynamic precision shooters - differ significantly from activities that can be defined as routine, such as office work. The authors conclude that studying HRV parameters is a useful tool when used by occupational physicians for addressing work suitability assessments. Additionally, their results provided insights on how the use of wearable devices can help control stressful stimuli in order to reach a sustainable development considering all worker categories when exposed to various work-related stressful conditions.

Izzah et al. [118] developed a machine learning model to detect cognitive load using HRV. The HRV features were evaluated from 30 subjects during rest, two cognitive tests (d2 Attention and Featuring Switcher Task both labeled as cognitive load), and recovery. Both rest (baseline) and recovery period are labeled as rest. HRV features derived from RR Intervals which is the time between two successive heartbeats measured in milliseconds (ms) based on five minutes recording for both time and frequency domains include; Mean HR, standard deviation of normal-normal intervals (SDNN), root mean square of successive differences (RMSSD), Percentage of successive normal-normal intervals that differ by more than 50 ms (pNN50), power in low frequency range (LF) ms^2, power in high frequency range (HF) ms^2, and LF/HF ratio. Five classifier algorithms were selected: linear support vector machine (LSVM), radial basis function (RBF) kernel (RBF-SVM), k-nearest neighbor (KNN), random forest (RF), and Naïve Bayes. The accuracy to predict two classes (distinguish between rest and cognitive state) range from 0.54 to 0.62, with LSVM, showing the best.

Matuz et al. [119] used dataset of three experiments (85 participants total) where each applied different cognitive task for fatigue induction to develop a machine learning model to predict fatigue level (regression model) and status (classification model) based on HRV. The

HRV features (20 features including SDNN, RMSSD, pNN50, LF, HF, and LF-HF) were evaluated based on different time windows ranging from 1 minute to 5 minutes. Visual analogue scale (VAS) was used to assess the fatigue level where participants indicated their actual experience of fatigue on a 100 mm horizontal line with “No fatigue at all” and “Very severe fatigue” printed on the left and right end of the line, respectively. For fatigue status (fatigue and non-fatigue), this binary classification problem was formed in two ways; resting data and task-related data. In resting data, the pre-experiment resting HRV data was labelled as “non-fatigue” and the post-experiment resting HRV as the “Fatigue”. In task-related data, the beginning period of task (the first 1–5 minutes) performance was labelled as “non-fatigue” and the end of the task (the last 1–5 minutes) was labelled as “Fatigue”. Elastic net regression and the least absolute shrinkage and selection operator (LASSO) were selected as regression algorithms where Elastic net achieve the best performance with Root Mean Squared Error (RMSE) = 18.73 in 5 minutes HRV time windows. For classification at resting data, both SVM and KNN achieved an accuracy of 70.1% in 5 minutes time windows compared to RF with 69.1% accuracy. In task-related data, SVM model achieved an accuracy of 76.1% in 5 minutes time windows compared to 74.7% and 73.3% accuracy for KNN= and RF =, respectively.

2.3.2 Heart Rate Variability to Predict Performance

Tsunoda et al. [120] developed a framework to predict cognitive performance using only HRV features. An experiment conducted for 45 participants to perform ATMT (Advanced Trail Making Test Task) task where they asked to use a computer and click numbered circles in ascending order from 1 to 40 as soon as they as can. During the task, HRV of the participants were collected in addition to their score in the task. In their previous study [121], the authors found a correlation between changes in cognitive performance and changes in HRV features.

In another study, they also discovered a correlation between the level of cognitive performance and the level of HRV features [122]. These findings suggested that changes in HRV features contribute to changes in cognitive performance. Therefore, the authors developed a prediction model that uses HRV features to predict ATMT scores level (two classes; higher or lower than the first trial). They used 17 HRV features based on 5 minutes recording at rest and then moving the window of 5 minutes during the task. The model accuracy was reported at 84.4 % using SVM algorithm combined with Principal Component Analysis (PCA) as a technique for reducing data dimensionality to find the most significant features in a dataset). It was concluded that the framework has the potential to predict worker's performances and assist managers in proactively suggesting break intervals or adjustments in job assignments to avoid productivity losses due to cognitive work performance declines.

Schaich et al. [123] examined the association between HRV and cognitive performance using three standardized tests: the cognitive abilities screening instrument (CASI), processing speed (Digit Symbol Coding-DSC), and working memory (Digit Span-DS). For CASI, DSC, and Digit Span, the score ranges are: 0–100, 0–133, and 0–28, respectively. HRV was computed as SDNN and RMSSD. The authors utilized linear regression models to analyze the data from 3018 participants in the study. The results indicated that SDNN was correlated with better CASI performance. RMSSD was correlated with better DSC performance, and SDNN was also correlated with better Digit Span performance. The authors concluded that higher HRV is generally correlated with better cognitive performance.

Kadoya et al. [124] investigated the relationship between worker's emotions and productivity by determining emotion status during work using wearable devices. The device could measure heart rate variability and used an algorithm to differentiate various emotions based on HRV pattern. According to Choi et al. [125], heart rate variability may be utilized to accurately assess human emotions.

Heart rate variability might be an objective technique to measure emotional reactions [126][127]. Fifteen plastic toy painters participated in this study where their physiological responses were captured, and their productivity measured by a manager who tracks the number of finished toys by each worker at each shift. The result indicated that happiness among other emotional states such as anger, relaxation, sadness, and neutrality, was found to be strongly and positively associated with productivity. To assure high productivity, the authors suggested that workers' emotional states should be addressed as a part of an organization's operational plan.

Grässler et al. [128]conducted a systematic literature review of 27 studies filtered from 493 studies over five databases (Scopus, Cochrane Library, Web of Science, Medline, and Pubmed). The aim of their study is to review the association between resting HRV and the cognitive performance where five cognitive functions were analyzed separately: executive functions, memory and learning, language abilities, visuospatial functioning, and processing speed in addition to the global cognitive function. In their studies they only considered two HRV features at reset period; RMSSD and HF. Except for processing speed, their result revealed that a significant correlation in each cognitive function was found, indicating a positive association between resting HRV and cognitive performance.

Forte et al. [129] reviewed fifteen studies with a total of 1051 healthy participants filtered from 438 studies. The aim was to analyze the relationship between HRV and decision-making. Their study indicates an association between higher vagally mediated HRV and better decision-making performance while a lower HRV is associated with lower performance in developing good decisions. HRV was evaluated based on five minutes recording except one study based on 3 minutes and at resting periods for seven studies while four studies at both resting and reactivity (during task). HRV features used in these studies are SDNN, RMSSD, and pNN50 in time domain and HF, LF, and LF/HF in frequency domain. Vagal nerves are the main nerves of parasympathetic nervous system and these four features of time domain indicate the vagal contribution to HRV. In frequency domain, LF is influenced by sympathetic and parasympathetic branches of the autonomic nervous system (ANS) reflecting a mix between the sympathetic and vagal influences. HF reflects vagal tone and the LF/HF-HRV ratio is possible to be considered as an index of the Sympathovagal balance.

Alharbi and Alomainy [130] developed a machine learning model to predict cognitive performance of different cognitive domains using four cognitive testes. First, Trail Making Test was used to assess cognitive flexibility, processing speed, and executive functions. Second, Fluid Intelligence Test is used to assess verbal and numerical reasoning. Third, Symbol Digit Substitution Test is used to assess processing speed. Fourth, Numeric Memory Test is used to assess working memory. To convert each test scores to categorized labels (poor, mediocre or high – or just two classes of poor or high), percentiles of each test score were used so that the number of samples in all categories was equal and the classification was balanced. Two HRV features of time domain were used; RMSSD and SDNN in 10-second Electrocardiograph (ECG) recording of resting period for 25,000 participants. Six machine learning algorithms were selected: LSVM, KNN, linear discriminant analysis (LDA), decision trees (DT), RF, and extra trees using SKLearn library.

The highest accuracy achieved was 53% using LSVM to predict three class of Numeric Memory and 66% using LDA to predict two class of Fluid Intelligence. After filtering dataset to include male who are 65 years old and above with tobacco and alcohol usage, the data was limited to 120 samples and the highest accuracy achieved was 82% using SVM to predict three class of Symbol Digit Substitution and 91% using LDA to predict two class of Trail Making part A. The authors believe that their work to employ machine learning to predict cognitive performance based on HRV data is the first in this field.

By reviewing the published literature in this area, it is clear that HRV is a proven marker that affect worker performance and also can be used to predict worker's performance as presented in [120], [123], [124] [130]. In this respect, there have been several studies about the HRV correlation to worker's performance. In addition, the most frequent HRV features used in these studies are SDNN, RMSSD, and pNN50 in time domain and HF, LF, and LF/HF in frequency domain at the resting period. A summary of studies that applied machine learning (ML) for prediction based on HRV data is presented in Table 5. As mentioned in previous sections, there are also many studies that provide job assignment model based on workers' skills. However, the studies that consider assignment model based on different skills with different levels combined real-time workers performance are absent in the literature.

Table 5. A summary of studies used ML for prediction based on HRV data.

Ref	Participants	HRV features	HRV evaluation period	Target	Classifier Accuracy
[130]	25 K	RMSSD and SDNN	10-second recording of resting period.	Cognitive performance of 4 different cognitive testes	SVM (53 %) for three levels and LDA (66 %) for two levels.
	120 (Above 65 years old)				SVM (82 %) for three levels and LDA (91 %) for two levels.
[120]	45	17 features of HRV including SDNN, RMSSD, pNN50, LH, HF, LF/HF	5 min At rest and moving window during the task	Performance level of ATMT (Higher or lower than the first trail)	SVM (84.4%) RF (63.9%)

[111]	10 Construction Workers	24 features of HRV including SDNN, RMSSD, LH, HF And other features from EDA and ST	1to 60 seconds	workers' stress level	SVM (84.48 %) for two levels (low& High) and (73.28 %) for three levels.
[118]	30	Mean HR, SDNN, RMSSD, pNN50, LF, HF, and LF-HF	five minutes recording for: (before and after task labeled as rest) and (during task labeled as cognitive)	Detect cognitive load (two classes; rest or cognitive state)	LSVM (62%)
[119]	85	20 HRV features including SDNN, RMSSD, pNN50, LF, HF, and LF-HF	Different time windows ranging from 1 minute to 5 minutes:	predict fatigue status (before task labeled as non-fatigue and after task labeled as fatigue)	Both SVM and KNN (70.1%) RF (69.1%) in 5 minutes, time windows
				predict fatigue level using Visual analogue scale (VAS) rand 0 to 100.	**(Regression model)** Elastic net RMSE = 18.73 in 5 minutes time windows

3. Methodology

This section presents the methodology to achieve the research objectives of developing an efficient job allocation model in the context of industry 4.0. This was initiated by identifying the right skills needed for I4.0. The methodology of matching workers who possess certain skills in order to achieve optimum production performance level is presented in a three-phases approach and elaborated in the following sections.

- Phase 1: Identifying skills and competencies selection model based.
- Phase 2: Developing a job assignment optimization model to maximize profit by matching workers various skills (and different skill levels) to available tasks. The model was coded in Python and is solved using Gurobi solver. Model limitations and possible room for improvements are identified.
- Phase 3: Transforming the job assignment model for I4.0 applications by engaging a real-time factor that affect workers performance and using machine learning prediction model. The objective in this phase is to develop a model that is able to use real-time monitoring data about workers status that would affects worker's performance. In this respect, the following are performed:

- Enhancing the model based on the identified limitations.
    - Engaging worker's performance factor based on real-time data.
    - Solving the model using GEKKO solver.
    - Building a prediction model to predict the worker's performance factor based on real-time data using machine learning prediction algorithms.
    - Re-feed the model with updated data and predicted values of the worker's performance factor. This task will be continued by identifying limitations and possible improvements.

3.1 Skills and competencies selection model for Industry 4.0

The objective in this phase is to identify the skills needed for each production job and the level of proficiency required to match the right workers to a job. After reviewing the published literature, there is one suitable model that categorizes different skills along with the proficiency level that is appropriate for industry 4.0 requirements. The study conducted by Hecklau et al. [24] and resulted in developing a model that uses radar charts to help managers to visualize the core skills, their categories and the minimum required competence level for each competence as shown in Figure 1.

3.2 Jobs assigning models based on skills matching

The goal of this phase is to develop an optimization model that assigns workers to jobs based on skills matching and control that by the production measures, production capacity, and/or production cost. As a result of the first phase, a production manager can define the skills needed for each job and the required skill level. Hence, there will be a set of jobs to be performed with specific skills required. Furthermore, there will be a set of available workers who have different levels of the required skills. Moreover, there are certain production measures to control the model as illustrated in Figure 6.

The model assumptions are as follows:

- Workforce is NOT subject to variation.
- Overtime allocation is NOT considered for this stage.
- The modeling starts for a single-product case.
- There are various workstations in the production system (one job per station).
- Capacity and cost limitation levels are subject to change in each time period.
- The minimum expected qualification required is determined by managers using the model discussed in phase one.

- There is a cost associated with using high-skilled workers. In other words, cost of performing a job is a function of workers skills.

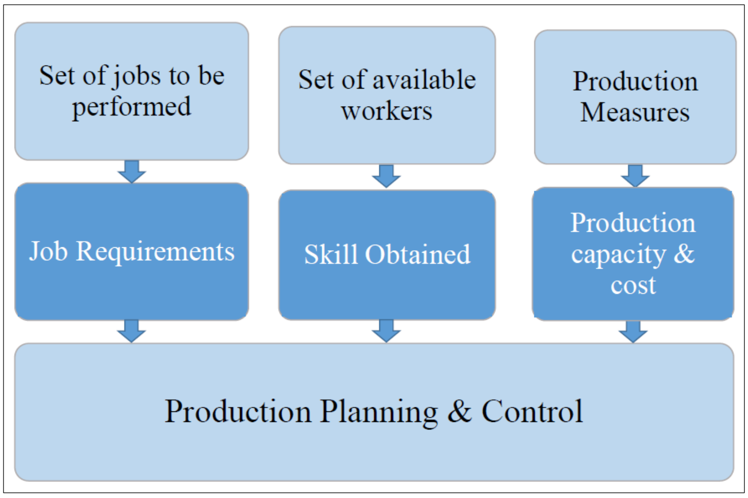

Figure 6. Model development illustration

Next, the notations and the model are presented.

- Input Parameters:

  LC_{jt}: Baseline Capacity of a work station to produce *j* units at time period *t*.

  UC_{jt}: Upper limit Capacity of a work station to produce *j* units at time period *t*.

  Vj: Unit value ($) of producing job *j* ($j = 1, \ldots, J$).

  LZ_{jt}: Baseline for the cost ($) of producing *j* units at time period *t* ($t=1, \ldots, T$).

  UZ_{jt}: Upper limit for the cost ($) of producing *j* units at time period *t* ($t=1, \ldots, T$).

  Q_{il}: Level of worker ***i*** ($i = 1, \ldots, I$) in skill *l* (l=1,…,L).

  S_{jl}: Level of skill *L* required by job j.

- Decision variables:

  X_{ijt}: this decision variable, will be 1 if worker *i* is assigned to perform job *j* in time period *t*; otherwise, will be zero.

$$\text{Max} \sum_{t=T} \sum_{i=I} \sum_{j=J} \left[x_{ijt} * Vj * LCjt * \left(\left(\sum_{l=L} Q_{iL} \ / \ S_{jL} \right) / L \right) \right] - \left[x_{ijt} * LZjt * \left(\left(\sum_{l=L} Q_{iL} \ / \ S_{jL} \right) / L \right) \right] \quad (1)$$

Subject to:

$$Xijt * \left(Q_{iL} - S_{jL} \right) \geq 0 \quad \forall \, l \in L \quad (2)$$

$$Xijt * \left[LCjt * \left(\sum_{l=L} Q_{iL} \ / \ S_{jL} \right) / L \right] \leq UCjt \quad \forall \, j \in J \text{ and } \forall \, t \in T \quad (3)$$

$$Xijt * \left[LZjt * \left(\sum_{l=L} Q_{iL} \ / \ S_{jL} \right) / L \right] \leq UZjt \quad \forall \, j \in J \text{ and } \forall \, t \in T \quad (4)$$

$$\sum_{i=I} Xij = 1 \quad \forall \, j \in J \text{ and } \forall \, t \in T \quad (5)$$

$$\sum_{j=J} Xij \leq 2 \quad \forall \, i \in I \text{ and } \forall \, t \in T \quad (6)$$

The objective function in equation (1) intends to maximize the production profit. It has two main components: production rate (representing revenue) and cost of job assignment. In the first term (i.e. production rate), there is an average ratio of worker skills level (Q_{il}) divided by skills level required by the job (S_{il}) which can be calculated for each job assigned to each worker. The model attempts to select the best matchings that satisfy the constraints and maximize the profit. The average ratio will be multiplied by LC_{jt}. The parameter can be regarded as manager input. For example, $LC_{jt} = 7$ indicates that a worker who has certain level of skills required by this job would be able to produce 7 units in unit time.

By using the average ratio to determine worker production rate based on the baseline defined by managers, it is possible to assume a linear relationship. For instance, when a manager defines the skills required for a job to be 3 skills and the level of each one is level 2, the average ratio (Q_{il} / S_{iL}) for a worker who obtained level 4 at all these three skills is two. Therefore, when a manager defines the baseline for that job to be 5 units, the worker is expected to produce 10 units. The production rate will be multiplied by that job unit value V_j to get the production value of that worker/job combination. Next, the model will select this combination by setting the decision variable X_{ijt} equal to one or discard that combination by setting X_{ijt} equal to zero.

In the assignment cost term, the average ratio of worker skills level to the skills level required by the job (Q_{il} / S_{iL}) is multiplied by the cost baseline, LZ_{jt}. The manager will define the baseline cost that is going to be paid for a worker who produced the baseline number of J units. For instance, a worker who produces $LC_{jt} = 7$ units will be paid $LZ_{jt} = 10\$$. This value will be multiplied by the same decision variable X_{ijt} which will be determined by the model.

Constraint (2) ensures that each skill level of the selected worker is equal or greater than the skill level required by that job. Constraint (3) ensures that the selected worker's production rate is not exceeding the station production (job) capacity limit at each specific time period. Constraint (4) ensures that the selected worker's assignment cost for performing a specific job will not exceed the job labor cost upper limit at that time period. Constraint (5) ensures that each job is assigned to a worker. Constraint (6) ensures that each worker is performing no more than specific number of jobs per time period (assumed to be 2 jobs).

3.3 Job assigning model based on skills matching and real time factor

The goal in this phase is to transform the assignment model for I4.0 applications and feed the model with real-time monitoring data about workers' status that indicates worker's performance. For this, a new factor is introduced which is called the worker's performance factor to represent real-time data about workers' status. This factor will be predicted by a machine learning model, and as a result, there will be a performance factor for each worker for each job. Developing a suitable machine learning model is a subtask of this phase. Moreover, considering previous model limitations (which will be discussed in Chapter 4), there is a need for relaxing the model. In this way, a variable to determine the number of working hours performed by workers is needed. This variable will relax the model by accommodating workers who can perform the job, but their production rate would result in exceeding production limits or cost limits. Thus, the new variable can determine the number of hours which does not exceed the corresponding limits and at same time accommodate that worker to perform the job. Accordingly, some input parameters will be modified as explained below.

Since the performance factor is based on real-time data and worker's status might vary at different period of time, the model should adopt this dynamic situation. Therefore, the concept of job rotation is adopted where the manager runs the assignment model at certain periods to get new worker/job combination as the workers' status changed and the performance factor will be updated accordingly. The model can be run for each shift where the number of hours per shift is pre-determined. Therefore, the updated assignment model will consider single time period instead of multiple time periods.

The model assumptions are as follows:

- Workforce is NOT subject to variation.
- Overtime allocation is NOT considered for this stage.
- The modeling starts for a single-product case.
- There are various workstations in the systems (one job per station).
- Station production capacity limitation level is subject to change in each rotation period.
- Minimum expected qualification required is determined by managers (i.e. it is a management constant).
- The model will determine number of working hours for each selected worker, so the labor cost can be calculated based on worker's hourly wage.

Next, the notations and the model are presented as follows:

- Input parameters:

  LC_j: Baseline Capacity of station to produce *j* units per unit time.

  UC_j: Upper limit Capacity of station to produce *j* units per unit time.

  Vj: Unit value ($) of producing *j*.

Zi: wage ($) of worker *i* per unit time.

UZj: Upper limit for the cost ($) of producing *j* units per unit time.

Q_{il}: Level of worker *i* in skill *l*.

S_{jl}: Level of skill *l* required by job *j*.

PFij: Predicted worker's performance/efficiency factor based on worker's real time data status.

- Decision variables:

Yij : Number of working hours performing job *j* by worker *i*.

Xij: this decision variable, will be 1 if worker *i* is assigned to perform job *j*; otherwise, will be zero.

$$\text{Max} \sum_{i=I} \sum_{j=J} [\, x_{ij} * Y_{ij} * V_j * (\text{LCj} * [(\sum_{l=L} Q_{iL} \;/\, S_{jL})/L] * PF\text{ij}\,)\,] - [\, x_{ij} * \text{Yij} * \text{Zi}\,] \qquad (7)$$

Subject to:

$$Xij * \left(Q_{iL} - S_{jL}\right) \geq 0 \quad \forall\, l \in L \qquad (8)$$

$$Xij * Yij * (LCj * [(\sum_{l=L} Q_{iL} \;/\, S_{jL})/L] * PFij) \leq UCj \quad \forall\, j \in J \qquad (9)$$

$$Xij * (Zi * Yij) \leq UZj \quad \forall\, j \in J \qquad (10)$$

$$\sum_{i=I} Xij = 1 \;\; \forall\, j \in J \qquad (11)$$

$$\sum_{j=J} Xij = 1 \;\; \forall\, i \in I \qquad (12)$$

$$\sum_{j=J} Yij \leq 4 \;\forall\, i \in I \qquad (13)$$

Since there are two variables multiplied by each other in the objective function as well in the second and third constraint, the model becomes nonlinear. The model objective function (equation 7) is about to maximize the production profit. There are two main components in this equation: production rate and assignment cost. In the production term, similar to the original model, there is the average ratio of worker skills level to the skills level required by the job (Q_{il} /S_{il}) that can be calculated for each job. Hence, the model will select the optimal matchings that satisfy the objective. The average ratio will be multiplied by the LC_{jt}, which is a managerial input. To get the number of units produced by this worker per unit time, the average ratio will be multiplied by the LC_{jt} and by the performance factor PF_{ij} of this worker for this job. The model then will determine number of hours Y_{ij} for this worker to perform this job, so that production rate will be multiplied by that job unit value V_j to get the production value of that worker/job combination.

Next, the model will select this combination by setting the decision variable X_{ijt} equal to one or discard that combination by setting X_{ijt} equal to zero. The performance factor PF_{ij} will control the number of units that worker produces based on the real time data.

In the cost assignment term, since the number of working hours for a worker to perform a job is determined by the model (Y_{ij}), that number will be multiplied by the wage of the worker per hour (Z_i).

Constraint (2) ensures that each skill level of the selected worker is equal or greater than skills level required by that job. However, in this constraint the PF_{ij} is not used because the purpose of PF_{ij} is to represent the worker performance not determining whether he is able to do the job or not. Constraint (3) ensures that the selected worker's production rate is not exceeding the station capacity limit. Constraint (4) ensures that the selected worker's assignment cost to produce that job is not exceeding the cost limit. Constraint (5) ensures that each job is assigned to a worker. Constraint (6) ensures that each worker is performing no more than one job. Constraint (7) ensures that number of working hours of each worker does not exceed the shift hours (the rotation period hours) which is defined by the production manager (it is assumed to be 6 hours for this model in the numerical example).

3.3.1 Machine Learning Model development.

The objective at this stage is to develop a machine learning (ML) model to predict worker's performance based on real time data of worker's status. Accordingly, input features to be used in the model must be defined. Other aspects of the input features such as the time period to evaluate these features should be also defined. As pointed out in the literature review, HRV is a proven indicator of worker performance which can be used to predict worker's performance. There have been several studies about the HRV correlation to employee's performance and studies that predict workers' performance using HRV [120], [123], [124], [130]. The most frequent HRV features used in these studies are SDNN, RMSSD, and pNN50 in time domain and HF, LF, and LF/HF in frequency domain at resting period. Therefore, these six features will be used as input features to the ML model. Regarding the time period to evaluate HRV features; according to a standard of measurement developed by Task Force of The European Society of Cardiology and The North American Society of Pacing and Electrophysiology, "short-term 5-min recordings and nominal 24 hours long-term recordings seem to be appropriate options" [131]. Hence, 5-min recording period will be used in this study. The HRV features evaluation will be at the rest situation (before the task) because in the assigning model, assigning workers to jobs will be based on predicting their performance using their HRV reading at the assigning time.

As per the assigning model, PF_{ij} will be predicted by a ML model, and as a result, there will be performance factor for each worker for each job. In other word, there will be a machine learning model for each worker for each job. The best way to satisfy this goal is to collect data (job performance level and HRV features) for each worker continuously and build an automated machine learning model that selects the best algorithm that fits the dataset. This can be done by using tool such as DataRobot [132]. However, the purpose here is to develop a machine learning model based on published dataset to merely test the assignment model using the predicted worker's performance as a continuous value (specific value) and also as categorical value (i.e. level).

After reviewing the available published datasets, it appears that there is no dataset that captures data for different individuals (i=1,2,3,4), each performing different jobs (j=1,2,3) multiple times (to obtain a reasonable number of instances) for each job while recording their job performance results and HRV. This type of dataset would be applicable for developing a machine learning model for each worker for each job. Therefore, the other option is to find dataset that suits developing a model for each job using data from all workers who performed that job. In December 2021 Gao et al. [133] published a dataset of an experiment capturing data on emotion, cognition, sleep, and multi-model physiological signals. An overview of the dataset is presented in Table 6.

Table 6. Dataset Overview

Participants	Location	Measures
89 healthy college students. (32 M and 57 F, age: 23.68 ± 2.12 Y)	Southeast University, Nanjing, China	1- Physiological signals: -ECG during sleeping time. -EDA and PPG recorded during emotion induction and cognitive ability assessment. 2- Cognitive ability assessment tests results.

Figure7 . Empatica E4 wristband [134]

For each participant, the photoplethysmography (PPG) signals were recorded using Empatica E4 wristband [134]. The Time between individuals heart beats in seconds (Inter-beat interval (RR Intervals)) was extracted from the PPG signals along with the time (respect to the initial time (Unix timestamp in Universal Time Coordinated - UTC)) of the detected inter-beat interval expressed in seconds as shown in Figure 8 column C and B. Cognitive tests were performed using Cambridge neuropsychological test automatic battery (CANTAB). An overview of the CANTAB tests used in the experiment is presented in Table 7. For each participant, the starting time and duration for all 4 tests were reported in addition to each test outcome.

Table 7. Overview of The CANTAB Tests

Test	Description	Duration	Outcome Measures
MOT: Motor Screening	An assessment of sensorimotor deficits or lack of comprehension	2 minutes	Mean Latency (ML) Mean Error (ME)
RVP: Rapid Visual information Processing	A measure of sustained attention	7 minutes	Probability of Hit (PH) Mean Latency (ML)
SST: Stop Signal Task	Stop signal response inhibition test, which uses staircase functions to generate an estimate of stop signal reaction time.	Up to 14 minutes	SSRT (stop signal reaction time) Mean correct Reaction Time on go trials (MRT) Proportion of successful stops (PSS)
SWM: Spatial Working Memory	A test of the subject's ability to retain spatial information and to manipulate remembered items in working memory. It has notable executive function demands and provides a measure of strategy as well as working memory errors	4 minutes	Mean Error (ME) Strategy (ST)

	A	B	C	D	E
1	unix timestamp in UTC (initial time)	Detection time of inter-beat interval expressed in seconds respect to the initial time	Inter-Beat Interval (IBI) in second (the distance in seconds from the previous beat)	Detection time in HH:MM:SS	IBI (millisecond)
2	1529492453	102.426563	0.703157	19:02:35	703.157
3		103.301603	0.87504	19:02:36	875.04
4		114.520867	0.921917	19:02:48	921.917
5		115.442784	0.921917	19:02:48	921.917
6		116.44283	1.000046	19:02:49	1000.046
7		117.270993	0.828163	19:02:50	828.163
8		118.161659	0.890666	19:02:51	890.666
9		119.114827	0.953169	19:02:52	953.169
10		119.94299	0.828163	19:02:53	828.163

Figure 8. Participant's Inter-beat Interval

By transferring the Unix timestamp to normal time, the time of the detected inter-beat interval are expressed in HH:MM:SS format as shown in column D in Figure 8. As mentioned earlier the HRV features should be calculated at resting time based on 5 minutes periods. Unfortunately, all the 4 tests were performed in a continuous way and there was no resting period between each test. Therefore, the 5 minutes period prior to the first test was selected to calculate the HRV features. Out of 89 participants only 67 participants completed the CANTAB with E4 recording and their data was provided by the author. Of those 67 participants, only 39 participants have a clean recording of 5 minutes period before the CANTAB test starts (No recording was provided by E4 wristband before the test for the other 27 participants).

Using HRV-analysis library in Python developed by Aura Healthcare, the six most frequently HRV were calculated. For each participant, these six HRV features represent the independent variable (input features). Regarding the dependent variable (target), there are 9 outcome measures as shown in Table 7 (each outcome measures represent a job for the assigning model), so there will be 9 prediction models. Figure 9 shows the independent and dependent variables for the first 23 participant out of 38 participants. One participant was removed due to having outlier values in his input features. Interquartile Range (IQR) method was applied for outliers' detection [135].

#	ID	X						Y								
								MOT		RVP		SST			SWM	
		SDNN	PNN50	RMSSD	LF	HF	LF-HF	Mean latency (ms)	Mean error	Probability of hit	Mean latency (ms)	SSRT (stop signal reaction time) ms	Mean correct RT on GO trials	Proportion of successful stops	Mean Error	Strategy
1	5	0.683	0.703	0.587	0.884	0.591	0.473	648.3	11.37	0.63	429.47	206.5	576.33	0.58	1	32
2	7	0.139	0.130	0.047	0.218	0.044	1.000	603.3	8.08	0.7	400	259.02	511.73	0.4	25	29
3	17	0.265	0.402	0.180	0.032	0.190	0.097	528.4	11.37	0.91	365.17	274.48	664.68	0.65	41	36
4	21	0.713	0.886	0.822	0.721	1.000	0.199	504.3	11.85	0.78	523.86	502.78	1336.46	0.92	75	42
5	24	0.329	0.370	0.293	0.313	0.181	0.539	610.6	9.01	1	416.19	257.92	602.03	0.4	1	25
6	28	1.000	0.483	0.343	0.320	0.353	0.285	717.3	16.28	0.89	482.33	291.85	924.31	0.62	27	35
7	31	0.175	0.171	0.170	0.107	0.078	0.447	806	8.56	0.81	552.23	579.35	400.97	0.5	31	39
8	33	0.606	0.536	0.601	0.734	0.963	0.214	723.5	10.1	0.81	440.68	221.58	757.45	0.7	11	33
9	34	0.937	0.753	0.672	1.000	0.515	0.623	885.8	8.43	0.7	440.37	216.48	479.06	0.48	4	32
10	41	0.594	0.869	0.732	0.501	0.564	0.268	650.3	10.54	0.81	370.77	223.8	509.34	0.58	9	33
11	45	0.544	0.280	0.150	0.532	0.246	0.672	826.5	6.06	0.89	358.46	261.45	614.16	0.42	32	32
12	59	0.079	0.043	0.114	0.000	0.012	0.379	425.1	14.71	0.67	421.33	228.6	506.88	0.45	1	31
13	62	0.690	0.606	0.491	0.544	0.480	0.353	574.5	12.45	0.67	394.33	269.08	571.44	0.45	10	34
14	73	0.000	0.000	0.010	0.044	0.025	0.496	722	13.97	0.56	344.87	266.7	671.28	0.6	13	28
15	77	0.443	0.433	0.419	0.516	0.219	0.724	1009.7	4.73	0.67	447.78	281.82	613.18	0.45	25	30
16	78	0.079	0.043	0.114	0.000	0.012	0.379	664	8.52	0.63	509.88	459.6	1000.46	0.4	1	30
17	8	0.360	0.130	0.387	0.038	0.000	0.693	1145.6	6	0.93	354.52	277.95	1116.24	0.88	36	41
18	10	0.272	0.299	0.216	0.060	0.152	0.180	540	12.39	0.67	429.83	247.32	601.78	0.45	13	29
19	12	0.642	0.227	0.408	0.260	0.703	0.085	1091.9	4.93	0.67	361.33	330	623.77	0.6	1	27
20	15	0.089	0.209	0.168	0.096	0.050	0.538	742.9	7.39	0.81	373.09	255.42	932.78	0.75	11	33
21	19	0.618	0.719	0.735	0.025	0.435	0.000	677.3	8.63	0.81	268.5	174.23	605.64	0.55	14	28
22	27	0.664	0.811	0.619	0.184	0.916	0.017	710.1	5.26	0.74	411.3	381.88	1123.22	0.65	1	31
23	30	0.312	0.375	0.288	0.019	0.241	0.050	733.8	10.2	0.93	361.68	252.73	847.7	0.6	17	35

Figure 9. independent and dependent variables for the first 23 participant

Since the output (dependent variables) are known and available, the model is supervised machine learning. There are two main groups of supervised machine learning: regression (predicting continuous value) and classification (predicting categorical value). Correlation matrix was developed to assess the correlation between the target variable (each outcome measure of the CANTAB tests) and the input variables and the correlation between the input variables. As shown in Figure 10 , the correlation between the target variables and input features is considered weak, with the highest value is 0.31. Moreover, there were high correlations (reached 0.92) between some of the input features which lead to multicollinearity issue (it undermines the statistical significance of an independent variable for the modeling). Based on the correlation matrix, the possible ML algorithms that can suit the dataset were selected. Table 8 presents an overview of some ML algorithms and their suitability to the applied dataset. Input features are normalized using Min-Max method, so each feature has a range between 0 to 1 [136]. As some ML algorithms are sensitive to input features value range, specifically the instance-based algorithms.

	SDNN	PNN50	RMSSD	LF	HF	LF-HF
SDNN	1					
PNN50	0.732	1				
RMSSD	0.726	0.922	1			
LF	0.630	0.511	0.502	1		
HF	0.721	0.775	0.801	0.516	1	
LF-HF	-0.211	-0.414	-0.408	0.314	-0.520	1
Mean latency (ms)	0.107	-0.107	-0.056	0.099	-0.121	0.320
Mean error	-0.017	-0.097	-0.170	-0.087	-0.127	-0.145
Probability of hit	0.164	0.202	0.196	-0.078	0.097	-0.165
Mean latency (ms)	0.155	0.113	0.131	0.309	0.063	0.102
SSRT(stop signal reaction time) ms	-0.056	-0.197	-0.148	-0.031	-0.013	0.124
Mean correct RT on GO trials	0.104	-0.085	-0.071	-0.003	0.083	0.108
Proportion of successful stops	0.169	0.035	0.065	0.122	0.179	0.094
Mean Error	-0.081	-0.017	-0.051	0.032	-0.009	0.091
Strategy	-0.109	0.080	0.117	-0.015	0.053	-0.156

Figure 10. Correlation Matrix

Table 8. Overview of Some Machine Learning Algorithms

Algorithm	Description	Assumptions	Suitability to The Dataset
Linear Regression Regression	A relationship is established between independent and dependent variables by fitting them to a line	- Linear relation between the input features and the target variable[137] - Sensitive to correlated features, Ridge regression can be used to handle multicollinearity [138]	Not Suitable: No relation between the input features and the target variable
Decision Trees (DT) Regression & Classification	A flow-chart-like tree structure that uses a branching method to illustrate every possible outcome of a decision. Each node within the tree represents a test on a specific variable (Feature) and each branch is the outcome of that test.	Does not make assumptions about the functional form of the relationship between the input features and the target variable and the distribution (non-parametric Algorithm) [139], [140].	Suitable

	To decide which feature to select at which value as root node and then other internal nodes, (Gini Impurity, Entropy, or Information Gain).		
Random Forest (RF) tree-based Regression & Classification	A collective of decision trees is called a Random Forest. To classify a new object based on its features, each tree classify, and the tree “votes” for that class. The forest chooses the classification having the most votes (over all the trees in the forest- equal vote). Budling a tree by randomly selecting subset of the dataset and part of the features (not all as DT) to decide the root node to separate the sample and same apply to decide the next internal node.		Suitable
k-nearest neighbors’ algorithm (KNN) Instance-based Regression & Classification	Stores all available cases and classifies any new cases by taking a majority vote of its k neighbors. The case is then assigned to the class with which it has the most in common. A distance function performs this measurement (Euclidean Distance).		Suitable
Gradient Boosting (GBoost) tree-based Regression & Classification	Gradient boosting is a type of machine learning boosting. It relies on the intuition that the best possible next model, when combined with previous models, minimizes the overall prediction error. The key idea is to set the target outcomes for this next model in order to minimize the error.	Does not make assumptions about the functional form of the relationship between the input features and the target variable and the distribution (non-parametric Algorithm) [139], [140].	Suitable

Adaptive Boosting (AdaBoost) tree-based Regression & Classification	Forest of stump (a tree with one node and two leaves only). In random forest each tree has equal vote but in AdaBoost some stomps have more weight in voting than others. In random forest building a tree is built independently from other trees (order is not important-which tree is first), in AdaBoost order of building stumps is important where the next stump will consider the error of the first stum as GBoost.		Suitable
Support vector machine (SVM) Instance-based Regression & Classification	Plot raw data as points in an n-dimensional space (where n is the number of input features). The value of each feature is then tied to a particular coordinate, making it easy to classify the data. Lines called classifiers can be used to split the data and plot them on a graph.		Suitable
Naive Bayes Classification	Applying the Bayes' theorem on data with a naive assumption of conditional independence between every pair of features, given the value of the class variable	Naive Bayes classifiers assume that the value of a particular feature is independent of the value of any other feature, given the class variable.	Not Suitable: high correlation between input features

For each target variable, a regression model was developed (using Sklearn library in Python) by applying each one of the suitable algorithms using K-fold cross validation data split technique which is more reliable to the small dataset (38 rows) compared to the traditional data spilt technique as 30% for test and 70% for train for example. Evaluation metrics used for the regression to compare the result of these different algorithms was Mean Absolute Percentage Error (MAPE) which applied to the test sets. Table 9 presents an overview of some metrics used in regression model. The highest performance algorithm for each target variable (9 models in total) was selected to apply tuning in order to improve its performance. Tuning parameter values for each algorithm are presented in Table 10. Grid-Search by Sklearn was used as a tuning technique that computes the optimum values of hyperparameters of an algorithm. Therefore, there will be a regression model for each target variable build based on the highest performance algorithm (e.g., the model of target variable A was built based on KNN while the model of target variable B was built based on Decision Tree). Figure13 represents steps of applying the machine learning modeling.

Table 9. Overview of Some Evaluation Metrics for Regression Model

Metric	Definition	Formula
R Squared	The coefficient of determination or R-squared represents the proportion of variation of data points explained by the model.	$R^2 = 1 - \frac{\sum_{i=1}^{n}(Yi - \widehat{Yi})^2}{\sum_{i=1}^{n}(Yi - \bar{Y})^2}$
MAE	The Mean Absolute Error represents the average of the absolute difference between the actual and predicted values in the dataset. It measures the average of the residuals in the dataset.	$MAE = \frac{1}{n}\sum_{i=1}^{n}\lvert Yi - \widehat{Yi}\rvert$
MAPE	Mean Absolute Percentage Error is the mean of all absolute percentage errors between the predicted and actual values.	MAPE = $\frac{1}{n}\sum_{i=1}^{n}\left\lvert\frac{(Yi - \widehat{Yi})}{Yi}\right\rvert 100 \;\; , yi \neq 0$
MSE	Mean Squared Error represents the average of the squared difference between the original and predicted values in the data set. It measures the variance of the residuals.	$MSE = \frac{1}{n}\sum_{i=1}^{n}(Yi - \widehat{Yi})^2$
RMSE	Root Mean Squared Error is the square root of Mean Squared error. It measures the standard deviation of residuals.	$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Yi - \widehat{Yi})^2}$
Normalized RMSE	A normalization of RMSE by dividing it to the range of the actual values [141].	$NRMSE = \frac{RMSE}{Range}$

Table 10. Tuning parameter values for each algorithm

Algorithm	Parameter Values
DT	Max Depth = [5,8,10,13,15,18,20,25,27,32,35,37,40,44,46,50]
RF	Number of Estimators (tree) = [80,90,100,110,115,120,123,125,130,135,138,140] Max Depth = [5,8,10,13,15,18,20,25,27,32,35,37,40,44,46,50]
KNN	K range = range (1, 12) Weight Options = ['uniform', 'distance'] Uniform: All points in each neighborhood are weighted equally. Distance: weight points by the inverse of their distance (closer neighbors of a query point will have a greater influence than neighbors which are further away).
GBoost	Number of Estimators (tree) = [80,90,100,110,120,130,140,150,160,170,180,200] Max Depth = [2,3,5,8,10,13,15,18,20,25,27,32,35,37,40,44,46,50] Learning Rate = [0.01,0.1,1]
AdaBoost	Number of Estimators (tree) = [20,30,50,80,90,100,120,150,160,170,180,200] Learning Rate = [0.01,0.1,1,1.1]
SVM	C = [0.25,0.5,1, 10, 100, 1000] kernel = ['linear', 'rbf'*] * Radial basis function

For classification models, each target variable is discretized into categories to transform the problem from regression to classification [142]. Equal Width Binning (EWB) method was used to discretize target value [143] to three and two performance levels; {High= 100% to 66.67, Medium = 65% to 33.34%, Low = 32% to 0} and {High= 100% to 50%, and Low = 49% to 0}. EWB has resulted in imbalanced classes, i.e. number of instances of high level does not equal the instances of low level. For each target variable, a classification model was developed (using Sklearn library in Python) applying each one of the suitable algorithms using stratified K-fold cross validation data split technique instead of the regular K-fold to ensure proportion of each class in the original dataset maintained in taring and test sets as possible. Table 11 presents an overview of the metrics that can be used with classification models. Balanced accuracy was chosen as the evaluation metrics to compare the result of these different algorithms. Balanced accuracy is more suitable to imbalanced classes compare to accuracy (Figure 11 and Figure 12 illustrate this comparison in the case of imbalanced classes). The highest performance algorithm for each target variable (total 9 classification models) was selected to apply tuning in order to improve its performance. Tuning parameter values for each algorithm are presented in Table 10. Grid-Search by Sklearn was used as a tuning technique that computes the optimum values of hyperparameters of an algorithm. Therefore, there will be a classification model for each target variable developed based on the highest performance algorithm (e.g., the model of target variable A was built based on KNN while the model of target variable B was built based on Decision Tree). Figure13 represents steps of applying the machine learning modeling. In the case of use classification to predict worker's performance, there will be fixed values for PFij in the assigning model (1.0, 0.667, and 0.334 in case of three class and 1.0, and 0.5 in case of two classes).

Confusion Matrix (2 classes)	Predicted (Positive)	Predicted (Negative)
Actual (Positive)	TP	FN
Actual (Negative)	FP	TN

TP=True Positive, FP=False Positive, TN=True Negative, Fales Negative

Figure 11. Confusion Matrix

Table 11. Overview of Some Evaluation Metrics for Classification Model

Metric	**Definition**	**Formula**
Accuracy	The percentage of correct predictions made by a model.	$\frac{(TP+TN)}{(TP+FN+FP+TN)}$
Recall	Percentage of correct predictions of a class among all that class samples.	$\frac{(TP)}{(TP+FN)}$
Precision	Percentage of correct predictions of a class among all predictions for that class.	$\frac{(TP)}{(TP+FP)}$
Balanced Accuracy	The average of correct predictions for each class was made by a model. It can be a better judge of performance in the imbalanced classes. "The average of recall obtained on each class".	$\frac{1}{2}\left(\frac{TP}{(TP+FN)}+\frac{TN}{(TN+FP)}\right)$

Confusion Matrix	Predicted (High)	Predicted (Low)	
Actual (High)	5	15	20
Actual (Low)	15	65	80

$$\text{Accuracy} = \frac{70}{100} = 0.7$$

$$\text{Balanced Accuracy} = \frac{1}{2}\left(\frac{5}{20}+\frac{65}{80}\right) = 0.53$$

Figure 12. Confusion Matrix Example

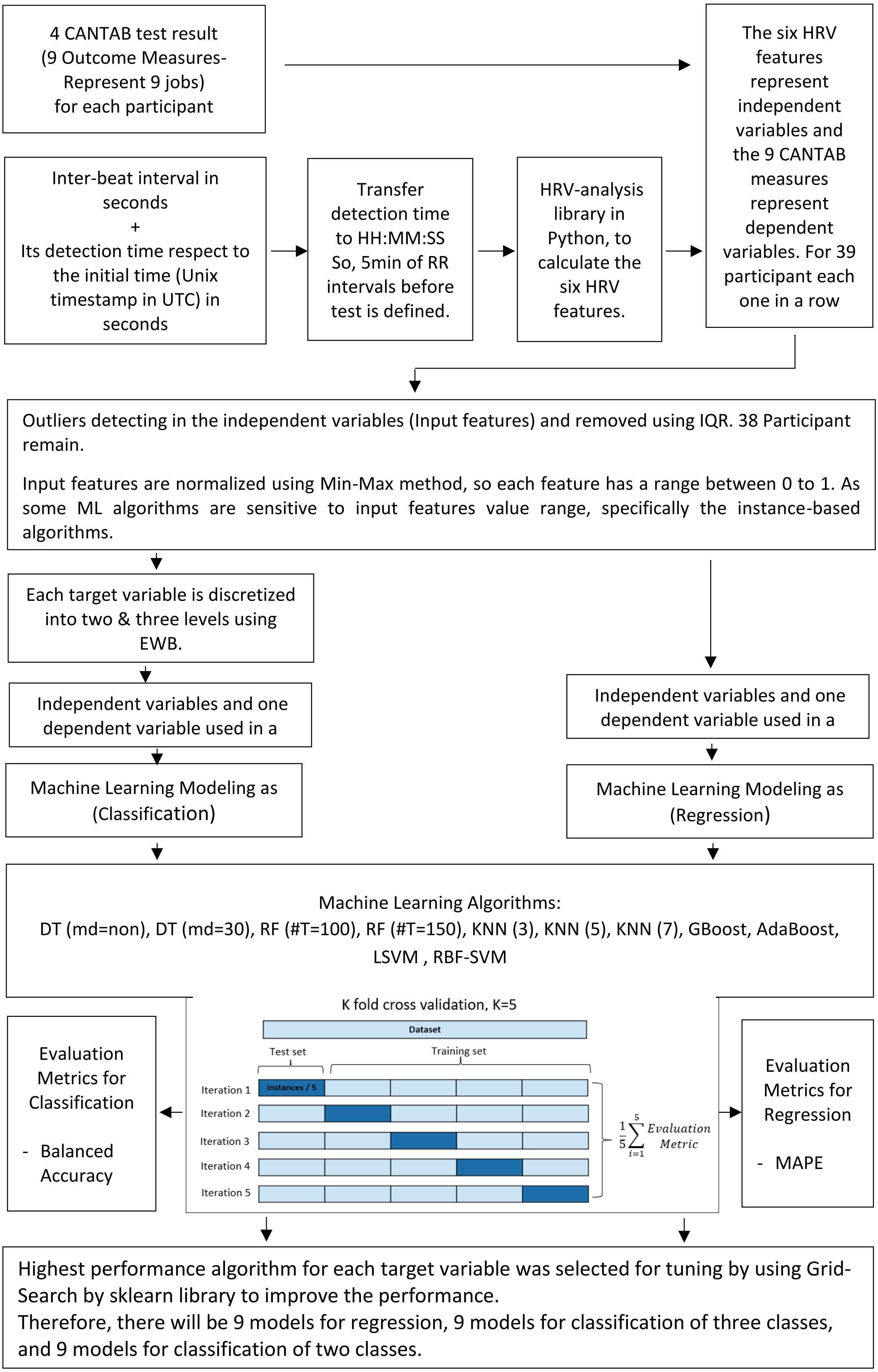

Figure13 . Machine learning modeling steps.

4. Result and Discussion

This section discusses the results of the numerical example using the two models mentioned in phase two and three. The models were tested for different scenarios.

4.1 Jobs assigning models based on skills matching (phase two)

The model was written in Python language and was solved by using Gurobi solver. Then it was tested using the following input data;

- number of workers = 5
- number of jobs 7
- number of skills = 4
- time periods = 2

Each Job requires a certain level of skill (1 to 5). Each worker has a certain level of each skill (1 to 5). For each time period, there was a specific production capacity and a worker assignment cost limit for each station (Job). To see how the model responses, two scenarios were considered.

The first scenario involves changing the jobs required skills level which resulted in the outputs presented in Figure *14* while the workers skills level and other parameters remain constant. The production capacity per station *UCjt* was set to be between 40 and 90. The production value associated with the average Q/S ratio of the selected workers was recorded. In this respect, Q/S ratio is the average of the individual ratios of a worker skill level to the job required level of that skill. In Figure *14*, the horizontal axis represents the average Q/S for the selected workers to perform 5 jobs. The Y axis represents the production revenue. For example, the second data point where Q/S equals 1.650 and production revenue is $1,674. The selected workers are shown in Table 12 where worker 1 will do job 1 in the first time period and worker

2 will do this job in the second time period. This is due to the different production capacity limit per station in each time period. In the third data point, where the required skills levels for certain jobs have changed, a new worker/job combination with different average Q/S (like 1.758) and its associated production revenue ($1,777) have generated.

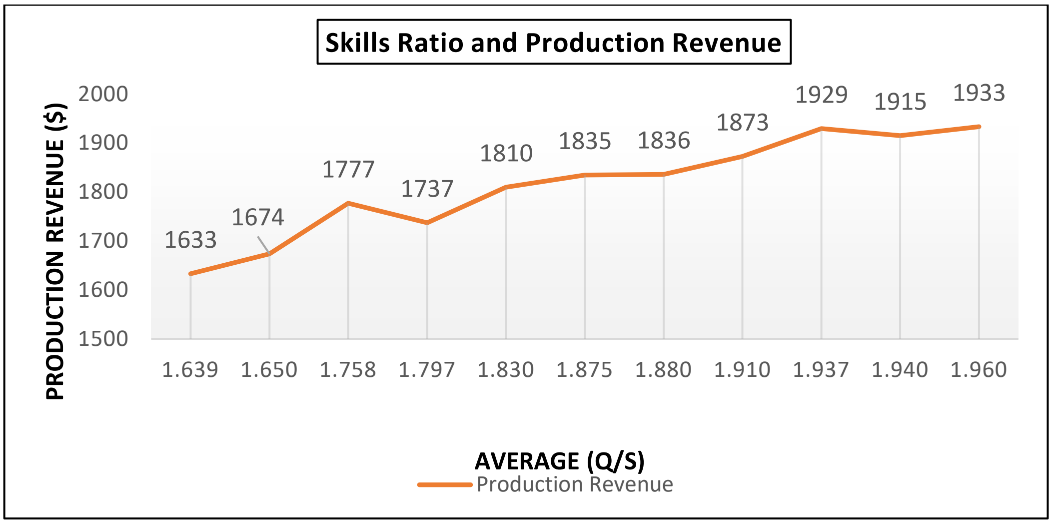

Figure 14. Production revenue associated with average Q/S

Table 12. Workers/Jobs combination

Xijt	Average ∑Qil/Sjl	Production Revenue $
111	1.65	1,674
171		
172		
212		
251		
252		
261		
321		
322		
441		
442		
462		
531		
532		

In general, production revenue has an increasing rate. In certain cases, the production revenue would decrease depending on the combination of available workers. Thus, the model must satisfy all constraints and maximize the production revenue with the available workers.

The second scenario comprises using the same data in Scenario 1, except the production capacity per station (UC_{jt}). In this scenario, the production capacity limit was decreased to be between 23 to 38; and then the production revenue associated with the average Q/S ratio for the selected workers was recorded. Table 13 and Table 14 are the results for the same job and same available workers but for different production capacity limits. Accordingly, different workers to job combination as well as different Q/S and production revenue were obtained.

In general, in this scenario, production revenue would decrease, and workers to job combination changed for the same required skills and available skills. This is due to differences in production capacity as compared in Figure *15*

Table 13. workers/ jobs combinations with high capacity

Xijt	Average ∑Qil/Sjl	Production Revenue 'High C'
131	1.93	1,929
132		
171		
172		
211		
212		
251		
252		
321		
322		
342		
441		
461		
462		

Table 14. workers/ jobs combinations with low capacity

Xijt	Average ∑Qil/Sjl	Production Revenue 'Low C'
131	1.90	1,857
152		
171		
172		
211		
212		
232		
241		
321		
322		
351		
442		
461		
462		

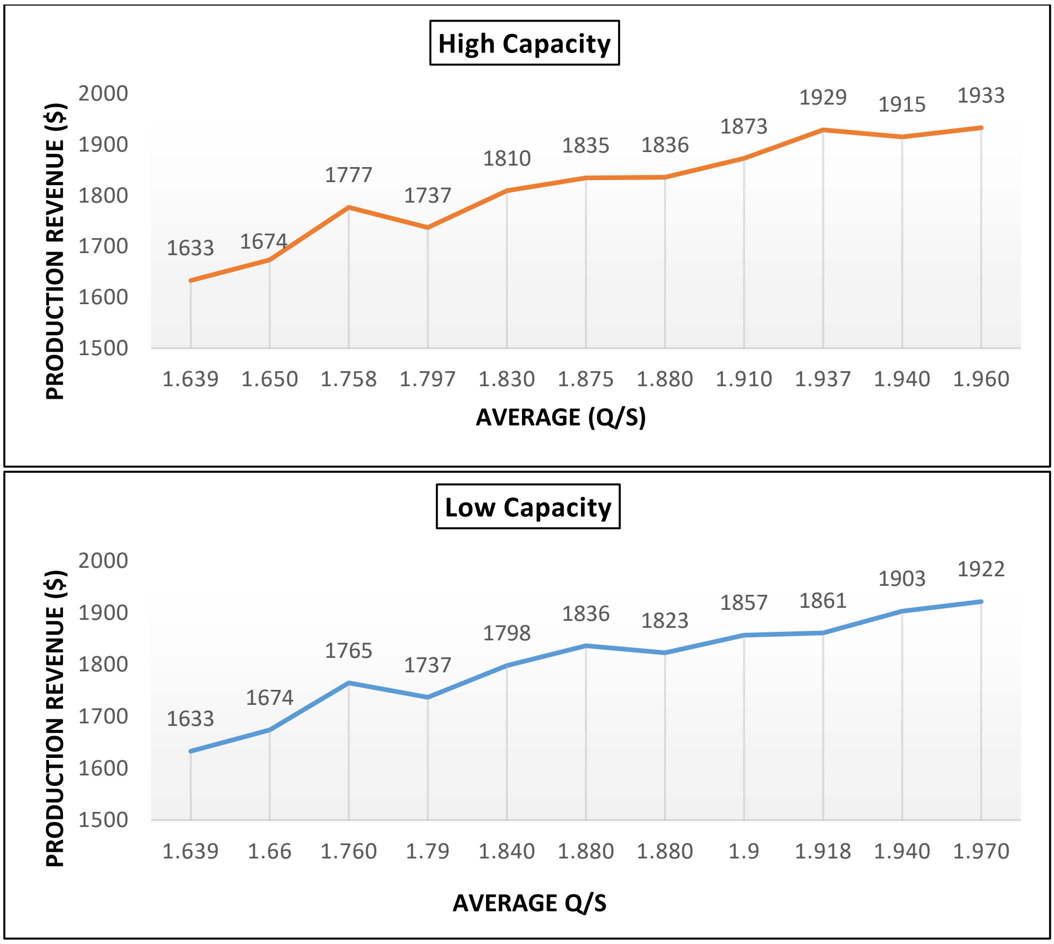

Figure 15. Production revenue associated with average Q/S for low and high production capacity.

4.1.1 Enhancing the Model

By changing the level of jobs' required skills or the production limits, the model sometimes does not find a solution due to some of the existing workers' production rate exceeding the capacity limits. To relax the model, a variable to determine the number of hours performed by the worker is needed. This variable (Yij : Number of working hours performing job j by worker i) will relax the model by accommodating workers who can perform the job, but their production rates would result in exceeding production limits or cost limits. Thus, the new variable can determine the number of hours not to exceed the limits and at the same time accommodate that worker to perform the job.

4.2 Job assigning model based on skills matching and real time Factor

The model was solved by using Gekko solver in Python (and it is nonlinear model). The number of working hours Yij introduced as another decision variable. In addition, predicted performance factor PFij value for each worker in each job is obtained from the developed machine learning model that predicts worker's performance based on worker's real time data status (HRV features as input data for the prediction models). The next section provides the relevant results followed by another section that shows the response of the assignment model in four scenarios.

4.2.1 Machine Learning Model Results

Nine prediction models were developed to predict workers performance in 9 jobs to be used in the assignment model. Two types of prediction models were developed. The first type includes regression models to predict a continuous value of the cognitive measure where MAPE was used as evaluation metrics because of easy interpretation. The second group consists of classification models to predict the level of the cognitive measure with two classes (High =1 and Low = 0.5) or three classes (High =1, Medium= 0.67, and Low = 0.34). In this way, 9 classification models were developed to predict two classes, and 9 models to predict three classes for nine jobs. Six different algorithms (DT, RF, KNN, GBoost, AdaBoost, and SVM) were used to train the models for both regression and classification. Table 15 presents the value of MAPE of each algorithm for each performance measure. Next, the algorithm with highest MAPE was tuned. Table 18 presents the algorithm with lowest MAPE after tuning for each performance measure. For two-classes classification, Table 16 presents the value of Balanced accuracy of each algorithm for each performance measure. Then, the algorithm with highest balanced accuracy was tuned. Table 18 present the algorithm with highest balanced accuracy after tuning for each performance measure. For three classes classification, Table 17 presents the value of balanced accuracy of each algorithm for each performance measure. No tuning was done for the three-class classification model due to very low balance accuracy.

Table 15.Value of MAPE of each algorithm for each performance measure

Algorithm	MAPE %								
	MOT		RVP		SST			SWM	
	ML	ME	PH	ML	SSRT	MRT	PSS	ME	ST
DT (md=non)	30.04	31.37	23.12	19.20	25.69	36.71	35.71	484.80	19.19
DT (md =30)	30.68	31.93	23.34	19.73	26.71	38.52	38.05	544.31	16.45
RF (# tree =100)	22.09	30.24	18.54	**14.02**	24.74	35.19	28.63	516.37	16.38
RF (# tree =150)	21.99	30.56	18.27	14.54	24.41	34.77	28.01	486.36	15.98
KNN (k=5)	19.41	31.22	**17.41**	15.19	24.97	31.94	26.19	408.94	16.92
KNN (k=3)	20.23	31.64	18.74	15.73	25.57	31.58	27.70	380.33	16.23
KNN (k=7)	19.23	31.49	19.00	15.73	22.84	32.33	**23.61**	436.07	16.16
GBoost	22.54	**28.72**	20.67	15.96	21.21	37.52	34.46	509.60	17.86
AdaBoost	**18.22**	31.48	18.58	16.55	22.10	**29.81**	31.71	451.66	**15.17**
SVM_L	19.64	29.39	19.25	16.54	**18.87**	31.57	29.27	**258.08**	16.08
SVM (RBF)	19.48	29.33	19.69	16.56	19.00	31.55	28.61	261.37	16.47

Table 16.Value of Balanced Accuracy of each algorithm for each performance measure_ 2 classes

Algorithm	Balanced Accuracy % for 2 classes								
	MOT		RVP		SST			SWM	
	ML	ME	PH	ML	SSRT	MRT	PSS	ME	ST
DT (md=non)	51.67	44.33	**62.33**	51.67	50.95	59.33	39.67	54.05	47.00
DT (md =30)	50.00	44.67	53.17	49.67	**54.29**	**64.67**	43.33	54.05	47.00
RF (# tree =100)	48.33	52.00	54.83	59.67	45.48	45.00	46.67	**70.00**	54.83
RF (# tree =150)	48.33	45.33	47.83	54.67	45.48	46.33	48.33	68.57	54.83
KNN (k=5)	50.00	44.67	44.83	46.67	50.00	51.67	45.00	70.00	60.67
KNN (k=3)	45.00	47.33	45.83	44.67	48.33	46.00	45.00	70.00	62.00
KNN (k=7)	50.00	42.67	43.00	53.33	50.00	48.33	48.33	70.00	54.00
GBoost	**70.00**	47.00	46.67	**62.33**	49.29	54.33	43.33	56.90	51.33
AdaBoost	61.67	**55.33**	46.33	57.33	49.29	51.33	39.67	61.19	**62.00**
SVM_L	50.00	50.00	50.00	50.00	50.00	50.00	**50.00**	70.00	50.00
SVM (RBF)	50.00	48.33	47.50	48.33	50.00	48.33	50.00	70.00	50.33

Table 17.Value of Balanced Accuracy of each algorithm for each performance measure_ 3 classes

Algorithm	Balanced Accuracy % for 3 classes								
	MOT		RVP		SST			SWM	
	ML	ME	PH	ML	SSRT	MRT	PSS	ME	ST
DT (md=non)	27.67	43.89	37.22	32.50	32.56	41.00	28.33	39.11	30.67
DT (md =30)	29.33	47.78	47.50	35.00	29.22	43.00	28.33	38.22	30.67
RF (# tree =100)	43.67	36.67	35.56	44.17	36.44	31.00	31.67	38.22	31.33
RF (# tree =150)	40.33	37.22	35.28	40.56	36.44	32.67	27.78	38.22	32.67
KNN (k=5)	37.33	32.22	28.06	43.61	40.00	35.33	35.56	40.22	36.67
KNN (k=3)	31.33	44.44	29.44	40.00	38.33	26.67	37.22	41.11	32.00
KNN (k=7)	38.00	39.44	31.94	43.06	40.00	35.33	45.56	42.22	31.33
GBoost	39.67	45.00	37.50	50.28	30.89	40.00	21.67	44.22	31.33
AdaBoost	42.67	29.44	32.78	41.11	35.33	38.67	25.56	39.56	28.67
SVM_L	40.00	31.67	29.72	40.56	40.00	36.67	35.00	43.33	36.67
SVM (RBF)	40.00	42.78	30.28	49.44	40.00	36.67	35.56	43.33	36.67

Table 18. Highest performance algorithms for regression and classification

<table>
<tr><th colspan="2" rowspan="2">Cognitive Measures</th><th colspan="2">Prediction Models</th></tr>
<tr><th>Regression</th><th>Classification_2 Classes</th></tr>
<tr><td rowspan="2">MOT</td><td>ML</td><td>AdaBoost (learning rate = 0.1, # of estimators = 90)
MAPE = 17.5%</td><td>- GBoost (learning rate = 0. 1, # of estimators = 80, max depth = 3)
- Balanced Accuracy = 70%</td></tr>
<tr><td>ME</td><td>- GBoost (learning rate = 0.01, # of estimators = 120, max depth = 5)
- MAPE = 26.6%</td><td>- AdaBoost (learning rate = 1.1, # of estimators = 120)
- Balanced Accuracy = 64%</td></tr>
<tr><td rowspan="2">RVP</td><td>PH</td><td>- KNN (k =5, weights = distance)
- MAPE = 17.7%</td><td>DT (max depth =32)
Balanced Accuracy = 64.3%</td></tr>
<tr><td>ML</td><td>- RF (# of estimators = 90, max depth = 10)
- MAPE = 13.4%</td><td>- GBoost (learning rate = 1, # of estimators = 80, max depth = 5)
- Balanced Accuracy = 64.3%</td></tr>
<tr><td rowspan="3">SST</td><td>SSRT</td><td>- SVR (C=100, kernel=RBF)
- MAPE = 17.4%</td><td>DT (max depth =50)
Balanced Accuracy = 55%</td></tr>
<tr><td>MRT</td><td>- AdaBoost (learning rate = 0.1, # of estimators = 90)
- MAPE = 28%</td><td>- DT (max depth =3)
- Balanced Accuracy = 66.6%</td></tr>
<tr><td>PSS</td><td>- KNN (k =11, weights = uniform)
- MAPE = 23.1%</td><td>- SVC (C=1, kernel='linear')
- Balanced Accuracy = 50%</td></tr>
<tr><td rowspan="2">SWM</td><td>ME</td><td>- SVR (C=0.5, kernel='linear')
- MAPE = 255.6%</td><td>- RF (# of estimators = 80, max depth = 2)
- MAPE = 70%</td></tr>
<tr><td>ST</td><td>- AdaBoost (learning rate = 1, # of estimators = 20)
- MAPE = 14.4%</td><td>- AdaBoost (learning rate = 0.01, # of estimators = 100)
- MAPE = 64.6%</td></tr>
</table>

Accordingly, SWM (Mean Error) model was excluded due to its very high MAPE value. For two-classes classification models, SST and Proportion of successful stops models were excluded due to their low balance accuracy values. All three-classes models were disregarded due to their low prediction accuracies. Thus, six regression models and six two-classes classification models were used for predicting the performance factor in the assignment model. These models were exported using pickle package (file_name.pkl) in Python to be ready for recall by Python and performing the prediction when it is needed.

4.2.2 Assigning model Response.

The assignment model was tested using the following inputs:

- number of workers = 6
- number of jobs = 6,
- number of skills = 4, with specific rotation periods

Each Job requires a certain level of skill (1 to 5). Each worker has a certain level of each skill (1 to 5). There was a specific production capacity and worker assignment cost limits for each station (Job). Predicted performance factor $PFij$ for each worker in each job (6 jobs) obtained from the machine learning models. Regression models predict the original target value of each measure. This value normalized to be between 0 and 1 using Max-Min method based on the original value. The classification model predicts the level as high or low where high set to be equal 1 and low equal 0.5. Table *19* presents an example of input data for the prediction models and the predicted value (regression and classification) for 6 workers in the first 3 jobs. Hence, each job has its own prediction model as mentioned in Table 18.

Table 19 :An Example of input data for the prediction models and the predicted value

#	Input Features						MOT								RVP			
							425.1	ML		1313	4.73	ME		16.28	0.41	PH		1
	SDNN	PNN50	RMSSD	LF	HF	LF-HF	**Predicted Value**	Normalized	**Predicted Class**	level	Predicted Value	Normalized	Predicted Class	level	Predicted Value	Normalized	Predicted Class	level
1	0.805	0.827	0.131	0.201	0.316	0.478	683.2	0.709	H	1.0	9.177	0.615	H	1.0	0.730	0.543	H	1.0
2	0.484	0.481	0.406	0.290	0.420	0.210	865.4	0.504	L	0.5	7.763	0.737	H	1.0	0.510	0.169	L	0.5
3	0.077	0.474	0.197	0.209	1.000	0.329	698.1	0.693	H	1.0	7.925	0.723	H	1.0	0.679	0.456	H	1.0
4	0.383	0.659	0.425	0.084	0.224	0.143	985.5	0.369	L	0.5	9.524	0.585	H	1.0	0.810	0.678	H	1.0
5	0.837	1.000	0.476	0.236	0.000	0.071	885.8	0.481	H	1.0	7.465	0.763	L	0.5	0.725	0.535	H	1.0
6	0.315	0.773	0.022	0.649	0.041	0.489	693.12	0.698	H	1.0	10.368	0.512	L	0.5	0.791	0.646	H	1.0

To better evaluate the model performance, four scenarios were considered. The first scenario involves changing the required skills level which creates the results in Figure 16. Again, we keep the workers' skills level and other parameters remain as they are where the rotation period below or equal 8 hours. Then, we record the production revenue associated with the average Q/S ratio of the selected workers. The last data point where Q/S equals 2.42 and production revenue $5,380, the workers/jobs combination and their optimal working hours are presented in Table 20.

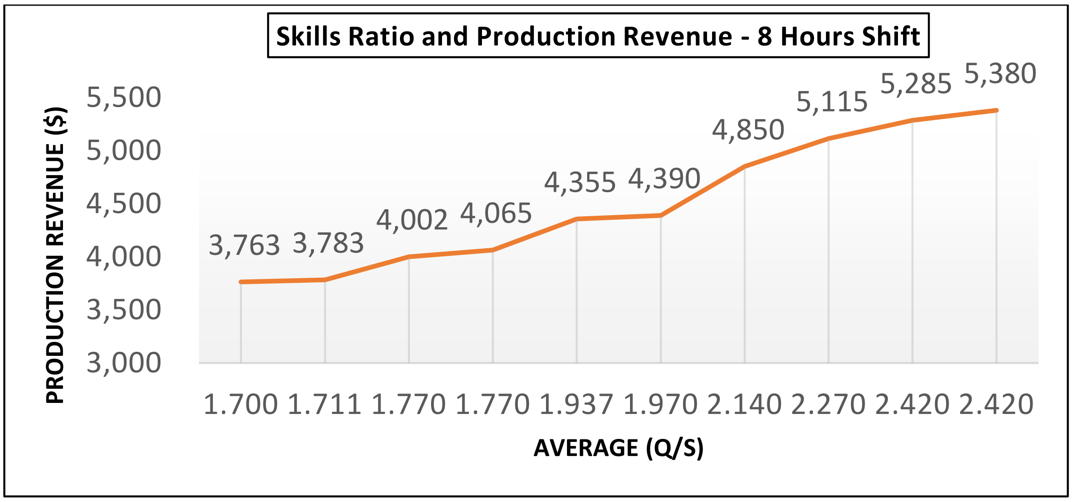

Figure 16. Production revenue associated with average Q/S for 8-hours shift.

In general production revenue would increase, but in some cases, we could detect a decrease in production revenue. This is mainly due to the combination of available workers and the number of working hours determined by the model (not to exceed the production and cost limits). Hence, the model could satisfy all the constraints and maximize the production revenue with the available workers.

Table 20. Workers/Jobs combination

Xij	Average ∑Qil/Sjl	Production Revenue	Yij
13	2.42	5,380	8
22			8
36			7.2
41			7.6
54			7.5
65			8

The second scenario comprises using the same data in the Scenario 1 (except the rotation period). In this scenario 2, the rotation period is 6 hours or less. Then the production revenue associated with the average Q/S ratio for the selected workers was recorded. The results in Table 21 and Table 22 are for the same job and same available workers. Here, due to different rotation period, we obtained different workers to job combination as well different Q/S and production revenue. In this scenario, production rate would decrease, and workers combination changed for the same job required skills and worker skills due the different production capacity as compared in Figure 17.

The third scenario involves testing the model response by applying the worker's predicted performance factor $PFij$ obtained from the regression models. In this scenario, same data in Scenario 2 was used but without changing worker skills level or jobs requirement level data. The assignment model was fed with new prediction revenue of each worker in each job. The model then determines worker to job assignments and the production revenue associated with the average Q/S ratio for the selected workers was recorded and presented in Figure 18. Table 23 and Table 24 show the results for the same job and same available workers but due to difference in worker's predicted performance factor $PFij$, different job assignment as well as different Q/S and production revenue were obtained.

Table 21. workers/ jobs combinations with 8 Hours shift

Xij	Average ∑Qil/Sjl	Production Revenue	Yij
11	2.42	5,285	7.1
23			8
32			8
46			7.2
45			7.5
65			8

Table 22. workers/ jobs combinations with 6 Hours shift

Xij	Average ∑Qil/Sjl	Production Revenue	Yij
13	2.56	4,539	6
62			6
31			6
44			5.6
52			6
65			6

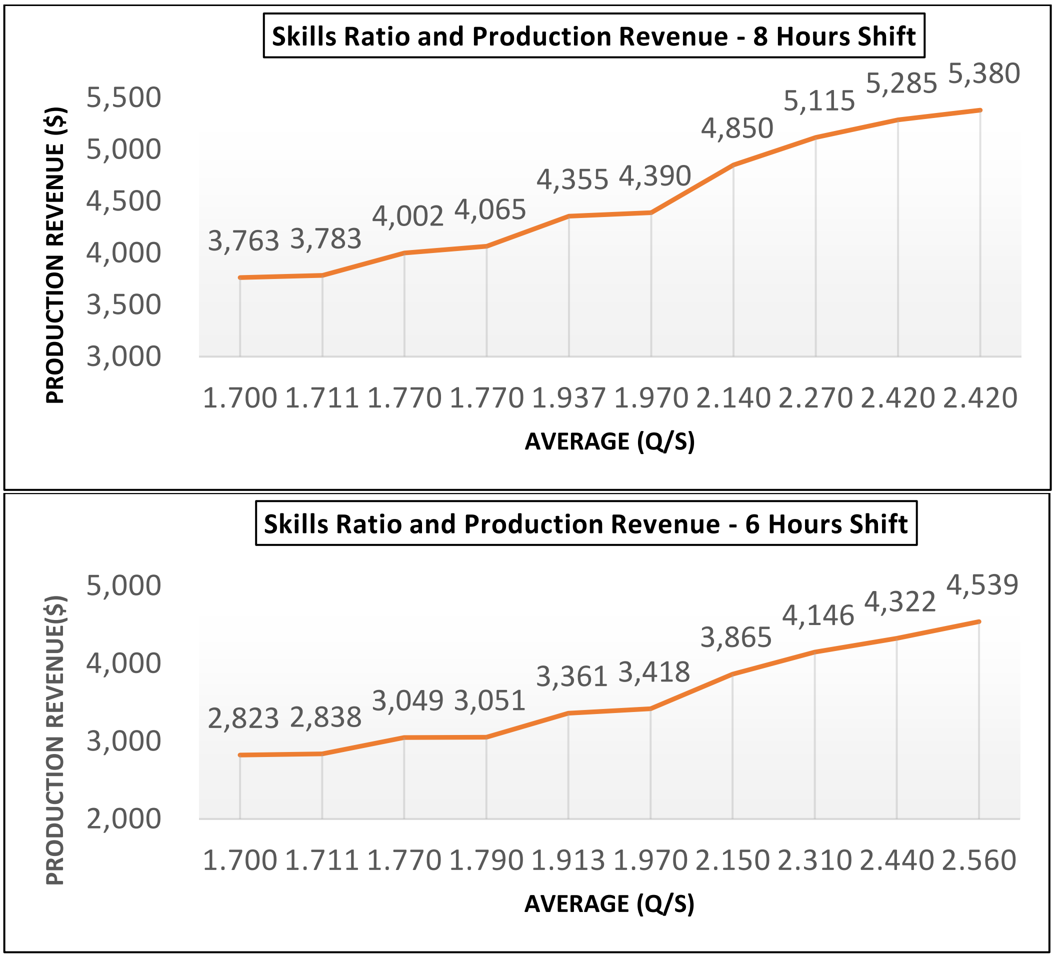

Figure 17. Production revenue associated with average Q/S for 8- and 6-hours shift.

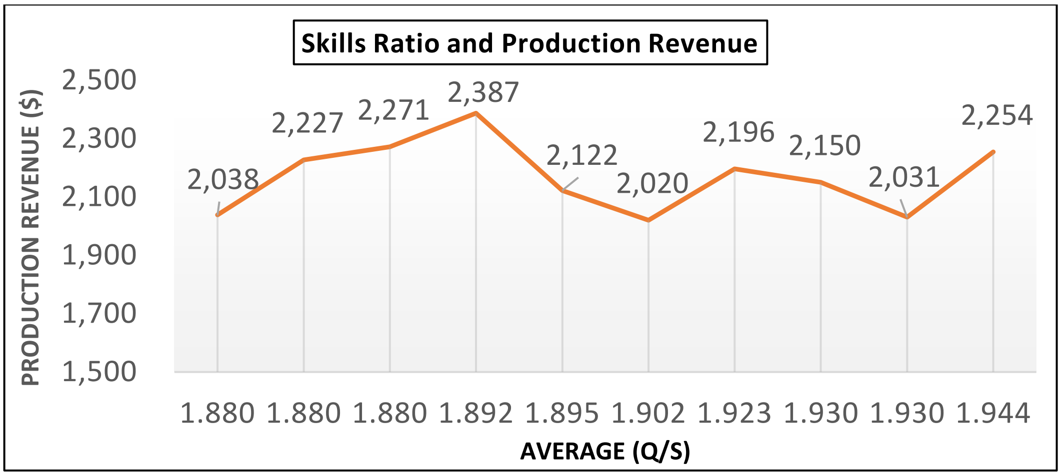

Figure 18. Production revenue associated with average Q/S for same workers and jobs with different predicted PFij (Regression)

Table 23. workers/ jobs combinations with predicted PFij (Regression)

Xij	Average ∑Qil/Sjl	Production Revenue	PFij	Yij
11	1.88	2,038	0.709	6
24			0.687	6
32			0.723	6
45			0.810	6
56			0.489	6
63			0.646	6

worker's predicted performance factor for each job

	j1	j2	j3	j4	j5	j6
i1	0.709	0.615	0.543	0.482	0.661	0.419
i2	0.504	0.737	0.169	0.687	0.801	0.413
i3	0.693	0.723	0.456	0.557	0.700	0.339
i4	0.369	0.585	0.678	0.498	0.810	0.435
i5	0.481	0.763	0.535	0.475	0.601	0.489
i6	0.698	0.512	0.646	0.474	0.728	0.435

Table 24. workers/ jobs combinations with predicted PFij (Regression)

Xij	Average ∑Qil/Sjl	Production Revenue	PFij	Yij
15	1.93	2,150	0.617	6
21			0.852	6
32			0.512	6
44			0.456	6
56			0.793	6
63			0.451	6

worker's predicted performance factor for each job

	j1	j2	j3	j4	j5	j6
i1	0.481	0.349	0.482	0.232	0.617	0.685
i2	0.852	0.453	0.627	0.228	0.224	0.174
i3	0.672	0.512	0.548	0.467	0.627	0.391
i4	0.698	0.176	0.814	0.456	0.645	0.370
i5	0.675	0.350	0.604	0.478	0.626	0.793
i6	0.667	0.440	0.451	0.645	0.706	0.391

The fourth scenario involves testing the model by applying the worker's predicted performance factor $PFij$ obtained from the classification models. In this scenario, same data in Scenario 3 was used. The model could assign worker to job and the production revenue associated with the average Q/S ratio for the selected workers was recorded as presented in Figure19 Figure19 . Table 25 and Table 26 provide the results for the same job and same available workers but due to difference in worker's predicted performance factor $PFij$, different workers to job assignments as well as different Q/S and production revenue were obtained. To this end, it is possible to realize that the assignment model could acceptably deal with both prediction type regression and classification. However, classification shows a higher production revenue curve. Since the classification performance prediction results is either 1 or 0.5, the model will select the worker with high performance whenever it satisfied the constraints and lead to the higher optimal production revenue.

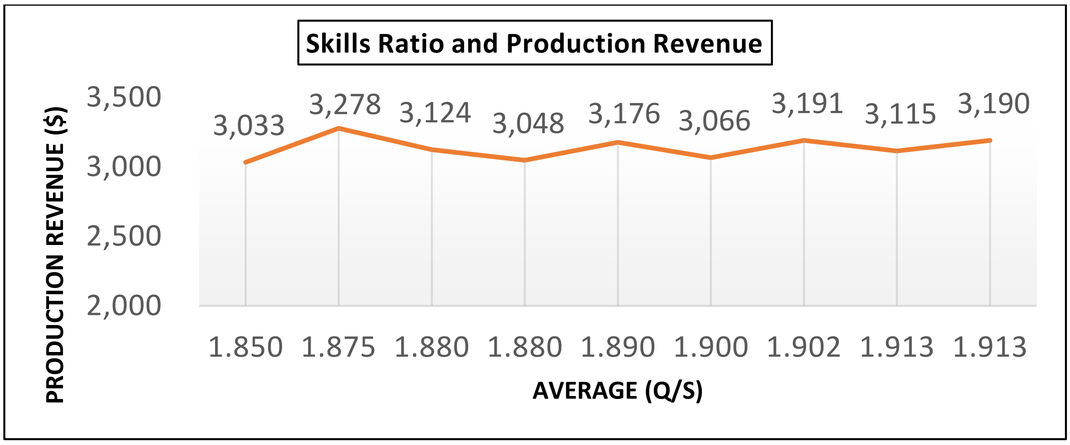

Figure19 . Production revenue associated with average Q/S for same workers and jobs with different predicted PFij (Classification)

Table 25 . Workers/ jobs combinations with predicted PFij (Classification)

Xij	Average ∑Qil/Sjl	Production Revenue	PFij	Yij
16	1.902	3,191	1	6
24			1	6
35			1	6
42			1	6
51			1	6
63			1	6

worker's predicted performance factor for each job

	j1	j2	j3	j4	j5	j6
i1	1	1	1	0.5	1	1
i2	0.5	1	0.5	1	1	0.5
i3	1	1	1	1	1	0.5
i4	0.5	1	1	0.5	1	0.5
i5	1	0.5	1	0.5	0.5	1
i6	1	0.5	1	0.5	1	0.5

Table 26 . Workers/ jobs combinations with predicted PFij (Classification)

Xij	Average ∑Qil/Sjl	Production Revenue	PFij	Yij
12	1.89	3,176	1	6
24			1	6
36			1	6
41			1	6
55			0.5	6
63			1	6

worker's predicted performance factor for each job

	j1	j2	j3	j4	j5	j6
i1	1	1	0.5	1	0.5	0.5
i2	1	0.5	1	1	1	0.5
i3	0.5	0.5	1	0.5	0.5	1
i4	1	1	1	0.5	0.5	0.5
i5	0.5	0.5	0.5	0.5	0.5	0.5
i6	1	0.5	1	1	1	0.5

5. Conclusion

5.1 Summary of results

Industry 4.0 aims to empowering customers with the so-called individualized products, so the manufacturing tasks are changing to accommodate that individual customized products and its associated demands which lead to a need for adaptive employees who are able to perform different types of jobs. This transformation will alter job profiles, necessitating the acquisition of a new and diverse set of worker skills which creates the need for multi-skills workers. A production manager in the era of I4.0 is in need for developing and utilizing an appropriate workforce planning that considers workers with different types of skills to cope with I4.0 requirements and achieve efficient production. In this research an assignment model is developed which is able to assign workers to jobs based on both skills matching considering various skills level and workers' physiological status which impact worker's performance. The ultimate goal of the model is to maximize production profit by selecting worker to job combination that achieves the optimal production value.

Based on a thorough literature review, a model to identify skills needed for industry 4.0 along with their level is presented too. This skills identification model can be used by the production manager to identify skills needed for jobs and possessed by workers to be used as inputs in the assignment model. The model was then tested with fixed worker's skills level and varying skills level. The model properly responded to these changes and assigned workers to certain jobs that resulted in optimal production value. Furthermore, the assignment model responded acceptably to the varying production capacity limit. However, in some situations the model does not provide a solution due to the fact that workers' production level exceed the production capacity. Thus, a new variable was introduced (working hours) to accommodate those workers' production level. The model was tested with different working hour values and provided acceptable performance and resulted in optimal production revenue.

I4.0 technologies provide means to collect workers' physiological status in real-time. Since the worker's status is proven to affect his/her performance. In this study, worker's performance data was applied in the job assignment model. The performance values are predicted by machine learning models using workers' physiological status data, mainly HRV features. To test the assignment model response with real data, machine learning models were developed using available published dataset. The dataset is for number of students who performed 4 cognitive tests which have in total 9 cognitive measures (for our assigning model, each cognitive measure represents a job). Inter beat interval (where HRV features derived) was recorded for each student before starting the tests and the result of these 9 measures were reported. This dataset was used to train prediction models where each cognitive measure considered as target variable and the HRV features are the dependents variables. For each cognitive measures (job) Two types of prediction models were developed, regression and classification. Out of the 9 regression models, 8 models have an acceptable performance. For classification, 7 two-classes classification models have acceptable performance. Balanced accuracy was used to evaluate the classification model. All three-classes classification models were disregarded due to their low balanced accuracy.

To test the assignment model response to the predicted workers' performance value, 6 mutual models were selected (6 for regression and 6 for classification) to predict performance value of 6 workers in 6 jobs. The model responded acceptably to the predicted workers' performance and could provide a job assignment that resulted in optimal production revenue.

To sum up, skills and skill levels need be considered in the assignment model in the area of industry 4.0. along with real-time data about worker status. This assignment model is able to assign workers to jobs based on skills matching considering the skills level and engaging workers' performance factor that is predicted by machine learning models either as regression or classification.

5.2 Limitations

This section describes some limitations in the assigning model and in the machine learning models. For the assigning models, one of its limitations is that the skills required by a job were treated evenly. Suppose a job required three skills with certain level to be performed. In particular, one of these skills is more important and has higher impacts on the job performance compared to the other two. Therefore, skills weight can be introduced to the model where a job requires different skills with certain levels and also each skill has a weight (skills weight summation equal 1). Doing this, the assignment model will treat skills differently.

Another limitation in the assignment model is that both predicted performance factor *PFij* and the Q/S ratio contribute evenly to the worker production. Suppose the machine learning model performance is low. Thus, the predicted performance factor contribution to the workers production rate should be minimized. One way to solve that is by giving a weight for the Q/S ratio and *PFij* (both weight summation equal 1, ws + wp =1), therefore the left side of the objective function could be changed as follow:

$$[\, \mathrm{x}_{ij} * Y_{ij} * V_j * (\mathrm{LCj} * \{\mathrm{ws} * [(\sum_{l=L} \mathrm{Q_{iL}} \;/\; \mathrm{S_{jL}})/L\,] + \mathrm{wp} * PF\mathrm{ij}\})\,]$$

The main limitation of the machine learning process was the dataset size. The dataset size is relatively small, only 38 data rows were used to train the prediction models. Larger dataset would provide more reliable prediction.

Another considerable limitation is that the resting period (where the HRV features were evaluated based on) which is defined before the first test. In the sued dataset, all the tests were taken without having a resting period between them. That could result in inaccurate performance, for example in second test because the participant got effect by the first test. that applicable for third test performance affect by the second test.

5.3 Future Work

This section provides directions for future work regarding the assignment model and the associated prediction model. The presented assignment model in this research focused on worker's performance that was predicted based on real-time data using workers' status. However, other types of real-time data about workers' status such as stress level can be also utilized. For certain jobs, stress can have a huge impact (e.g. jobs that involve decision making). In doing so, a manager can define for each jobs the accepted stress level (STj; High =3,Med=2, Low=1), hence the assignment model considers that in the assignment process. The stress level can be treated as constraint in the model:

$$\text{Xij} * \ (STj - STi) \ \geq 0$$

To predict stress level based on HRV feature, several studies have been performed in this area and couple of studies published the dataset used in their prediction models. Nkurikiyeyezu et al.[144] develop a generic stress prediction model to predict stress level of any person instead of person-specific prediction model. They used a dataset of 15 participant that reaches to 81,892 data rows with approximately 5,300 – 5,460 observations for each participant. Their prediction model accuracy reached 92.5%.

In the machine learning models, the input features (independent variables) for the models are mainly six HRV features (SDNN, RMSSD, pNN50, HF, LF, and LF/HF). Other HRV features could be added and that might improve the model's performance. In addition to the HRV, Electro-Dermal Activity (EDA), and skin temperature (TS) could be also used as input features. According to Sharma et al. [145], cognitive performance naturally involves cognitive workload, which is known to influence physiological responses, such as HRV, EDA, and TS. Therefore, the use of automated methods that rely on measuring these physiological responses is becoming increasingly popular to assess cognitive performance and cognitive workload, which are known to influence an individual's physiological responses.

Other direction for future work to improve the machine learning models can be the use of feature selection technique such as LASSO and RF instead of using predefine input features. Moreover, trying different time window to evaluate HRV features could lead to different results. Hence, instead of using 5 minutes recording period, 1 or 2 minutes could be used or even 30 or 10 seconds. all these time windows have been reported in the literature, however, they are not recommended by the Task Force of The European Society of Cardiology and The North American Society of Pacing and Electrophysiology [131].

References

[1] L. Prifti, M. Knigge, H. Kienegger, and H. Krcmar, “A Competency Model for ‘Industrie 4.0’ Employees,” *Proceedings der 13. Interantionalen Tagung Wirtschaftsinformatik*. pp. 46–60, 2017.

[2] N. Galaske, A. Arnt, H. Friedrich, K. Bettenhausen, and R. Anderl, “Workforce Management 4.0 - Assessment of Human Factors Readiness Towards Digital Manufacturing,” *Advances in Intelligent Systems and Computing*, vol. 606, no. October, pp. 1–5, 2018, doi: 10.1007/978-3-319-60474-9.

[3] S. Jaschke, “Mobile learning applications for technical vocational and engineering education: The use of competence snippets in laboratory courses and industry 4.0,” *Proceedings of 2014 International Conference on Interactive Collaborative Learning, ICL 2014*, no. December, pp. 605–608, 2014, doi: 10.1109/ICL.2014.7017840.

[4] H. Kagermann, W. Wahlster, and J. Helbig, “Securing the future of German manufacturing industry: Recommendations for implementing the strategic initiative INDUSTRIE 4.0,” *Final Report of the Industrie 4.0 Working Group*, no. April, pp. 1–84, 2013.

[5] A. Baron, R. Clake, P. Turner, and S. Pass, “Workforce planning: Right people, right time, right skills,” 2010.

[6] S. Mattsson, Å. Fast-Berglund, and M. Åkerman, “Assessing Operator Wellbeing through Physiological Measurements in Real-Time—Towards Industrial Application,” *Technologies (Basel)*, vol. 5, no. 4, p. 61, 2017, doi: 10.3390/technologies5040061.

[7] HSE, *Understanding Human Failure: Human Failure and Accidents HSG48*, vol. HSG48. 1999. doi: 10.1108/ss.1999.11017cae.007.

[8] S. Mattsson, Å. Fast-Berglund, and P. Thorvald, “A Relationship between Operator Performance and Arousal in Assembly,” *Procedia CIRP*, vol. 44, pp. 32–37, 2016, doi: 10.1016/j.procir.2016.02.158.

[9] T. A. Wright and R. Cropanzano, “The role of psychological well-being in job performance: A fresh look at an age-old quest,” *Organ Dyn*, vol. 33, no. 4 SPEC.ISS., pp. 338–351, 2004, doi: 10.1016/j.orgdyn.2004.09.002.

[10] M. T. Ford, C. P. Cerasoli, J. A. Higgins, and A. L. Decesare, “Relationships between psychological, physical, and behavioural health and work performance: A review and meta-analysis,” *Work Stress*, vol. 25, no. 3, pp. 185–204, 2011, doi: 10.1080/02678373.2011.609035.

[11] T. A. Wright, R. Cropanzano, and D. G. Bonett, “The moderating role of employee positive well being on the relation between job satisfaction and job performance,” *J Occup Health Psychol*, vol. 12, no. 2, pp. 93–104, 2007, doi: 10.1037/1076-8998.12.2.93.

[12] A. Muaremi, B. Arnrich, and G. Tröster, “Towards Measuring Stress with Smartphones and Wearable Devices During Workday and Sleep,” *Bionanoscience*, vol. 3, no. 2, pp. 172–183, 2013, doi: 10.1007/s12668-013-0089-2.

[13] D. Li, A. Landström, S. Mattsson, and M. Karlsson, “How changes in cognitive automation can affect operator performance and productivity,” *The 6th International Swedish Production Symposium*, no. 2013, 2014, doi: 10.13140/2.1.5110.9128.

[14] P. U. Mats Bohgard , Stig Karlsson , Eva Lovén , Lars-Åke Mikaelsson , Lena Mårtensson, Anna-Lisa Osvalder, Linda Rose, *Work and Technology on Human Terms*. 2009.

[15] A. Schumacher, S. Erol, and W. Sihn, “A Maturity Model for Assessing Industry 4.0 Readiness and Maturity of Manufacturing Enterprises,” *Procedia CIRP*, vol. 52, no. December, pp. 161–166, 2016, doi: 10.1016/j.procir.2016.07.040.

[16] Shiyong Wang, Jiafu Wan, Di Li, and Chunhua Zhang, “Implementing Smart Factory of Industrie 4.0: An Outlook,” *Int J Distrib Sens Netw*, vol. 2016, no. 1, pp. 1–10, 2016.

[17] S. I. Tay, T. C. Lee, N. Z. A. Hamid, and A. N. A. Ahmad, “An overview of industry 4.0: Definition, components, and government initiatives,” *Journal of Advanced Research in Dynamical and Control Systems*, vol. 10, no. 14, pp. 1379–1387, 2018.

[18] P. Shaw PGDM Student, X. Kochi Ranjana Mary Varghese Assistant Dean, and X. Kochi, “Industry 4.0 and Future of Hr,” *Journal of Management (JOM*, vol. 5, no. 6, pp. 96–103, 2018.

[19] H. Balalle and R. Balalle, “Fourth industrial revolution and future of workforce,” *International Journal of Advance Research, Idea and Innovations in Technology*, vol. 4, no. 5, pp. 151–153, 2018.

[20] G. Rehe, R. Schneider, and D. Thimm, “Transformation towards human workforce 4.0 through the qualification model LTA-FIT,” in *Proceedings of the Annual Hawaii International Conference on System Sciences*, 2020, pp. 4601–4610. doi: 10.24251/hicss.2020.563.

[21] M. Schinner, A. C. Valdez, E. Noll, A. K. Schaar, and P. Letmathe, “Psychological and a Managerial Perspective Pervasive digitalization Integrated cyber-physical systems,” 2017.

[22] F. Baena, A. Guarin, J. Mora, J. Sauza, and S. Retat, “Learning Factory: The Path to Industry 4.0,” *Procedia Manuf*, vol. 9, pp. 73–80, 2017, doi: 10.1016/j.promfg.2017.04.022.

[23] L. Bonekamp and M. Sure, “Consequences of Industry 4.0 on human labour and work organisation,” *Journal of Business and Media Psychology*, vol. 6, no. 1, pp. 33–40, 2015.

[24] F. Hecklau, M. Galeitzke, S. Flachs, and H. Kohl, “Holistic Approach for Human Resource Management in Industry 4.0,” *Procedia CIRP*, vol. 54. pp. 1–6, 2016. doi: 10.1016/j.procir.2016.05.102.

[25] S. Fareri, F. Chiarello, E. Coli, D. Teloni, G. Dente, and G. Fantoni, “Workers 4.0: skills, profiles and jobs in different business functions,” *Fondazione Giacomo Brodolini*, pp. 95–107, 2018.

[26] Mindtools.com, “Porter’s value chain: understanding how value is created within organizations,” 2015. https://www.mindtools.com/pages/article/newSTR_66.htm

[27] G. Fantoni, F. Chiarello, S. Fareri, S. Pira, and A. Guadagni, “Defining industry 4.0 professional-data-driven approach 75-95.pdf,” *Fondazione Giacomo Brodolini*, pp. 75–94, 2018.

[28] Yigit Kazancoglu and Y. O.-O. Deniz, “Analyzing Workforce 4.0 in the Fourth Industrial Revolution and proposing a road map from operations management perspective with fuzzy DEMATEL,” *Journal of Enterprise Information Management*, vol. 31, no. 6, pp. 891–907, 2018.

[29] B. Dworschak and H. Zaiser, “Competences for cyber-physical systems in manufacturing-First findings and scenarios,” *Procedia CIRP*, vol. 25, no. C, pp. 345–350, 2014, doi: 10.1016/j.procir.2014.10.048.

[30] F. Ferrier, “and training journey into the future going boldly national vocational system education innovation future training boldly vocational education journey going boldly national innovation system training Going boldly into the future A VET journey into the nati”.

[31] E. A. Hartmann and M. Bovenschulte, “Skills Needs Analysis for ‘Industry 4.0’ based on Roadmaps for Smart Systems,” *In: SKOLKOVO Moscow School of Management & International Labour Organization (ed.) (2013): Using Technology Foresights for Identifying Future Skills Needs. Global Workshop Proceedings*, pp. 24–36, 2013.

[32] EPoSS, “Strategic Research Agenda,” no. June, 2009.

[33] S. Adolph, M. Tisch, and J. Metternich, “Challenges and approaches to competency development for future production,” *Journal of International Scientific Publications*, vol. 12, no. September, pp. 1001–1010, 2014.

[34] A. Mittelmann, “Competence Development for Work 4.0,” pp. 263–275, 2018, doi: 10.1007/978-3-319-73546-7_16.

[35] European Commission, “The Future of Work Skills and Resilience for a World of Change,” *EPSC Strategic Notes*, no. 13, p. 12, 2016.

[36] R. Anderl and J. Fleischer, “Guideline Industrie 4.0,” *VDMA Verlag*, p. 32, 2015.

[37] Y. Wang, G. Wang, and R. Anderl, “Generic Procedure Model to introduce Industrie 4.0 In Small and Medium-sized Enterprises,” *Lecture Notes in Engineering and Computer Science*, vol. 2226, pp. 971–976, 2016.

[38] D. Mourtzis, *Development of Skills and Competences in Manufacturing Towards Education 4.0: A Teaching Factory Approach*, no. 9783319895628. Springer International Publishing, 2018. doi: 10.1007/978-3-319-89563-5.

[39] K. L. Kusmin, T. Ley, and P. Normak, “Towards a data driven competency management platform for industry 4.0,” *CEUR Workshop Proc*, vol. 2025, 2017.

[40] D. Buhr, “Social Innovation Policy,” no. January 2015, 2017.

[41] A. Łupicka and K. Grzybowska, “Key Managerial Competencies for Industry 4.0 - Practitioners’, Researchers’ and Students’ Opinions,” *Logistics and Transport*, vol. 3, pp. 39–46, 2018, doi: 10.26411/83-1734-2015-3-39-4-18.

[42] F. Acerbi, S. Assiani, and M. Taisch, “A Methodology to Assess the Skills for an Industry 4.0 Factory,” in *IFIP Advances in Information and Communication Technology*, 2019, pp. 520–527. doi: 10.1007/978-3-030-29996-5_60.

[43] M. Pinzone, P. Fantini, S. Perini, S. Garavaglia, M. Taisch, and G. Miragliotta, "Jobs and skills in industry 4.0: An exploratory research," *IFIP Adv Inf Commun Technol*, vol. 513, pp. 282–288, 2017, doi: 10.1007/978-3-319-66923-6_33.

[44] European Committee for Standardization (CEN), "User guide for the application of the European e-Competence Framework 3.0," 2014.

[45] "O*NET Resource Center." https://www.onetcenter.org/

[46] T. K. Bensghir and U. Turen, "Industry 4.0 from the MIS Perspective," *Industry 4.0 from the MIS Perspective*, no. November, 2019, doi: 10.3726/b15120.

[47] L. Gehrke, "A Discussion of Qualifications and A German and American Perspective," *ASME American Society of Mechanical Engineers, VDI The Association of German Engineers publications*, no. April, p. 29, 2015.

[48] C. B. Frey and M. A. Osborne, "The Future of employment Computerisation?," pp. 1–72, 2013.

[49] M. D. Bermúdez and B. F. Juárez, "Competencies to adopt Industry 4.0 for operations management personnel at automotive parts suppliers in Nuevo Leon," *Proceedings of the International Conference on Industrial Engineering and Operations Management*, vol. 2017, no. OCT, pp. 736–747, 2017.

[50] M. Lorenz, M. Rüßmann, R. Strack, K. L. Lueth, and M. Bolle, "Man and Machine in Industry 4.0," *Boston Consulting Group*, p. 18, 2015.

[51] O. Bongomin, G. Gilibrays Ocen, E. Oyondi Nganyi, A. Musinguzi, and T. Omara, "Exponential Disruptive Technologies and the Required Skills of Industry 4.0," *Journal of Engineering (United Kingdom)*, vol. 2020, no. October, 2020, doi: 10.1155/2020/4280156.

[52] M. J. Sousa and Á. Rocha, "Skills for disruptive digital business," *J Bus Res*, vol. 94, no. December 2017, pp. 257–263, 2019, doi: 10.1016/j.jbusres.2017.12.051.

[53] M. A. Peters, "Technological unemployment: Educating for the fourth industrial revolution," *Educational Philosophy and Theory*, vol. 49, no. 1, pp. 1–6, 2017, doi: 10.1080/00131857.2016.1177412.

[54] A. Benešová and J. Tupa, "Requirements for Education and Qualification of People in Industry 4.0," *Procedia Manuf*, vol. 11, no. June, pp. 2195–2202, 2017, doi: 10.1016/j.promfg.2017.07.366.

[55] S. M. Sackey and A. Bester, "Industrial engineering curriculum in industry 4.0 in a South African context," *South African Journal of Industrial Engineering*, vol. 27, no. 4, pp. 101–114, 2016, doi: 10.7166/27-4-1579.

[56] S. Coşkun, Y. Kayıkcı, and E. Gençay, "Adapting Engineering Education to Industry 4.0 Vision," *Technologies (Basel)*, vol. 7, no. 1, p. 10, 2019, doi: 10.3390/technologies7010010.

[57] W. Maisiri, H. Darwish, and L. van Dyk, "An investigation of industry 4.0 skills requirements," *South African Journal of Industrial Engineering*, vol. 30, no. 3, pp. 90–105, 2019, doi: 10.7166/30-3-2230.

[58] G. B. Cotet, B. A. Balgiu, and V. C. Z. Negrea, "Assessment procedure for the soft skills requested by Industry 4.0," *MATEC Web of Conferences*, vol. 121, pp. 1–8, 2017, doi: 10.1051/matecconf/201712107005.

[59] H. James Wilson and P. R. Daugherty, "Collaborative intelligence: Humans and AI are joining forces," *Harv Bus Rev*, vol. 2018, no. July-August, pp. 1–12, 2018.

[60] J. Guszcza, H. Lewis, and P. Evans-Greenwood, "Cognitive collaboration: Why humans and computers think better together," no. 20, p. 30, 2016.

[61] P. Tsarouchi, A. Matthaiakis, and S. Makris, "On a human-robot collaboration in an assembly cell," *Int J Comput Integr Manuf*, vol. 30, no. 6, pp. 580–589, 2017, doi: 10.1080/0951192X.2016.1187297.

[62] M. Fiasche, M. Pinzone, P. Fantini, A. Alexandru, and M. Taisch, "Human-centric factories 4.0: A mathematical model for job allocation," in *2016 IEEE 2nd International Forum on Research and Technologies for Society and Industry Leveraging a Better Tomorrow, RTSI 2016*, 2016, pp. 1–4. doi: 10.1109/RTSI.2016.7740613.

[63] G. M. Campbell, "A two-stage stochastic program for scheduling and allocating cross-trained workers," *Journal of the Operational Research Society*, vol. 62, no. 6, pp. 1038–1047, 2011, doi: 10.1057/jors.2010.16.

[64] W. J. Abernathy, N. Baloff, J. C. Hershey, and S. Wandel, "THREE-STAGE MANPOWER PLANNING AND SCHEDULING MODEL - A SERVICE-SECTOR EXAMPLE.," *Oper Res*, vol. 21, no. 3, pp. 693–711, 1973, doi: 10.1287/opre.21.3.693.

[65] G. M. Campbell, "Cross-utilization of workers whose capabilities differ," *Manage Sci*, vol. 45, no. 5, pp. 722–732, 1999, doi: 10.1287/mnsc.45.5.722.

[66] S. Digiesi, F. Facchini, G. Mossa, and G. Mummolo, "A RULA-Based Optimization Model for Workers ' Assignment to an Assembly Line," pp. 8–13, 2017.

[67] B. A. Norman, W. Tharmmaphornphilas, K. L. Needy, B. Bidanda, and R. C. Warner, "Worker assignment in cellular manufacturing considering technical and human skills," *Int J Prod Res*, vol. 40, no. 6, pp. 1479–1492, Jan. 2002, doi: 10.1080/00207540110118082.

[68] S. D. Akyol and A. Baykasoğlu, "A multiple-rule based constructive randomized search algorithm for solving assembly line worker assignment and balancing problem," *J Intell Manuf*, vol. 30, no. 2, pp. 557–573, 2019, doi: 10.1007/s10845-016-1262-6.

[69] C. Miralles, J. P. García-Sabater, C. Andrés, and M. Cardós, "Branch and bound procedures for solving the Assembly Line Worker Assignment and Balancing Problem: Application to Sheltered Work centres for Disabled," *Discrete Appl Math (1979)*, vol. 156, no. 3, pp. 352–367, 2008, doi: 10.1016/j.dam.2005.12.012.

[70] A. Baykasoğlu, "Multi-rule multi-objective simulated annealing algorithm for straight and U type assembly line balancing problems," *J Intell Manuf*, vol. 17, no. 2, pp. 217–232, 2006, doi: 10.1007/s10845-005-6638-y.

[71] M. C. O. Moreira, M. Ritt, A. M. Costa, and A. A. Chaves, "Simple heuristics for the assembly line worker assignment and balancing problem," *Journal of Heuristics*, vol. 18, no. 3, pp. 505–524, 2012, doi: 10.1007/s10732-012-9195-5.

[72] C. Weckenborg, K. Kieckhäfer, C. Müller, M. Grunewald, and T. S. Spengler, “Balancing of assembly lines with collaborative robots,” *Business Research*, vol. 13, no. 1, pp. 93–132, 2020, doi: 10.1007/s40685-019-0101-y.

[73] E. Helms, R. D. Sehraft, and M. Hägele, “Rob@work: Robot assistant in industrial environments,” *Proceedings - IEEE International Workshop on Robot and Human Interactive Communication*, pp. 399–404, 2002, doi: 10.1109/ROMAN.2002.1045655.

[74] G. Michalos, S. Makris, J. Spiliotopoulos, I. Misios, P. Tsarouchi, and G. Chryssolouris, “ROBO-PARTNER: Seamless human-robot cooperation for intelligent, flexible and safe operations in the assembly factories of the future,” *Procedia CIRP*, vol. 23, no. C, pp. 71–76, 2014, doi: 10.1016/j.procir.2014.10.079.

[75] Y. K. Kim, W. S. Song, and J. H. Kim, “A mathematical model and a genetic algorithm for two-sided assembly line balancing,” *Comput Oper Res*, vol. 36, no. 3, pp. 853–865, 2009, doi: 10.1016/j.cor.2007.11.003.

[76] R. Esmaeilbeigi, B. Naderi, and P. Charkhgard, “New formulations for the setup assembly line balancing and scheduling problem,” *OR Spectrum*, vol. 38, no. 2, pp. 493–518, 2016, doi: 10.1007/s00291-016-0433-3.

[77] J. E. Gomar, C. T. Haas, and D. P. Morton, “Assignment and Allocation Optimization of Partially Multiskilled Workforce,” *J Constr Eng Manag*, vol. 128, no. 2, pp. 103–109, 2002, doi: 10.1061/(asce)0733-9364(2002)128:2(103).

[78] Rodrigues Ana Maria, “Planningandschedulingmultiskilledworkforce.Pdf.” Univ. of Texas at Austin, Austin, Tex, 1998.

[79] J. Lian, C. Liu, W. Li, and Y. Yin, “Computers & Industrial Engineering A multi-skilled worker assignment problem in seru production systems considering the worker heterogeneity,” *Comput Ind Eng*, vol. 118, no. January, pp. 366–382, 2018, doi: 10.1016/j.cie.2018.02.035.

[80] F. Niakan, A. Baboli, T. Moyaux, and V. Botta-Genoulaz, “A bi-objective model in sustainable dynamic cell formation problem with skill-based worker assignment,” *J Manuf Syst*, vol. 38, pp. 46–62, 2016, doi: 10.1016/j.jmsy.2015.11.001.

[81] M. M. Savino, A. Brun, and A. Mazza, “Dynamic workforce allocation in a constrained flow shop with multi-agent system,” *Comput Ind*, vol. 65, no. 6, pp. 967–975, 2014, doi: 10.1016/j.compind.2014.02.016.

[82] G. Neubert and M. Savino, “Flow shop operator scheduling through constraint satisfaction and constraint optimisation techniques,” *International Journal of Productivity and Quality Management*, vol. 4, no. 5–6, pp. 549–568, 2009, doi: 10.1504/IJPQM.2009.025185.

[83] A. Ayough, M. Zandieh, and H. Farsijani, “GA and ICA approaches to job rotation scheduling problem: Considering employee’s boredom,” *International Journal of Advanced Manufacturing Technology*, vol. 60, no. 5–8, pp. 651–666, 2012, doi: 10.1007/s00170-011-3641-7.

[84] E. J. Lodree, C. D. Geiger, and X. Jiang, “Taxonomy for integrating scheduling theory and human factors: Review and research opportunities,” *Int J Ind Ergon*, vol. 39, no. 1, pp. 39–51, 2009, doi: 10.1016/j.ergon.2008.05.001.

[85] J. Bhadury and Z. Radovilsky, “Job rotation using the multi-period assignment model,” *Int J Prod Res*, vol. 44, no. 20, pp. 4431–4444, 2006, doi: 10.1080/00207540500057621.

[86] G. Mossa, F. Boenzi, S. Digiesi, G. Mummolo, and V. A. Romano, “Productivity and ergonomic risk in human based production systems: A job-rotation scheduling model,” *Int J Prod Econ*, vol. 171, pp. 471–477, 2016, doi: 10.1016/j.ijpe.2015.06.017.

[87] P. P. F. M. Kuijer, B. Visser, and H. C. G. Kemper, “Job rotation as a factor in reducing physical workload at a refuse collecting department,” *Ergonomics*, vol. 42, no. 9, pp. 1167–1178, 1999, doi: 10.1080/001401399185054.

[88] S. E. Moussavi, M. Mahdjoub, and O. Grunder, “Productivity improvement through a sequencing generalised assignment in an assembly line system,” *Int J Prod Res*, vol. 55, no. 24, pp. 7509–7523, 2017, doi: 10.1080/00207543.2017.1378828.

[89] S. E. Moussavi, M. Mahdjoub, and O. Grunder, “Reducing production cycle time by ergonomic workforce scheduling,” *IFAC-PapersOnLine*, vol. 49, no. 12, pp. 419–424, 2016, doi: 10.1016/j.ifacol.2016.07.642.

[90] F. Bentefouet and D. A. Nembhard, “Optimal flow-line conditions with worker variability,” *Int J Prod Econ*, vol. 141, no. 2, pp. 675–684, 2013, doi: 10.1016/j.ijpe.2012.10.008.

[91] B. Xin, Y. Li, J. Yu, and J. Zhang, “An adaptive BPSO algorithm for multi-skilled workers assignment problem in aircraft assembly lines,” *Assembly Automation*, vol. 35, no. 4, pp. 317–328, 2015, doi: 10.1108/AA-06-2015-051.

[92] S. Ghinato, Paulo;Fujii, “a Basic Approach To the Multifunction Workers Assignment Problem in U-Shaped Production Lines,” pp. 1–8.

[93] L. Cao and B. Wang, “Personnel assignment for assembly production with worker differences,” *Proceedings of the 2010 10th International Conference on Intelligent Systems Design and Applications, ISDA’10*, pp. 1476–1480, 2010, doi: 10.1109/ISDA.2010.5687119.

[94] R. W. Y. Yee, A. C. L. Yeung, and T. C. E. Cheng, “The impact of employee satisfaction on quality and profitability in high-contact service industries,” *Journal of Operations Management*, vol. 26, no. 5, pp. 651–668, Sep. 2008, doi: 10.1016/j.jom.2008.01.001.

[95] M. Liu and X. Yang, “Bi-objective optimization for Scheduling and Multi-skilled Worker Assignments in the Hybrid Flow Shop,” *IFAC-PapersOnLine*, vol. 52, no. 13, pp. 2128–2133, 2019, doi: 10.1016/j.ifacol.2019.11.520.

[96] L. Wu, F. T. S. Chan, B. Niu, and L. Li, “Cross-trained worker assignment and comparative analysis on throughput of divisional and rotating seru,” *Industrial Management and Data Systems*, vol. 118, no. 5, pp. 1114–1136, 2018, doi: 10.1108/IMDS-07-2017-0303.

[97] R. R. Inman, W. C. Jordan, and D. E. Blumenfeld, “Chained cross-training of assembly line workers,” *Int J Prod Res*, vol. 42, no. 10, pp. 1899–1910, 2004, doi: 10.1080/00207540410001666224.

[98] C. Liu, J. Lian, Y. Yin, and W. Li, “Seru Seisan- an innovation of the production management Mode in Japan,” *Asian Journal of Technology Innovation*, vol. 18, no. 2, pp. 89–113, Jan. 2010, doi: 10.1080/19761597.2010.9668694.

[99] C. Liu, W. Li, J. Lian, and Y. Yin, “Reconfiguration of assembly systems: From conveyor assembly line to serus,” *J Manuf Syst*, vol. 31, no. 3, pp. 312–325, 2012, doi: 10.1016/j.jmsy.2012.02.003.

[100] L. Wu, F. Cai, L. Li, and X. Chu, “Cross-trained worker assignment problem in cellular manufacturing system using swarm intelligence metaheuristics,” *Math Probl Eng*, vol. 2018, 2018, doi: 10.1155/2018/4302062.

[101] Ö. Mutlu, O. Polat, and A. A. Supciller, “An iterative genetic algorithm for the assembly line worker assignment and balancing problem of type-II,” *Comput Oper Res*, vol. 40, no. 1, pp. 418–426, 2013, doi: 10.1016/j.cor.2012.07.010.

[102] P. Samouei and J. Ashayeri, “Developing optimization & robust models for a mixed-model assembly line balancing problem with semi-automated operations,” *Appl Math Model*, vol. 72, pp. 259–275, 2019, doi: 10.1016/j.apm.2019.02.019.

[103] Y. S. Can *et al.*, “Real-Life Stress Level Monitoring Using Smart Bands in the Light of Contextual Information,” *IEEE Sens J*, vol. 20, no. 15, pp. 8721–8730, 2020, doi: 10.1109/JSEN.2020.2984644.

[104] Y. S. Can, B. Arnrich, and C. Ersoy, “Stress detection in daily life scenarios using smart phones and wearable sensors: A survey,” *J Biomed Inform*, vol. 92, p. 103139, 2019, doi: 10.1016/j.jbi.2019.103139.

[105] Jø. Winkel and S. E. Mathiassen, “Assessment of physical work load in epidemiologic studies: Concepts, issues and operational considerations,” *Ergonomics*, vol. 37, no. 6, pp. 979–988, 1994, doi: 10.1080/00140139408963711.

[106] D. Romero, S. Mattsson, Å. Fast-Berglund, T. Wuest, D. Gorecky, and J. Stahre, “Digitalizing occupational health, safety and productivity for the operator 4.0,” *IFIP Adv Inf Commun Technol*, vol. 536, no. August, pp. 473–481, 2018, doi: 10.1007/978-3-319-99707-0_59.

[107] J. Khakurel, S. Pöysä, and J. Porras, “The use of wearable devices in the workplace - A systematic literature review,” *Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, LNICST*, vol. 195 LNICST, pp. 284–294, 2017, doi: 10.1007/978-3-319-61949-1_30.

[108] PwC, “The Wearable Future,” *Consumer Intelligence Series*, pp. 1–50, 2014.

[109] D. Battini, S. Finco, and F. Sgarbossa, “Human-oriented assembly line balancing and sequencing model in the industry 4.0 era,” *International Series in Operations Research and Management Science*, vol. 289, no. March 2021, pp. 141–165, 2020, doi: 10.1007/978-3-030-43177-8_8.

[110] M. Calzavara, A. Persona, F. Sgarbossa, and V. Visentin, “A device to monitor fatigue level in order-picking,” *Industrial Management and Data Systems*, vol. 118, no. 4, pp. 714–727, 2018, doi: 10.1108/IMDS-05-2017-0182.

[111] H. Jebelli, B. Choi, and S. Lee, “Application of Wearable Biosensors to Construction Sites. I: Assessing Workers’ Stress,” *J Constr Eng Manag*, vol. 145, no. 12, p. 04019079, 2019, doi: 10.1061/(asce)co.1943-7862.0001729.

[112] R. Castaldo, P. Melillo, U. Bracale, M. Caserta, M. Triassi, and L. Pecchia, “Acute mental stress assessment via short term HRV analysis in healthy adults: A systematic review with meta-

analysis," *Biomed Signal Process Control*, vol. 18, no. January 2019, pp. 370–377, 2015, doi: 10.1016/j.bspc.2015.02.012.

[113] L. Salahuddin and K. Desok, "Detection of acute stress by heart rate variability using a prototype mobile ECG sensor," *Proceedings - 2006 International Conference on Hybrid Information Technology, ICHIT 2006*, vol. 2, pp. 453–459, 2006, doi: 10.1109/ICHIT.2006.253646.

[114] G. Tan, T. K. Dao, L. Farmer, R. J. Sutherland, and R. Gevirtz, "Heart rate variability (HRV) and posttraumatic stress disorder (PTSD): A pilot study," *Applied Psychophysiology Biofeedback*, vol. 36, no. 1, pp. 27–35, 2011, doi: 10.1007/s10484-010-9141-y.

[115] A. Levine, O. Zagoory-Sharon, R. Feldman, J. G. Lewis, and A. Weller, "Measuring cortisol in human psychobiological studies," *Physiol Behav*, vol. 90, no. 1, pp. 43–53, 2007, doi: 10.1016/j.physbeh.2006.08.025.

[116] K. Vedhara *et al.*, "An investigation into the relationship between salivary cortisol, stress, anxiety and depression," *Biol Psychol*, vol. 62, no. 2, pp. 89–96, 2003, doi: 10.1016/S0301-0511(02)00128-X.

[117] G. Gancitano *et al.*, "HRV in active-duty special forces and public order military personnel," *Sustainability (Switzerland)*, vol. 13, no. 7. 2021. doi: 10.3390/su13073867.

[118] N. Izzah, A. P. Sutarto, and M. Hariyadi, "Machine Learning models for the Cognitive Stress Detection Using Heart Rate Variability Signals," *Jurnal Teknik Industri*, vol. 24, no. 2, pp. 83–94, Nov. 2022, doi: 10.9744/jti.24.2.83-94.

[119] A. Matuz, D. van der Linden, G. Darnai, and Á. Csathó, "Machine learning models in Heart Rate Variability based mental fatigue prediction: training on heterogeneous data to obtain robust models," 2022. doi: 10.21203/rs.3.rs-1633975.

[120] K. Tsunoda, A. Chiba, K. Yoshida, T. Watanabe, and O. Mizuno, "Predicting changes in cognitive performance using heart rate variability," *IEICE Trans Inf Syst*, vol. E100D, no. 10, pp. 2411–2419, 2017, doi: 10.1587/transinf.2016OFP0002.

[121] K. Tsunoda, A. Chiba, H. Chigira, T. Ura, and O. Mizunq, "Estimating changes in a cognitive performance using heart rate variability," *2015 IEEE 15th International Conference on Bioinformatics and Bioengineering, BIBE 2015*, pp. 5–10, 2015, doi: 10.1109/BIBE.2015.7367712.

[122] K. Tsunoda, A. Chiba, H. Chigira, K. Yoshida, T. Watanabe, and O. Mizuno, "Online estimation of a cognitive performance using heart rate variability," *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, vol. 2016-Octob, pp. 761–765, 2016, doi: 10.1109/EMBC.2016.7590813.

[123] C. L. Schaich *et al.*, "Association of heart rate variability with cognitive performance: the multi-ethnic study of atherosclerosis," *J Am Heart Assoc*, vol. 9, no. 7, 2020, doi: 10.1161/JAHA.119.013827.

[124] Y. Kadoya, K. Mostafa Saidur Rahim, S. Watanapongvanich, and P. Binnagan, "Emotional status and productivity: Evidence from the special economic zone in Laos," *Sustainability (Switzerland)*, vol. 12, no. 4, pp. 1–15, 2020, doi: 10.3390/su12041544.

[125] K. H. Choi, J. Kim, O. S. Kwon, M. J. Kim, Y. H. Ryu, and J. E. Park, “Is heart rate variability (HRV) an adequate tool for evaluating human emotions? – A focus on the use of the International Affective Picture System (IAPS),” *Psychiatry Res*, vol. 251, no. January, pp. 192–196, 2017, doi: 10.1016/j.psychres.2017.02.025.

[126] G. Valenza, A. Lanata, and E. P. Scilingo, “The role of nonlinear dynamics in affective valence and arousal recognition,” *IEEE Trans Affect Comput*, vol. 3, no. 2, pp. 237–249, 2012, doi: 10.1109/T-AFFC.2011.30.

[127] R. D. Lane, K. McRae, E. M. Reiman, K. Chen, G. L. Ahern, and J. F. Thayer, “Neural correlates of heart rate variability during emotion,” *Neuroimage*, vol. 44, no. 1, pp. 213–222, 2009, doi: 10.1016/j.neuroimage.2008.07.056.

[128] B. Grässler, A. Hökelmann, and R. H. Cabral, “Resting heart rate variability as a possible marker of cognitive decline,” *Kinesiology*, vol. 52, no. 1. University of Zagreb - Faculty of Kinesiology, pp. 72–84, Jun. 01, 2020. doi: 10.26582/k.52.1.9.

[129] G. Forte, M. Morelli, B. Grässler, and M. Casagrande, “Decision making and heart rate variability: A systematic review,” *Appl Cogn Psychol*, vol. 36, no. 1, pp. 100–110, Jan. 2022, doi: 10.1002/acp.3901.

[130] E. Alharbi and A. Alomainy, “Machine Learning approach to Predict Cognitive Performance using HRV,” in *Proceedings of 2022 2nd International Conference on Computing and Information Technology, ICCIT 2022*, Institute of Electrical and Electronics Engineers Inc., 2022, pp. 379–382. doi: 10.1109/ICCIT52419.2022.9711625.

[131] T. Force of The European Society of Cardiology and T. North American Society of Pacing, “Guidelines Heart rate variability Standards of measurement, physiological interpretation, and clinical use,” 1996.

[132] “DataRobot.” https://www.datarobot.com/platform/ (accessed Mar. 03, 2023).

[133] Z. Gao, X. Cui, W. Wan, W. Zheng, and Z. Gu, “Long-range correlation analysis of high frequency prefrontal electroencephalogram oscillations for dynamic emotion recognition,” *Biomed Signal Process Control*, vol. 72, Feb. 2022, doi: 10.1016/j.bspc.2021.103291.

[134] “Empatica E4 wristband,” *empatica*. https://www.empatica.com/research/e4/ (accessed Apr. 08, 2023).

[135] H. P. Vinutha, B. Poornima, and B. M. Sagar, “Detection of outliers using interquartile range technique from intrusion dataset,” in *Advances in Intelligent Systems and Computing*, Springer Verlag, 2018, pp. 511–518. doi: 10.1007/978-981-10-7563-6_53.

[136] P. Jamal, M. Ali, R. H. Faraj, P. J. M. Ali, and R. H. Faraj, “1-6 Data Normalization and Standardization: A Technical Report,” 2014. [Online]. Available: https://docs.google.com/document/d/1x0A1nUz1WWtMCZb5oVzF0SVMY7a_58KQulqQVT8LaVA/edit#

[137] S. Ray, “A quick review of machine learning algorithms.,” in *International conference on machine learning, big data, cloud and parallel computing (COMITCon).IEEE*, 2019, pp. 35–39.

[138] X. Su, X. Yan, and C. L. Tsai, “Linear regression,” *Wiley Interdiscip Rev Comput Stat*, vol. 4, no. 3, pp. 275–294, May 2012, doi: 10.1002/wics.1198.

[139] G. E. Atteia, H. A. Mengash, and N. A. Samee, “Evaluation of using Parametric and Non-parametric Machine Learning Algorithms for Covid-19 Forecasting,” *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 10, pp. 647–657, 2021, doi: 10.14569/IJACSA.2021.0121071.

[140] F. Siddiqui and Q. M. Ali, “Performance of non-parametric classifiers on highly skewed data,” 2016. [Online]. Available: http://www.ripublication.com/gjpam.htm

[141] M. V. Shcherbakov, A. Brebels, N. L. Shcherbakova, A. P. Tyukov, T. A. Janovsky, and V. A. evich Kamaev, “A survey of forecast error measures,” *World Appl Sci J*, vol. 24, no. 24, pp. 171–176, 2013, doi: 10.5829/idosi.wasj.2013.24.itmies.80032.

[142] S. García, J. Luengo, J. A. Sáez, V. López, and F. Herrera, “A survey of discretization techniques: Taxonomy and empirical analysis in supervised learning,” *IEEE Trans Knowl Data Eng*, vol. 25, no. 4, pp. 734–750, 2013, doi: 10.1109/TKDE.2012.35.

[143] Y. Yang, G. I. Webb, and X. Wu, “Discretization Methods,” in *Data Mining and Knowledge Discovery Handbook*, Springer US, 2009, pp. 101–116. doi: 10.1007/978-0-387-09823-4_6.

[144] K. Nkurikiyeyezu, A. Yokokubo, and G. Lopez, “The Effect of Person-Specific Biometrics in Improving Generic Stress Predictive Models,” Oct. 2019, [Online]. Available: http://arxiv.org/abs/1910.01770

[145] K. Sharma, E. Niforatos, M. Giannakos, and V. Kostakos, “Assessing Cognitive Performance Using Physiological and Facial Features,” *Proc ACM Interact Mob Wearable Ubiquitous Technol*, vol. 4, no. 3, Sep. 2020, doi: 10.1145/3411811.