SUMMER
2024

JOURNAL OF MANAGEMENT & ENGINEERING INTEGRATION

AIEMS
VOL 17
NO 1

ISSN: 1939-7984

JOURNAL OF MANAGEMENT AND ENGINEERING INTEGRATION

Scope: The Journal of Management and Engineering Integration (JMEI) is a double-blind refereed journal dedicated to exploring the nexus of management and engineering issues of the day. JMEI publishes two issues per year, one in the Summer and another in Winter. The Journal's scope is to provide a forum where engineering and management professionals can share and exchange their ideas for the collaboration and integration of Management and Engineering research and publications. The journal will aim on targeting publications and research that emphasizes the integrative nature of business, management, computers and engineering within a global context.

EDITORIAL REVIEW BOARD

REVIEWERS

The Journal Editorial Team would like to thank the reviewers for their time and effort. The comments that we received were very constructive and detailed. They have been very helpful in our effort to continue to produce a top-quality journal. Your participation and timely response are very important for providing a distinguished outlet for original articles. In this issue, articles are assigned digital object identification (DOI) numbers to make sure readers and researchers can reliably find your work and to help track the ways your work is cited by others.

Edwin Sawan, Ph.D., P.E.
Editor-in-Chief

Parshuram Aarotale
Sue Abdinnour
Esther Omotola Adeyemi
Cintia Buffon
Deborah S. Carstens
Tristan Davison
Palmer Frye
Andrzej Gapinski
Abdelnasser Hussein
Sai Lakshmi Jashti
Robert Keyser
Shruti Kshirsagar
Roger Merriman
Alexandra Schönning
Scott Swain
John Wang
Gamal Weheba
Brooke Wheeler
Fujian Yan
Bayram Yildirim

TABLE OF CONTENTS

Esther Omotola Adeyemi and Sharon Bommer
EVALUATING GENDER-BASED EFFECTIVENESS WITHIN MULTI-ATTRIBUTE TASK BATTERY STUDIES 1

Abdelhakim A. Al Turk and Gamal S. Weheba
RESTORING HISTORICAL BUILDINGS USING ADDITIVE MANUFACTURING 10

John Wang, Jeffrey Hsu and Zhaoqiong Qin
EXPLORING NVIDIA'S EVOLUTION, INNOVATIONS, AND FUTURE STOCK TRENDS 21

Cintia Zuccon Buffon, Elizabeth A. Cudney and Ahmad Elshennawy
THE APPLICATION OF THE KANO MODEL IN EDUCATION – A SYSTEMATIC LITERATURE REVIEW 34

Bassam Jaradat and Gamal Weheba
A PRACTICAL GUIDE TO ISO 25022: MEASUREMENT OF SOFTWARE QUALITY IN USE 47

Bing Hu and Nicholas Mason
MATCHING WORKLOADS TO SYSTEMS WITH DEEP REINFORCEMENT LEARNING 55

https://doi.org/10.62704/10057/28080

EVALUATING GENDER-BASED EFFECTIVENESS WITHIN MULTI-ATTRIBUTE TASK BATTERY STUDIES

Esther Omotola Adeyemi[1]

Sharon Bommer[1]

[1]*University of Dayton*

Adeyemie1@udayton.edu; sbommer1@udayton.edu

Abstract

The prevalent notion that females outperform males in multitasking is challenged by scrutinizing the cognitive abilities of each gender across various demand scenarios (Buser & Peter, 2012; Lui et al., 2020; Szameitat et al., 2015). Investigating cognitive performance at two levels of workload defined as high- and low-demand situations, this research delves into task-specific variations, cognitive strategies, and adaptability to multitasking challenges. Sixty participants in total, with thirty participants in each gender group, engaged in tasks encompassing resource management, tracking, monitoring, and communication using the U.S. Army Aeromedical Research Laboratory's Multi-Attribute Task Battery simulation software. Statistical variance analysis demonstrated notable task discrepancies across different difficulty levels (DL). However, the gender-based analysis indicated non-significant differences in given tasks, except communication tasks. Findings indicate that men tend to perform well in tasks requiring communication and target tracking, benefiting from auditory cues, while women excel in fuel management, displaying strong decision-making abilities. In assessments of visual and spatial perception, females outperformed males at low difficulty levels, whereas males showed greater proficiency at high difficulty levels. This paper advocates for a balanced understanding and approach of cognitive abilities in diverse scenarios.

Keywords: *USAARL Multi-Attribute Task Battery; Cognitive Abilities; Gender Differences; Performance; Human Factor*

1. Introduction

Gender stereotyping has long influenced societal perceptions of individuals' capabilities and roles, shaping expectations, and behaviors in various domains. In the research by Bosak et al. (2018), females were characterized as intuitive, supportive, creative, and affectionate, whereas males were depicted as dominant, analytical, adept problem solvers, and competitive. In the context of multitasking, these stereotypes play a significant role in dictating gender roles and influencing how individuals navigate complex tasks. Based on many relevant references, it is evident that gender stereotypes in multitasking and performance have been the subject of extensive research. Traditional gender roles have often portrayed women as adept multitaskers, attributed to their assumed roles in managing household duties alongside professional responsibilities. Conversely, men have been stereotyped as more suited for singular, focused tasks, reflecting societal expectations of leadership and assertiveness. For instance, Lui et al. (2020) found that over 80% of stereotype holders believed that women performed better in multitasking situations than men, indicating the existence of gender stereotypes in multitasking. The influence of gender stereotypes on multitasking behaviors is profound. Women, burdened by the expectation of excelling in multitasking roles, may experience heightened stress and pressure to balance numerous responsibilities simultaneously. This pressure often stems from societal narratives that portray

Submitted: March 13, 2024
Revised: June 16, 2024

multitasking as a natural extension of femininity. Furthermore, men may face stigma or dismissal when engaging in multitasking activities, as it may be perceived as a deviation from traditional masculine roles. However, these traditional positions are non-absolute and only relative to mindset, upbringing, and natural talents. For example, Strobach and Woszidlo (2015) highlighted the uncertainty regarding whether the gender stereotype for multitasking aligns with actual multitasking performance differences. This uncertainty is further supported by Hirsch et al. (2019), who stated that empirical evidence for gender differences in multitasking performance is mixed. Kray et al. (2001) demonstrated that men outperformed women in negotiations when the negotiation was perceived as diagnostic of ability or linked to gender-specific traits, indicating the negative impact of stereotype confirmation on women's performance. Additionally, Kahalon et al. (2018) highlighted the detrimental consequences of activating negative stereotypes on performance, emphasizing the need to address gender stereotypes in various contexts.

2. Literature Review

The core of human factors lies in understanding and optimizing the interactions between humans and various elements within a system to enhance performance. This involves considering factors such as the individual's skills, the work environment, processes, and organizational culture on overall performance (Chrimes et al., 2018). Gender performances are shaped by people's experiences, interactions, and wider social expectations (Godec, 2020). Gender differences in performance are influenced by a multitude of factors, including socioeconomic background, gender equality, and cultural expectations. Similarly, researchers argued that gender differences exist at the level of cognitive functioning in the academic environment, indicating variations in academic performance between male and female students (Parajuli & Thapa, 2017). Another study examined gender differences in personality traits in relation to academic performance, finding no significant gender differences in personality characteristics and academic achievement (Khan, 2021).

The discourse surrounding the context of gender and multitasking has been perceived as mysterious event by many cognitive scientists over the past decades until the advent of the multiple resource theory (Wickens, 2002). Multiple Resource Theory offers a valuable framework for assessing and demystifying these popular beliefs. By examining how individuals allocate cognitive resources across various tasks, this theory provides insights into the complexities of multitasking and challenges simplistic gender- based assumptions. One of the prominent products of multiple resource theory is the innovation of a multi-attribute task battery. The Multi-Attribute Task Battery (MATB) is a widely used tool for assessing human performance in multitasking environments and was originally developed at NASA Langley Research Center in 1992 to provide a comprehensive behavioral metric for assessing complex operator performance (Mills & Gilliland, 1994). Since then, MATB has been applied in various fields, including pilot workload assessment (Adeyemi et al., 2023; Mortazavi et al., 2018), multitasking performance and workload evaluation (Adeyemi & Bommer, 2023; Kim et al., 2016), and mental workload detection (Chu et al., 2022).

Mental workload assessment is crucial given the intricate decisions pilots must make amidst complex interfaces. Yet, there is a gap in studying gender effects; thus, MATB is utilized to explore mental workload perception by gender.

3. Methodology

In this study, participants were given a consent form to review and sign, indicating their voluntary participation in the experiment. Participants received a brief introduction and tutorial covering software navigation and interaction to ensure they understood the software's basic functions.

Before commencing the actual experimental scenarios, participants had the opportunity to familiarize themselves with the simulation software through practice sessions. Among the female participants, 76.67% had prior experience with gaming software, while 96.67% of male participants reported the same.

3.1. Participants & Materials

The experimental procedure obtained approval from the Institutional Review Board (IRB) at The University of Dayton, Ohio. The participants consisted of University of Dayton students aged between 18 and 34, representing diverse ethnic backgrounds. The study involved 60 individuals in total, comprising 30 males and 30 females. Before advancing to the main experiment, all participants underwent a color blindness test to verify their ability to discern different color stimuli within the simulation. Additionally, a demographic survey was administered to collect information on participants' age, gender, ethnicity, and other pertinent personal details. The entire experimental procedure lasted approximately 20 minutes and the experimental scenarios were randomized. The subjective measures of mental workload (MWL) were also collected using the NASA task load index (NASA TLX) and Instantaneous Self-assessment (ISA) tool at the end of each experimental scenario. During the experiment, an Acer Aspire 5 Notebook Laptop equipped with specifications such as an Intel Core i7-1165G7 processor and 36GB of RAM is utilized to operate the USAARL MATB software. This configuration interfaces with a TECKNET Bluetooth wireless mouse and a Logitech G Extreme 3D Pro joystick for control and interaction.

3.2. The Experiment/Tasks

In the USAARL MATB simulation depicted in Figure 1, operators engage in four distinct tasks aimed at testing their cognitive and operational abilities. For the monitoring task [1], operators must swiftly react to the activation of light by pressing the corresponding button using the joystick to extinguish it before it automatically switches off. However, for the tracking task [2], operators are tasked with keeping a jittering circle centered within a designated square on the plot utilizing a joystick. Operators must maintain the circle's position for optimal scoring, which increases as the circle approaches the center.

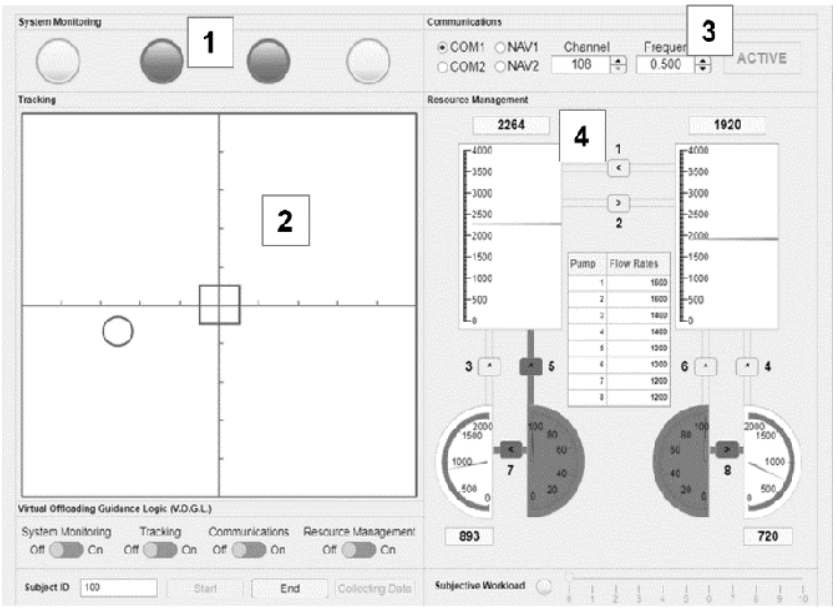

Figure 1. USAARL MATB user interface

In the communication task [3] operators receive auditory cues prompting adjustments to radio settings, including channel and frequency changes. Their attention is focused on responding specifically to the call name "NASA 504," disregarding any other call names and adjusting radios accordingly based on provided instructions. Finally, in the resource management task [4], operators must maintain fuel levels in two large tanks at a steady 2000 units, utilizing eight numbered fuel pumps to regulate flow rates between tanks. Throughout the simulation, operators are tested on their ability to react promptly, maintain focus amidst auditory distractions, and strategically manage resources for optimal performance across all tasks.

4. Experimental Results & Discussion

Figure 2 depicts an interaction plot illustrating the relationship between demand levels (low and high) and genders, indicating that males tend to excel in tasks related to communication, particularly those involving auditory cues and target tracking. On the other hand, females showcase strengths in fuel management strategy, underscoring their ability to make effective decisions. When it comes to assessing visual and spatial perception in monitoring tasks, females display superior performance at lower difficulty levels, while males show greater proficiency at higher difficulty levels. These results highlight gender-specific strengths in various cognitive domains, emphasizing the importance of considering diverse abilities in task design and evaluation.

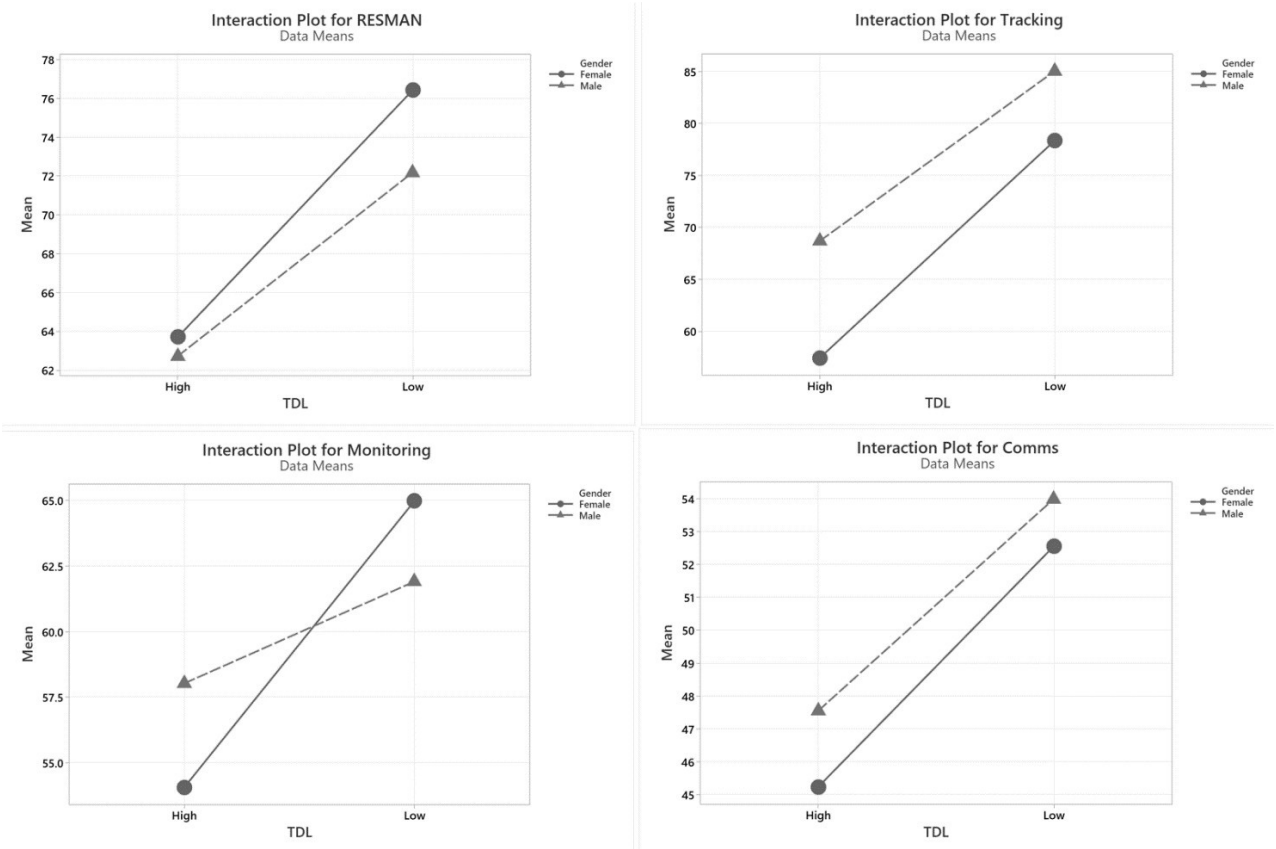

Figure 2. Rela:onship in performance between individual given tasks, gender & TDL

However, Table 1 reflects the overall statistical data comparing the performance of participants with respect to gender and task difficulty level. The statistical analysis of variance uncovered notable variations across all given tasks, particularly based on their difficulty levels (DL). Moreover, it was observed that the communication task exhibited a significant divergence based on gender,

contrasting with other tasks. These findings underline the relationship between task complexity and participant demographics. The analysis highlights the importance of considering both task difficulty and gender dynamics in understanding task performance variability.

Table 1. Sta:s:cal analysis of variance for gender & demand level

Parameters	Male (Low)	Female (Low)	Male (High)	Female (High)	*p* values	
	Mean (SD)	*Mean (SD)*	*Mean (SD)*	*Mean (SD)*	*TDL*	*GENDER*
RESMAN	72.17 (11.11)	76.42 (9.39)	62.71 (15.18)	63.71 (16.87)	0.000*	0.288
TRACKING	85.04 (4.389)	78.35 (1.48)	68.68 (6.40)	57.44 (11.31)	0.000*	0.000*
MONITORING	61.90 (9.59)	64.97 (7.77)	58.02 (9.76)	54.07 (11.08)	0.000*	0.804
COMMS	54.65 (18.43)	52.55 (17.32)	47.88 (15.78)	45.57 (13.47)	0.021*	0.433

** Indicates significant differences at α = 0.05*

Table 2 showcases participants' overall subjective assessment using the NASA Task Load Index (NASA TLX) measure and the Instantaneous Self-Assessment (ISA) tool integrated into the USAARL MATB. The NASA TLX interpretation classifies workload scores from 0-9 as low, 10-29 as medium, 30-49 as somewhat high, 50-79 as high, and 80-100 as very high (Prabaswari et al., 2019). Consequently, it is logical to infer that both genders perceived the workload of both low and high task difficulty levels as 'High' based on the unweighted/raw NASA TLX scores recorded. The USAARL ISA scale ranges from 1 to 10. For both genders, low difficulty levels were rated at approximately 6, while high difficulty levels received a rating of 7, indicating a 'High' perception of Mental Workload (MWL).

Table 2. Gender-based subjec:ve ratings

Parameters	Average ISA (scale 1-10)		NASA TLX (unweighted)	
	Male	*Female*	*Male*	*Female*
Low	5.9	5.7	67.08	66.25
High	7.0	6.6	60.00	55.83

Table 3. Sta:s:cal ANOVA for subjec:ve measures

Parameters	*p* values	
	Gender	*TDL*
NASA TLX	0.375	0.120
ISA	0.205	0.063

Interestingly, both subjective assessments align with each other but contradict the performance evaluation, which suggests significant differences between Task Difficulty Levels (TDL) as indicated in Table 1. A statistical analysis of variance was conducted to validate the interpretation provided, further supporting our findings as depicted in Table 3, revealing no statistically significant differences across gender or task difficulty levels for subjective measures.

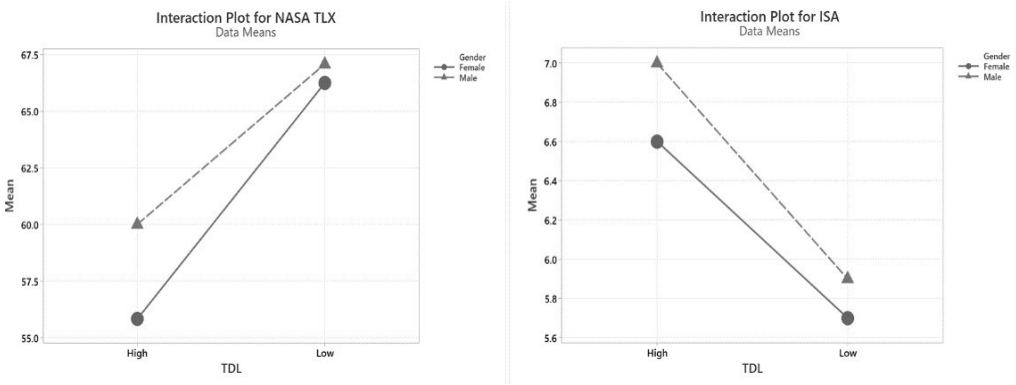

Figure 3. Rela:onship in subjec:ve measure, gender & TDL

Before initiating the experiment, a demographic survey was conducted to gather background information from each participant and to enhance the study's insights. The survey included questions about personal data and participants' experience with gaming software, categorized as "always/often" or "rarely/sometimes." Among the participants, 23 out of 30 females and 29 out of 30 males reported having experience with gaming software, representing 76.67% of females and 96.67% of males, as illustrated in Figure 4. Additionally, during the initial training phase aimed at familiarizing participants with the simulation, it was observed that males exhibited quicker learning abilities compared to their female counterparts. The data suggests a notable difference in the experience with gaming software between male and female participants. A significant majority of females, accounting for 87%, reported having experience "rarely/sometimes," whereas a smaller proportion, 13%, indicated "always/often." In contrast, among males, a substantial 66% reported "always/often," while 34% fell into the "rarely/sometimes" category. This discrepancy implies varying levels of familiarity and exposure to gaming software between the two genders.

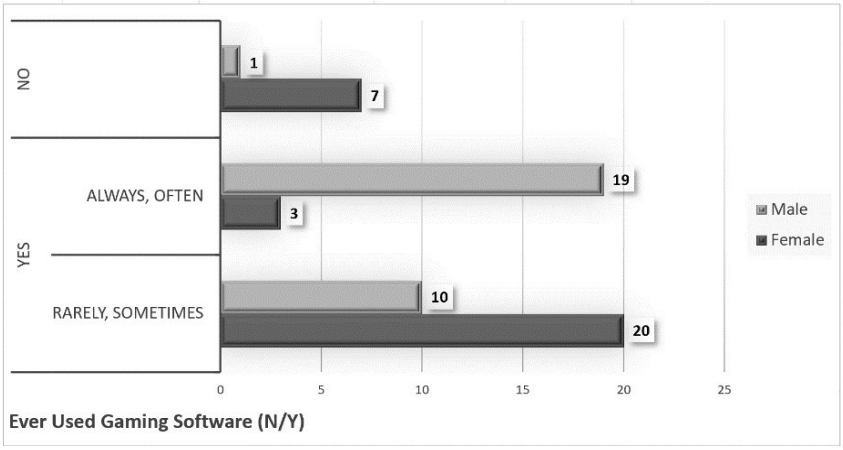

Figure 4. Par:cipant's experience with gaming software

While individual tasks allocated within the MATB simulation have been scrutinized as shown in Figure 2 and Table 1, Figure 5 delineates the overall performance correlation between gender and task difficulty level. The illustration indicates that males exhibited notably superior performance in high-difficulty tasks compared to females. Even in low-difficulty tasks, although the variance seems slight, males consistently outperformed females.

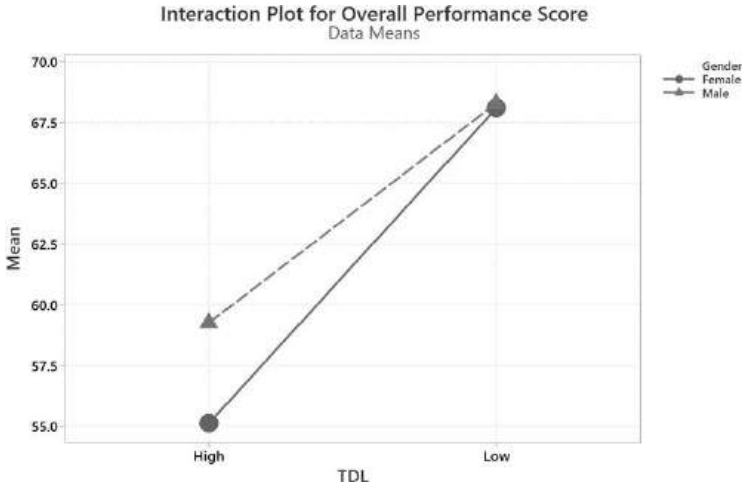

Figure 5. The rela:onship with overall performance score

5. Conclusion

This study has delved into the vital realm of mental workload assessment particularly focusing on gender differences using the Multitasking Assessment Tool (USAARL MATB). The results indicate that male participants outperform their female counterparts in communication and tracking tasks. However, females excel in fuel management strategy, showcasing their decision-making prowess. Visual and spatial perception tasks unveil gender-specific strengths, with females demonstrating superior performance at lower difficulty levels and males excelling at higher difficulty levels. The study's overall findings indicate that men demonstrate higher performance than women in MATB studies, possibly attributable to the male participants in this study having greater exposure to gaming activities.

Furthermore, the subjective assessment using the NASA TLX and Instantaneous Self-Assessment (ISA) tools reveals that both genders perceive high mental workload for both low and high task difficulty levels. This discrepancy between subjective perception and objective performance evaluation raises intriguing questions about the subjective experience of mental workload and its implications for task design. These findings advocate for a measurable approach to designing tasks and engaging participants, recognizing the diversity in cognitive strengths among individuals in the industry. The insights gained from this study can inform tailored approaches to task design and participant engagement, fostering a more inclusive and effective interface.

Future research should consider incorporating more female gamers or providing comprehensive training to all participants to ensure a more balanced representation across genders. The general conclusions of this study suggest that men tend to outperform women in MATB studies, potentially linked to their higher exposure to gaming activities. This study accentuates the importance of understanding how gender-specific strengths and perceptions of workload impact task performance within engineering teams for more effective engineering management. By recognizing and accommodating these differences in task design and team dynamics, engineering managers can enhance productivity, foster inclusivity, and optimize project outcomes.

6. Acknowledgments

I am profoundly thankful to my advisor, Dr. Sharon Bommer, for her outstanding mentorship and steadfast encouragement.

7. Disclaimer Statements

Funding: None

Conflict of Interest: None

8. References

Adeyemi, E. O., Bommer, S., & Appiah-Kubi, P. (2023). Integration of mental workload analysis in human systems design for engineering management: A literature review on the effect of colors on human performance. *IISE Annual Conference Proceedings*, 1-6.

Adeyemi, E., & Bommer, S. (2023, October). A pilot study on the impact of colors on human performance within a multitasking simulation environment. In *1st International Conference on Smart Mobility and Vehicle Electrification*. https://doi.org/10.46254/EV01.20230069

Bosak, J., Eagly, A., Diekman, A., & Sczesny, S. (2018). Women and men of the past, present, and future: Evidence of dynamic gender stereotypes in Ghana. *Journal of Cross-Cultural Psychology*, *49*(1), 115–

129. https://doi.org/10.1177/0022022117738750

Buser, T., & Peter, N. (2012). Multitasking. *Experimental Economics, 15*(4), 641–655. https://doi.org/10.1007/s10683-012-9318-8

Chrimes, N., Bradley, W. P., Gatward, J., & Weatherall, A. (2018). Human factors and the 'next generation' airway trolley. *Anaesthesia*. https://doi.org/10.1111/anae.14543

Chu, H., Cao, Y., Jiang, J., Yang, J., Huang, M., Li, Q., Jiang, C., & Jiao, X. (2022). Optimized electroencephalogram and functional near-infrared spectroscopy-based mental workload detection method for practical applications. *Biomedical Engineering Online*. https://doi.org/10.1186/s12938- 022-00980-1

Godec, S. (2020). Home, school and the museum: Shifting gender performances and engagement with science. *British Journal of Sociology of Education, 41*(2), 147–159. https://doi.org/10.1080/01425692.2019.1700778

Hirsch, P., Koch, I., & Karbach, J. (2019). Putting a stereotype to the test: The case of gender differences in multitasking costs in task-switching and dual-task situations. *PLoS ONE, 14*(8), 1–16. https://doi.org/10.1371/journal.pone.0220150

Kahalon, R., Shnabel, N., & Becker, J. C. (2018). Positive stereotypes, negative outcomes: Reminders of the positive components of complementary gender stereotypes impair performance in counter-stereotypical tasks. *British Journal of Social Psychology,* 57(3), 482-502. https://doi.org/10.1111/bjso.12240

Khan, D. (2021). Gender differences in personality traits in relation to academic performance. *Mier Journal of Educational Studies Trends & Practices*. https://doi.org/10.52634/mier/2020/v10/i1/1356

Kim, J. H., Yang, X., & Putri, M. (2016). Multitasking performance and workload during a continuous monitoring task. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. https://doi.org/10.1177/1541931213601153

Kray, L. J., Thompson, L., & Galinsky, A. D. (2001). Battle of the sexes: Gender stereotype confirmation and reactance in negotiations. *Journal of Personality and Social Psychology*. https://doi.org/10.1037/0022-3514.80.6.942

Lui, K. F. H., Yip, K. H. M., & Wong, A. C.-N. (2020). Gender differences in multitasking experience and performance. *Quarterly Journal of Experimental Psychology, 74*(2), 344–362. https://doi.org/10.1177/1747021820960707

Mills, S. H., & Gilliland, K. (1994). An improved version of the multi-attribute task battery (MATB) and data reduction program (MATPROC). *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. https://doi.org/10.1177/154193129403801506

Mortazavi, M., Raissi, K., & Mehne, H. H. (2018). Pilot workload assessment under different levels of autopilot failure. *Scientia Iranica*. https://doi.org/10.24200/sci.2018.50592.1777

Parajuli, M., & Thapa, A. (2017). Gender differences in the academic performance of students. *Journal of Development and Social Engineering, 3*(1), 39–47. https://doi.org/10.3126/jdse.v3i1.27958

Prabaswari, A. D., Basumerda, C., & Utomo, B. W. (2019). The mental workload analysis of staff in study program of private educational organization the mental workload analysis of staff in study program of private educational organization. *IOP Conference Series: Materials Science and Engineering*, *528*(1), 012018. https://doi.org/10.1088/1757-899X/528/1/012018

Strobach, T., & Woszidlo, A. (2015). Young and older adults' gender stereotype in multitasking. *Frontiers in Psychology*. https://doi.org/10.3389/fpsyg.2015.01922

Szameitat, A. J., Hamaida, Y., Tulley, R. S., Saylik, R., & Otermans, P. C. J. (2015). "Women are better than men"– Public beliefs on gender differences and other aspects in multitasking. *PLoS ONE*, *10*(10), e0140371. https://doi.org/10.1371/journal.pone.0140371

Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science*, *3*(2), 159–177. https://doi.org/10.1080/14639220210123806

https://doi.org/10.62704/10057/28081

RESTORING HISTORICAL BUILDINGS USING ADDITIVE MANUFACTURING

Abdelhakim A. Al Turk [1]
Gamal S. Weheba[1]
[1]Wichita State University
axalturk@shockers.wichita.edu

Abstract

This paper considers using additive manufacturing in restoration projects. This research proof that Polyblend Calcium Carbonate (Grout) powder can be used as a building material on the ZPrinter 450 Binder Jetting 3D printer. Such material has promising applications in restoring historical structures. Using statistically designed experiments, this research investigated the effect of the layer thicknesses, binder saturation, curing solution, and curing time on 3D-printed tiles. The breaking strength, deviation from the nominal dimension, and surface roughness were measured as the key quality characteristics of printed tiles. The results were confirmed using additional runs. Research findings could be instrumental in offering unprecedented capabilities to preserve and revitalize historical structures and architectural treasures. These innovative research efforts will be extremely instrumental in the field of 3D printed customized tiles, façades restoration, and decorative calligraphy tiles.

Keywords: *Additive Manufacturing; 3D Printing; Polyblend Grout; Design of Experiments; Restoration; Tiles*

1. Introduction

Additive Manufacturing (AM) represents a type of layered manufacturing, employing a range of machines tailored to replicate a physical structure from a 3-dimensional Computer-Aided Design (CAD) model swiftly and with high automation. According to the American Society for Testing and Materials, AM is defined as "the process of joining materials to make objects from 3D model data, usually layer upon layer." Although there exists variability in materials and methodologies, the fundamental process involves segmenting the CAD model into layers and constructing them as prototypes or final products. Initially, AM technology was primarily utilized for prototyping to aid new product design endeavors. In the construction industry, AM is used in a wide range of application areas. As mentioned by Al Turk & Weheba (2021), the application of AM in the construction industry is a fast-growing field and has a lot of great advantages over traditional systems. In the research, binder jetting technology was used to 3D print grout tiles. The research considered the effects of binder saturation level, layer thickness, post processing curing solution, and curing time on the quality of 3D printed samples. Numerical optimization design was used to quantify the effect of selected process factors.

2. Literature Review

Castilho et al. (2015) studied the effect of shell and core saturation on the properties of 3Dprinted parts. They tied the bulk density, the apparent porosity, the total porosity, and the compressive strength of 3D printed specimens. The results indicated that as the shell/core saturation increases, both the bulk density and the compressive strength increase, and the porosity decreases. At maximum saturation values, the strength can reach up to 20 MPa. Shakor et al. (2019), utilize

Submitted: April 8, 2024
Revised: June 16, 2024

geopolymer cement and studied the effect of post-processing techniques and binder levels on the compressive strength of 3D printed specimens. They reported that 3D printed samples with 170% binder saturation resulted in compressive strength of 14 MPa when cured for 28 days in water. Ingaglio et al. (2019), studied the effect of changing the binder saturation level on the compressive strength of the 3D printed samples. The results shows that specimens printed using calcium-sulpho-aluminate (CSA) powder and 150% binder saturation level resulted in 9 MPa compressive strength. Odaglia et al. (2020), studied the effect of changing the powder material on the compressive strength of 3D printed samples. As a result, the Gneiss sand samples resulted in compressive strength of 5.5 MPa, and the Brick Sand samples were 1.5 MPa.

3. Experimental Work

As indicated in the flow chart in Figure 1, the first step in this research is to identify the process factors and response variables related to 3D tiles printing. A design of experiment with the appropriate order and methodology will be selected to determine build process parameters that can lead to maximum strength and the best dimensional accuracy. The ZPrinter 450 (binder jetting 3D printer) was used to print the 2.5 x 2.5 x 0.4 inches (63.5 x 63.5 x 10.1 mm) tile samples. The original ZP151 powder will be replaced by polyblend calcium carbonate (Grout) powder. The polyblend powder is a mix of Calcium Carbonate, Portland Cement Titanium Dioxide, and Crystalline Silica. The same Zb 63 binder recommended by 3D Systems is used. Once the build process is completed, the printed tiles were divided into two groups. For the first group, the specimens were immersed in tap water at room temperature. For the second group, the printed specimens were immersed in an alkaline solution at room temperature. Finally, optimization techniques were used to identify the maximum breaking strength and the minimum of all other reaming responses. Confirmation experiments were performed to verify the results.

4. Experimental Factors

This research was focused on quantifying the effect of four factors. The layer thickness, binder saturation, curing solution, and curing time. These factors and their levels are listed in Table 1. The ZPrinter 450 allows users to select four levels of layer thickness, as mentioned in the user manual. These measurements are employed to regulate the platform's descent after the completion of each layer. Consequently, this parameter was considered to exert a significant influence on the strength of the printed specimens. The layer thickness was tested at two levels: 0.0034” (0.088mm) and 0.005” (0.125 mm). As stated by Xia et al. (2018), the binder saturation is defined as the ratio of the volume of jetted binders to the pore volume of the powder bed in each layer. The lowest and highest core binder saturation levels used are 100 and 272%, respectively. Curing is the process of wetting the samples, allowing them to gain strength during their construction period and to prohibit the chances of shrinkage. Two curing solutions were considered. One solution was the regular tap water and the other was an alkaline solution. The alkaline solution used is composed of N Grade sodium silicate solution with a SiO2/Na2O ratio of 3.22 (71.4% w/w) and 8.0 M NaOH solution (28.6% w/w). According to the ASTM C 192 standards, cementitious samples must be cured for 28 days to ensure the full development of breaking strength. According to ACI 318-19, the breaking strength of the cement-based sample increased with time, gaining 70% of its strength after 7 days, 90% after 14 days, and 99% after 28 days. For these reasons, two curing durations were considered: 14 and 28 days.

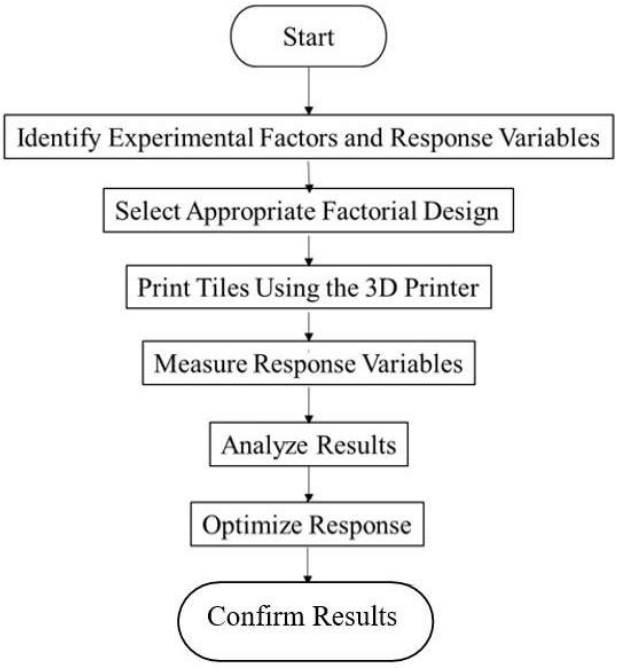

Figure 1. Research Methodology

Table 1. Selected Factors and Their Levels

Factor	Type	Low level (-)	High level (+)
A- Binder Saturation (%)	Numeric	100	272
B- Layer Thickness (mm)	Numeric	0.0034" (0.088mm)	0.005" (0.125 mm)
C- Curing time (Days)	Numeric	14	28
D- Curing Solution	Categoric	Tap Water	Alkaline Solution

5. Response Variables

The effects of the experimental factors were quantified by using four response variables. The deviation from the nominal of the printed specimens, the surface roughness, and the breaking strength. To evaluate the surface roughness (Ra), all samples were scanned using an optical 3D measurement system (Alicona G4 InfiniteFocus). The surface roughness data in the direction perpendicular to the deposition path were collected using a profile length of 0.1946" (4.945 mm) and a cutoff wavelength of 31.49" (800 mm). The deviation from nominal Z direction of the 3D printed samples were measured using a coordinate measuring machine (CMM) and recorded in terms of the deviation from the nominal from the CAD model. The MTS Criterion compressive strength machine was used to obtain the breaking strength.

6. Experimental Design

A factorial design was used to quantify the effect of selected build process parameters on the breaking strength, surface roughness, and dimensional accuracy of the constructed specimens. The approach for designing industrial experiments recommended by Coleman and Montgomery (1993) was followed. A second-order design was used as we anticipate that the effect is non-linear. In this

research, a replicated 2^4 factorial design replicated twice was used to quantify the effect of the four experimental factors. Four samples were printed in each run, resulting in printing a total of 128 specimens. The average of four samples were reported as response variable. The data was analyzed using the Design Expert V13 software. The design matrix shown in Table 2 indicates the run order for each of the factor level combinations and recorded values of the average three response variables.

Table 2. Design Matrix and Measured Responses

Run	Factor A: Binder Saturation	Factor B: Layer Thickness	Factor C: Curing Time	Factor D: Curing Saturation	Response 1: Breaking Strength (N)	Response 2: Deviation Z (Inch)	Response 3: Ra (µm)
1	272	0.088	28	Solution	2592.77	0.01	3.311
2	100	0.088	14	Water	792.18	0.022	3.80525
3	272	0.088	14	Water	846.66	0.003	2.79825
4	272	0.125	14	Solution	1278.66	0.01	3.5215
5	100	0.088	14	Water	818.98	0.032	3.549
6	272	0.125	28	Water	1148.86	0.02	3.152
7	272	0.088	14	Water	1050.49	0.006	3.798
8	100	0.125	28	Solution	1391.91	0.013	2.834
9	100	0.125	28	Water	789.54	0.028	3.532
10	100	0.125	14	Water	842.47	0.025	3.23125
11	272	0.088	14	Solution	1859.63	0.014	2.19175
12	272	0.125	14	Water	953.16	0.012	3.6165
13	100	0.088	28	Water	1017.01	0.022	4.593
14	100	0.125	28	Solution	2267.71	0.026	2.476
15	100	0.125	14	Water	629.29	0.007	3.2515
16	100	0.125	28	Water	791.3	0.02	4.445
17	100	0.088	28	Solution	1810.13	0.012	3.462
18	100	0.088	14	Solution	1593.51	0.021	2.97025
19	272	0.088	28	Water	1000.59	0.01	2.735
20	272	0.125	14	Solution	2007.42	0.033	3.77925
21	272	0.125	28	Solution	1859.94	0.006	2.455
22	272	0.125	14	Water	858.39	0.005	3.876
23	100	0.088	14	Solution	1247.86	0.027	2.846
24	100	0.125	14	Solution	1245.49	0.02	3.412
25	100	0.088	28	Water	1069.16	0.02	3.593
26	272	0.125	28	Solution	1905.05	0.011	3.104
27	272	0.088	14	Solution	1780.78	0.016	3.41325
28	272	0.088	28	Solution	2985.12	0.005	3.192
29	100	0.088	28	Solution	1748.01	0.01	2.765
30	272	0.125	28	Water	2391.84	0.01	3.679
31	272	0.088	28	Water	1903.65	0.018	4.634
32	100	0.125	14	Solution	1049.23	0.015	3.149

7. Results & Analysis

The following section covers the results and the analysis of the three main response variables, breaking strength, deviation from the nominal Z direction, and surface roughness (Ra). In addition to that, it shows the numerical optimization, and the performed confirmation runs.

7.1 Analysis of the Breaking Strength

Initial investigation of the ANOVA results indicates that the three factors of binder saturation (A), curing time (C), and curing solution (D) have significant effects on the response. However, the statistical analysis of the residuals suggested the need for power transformation with Y= -0.50 to

stabilize the error variance. The ANOVA results for the transformed values of breaking strength are presented in Table 3. The lack of fit is not significant, with a P-value of 0.66, indicating that all other interactions can be ignored. Overall, the model is significant with a P-value < 0.0001.

Table 3. ANOVA of the Breaking Strength

Source	Sum of Squares	df	Mean Square	F-value	p-value
Model	7.84E-04	3	2.61E-04	32.1	< 0.0001
A-Binder Saturation	1.45E-04	1	1.45E-04	17.8	2.3E-04
C-Curing time	1.51E-04	1	1.51E-04	18.5	1.87E-04
D-Curing Solution	4.88E-04	1	4.88E-04	60.0	< 0.0001
Residual	2.28E-04	28	8.141E-06		
Lack of Fit	8.41E-05	12	7.006E-06	0.779	0.6645
Pure Error	1.44E-04	16	8.993E-06		
Cor Total	1.01E-03	31			

Figure 2 depicts the effect of the binder saturation on the breaking strength. The maximum breaking strength was obtained when the samples were printed using the maximum binder saturation level. Using the high level of binder saturation (272%) resulted in a 31% increase in the strength of the printed specimens.

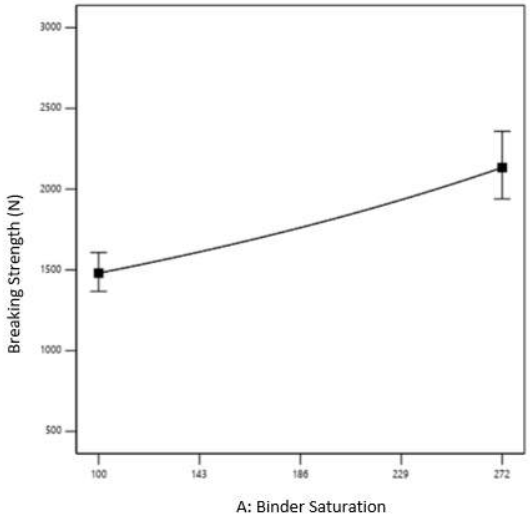

Figure 2. Effect of Binder Saturation (A)

An examination of the effect of the curing time on the breaking strength in Figure 3, shows that the higher the curing time, the higher the breaking strength. Using the high level of curing time (28 days) resulted in a 31.8% increase in the strength of the printed specimens. The one factor plot in Figure 4 suggests that the maximum breaking strength was obtained when the samples were cured in alkaline solution. Using the alkaline solution for curing resulted in an 77% increase in the strength of the printed specimens.

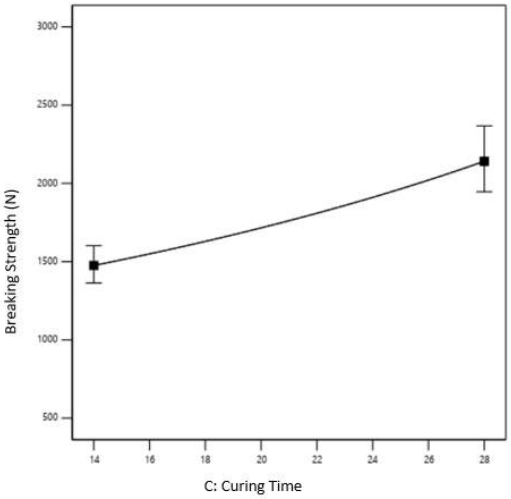

Figure 3. Effect of the Curing Time (C)

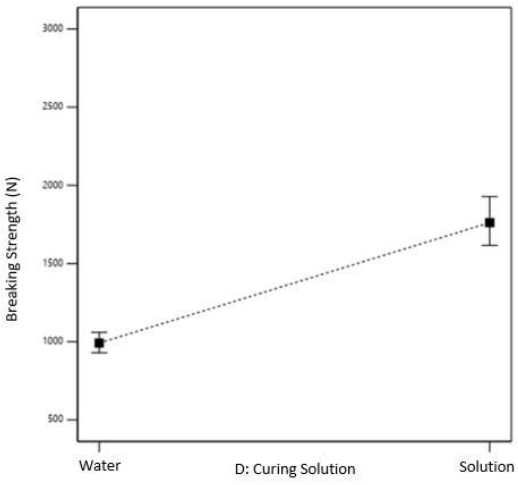

Figure 4. Effect of the Curing Solution (D)

7.2 Analysis of the Deviation from the Nominal Z Direction

Similar procedure was followed in analyzing changes in the specimens' height in terms of deviations from the nominal dimension. The statistical analysis of the suggested that a root transformation with Y= 0.55 is needed to stabilize the error variance. The ANOVA of the transformed deviation measurements are shown in Table 4. As shown, binder saturation (A), and CD combination between curing time (C), and curing solution (D) have significant effects on the response. The lack of fit is not significant, with a P-value of 0.575 indicating that all other interactions can be ignored. Overall, the model is significant with a P-value < 0.005. All diagnostics plots of the residuals supported the underlying assumptions of the ANOVA.

Table 4. ANOVA of the Deviation from Nominal

Source	Sum of Squares	df	Mean Square	F-value	P-value
Model	4.22E-04	4	1.05e-4	4.81	4.66E-03
A-Binder Saturation	2.60E-04	1	2.60E-04	11.9	1.89E-03
C-Curing time	1.08E-05	1	1.08E-05	0.492	0.489
D-Curing Solution	1.79E-06	1	1.79E-06	8.18E-02	0.777
CD	1.49E-04	1	1.49E-04	6.80	0.0147
Residual	5.92E-04	27	2.19E-05		
Lack of Fit	2.23E-04	11	2.03E-05	0.881	0.575
Pure Error	3.69E-04	16	2.31E-05		
Cor Total	1.01E-03	31			

The one factor plot in Figure 5 depicts the effect of the binder saturation on the average Z deviation from the nominal. As shown, the minimum Deviation in Z direction was obtained when the maximum binder saturation was used. Using the high level of binder saturation (272%) resulted in a 40.25% decrease in the average height deviation of the printed specimens. An examination of the CD interaction plot shown in Figure 6 reveals that changes in the curing solution have a significant effect on the average deviation in the Z direction when the samples are cured for 28 days.

Curing the samples in an alkaline solution for 28 days decreased the average deviation. The maximum deviation of 0.0195" (0.5 mm) is observed when specimens were cured in an alkaline solution for 14 days. The minimum deviation of 0.01165" (0.3 mm) was observed when specimens were cured in an alkaline solution for 28 days. While curing the samples in water for 28 days, the deviation height increased.

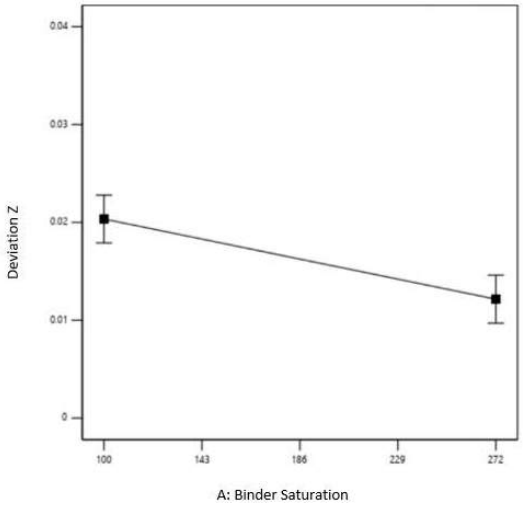

Figure 5. Effect of the Binder Saturation (A)

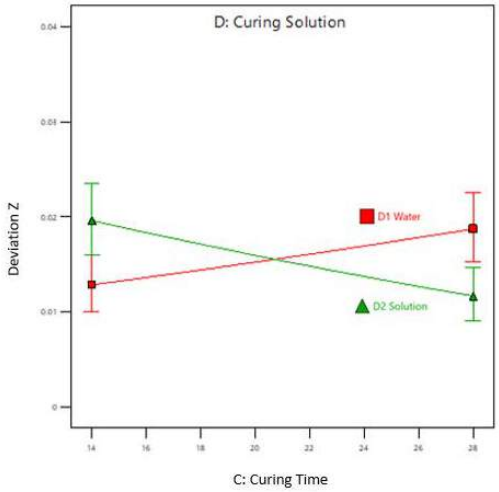

Figure 6. CD Interaction Plot

7.3 Analysis of the Surface Roughness (Ra)

A similar procedure was followed in analyzing changes in the Surface roughness (Ra). The ANOVA results are shown in Table 7. The ANOVA indicates that the curing solution (D) has significant effects on the response. The lack of fit is not significant, with a P-value of 0.7563, indicating that all other interactions can be ignored. Overall, the model is significant with a P-value < 0.005.

Table 7. ANOVA of the Surface Roughness

Source	Sum of Squares	df	Mean Square	F-value	P-value
D-Curing Solution	2.77	1	2.77	10.85	0.0025
Residual	7.64	30	0.2548		
Lack of Fit	2.87	14	0.2051	0.6879	0.7563
Pure Error	4.77	16	0.2982		
Cor Total	10.41	31			

The one factor plot in Figure 7 depicts the effect of the curing solution on the average surface roughness (Ra) from the nominal. As shown, the lower surface roughness (Ra) 3.3 µm was obtained when the maximum binder saturation was used.

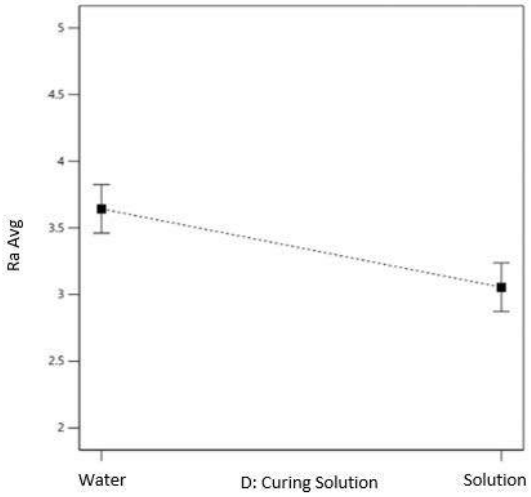

Figure 7. Effect of the Curing Solution (D)

7.4 Numerical Optimization and Confirmation Runs

Numerical optimization was utilized to determine settings that would allow for maximum breaking strength, minimize deviation Z, and minimize surface roughness Ra. This setting is to use layer thickness of 0.0034" (0.088mm), high level binder saturation level 272%, cure samples in alkaline solution for 28 days.

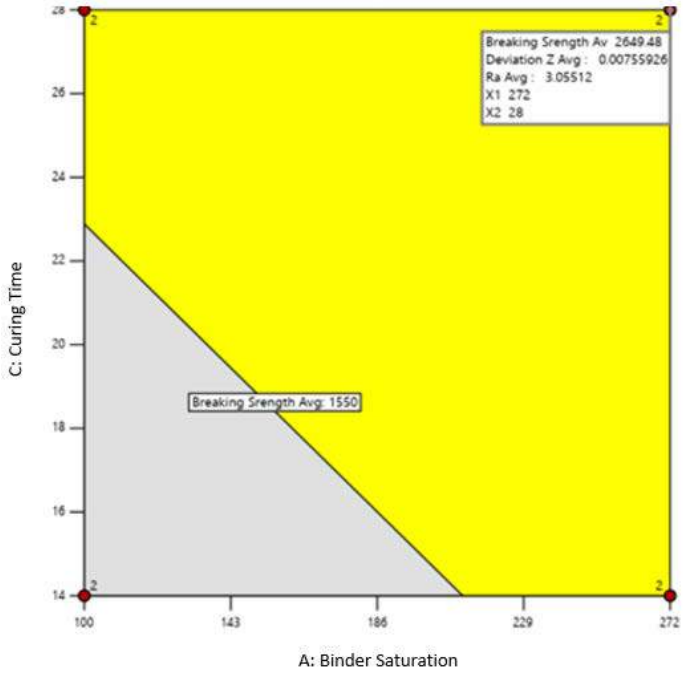

Figure 8. Overlay Plot With Recommended Settings using the Alkaline Solution

The overlay plot shown in Figure 8 was constructed by limiting the deviation from the nominal to 0.007" (0.1778 mm), limiting the surface roughness at 3 µm and setting the minimum breaking strength at 1550 N. The plot highlights the recommended area for selecting levels of layer thickness, binder saturation level, and curing time when the alkaline solution is used.

A predictive model may only be safely used in decision-making when confirmed. Confirmation is a critical step that shows the applicable range of the model and its performance limits, as mentioned by Montgomery (2012). Confirmation runs were performed by running five additional runs at a high level of binder saturation (272%) and at the lowest layer thickness of 0.0034" (0.088mm). Then the samples were cured in an alkaline solution for 28 days. The observed averages as shown in Table 7, are well within the 95% confidence interval (CI) of their expected values. This provides confirmation of the prediction capability of the model developed in this research.

Table 7. Results of Confirmation Runs

Analysis	Predicted Avg	Observed Avg	95% CI low	95% CI high
Breaking Strength	2649.48	2497.12	2014.37	3594.82
Deviation from Nominal	0.00755926	0.015721	-0.0006033	0.015743
Ra	3.05512	3.2326	2.52697	3.58328

8. Conclusions

This research was aimed at providing proof of the concept that polyblend calcium carbonate (grout) can be used as a material on a binder jetting 3D printer (ZPrinter 450). The statistical analysis of the results indicated that the maximum breaking strength that tile specimens can reach is 2.98 KN. This was obtained when the specimens were printed using a low layer thickness of 0.0035" (0.088 mm), a high binder saturation level of %272, and cured in the alkaline solution for 28 days. The minimum deviation in the Z direction of 0.01165" (0.3 mm) was observed when specimens were cured in an alkaline solution for 28 days. While curing the samples in water for 28 days, increase the deviation height. Moreover, a surface roughness (Ra) of 3.3 µm was obtained when the maximum binder saturation was used.

One of the applications of this research is to 3D print mosaics and decorative tiles (PEI 0). These 3D-printed tile products can be used as ornamentation on walls in hotels, historical structures, churches, and mosques. As listed in the current revision of the American National Standard Specifications for Ceramic Tile ANSI (A137.1:2022), the breaking strength of mosaic tiles should be greater than 1.11 KN, which was exceeded in the presented research. The research findings may be used to 3D print customized tiles with letter scripts, like the house number plaque or the decorative calligraphy tiles found in churches and mosques. This may eliminate the need for carving, as objects can be designed using CAD software and printed directly.

Using this research, architectural façadism can be applied. The façade of a building, especially the principal front that looks onto a street or open space, is architectural façadism, which is the practice of preserving a structure's façade, or face, while constructing a new building behind it. Restoring or recreating a historical façade is a perfect application for this research. By utilizing digital 3D scanners and computer-aided designs, the method developed in this research allows for the replication of precise replicas of delicate components by seamlessly integrating them into existing structures. These innovative research efforts will be extremely instrumental in the field of restoration by offering unprecedented capabilities to preserve and revitalize historical structures and architectural treasures.

9. Disclaimer Statements

Funding: None

Conflict of Interest: None

10. References

ACI CODE-318-19(22): Building Code Requirements for Structural Concrete and Commentary (Reapproved 2022). https://doi.org/10.14359/51716937

ANSI A137.1:2022 – Standard Specifications for Ceramic Tile. https://blog.ansi.org/ansi-a137-1-2022-standard-ceramic-tile/

Al Turk, A. A., & Weheba, G. S. (2021). 3D printing of Portland cement using a binder jetting system. *Journal of Management & Engineering Integration, 14*(1), 128-136. https://doi.org/10.62704/10057/24775

ASTM C192 / C192M-19. Standard practice for making and curing concrete test specimens in the laboratory. ASTM International, West Conshohocken, PA, 2019. Retrieved from www.astm.org

Castilho, M., Gouveia, B., Pires, I., Rodrigues, J., & Pereira, M. (2015). The role of shell/core saturation level on the accuracy and mechanical characteristics of porous calcium phosphate models produced by 3D printing. *Rapid Prototyping Journal, 21*(1), 43-51. https://doi.org/10.1108/RPJ-02-2013-0015

Coleman, D., & Montgomery, D. (1993). A systematic approach to planning for a designed industrial experiment. *Technometrics, 35*(1), 1-12. https://doi.org/10.1080/00401706.1993.10484984

Montgomery, D. C. (2012). *Design and analysis of experiments* (8th ed.). John Wiley & Sons.

Ingaglio, J., Fox, J., Naito, C. J., & Bocchini, P. (2019). Material characteristics of binder jet 3D printed hydrated CSA cement with the addition of fine aggregates. *Construction and Building Materials, 206*, 494-503. https://doi.org/10.1016/j.conbuildmat.2019.02.065

Odaglia, P., Voney, V., Dillenburger, B., & Habert, G. (2020, July). Advances in binder-jet 3D printing of non-cementitious materials. In *Second RILEM International Conference on Concrete and Digital Fabrication* (pp. 103-112). https://doi.org/10.1007/978-3-030-49916-7_11

Shakor, P., Nejadi, S., Paul, G., Sanjayan, J., & Nazari, A. (2019). Mechanical properties of cement-based materials and effect of elevated temperature on three-dimensional (3-D) printed mortar specimens in inkjet 3-D printing. *ACI Materials Journal, 116*(2), 55-67. https://doi.org/10.14359/51714452

https://doi.org/10.62704/10057/28082

EXPLORING NVIDIA'S EVOLUTION, INNOVATIONS, AND FUTURE STOCK TRENDS

John Wang[1]
Jeffrey Hsu[2]
Zhaoqiong Qin[3]
[1]Montclair State University
[2]Fairleigh Dickinson University
[3]South Carolina State University
wangj@montclair.edu; *jeff@fdu.edu*; *zqin@scsu.edu*

Abstract

This paper undertakes a thorough examination of Nvidia's stock market performance, intertwining historical analysis with forward-looking projections to illuminate the dynamic trajectory of this semiconductor industry giant. Commencing with a retrospective review, the authors delve into pivotal milestones, technological innovations, and strategic maneuvers that have shaped Nvidia's stock evolution. Utilizing advanced machine learning algorithms, including Random Forest and Support Vector Regression (SVR), alongside traditional statistical forecasting methods, we forecast future patterns. Through the incorporation of emerging industry dynamics, technological advancements, and market forecasts, our goal is to furnish stakeholders with insights pivotal for strategic decision-making. This dual perspective, encompassing both historical retrospection and future outlook, weaves a holistic narrative capturing the essence of Nvidia's stock market journey. It serves as a valuable resource for avid readers navigating the ever-changing landscape of the semiconductor market, especially considering the expanding role of Nvidia's GPUs in Artificial Intelligence (AI) and deep learning applications.

Keywords: *Innovations, AI; Machine Learning; Random Forest; Support Vector Machine; Deep Learning Applications*

1. Introduction

Nvidia, a behemoth in the technology industry, has earned its reputation as a leading competitor through a combination of organizational strength and technological innovation. Established in 1993 by Curtis Priem, Christopher Malachowsky, and Jen-Hsun Huang, the company has evolved into a powerhouse specializing in AI technology and computer science data. As its website indicates, brilliant minds make breakthrough discoveries. Nvidia pioneered accelerated computing to tackle challenges no one else can solve. Their work in AI and digital twins is transforming the world's largest industries and profoundly impacting society. AI is powering change in every industry, from generative AI and speech recognition to medical imaging and improved supply chain management, providing enterprises with the computing power, tools, and algorithms their teams need to do their life's work (NVIDIA, 2024).

Nvidia's journey is marked by transformative contributions to the gaming industry, with its graphics technology becoming a cornerstone of success. Widely adopted by companies for video game graphics, the impact of Nvidia's AI technologies is evident in their widespread use across over 40,000 companies and 15,000 startups, as of a 2020 announcement. While precise current numbers are higher, official updates have not been made public. The company's innovative timeline highlights its evolution from 3D graphics for games in 1993 to the development of GPUs and AI, culminating in the 2023 introduction of the omniverse, a boon for the metaverse. Strategic acquisitions have further bolstered Nvidia's standing in crucial markets.

Submitted: March 31, 2024
Revised: June 15, 2024

Nvidia's revenue growth narrative is a testament to its agility and ability to capitalize on evolving trends (See Fig. 1).

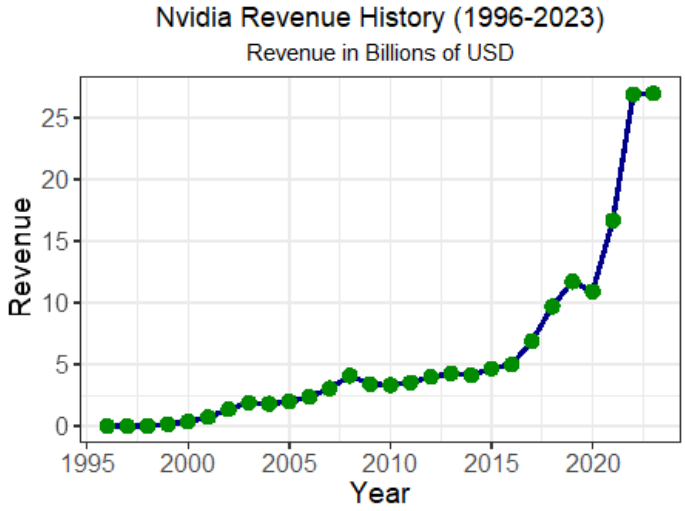

Figure 1. Nvidia Revenue History (1996-2023)

Operating across four categories – architecture, enterprise, gaming, and industry technology – Nvidia's product portfolio is extensive. Their architecture products focus on speeding up AI training, while the enterprise category offers software technologies for creating graphics and images. In gaming, products ensure smooth performance and quality graphics, while industry technology products address problem-solving, innovative technology perspectives, and AI-driven data collection. Nvidia's commitment to a healthy work environment and organizational culture, as articulated on its website, underscores its dedication to diversity, employee well-being, and compassionate management. This culture cultivates a space for continuous learning and growth, fostering a workforce capable of tackling the world's greatest challenges.

Nvidia harnesses cutting-edge technologies, including GPUs, CUDA, RTX, and AI, in their computing software. GPUs enhance graphics performance, positively impacting the entertainment and manufacturing industries. NVSwitch and NVLink improve GPU communication. CUDA accelerates computing processes, reducing programming time. RTX technology, designed for better 3D visuals, speeds up simulations. Nvidia's AI innovation tackles complex computing problems, enhancing graphics realism in gaming and aiding medical research. Their AI platform, comprising supercomputers, platform software, and models/services, powers solutions from disease prevention to the data analytics revolution. Nvidia's management emphasizes continuous testing, customer support, and data forecasting. With resolute teams focusing on software engineering, data research, robotics, and gaming systems, each member contributes to advancing technology in simulations and the metaverse. Nvidia's foremost achievement lies in AI, reshaping industries with high-performing software solutions (Shao et al., 2024).

Nvidia's most recent earnings report, released on May 8th, 2024, showed impressive year-over-year growth of 262%. Since its initial public offering (IPO) in January 1999, Nvidia's stock price has surged an impressive 38,544%. The company has emerged as a dominant force in both blockchain development and the AI revolution, leveraging its expertise in GPU design to create specialized chips optimized for cryptocurrency mining, machine learning, and artificial intelligence applications.

To navigate the competitive landscape and ensure continued growth, Nvidia has unveiled a series of strategic plans for the future. These include developing new AI chip platforms like Blackwell and Rubin, expanding into new markets, and creating new software offerings to support its AI ecosystem.

Nvidia has become synonymous with the artificial intelligence revolution. AI is rapidly penetrating most

industries, from healthcare and finance to manufacturing and retail, creating multiple multi-billion-dollar vertical markets. As businesses seek to leverage AI for automation, optimization, and data analysis, the demand for Nvidia's AI chips continues to grow rapidly.

Table 1. Nvidia's Revenue Growth from IPO

Period	Revenue	Absolute Growth from IPO	Percentage Growth from IPO
Fiscal Year 2024	$60.90 billion	$60.74 billion	38544%
Fiscal Year 2020	$10.92 billion	$10.762 billion	6911%
Fiscal Year 2010	$3.33 billion	$3.172 billion	2108%
IPO Year 1999	$0.158 billion	0	0

To stay ahead in the competitive AI landscape, Nvidia has unveiled a multi-pronged strategy. This includes developing next-generation AI chip platforms like Blackwell and Rubin, venturing into new markets, and bolstering its AI ecosystem with innovative software offerings.

Nvidia's role in the AI revolution is undeniable. As AI rapidly infiltrates diverse industries – from healthcare and finance to manufacturing and retail – it creates a landscape brimming with multi-billion-dollar opportunities. Businesses are increasingly turning to AI for automation, optimization, and data analysis, further fueling the surging demand for Nvidia's industry-leading AI chips.

2. Literature Review

Nvidia's state-of-the-art hardware and software consistently serve as the cornerstone for pioneering breakthroughs in AI research. Their unwavering commitment to delivering robust and efficient tools is instrumental in pushing the boundaries of possibilities across various domains, exemplifying a dedication to excellence in the field.

In line with Negi et al. (2023), Nvidia plays a crucial role in the development and implementation of Industry 4.0 technologies and tools, such as IoT, Cloud Computing, Artificial Intelligence, Edge Computing, Big data, Blockchain, Augmented Reality (AR), Virtual Reality (VR), Digital Twin, Metaverse, and Robotics, acting as a key enabler for various tools like, Nvidia's Jetson edge AI platform brings AI processing to the edge of the network, closer to sensors and actuators, enabling faster and more efficient responses to real-time data.

As Silva et al. (2024) indicated high-performance computing needs powerful resources to solve crucial problems. Growing demands, exascale advancements, and energy costs push for green HPC. Supercomputers and industry giants like Nvidia go beyond just saving energy, improving entire infrastructure, and collaborating with the energy sector for sustainability and cost benefits. Users must also participate for success, navigating the challenges and opportunities in this performance-focused field.

The advancement of autonomous driving (AD) technology, crucial for the next generation, hinges on the integration of intelligent perception, prediction, planning, and control. Despite progress, a performance bottleneck persists in AD algorithms. Addressing this challenge requires a focus on data-centric solutions. Li et al. (2024) reviewed state-of-the-art data-driven AD technologies, emphasizing a taxonomy of AD datasets, and examined benchmark closed-loop AD big data pipelines, leveraging Nvidia technologies. Future directions, applications, limitations, and concerns are explored to encourage collaborative efforts in AD advancement.

Ericson et al. (2024) highlighted the need for accurate risk forecasting in finance, particularly for VaR models. Building on HS, this study explores deep generative methods like Nvidia's NVDA for conditional time series generation. A Key Performance Indicators framework assesses distribution, autocorrelation, and backtesting. Models including HS, GARCH, and CWGAN are assessed on real and simulated data, displaying NVDA's role. Future research directions are discussed.

Six decades of evolution have seen computing systems revolutionize society through the internet and accessible tech. Adapting to evolving needs, systems like cloud, fog, edge, and the Internet of Things (IoT) emerge, driving economic opportunities. Gill et al. (2024) analyzed factors shaping this journey, examining architectures alongside nascent technologies like serverless, quantum, and on-device AI. Highlighting trends like rapid framework shifts, specialization, and decentralization, the review anticipates future challenges and directions. Nvidia's Jetson Series, with its efficient AI performance (32 Tera Operations Per Second), empowers robotics development.

In their 2023 work, Kurth et al. utilized Nvidia DALI's external source feature and integrated the Nvidia CuPy package in Python for direct host buffer pinning. Introducing SweepNet, a novel CNN architecture for precise identification of positive selection in population genomics, the authors achieved higher training efficiency and reduced epochs for superior validation accuracy through extensive hyperparameter exploration. Outperforming existing CNNs amidst confounding factors, SweepNet's constant trainable parameters ensure efficient training on large genomic datasets, establishing it as a robust solution.

Natural language processing (NLP) is an area of research and study that makes it possible for computers to comprehend human language by utilizing software engineering concepts from computer science and artificial intelligence. Dikshit et al. (2024) presented seven classifications and twenty-one state-of-the-art models with a survey of the operating principles. The authors also provided a comparative study of all models based on metrics such as accuracy, F1 score, exact match, squad scores, glue, and superglue dataset scores. Statistical learning models have transformed stock exchange practices, with quantitative analysts utilizing them for enhanced profit and risk predictions.

Chlebus et al. (2021) focused on creating an effective machine learning model for daily stock returns, using Nvidia Corporation's data from 07/2012 to 12/2018. Applying various models, including Nvidia as a key feature, the authors found superior predictive performance compared to naive and econometric models. This research contributes to the literature on model selection and emphasizes the importance of stationary exogenous variables in time series models.

The increasing demand for more accurate climate simulations in shorter durations prompts the exploration of Hardware Acceleration using GPUs or FPGAs. Vanderbauwhede & Takemi (2013) concentrated on the Weather Research and Forecasting Model (WRF) to assess the feasibility and advantages of GPU acceleration for Numerical Weather Prediction code. Profiling the WRF, the authors achieved up to a 7× speed improvement with a data-parallel kernel version on the GPU, demonstrating the feasibility and worthiness of GPU acceleration. The Nvidia GeForce GX480, with 15 CUs and 32 PEs each, was employed in our study. Uruguay is incorporating wind energy into its power generation, driving the development of a computational tool predicting electrical power injection.

Using the Weather Research and Forecasting (WRF) model, Silva et al. (2014) advanced GPU porting efforts, implementing sintb and bdy_interp1 routines on an Nvidia GeForce GTX 480 GPU. Achieving up to 33.9X speedup compared to sequential WRF and 9.2X compared to four-threaded WRF, the integrated routines significantly reduce total runtime, enhancing efficiency for both multi-core and single-threaded WRF versions.

The GPU/VGA, a vital computer component for video rendering and code compilation, has experienced price fluctuations since the COVID-19 pandemic, driven by heightened demand for remote work and online

activities. Hadi et al. (2023) employed deep learning-based time series forecasting, specifically the Transformer model, to predict future GPU prices, using the Nvidia RTX 3090 Founder Edition. Comparative analysis with RNN and LSTM models indicates the Transformer's superior effectiveness in forecasting GPU prices over an extended 30-day period, with higher correlation and lower error values.

According to Hu et al. (2023), ensemble forecasting with SwinVRNN is remarkably fast, taking only 44 seconds for a 100-member forecast using a single Nvidia Tesla V100 GPU. Classified as a learned distribution perturbation method, it samples perturbation noises from a learned distribution. Unlike traditional numerical weather prediction models, recent data-driven approaches like SwinVRNN offer promising speed for medium-range ensemble forecasting with impressive accuracy. This model, incorporating a Swin Transformer-based Variational Recurrent Neural Network, outperforms existing methods, particularly surpassing the operational ECMWF Integrated Forecasting System (IFS) on crucial atmospheric variables.

Campoy et al. (2023) introduced a CAIDE extension for distributed training and inference phases, optimizing performance with Nvidia capabilities. The framework and DL-based forecasting model are vital for solar energy prediction, now enhanced with Nvidia technology. Forecasting Global Horizontal Irradiance (GHI) for four horizons for up to one hour, the model traditionally relies on centralized predictive methods with inputs from deployed sensors. CAIDE, following Model Based System Engineering (MBSE) and Internet of Things (IoT) methodologies, supports solar plant deployment and analysis.

Zhang et al. (2023) analyzed Nvidia stockholders' discrete actions (A) and states (F) using transition probability distribution (P) and reaction function (R). Modular Neural Network (MNN) facilitated efficiency by decomposing complex problems into interconnected modules. Renowned for high-performance GPUs, Nvidia excels in gaming, AI, data centers, and autonomous vehicles. The company's expansion into AI and machine learning, with SoC solutions and a comprehensive software stack, reflects its commitment to innovation and sustainability in the computing industry.

Nvidia's GPUs empower deep learning models tackling the critical issue of human error (94%) in road crashes. This demands alternative solutions. Enter Autonomous Vehicles (AVs), replacing human drivers with advanced AI. Effective AVs predict nearby traffic behavior, mimicking human intuition. Bharilya & Kumar (2024) delve into AV trajectory prediction methods, focusing on machine learning (deep learning & reinforcement learning). Analyzing 200+ studies, it covers conventional methods, deep learning, and reinforcement learning approaches, highlighting strengths, weaknesses, datasets, and evaluation metrics. It identifies challenges, outlines research directions, and significantly advances AV trajectory prediction knowledge, shaping future developments.

Examining Nvidia's Decision Process using a Modular Neural Network (DNN Layer), Zhang et al. (2023) converged on discrete actions (A) by Nvidia stockholders and states (F). They utilized transition probability distribution (P) and reaction function (R). A Modular Neural Network (MNN) decomposes problems into interconnected modules, enhancing efficiency. Nvidia excels in GPU design for gaming, AI, data centers, and autonomous vehicles. Renowned for high-performance GPUs, Nvidia expanded into AI and machine learning, offering SoC solutions and a comprehensive software stack, emphasizing sustainability and driving innovation in gaming, data centers, and AI, shaping the computing industry's future.

Wang et al. (2024) provided a comprehensive analysis of Nvidia's technological innovations, market strategies, and future respects. Their study emphasizes the dynamics of GPU innovation, its disruptive influence on the market, and the robust innovation engine ingrained within Nvidia's culture of calculated risk-taking. The investigation thoroughly examines both internal and external factors contributing to Nvidia's remarkable success and its consequential industry dominance.

3. Methodology

Nvidia's stock is shaped by too many factors. Measuring these variables can be contentious, or even challenging to obtain. The authors have applied a few machine learning algorithms.

3.1. Random Forest

As per the findings by Ruseno et al. (2023), the Random Forest algorithm emerges as a robust ensemble technique comprised of multiple decision trees. While its tree adheres to a standardized structure, each training occurs on distinct subsets of the data, resulting in unique leaf nodes and predictions. The amalgamation of these individual tree predictions empowers the Random Forest to attain a more resilient and accurate overall outcome, embodying a "wisdom of the crowds" approach that mitigates the limitations of individual trees and minimizes the risk of overfitting.

Functioning as a supervised learning model, Random Forest utilizes labeled data to systematically "learn" the mapping of input features to desired outputs. This sets it apart from unsupervised learning algorithms such as K-means clustering, which endeavors to identify inherent patterns in unlabeled data devoid of predefined classes. The inherent flexibility of Random Forest endows it with versatility, enabling adept handling of both regression (predicting continuous values) and classification (predicting discrete categories) problems. Such versatility has positioned it as a preferred choice among engineers across diverse domains.

The Random Forest algorithm operates through three pivotal phases: input, training, and prediction.

A. Input: Sample training set (x): This dataset comprises Nvidia revenue parameters used for training. Response (y): This denotes the target parameter that the algorithm aims to predict.
B. Training: The training process involves a meticulous iterative construction of an ensemble of decision trees, comprising the following three distinct stages:
    1) Bootstrap Sampling:
    For each tree (t = 1 to T), a random sample with replacement (size N) is drawn from the original dataset ([x, y]), thereby creating a new training set ([xt, yt]).
    2) Decision Tree Growth:
    Employing [xt, yt] as the training data, the t-th decision tree is systematically constructed through recursive binary splits. A stopping criterion is rigorously applied at each unsplit node to govern the controlled growth of the tree.
    3) Ensemble Building:
    This entire process is systematically repeated for T iterations, resulting in the establishment of a forest comprising T independent decision trees. This ensemble forms the basis of the Random Forest algorithm, contributing to its predictive power and robustness.
C. Prediction: For novel test data (xi), the algorithm forecasts the target parameter (y) by consolidating the predictions from individual trees within the forest. This aggregation process involves averaging for regression tasks and employing a majority vote for classification scenarios. The resultant aggregated prediction serves as the ultimate output for the new data point.

The Random Forest algorithm finds application in two primary problem domains: Regression: When the target parameter is continuous, the model endeavors to minimize the mean squared error (MSE) as its designated loss function. Classification: In cases where the target parameter is categorical, the model strives to assign data points to the most probable class. This dichotomy in application demonstrates the adaptability of the Random Forest algorithm across both regression and classification tasks (Schott, 2019).

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(f_i - y_i)^2 \quad (1)$$

Where N denotes the aggregate number of data points; fi represents the model-derived value for a given data point, and yi corresponds to the actual value associated with data point i.

Turning attention to the classification problem, characterized by a discrete target parameter, the resolution entails the utilization of the GINI index. The computation of the GINI index is executed by Equation (2):

$$Gini = 1 - \sum_{i=1}^{C}(P_i)^2 \quad (2)$$

In the given context, pi represents the relative frequency of a specific class within the dataset, and C stands for the total number of classes. Entropy assumes a crucial role in guiding the branching of nodes within a decision tree structure.

$$Entropy = \sum_{i=1}^{C} - P_i * log_2 P_i \quad (3)$$

3.2. The Support Vector Machine (SVM)

The Support Vector Machine (SVM) stands as a formidable and versatile machine learning algorithm utilized for both classification and regression tasks. Introduced by Vladimir Vapnik and colleagues in 1995, SVM has garnered widespread popularity across diverse domains owing to its exceptional performance and solid theoretical underpinnings. The primary objective of an SVM is to discern a hyperplane—a multi-dimensional line or plane—that optimally segregates a dataset of points into their respective classes. This hyperplane is strategically positioned to maximize the margin between the closest points of each class, thereby establishing a distinct decision boundary for subsequent predictions. The elegance of SVM lies in its ability to efficiently oversee complex datasets and provide robust solutions for various real-world problems.

Diverging from certain algorithms, SVM regression does not presuppose a predefined model or data distribution. Instead, it relies on kernel functions to map data points into a higher-dimensional space where an effective hyperplane can be discerned. This non-parametric approach endows SVM with robustness against outliers and enables it to address intricate non-linear relationships within the data.

As per Genuer et al. (2020), in the context of Linear SVM Regression, the function used to predict new values is intricately dependent on the support vectors:

$$\boldsymbol{f(x) = \sum_{n=1}^{N}(a_n - a_n^*)(x_n' x) + b} \quad (4)$$

In the realm of nonlinear SVM Regression, the function utilized to predict new values is expressed as:

$$\boldsymbol{f(x) = \sum_{n=1}^{N}(a_n - a_n^*)G(x_n, x) + b} \quad (5)$$

4. Results

4.1. Machine Learning

After loading essential libraries and datasets, the data is divided into training and testing sets. The initial model is then trained, and its performance undergoes assessment. Following this, hyperparameters undergo tuning through nested loops, and the results are evaluated and stored. This process entails checking for superior combinations, printing the best one, training the ultimate model using the optimal hyperparameters, deploying the final model for predictions, and evaluating its performance. The historical and forecasted values are visualized (See Fig. 2).

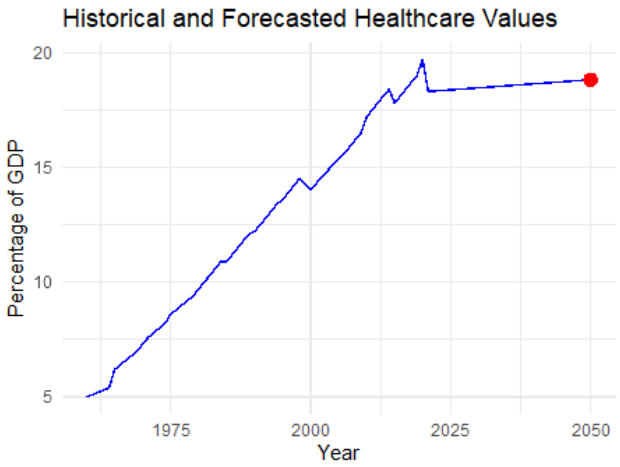

Figure 2. Historical and Random Forest forecasted values

The authors additionally conducted Support Vector Regression (SVR); a specialized variant of Support Vector Machine (SVM) designed for regression tasks. The outcomes are displayed in Table 2.

Table 2. Machine Learning Algorithms

Models	Periods	Forecast for 2050	RMSE
Random Forest	1996-2023	2.297492e+10	846738102
SVM Regression	1996-2023	7.389972e+09	902842632

```
print(forecast_2050)
Model       RMSE Forecast_2050
1  Random Forest  846738102  2.297492e+10
2 SVM Regression  902842632  7.389972e+09
```

Limited data presents a hurdle for Support Vector Regression (SVR). This method thrives on large datasets to grasp underlying patterns and perform well on unseen data. In scenarios like this one, with insufficient data, SVR is susceptible to overfitting and subpar performance. The model struggles to learn effectively due to the scarcity of data points. Therefore, a substantial dataset is essential for training a robust and accurate SVR model.

4.2. Standard Linear Regression

Furthermore, the authors evaluated a single standard linear regression model across the entire time horizon (refer to Table 3 and Fig. 3). Due to the constraints of the limited dataset, no further comparisons were made concerning different periods.

Table 3. A standard regression models

Observations	Time Period	Slope	Intercept	P-value	t Stat	Adjusted R^2
28	1996-2023	0.248992	4.53871	1.18E-38	102.2585	0.997171

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	-4.5E+09	1.69E+09	-2.6533	0.013412	-7.9E+09	-1E+09	-7934794233	-1007296085
Year	7.03E+08	1.02E+08	6.923208	2.38E-07	4.94E+08	9.12E+08	494178422	911543028

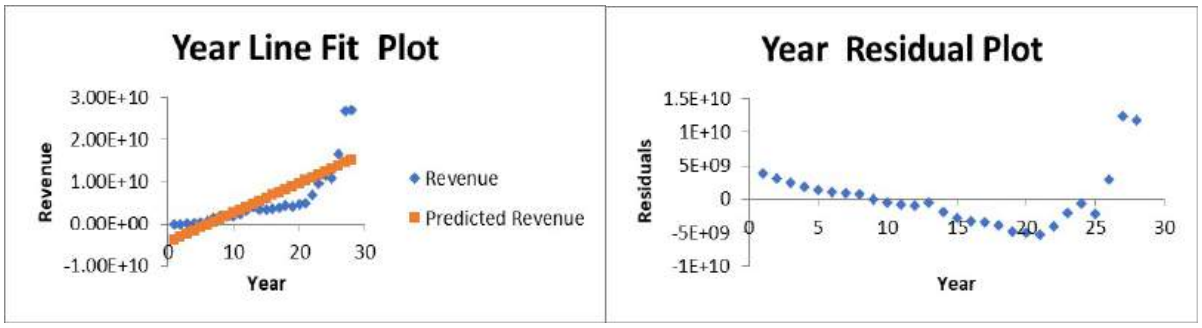

Figure 3. Line Fit Plots and Residual Plots

Rather than thinking in terms of one being inherently more dependable than the other, it is crucial to understand that both Machine Learning (ML) and Regression models have their strengths and weaknesses, making them suitable for different scenarios. Even though Regression models have a strong statistical foundation and nice interpretability, they are good for linear relationships and excel at capturing that trend and making accurate predictions. Based on the current dataset shape and trend, the authors prefer Machine Learning algorithms to oversee complex, non-linear relationships and adapt to diverse data types.

5. Discussions and Limitations

Acknowledging the prevailing emphasis in contemporary studies on the broader landscape of AI potentials and perils, our research delves into a focused analysis of Nvidia—an industry pioneer and primary innovator—providing insights into a sustained long-term trajectory. The consistent annual revenue growth of Nvidia, marked by an impressive rate of 7.03E+08 from 1996 to 2023, signals a robust and promising trend. Significant is the exponential surge in Nvidia's revenues following the heightened international attention and impact catalyzed by ChatGPT since November 30, 2022. This upward trajectory is statistically significant, as validated by both the P-value and t Stat. What distinguishes our research from other publications are its unique attributes, encompassing advanced Machine Learning algorithms and the integration of traditional mathematical models.

Forecasting Nvidia's revenue for 2050 using the Random Forest model yields a value of 2.297492e+10 while employing regular linear regression results in a figure of 3.4165E+10. The two methods show some differences. It is important to acknowledge the inherent complexity and unpredictability of the financial system, which makes it challenging to draw definitive conclusions. The analysis is further compounded by numerous factors, including evolving elements and intricate interactions such as demographics, technology, and political landscapes. These factors contribute to the complexity of future predictions, making them a challenging endeavor.

While both Machine Learning and Regression models provide comprehensive forecasts, their accuracy tends to diminish over longer periods. Predictions are susceptible to unforeseen and significant shifts, such as changes in government policies, economic fluctuations, or major technological breakthroughs. When extending projections, it is prudent to adopt a cautious approach. Considering the integration of additional methods or seeking expert perspectives can enhance the robustness of the analysis. The dynamic nature of the financial landscape underscores the importance of adapting methodologies and incorporating diverse viewpoints to navigate the uncertainties associated with long-term forecasting.

Nvidia confronts escalating competition from industry stalwarts AMD, Intel, and Qualcomm. AMD's Radeon RX 7000 GPUs display notable performance enhancements, complemented by strategic partnerships with Microsoft for cloud gaming and the acquisition of Pensando for bolstering data center solutions. Intel, meanwhile, has entered the fray with the release of Arc Alchemist GPUs, an expanded foundry business, and

the strategic acquisition of Mobileye for advancements in self-driving technology. Qualcomm's Snapdragon 8 Gen 2, prioritizing AI enhancement, collaboration with Samsung for future chip development, and increased forays into the automotive sector, add further complexity to the competitive landscape. While Nvidia maintains its leadership status, the relentless advancements made by these competitors in AI, graphics, and data center solutions necessitate ongoing innovation to uphold market dominance amid the broader diversification driven by tech giants like Apple, Google, and Huawei.

The future of computing is poised for groundbreaking progress, with quantum computing tackling once-impossible problems, neuromorphic computing mimicking the brain's efficiency, and edge computing empowering real-time applications. Artificial general intelligence promises to revolutionize every facet of life, while ongoing advances in materials science pave the way for smaller, faster, and more efficient devices. The integration of these diverse technologies holds the potential for unimaginable breakthroughs. However, ethical considerations and responsible development are paramount in navigating this transformative landscape. Despite uncertain timelines, the future of computing presents an exciting journey marked by innovation and the need for ethical stewardship.

Regression and machine learning (ML) models each play distinct roles. Regression excels in providing interpretability and managing linear relationships, making it ideal for scenarios with clear trends and a focus on understanding underlying patterns. On the other hand, ML models are adept at managing complexity and overseeing diverse tasks, especially in the presence of non-linear data. Given the potential non-linear relationships within Nvidia's datasets, exploring methodologies beyond traditional linear regression is essential. With the rapid advancements in computing, prioritizing models like Random Forests could unlock valuable insights from Nvidia's data. Continuously reassessing method appropriateness and aligning with the evolving data landscape is crucial. By embracing innovation and staying abreast of advancements, Nvidia can harness the power of computing to uncover profound insights and achieve success amidst intricate relationships.

6. Conclusion

In conclusion, our exploration into the historical review and future projections of Nvidia's stock market prowess illuminates the multifaceted dynamics that have propelled this semiconductor giant to its current standing. Through an examination of key milestones, technological innovations, and strategic decisions, we have traced the remarkable evolution of Nvidia's stock performance. The historical retrospective underscores the company's resilience in navigating industry shifts, its adeptness in capitalizing on emerging technologies, and its strategic foresight in forming crucial partnerships. These factors, among others, have contributed significantly to Nvidia's historical success in the stock market. Our forward-looking analysis integrates current market trends, technological advancements, and industry forecasts to project potential future scenarios for Nvidia's stock. This foresight serves as a valuable tool for stakeholders, offering insights into potential opportunities and challenges that may shape the company's trajectory in the coming years.

Nvidia's rise to prominence is due to its strategic planning, internal excellence, and commitment to innovation. As it keeps pushing the boundaries of technology, Nvidia's impact is felt across various industries, making it a key player in the ever-changing tech landscape. Nvidia strategically integrates GPU leadership, AI innovation, expansion into data centers and automotive markets, ecosystem development, and a dedication to sustainability. This comprehensive strategy is designed to foster growth, influence industry trends, and introduce transformative technologies across various sectors. The CEO, Huang, asserted that the artificial intelligence industry, in which Nvidia occupies a pivotal position as a major supplier of crucial chips, has reached a "tipping point," signifying the transition to mainstream adoption of the technology.

As we peer into the future of Nvidia's stock market journey, the semiconductor landscape will continue to evolve, presenting both risks and opportunities. Our hope is that this study serves as a foundation for

informed decision-making, aiding investors, analysts, and industry participants in navigating the complex and dynamic terrain of the semiconductor market. In the face of constant innovation and market fluctuations, Nvidia's ability to adapt, innovate, and capitalize on emerging trends will be pivotal in sustaining and enhancing its stock market prowess. As such, continuous monitoring of industry shifts and initiative-taking strategic decision-making will be essential for stakeholders aiming to stay abreast of Nvidia's ongoing narrative in the ever-evolving world of technology and finance.

7. Disclaimer Statements

Funding: None

Conflict of Interest: None

8. References

Bharilya, V., & Kumar, N. (2024). Machine learning for autonomous vehicle's trajectory prediction: A comprehensive survey, challenges, and future research directions. *Vehicular Communications*, 100733. https://doi.org/10.1016/j.vehcom.2024.100733

Campoy, J., Prado-Rujas, I. -I., Risco-Martin J. L., Olcoz, K., Perez, M.S.(2023) Distributed training and inference of deep learning solar energy forecasting models. In *31st Euromicro International Conference on Parallel, Distributed and Network-Based Processing (PDP)*, Naples, Italy, 2023, pp. 173-176, https://doi.org/10.1109/PDP59025.2023.00035

Chlebus, M., Dyczko, M., & Woźniak, M. (2021). Nvidia's stock returns prediction using machine learning techniques for time series forecasting problem. *Central European Economic Journal, 8*(55), 44-62. https://doi.org/10.2478/ceej-2021-0004

Dikshit, S., Dixit, R., & Shukla, A. (2024). Review and analysis for state-of-the-art NLP models. *International Journal of Systems, Control and Communications, 15*(1), 48-78. https://doi.org/10.1504/IJSCC.2024.135183

Ericson, L., Zhu, X., Han, X., Fu, R., Li, S., Guo, S., & Hu, P. (2024). Deep generative modeling for financial time series with application in VaR: A comparative review. *arXiv preprint* arXiv:2401.10370. https://doi.org/10.48550/arXiv.2401.10370

Genuer, R., Poggi, J.M. (2020). Random Forests. In: *Random Forests with R.* (pp.33-55) Springer, Cham. https://doi.org/10.1007/978-3-030-56485-8_3

Gill, S. S., Wu, H., Patros, P., Ottaviani, C., Arora, P., Pujol, V. C., ... & Buyya, R. (2024). Modern computing: Vision and challenges. *Telematics and Informatics Reports, 13*, 100116. https://doi.org/10.1016/j.teler.2024.100116

Hadi, R. F., Sa'adah, S., & Adytia, D. (2023). Forecasting of GPU prices using transformer method. *Jurnal Sisfokom (Sistem Informasi dan Komputer), 12*(1), 136-144. https://doi.org/10.32736/sisfokom.v12i1.1569

Hu, Y., Chen, L., Wang, Z., & Li, H. (2023). SwinVRNN: A data-driven ensemble forecasting model via learned distribution perturbation. *Journal of Advances in Modeling Earth Systems, 15*(2), e2022MS003211. https://doi.org/10.1029/2022MS003211

Kurth, T., Subramanian, S., Harrington, P., Pathak, J., Mardani, M., Hall, D., ... & Anandkumar, A. (2023, June). FourCastNet: Accelerating global high-resolution weather forecasting using adaptive Fourier neural operators. In Proceedings of the *Platform for Advanced Scientific Computing Conference* (pp. 1-11). https://doi.org/10.1145/3592979.3593412

Li, L., Shao, W., Dong, W., Tian, Y., Yang, K., & Zhang, W. (2024). Data-centric evolution in autonomous driving: A comprehensive survey of big data system, data mining, and closed-loop technologies. *arXiv preprint* arXiv:2401.12888. https://doi.org/10.48550/arXiv.2401.12888

Negi, P., Singh, R., Gehlot, A., Kathuria, S., Thakur, A. K., Gupta, L. R., & Abbas, M. (2024). Specific soft computing strategies for the digitalization of infrastructure and its sustainability: A comprehensive analysis. *Archives of Computational Methods in Engineering*, *31,* 1341-1362. https://doi.org/10.1007/s11831-023-10018-x

NVIDIA. (2024). NVIDIA in Brief. Retrieved on June 9, 2024, from https://www.nvidia.com/content/dam/en-zz/Solutions/about-nvidia/corporate-nvidia-in-brief.pdf

NVIDIA. (2023). NVIDIA revenue 1995-2023. Retrieved on June 9, 2024, from https://finance.yahoo.com/quote/NVDA/analysis/

Ruseno N, Lin CY, Guan WL. (2023). Flight test analysis of UTM conflict detection based on a network remote ID using a random forest algorithm. *Drones, 7*(7):436. https://doi.org/10.3390/drones7070436

Schott M. (2019, April 25) Random forest algorithm for machine learning. *Capital One Tech*. Retrieved from: https://medium.com/capital-one-tech/random-forest-algorithm-for-machine-learning-c4b2c8cc9feb

Shao, Y., Li, H., Gu, X., Yin, H., Li, Y., Miao, X., ... & Chen, L. (2024). Distributed graph neural network training: A survey. *ACM Computing Surveys*, *56*(8), 1-39. https://doi.org/10.1145/3648358

Silva, C. A., Vilaça, R., Pereira, A., & Bessa, R. J. (2024). A review on the decarbonization of high-performance computing centers. *Renewable and Sustainable Energy Reviews*, 189, 114019. https://doi.org/10.1016/j.rser.2023.114019

Silva, J. P., Hagopian, J., Burdiat, M., Dufrechou, E., Pedemonte, M., Gutiérrez, A., ... & Ezzatti, P. (2014). Another step to the full GPU implementation of the weather research and forecasting model. *The Journal of Supercomputing, 70,* 746-755. https://doi.org/10.1007/s11227-014-1193-y

Vanderbauwhede, W., & Takemi, T. (2013, July). An investigation into the feasibility and benefits of GPU/multicore acceleration of the weather research and forecasting model. In *2013 International Conference on High Performance Computing & Simulation (HPCS*) Helsinki, Finland, 2013, pp. 482-489, https://doi.org/10.1109/HPCSim.2013.6641457

Vapnik V. (1995). *The nature of statistical learning theory*. Springer.

Wang, J., Hsu, J., & Qin, Z. (2024). A comprehensive analysis of Nvidia's technological innovations, market strategies, and prospects. *International Journal of Information Technologies and Systems Approach (IJITSA), 17*(1), 1-16. https://doi.org/10.4018/IJITSA.344423

Zhang, Y. X., Haxo, Y. M., & Mat, Y. X. (2023). Nvidia (NASDAQ: NVDA). *AC Investment Research Journal*, *220*(44).

https://doi.org/10.62704/10057/28083

THE APPLICATION OF THE KANO MODEL IN EDUCATION – A SYSTEMATIC LITERATURE REVIEW

Cintia Zuccon Buffon[1]
Elizabeth A. Cudney[2]
Ahmad Elshennawy[1]
[1]*University of Central Florida*
[2]*Maryville University*
cintiazbuffon@ucf.edu

Abstract

As markets become more competitive, customer satisfaction becomes one of the most critical factors of success for any organization, especially in the service industry, where the opportunity for customer delight can be limited. Customer requirements and their perceptions of quality can differ significantly from organizations' views. Therefore, employing tools to bridge this gap and gaining insights into the voice of customers is imperative for organizations. The Kano model provides a framework for understanding the impact of quality attributes on customer satisfaction. The Kano model has been used in many different industries over the years. Multiple case studies have applied the Kano model to various service industry sectors from healthcare to tourism. However, the use of the Kano model in education remains limited. This study provides a systematic literature review of the use of the Kano model in education. This study aims to examine the extent to which the Kano model is applied as a quality improvement tool in the education sector and to identify research gaps in the literature for future research.

Keywords: *Kano Model; Student Satisfaction; Education; Voice of the Student*

1. Introduction

The Kano model is a widely used tool to improve customer satisfaction. It classifies quality attributes based on customer perspective. The Kano model has been used in various sectors to improve services and products. Multiple applications have been found in the literature, including hospital and medical services, product development, education, website management, and government services (Qiu & Li, 2022). Although the Kano model has been applied across various industries, its application in education is minimal. This study uses a systematic literature review approach to investigate the implementation of the Kano model in education and identify research gaps.

As with any other organization with processes, the education system can be seen as a production system with unique inputs, processes, and outputs. Therefore, quality tools used in manufacturing industries can be adapted for education. The Kano model is one of the tools that can be explored and used to better understand students' voices and identify which attributes are relevant to their experiences (Grapragasem et al., 2014; Ku and Shang, 2020).

As defined by Kano et al. (1984), the Kano model is a tool for identifying the correlation between objective quality, which relates to conformance to specified requirements, and subjective quality, which relates to customer satisfaction. The model proposed by Kano et al. (1984) comprises a two-dimensional customer satisfaction method based on the correlation between the fulfillment of

Submitted: April 26, 2024
Revised: June 10, 2024

requirements and customer satisfaction. Kano et al. (1984) presented five categories of quality elements (or attributes): attractive, one-dimensional, must-be, indifferent, and reverse. Figure 1 was adapted from Kano et al. (1984) and illustrates how these attributes are arranged in a two-dimensional plane.

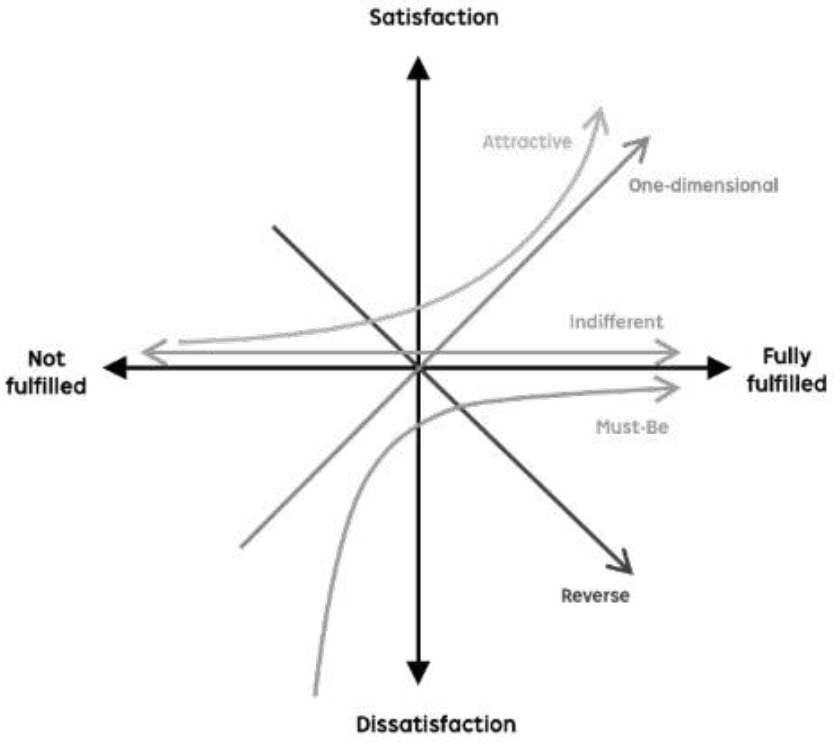

Figure 1. The Kano model (Mindmesh, n.d.)

Attractive attributes satisfy customers when present but do not lead to customer dissatisfaction if absent. One-dimensional attributes satisfy customers when included, but cause dissatisfaction if missing. Must-have attributes do not increase satisfaction when present but lead to significant dissatisfaction if absent. Indifferent attributes neither satisfy nor dissatisfy customers regardless of their presence. Reverse attributes cause dissatisfaction when present, and satisfaction when absent.

The Kano model provides a systematic framework for categorizing attributes based on their impact on customer satisfaction and dissatisfaction. This model guides the implementation of features that enhance customer satisfaction and retention. Applying the Kano model to higher education can determine the features and attributes that students consider significant for their educational experience. However, the literature review revealed the limited use of the Kano model in education, particularly in the United States. This limited research highlights the need to better understand students' preferences, especially as a new generation of students entering higher education. Universities must understand what makes them a preferred choice among students because of increased competition and diverse options.

2. Research Methodology

Research regarding the application of the Kano model in education is still scarce compared with other industries. This study used the preferred reporting items for systematic reviews and meta-analyses (PRISMA) protocol to produce a comprehensive systematic literature review of the material published in this area. This study aimed to answer the following research questions (RQ):

RQ1: What is the current literature on applying the Kano model in education?

RQ2: What are the diverse applications and implementations of the Kano model within the education sector and how do these uses contribute to enhancing educational processes and outcomes?

This systematic literature review searched for articles in the following databases: Education Resources Information Center (ERIC), ProQuest, Engineering Village (Compendex and Inspec), Web of Science, Science Direct, and IEEE Xplore. An advanced search of these databases was performed using different combinations of keywords. When searching for Kano model applications in education, the following key terms were used: ("Kano model" AND "education") OR ("Kano model" AND "student satisfaction") OR ("Kano model" AND "teaching effectiveness") OR ("Kano model" AND "teaching quality"). Figure 2 shows the process flow diagram of this study.

The articles were excluded from the literature review, if found to be duplicate from different sources or not related to education, and / or did not use the Kano Model in their study. This study identified various applications of the Kano model in education, categorizing the studies into four main groups. The first group focuses on applying the Kano model to assess the quality attributes for an institution's overall quality. The second group examined E-learning and virtual platforms. The third group explored the application of the Kano model to understand teaching quality. Finally, the fourth group encompasses other applications, including studies examining students' perceptions of quality in areas such as university furniture or library services.

3. Systematic Literature Review

Although the Kano model was developed in 1984, its use is still limited when considering educational settings. This literature review found that most studies were related to the institution's overall quality, with 15 papers published from 2010 to 2023. Online learning also had the highest number of papers, with many published after the pandemic. There have been 12 published papers on teaching quality. The last category is related to papers that do not fit into other categories.

It is possible to understand from this research that most studies had small sample sizes and investigated a large number of attributes. The majority of papers published in English are from the Asia-Pacific region (26), and only a few studies have been conducted in the United States. This finding highlights the need for more studies to understand the current needs of students, particularly in higher education.

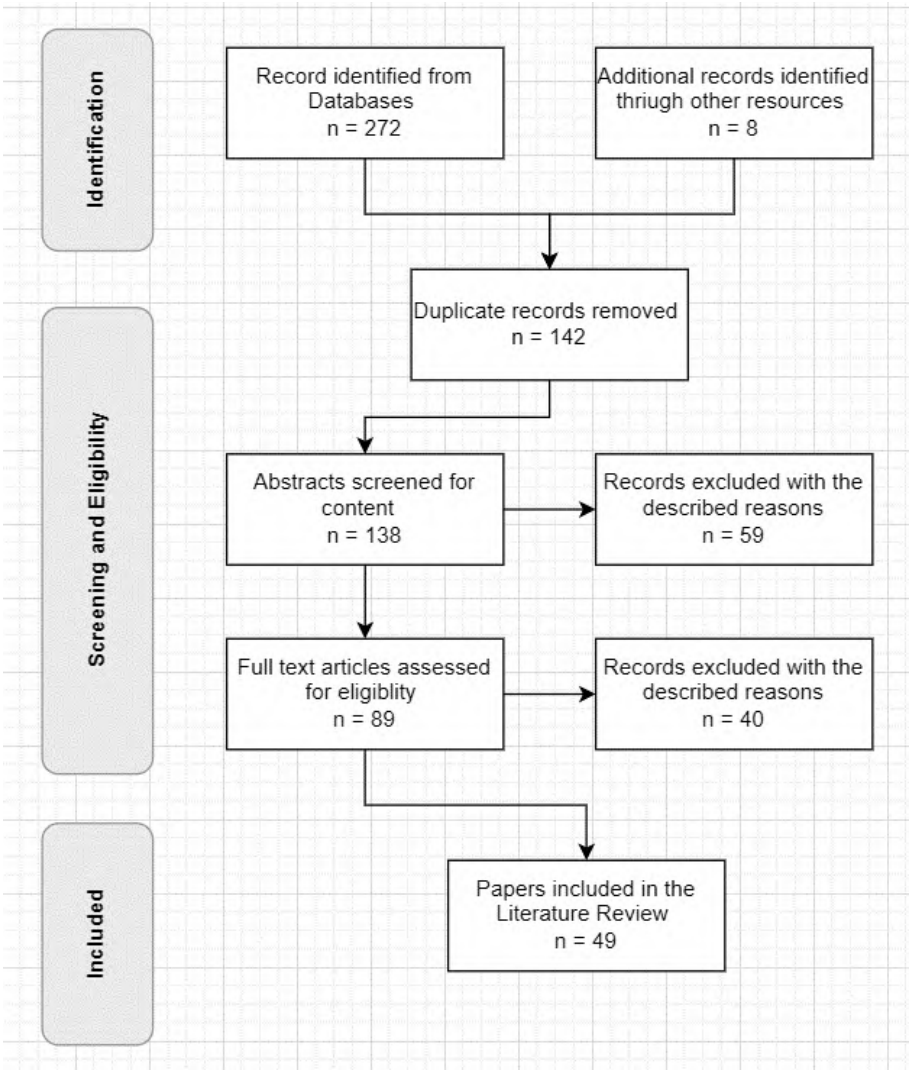

Figure 2. PRISMA flow diagram

3.1 Overall Quality of the Institution

Several studies have applied the Kano model to assess attributes related to an institution's overall quality. Kinker et al. (2023) used the Kano questionnaire to understand the differences between the perceptions of students and professors. Their results showed significant differences in the educational quality attributes that students consider important compared to how professors classify them. In this study, the only common attribute was green campus initiatives, which were considered indifferent by both groups.

Many studies in this area have utilized other tools incorporated into the Kano model. For example, Ganbold et al. (2022) used the five dimensions of SERVQUAL to categorize forty attributes. The study found that students considered online course evaluations and university office hours to be must-be attributes. Tehranineshat et al. (2021) categorized the attributes identified in the five dimensions of SERVQUAL: tangibles, reliability, responsiveness, assurance, and empathy. This study found that nursing and midwifery students considered 27 one-dimensional attributes, including high-quality curriculum and standardized methods.

SERVQUAL is typically used to measure service quality, and it works well with the Kano model, as illustrated in different studies. Sukwadi and Yang (2012a, b) used SERVQUAL in their study. Swakwadi and Yang (2012a) used an integrated SERVQUAL model, fuzzy logic, Kano model, and blue ocean strategy to categorize their findings. Of the 29 quality attributes identified, six were considered indifferent, among them "employees professionally dressed." Swakwadi and Yang (2012b) separated the attributes into service quality, student satisfaction, and student loyalty. The results showed that "library provides up-to-date learning sources" was a must-be attribute.

Sahney (2011) applied the Kano model with SERVQUAL to use its results to develop the Quality Functional Deployment (QFD). This study found that "sufficient staff" and "responsiveness" are one-dimensional attributes. Kelesbayev et al. (2020) also used the Kano model integrated into the QFD. Thirty attributes related to features, administrative staff, courses and exams, practice and career, qualifications, and opportunities were identified. Rabaiei et al. (2021) used a novel combination of the Kano model and data-mining algorithms to understand student satisfaction. After classifying attributes into Kano categories, different data mining algorithms were used for feature selection to prioritize the one-dimensional, must-be, and attractive attributes to drive customer satisfaction.

Mohammad (2020) focuses on studying the quality attributes of students at a university in Jordan. He found that the majority of the 37 attributes were considered one-dimensional, including "updated material." This study identified a reverse attribute related to course offerings in the scientific stream. Nzumile and Taifa (2021) also found reverse attributes. Students are dissatisfied if "classes are closed after lectures" and "student's VIP section in the cafeteria". In their study, out of the 46 attributes, half of them were considered one-dimensional, among them "higher number of classes," "well-equipped laboratories," and "affordable food price."

Chan et al. (2014) used data from open-ended and multiple-choice surveys, which did not follow the Kano questionnaire. They performed the attribute selection, refinement, and classification. Based on the results of these steps, they compared their classification to the existing work that used the Kano questionnaire. Overall, their classification method worked well for the attributes that appear in other studies. For example, their study identifies that the tutorial structure is considered one-dimensional, the same as that identified in another study.

Madzík et al. (2019) investigated seven critical requirements in different university areas. They found that "research orientation" was classified as an attractive attribute with the most influence on student satisfaction. Yang et al. (2010) performed multiple iterations using a customer-oriented reputation model to reduce the number of analyzed attributes. The results showed that "teaching quality" was considered one-dimensional. Ghosh (2014) explores attributes related to management and university facilities, showing differences in students' preferences across the institutions where the study was performed.

3.2 Online Education

Another category identified in multiple articles through the literature review was online learning, in which articles focused on student satisfaction regarding the online methods and resources used. Online learning has become even more critical during the COVID-19 pandemic, when it is the only solution to continue the learning process (Qiao & Zhang, 2022). Chen et al. (2022) highlighted that the success of an e-learning environment is highly dependent on user satisfaction. Their study used the Kano model to understand e-learning in film majors. They also used the decision tree algorithm to understand the influence of student demographics on the classification of attributes. However, the decision tree was not accurate enough to predict that influence mentioned above, probably because of the small size of the data points gathered from the survey.

Um et al. (2021) used SERVQUAL to identify seven main quality attributes and the modified Kano model to understand how those attributes affect satisfaction and dissatisfaction with blended learning – virtual and non-virtual learning. The Kano Model was modified to consider only the main three categories attractive, one-dimensional and must be. Similarly, Bauk et al. (2014) studied students' satisfaction with courses taught in a blended environment for students from a Mediterranean University. They identified ten attributes based on the five quality components of the DeLone and Mclean (D&M) model. This study was important because most respondents classified the attributes as indifferent, which could be due to an incorrect understanding of the questionnaire, as many responses were categorized as questionable. The same occurred with the work of Bauk and Jusufranic (2014), where except for "available access to the system at any time," all the other nine attributes were classified as indifferent by the students.

Violante and Vezzetti (2013) used the Kano model to understand e-learning and to train students to use medical devices. This study identified three categories of attributes: learning interface, content, and personalization. In total, 18 quality attributes were evaluated using the Kano model. González-Yebra et al. (2019) studied the incorporation of alternative spaces into learning interactions. The use of a 3D virtual campus as a learning element was studied. There were 22 attributes were identified in these four categories. No must-be attributes were identified, but one attribute was considered reverse, "advisable for students with a lot of patience," which makes sense, as all students want to be able to use the feature.

Qiao and Zhang (2022) identified 18 quality attributes distributed into five categories related to massive open online courses (MOOCs). Interestingly, most of the attributes were considered attractive. This finding indicates that, since this is a relatively new approach, students will consider all the course features as delighters; as time passes and more of those courses are available, the scenario might change to where some of those attributes are considered basic attributes. Kang and Chen (2016) reported a similar finding. They applied the Kano model to understand customer perspectives on mobile learning. Their study identified that most attributes (16 out of 19) were considered attractive by Taiwanese students. These attributes include "course-time arrangement", "flexibility", "task content," and "efficiency".

Fujs et al. (2022) used focus groups to identify the attributes in their research. They prepared two Kano questionnaires: one for students and one for instructors. They implemented their approach in three case studies that were conducted at various universities in Slovenia. Studies that consider students' and professors' perspectives are essential because professors' ideas of a factor may differ significantly from those of their students. Chen (2023) utilized data from user comments from an advanced mathematics course. They did not use the Kano questionnaire to collect data. However, they used the Kano model classification and an exponential algorithm to calculate better and worse coefficients based on the data collected. The results show one reverse attribute related to the length of the video, indicating that if the length is not moderate, it can cause dissatisfaction among users. This finding corroborates the work of Kotler et al. (2019), who expressed that people are less attentive nowadays and that the need for shorter videos is a reality.

Marutschke and Hayashi (2022) investigated students' perceptions of an online course by comparing the results of applying the Kano model before and after students took the online course. Marutschke and Hayashi (2021) used the same approach. They investigated 12 features by using pre- and post-questionnaires. Based on the results, it was possible to observe changes in students' perceptions of some of the attributes investigated. The shift in perception occurred for most of the attributes, except for "user-friendly platform", "quizzes and exercises", "interactive features," and "text," which maintained their importance on both results. Dorner and Kumar (2017) used the same

approach. They used the Kano model to understand the five attributes related to online learning, considering pre- and in-service teaching. The study identified that both surveys classified computer skills and Internet abilities as must-be, whereas mentor activity and mentoring were identified as indifferent.

Dominici and Palumbo (2013) focused on understanding the importance of e-learning for students. This study is one of the few that used the Kano model to assess virtual features before the pandemic and highlights the importance of virtual learning for higher education. Fazlollahtabar et al. (2012) also conducted a study to investigate which aspects of virtual learning most influenced student satisfaction before the pandemic. This study surveyed students from Iran to shed light on six attributes. This indicates that most attributes must be considered when designing e-learning systems. The "user interface", "infrastructure", and "content" are all basic requirements for students to be satisfied with online learning.

3.3 Teaching Quality

Teaching quality or effectiveness was the only category of studies performed in the US. The work of Cudney et al. (2023) is a pilot study that investigates the main attributes of teaching effectiveness based on students' perceptions. Based on the literature, they identified 25 quality attributes for their research. Their results show that 19 attributes were identified as one dimensional, among them: "inclusive learning environment," "ideas and concept communicated clearly," "provides quick turnaround on grading," "shows respect," "use grading rubrics," and "demonstrates genuine interest in my success."

Goerdt et al. (2019) applied the Kano model to social work students in five different US universities. One main difference in this study is that the authors deployed the Kano questionnaire at the beginning of the semester, and another questionnaire after the test. The study identified four must-be factors: "expertise on the subject", "helpfulness", "logical structure of lecture," and "respect," and six attributes were considered attractive: "use of humor", "fostering teamwork," and "friendliness" are some examples. Gruber et al. (2010) conducted a study using the Kano model at two universities in the United States, specifically for marketing majors. The authors highlight the importance of professors' understanding of how to better deliver their courses to students to ensure successful outcomes. The two universities have the same attribute classification, except for one factor – "humor" – which was considered an attractive attribute by the Midwest university but as one-dimensional by the Southwest university.

Chen and Zheng (2023) used the Kano model to explore how nursing students in China perceived 16 quality attributes. Some of the attributes identified by the authors were: "ethics", "manners", "professional knowledge", "teaching ability," and "teaching methods". Their study found that all 16 attributes studied were considered one-dimensional by the students. Strano et al. (2018) identify three dimensions of quality attributes: personal, external, and institutional. Their survey contained 17 attributes related to the personal dimension, 13 to the institutional dimension, and 13 to the external dimension. The main differentiator of their study was that they applied the survey to both the students and teachers. While students identified that it is important for teachers to "promote the integration of students into new teaching and learning methodologies and evaluations," teachers classified this attribute as indifferent.

Tetteh (2015) raised an important question regarding the role of teachers in learning outcomes for students. During the earlier times, eras 1.0 to 3.0, the learning responsibility was solely on the students, where teaching characteristics were not considered if the student had a poor outcome. This finding has changed during the latest era, when teachers are found to be an essential part of

the learning path, and they act more as facilitators. Their study found that, out of the 35 attributes investigated, nine were classified as must-be, including: “fair principles,” “encourages teamwork,” and “flexibility.” Gruber et al. (2010) used the Kano model to understand teachers’ characteristics as quality attributes in higher education for marketing students in the UK. Their study identified 19 attributes of teaching quality based on previous studies and focus groups. This study emphasizes the importance of viewing students as active contributors to the services offered. It also notes that they should not anticipate having a direct influence over the course content, but more on its delivery.

Sung (2009) employed the Kano model to identify the quality attributes in English instruction among university students from Indonesia and Taiwan. Both groups underscored the importance of incorporating culturally relevant content into language learning, advocating for a more holistic educational approach. The study identified that Indonesian students give more importance to cultivating a supportive learning atmosphere, while Taiwanese students prioritize instructors’ oral communication proficiency. These results highlight the need for tailored English courses to address the distinct needs of students from diverse cultural backgrounds. Chien (2007) identified five dimensions of teaching quality based on literature: individual characteristics, teaching attitude and ability, course quality, learning environment, and teaching purpose. In their study, they found that eight quality attributes were identified as must-be by Taiwanese students: “the aim of the course”, “lecturing ability”, “interactivity,” and “personal attainment”.

The Kano model can be combined with other quality and customer satisfaction tools to identify quality attributes and improve the overall student satisfaction. This can be an enhancement to a better understanding of customer satisfaction. Ku and Shang (2020) integrated the Kano model with the Revised Importance-Performance Analysis (RIPA) framework to understand students’ preferences in physical education. Kuo et al. (2011) integrated the use of Importance-Performance Analysis (IPA) and the Kano model to better understand students’ preferences in tourism and hospitality. Takagi and Uesu (2014) utilized the Kano model with fuzzy logic to uncover the needs and preferences of undergraduate students regarding mathematics lectures.

3.4 Other Applications

Seven other studies did not fall into prior categories and were classified as other applications. Kapuria et al. (2019) used the Kano model to understand student perceptions of university classroom furniture. This study is an interesting use of the Kano model in education because it identifies an area that has not yet been explored in terms of student satisfaction. Taifa and Desai (2016) also studied student satisfaction with classroom furniture. They developed 23 attributes and found that the majority of them were indifferent to the students, including durable furniture, adjustable seats, easy-to-move, adjustable height, and desk material.

Three other studies have used the Kano model to understand students’ preferences in course development. Thongoiam and Phumchusri (2022) used two different methods to understand how industrial engineering students felt about ten attributes related to the course development of a master’s program. This study exemplifies one way of using the Kano model before a product/service design to better understand customer preferences. It was found that all four attributes investigated were considered indifferent for the students, and that they would not interfere with their level of satisfaction if present or not in the new master’s program. The study by Song et al. (2021) had the largest sample size in this literature review. Their study had an 89% response rate, with 976 responses. This study used the Kano model to better understand the attributes related to the art appreciation curriculum. Ömürgönülşen et al. (2021) used the Kano model to enhance the quality of education by understanding students’ requirements for course development. They utilized the results from the Kano model as part of the students’ need to build a house of quality through quality

function deployment (QFD).

Cahyana et al. (2019) applied the Kano model to understand student preferences regarding academic information services. This study was performed in the Department of Informatics of an Indonesian university. The results show that three attributes were considered one-dimensional: "availability of academic information", "conventional academic information services," and "enhancement of the platform according to students' choices". The other 13 attributes were considered attractive, including "unlimited access to the academic information", "availability of digital information boards," and "fast process in adding academic information".

Mean-Shen (2010) used the Kano model to analyze the quality attributes of university libraries in Taiwan. They identified 34 quality attributes in three categories: equipment, services, and environment. That study found multiple attributes that students considered must-be at a library university. Some of them are "network connections," "availability to renew a book online," "clear organization," and others.

4. Conclusions and Future Work

The systematic literature review presented in this paper sheds light on the current state of research on the application of the Kano model in education. Despite the widespread use of the Kano model in various industries, its application to education remains relatively limited. Although its use has been consistent over the years, its adoption remains in its initial stages.

This review highlights the areas in which the Kano model has been applied in education. These include assessing the overall quality of educational institutions, evaluating online learning, understanding teaching quality, and other applications, such as student preferences for classroom furniture and course development. The reviewed studies demonstrate the importance of understanding quality attributes from the students' perspective to enhance their satisfaction.

This study revealed variations in methodologies used across studies, including integrating the Kano model with other frameworks, such as SERVQUAL and QFD. Moreover, pre- and post-course surveys provide valuable insights into how students' perceptions evolve. Most studies originate from the Asia-Pacific region, indicating a geographical disparity in the research focus. Additionally, many studies had small sample sizes, suggesting the need for larger-scale investigations to validate the findings and ensure broader applicability.

Further research should explore understudied areas within education, such as applying the Kano model to specific disciplines, student services, and administrative processes. Additionally, comparative studies across different cultural and educational contexts can provide valuable insights into how student preferences and satisfaction vary globally. Longitudinal studies that track student satisfaction and perceptions over an extended period can offer insights into evolving trends and the impact of interventions.

5. Disclaimer Statements

Funding: None

Conflict of Interest: None

6. References

Bauk, S., & Jusufranic, J. (2014). Competitiveness in higher education in terms of the level of students satisfaction with e-learning in blended environment. *Montenegrin Journal of Economics, 10*(1), 25–42. Retrieved from https://www.mnje.com/sites/mnje.com/files/025-042__bauk_and_jusufranic.pdf

Bauk, S., Šćepanović, S., & Kopp, M. (2014). Estimating students' satisfaction with e-learning system in blended learning. In *INTED2014 Proceedings*: *8th International Technology, Education and Development Conference* (pp. 263–271). Valencia, Spain: IATED. https://library.iated.org/view/BAUK2014EST

Cahyana, R., Rahayu, S., & Satria, E. (2019). Revealing student satisfaction related to academic information services using the Kano model. *Journal of Physics: Conference Series, 1402*(6), 066106. https://doi.org/10.1088/1742-6596/1402/6/066106

Chan, T., Rosemann, M., & Shiang-Yen, T. (2014). Identifying satisfaction factors in tertiary education: The case of an information systems program. In B. Tan, E. Karahanna, & A. Srinivasan (Eds.), *Proceedings of the 35th International Conference on Information Systems* (pp. 1-17). Association for Information Systems (AIS). Retrieved from https://eprints.qut.edu.au/84096/

Chen, H. (2023). Influencing factors of the quality of MOOCs based on the KANO model. *International Journal of Emerging Technologies in Learning, 18*(17), 20–32. https://doi.org/10.3991/ijet.v18i17.42507

Chen J-L, Zheng L-N. (2023). The satisfaction of the undergraduate nursing classroom teaching quality based on the Kano model in China. *SAGE Open Medicine*, *11*. https://doi.org/10.1177/20503121231157207

Chen, W., Chang, J., Chen, L., & Hsu, R. (2022). Using refined Kano model and decision trees to discover learners' needs for teaching videos. *Multimedia Tools and Applications*, *81*(6), 8317–8347. https://doi.org/10.1007/s11042-021-11744-9

Chien, T. (2007). Using the learning satisfaction improving model to enhance the teaching quality. *Quality Assurance in Education, 15*(2), 192–214. https://doi.org/10.1108/09684880710748947

Cudney, E. A., Anderson, S., Beane, R., Furterer, S., Mohandas, L., & Laux, C. (2023). Using the voice of the student to identify perceptions of teaching effectiveness attributes: a pilot study. *Quality Assurance in Education, 31*(3), 485–583. https://doi.org/10.1108/QAE-10-2022-0187

Dominici, G., & Palumbo, F. (2013). How to build an e-learning product: Factors for student/customer satisfaction. *Business Horizons, 56*(1), 87–96. https://doi.org/10.1016/j.bushor.2012.09.011

Dorner, H., & Kumar, S. (2017). Attributes of pre-service and inservice teacher satisfaction with online collaborative mentoring. *Online Learning Journal, 21*(4), 283–301. https://doi.org/10.24059/olj.v21i4.1020

Fazlollahtabar, H., Rezaie, M., & Nosratabadi, H. E. (2012). Applying Kano model for users' satisfaction assessment in e-learning systems. *International Journal of Information and Communication Technology Education (IJICTE) 8(3)* 1–12. https://doi.org/10.4018/jicte.2012070101

Fujs, D., Vrhovec, S., Žvanut, B., & Vavpotič, D. (2022). Improving the efficiency of remote conference tool use for distance learning in higher education: A kano based approach. *Computers & Education, 181*, 104448. https://doi.org/10.1016/j.compedu.2022.104448

Ganbold, B., Park, K., & Hong, J. (2022). Study of educational service quality in Mongolian universities. *Sustainability, 15*(1), 580. https://doi.org/10.3390/su15010580

Ghosh, K. (2014). Institutional effectiveness in higher education: a comparative analysis of students' perceived responses in Indian context. *International Journal of Indian Culture and Business Management, 9*(2), 181. https://doi.org/10.1504/ijicbm.2014.064188

Goerdt, L. A., Blaalid, B., Elswick, S., Johnson, D. H., Delavega, E., Kindle, P. A., & Granruth, L. B. (2019). Teaching characteristics and student satisfaction: Impact on social work students' interest in policy. *Journal of Social Work Education, 55*(4), 767–776. https://doi.org/10.1080/10437797.2019.1611512

González-Yebra, Ó., Aguilar, M. A., Aguilar, F. J., & De Jesus Lucas, M. (2019). Co-design of a 3D virtual campus for synchronous distance teaching based on student satisfaction: experience at the University of Almería (Spain). *Education Sciences, 9*(1), 21. https://doi.org/10.3390/educsci9010021

Gruber, T., Fuß, S., Voss, R., & Gläser-Zikuda, M. (2010). Examining student satisfaction with higher education services: Using a new measurement tool. *International Journal of Public Sector Management, 23*(2). https://doi.org/10.1108/09513551011022474

Gruber, T., Reppel, A., & Voss, R. (2010). Understanding the characteristics of effective professors: The student's perspective. *Journal of Marketing for Higher Education, 20*(2), 175–190. https://doi.org/10.1080/08841241.2010.526356

Kang, Y., & Chen, J. (2016). Applying Kano model to evaluate the quality for mobile experiential learning model. In *2016 International Conference on Applied System Innovation (ICASI).*

https://doi.org/10.1109/icasi.2016.7539882
Kano, N., Seraku, N., Takahashi, F., & Tsuji, S. (1984). Attractive quality and must-be quality. *Journal of the Japanese Society for Quality Control, 14*(2), 147-156. https://doi.org/10.20684/quality.14.2_147
Kapuria, T. K., Rahman, M., & Doss, D. A. (2019). Users' satisfaction attributes analysis toward betterment of classroom furniture design for Bangladeshi university students. *Journal of the Institution of Engineers India Series C, 101*(1), 175–184. https://doi.org/10.1007/s40032-019-00541-x
Kelesbayev, D., Alibekova, Z., Izatullayeva, B., Dandayeva, B., & Mombekova, G. (2020). Establishing a quality planning scheme with Kano model and a case study. *Calitatea: Acces La Success*, *21*(176), 56–64.
Kinker, P., Singh, R. K., Singh, A., & Jain, R. (2023). An approach to evaluate service quality in polytechnic education institutes: a case study. *International Journal of Applied Systemic Studies, 10*(1), 16. https://doi.org/10.1504/ijass.2023.129064
Kotler, P., Kartajaya, H., & Setiawan, I. (2019). Marketing 4.0: Moving from traditional to digital. In P. Kotler, H. Kartajaya & D. Hooi (Eds.)*, Asian Competitors: Marketing for Competitiveness in the Age of Digital Consumers* (pp. 99–123). New Jersey: World Scientific https://doi.org/10.1142/9789813275478_0004
Ku, G. C., & Shang, I. (2020). Using the Integrated Kano–RIPA model to explore teaching quality of physical education programs in Taiwan. *International Journal of Environmental Research and Public Health*, *17*(11), 3954. https://doi.org/10.3390/ijerph17113954
Kuo, N., Chang, K., & Lai, C. (2011). Identifying critical service quality attributes for higher education in hospitality and tourism: Applications of the Kano model and importance-performance analysis (IPA). *African Journal of Business Management, 5*(30). https://doi.org/10.5897/AJBM11.1078
Madzík, P., Budaj, P., Mikuláš, D., & Zimon, D. (2019). Application of the Kano model for a better understanding of customer requirements in higher education—a pilot study. *Administrative Sciences, 9*(1), 11. https://doi.org/10.3390/admsci9010011
Marutschke, D. M., & Hayashi, Y. (2021). Ex-ante and ex-post feature evaluation of online courses using the Kano model. In *In A. I. Cristea & C. Troussas (Eds.), Intelligent tutoring systems* (Vol. 12677, pp. 310-320). Springer. https://doi.org/10.1007/978-3-030-80421-3_34
Marutschke, D.M., Hayashi, Y. (2022). Kano model-based macro and micro shift in feature perception of short-term online courses. In: Wong, LH., Hayashi, Y., Collazos, C.A., Alvarez, C., Zurita, G., Baloian, N. (Eds.), *Collaboration Technologies and Social Computing: CollabTech 2022.*(Vol. 13632, pp. 112-125) Springer, Cham. https://doi.org/10.1007/978-3-031-20218-6_8
Mean-Shen, L. (2010). A refined and integrated Kano model and the implementation of quality function deployment - Research on the library of a vocational and technical school in Southern Taiwan. *International Journal of Organizational Innovation,* 2(3), 305–335. Retrieved from https://www.ijoi-online.org/index.php/back-issues/22-volume-2-number-3-winter-2010
Mindmesh. (n.d.). *The Kano Model*. Mindmesh. https://www.mindmesh.com/glossary/what-is-kano-model
Mohammad, A. (2020). Applying the Kano model to quality improvement within higher education: An applied study in the World Islamic Science and Education University - Jordan. *Calitatea: Acces La Success, 21*(175), 62–67.
Nzumile, J. M., & Taifa, I. W. R. (2021). Stratification of students' satisfaction requirements using the Kano model. *The Journal of Business Education, 1*, 1–16.
Ömürgönülşen, M., Eryiğit, C., Tektaş, Ö. Ö., & Soysal, M. (2021). Enhancing the quality of a higher education course: quality function deployment and Kano model integration. *YüKseköĞRetim Dergisi, 10*(3), 312–327. https://doi.org/10.2399/yod.19.560956
Qiao, L., & Zhang, Y. (2022). Analysis of MOOC quality requirements for landscape architecture based on the Kano model in the context of the COVID-19 epidemic. *Sustainability, 14*(23), 15775. https://doi.org/10.3390/su142315775
Qiu, Y., & Li, W. (2022). The application of connotative teaching in private colleges based on Kano model. *Mathematical Problems in Engineering*, *2022*(1), 1–8. https://doi.org/10.1155/2022/8269083
Rabaiei, K. A., Alnajjar, F., & Ahmad, A. (2021). Kano model integration with data mining to predict customer satisfaction. *Big Data and Cognitive Computing, 5*(4), 66. https://doi.org/10.3390/bdcc5040066
Sahney, S. (2011a). Delighting customers of management education in India: A student perspective, part II. *The TQM Journal, 23*(5), 531–548. https://doi.org/10.1108/17542731111157635
Sahney, S. (2011b). Delighting customers of management education in India: A student perspective, part I.

The TQM Journal, 23(6), 644–658. https://doi.org/10.1108/17542731111175257

Song, B., Lin, R., Wang, Z., Cai, Y., & Wang, F. (2021). Curriculum design of art appreciation in colleges based on Kano model demand analysis. In *Proceedings of the 2021 International Conference on Digital Society and Intelligent Systems (DSInS)* (pp. 67-71). IEEE. https://doi.org/10.1109/DSInS54396.2021.9670590

Strano, M., Gabriel, B., Moreira, G., Dias-De-Oliveira, J., Valente, R. a. F., & Neto, V. (2018). Towards excellence in engineering education: The profile of higher education teachers on engineering programs. In *Proceedings of the 2018 3rd International Conference of the Portuguese Society for Engineering Education (CISPEE)* (pp. 1-8) IEEE. https://doi.org/10.1109/cispee.2018.8593382

Sukwadi, R., & Yang, C. (2012). Determining critical service attributes and appropriate improvement actions in Indonesian HEIs. *Industrial Engineering and Management Systems, 11*(3), 241–254. Korean Institute of Industrial Engineers. https://doi.org/10.7232/iems.2012.11.3.241

Sung, D. (2009). Attractive quality attributes of English language teaching at two east Asian universities. *The journal of Asia TEFL, 6*(4), 131–149.

Taifa, I. W. R., & Desai, D. A. (2016). Student-defined quality by Kano model: A case study of engineering students in India. *International Journal for Quality Research, 10*(3), 569-582 https://doi.org/10.18421/ijqr10.03-09

Takagi, S., & Uesu, H. (2014). A questionnaire analysis by fuzzy reasoning for undergraduate students in mathematics lectures applying the Kano model. In *Proceedings of the 2014 Joint 7th International Conference on Soft Computing and Intelligent Systems (SCIS) and 15th International Symposium on Advanced Intelligent Systems (ISIS)* (pp. 851-854). IEEE. https://doi.org/10.1109/scis-isis.2014.7044903

Tehranineshat, B., Naderi, Z., Momennasab, M., & Yektatalab, S. (2021). Assessing the expectations and perceptions of nursing students regarding the educational services in a school of nursing and midwifery based on the SERVQUAL and Kano models: A case study. *Hospital Topics, 100*(1), 26–34. https://doi.org/10.1080/00185868.2021.1913080

Tetteh, G. A. (2015). Improving learning outcome using Six Sigma methodology. *Journal of International Education in Business, 8*(1), 18–36. https://doi.org/10.1108/jieb-10-2013-0040

Thongoiam, M., & Phumchusri, N. (2022). Identifying customer needs for a master's degree program in industrial engineering: A case study from prospective students' insights. In *Proceedings of the 4th International Conference on Management Science and Industrial Engineering (MSIE '22)* (pp. 206–213. ACM. https://doi.org/10.1145/3535782.3535810

Um, K.-H., Fang, Y., & Seo, Y.-J. (2021). The ramifications of COVID-19 in education: Beyond the extension of the Kano model and unipolar view of satisfaction and dissatisfaction in the field of blended learning. *Educational Measurement: Issues and Practice, 40*(3), 96-109. https://doi.org/10.1111/emip.12432

Violante, M. G., & Vezzetti, E. (2015). Virtual interactive e-learning application: An evaluation of the student satisfaction. *Computer Applications in Engineering Education, 23*(1), 72–91. https://doi.org/10.1002/cae.21580

Yang, C.-C., Sukwadi, R., & Mu, P.-P. (2010). Integrating Kano model into customer-oriented reputation model in higher education. In *Proceedings of the 2010 International Conference on Management and Service Science* (pp. 1-4). IEEE. https://doi.org/10.1109/ICMSS.2010.5576618

The Journal of Management and Engineering Integration Vol. 17, No. 1 | Summer2024
https://doi.org/10.62704/10057/28085

A PRACTICAL GUIDE TO ISO 25022: MEASUREMENT OF SOFTWARE QUALITY IN-USE

Bassam Jaradat[1]
Gamal Weheba[1]
[1]*Wichita State University*
gamal.weheba@wichita.edu

Abstract

This research investigated the challenges in measuring software Quality in-use following the ISO/IEC 25022:2016 standard. By highlighting these implementation challenges, the research proposed an implementation guide to simplify the evaluation process and recommended adjustments for future iterations of the standard. These recommendations are geared towards enhancing the precision of software quality measurement practices.

Keywords: *Software Quality; Quality in Use; Quality Measures; ISO/IEC 25022*

1. Introduction

Software quality-in-use is a key factor in the success of any software from the users' perspective based on their preferences. With the large amount of software published online and the development of mobile applications, users need to find the most suitable software to fulfill their needs. The concept of software quality has evolved, initially defined as the adherence to a standard or specification. According to www.statista.com, revenue in the Software market is projected to reach US$698.80bn by the end of 2024, and enterprise Software dominates the market with a projected market volume of US$292bn in 2024. Revenue is expected to show an annual growth rate (CAGR 2024-2028) of 5.27%.

In 1991, the International Organization for Standardization (ISO) introduced the standard for evaluating software quality ISO 9126. This standard was replaced by the ISO/IEC 25022:2016, focusing on quality-in-use within the ISO/IEC 25000:2014 SQuaRE series of standards. These revisions aim to enhance the effectiveness and relevance of quality assessment frameworks in software evaluation processes.

This research investigated the ISO/IEC 25022:2016 standard and identified implementation challenges in measuring software quality in-use. It proposed an Implementation guide to overcome these challenges and provided recommendations for future standard revisions. The following section presents a review of the literature on software quality models; this review aims to investigate how software quality models have evolved significantly in measuring quality. Section 3 presents a discussion highlighting the implementation challenges in implementing the standard. Conclusion and future research are provided in Section 4.

2. Literature Review

Since the pioneering work of McCall et al. (1977), several models have been developed and used to evaluate software quality and usability. Miguel et al. (2014) provided a comprehensive review of these models with their strengths and deficiencies. The ISO/IEC 25022:2016 standard replaced and expanded upon ISO/IEC TR 9126-4:2004 by offering a more structured and modular approach to

Submitted: April 26, 2024
Revised: June 16, 2024

software quality evaluation shown in Figure 1. The standard defines quality-in-use as "the degree to which specific users can use a product or system to meet their needs and achieve specific goals with effectiveness, efficiency, satisfaction, and freedom from risk in specific contexts of use."

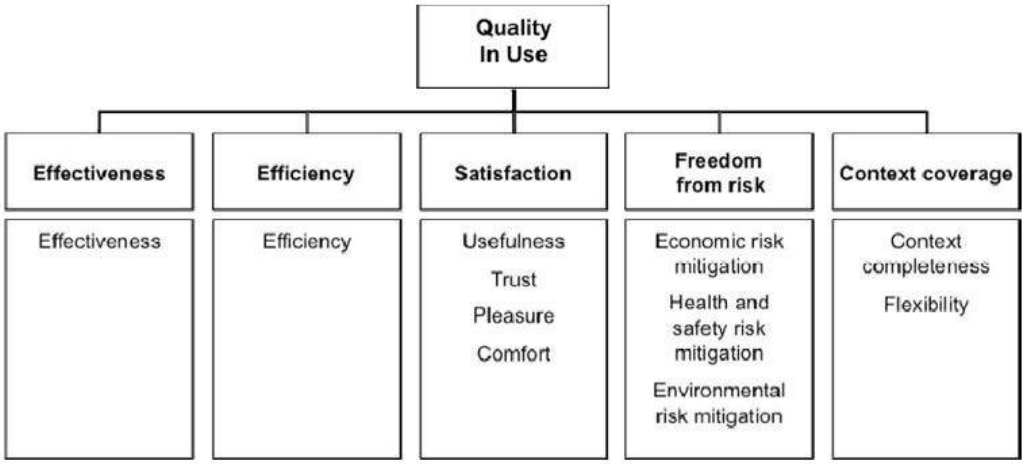

Figure 1. Quality-in-use model

Zhao, et al. (2017) analyzed quality evaluation based on ISO/IEC SQuaRE series standards and its considerations. They pointed out that the Quality in-use model defines five characteristics - effectiveness, efficiency, satisfaction, freedom from risk, and context coverage. These are essential measures in assessing software from the perspective of its users. It is essential to tailor characteristics and sub-characteristics based on project goals, focusing on the most pertinent ones to avoid impracticality.

Gong, Lu, and Cia (2016) indicated that the inconsistency in measurement units for quality measures may pose a challenge. Different measurement units within the same characteristic and sub-characteristic can create difficulties during the evaluation phase. They also stated that the units of the quality measures could be more consistent in both ISO/IEC 25022:2016 and ISO/IEC 25023:2016. They indicated that the measurement results have different variation tendencies.

Zhao et al. (2017) pointed out that the measurement functions in ISO/IEC 25022:2016 are linear functions, such as X=A/B or X=1-A/B. They indicated that nonlinear functions may need to represent the quality measure adequately in some cases. Unlike linear functions that exhibit uniform adjustments with variations in A, nonlinear functions introduce changes in the resulting variable X and offer the advantage of mapping the independent variable into a recognizable range or level, which can be beneficial in certain situations.

Pradanita and Rochimah (2020) completed an assessment based on the ISO/IEC 25022:2016 Standard. Their research aimed at demonstrating the impact of effectiveness, efficiency, satisfaction, and freedom from risk on the overall quality-in-use of a digital wallet product. They concluded that satisfaction demonstrated the strongest correlation (0.939), whereas freedom from risk exhibited the weakest correlation (0.730). They used these results to explain the high level of user satisfaction with the product despite the presence of potential risks. They developed a conceptual framework for explaining observed correlations, making it more straightforward to develop hypotheses.

Robles and Torres (2023) evaluated the quality-in-use of a human talent management information system based on the ISO/IEC 25022:2016 standard. A specialized quality evaluation

model, structured and applied following the standard, was utilized as a practical instrument to identify shortcomings and initiate continual improvements in the organization's processes. They evaluated efficiency, effectiveness, and satisfaction. They calculated the proportion of tasks completed and the time efficiency. The evaluation indicated that the levels achieved for each characteristic were enough to determine how well the system aligns with the user's perspective and measure the overall quality.

3. Implementation Challenges

The ISO/IEC 25022:2016(E) Standard contains 13 tables and provides a total of 36 measures (24 general and 12 special). General measures labeled with the letter G are generally applicable and could be used in a wide range of situations. Special measures are labeled with the letter S and used for specific needs. The format for each table is based on Clause 8 of the standard. It provides the identification code (ID), the name of the quality measure, a description of the information provided, the measurement function, and the method that can be used to obtain its value. Usability has been defined as quality-in-use, focuses on users' perceptions, and refers to the subset composed of effectiveness, efficiency, satisfaction, freedom from risk, and context coverage. The standard avoids subjective assessments by proposing objective evaluation for some quality characteristics. However, general satisfaction, which can reflect users' overall perception and feeling toward the software, is considered an independent measure that does not correlate with the other measures. Surprisingly, the standard does not explicitly encourage organizations to quantify the extent of this correlation.

Furthermore, the standard does not provide a clear distinction between independent and dependent tasks and errors. Without this distinction, it could be challenging to prioritize tasks or understand the impact of errors on the overall software assessment.

Overall, the standard is complex and challenging to understand, especially for individuals and organizations without prior experience in software quality evaluation. In addition, conducting thorough evaluations according to the standard demands substantial resources, including time, expertise, and possibly financial investment. This could be a barrier, particularly for small organizations with limited resources, as they may need help to allocate the necessary resources to adhere to the standard effectively. These authors have identified the following challenges in implementing the standard. These are highlighted in the following subsections.

3.1. Effectiveness Measures

The ISO/IEC 25022:2016 (E) standard defines effectiveness as the "accuracy and completeness with which users achieve specified goals." Table 1, page 11 of the standard contains four general measures and one special measure for effectiveness.

3.1.1. Using correctly as a synonym for accurately may present a challenge in describing Ef-1-G, and Ef-2-S effectiveness in some contexts. For example, in data processing software, if an output of 0.529 represents the accurate outcome, an output of 0.53 may be considered incorrect. In this scenario, a threshold or tolerance may be needed to determine how close the output is to the true value (accurate) to be considered a correct output. For consistency, these two measures should have been described as the proportion of tasks/objectives that are completed accurately.

3.1.2. Accomplishing a task or objective correctly without assistance suggests that the users may not receive any help or assistance. Under Clause 6, page 8, the standard states, "It is important that the users are only given the type of help and assistance that would be available to them in the operational environment." Internal help functions are typically

available to users during normal operations. However, the standard does not distinguish between users utilizing internal help functions and users asking for external assistance from the evaluation administrator or technical personnel. It would be more appropriate to describe these measures as the proportion of tasks/objectives that are accurately completed without external assistance.

3.1.3. While the standard allows for assigning different weights to tasks based on their complexity, the standard does not account for the impact of users' experience in correctly (accurately) completing the tasks. No guidelines were offered to account for the users' skills in calculating the measures. The standard could have suggested a similar weighing scheme to account for users' skills and familiarity with the system.

3.1.4. When counting the number of errors in a task in Ef-3-G, the standard does not distinguish between corrected and uncorrected errors. In situations where the user continues to perform the task using erroneous inputs, it will be difficult to identify the number of errors made. The errors are counted without consideration of their severity or impact. If errors have different severity, a seriousness classification may be applied. Serious errors that can go undetected by the system should be penalized the most. Measure EF-5-G, shown in Figure 3.5 accounts only for the frequency a specific error is made by the users.

3.2. Efficiency Measures

Table 2 page 12 in the standard represents all efficiency measures as recommended in Claus 8 of the ISO/IEC 25022:2016. Efficiency is defined by the "resources expended about the accuracy and completeness with which users achieve goals." This set of measures consists of one general and five special measures.

3.2.1. In measuring the task time Ey-1G, the standard was used "successfully" as a synonym for accuracy in describing the measure. This may represent a challenge in calculating the measure. It would have been more appropriate to describe the measure as the time taken to accurately complete a task.

3.2.2. According to the standard, time efficiency (Ey-2-S), shown in Figure 3.7, Note 1 acknowledged the correlation between efficiency, effectiveness, and task time. This indicates redundancy in measuring user performance. It would have been more appropriate for the standard to specify one measure (Task time) only or encourage users to evaluate the strength of the correlation between these measures and their impact on overall user satisfaction.

3.2.3. Fatigue consequences (Ey-6-S) is one of the efficiency measures shown in Figure 3.8. As indicated, a combination of user performance methods and automated data collection can be utilized. According to Note 1, the measure applies to continuous use by an experienced user only. However, the standard does not suggest a period for the assessment. Some multiple of the task time could have been suggested to facilitate the assessment.

3.3. General Satisfaction Measures

Per the standard, general satisfaction is defined as "the degree to which user needs are satisfied when a product or system is used in a specified context of use." Users' needs include their desires and expectations associated with using a product, system, or service. SUs-1-G in Table 3, page 14 is a highly subjective measure of software quality. The standard ignores the need to report the reliability of questionnaires when evaluating satisfaction. Failure to address the necessity of reporting the reliability is a significant oversight. To rectify this, common methods for reporting

internal consistency (e.g., Cronbach's alpha), test-retest reliability, and inter-rater reliability need to be utilized.

3.4. Economic Risk Mitigation Measures

Economic risk mitigation measures per the standard to "assess the impact of quality on economic objectives related to financial status, efficient operation, commercial property, reputation, or other resources that could be at risk or provide opportunities." This set of measures is listed in Table 9, pages 18 and 19 of the standards. The list includes five general measures and three specials. business analytics is the method proposed for calculating the measures.

3.4.1. In calculating return on investment (ROI) under Rec-1-G, the standard requires the evaluation of "additional benefits". The calculation may result in a negative or positive value. As stated, software users need to identify benefits beyond what is expected. It is common practice in business analytics to calculate ROI based on the net profit generated. To align with established practices, it would be more appropriate for the standard to require the assessment of the total return or net profit resulting from using the software.

3.4.2. In the calculation REc-2-G: the time to achieve a return on investment, the standard relies on the concept of expected return on investment. Note 2 under section 8.5.2, states "the expected values of the measures can be estimated, based on actual values from historical data." From a business perspective, it would have been more appropriate for the standard to adopt measures such as the payback period, time to achieve zero return on investment, or time to break even. This would provide a clearer indication of the investment's profitability and align with common business practices.

3.4.3. The business performance measure (Rec-3-G) requires the evaluation of profitability or sales compared to target. This measure can be very difficult to calculate in large organizations where several software systems are utilized. It is the collective integration of these software products that is expected to have an impact on sales or profitability. It will not be easy to pinpoint the impact of one specific system.

4. Practical Guide

To streamline the implementation of the ISO/IEC 25022:2016 standard, it is important to start by carefully specifying each component of the context of use. This includes users, their goals, and the environment of use. The evaluation should include a sample of representative user groups and tasks. In measuring effectiveness, efficiency, and satisfaction, start by classifying selected tasks into two groups. Independent tasks that cannot be partially completed can be evaluated by following the steps shown in Figure 2. Under these conditions, different weights may be assigned to each task depending on its complexity. Users unable to complete the task should be excluded from the sample. In this proposed implementation guide, it is assumed that all tasks are independent. However, if dependencies exist among tasks, establish an appropriate sequence for evaluating each task separately. This would ensure manageable execution and successful outcomes. Following the completion of each task, the results should be checked for accuracy. Only tasks completed accurately (correctly) without assistance should be included in calculating the number of unique tasks completed (Ef-1-G).

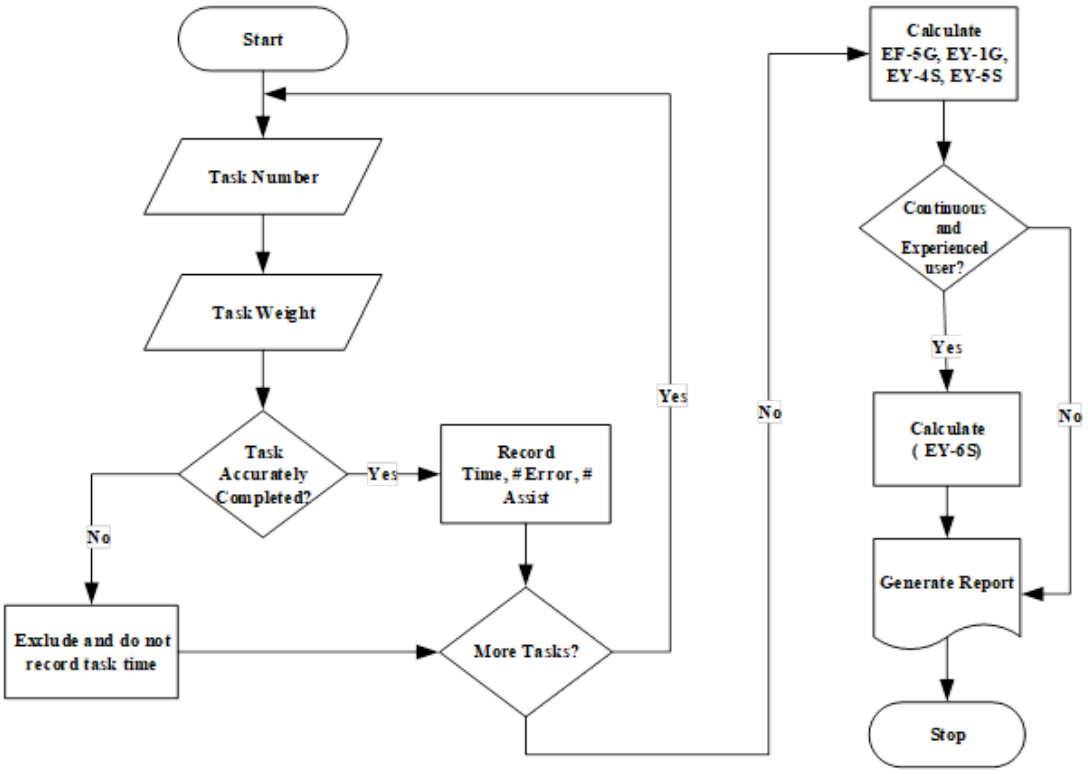

Figure 2. Tasks cannot be partially completed

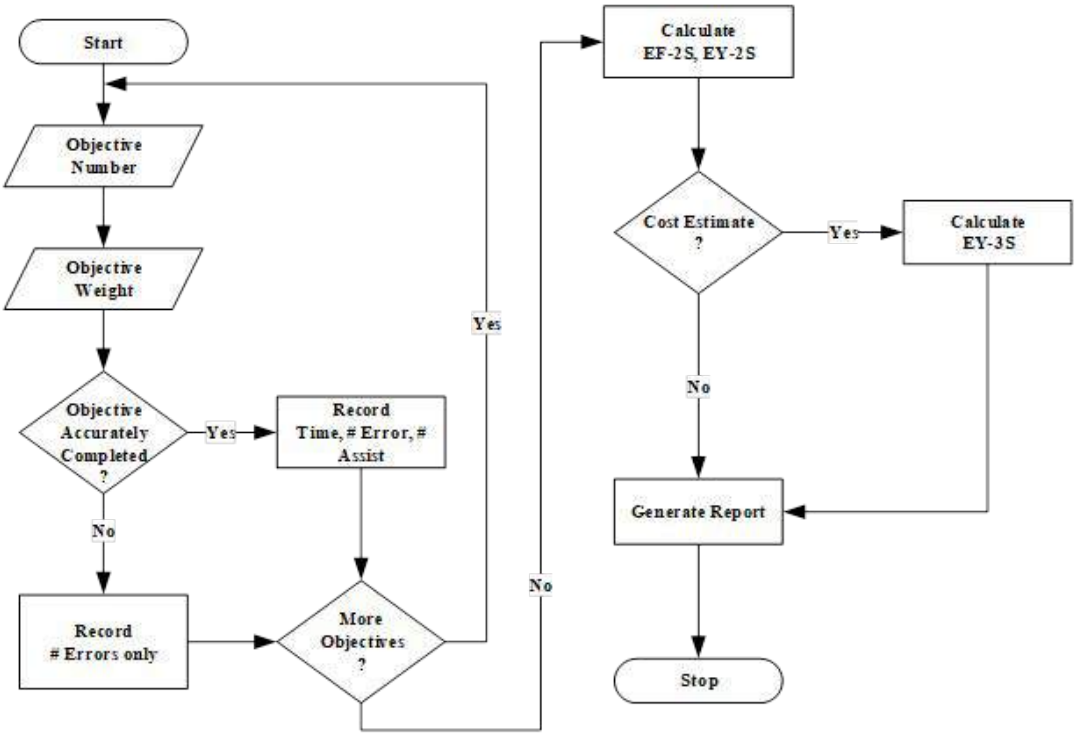

Figure 3. Tasks can be partially completed

Figure 3 presents the steps to follow when selected tasks can be partially completed. Here, each potential missing or incomplete objective is given a weight. Following the successful completion of an objective, the results should be checked for correctness (accuracy). Only objectives completed correctly without assistance should be included in calculating the proportion of objectives achieved (Ef-2-S) and time efficiency (Ey-2-S). For all objectives, the number of errors encountered should be recorded. If cost estimates are available, the cost-effectiveness (Ey-3-S) measure may be obtained. Subsequently, a report capturing measured values can be prepared following the recommended format in the ISO/IEC 25062.

5. Conclusions

This research provided the challenges identified in measuring user performance associated with quality-in-use according to the ISO/IEC 25022:2016 standard. While the standard offers a foundation for evaluating the quality-in-use, the challenges identified in Section 3 need to be addressed by the development committee CS 7/ WG 6 and considered in the upcoming revisions of the standard. These authors are in contact with the CS 7/ WG 6 committee requesting clarification of all the challenges identified and confirmation of the proposed guidelines. With input from the CS 7/ WG 6 committee, these guidelines can be used to develop a user-friendly web application to allow for automated data collection and analysis. This type of research is expected to continue as new versions of the standard are published. Future research may target the incorporation of Artificial Intelligence (AI) in evaluating software quality and analyzing user feedback. This will allow organizations to track trends in user satisfaction and quantify the correlation between the measures.

6. Disclaimer Statements

Funding: None

Conflict of Interest: None

7. References

Gong, J., Lu, J., & Cai, L. (2016). An induction to the development of software quality model standards. In *Proceedings of the third International Conference on Trustworthy Systems and their Applications (TSA)* (pp. 117-122). IEEE. https://doi.org/10.1109/TSA.2016.28

ISO/IEC TR 9126-4. (2004). *Software engineering - Product quality - Part 4: Quality-in-use metrics*. International Organization for Standardization.

ISO/IEC 25000:2014. (2014). *Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE)*. International Organization for Standardization.

ISO/IEC 25022:2016. (2016). *Systems and software engineering — Systems and software quality requirements and evaluation (SQuaRE) — Measurement of quality-in-use*. International Organization for Standardization.

ISO/IEC 25023:2016. (2016). *Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — Measurement of system and software product quality*. International Organization for Standardization.

McCall, J. A., Richards, P. K., Walters, G. F. (1977). *Factors in software quality, Volumes I, II, and III*. US Rome Air Development Center Reports, US Department of Commerce, USA.

Miguel, J. P., Mauricio, D., & Rodriguez, G. D. (2014). A review of software quality models for the evaluation of software products. *International Journal of Software Engineering & Applications (IJSEA), 5*(6), 31-54. https://doi.org/10.5121/ijsea.2014.5603

Pradanita, W. R., & Rochimah, S. (2020). Quality-in-use of digital wallet based on ISO/IEC 25022. In *Proceedings of the 7th International Conference on Electrical Engineering, Computer Sciences, and Informatics (EECSI)* (pp. 282-286). IEEE. https://doi.org/10.23919/EECSI50503.2020.9251300

Zhao, Y. J., Gong, J., Hu, Y., Liu, Z., & Cai, L. (2017). Analysis of quality evaluation based on ISO/IEC SQuaRE

series standards and its considerations. In *Proceedings of the IEEE/ACIS 16th International Conference on Computer and Information Science (ICIS)* (pp. 245-250). IEEE. https://doi.org/10.1109/ICIS.2017.7960001

https://doi.org/10.62704/10057/28084

MATCHING WORKLOADS TO SYSTEMS WITH DEEP REINFORCEMENT LEARNING

Bing Hu[1]
Nicholas Mason[1]
[1]*Intel Corporation*
bing.hu@intel.com; nicholas.mason@intel.com

Abstract

Along with the evolution of computer microarchitecture over the years, the number of dies, cores, and embedded multi-die interconnect bridges has grown. Optimizing the workload running on a central processing unit (CPU) to improve the computer performance has become a challenge. Matching workloads to systems with optimal system configurations to achieve the desired system performance is an open challenge in both academic and industrial research. In this paper, we propose two reinforcement learning (RL) approaches, a deep reinforcement learning (DRL) approach and an evolutionary deep reinforcement learning (EDRL) approach, to find an optimal system configuration for a given computer workload with a system performance objective. The experimental results demonstrate that both approaches can determine the optimal system configuration with the desired performance objective. The comparison studies illustrate that the DRL approach outperforms the standard RL approaches. In the future, these DRL approaches can be leveraged in system performance auto-tuning studies.

Keywords: *Deep Reinforcement Learning; Computer Workload; System Configuration*

1. Introduction

In recent decades, microprocessors have rapidly evolved in terms of both size and complexity. The purpose of this evolution is to achieve a higher performance. Despite many efforts in this area, finding the best way to optimize the system performance remains an open challenge. Optimizing computer system performance is a complicated research problem that involves hardware, software, and hardware-software-co-related factors.

In this study, we attempted to solve a simplified version of the computer system performance optimization challenge by using a single computer workload and various system configurations. In particular, our research problem can be stated as follows: For a given application workload running on a computer system, how can we choose an optimal system configuration to help the system achieve the desired computing performance requirement or objective?

In the literature, researchers have used mapping techniques to solve computer system performance optimization challenges. These studies lead to cost benefits, such as millions of dollars in annual-savings for data centers (Ilyas et al., 2014). Among these studies, there are two main methodologies: static mapping (offline or design-time), and dynamic mapping (online or run-time) (Kundu & Chattopadhyay, 2015; Sahu & Chattopadhyay, 2013). Hybrid application mapping (HAM) was later developed (Pourmohseni et al., 2020) to find a balance between offline and online learning to reduce unnecessary retraining phases. In addition to these classical mapping approaches, reinforcement learning (RL) is another widely adopted mapping technique, especially for network mapping problems (Guerra et al., 2023). These mapping methods work well for problems with small

Submitted: April 28, 2024
Revised: June 15, 2024

datasets that do not require a large amount of computing resources. Therefore, researchers have expanded optimization studies using non-traditional mapping techniques, such as deep learning (DL).

DL is a representation learning tool in the machine learning (ML) toolbox (Goodfellow et al., 2016). Recently, many researchers have adopted DL in workload studies because of DL's promising learning ability in complex problems (Al-Asaly et al., 2022; Wang et al., 2023). Among these DL studies, only a few have explored the system performance optimization problem. For example, Rawat et al. studied a resource optimization problem (Rawat et al., 2020). It applies a genetic algorithm (GA) and artificial neural network techniques to the cloud infrastructure resource provision problem. However, these studies have not yet touched on benchmark workloads.

To fill this gap, we propose a deep reinforcement learning (DRL) and an evolutionary deep reinforcement learning (EDRL) approach for a computer system performance optimization challenge: choosing a system configuration for a given Standard Performance Evaluation Corporation (SPEC) central processing unit (CPU) workload or game workload to achieve an optimal system performance objective. The experimental results indicate that both approaches had an accuracy of approximately 100% in determining the optimal system configuration. Comparison studies illustrate that the DRL approach outperforms state-of-the-art reinforcement learning (RL) approaches, such as the Q-learning method (Guerra et al., 2023).

2. Tools and Techniques

We have three software tools written in Python: The Workload Analytics Platform (WAP), the DRL tool, and the EDRL tool (Figure 1).

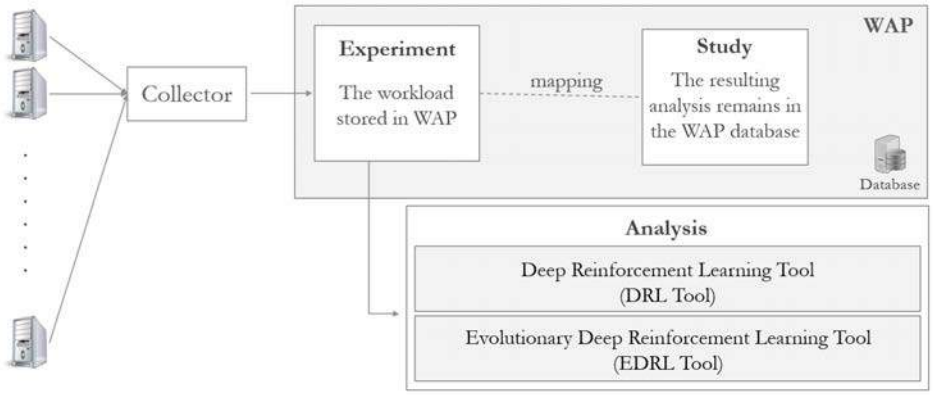

Figure 1. WAP, DRL tool and EDRL tool

2.1 Workload Analytics Platform (WAP)

WAP is a single unified framework for sharing, storing, and analyzing large-scale data in one place. Data collection was performed using Intel teams. The WAP collector stores data and relevant files, such as EMON files (EMON User's Guide,) and logs from Intel teams in the WAP database. The collector stores the data in JavaScript Object Notation (JSON) format. The architecture of WAP is shown in Figure 1. The main role of the WAP is to store and query data by setting a common standard across many teams. Data engineers can use analytical tools to "study" the workload data residing in the WAP. The WAP collects experimental data for the DRL and EDRL tools.

2.2 Deep Reinforcement Learning (DRL) Tools

The DRL tool is an analysis library that leverages data retrieved from WAP. DRL techniques are applied to determine the optimal system configurations for a given set of computer workloads. The DRL tool has an RL agent, a DRL agent, or an EDRL agent (Figure 2). DRL combines deep learning with reinforcement learning. EDRL adds an evolutionary algorithm, such as GA, to the DRL architecture. In this study, the EDRL approach was viewed as an extension of the DRL method.

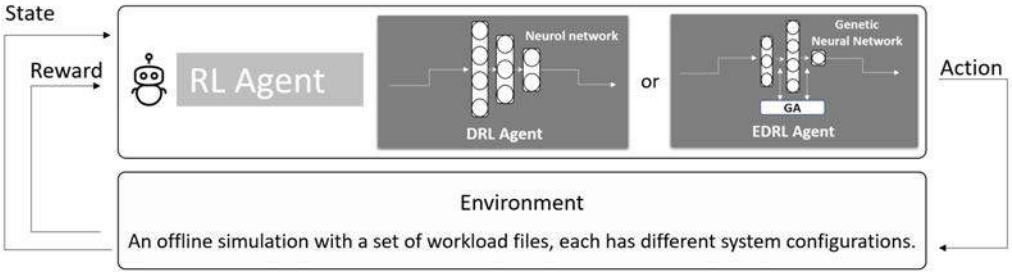

Figure 2. DRL analysis tools

2.2.1 Deep Reinforcement Learning (DRL) Method:

In the DRL tool, an agent uses cross-entropy for the loss function of a neural network (NN). Compared with other RL methods, cross-entropy has good convergence strength (Lapan, 2018). The DRL cross-entropy method aims to train a good generation. We used 80% and 20% of the data for the training and testing, respectively. The DRL cross-entropy method involves the following steps:

Step 1. Run N episodes in one batch iteration, where N is a natural number.

Step 2. Calculate an episode-reward for each episode.

Step 3. Select an "elite" set of episodes from the batch iteration with rewards > 70% of the N episodes.

Step 4. Train an NN on the observations stored in "elite" set. Calculate the cross-entropy loss value by comparing the NN outputs with action-data stored in the "elite" set (called "elite" actions). This step creates NN outputs reinforced by the "elite" actions.

Step 5. Repeat Step1 for another batch iteration until the desired reward threshold is reached.

2.2.2 Deep Reinforcement Learning (DRL) Architecture:

The DRL tool employs a standard RL framework that includes state, action, reward, agent, and environment (Figure 2). Next, we describe each of these concepts separately.

- The state represents the current situation. We used a measurement value or observation of the metrics to represent a state.
- The action is a task in which the RL agent decides to perform the next task based on the state and reward. For our experiments, we had three action options: "Step", "Select", and "Stay". The "Step" action advances the simulation to capture the next observation from the currently running data file. The "Select" action selects another system configuration to run next and ends the current episode of a batch iteration. The "Stay" action runs the current data file until its end. An episode is a sequence of tuple data structures that contains observations, actions, and action rewards.
- The reward provides the agent with information regarding the benefits of the previous action. The reward policy was changeable, based on the objectives of the experiment. Detailed

reward policies for the experiments are described in Section 3.2.

- The agent is an intelligent core in the RL framework. It observes the previous state and current state, and is rewarded based on the state and action. The lifetime of an agent lies within an episode. DRL agents use the neural network (NN) during the learning process.
- The environment can start or restart a SPEC CPU benchmark workload based on the action option that the RL system decides to perform.

2.2.3 Deep Reinforcement Learning (DRL) Agent:

The DRL system contains an agent with a simple NN. The NN has three layers: input, hidden, and output layers. The NN learns to obtain a better reward through repeated actions. The NN is a multilayer perceptron (MLP). The input layer of the NN receives the state inputs. The hidden layer of MLP is a fully connected network. The output layer of the MLP outputs the action to be performed.

2.2.4 Evolutionary Deep Reinforcement Learning (EDRL) Agent:

The EDRL agent contains a genetic neural network (GNN), that combines a GA with an NN. Similar to the DRL component, the NN is an MLP. The difference is that the EDRL agent uses a GA to tune the hyper parameters of the neural network, whereas the DRL agent uses parameters trained in a gradient-based back propagation algorithm. Each generation of GA represents a set of NN parameters. During the GA selection process, only the child with the highest fitness score was directly passed on to the next generation. Hence, the EDRL agent has an auto-turning hyper-parameter capability.

3. Experimental Setup

The experiments were run on a computer with an 11th Gen Intel core i9-11950H processor (@ 2.60 GHz), 128GB of memory, and a 1TB hard drive. No graphics processing unit (GPU) was used during the experiments. Next, we introduce the experimental data and settings.

3.1 Experiment Data

The SPEC CPU benchmark family had multiple workload suites. Among these workloads, SPEC CPU2006 and CPU2017 are the most popular ones (Henning, 2006; SPEC CPU® 2006; SPEC CPU® 2017). Although recognized as a performance benchmarking tool, researchers may forget that these benchmark workloads are valuable datasets. To fill this gap, we used the SPEC CPU benchmark workloads for the experiments. In addition to the SPEC CPU2006 and CPU2017 workloads, we adopted two game workloads (total-war three-kingdom and brigade games) for DRL experiments.

3.1.1. SPEC CPU 2006 and CPU 2017 Workloads: namd, mcf, and povray

Based on Yasin's paper (Yasin et al., 2014), we selected three representative benchmark workloads for our studies: namd [floating-point (fp)], mcf [integer (int)], and povray [fp]. Namd is the core-bound backend workload, mcf is the memory-bound backend workload, and povray is the frontend workload. These three benchmark workloads exist in the SPEC CPU2006 and CPU2017 benchmark suites, respectively.

For the simplicity of the experiments, we have only selected CPU count (CPU), clock frequency (Freq), and cache quality of service (CQoS) three popular system configurations. From these system configurations, we have picked 12 "CPU-Freq-CQoS" system configurations based on SPEC CPU2006 and CPU2017 datasets collected by WAP tool respectively. Further details are provided in Tables 1 and 2.

Table 1. CPU-Freq-CQoS configurations for SPEC CPU2006.

CPU	Freq	CQoS
CPU16	1	CQoS3f
CPU16	1	CQoS7ff
CPU16	2	CQoS3f
CPU16	2	CQoS7ff
CPU32	1	CQoS3f
CPU32	1	CQoS7ff
CPU32	2	CQoS3f
CPU32	2	CQoS7ff
CPU64	1	CQoS3f
CPU64	1	CQoS7ff
CPU64	2	CQoS3f
CPU64	2	CQoS7ff

Table 2. CPU-Freq-CQoS configuration for SPEC CPU2017.

CPU	Freq	CQoS
CPU16	1	CQoS7
CPU16	1	CQoSfffff
CPU16	2	CQoS7
CPU16	2	CQoSfffff
CPU32	1	CQoS7
CPU32	1	CQoSfffff
CPU32	2	CQoS7
CPU32	2	CQoSfffff
CPU64	1	CQoS7
CPU64	1	CQoSfffff
CPU64	2	CQoS7
CPU64	2	CQoSfffff

Table 3. Available system configurations for total war game.

Game mode	Resolution	Quality
battle	720p	low
battle	720p	high
battle	1080p	low
battle	1080p	high
battle	1440p	low
battle	1440p	high
campaign	720p	low
campaign	720p	high
campaign	1080p	low
campaign	1080p	high
campaign	1440p	low
campaign	1440p	high

Table 4. Available system configurations for brigade game.

Graphic API	Resolu-tion	Quality
vulkan	720p	low
vulkan	720p	high
vulkan	080p	low
vulkan	080p	high
vulkan	440p	low
vulkan	440p	high
dx12	720p	low
dx12	720p	high
dx12	1080p	low
dx12	1080p	high
dx12	440p	low
dx12	1440p	high

3.1.2. Two Game Workloads: Total War Three Kingdom Game and Brigade Game

We selected three system configurations (game mode, resolution, and quality) for the three-kingdom game. The available system configuration combinations are listed in Table 3. For the brigade game, we chose three system configurations (graphic API, resolution, and quality). The possible system configuration combinations are presented in Table 4.

3.2 Experiments Settings

We applied the DRL or EDRL model to determine the optimal system configuration (such as a CPU-Freq-CQoS system configuration) for a given category of workloads to obtain the desired system performance objective. We designed three experiments with different optimal performance objectives: shortest execution time, smallest accumulated cycle per instruction (CPI) and smallest accumulated backend metric of top-down microarchitecture analysis metric (TMAM) (in short, we call it as the smallest-accumulated-TMAM-backend objective). TMAM is a subset of EMON metrics (Yasin et al., 2014) designed to determine system performance constraints or bottlenecks (Roodi & Moshovos, 2018). Smaller TMAM measurements indicate a better system performance.

In addition to the performance objective, a reward policy is another setting that must be established. Both DRL and EDRL experiments used the same reward policy. The reward policy has two components: action reward and episode reward. An action reward is based on the actions performed by an agent. The action rewards in an episode are summed as an episode reward. The DRL and EDRL models use an NN to learn important episode rewards in order to make future batch-iteration-level decisions (Figure 3). The smallest-accumulated TMAM-Backend experiment is used as an example. The action-reward policy function of the experiment was as follows:

Reward = 0 when the end of the batch iteration simulation is reached. The episode ends.

Reward = – TMAM metric value, when the end of a batch iteration simulation is not reached, and the action option is "Step."

Reward = – sum of TMAM metric values of the selected new configuration data when the action option is "Select." The episode ends.

Reward = – sum of the remaining TMAM metric values of the current configuration data when the action option is "Stay." The episode ends.

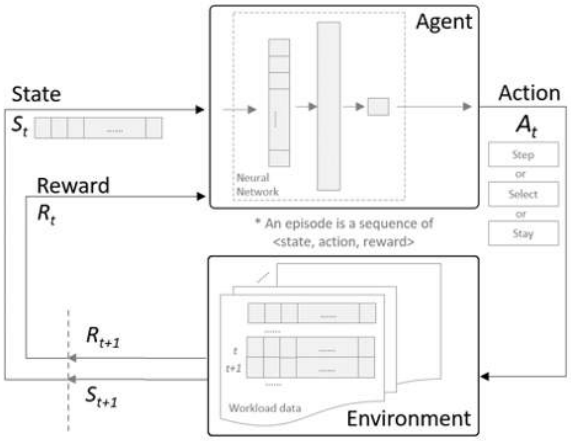

Figure 3. DRL model

4. Experiment Results

4.1 Experiments with SPEC CPU2006 and CPU2017

Both DRL and EDRL models can determine the expected optimal system configuration for the experiment. The optimal system configurations are listed in Tables 5 and 6. The execution times of the DRL and EDRL models are presented in Table 7.

Table 5. Optimal "CPU-Freq-CQoS" system configurations for SPEC CPU2006 workloads

(Optimal system configuration for achieving the experimental performance objective)

Experiment objective \ Workload	namd	mcf	povray
Shortest Execution Time	CPU16-Freq2-CQOS3f	CPU16-Freq2-CQOS7ff	CPU16-Freq2-CQOS7ff
Smallest accumulated CPI	CPU16-Freq2-CQOS3f	CPU16-Freq1-CQOS7ff	CPU16-Freq2-CQOS7ff
Smallest accumulated TMAM Backend	CPU16-Freq2-CQOS7ff	CPU16-Freq2-CQOS7ff	CPU16-Freq2-CQOS7ff

Table 6. Optimal "CPU-Freq-CQoS" system configurations for SPEC CPU2017 workloads

Experiment objective \ Workload	namd	mcf	povray
Shortest Execution Time	CPU16-Freq2-CQoSfffff	CPU16-Freq2-CQoSfffff	CPU16-Freq2-CQoSfffff
Smallest accumulated CPI	CPU16-Freq2-CQoSfffff	CPU16-Freq1-CQoSfffff	CPU16-Freq2-CQoSfffff
Smallest accumulated TMAM Backend	CPU16-Freq2-CQoSfffff	CPU16-Freq1-CQoSfffff	CPU32-Freq2-CQoSfffff

Table 7. Smallest TMAM backend experiment duration time: DRL vs. EDRL

Benchmark \ Workload	DRL execute time (seconds)			EDRL execute time (seconds)		
	namd	mcf	povray	namd	mcf	povray
SPEC CPU2006	5.4	4.2	4.0	29.3	27.8	30.5
SPEC CPU2017	4.1	3.6	3.9	14.8	15.0	15.6

In summary, the DRL and EDRL models determine the optimal system configurations for the SPEC CPU benchmark workload. Both DRL and EDRL run faster with the SPEC CPU2017 dataset than with the SPEC CPU2006 dataset. The size of the dataset could be the reason for this. In general, the SPEC CPU2006 data file size was larger than that of the CPU2017. Another observation is that DRL runs faster than EDRL (Table 7).

4.2 Comparison Studies

We used the median of the reward (reward median) to evaluate the performance of the DRL model using the state-of-the-art reinforcement learning (RL) methods: state action reward state action (SARSA), actor-critic (AC), and Q-learning (Guerra et al., 2023). A higher median reward indicates better performance of the method. A comparison of the results is presented in Table 8. For example, in the experiment to achieve the maximum micro-operations (UOPS) performance objective with Total War workloads, the DRL method had the highest reward median (5.85) compared with the SARSA, Q-learning, and AC methods. The results showed that the DRL model outperformed other RL methods.

Table 8. Comparison studies by using the reward median

Workload / Method	SPEC CPU2006 (namd) (Performance objective: smallest CPI)	SPEC CPU2017 (namd) (Performance objective: smallest CPI)	Total War (Performance objective: max UOPS)	Brigade (Performance objective: max UOPS)
SARSA	-455.50	-40.80	4.60	53.01
Q-Learning	-437.90	-41.50	4.61	53.02
AC	-435.70	-28.50	5.30	51.90
DRL	-388.06	-17.08	5.85	58.34

5. Conclusions and Future Work

We demonstrated that these two DRL tools are valuable for selecting a performance-optimal system configuration for a given set of workloads. Both tools had 100% accuracy in determining the optimal solutions. Compared with state-of-the-art RL methods, such as Q-learning (Guerra et al., 2023), DRL exhibits a better media-reward performance.

We also observed that the DRL method required shorter execution time than the EDRL method. EDRL has a promising accuracy for complex real-life systems, particularly for reinforcement models that contain multiple agents. The GA is an evolutionary algorithm adopted in EDRL because of its high flexibility and ability to self-adapt to search for optimal solutions. However, it is also well known that GA has a longer execution time than non-evolutionary algorithms. Unsurprisingly, the EDRL method runs slower than the DRL method does. The size of the data files also explains why the EDRL method did not perform well. The EMON data file has more than 300 metrics, with thousands of rows of observations. For this type of large dataset and only one RL agent in the experiments, DRL is a better choice for determining the optimal system configuration for the SPEC CPU benchmark workloads. In the future, we will explore the methods to improve the performance of the GA algorithm. In the meantime, we will research other evolutionary algorithms that run faster than GA for better EDRL performance with large datasets, such as the generalized reduced gradient algorithm.

Currently, the DRL tool has been used on SPEC CPU benchmark workloads and game workloads; it has not yet been extended to other large-scale workloads, such as security workloads. We foresee a potential opportunity to study more complex workloads using DRL technology. The DRL tool extends the Intel Workload Analytic Platform (WAP) capability to analyze complex workloads and provides insights from a machine learning perspective. In the future, the DRL tool can be migrated to a portable stand-alone library to enrich the Intel software community.

6. Acknowledgments

The authors thank Dr. Karl Kempf for his insightful comments and unwavering support of the foundational research required for this study. We also thank the Intel decision engineering (DE) team, particularly Harry Travis, for their leadership. We thank Andy Anderson and Anirudha Rahatekar for providing valuable data and expertise throughout the lifetime of this project; Thong Duong for supporting this study; and Joyce Weiner for reviewing this paper and providing valuable feedback.

7. Disclaimer Statements

Funding: None

Conflict of Interest: None

8. References

Al-Asaly, M. S., Bencherif, M. A., Alsanad, A., & Hassan, M. M. (2022). A deep learning-based resource usage prediction model for resource provisioning in an autonomic cloud computing environment. *Neural Computing and Applications, 34*(13), 10211–10228. https://doi.org/10.1007/s00521-021-06665-5

Intel. (2023, July 3). *EMON User's Guide.* https://www.intel.com/content/www/us/en/content-details/686077/emon-user-s-guide.html

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning.* The MIT Press.

Guerra, A., Guidi, F., Dardari, D., & Djuric, P. M. (2023). Reinforcement Learning for Joint Detection & Mapping using Dynamic UAV Networks. *IEEE Transactions on Aerospace and Electronic Systems*, 1–16. https://doi.org/10.1109/TAES.2023.3300813

Henning, J. L. (2006). SPEC CPU2006 Benchmark Descriptions. *ACM SIGARCH Computer Architecture News.*

Ilyas, M. S., Raza, S., Chen, C.-C., Uzmi, Z. A., & Chuah, C.-N. (2014). RED-BL: Evaluating dynamic workload relocation for data center networks. *Computer Networks*, *72*, 140–155. https://doi.org/10.1016/j.comnet.2014.07.001

Kundu, S., & Chattopadhyay, S. (2015). Application Mapping on Network-on-Chip. In *Network-on-Chip* (pp. 119–154). CRC Press.

Lapan, M. (2018). *Deep Reinforcement Learning Hands-On*. Packt Publishing.

Pourmohseni, B., Glaß, M., Henkel, J., Khdr, H., Rapp, M., Richthammer, V., Schwarzer, T., Smirnov, F., Spieck, J., Teich, J., Weichslgartner, A., & Wildermann, S. (2020). Hybrid Application Mapping for Composable Many-Core Systems: Overview and Future Perspective. *Journal of Low Power Electronics and Applications, 10*(4), 38. https://doi.org/10.3390/jlpea10040038

Rawat, P. S., Gupta, P., Dimri, P., & Saroha, G. P. (2020). Power efficient resource provisioning for cloud infrastructure using bio-inspired artificial neural network model. *Sustainable Computing: Informatics and Systems*, *28*, 100431. https://doi.org/10.1016/j.suscom.2020.100431

Roodi, M., & Moshovos, A. (2018). Gene Sequencing: Where Time Goes. *2018 IEEE International Symposium on Workload Characterization (IISWC)*, 84–85. https://doi.org/10.1109/IISWC.2018.8573474

Sahu, P. K., & Chattopadhyay, S. (2013). A survey on application mapping strategies for Network-on-Chip design. *Journal of Systems Architecture, 59*(1), 60–76. https://doi.org/10.1016/j.sysarc.2012.10.004

SPEC. (2022, June 16). *SPEC CPU® 2006*. https://www.spec.org/cpu2006

SPEC. (2022, June 15). *SPEC CPU® 2017*. https://www.spec.org/cpu2017

Wang, T., Fan, X., Cheng, K., Du, X., Cai, H., & Wang, Y. (2023). Parameterized deep reinforcement learning with hybrid action space for energy efficient data center networks. *Computer Networks*, *235,* 109989. https://doi.org/10.1016/j.comnet.2023.109989

Yasin, A., Ben-Asher, Y., & Mendelson, A. (2014). Deep-dive analysis of the data analytics workload in CloudSuite. In, *2014 IEEE International Symposium on Workload Characterization (IISWC)* (pp.202–211). https://doi.org/10.1109/IISWC.2014.6983059

SUMMER
2024

AIEMS VOL 17 NO 1

AIEMS

AIEMS is committed to furthering the integration between business management, engineering, & industry. Our objective is to promote research, collaboration, & practice in these multidisciplinary areas. AIEMS seeks to encourage local, national, & international communication & networking via conferences & publications open to those in both academia & industry. We strive to advance professional interaction & lifelong learning via human & technological resources, and to influence and promote the recruitment and retention of young faculty and industralists.