Learning Machine Learning:
A Comparative Study of Industrial Design and Computer Science and Engineering Students

Exploring the Role of Mathematics Backgrounds in Foundational ML Education

Beopgi Jo[1]

Supervisors: Gosia Migut[1]**, Ilinca Rențea**[1]

[1]EEMCS, Delft University of Technology, The Netherlands

A Thesis Submitted to EEMCS Faculty Delft University of Technology,
In Partial Fulfilment of the Requirements
For the Bachelor of Computer Science and Engineering
January 25, 2025

Name of the student: Beopgi Jo
Final project course: CSE3000 Research Project
Thesis committee: Gosia Migut, Ilinca Rențea, Jesse Krijthe

Abstract

Machine learning (ML) has become a critical skill across various disciplines, yet teaching it to students outside Computer Science and Engineering (CS) remains challenging due to differing academic backgrounds. This study investigates the differences in learning outcomes between Industrial Design (ID) and CS students when introduced to foundational ML topics, focusing on the influence of prior mathematical knowledge.

Through initial surveys on mathematical proficiency, structured ML tutorials, and final assessments on learning outcomes, the research examines correlations between mathematical proficiency and ML performance, faculty-specific challenges, and qualitative feedback on learning experiences. Results reveal that prior mathematics knowledge significantly impacts performance on mathematics-intensive topics such as Bayes' Rule, while its influence is minimal on less math-relevant topics like ML pipelines. Furthermore, ID students emphasized creative and interactive teaching methods, contrasting with the programming-oriented preferences of CS students.

These findings highlight the need for interdisciplinary instructional strategies that cater to diverse learner strengths. By uncovering faculty-specific patterns in ML learning, this study contributes to the design of more inclusive and effective educational practices, fostering a broader understanding and application of ML across disciplines.

1 Introduction

Machine learning (ML) has emerged as a key component of contemporary technology, enabling applications ranging from self-driving automobiles to voice assistants. Its pervasive influence means that people from diverse academic and professional backgrounds increasingly encounter ML, regardless of their technical expertise. Recognizing this, many universities now integrate ML education into interdisciplinary programs, emphasizing its relevance beyond computer science and engineering (CS). Combined ML degrees, such as Data Science for Business, Creative Practices, and other interdisciplinary programs, reflect this growing trend, offering ML education to students from fields such as business, social sciences, and design. [27][11] Industrial design (ID) is one of these fields which stands out for having a more design-based, visual, and user-centric point of view in problem solving. [12] In contrast to computer science (CS), which places a strong emphasis on algorithmic problem solving, ID fosters conceptual and visual skillsets that may offer alternative methods for comprehending and using machine learning. This discrepancy presents an opportunity to investigate how ML might be taught to ID students in order to provide them applicable abilities and to find new ways where ML may improve human experiences through design.

A recent study on the use of artificial intelligence (AI) in industrial design states that although AI can be used in various areas of industrial design from product design to development and manufacturing, its role has not yet been extensively explored. [24]. As there are increasing number of examples where ML has worked as an important factor in solving technical challenges faced in the fourth industrial revolution [14], the demand for higher-level education on this topic for practitioners in related fields is increasing. This is also true for ID students as they explore user-friendly solutions in various contexts, where ML could play an important role. [25] However, although recent studies have addressed ML education for non-CS students, most have focused on STEM disciplines rather than design-driven fields. For instance, Cheong examined how recommendation algorithms could guide non-CS students in course selection, demonstrating that structured pathways can mitigate the cognitive challenges of technical prerequisites. [9] Similarly, Banadaki explored how supervised research experiences could improve engagement and skill acquisition for non-CS students. [2] However, these strategies have been tested among broad student populations without specific focus on ID students, who may leverage their visual and conceptual strengths in learning ML. This gap in the literature suggests the need to investigate how ID students, with their distinct skillsets, approach ML compared to CS students, who often have stronger foundations in programming and mathematics. This research addresses two centric questions: **What are the differences in learning outcomes between industrial design and computer science students when introduced to foundational machine learning topics? How are these outcomes influenced by prior mathematics knowledge?** To answer these, I conducted a comparative study, guiding ID and CS students through structured ML tutorials and assessing their performances. By analyzing the learning outcomes of these two groups, I aim to evaluate whether a student's mathematics background influences their ability to successfully learn foundational machine learning topics and uncover any patterns that are present across different faculties.

This paper will be presented in the following structure. Chapter 2 provides the background of the research with reviews of relevant literature, contextualizing the study within existing research. Chapter 3 outlines the research methodology, including participant recruitment, tutorial design, and data collection and analysis methods. Chapter 4 reflects on the ethical aspects of the research and discusses the reproducibility of the methods. Chapter 5 presents the results, focusing on comparative analysis of learning outcomes between ID and CS students. Chapter 6 discusses the implications of these findings for interdisciplinary ML education by providing interpretations of the results and comparisons with other literature, and Chapter 7 concludes the paper with key insights and recommendations for future research.

2 Background of Research

Machine Learning and Industrial Design

With increasing importance of ML, universities are increasingly embedding ML courses in diverse curricula to prepare students for the challenges of a data-driven world. [25] However, teaching ML effectively to students from non-CS disciplines presents significant challenges. Many of these students lack the mathematical and programming foundations typically assumed in ML education. [13]

At TU Delft, these challenges are amplified by the distinct skillsets cultivated between faculties. Specifically, with increasing popularity and usage in generative AI, changes have been made to existing course curriculum and new courses were created in ID faculty in order to provide insights on this topic. [25] However, ID students excel in visual and conceptual thinking, user-centered design, and creative problem solving, [12] and these strengths stand in contrast to the mathematically rigorous and algorithmic problem solving approaches emphasized in CS education. This divergence in skillsets raises compelling questions about how ML education should adapt to accommodate such diverse learning needs.

Working as an ID teaching assistant, I observed how ID students approached technical topics differently compared to CS students. Although ID students demonstrated ingenuity in conceptual and user-focused tasks, they often struggled with algorithmic abstractions and programming. In contrast, CS students excelled in these areas, but tended to approach tasks with less emphasis on user experience or contextual design. These experiences motivated this research to investigate how such differences influence ML learning outcomes and to develop teaching strategies that address these interdisciplinary challenges.

Recent research supports these observations, emphasizing the higher-order cognitive challenges in ML education for non-majors. Sulmont et al. interviewed instructors of ML courses for non-majors and found that while algorithms are often considered the difficult part of ML, instructors actually found that higher-level design decisions and model comparisons — tasks aligned with the upper levels of the SOLO taxonomy (Structure of Observed Learning Outcomes taxonomy)[5] — were the most challenging to teach. Conversely, procedural knowledge, such as following steps aligned with lower-SOLO taxonomy levels, was described as easier to convey. This distinction provides a framework for understanding the difficulties faced by students, particularly those in design-focused disciplines, and highlights the need for courses to emphasize conceptual and decision-making skills to address these challenges effectively. [22]

Research Themes in Literature

The literature on ML education provides a foundation for addressing the challenges of teaching ML to diverse student populations, offering valuable insights into interdisciplinary learning and pedagogical design. Cheong highlighted the importance of structured learning pathways to reduce cognitive load for non-CS students. By providing step-by-step guidance, such pathways help bridge gaps in technical understanding, enabling students from diverse backgrounds to grasp foundational ML concepts effectively. [9] Similarly, Ko underscored the prevalence of misconceptions in ML learning, such as the over-reliance on default algorithm configurations. This research emphasized the need for foundational clarity in ML education to correct these misconceptions and enhance understanding. [13]

Bloom's Taxonomy, as revised by Anderson and Krathwohl, provides a hierarchical framework for designing learning objectives across cognitive levels, making it particularly valuable for interdisciplinary education. By starting with foundational knowledge and progressively building toward complex problem-solving, educators can accommodate the diverse strengths of ID and CS students. In this research, Bloom's Taxonomy was used to design tutorials and assessments that align with the varied learning needs of these student groups, ensuring a clear progression from basic to advanced ML concepts. [1]

Interdisciplinary approaches to learning have also been shown to be instrumental in ML education. Banadaki demonstrated that interactive and visual teaching methods significantly improve comprehension for non-STEM students, showcasing the benefits of visual aids and real-world analogies in enhancing understanding. [2] Tokuta et al. emphasized the importance of designing educational frameworks that resonate with non-CS audiences, such as those in data science. [23] These strategies, which incorporate relatable examples and foster conceptual understanding, align with the design-focused nature of ID students and offer pathways to efficient learning. Together, these studies underscore the value of interdisciplinary approaches to ML education, providing a foundation for this research.

Connecting to the Research Question

This research seeks to build on these themes by investigating how differences in faculty-specific skillsets and prior mathematical knowledge affect learning outcomes in foundational ML topics. The central research questions are:

What are the differences in learning outcomes between industrial design and computer science students when introduced to foundational machine learning topics? How are these outcomes influenced by prior mathematics knowledge?

To answer these questions, the research addresses four subquestions:

1. How do industrial design and computer science students differ in their prior knowledge in mathematics?
2. How does prior proficiency in mathematics correlate with performance on foundational ML topics?
3. How do students from these faculties perform on ML topics with varying levels of relevance in mathematics?
4. What qualitative patterns emerge in the challenges students face while learning ML?

By aligning tutorial design with Bloom's Taxonomy and incorporating interdisciplinary teaching methods, this research aims to uncover actionable insights into ML education, contributing to more inclusive and effective strategies for diverse student populations.

3 Methodology and Rationale

This study employed a multi-step design to explore the impact of prior mathematics knowledge and faculty-specific factors on learning foundational ML concepts. Participants from ID and CS faculties at TU Delft were recruited, emphasizing voluntary and anonymous participation. All data from participants who followed the full procedure - 10 and 13 from ID and CS respectively - were used. This section introduces each step that was taken to collect these data and draw conclusions from them.

3.1 Study Design

The research consisted of three main sections: an initial mathematics survey, tutorials, and assessments. This structure was inspired from other studies that used a methodology involving pre-testing one variable, followed by an intervention, then post-testing a second variable, to explore correlations or outcomes. [6] [21] The materials used in these steps can be found in the Github repository[1].

Initial Mathematics Survey

The initial survey[2] assessed participants' prior mathematics experiences, including coursework, extracurricular activities, and self-rated confidence in and mathematics. Participants' mathematic proficiency was also measured with mathematics questions. These questions, aligned with ML tutorial topics, spanned three domains — calculus, probability, and linear algebra - and each question addressed one specific domain. The questions also spanned varying levels of difficulty, categorized as easy, medium, and hard, and each of these questions contributed different number of points towards the total score. This approach enabled nuanced classifications of participants' mathematics proficiency and provided a baseline for correlating performance with learning outcomes. [8]

This initial survey serves as more than a tool for classifying participants by their faculties; it enables a nuanced exploration of the influence of prior mathematics knowledge. By collecting detailed information on participants' academic and extracurricular mathematics experiences, the survey allows for a double-layered analysis. Firstly, faculty-based insights can be gained, comparing learning outcomes by grouping students into their respective disciplines. Secondly, mathematics-based insights can also be gained, identifying correlations between participants' performance on the final assessment and their specific levels of mathematics proficiency. This dual classification allows for more detailed observations. For example, if a student excels in a particular ML topic, the analysis can explore whether this success stems from their faculty's general approach to problem-solving or their individual mathematics background. Furthermore, it allows for testing whether mathematics proficiency significantly affects ML learning outcomes. If students with similar levels of mathematics knowledge from different faculties show varying performances, this could highlight faculty-specific factors influencing ML learning.

ML Tutorials

Three tutorials hosted on a Notion website introduced key ML topics:

1. **Machine Learning Pipelines**[3]: Covered conceptual topics such as data preparation, model training, and evaluation.
2. **Bayes' Rule**[4]: Introduced probabilistic reasoning and classification concepts.
3. **Perceptrons**[5]: Explained basics of multilayered perceptrons and training processes.

These tutorials were designed to incorporate varying levels of mathematics relevance, ensuring the assessment could differentiate between participants' familiarity with and ability to grasp mathematics-heavy and mathematics-light topics. Machine Learning Pipelines was included as a topic with the least mathematics relevance, emphasizing conceptual understanding and process logic. Bayes Rule, on the other hand, required mathematical reasoning and probabilistic thinking, making it a strong indicator of mathematics proficiency. Perceptron topic was included as a topic with a moderate level of mathematics relevance; while still rooted in mathematical and algorithmic concepts, this tutorial balanced theoretical explanations with practical applications.

This progression allowed for a detailed examination of how participants performed across topics with varying demands on mathematics knowledge. Coupled with questions that tested these topics, this design could reveal topic-specific strengths - faculties or groups excelling in particular ML topics - and sensitivity to mathematics relevance - whether students perform better on topics aligned with their prior mathematics training.

These tutorials were developed following the constructive alignment, where the materials for learning and assessments are constructed around the learning objectives. [3] For each topic, intended learning objectives (ILO's) were selected first following Bloom's taxonomy. The tutorials were then built to accommodate these ILO's, making sure that students were provided with the appropriate level and amount of information to achieve them. [4] The tutorials included ILO's in the beginning, making sure that students are aware of the contents that the tutorials have and what they should focus on, aiming for a more engaging and effective learning process. [26] The ILO's for each topic and how they were addressed in the tutorials can also be found in Appendix A.

Assessments

After learning the tutorials, participants took an assessment[6] to evaluate their learning outcomes with questions on the learned ML topics. According to the constructive alignment, the questions are formulated around the ILO's in order to synchronize learning and assessment. [15] This can also be found in Appendix A.

[1] github.com/JustinBJo/CSE3000/
[2] github.com/JustinBJo/CSE3000/blob/main/forms/initial_survey.md
[3] github.com/JustinBJo/CSE3000/blob/main/tutorials/tutorial1_ml_pipeline.md
[4] github.com/JustinBJo/CSE3000/blob/main/tutorials/tutorial2_bayes_rule.md
[5] github.com/JustinBJo/CSE3000/blob/main/tutorials/tutorial3_perceptrons.md
[6] github.com/JustinBJo/CSE3000/blob/main/forms/final_assessment.md

The questions in the final assessment were designed to feature questions of varying difficulty levels (foundational and critical), allowing for a more nuanced assessment of learning outcomes. Foundational questions tested basic recall and understanding of tutorial content, while critical questions assessed the ability to apply learned concepts in complex scenarios. Similar to the initial mathematics survey, the number of points that can be gained from a question differed by their difficulty level. This tiered approach enabled a granular evaluation of participants' learning. [8]

In addition to the quantitative data, qualitative feedback gathered from participants provided valuable insights into their learning experiences. This feedback revealed patterns in the challenges faced by students from different faculties or with varying levels of mathematics proficiency, offering a deeper understanding of the factors influencing their learning outcomes. Participants were also asked to share their perceptions of the tutorials and assessment questions, indicating whether these materials were accessible, engaging, or overly challenging. Furthermore, the feedback highlighted specific areas where students struggled the most, offering actionable insights for improving ML education strategies. [17]

The thematic analysis of this qualitative data supports the broader goals of fostering inclusivity and effectiveness in teaching ML concepts. [7] By identifying barriers and areas for enhancement, the findings align with calls for pedagogical content knowledge in ML education. [13]

3.2 Data Analysis

In this section, the methods used to analyze the data and address the research questions will be explained. The analysis was based on both quantitative and qualitative approaches, employing various statistical tests and methods to answer each of the sub-questions. The chosen methods were implemented with Python[7].

Sub-question 1: Comparison of Initial Mathematics Scores

In order to answer the first sub-question "How do industrial design and computer science students differ in their prior knowledge in mathematics?", the initial mathematics scores of ID students and CS students were compared using the **Mann-Whitney U test**. [16] This non-parametric test was selected because it does not require assumptions of normality, making it a more appropriate choice given the small size of the sample. The **p-value** from the test helps determine whether there is a statistically significant difference between the two groups in terms of their initial mathematics knowledge. In addition, number of mathematics-related courses the students took and activities related to mathematics were counted to compare the students' experience with mathematics. [10]

Sub-question 2: Correlation Analysis

The second sub-question "How does prior proficiency in mathematics correlate with performance on foundational ML topics?" was addressed by investigating the correlation between initial survey scores and performance across different machine learning topics. This was done by using a **Pearson correlation** analysis. **The Pearson correlation coefficient** measures the strength and direction of the linear relationship between two continuous variables. The results of the analysis help identify how strongly the initial mathematics scores correlate with the final performance on the machine learning topics. [20]

Sub-question 3: Performance across ML Topics

To answer the third sub-question "How do students from these faculties perform on ML topics with varying levels of relevance in mathematics?", the performance of ID students and CS students across the three machine learning topics were compared. Since the data from both groups were not normally distributed, the **Kruskal-Wallis H test** was chosen for this analysis. This method is designed to compare the distributions of more than two groups; although there are only two faculties, the Kruskal-Wallis test is appropriate for comparing performance across multiple topics in each group. This evaluates whether there are significant differences in the score distribution between the topics for both faculties. [18]

Sub-question 4: Qualitative Patterns

To address the fourth sub-question "What qualitative patterns emerge in the challenges students face while learning ML?", a thematic analysis was conducted on the qualitative responses provided by the students. The analysis involved coding the data by identifying key themes and categorizing them into broader topics. These themes were then analyzed to understand students' experiences, challenges, and the effectiveness of different learning materials. [19] [7]

4 Responsible Research

This research adheres to the principles of responsible research, prioritizing ethical considerations and ensuring the reproducibility of methods. In this section, the key aspects of responsible research and the steps taken during the study to align with these principles are reflected.

4.1 Dealing with Research Data

Each step of the research was conducted with a commitment to ethical principles. Below is the outline of potential ethical issues that could arise and the steps taken to address them.

Data Accuracy and Integrity

All data were handled with care to maintain integrity. Responses were recorded exactly as submitted, and no alterations were made to the data. A robust data storage protocol - Microsoft OneDrive - ensured that raw data were securely stored in their original form for verification purposes.

Participant Anonymity and Confidentiality

Participant anonymity was maintained by using participant-generated codes rather than personal identifiers. This approach allowed for linking pre and post-tutorial responses without compromising confidentiality. No identifiable information was collected, and all data were securely stored in password-protected systems accessible only to the researcher.

[7] github.com/JustinBJo/CSE3000/blob/main/analysis/analysis.py

Informed Consent
Participants were fully informed about the study's purpose, their rights, and how their data would be used. They were provided with a detailed consent form outlining the goals of the study, procedures, and data usage. They were informed of their right to withdraw at any stage without penalty. Participation was entirely voluntary.

Reproducibility
The methodology was thoroughly documented, including the design and structure of the survey, tutorials, and assessment. This detailed documentation ensures that the study can be replicated under similar conditions, promoting transparency and reproducibility.

Avoiding Bias
In order to avoid influences of personal relationships in the outcome, participants were recruited from a pool of individuals unknown to the researcher, and no direct contact was made during or after participation. Recruitment was conducted via neutral methods, such as visiting faculty areas and distributing online links.

Psychological Well-being
The assessments and surveys were designed to be straightforward and non-intimidating. Participants were informed that their performance would not be judged or shared and only the aggregated results would be used, ensuring a relaxed and supportive environment.

4.2 Contributions

Many materials prepared for the research were created collaboratively with two other co-researchers: Junwon Yoon and Oisín Hageman. These researchers conducted similar researches on different faculties, and the initial survey, tutorials, and final assessment were formulated in collaboration and shared. The final results were shared, but the processing and analyses were done individually.

Use of Generative Artificial Intelligence
Throughout the scope of the research, generative AI ChatGPT was used in areas where the efficiency could be improved. It was used for a more professional sentence structuring in writing tutorials and the report and to check for any spelling or grammar mistakes, but not for generating contents. It was also used to obtain support in finding the suitable libraries for data analysis and learning how to use them. The prompts used for these can be found in Appendix B. In addition, automated writing and proofreading AI Writefull was also used in writing report for checking spelling and grammar mistakes. Any areas in the research that require originality and creativity, such as generating ideas for methodology, writing the report, and analyzing the results, were done manually without any support from artificial intelligence.

5 Results and Analyses

In this section, the results of the experiments are presented. The analysis covers four sub-questions, and each is addressed using relevant statistical tests and descriptive methods. The figures included in this section can also be found in Appendix C in a larger scale.

5.1 Sub-question 1: Initial Mathematics Scores Comparison

The descriptive statistics for the initial mathematics scores of both groups are summarized in Table 1, and visualized in the scatterplot in Figure1. The comparison between the two groups is further illustrated in the boxplot in Figure 2.

Group	Mean	Standard Deviation
ID	35.10	28.38
CS	43.08	16.04

Table 1: Descriptive Statistics for Initial Mathematics Scores

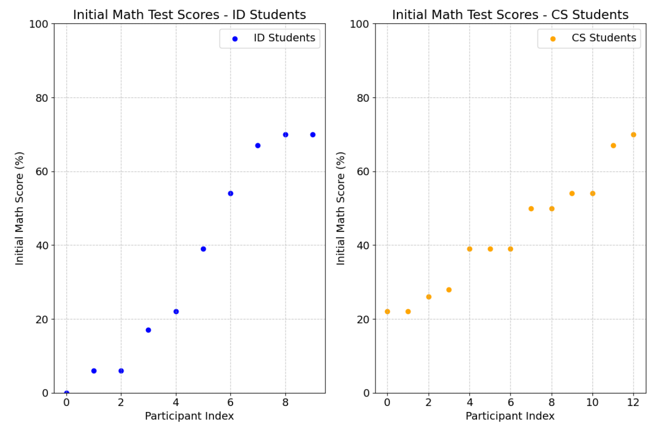

Figure 1: Scatterplot of Initial Mathematics Scores, Sorted by Scores

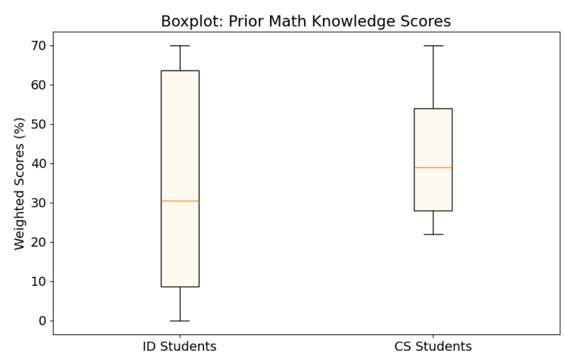

Figure 2: Boxplot of Initial Mathematics Scores

In addition, the total counts of previous mathematics-related courses or activities for both faculty students are recorded. These can be found in Table 2.

To compare the initial mathematics scores between ID and CS students, a **Mann-Whitney U test** was conducted. The results of the Mann-Whitney U test yielded a **U-statistic** of **53.00** and a **p-value** of **0.4730**. The p-value greater than conventional threshold p-value 0.05 suggests that there was no statistically significant difference in the initial mathematics scores between ID and CS students. [16]

Group	Count
ID	12
CS	28

Table 2: Number of Mathematics-Related Courses or Activities

5.2 Sub-question 2: Correlation Analysis

The correlation between the initial mathematics scores and performance on three topics was assessed for all students. These results were visualized on scatterplots, which can be found in Figures 3, 4, and 5, to find patterns between the two factors. The results of the **Pearson correlation analysis** are presented in Table 3.

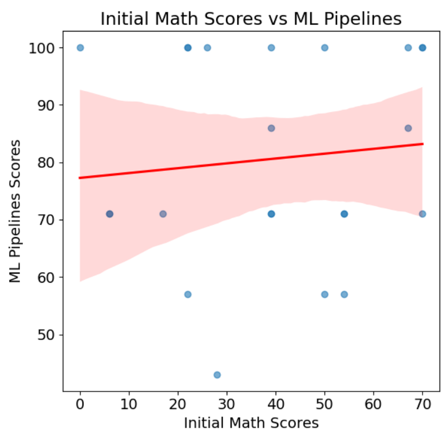

Figure 3: Scatterplot of Initial Mathematics Score to ML Pipeline Scores

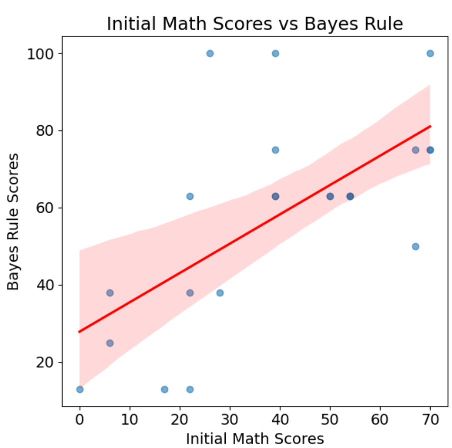

Figure 4: Scatterplot of Initial Mathematics Score to Bayes' Rule Scores

From the analysis, it was found that there is a strong, statistically significant positive correlation between the initial mathematics scores and performance on the Bayes' Rule topic with the value of r greater than 0.5 and p significantly smaller than 0.05. In contrast, there was no significant correlation with either the ML Pipelines ($r = 0.10$, $p = 0.6422$) or Perceptrons ($r = 0.03$, $p = 0.8869$) topics. [20]

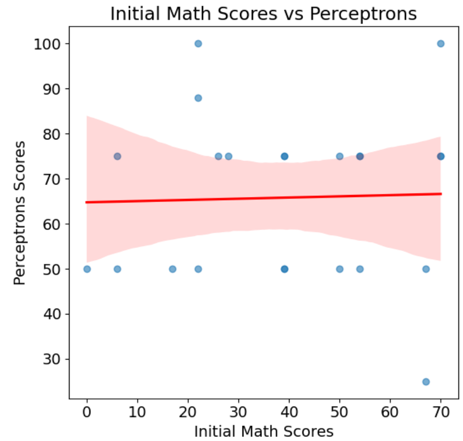

Figure 5: Scatterplot of Initial Mathematics Score to Perceptrons Scores

Topic	Pearson Coefficient (r)	P-value
ML Pipelines	0.10	0.6422
Bayes' Rule	0.64	0.0010
Perceptrons	0.03	0.8869

Table 3: Correlation Analysis between Initial Scores and Topic Performance

5.3 Sub-question 3: Faculty Performance on ML Topics

The performance of faculty members in learning the three ML topics was assessed using the **Kruskal-Wallis H test**. First, the results of each topic for both faculty students are summarized in Table 4, and visualized in Figure 6.

Group	ML Pipelines		Bayes' Rule		Perceptrons	
	Mean	Std	Mean	Std	Mean	Std
ID	81.20	16.73	47.80	31.47	60.00	21.08
CS	80.15	19.88	65.69	19.04	70.23	15.81

Table 4: Performance on each Topics by Different Faculties

The results of analyzing scores for each topic using Kruskal-Wallis H test are summarized in Table 5.

Topic	H-statistic	P-value
ML Pipelines	0.01	0.9217
Bayes Rule	1.62	0.2026
Perceptrons	1.79	0.1807

Table 5: Faculty Performance on ML Topics

The results of the Kruskal-Wallis H test showed no significant differences in faculty performance for any of the topics. The **p-values** for all topics (ML Pipelines: $\mathbf{p = 0.9217}$, Bayes' Rule: $\mathbf{p = 0.2026}$, Perceptrons: $\mathbf{p = 0.1807}$) were all above the

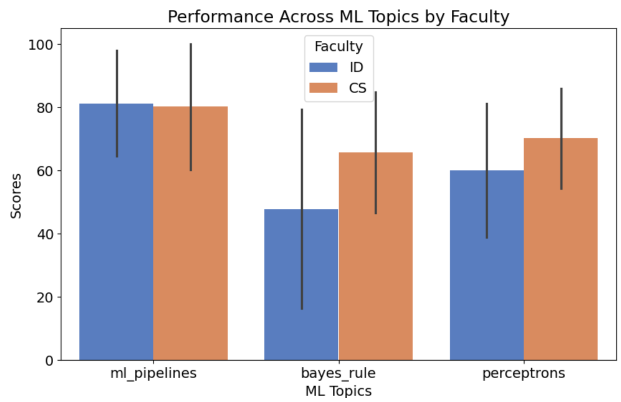

Figure 6: Performance on each Topics by Different Faculties

conventional threshold of 0.05, suggesting that faculty performance did not significantly vary across the different topics. [18]

5.4 Sub-question 4: Qualitative Patterns

Using thematic approaches, the qualitative responses from the final assessments were analyzed. Students from both faculties reported ”Easy” on ML pipelines, ”Difficult” on Bayes' Rule, and ”Moderate” to ”Difficult” on the test. The only topic that differed was perceptrons; ID students mostly reported ”Difficult”, whereas CS students mostly reported ”Moderate”. Both faculties reported ML pipelines to be the easiest due to being ”fundamental” and ”straightforward”. For the hardest topic, CS students had an even division between Bayes' Rule and Perceptrons. On the other hand, 80% of ID students reported Bayes' Rule for this question. Students identified similar resources helpful from the tutorial: the items that were the most frequent were real life examples and diagrams. When asked how they would teach the material to their peers, ID students suggested videos, lectures, interactive examples, and **prototyping** demo with relevant topics. CS students also suggested videos, lectures, and tutorials, but they also suggested **programming** demo with relevant topics.

6 Interpretation and Discussion

This section interprets the results of the study, relating them to the research questions and sub-questions. The discussion includes a synthesis of findings, comparisons with prior literature, identification of limitations, and a cohesive response to the overarching research question.

6.1 Interpretation

While CS students had significantly more previous mathematics related courses or activities than ID students, the results of the Mann-Whitney U test indicated no statistically significant difference between the groups. This could be because the chosen topics are usually not taught extensively in high school level mathematics, and students in both faculties had little experience with them. However, the high number of STEM-related courses taken by CS students, which was more than twice that of ID students, implies that there could be a discrepancy in mathematics proficiency that was not captured by the survey. There was a clear division of the questions that most students answered correctly to the ones answered incorrectly; this suggests the possibility of easy questions being too easy and hard questions being too hard, leading the test being less representative than expected.

Pearson correlation analysis revealed a strong positive correlation between initial mathematics scores and Bayes' Rule performance, but no significant correlation for ML pipelines or perceptrons. The correlation for Bayes' Rule is consistent with expectations, as understanding Bayes' Rule inherently involves applying probability theory, which requires strong mathematical reasoning. On the other hand, perceptrons, despite being foundational in machine learning, involve concepts such as linear separability and weight updates, which require mathematical intuition but are often taught with a visual and procedural manner. In the final assessment, most students answered questions asking about theories of perceptrons and simple output calculation correct. However, it was also found that the question about weight update was mostly answered incorrectly. This could imply the lack of higher mathematical explanation the tutorial had, or the possibility that a passive learning method, such as a tutorial, might not be adequate for teaching such a topic that requires algorithmic thinking. [22]

The Kruskal-Wallis H test found no statistically significant differences between ID and CS students across the three ML topics. This finding aligns with the first paragraph's observation that there were no significant differences in initial mathematics scores between the groups. It may indicate that the tutorials were effective in providing a level playing field, mitigating the potential impact of differences in STEM-related educational backgrounds. Additionally, the accessibility of the ML pipelines topic and the structured nature of the tutorials may have contributed to reducing disparities. However, the lack of faculty-specific differences could also suggest that factors other than mathematics proficiency, such as motivation, learning preferences, or instructional design, played a significant role in determining performance outcomes.

Students from both faculties found ML pipelines the easiest topic and Bayes' Rule the most challenging. However, perceptrons were perceived as ”difficult” by most ID students, while CS students found them ”moderate”. This could be due to the topic requiring more algorithmic thinking during the learning phase, which is a skill that CS students are taught more often. Both groups valued real-life examples and diagrams as effective learning tools but diverged in teaching recommendations, with ID students favoring interactive and prototype-based learning and CS students emphasizing programming demos. This could be due to their previous learning practices, which makes them comfortable and efficient in acquiring new skills and knowledge.

6.2 Comparison with Prior Studies

The findings of this study align and diverge from previous research in machine learning education. The most interesting finding that diverged from prior research was that there were no notable differences in level of mathematics proficiency between ID and CS students. This can imply the importance of students' different ways of understanding a topic, as suggested by Sulmont et al., rather than the prior mathematics or STEM knowledge.

Valuable insights could also be gained from comparing the findings with those of previous research. Sulmont et al. observed that higher-order cognitive tasks, such as designing and comparing models, pose significant challenges for non-majors. This is consistent with the difficulties faced by ID students in topics such as Bayes' Rule and perceptrons, which demand algorithmic reasoning and mathematical understanding. Similarly, Cheong emphasized the importance of structured pathways for non-CS learners, highlighting the role of scaffolding in reducing cognitive load. This study's findings support this view, as no significant faculty differences were observed in the ML pipelines topic, which was presented through well-structured tutorials. Additionally, Banadaki's work underscored the value of visual and interactive methods for non-STEM learners, aligning with the preferences of ID students for diagrams, real-life examples, and prototype-based teaching approaches. Lastly, Tokuta et al. advocated for educational frameworks tailored to non-CS audiences. ID students performing better on less math-intensive topics suggests that interdisciplinary teaching strategies can effectively bridge the gap between technical and design-oriented learners.

6.3 Limitations

While this study offers valuable insights into ML education for students from diverse academic backgrounds, it has several limitations. First, the small sample size, consisting of only 10 ID and 13 CS students, limits the generalizability of the findings and raises concerns about statistical power. Second, the scope of the study was restricted to three ML topics, which may not fully capture the range of concepts required for a comprehensive understanding of ML. Expanding the topic scope could provide a more holistic analysis of learning outcomes. Third, despite efforts to standardize the tutorial content, variations in the design and emphasis of each topic may have inadvertently influenced the results. Moreover, the foundational focus of the tutorials, necessitated by the short time available for learning, constrained the ability to assess higher-level cognitive outcomes. Finally, the reliance on self-reported data for prior mathematics experiences introduces potential inaccuracies, which may have affected the analysis of initial proficiency levels. Not having participants in a controlled environment for the initial survey and assessment also could have introduce external variables that could affect the data.

6.4 Answering the Research Question

The findings suggest that prior mathematics knowledge plays a significant role in learning math-intensive ML topics, such as Bayes' Rule, but has minimal influence on less math-relevant topics, such as ML pipelines. While CS students performed slightly better overall, especially on Bayes' Rule, ID students demonstrated comparable proficiency on less math-intensive topics, indicating the potential for interdisciplinary teaching approaches to bridge gaps. Challenges faced by each faculty highlight their distinct strengths and weaknesses, underscoring the need for tailored teaching strategies that align with their respective skillsets.

7 Conclusion and Future Work

7.1 Conclusion

This study explored the differences in learning outcomes between industrial design (ID) and computer science (CS) students when introduced to foundational machine learning (ML) topics, focusing on the influence of prior mathematics knowledge. By conducting structured tutorials and assessments, the research provided insights into how diverse academic backgrounds shape the learning experience.

The findings revealed that while prior mathematics proficiency significantly impacts performance on math-intensive ML topics such as Bayes' Rule, it has less influence on less math-relevant topics like ML pipelines. CS students generally performed better on quantitative topics, consistent with their stronger mathematical backgrounds. However, ID students demonstrated comparable proficiency on less mathematics-intensive topics, highlighting their adaptability and potential to learn ML through interdisciplinary approaches. Qualitative responses also underscored the value of interactive and visual teaching methods, particularly for ID students, who emphasized creativity and practical application.

By integrating these findings, this study contributes to a deeper understanding of how faculty-specific skillsets and prior knowledge influence ML learning outcomes. It highlights the importance of designing tailored instructional strategies that accommodate diverse learner needs, paving the way for more inclusive and effective educational practices.

7.2 Future Work

While this research sheds light on key aspects of ML education for diverse faculties, several avenues for further exploration remain. For the future work, increasing the number of participants and including students from additional disciplines could provide more generalizable results and uncover broader trends. Because the participants in this study only had two weeks to learn the contents, the contents in the tutorial had to be limited and more advanced topics could not be included. Exploring a wider range of ML concepts may also provide deeper insights into how different faculties approach complex topics, and tracking student performance over a longer period could allow for a more valid comparison on the learning outcomes of different faculties, and students of different mathematics proficiencies. Moreover, it could also reveal how interdisciplinary approaches influence sustained learning and application of ML concepts. Lastly, in order to provide actionable insights for educators, designing and testing teaching methods specifically adapted for non-STEM students such as interactive games or hands-on projects could be useful.

References

[1] L. W. Anderson and D. R. Krathwohl. *Taxonomy for Learning, Teaching, and Assessing*. Pearson, 2001. Accessed on 21.11.2024.

[2] Y. M. Banadaki. Enhancing the role of machine learning in stem disciplines through supervised undergraduate research experiences. *Infonomics Society*, 2020. Accessed on 25.11.2024.

[3] J. Biggs. Enhancing teaching through constructive alignment. *Higher Education: The International Journal of Higher Education Research*, 1996. Accessed on 21.01.2025.

[4] J. Biggs and C. Tang. Applying constructive alignment to outcomes-based teaching and learning. *Training material for "quality teaching for learning in higher education" workshop for master trainers, Ministry of Higher Education, Kuala Lumpur. Vol. 53. No. 9*, 2010. Accessed on 21.01.2025.

[5] J. B. Biggs and K. F Collis. *Evaluating the quality of learning: The SOLO taxonomy (Structure of the Observed Learning Outcome)*. Academic Press, 2014. Accessed on 12.12.2024.

[6] A.Y. Binsaleh, M.S. Abdallah, B.M. Osman, M.M. Bahaa, N. Alsubaie, T.A. Elmasry, M. Yasser, M. Eldesoqui, A.-N.A.-A. Gadallah, and M.A. et al Hamouda. Antimicrobial stewardship intervention for the family caregiver attending primary health care setting: A quasi-experimental study. *Antibiotics 2024, 13, 1145*, 2024. Accessed on 25.01.2025.

[7] V. Braun and V. Clarke. Thematic analysis. *Maggino, F. (eds) Encyclopedia of Quality of Life and Well-Being Research*, 2023. Accessed on 12.12.2024.

[8] S. M. Case and D. B. Swanson. *Constructing Written Test Questions For the Basic and Clinical Sciences*. National Board of Medical Examiners, 2002. Accessed on 12.12.2024.

[9] N. Cheong. Machine learning algorithms for recommendation of learning cs courses in e-learning systems. In *Proceedings of the 5th International Conference on Learning Analytics and Knowledge*. ACM, 2022. Accessed on 26.11.2024.

[10] M. P. Fay and M. A. Proschan. Wilcoxon-mann-whitney or t-test? on assumptions for hypothesis tests and multiple interpretations of decision rules. *Stat Surv. 2010;4:1–39*, 2010. Accessed on 14.01.2025.

[11] R Fiebrink. Machine learning education for artists, musicians, and other creative practitioners. In *ACM Transactions on Computing Education (TOCE), Volume 19, Issue 4*, 2019. Accessed on 30.11.2024.

[12] J. R. Goldberg and P. Malassigné. Lessons learned from a 10-year collaboration between biomedical engineering and industrial design students in capstone design projects. *Int J Eng Educ.2017;33(5):1513-1520.*, 2017. Accessed on 21.01.2025.

[13] A. J. Ko. We need to learn how to teach machine learning. *Bits and Behavior*, 2017. Accessed on 18.11.2024.

[14] P. Larrañaga, D. Atienza, J. Diaz-Rozo, A. Ogbechie, C.E. Puerto-Santana, and C. Bielza. *Industrial Applications of Machine Learning (1st ed.)*. CRC Press, 2018. Accessed on 25.01.2025.

[15] A Liaqat. The design of curriculum, assessment and evaluation in higher education with constructive alignment. *Journal of Education and e-Learning Research, Vol. 5 No. 1 p72-78*, 2018. Accessed on 21.01.2025.

[16] P. E. McKnight and J. Najab. Mann-whitney u test. *The Corsini Encyclopedia of Psychology*, 2010. Accessed on 14.01.2025.

[17] A. G. Mubuuke, A. J. N. Louw, and S. V Schalkwyk. Utilizing students' experiences and opinions of feedback during problem based learning tutorials to develop a facilitator feedback guide: an exploratory qualitative study. *BMC Med Educ 16, 6*, 2016. Accessed on 21.01.2025.

[18] E. Ostertagová, O. Ostertag, and J. Kováč. Methodology and application of the kruskal-wallis test. Master's thesis, Faculty of Mechanical Engineering, Technical University of Košice, 2014. Accessed on 14.01.2025.

[19] K. L Peel. A beginner's guide to applied educational research using thematic analysis. *Practical Assessment Research and Evaluation. 25 (1)*, 2020. Accessed on 12.12.2024.

[20] P. Schober, C. Boer, and S. A. Lothar. Correlation coefficients: Appropriate use and interpretation. *Anesthesia Analgesia 126(5):p 1763-1768*, 2018. Accessed on 14.01.2025.

[21] M. Sharafi and S. A. Sardareh. The effect of dynamic assessment on elementary efl students' l2 grammar learning. *Journal of Applied Linguistics and Language Research Volume 3, Issue 3, 2016, pp. 102-120*. Accessed on 25.01.2025.

[22] E. Sulmont, E. Patitsas, and J. R Cooperstock. What is hard about teaching machine learning to non-majors? insights from classifying instructors' learning goals. *Transactions on Computing Education*, 2019. Accessed on 12.12.2024.

[23] A. Tokuta, R. Z. Lowe, A. Smith, and R. N. Uma. A framework for introductory data science experiences for non-computer science majors. *IEEE Transactions on STEM Education*, 2024. Accessed on 28.11.2024.

[24] Y. P. Tsang and C. K .M Lee. Artificial intelligence in industrial design: A semi-automated literature survey. *Engineering Applications of Artificial Intelligence*, 112:104884, 2022. Accessed on 25.01.2025.

[25] A. van der Helm, Private Conversation, 2024.

[26] X. Wang, Y. Su, S. Cheung, E. Wong, and T. Kwong. An exploration of biggs' constructive alignment in course design and its impact on students' learning approaches. *Assessment Evaluation in Higher Education, 38(4), 477–491*, 2012. Accessed on 21.01.2025.

[27] J. Wunderlich, L. Higgins, and Y. Lichtenstein. Machine learning for business students: An experiential learning approach. In *ITiCSE '21: Proceedings of the 26th ACM*

Conference on Innovation and Technology in Computer Science Education V. 1, 2021. Accessed on 22.01.2025.

Appendix

A Constructive Alignment

A.1 Machine Learning Pipeline

Intended Learning Objectives:

- Understand the general machine learning pipeline
  - addressed in Chapter 1
- Explain the purpose of training, test and validation sets
  - addressed in Chapter 3.1, 3.2, 4.1, and 4.2
  - assessed in question 3
- Compare the performance of a model trained on different dataset splits
  - addressed in Chapter 3.3 and 4.3
  - assessed in questions 4 and 7
- Identify overfitting and underfitting of a machine learning model
  - addressed in Chapter 2
  - assessed in questions 5 and 6

A.2 Bayes' Rule

Intended Learning Objectives:

- Apply Bayes' rule to solve probability problems
  - addressed in Chapter 1
- Understand the relationship between prior, likelihood, and posterior probabilities
  - addressed in Chapter 1 and 2
  - assessed in question 8
- Calculate conditional probabilities in real-world scenarios
  - addressed in Chapter 1 and 2
  - assessed in questions 10, 11, and 12
- Evaluate decisions with Bayesian reasoning
  - addressed in Chapter 3
  - assessed in question 9

A.3 Perceptrons

Intended Learning Objectives:

- Understand the pipeline of training artificial neural networks
  - addressed in Chapter 1, 2.2, and 3.1
  - assessed in questions 14 and 17
- Calculate the output of a single perceptron
  - addressed in Chapter 2.1
  - assessed in question 15
- Explain the advantages/disadvantages of neural networks
  - addressed in Chapter 3.2
  - assessed in questions 13 and 16

B Prompts used for Generative AI

For checking grammars and spellings and reformulating sentences to an appropriate style

Prompt: ”Can you check grammar and spellings and reformulate the sentence below so that it becomes appropriate for an academic report? [Input sentence]”

This prompt was used in several places in the report and in the tutorial.

Example prompt: ”Can you check grammar and spellings and reformulate the sentence below so that it becomes appropriate for an academic report? Machine learning (ML) is becoming a crucial technology of modern society, giving life to various gadgets from selfdriving cars to voice assisstants”

Output: ”Machine learning (ML) has emerged as a key component of contemporary technology, enabling applications ranging from self-driving automobiles to voice assistants.”

For finding an appropriate Python library to perform Mann-Whitney U Test and learning how to use it

Prompt: ”I am trying to perform Mann-Whitney U test to compare test scores of two different student groups. Can you suggest me an appropriate python library to do this, with explanation on how to use it?”

For finding an appropriate Python library to perform Pearson Correlation Analysis and learning how to use it

Prompt: ”I want to use pearson correlation analysis for finding the correlation between initial test scores to test results of 3 different topics, finding coefficients for each topic. Can you suggest me an appropriate python library to do this, with explanation on how to use it?”

For finding an appropriate Python library to perform Kruskal-Wallis H Test and learning how to use it

Prompt: ”I am trying to perform Kruskal-Wallis H Test to compare three different test scores of two different student groups, making three different analysis. Can you suggest me an appropriate python library to do this, with explanation on how to use it?”

C Figures in Results and Analyses

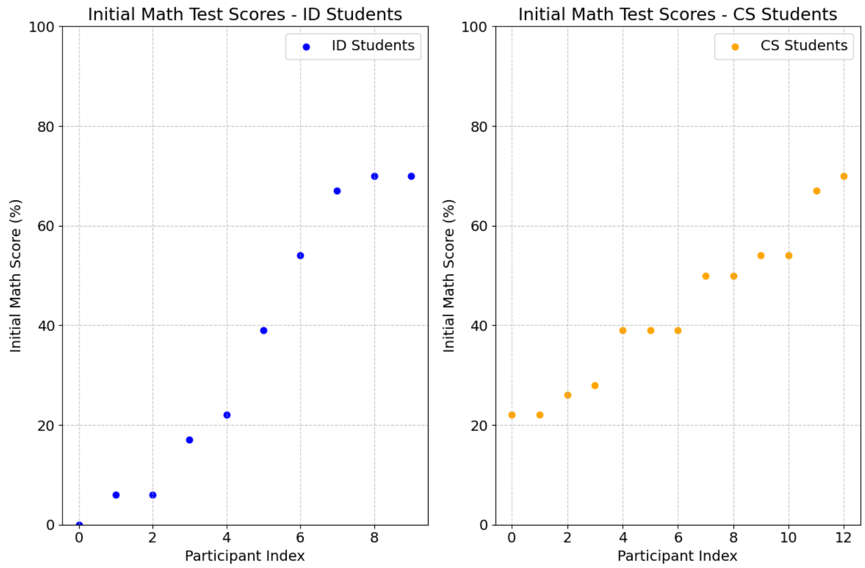

Figure 7: Scatterplot of Initial Mathematics Scores

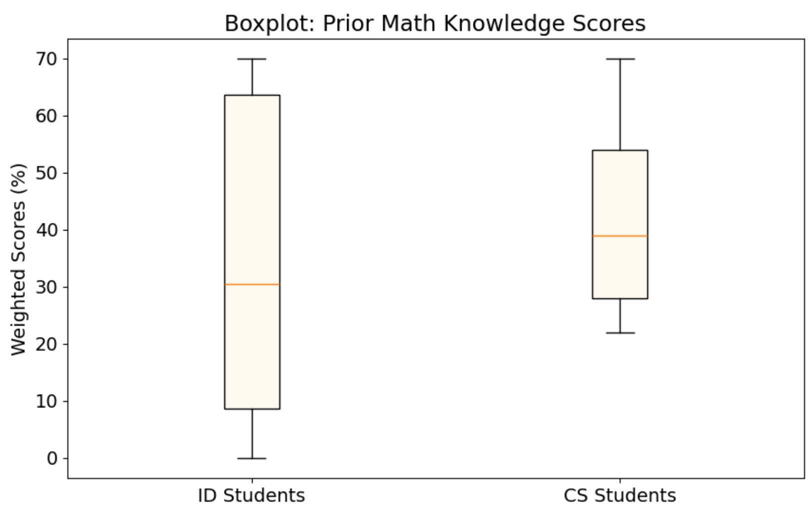

Figure 8: Boxplot of Initial Math Scores

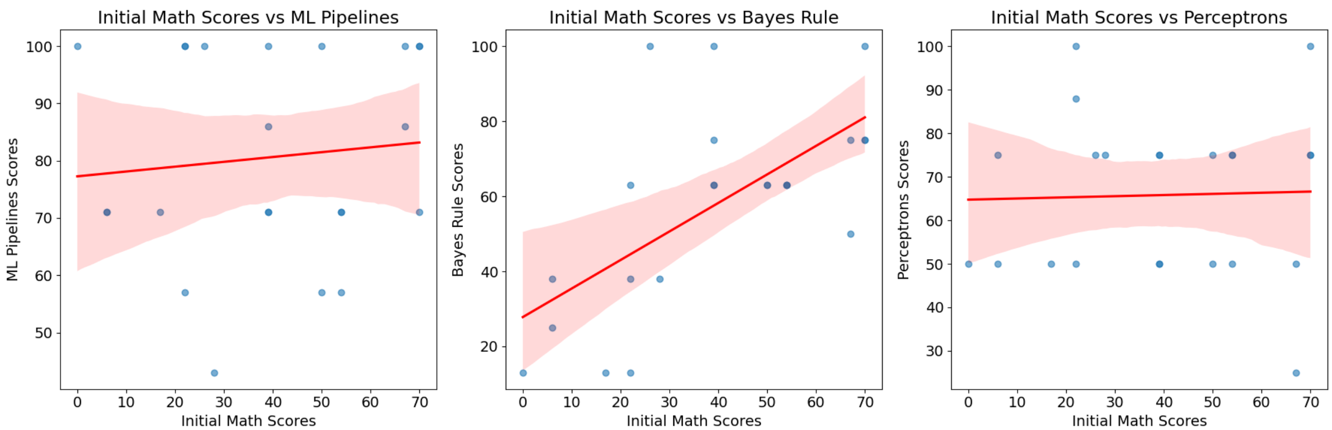

Figure 9: Scatterplot of Initial Math Score to Individual Topic Scores

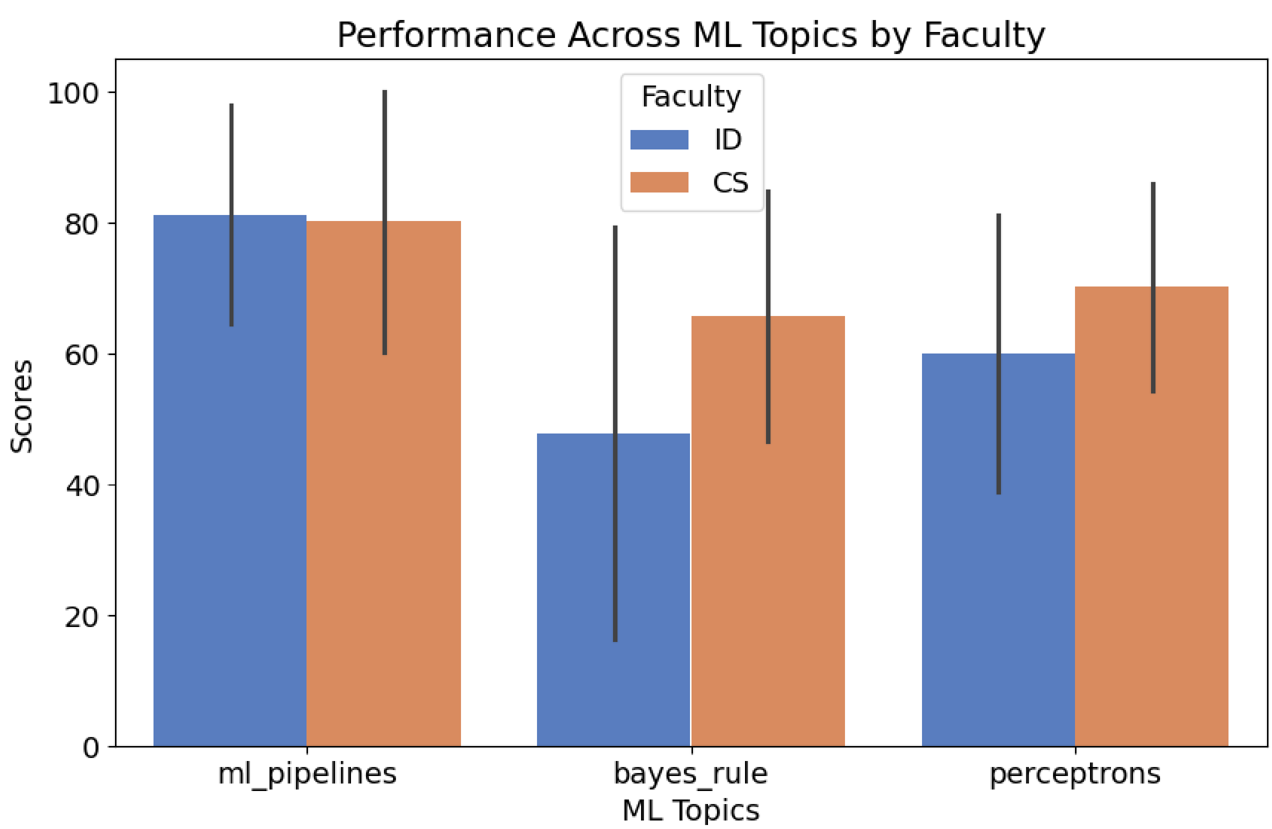

Figure 10: Performance on each Topics by Different Faculties