We need to learn how to teach Machine Learning

Machine Learning for Everyone: Exploring Diverse Pedagogical Approaches for Non-CS Students.

Ghiyath Alaswad

Supervisor(s): Responsible Professor: Prof. dr. ir. G. Migut,
Supervisor: Ilinca Rențea

[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 26, 2025

Name of the student: Ghiyath Alaswad
Final project course: CSE3000 Research Project
Thesis committee: Responsible Professor: Prof. dr. ir. G. Migut, Supervisor: Ilinca Rențea, Examiner: prof. Jesse Krijthe

Abstract

Machine learning (ML) has become a vital skill across various disciplines, driving innovation and transforming industries. This growing demand emphasizes the need for effective teaching methods tailored to students with diverse academic and technical backgrounds. Teaching ML to non-majors presents significant challenges, as many students lack foundational knowledge in mathematics and programming. This study explores how university instructors addressing these challenges in bachelor's and master's level courses, based on semi-structured interviews.

The analysis uncovered several key strategies used by instructors. Many emphasized connecting ML concepts to real-world applications, making the subject more relatable. Visualization tools were commonly employed to simplify abstract concepts and improve comprehension. Hands-on activities, including live coding and interactive assignments, were highlighted as effective methods to engage students and bridge theory with practice. However, instructors faced challenges as well such as accommodating diverse student backgrounds, correcting misconceptions, and designing assignments that balanced accessibility and depth.

Although tracking student progress was not the focus of this study, some insights were provided. Some instructors mentioned using formative assessments, such as quizzes and project-based evaluations, to measure understanding.

These findings highlight the importance of adaptable teaching methods and inclusive learning experiences in making ML education more accessible and effective for non-majors, while also providing actionable advice to improve the educational process and course design.

1 Introduction

Machine learning (ML) is one of the fastest-growing fields in science, driving transformative changes across industries by solving complex tasks through data-driven approaches (Jordan & Mitchell, 2015). While traditionally rooted in computer science and statistics, ML has drawn attention from diverse audiences, including non-CS individuals seeking accessible learning methods (Rashid & Kausik, 2024). This shift underscores the need for tailored pedagogical strategies to make ML concepts accessible to learners from various academic backgrounds, especially those lacking formal programming or mathematical training (Sulmont et al., 2019a).

Teaching ML to non-majors presents unique challenges. Foundational gaps in mathematics and programming often hinder students' engagement and comprehension. Additionally, misconceptions, such as viewing ML as an entirely automated process, increased these difficulties (Sulmont et al., 2019b). Addressing these barriers requires innovative teaching methodologies, including real-world applications, visualization tools, and hands-on activities to bridge theory and practice effectively (Mike & Hazzan, 2022).

> ”Lowering the barrier for teaching is as important as lowering the barrier for learning.”

This perspective, highlighted by Acquaviva (Acquaviva, 2022), motivates the need to focus not only on student challenges but also on the obstacles educators face in designing and delivering ML courses. By addressing these dual barriers, educators can empower students while refining their teaching practices to achieve broader accessibility and impact.

This study explores how university instructors address these challenges in ML education for non-majors. Through semi-structured interviews with educators, we analyze effective teaching strategies, common challenges faced by students and instructors, methods for tracking student progress, and actionable advice to improve ML curricula. The findings aim to support the development of inclusive, interdisciplinary teaching approaches that make ML education more accessible and impactful for learners outside the traditional computer science domain.

To provide a clear outline, the remainder of this paper is structured as follows: related work on teaching machine learning to non-majors is reviewed in Section 2, summarizing key findings and gaps in the literature. The methodology, including participant selection and the thematic analysis approach, is detailed in Section 3. The main results are presented in Section 4, covering teaching strategies, challenges for both students and teachers, tracking progress strategies, and advice from instructors. Ethical considerations and responsible research practices are outlined in Section 5. The findings are discussed in Section 6, highlighting their implications for ML education, including limitations and suggestions for future research. Finally, Section 7 concludes the paper by summarizing the contributions and findings.

1.1 Research Questions

To better understand how machine learning can be effectively taught to non-majors, this study focuses on the following research questions:

1. What pedagogical strategies do instructors use to teach machine learning concepts, simplify complex topics, and track student progress and understanding?
2. What challenges do non-majors face when learning machine learning, and how do these challenges impact their engagement and understanding?
3. What challenges do instructors face when teaching machine learning to non-majors, and how do they address these challenges?
4. What actionable advice can be drawn from instructors' experiences to guide educators in designing and delivering effective ML courses for non-majors?

2 Related Works

The growing interdisciplinary relevance of machine learning has driven research into effective pedagogical frameworks for teaching non-CS students. Historically, ML education has focused on computer science students, emphasizing mathematical depth and algorithmic understanding. However, as ML applications extend to fields like business, creative arts, and engineering, teaching strategies must evolve to accommodate diverse learner needs (Fiebrink, 2019; Linos & Chenoweth, 2020).

Bloom's taxonomy, a general educational framework, provides a structured approach for designing curricula. While not specific to ML, it enables progression from foundational knowledge (e.g., Remember) to higher-order skills such as creating and evaluating. (Mike & Hazzan, 2022) highlights the value of focusing on conceptual understanding before delving into complex mathematics, aligning well with the early stages of Bloom's taxonomy. Similarly, project-based learning, linked to higher levels of Bloom's taxonomy (e.g., Apply, Create), enhances student engagement and encourages interdisciplinary problem-solving skills (Chenoweth & Linos, 2018).

A significant challenge in teaching ML to non-majors involves addressing misconceptions and knowledge gaps. For instance, (Sulmont et al., 2019a) highlights that students often perceive ML systems as entirely autonomous, overlooking the critical role of human decision-making in tasks like feature selection and parameter tuning. Active learning strategies, such as interactive tools and real-world datasets, help dispel these misconceptions by illustrating ML workflows in relatable contexts (Fiebrink, 2019).

The use of interactive platforms like Jupyter Notebooks has also emerged as a key enabler for teaching ML concepts to students with limited programming experience. Such tools provide an intuitive environment for experimentation, fostering confidence and engagement among non-CS learners (Locke & Rainer, 2024). Tailored teaching approaches, including domain-specific case studies and ethical considerations, further enhance relevance and accessibility, as exemplified by creative ML courses designed for practitioners in the arts (Fiebrink, 2019). As another example of an alternative analysis approach, (Ishikawa & Yoshioka, 2019) employed a questionnaire-based methodology to investigate challenges in ML engineering, offering a broader view of the difficulties practitioners face. Despite the progress in ML education research, significant gaps remain. Existing studies often overlook strategies for tracking and evaluating student progress comprehensively, as well as the challenges faced by educators. While prior research, such as (Fiebrink, 2019) and (Locke & Rainer, 2024), provides valuable insights into creative and practical approaches, this paper focuses exclusively on insights gathered through interviews with active instructors. These interviews offer actionable strategies to refine instructional practices, enhance student engagement, and promote collaboration among educators.

3 Methodology

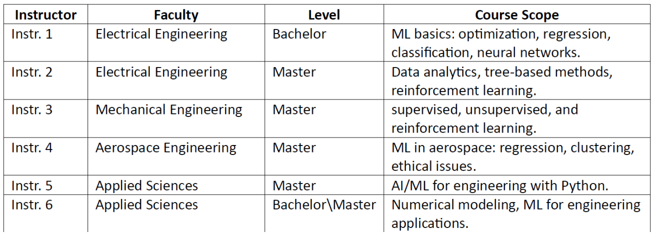
This section outlines the methodological approach used in this study, including participant selection, interview design, data collection, and the thematic analysis process. It describes how semi-structured interviews with instructors teaching ML to non-majors were conducted and analyzed to uncover key teaching strategies and challenges in ML education.

3.1 Participant Selection

To identify suitable participants, TU Delft study guide was reviewed to compile a list of ML courses and their respective instructors.The study guide included detailed information about each course, such as descriptions, learning objectives, content, study goals, educational methods, and assessment techniques. This review provided a foundational understanding of the course contexts and allowed for the development of a comprehensive list of potential instructors.

Participants were chosen through purposeful sampling, a strategy described by Creswell (Creswell, 2012), which involves selecting individuals who are especially knowledgeable about the topic of interest. This method ensured that the instructors selected for the study had direct experience teaching ML to non-majors, allowing for rich and relevant insights.

Table 1 provides an overview of the instructors, detailing their respective faculties, teaching levels, and course focus. All participants were instructors at Delft University of Technology (TU Delft), representing an adequate range of faculties and teaching levels, including both bachelor's and master's courses. This diversity ensured a variety of perspectives on teaching ML to non-majors, enriching the data collected. Once the list of potential participants was finalized, they were contacted via email, provided with an overview of the study, and invited to participate in a one-on-one interview. Consent to participate was obtained prior to scheduling interviews, ensuring that participants were fully informed about the study's purpose and expectations.



Instructor	Faculty	Level	Course Scope
Instr. 1	Electrical Engineering	Bachelor	ML basics: optimization, regression, classification, neural networks.
Instr. 2	Electrical Engineering	Master	Data analytics, tree-based methods, reinforcement learning.
Instr. 3	Mechanical Engineering	Master	supervised, unsupervised, and reinforcement learning.
Instr. 4	Aerospace Engineering	Master	ML in aerospace: regression, clustering, ethical issues.
Instr. 5	Applied Sciences	Master	AI/ML for engineering with Python.
Instr. 6	Applied Sciences	Bachelor\Master	Numerical modeling, ML for engineering applications.

Figure 1: Overview of Instructors Interviewed.

3.2 Interview Design

The interviews were designed to align with the study's objectives and Creswell's recommendations for qualitative research (Creswell, 2012). Open-ended questions were developed to encourage detailed responses and allow participants to elaborate on their experiences and teaching methods. These questions were carefully structured to focus on six key areas: teaching strategies, which explored the methods and tools instructors use to simplify ML concepts for non-majors; student challenges, addressing the barriers students face and how instructors help overcome them; tracking progress, examining how instructors track students' progress and evaluate their understanding of ML concepts and their ability to apply them; preconceptions, investigating common misunderstandings students have about ML and the strategies instructors use to address these issues; challenges teachers faced, which delved into the most rewarding and challenging aspects of teaching ML to non-majors; and advice from teachers, capturing practical recommendations for improving ML education. This design ensured that the questions captured a comprehensive range of insights relevant to the study's objectives.

A detailed list of the interview questions can be found in Appendix A. To tailor the interviews to the specific contexts of each course, I analyzed the course descriptions, learning objectives, and assessment methods from the study guide prior to conducting the interviews. This preparation ensured that the questions were relevant and reflective of the unique characteristics of each course.The interview design followed the guidelines outlined by Creswell, which emphasize flexibility while maintaining a focus on the study's core objectives.

3.3 Conducting the Interviews

Each interview commenced with an overview of the research objectives, methodology, and ethical considerations to ensure participants were fully informed. Participants signed a consent form in advance, explicitly granting permission for recording of the session. At the start of the interview, participants were provided an opportunity to ask any questions about the research, ensuring clarity and addressing potential uncertainties.

Five out of six interviews were conducted online in a one-to-one format, facilitating flexibility and accessibility for participants, while one interview was conducted in person. The semi-structured format guided the discussion and was structured around open-ended questions designed to explore key areas, including teaching strategies, challenges, and progress tracking. Each session lasted approximately 20–30 minutes, allowing sufficient time for participants to share their perspectives. At the conclusion of each session, participants were given space to offer any additional insights or thoughts they felt were relevant but had not been discussed.

3.4 Data Analysis

The analysis of the collected interview data followed a structured thematic analysis approach, with the goal of identifying key themes related to teaching strategies, challenges faced by students and teachers, methods for tracking student progress, and practical advice shared by instructors.

Initial Coding

The process began with a thorough review of the interview transcripts. Each transcript was carefully read multiple times to gain a deep understanding of the content. Initial codes were then generated inductively, meaning they emerged directly from the data rather than being predetermined. This approach, recommended by (Creswell, 2012), ensures that the analysis remains grounded in participants' actual experiences. Examples of initial codes included:

- *Active Learning*
- *Real-World Applications*
- *Challenges with Mathematical and Programming Backgrounds*
- *Designing Effective Assessments*
- *Tailored Content*
- *Collaboration Among Instructors*

A complete list of codes generated during this process is provided in the appendix (Appendix C). Following the guidelines of (Miles, 1994), the initial coding process helped break down the data into manageable units, ensuring that all relevant insights were captured effectively.

Identifying and Refining Themes

After completing the initial coding, related codes were grouped into broader themes to represent significant aspects of the data. This step involved reviewing the codes to identify patterns and clustering them into meaningful categories. The themes were refined iteratively, ensuring they captured the essence of the data accurately. This process included recoding, comparing themes with initial categorizations, and ensuring consistency across the dataset.

The main themes identified were **Teaching Strategies**, focusing on methods instructors use to simplify complex ML concepts and engage students; **Challenges Faced by Students**, which highlighted obstacles such as conceptual difficulties and limited programming skills; **Challenges Faced by Teachers**, addressing issues like variability in student preparedness and the impact of AI tools on assessment; **Progress Tracking**, outlining techniques for assessing and monitoring student learning; and **Advice from Teachers**, summarizing practical recommendations for improving ML education for non-majors. These themes provide a structured basis for analyzing the complexities of teaching ML to non-majors. They are further explored in the Results section, where each theme is elaborated with supporting evidence and illustrative quotes from the interview data.

This approach aligns with the recommendations of (Dey, 2003), who emphasizes the importance of iterative refinement in thematic analysis to ensure rigor and accuracy.

Tool Usage

The qualitative data analysis software **Atlas.ti** was used to facilitate the systematic organization and management of codes and themes. Atlas.ti allowed for efficient data segmentation, linking of codes to specific quotes, and visualization of relationships between themes. The software's capabilities enhanced the rigor of the analysis process and ensured that findings were traceable and well-documented. Atlas.ti's scalability also makes it suitable for future research involving larger datasets (Miles, 1994).

Separately, **GenAI**, e.g., ChatGPT, were used at various stages of this research. These tools helped in generating ideas that served as inspiration, helping with guiding the interviews and performing the thematic analysis, and rephrasing parts of the text for better flow and more solid sentence construction. This integration of advanced tools complemented the traditional analysis methods, improving the overall clarity and coherence of the research outcomes. Additionally, the specific prompts are provided in Appendix B for transparency and reproducibility.

Reviewing and Validating Themes

To ensure the validity of the identified themes, the final step involved revisiting and cross-checking them against the original interview data. This iterative process ensured consistency, minimized biases, and refined the themes to accurately reflect the data. This approach follows the recommendations of (Dey, 2003), which emphasize the importance of iterative validation in thematic analysis. These validated themes form the basis of the findings presented in the **Results** section, contributing to a deeper understanding of how machine learning is taught to non-majors.

4 Results

This section presents the findings from the interviews. The results are organized around five main themes: teaching strategies, challenges faced by students, methods for tracking student progress, challenges faced by instructors, and advice from instructors.

To give an overview of the relative importance and focus on each theme and its associated subcategories, Figure 2 illustrates the frequency with which these themes were mentioned in the interviews. This visualization offers a quantitative perspective on how each theme is reflected in the research, helping to identify areas that instructors considered particularly significant in the context of teaching machine learning to non-majors.

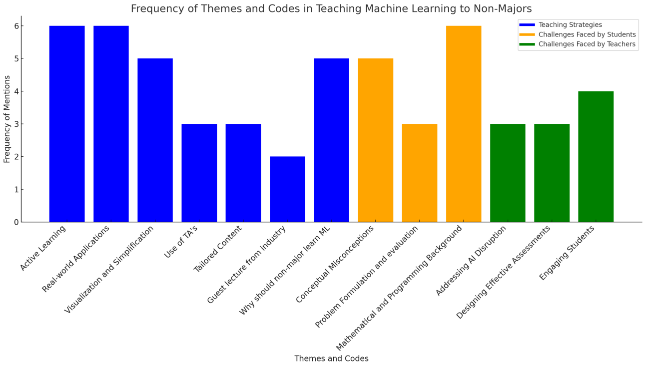


Figure 2: Frequency of Themes and Codes in Teaching Machine Learning to Non-Majors

For a deeper understanding of the analysis, the appendix includes additional resources: a comprehensive table of codes used in the study (Appendix C.2) and a diagram illustrating the relationships between the themes and their related codes (Appendix C.1).

4.1 Teaching Strategies

Real-world Applications and Guest Lectures: Instructors highlighted the importance of making machine learning concepts tangible by connecting them to practical, domain-specific problems. To enhance this, they often invited industry professionals as guest speakers to provide students with real-world insights and applications. These strategies aimed to help students understand the relevance of machine learning in their fields while showing them its impactful use cases. One approach involved using datasets tailored to the students' domains. For example, tasks such as predicting chemical properties were designed to demonstrate how machine learning can solve engineering problems in a meaningful and accessible way.

> ”We provide datasets of chemical properties of compounds and ask students to predict the properties of unseen compounds, connecting the tasks to real-world engineering problems.”

Guest lectures served as an effective way to bridge classroom learning with real-world practice. Industry professionals shared how machine learning is applied in various scenarios, such as automated failure detection and process optimization, offering students a practical perspective on the field.

> ”We also invite guest speakers from the industry to discuss practical applications of machine learning, such as its use in automated failure detection and process optimization.”

Why Should Non-Majors Learn ML? Instructors emphasized the importance of motivating students to learn machine learning by showcasing its transformative potential in various fields. Machine learning was presented as a powerful tool capable of automating repetitive tasks, uncovering patterns in data, and solving complex problems that traditional methods cannot easily address. By showing students real-world successes of ML, instructors attempted to convince them of its relevance and value in their disciplines. Instructors also highlighted that while machine learning tools and models are widely available, their correct application requires domain knowledge. Students from non-technical fields bring essential expertise to ensure that ML is applied effectively and ethically within their respective areas. This underscores the interdisciplinary nature of machine learning and the critical role non-majors play in harnessing its full potential.

> ”My main strategy is to motivate students by showing them how machine learning can be applied to real-world control problems and by highlighting its recent successes across disciplines.”

Visualization and Simplification: Strategies like using geometric visualizations and plots were cited as effective methods to make complex concepts intuitive. Simplified datasets and Python assignments were used to bridge the gap between theory and practice.

> ”We follow a geometrical approach, translating intuition into mathematical formulas and then into code.”

Active Learning: Most instructors adopted hands-on approaches, such as providing Jupyter notebooks with pre-written templates for students to complete during labs.

Tailored Content: Instructors recognized that students come from diverse academic backgrounds, each with unique strengths and challenges. Tailoring course content to meet these varied needs was highlighted as an essential strategy for fostering inclusivity and engagement. For students with little or no programming experience, instructors provided resources designed to build confidence and close knowledge gaps. Tools such as Python primers and step-by-step Jupyter notebooks were particularly effective in helping students understand the basics of libraries and object-oriented programming. Instructors also adapted their teaching to match the academic disciplines of their students. By choosing examples and exercises relevant to the students' fields, they made abstract machine learning concepts more relatable and practical.

> ”Another strategy is to adapt the material depending on the audience. For example, I don't go into advanced neu-

> ral network concepts with biology students, but I might with control students who have a stronger mathematical background."

Recognizing that not all students had the same level of computational ability, instructors designed smaller datasets and simplified exercises to align with students' individual capabilities. This approach ensured that students with varying levels of mathematical and programming skills could still actively engage with the material.

> "The features we gave and the datasets are really toy and tiny because not all of them have huge computational capability."

By addressing these diverse needs, tailored content not only bridged knowledge gaps but also created a more equitable and engaging learning environment. Students felt supported and motivated, which ultimately enhanced their learning experience and outcomes.

Use of Teaching Assistants (TAs): Teaching Assistants were widely acknowledged as essential in supporting students during labs, guiding them through assignments, and addressing their individual questions.

4.2 Challenges Faced by Students

Mathematical and Programming Background: Many students struggled with the mathematical foundations of machine learning, such as optimization, probability, and linear algebra. Basic programming skills were also a common barrier.

> "The biggest barrier isn't machine learning itself, but it's programming prerequisites."

Conceptual Misconceptions: Students often perceived ML as a "magic box," lacking a clear understanding of the decision-making processes and data requirements.

> "Students assume ML is automated and fail to grasp its human-driven aspects."

Problem Formulation and Evaluation: Formulating well-defined problems and evaluating ML models effectively emerged as significant challenges. Students struggled to recognize how to structure their datasets, choose appropriate metrics, and validate their results.

> "Students struggled with identifying which data preprocessing steps to take and how to split datasets properly for training and testing."

4.3 Tracking Student Progress

Formative Feedback: Weekly assignments provided students with immediate feedback from instructors or automated systems. This approach allowed students to identify and correct misconceptions early.

Practical Assessments: Instructors emphasized using practical coding tasks during exams and projects to assess students' understanding.

4.4 Challenges Faced by Instructors

Addressing AI Disruption: Tools like ChatGPT and Copilot posed challenges for maintaining academic integrity. Instructors encouraged students to use these tools for troubleshooting rather than directly solving assignments.

> "Students assume large language models always generate correct solutions, which reduces critical thinking."

Designing Assignments, Projects, and Exams: Instructors highlighted the difficulty of creating assignments, projects, and exams that strike the right balance between being domain-relevant, educationally valuable, and accessible for students with diverse skill levels.

> "We struggle to find assignments that are simple enough for beginners yet still relevant to their disciplines."

Engaging Students: Instructors faced difficulties motivating students who perceived ML as irrelevant to their fields.

> "Students don't see the relevance of ML in their field, so we ensure every example is tied to their domain."

4.5 Advice from Instructors

Adaptability: Instructors emphasized the need to adapt the curriculum to students' backgrounds and abilities. They suggested starting with broad concepts to give students a broad understanding before diving into complex technical details. This approach was seen as effective in reducing student anxiety and ensuring engagement.

> "Starting with broad concepts before diving into PCA helped reduce student anxiety."

This strategy allows students to build confidence in their understanding of machine learning without being overwhelmed by mathematical complexities early in the course.

Collaboration: Sharing teaching resources and strategies across institutions was highlighted as a way to improve teaching efficiency and quality. Instructors appreciated the value of building a network of educators to exchange materials and ideas, reducing the workload and enhancing the learning experience for students.

> "We don't need to reinvent the wheel. Sharing teaching materials across universities is a great way to save time and improve quality."

Additionally, instructors emphasized the importance of collaborating with domain experts to bridge the gap between subject-specific knowledge and machine learning expertise.

> "My solution is to partner up, talk to other teachers who are teaching courses in machine learning but have a different background. This allows us to combine perspectives and improve the material for students."

Collaboration fosters innovation in pedagogy, ensuring instructors can leverage collective expertise to create courses that are both engaging and interdisciplinary.

Focus on Practical-Oriented Content: Instructors emphasized the need to frame theoretical topics, such as neural networks or optimization, in the context of specific applications to maintain student engagement and motivation.

> "For students with a more practical mindset, it's critical that they see what applications they are working toward. If you start with just a theoretical framework, like neural networks, they often lose interest. Instead, you need to outline specific applications first, even if it's a math-heavy topic, so they understand why it matters."

This advice underscores the importance of contextualizing theoretical concepts to help students connect abstract topics with real-world use cases and maintain their interest throughout the course.

5 Responsible Research

5.1 Ethical Considerations

This study follows strict ethical guidelines to ensure that all research involving human participants is conducted responsibly. Before initiating the research, approval was obtained from the **Human Research Ethics Committee (HREC)** at TU Delft. The application involved a thorough assessment of potential risks, participant consent procedures, and data security measures.

Participants were provided with detailed information about the study, including its purpose, procedures, and expected duration. The informed consent process ensured that participants voluntarily agreed to take part in the study after understanding all relevant details. Consent forms explicitly outlined participants' rights, such as the ability to withdraw at any time without consequences. Recordings of the interviews were only conducted with explicit permission, and participants were assured that their data would be anonymized.

5.2 Data Security and Privacy

To mitigate the risk of data breaches, all collected data, including recordings and transcriptions, were securely stored on password-protected devices. Access to the data was restricted to the research team. During transcription, personally identifiable information (PII) was removed to ensure anonymity. Furthermore, personal data, such as names and contact information, will be permanently deleted after the completion of the research.

Anonymized data will only be used in published reports, ensuring that no participant can be re-identified. Additionally, participants agreed that their anonymized responses could be quoted in research outputs.

5.3 Reproducibility and Transparency

In line with the principles of responsible research, efforts have been made to ensure reproducibility and transparency in both methods and results. While the dataset used for this study is relatively small, the analysis and coding of interview responses were conducted using **Atlas.ti**, a qualitative data analysis software. This decision was made to facilitate better reproducibility in case a larger dataset is used in future research. Using a well-established tool like Atlas.ti allows for systematic and traceable coding of qualitative data, enhancing the reliability of the findings.

Additionally, a clear methodology, including detailed documentation of the data collection and analysis processes, accompanies the research to ensure that other researchers can replicate or build upon this work. Metadata for anonymized datasets will be registered in **4TU.ResearchData**.

6 Discussion

This section reflects on the study's key findings, situates them within the broader context of existing literature, and identifies practical implications and areas for future research.

6.1 Key Findings

The study revealed several effective strategies and notable challenges in teaching machine learning (ML) to non-majors. Instructors employed strategies such as tailoring content to students' diverse academic backgrounds, emphasizing real-world applications, promoting active learning, and using visualization techniques to make abstract concepts more comprehensible. These approaches reflect a student-centered approach to teaching, aimed at making ML both accessible and engaging.

Tailoring content emerged as a critical approach to addressing the diverse academic backgrounds of students. By incorporating discipline-specific examples, instructors bridged the gap between abstract ML concepts and practical applications, helping students connect the material to their own fields and careers. This approach motivated students while making the subject matter more relatable.

Real-world applications played a key role in enhancing engagement and understanding. Instructors designed tasks that connected ML concepts to tangible problems, such as predicting chemical properties or optimizing engineering processes. These examples demystified ML and demonstrated its potential to solve real-world challenges, enabling students to see the practical value of the subject.

Active learning through hands-on assignments and guided projects was another effective strategy. Activities like collaborative projects and live coding sessions encouraged students to experiment with ML techniques, fostering deeper engagement and critical thinking.

Visualization techniques also played an important role in making abstract concepts more comprehensible. By using graphical representations and step-by-step demonstrations of algorithms, instructors helped students overcome conceptual barriers, particularly those with limited mathematical backgrounds. This approach reinforced students' understanding by providing intuitive, accessible teaching materials.

Despite these successes, instructors encountered significant challenges. Students' varying levels of programming and mathematical proficiency often necessitated additional support, such as preparatory tutorials or the involvement of teaching assistants. Conceptual misconceptions about topics like the probabilistic nature of ML, as well as difficulties in problem formulation and evaluation, were common hurdles.

Another emerging challenge was the reliance on generative AI tools like ChatGPT. While these tools have potential benefits, instructors raised concerns about their impact on authentic learning. They worried that students might use these tools to bypass real engagement with the material, underscoring the need for innovative assessments to address this issue.

6.2 Comparison with Existing Literature

The strategies and challenges identified in this study align closely with prior research while also highlighting new areas of concern in teaching ML to non-majors.

The use of real-world applications has been widely recognized as an effective strategy for engaging students in technical disciplines. Sulmont et al. emphasize the importance of contextualizing ML concepts through domain-specific datasets, such as those used for predicting cancer outcomes or spam classification (Sulmont et al., 2019a). Instructors in this study similarly designed projects tied to practical problems, helping students connect abstract ML concepts to tangible, discipline-relevant tasks.

Visualization techniques are another well-documented approach to simplifying complex ideas. Fiebrink highlights the effectiveness of visual tools, such as the Regression Explorer, in making abstract ML concepts more accessible to non-majors by illustrating geometric interpretations and algorithmic processes (Fiebrink, 2019). Instructors in this study echoed this by employing graphical representations and simplified datasets to demystify challenging concepts for students with limited technical backgrounds.

Active learning, particularly through hands-on assignments, is a cornerstone of effective ML education. Linos et al. demonstrated that project-based learning, where students analyze datasets and experiment with algorithms, significantly enhances their ability to apply theoretical knowledge in practical contexts (Linos & Chenoweth, 2020). The instructors in this study adopted similar practices, integrating collaborative projects and live coding sessions to foster engagement and critical thinking.

Challenges related to diverse academic backgrounds align with findings by Sulmont et al. and Mike et al., who observed that non-majors often lack foundational programming and mathematical skills (Mike & Hazzan, 2022; Sulmont et al., 2019a). Instructors in this study responded to these challenges by tailoring teaching materials, offering supplementary resources such as preparatory tutorials, and relying on teaching assistants for additional support.

A newer issue identified in this study is the reliance on AI tools like ChatGPT, reflecting emerging concerns in educational settings. Mike et al. highlight the ethical and educational challenges posed by AI technologies, particularly in ensuring that students engage deeply with course content rather than relying on automated tools (Mike & Hazzan, 2022). This finding underscores the growing need for innovative assessment methods that prioritize critical thinking and authentic learning.

6.3 Implications for Educators

The findings from this study offer actionable strategies and considerations for educators seeking to improve ML education for non-majors.

First, tailoring course content to align with students' academic backgrounds and fields of interest is crucial. Discipline-specific examples and projects not only make ML concepts more relatable but also empower students to see the relevance of ML in solving real-world problems. Educators should collaborate with domain experts to design assignments that reflect practical challenges within students' disciplines.

Second, integrating real-world applications and visualization techniques into teaching can demystify complex ML concepts. By using accessible datasets, graphical representations, and step-by-step demonstrations, educators can make ML more intuitive, particularly for students with limited technical skills. Interactive tools, such as Jupyter notebooks with embedded visualizations, can further enhance student comprehension and engagement.

Third, active learning remains a cornerstone of effective ML education. Hands-on projects, guided coding exercises, and collaborative assignments encourage experimentation and foster deeper understanding. Providing students with structured but open-ended tasks allows them to apply theoretical concepts in meaningful ways, enhancing both engagement and problem-solving skills.

In light of the growing use of AI tools like ChatGPT, educators must also rethink assessment strategies. Traditional methods may no longer be sufficient to ensure authentic learning. Instead, innovative assessments such as in-class coding tasks, oral examinations, and project-based evaluations can promote critical thinking and ensure students genuinely engage with the material.

Finally, addressing students' diverse backgrounds requires targeted support. Offering preparatory courses in programming and mathematics can bridge foundational gaps, while leveraging teaching assistants can provide personalized guidance during practical tasks. Collaboration among educators from different disciplines can also lead to more inclusive teaching practices, ensuring that ML courses meet the needs of all students.

6.4 Limitations and Future Work

While this study provides valuable insights, several limitations must be acknowledged. First, the sample size was small, consisting of six instructors from a single technical university in the Netherlands. This narrow focus limits the generalizability of the findings and raises questions about whether the strategies and challenges identified would hold true in other contexts, such as non-technical universities, or institutions outside the Netherlands. Future studies should expand the sample to include a more diverse range of institutions, including universities of applied sciences and organizations in different cultural and educational settings.

Second, this study relied solely on the perspectives of instructors, which, while valuable, represent only one side of the teaching and learning process. Including student perspectives in future research could provide a more balanced understanding of how teaching strategies impact learning outcomes and reveal areas where student needs and instructor priorities may diverge.

Third, the growing role of generative AI tools like ChatGPT introduces both opportunities and challenges for ML education. While these tools have the potential to enhance learning, they also pose risks to academic integrity and authentic engagement. Future research should explore how such tools influence student behavior, evaluate their impact on learning outcomes, and develop strategies for integrating them responsibly into ML curricula. Additionally, studies could investigate innovative assessment methods that ensure students actively engage with course material despite the availability of AI-driven solutions.

Finally, future research could explore differences in pedagogical strategies between courses designed for computer science students and those tailored for non-majors. Such a comparative analysis would provide valuable insights into how teaching strategies adapt to different student populations, enriching our understanding of ML education across disciplines.

By addressing these limitations, future research can build on the findings of this study to provide a more comprehensive understanding of how to effectively teach ML to non-majors in an increasingly interdisciplinary and technology-driven world.

7 Conclusion

This study examined how instructors teach machine learning to non-majors, offering insights into effective strategies and the challenges they face. By tailoring content to students' fields, using real-world applications, fostering active learning, and leveraging visualization techniques, educators create student-centered learning environments that make ML more accessible and relevant. These strategies, while effective, are often accompanied by challenges such as varied student preparedness, conceptual misunderstandings, and the complexities introduced by tools like ChatGPT.

The contributions of this study lie in its focus on the unique context of teaching ML to non-majors. By highlighting both effective strategies and persistent barriers, it provides practical guidance for educators and curriculum developers. Tailored teaching, targeted support for foundational gaps, and innovative assessments are key to making ML education more inclusive and impactful for diverse learners.

Future research should expand beyond the limitations of this study, exploring ML pedagogy in varied institutional, disciplinary, and geographic contexts. Incorporating student perspectives would also offer a richer understanding of how teaching strategies are experienced in practice. Moreover, as AI tools become increasingly integrated into education, research is needed to explore their potential to enhance learning while ensuring academic integrity.

Equipping non-majors with ML knowledge is essential in today's interdisciplinary landscape. By addressing these challenges and building on strategies identified in this study, educators can empower students to apply ML effectively in their respective fields, driving innovation and solving real-world problems.

References

Acquaviva, V. (2022). Teaching machine learning for the physical sciences: A summary of lessons learned and challenges. *The Second Teaching Machine Learning and Artificial Intelligence Workshop*, 35–39.

Chenoweth, S., & Linos, P. K. (2018). Qualitative findings from an online course on machine learning. *2018 IEEE Frontiers in Education Conference (FIE)*, 1–9.

Creswell, J. W. (2012). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (4th). Pearson.

Dey, I. (2003). *Qualitative data analysis: A user friendly guide for social scientists*. Routledge.

Fiebrink, R. (2019). Machine learning education for artists, musicians, and other creative practitioners. *ACM Transactions on Computing Education (TOCE)*, *19*(4), 1–32.

Ishikawa, F., & Yoshioka, N. (2019). How do engineers perceive difficulties in engineering of machine-learning systems? - questionnaire survey. *2019 IEEE/ACM Joint 7th International Workshop on Conducting Empirical Studies in Industry (CESI) and 6th International Workshop on Software Engineering Research and Industrial Practice (SERIP)*, 2–9. https://doi.org/10.1109/CESSER-IP.2019.00009

Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science*, *349*(6245), 255–260.

Linos, P. K., & Chenoweth, S. (2020). Teaching machine learning: A step towards introducing computing in undergraduate education. *2020 IEEE 32nd Conference on Software Engineering Education and Training (CSEE&T)*, 1–2.

Locke, J., & Rainer, K. (2024). Teaching artificial intelligence and machine learning to” everyone else”.

Mike, K., & Hazzan, O. (2022). Machine learning for non-majors: A white box approach. *Statistics Education Research Journal*, *21*(2), 10–10.

Miles, M. B. (1994). Qualitative data analysis: An expanded sourcebook. *Thousand Oaks*.

Rashid, A. B., & Kausik, M. A. K. (2024). Ai revolutionizing industries worldwide: A comprehensive overview of its diverse applications. *Hybrid Advances*, *7*, 100277. https://doi.org/https://doi.org/10.1016/j.hybadv.2024.100277

Sulmont, E., Patitsas, E., & Cooperstock, J. R. (2019a). Can you teach me to machine learn? an exploration of pedagogical content knowledge for teaching machine learning to non-majors. *Proceedings of the 50th ACM Technical Symposium on Computer Science Education (SIGCSE ’19)*, 948–954. https://doi.org/10.1145/3287324.3287392

Sulmont, E., Patitsas, E., & Cooperstock, J. R. (2019b). What is hard about teaching machine learning to non-majors? insights from classifying instructors’ learning goals. *ACM Transactions on Computing Education (TOCE)*, *19*(4), 1–16. https://doi.org/10.1145/3336124

A Appendix: Interview Questions

The following questions were used during the semi-structured interviews conducted with instructors teaching machine learning to non-majors:

1. Could you briefly describe the context in which you teach machine learning?
2. What preconceptions do students typically have about AI and ML when they begin the course?
3. What pedagogical strategies do you use to teach AI and ML concepts to your students?
4. What specific techniques do you use to simplify ML concepts for students?
5. How do you relate AI and ML concepts to applications in your field?
6. What are the main challenges or barriers students face when learning and applying ML concepts in your course? (conceptual, practical, or domain-specific)
7. What strategies have been most effective in helping students overcome these challenges?
8. What are the most common mistakes students make when learning ML?
9. How do you track and evaluate whether students are progressing in their understanding and application of ML concepts?
10. How do you gather and use feedback from students to refine your teaching strategies or course content? Can you provide an example of a change you made based on student feedback?
11. What has been the most rewarding or challenging aspect of teaching this course?
12. What advice would you give to other instructors teaching ML to non-majors?

B Appendix: Prompts Used with ChatGPT

The following prompts were used during various stages of the research to generate ideas, guide interviews, rephrase text, and support thematic analysis:

1. What are some good interview questions to ask instructors about teaching strategies, challenges, and progress tracking in machine learning education for non-majors?
2. Can you give me tips for conducting interviews with university instructors to get meaningful insights into their teaching experiences?
3. What steps should I follow to analyze interviews and identify meaningful patterns or themes?
4. What are some commonly used qualitative analysis tools for thematic analysis, and how do they compare in terms of features and usability?
5. How can I refine and organize my initial codes into themes for a qualitative analysis?
6. Suggest ways to improve the flow and clarity of this text....
7. Help me rephrase this sentence to make it more formal....

C Appendix: Thematic Analysis

C.1 network of Themes and Codes

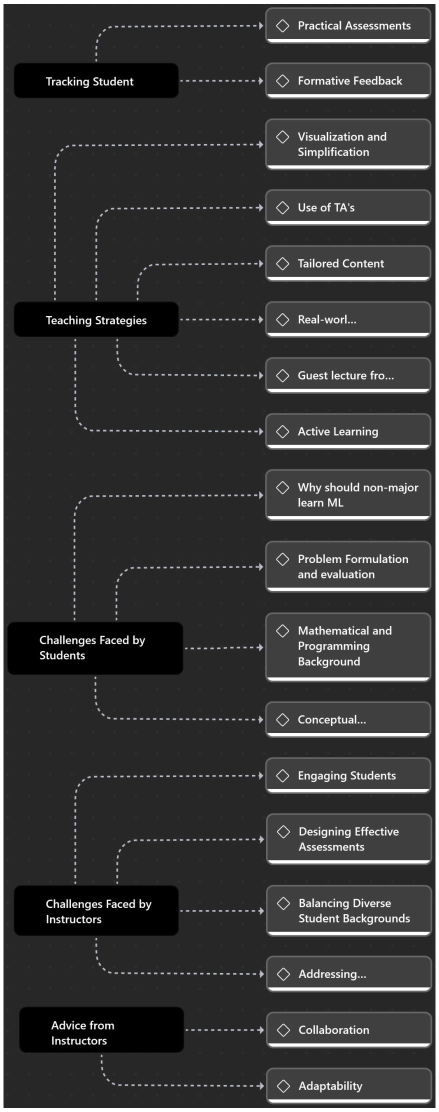
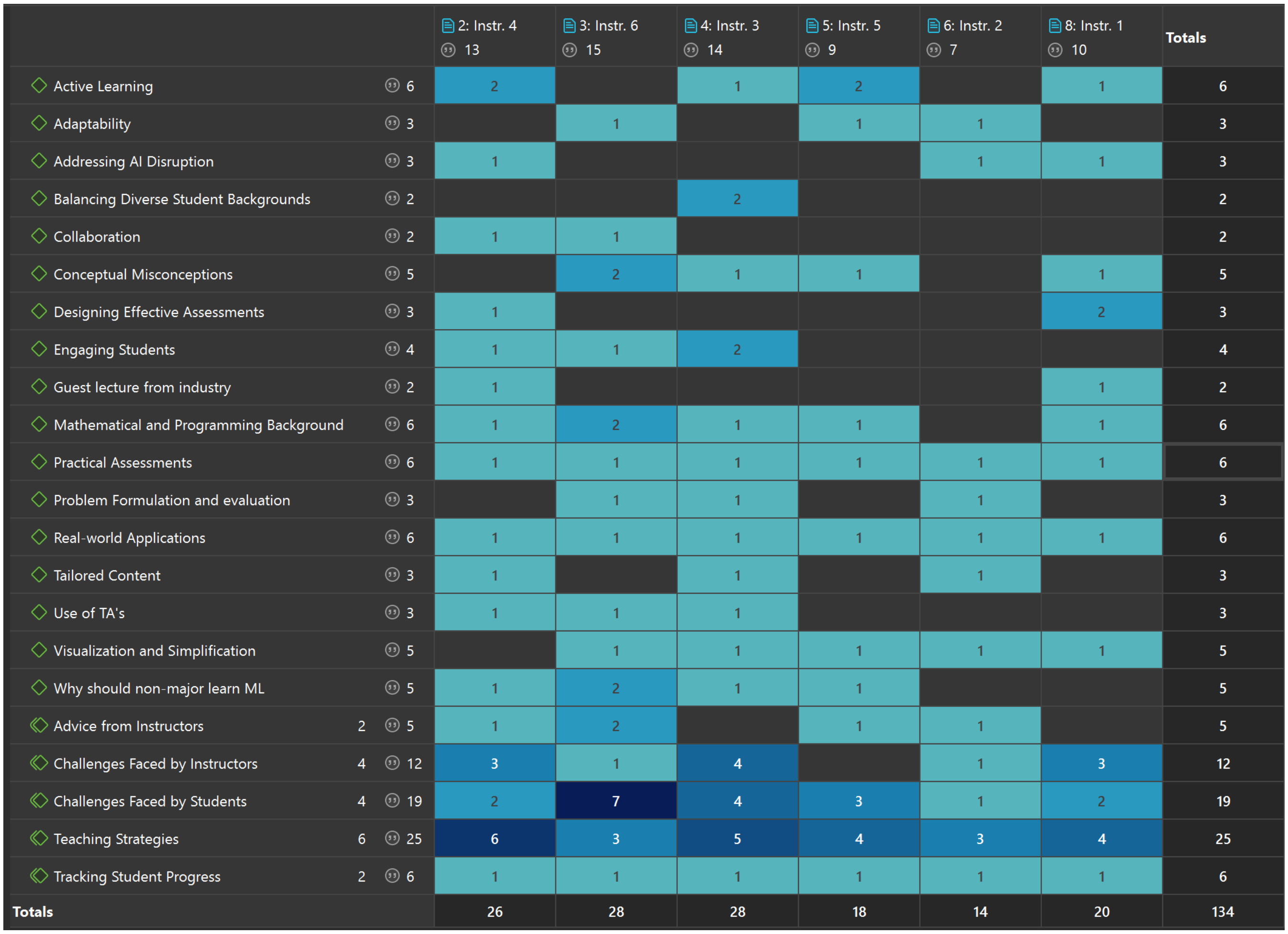


Figure 3: Diagram Showing the network of Themes and Codes

C.2 Table of Codes and themes frequencies



			2: Instr. 4 13	3: Instr. 6 15	4: Instr. 3 14	5: Instr. 5 9	6: Instr. 2 7	8: Instr. 1 10	**Totals**
Active Learning		6	2		1	2		1	6
Adaptability		3		1		1	1		3
Addressing AI Disruption		3	1				1	1	3
Balancing Diverse Student Backgrounds		2			2				2
Collaboration		2	1	1					2
Conceptual Misconceptions		5		2	1	1		1	5
Designing Effective Assessments		3	1					2	3
Engaging Students		4	1	1	2				4
Guest lecture from industry		2	1					1	2
Mathematical and Programming Background		6	1	2	1	1		1	6
Practical Assessments		6	1	1	1	1	1	1	6
Problem Formulation and evaluation		3		1	1		1		3
Real-world Applications		6	1	1	1	1	1	1	6
Tailored Content		3	1		1		1		3
Use of TA's		3	1	1	1				3
Visualization and Simplification		5		1	1	1	1	1	5
Why should non-major learn ML		5	1	2	1	1			5
Advice from Instructors	2	5	1	2		1	1		5
Challenges Faced by Instructors	4	12	3	1	4		1	3	12
Challenges Faced by Students	4	19	2	7	4	3	1	2	19
Teaching Strategies	6	25	6	3	5	4	3	4	25
Tracking Student Progress	2	6	1	1	1	1	1	1	6
Totals			26	28	28	18	14	20	134

Figure 4: Table of Codes and themes frequencies