

Embedding Responsible AI in a Technical Machine Learning Course

Stefania Zourlidou and Frank Hopfgartner

Institute for Web Science and Technologies, University of Koblenz, Koblenz, Germany

zourlidou@uni-koblenz.de

hopfgartner@uni-koblenz.de

Abstract: Machine learning (ML) courses in higher education usually emphasise mathematical foundations and algorithmic methods, while ethical and societal questions are often treated as separate or optional topics. This paper reports a small-scale practice-based evaluation of a blended design that embedded responsible AI within a Master's-level ML and Data Mining course at a German university in winter semester 2025/26. The design combined fifteen on-campus lectures on core ML topics, including one lecture on fairness as a technical topic, fifteen optional weekly assignments discussed in tutorials, and an eight-module microlearning course on AI literacy and responsible AI. Completion of at least half of the microlearning and a mandatory practical assignment were required for exam eligibility. The study analyses an anonymous end-of-course survey, using only responses from students who explicitly consented to research use (n=35). Student perceptions were generally positive: 78% reported a good balance between theory and practice, 75% found the weekly assignments important for understanding, and 59% judged the microlearning to fit well with the on-campus lectures. For responsible AI, around two thirds reported greater confidence in identifying fairness issues, feeling better prepared to discuss ethics professionally, or perceiving ethics as integrated rather than added on. 75% intended to consider fairness and ethical implications in future ML projects, with no respondent disagreeing. Open responses suggested perceived transfer, with students naming hiring systems, facial recognition, student-risk prediction and credit or loan decision-making, and often using technical vocabulary such as training data, optimisation objectives and evaluation metrics. The paper does not claim causal impact, and findings should be interpreted in light of the small voluntary sample, the single-cohort single-institution setting, and the possibility of social-desirability bias on an end-of-course survey on ethics. Instead, it offers descriptive evidence that a lightweight microlearning layer, combined with a technical fairness lecture, can be a feasible way to connect technical ML learning with ethical reflection in practice. It also identifies revisions for the next iteration: reduce workload tension, add explicit ethical-reflection prompts to assignments, and include more practical coding work.

Keywords: Machine Learning Education, Responsible AI, AI Literacy, Microlearning, Blended Learning, Higher Education

1. Introduction

Machine learning (ML) is now a core component of computer science and data science curricula in higher education. Master's-level ML courses often prioritise mathematical foundations and algorithms such as regression, classification, clustering and neural networks. Ethical and societal questions may be acknowledged as important, but are still frequently relegated to a guest lecture, optional reading list or separate elective. This separates technical competence from ethical awareness. The rapid deployment of ML in consequential domains such as hiring, lending, healthcare and education makes this separation difficult to defend. Documented intersectional accuracy disparities in commercial gender classification (Buolamwini and Gebru, 2018), algorithmic bias in educational ML systems (Baker and Hawn, 2022) and harms in public-service decision-making (Eubanks, 2018) show that fairness, transparency and accountability are not optional add-ons. Policy frameworks such as the EU AI Act (European Parliament and Council of the European Union, 2024), together with AI-literacy frameworks (Long and Magerko, 2020; Ng et al, 2021), increase the importance of preparing graduates to reason about these properties as well as model accuracy. Computer science education likewise calls for ethics to be integrated into technical programmes (Grosz et al, 2019; Saltz et al, 2019; Ferrer et al, 2021). A parallel development in e-learning offers a delivery vehicle. Blended learning combines online and face-to-face elements (Garrison and Vaughan, 2008), while microlearning uses short, focused online units (Hug, 2005; Leong et al, 2021). In this design, microlearning was used to add case studies, societal context and AI literacy while keeping the technical lecture sequence intact. Generative AI has further intensified discussion of responsible AI education (Kasneci et al, 2023).

This paper reports on a blended course design that takes these developments seriously. In the winter semester 2025/26, a Master's-level ML and Data Mining course at a German university combined three components: (i) fifteen on-campus lectures covering technical ML topics from optimisation to neural networks, including a dedicated lecture on fairness, (ii) fifteen optional weekly assignments combining knowledge questions, practical problems and programming tasks, discussed in weekly tutorials, and (iii) an eight-module microlearning course on AI literacy and responsible AI. Partial completion of the microlearning was required for exam eligibility, alongside a mandatory practical assignment. We report on three questions: how students perceive the three-

part design, how they report their framing of ML and confidence in identifying fairness issues, and what ethical examples they describe. The contribution is descriptive and practice-based, not causal.

2. Related Work

Ethics has been recognised in formal computing-curriculum guidance (Joint Task Force on Computing Curricula, 2013), but an isolated "ethics module" is widely seen as insufficient. Grosz et al (2019) describe Harvard's embedded ethics model, in which short ethics sessions are co-taught within technical CS courses. Other authors argue that judgements about data, evaluation metrics and deployment cannot be separated from technical design (Saltz et al, 2019; Ferrer et al, 2021; Borenstein and Howard, 2021). In ML, the fair-ML literature provides quantitative fairness definitions, bias sources and mitigation approaches that can be taught as mathematical and engineering content (Mehrabi et al, 2021), and course reports show how ethics activities can be integrated into programming and CS classrooms (Fiesler et al, 2021; Skirpan et al, 2018).

Blended learning combines online and face-to-face elements within one course design (Garrison and Vaughan, 2008). Microlearning describes short, focused online units, typically a few minutes of video followed by a brief assessment (Hug, 2005; Leong et al, 2021). Such units can support focused objectives when aligned with cognitive-load principles (Sweller, van Merriënboer and Paas, 2019), and they offer a low-friction way to add content layers to existing courses (Sankaranarayanan et al, 2023). AI literacy - the ability to critically evaluate, communicate and collaborate with, and use AI (Long and Magerko, 2020) - has emerged as a distinct competence, with later work refining its operationalisation in formal education (Ng et al, 2021; Casal-Otero et al, 2023). General-public AI literacy does not substitute for the technical literacy needed by CS students, and this design attempts to bridge that gap.

Recent work has moved from broad calls for responsible-AI education to concrete curricular interventions. Aler Tubella, Mora-Cantalops and Nieves (2024) identify competencies and implementation challenges for responsible-AI teaching in higher education. Weichert and Eldardiry (2025) evaluate a two-lecture AI-policy module in a graduate introductory machine-learning course, while Jayapal, Hammerman and Chaspari (2025) report a pilot AI-ethics curriculum embedded in graduate-level machine learning. Although the two curricular interventions focus mainly on discrete modules, the present study examines a broader blended configuration combining a technical fairness lecture with sustained microlearning.

Evaluating educational designs in authentic course settings is methodologically constrained: experimental designs are rarely feasible because of small cohorts, ethical concerns about withholding instruction, and the entanglement of multiple course components. Practice-based and design-based approaches (Reeves, 2006; Anderson and Shattuck, 2012) instead emphasise iterative description of what was done and how it was perceived. We adopt this stance by describing a concrete design and reporting what students said about it, while being explicit about the limits of inference.

3. Course Design

3.1 Course Context

The course is the Master's-level Machine Learning and Data Mining course taught in winter semester 2025/26 at a German university. Its main target programme is a data-science-oriented Master's programme. It also serves students from other Master's programmes, including an applied mathematical modelling programme. Prerequisites include undergraduate linear algebra, calculus, probability and statistics, and basic Python. Assessment is by a written examination. Exam eligibility depends on two requirements, which are passing a mandatory practical assignment and completing at least half of the microlearning course described below.

3.2 Component A: Lectures

Fifteen 90-minute on-campus lectures span the technical ML curriculum: introduction, mathematical optimisation (two lectures), linear regression, linear two-class classification (two lectures), linear multi-class classification, clustering, feature engineering, non-linear feature engineering, dimensionality reduction and principal components, tree-based learners, optimising ML models, fairness, considerations and limits in ML, and neural networks. Lecture 14 ("Fairness, considerations and limits in ML") is positioned within the technical sequence and not as a separate ethics day and treats fairness as a quantitative topic: group-fairness metrics, sources of bias in training data, mitigation strategies, and limitations.

3.3 Component B: Assignments and Tutorials

Fifteen weekly assignments accompany the lectures. They are optional in the sense that they are not marked individually, but students are recommended to complete at least 50% of each. Each assignment combines knowledge questions, conceptual problems, and small programming tasks. Solutions are discussed in weekly tutorials led by teaching assistants. A separate mandatory practical assignment, distinct from the optional weekly ones, must be passed at 50% to qualify for the exam.

3.4 Component C: Microlearning Course “Machine Learning Made Easy”

Component C was a microlearning course hosted on the institution’s learning management system. It comprised a pre-course questionnaire, eight content modules, and a post-course assessment. Each module paired a short video with an assessment (about 30-40 minutes’ work), for a total of about six to seven hours. The modules were: (1) Everyday Intelligence: What is AI, (2) Data Matters: The Fuel of AI, (3) How Machines Learn from Data, (4) Teaching the Machine, (5) When Machines Judge, (6) Personalisation and Evaluation, (7) Truth, Trust, and Misinformation, and (8) Responsible AI. Content emphasised AI literacy and the societal dimensions of ML, i.e., how data shapes models, where bias enters, how recommendation systems work, how misinformation can be amplified, and what responsible AI practice entails. Completion of at least 50% of the modules was required for exam eligibility.

3.5 Design Rationale

The design uses two entry points to responsible AI. Lecture 14 treats fairness as part of technical modelling. The microlearning adds AI literacy and societal context without crowding out the technical curriculum. The assignments and tutorials provide practice and feedback. Together, the components form a lightweight model for embedding responsible AI in a technical ML course.

4. Methodology

An end-of-course questionnaire comprising 52 numbered items plus an initial research-consent item was administered through the institutional learning management system from late January to mid-March 2026. Responses should therefore be read as end-of-course perceptions and not as a single post-exam measurement. The questionnaire covered demographics, engagement with lectures, tutorials, assignments and microlearning, generative-AI use and perceptions, course design, responsible-AI learning, overall satisfaction, and four open-ended items. Likert items used a five-point scale (Strongly disagree to Strongly agree). The survey was voluntary and anonymous. As the initial item, respondents were asked whether they consented to research use and possible publication of their fully anonymised responses. Only respondents who explicitly consented are analysed here. The survey received 51 valid submissions: 35 consented, 8 did not consent, and 8 left the consent item blank. All analyses are restricted to the consenting subset ($n=35$). Missing or invalid export values were treated as missing, and percentages were calculated item-wise from available consenting responses. Consequently, valid n varies between items. No direct identifiers were collected and demographic reporting is aggregated.

Of the 35 consenting respondents, 26 were enrolled in the main data-science-oriented Master’s programme, 3 in an applied mathematics/modelling programme, 5 in other Master’s programmes, and 1 did not answer the programme item. Twenty-five were in their first or second semester. Eighteen had taken at least one prior AI or ML course, fifteen had not, and two did not answer. The modal age category was 23-25 ($n=15$), followed by 26-30 ($n=11$), 20-22 ($n=6$), 31+ ($n=2$), and one non-response. Twenty-three identified as men, eight as women, three preferred not to self-describe or not to say, and one did not answer. As with any voluntary end-of-course survey, the consenting subset is likely affected by self-selection: students who disengaged from the course or held strongly negative views may be under-represented, and we do not have administrative data to compare the subset against the full active cohort.

Quantitative analysis is descriptive. For each Likert item we report the percentage agreeing (Agree and Strongly agree), neutral, and disagreeing on a 1-5 scale. Percentages may not sum to 100 because of rounding. Where means are reported, they are used as descriptive shorthand only. Likert responses are treated as ordinal throughout. We do not perform inferential tests, both because the descriptive question is the right one for practice-based evaluation at this scale and because the sample is too small for robust subgroup inference. Q40 (“Before this course, I mainly saw ML as a technical field”) and Q41 (“After this course, I see ML as a socio-

technical field") are presented side by side as a retrospective comparison of two differently worded constructs. No within-person pre/post effect is inferred.

Qualitative analysis used a lightweight inductive coding approach informed by Braun and Clarke (2006) for Q47 and Q52. The first author coded responses in two passes: an initial pass identified categories close to students' wording, and a second pass consolidated them into the domain and workload/integration groupings reported in Section 5.4. Responses such as "no", "none" and "N/A" were classified as non-substantive and were not included in the thematic categories. No inter-rater coding was undertaken.

5. Findings

5.1 Engagement with the Three Components

Self-reported engagement is summarised in Figure 1. Microlearning completion was concentrated at the top end: 22 of 34 respondents (65%) completed all eight modules and 26 of 34 (76%) completed at least six. Two reported 0-2 modules, suggesting that not all survey respondents had completed the exam-eligibility requirement at the time of response or intended to sit the exam. Weekly assignments showed a wider spread: 11 of 32 respondents completed 12-15, 5 completed 8-11, 11 completed 4-7, and 5 completed 1-3. Weekly hours ranged from 2-4 hours (10 respondents) to 5-7 (13), 8-10 (5), and more than 10 (6). All 33 respondents who answered the mandatory-assignment question reported submitting it (94% of the consenting sample). Lecture and tutorial attendance were mixed.

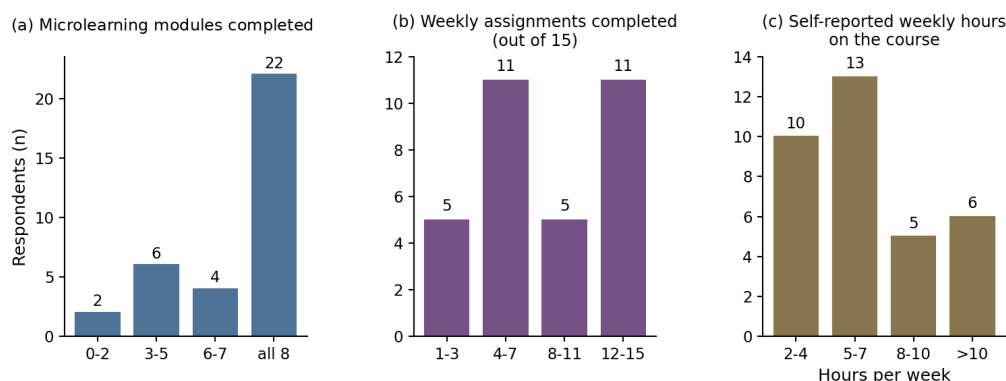


Figure 1: Self-reported engagement with the three course components (consenting respondents, n=35, valid responses per item shown in panels). The microlearning, gated by a 50% exam-eligibility threshold, generated high completion rates, with most respondents completing six or more modules. The optional weekly assignments show the diverse engagement profile typical of non-mandatory work.

5.2 Perceptions of the Course Design (RQ1)

Likert distributions for nine course-design items are shown in Figure 2. The design is broadly endorsed. 78% agreed or strongly agreed that the course provided a good balance between theory and practice (Q26). 75% agreed that the optional weekly assignments were important for their understanding (Q31), the second-highest endorsement of any design item, and 69% agreed that the practical components connected theory and practice (Q28). The microlearning fit item was positive but less emphatic: 59% agreed that the microlearning fitted well with the on-campus lectures (Q33). The tutorials received more measured endorsement (53% agree, 44% neutral, Q32), consistent with the variable attendance reported in Section 5.1.

Two additional items addressed balance directly: Q29 ("too theoretical") split the cohort (44% agree, 41% neutral, 16% disagree), while Q30 ("too tool-focused") was disagreed with by 41% and agreed with by 16%, suggesting that imbalance was felt mainly in the theory direction. Workload (Q34) had the highest disagreement rate among the positively framed design items: 56% agreed it was manageable but 12% disagreed, with qualitative feedback linking this to the lecture-plus-microlearning combination. Confidence outcomes were positive: 62% agreed their confidence in applying ML to new datasets had increased (Q35) and 69% agreed their confidence in explaining ML to non-experts had increased (Q36), with no disagreement on Q35 and only 3% on Q36.

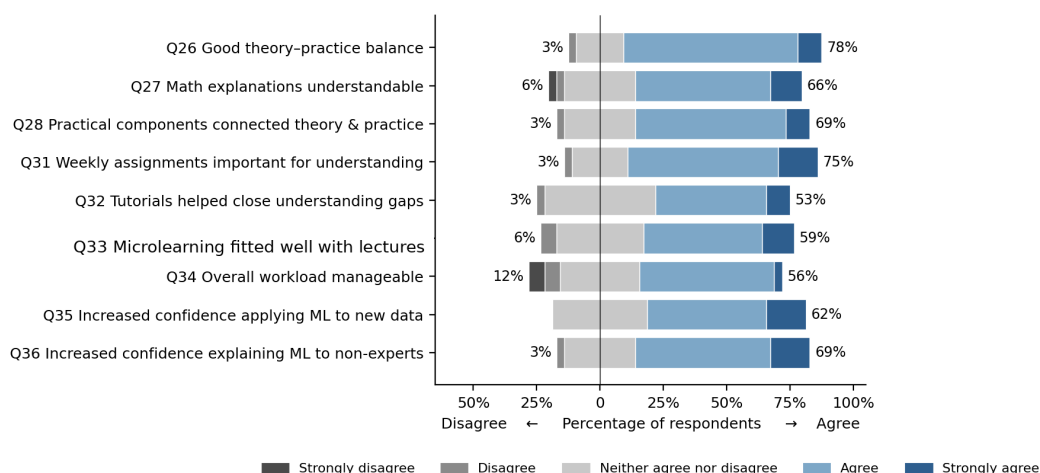


Figure 2: Student perceptions of the blended course design. Percentages on the right show agree and strongly agree. Percentages on the left show strongly disagree and disagree when not negligible (n=32 valid responses per item).

5.3 Perceptions of Responsible AI Learning (RQ2)

Likert distributions for the seven core responsible-AI items are shown in Figure 3. The retrospective comparison of Q40 and Q41 is presented separately in Figure 4. Three observations stand out.

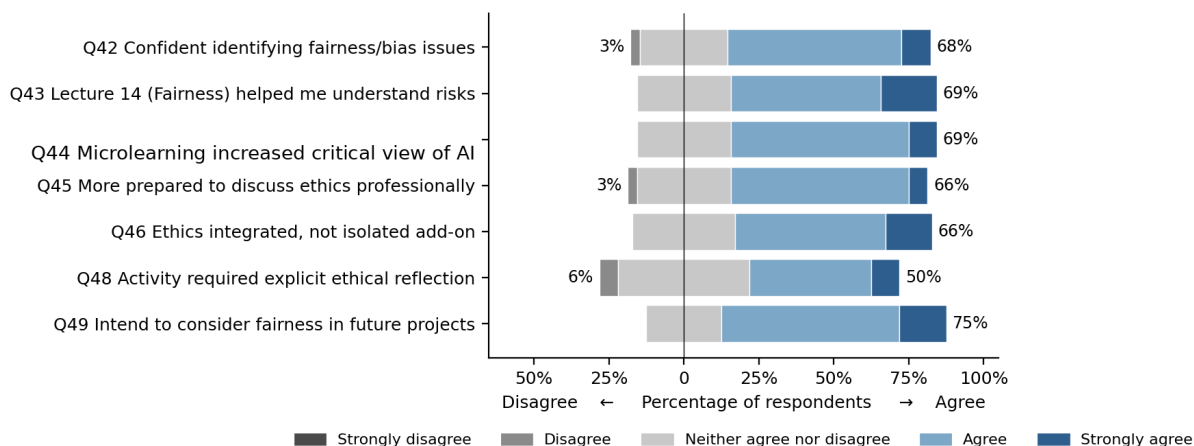


Figure 3: Student perceptions of responsible-AI learning. n=31-32 valid responses per item.

First, the strongest single result in the survey is the future-intention item (Q49): 75% of respondents intend to actively consider fairness and ethical implications in future ML projects, with not a single respondent disagreeing. While stated intentions are not the same as future behaviour, the absence of any disagreement on a voluntary survey suggests that the ethical framing was broadly accepted rather than resisted.

Second, the two responsible-AI components received comparably high endorsement. 69% agreed that Lecture 14 helped them understand ethical risks (Q43), and 69% agreed that the microlearning made them more critical of everyday AI systems (Q44). Around two thirds agreed they were more confident identifying fairness or bias issues (Q42, 68%), better prepared to discuss ethics professionally (Q45, 66%), and that ethics felt integrated rather than isolated (Q46, 66%).

Third, the weakest item in the ethics block is Q48 ("at least one assignment or activity required me to explicitly reflect on ethical or fairness aspects of ML"): only 50% agreed, with 44% neutral. This is a useful diagnostic. Although Lecture 14 and the microlearning carry the ethical content, the weekly assignments do not currently include an explicit reflection prompt. Half the cohort therefore does not perceive an activity that required ethical reflection, even though many of those same respondents endorse Q46 and Q49. The disconnect suggests a concrete revision target: add a short ethical-reflection prompt to one or two weekly assignments.

Figure 4 reports the two retrospective ML-framing items descriptively. Before the course, 53% of respondents agreed they had mainly seen ML as a purely technical field, 31% were neutral and 16% disagreed. After the course, 53% agreed they now see ML as a socio-technical field, 41% were neutral and 6% disagreed. Because Q40 and Q41 assess differently worded constructs, no within-person pre/post change is inferred; the distributions are reported as complementary indicators of students' perceived framing of ML. Both items are also vulnerable to recall bias and socially desirable responding on an end-of-course survey on ethics. The qualitative responses in Section 5.4 provide a separate line of descriptive evidence.

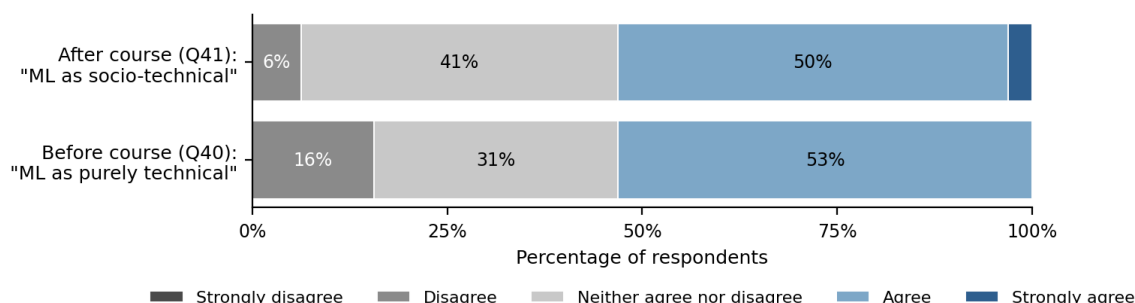


Figure 4: Retrospective comparison of two differently worded ML-framing items (Q40 and Q41), n=32 valid responses per item.

5.4 Qualitative Evidence of Perceived Transfer (RQ3)

Q47 asked students to describe one example, from the course or from their own life, where they now see fairness or ethical issues in an ML or AI system. This open-ended question was used to examine whether students could transfer the responsible-AI themes from the course to concrete examples. Thirteen respondents provided a non-blank entry. Nine entries contained substantive examples, one was unrelated, and three were non-substantive responses or declines, such as "No" and "N/A". Only the nine substantive responses were included in the thematic categorisation. They clustered into four domains: hiring and CV screening (five responses), student-risk prediction in higher education (two), facial recognition (one), and credit or loan decision-making (one). Students described, for example, how historical hiring bias can enter training data and how attendance or platform-activity features can disadvantage students with jobs, caregiving responsibilities or unreliable internet access. Three features are worth noting: first, several examples were linked to course components, including the tutorial hiring case and the microlearning module on personalisation and judgement. Second, the facial-recognition example suggests transfer to public discourse. Third, several responses used technical language from the course, such as training data, optimisation objectives, evaluation metrics, fairness and transparency.

Q52, the open feedback item, drew thirteen responses, seven of them substantive. Of these, five mentioned workload as too high, with three of the five linking this partly to the lecture-plus-microlearning combination. One called for more practical coding work, in line with Q37, and one clearly praised the integration of ethical content. Negative feedback concentrated on workload and the desire for more practical coding work.

5.5 Overall Satisfaction

Overall satisfaction with what was learned was high: 72% agreed or strongly agreed (Q50, mean 3.97 on a 1-5 scale), and 75% would recommend the combination of on-campus lectures, assignments, and microlearning to other students (Q51, mean 4.09). The recommendation item received the highest mean of any Likert item in the survey, with twelve respondents choosing "strongly agree".

6. Discussion

Three results form the core of what was perceived as working. First, 66% agree that ethics felt integrated rather than added on, with zero disagreement. Second, the future-intention finding (75% intend to consider fairness in future projects; 0% disagreement) suggests broad acceptance of the ethical framing, while not demonstrating future behaviour. Third, the qualitative examples (hiring systems, student-risk prediction, facial recognition, and credit or loan decisions) show perceived transfer beyond the specific course examples.

Workload is the clearest design problem. Although a majority report the workload as manageable, 12% disagree (the highest rate among the positively framed course-design items), and qualitative responses attribute the issue partly to the lecture-plus-microlearning combination. The microlearning enabled AI-literacy and societal coverage but added load on top of fifteen lectures. Possible mitigations include reducing the microlearning module count, aligning modules more tightly with the lecture schedule, or making one or two modules optional. A second issue is the perception that no activity required explicit ethical reflection (Q48: only 50% agree). The current design relies on Lecture 14, the microlearning and tutorial discussion to carry ethical reflection. The optional assignments do not include reflection prompts. Adding a short prompt, for example, "Identify one fairness consideration in your solution", to one or two assignments is a low-cost change that targets this gap directly. A third issue, recurring across Q37 and Q52, is the desire for more practical coding work, especially in service of the mathematical material - this is a familiar tension that deserves attention in any next iteration.

Three limitations bound what can be inferred. First, the sample is small ($n=35$), voluntary and self-selected. Second, Q40 and Q41 are retrospective, differently worded items and are vulnerable to recall bias; they are reported descriptively only. Third, the study is single-cohort, single-institution and single-course, so we cannot claim causal impact or generalisability. A further important caveat is social-desirability bias: an end-of-course survey explicitly framed around ethics may encourage respondents to over-report agreement with ethically aligned statements such as Q49. Open-ended responses were completed outside a controlled setting, and possible use of generative-AI tools could not be ruled out; they are therefore interpreted as submitted examples of perceived transfer rather than independently verified evidence of unaided competence. Additional limitations include instructor effects (the first author was involved in both course delivery and qualitative coding, which may have influenced both the design and the interpretation of responses) and possible LMS-related response bias. Future iterations could add longitudinal follow-up, cross-cohort comparison and interviews on how technical and ethical content interact in students' reasoning.

7. Conclusion

This paper described a blended Machine Learning course design embedding responsible AI through a technical fairness lecture, an eight-module AI-literacy microlearning course, and weekly optional assignments. Among 35 consenting respondents, perceptions were broadly positive and students generated ethical examples from hiring, education, identification, and credit or loan systems using technical vocabulary from the course. We do not claim causal impact. The findings are nevertheless consistent with a workable approach in this context: a fairness lecture plus microlearning wrap-around appears feasible as a way to support integrated responsible-AI teaching inside a technical ML course. The most actionable revision targets, reducing workload tension, adding explicit ethical-reflection prompts, and increasing practical coding work, are concrete and low-cost. Generalisability beyond one course and cohort remains an open question.

Ethics Declaration

The survey was voluntary and anonymous, and explicit research-use consent was gathered as the first item. Only consenting responses are analysed. No direct identifiers were collected; demographic results are reported only in aggregate. Formal institutional ethical review was not sought for this study, which was conducted as a voluntary, anonymous end-of-course evaluation with explicit research-use consent collected as the first survey item.

AI Declaration

An AI language model was used solely for language revision and copy-editing. All substantive arguments, interpretations, source selections, and final wording were reviewed and approved by the authors, who take full responsibility for the final manuscript.

References

- Aler Tubella, A., Mora-Cantalops, M. and Nieves, J.C. (2024) "How to teach responsible AI in Higher Education: challenges and opportunities", *Ethics and Information Technology*, Vol 26, Article 3. doi:10.1007/s10676-023-09733-7.
- Anderson, T. and Shattuck, J. (2012) "Design-based research: A decade of progress in education research?", *Educational Researcher*, Vol 41, No. 1, pp 16–25. doi:10.3102/0013189X11428813.
- Baker, R.S. and Hawn, A. (2022) "Algorithmic bias in education", *International Journal of Artificial Intelligence in Education*, Vol 32, pp 1052–1092. doi:10.1007/s40593-021-00285-9.

- Borenstein, J. and Howard, A. (2021) "Emerging challenges in AI and the need for AI ethics education", *AI and Ethics*, Vol 1, pp 61–65. doi:10.1007/s43681-020-00002-7.
- Braun, V. and Clarke, V. (2006) "Using thematic analysis in psychology", *Qualitative Research in Psychology*, Vol 3, No. 2, pp 77–101. doi:10.1191/1478088706qp063oa.
- Buolamwini, J. and Gebru, T. (2018) "Gender shades: Intersectional accuracy disparities in commercial gender classification", in *Proceedings of the 1st Conference on Fairness, Accountability and Transparency*, *Proceedings of Machine Learning Research*, Vol 81, pp 77–91.
- Casal-Otero, L. et al (2023) "AI literacy in K-12: a systematic literature review", *International Journal of STEM Education*, Vol 10, Article 29. doi:10.1186/s40594-023-00418-7.
- Eubanks, V. (2018) *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*, St. Martin's Press, New York.
- European Parliament and Council of the European Union (2024) "Regulation (EU) 2024/1689 of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)", *Official Journal of the European Union*, L 2024/1689, 12 July. Available at: <https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng>.
- Ferrer, X. et al (2021) "Bias and discrimination in AI: A cross-disciplinary perspective", *IEEE Technology and Society Magazine*, Vol 40, No. 2, pp 72–80. doi:10.1109/MTS.2021.3056293.
- Fiesler, C. et al (2021) "Integrating ethics into introductory programming classes", in *Proceedings of the 52nd ACM Technical Symposium on Computer Science Education (SIGCSE '21)*, ACM, New York, pp 1027–1033. doi:10.1145/3408877.3432510.
- Garrison, D.R. and Vaughan, N.D. (2008) *Blended Learning in Higher Education: Framework, Principles, and Guidelines*, Jossey-Bass, San Francisco.
- Grosz, B.J. et al (2019) "Embedded EthiCS: integrating ethics across CS education", *Communications of the ACM*, Vol 62, No. 8, pp 54–61. doi:10.1145/3330794.
- Hug, T. (2005) "Micro learning and narration. Exploring possibilities of utilization of narrations and storytelling for the designing of micro-units and didactical micro-learning arrangements", in *Proceedings of the Fourth Media in Transition Conference*, MIT, Cambridge, MA.
- Jayapal, I.V., Hammerman, J.K.L. and Chaspari, T. (2025) "Expanding AI Ethics in Higher Education Technical Curricula: A Study on Perceptions and Learning Outcomes of College Students", in *Proceedings of the 2025 ASEE Annual Conference and Exposition*, Montreal, Quebec, Canada. doi:10.18260/1-2--56476.
- Joint Task Force on Computing Curricula (2013) *Computer Science Curricula 2013: Curriculum Guidelines for Undergraduate Degree Programs in Computer Science*, ACM and IEEE Computer Society, New York. doi:10.1145/2534860.
- Kasneci, E. et al (2023) "ChatGPT for good? On opportunities and challenges of large language models for education", *Learning and Individual Differences*, Vol 103, 102274. doi:10.1016/j.lindif.2023.102274.
- Leong, K. et al (2021) "A review of the trend of microlearning", *Journal of Work-Applied Management*, Vol 13, No. 1, pp 88–102. doi:10.1108/JWAM-10-2020-0044.
- Long, D. and Magerko, B. (2020) "What is AI literacy? Competencies and design considerations", in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, ACM, New York, pp 1–16. doi:10.1145/3313831.3376727.
- Mehrabi, N. et al (2021) "A survey on bias and fairness in machine learning", *ACM Computing Surveys*, Vol 54, No. 6, pp 1–35. doi:10.1145/3457607.
- Ng, D.T.K. et al (2021) "Conceptualizing AI literacy: An exploratory review", *Computers and Education: Artificial Intelligence*, Vol 2, 100041. doi:10.1016/j.caeai.2021.100041.
- Reeves, T.C. (2006) "Design research from a technology perspective", in van den Akker, J. et al (eds.) *Educational Design Research*, Routledge, London, pp 52–66.
- Saltz, J. et al (2019) "Integrating ethics within machine learning courses", *ACM Transactions on Computing Education*, Vol 19, No. 4, Article 32, pp 1–26. doi:10.1145/3341164.
- Sankaranarayanan, R. et al (2023) "Microlearning in diverse contexts: a bibliometric analysis", *TechTrends*, Vol 67, No. 2, pp 260–276. doi:10.1007/s11528-022-00794-x.
- Skirpan, M. et al (2018) "Ethics education in context: a case study of novel ethics activities for the CS classroom", in *Proceedings of the 49th ACM Technical Symposium on Computer Science Education (SIGCSE '18)*, ACM, New York, pp 940–945. doi:10.1145/3159450.3159573.
- Sweller, J., van Merriënboer, J.J.G. and Paas, F. (2019) "Cognitive architecture and instructional design: 20 years later", *Educational Psychology Review*, Vol 31, pp 261–292. doi:10.1007/s10648-019-09465-5.
- Weichert, J. and Eldardiry, H. (2025) "Educating a Responsible AI Workforce: Piloting a Curricular Module on AI Policy in a Graduate Machine Learning Course", in *Proceedings of the 2025 ASEE Annual Conference and Exposition*, Montreal, Quebec, Canada. doi:10.18260/1-2--56314.