Navigating the Pedagogical Landscape
An Exploration of Machine Learning Teaching Methods

Andreea Zlei

Supervisor(s): Gosia Migut

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 31, 2024

Name of the student: Andreea Zlei
Final project course: CSE3000 Research Project
Thesis committee: Gosia Migut

Abstract

This study delves into machine learning (ML) education by conducting a comprehensive literature review, a targeted survey of ML lecturers in Dutch universities, and a comparative experiment. These methods aid in addressing the challenges of aligning teaching methods with the evolving nature of ML and the growing demands of the field, and fill in knowledge gaps on the success of different teaching methods in ML education. The paper investigates whether traditional methods are effective in equipping future engineers with the necessary skills for tomorrow's challenges, amidst the rapid advancement of ML and its applications. The literature review explores the range of teaching methods in ML education and not only, emphasizing a shift towards technology-enhanced and active learning approaches when teaching ML. A survey of ML lecturers explores the landscape of ML education in Dutch universities. The study investigates teaching methodologies, tools, and challenges, providing valuable insights into the evolving practices of ML instruction. Findings indicate a predominant trend towards adopting a blended approach, with lectures, projects, and group work forming core instructional methods. Virtual environments, active learning strategies, and staying informed through community engagement are highlighted. Word frequency and thematic analyses reveal key themes, emphasizing student-centric learning, practical application, and the integration of diverse teaching methods. Additionally, an experimental comparison of two teaching methods, lecture and jigsaw, sheds light on their seemingly similar efficacy when applied to the domain of ML education. The research contributes to the optimization of ML education practices, offering comprehensive insights for educators and policymakers.

1 Introduction

Are we laying a robust foundation for future engineers in machine learning education, or do we risk leaving them ill-equipped for the challenges of tomorrow? Despite the ubiquity of machine learning (ML), the escalating demand for skilled professionals in artificial intelligence, ML, and data science, and universities dedicating substantial resources to meet this growing need, machine learners' commitment to data-driven practices may not be evident in their teaching and educational approaches (Steinbach, Seibold & Guhr, 2021; Gelman & Loken, 2012). The historical reliance on traditional lecture-based teaching methods in ML courses, characterized by extensive slide presentations, reflects a choice often driven by intuitive judgment rather than research-backed pedagogical practices.

The current body of research on machine learning education outlines diverse methods and approaches. Notable contributions include Schiendorfer, Gajek, & Reif (2021), discussing the adaptation of computer science teaching techniques for ML, and Steinbach, Seibold, and Guhr (2021), delving into challenges and strategies in ML education in 2020. Technological solutions in literature emphasize a shift toward integrating technology and promoting active learning (Rattadilok, Roadknight, & Li, 2018; Chow, 2019; Kaspersen, Bilstrup, & Petersen, 2021). Despite these contributions, further research is needed to comprehend how individuals learn to create, evaluate, and improve ML-based systems (Shapiro & Fiebrink, 2019).

The identified gaps in ML education research encompass a broad spectrum of knowledge domains. Amy J. Ko (2017) identifies a content knowledge gap, highlighting the need for well-defined analogies, examples, representations, and explanations to aid learners in understanding ML concepts. Additionally, a lack of systematic understanding exists regarding challenging ML concepts for learners and the reasons behind these difficulties. Under-explored areas include insight into learners' pre-existing conceptions of ML, informal assessment methods, and awareness of common mistakes during ML application. Steinbach, Seibold & Guhr (2021) further define and classify the knowledge needed for ML education research. Filling these knowledge gaps will be essential for cultivating a comprehensive and effective ML education framework.

In order to contribute to filling these knowledge gaps, the main question this research project aims to answer is:

What are the teaching techniques documented in literature and used in machine learning courses of Dutch universities, and how do two of these techniques compare?

This research question has been broken down into the following sub-questions:

- What are the teaching methods and techniques documented in existing literature (in general, for computer science, and for ML)?
- What are the teaching methods and techniques used in Dutch universities in machine learning courses?
- How do two of these methods compare when applied to machine learning education?

The need for this research is underscored by the lack of structured, research-based approaches for ML instruction (Ko, 2017). While various teaching methods have been explored in higher education and computer science, similar comprehensive insights are notably lacking in ML.

Two central hypotheses guide this research endeavor: first, that ML lecturers in the Netherlands use a variety of teaching methods in their courses, but there is a lack of comprehensive knowledge regarding their specific approaches; and second, that the practical applications identified in the literature and in ML courses throughout Dutch universities can provide knowledge on effective ML teaching.

The research presented here offers several key conclusions that directly address the research questions posed. The literature review explores the range of teaching methods and underscores a growing preference for active learning strategies in the field of ML education. While the jigsaw method outperforms lectures according to Carpenter (2006), it hasn't been studied in the context of ML. A survey of Dutch university ML courses indicates a move towards blended learning, yet it appears that the jigsaw and other successful active learning methods are not widely implemented in practice. We propose that it is interesting to investigate whether the jigsaw method is also the most useful in the context of ML, thus we conducted an experiment that aims to compare it with lectures. This comparative experiment showed that, if we assume the small sample size to be representative of the student population, both approaches might be similarly effective in teaching ML concepts. This outcome suggests that the choice of teaching method can be flexible, tailored to specific educational contexts, and need not adhere strictly to traditional models. Collectively, these conclusions point to a dynamic educational environment in ML where innovative teaching methods are gaining traction and proving effective.

In order to answer the research question, the rest of the paper fits together as follows: section 2 describes the methodologies used in the paper, section 3 contains a literature review, in section 4 a survey with Dutch lecturers is described, section 5 describes an experiment that compares two teaching methods, section 6 is about responsible research, section 7 is a discussion on the findings, and finally, section 8 contains the conclusions drawn from the study and outlines avenues for future research.

2 Methodology

This section outlines the methodology employed to address the research questions, focusing on the two main components: literature review and empirical study, which entails a survey and an experiment. The combination of these methods allows for triangulation, enhancing the validity and reliability of the study's findings. Thus, a more nuanced and well-rounded perspective on teaching methods in ML education is offered.

2.1 Literature Review

To address the first sub-question, academic databases, journals, and relevant publications will be systematically explored to identify teaching methods and techniques in the field of machine learning education. Thus, the literature review provides a foundational understanding of documented methods and techniques. This allows for the identification of trends, emerging practices, and theoretical frameworks.

2.2 Survey on Dutch ML Lecturers

To address the sub-question pertaining to the teaching methods used in Dutch ML courses, a survey will be conducted. This allows for the extraction of real-world, context-specific data on current teaching practices in the context of Dutch universities, providing a nuanced understanding of the landscape. The reliability of the survey will be ensured through participants' anonymity in order to reduce response bias, and clear instructions minimize ambiguity and misinterpretation.

The data will be analysed by statistical means and visualisations. This approach ensures a thorough exploration of both quantitative and qualitative aspects of the collected data, contributing to the robustness and depth of the study's conclusions.

2.3 Comparative Experiment

An experiment with two groups will be conducted, in order to compare two teaching methods, one, lectures, identified as the most prominent method used in Dutch universities, and the other, jigsaw, identified as effective for ML through the literature review. This aims to answer the third sub-question. By exposing participants to different teaching methods and collecting both quantitative performance data and qualitative feedback, the practical implications and effectiveness of the instructional techniques will be assessed and compared.

3 Literature Review

Effective teaching methods play a pivotal role in shaping students' learning experiences across diverse educational settings. An effective teaching method encourages students to challenge their preconceived notions, enhancing their overall academic success, and fostering motivation to learn by placing them in scenarios where they perceive themselves as the creators of solutions and as responsible agents for instigating change (Bidabadi, Isfahani, Rouhollahi, & Khalili, 2016). As educational practices continue to evolve, understanding the significance of diverse teaching methods becomes crucial for educators and instructional designers.

This literature review gathers diverse literature on techniques in higher education, computer science, and machine learning. Moreover, Beck's (1998) taxonomy is employed for classifying the instructional methods identified during the literature research, offering a systematic approach to categorising the literature review findings. Finally, there is a comparison of teaching methods based on literature.

3.1 Teaching Methods in Higher Education

Established learning theories need to be considered in order to understand the basis of teaching methods in any field (Ertmer and Newby, 1993). The three leading human learning theories are behaviourism, cognitivism, and constructivism. Behaviorism sees learners as passive, influenced by external factors. Cognitivism, however, considers learning as an internal process focused on memory, thinking, and problem-solving. On a different note, constructivism posits that learners actively create their own understanding by building on prior knowledge and experiences, interpreting and deriving meaning from their surroundings to form mental models for new information, and actively engaging in the learning process. More recently, the constructivist philosophy has acted as a driving force behind numerous significant shifts in educational practices (Greening, 2000), and lies at the foundation of many modern teaching methods, as it challenges the traditional view of knowledge and asserts that it is an evolving understanding rather than a final answer (Ma, 2021). Teaching, from this perspective, involves guiding students to actively construct and analyze knowledge based on their experiences.

3.2 Computer Science Education

In their work presented at the Central European Conference on Information and Intelligent Systems (2011), Mohorovicic and Strcic provide a comprehensive overview and comparison of various teaching methods in computer science, addressing both their benefits and challenges. The paper underscores the impact of learning styles and motivation on the success of programming education, distinguishing between deep understanding and surface memorization, the latter of which is deemed essential in ML. The authors note that in computer science, teachers often employ their own blend of methods, with one method typically dominating, in an attempt to create a more engaging and motivating learning experience.

Several papers analyze diverse instructional methods for computer science, and discuss the importance of problem-solving, active engagement, exploration and discovery. Constructivist approaches are particularly recommended (Ben-Ari, 2001; Greening, 2000). Mohorovicic and Strcic (2011) highlight methods like problem and puzzle-based learning, pair programming, and game-themed programming. Zendler and Klaudt (2015) evaluated 20 methods, including project work and computer simulation, for their efficacy in teaching computer science. Shahid et al. (2019) explore a range of instructional techniques and technological tools, such as blended learning and online courses. Pucher and Lehner (2011) discuss traditional and self-organized learning methods, while Liu, Tong, and Yang (2018) focus on the utility of mind mapping in teaching programming.

A teaching method mentioned extensively in studies on computer science education is project-based learning. Project-based learning involves students engaging in experiential and interdisciplinary work on temporary projects where they contribute their skills and preferences (Pucher, & Lehner, 2011). The method is rooted in constructivist and situated/socio-cultural learning theories, which highlight context, knowledge construction, and peer collaboration (Schilling & Klamma, 2010). Situated learning involves novices joining a community of practice, fostering shared problem-solving practices through shared interests, joint activities, and collaborative engagement in problem-solving. This highlights the importance of authentic problem-solving and meaningful activities in a community-oriented learning environment.

3.3 Teaching Machine Learning

Adapting computer science teaching techniques to ML requires an examination of conventional software engineering approaches, according to Schiendorfer, Gajek, & Reif (2021). Software engineering students often struggle with moving beyond designing ML models to adopting a systematic, experiment-based approach. For novices in ML, grappling with both new algorithms and a novel developmental style based on experimental methodology, can be overwhelming. Deciding what to treat as a black box versus understanding in more detail adds another layer of complexity. Moreover, in transforming software engineers into proficient ML engineers, a common mistake is relying solely on available libraries and frameworks. Additionally, there are three fundamental skills software engineers need to acquire in order to become ML engineers: a systematic methodology for hyper-parameter tuning, knowledge of proper data splitting, and understanding gradient signals.

Key instructional strategies specific for teaching ML have been uncovered in literature, emphasizing the need for instructors to understand diverse learner backgrounds and mathematical readiness. Steinbach, Seibold & Guhr (2021) note that practical approaches like staying close to applications, live coding, and minimizing mental load aim to enhance accessibility and engagement. Community preferences underscore the importance of central plot-driven instructional material and optional modules for diverse learners. The paper highlights feedback, modular lesson structures, and low-stakes testing using sticky notes as contributors to effective learning. Long-term assessment challenges highlight the necessity of monitoring teaching quality and learning outcomes, with open educational resources playing a pivotal role in providing accessible and up-to-date material. The collaborative use of resources view the role of educators as curators, tailoring content to students' needs and fostering a collaborative learning environment. In addition to these teaching methods, Schiendorfer, Gajek, & Reif (2021) describe exercises designed for software engineering students to enhance the skills needed for ML, covering topics like tuning the learning rate, managing data splitting for model hyper-parameter tuning, and the importance of back-propagation and auto-differentiation in training deeper models.

Numerous studies highlight technology's role in enhancing machine learning (ML) education. Kaspersen, Bilstrup, and Petersen (2021) developed a user interface for students to build and evaluate ML models, improving their understanding of ML's practical aspects and promoting analytical thinking about data and model interaction. Rattadilok, Roadknight, and Li (2018) implemented a gamified teaching method, adapting Clash of Clans to teach ML concepts and create engaging datasets for students. Chow (2019) incorporated a Kaggle competition into ML education, combining game-based learning and social constructivism to actively involve students in ML studies.

The literature underscores the diverse landscape of teaching methods, with a shift toward technology-enhanced and active learning approaches. In machine learning education, bridging the gap between traditional software engineering and empirical ML methodologies is critical, as well as optimizing teaching practices, fostering engagement, and addressing the evolving needs of learners.

3.4 Classification of Teaching Methods

Beck (1998) notes the lack of uniformity in lists of teaching strategies found in 25 teacher education textbooks, with authors offering varying numbers and classifications. Despite the pivotal role of teaching strategies in instruction, there is a surprising lack of attention to establishing a more uniform and common classification system. A classification system is seen as valuable for educators, providing distinct categories to relate similar strategies and differentiate between them. The taxonomy proposed in this paper aims to provide a comprehensive view on the different categories of teaching methods available.

The eight categories proposed are: associative, deliberative, expositive, individualistic, interrogative, investigative, performative, and technological. Educators use various strategies to enhance learning: Associative strategies group students by skills and interests for overall growth. Deliberative approaches promote thoughtful discussions to improve cognitive and communication skills. Expositive methods orderly deliver information to audiences. Individualistic strategies are tailored to each student's unique needs. Interrogative techniques use questions to engage students and develop critical thinking. Investigative methods focus on problem-solving through data analysis. Performative approaches encourage creative expression in arts and physical skills. Technological strategies incorporate devices like computers for information access and recording.

For the purpose of the current research paper, this taxonomy has been used to categorize all of the teaching methods identified during the literature review. This is a new undertaking in the domain of ML education. The categorisation can be found in Appendix A.

Machine learning involves a combination of theoretical understanding, practical application, and problem-solving. Therefore, a diverse set of teaching methods that engage students actively, encourage critical thinking, and provide hands-on experience with real-world problems could be effective in the ML education context. A combination of expositive methods, such as lecture, direct instruction, and live coding, associative methods, such as team-based learning and collaborative approaches, investigative methods, such as project-based learning and problem-based learning, and technological methods, such as the Kaggle ML comptetition, virtual reality technology and information and communication technology integration, could be beneficial to students in learning ML. In addition, other categories could serve as a supplementary way of increasing engagement and encouraging exploration through deliberative methods like peer feedback, interrogative methods such as class discussions, and performative methods like storytelling and educational games. Finally, the individualistic approach, including prerecorded lectures and learning tasks, caters to the diverse skills, needs, and interests of each student, allowing for a personalized learning experience that takes into account varying levels of prior knowledge and distinct learning paces.

3.5 Comparison of Teaching Methods

Recent studies increasingly endorse constructivist, active learning approaches, highlighting the benefits of discussions in enhancing comprehension and fostering participation, self-confidence, and leadership skills. Team learning methods show improved student performance compared to traditional lecture-based approaches (Ben-Ari, 2001; Bidabadi et al., 2016; Ertmer & Newby, 1993; Greening, 2000; Sadeghi et al., 2014; Sivarajah et al., 2019). However, conflicting findings exist in the literature; some studies suggest that active, discussion-based methods are inferior, while others demonstrate their superiority in terms of material retention.

A qualitative analysis performed by Bidabadi, Isfahani, Rouhollahi, & Khalili (2016) revealed that the most effective teaching approach involves a mixed method, a combination of student-centered and teacher-centered methods, along with strategic educational planning and prior readiness.

Another study by Sadeghi, Sedaghat, & Ahmadi (2014) indicates that the blended method is effective in increasing the students' learning rate. Moreover, students expressed higher satisfaction with the blended learning method compared to the lecture method.

In contrast, Carpenter (2006) evaluates five teaching methods (lecture, lecture/discussion combination, jigsaw, case study, team project) in a large class setting. Student performance significantly improves under the lecture method compared to both lecture/discussion and team project methods. The jigsaw method, a cooperative learning strategy where each student becomes an "expert" on one aspect and then teaches others, outperforms the lecture method. Lecture/discussion shows improvement compared to the team project method but is surpassed by the jigsaw and case study methods. The jigsaw method demonstrates significant improvement compared to both case study and team project methods, while the case study method outperforms the team project method.

The most effective teaching methods from these three comparative studies seem to be a mixed method, blended learning, and jigsaw. This seems to fit with the findings of the literature review on effective ML education, findings which prioritize active learning, experiment-based approaches, live coding, and staying close to applications. However, these teaching methods' effectiveness has not been compared in the field of machine learning education, and further research is necessary in order to draw conclusions pertaining to the domain of ML.

4 Survey Setup and Results of Dutch ML Teaching Methods

This chapter delves into the design and analysis of the survey for gathering data about teaching Dutch ML courses.

4.1 Survey Design

The target group of the survey is machine learning lecturers all across Dutch universities. 98 participants have been invited over email, of which 24 responded.

The survey description contains information on the research and includes the ability to opt out at any time and to give consent by submitting the responses. This is in line with the ethical research standards.

The survey has 16 questions in total: 5 closed questions and 11 open ones. The aim is to gather data on the lecturer's educational background, the teaching methodologies and tools used in the ML course, the course curriculum and materials, and challenges in teaching ML. The questions can be viewed in Appendix B: Questionnaire.

4.2 Data Analysis

The survey responses reveal a predominant trend in the adoption of a mixed-method approach to teaching ML. Lectures, often complemented by labs or workshops, form the core of instructional methods. Additionally, respondents frequently incorporate projects and group work into their courses. Experimenting with other teaching methods includes leveraging virtual environments such as Gather Town for group work, or utilizing Kaggle for competitions.

The teaching methods used by the survey participants can be seen in Figure 1.

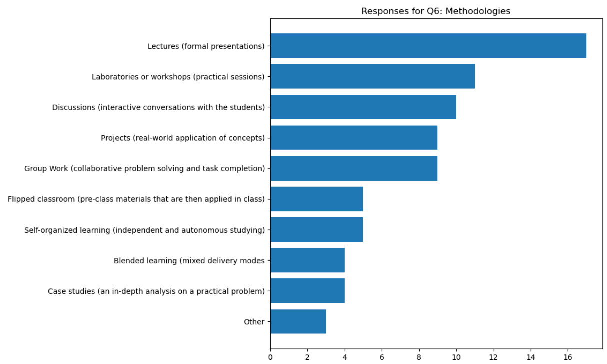


Figure 1: Teaching methods

Word Frequency Analysis

Word frequency analysis on the survey responses reveals key themes in ML education, emphasizing fundamental concepts like ”learning,” ”machine,” ”data,” and ”ML.” Terms like ”exam” and ”project” highlight a focus on assessments and practical application, while ”lectures,” ”Python,” and ”exercises” indicate diverse teaching methods. Words like ”methods” and ”teaching” stress the importance of effective teaching strategies. The use of specific platforms like ”brightspace” and terms such as ”regression” and ”classification” point to the use of varied technological tools and a focus on statistical methods in education. A word cloud in Figure 2 visualizes these frequent words.

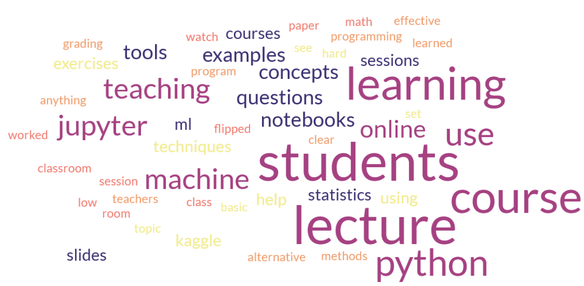


Figure 2: Word cloud of most frequent words

Thematic Analysis

To draw insights from the qualitative data of the survey, thematic analysis has been performed. Answers have been assigned codes, and we have revealed some common themes that occur in the data.

Instructors in ML education employ a variety of technological tools such as Microsoft Teams, Google Colab, Python, Git, Kaggle, GatherTown, GitHub, Vocareum, and Jupyter notebooks. A key pedagogical focus is on effective strategies like visualization, hands-on exercises, and integrating statistics with programming, highlighting the need for approaches that extend beyond theoretical knowledge to practical skills and visual understanding. Student engagement is a priority, with instructors incorporating interactive elements, collaborative projects, and practical exercises. Practical application of ML concepts is emphasized through real-world applications and coding environments, preparing students for industry relevance. Interactive elements like widgets and interactive Jupyter notebooks play a significant role in enhancing engagement and interaction. Collaboration is fostered through group projects and online platforms. Some instructors explore innovative methods such as flipped classrooms, while others prefer established practices. Continuous learning among educators through seminars, workshops, and online communities is notable, with those engaged in research often adopting more unconventional methods. Challenges in implementing new methods, particularly flipped classrooms, highlight the complexities and need for a balance between innovation and traditional approaches. Finally, there’s a degree of skepticism about the effectiveness of new tools and methods, underscoring the importance of evidence-based practices in education.

The identified themes collectively paint a comprehensive picture of the diverse and innovative approaches employed by ML instructors. These insights provide valuable benchmarks for effective teaching methodologies, technological tool integration, and strategies to enhance student engagement in machine learning education.

Co-occurence of Teaching Methods

In the context of the provided responses on teaching methods in ML education, co-occurrence analysis reveals which methods tend to be mentioned together, providing insights into the interconnectedness of various instructional approaches.

The co-occurence of codes applied to the responses on the teaching methods can be seen in Figure 3. Most interestingly, it can be seen that lectures often co-occur with discussions, group work, labs or workshops, and projects, whereas blended learning co-occurs the most with the flipped classroom method.

	blended learning	case study	discussions	Flipped classroom	group work	labs or workshops	lectures	projects	self-organised learning
blended learning		2	3	4	3	1	4	2	2
case study	2		4	2	2	0	4	1	1
discussions	3	4		3	6	6	11	6	3
Flipped classroom	4	2	3		4	2	5	2	3
group work	3	2	6	4		7	11	5	2
labs or workshops	1	0	6	2	7		12	7	3
lectures	4	4	11	5	11	12		10	5
projects	2	1	6	2	5	7	10		2
self-organised learning	2	1	3	3	2	3	5	2	

Figure 3: Co-occurence of teaching methods

5 Experimental Setup and Results of Comparing Lectures and the Jigsaw Method

This chapter describes the experimental setup and results of comparing two teaching methods, lecture and jigsaw, in a small experiment. A lecture is a traditional teaching approach where an instructor delivers information to a large group of students, typically in a formal and structured manner. The jigsaw method is a cooperative learning technique where students are divided into groups and each member is assigned a different piece of a topic to learn and teach to their peers.

5.1 Experiment Design

The methods of lectures and jigsaw have been selected for comparison because jigsaw is a method that, according to literature (Carpenter, 2006), while not having been studied in ML education, is found to be the most efficient out of five widely used methods, whereas the lecture method is the most used method in ML courses of universities throughout the Netherlands, according to the survey findings.

Aligning with Campbell and Stanley's (2015) insights into the significance of pre-testing and post-testing for evaluating learning outcomes, pre- and post-tests were implemented, which allows for a comprehensive measurement of the impact of the two chosen teaching techniques on students' performance. A pre-test at the experiment's start measures participants' baseline knowledge, aiding even distribution across groups for stronger internal validity. La Barge (2007) takes into account the possibility of the student guessing by asking them directly on each question if they know the answer or if they are guessing. The experiment adopted this method in order to have more reliable performance insights. Furthermore, drawing from Locke's (2013) perspective on the importance of maintaining a consistent environment to prevent bias, all experimental sessions were conducted in a homogeneous and stable setting. To enhance comparability, minimize bias, and ensure experiment reliability, participants were randomly assigned to groups in the experiment.

5.2 Sample Size Determination

The sample size n for the experiment is determined using Cohen's formula (1992), which is expressed as

$$n = \frac{(Z_\alpha + Z_\beta)^2 \cdot \sigma^2}{\delta^2}$$

The parameters were set as follows: $\alpha = 0.05$, representing the significance level; $\beta = 0.2$, indicating the probability of a Type II error; $\sigma = 1$, as the standard deviation of the population; and $\delta = 0.5$, signifying the anticipated effect size. These parameter choices are informed by a review of relevant literature and guided by Cohen's established benchmarks. Calculations suggest a required sample size of approximately 31.4, which is rounded up to 32. Thus, to achieve the intended statistical power and significance, the experiment requires a minimum of 32 participants per group. Due to time constraints, however, 13 participants were recruited for the study, 6 in one group and 7 in the other. While this limitation affected the sample size, the experiment still yielded interesting insights. Nevertheless, future research should endeavor to include larger sample sizes to enhance the generalizability and robustness of the findings.

5.3 Experiment Procedure

Before the experiment commenced, the 13 participants were provided with a detailed explanation of the study's purpose, procedures, and any potential risks. Informed consent was obtained from each participant.

Participants were selected uniformly from the target group, computer science students, and were randomly assigned to one of two groups. Participants in group A learned K-Nearest Neighbours, a simple ML concept, through a lecture. Participants in group B learned the same concept through the jigsaw method. Prior to the lesson, pre-tests were administered. Following the learning phase, both groups took a post-test to evaluate their performance. Finally, the participants were asked to rate their satisfaction with the teaching method on a scale from 1 to 5, and to give open feedback on their experience with learning the concept through this method compared to other methods. Personal identifiers were removed from all collected data to ensure participant confidentiality.

The instructions for Group B, the jigsaw group, are found in Appendix C. The study material given to Group B can be found in Appendix D. The lecture slides presented to Group A are found in Appendix E. The test questions given to both groups as a pre-test and post-test can be found in Appendix F, and the answers to the test are in Appendix G.

5.4 Data Analysis

The experimental results from the lecture and jigsaw methods in teaching k-Nearest Neighbors (k-NN) have been analysed in order to provide insights into the effectiveness of these teaching strategies. The analysis of pre-test and post-test scores reveals a significant improvement in participant understanding across both methods.

For the lecture method, the pre-test scores were notably low, with a mean of 0.42 and a median of 0.00, suggesting limited prior knowledge of the subject. The post-test scores showed a dramatic increase, with a mean of 5.00 and a median of 5.00, indicating a substantial improvement in understanding. This improvement was statistically validated by a paired sample t-test, yielding a T-Statistic of 8.76 and a very low P-Value of 0.00032, far below the standard alpha level of 0.05. This significant statistical result, coupled with a high Cohen's d value of 3.58, suggests that the lecture method had a profound and practically significant impact on the participants' learning.

Similarly, the jigsaw method demonstrated its effectiveness. The pre-test scores, with a mean of 1.64 and a median of 2.00, were higher than those for the lecture method, indicating a slightly better initial understanding. The post-test scores improved significantly, achieving a mean of 5.50 and a median of 6.00. The paired sample t-test for the jigsaw method also indicated statistical significance, with a T-Statistic of 6.38 and a P-Value of 0.0007. The Cohen's d value of 2.41 for the jigsaw method further confirms its practical significance on the participants' learning.

The box plot on the left of Figure 4 shows the distribution of scores for both the pre-test and post-test for the lecture group. The central tendency and variability of scores in each test can be observed. The plot clearly indicates an upward shift in scores from the pre-test to the post-test. The paired bar chart visible on the right of Figure 4 displays the pre-test and post-test scores side-by-side for each participant.

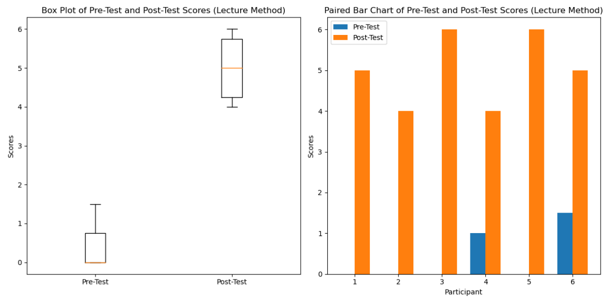


Figure 4: Pre- and Post-test scores for Lecture

The pre- and post-test results for the jigsaw group have also been visualised in a box plot and a bar chart. The box plot, found in Figure 5 on the left, provides a summary of the pre-test and post-test scores for jigsaw, showing the median, quartiles, and potential outliers. A noticeable upward shift in scores from the pre-test to the post-test can be seen. The bar chart, found in Figure 5 on the right, shows the scores of each participant in both the pre-test and post-test.

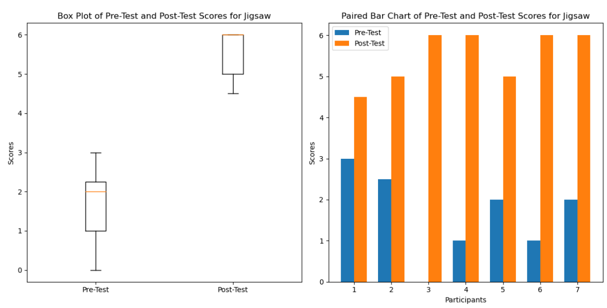


Figure 5: Pre- and Post-test scores for Jigsaw

An ANOVA test was conducted to compare the effectiveness of the jigsaw and lecture methods in teaching the k-NN algorithm. The test assessed whether there was a statistically significant difference in first the post-test scores, and secondly, in the improvement from pre-test to post-test scores between the two methods. The p-values obtained from the ANOVA tests for both the post-test scores (approximately 0.267) and the differences between pre-test and post-test scores (approximately 0.392) were higher than the conventional alpha level of 0.05. This might indicate that there is no statistically significant difference in the effectiveness of the jigsaw method compared to the lecture method in terms of both the final scores achieved by the participants and the magnitude of improvement observed.

After the experiment, the participants from both groups were asked the following question: ”How do you compare the method to other methods you have seen, such as labs, flipped classroom, blended learning, or case studies?”. Those in the jigsaw group appreciated the interactive and discussion-based nature of the method, highlighting its effectiveness in engaging students and deepening their understanding. This preference underscores a growing trend towards active learning, where students are not just passive recipients of information but active participants in the learning process. However, some concerns were raised about the accuracy of information in these less structured settings, suggesting the need for a careful balance between structure and interaction in teaching methods. On the other hand, participants from the lecture group valued the structured and direct content delivery of traditional lectures, particularly when complemented by opportunities for active participation and self-study. The practical aspects of laboratories were also mentioned for reinforcing theoretical knowledge, indicating a preference for a blended approach that combines lectures with practical application. The feedback overall suggests a growing preference for concise, focused teaching methods that directly address key concepts, and a desire for more interactive elements, such as animations and detailed examples, to enhance engagement.

The participants were also asked to rate, on a scale from one to five, their perceived effectiveness of the lectures and jigsaw methods in achieving the learning goals. The lecture method received a high overall effectiveness rating, with both the mean and median at 4.25 out of 5, indicating strong participant satisfaction. The standard deviation of 0.76 suggests moderate variability, meaning that while most participants found lectures highly effective, there were some variations in perceptions of its effectiveness. For the jigsaw method, the mean effectiveness rating was slightly lower at approximately 3.86, with a median of 4.00, suggesting that on average, participants found it to be effective, but with some variability in their experiences. The standard deviation of 0.90 points to a slightly broader range of opinions about the jigsaw method compared to lectures. The 95% confidence interval for the jigsaw method’s effectiveness rating, ranging from approximately 3.03 to 4.69, indicates that while there is a general satisfaction with the method, it might not be as universally favored as traditional lectures. The confidence interval for the lecture method, spanning from approximately 3.45 to 5.05, reinforces the idea that lectures are generally well-regarded by participants in terms of effectiveness.

6 Responsible Research

The experiment’s participants were provided with and signed an informed consent form. The experiment has undergone a review by the Human Research Ethics Committee, which included a risk assessment detailing potential hazards and their mitigation strategies, as well as a data management plan describing the secure processing and storage of data to ensure reproducibility.

The research adhered to the Netherlands Code of Conduct for Research Integrity, ensuring ethical standards in design, data collection, analysis, and reporting. All phases of the study emphasized transparency, honesty, and participant welfare, maintaining integrity and objectivity. Additionally, the research followed TU Delft’s FAIR data principles, focusing on transparent data management, methodology documentation, and data sharing to enhance reproducibility and contribute to scientific validation and knowledge advancement.

Moreover, this research is reproducible as it can be replicated by redoing the survey and the experiment. The step-by-step instructions on how the experiment was set up are present in the Survey Design and Experiment Design subsections, and the questions and materials used for executing the survey and experiment are found in the appendices.

7 Discussion

This study's investigation into the effectiveness of various teaching methods in machine learning education, particularly within Dutch universities, presents several noteworthy findings.

The literature review, along with our survey findings, underscores a significant shift towards technology-enhanced, active learning approaches in ML education, and highlights the importance of an interactive, student-centric learning environment. While literature review insights prioritize active learning, experiment-based approaches, live coding, and staying close to applications, the survey findings indicate a shift towards mixed-method and blended learning approaches integrating traditional lectures, hands-on projects, and group work. This shift caters to the dual needs of theoretical understanding and practical application in ML, with traditional lectures laying the theoretical groundwork, and project-based methods providing avenues for applying this knowledge in real-world scenarios. Combining different instructional strategies not only accommodates diverse learning needs, but also adds relevance and depth to the learning process by mirroring real-world applications.

Insights from the survey, such as the diverse approaches to teaching, indicate a field in flux, with educators balancing tried-and-tested methods against emerging pedagogical trends. The challenges and skepticism expressed by some instructors highlight the ongoing debate within the educational community regarding the most effective teaching strategies, suggesting that continuous adaptation and rigorous evaluation of teaching methods are crucial for the future development of ML education.

Our research suggests a need to reevaluate the dominance of traditional lecture-based teaching in ML education, considering the potential of more interactive methods like the jigsaw approach. Although traditional lectures remain prevalent, integrating interactive methodologies could enhance engagement and effectiveness. The jigsaw method, in particular, has shown promise in our experiment, fostering active participation and peer-teaching, aligning well with the collaborative nature of ML and echoing the literature review findings. This preliminary evidence, albeit the small sample size, suggests that such interactive methods could be as effective as traditional lectures in promoting deep understanding and engagement in ML topics. Furthermore, feedback on the experiment could reflect a broader educational trend, emphasizing the importance of accommodating diverse learning styles through a variety of teaching methods, and the potential benefits of hybrid or blended learning approaches in ML education.

The experiment indicates that both the jigsaw and lecture methods effectively enhanced participants' understanding of k-NN, with no significant difference in their overall effectiveness. This suggests that the choice between these methods in an educational setting may depend more on factors like educator preference, resource availability, or individual learning styles, rather than on the efficacy of one method over the other. Interestingly, participant satisfaction tended to favor lectures, possibly due to their structured format and familiarity. However, while traditional lecture methods are still highly valued for their effectiveness in achieving learning goals, the jigsaw method also holds considerable merit, particularly in terms of engaging participants and offering a more interactive learning experience. The variability in ratings for both methods highlights the importance of considering different learning styles and preferences when designing educational experiences, underscoring the potential benefits of incorporating a variety of teaching methods in the learning process.

All things considered, it is crucial to acknowledge the limitations of our study. The small sample size and the specific context of Dutch universities may limit the generalizability of our findings. Additionally, the selection of the jigsaw method was based on its documented success in other fields but may not capture the full spectrum of innovative teaching practices applicable to ML education.

Based on the insights of this research, we advocate for the integration of project-based and active learning strategies in ML teaching, while incorporating technology. We also think that it is important for educators to leverage existing educational research to inform their ML courses. Furthermore, we advocate for more research in the domain of ML education to fill in the existing knowledge gaps.

8 Conclusions and Future Work

This study delved into the effectiveness of various teaching methods in machine learning education, particularly within the context of Dutch universities. It involved a comprehensive literature review, a targeted survey of Dutch ML lecturers, and a comparative experiment between the traditional lecture method and the jigsaw method. The primary research questions centered on identifying effective teaching techniques in ML, examining the alignment of traditional methods used in Dutch universities with the evolving nature of ML, and comparing these methods' effectiveness.

The conclusions drawn from this study highlight the effectiveness of mixed-method and blended learning approaches, which combine traditional lectures with active learning strategies, in meeting diverse learning needs and balancing theoretical knowledge with practical application. Traditional lectures continue to play a crucial role in providing foundational theoretical knowledge. Interestingly, the study found that interactive methods like the jigsaw technique, though less common, show promise in enhancing student engagement and understanding in ML. A key finding from the survey of Dutch university lecturers revealed a trend towards mixed-method approaches. The study also emphasized the significant role of technology integration in teaching, enhancing practical skill development in ML.

The limitations of the study open up avenues for future research, suggesting the need for larger-scale studies across a broader range of educational settings to enhance the robustness of the findings. Investigating the long-term impact of different teaching methods on information retention and the practical application of ML skills could provide deeper insights. Further exploration of a wider range of interactive teaching methods in ML education is warranted. While papers such as Schiendorfer, Gajek, & Reif (2021) or Steinbach, Seibold & Guhr (2021) establish a good baseline, comparative studies such as Carpenter's (2006) are needed in the context of ML. Our paper provides an example of how such an experiment can be conducted in ML education, but its results may not be representative when scaled to a larger classroom. Additionally, aligning teaching methods with industry requirements and real-world applications remains an area ripe for further investigation. Comparative studies between ML education and other technical disciplines could yield valuable insights, and exploring the psychological impact of different teaching methods on student motivation and engagement in ML could offer new perspectives.

In conclusion, this research contributes significantly to the understanding of ML education, underscoring the effectiveness of blended learning and the potential of interactive methods. It emphasizes the need for continuous adaptation and innovation in teaching strategies to keep pace with the rapidly evolving field of ML.

References

Beck, C.R. (1998). "A taxonomy for identifying, classifying, and interrelating teaching strategies". In *The Journal of General Education, 47(1), 37-62*.

Ben-Ari, M. (2001). "Constructivism in computer science education". In *Journal of computers in Mathematics and Science Teaching, 20(1), 45-73*.

Bidabadi, N. S., Isfahani, A. N., Rouhollahi, A., Khalili, R. (2016). "Effective teaching methods in higher education: requirements and barriers". In *Journal of advances in medical education professionalism, 4(4), 170*.

Campbell, D.T., & Stanley, J.C. (2015). "Experimental and quasi-experimental designs for research". In *Ravenio books*.

Carpenter, J.M. (2006). "Effective teaching methods for large classes". In *Journal of Family & Consumer Sciences Education, 24(2)*.

Chow, W. (2019). "A Pedagogy that Uses a Kaggle Competition for Teaching Machine Learning: an Experience Sharing". In *2019 IEEE International Conference on Engineering, Technology and Education (TALE), Yogyakarta, Indonesia, 2019, pp. 1-5*. doi: 10.1109/TALE48000.2019.9226005.

Ertmer, P.A., & Newby, T.J. (1993). "Behaviorism, cognitivism, constructivism: Comparing critical features from an instructional design perspective". In *Performance improvement quarterly, 6(4), 50-72*.

Gelman, A. and Loken, E. (2012). "Ethics and statistics: Statisticians: When we teach, we don't practice what we preach". In *CHANCE, 25(1):47–48, 2012*. doi: 10.1080/09332480.2012.668469. URL https://doi.org/10.1080/09332480.2012.668469.

Greening, T. (2000). "Emerging constructivist forces in computer science education: Shaping a new future?". In *Computer science education in the 21st century, 47-80*.

Kaspersen, M.H., Bilstrup, K.E.K., & Petersen, M.G. (2021). "The machine learning machine: A tangible user interface for teaching machine learning". In *Proceedings of the fifteenth international conference on tangible, embedded, and embodied interaction (pp. 1-12)*.

Ko, A.J. (2017). "We need to learn how to teach Machine Learning". *Medium*. https://medium.com/bits-and-behavior/we-need-to-learn-how-to-teach-machine-learning-acc78bac3ff8

La Barge, G. (2007). "Pre-and post-testing with more impact". In *The Journal of Extension*, 45(6), 17.

Liu, Y., Tong, Y., & Yang, Y. (2018). "The application of mind mapping into college computer programming teaching". In *Procedia Computer Science, 129, 66-70*.

Locke, L.F., Spirduso, W.W., & Silverman, S.J. (2013). "Proposals that work: A guide for planning dissertations and grant proposals". In *SAGE Publications*.

Ma, L. (2021). "An immersive context teaching method for college English based on artificial intelligence and machine learning in virtual reality technology". In *Mobile Information Systems, 2021, 1-7*.

Martins, R.M., & Gresse Von Wangenheim, C. (2023). "Findings on teaching machine learning in high school: A ten-year systematic literature review". In *Informatics in Education*, 22(3), 421-440.

Mohorovicic, S., & Strcic, V. (2011). "An overview of computer programming teaching methods". In *Central European Conference on Information and Intelligent Systems (p. 47)*. Faculty of Organization and Informatics Varazdin.

Pucher, R., & Lehner, M. (2011). "Project based learning in computer science–a review of more than 500 projects". In *Procedia-Social and Behavioral Sciences, 29, 1561-1566*.

Rattadilok, P., Roadknight, C., and Li, L. (2018). "Teaching Students About Machine Learning Through a Gamified Approach". IN *2018 IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE)*, Wollongong, NSW, Australia, 2018, pp. 1011-1015, doi: 10.1109/TALE.2018.8615279.

Sadeghi, R., Sedaghat, M. M., & Ahmadi, F. S. (2014). "Comparison of the effect of lecture and blended teaching methods on students' learning and satisfaction". In *Journal of advances in medical education professionalism, 2(4), 146*.

Schiendorfer, A., Gajek, C., & Reif, W. (2021). ”Turning software engineers into machine learning engineers”. In *European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (pp. 36-41). PMLR*.

Schilling, J., & Klamma, R. (2010). ”The difficult bridge between university and industry: a case study in computer science teaching”. In *Assessment & Evaluation in Higher Education, 35(4), 367-380*.

Shahid, F., Aleem, M., Islam, M.A., Iqbal, M.A., & Yousaf, M.M. (2019). ”A review of technological tools in teaching and learning computer science”. In *Eurasia Journal of Mathematics, Science and Technology Education, 15(11)*, em1773.

Shapiro, R.B., & Fiebrink, R. (2019). ”Introduction to the special section: Launching an agenda for research on learning machine learning”. In *ACM Transactions on Computing Education (TOCE)*, 19 (4), 1-6.

Sivarajah, R.T., Curci, N.E., Johnson, E.M., Lam, D.L., Lee, J.T., & Richardson, M.L. (2019). ”A review of innovative teaching methods”. In *Academic radiology, 26(1), 101-113*.

Steinbach, P., Seibold, H. & Guhr, O. (2021). ”Teaching Machine Learning in 2020. Proceedings of the First Teaching Machine Learning and Artificial Intelligence Workshop”. In *Proceedings of Machine Learning Research* 141:1-6. Available from https://proceedings.mlr.press/v141/steinbach21a.html.

Yakovleva, N.O., & Yakovlev, E.V. (2014). ”Interactive teaching methods in contemporary higher education”. In *Pacific Science Review, 16(2), 75-80*.

Zendler, A., & Klaudt, D. (2015). ”Instructional methods to computer science education as investigated by computer science teachers”. In *Journal of Computer Science, 11(8), 915*.

Appendix A: Categorisation of Teaching Methods

The associative category encompasses teaching methods that prioritize grouping students based on their skills, needs, and interests, fostering mental, emotional, and social growth. Context Teaching immerses learners in a dynamic environment, encouraging independent exploration and collaboration. Random Access Teaching guides learners with distinct purposes, emphasizing collaborative learning and the importance of context in constructing meaning. Team-based Learning involves fixed groups, accountability, and frequent feedback to enhance both learning and team development. Collaborative methods emphasize collaboration and group work in the learning process, fostering a sense of community and shared knowledge.

Deliberative teaching methods focus on stimulating thoughtful exchange of ideas, in order to enhance cognitive, social, and verbal communication skills. Peer Feedback involves students providing ongoing feedback to their peers, fostering self-awareness and competency development. The Basket Method simulates common specialist situations, requiring students to perform unplanned activities efficiently. Audience Response System utilizes technology for presenters to engage with the audience, encouraging interactivity and participation. Flipped Classroom involves students engaging with educational content outside of class, with in-class time dedicated to active learning. Active Learning shifts the focus from teacher-centered to learner-centered education, encouraging students to actively engage with the material. Various strategies such as Think-pair-share, Jigsaw, Pair Programming, and Buzz Groups promote collaboration, critical thinking, and problem-solving skills.

Expositive teaching methods involve delivering information in an organized, authoritative manner, whether orally or in writing, to a receptive audience. Lecture, a traditional method, has the instructor presenting information in a spoken format. Long-Distance Teaching utilizes remote conferencing systems to overcome geographical challenges. Mind Mapping is employed as a diagramming tool to generate, visualize, structure, and classify ideas. Direct Instruction provides clear and structured teaching methods to convey information directly to students. Blended Learning mixes online and in-person learning for a flexible educational experience. Micro Content Modules, Multi-Agent Micro Content Modules, Web MOOCs, and live coding are variations that leverage technology and different formats for content delivery. The strategies aim to convey information effectively while adapting to diverse learning styles and preferences.

Individualistic teaching methods focus on personalized learning experiences, in order to provide tailored instruction that meets the individual skills, needs, and interests of students. Prerecorded Lectures are recorded in advance, allowing students to access them at their convenience. Learning Tasks emphasize active engagement and task-oriented learning. Discovery Learning fosters exploration and self-directed discovery by allowing students to explore concepts independently. Computer Simulation utilizes virtual environments for practical learning experiences. Trainings/Workshops involve hands-on sessions for skill development. Game-themed Programming incorporates game-like elements into programming activities to make learning more engaging. These methods recognize the diversity of learners and aim to provide personalized and self-paced learning experiences.

Interrogative teaching methods involve employing questioning skills to foster participation, clarify understanding, and promote higher-level thinking. One-minute Papers are brief written reflections or summaries by students on a topic, typically taking one minute to complete. Class Discussion combines traditional lecture-style teaching with a discussion session to promote interaction and deeper understanding. Interactive Lecture is a format that encourages active participation and engagement from students through discussions, questions, or activities. These methods enhance critical thinking, communication skills, and deeper understanding through active participation.

Investigative teaching methods focus on problem-solving and real-world applications, in order to solve problems using inductive reasoning, data collection, analysis, and drawing conclusions. Case Study involves problem-solving through the examination of specific cases, fostering teamwork and independent solution modeling. Project-Based Learning focuses on students working on temporary projects to gain experiential and interdisciplinary learning. Problem-Based Learning encourages the development of problem-solving skills by presenting real-world problems for students to solve. Gamified Approach uses game-like elements to make learning more engaging, while Kaggle Competition involves students in machine learning competitions, incorporating principles of game-based learning and social constructivism. Machine Learning Machine proposes a user interface for students to build and test Machine Learning models, incorporating hands-on investigative learning.

Performative teaching methods involve active demonstrations and creative expression, to encourage creative, aesthetic, and psycho-motor expression through dramatic/fine arts and physical skills. Behavioral Modeling focuses on teaching interpersonal skills and professional conduct by presenting and reproducing a model of desired behavior. Metaphor Game encourages creative problem-solving by applying metaphors to professional situations. Storytelling uses myths and stories to teach the rules of work and help students understand the specifics of their future profession. Educational Games, including game-themed programming, incorporate game elements to make learning more engaging, interactive, and enjoyable.

Technological teaching methods leverage digital tools and platforms that enable students to access and record information. VR Technology utilizes virtual reality to create immersive learning experiences. ICT (Information and Communication Technology) integrates technology, such as computers and the internet, in teaching and learning. CAI (Computer-Aided Instruction) and CBI (Computer-Based Instruction) use computers to aid in the instruction and learning process. Serious games, web games, micro-games, video games, and videos incorporate various digital media for educational purposes. These methods harness the power of technology to enhance engagement, accessibility, and interactivity in the

Teaching Method	Category	Description	Source
Context Teaching	ASSOCIATIVE	Teaching method prioritizing authenticity and association by providing a dynamic learning environment with independent exploration and collaboration.	(Ma, 2021)
Random Access Teaching	ASSOCIATIVE	Guides learners with distinct purposes, emphasizing collaborative learning and the importance of context in constructing meaning.	(Ma, 2021)
team-based learning	ASSOCIATIVE	An approach involving fixed groups, accountability, and frequent feedback to enhance both learning and team development.	(Sivarajah, Curci et.al., 2019)
Collaborative	ASSOCIATIVE	Emphasizes collaboration and group work in the learning process.	Shahid et.al. (2019)
Peer Feedback	DELIBERATIVE	Involves students providing ongoing feedback to their peers, fostering self-awareness and competency development.	(Yakovleva & Yakovlev, 2014)
Basket Method	DELIBERATIVE	Simulates common specialist situations, requiring students to perform unplanned activities efficiently.	(Yakovleva & Yakovlev, 2014)
Audience Response System	DELIBERATIVE	Utilizes technology for presenters to engage with the audience, encouraging interactivity and participation.	(Sivarajah, Curci et.al., 2019)
Flipped Classroom	DELIBERATIVE	Involves students engaging with educational content outside of class, with in-class time dedicated to active learning.	(Sivarajah, Curci et.al., 2019)
Active Learning	DELIBERATIVE	Instructional approach where students actively engage with material through various activities, shifting focus from teacher-centered to learner-centered	(Sivarajah, Curci et.al., 2019; Shahid et.al., 2019)
think-pair-share	DELIBERATIVE	A collaborative learning strategy where students think individually, discuss in pairs, and then share with the larger group.	(Sivarajah, Curci et.al., 2019)
Jigsaw	DELIBERATIVE	A cooperative learning strategy where each student becomes an "expert" on one aspect and then teaches others.	Carpenter, Sivarajah
Pair Programming	DELIBERATIVE	Programming approach where two individuals work together at one computer, with one typing and the other reviewing.	(Mohorovicic and Strcic, 2011)
Buzz groups	DELIBERATIVE	Small groups formed to discuss or brainstorm specific topics or ideas.	Pucher, R., & Lehner, M. (2011)
group sessions moderated by the students	DELIBERATIVE	Collaborative sessions where students take on the role of moderators to guide discussions or activities.	Pucher, R., & Lehner, M. (2011)
low-stakes testing	DELIBERATIVE	Assessments that carry lower consequences, often used for formative purposes.	Steinbach, Seibold \& Guhr (2021)
Class Discussion	DELIBERATIVE, ASSOCIATIVE	Combination of traditional lecture-style teaching followed by a discussion session to promote interaction and deeper understanding.	(Carpenter, 2006)
Action Learning Method	DELIBERATIVE, ASSOCIATIVE	Promotes self-learning environments through collaborative group work on practice-focused problems.	(Yakovleva & Yakovlev, 2014)
Behavioral Modeling	DELIBERATIVE, PERFORMATIVE	Focuses on teaching interpersonal skills and professional conduct by presenting and reproducing a model of desired behavior.	(Yakovleva & Yakovlev, 2014)
Metaphor Game	DELIBERATIVE, PERFORMATIVE	Encourages creative problem-solving by applying metaphors to professional situations.	(Yakovleva & Yakovlev, 2014)
Lecture	EXPOSITIVE	Traditional teaching method where an instructor presents information to students in a spoken format.	(Carpenter, 2006)
Long-Distance Teaching	EXPOSITIVE	Utilizes remote conferencing systems for real-time online meetings to overcome geographical challenges.	(Sivarajah, Curci et.al., 2019)
Mind Mapping	EXPOSITIVE	Utilized as a diagramming tool in teaching to generate, visualize, structure, and classify ideas.	(Liu, Tong, & Yang, 2018)
Direct instruction	EXPOSITIVE	Provides clear and structured teaching methods to convey information directly to students.	Zendler and Klaudt (2015)
Blended learning	EXPOSITIVE	Mixes online and in-person learning to create a flexible and diverse educational experience.	Pucher, R., & Lehner, M. (2011)
MCM (Micro Content Modules):	EXPOSITIVE	Short, focused learning modules that deliver specific content.	Shahid et.al. (2019)
MAMCM (Multi-Agent Micro Content Modules	EXPOSITIVE	Learning modules that involve multiple agents for delivering content.	Shahid et.al. (2019)
Web MOOC (Massive Open Online Course):	EXPOSITIVE	A large-scale online course accessible over the web.	Shahid et.al. (2019)
live coding	EXPOSITIVE	Involves instructors coding in real-time during a lecture or presentation.	Steinbach, Seibold \& Guhr (2021)
minimizing mental load	EXPOSITIVE	Teaching strategy aimed at reducing cognitive load on students to enhance learning.	Steinbach, Seibold \& Guhr (2021)
plot-driven instructional material	EXPOSITIVE	Educational material designed around a central plot or narrative to engage students.	Steinbach, Seibold \& Guhr (2021)
optional modules	EXPOSITIVE	Additional, elective components within a course that students can choose based on their interests.	Steinbach, Seibold \& Guhr (2021)
feedback	EXPOSITIVE	Information provided to students about their performance to facilitate improvement.	Steinbach, Seibold \& Guhr (2021)
modular lessons	EXPOSITIVE	Instructional units organized into separate modules for flexibility and focused learning.	Steinbach, Seibold \& Guhr (2021)
Interactive lecture	EXPOSITIVE, DELIBERATIVE	A lecture format that encourages active participation and engagement from students through discussions, questions, or activities.	(Sivarajah, Curci et.al., 2019)
Anchored Instruction	EXPOSITIVE, INVESTIGATIVE	Focuses on real-world scenarios, emphasizing problem-solving within practical situations.	(Ma, 2021)
Prerecorded lectures	INDIVIDUALISTIC	Lectures that are recorded in advance and can be accessed by students at their convenience.	(Mohorovicic and Strcic, 2011)
Learning tasks	INDIVIDUALISTIC	Methodology focused on promoting active engagement and task-oriented learning.	Zendler and Klaudt (2015)
Discovery learning	INDIVIDUALISTIC	Fosters exploration and self-directed discovery by allowing students to explore concepts independently.	Zendler and Klaudt (2015)
Computer Simulation	INDIVIDUALISTIC	Utilizes virtual environments for practical learning, often in the form of simulations.	Zendler and Klaudt (2015)
Trainings/Workshops	INDIVIDUALISTIC	Involves hands-on training sessions or workshops for skill development.	Pucher, R., & Lehner, M. (2011)
ONe-minute papers	INTERROGATIVE	Brief written reflections or summaries by students on a topic, typically taking one minute to complete.	Pucher, R., & Lehner, M. (2011)
Case Study	INVESTIGATIVE	Involves problem-solving through the examination of specific cases, fostering teamwork and independent solution modeling.	(Yakovleva & Yakovlev, 2014; Carpenter, 2006)
Project-Based Learning	INVESTIGATIVE	Focuses on students working on temporary projects to gain experiential and interdisciplinary learning.	(Pucher & Lehner, 2011; Schilling, & Klamma, 2010)
Problem-Based Learning	INVESTIGATIVE	Encourages the development of problem-solving skills by presenting real-world problems for students to solve.	(Rattadilok, Roadknight, & Li, 2018; Mohorovicic and Strcic, 2011; Zendler and Klaudt, 2015; Sc
Gamified Approach	INVESTIGATIVE	Uses game-like elements and principles to make learning more engaging and enjoyable.	(Rattadilok, Roadknight, & Li, 2018)
Kaggle Competition	INVESTIGATIVE	Involves students in a machine learning competition on Kaggle, incorporating principles of game-based learning and social constructivism.	(Chow, 2019)
Machine Learning Machine	INVESTIGATIVE	propose a user interface for students to build and test Machine Learning models	Kaspersen, Bilstrup and Petersen (2021)
Puzzle-based learning	INVESTIGATIVE	Involves learning through solving puzzles, fostering critical thinking and problem-solving skills.	(Mohorovicic and Strcic, 2011)
Inductive teaching methods	INVESTIGATIVE	Teaching approaches that involve presenting specific examples before deriving general principles.	Shahid et.al. (2019)
staying close to applications	INVESTIGATIVE	Teaching approach that emphasizes practical applications and real-world examples.	Steinbach, Seibold \& Guhr (2021)
Storytelling	PERFORMATIVE	Uses myths and stories to teach the rules of work and help students understand the specifics of their future profession.	(Yakovleva & Yakovlev, 2014)
educational games	PERFORMATIVE	Games designed for educational purposes to supplement traditional lectures and make the learning process engaging.	(Sivarajah, Curci et.al., 2019)
Game-themed programming	PERFORMATIVE, INDIVIDUALISTIC	Incorporates game-like elements into programming activities to make learning more engaging.	(Mohorovicic and Strcic, 2011)
VR Technology	TECHNOLOGICAL	Utilization of virtual reality technology to create immersive learning experiences.	Ma
ICT (Information and Communication Techno	TECHNOLOGICAL	Integration of technology, such as computers and the internet, in teaching and learning.	Shahid et.al. (2019)
CAI (Computer-Aided Instruction):	TECHNOLOGICAL	Utilizes computers to aid in the instruction and learning process.	Shahid et.al. (2019)
CBI (Computer-Based Instruction):	TECHNOLOGICAL	Incorporates computer technology for delivering educational content.	Shahid et.al. (2019)
serious games	TECHNOLOGICAL	Games designed for a primary purpose other than entertainment, often for educational or training purposes.	Shahid et.al. (2019)
web games	TECHNOLOGICAL	Games that are accessible and playable through web browsers.	Shahid et.al. (2019)
micro-games	TECHNOLOGICAL	Small-scale games that focus on specific skills or learning objectives.	Shahid et.al. (2019)
video games	TECHNOLOGICAL	Incorporates elements of video games into the learning process.	Shahid et.al. (2019)
videos	TECHNOLOGICAL	Visual content used for educational purposes.	Shahid et.al. (2019)

Figure 6: Categorisation of teaching methods found in literature review

learning process.

Appendix B: Questionnaire

Machine Learning Education in Dutch Universities - Survey Questionnaire

This questionnaire aims to gather insights into the teaching methodologies employed by educators in machine learning courses throughout Dutch universities. There is a lack of knowledge on how different universities in the Netherlands conduct their Machine Learning courses, and gathering such data aims to close this gap and provide a comprehensive overview of the current landscape. Your input will contribute to a better understanding of the diverse approaches and strategies used in machine learning education.

This survey is part of the research paper "Navigating the Pedagogical Landscape: An Exploration of Machine Learning Teaching Methods" by Andreea Zlei, under the supervision of Gosia Migut. This research is conducted as part of the Computer Science and Engineering bachelor course CSE3000 Research Project at TU Delft.

This survey will take about 20 minutes to complete. Your participation in this study is entirely voluntary and you can withdraw at any time by closing the questionnaire before submitting the responses. The data collected is completely anonymous. By submitting the responses, you give your informed consent to participate in the research.

If you are interested in the results, the research will be available in February 2024 at https://cse3000-research-project.github.io/2024/Q2.

This section gathers basic information about your background and role in machine learning education.

Academic role with regards to Machine Learning (select all that may apply)

☐ Lecturer
☐ Researcher

Educational institution

○ Maastricht University
○ Tillburg University
○ TU Delft
○ TU Eindhoven
○ University of Amsterdam
○ University of Twente
○ University of Groningen
○ Vrije Universiteit Amsterdam
○ Wageningen University
○ ______ Other...

What is the name of the ML course you teach, and what program is it part of?

______

How many years of teaching experience do you have?

○ Less than 1 year
○ 1-5 years
○ 6-10 years
○ More than 10 years

Are you involved in the ML course's curriculum development? (course contents, course structure)

○ Yes
○ No

This section explores the various teaching methodologies employed in machine learning courses.

What teaching methodologies do you employ in your machine learning courses? (select all that may apply)

☐ Lectures (formal presentations)
☐ Discussions (interactive conversations with the students)
☐ Laboratories or workshops (practical sessions)
☐ Projects (real-world application of concepts)
☐ Flipped classroom (pre-class materials that are then applied in class)
☐ Case studies (an in-depth analysis on a practical problem)
☐ Self-organized learning (independent and autonomous studying)
☐ Blended learning (mixed delivery modes, s.a. online and in-person instruction)
☐ Group Work (collaborative problem solving and task completion)
☐ ______ Other

What technological tools do you use as part of the course?

______

Are there specific pedagogical strategies or techniques you find effective in conveying

complex machine learning concepts to students?

______

Have you experimented with alternative teaching methods in your machine learning courses? If so, what were the outcomes?

______

Do you stay informed about new teaching methods or tools in machine learning education? How?

______

In this section you are asked to dive into the structure and content of your machine learning courses, describe lecture formats, labs, and projects, and share the curriculum contents.

Can you describe the structure of your machine learning courses? (Or provide a link to the structure)

______

What is the curriculum (contents) of your Machine Learning course? (Or provide a link to the curriculum)

______

Do you use any course books? Which ones?

______

How do you incorporate hands-on experiences, such as projects or practical exercises, into your machine learning courses?

What methods of assessment/examination does the course use?

This section explores challenges faced in teaching machine learning.

What are some challenges you face in teaching machine learning, and how do you address them?

You have reached the end of the survey. Press 'Next' to submit your responses.

Powered by Qualtrics

Appendix C: Jigsaw Instructions

- Divide participants into two groups [1 min]
- **Research and Discussion.** Participants research and discuss their assigned aspect of KNN. Collaboration and sharing of insights within the expert groups. [10 min]
- Reorganise participants into new groups [1 min]
- **Teach Other Members**. Participants share their expertise with their new group members. Discussions, questions, and collaborative problem-solving related to KNN. [10 min]
- **Application**. Solve the practice exercise together in groups. [5 min]
- **Reflection**. The teacher assesses the students' understanding and knowledge of the entire topic and facilitates a reflection session where students can discuss the process and what they learned. [10 min]

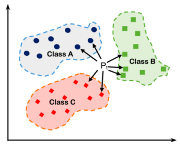
Appendix D: Jigsaw Study Material

K-Nearest Neighbors

Recap

Machine learning is like teaching a computer to recognize patterns and make decisions on its own. It's about enabling systems to learn from experience, adapt to new information, and perform tasks without being explicitly programmed.

Now, one integral part of machine learning is classifiers. These are algorithms that learn to assign categories or labels to items based on their features. Think of them as virtual decision-makers that can sort data into different buckets.

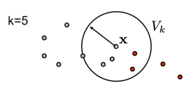


Let's consider a real-life example to grasp this concept. When you receive an email, your email provider uses a classifier to determine whether it's spam or not. The system looks at various features of the email, such as sender, content, and links, to make this classification.

Today, we're going to delve into the world of one such classifier, the K-Nearest Neighbors. KNN is a versatile and intuitive classification algorithm that relies on the proximity of data points. Let's begin by understanding the basic intuition behind KNN.

Intuition

- Use the intuition to classify a new point x:
  - Locate the cell on the new point x
  - Do **not** fix the volume of the cell: grow the cell until it covers k objects: find the k-th neighbours
  - predict the class y of new point x



Imagine you have a dataset where you know the class of each point. We are given a new point, let's call it x and we want to classify it.

A "cell" refers to the region or space in the feature space that encompasses a set of data points. The cell is essentially a grouping of nearby points.

To classify x, we locate the cell on which x falls. Here's the catch: we don't rigidly fix the volume of the cell. Instead, we grow the cell until it covers k objects, finding the k-th neighbours. The class y of x is then predicted based on the majority class of these k neighbours.

Algorithm

- Given:
  - training examples $\{x_i, y_i\}$
    - x_i attribute-value representation of examples
    - y_i class label: {male, female}, digit {0,1, ... 9} etc.
  - testing point x that we want to classify

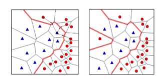
xi represents how we express the features or attributes of data point i in our dataset. Each data point is represented as a vector of attribute-value pairs. For example, we could have a set of data points about people. Each data point represents a person, and the attributes are height and weight. yi is the class label. In our example, the class label could be healthy or unhealthy.

We have a point x that we want to classify.

- Algorithm:
  - compute distance $D(x, x_i)$ to every training example x_i
  - select k closest instances $x_{i1} \dots x_{ik}$ and their labels $y_{i1} \dots y_{ik}$
  - output the class y^* which is most frequent in $y_{i1} \dots y_{ik}$ **(majority vote)**

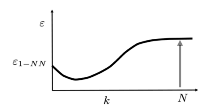
Influence of k

- Value of k has strong effect on k-nn performance
  - Large value ➔ everything classified as the most probable class
  - Small value ➔ highly variable, unstable decision boundaries, eg. for 1-nn:



Choosing k

- set aside a portion of the training data (validation set)
- vary k
- Pick k that gives best generalization performance



Distance measures

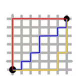
- The key component of the k-NN algorithm
  - defines which examples are similar and which aren't
  - can have strong effect on performance

Euclidean (numeric features):

$$D(x, x') = \sqrt{\sum_d |x_d - x'_d|^2}$$

Manhattan distance

$$D(x, x') = \sum_d |x_d - x'_d|$$

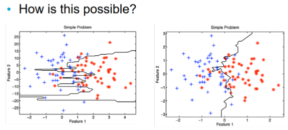


Hamming (categorical features):
- number of features where x and x' differ

$$D(x, x') = \sum_d 1_{x_d \neq x'_d}$$

Sometimes strange results

- How is this possible?



Scale your features!

Sometimes, you get strange results such as the decision boundary on the left. This is why it's important to scale your features when doing k-NN classification. Doing this, we get the result on the right, which looks much better.

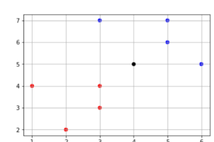
K-NN Pros and cons

- simple and flexible classifier
- often a very good classification performance
- it is simple to adapt the complexity of the classifier

- relatively large training sets are needed
- the complete training set has to be stored
- distances to all training objects have to be computed
- the features have to be scaled sensibly
- the value for *k* has to be optimized

Practice exercise

Given a labelled two-dimensional data set:
– Red label: (1,4); (2,2); (3,3); (3,4);
– Blue label: (3,7); (5,7); (5,6); (6,5);
Predict the label of a new black point (4, 5) using a 3-nn classifier with Manhattan distance.



Appendix E: Lecture Slides

k-Nearest Neighbors (k-NN)

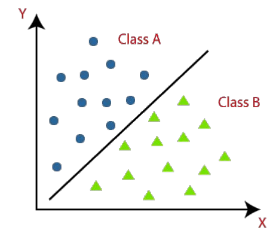
Based on the slides from CSE2510 Machine Learning @ TUDelft

Recap - Machine Learning and Classifiers

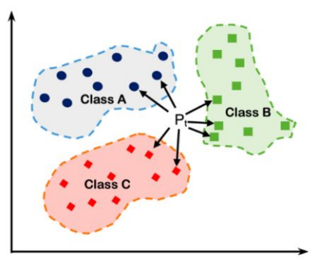
ML = learning from data and improving over time without explicit programming.

Classifiers = Algorithms assigning categories or labels based on features.

Example: Email spam filters determining whether an email is spam or not.

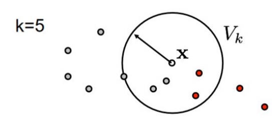


K-Nearest Neighbors classifier



k-NN - Intuition

- Use the intuition to classify a new point x:
  - Locate the cell on the new point x
  - Do **not** fix the volume of the cell: grow the cell until it covers k objects: find the k-th neighbours
  - predict the class y of new point x

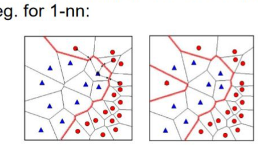


k-NN - Algorithm

- Given:
  - training examples $\{x_i, y_i\}$
    - x_i attribute-value representation of examples
    - y_i class label: {male, female}, digit {0,1, ... 9} etc.
  - testing point x that we want to classify
- Algorithm:
  - compute distance $D(x, x_i)$ to every training example x_i
  - select k closest instances $x_{i1} \dots x_{ik}$ and their labels $y_{i1} \dots y_{ik}$
  - output the class y^* which is most frequent in $y_{i1} \dots y_{ik}$ **(majority vote)**

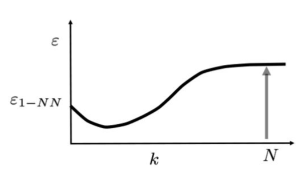
k-NN - Influence of k

- Value of k has strong effect on k-nn performance
  - Large value ➔ everything classified as the most probable class
  - Small value ➔ highly variable, unstable decision boundaries, eg. for 1-nn:



k-NN - Choosing k

- set aside a portion of the training data (validation set)
- vary k
- Pick k that gives best generalization performance



k-NN - Distance measures

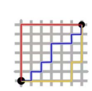
- The key component of the k-NN algorithm
  - defines which examples are similar and which aren't
  - can have strong effect on performance

Euclidean (numeric features):

$$D(x, x') = \sqrt{\sum_d |x_d - x'_d|^2}$$

Manhattan distance

$$D(x, x') = \sum_d |x_d - x'_d|$$



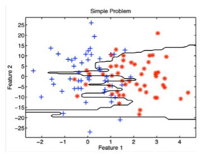
Hamming (categorical features):
- number of features where x and x' differ

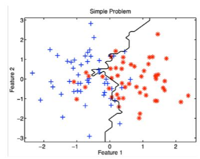
$$D(x, x') = \sum_d 1_{x_d \neq x'_d}$$

k-NN - Scaling is important

Sometimes strange results

- How is this possible?





Scale your features!

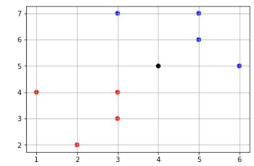
TU Delft

k-NN - Practice exercise

- Given a labeled two-dimensional data set:
  - Red label: (1,4); (2,2); (3,3); (3,4);
  - Blue label: (3,7); (5,7); (5,6); (6,5);
- Predict the label of a new black point (4, 5) using 3-nn classifier with Manhattan distance.

A. Red label

B. Blue label



TU Delft

k-NN - Pros and cons

- simple and flexible classifier
- often a very good classification performance
- it is simple to adapt the complexity of the classifier

- relatively large training sets are needed
- the complete training set has to be stored
- distances to all training objects have to be computed
- the features have to be scaled sensibly
- the value for *k* has to be optimized

TU Delft

Q & A

Appendix F: Test

Based on CSE2510 Machine Learning

k-NN Test

1. Suppose we have a training set consisting of 5 points in 2D space: [(0,0), (0,1), (0.5, 0.5), (1,0), (1,1)], with the corresponding labels: [1,1,2,3,3]. We train a k-NN classifier with Euclidean distance on this training set. What are predicted labels for the following three points: [(0,0.25), (0.25, 0.5), (0.75, 0.5)] for k=1 and k=3?

☐ k=1 -> [2, 2, 2], k=3 -> [1, 1, 3]
☐ k=1 -> [1, 1, 3], k=3 -> [1, 1, 3]
☐ k=1 -> [1, 2, 2], k=3 -> [1, 1, 3]
☐ k=1 -> [1, 2, 2], k=3 -> [1, 2, 2]

☐ I guessed this answer
☐ I answered according to my calculations

2. What is an advantage of k-NN?

☐ Features don't need scaling
☐ Often a good classification performance
☐ Only parts of the training set have to be stored at one time
☐ Small training sets are enough to get a good classification

☐ I guessed this answer
☐ I answered according to my knowledge

3. What effect does the chosen k have on k-nn performance?

4. Describe the steps of the k-NN algorithm

5. Describe how you choose the value of k

6. Which of the following statements is true for k-NN classifiers?

- ☐ k-NN does not require an explicit training step.
- ☐ The classification accuracy is always better with larger values of k.
- ☐ The decision boundary is linear.
- ☐ k-NN is not sensitive to feature scaling.

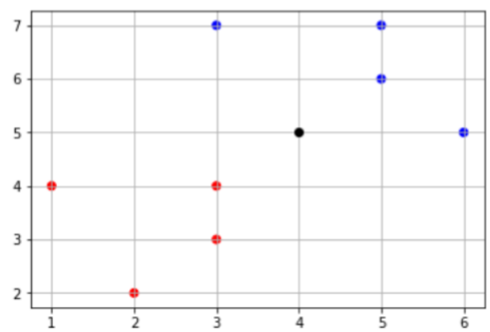
- ☐ I guessed this answer
- ☐ I answered according to my knowledge

Appendix G: Test Answers

Based on CSE2510 Machine Learning

Practice exercise - Answer

Given a labeled two-dimensional data set:
– Red label: (1,4); (2,2); (3,3); (3,4);
– Blue label: (3,7); (5,7); (5,6); (6,5);
Predict the label of a new black point (4, 5) using 3-nn classifier with Manhattan distance.



- ☐ Red label
- ☐ **Blue label**

KNN Test - Answers

1. Suppose we have a training set consisting of 5 points in 2D space: [(0,0), (0,1), (0.5, 0.5), (1,0), (1,1)], with the corresponding labels: [1,1,2,3,3]. We train a k-NN classifier with Euclidean distance on this training set. What are predicted labels for the following three points: [(0,0.25), (0.25, 0.5), (0.75, 0.5)] for k=1 and k=3?

- ☐ k=1 -> [2, 2, 2], k=3 -> [1, 1, 3]
- ☐ k=1 -> [1, 1, 3], k=3 -> [1, 1, 3
- ☐ **k=1 -> [1, 2, 2], k=3 -> [1, 1, 3]**
- ☐ k=1 -> [1, 2, 2], k=3 -> [1, 2, 2]

2. What is an advantage of k-NN?

- ☐ Features don't need scaling
- ☐ **Often a good classification performance**
- ☐ Only parts of the training set have to be stored at one time
- ☐ Small training sets are enough to get a good classification

3. What effect does the chosen k have on k-nn performance?

Large value ➔ everything classified as the most probable class

Small value ➔ highly variable, unstable decision boundaries

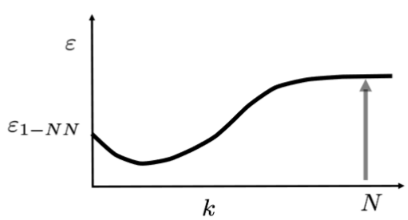
4. Describe the steps of the k-NN algorithm

- Given:
  - training examples $\{x_i, y_i\}$
    - x_i attribute-value representation of examples
    - y_i class label: {male, female}, digit {0,1, ... 9} etc.
  - testing point x that we want to classify
- Algorithm:
  - compute distance $D(x, x_i)$ to every training example x_i
  - select k closest instances $x_{i1} \ldots x_{ik}$ and their labels $y_{i1} \cdots y_{ik}$
  - output the class y^* which is most frequent in $y_{i1} \cdots y_{ik}$ (**majority vote**)

5. Describe how you choose the value of k

Choosing the value of k

- Selecting the value of k
  - set aside a portion of the training data (validation set)
  - vary k
  - Pick k that gives best generalization performance



6. Which of the following statements is true for k-NN classifiers?

☐ **k-NN does not require an explicit training step.**

☐ The classification accuracy is always better with larger values of k.

☐ The decision boundary is linear.

☐ k-NN is not sensitive to feature scaling.