VotestratesML: A High School Learning Tool for Exploring Machine Learning and its Societal Implications

MAGNUS HØHOLT KASPERSEN, Department of Computer Science, Aarhus University, Denmark
KARL-EMIL KJÆR BILSTRUP, Department of Computer Science, Aarhus University, Denmark
MAARTEN VAN MECHELEN, Department of Computer Science, Aarhus University, Denmark
ARTHUR HJORT, Department of Computer Science, Aarhus University, Denmark
NIELS OLOF BOUVIN, Department of Computer Science, Aarhus University, Denmark
MARIANNE GRAVES PETERSEN, Department of Computer Science, Aarhus University, Denmark

Fig. 1. (a) VotestratesML's interface for building ML models based on voter profile data; (b) Students engaged in building models with VotestratesML (Note: all necessary permissions were collected to feature photos of students).

The increased use of Artificial Intelligence, and in particular Machine Learning (ML) raises the need for widespread AI literacy, in three particular areas related to ML; understanding how ML works, the process behind creating ML models, and the ability to reflect on its personal and societal implications. Existing ML learning tools focus primarily on the first two areas, and to a lesser degree the third. In order to address this, we designed VotestratesML; a tool allowing K-12 students to build models and make predictions based on real world voting data. Based on in-situ deployments of VotestratesML, we reflect on opportunities and challenges for engaging K-12 students in understanding and reflecting on ML. We find that the design of VotestratesML supports students' engagement in all three areas of ML, through grounding ML in a known subject area and allowing for collaboration and competition.

CCS Concepts: • **Human-centered computing** → *Interactive systems and tools.*

Additional Key Words and Phrases: Computational empowerment, computational thinking, learning tools, machine learning, AI literacy

Manuscript submitted to ACM

ACM Reference Format:
Magnus Høholt Kaspersen, Karl-Emil Kjær Bilstrup, Maarten Van Mechelen, Arthur Hjort, Niels Olof Bouvin, and Marianne Graves Petersen. 2021. VotestratesML: A High School Learning Tool for Exploring Machine Learning and its Societal Implications. In *FabLearn Europe / MakeEd 2021 - An International Conference on Computing, Design and Making in Education (FabLearn Europe / MakeEd 2021), June 2–3, 2021, St. Gallen, Switzerland*. ACM, New York, NY, USA, 14 pages. https://doi.org/10.1145/3466725.3466728

1 INTRODUCTION

The increased use of Machine Learning (ML) in almost all aspects of our lives [26, 28] raises the necessity of a widespread AI literacy [9, 22]. AI, and subsequently ML, is part of the infrastructure of many everyday technologies in ways that are often opaque and difficult to comprehend [6, 30]. Understanding how AI and ML work is important not only for those pursuing a career in the STEM (Science, Technology, Engineering and Maths) fields but, arguably, for all children, as these technologies are increasingly pervading all aspects of our lives (i.e., education, leisure and work). Thus, AI literacy becomes a precondition to fully participate in society, whereby the term 'literacy' denotes that the goal is not simply to develop children's instrumental skills, but also a critical understanding of manifestations of power and ideology in AI technologies, and consequently, its human and societal implications [36, forthcoming].

The importance of a critical understanding of technologies has also been echoed in recent studies on Computational Thinking (CT). Researchers (e.g., [5, 12, 14, 16, 33]) have argued that traditional CT competencies (e.g., decomposition, abstraction, automation, etc.) should be complemented with critical perspectives on the personal, social and political consequences of digital technology. In these broader and more comprehensive approaches to CT, students learn about technologies and computational infrastructures, not just in order to become better programmers or designers, but also to enable them to engage with political and ethical questions about technology in the real world. As Kafai et al. [16] put it, these approaches see the *"cognitive understanding of underlying concepts of CT and its uses in the world as key to becoming a more critical practitioner of computation"*. We agree with the critical approaches above, and align ourselves in particular with the notion of Computational Empowerment (CE) [7, 14]; a perspective on computing education rooted in Scandinavian participatory design values of democracy and skilfulness that aims to develop children's critical understanding and informed decision-making with regards to the role of digital technology in their lives and society more broadly [14]. Using CE as an approach to AI literacy, raises the question as to how ML learning tools can qualify and support children's critical reflection and understanding of ML, including its human and societal implications.

As a result of the recent awareness about the need for students and the general public to understand AI and ML, researchers have started exploring what K-12 ought to know about ML [34] in order to become AI literate [22], which can be summarised as three main areas or learning objectives; *1)* what is a ML system, *2)* how are ML models developed, *3)* and what are possible implications of ML applications? An important aspect to consider here is what Blikstein [4] terms *selective exposure*, i.e., which aspects of the technology should be foregrounded to children, and how to maximise what can be achieved with the technology through a certain tool. In current research, we often see the former two areas emphasised in educational tools and practices [17, 37, 39], while the third is either neglected or only briefly touched upon (see e.g., [39], which seeks to engage ML novices in interactive model building, but leaves notions of empowerment to future work). This trend is also demonstrated by Giannakos et al. [11], who identify several projects/games for teaching AI/ML, where most focus on the first two areas (e.g., ML4Kids[1] and AI4Children[2]), while a few include the third area (e.g., The Moral Machine [3]). This aligns with findings of Van Mechelen et al. [35], who, based on a systematic review of

[1]https://machinelearningforkids.co.uk/
[2]https://www.ai4children.org/
[3]https://moralmachine.mit.edu

child-computer interaction research, concluded that design ethics (i.e., the human/societal impacts of technology and its qualities in use) has been been underdeveloped in literature, both as an overall concern and an explicit learning goal for children. Thus, while interesting work has been done that can inform the design of tools and activities to teach students about ML, to our knowledge, no studies have taken a CE approach to learning about ML, which entails a more equally distributed focus on all three main areas. With the aim to do so in this paper, we explore the following questions: a) *what should students learn about ML in order to support CE?* and b) *how do we design learning tools and -activities that support students in engaging with ML?* and c) *what are opportunities and challenges for a CE approach to meaningfully engage students in understanding and reflecting on ML?* To explore these questions, we designed VotestratesML (see Figure 1); a prototype of an interactive, collaborative learning tool for ML, aimed at supporting high school students in exploring and reflecting on the role of data and ML in political campaigns. VotestratesML was iteratively developed in a Constructive Design Research (CDR) process [21, 32], and serves both as a proposal for a high school learning tool for ML, and as a means to explore the design space of ML learning tools in high schools. The process was realised through six user-interventions in two social studies high school classes. To inform the future design of learning tools for ML, we present our findings and how they formed the design of VotestratesML.

With this paper, we address Eriksson et al.'s [10] call to open up new frontiers for research that carries FabLearn Europe / MakeEd beyond the foundation and legacy of Papert and Harel [27] in order to embrace the entire education system with a broader perspective of 21st century learning, and encompass diverse issues related to children's use and understanding of emergent technologies [10]. More specifically, our contribution is a synthesis of what students should know about ML as seen through the lens of CE, as well as the design of VotestratesML and the ways in which this tool can engage students in learning about ML and critically reflect on its impacts.

2 WHAT STUDENTS SHOULD LEARN ABOUT MACHINE LEARNING

In this section we address RQ a) *what should students learn about ML in order to support CE?*, by reviewing existing literature about teaching ML to K-12 students, divided into three main areas that correspond to current research trends about what these students should know about ML [22, 34]; *1)* what a ML system is, *2)* how ML models are created, and *3)* what possible implications of ML applications are. Often learning tools for ML address the first two areas and to a lesser extend the third, but we argue that a focus on CE requires a specific focus on the third area in combination with the former two.

2.1 What a ML System is

In the literature, we see a focus on teaching what a ML system is. Along these lines, Touretzky et al. [34] emphasise how computers learn from data and maintain models/representations of the world based on this data. Similarly, Long and Magerko [22] state competencies such as distinguishing between artefacts that use and do not use ML, understanding the strengths and weaknesses of ML system, and recognising how ML systems reason and make decisions as core in developing a literacy about ML and AI. This area is reflected in most tools for teaching ML. Many block programming languages include this such as AI Snap! blocks [17] that allow students to build ML applications in the Snap! block-programming language, using predefined ML blocks. Similarly, *Cognimates* [8] and *Machine Learning For Kids* from IBM allow children to experiment with speech recognition, object recognition, etc., through the *Scratch* programming language. Other tools specifically designed for teaching ML also focus on this area. Popbots by Williams et al. [37] introduce the basics of ML systems, such as training a classifier by labelling data input. *Cozmo*, a robot from Anki, allows children to play with built-in object and marker detection, face recognition etc.

2.2 The ML Process

Another current focus is on the process of creating ML models. Here, Long and Magerko emphasise the different steps involved in creating ML models and the role humans play in doing so. Several learning tools focus explicitly on this aspect of ML. Hitron et al. [13] introduce the *Gest* system, a ML-based gesture recognition system designed to teach children aged 10-13 about data labelling and evaluation of ML models. They conclude that children are able to understand ML processes, specifically data labelling and evaluation of models, and that children are able to apply this knowledge in other contexts. Zimmermann-Niefield et al. [39] deploy learning tools for ML into a high school context. They design and evaluate *AlpacaML*, an iOS application for building models of athletic moves through sensor-input, for use in physical education in high school. They found that students developed their own theories on what constituted a "good" model and were focused on the performance of their model rather than whether the model was true to what it modelled. Kaspersen et al. [18] and their Machine Learning Machine try to encompass the entire process of building ML systems, from collecting data to training and later evaluating a model, and the iterative nature of doing so.

2.3 Implications of ML

Finally, there is a focus on the implications that ML systems might have on society and in our personal lives. Long and Magerko [22] stress that students should understand that "data cannot be taken at face-value" and that some systems have the ability to "physically act on the world". Finally, they point out that students should be able to "[i]dentify and describe different perspectives on the key ethical issues surrounding AI". Similarly, Touretzky et al. [34] present the idea that AI and ML systems "can impact society in both positive and negative ways" and that this is a core focus to teach. We also see this in research on learning tools, albeit to a lesser extend. Bilstrup et al. [3] engages high school students in designing ML systems in a card-based design workshop which engages them in discussions of the ethical dilemmas in applying ML to solve real-world issues. Similarly, Skinner et al. [31] explore children of colors' perception of fairness in ML/AI systems in a co-design workshop in which 9-14 year old children design a "fair" AI librarian.

As mentioned, these areas are not mutually exclusive. For example, understanding the ML process requires some knowledge of how ML works and vice versa. This is seen in most tools targeting these aspects. E.g., in a later paper, Zimmermann-Niefield et al. [38] expand AlpacaML to integrate with Scratch allowing students to build gesture-controlled interactive media and to engage more with understanding different components of a ML system and how they fit together. Similarly, tools targeting the former two areas, sometimes include reflections on ML implications. Hitron et al. [13] also investigate if children after using Gest are able to reflect on situations where it is inappropriate to use ML and find that about half of the participants were able to identify issues about privacy, intimacy, and safety. Kaspersen et al. [18] propose scenarios explaining how the Machine Learning Machine could be used to explore ethical issues of ML, such as data representation and biases, but do not evaluate this aspect. These aspects are, however, most often not presented as a core aspect of the learning tool design, but as an addition that requires a separate discussion.

From a CE perspective, understanding and being able to reflect and act on the implications of ML is the most important of the three areas, and should be at the core when designing ML learning tools. However, to be able to do so requires some insight into what ML is and the process of creating ML models and humans role in doing so. Thus, all three areas must be covered, at least to some extend, by learning tools aiming for CE, and research is needed that explores how to do so. In the following section, we present VotestratesML, which as a learning tool embodies our exploration of how a balance between the three areas could be struck.

3 VOTESTRATESML: A COLLABORATIVE LEARNING TOOL FOR TEACHING ML, ITS PROCESS, AND ITS IMPLICATIONS

VotestratesML is a web application enabling students to collaborate in real time on iteratively building models for predicting voter behaviour using voter profile data, and which aims to scaffold class-wide discussions about ML. Data from a survey of the Danish national election in 2015 [24] is used in the prototype to make predictions about real world voters. This section addresses RQ b) *how do we design learning tools and -activities that support students in engaging with ML?*, as VotestratesML and the rationale behind it embodies our suggestion for how such a tool might look. First, we describe a typical use scenario in a classroom, then we detail the VotestratesML prototype and user interface (UI) and highlight certain aspects of the design of the tool. Finally, we provide the overall design rationale behind the tool.

3.1 A VotestratesML Use Scenario

In a typical use situation of VotestratesML, students are introduced to the tool and how it supports building ML models. Next, students are divided into groups of 3-4 people. Each group member logs into VotestratesML on their laptop gaining access to the collaborative ML tool, see Figure 2a. The teacher then asks groups to create the best possible model for, e.g., predicting if a person will vote for The Social Democratic Party. Each group chooses the *Votes for The Social Democratic Party*-label available in VotestratesML. The groups go on to discuss which features to include in their model. Here, students can draw on their existing knowledge from social studies class. For example, according to the *Michigan Model* for voter behaviour [15], which students are taught in advanced social studies, voting behaviour is affected by family, and thus, it might be important to choose features describing how the parents of a person voted. Once students agree on a set of features, they need to choose an algorithm, and to determine its parameters. Once the model is trained and tested, students can change their model parameters, features, shuffle the data or change the training-data/test-data ratio to try to improve the model. This workflow is summarised in Figure 3. When they are satisfied, students push their model to a shared view projected on the wall in the classroom as seen in Figure 2b. From here, they can compare it to models from other groups and discuss how to improve their model. Finally, when the teacher ends the exercise, a more detailed view is projected on the wall (see Figure 2c) and used in discussions of students' choice of features and model parameters, the predictions of the models, and the implications of this way of working on politics and society. The topic of these discussions depend on the task given to the students, but might include how these predictions can be used to target political advertisements to specific demographics and what the consequence of this is.

3.2 Collaboratively Building, Evaluating and Reflecting on ML Models with VotestratesML

To support real time collaboration, VotestratesML is built in Webstrates [19], which synchronises DOM elements between multiple clients in real time. Each student works on their own laptop, but collaborate with their group through a shared Webstrates-based website, making the interface of VotestratesML inherently collaborative between group members' laptops. Every group member has full control over the application and they must communicate to coordinate their work, the intention being that students will divide tasks with different focus points between them.

VotestratesML consists of three interdependent components that support different types of class activities through providing different functionalities and model views. The relation between components is illustrated in Figure 2.

Using the *Collaborative Component*, see Figure 1a, students collaborate in groups by building, testing and improving the group's ML model for predicting voter behaviour. VotestratesML conceptualises the process of building ML models in five steps as illustrated in Figure 3. First students process the data-set by shuffling the data and splitting it into

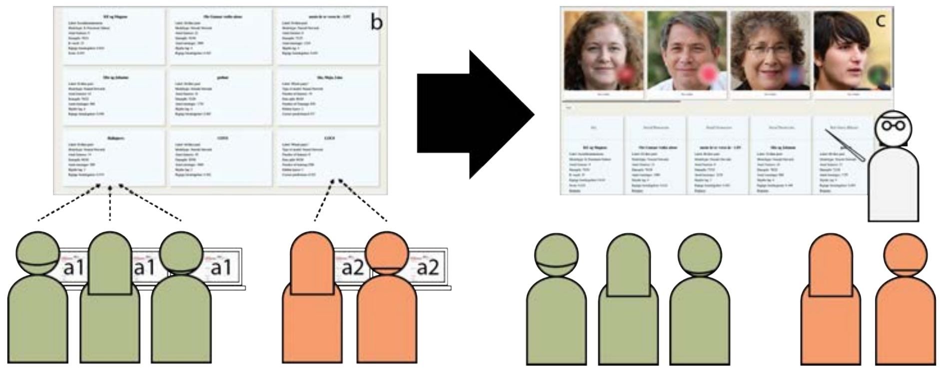

Fig. 2. The different components of VotestratesML and their use: *a* is the *Collaborative Component*, *b* the *Competitive Component*, and *c* the *Discussion Component*. By using the Collaborative Component, groups of students build models, which are pushed to the Competitive Component, allowing groups to compare their models. The Discussion Component is used afterwards to discuss ML, mediated by a teacher.

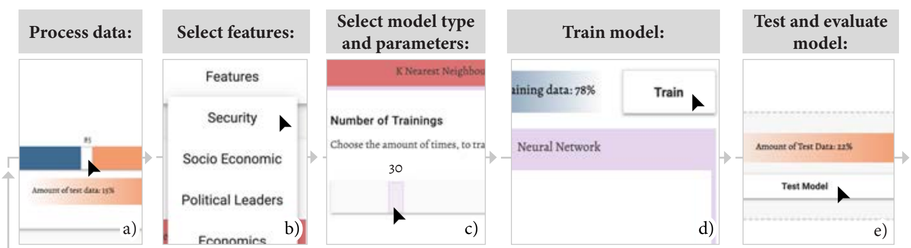

Fig. 3. The iterative process of building ML models (cropped): First, in *a)* basic data processing can be made; in *b)* features are chosen; In *c)* parameters of the algorithm are chosen; in *d)*, the model is trained, and finally, in *e)* it is tested. Students can then start the process over improving the model, e.g. adding more features or changing model parameters, or they can share it to the Competitive view, if they are satisfied.

Fig. 4. The competitive aspect of Votestrates ML (cropped): *a)* shows how students can compare their models while working to improve them; *b)* shows how the teacher can use students' models on personas; *c)* shows how predictions can be discussed and how the models can be inspected in detail.

a training dataset and a test dataset. Second, they explore the properties of the data-set, which range from age and gender to voters' attitudes towards environmental issues and tax policies, and choose one label and as many features as they like. Third, they choose between two ML algorithms; a K-Nearest Neighbour (KNN) or a Feedforward Neural

Network (FNN) and set the parameters for the chosen model (For KNN: value of k, for FNN: layers and number of iterations at the training step). Fourth, they train a model, and last, they evaluate the model to see how it performs on the test dataset and get an *accuracy* and an *f1* score, which they can use to compare their model with earlier attempts or other groups' models. Students can freely jump between these five steps and group members can work on different steps simultaneously. The interface is designed to support students in following this process, even if they do not fully understand ML. It does so, by only allowing actions that bring the application to meaningful states, greying out buttons and hiding UI elements when not relevant. Thus, students can work on creating functional ML models from the beginning and can explore ML further by tinkering with data and model parameters to improve the predictive ability of their models. The *Competitive Component* has a single view, see Figure 4a, which is projected on a shared screen in the classroom and provides information about the scores of each group's model. Models are pushed to the Competitive Component by each group as they iteratively try to improve them. Using this component, students can share their models with the classroom, and see how well they fare in comparison with other groups' models. Finally, the *Discussion Component* also has a single view which is projected on a shared screen and used to mediate a joint discussion in the class. This component is used to inspect students finished models when the teacher ends the activity of building them. The Discussion Component provides full information about each model, see Figure 4c, and lets the teacher use students' models to predict the behaviour of different personas, see Figure 4b. This is used as an anchor point for discussions about the implications of ML for elections and broader discussions about ML in society.

3.3 Design Rationale: Machine Learning for Social Science

VotestratesML differentiates itself from similar learning tools in a number of ways. First, it combines the three areas of what students should know about ML; *what a ML system is, the ML process* and *the implications of ML* (with an emphasis on the latter). Second, VotestratesML takes a top-down approach to introducing students to ML. Instead of starting by introducing ML concepts, it takes departure in the social studies subject as a tool for analysing voter behaviour. The tool is designed to support typical social studies class activities; Students work in groups on tasks assigned by the teacher, followed by a discussion based on the group work. The available data categories were chosen based on models of voter behaviour taught in Danish social studies classes [15], allowing students to explore how ML models can predict voter behaviour, compare theoretical models with ML models, and discuss ML models from a social studies perspective and how they are already used actively to affect the outcome of democratic elections [1, 25, 29]. By focusing the design of VotestratesML on making predictions about people's voting behaviour based on properties like age, income, gender, etc., students can draw on their prior knowledge to explore the extent to which their knowledge and preconceived notions are supported by real-life data. Thus, their prior knowledge and interest in the topic of voter behaviour can bootstrap their exploration of machine learning methods and the implications of these methods.

In addition, the design of the interface aims to scaffold students in building ML models, even with a very limited understanding of ML. Through experimentation with building models for predicting voter behaviour and iteratively improving these models, students gain experience with the trial-and-error process and begin to gain an understanding of ML models. The KNN and FNN algorithms were chosen because they exemplify the extremes with regards to ML complexity; KNN being simple and FNN being complex. KNN was used to illustrate the basics of ML, while FNN was used to illustrate how ML can become almost too complex to understand. As such, students were not expected to understand the FNN algorithm, rather they could experience how difficult it can be to explain its predictions.

By providing a collaborative interface in which groups are forced to negotiate design choices internally because all group-members have full control of the application, VotestratesML aims to scaffold reflections on the process of

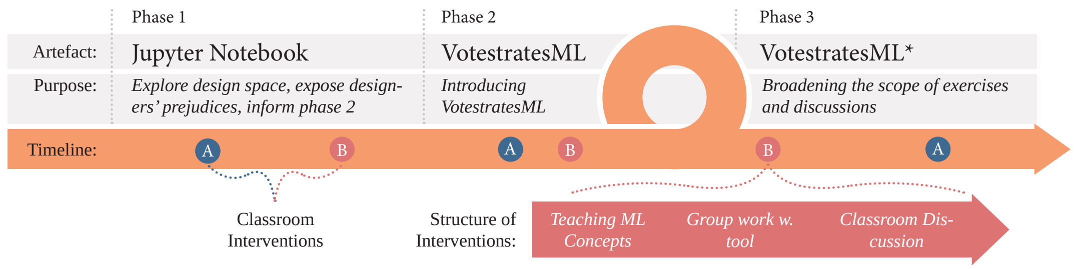

Fig. 5. An overview of the design process of VotestratesML. The study spanned approximately four months. Each class was involved three times, and each intervention had the same basic structure.

building ML models. Similarly, the competitive aspect of VotestratesML aim to scaffold students in considering how their models can be improved, and VotestratesML allow students to do so through trial-and-error, which is analogous to how ML models are often improved in real-world situations. Thus, students experience on first-hand the design choices and potential issues embed in ML models. The design of VotestratesML aims to actively scaffold insightful discussions about the implications of ML and to motivate students to further explore ML by interacting with the tool.

4 METHOD

VotestratesML is the result of a constructive design-research (CDR) process and was evaluated and interated on through several real world deployments. In this section, we present the methodology behind our work, the design process leading to the design of VotestratesML as well as how it was evaluated in the field.

We adopted a Constructive Design Research methodology [20, 21] in order to investigate how ML learning tools can be designed for K-12 classes and how they can support critical discussions and reflections around the use of ML. According to Bang et al. [2], CDR can be seen as a way of iteratively making and testing hypotheses, where knowledge-generation revolves around the construction of an artefact (e.g. a product, a service, media etc.) [20, 32] and experiments with this [2, 21]. In this work, VotestratesML is the central artefact, and our findings are based on the different experiments leading to the version of VotestratesML presented above as well as the lessons learned from the in-situ deployment of VotestratesML, which are presented below. The hypotheses central to the development of VotestratesML are similar to the two central arguments of our design rationales; that grounding ML in an existing subject (i.e. social studies) can engage students already interested in the subject but not in ML, and second that collaborating in groups and competing against other groups can be an effective way of scaffolding students' reflections about ML. These aspects will be discussed further below. An overview of the entire design process can be seen in Figure 5.

4.1 Participants

We collaborated with two social studies teachers from a public high school, located in an upper middle class area in a mid-sized Danish city, each inviting us to work with one of their classes. We consented 30 students in one class (A), and 31 in the other (B), aged 17 to 20. All participants under the age of 18 were asked for the consent of their parents or legal guardians. Consent forms were collected by the researchers at the beginning of the interventions. All students from both classes had elected into a 3-year social studies class. Students from class A were in their third year, while those from class B were in their second year. In both classes, the gender ratio was approximately equally distributed.

4.2 Design Process

VotestratesML was designed in an iterative design process with three main phases (as illustrated in Figure 5). In each phase of the design process, each class participated in an intervention with the dual purpose of teaching students about ML and informing the design of VotestratesML, which was introduced in phase 2 and iterated on in phase 3. This resulted in a total of six interventions. All interventions took place during regularly scheduled social studies class, with each intervention lasting 90 minutes. Two researchers and a teacher were present during each intervention. The researchers took over teaching during the intervention, while the teacher focused on helping students with exercises and contributing during class discussions.

The first phase of the design process focused on introducing ML to social studies students and informing the first version of VotestratesML. We did so by introducing students to central ML-concepts and letting them working with these in a collaborative Jupyter Notebook, based on Google Collaboratory. In the second phase we explored how to contextualise ML for a social studies class, by deploying the first iteration of VotestratesML. We introduced new concepts integral to using VotestratesML, in particular the K-Nearest Neighbours (KNN) algorithm, and let students work in groups using VotestratesML to build social-studies specific ML models, such as creating a model for predicting who will be voting for the Social Democrats. This was followed by a discussion of how such models could be used during elections. Finally, in the third phase focused on making the complexity and ambiguity of ML explorable for students and to support discussions regarding more general implications of ML. We introduced students to neural networks and asked them to work in groups to explore these in their group work with VotestratesML. As a last exercise, we asked students to build models for predicting the voting behaviour of the present researchers. Finally, we facilitated a closing discussion on the role of AI in society and how it will evolve in the future.

4.3 Data Gathering and Analysis

Throughout the interventions data was collected in the form of observations and field notes, sound recordings and photography. Following each intervention, a write-up [23] of observations and field notes was produced and discussed between the two researchers present at the interventions. Audio recordings of students' intra-group discussions during group work and of class-wide discussions were reviewed by two of the authors, and interesting excerpts were selected and later transcribed by two researchers. We analysed data with regards to themes and critical incidents informing our hypotheses. With regards to grounding ML in social studies, we looked for ways in which students' used ML concepts and their experiences with VotestratesML to identify societal implications of ML in discussions internally in the groups and in class-wide discussions. Regarding collaboration and competition we focused on students' engagement and frustrations, how they discussed design-decisions while using VotestratesML and for breakdowns with the prototype.

5 OPPORTUNITIES AND CHALLENGES FOR A CE APPROACH TO MEANINGFULLY ENGAGE STUDENTS IN UNDERSTANDING AND REFLECTING ON ML

As argued, we see a need for CE learning tools that support students in understanding ML, the process of creating ML models, and in reflecting on the implications of ML and how it should form our future lives and societies. Here, we present and discuss the findings from designing and evaluating VotestratesML as they relate to RQ c) *what are opportunities and challenges for a CE approach to meaningfully engage students in understanding and reflecting on ML?.*

5.1 Implications of Machine Learning: Grounding ML in a Well-known Subject Area

VotestratesML frames ML as a social studies tool for predicting voter behaviour, which can be used to achieve goals in the social studies subject. With the tool, students can utilise theoretical models from social studies or test their own intuition about who votes for whom. In this way, the students can engage with ML both from the perspective of the ML technology and its concepts and processes, but also approach it from a well-known subject area. In particular the narrative around ML and the models which students build, modelling voter behaviour, inherently frames the discussions in terms of potential societal implications of ML.

This design is based on our experiences in phase 1, where we observed that many students did not see themselves as individuals who could or should understand the technical aspect of ML. The exercises in Jupyter notebooks were centred around basic components of ML, such as features, labels and training- and test-data. Many students quickly distanced themselves from the exercises with comments such as *"I'm not very good at math"* and *"This looks very complex"* before having invested any time with the prototype. As one student put it: *"For this to be really exciting, you probably need to be interested in the subject [ML]"*. Although students were very eager to discuss ML, they did not draw on ML concepts in their discussions. In contrast, in the interventions with VotestratesML, students' models and their predictions were used as basis for class discussions in which students' argumentation was grounded in social studies theory and in their existing knowledge of voter groups, or based on their experiences from trial-and-error processes with the prototype. Different types of arguments led to different discussions concerning the use of ML. For instance, in phase 2, students role-played as advisers for the Social Democratic party, and were asked to advice, based in ML models, how to best persuade voters from another party. The exercise spurred discussions about the ethics of using ML in political campaigns, e.g.: *"The Social Democrats do vote-seeking [in the role-playing exercise]. But I don't have much respect for that. I would rather go for policy-seeking..."*. In another discussion about how micro-targeting could make democratic parties focus on single issues instead of prioritising coherent ideologies, one student argued: *"It is a democratic problem, if you [democratic parties] focus on single issues rather than a coherent ideology or harmony in their policy.. it will lack coherence and will not work"*. Furthermore, the predictions of the personas' behaviours spurred technical questions about ML, which could be verbalised using students' models about, e.g., how models with high accuracy sometimes make poor predictions or why models with more features are not always better.

Instead of requiring students to "be good at math" or have an initial interest in ML in and of itself, this approach allows students to utilise ML for achieving meaningful objectives and advance their proficiency as e.g. social studies students, while learning about ML along the way. We argue, that the framing of CE learning tools as belonging in existing subjects is an effective way to engage more students with different subject-related backgrounds, interests and knowledge in working with ML-technologies.

5.2 The ML process: Supporting Reflection through Collaboration & Competition

VotestratesML aims at exposing the "messiness" of ML and to have students experience how ML processes are based on human choices, judgements and trial-and-error. To achieve this, VotestratesML guides students through the iterative and messy process of building ML models. It black boxes many of the underlying ML mechanisms, but exposes many of the ambiguous choices and the complexity embedded in ML models. Students are, e.g., encouraged to experiment with features, parameters and models through an interface that allows for quick iteration and collaboration, see Figure 3.

The iterative process of building ML models combined with collaborative and competitive aspects of VotestratesML spurred reflection on and discussion about ML during the activities. While competing in small groups in phase 2 and

3, students collaborated to build the best possible model in order to beat the other groups. Students were observably motivated by the competitive aspect of the exercises and the option for each group to compare themselves to the other groups seemed to give students a sense of ownership, and pride in their models, motivating them to keep improving it. For example, groups regularly cheered after having tested their new model and some students would make exclamations such as *"Share it, share it, share it! We need a group name! It's a great start, nobody is close to us"* or *"We should choose this features, nobody else has used this one!"*.

To beat the models of other groups, students would eagerly discuss how to gain an advantage over their competitors. These discussions were, although based on students' understanding of social studies and their personal prejudices, discussions about ML, e.g., whether a new feature was important prior to selecting it:

Student A: *"What about this one [a feature]: 'Supports a political party.' It could be good."*
Student B: *"Yes.. But isn't it very typical for young people to issue-vote? Like, just voting for what you feel."*
Student A: *"Yeah.. You might be right about that."*
Student B: *"Can you try train it? [the model with the new feature]"*

The discussions and experiments with different features spurred students' curiosity about ML, leading them to ask questions such as *"What is a good score?"* and to ideas about why other groups' models performed better than theirs; *"What if we have trained our model too much, so it now has become bad?"*. On the other hand, the competitive elements of VotestratesML made some reflections superficial due to the high pace of the competition. Often, students would brush over curious results in order to have time for another attempt to beat their peers. This indicates that while competition is a strong motivator, it is important to strike a balance between encouraging competition and encouraging reflection.

The experience of building and optimising models by any means, however, also led to nuanced reflection on the issues around ML afterwards and students taking critical stances towards the technology. For example, in a classroom discussion in phase 3 one student argued: *"It is against our culture, where you are not allowed to generalise by any means. That is outright what this is built on."*, while another argued: *"Machines only look at what we ask them to look at. If they have made a model, it is us who decide, what they should be looking for. Humans are better at marginalising than machines are. If we do not make them marginalise groups, they won't do it!."*. Another student recognised problems with ML models on how to choose which norms should guide the models decision: *"It is very different from person to person, which values you find most important. [...] It is just very different, and it [an ML model] may not be able to predict this"*.

5.3 What a ML system is: Conceptualising ML for K-12 students from a CE Perspective

Where other tools mainly focus mainly on what ML is, and the process of creating ML models, we have worked towards exposing the implications ML through the interaction with and activities around VotestratesML.

This approach worked well at engaging students in understanding ML and reflect on its implications, but we also experienced that students formed misunderstandings about ML, which hindered their reflection: During phase 2, we observed that students had difficulties understanding the connection between the data and the voter behaviour predictions of their models. If choosing a feature did not have the expected effect or a model made an unexpected prediction, most students rejected what the model actually stated, and instead wrote it off as a problem with the data, rather than reconsider the veracity of their expectations. Especially the generalisability of predictions seemed to be difficult for students. To explore this, in phase 3, we asked students to use VotestratesML to predict how the researchers present at the interventions would vote. While working on the models, students asked the researchers questions such as: *"What do you [the researcher] think of this?"* or *"How do we know, what topics you [the researcher] care about"*, indicating that students thought of the models as able to predict the voting behaviour of a specific person or group of persons, rather than as general models for predicting voting behaviour. We believe this misunderstanding about the role of data stems from the data-set being relatively black-boxed in VotestratesML. As argued above, this was done to scaffold students in quickly building ML models, but might have hindered their understanding of the data-set and what each row in the data-set represents. This highlights the importance of finding the right concepts to black box and the right ones to glass box, something also noted by prior research in ML learning tools [13], and that the balance between the black and glass boxes is not trivial to achieve.

Furthermore, we chose to use the terminology used in typical ML activities (features, labels etc.) and while most students were able to use many of these terms in discussions, these also revealed misapprehensions in students' understanding of the terms. We wonder if the use of these terms is an appropriate way to conceptualise ML in high school. It would perhaps be more apt to draw on concepts already known to students, such as dependent and independent variables, or it might be better to simply describe them as input and output. We encourage future research to explore what terminology is most appropriate for introducing ML in high schools.

Although students' conceptions of ML were challenged, we also observed how their understanding of ML helped them reflect on, discuss, and form opinions about the use of ML in society, and argue for these opinions using terminology from social studies theory and from ML. We see VotestratesML as a response to recent calls on "pulling back the curtain" of intangible and abstract computational systems [16], by aiming to provide a concrete starting point for discussing ML. VotestratesML achieves this by having students build interactive ML models, explore how data and computation can be used to predict people's behaviour, and by using students' own models to discuss the societal and personal implications from a social studies perspective. Furthermore, we have aimed to illustrate for High School students, that ML is not just a powerful tool that can improve their lives; ML models have both positive and negative implications, and it is important that we, as a society, actively engage in discussions about, and take a critical stance towards it.

6 CONCLUSION & FUTURE WORK

We have in this paper investigated how to design machine learning tools for high schools students aiming towards a general Machine Learning (ML) literacy by asking three questions: a) *what should students learn about ML in order to support CE?* and b) *how do we design learning tools and -activities that support students in engaging ML?*, and c) *what are opportunities and challenges for a CE approach to meaningfully engage students in understanding and reflecting on ML?* Based on related work on what children should know about ML, we have argued for what students should know about ML when the aim is CE. Based on this, we designed VotestratesML; a collaborative learning tool which embodies our approach for engaging students in learning about ML by allowing them to explore ML and its implications for democratic elections. With VotestratesML, we explored how to design ML-learning tools that qualify and support students in reflections and discussions about the implications of ML. Based on the design process and the deployment of VotestratesML in two social studies classes in a Danish high school, we have identified key opportunities and challenges. We found that VotestratesML provided a concrete starting point for discussing ML and its implications and that by framing ML as a social studies tool for predicting voter behaviour, VotestratesML helped to engage students in understanding ML by allowing them to utilise ML for achieving meaningful objectives and to advance their proficiency as social studies students. Further, we found that collaborating and competing about building ML models supported students discussing ML and sparked students' curiosity about how ML functions. We argue that VotestratesML exemplifies an approach of 'pulling back the curtain' of otherwise intangible and complex computational systems that can be used to engage students with different subject-related backgrounds and knowledge in working with ML as well as other emerging technologies. However, more research is needed on conceptualising ML for high school students to allow them to understand and discuss its implications, especially with regards to which concepts to black box and which to glass box and with regards to the terminology used for introducing ML.

ACKNOWLEDGMENTS

A warm thanks to the participating teachers and students for letting us into their classroom. Thanks to The Danish National Archives for allowing us access to the data-set. Thank you to our colleagues at the Department of Computer Science and the Center for Computational Thinking and Design at Aarhus University for feedback and comments. Specifically, thanks to Clemens Nylandsted Klokmose and the Webstrates community for providing technical support during development of VotestratesML. This work was supported by a research grant (#28831) from VILLUM FONDEN.

REFERENCES

[1] Nick Anstead. 2017. Data-Driven Campaigning in the 2015 United Kingdom General Election. *The International Journal of Press/Politics* 22, 3 (2017), 294–313. https://doi.org/10.1177/1940161217706163

[2] Anne Louise Bang, Peter Gall Krogh, Martin Ludvigsen, and Thomas Markussen. 2012. The Role of Hypothesis in Constructive Design Research. In *The Art of Research 2012: Making, Reflecting and understanding*. Helsinki, 1–11.

[3] Karl-Emil Kjær Bilstrup, Magnus H. Kaspersen, and Marianne Graves Petersen. 2020. Staging Reflections on Ethical Dilemmas in Machine Learning : A Card-Based Design Workshop for High School Students. In *DIS '20: Proceedings of the 2020 ACM Designing Interactive Systems Conference*. Association for Computing Machinery, New York, NY, United States, Eindhoven, Netherlands, 1211–1222. https://doi.org/10.1145/3357236.3395558

[4] Paulo Blikstein. 2013. Gears of our childhood. In *IDC '13 Proceedings of the 12th International Conference on Interaction Design and Children*. ACM Press, New York, New York, 173–182. https://doi.org/10.1145/2485760.2485786

[5] Karen Brennan and Mitchel Resnick. 2012. New frameworks for studying and assessing the development of computational thinking. In *Preceedings of the 2012 annual meeting of the American Educational Research Association*. Vancouver, Canada, 1–25. https://doi.org/10.1.1.296.6602

[6] Jenna Burrell. 2016. How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big Data and Society* 3, 1 (2016), 1–12. https://doi.org/10.1177/2053951715622512

[7] Christian Dindler, Rachel Smith, and Ole Sejer Iversen. 2020. Computational empowerment: participatory design in education. *CoDesign* 16, 1 (1 2020), 66–80. https://doi.org/10.1080/15710882.2020.1722173

[8] Stefania Druga. 2018. Growing up with AI - Cognimates: from coding to teaching machines. , 204 pages. http://hdl.handle.net/1721.1/120691

[9] Stefania Druga, Sarah T. Vu, Eesh Likhith, and Tammy Qiu. 2019. Inclusive AI literacy for kids around the world. In *Proceedings of ACM Fablearn conference (Fablearn'19)*. 104–111. https://doi.org/10.1145/3311890.3311904

[10] Eva Eriksson, Ole Sejer Iversen, Gökçe Elif Baykal, Maarten Van Mechelen, Rachel Smith, Marie-louise Wagner, Bjarke Vognstrup Fog, Clemens Klokmose, Bronwyn Cumbo, Arthur Hjorth, Line Have Musaeus, Marianne Graves Petersen, and Niels Olof Bouvin. 2019. Widening the scope of FabLearn Research : Integrating Computational Thinking , Design and Making. In *FabLearn Europe '19*. https://doi.org/10.1145/3335055.3335070

[11] Michail Giannakos, Iro Voulgari, Sofia Papavlasopoulou, Zacharoula Papamitsiou, and Georgios Yannakakis. 2020. *Games for Artificial Intelligence and Machine Learning Education: Review and Perspectives*. Springer Singapore, Singapore, 117–133. https://doi.org/10.1007/978-981-15-6747-6_7

[12] Shuchi Grover and Roy Pea. 2013. Computational Thinking in K–12. *Educational Researcher* 42, 1 (2013), 38–43. https://doi.org/10.3102/0013189X12463051

[13] Tom Hitron, Yoav Orlev, Iddo Wald, Ariel Shamir, Hadas Erel, and Oren Zuckerman. 2019. Can children understand machine learning concepts? The effect of uncovering black boxes. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems - CHI '19*. ACM Press, New York, New York, USA. https://doi.org/10.1145/3290605.3300645

[14] Ole Sejer Iversen, Rachel Charlotte Smith, and Christian Dindler. 2018. From computational thinking to computational empowerment. In *PDC '18 Proceedings of the 15th Participatory Design Conference: Full Papers - Volume 1*. ACM Press, Hasselt and Genk, Belgium, 1–11. https://doi.org/10.1145/3210586.3210592

[15] Jakob G. Jensby, Anders E. Pedersen, and Peter Brøndum. 2017. *Politikbogen* (3 ed.). Forlaget Columbus. 318 pages. https://forlagetcolumbus.dk/produkt/politikbogen/produkt/vis/

[16] Yasmin Kafai, Chris Proctor, and Debora Lui. 2019. From theory bias to theory dialogue: Embracing cognitive, situated, and critical framings of computational thinking in K-12 Cs education. In *ICER 2019 - Proceedings of the 2019 ACM Conference on International Computing Education Research*. 101–109. https://doi.org/10.1145/3291279.3339400

[17] Ken Kahn and Niall Winters. 2018. AI Programming by Children. In *Proceedings of the Constructionism 2018 Conference*, Valentina Dagienè and Eglè Jasutè (Eds.). Vilnius, Lithuania, 315–324.

[18] Magnus Høholt Kaspersen, Karl-Emil Kjær Bilstrup, and Marianne Graves Petersen. 2021. The Machine Learning Machine: A Tangible User Interface

for Teaching Machine Learning. In *Fifteenth International Conference onTangible, Embedded, and Embodied Interaction (TEI '21)*. ACM New York, NY, USA, 12.

[19] Clemens Klokmose, James Eagan, Siemen Baader, Wendy Mackay, and Michel Beaudouin-Lafon. 2016. Webstrates: demonstrating the potential of Shareable Dynamic Media. In *Proceedings of the 19th ACM Conference on Computer Supported Cooperative Work and Social Computing Companion - CSCW '16 Companion*. ACM Press, New York, New York, USA, 61–64. https://doi.org/10.1145/2818052.2874325

[20] Ilpo Koskinen, John Zimmerman, Thomas Binder, Johan Redstrom, and Stephan Wensveen. 2011. *Design Research Through Practice: From the Lab, Field, and Showroom*. Elsevier.

[21] Peter Gall Krogh and Ilpo Koskinen. 2020. *Drifting by Intention: Four Epistemic Traditions from within Constructive Design Research* (1 ed.). Springer International Publishing. 162 pages.

[22] Duri Long and Brian Magerko. 2020. What is AI Literacy? Competencies and Design Considerations. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery (ACM), New York, NY, USA, 1–16. https://doi.org/10.1145/3313831.3376727

[23] Matthew B. Miles and A. Michael Huberman. 1994. *Qualitative Data Analysis: An expanded sourcebook (3rd ed.)*. SAGE Publications, London, UK.

[24] Kasper Møller Hansen. 2017. Danish National Election Study 2015. https://doi.org/10.5279/DK-SA-DDA-31083

[25] David W. Nickerson and Todd Rogers. 2014. Political Campaigns and Big Data. *Journal of Economic Perspectives* 28, 2 (5 2014), 51–74. https://doi.org/10.1257/jep.28.2.51

[26] Cathy O'Neil. 2017. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Vol. 52. Penguin Books.

[27] Seymour Papert and Idit Harel. 1991. Situating Constructionism. *Constructionism* 36, 2 (1991), 1–11.

[28] Frank Pasquale. 2015. *The Black Box Society: The Secret Algorithms That Control Money and Information*. Harvard University Press. http://www.jstor.org/stable/j.ctt13x0hch

[29] Filipe N. Ribeiro, Koustuv Saha, Mahmoudreza Babaei, Lucas Henrique, Johnnatan Messias, Fabricio Benevenuto, Oana Goga, Krishna P. Gummadi, and Elissa M. Redmiles. 2019. On Microtargeting Socially Divisive Ads. In *Proceedings of the Conference on Fairness, Accountability, and Transparency - FAT* '19*. ACM Press, New York, New York, USA, 140–149. https://doi.org/10.1145/3287560.3287580

[30] Ben Shneiderman, Catherine Plaisant, Maxine Cohen, Steven Jacobs, Niklas Elmqvist, and Nicholoas Diakopoulos. 2016. Grand challenges for HCI researchers. *Interactions* 23, 5 (2016), 24–25. https://doi.org/10.1145/2977645

[31] Zoe Skinner, Stacey Brown, and Greg Walsh. 2020. Children of Color's Perceptions of Fairness in AI: An Exploration of Equitable and Inclusive Co-Design. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems (CHI EA '20)*. Association for Computing Machinery, New York, NY, USA, 1–8. https://doi.org/10.1145/3334480.3382901

[32] Pieter Jan Stappers and Elisa Giaccardi. 2017. Research through Design. In *The Encyclopedia of Human-Computer Interaction, 2nd edition*, M. Soegaard and R. Friis-Dam (Eds.).

[33] Mike Tissenbaum, Josh Sheldon, Lissa Seop, Clifford H. Lee, and Natalie Lao. 2017. Critical computational empowerment: Engaging youth as shapers of the digital future. *IEEE Global Engineering Education Conference, EDUCON* April (2017), 1705–1708. https://doi.org/10.1109/EDUCON.2017.7943078

[34] David Touretzky, Christina Gardner-McCune, Fred Martin, and Deborah Seehorn. 2019. Envisioning AI for K-12 : What should every child know about AI ?. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 33. 9795–9799. https://doi.org/10.1609/aaai.v33i01.33019795

[35] Maarten Van Mechelen, Gökçe Elif Baykal, Christian Dindler, Eva Eriksson, and Ole Sejer Iversen. 2020. 18 Years of Ethics in Child-Computer Interaction Research: A Systematic Literature Review. In *Proceedings of the Interaction Design and Children Conference* (London, United Kingdom) *(IDC '20)*. Association for Computing Machinery, New York, NY, USA, 161–183. https://doi.org/10.1145/3392063.3394407

[36] Maarten Van Mechelen, Marie-Louise Wagner, Gökçe Elif Baykal, Rachel Smith, and Ole Sejer Iversen. 2021. Digital Design Literacy in K9 Education: Experiences from Pioneer Teachers. In *Proceedings of the Interaction Design and Children Conference* (Athens, Greece) *(IDC '21)*. Association for Computing Machinery, New York, NY, USA, 11 pages pages. https://doi.org/10.1145/3459990.3460696

[37] Randi Williams, Hae Won Park, and Cynthia Breazeal. 2019. A is for Artificial Intelligence: The Impact of Artificial Intelligence Activities on Young Children's Perceptions of Robots. In *Conference on Human Factors in Computing Systems - Proceedings*. ACM Press, New York, New York, USA, 1–11. https://doi.org/10.1145/3290605.3300677

[38] Abigail Zimmermann-Niefield, Shawn Polson, Celeste Moreno, and R. Benjamin Shapiro. 2020. Youth making machine learning models for gesture-controlled interactive media. In *Proceedings of the Interaction Design and Children Conference, IDC 2020*. 63–74. https://doi.org/10.1145/3392063.3394438

[39] Abigail Zimmermann-Niefield, Makenna Turner, Bridget Murphy, Shaun K Kane, and R Benjamin Shapiro. 2019. Youth Learning Machine Learning through Building Models of Athletic Moves. In *IDC '19 Proceedings of the 18th ACM International Conference on Interaction Design and Children*. ACM Press, New York, New York, USA, 121–132. https://doi.org/10.1145/3311927.3323139