2020

Critical Media, Information, and Digital Literacy: Increasing Understanding of Machine Learning Through an Interdisciplinary Undergraduate Course

Barbara R. Burke
University of Minnesota, Morris, burkebr@morris.umn.edu

Elena Machkasova
University of Minnesota - Morris, elenam@morris.umn.edu

Recommended Citation

Burke, Barbara R. and Machkasova, Elena (2020) "Critical Media, Information, and Digital Literacy: Increasing Understanding of Machine Learning Through an Interdisciplinary Undergraduate Course," *Irish Communication Review*: Vol. 17: Iss. 1, Article 1. Available at: https://arrow.tudublin.ie/icr/vol17/iss1/1

Irish Communication Review

Volume 17 | Issue 1 Article 1

July 2020

Critical Media, Information, and Digital Literacy: Increasing Understanding of Machine Learning Through an Interdisciplinary Undergraduate Course

Barbara R. Burke
University of Minnesota, Morris, burkebr@morris.umn.edu

Elena Machkasova
University of Minnesota, Morris, elenam@morris.umn.edu

Critical Media, Information, and Digital Literacy: Increasing Understanding of Machine Learning Through an Interdisciplinary Undergraduate Course

Cover Page Footnote

Acknowledgements: We appreciate the cooperation of the students in IS 3255H--Honors course: Machine learning and society: Who trains whom? and especially those who allowed their papers to be shared: Nicole Abraham, Nathan Beneke, Corinne McCumber, and Nick Weninger. The course was designed with support from a University of Minnesota Educational Development Grant. An earlier version of this paper was presented by Burke at the Technological University Dublin's 3rd annual conference of the Centre for Critical Media Literacy: "Emerging technologies, social media and the politics of the algorithm" in November 2019. Burke's travel to the conference was supported with funds from the Faculty Research Enhancement Fund (FREF IIIi).

Critical media, information, and digital literacy: Increasing understanding of Machine Learning through an interdisciplinary undergraduate course

Barbara R. Burke and Elena Machkasova,
University of Minnesota, Morris

Abstract

Widespread use of Artificial Intelligence in all areas of today's society creates a unique problem: algorithms used in decision-making are generally not understandable to those without a background in data science. Thus, those who use out-of-the-box Machine Learning (ML) approaches in their work and those affected by these approaches are often not in a position to analyse their outcomes and applicability. Our paper describes and evaluates our undergraduate course at the University of Minnesota Morris, which fosters understanding of the main ideas behind ML. With Communication, Media & Rhetoric and Computer Science faculty expertise, students from a variety of majors, most with no prior background in data science or computing, reviewed the scope of applicability of algorithms and became aware of possible biases, 'politics' and pitfalls. After discussing articles on societal attitudes towards technology, explaining key concepts behind ML algorithms (training and dependence on data), and constructing a decision tree as an example of an algorithm, we attempted to develop guidelines for 'best practices' for use of algorithms. Students presented a 'case analysis' capstone paper on an application of machine learning in society. Paper topics included: use of algorithms by child protection services, 'deepfake' videos, genetic testing. The level of papers was indicative of students' strong interest in the subject and their ability to understand key terms and ideas behind algorithms, societal perception and misconceptions of use of algorithms, and their ability to identify good and problematic practices in use of algorithms.

Introduction

As Machine Learning (ML) spreads to more and more aspects of society, be it financial, medical, educational, artistic, or daily minutiae, it is essential that society is educated about its key foundations, capabilities, limitations, and effects. Given a large number of misconceptions about Machine Learning and artificial intelligence, decision-makers may put unwarranted trust or distrust into Machine Learning methods, which may have drastic unanticipated consequences for the society. Not understanding the role training data plays in an algorithm's behaviour may lead to perpetuating problematic practices and biases via an algorithm. It is crucial that educators take on the challenge of teaching machine learning concepts and implications to a wide range of learners, not just to statisticians, computer scientists, or data scientists. Such teaching would require explaining the key ideas of machine learning, exploring perceptions of algorithms in society, reasoning about implications of their use, and being able to form an informed opinion about the use of machine learning in various areas of society. These conversations must span multiple disciplines, and it is essential that the terms employed in such discourse are understood the same way by all participants. While the developers of machine learning methods usually have graduate degrees, machine learning is involved in such a wide range of areas that one cannot limit education about the basics of machine learning to graduate programs.

With this in mind, the two authors of this paper have developed and co-taught an interdisciplinary undergraduate course, 'Machine Learning in Society: Who trains whom?' that was offered in the curriculum of the selective, Honors program (like an Honors College) at the University of Minnesota Morris. The two faculty brought in expertise in Communication, Media, and Rhetoric (CMR) (Burke) and Computer Science (Machkasova). There were eleven students in the class, with majors including Art History, Biology, CMR, Computer Science, Elementary Education (Primary School Teacher training), English, History, Mathematics, Medieval Studies, Political Science, and Statistics. Some students double-majored in different areas. The students were juniors (3rd year) and seniors (4th year).

Relevant background

Studies and arguments regarding the uses and misuses of technologies of communication have simplistically been sorted into works which are technologically wary/ distrustful (Czitrom, 1983; Postman, 1992) and those which are technologically enthusiastic (e.g., Walther, 1996). Within the latter category expressions of 'revolutions' and 'social consequences' embrace change as inevitable (Innis, 1951; McLuhan, 1964) and the necessities of a 'marketplace' which responds to a social demand/ push approach to new products and processes (e.g., authors in Neuman, 2010, who assert that understanding older forms of media technology development and adoption can guide us in understanding current and future patterns). With celebrations of progress and suggestions of kinds of improvements for society, communications and computer technologies have been championed as tools for increasing equality (Dahlberg, 2001), scrutinizing the processes of modern democracy (Mumford, 1970; Castells, 1996, Aikens, 1997), and enhancing engagement with communities (Wellman, 1984; Baym, 1995; Kavanaugh, et. al, 2005).

Computer-mediated decision-making, as shown within popular media, is also often offered as an improvement on human practices, for machines are generally viewed as neutral, logical and not swayed by emotions (Mackenzie, 2006). Examples include books, television programmes, films, web-series and others that readily show 'intelligent robots' as emerging at the cusp of tomorrow (Sawyer, 2012). Nonetheless, we suggest and thus teach in our course, that understanding the current reality of AI, machine learning and algorithms (as applied for social practices) requires sceptical and nuanced understandings that distinguish between the potentially possible, and the realized technologies of the present era. Additionally, critical media literacy – the place where communication studies, media literacy, digital literacy and information literacy intersect – is the location for which we assert scholars and students may be particularly able to explore what machine learning can and cannot be. As noted by the 2019 conference on 'Emerging technologies, social media & the politics of the algorithm' at TU Dublin, now is a good time to consider ways we can deliberately and strategically have these important concerns addressed by scholars and students.

Below, in describing our work teaching undergraduate students about the applications and consequences of Machine Learning, we explain: our objectives; course format, activities and assignments; selected course materials and responses; and significant assignments. We then offer our conclusions and ideas for future work.

Course objectives

The goal of the interdisciplinary course was for the students to learn to do the following:

- Question public attitudes about computing and new technologies. Specifically, students were introduced to ideas of a techno-optimistic view and a techno-sceptical view.
- Challenge the assumption that algorithms are fair and unbiased.
- Understand key machine learning mechanisms and the role of training data in determining algorithms' behaviour.
- Expose how data rooted in current social practices may introduce (often unintended) biases into machine learning, which perpetuates these biases in society.

Course format

The course was scheduled to run for 15 weeks, one class meeting (of 100 minutes) per week. Because of two weather-related cancellations, it actually had 13 class meetings. In addition to the class meetings, the course utilized an online course management system (Canvas) that was used for posting readings, submission of assignments, and some amount of discussion (mostly, students submitting additional resources that they thought were relevant).

Students were given weekly assignments which included, reading: some introductions to key terminology in communication studies; popular news and web articles on various uses of machine learning; and selected chapters from Cathy O'Neil's book Weapons of Math Destruction (O'Neil, 2016).

Classroom activities included:

- Lecture-like presentations by instructors based on the readings and students' questions. For instance, the students asked us to provide more information about the 2008 housing crisis mentioned in one of the readings, or background on neural networks.
- discussions of the readings by students, sometimes first as small group discussions.
- group activities, such as a hands-on exercise on a mathematical model called decision trees, and a brainstorming of ideas for an unbiased hiring algorithm for a company.

- Students’ presentations.

These assignments and their takeaways are detailed later in the paper.

The writing assignments included:

- Weekly feedback on assigned readings. We asked students for a short (four to six sentences) feedback in which they addressed the following:
  - what were the most important things they’ve learned from the readings;
  - what was the most surprising or unfamiliar part of the readings;
  - the students were required to ask at least two questions about the readings;
  - optionally, students could add any other feedback they felt relevant.

- Essay-style written assignments. The three assignments in this category stressed critical thinking and increased the analysis component as the semester progressed:
  - Based on an article about using ML in ‘Precision agriculture’ (PrecisionAg.com, 2019) the students were asked to answer questions about key ML terminology (such as: ‘Would you consider the algorithm to be inference or prediction?’), understanding data used in the algorithms, identifying the key stakeholders of the algorithms in question, and identifying its potential benefits and downsides. The purpose of the assignment is to assess students’ ability to identify technical and societal elements of an algorithm.
  - The students were asked to analyse an article about an application of ML (they were given a choice among several articles). While the questions for this assignment addressed the key elements of algorithms, such as training data selection and potential bias, as well as its key stakeholders, the students were also asked to discuss how the algorithm fits into the democratic society, and additionally to find potential gaps in the article’s description of the algorithm that make it difficult to evaluate it fully.
  - The final assignment in this category asked students to discuss what a person whose area of expertise is not data science can do to increase or support justice in a modern society that uses AI.

- Final, capstone papers in which students analysed an application of ML to a big-data problem in a society. The students also presented their chosen topics in class to get feedback before their papers were completed. A description and examples of final papers and presentations are detailed later in the paper.

Course materials and important readings

Presentation of media and communication studies theories

The course began with an introduction to the range and variety of perspectives related to media technology adoption, and subsequent social responses to new ways of communicating. The students were assigned an assortment of excerpted book chapters and essays from our 'Mass Media and Society' curriculum and listened to a supplemental mini-lecture with core definitions, and the names of some key scholars. We had general discussions about technology, with critical arguments about why scholars do not see systems and technologies as neutral. Students then investigated social assumptions about 'progress' (as inevitable and/or beneficial, or problematic for various people and groups) and the assertions of media theorists who embraced or rejected technological determinism. Included in the class lessons were some interpretative, historical arguments about early uses of mass media, and explanations of ways we could, and could not, use the theories that have been applied to telegraphy, radio and broadcast television studies to digital media and systems, and we thus established a critical and grounded perspective to explicate the ways the course's digital and information literacy was connected to, but not synonymous with the basic media literacy skills the students had previously known.

Presentation of technical ideas

Presenting technical ideas behind machine learning to an audience with varying levels of mathematical background requires a careful selection of resources, examples, and the terminology used. Fortunately, many experts have provided a light-weight introduction to their areas in the form of blog posts or public articles. The process of using these resources worked as follows:

- Students were provided with these materials ahead of the class meeting.
- A day or two before the class meeting students would submit their weekly reflections that included their key takeaways from the materials as well as their questions.
- The instructors used these write-ups to put together an in-class presentation clarifying the confusing points. The presentation would mostly use a whiteboard rather than power-point slides, to make it more interactive and flexible, allowing the instructors to adjust based on the students' real-time questions.

To give students a first-hand experience with the approaches used in ML an in-class activity – a walk-through a decision tree algorithm – was used. Since it was quite a significant in-class time investment, only one such activity was used. We discuss it in more detail at the end of this section.

Below are some of the topics and key resources that were covered as a part of technical introduction to ML:

- 'A Non-Technical Introduction to Machine Learning' by Noah Yonack https://blog.safegraph.com/a-non-technical-introduction-to-machine-learning-b49fce202ae8 This resource introduces terms, such as machine learning, artificial intelligence, model, algorithm, training data, testing data, feature, supervised vs unsupervised learning, and others. It also describes the process of machine learning without going into mathematical details.
- 'Introduction to Unsupervised Learning' by a group 'Algorithmia' https://algorithmia.com/blog/introduction-to-unsupervised-learning introduces the ideas behind the main methods of unsupervised learning, such as clustering and data compression. The students were shown an example of K-means clustering.
- 'Neural networks' by Chris Woodford https://www.explainthatstuff.com/introduction-to-neural-networks.html discusses the basics of neural networks. To supplement this resource, the students were walked through an illustration of backpropagation: updates of weights that happen during a network training process.
- 'Introduction to Convolutional Neural Networks' by a company Rubik's cube https://rubikscode.net/2018/02/26/introduction-to-convolutional-neural-networks illustrates how the structure of a convolutional neural network is modeled after the process of human vision. It demonstrates interactions between network layers and helps understand how numeric encoding of data (in this case an image) can be used to recognize its features.
- 'What Is Natural Language Processing And What Is It Used For?' by Terence Mills https://www.forbes.com/sites/forbestechcouncil/2018/07/02/what-is-natural-language-processing-and-what-is-it-used-for/#cb5dfc35d71f introduces students to the key approaches behind natural language processing, including the Hidden Markov Models (without going into the mathematical details) and semantic analysis. The role of context is emphasised, and the idea that two words are related in meaning if they often appear in the same context is illustrated with examples.
- Weapons of Math Destruction by Cathy O'Neil. The book provides an excellent discussion of issues with using ML in society, such as codifying and perpetuating biases in decision making. Four weeks of the course were devoted to discussing selected chapters of the book. Students strongly

related to many examples of social issues affected by ML, including criminal justice, effects of college rankings on college affordability, health insurance price discrimination, hiring discrimination, and targeted social media campaigns. By the time the students were reading the book they were already familiar with the foundational ideas behind machine learning, which made it easier for them to understand how the choice of training data and the criteria for the algorithm's 'success' used in training may be a source of bias in the working of the algorithms. The students' familiarity with societal attitudes about technology (from the readings in Communication theory) helped them understand how algorithms' biases may easily remain unnoticed and unchallenged.

Technical material was challenging to students, especially the ones with less mathematical background. However, such information is necessary to make sure that students have the ability to 'un-magic' ML applications, such as the ones mentioned in news articles.

To illustrate how students dealt with technical material, we include a few questions and comments from their weekly written feedback. This set is for the reading materials on neural networks and convolutional neural networks. This list illustrates the most common themes in the students' write-ups.

1. I did not know that networks as advanced as neural networks existed. I also was unaware that brains and computers 'think' in completely different ways, which is a difference neural networks try to bridge (Art History & CMR major).
2. The reading mentions that no one has attempted to build a neural network in the same way a brain is built with parallel structures, and I'm curious as to why (Biology major).
3. [H]ow does the adjusting of weights during backpropagation work? (Biology major).
4. Why are multiple non-linearity layers (or ReLU layers) necessary in convolutional neural networks? (English major).
5. I was surprised to understand that the computer simply shaves off unnecessary detail and emphasizes key parts of the image until a neural network can identify it (Math major).
6. The article mentions that deep learning systems need feedback, and therefore this type of network is supervised machine learning. I'm curious, would it be possible to create a type of neural network that worked with unsupervised machine learning? (Statistics major). A related question was asked by a Biology major: How does backpropagation work in cases where you don't know what the final output is supposed to be?

7. Can feedback from a neural network trick a human into believing that the computer can feel emotions and make humanlike decisions? (Elementary Education major).

Now we illustrate how this feedback would inform our in-class discussions and general class progress. Comment 1 and similar comments indicate that the students understood the part of the article that discusses the similarities and differences between the computer-based ‘learning’ and the brain quite well, and learned something new from it (even Biology majors mentioned things that they didn’t know about how the brain functioned). The follow-up question 2 prompted an interesting discussion of how the goal of ML (and computing in general) is quite different from actually simulating a human brain, and also of the research focus on creating usable systems for concrete tasks, rather than explore a more challenging field of creating a ‘thinking’ device.

Questions 3 and 4 prompted a detailed whiteboard example of updating weights in a very small neural network with mock-up numbers as weights (i.e. without the actual weight computation). The main point of the example was to show how, starting with random weights, we can use backpropagation to increase or decrease weights where the final label of the outcome is incorrect. For the next offering of the course, it may be a good idea to find an interactive demo of this process. The example also showed how a ReLU allows one to model a threshold so that all values below it is ignored. This motivates the use of non-linearity and allows one to connect question 4 to what is essentially an answer to it, in the comment 5 above (using one student’s comment to answer another student’s question is a great motivational tool).

Question/comment 6 is another place to connect two seemingly different items of feedback: one question is technical, the other one gives the intuition behind it. Explaining how the two are similar is a fruitful teaching moment. The answer to this question also illuminates an important idea: that unsupervised ML can be turned into supervised by using one of the data fields (or a combination of several) as a ‘label’. However, without selecting such a label neural networks are not helpful for unsupervised learning since they look for complex patterns in data, whereas unsupervised methods group data, and there are just too many complex ways in which data could be grouped.

Finally, question 7 opens a door to discussion of the Turing test, an anthropocentric view of computers, and/or whether human needs for interaction can be satisfied by a computer – all of which is within the scope of the course since we are looking at machine learning in society, and social and psychological aspects of interacting with a computer are an important aspect of this topic.

The above example illustrates our weekly preparation process after going over students' write-ups the day before a class meeting. The process allowed us to shape in-class discussions according to the students' needs and interests. We feel that the requirement to write weekly feedback engages students more with the material. We should point out two challenges, though:

- Not all students were submitting their feedback on all readings in a timely fashion. There was a small grade penalty for late work (with the rationale that late feedback is better than no feedback at all, since the students still have a motivation to complete the readings). Given that this group of students, being in the Honors program, was probably more motivated than most other students, the lack of responses (especially from students struggling with the readings) may become a problem if the course were to be expanded to a regular curriculum offering.
- There is not much time between getting students' feedback and the class, so instructors need to be committed to allocating a significant amount of time the day before the class for preparing. Sometimes unexpected questions or comments arise, and this makes the preparation challenging (although quite rewarding since it allows for a dynamic 'dialogue' with students). We also expect that if the class is taught several times, the element of surprise diminishes, and instructors accumulate a collection of examples and resources that can be used in class in response to students' confusions or thoughts.

Significant course activities and outcomes

Decision tree exercise

We felt that students' understanding of the mechanical nature of ML algorithms would be more solid if they had an opportunity to 'walk through' the workings of an algorithm. We chose a simple instance of the decision tree algorithm (Shalev-Shwartz, et al., 2014) for this activity. The students were given a small (35 records) fragment of a dataset of films from a public datasets repository (https://perso.telecom-paristech.fr/eagan/class/igr204/datasets) that was simplified to include only three variables:

1. Film length in minutes
2. Year released
3. Popularity, as a binary: popular/unpopular. In the original dataset there was a numeric score between 0 and 100; it was not specified how it was determined. In order to apply the decision tree procedure, we introduced a cut-off point, arbitrarily set to 50, to use a binary classification.

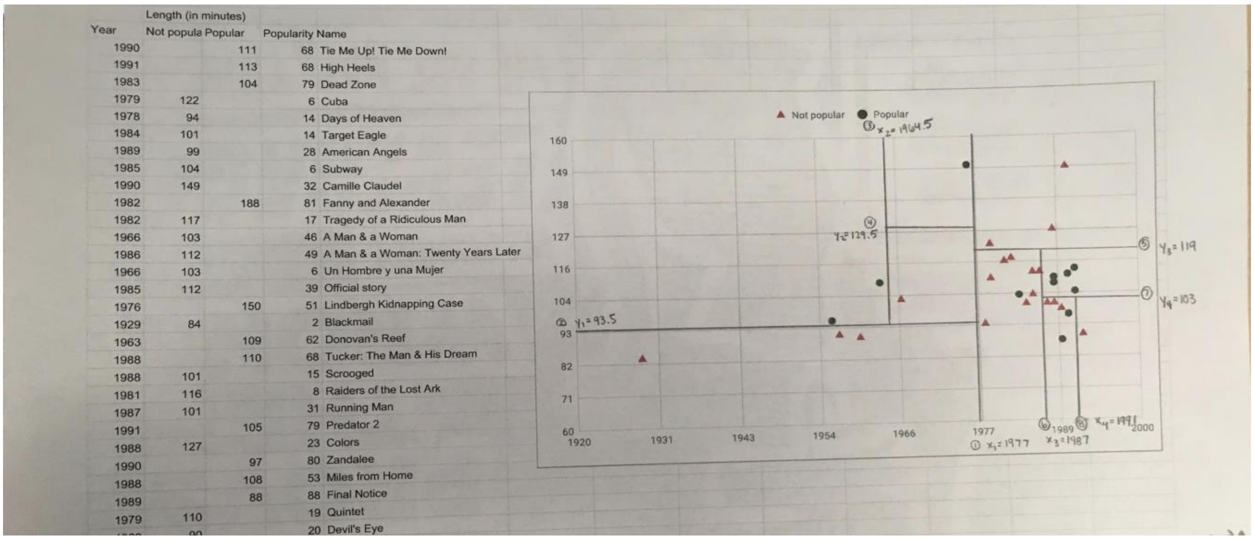
The table below shows three data points. It gives the popularity score, as well as the binary “popular/unpopular” classification.

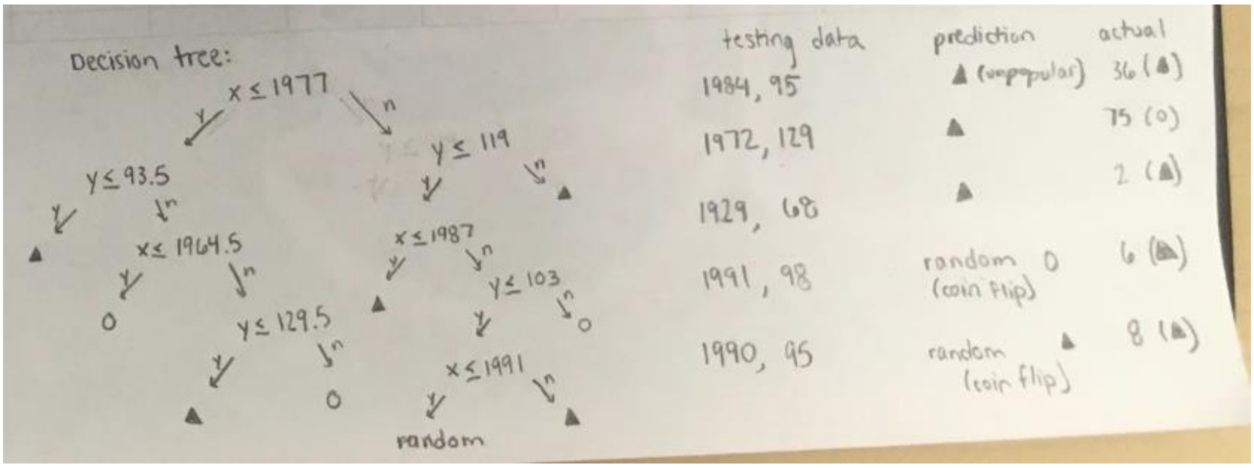
Year	Length (min)	Popularity	Name
1990	111	Popular (68)	Tie me up! Tie me down!
1983	104	Popular (79)	Dead zone
1979	122	Unpopular (6)	Cuba

The names of the films were given just to be able to identify data points. All other fields in the original dataset were omitted.

The decision tree algorithm partitions the problem space (in this case, a two-dimensional space of length and year) into rectangles, each rectangle containing only one class of points (all popular or all unpopular). If it cannot arrive at a perfect separation within a certain number of steps, it produces rectangles in which most points belong to the same class. These rectangles are then used to predict whether a new point, given its coordinates in the problem space (in this case its length and year) would be popular or unpopular, based on the rectangle in which it is positioned.

The students were asked to create the rectangular partitions based on the given data and then try to predict the popularity of five films that were not a part of the training data – the data they used to create the partitions. The five test films were given to them only after they settled on the partitions. While the decision tree algorithm is deterministic if computations for optimal partitions are performed strictly, the students were working on paper and thus were given some leeway in approximation and guessing, so the results produced by different groups were slightly different. Below is one of the four group’s submissions, one diagram showing the partitions, and the other one the corresponding decision 'tree' procedure:





Not surprisingly - given the small number of training data points and the omission of several fields in the original dataset, not to mention the impossibility to perfectly predict a film popularity based on any set of parameters – the classification of the new points was only slightly better than random (two groups correctly predicted three out of five, and the other two groups four out of five). The point of this exercise, however, was not to make a perfect prediction, but to expose students to the process and then discuss takeaways. The most important takeaway was that the algorithm performed better than random guessing but did not actually 'learn' anything about the nature of film. The other takeaway was that the process of creating the partitions was fairly simple mathematically, although quite tedious, especially if it were to be done strictly, instead of approximating. This allowed students to glean into how ML just extrapolates from statistical properties of the given data, and can detect correlations of data parameters to the output it aims to predict, but does not 'know' or 'understand' anything about the actual nature of the data it is given.

Brainstorming a fair hiring algorithm

Another important classroom activity was conducted close to the end of the semester, with the goal of giving students an opportunity to synthesize their knowledge of multiple aspects of the material. Students worked in groups of three, and were prompted to suggest ideas for a fair algorithm that selects a group of candidates to be interviewed by a company, given the candidates' CVs. Each group brainstormed and proposed ideas, including what kind of data they would use to train the algorithm. Other groups then were given a few minutes to ask questions about the proposals and provide feedback. Then, the class as a whole was asked to combine and summarize promising ideas. Some of the interesting takeaways from this exercise were as follows:

1. Many students demonstrated distrust of automated selection of promising candidates altogether and suggested algorithms for elimination of unqualified candidates, and then evaluating the resulting pool manually. While this demonstrates that the students understood the concerns with trusting ML to perform an objective selection, this approach is not realistic, thus we tried to steer the students away from proposals that don't significantly utilize automation. A milder version of this approach was to disregard a large number of parameters in a CV that are strongly correlated with economic and social status, gender, or race, such as the name of the school candidates graduated from.
2. Students proposed various ways of detecting and reducing current biases in the hiring practices which may be contained in the training data. Specifically, they proposed stratified criteria for different groups of employees (to counteract the fact that certain groups, such as women, tend to use more modest phrasing for their accomplishments, or that certain underprivileged groups are less likely to have graduated from 'top' schools). Other ideas included randomizing some fields of the data (for instance, randomly switching genders and gender-correlated parameters), to make sure that the algorithm does not develop a gender bias. Another idea along the same lines was to add to the training data made-up data points of strong candidates from underrepresented groups (or, if such candidates are already present, but only in small numbers, increase their percentages). These proposals show

that students have a good understanding of how training works and how biases in training data lead to a biased algorithm and are comfortable speculating about the effects of using different (including artificially adjusted) data sets.

3. Another important point that was brought up by several students was the need for transparency. Open-sourcing algorithms and making anonymized training data public were seen as essential for creating a fair algorithm since it allows others to scrutinize the methods and point out biases and imperfections when they exist.

Final capstone papers and presentations

For the final capstone papers and presentations students were asked to identify an instance wherein ML is applied to address a big-data problem in society, find a recent (within the last 6 months) news story addressing the instance, do additional research via related news stories and other available sources, and then summarize their findings, specifically focusing on potential consequences of the use of the technology. The presented work was of a very high quality, showing that all students satisfied the course learning outcomes.

The students chose a wide range of topics, including applications of ML to such varied areas as Child Protection Agencies work, legal profession, combating opioid crisis, and automated driving. Several students addressed various issues in genetic testing and health insurance. Some students focused on social media, addressing topics of content regulation (such as restricting inappropriate content), 'deepfake' videos, and social media ranking and advertising.

All of the students were successfully able to present the ideas-level overview of the algorithms they were describing and the data that the algorithms were trained on (to the degree to which it was possible to get this information from their sources). This was especially impressive for students whose majors were not in mathematical fields. For instance, an Elementary Education major was able to describe the mechanisms by which one person's speech is combined with another person's image in 'deepfakes', and a biology major described input-gathering and algorithms used in self-driving cars.

We were also very happy to see that students with more background in computer science and mathematics didn't just try to be on par with their classmates from other majors, but used their expertise to go more in-depth into the material. In particular, a CS/Mathematics major chose a topic of automated bias detection and removal in algorithms, using, among other sources, a recent CS journal publication. This student's final paper also used psychology resources on unconscious bias training, showing that this student also went beyond their major subjects.

Another important observation is that several students choose topics that they were personally connected to, such as choosing 'deepfake' videos because of concerns about online bullying and revenge or choosing issues with DNA testing because of medical situations of family or friends.

Overall, the final capstone papers a good grasp on all aspects of the material, including a solid understanding of the core principles of ML, the role of training data, and potential implications of using the algorithm in a social context. We were also very pleased to see the students' high level of interest in, and engagement with, the material. We are confident that the students will bring this interest and understanding into their future occupations, whatever those might be.

Future work

Herein we described the first offering of a new course at our institution. Leveraging the collaborative, simple structure of a small, selective, American liberal-arts college and the interdisciplinary honours programme, we did not have to create an elaborate argument for designing a learning experience that combines the skills of a media and communication specialist and a computer scientist. We hope that this experimental course design can help others develop similar courses at other larger colleges, with more separated faculty, where collaborations may be more difficult to form, and interdisciplinary classes are difficult to place within a class schedule. We believe the effort is worth attempting, for during the teaching of the course we learnt many things about not only the topics assigned and about our students, but also about ways we could both expand our teaching and research agendas.

Based on student and administrative responses, we are eager to revisit and rerun (and improve) the course, perhaps expanding it to a larger interdisciplinary audience if/when logistics allow. In the next offering, we will again keep a weekly schedule, and will again keep the cycle of assigning pre-reading and student-postings of study questions, and then meeting together to consider and plan the lecture and discussion. This teaching preparation time is a necessary element to keep us connected to the specifics of the student needs and helped us greatly with planning who would be 'lead instructor' for the lesson. In the future offering of the course we may also include more automated ML and data visualization tools, to allow students the opportunity to experiment with some ML approaches more directly.

Several of the readings selected and described in this paper were great for discussions and we hope to use them again, but in a rapidly changing field such as Machine Learning we anticipate that there will be several new examples and concerns which we will likely need to address. We recognize the range and the variety of new papers, such as those presented at the 'Emerging technologies….' conference, can serve as an excellent source for case studies for future offerings of the course.

The activities which captured student interest, such as building decision trees and thinking about CVs sorting algorithm decisions provided an excellent learning opportunity for students and should be included in future course offerings. Students' course feedback indicated desire for more activities, so we may try to expand the range of those. The students also enjoyed writing the capstone final papers, which has not always been the instructors' experience. We strongly believe that students of all academic majors should write papers that integrate research with new inquiry and are extremely pleased with the willingness of the students to give the papers in class, and to share the final works with each other (and future students/ readers). Examples of student papers can be found at Machkasova's website: http://cda.morris.umn.edu/~elenam/3255Spring2019/index.html).

Conclusions

The course description stated: The course reviews and questions public attitudes about computing and new technologies. It challenges the view that algorithms are fair and unbiased. It discusses key Machine Learning mechanisms and the role of training data in determining algorithms' behaviour. It exposes how data rooted in current social practices may introduce biases into Machine Learning, which perpetuates these biases in society.

In evaluating the ways we met our institution's student learning outcomes which we selected for the course: Engagement with big questions, both contemporary and enduring; Critical thinking and problem-solving; Information and technology literacy; and Ethical reasoning and action, we found the assignments and discussion notes gave us several instances/ data points to make the claim we had met all of these goals. Additionally, in a review of the ways we met the Honors programme student learning outcome: demonstrating interdisciplinary thinking in scholarly and/or creative ways, students' final capstone papers and presentations indicated that every student (regardless of their major) was able to address both technical ideas behind machine learning applications and representations of Machine Learning applications in the media. Students were able to point out benefits and concerns in regard to specific uses of Machine Learning. It was great to see how students' topics for papers and presentations in several cases extended much beyond their backgrounds in their majors.

The discussions started in our classes have important social outcomes for our students and will shape the ways they will think about Machine Learning in the future. Anecdotally, several students have already been taking the ideas and using them in other courses, and we expect the integration of ideas of the course will more-importantly aid the students in future experiences in reading, listening, discussing, and critically thinking about technology.

References

Agricultural Data Coalition (2019, January 30) How Does Artificial Intelligence Really Work in Agriculture? https://www.precisionag.com/digital-farming/data-management/how-does-artificial-intelligence-really-work-in-agriculture/ .

Aikens, S. G. (1997). American democracy and computer-mediated communication: A case study in Minnesota. Doctoral dissertation in social and political sciences. Cambridge University.

Baym, N. K. (1995). 'The emergence of community in computer-mediated communication', pp. 138–163 In S. G. Jones (ed.), CyberSociety: Computer-mediated communication and community. Newbury Park, CA: Sage.

Castells, M. (1996). The information age: Economy, society and culture (3 volumes). Blackwell, Oxford, 1997, 1998.

Czitrom, D. (1983). Media and the American Mind. Chapel Hill: University of North Carolina Press.

Dahlberg, L. (2001). 'Computer-mediated communication and the public sphere: A critical analysis', Journal of computer-mediated communication, 7:1, JCMC714. https://doi.org/10.1111/j.1083-6101.2001.tb00137.x .

Innis, H. (1951). The Bias of Communication. Toronto: University of Toronto Press.

Kavanaugh, A.L., D.D. Reese., J.M. Carroll, and M.B. Rosson, (2005). 'Weak ties in networked communities', The Information Society, 119-131.

Mackenzie, A. (2006). Cutting Code: Software and Sociality. (Vol. 30: Digital Formations). New York: Peter Lang Publishers.

McLuhan, M. (1964). Understanding media. New York: McGraw-Hill.

Mumford, L. (1970). Technics and human development: the myth of the machine, vol. I (pp. 381-410). New York: Harcourt, Brace & World.

Neuman, W.R. (Ed.). (2019). Media, Technology, and Society: Theories of Media Evolution. The University of Michigan Press, http://www.press.umich.edu/titleDetailDesc.do?id=293114 2010

O'Neil, C. (2016). Weapons of math destruction: How big data increases inequality and threatens democracy. NY: Crown Publications.

Postman, N. (1992). Technopoly: The Surrender of Culture to Technology. New York: Knopf.

Shalev-Shwartz, S. and S. Ben-David (2014). Understanding Machine Learning. Cambridge, MA: Cambridge University Press.

Sawyer, R. J. (2012). 'AI and Sci-Fi: My, Oh my!' Keynote Address presented September 29, 2012, at the 10th anniversary celebration for The Institute for Quantum Computing. Waterloo, Ontario, https://www.sfwriter.com/precarn.htm .

Walther, J. B. (1996) ‘Computer-mediated communication: Impersonal, interpersonal, and hyperpersonal interaction’, Communication research, 23(1): 3-43.

Wellman, B., P. Carrington, and A. Hall, (1984). ‘Networks as personal communities’ in Berkowitz, S.D. and B. Wellman, (eds.), Structural Sociology. Cambridge, MA: Cambridge University Press.