CRITICAL MACHINE LEARNING: AN ONLINE RESOURCE

Achim Koh
Interactive Technology and Pedagogy: Independent Study Project
December 7, 2017

Abstract

Critical Machine Learning (CML) is an online resource that targets researchers and educators of various backgrounds and attempts to bridge technical understanding and critical discussion surrounding machine learning. CML attempts to encourage interdisciplinary thinking that researchers and educators can apply in their respective disciplines. The non-technical user can use CML to familiarize oneself with fundamental machine learning terms and avoid misconceptions when applying them in one's respective field; users already familiar with the technology can use the resource to explore diverse perspectives about the field, especially those that come from outside the mainstream industry. The website contains lists of selected material that covers topics such as introductory concepts of machine learning, the black box metaphor, and the potential bias inherent in both data and process; it also serves as a syllabus for pedagogical practices such as workshops. CML aims at providing a middle ground where the user can find resources for both technical understanding of and critical thinking about machine learning.

Introduction

Machine learning is a central part in the recent development of artificial intelligence technology, a field under much attention regarding its potential to cause or accelerate social changes. The IT industry has been quick to apply machine learning to a vast range of products, and other social areas are increasingly participating in the trend. Machine learning algorithms are becoming important elements of the computational processes operating in the contemporary society such as finance, journalism, advertising, medicine, education, criminal justice, labor, translation, robotics, and much more. Increases in both computational power and the amount of available data across social fields opened the door for notable developments in digital technology. One of them is the rise of big data; another is a large improvement in the automated application of statistical methods in order to detect patterns, predict future data and make decisions—in other words, in the field of machine learning.

Along with the prospect of efficiency through automation, well described by Gideon Lewis-Kraus' reporting on Google Translate in "The Great A.I. Awakening," this technological change also raises sociopolitical concerns.[1] The increasing influence of machine learning methods as algorithmic social mechanism calls for discussions about the social implications, not only by technology experts but also by members of other fields which are affected by the technological change. Such broad discussions call for both accessible technical knowledge and critical thinking.

However, the technical understanding of machine learning can be difficult, especially to people outside the computer science academia and industry. Computer science, like many other fields, operates on a relatively exclusive expertise, resulting in a type of illiteracy for outsiders

1. Gideon Lewis-Kraus, "The Great A.I. Awakening," *The New York Times*, December 14, 2016, sec. Magazine.

lacking the mathematical and computer-scientific background. Moreover, in a real-life application, the complexity of not only the data and code involved but also the machine learning models themselves can reach a level where it is simply not feasible to provide a clear explanation. The fast-changing nature of the field only adds to the difficulties of a machine learning literacy.[2]

Nevertheless, one can benefit from the relative openness of machine learning research, much of which is based on publicly available datasets and published on open platforms such as *arxiv*. There are also considerable amounts of machine-learning related educational resources that keep increasing, from short toolkit tutorials to graduate-level MOOCs. This openness, while being a good thing, is perhaps best appreciated with a grain of salt. Debuting a series of instructional videos about machine learning, Facebook's Director of AI Research does not bother hiding their hope to cultivate a young generation of engineers which will feed the industry's need.[3] Similarly, a large part of the cutting-edge research and publicly available tools is driven by industrial players or with their support—which invites a critique on the direction of technological developments. Nevertheless, openly available technical resources are valuable to non-experts, especially since the increasingly high levels of abstraction involved in machine learning methods make it harder to understand the mechanisms of the technology without directly engaging it.

---

2. There are several factors contributing to a "Machine learning information overload:" the field is growing both academically and industrially; the increasing trend of open-sourcing studies through the use of platforms such as arxiv, which adds speed to the dissemination of research; increasing media attention; social or personalized feed algorithms, which ingests more news as one explores the subject; etc. It is not coincidental that an aggregator for arxiv researchers is named *Arxiv Sanity Preserver* (arxiv-sanity.com).

3. Yann Lecun and Joaquin Quiñonero-Candela, "Artificial Intelligence, Revealed," Facebook Code, December 1, 2016, https://code.facebook.com/posts/384869298519962/artificial-intelligence-revealed/.

Conversely, scholars in social sciences and other fields have been putting efforts to address issues that concern the social implications of machine learning technologies. These efforts include interdisciplinary discussions that question biases in the algorithmic society, envision the future and contextualize the technological progress. The *Data & Society*[4] and *AI Now*[5] research institutes are just some examples of these critical efforts that tend to be open and inclusive with regards to their research. However, these resources—technical and critical—still tend to exist somewhat separately, which makes it difficult to navigate between the two types.

The proposed project, *Critical Machine Learning* (CML), attempts to bridge technical understanding and critical discussion; by doing so, it aims at lowering the barrier for researchers and educators to understand technical concepts and consider critical issues around machine learning. CML's goal is to become a resource that encourages interdisciplinary thinking, and that researchers and educators can apply in their respective disciplines. The project operates under the assumption that hands-on technical experience and critical thinking about technology complement each other, and that such an approach is applicable in multiple fields of research and education.

The title, *Critical Machine Learning*, refers to an attempt at a critical understanding of the implications of machine learning technology, which includes both observations of its current direction and projections of where we should take it. In this sense the project attempts to support an informed critique of technology that requires some strategic understanding of the subject in order to be effective, as pointed out by danah boyd in "When Good Intentions Backfire."[6]

---

4. http://datasociety.net/.

5. https://ainowinstitute.org/.

6. danah boyd, "When Good Intentions Backfire," Data & Society: Points, February 15, 2017, https://points.datasociety.net/when-good-intentions-backfire-786fb0dead03.

Moreover, CML also aims to widen the middle ground between practical knowledge and humanities- and social science-based critique. It can also refer to speculative works that challenge assumptions about computational processes operating within contemporary societies, and imagine what their use should be; in this sense it is also loosely inspired by Critical Design, a term popularized by designers Dunne & Raby and referring to a critical theory-based approach that makes use of speculative and fictional narratives that question the status quo.[7]

Points of interest

Machine learning, as any technology that is undergoing fast technical developments and which usage is explosively increasing, exists in relation to the many social contexts it is embedded in. Cathy O'Neil's *Weapons of Math Destruction* provides a glimpse into some of these contexts, from both the side of people who own and deploy the tools and the side of those who are subject to them. O'Neil describes current applications of machine learning-based models in the many aspects of contemporary life including finance, education, advertising, criminal justice, labor, credit loans, insurance and civic life, and refers to especially harmful ones as Weapons of Math Destruction. (WMDs) These harmful applications tend to be opaque in their working, to operate in a large scale and to work against the interest of those subject to the models. Moreover, the algorithms powering these applications are transferable, i.e. tools developed for one purpose can serve another—a characteristic pointing to the current expanding usage of machine learning-based models.[8]

7. Anthony Dunne, *Hertzian Tales: Electronic Products, Aesthetic Experience, and Critical Design* (Cambridge, Mass: The MIT Press, 2008).

8. Cathy O'Neil, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy* (New York: Crown, 2016).

An important portion of this project addresses the notion of opacity, as the project aims to provide a ground knowledge for non-experts in the field. As addressed below, some aspects of machine learning's opacity are native to the technology and its current state, such as the scale in which it is being used, while some are much more social and based on human decisions. CML also interests itself in the current and potential damages that machine learning can cause as an increasingly influential technology. Such damages are closely tied to the value systems and ethics embedded in the design and applications of algorithms; the technology does not exist in a vacuum, and therefore invites interdisciplinary discussion.

If machine learning-related technology is increasingly being used in ways that affect people's lives—for example to make decisions on whether someone should be eligible for a loan, be granted a bail, or be diagnosed of a certain disease—the social implementations will benefit from a transparent knowledge of the technology. Fittingly, a recurring concern in the critical discourse around machine learning points to its opacity, its being a 'black box': you feed some data to the machine, it outputs some result, but we do not really know what happens inside. One thing to note here is that a black box is not by itself problematic. As Alex Galloway points out in "The Computational Decision," black boxes are often the result of a trade-off between a technical explanation of a certain process and the convenience and efficiency of being able to utilize said process without explicitly knowing its mechanism.[9] The different social contexts within which a technology operates must be considered in order to critique or problematize its trait as black box.

The black box metaphor can apply to different types of phenomena, such as undisclosed governmental administrative procedures or electronic devices. The observer does not have access

9. Alexander R. Galloway, "The Computational Decision," December 13, 2015, http://cultureandcommunication.org/galloway/the-computational-decision.

to the inner workings of the system, or it is not feasible to have enough knowledge, time and/or resources to understand them. In the case of machine learning, there exists another layer of opacity which adds complexity to our current concern. To set up a perspective, here I will borrow from Jenna Burrell's article "How the machine 'thinks'" which provides a useful framework to articulate what opacity is. Burrell distinguishes three forms of opacity with regards to algorithms in general, and especially machine learning algorithms:

- intentional concealment by (mostly) companies and governments;
- technical illiteracy; and
- complexity of machine learning models themselves, in addition to that of data and code.[10]

Opacity as secrecy

One reason that algorithms can be opaque is because their owners, often corporations, are willingly keeping it a secret against adversaries. Such a strategy could serve to protect intellectual properties, evidently, but also to maintain a certain level of product quality. For example, search engines deliberately do not reveal their exact ranking mechanisms for search results; otherwise it would be too easy for the adversary (in this case, a content provider) to 'game' the system to draw more attention to their web page.

Such a secrecy could also be a means of concealing illegal or unethical activities. For example, a company might write in their code some tricks to manipulate consumers and unjustifiably discriminate against them, and have such activities undiscovered under the protection of intellectual property laws.

10. Jenna Burrell, "How the Machine 'Thinks': Understanding Opacity in Machine Learning Algorithms," *Big Data & Society* 3, no. 1 (January 5, 2016): 2053951715622512, https://doi.org/10.1177/2053951715622512.

Opacity as illiteracy

Coding and designing algorithms is a specialized skill. Open-sourced code can still be very hard to understand, especially to a non-programmer. Somewhat like how a legal document full of legal jargon would be impenetrable by non-experts, the difficulty of understanding the computer science adds another layer of opacity.

Opacity caused by the scale and complexity at which we use algorithms

Some algorithms are hard to understand, even when we have access and the expertise. Look at Google search's page source and you will be dazzled by the amount of code that runs the seemingly simple white page with not much else than a logo and a search box. And this is only the front end; the code running on Google's servers, that actually does the information retrieval, is a whole body of complexity on its own. It is also not a single program but a combination of multiple components, too, which would make understanding the code even harder.

But the problem does not end there. If such complexity can be achieved through traditional, explicit rule-based programming, the statistical approaches of machine learning models take the problem a few steps further. Machine learning algorithms learn on training data and in the process modify their internal logic. This process can be automated and run in a large scale very quickly, which makes machine learning algorithms very effective in tasks near impossible to do by hand like classification, i.e. relying on mathematical functions that take into account many variables to categorize unknown data points.

The process of a machine learning algorithm making its decision based on some data is perhaps not too difficult to understand as a whole. But its internal specifics, such as how a certain variable (also referred to as feature or dimension, disregarding the technical differences for now) influences the result, is not necessarily friendly to human comprehension. Human explanation tends to rely on semantics; we assign meanings to why and how some things happen. But the mathematics of optimization that underlie machine learning does not really care for meanings—what it wants is to get to the right result. The neural network that recognizes dog faces does not look for their eyes or noses, at least not in the same way a human does.

The problem of dimensionality adds to the difficulty of understanding machine learning algorithms. Because there are so many variables that an algorithm takes into account, it is both inefficient to calculate and hard to understand the specific inner mechanics. Also, the relation between a variable and the decision made might not be linear. Some variables may seem to have no influence when looked at separately, but in fact have an influence in a combined manner. When we train machine learning algorithms, data points (e.g. emails) are represented using a number of variables (e.g. words in the email dataset). And we can calculate what variables have the biggest influence on the decision (e.g. is this email spam or not?). But this quantified influence may not tell you much (or could even be misleading) in terms of understanding what the algorithm is doing with its thousands of features as an aggregate.

Against opacity

The different types of opacity call for different, or mixed modes of engagement. Ways to deal with abusive intentional secrecy may include auditing the source code, possibly via regulations; auditors independent from conflicts of interests could provide a way to protect intellectual properties in spite of the audit. In "Auditing Algorithms," Sandvig et al. also lay down audit designs that would not require access to the code, for example by scraping the platform, employing dummy user accounts or hiring some users to participate in the research. Such efforts should be concurrent with an education for computational literacy, and efforts from the field to write 'good' and legible code.[11]

Machine learning also has a sort of "fundamental opacity," that cannot be easily overcome by just having access to source data and code.[12] Burrell notes some approaches that the designers of the machine learning models could take: not using machine learning in specific domains is actually one of them, as well as techniques to simplify the models such as feature reduction.[13] In addition to these approaches that aim at making it easier to understand why the model is doing certain things, she also notes other approaches that focus more on evaluating what the model does from an outside point of view. To look at whether a classifier has discriminatory effects, one could feed a wide range of different data and see how it performs, without having to explicitly interpret the inner logics in a human-comprehensible way. Different methods of algorithmic auditing are being researched, such as the FairML toolkit included in the CML collection.[14]

Project Outline

CML consists of two components. The first is a set of online resources that include curated material for both technical understanding of and critical thinking about machine learning.

---

11. Christian Sandvig et al., "Auditing Algorithms: Research Methods for Detecting Discrimination on Internet Platforms," *Data and Discrimination: Converting Critical Concerns into Productive Inquiry*, 2014.

12. Burrell, "How the machine 'thinks.'"

13. Ibid.

14. https://github.com/adebayoj/fairml.

The resources will cover various machine learning methods and computational concepts, along with diverse social issues impacted by the technology. The second component is an offline workshop that puts the research material into practice. While the online resource covers different research areas and computational methods, the workshop focuses on a more specific sub-topic—in this case, an introduction to basic concepts around machine learning. Since the workshop is aimed more at a general understanding of machine learning and discussions around it and less at highly technical skills, participants from all disciplines and technical backgrounds are welcome.

CML was first drafted as an independent studies proposal for the Interactive Technology and Pedagogy Core II course at the Graduate Center, CUNY in Spring 2016. I conducted preliminary research throughout the remainder of 2016. I conducted two related events, a presentation at CUNY IT Conference 2016 (Dec 2016) and a workshop at NYCDH Week 2017 (Feb 2017), where I presented the project as well as basic concepts of machine learning. The online resource was built using *Github* and *Are.na*, and made public on criticalml.net in April 2017.

The main repository of the resources is hosted on Github, which offers free service and can be expected to reliably continue to do so in the near and mid-term future considering its popularity. A near-term stability should suffice for the purpose of this project that addresses an up-and-coming development, partially in a technological literacy perspective. On the other hand, Are.na is a relatively new service and its prospect as a long-lasting platform is yet to be tested. But fortunately the platform offers a simple export functionality which allows the user to save the content of a specific channel in diverse formats such as JSON, PDF or HTML. The exported collections can then be backed up within the Github repository.

The presentation slides from the NYCDH workshop are included in the online repository as well. This workshop was designed as an introduction to machine learning's definition, its key concepts and applications, critical issues in data collection, as well as other ML-related resources. In short, it was a condensed version of the online project. In hindsight, some of my presentation suffered from rather a shallow and vague explanation, in an attempt to fit all material in a two-hour time slot. The workshop lasted less than one and a half hour, so future iterations of this introductory session will benefit from a more thorough overview.

The registration form specified no prerequisite skill or knowledge, which was consistent to the project's aim, but resulted in a relatively varied audience in terms of familiarity to the subject. The first part of the workshop, which was dedicated to key definitions and concepts, seems to have been too vague for some participants who were not familiar at all with the topic, while being too mundane for more experienced ones. This was also the case in the previous presentation at the CUNY IT Conference. The project's target being participants from diverse backgrounds, a way towards improvement here would be to announce more specifically the technical level of the workshop and, again, to provide more thorough overviews of basic concepts.

The Online Resource

The *Critical Machine Learning* online resource is located at http://criticalml.net/. It includes Are.na channels that are sub-collections of selected material under a specific topic, as well as a Github repository of markdown files that further explore the topics and point to more material. The Github repository also comprises more generic material helpful for hands-on technical learning of machine learning techniques. A technical overview of Are.na and Github are provided in the Appendix, along with the rationale behind choosing these services as platform. The Github repository, located at https://github.com/achimkoh/critical-ml-resources, serves as source for the website; the static website generator *Jekyll* is used to turn the markdown files into webpages, including links to Are.na channels. Below are descriptions of the Are.na sub-collections.

The Primers[15]

A number of introductory materials that explain the definition and applications of machine learning are provided as primers. These were chosen under two criteria: 1) the material should be produced in a way that assumes no prior in-depth technical knowledge, nor the intention to learn it in the future, from the readers; and 2) each material, or some materials combined, should provide an overview on the social implications that this technological trend represents.

In this respect, the primers can be grouped into three main sub-categories. "AI Literacy: The Basics of Machine Learning" and "Artificial Intelligence, Revealed" are short introductions to the potentially confusing terminology such as machine learning, artificial intelligence, neural networks and deep learning. "What is Code?" and "The End of Code" are meant to provide context on the role of software code in the technology industry, and contemporary society in general, as well as the difference between the traditional way of programming and machine learning. "The Great A.I. Awakening" and "A.I. Versus M.D." give closer looks on the history and recent rise of machine learning, respectively focusing on translation and medicine.

15. https://www.are.na/critical-machine-learning/primers.

Black Box[16]

The items in this collection are centered around the notion of opacity and why it matters. The aforementioned *Weapons of Math Destruction* and “How the machine ‘thinks’” are the centerpieces that lay the ground for discussion. These are accompanied with complementary readings, as well as technical resources that address the problem of how to look inside the black box. In her article, Burrell points to some strategies against the black box, including algorithmic audits, “adversarial” machine learning techniques, education and open sourcing.[17] FairML is one example of algorithmic audit; it is a Python toolkit that allows one to test predictive models such as the recidivism model addressed by Propublica’s now-famous “Machine Bias” series.[18]

Future Work

Things move fast in the field of machine learning; the technical research and application, as well as the critical discourse about their impact, have quickly developed since this project was first conceived. Luckily, efforts to mediate the space in-between are also more and more present. Therefore, while it is fair to say that this project will probably lose much of its relevance within a few years, it attempts to serve as a temporary meta-resource amidst a fast-paced environment. Within this short- to mid-term lifespan of the project, CML’s online resource will be maintained and updated in order to reflect the aforementioned developments.

Another way I hope the project will contribute to interdisciplinary thinking on machine learning is through additional workshops, either introductory or focused on specific sub-topics.

16. https://www.are.na/critical-machine-learning/black-boxes.

17. Burrell, “How the machine ‘thinks’”.

18. Julia Angwin et al., “Machine Bias,” *ProPublica*, May 23, 2016, https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.

These workshops could be held through the Graduate Center, other educational institutions, or research networks such as NYCDH. Based on the material for previously held workshops, I expect to write up a two-hour course plan for introducing a few main concepts and issues by February 2018. This will help me and others to offer similar workshops more easily.

An additional workshop session could further explore the notion of opacity with regards to machine learning. This will be well complemented by a hands-on tutorial of existing machine learning tools, in which one can get an overview of the practitioner's side. The hands-on practice will help participants understand the nuances of opacity in machine learning. This workshop could lead into another sub-topic, the concept of fairness; one way to approach this concept is to take existing research, such as Angwin et al.'s investigative report on racial algorithmic biases in criminal justice, and to examine and/or replicate the inner workings of the algorithms in question.

The hands-on section of the workshops will be designed to use popular and open-sourced machine learning toolkits. One approach is to use the Python programming language, with which toolkits such as Scikit-learn and Tensorflow are available. Another option is to use JavaScript, using libraries such as deeplearn.js. The advantage of using Python packages is the larger and older userbase, which is connected to a better documentation and more resources available. On the other hand, JavaScript libraries are easier to run on the web browser, which can be more familiar to the programming novice.

The workshop can also rely on cloud-based environments such as the *DH Box* platform that provides open source tools such as a command line shell and iPython in a virtual Linux environment accessible via a web browser.[19] While participants will still preferably bring their own laptop, using pre-configured tools like DH Box has advantages over installing tools directly

19. http://dhbox.org.

on students' devices or cloud computing machine instances; it eliminates most of the configuration process, which can be time-consuming and distracting away from the content. Assuming a stable network connection, the ease of preparation is a great convenience. Also, having students bring their own devices allows the workshop to be held in a setting other than computer labs; this will be beneficial in conducting a discussion.

Machine learning is an important component of the fast-changing technological landscape. By covering its basic concepts as well as social implications, this online resource attempts to be of service to technological literacy which will in turn benefit the social discourse in general. In addition, CML will also hopefully be useful in promoting critical thinking among machine learning practitioners. The wide application of and different types of opacity inherent in machine learning provide favorable conditions for intended or unintended biases that have real consequences. As one engages the technology in one's respective field, it is important to consider how the technology is situated within diverse social contexts. Moreover, one will be best served by remembering that technology and social contexts are in flux and critical thinking is an ever-ongoing work.

Appendix

Github is a web-based version control repository hosting service that uses the Git system. The largest service of its kind, it is mostly used for software code.[20] However, its use cases also extend to writing and displaying text as well as curating and maintaining lists.[21] The possibility to quickly update contents in a repository and keep track of changes, as well as its web interface that favors the Markdown text format, makes Github a useful tool for managing short text and lists that are evolving.

For example, one can find hundreds of lists about specific subjects, from programming languages such as JavaScript and Python to productivity methods and civic activism, all entitled "*Awesome [topic]*." These curated lists of tools and resources relevant to the given subject often contain not much more than hyperlinks and short descriptions, but are extremely popular; the *Awesome* repository, which is a meta-list of said *Awesome* lists, contains almost 400 items and is the eighth-most starred repository on Github as of this writing.[22] When considering the numerous other lists also curated on the platform but not titled or tagged with "Awesome," it is safe to say that list curation is an important use case of Github.

Therefore, hosting the CML resource list on the platform, a popular choice for hosting computer science and software-related lists, was a reasonable choice. Moreover, this project targets not only researchers in non-technical fields but also technologists. In such spirit, Github—itself used mainly by technologists but increasingly by practitioners in all kinds of

---

20. Georgios Gousios et al., "Lean GHTorrent: GitHub Data on Demand," in *Proceedings of the 11th Working Conference on Mining Software Repositories*, MSR 2014 (New York, NY, USA: ACM, 2014), 384–387, https://doi.org/10.1145/2597073.2597126.

21. https://github.com/showcases/writing.

22. https://github.com/sindresorhus/awesome.

disciplines—appeared to be an appropriate middle ground where technical resources and critical readings could meet.

However, while Github was great for a full list, I also needed to have sub-collections of list items under a specific topic; creating separate files on Github would be one solution, but I felt that one of the good qualities of these lists was that they offer an overview at a glance by having everything in one page, and having separate files would strip this trait away. Accordingly, I looked for an additional tool that would allow an easier navigation of specific sub-topics, and chose to use Are.na.

Are.na is a social bookmarking platform with some traits of content management systems, notably the ability to upload files and to post the user's own content. Each user can create "channels," collections in which they can add items such as links and files that are referred to as individual "blocks." Channels are either open (everyone can view and add blocks), closed (everyone can view but only permitted users can add blocks), or private (only permitted users can view and add blocks). A block must be added in a specific channel, and can be connected to other channels afterwards. A feature of Are.na's channel-block units, which differentiates the platform from typical tag-link or category-article systems, is that any channel can itself be added to another channel as a block; also, this does not involve a hierarchy among channels, in the likes of a parent channel and corresponding sub-channels. While seemingly a small difference, this taxonomy facilitates a fluid and modular organization of items; not only individual items but also collections of items can be re-categorized in another collection, without interfering with the existing categorization.

The ease of re-organization, along with the minimal and responsive (in the sense that it reacts quickly in the browser) web interface provided an advantage in my case of trying to organize resources while simultaneously learning about the subject. Such a characteristic appears explicitly intended; the website describes its goal as becoming "a platform that helps people continually recontextualize information into new ideas and help us all understand the vast amount [of] information we face on a daily basis."[23] Moreover, as the quote points out, the amount of information was non-trivial for this project—not only because any research typically deals with lots of information, but also because of the aforementioned information overload pertinent to machine learning.

23 "About," Are.na, https://www.are.na/about. Accessed March 27, 2017.

Bibliography

"About." Are.na, March 27, 2017. https://web.archive.org/web/20170327191952/https://www.are.na/about.

Angwin, Julia, Jeff Larson, Surya Mattu, and Lauren Kirchner. "Machine Bias." *ProPublica*, May 23, 2016. https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.

boyd, danah. "When Good Intentions Backfire." *Data & Society: Points*, February 15, 2017. https://points.datasociety.net/when-good-intentions-backfire-786fb0dead03.

Burrell, Jenna. "How the Machine 'Thinks': Understanding Opacity in Machine Learning Algorithms." *Big Data & Society* 3, no. 1 (January 5, 2016): 2053951715622512. https://doi.org/10.1177/2053951715622512.

Daly, Liza. "AI Literacy: The Basics of Machine Learning." World Writable, April 11, 2017. https://worldwritable.com/ai-literacy-the-basics-of-machine-learning-2e20f93e34b4.

Dunne, Anthony. *Hertzian Tales: Electronic Products, Aesthetic Experience, and Critical Design*. Cambridge, Mass: The MIT Press, 2008.

Ford, Paul. "What Is Code? If You Don't Know, You Need to Read This." *Bloomberg.Com*. Accessed November 11, 2017. http://www.bloomberg.com/graphics/2015-paul-ford-what-is-code/.

Galloway, Alexander R. "The Computational Decision," December 13, 2015. http://cultureandcommunication.org/galloway/the-computational-decision.

Gousios, Georgios, Bogdan Vasilescu, Alexander Serebrenik, and Andy Zaidman. "Lean GHTorrent: GitHub Data on Demand." In *Proceedings of the 11th Working Conference on Mining Software Repositories*, 384–387. MSR 2014. New York, NY, USA: ACM, 2014. https://doi.org/10.1145/2597073.2597126.

Lecun, Yann, and Joaquin Quiñonero-Candela. "Artificial Intelligence, Revealed." Facebook Code, December 1, 2016. https://code.facebook.com/posts/384869298519962/artificial-intelligence-revealed/.

Lewis-Kraus, Gideon. "The Great A.I. Awakening." *The New York Times*, December 14, 2016, sec. Magazine. https://www.nytimes.com/2016/12/14/magazine/the-great-ai-awakening.html.

Mukherjee, Siddhartha. "A.I. Versus M.D." *The New Yorker*, April 3, 2017. https://www.newyorker.com/magazine/2017/04/03/ai-versus-md.

O'Neil, Cathy. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. New York: Crown, 2016.

Peter Stone, Rodney Brooks, Erik Brynjolfsson, Ryan Calo, Oren Etzioni, Greg Hager, Julia Hirschberg, Shivaram Kalyanakrishnan, Ece Kamar, Sarit Kraus, Kevin Leyton-Brown, David Parkes, William Press, AnnaLee Saxenian, Julie Shah, Milind Tambe, and Astro Teller. "Artificial Intelligence and Life in 2030." *One Hundred Year Study on Artificial Intelligence: Report of the 2015-2016 Study Panel*. Stanford, CA: Stanford University, September 2016. https://ai100.stanford.edu/2016-report.

Sandvig, Christian, Kevin Hamilton, Karrie Karahalios, and Cedric Langbort. "Auditing Algorithms: Research Methods for Detecting Discrimination on Internet Platforms." *Data and Discrimination: Converting Critical Concerns into Productive Inquiry*, 2014.

Tanz, Jason. "Soon We Won't Program Computers. We'll Train Them Like Dogs." WIRED, May 17, 2016. https://www.wired.com/2016/05/the-end-of-code/.