AI IN THE LIBRARY

ANDROMEDA YELTON

SJSU INFO 287-15

Spring 2020

MACHINE LEARNING TECHNIQUES 1: INTRODUCTION TO ML

Overview

What is machine learning? Depending on the sources you read, it might be:

- A technology revolutionizing everything from consumer technology to healthcare to finance;
- A meaningless marketing buzzword;
- The rise of robots who will take our jobs, develop superintelligence, and kill us all.

In this course, you'll learn what machine learning (the most important subset of AI today) really is, beyond the hype. What do ML programmers actually do? What's going on in ML systems, and what does that mean for how we should use (or not use), understand, and question them? How is ML useful — or worrisome — for libraries in particular? You'll practice communicating your understanding so you can take it out into the world and be the machine learning expert in your library. You'll ask lots of questions, and work with your classmates to develop answers.

This week we'll start at the beginning: what is machine learning?

Learning Outcomes

- Understand that machine learning means algorithms that can predict things about data.
- Recognize fundamental machine learning terminology.
- Preview ideas that will recur throughout the course (such as data and interpretability).

This suppports:

- Course Learning Objective (CLO) #1: Understand and explain the basics of AI: both its underlying principles and common machine learning techniques.

Readings

REQUIRED

- Meredith Broussard, *Artificial Unintelligence.* Ch 7.
- http://www.r2d3.us/visual-intro-to-machine-learning-part-1/
- http://www.r2d3.us/visual-intro-to-machine-learning-part-2/

OPTIONAL

- Jason Tanz. “Soon We Won't Program Computers. We'll Train Them Like Dogs.” https://www.wired.com/2016/05/the-end-of-code/
  - A historical and philosophical discussion
- Kaggle Python Tutorial on Machine Learning. https://www.datacamp.com/community/open-courses/kaggle-python-tutorial-on-machine-learning.
  - The datacamp tutorial discussed by Broussard
- John McCarthy, Marvin Minsky, Nathaniel Rochester, & Claude Shannon. “A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence.” http://jmc.stanford.edu/articles/dartmouth/dartmouth.pdf
  - The “AI summer” is a foundational moment in computing history; comparing their agenda for the summer of 1956 to where we’re at with AI today is illuminating. Pages 1-4 are the most important part.
- Andriy Burkov. The Hundred-Page Machine Learning Book. http://themlbook.com/
  - An action-packed but clear reference guide to the major math and programming techniques in ML.

MACHINE LEARNING TECHNIQUES 2: INTRODUCTION TO ML

Overview

Last week, we learned what machine learning is in a general sense. This week, we’ll dig into specifics. What are some actual techniques that people use when they do machine learning? What kinds of data sets do these techniques operate on, and what sorts of predictions do they make?

We’ll explore this in the context of a real-world library application: machine learning techniques as applied to the dime novels collection of Northern Illinois University. (This is a pattern we’ll continue throughout the course: a rough alternation of theory and cultural heritage applications.) This will give you a more concrete understanding of machine learning techniques, as well as showing you opportunities and limitations of ML in a library context.

Finally, you'll get to work with actual machine learning code, so that you can see concretely what the process is like. For those of you without programming knowledge, don't worry; I'll supply working code and point to specific changes you can make to see different effects. If you do have programming knowledge, you can use this code as a jumping-off point for more radical exploration.

Learning Outcomes

- See, and edit, real machine learning code.
- Understand that there are different techniques for machine learning that apply best to different data sets and goals.
- Explore use cases for machine learning with library collections.
- Start to form your own evidence-based opinions on whether, and how, ML can be useful in libraries.

This suppports:

- CLO #1: Understand and explain the basics of AI: both its underlying principles and common machine learning techniques.
- CLO #2: Discuss realistic ways that AI can be a part of library services.

Readings

REQUIRED

- Matthew Short. "Text Mining and Subject Analysis for Fiction; or, Using Machine Learning and Information Extraction to Assign Subject Headings to Dime Novels."
- https://dimenovels.lib.niu.edu/
  - Definitely don't read every dime novel! But do take a look around and familiarize yourself with the site and the collection.
- Kaggle machine learning exercise. https://www.kaggle.com/thatandromeda/sjsu-ai-in-the-library-week-2/

OPTIONAL

- Wikipedia pages for software topics are generally quite good, albeit technical; feel free to look there for more detail on any of the machine learning techniques or terminology we encounter.
- Tutorials on software and data science topics, aimed at a non-software-engineer audience:

        - The Carpentries (Data Carpentry, Library Carpentry, Software Carpentry). https://carpentries.org/
            - Especially the lessons on using spreadsheets and OpenRefine to manage data; also look for in-person workshops
        - Programming Historian. https://programminghistorian.org/en/
            - Examples relevant to this week include TF-IDF (https://programminghistorian.org/en/lessons/analyzing-documents-with-tfidf), Topic modeling and MALLET (https://programminghistorian.org/en/lessons/topic-modeling-and-mallet), or Text mining (https://programminghistorian.org/en/lessons/text-mining-with-extracted-features, using Hathi Trust data!)
    - Ted Underwood. "Seven ways humanists are using computers to understand text." https://tedunderwood.com/2015/06/04/seven-ways-humanists-are-using-computers-to-understand-text/
        - Lots of use cases for computational techniques in the humanities, including but not limited to machine learning; the links are worth following
    - Emily Higgs. "Natural Language Processing for Discovery of Born-Digital Records." https://youtu.be/neeM_o63h7k?t=360
        - Code4Lib talk on natural language processing for named entity recognition to improve archival metadata.
    - Karen Spärck Jones. "A statistical interpretation of term specificity and its application in retrieval." https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.115.8343
        - The 1972 paper inventing TF-IDF

DATA AND ITS DISCONTENTS 1: THE CHARLES TEENIE HARRIS ARCHIVE AND MESSY DATA

Overview

We've seen how to train a machine learning system on a data set. But where does that data set come from, and what's in it? Data sets in the real world — including library data — tend to be messy. Their holes and inconsistencies affect what machine learning systems learn during training, and data cleanup can be a surprisingly large percentage of the work programmers do. At the same time, ML can sometimes be used to remediate data problems, since its ability to find patterns means it can also highlight things which don't fit those patterns.

This week, we'll see an example of a cultural heritage institution using ML to remediate some of its data problems, while also being limited by challenges with its data set. We'll read about data problems encountered in libraries and the work it takes to fix them. And you'll pick a data set of interest to you and start evaluating it.

Learning Outcomes

- Locate an interesting data set.
- Identify key properties of your data set.
- Discuss data messiness: both types of errors that may arise and tools for remediating them.
- Describe the challenges that messy data can pose for AI projects in the cultural heritage space.

This suppports:

- CLO #3: Critically analyze potential pitfalls of AI systems, including the role of their underlying data sets and their ramifications in society.

Readings

REQUIRED

- Dominique Luster, "Machine Learning and Metadata with the Charles 'Teenie' Harris Archive" (starts at https://youtu.be/7mdMtukvtxc?t=7980)
- Any **one** of the following articles on data cleaning:
  - Alicia Detelich, "Large-Scale Date Normalization in ArchivesSpace with Python, MySQL, and TimeTwister." https://journal.code4lib.org/articles/14443
  - Christina Harlow, "Data Munging Tools in Preparation for RDF: Catmandu and LODRefine." https://journal.code4lib.org/articles/11013
  - Thomas Johnson and Karen Estlund, "Recipes for Enhancing Digital Collections with Linked Data." https://journal.code4lib.org/articles/9214
  - Andy Meyer, "Using R and the Tidyverse to Generate Library Usage Reports." https://journal.code4lib.org/articles/13282
  - Mark Phillips, Hannah Tarver, and Stacy Frakes, "Implementing a Collaborative Workflow for Metadata Analysis, Quality Improvement, and Mapping." https://journal.code4lib.org/articles/9199
- https://rise.articulate.com/share/6GSXXSrn2BhRTkg_5t3SJJ7enE5WAWIC#/lessons/GO5qCgbGIJBVTwurFpOEN802HbkEyoHR

- The “Where can I find open data?” lesson

OPTIONAL

- "Teenie Week of Play” and “Teenie Search”. https://github.com/cmoa/teenie-week-of-play, https://github.com/cmoa/teenie-search
  - Machine learning code used by the Charles Teenie Harris archive.
- Any of the other articles from data cleaning list above.
- “Data Equity for Main Street”. https://data-equity.org/
  - An open source curriculum about finding and using open data, developed in part by the California State Library. The source of the “Where can I find open data?” lesson above.
- Catherine D’Ignazio and Rahul Bhargava. Data Basic. https://databasic.io/
  - A suite of tools for analyzing data in a user-friendly way. Includes K12 lesson plans. Note that the Word Counter tool includes “ignore case” and “ignore stopwords” options, and now you know why!
- Hadley Wickham. “Tidy Data”. http://vita.had.co.nz/papers/tidy-data.pdf

DATA AND ITS DISCONTENTS 2: ALGORITHMIC BIAS AND WHAT MACHINES LEARN

Overview

Machine learning systems are exceptional at detecting patterns in data. But what if the pattern is one of unfairness? Even when system designers don’t intentionally encode particular biases into systems, or support those biases, their systems’ outputs tend to replicate and magnify patterns of bias in the data. Depending on how the system is used, the outcomes range from awkward to hazardous.

This week, we’ll learn to look for biases in data sets and machine learning systems, and contemplate where they came from. We’ll see specific consequences of those biases, both in and out of libraries. We’ll grapple with options for reducing bias.

Content warning: necessarily, this week’s work includes specific examples of racism, sexism, and anti-Semitism. Please care for yourself as needed. As you participate in discussions, please be especially thoughtful about the impact of your words on your classmates.

Learning Outcomes

- Identify potential sources of bias in data sets.
- Discuss possible strategies for fixing bias in machine learning systems.
- Draw connections between biases in data sets and real-world risks to people.
- Reflect on possible biases contained within a data set and their implications.

This suppports:

- CLO #3: Critically analyze potential pitfalls of AI systems, including the role of their underlying data sets and their ramifications in society.

Readings

REQUIRED

- Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner. "Machine Bias." https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing
- Matthew Reidsma. "Algorithmic Bias In Library Discovery Systems." https://matthew.reidsrow.com/articles/173
  - The systems discussed here may not use machine learning techniques, just traditional programming, but the problems Reidsma uncovers and the approach he takes are relevant to either case.
- Any one of the optional readings (if it's one of the books, a single chapter is fine).

OPTIONAL

- We are living in a golden age for books about algorithmic bias. Examples include:
  - Meredith Broussard. *Artificial Unintelligence: How Computers Misunderstand the World.*
    - We read Chapter 7 in Week 1, but the whole book is relevant to this week's theme.
  - Catherine D'Ignazio and Lauren Klein. *Data Feminism*. Open review draft edition available at https://bookbook.pubpub.org/data-feminism.
  - Virgina Eubanks. *Automating Inequality*.
    - Prefer video? She gave a talk on the topics in her book, available here: https://cyber.harvard.edu/events/automating-inequality
  - Safiya Noble. *Algorithms of Oppression*.
  - Cathy O'Neil. *Weapons of Math Destruction*.

- Cynthia Rudin, Caroline Wang, and Beau Coker. "The Age of Secrecy and Unfairness in Recidivism Prediction". https://hdsr.mitpress.mit.edu/pub/7z10o269
- Noah Geraci. "Programmatic approaches to bias in descriptive metadata". https://2019.code4lib.org/talks/Programmatic-approaches-to-bias-in-descriptive-metadata
  - How can you clean data when the thing you want to clean isn't typographical errors, but bias?
- Find contemporary news coverage of AI bias in the following stories:
  - Google photos and its problems identifying black people (e.g. https://www.wired.com/story/when-it-comes-to-gorillas-google-photos-remains-blind/)
  - The Wii (same problem)
  - Microsoft's Tay twitter account (every -ism you can imagine)
- James Mickens, "Why Do Keynote Speakers Keep Suggesting That Improving Security Is Possible?" (Usenix '18 keynote): https://www.youtube.com/watch?v=ajGX7odA87k
  - Hilarious but thoughtful critique of machine learning from a computer science professor.
- Additional articles from the Pro Publica machine bias series (https://www.propublica.org/series/machine-bias)
  - Especially these followups to the required reading: https://www.propublica.org/article/technical-response-to-northpointe, https://www.propublica.org/article/propublica-responds-to-companys-critique-of-machine-bias-story
- Data 4 Black Lives. https://www.youtube.com/channel/UCHpUM-G3uVBZhCLarz2qOWA
  - This organization has put on several (extraordinary) conferences on the theme of big data and black lives, with topics ranging from health inequality through political organizing to the arts. Many of the conference videos are available on YouTube.
- Megan Garcia. "How to Keep Your AI From Turning Into a Racist Monster." https://www.wired.com/2017/02/keep-ai-turning-racist-monster/
- Robyn Speer. "ConceptNet Numberbatch 17.04: better, less-stereotyped word vectors." https://blog.conceptnet.io/posts/2017/conceptnet-numberbatch-17-04-better-less-stereotyped-word-vectors/
  - Overview of a technical approach to addressing biased machine learning outputs. This is one of several options — there is not consensus as to the best way to address this problem — and it references other options; the links are worth following.
- Tolga Bolukbasi, Kai-Wei Chang, James Zou, Venkatesh Saligrama, Adam Kalai. "Man is to Computer Programmer as Woman is to Homemaker? Debiasing Word Embeddings." https://arxiv.org/abs/1607.06520

- The quantitative underpinnings of Speer's blog post.

THE SEARCH FOR MEANING 1: HAMLET AND NEURAL NETS

Overview

This week we'll dig into the most popular and important machine learning technique today: the neural net. We'll also explore unsupervised learning.

So far we've mostly seen examples of supervised learning, which is when you train a machine learning system on data that you have labeled with known correct labels, in hopes you will be able to use it to label other similar data in the future. Supervised learning is good for identifying known entities, or classifying using known categories: "Is this email spam or not?" "Who is the person in this photograph?" "What genre is this novella?"

However, we have data sets that haven't been labeled. Maybe labeling them would be terribly costly. Or maybe there isn't an obvious way to label them, because the questions we're asking don't have right or wrong answers. ("I liked these books; what other ones should I try?") So sometimes we throw unlabeled data into a system and ask it to identify whatever structures it can find. This is called unsupervised learning.

Finally, we'll see how this can be applied in a library setting via HAMLET, a set of prototype discovery tools I made using a neural net trained on an institutional repository.

Learning Outcomes

- Characterize neural nets, a particularly important technique for machine learning.
- Grapple with the differences in knowledge organization between traditional library taxonomic structures and unsupervised learning.
- Add to your knowledge of the application of machine learning to library services.

This suppports:

- CLO #1: Understand and explain the basics of AI: both its underlying principles and common machine learning techniques.
- CLO #2: Discuss realistic ways that AI can be a part of library services.

Readings

REQUIRED

- Janelle Shane. *You Look Like A Thing And I Love You*, pp. 62-82.
  - On her blog AI Weirdness, laser physicist Janelle Shane trains neural nets on all sorts of strange data sets (knitting patterns, beer names, D&D spells, recipes, H.P. Lovecraft, recipes and H.P. Lovecraft at the same time…) Hilarity ensues…but also, Shane has the best general-audience explanation of AI that I've seen. It's funny and readable, but still accurate and thoughtful about the technical aspects.
- Andromeda Yelton. "HAMLET: Neural-Net-Powered Prototypes for Library Discovery." HTML https://journals.ala.org/index.php/ltr/article/view/6909/9301, PDF https://journals.ala.org/index.php/ltr/article/view/6909/9307
- https://hamlet.andromedayelton.com
- https://www.propublica.org/article/breaking-the-black-box-how-machines-learn-to-be-racist

OPTIONAL

- More by Janelle Shane!
  - The rest of her book.
  - The blog: https://aiweirdness.com.
  - If you're a knitter, you may enjoy the time Shane (who is not a knitter) worked with a Ravelry community to have her neural net generate knitting patterns, which the Ravelers then tried their level best to knit: https://www.ravelry.com/discuss/lazy-stupid-and-godless/3718985/1-25, https://www.theatlantic.com/technology/archive/2018/03/the-making-of-skyknit-an-ai-yarn/554894/
- HAMLET source code: https://github.com/thatandromeda/hamlet
- Gensim: https://radimrehurek.com/gensim/index.html
  - This popular Python library implements word2vec and doc2vec; I used it to build HAMLET.
- Tomas Mikolov, Kai Chen, Greg Corrado, Jeffrey Dean. "Efficient Estimation of Word Representations in Vector Space." https://arxiv.org/abs/1301.3781
  - The original word2vec paper.

- Quoc V. Le, Tomas Mikolov. "Distributed Representations of Sentences and Documents." https://arxiv.org/abs/1405.4053
  - This paper generalizes word2vec to create doc2vec.
- https://multithreaded.stitchfix.com/blog/2015/03/11/word-is-worth-a-thousand-vectors/, but see also https://multithreaded.stitchfix.com/blog/2017/10/18/stop-using-word2vec
  - Reasons to use, and not use, word2vec, from the StitchFix engineering blog.
- Warren S. McCulloch and Walter Pitts. "A Logical Calculus of the Ideas Immanent in Neural Activity."
  - The 1943 paper, by a neuroscientist and a mathematician, which invented neural nets — many decades before we would have enough computing power to make them, and indeed before either ENIAC (the first general-purpose digital computer) or programming languages. One of the most challenging papers I've ever read.

THE SEARCH FOR MEANING 2: INTERPRETABILITY AND INTENT

Overview

As we saw last week, neural nets can be weirdly effective, but they can also be just weird. Computers don't work like human brains. Their outputs can seem alien.

Naturally, people have tried to understand why machine learning systems reach the decisions they do. It turns out this is a hard question. We have some ways to look inside machine learning systems, but we can't, in general, say why a system reached a given answer.

Is this a problem? People certainly find it disquieting — especially when machine learning systems are applied to decisions with real impact on people's day-to-day lives. Then again, do humans always know why they make the decisions they make?

This week, we're going to look into the abyss and see what we find.

Learning Outcomes

- Discuss interpretability: what does it mean, in an ML context, and why does it matter.

- Abandon hope of analyzing machine outputs as if they were human outputs; recognize machine decision-making as a distinct kind of process.
- Consider contexts in which the strangeness or uninterpretability of ML decisionmaking is salient; analyze whether it is a problem.

This suppports:

- CLO #1: Understand and explain the basics of AI: both its underlying principles and common machine learning techniques.
- CLO #3: Critically analyze potential pitfalls of AI systems, including the role of their underlying data sets and their ramifications in society.

Readings

REQUIRED

- Janelle Shane. "When algorithms surprise us". https://aiweirdness.com/post/172894792687/when-algorithms-surprise-us
  - Sometimes machine learning gets it wrong and it isn't some kind of racist sexist nightmare dystopia. It's just really funny. (But also serious: understanding how the computer got to the startling answers helps us understand the differences between human intelligence and machine learning, and develop better expectations about what the machines are doing.)
- https://slate.com/technology/2015/07/google-deepdream-its-dazzling-creepy-and-tells-us-a-lot-about-the-future-of-a-i.html
- Brent Mittelstadt, Chris Russell, and Sandra Wachter. "Explaining Explanations in AI." PDF available via https://dl.acm.org/doi/10.1145/3287560.3287574.
- Finale Doshi-Velez, Been Kim. "Towards A Rigorous Science of Interpretable Machine Learning." https://arxiv.org/abs/1702.08608

OPTIONAL

- Joel Lehman et al. "The Surprising Creativity of Digital Evolution: A Collection of Anecdotes from the Evolutionary Computation and Artificial Life Research Communities." https://arxiv.org/pdf/1803.03453v1.pdf
  - This is Janelle Shane's source for the blog post above, aka the more detailed and technical version.
- Any of last week's optional readings.

- Finale Doshi-Velez et al. “Accountability of AI Under the Law: The Role of Explanation.” https://arxiv.org/abs/1711.01134
- https://www.fatml.org/
    - In the machine learning context, FAT stands for “Fairness, Accountability, and Transparency”. There’s a whole conference about it. The papers in its archives are interesting.
- https://facctconference.org/
    - In fact, there are multiple conferences.
- Os Keyes, Jevan Hutson, Meredith Durban. “A Mulching Proposal.” https://ironholds.org/resources/papers/mulching.pdf
    - On the limits of the popular FAT framework.
- https://playground.tensorflow.org/
    - This lets you build a tiny neural network in your browser. You can decide how many hidden layers there are; how many neurons per level; and what kinds of things your input neurons should be able to notice about your data (like distinctions between right and left). You can train it on a variety of data sets (by pressing play) and see how good a job it does (or doesn’t do) making a boundary between blue and yellow dots. You will see the functions — i.e. the decision boundaries — that each neuron learns. The thickness of the lines connecting them shows how much each neuron contributes to the neuron in the next layer. Does this help you to understand how the neural net makes a decision? Maybe. Maybe not.
- https://www.tensorflow.org/tutorials/generative/deepdream
    - Make your own deepdream images with Python programming.

AI AND YOU 1: THE UNIVERSITY OF RHODE ISLAND AI LAB: AI AS SERVICE MODEL

Overview

By now, you know all about different techniques for AI; how those technologies can be applied in a library service context; and how to look at AI from a critical and humanistic perspective. What if you’re not a developer, though? What if you work in a library without developers? Can you still use AI as part of your service model?

Yes! Library patrons — whether researchers, teachers, students, or engaged readers of the news — may have questions and projects related to AI. With your shiny new humanistic and technical skills, you probably have a pretty good idea of what those questions might be, and what sources might help your patrons advance their interests. The University of Rhode Island has a model of how this can work in practice with its AI lab.

You'll learn about the services provided through this lab. You'll also have the opportunity to talk to one of the people responsible for the AI lab in the discussion forums.

Learning Outcomes

- Imagine non-software-development-based AI services that libraries can offer to patrons.
- Reflect on what you know that could be useful in developing these services.
- Consider the breadth of AI-related interests that library patrons may have.

This suppports:

- CLO #2: Discuss realistic ways that AI can be a part of library services.

Readings

REQUIRED

- Bohyun Kim. "AI and Creating the First Multidisciplinary AI Lab." https://journals.ala.org/index.php/ltr/article/view/6910
- University of Rhode Island AI lab web site: https://web.uri.edu/ai/
- Either of these writeups of events where a public library collaborated with other cultural institutions to develop and provide space for an AI-themed art experience:
  - "The Laughing Room". https://metalabharvard.github.io/projects/laughingroom/
    - Cambridge (MA) Public Library. It's worth reading some of the press coverage linked from this site, too.
  - "The Library's Other Intelligences." https://fciny.org/projects/the-librarys-other-intelligences
    - Helsinki (Finland) Public Library.
- https://labs.loc.gov/
  - Library of Congress Labs encourages innovation with LC digital collections. This includes LC-sponsored machine learning projects, but also rests on digitization, rights work, and metadata work to enable public use (including machine learning use) of digital collections.

    - There's no need to read everything on this site, but explore the parts that most interest you, and make sure to check out LC for Robots.

OPTIONAL

- None this week.

AI AND YOU 2: FINAL PROJECT PRESENTATIONS

Overview

This week is all about you, and driven by your interests. Share what you've learned in your final project; explore what your classmates have learned; celebrate each others' accomplishments; and discuss all the ideas you've brought to the table.

Thank you for being part of this journey.

Learning Outcomes

- Communicate clearly about AI to a library audience.
- Whatever you wanted to learn!

This suppports:

- One or more of CLO #1, 2, or 3, depending on your project topic.