Research Project Proposal

GA TECH CS6460 FALL 2018
SCOTT LILLEBOE

Introduction

Artificial intelligence (AI) and machine learning (ML) has seen an increase in investment from industries worldwide. In 2018, there will be an estimated $19.1 billion invested in AI systems. That is a 54.2% increase from 2017. IDC predicts AI spending to increase to $52.2 billion by 2021 (Shirer & Daquila, 2018). This investment and interest in AI is increasing at a phenomenal pace with little indication it is going to slow down in the short term. Deloitte anticipates the number of AI/ML projects coming from industry to double in the next year (Choudhury, 2018).

The number of AI/ML courses has increased along with interest in current offerings (Beckett, 2016). The investment in AI/ML from industry would appear to correlate with the increase in interest in AI/ML related courses. The need to train and find employees with the skills necessary to create the AI/ML projects may have helped drive the increased interest in related courses.

There is risk in those creating and interpreting AI/ML algorithms not understanding algorithm bias. Bias in this context is the bias found in datasets and how the algorithms results are interpreted as part of the decision processes that utilize those results. These algorithms are making decisions, biased in some cases, that have a large and lasting impact on people's lives financially through loan applications (DeBrusk, 2018) and criminal sentencing (Kirchner, 2017). There is a rapid industry investment in AI/ML along with courses teaching the subject matter; however, Is AI/ML algorithm bias along with legal and ethical issues of their use being taught to bring awareness of the potential dangers? Are employers asking candidates in AI/ML related jobs their knowledge of these topics when screening potential hires? These are open questions that may have value in their answering.

Related Work

There has been work done to evaluate dataset bias in algorithms. This work involves identifying biased features of datasets (Hajian & Domingo-Ferrer, 2013), bias in recognition datasets (Buolamwini, 2016), how the datasets are gathered and sampled (Cortes, Mohri, Riley, & Rostamizadeh, 2008), and bias introduced through word embedding (Bolukbasi, 2016). Bias in interpreting algorithm results has seen some recent work done. This is especially true in identifying cognitive biases in the decision processes involved in interpreting the algorithms results (Kliegr, 2018). These works have focused on identifying and introducing de-biasing techniques. The work hasn't focused on if we are teaching bias in this context.

There has been discussions on the ethical and legal issues surrounding AI/ML algorithms to generate decisions. Some of those discussions have been on imposing government policy to open up proprietary black box algorithms, so they can be evaluated for potential bias (Savage, 2016). There was no research discovered that shows that legal and/or ethical issues were being taught as part of AI/ML courses.

No research was found that looks to determine if industry is asking new hires their knowledge of AI/ML algorithm bias or legal/ethical issues with those algorithms.

Proposed Work

Hypothesis

The research hypothesis claims AI/ML algorithm bias, legal and ethical issues is not being taught by MOOC educational vehicles and employers are not anticipating new hires to have this knowledge.

Data Gathering & Analysis

There is a lack of information on the extent of AI/ML algorithm bias, legal, and ethical education. The research being proposed is to explore and answer the following topics:

- Is AI/ML algorithm bias, legal and ethical issues being taught in MOOCs?
- Are employers asking new hires their knowledge of AI/ML algorithm bias, legal and ethical issues?
- How can AI/ML algorithm bias, legal and ethical be introduced into a curriculum?

The research would like to identify any correlation between questioning new hires on AI/ML algorithm bias, legal and/or ethical knowledge and if the topics are being presented to students. If those topics are not being taught, then there might be a correlation to why job candidates are not being asked by employers. This correlation may stem from those doing the hiring not understanding the topics themselves as they were not taught the topics. This could produce a feedback loop that is preventing the education from happening. If employers are not demanding employees to know about these topics then this may be causing educational institutions to not teach the topics as no demand has been expressed.

Survey Questions

The data to support the hypothesis will come from a quantitative survey. The survey question topics are broken down as follows:

1. Is bias in the datasets being taught in the following domains?
    - MOOCs (including corporate training and excluding OMSCS)
    - OMSCS
2. Is bias in algorithm result interpretation being taught in the following domains?
    - MOOCs (including corporate training and excluding OMSCS)
    - OMSCS
3. Are legal and ethical issues of AI/ML algorithms being taught in the following domains?
    - MOOCs (including corporate training and excluding OMSCS)
    - OMSCS
4. Are companies hiring for AI/ML related jobs asking new hires about algorithm bias and/or legal/ethical issues of algorithms?

The proposed survey outline serves the purpose of determining if AI/ML algorithm bias, legal and ethical topics are being taught in various educational formats. The research proposal would like to separate MOOCs and the OMSCS program by asking about both separately in the survey. The survey question(s) regarding employer inquiry look to find some correlation between lack of employer inquiry and lack of education in the topics. Publicly available information on syllabi from MOOC courses on AI/ML will also be gathered to supplement the hypothesis of the topics being taught. Syllabi can be outdated and not always available; therefore, the survey is intended to provide data from a student perspective.

Survey Target Audience

The target audience of the survey will be from the OMSCS student body. The rational for using the OMSCS student body was based on these assumptions:

1. Currently taking OMSCS classes or have graduated
2. Have a higher than average likelihood of previously taking traditional AI/ML courses (undergraduate, graduates degrees)
3. Have a high likelihood of taking AI/ML MOOC courses outside of OMSCS (Udacity, Coursera, EDX, Udemy)
4. Have a higher than average likelihood of being in a AI/ML occupational field

2-4 are strictly based on anecdotal evidence gathered from conversations by fellow OMSCS students in AI/ML related courses.

Survey Instrument & Delivery

The survey will be administered through Survey Monkey or Peer Survey (http://peersurvey.cc.gatech.edu/). The survey will be promoted in the CS64060 Piazza class page and on Slack. The Slack channels to be targeted will be the general and AI/ML related channels.

Fall Back Plan

If the survey is unable to be sent to the OMSCS student body or there is not adequate participation the use of syllabus information that can be found from public resources will be used to base bias, legal and ethical education. The research paper may have to omit some of the employer discussion of asking new hires about bias, legal and ethical knowledge if either the survey participation is inadequate or the survey is unable to be performed. There may be available information concerning common interview questions for ML/AI fields that can be parsed and utilized if the survey fails to deliver adequate information. Discussion of this topic may be limited to a section on future research.

Deliverables

Schedule

Deliverable	**Scheduled Delivery**
8 Weekly Status Reports	Every Sunday starting Oct. 14th until Dec 2nd 2018
Intermediate Milestone 1	October 28, 2018
Intermediate Milestone 2	November 19, 2018
Final Paper	December 9, 2018
Final Presentation	December 9, 2018
Final Project	December 9, 2018

Intermediate Milestone 1

The first milestone is to create the survey questions and format as a functional prototype. The goal is to have the questions prepared so feedback from both the mentor and peers can be generated and used to refine the survey prior to release to the general OMSCS student body. During this milestone research will be done to gather any public syllabi of popular MOOC providers to determine if bias/legal/ethical issues in AI/ML algorithms is being addressed. The MOOCs to be looked into include: OMSCS, Udacity, Coursera, EDX and Udemy. These syllabi will act as both a fallback and supplemental information for the research paper.

Intermediate Milestone 2

The preliminary results of the survey that was sent to the OMSCS student body will be presented along with the initial analysis. If the survey results proved to be inadequate (lack of participation or unable to perform the survey), what information can be gathered from public resources about current MOOC offerings in AI/ML will be presented for feedback. The milestone will also look to gather feedback about bias, legal and ethical issues and how it can be introduced into an academic setting based on research found.

Final Project

The following will be archived as part of the final project:

- The survey questions
- The data from the surveys
- Documentation of the survey design chosen
- Syllabi/information from MOOCs that were found
- Research conclusions and recommendations (delivered through a final research paper)

Task List

Week #	Task #	Task Description	Estimated Time (Hours)	Member Responsible
8	1	Create research paper outline	1	Scott
8	2	Complete CITI Training	8	Scott
8	3	Research survey creation software	1	Scott
8	4	Research quantitative survey design	2	Scott
8	5	Weekly status update	1	Scott
9	6	Research quantitative survey design	5	Scott
9	7	Create survey questions	5	Scott
9	8	Weekly status update	1	Scott
10	9	Create survey prototype	2	Scott
10	10	Research public syllabi for current bias/legal/ethical education	6	Scott
10	11	Weekly status update	1	Scott
10	12	Prepare intermediate milestone information	2	Scott
INTERMEDIATE MILESTONE 1 DUE				
11	13	Evaluate milestone feedback	3	Scott
11	14	Create survey final design	3	Scott
11	15	Publish survey to OMSCS (Piazza, Slack)	1	Scott
11	16	Weekly status update	1	Scott
12	17	Analyze survey results	4	Scott
12	18	Determine if survey results are viable	1	Scott
12	19	Research teaching bias topics	8	Scott
12	20	Weekly status update	1	Scott

13	21	Research legal/ethic teaching topics	8	Scott
13	22	Weekly status update	1	Scott
13	23	Prepare intermediate milestone information	2	Scott
INTERMEDIATE MILESTONE 2 DUE				
14	24	Evaluate milestone feedback	2	Scott
14	25	Write final paper draft	5	Scott
14	26	Create final presentation draft	4	Scott
14	27	Weekly status update	1	Scott
15	28	Work on final paper	5	Scott
15	29	Create final presentation	5	Scott
15	30	Weekly status update	1	Scott
16	31	Work on final presentation	1	Scott
16	32	Work on final paper	5	Scott
16	33	Create final project deliverables	5	Scott
FINAL PROJECT DUE				

The task list document can be found here:
https://drive.google.com/open?id=1B9q4xIv2IJcaSlbqKnSQmF0BS2uNTy2b

References

Beckett, J. (2016, February 24). *Why Enrollment Is Surging in Machine Learning Classes*. Retrieved from NVIDIA: https://blogs.nvidia.com/blog/2016/02/24/enrollment-in-machine-learning/

Bolukbasi, T. C. (2016). Man is to computer programmer as woman is to homemaker? debiasing word embeddings. *Advances in Neural Information Processing Systems*, 4349-4357.

Buolamwini, J. (2016, November). *How I'm fighting bias in algorithms [Video file]*. Retrieved from TED: https://www.ted.com/talks/joy_buolamwini_how_i_m_fighting_bias_in_algorithms

Choudhury, S. R. (2018, April 17). *Machines will soon be able to learn without being programmed*. Retrieved from CNBC: https://www.cnbc.com/2018/04/17/machine-learning-investing-in-ai-next-big-thing.html

Cortes, C., Mohri, M., Riley, M., & Rostamizadeh, A. (2008). Sample Selection Bias Correction Theory. *Algorithmic Learning Theory, 5254*.

DeBrusk, C. (2018, March 26). *The Risk of Machine-Learning Bias (and How to Prevent It)*. Retrieved from MIT Sloan Management Review: https://sloanreview.mit.edu/article/the-risk-of-machine-learning-bias-and-how-to-prevent-it/

Hajian, S., & Domingo-Ferrer, J. (2013). A Methodology for Direct and Indirect Discrimination Prevention in Data Mining. *IEEE Transactions on Knowledge and Data Engineering, 25*(7), 1445-1459.

Kirchner, L. (2017, December 18). *New York City Moves to Create Accountability for Algorithms*. Retrieved from ProPublica: https://www.propublica.org/article/new-york-city-moves-to-create-accountability-for-algorithms

Kliegr, T. B. (2018). A review of possible effects of cognitive biases on interpretation of rule-based machine learning models. *arXiv preprint arXiv:1804.02969*.

Savage, N. (2016). When computers stand in the schoolhouse door. *Communications of the ACM*, 19-21.

Shirer, M., & Daquila, M. (2018, March 22). *Worldwide Spending on Cognitive and Artificial Intelligence Systems Will Grow to $19.1 Billion in 2018, According to New IDC Spending Guide*. Retrieved from IDC: https://www.idc.com/getdoc.jsp?containerId=prUS43662418