EXPLORATION OF ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING ALGORITHM BIAS, LEGAL AND ETHICAL TOPICS IN ONLINE CLASSES

Scott Lilleboe

Georgia Institute of Technology

Atlanta, Georgia

scott.lilleboe@gatech.edu

INTRODUCTION

AI/ML has seen an increase in investment from industries worldwide

2018 - estimated $19.1 billion invested in AI systems, a 54.2% increase from 2017

The number of AI/ML courses has increased along with interest in current offerings

INTRODUCTION

Known bias issues

- Pokémon Go
- LinkedIn
- Google face recognition

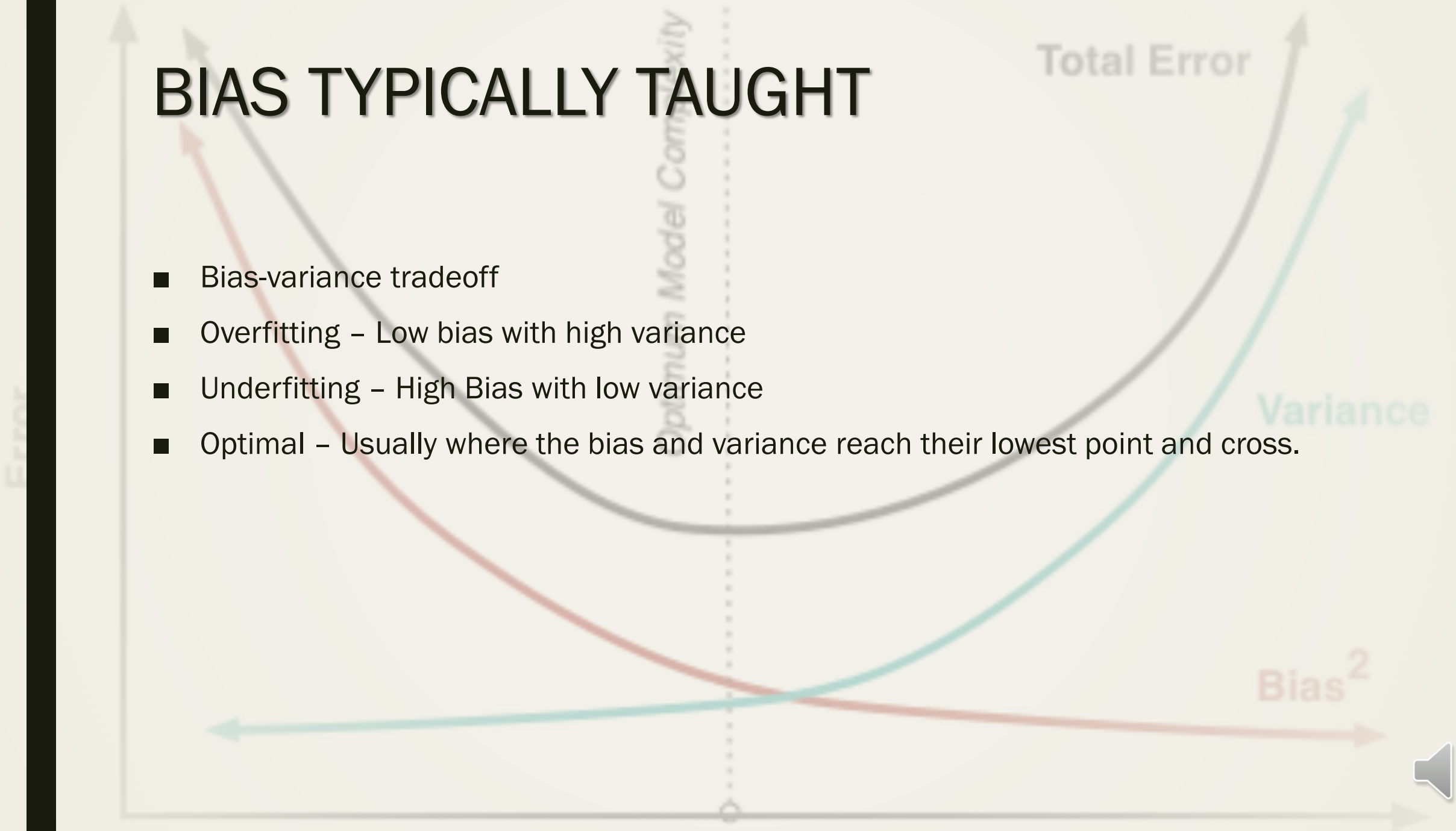
AI/ML algorithms are making decisions which can have a large and lasting impact on people's lives

- Loan applications
- Criminal sentencing
- Employment

Are we teaching the topics of AI/ML algorithm bias, legal and ethical issues?

BIAS TYPICALLY TAUGHT

- Bias-variance tradeoff
- Overfitting – Low bias with high variance
- Underfitting – High Bias with low variance
- Optimal – Usually where the bias and variance reach their lowest point and cross.



INPUT DATA BIAS

Recognition	Sample	Features	Word Embedding
• Image data can suffer from various forms of bias such as selection, capture, category/label, and negative set biases	• AI/ML datasets can be created with sampling selection bias due to the dataset not correctly representing the underlying populations	• Bias can be introduced from features of the dataset that should be omitted prior to the algorithms modeling the data. • Discrimination of certain minority or disadvantaged groups. Direct discrimination can be features explicitly mentioning those groups (i.e. race, gender, sexual orientation).	• Encodes words into a low dimensional continuous space to preserve semantic and syntactic information • Gender bias is often undesirable and can be introduced through word embedding

RESULT INTERPRETATION BIAS

Can be caused by the datasets and algorithm design

Even with correct data and proper algorithms their interpretation can be cognitively biased

There is no generally accepted measure of interpretability of AI/ML model results

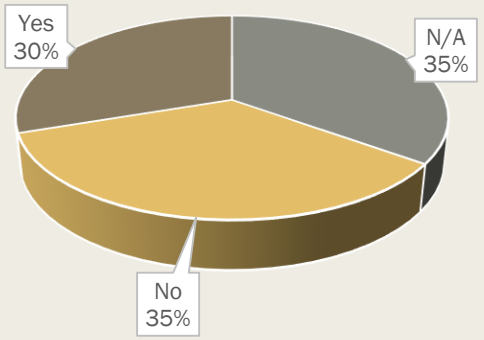
SURVEY

The survey served the purpose of helping to determine if AI/ML algorithm bias, legal and ethical topics are being taught in various MOOCS

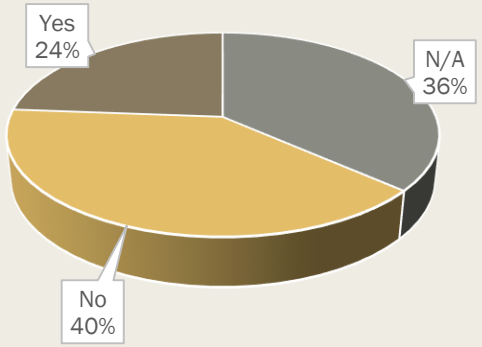
Only publically available syllabi and descriptions could be viewed due to time and financial constraints. Survey was designed to query the OMSCS student body to help identify information.

SURVEY RESULTS

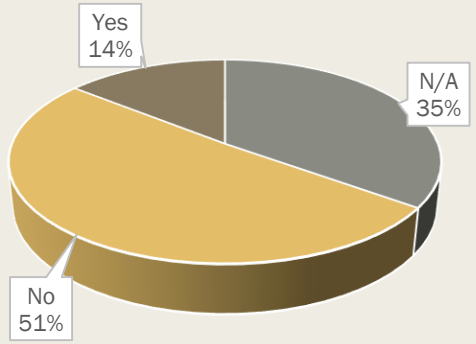
Question 1: If you have taken any AI/ML MOOC courses outside of OMSCS have any discussed how to identify features in the input data that could introduce bias?



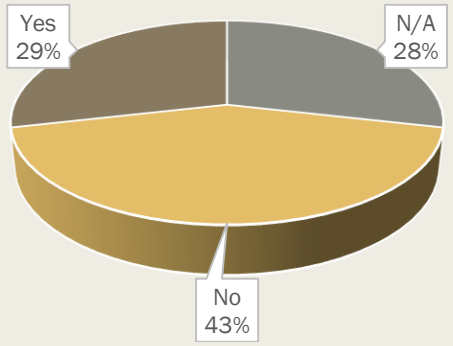
Question 2: If you have taken any AI/ML MOOC courses outside of OMSCS have any discussed how to identify bias when interpreting the AI/ML algorithm's results?



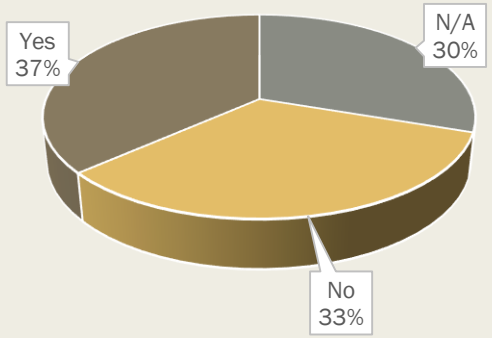
Question 3: If you have taken any AI/ML MOOC courses outside of OMSCS have any discussed legal or ethical implications of algorithm bias?



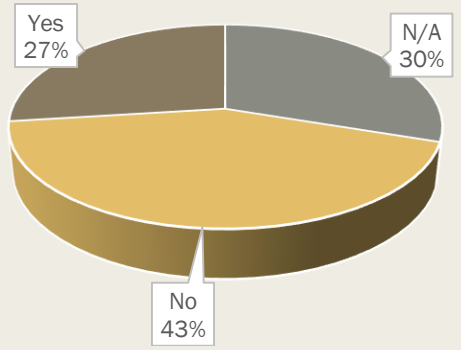
Question 4: If you have taken any AI/ML OMSCS courses have any discussed how to identify features in the input data that could introduce bias?



Question 5: : If you have taken any AI/ML OMSCS courses have any discussed how to identify bias when interpreting the AI/ML algorithm's results?

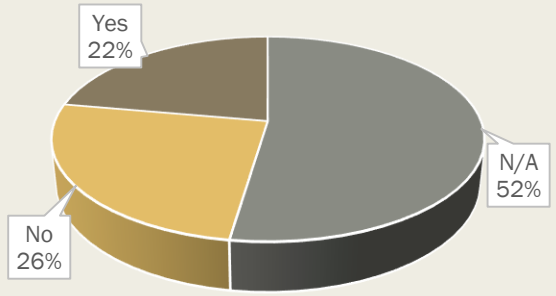


Question 6: If you have taken any AI/ML OMSCS courses have any discussed legal or ethical implications of algorithm bias?

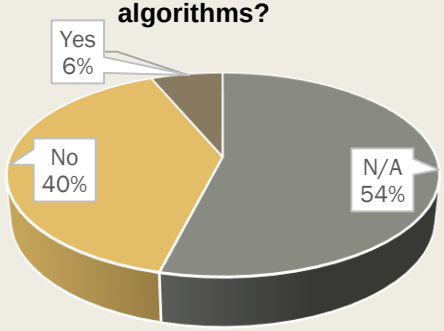


SURVEY RESULTS CONTINUED

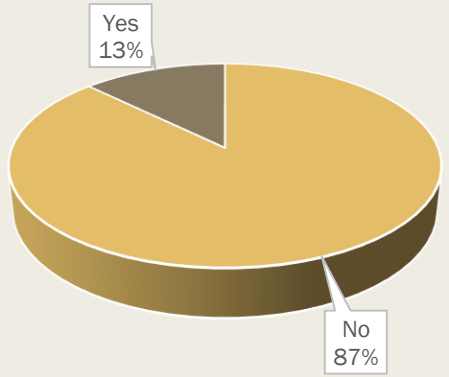
Question 7: If you have been interviewed for an AI/ML related position, were you asked any questions associated with algorithm bias (input data or results interpretation)?



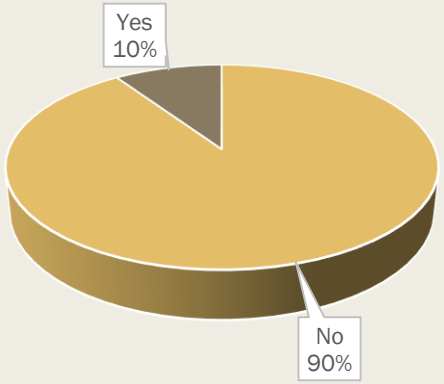
Question 8: If you have been interviewed for an AI/ML related position, were you asked any questions associated with legal or ethical issues regarding AI/ML algorithms?



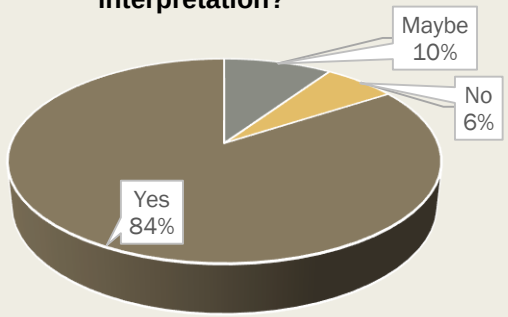
Question 9: Has any employer provided training regarding AI/ML algorithm bias?



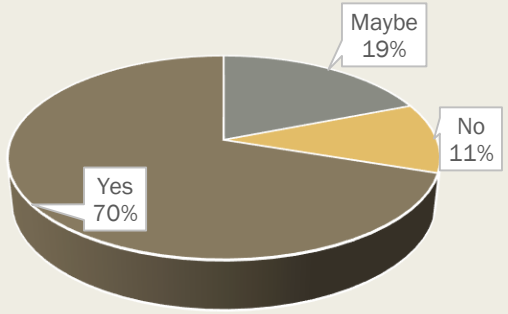
Question 10: Has any employer provided training regarding legal or ethical implications of AI/ML algorithms?



Question 11: Would you be interested in taking a course that teaches bias in AI/ML algorithm input data and results interpretation?



Question 12: Would you be interested in taking a course that teaches AI/ML algorithm legal and ethical issues?



SUPPLEMENTAL SURVEY

- Intent of finding specific examples of MOOCs that had taught algorithmic bias, legal and/or ethical issues based on the respondent's experience

MOOC Provider	Course Title	Bias	Ethical / Legal
Coursera	Machine Learning	3	2
EDX	Python for Machine Learning	1	1
Google	Introduction to Machine Learning		
OMSCS	Artificial Intelligence	1	1
OMSCS	Machine Learning	11	8
OMSCS	Knowledge-Based Artificial Intelligence: Cognitive Systems	6	9
OMSCS	Big Data for Health	1	1
OMSCS	Reinforcement Learning	3	2
OMSCS	Machine Learning for Trading	4	3
OMSCS	AI for Robotics	3	2
Udacity	Deep Learning	1	1
Udacity	AI Foundations	1	1

SUPPLEMENTAL SURVEY RESULTS

MOOC Provider	Course Title	Bias	Ethical / Legal
Coursera	Machine Learning	3	2
EDX	Python for Machine Learning	1	1
Google	Introduction to Machine Learning		
OMSCS	Artificial Intelligence	1	1
OMSCS	Machine Learning	11	8
OMSCS	Knowledge-Based Artificial Intelligence: Cognitive Systems	6	9
OMSCS	Big Data for Health	1	1
OMSCS	Reinforcement Learning	3	2
OMSCS	Machine Learning for Trading	4	3
OMSCS	AI for Robotics	3	2
Udacity	Deep Learning	1	1
Udacity	AI Foundations	1	1

SUPPLEMENTAL SURVEY RESULTS

CONCLUSION

- The survey results imply that AI/ML algorithm bias, legal and ethical issues are being covered in some courses
- There was not a substantial amount of employers providing training to employees
- Better designed survey or better access to the actual course material would provide a clearer picture

WHERE TO GO FROM HERE?

- A substantial interest in people wishing to take courses on AI/ML algorithm bias, legal and ethical issues.
  - *Creation of a MOOC*
  - *Employer training*