

Please note the participation  
at the end of these slides.

# **Data Analytics**

## **CS301**

### **Learn Machine Learning**

Week 13: 7<sup>th</sup> April  
Spring 2020  
Oliver BONHAM-CARTER



# Participation 6 and 7

- Part 6: Coding from two tutorials
- Part 7: Writing own reflections after reading an article
- These participation activities (6 and 7) use the same repository and are each worth a check mark
- Submit to your GitHub repository by Friday 17<sup>th</sup> April, 11:50pm
- **GitHub repository:**  
<https://classroom.github.com/a/qYd0LvAi>

THINK



# Participation 6: Machine Learning!

- On the next slide, you are given several different tutorials to follow online that provide key theory and code to complete a lesson in Machine Learning. Choose two different tutorials from the list, follow the lesson and complete the code that is shown in the tutorial.
- For each code block, leave descriptive comments to indicate what is happening at each stage of code. Remember, *Future-You* will be reading this code someday and so please leave language that is meaningful.
- Place the code of each tutorial in its own file: `src/ML1.R` and `src/ML2.R`
- Please be sure to cite where the code comes from by providing the tutorial title and the tutorial's link in each of your submission source code files.

**THINK**



# Participation 6: Machine Learning Tutorials!

- **Choose two of the following:**
- **Machine Learning in R for beginners**
  - <https://www.datacamp.com/community/tutorials/machine-learning-in-r#five>
- **Your First Machine Learning Project in R Step-By-Step**
  - <https://machinelearningmastery.com/machine-learning-in-r-step-by-step/>
- **Intro to Machine Learning with R & caret**
  - <https://www.youtube.com/watch?v=z8PRU46I3NY>
- **Machine learning with the "diabetes" data set in R**
  - <https://towardsdatascience.com/machine-learning-with-the-diabetes-data-set-in-r-11fa7ae944d0>



# Participation 7: Ethics!

- Obtain and read an article (see following slides for link)
- Read the article and respond to the **Questions in Blue** on the next slide which can be found in your repository: `writing/reflections.md`.
- **Article Citation:** Hagendorff, Thilo. "*From privacy to anti-discrimination in times of machine learning.*" *Ethics and Information Technology* 21.4 (2019): 331-343.

THINK



# Participation 7: Obtaining the Article

- Please follow the below link to the Allegheny's Pelletier Library.
- <https://allegheny.summon.serialssolutions.com/?#!/search?ho=t&fvf=IsFullText,true,f%7CContentType,Journal%20Article,f&l=en&q=From%20privacy%20to%20anti-discrimination%20in%20times%20of%20machine%20learning>
- Log in with your Allegheny User and password when prompted by the library server to get to a page where you can download Hagendorff's article. Note: if the above link does not work for you, then you can still search for the article at the library's online catalog and reach the same download page.

**THINK**



# Participation 7:

## Questions in Blue

- Q0: In a few sentences, summarize the article.
- Q1: What are some of the major benefits of machine learning, according to the article.
- Q2: How can discrimination become involved in a machine learning study?
- Q3: What is “Interdependent Privacy”?
- Q4: Concerning “Interdependent Privacy” how could a user’s data be violated?