**Case Study 1**

Thinking Through Classification

Teaching Resources

**Summary**

This unit is focused on getting the students thinking about classification. The teacher should engage the students by introducing them to many different kinds of classifiers, especially ones that they have likely encountered before. Eventually, there should be a class discussion on the ways that classification can be harmful. Begin by introducing the students to classifiers by asking them to discuss when they might have been classified by an algorithm. Try to think of real-world applications (e.g., Facebook classifying people for targeted advertisements). Then, assign them to read/watch the resources on machine bias and fake news. Then the students should be equipped to tackle the coding homework for Case Study 1: Thinking Through Classification. It is important to make sure that the students have the opportunity to reflect on what they think could be harmful about their recidivism risk algorithm and their fake news algorithm. Consider conducting a debrief class discussion, having them write a reflection, or facilitating group discussions after they have completed the case study.

**CS Topics That Are Covered in This Unit**

1. if/else/elif
2. Writing functions
3. print() versus return
4. Creating a counter variable and updating it with the += 1 operator
5. Data Types (lists, integers, and strings)

**Ethics Topics That Can Be Covered in This Unit**

1. Classifiers
2. Algorithmic Bias
3. Fake News
4. Harmful Uses of Classification
   1. Predictive Policing
   2. Recidivism Risk
   3. Gender Classification
   4. Etc…

**Resources (Reading / Watching To Assign the Students)**

1. [The Science of Fake News](https://www.youtube.com/watch?v=BIv9054dBBI)
2. [Machine Bias](https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing)

**Discussion Questions**

1. Who knows what a classifier is?
2. In what ways do you think you have been classified with your data?
3. How much data do you think you would need to accurately classify your age/gender/interests/etc….?
4. Do you think that it is possible to accurately predict a future human action based off of data about their past? Why or why not?
5. In what ways could classification be used for good?
6. In what ways could classification be used for bad?