Date: 17 October 2018
To: CMDA Capstone Teams
From: Mark Embree and Fred Faltin
Subject: Assessing the Ethical Implications of Your Project

1 Summary

Any interesting data science/mathematical modeling project carries with it nontrivial ethical considerations. In this individual memo, each team member will reflect on how such issues arise in the context of his or her own project.

Each team member should upload a memo to Canvas by 11:59pm on Friday 26 October 2018.

2 Purpose and Contents

Quantitative models bear significant authority: when quantified, a physical, social or economic problem immediately appears more precise. Yet that precision is illusory if the model poorly characterizes reality. Inherent assumptions can be overlooked; input data might be flawed; models calibrated for one parameter regime can be exercised well beyond their range of validity. Such problems are magnified if the model is used by those who do not understand its underlying mechanics.

Moreover, quantitative models often serve as the basis for important decisions: What wing is best for your new jet? Did the foam insulation that hit Space Shuttle Columbia’s wing do significant harm? Should you trade a particular stock? What subset of the population will most benefit from a government program? Bad models lead to bad decisions, resulting in economic loss, design failure, injustice, even death.

Ethical considerations extend beyond the accuracy of the underlying model. Can your model be used for illicit ends? Financial models can help set rational public policy, or reveal opportunities that individuals can exploit for short-term gain at the general public’s expense. Population models predict future resource needs, or can be used coercively by governments. The questions can become quite nuanced. Does an improved missile targeting system cooperate with the agents of mass destruction, or reduce civilian casualties in times of just war? You must cultivate your own carefully calibrated moral compass.

Applications of “big data” often raise deep ethical considerations. Our tastes are shaped by algorithms that feed on the data trail of our electronic lives. As Cathy O’Neill points out in her book Weapons of Math Destruction (2016), models that appear, on first blush, to be morally neutral can in fact encode inherent bias embedded in the data from which the models are constructed. As a result, social prejudice can become ever more deeply ingrained, harder to question as it becomes enmeshed in proprietary algorithms we are unable to scrutinize. The model can become the guardian of the bias it was designed to overcome.

This present assignment, to be completed individually by each team member, should assess the ethical questions surrounding your project. (You are encouraged to discuss the general ideas with your teammates, but the write-up and opinions must be your own.) Please consider issues of
(1) model validity; (2) consequences of policy decisions based on your model; (3) ill uses to which your model might be put. Are there any safeguards you can envision to limit these concerns?

Aim for one or two pages of content: be thoughtful rather than verbose. Think broadly about potential issues, but also be realistic: do not suggest ridiculous or even outlandish scenarios. Be specific about detailed aspects of your project, where appropriate. If your project does not seem to raise major issues (e.g., because it studies a historical event), consider the questions that would arise if similar analysis were applied to the contemporary world. The main goal is to demonstrate that you have thought seriously about the ethics of modeling and data science, in a context that hits home more closely than an abstract case study would.

Each student should post his or her report to Canvas by 11:59pm on Friday 26 October 2018.

3 Grading Rubric

This assignment will be graded on a 100 point scale, according to the following rubric.

- one-paragraph summary of the ethical considerations of your project: 20 points
- specific discussion of ethical considerations: 60 points
  Include (1) model validity; (2) decisions based on model; and (3) potential misuse of the model.
- formatting, style, grammar: 20 points