MATH 342W / 650.4 Spring 2021
Philosophy of Modeling Draft and Final Paper

Professor Adam Kapelner

Draft Due 11:59PM Sunday by email, Mar 21, 2021

(this document last updated 12:16 Noon on Monday 19th April, 2021)

Pick one of the prompts below or email me if you have your own topic of interest so I can approve it prior to you beginning the project. Instructions on titling your paper is given after the prompts.

(a) The laws of physics are absolutely true /
The laws of physics are *not* absolutely true /
The laws of physics are *perhaps* absolutely true.

(b) People who smoke contract lung cancer /
People who smoke do *not* contract lung cancer /
People who smoke *perhaps* contract lung cancer.

(c) Credit scores like those given by Experian, TransUnion and Equifax are useful /
Credit scores like those given by Experian, etc. are *useless* /
Credit scores like those given by Experian, etc. are *perhaps* useful.

(d) Autonomous vehicles will navigate perfectly /
Autonomous vehicles will *not* navigate perfectly /
Autonomous vehicles will *perhaps* navigate perfectly.

(e) The optimal human diet for health is [insert a diet here] /
The optimal human diet for health is *not* [insert a diet here] /
The optimal human diet for health is *perhaps* [insert a diet here].

(f) The medical establishment understands the human body /
The medical establishment does *not* understand the human body /
The medical establishment *perhaps* understands the human body.

(g) What makes a virus a dangerous pandemic is understood /
What makes a virus a dangerous pandemic is *not* understood /
What makes a virus a dangerous pandemic is *somewhat* understood.

(h) Global warming is real /
Global warming is *not* real /
Global warming is *perhaps* real.

(i) The government understands the optimal level of taxation /
The government does *not* understand the optimal level of taxation /
The government *perhaps* understands the optimal level of taxation.

(j) What makes a marriage successful is understood /
What makes a marriage successful is *not* understood /
What makes a marriage successful is *somewhat* understood.

The common theme among the prompts are the existence of some natural phenomenon which is then modeled by humans. Sometimes it is difficult to separate the phenomenon from the model since the model is so common and we are so brainwashed to think it is same as the phenomenon (it is not). Before you begin, ensure you understand what the phenomenon is you are modeling. If this is not straight, then your entire essay will be a failure.

The following must be written in your essay

- A title that sums what the phenomenon is and how accurately you believe you can model it e.g. "The [phenomenon] can be Modeled Well" or "The [phenomenon] Remains Poorly Understood" or "The [such-and-such model] for [phenomenon] a [good/bad] Model".
- An introduction of no more than 1.5 pages that talks about the phenomenon, how important it is to humans and history of attempts to model it.
- Definition of a model. Definition of a mathematical model.
- Description of what a model means in the context of your phenomenon.
- Description of your phenomenon and how it is measured exactly. Do you believe this measurement can be made accurately?
- Definition of supervised learning and what it would mean in the context of your phenomenon.
- Description of your independent variables and how they are measured exactly. Do you believe your measurements can be made accurately?
- Definition of historical data observations (training data). How would you obtain a dataset in your context?
- Definition of the three sources of error and what they are in the context of your modeling effort. Which of these sources would be large and why? Which are small and why?
- Definition of prediction. What prediction means in the context of your model. Who will be predicting and for what purpose?

- Definition of an error metric for predictions. What error metric would you use and what is the threshold for "usefulness" in your context?
- Definition of what stationarity means and what it means in the context of your phenomenon.
- Definition of interpolation and extrapolation and what they mean in the context of your phenomenon.
- What is an algorithm and candidate set? What is machine learning? What are some algorithm choices would be to build this model. Which choice would you choose in your context? Do you have enough sample in your historical data to do so?
- Define underfitting. Could your model be underfit?
- Define overfitting. Could your model be overfit?
- Define validation, in-sample and out of sample. How would you validate your model?
- Throughout the essay you must use all the following notation where appropriate:

  $$t, f, g, g_0, h^*, \delta, \mathcal{E}, e, \mathbb{D}, \mathcal{H}, \mathcal{A}, t, z_1, \ldots, z_t, n, p, X, x_{\cdot 1}, \ldots, x_{\cdot p}, x_{1\cdot}, \ldots, x_{n\cdot}, \mathcal{X}, y, \mathcal{Y}$$

  and explain these concepts in the context of your model and phenomenon.
- Conclude: is the title of your essay correct and why? Tie the answer to what you believe will be your model's predictive performance.

You are welcome to bring outside sources about philosophy of modeling as well as sources which help make your arguments in support of a prompt. Please cite them appropriately using natural text citations e.g. "The measurement device is accurate (Johnson et al., 1999)" or "Johnson et al. (1999) demonstrate the measurement device is accurate" and enter them into a bibilography in APA style where the reference elements are formatted in APA style.

Specs: Your essay must be typed and must be at least 10 pages double-spaced with one inch margin, 12pt Times (or Computer Modern if using LaTeX) and be appropriately organized. You choose a title which will be atop the first page. No need for a title page. Sectioning is at your preference. But, one of the choices in one of the prompts must be a powerful sentence in either the beginning or the end of the essay. The bibliography does not count towards the page limit. Keep footnotes to a minimum and do not use endnotes. You must email me a PDF of your paper (Microsoft Word allows saving as PDF).

Revision is Due 10d after draft is graded (as a PDF by email))

You will receive comments on your draft after your submission. You will have the opportunity to revise your draft once and resubmit as a PDF. Resubmit by responding to my email with your completed PDF. You will receive a grade for this assignment assessed by the performance on your revision (not your initial submission).