

# Answer\_PS1

January 19, 2020

## 1 Perspectives on Computational Social Science

### 1.1 Problem Set 1

#### 1.1.1 1. Building Models

##### Deviant Aggressive Behavior

- Theory I: The theory states that individuals learn deviant aggressive behavior from experience, i.e. rewards and punishment. Correspondingly, a social policy should reduce the rewards that may incentive this type of behavior and increase the punishments that discourage it. If we want to target criminal behaviors, for example, the relatively easy way is to charge criminals with high fines and/ or give them long-term sentences. Other types of punishment such as death penalty may also be considered if social norms and current legislations permit. On the other hand, a potential policy may instead focus on enhancing the judicial system so that innocent individuals are not wronged in any form; that is, they should not be punished for doing the right things. If practical, social-policy-makers may even want to award those who have never done anything deviant by, for example, giving them higher credit limit. They may even try to broadcast these awards and punishments, so that individuals can learn from not only their own experience but also others’.
- Theory II: The theory interprets deviant aggressive behavior as a way to show anger toward personal authority figures. It is hard to imagine a social policy which can directly reduce a subjective attitude like anger, especially when the theory does not explain why individuals target personal figures rather than the social system as a whole or those who directly cause their frustration as the target. But in general, if we want to reduce, say, rude and even criminal behaviors against innocent authority figures, a social policy may fund mental health programs that help mitigate aggressive emotions to indirectly make certain differences.
- Theory III: The theory assumes a profound social conflict between those who oppress others and those who are oppressed and therefore defines deviant aggressive behaviors as an action the latter choose to challenge the established rules. In this sense, all deviant aggressive behaviors are by nature revolutionary. To reduce such revolutionary actions, the first policy option is to radically reform the social structure so that nobody feels oppressed. This sounds like a communist world in its extreme sense, where nobody has property or any power over others. However, this social policy is largely unpractical. An easier way is probably to compensate the oppressed. A social policy may instead tax heavily on the rich to fund certain social security programs that help the poor or other historically discriminated communities.

With this policy, oppressed individuals may find it rational to conform with the rules rather than risk their available subsidies to challenge the social system as whole.

- Theory IV: This theory considers deviant aggressive behavior as a type of deviant subculture. Under its assumption, certain individuals *choose* to be deviant and their subculture may influence other individuals. Therefore, a social policy should work to alienate these deviant individuals so that they cannot “contaminate” others and may even force them to give up this social role. In practice, this might mean publicly broadcasting the danger of this subculture, and even criminalize these deviant individuals, just like how many governments do to discourage people from joining pyramid scheme or street gangs. (word count: 519)

**Waiting until the last minute** This observation might be true, because personally, I sometimes do want to turn in papers and assignments quite late and I used to have professors who release scores right before the deadlines.

- First Model:  
People tend to procrastinate probably because they need to prioritize some more emergent tasks. To generalize this explanatory model is to assume that individuals are rational agents who want to maximize their utility. Facing multiple tasks at the same time, they rationally prefer a strategy that allows them to finish as many tasks as possible; that is, to prioritize the tasks that have closer deadlines. For example, a professor chooses to grade exams right before the deadline, so that she can submit her own paper to an important journal in time first.

Because this model assumes that people want to maximize utility and wait until the last minute only to prioritize some more important tasks, one can make two predictions. First, if, say, a student only has one assignment which she can easily finish in a week, and she is not busy at all, then she won't wait for a month until the deadline to submit this assignment. Second, if an important task requires time to progress, then a utility-maximizing person will start the task earlier so that she can have enough time to finish the task, rather than wait until the end and finish it in hurries. (Honestly, these are not always the true.)

- Second Model:  
Alternatively, we want to keep this rational choice assumption but assume a different logic. In this model, we assume that people do things at the last minute because the future is uncertain. Imagine a professor who has a history of always changing her assignment instructions, or a course which allows students to discuss about a hard assignment over Piazza, a rational student would want to wait until the last minute so that she won't waste time confusing with wrong instructions and can get inspirations from other students' post on Piazza.

If, abstractly speaking, people rationally procrastinate because of uncertainty, then we have another two predictions. First, if a student anticipates that the situation won't change significantly before the deadline she is facing — the uncertainty level is low — then she will not procrastinate to finish it. Second, if everybody is rational, then nobody will start to work on that hard assignment early, and therefore nobody is going to post anything on Piazza, and nobody is going to be benefited from others' posts by waiting until the last minute. A potential way to avoid this self-contradictory result is perhaps to assume that only few students are rational and they are those who want to wait rather than finish their assignments as soon as possible.

### 1.1.2 2. Selecting and fitting a model

1. large sample size, few predictors:

A flexible statistical learning method is better because 1) it avoids biases and 2) a large sample size and few predictors make it hard to overfit.

2. many predictors, small observations: An inflexible method is better because it avoids overfitting. In this case, a flexible method is likely to significantly increase variance without reducing significant bias.
3. Nonlinear relationship A flexible method is better because it is hard to anticipate how this nonlinear model is with human intuition. Therefore, it is better to let machines itself to detect the best non-linear association.
4. High error variance An inflexible method is better because a flexible method, trying to minimize the variance of errors as much as possible, is overly sensitive to such noise.

Think about the graph including bias, variance, training error, test error, and irreducible error curves, moving from less flexible statistical learning methods towards more flexible approaches

$$E(Y_0 - \hat{f}(X_0))^2 = Var(\hat{f}(X_0)) + [Bias(\hat{f}(X_0))]^2 + Var(\epsilon) \quad (1)$$

(2)

- Absolute/squared bias decreases as the model gets more flexible and is likely to be inversely proportional to flexibility. Higher flexibility means that the model can better fit the train data, and therefore gives less bias. However, the marginal benefits of higher flexibility shrinks quickly as the model starts to overfit the train data, and therefore it is a inversely proportional relationship.
- Variance increases with flexibility. As the model starts to fit the train data overly well, it starts to gain higher variance. Also, just like the formula states above, as biases decreases, variance goes high.
- $Var(\epsilon)$ , the irreducible error—is independent of model selection, and therefore it should be flat, i.e. constant, in this graph.
- Training error should decrease with flexibility because, again, more flexible model tend to fit the train data more closely. (overfitting)
- Test error should decrease at first with flexibility because a slightly flexible model better captures the characteristics of the ground-truth model and therefore gives better predictions. However, it should soon start to increase with flexibility because an overly flexible model overfits the train data and therefore gives bad predictions.

[ ]: