IN-STK5000 Project 2 - Project 5

Anders Bredesen Hatlelid
Even Tronstad
Jacob Nicolai Arthur Sjødin
Torgeir Ladstein Waagbø

December 8, 2021

Initial planning

- In this part of the project, you are supposed to construct a new policy for making either treatment or vaccination decisions.

    –

- It is your choice which one of the two you wish to work on.
- In either case, choose one and stick to it.
- The policies will be ran with a simulator to be provided later.
- No matter what you choose, be sure to define the utility that your policy should maximise.

    –

- Your specification of a utility function also has some ethical dimension, because that is how you choose what to do.

    –

- Life expectancy, success probability of treatment, cost of treatment, and above all, availability of treatment, are all important factors to consider.

    –

- You can start off the project with a simple utility for the first deadline, and refine it later.

    –

- The project requires you to specify policies from the outset.

    –

- A policy can be generated easily once you have a predictive model, but you can leave policy optimisation and model tuning for the last part of the project.

    –

Privacy analysis (Deadline 1: 5 November)

- Does the existence of this database raise any privacy concerns?
  - What are som particularily sentitive variables, f.ex genes?
  - The existence of the database does indeed raise privacy concerns. In particular, a persons genome might be used to identify individuals.
    * TODO: discuss different privacy mechanishms.
    * Be more specific?
    * Reflect on which of the variables are most important for identification.
  - Other points from brainstorming: Mention:
- If the database was secret, but your analysis public, how would that affect privacy?
  - This is actually the case.
  - Might be possible to identify individual from the summary statistics.
  - TODO: discuss the points Anonymity, Secrecy, side-information and Utility on page 76 in the notes.
  - From brainstorming: Outliers like high age? Perhaps state the most most critical features.
  - How I interpret the question: We assume an adverasry to have perfect knowledge and unlimited computer power. We then need to ensure than he cannot identify the individual that participate in the study.
- (a)
  - Explain how you would protect the data of the people in the training set.
    * Need a general discussion about how to make a database private.
  - In particular, given that your policy and model are obtained from some 'training' data set, how would you guarantee that release, or use, of the policy and model does not leak private information about the individuals?

    * Here we need to discuss how to to make the desition and model private. Not sure if the individual reponses to the vaccines need to be private.

- (b)
  - Explain how would protect the data of the people that obtain treatment.
    * Is this with regard to just the storage of the data in the database?
  - When you apply the policy or model to decide what treatment to give, this decision can be assumed to be publicly available.
    *
  - How would you then ensure that the private information of the treated individual is not leaked?
    * Here it is the case that we need to answer how to prevent the identification of the individual based on the information that is public. The public information we release is vaccine desi
- (c)
  - Implement a private decision making mechanism for (b).
    *
- (d)
  - Estimate the amount of loss in utility as you change the privacy guarantee.
    *

Fair Policies (Deadline 2: 19 November)

Sensitive varaibles

We do not want to discriminate with regards to gender, that is, we want our choice of action to be independent of this variable. Income is another variable we could have considered, but to simplify, we will choose gender as our sensitive variable z.

Concepts of fairness

We will look at the concept of fairness as independence. Our action a is independent of a variable z if $\mathbb{P}^{\pi}_{\mu}(a|z) = \mathbb{P}^{\pi}_{\mu}(a)$. Furthermore, given a response y, we can either calibrate or balance the policy with respect to the sensitive variable. We will choose the latter, such that the action a is independent of z given the true outcome y. This is according to the definition of balance (4.3.2):

$$\mathbb{P}^{\pi}_{\mu}(a|y, z) = \mathbb{P}^{\pi}_{\mu}(a|y) \quad \forall y, z$$

Measures of fairness

We need a metric to measure fairness. We will use

$$F_{\text{balance}}(\mu, \pi) := \sum_{y,z,a} |\mathbb{P}^{\pi}_{\mu}(a|y, z) = \mathbb{P}^{\pi}_{\mu}(a|y)|^2$$

To estimate $\mathbb{P}(a_i|y_j, z)$ and $\mathbb{P}(a_i|y_j)$ for $i \in \{\text{No vaccine, vaccine 1, vaccine 2, vaccine 3}\}$ and $j \in \{\text{Presence of critical symptom, Not presence of critical symptom.}\}$ we can use the rate we obtain after running the policy on a simulated population.

Tuning a policy to be fair

Increasing the degree of fairness measured by the fairness-metric, may result in a loss in utility. To balance utility and fariness we intriducte the value:

$$V(\lambda, \mu, \pi) = (1 - \lambda)U(\mu, \pi) - \lambda F(\mu, \pi)$$

where λ is a parameter that needs to be tuned. To find the optimal tradeoff, we can use stochastic gradient descent. We have not managed to complete this task.

The optimization task above is unconstrained with respect to the amount of unfairness. However, it might be the case that we can only allow a certain amount of fairness, say ϵ. The problem would then be

$$\max_{\pi} U(\pi, \mu) \text{ s.t } F(\pi, \mu) < \epsilon$$

Bias in the data collection

START OLD

- Choose one concept of fairness, e.g. balance of decisions with respect to gender.
  - We will look at the concept of fairness as independence.
    Parity: probability of action is independent of the sensitive varaible.
  - NEW: We have chosen balance as a fairness criterion. That is,

    $$\mathbb{P}(a|y, z) = \mathbb{P}(a|y).$$

    Here a is the action, y is the outcome and z is a sensitive variable. This means that the action chosen is independent of the sensitive variable given the true outcome y. We regard gender to be a sensitive variable.
  - QUESTIONS
    * It is contitioned on y. What about x is in the feedback?
  - We originally chose the fairness criterion equality of opportunity defined as

    $$\text{Equality of opportunity} = \min\left(\frac{P(\hat{y} = 1|z = 1, y = 1)}{P(\hat{y} = 1|z = 0, y = 1)}, \frac{P(\hat{y} = 1|z = 0, y = 1)}{P(\hat{y} = 1|z = 1, y = 1)}\right)$$

    However, we got feedback suggusting this is not a good measure of fairness in this case, since it does not account for the action. We insted were encuraged to use a criteria on the form $P(a|x, ?) = P(a|x)$. We would like to get some feedback on what ? should be. Is '?' f.ex the sensitive variables?
- How can you measure whether your policy is fair?

- NEW: We need a metric for measuring the degree of fairness with regards to our policy. To this end we can compare the distribution of $\mathbb{P}(a|y,z)$ and $\mathbb{P}(a|z)$. We can sum the squared difference over all y, z, a, i.e

$$F_{\text{balance}}(\theta, \pi) \triangleq \sum_{y,z,a} |\mathbb{P}_\theta^\pi(a|y,z) - \mathbb{P}_\theta^\pi(a|y)|^2$$

- Questions
  - There are two different setups, which should we choose?
- OLD: Given that $P(a|x, z = 1) = P(a|x, z = 0)$ is a good criteria, we will calcuate this probablity for all the sensitive variables. If we get values close to 1 we will say that our policy is fair. We would also like to have a treshold where values above the thresold are acceptible.

- How does the original training data affect the fairness of your policy?
  - NEW: TODO
  - What does this point mean?
  - OLD: In the simulator, when data is generated, the people are vaccinated. To measure faness in the original training data we will then see how vaccines are distributed among the senistive variables.
- To help you in this part of the project, here is a list of guiding questions.
- (P1)
  - Identify sensitive variables.
    - NEW: We view gender as a sensitive variable.
    - OLD: One sensitive variable with regards to fairness is gender, because we dont want to discriminate people based on their gender. Another sensitive variable is salary, beacuse we dont want to give people a treatment or vaccine based on their salary. For a variable to be considered sensitive we must belive that the variable should not be taken into account by the policy when it chooses an action.
  - Do the original features already imply some bias in data collection?

    * Here we need to present diagrams of how features are distributed in the population. Perhaps start with gender, income and age?
    * NEW: TODO: tips from AMG: we must state precisely which features in the data that seems relevant and which irrelevant to us. We are not asked to suggest methods for reducting bias. We must strategically find out if the given data might be biased. How?
    * Questios:
    * OLD: To reduce our bias in our data, we must collect data that represents the whole population, and our data collection should not be based on belives we already have. For example if we want to test if a vaccine increases the probabilty of a symptom given a comorbiditie, we should not only collect data from people with the comobordotie, but also from people without that comorbidite, such that our data reflects our entire population. Also, it is important to collect data with variables that is important for our outcome. For example it could be important to use the location of where people live to predict if a person gets infected with Covid-19. A final convern is the the variables must be logical. For instance, having a variables 'number of male children' does not make sense. It would make more sense to inlude the number of children instead.

- (P2)

  - Analyse the data or your decision function with simple statistics such as histograms.

    * NEW:
    * OLD: We have plotted histograms of some of the variables. The age histogram is unrealisitc. For on thing, the birth rate drops exponentially in recent years. Moreover, after the age of around 25 there is an exponentiall decrease in the survival rate. We see that the age of our data is centered between 20-50 years old, which can be a realistic assumption. One problem os however that there is not that much young people, and we also have some very old people (200 years old) which is not realsitic. We also see that our data is balanced in respect to the genders. When it comes to the income of our data, this also looks to reflect a general population. The distribution

look like a pareto distribution. We would have expected the histogram to peak the minimum wage and not at 0.

    * OLD: We also bootsrap our data to see if there is big variation, but it doesnt look to be any big variation in the data.

- (P3)
  - For balance (or calibration), measure the total variation of the action (or outcome) distribution for different outcomes (or actions) when the sensitive variable varies.
    * TODO: programming
- (P4)
  - Advanced: Using stochastic gradient descent, find a policy that balances out fairness and utility.
    * TODO: programming

FINISHED OLD

Experiment Design (Deadline 3: 3 December)

- (1)
  - Using the utility function you have specified, estimate the utility of policies on the historical data.
  - Measure the utility of the historical policy on the historical data.
  - Provide error bounds on the expected utility and explain how those were obtained.
- (2)
  - Find an improved policy, and calculate the expected utility of the improved policy on the historical data.
- (3)
  - Obtain an estimate of the historical policy.
- (4)
  - Simulate the historical policy and your improved policy on the obtained simulator.