

# IN-STK5000 Project 2 - Project 5

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

December 6, 2021

## 1 Our policy and model

We are uncertain in regards to whether we have chosen a good model, so here is a description of our model.

We have chosen the adaptive treatment allocation setup described in chapter 7. That is, when treating  $N$  patients with a vaccine, we want to reduce the symptoms experienced by the  $N$  patients as much as possible.

We view each pair of action and symptom to be a bandit. For example the probability of fever when giving vaccine 2 is unknown to us. We model this probability with a beta prior. When estimating the probability  $P(\text{fever}|\text{vaccine 2})$  we take the expectation of the beta distribution.

The drawback with this approach, we think, is that when patients number 100 comes in to get a vaccine, we base our decision not on the feature of that patient, but by the observed symptoms from the 99 earlier patients.

For our utility function we weigh each symptoms with its own negative constant, which can be viewed as a reward. For example when the symptoms are: Covid-Recovered, Covid-Positive, Taste, Fever, Headache, Pneumonia, Stomach, Myocarditis, Blood-Clots, Death, we can give it the rewards

$$\mathbf{r} = -(0, 0.2, 0.1, 0.1, 0.1, 0.5, 0.2, 0.5, 1.0, 100)$$

with the utility being  $U = \sum_{i=1}^{10} r_i$ . The expected utility is then  $EU = \sum_{i=1}^{10} r_i P(r_i)$ .

Our policy is the to maximise the expected utility.

## 2 Questions

1. Is this a good model? We still have some time before the postponed delivery, so we can change it if it is desirable to base the action given to a patient on the features of the patient.
  - We have another suggestion for a model we are considering instead. The current setup considers 10 symptoms and 4 choices (0 vaccine, or vaccine 1-3). This makes 40 bandits. What if we instead only made three bandits like so: We consider only the three critical symptoms: Pneumonia, Myocarditis, Blood-Clots. We construct a single binary variable 1: 'presence of at least one critical symptom' and 0: 'not critical symptom'. We disregard the action of not giving a vaccine, so we have three actions, one for each vaccine. Thus our goal for vaccinating  $N$  individuals is to maximize the number of given vaccines resulting in fewest critical symptoms.
2. Using this model, does it make sense to use balance as a fairness criterion? It seems that using the bandits we are measuring the calibration  $\mathbb{P}(y|a)$ . Also, we can talk about fairness with regards to gender if we don't use the features to decide on the action taken. A suggestion is in the next point.
3. If we use balance as a fairness criterion and we want to measure the degree of fairness

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

Is it okay if after giving vaccines to  $N$  people, we divide the individual into groups with similar vaccines and symptoms, and then look at the difference in the count of the two genders?

4. In your feedback on the point 'Bias in the collected data': 'How would you strategically find out if the given data might be biased?' Could you give us a push in the right direction for answering this point?
5. Could we get some input on how to approach the question 'How does the original training data affect the fairness of your policy?'.