

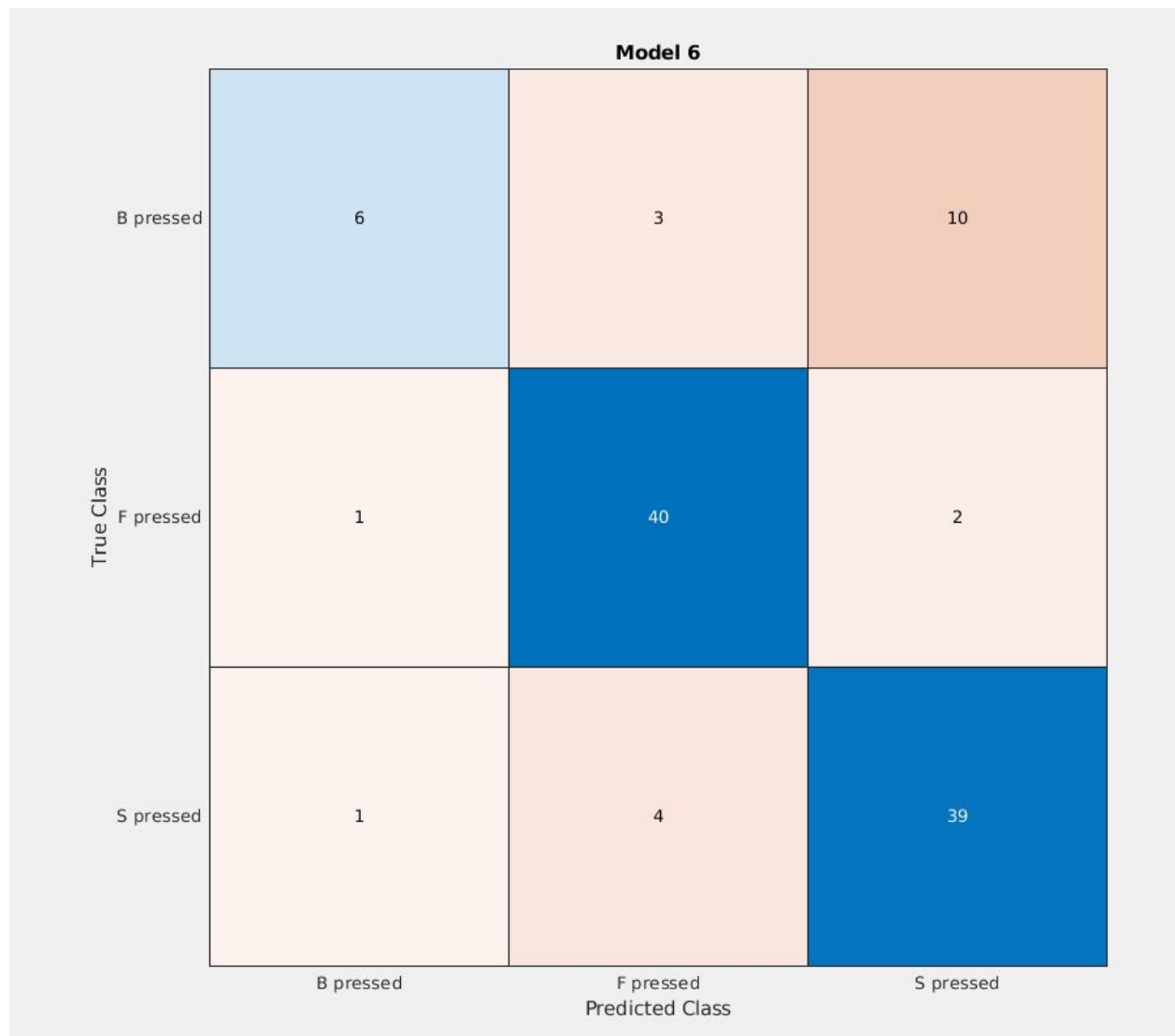
Katie Foster

Reflection:

My initial goals were to do part 1 over the weekend, and finish up part 2 (option B) during the week. I did get to do most of part 1 over the weekend. I watched the video, took notes, and listened to most of the podcast while baking (hoping to finish it tonight). I did all of part 2 today (Tuesday). I went through the notebook and did the exercises. I looked up what the different included features meant, but did not test out different ones. I opened up the classification learner and played around with it, but when I tried to export it to run on testing data, I ran into an error and did not spend much time trying to fix it. I did the assignment but not as well as I wish I had done. However, I am going to be forgiving of myself and recognize that I learned things, even if I did not do the assignment to my fullest potential. One good thing I did this time that I did not do last time was that I took notes on the information in part 1 (the video mostly). I really enjoyed how the video portrayed experts from some different fields, and how they each had a slightly different perspective. They broke down things for me in a way I had not really considered, eg: data security is a subset of data privacy, and how it is so important to have consistent regulation across states and even countries. I liked the sources in part 1 (at least the two that I engaged with: the video and the podcast). I also think the notebook was about the right amount guided. Next week I think I will try to engage more with the context/applications part of the assignment, if there is one. I have been saying to myself that I will do that if I get to it, but then not getting to it, so I would like to make it more of a priority next time.

Data Privacy Notes

- need to avoid so much complexity that people don't need and just opt in
- some services need to know certain data (amazon needs to know where you live), but others don't
- don't use DNA for hiring, but where do you draw line, and how do you deal with things in the middle? opt in? opt out? notices?
- don't treat all data the same - some is much more sensitive
- need to have consistent data laws across industry, sector, state lines
- lots of small inconsequential data can combine to allow you to make very consequential decisions and inferences
- Data privacy may be particularly important for socially marginalized groups
- data security (guard against breaches) is subset of data privacy
- standard/duty of care for data
- Right to be forgotten and begin again?
- make sure data privacy laws don't disproportionately hurt small businesses



I got 80.2% accuracy on the training data with a Fine Gaussian SVM, but unfortunately ran into an error trying to run the exported model on the testing set, so I don't have the info for that.