Extra Credit 1 Pts

- Listen to the AI Today podcast on Machine Learning Ops and provide your thoughts on the following questions:

1. Does knowing about Git and GitHub help you in understanding the podcast?

Git and GitHub are good tool for software and data engineers. They could easier manage and track the ongoing project status with their co-workers, especially Git and GitHub are convenient for remote collaboration (such as work from home due to COVID-19). Moreover, we learned devtools package on week 5. The aim of devtools is to make package development easier by providing R functions. Marsden has mentioned devtools in the podcast, data engineers may not have programming backgrounds, thus devtools is good for data engineers to develop new packages.

2. How do you think the ideas of ML OPs will affect your future data science projects? You may also want to check out this article on Towards Data Science.

As the code and data evolve independently, and typical Data scientists are not often as trained engineers, so they do not always follow good DevOps practices. “MLOps is a new discipline for collaboration and communication between data scientists and information technology (IT) professionals while automating and productizing machine learning algorithms” (Talagala, 2018). ML Ops is like a bridge that connects ML engineers and Data engineers having a smooth workflow, avoiding information asymmetry, and maintaining ML systems in production reliably and efficiently.