Possible changes to Bus241a syllabus, and commentary for Bus212:

Easy changes:

1. Python review goes to a quick recorded session. I have covered everything they should need in Bus215, so our Python stuff can be pretty much taken to almost zero.
2. Unsupervised learning: Eliminate completely. I don’t really do anything that would be different from Bus212a approaches here. Very much cookbook K-means clustering. Happy to drop this as long as Bus212a covers some of these important topics. (maybe also some lightweight PCA)
3. I have added recently some lectures on what I call data augmentation/generation. I include in this undersampling/oversampling, but most importantly SMOTE. I think these topics really belong mostly in 212, and ready to drop, but I find them super interesting/important. Have lots of debates with people on this. Obviously, the cutting edge of this is in generative network models, but I’m not sure I would ever have time for that, and also sort of too hard.

Harder changes:

1. Modeling preliminaries : Intro to some notation and measures. Shrink and speed up.
2. Overfitting/generalization/cross-validation. Shrink, but keep a lot of this. This course really needs to be a stress point for understanding how to test models and avoid overfitting. Then move right into regularization, ridge/lasso (shorter) and bias/variance tradeoffs. Operate under the assumption they know about these models.
3. Classification: Not sure what to do with this section. I cover a lot here in terms of various performance measures, confusion matrices, precision, recall, and ROC curves. These all need to be covered somewhere. Could move to 212, but we can talk on this. I need to talk about classification since much of ML is about this. (I think a majority.)
4. Data preprocessing: I do have a kind of strange section on preprocessing. This is a topic that belongs more in 212 in general, but I need to talk about scaling and hyperparameter searches, and how to implement in Python/Scikit-learn (pipelines, gridSearch). I probably should keep some of this, but shorten.

Little changes:

1. Training : I do a bunch on training. Batch gradient, stochastic gradients, minibatch, and early stopping. Would like to keep these with some new things I might add on convexity and why local mins are not much of a problem a lot of the time.
2. Support vectors: Keep all.
3. Nonlinear regression models (mostly kernel ridge): Keep, close to (2), and I don’t think these were ever in 212.
4. Trees/Ensemble learning/bagging/Random forests/Boosted trees, bootstrapping: I would like to keep all of this. I think I offer something on this. (I was bagging models before bagging was a thing, so I have lots of thoughts, and code, and examples in this area.) I also think this is one of the most important things in ML. If we both cover this, I don’t think a problem.
5. Neural networks: Will actually get to my notes on this. Will implement with Keras and Tensorflow on this. Only thing I worry about is that these models may not run well on their laptops, or that they work that well on some of my example data sets.
6. Dynamic networks: LSTM/GRU networks. Have these notes ready too. Have some examples, but maybe not so great. Should still do this. However, the computer implementation for these is not all that easy.

Additional topics:

1. Reinforcement learning: I never even have thought about how to cover this important topic. Also, not sure how generally useful it is for what MSBA’s do in the real world. I do know a bunch about this. Just not sure if I can bring it down to the masters level in a reasonable way, but could try. This could be tricky for many of our students.
2. Causality: Some sort of introductory material on causality might be good.
3. Thinking of some other “light weight” material to add, but no ideas yet