**Lesson 10 Code Talk**

What are the problems that require unsupervised learning that you encountered?

Have you come across a case of when you have to utilize a high dimensional technique or unsupervised learning to solve a problem?  If not can you think of a case where you might?

While working on the thermodynamics’ side for commercial turbofan engines very quickly we realized that if we wanted to have any prediction model in production looking at “on-wing” and/or “test-cell” data providing real time predictions we needed to somehow reduce the number of features entering the model while maintaining certain degree of control and accuracy. As you can imagine for turbofan engines we have many sensors while flying, plus extra measurements from extra sensors while testing the engine on the ground (“shop-test cell”), which in turn might lead to thousands of inputs for every single data point. During this time many different methodologies were implemented to try to solve this problem of dimensionality reduction and one of those techniques implemented in some of the models (especially for those with a high degree of complexity) was PCA to find a lesser set of features that could provide a comparable (or even better) performance, while improving the time response. Using this methodology we were able to find great insights on the contribution of the kind of sensors we were using, but ultimately for most of our models this technique was just the first step for feature selection and not the final answer to our high dimensionality issue. This mainly because our customers needed not only the prediction but a model that was easier to interpret.