**Arun Nair**

**Drill: Unsupervised or supervised?**

 Estimated Time: 15 minutes

*For each of these scenarios decide if you could use supervised or unsupervised techniques, or both!*

1. Define the likelihood that an individual will contract a specific disease
   1. We can use logistic regression to predict as benign or malignant.
   2. Bayes rule for conditional (correct prediction given that patient is sick/healthy).
2. Translate a set of images into variables for modeling
   1. This is part of unsupervised. Computer vision and image processing rely on pixels as vectors to learn the dataset.
   2. Instead of using image as an array of numbers, we can use convolutional neural networks to break the image into a number for tiles. The machine learning algorithm predicts each tile. The tiles are predicted in parallel.
3. An ecommerce company wants to identify power users
   1. Recommendation systems (supervised learning). We can classify our users using random forest, logistic regression or other supervised algorithms.
4. That same company wants to see shopping patterns in users
   1. NLP is unsupervised and can be used to see which user reviews affect our ratings or recommendation metrics (CTR, conversion rate, bounce rate etc.)
5. You want to reduce the number of variables inputting into your random forest model
   1. Tune your hyperparameters.
   2. Feature engineering.
   3. Max Depth.