

NOTES:

- The reason we need feature column is so that our model can get a list of each column and the vocabulary they contain
 - EX: 'Survived' column will contain the vocabulary of "Y" or "N"
- Training a model can take a lot of ram if we have a lot of data
 - So normally regular computers will not be able to load a huge amount of data so we will have to feed the data in batches.
 - The reason we don't feed one at a time is because it can take a lot of time we can increase our speed drastically if we just load in batches
- **EPOCH**
 - From my understanding an epoch is just the same batch of data but fed into the model in different orders.
 - The reason we need this is because we want our model to look at the same set of data but from different point of views (hence changing the order) so that it can start recognizing patterns within the dataset
- **OVERFITTING**
 - Overfitting is when we pass same data to our model too many times that it starts to just memorize the data, so it becomes really good at classifying a certain batch of data but when we pass it a new batch of data it performs horribly.

- To prevent this issue we start with a small number of epochs and then increment as we need

- **INPUT FUNCTION**

- The input function decides how your data set will be broken into epochs/batches that are going to be fed into your model
- The input function also encodes out data into `tf.data.object` objects because tensor flow requires this object type for their models