**Predicting Sentiments of Tweets using Azure Machine Learning Studio**

In this tutorial we will be building a sentiment classification model to predict, from the words used in a given tweet, whether the tweet will be positive or negative.

**What is sentiment?**

Sentiment in respect text is the emotional character, tone, or “feel” of the given message or text. For example, was this an angry Tweet? Was it happy, sad, solemn, or slanderous? Now defining the exact rules for sentiment is not so easy and in the supervised learning context, it will humans who have to read the text and classify as such.

Since this is an introductory tutorial, we will be building a binary classification model for sentiment. Wherein a tweet can only be “positive” or “negative” sentiment with no option of any other sentiment such as “neutral”. As such this model may not be practical for deployment in real world applications, but incremental development states that we all must start from a firm foundation.

**The Dataset**

We will be using a public dataset with 160,000 pre-classified tweets. There are two columns in the dataset, the text of the tweet, and a sentinment, “0” being negative and “4” being positive. The classification itself was scored by human judges (my guess is an army of linguistic grad students) where a judge would read the tweet and give it a score between 0 and 4; with 4 being highly positive and 0 being highly negative. This is a sampled down dataset where “1”, “2”, and “3” sentiments have been filtered out and between the remaining “1” and “4”

An even representation of both sentinments.