What are the differences between supervised, semi-supervised, and unsupervised text classification techniques?

**Supervised Text Classification**

Supervised techniques have labeled training datasets where each document / text sample is associated with a known category. A model learns from these training datasets and maps out the text features to the corresponding labels. These models include Naïve Bayes, Logistic Regression, and Support Vector Machines (SVMs).

**Semi-supervised Text Classification**

Semi-supervised Text Classification involves a smaller labelled data set with a larger pool of unlabeled data to improve a model performance. This can happen when labels are scarce, or it is simply too expensive to obtain labels.

**Unsupervised Text Classification**

Unsupervised classification does not have any labelled data. What a model does in the case of unsupervised text classification is to cluster similar text samples based on their structure. This is helpful in showing some common themes / topics. These models include K-Means and Hierarchical clustering.