

## **Assignment #2: Deliver on Monday - Tuesday**

**Given the following ham/spam dataset, you are required to design a naïve Bayes classifier model that classify efficiently any upcoming email into either ham /spam.**

**You are required to use the appropriate naïve Bayes classifier. Moreover, you need to use:**

- 1 -  $\chi^2$  test (Chi-squared) to check the probability of each word with each class category**
- 2- TFIDF to check the probability of each word with each class category**
- 3- Use laplacian smoothing in (1 & 2)**

**Which one is better, why?**