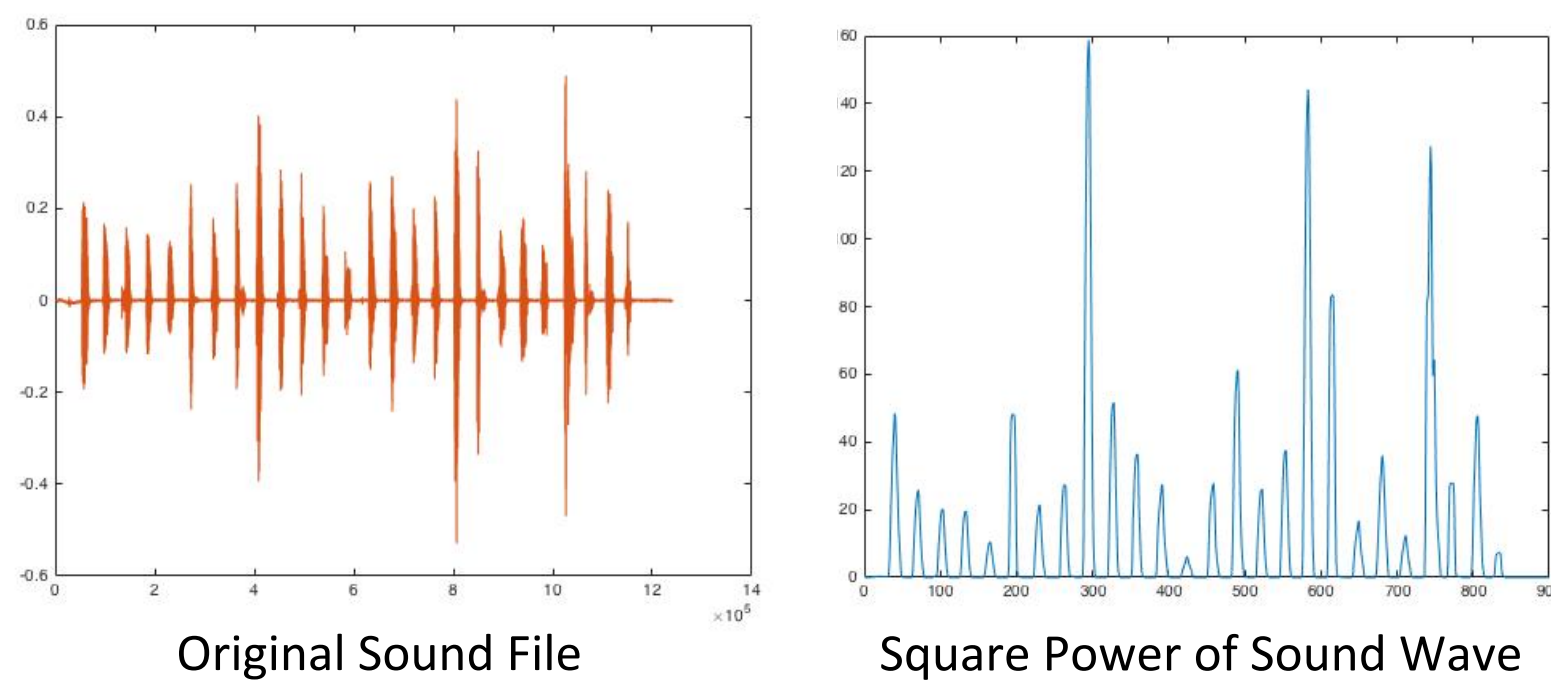
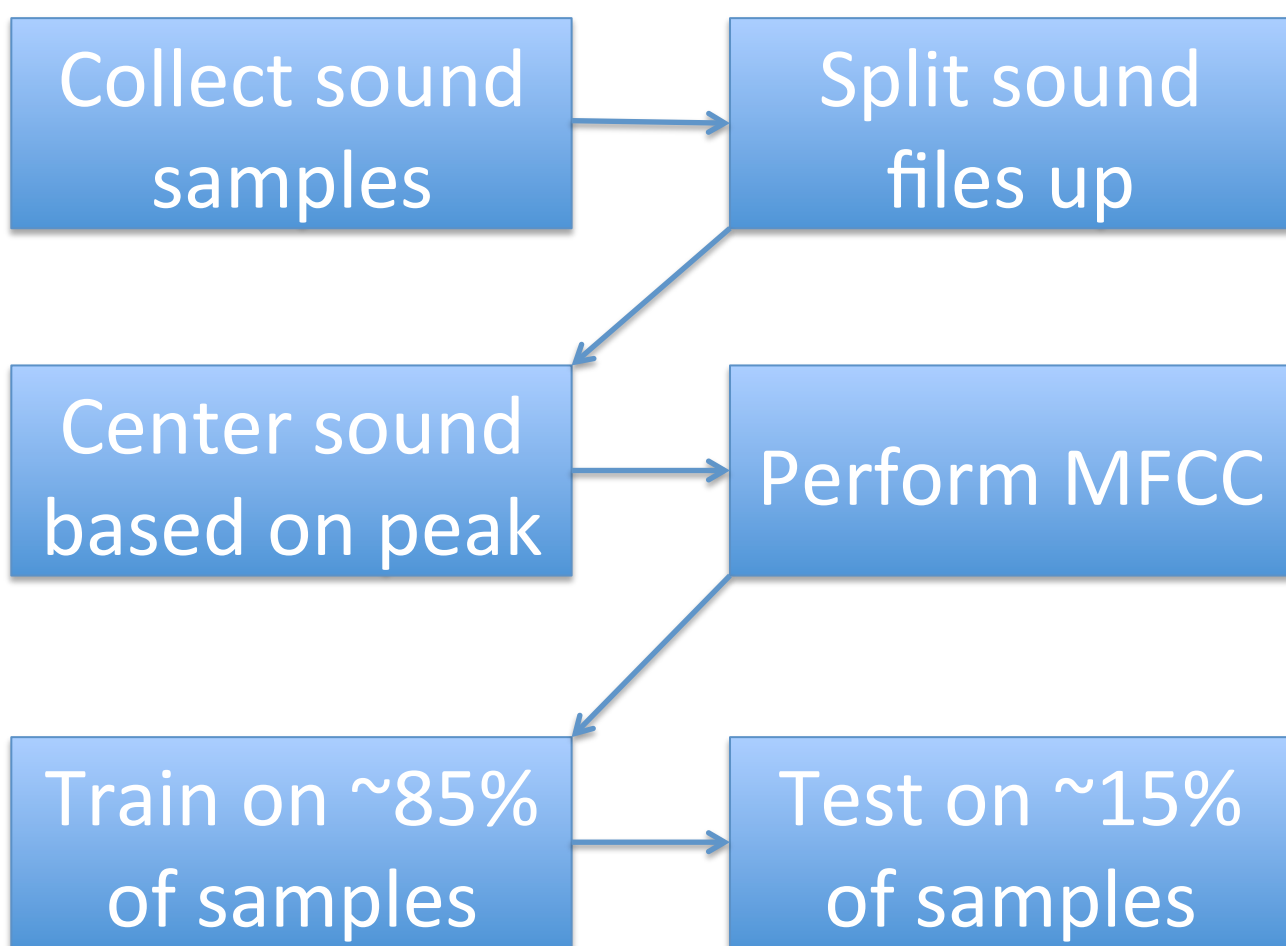


# Spoken Character Recognition

Yuki Inoue, Jason Liu, Allan Jiang

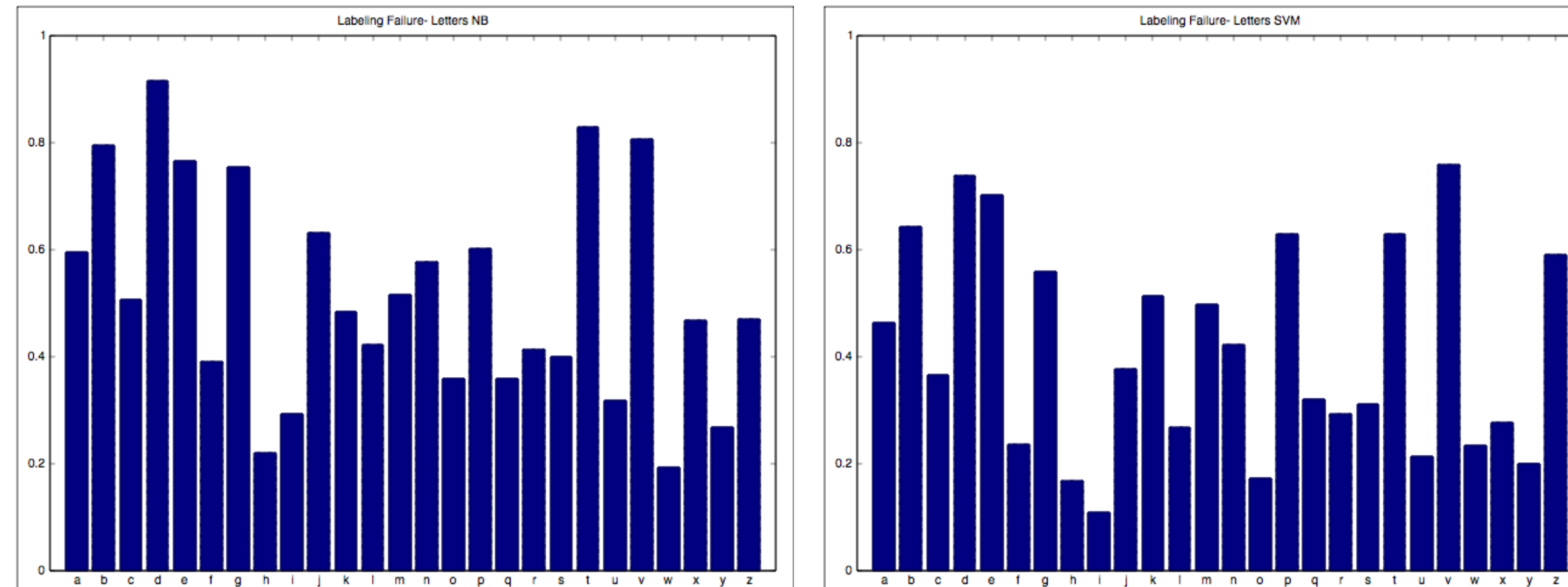
Motivation: Current research in voice recognition software emphasizes **word-by-word** data. A character recognition software would be particularly applicable in situations such as **randomized passwords**. It could also act as an **accessibility tool** for those who cannot type or are blind.



Mel-frequency cepstrum coefficients (MFCC) were used to characterize the soundbites. MFCC ignores individual differences in how the words are pronounced, but instead focuses on the short term power bands of the sound inputs.

## Letters

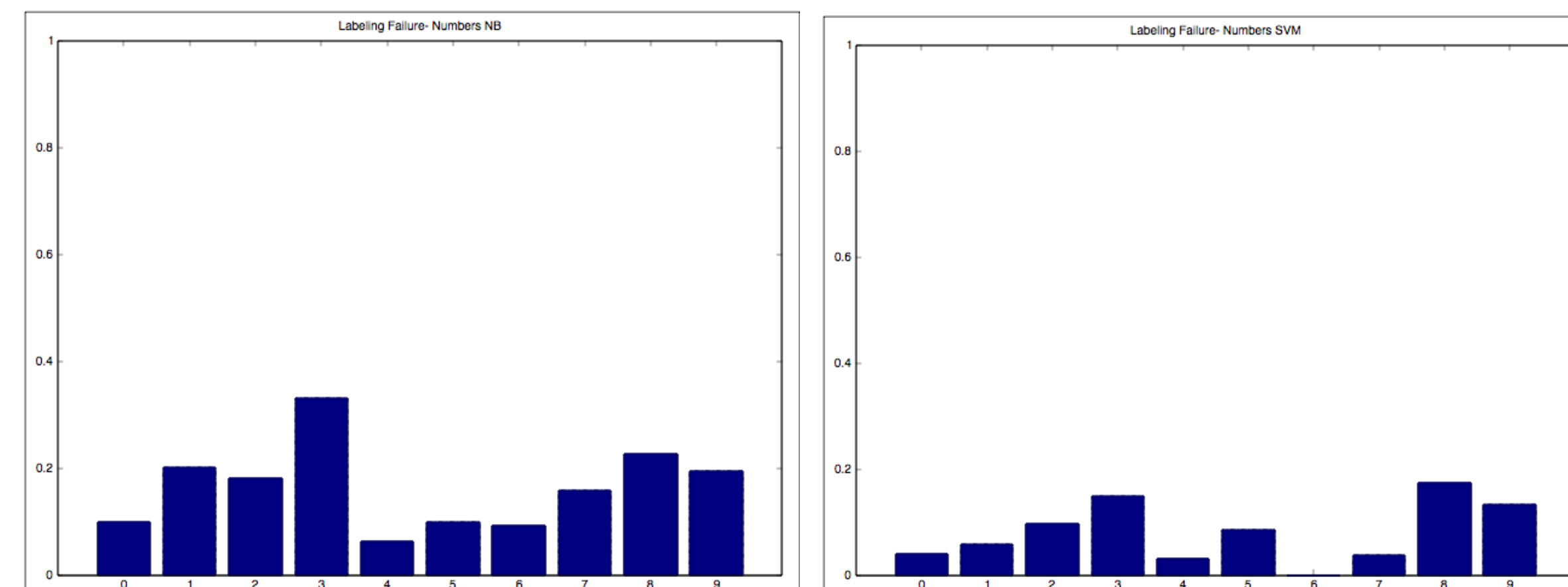
2000 Individual Samples



NB: 0.51322

## Numbers

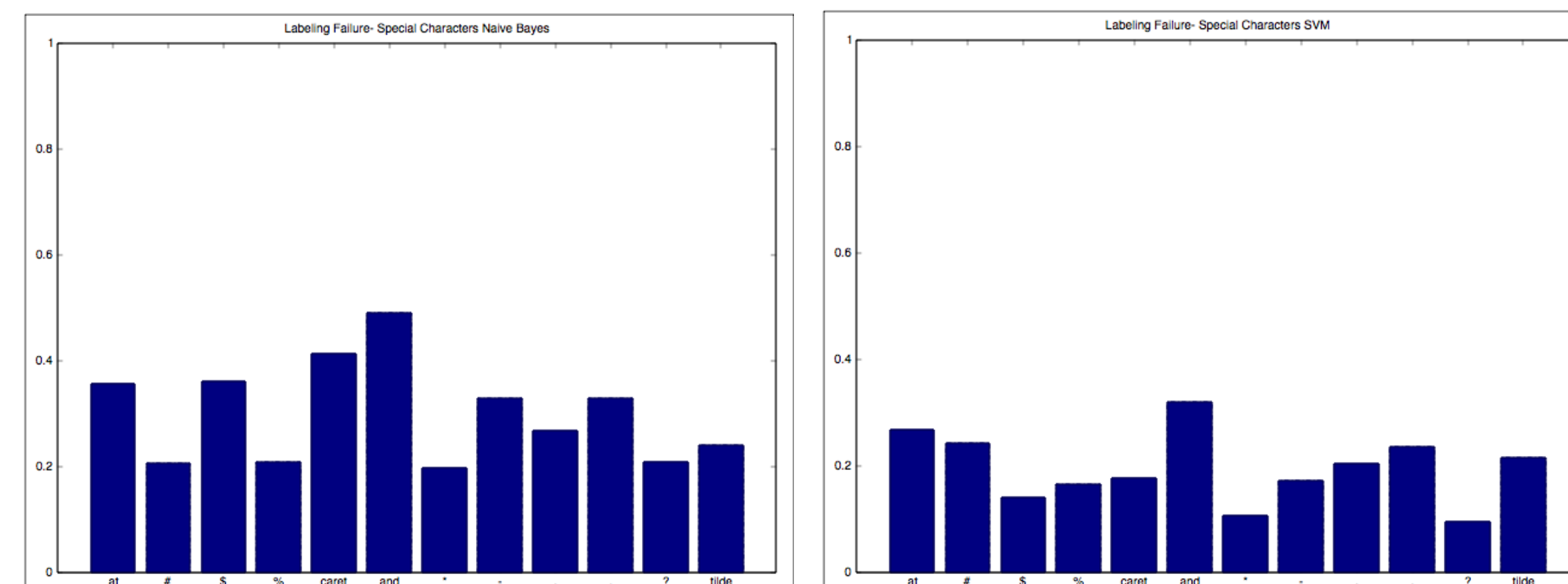
660 Individual Samples



NB- 0.16453

## Special Characters

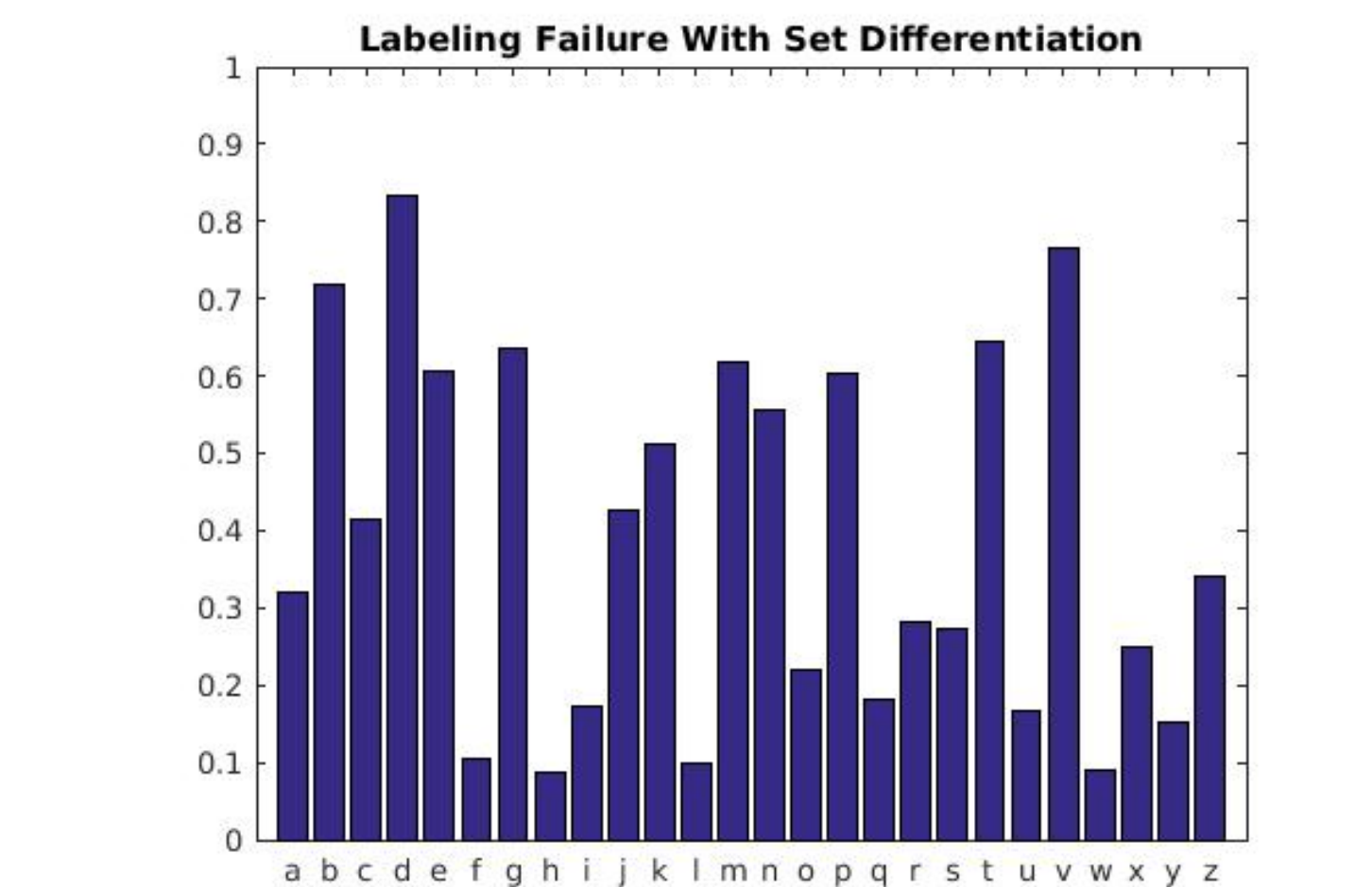
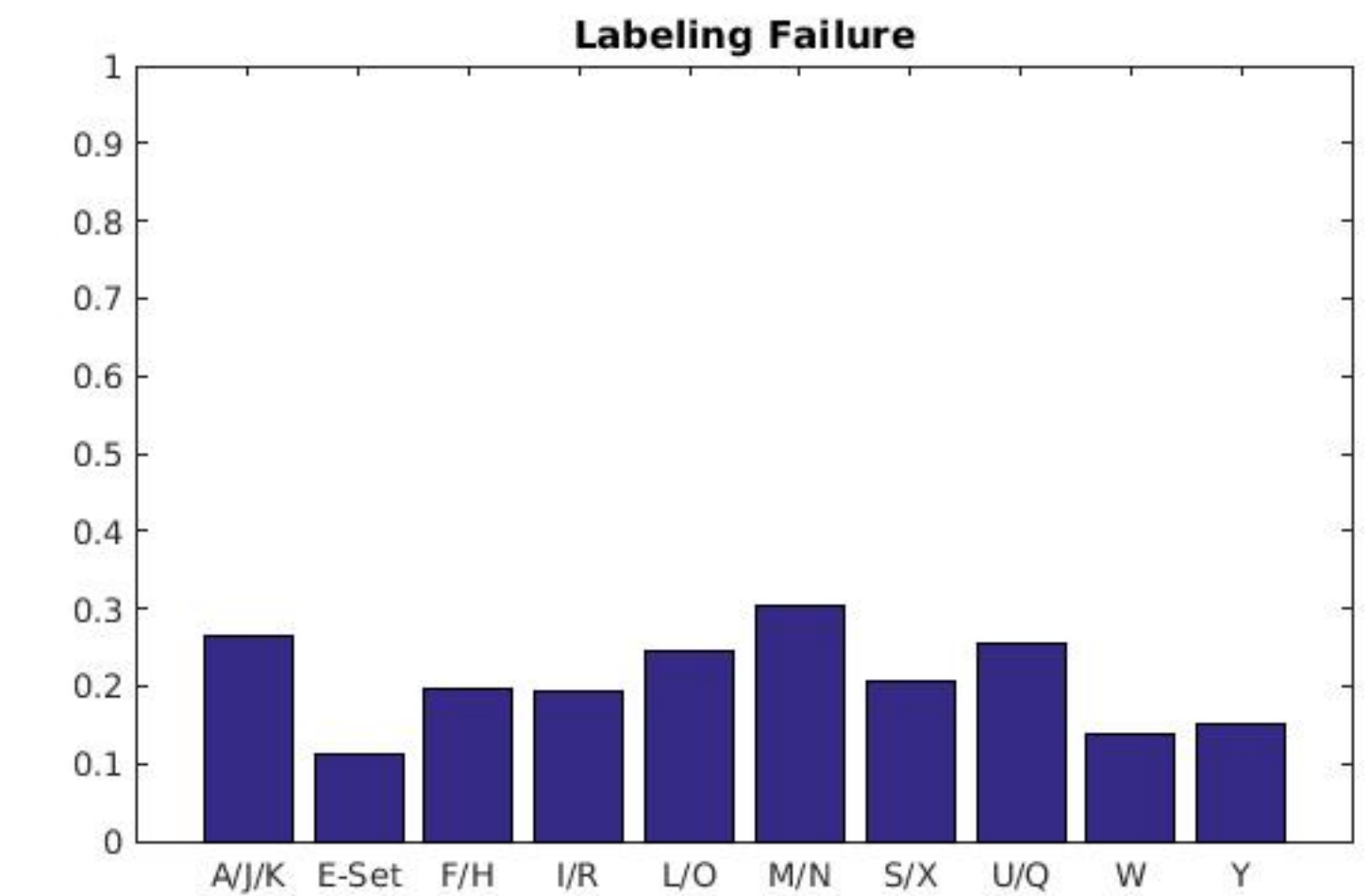
330 Individual Samples



NB- 0.28048

SVM- 0.18473

The E-Set Problem - It is difficult to distinguish between the 'E' letters, as their differences may not be captured by MFCC. Below is the accuracy with which we can classify certain letters as part of E-set, as well as others. Also below, is the classification failure given set differentiation



	N B*	SVM*	Log. Reg.
Letters	0.51322	0.41051	0.41928
Numbers	0.16453	0.08318	0.09245
Special	0.28048	0.18473	0.19536

\*Specifically, we used a Gaussian Naïve Bayes and a Support Vector Machine with a linear kernel