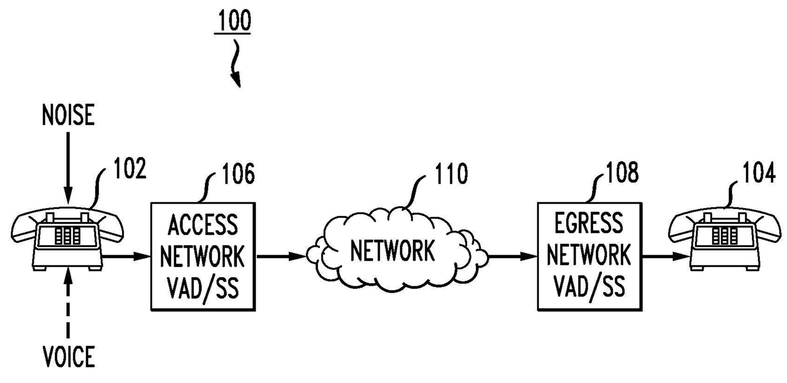
University of BRISTOL

Sam Phippen
supervised by Dr. Rafal Bogacz

Voice activity detection against keyboard noise

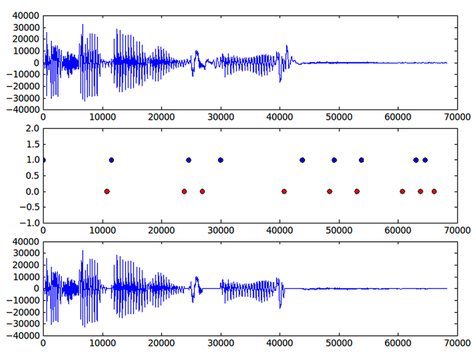
Basic question: does this audio contain voice or not?

Used in phones, VOIP and voice recognition systems

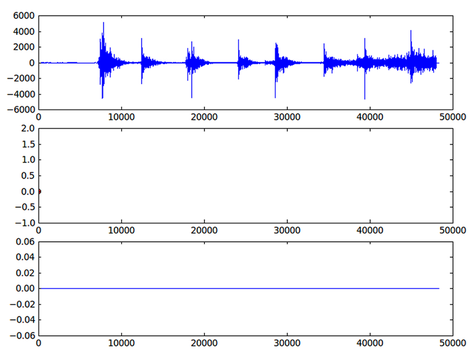


So far I have built a classifier to differentiate noise and voice, trained heavily against keyboard noise. Most VAD systems do not use advanced machine learning techniques or train against keyboard noise. The goal of this project is to build a system robust to users typing whilst talking in voip conversations.

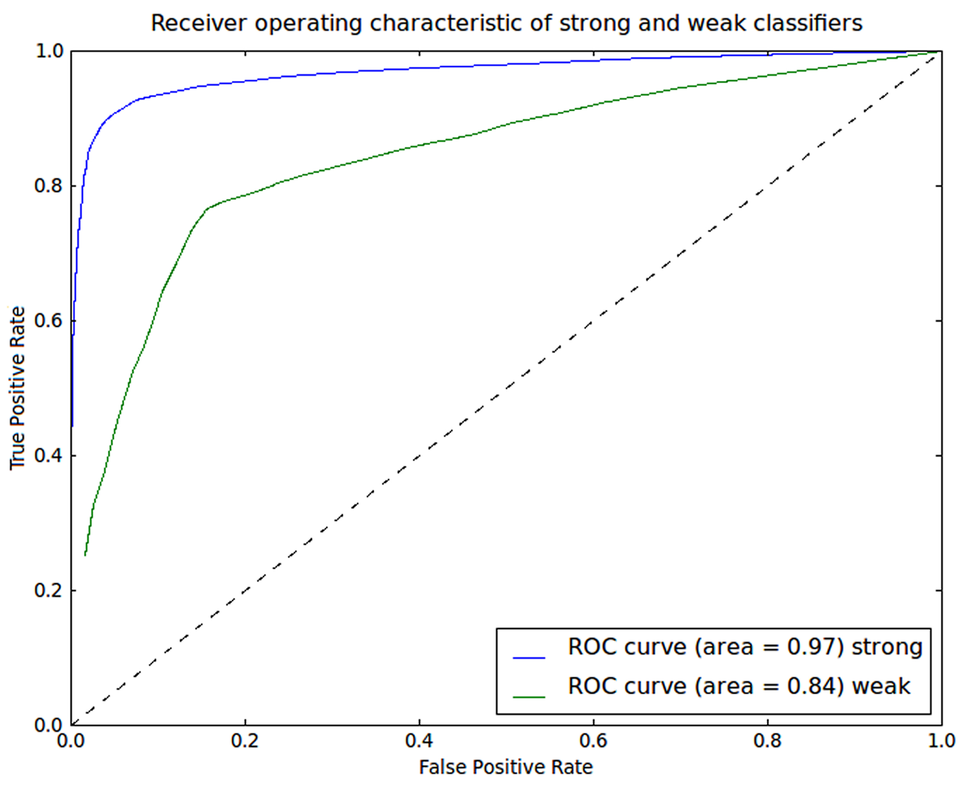
voice example



noise example



My classifier is essentially complete
it gets 90% accuracy on cross validation of the large (approximately 1 hour of audio) training set. Below is a ROC curve showing the difference in accuracy between my classifier and a baseline classifier trained on the energy of each frame of audio.

A “noisy” classification is given. To improve this most existing systems use a smoothing or hangover logic. This is my next aim.

For the business plan I am targetting voip systems as good VAD technology gives a good bandwidth saving, by not transmitting noise.