

Per l'implementazione dell'algoritmo Voice Activity Detection ho applicato un'analisi dell'energia sotto le ipotesi che:

- I primi istanti di una traccia (200ms) siano solamente di rumore
- Il rumore sia una variabile aleatoria gaussiana, e quindi assuma con alta probabilità valori attorno alla sua media
- L'energia del rumore sia minore dell'energia del parlato.

Sotto queste ipotesi ho implementato in c++ un programma che implementi il flow chart riportato sopra.

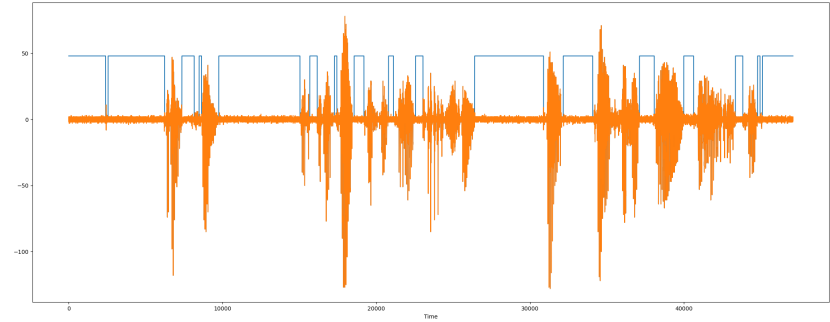
Ho scelto come soglia per catalogare i pacchetti due volte il valore dell'energia media calcolata nei primi 200ms.

Ho poi verificato sperimentalmente che questa scelta fornisca una buona catalogazione dei pacchetti.

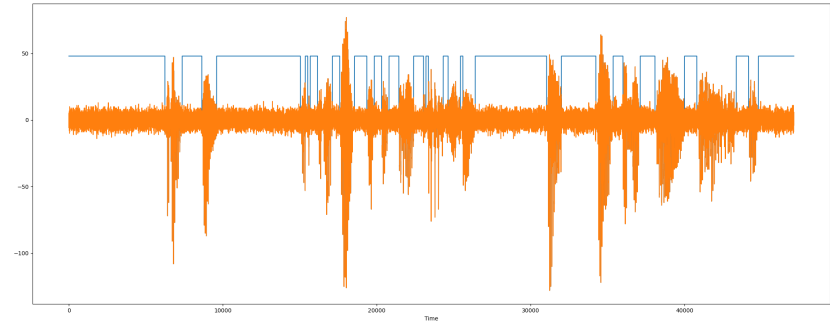
Nei grafici sotto riportati troviamo in arancione l'energia del segnale sorgente.

L'area sottesa dai rettangoli blu viene catalogata dalla mia implementazione come rumore, mentre quella compresa tra i rettangoli viene catalogata come parlato.

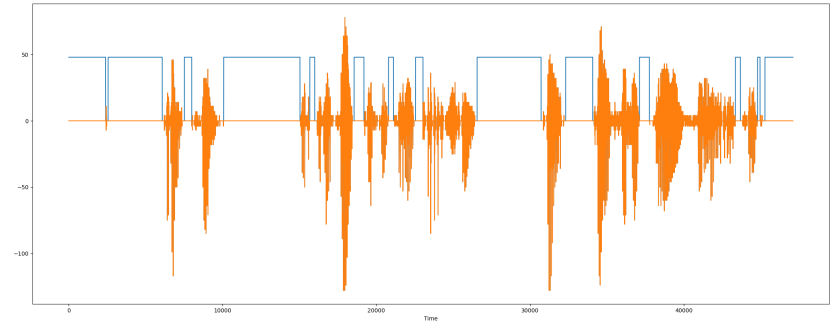
File inputaudio1.data



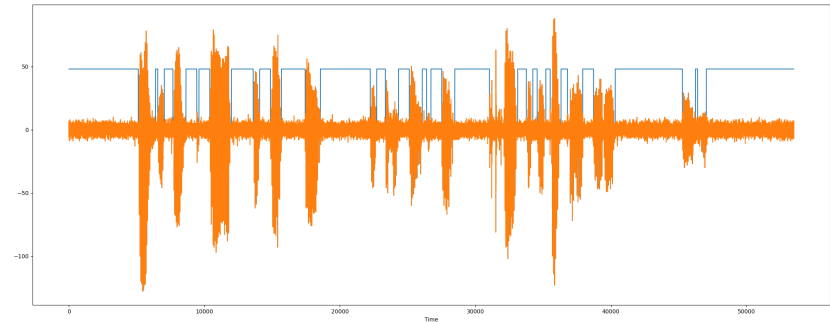
File inputaudio2.data



File inputaudio3.data



File inputaudio4.data



File inputaudio5.data

