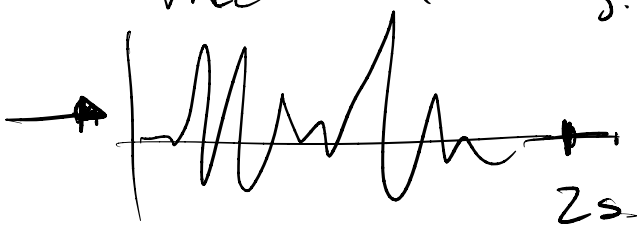
Señal de voz

machine learning.



Ruido.

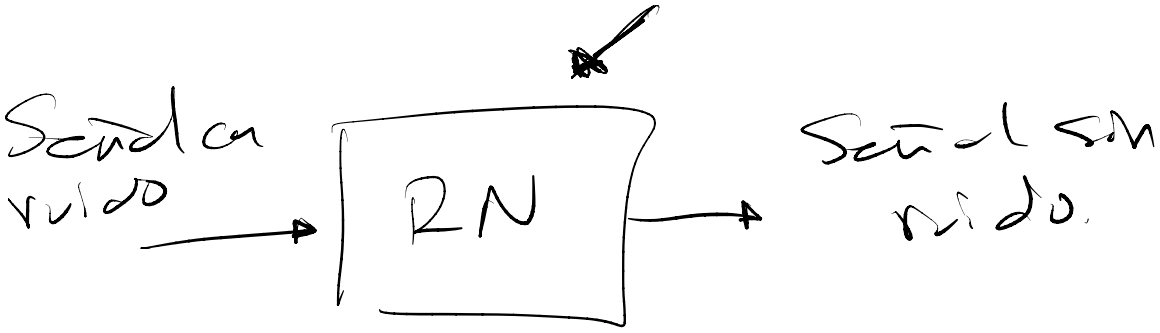
10 palabras.

Por cada palabra agregamos 20 veces

escojan (4) ruidos

- tráfico
- plastico
- lluvia.

} misma duración.



Generar Datos

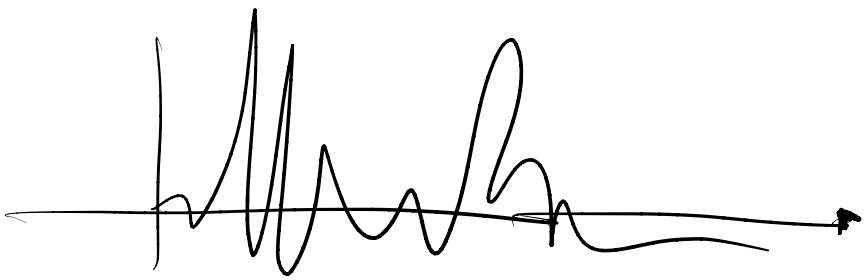
→ entradas →

{ Combinaciones entre señal sin ruido (palabra) y el ruido

2s → fs = 32647

↑ ↑ 32000 muestras por segundo

la señal tiene 64000 puntos.



P_1 y R_1

$S_{entrada\ 1} = 0.5P_1 + 0.5R_1$

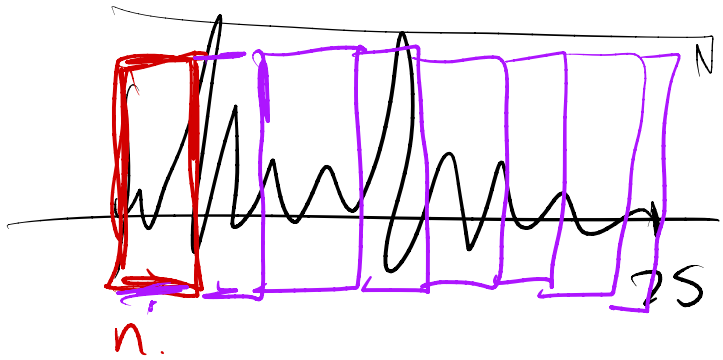
$S_{entrada\ 2} = 0.3P_1 + 0.7R_1$

$S_{entrada\ 3} = 0.7P_1 + 0.3R_1$

} misma salida

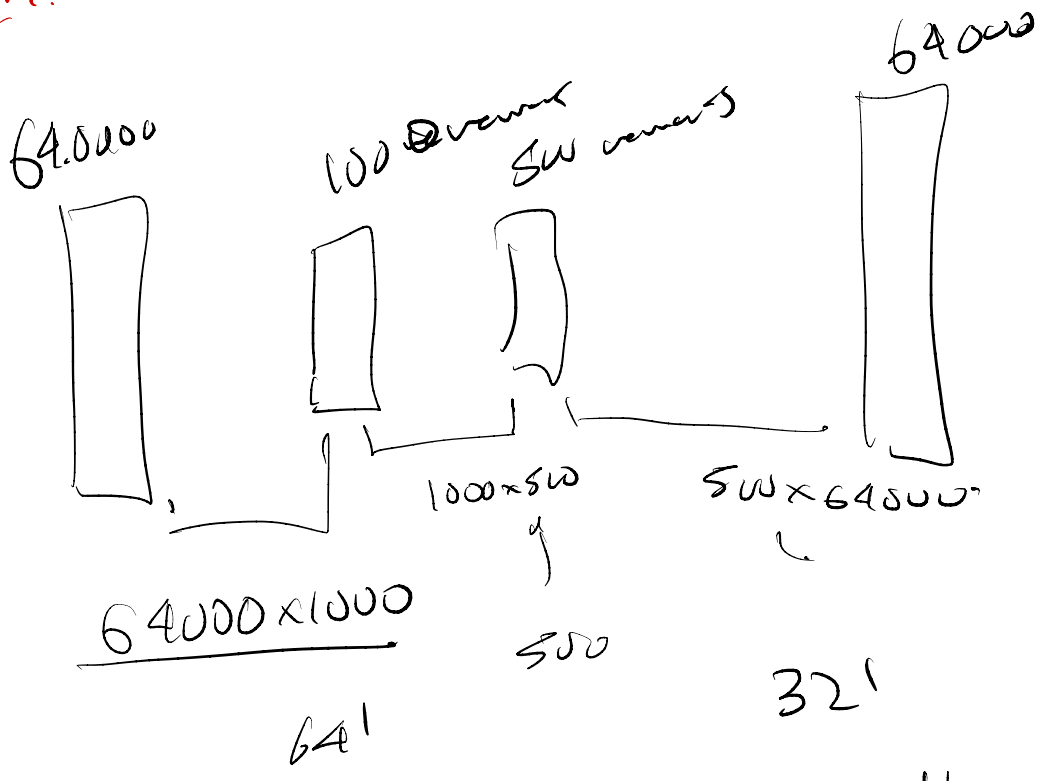
→ Salida

Palabra normal sin ruido.



32 kH

1s → 64000 entradas.



≈ 100 millones de puntos.

20 ms de la señal → 32kHz.

32000 → 1s.

n → 0.2s.

n = 32000 * 0.2

= 6400 puntos.

read wav. .wav.

y = read_wav('______')
{nombre del archivo

y → arreglo de numpy de 64000 puntos.

Primeros 20 ms.

$y_1 = y[0:6399]$; →

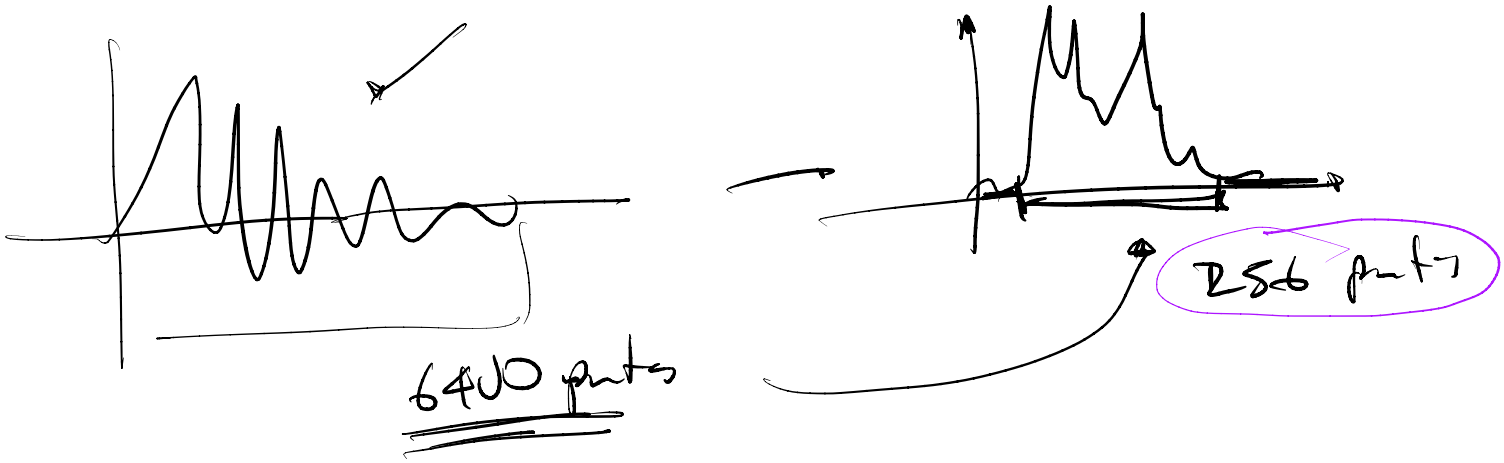
$y_2 = y[6400 : 6400 \times 2 - 1]$.

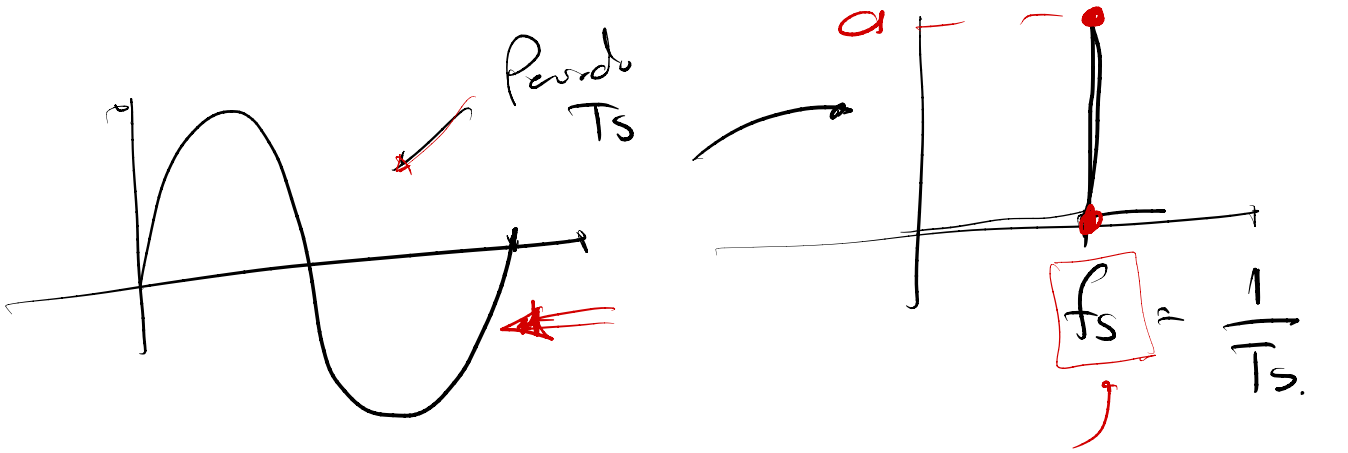
$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \end{bmatrix}$$

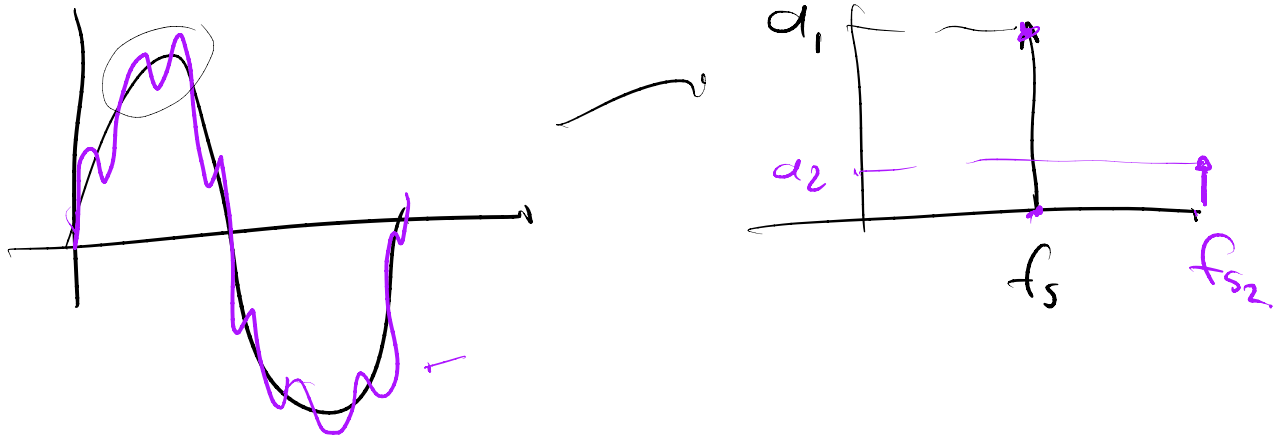
$m \times 6400$ chunks

$m \times 256$.

La densidad espectral de potencia!







Spectrogram python
densidad espectral de Potencia.

Red Neural que
modelo secuential
add_layer (Dense.