



**Data Science  
Academy**

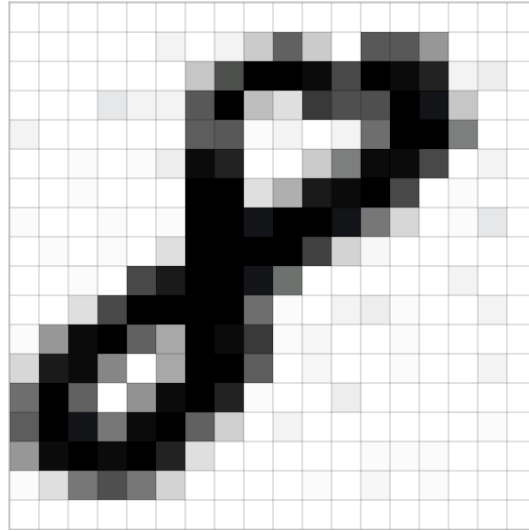
[www.datascienceacademy.com.br](http://www.datascienceacademy.com.br)

Processamento de Linguagem Natural

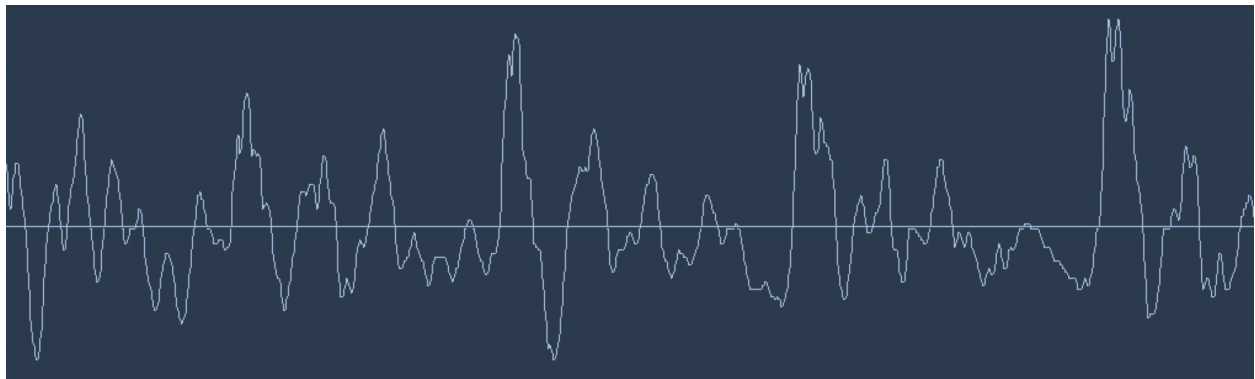
Transformando Som em Bits

O primeiro passo no reconhecimento de fala é óbvio - precisamos alimentar as ondas sonoras em um computador.

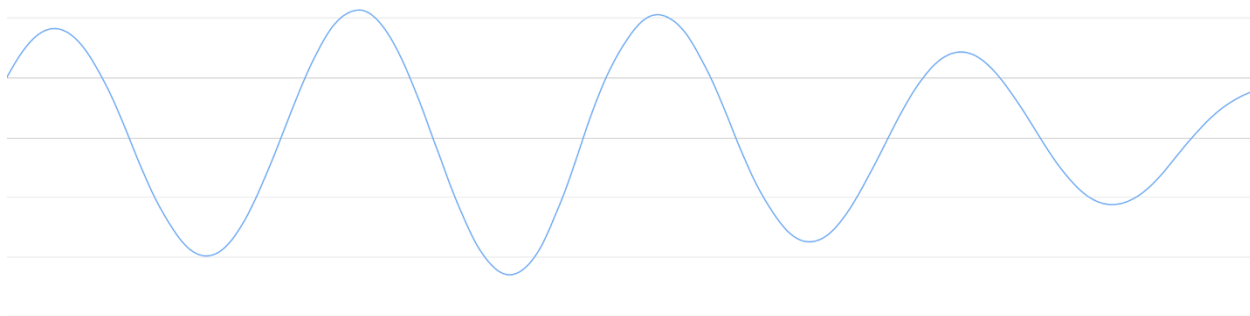
No curso de Visão Computacional aprendemos como receber uma imagem e tratá-la como uma matriz de números para que possamos alimentar diretamente uma rede neural para reconhecimento de imagem:



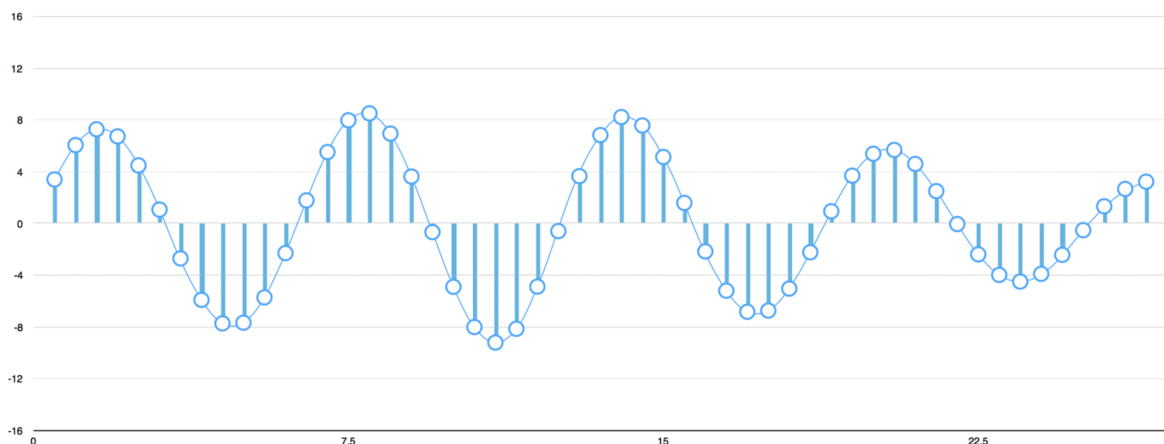
Mas o som é transmitido como ondas. Como transformamos as ondas sonoras em números? Considere esta imagem que representa o som da palavra "Olá":



As ondas sonoras são unidimensionais. A cada momento, eles têm um valor único baseado na altura da onda. Vamos dar zoom em uma pequena parte da onda sonora e dar uma olhada:



Para transformar essa onda sonora em números, registramos apenas a altura da onda em pontos igualmente espaçados:



Isso é chamado de amostragem. Estamos fazendo uma leitura milhares de vezes por segundo e gravando um número que representa a altura da onda sonora naquele momento. Isso é basicamente um arquivo de áudio .wav descompactado.

O áudio “Qualidade de CD” é amostrado a 44.1khz (44.100 leituras por segundo). Mas para o reconhecimento de fala, uma taxa de amostragem de 16khz (16.000 amostras por segundo) é suficiente para cobrir a faixa de frequência da fala humana.

Vamos experimentar a nossa onda sonora “Olá” 16.000 vezes por segundo. Aqui estão as primeiras 100 amostras:

```
[ -1274, -1252, -1160, -986, -792, -692, -614, -429, -286, -134, -57, -41, -169, -456, -450, -541, -761, -1067, -1231, -1047, -952, -645, -489, -448, -397, -212, 193, 114, -17, -110, 128, 261, 198, 390, 461, 772, 948, 1451, 1974, 2624, 3793, 4968, 5939, 6057, 6581, 7302, 7640, 7223, 6119, 5461, 4820, 4353, 3611, 2740, 2004, 1349, 1178, 1085, 901, 301, -262, -499, -488, -707, -1406, -1997, -2377, -2494, -2605, -2675, -2627, -2500, -2148, -1648, -970, -364, 13, 260, 494, 788, 1011, 938, 717, 507, 323, 324, 325, 350, 103, -113, 64, 176, 93, -249, -461, -606, -909, -1159, -1307, -1544]
```

Você não achou que seria fácil, não é mesmo?