

TIPOS DE SEÑALES

Los algoritmos de machine learning necesitan datos con un formato concreto: cierto número de observaciones, cada una de ellas con varias características.



Legs	Color	Weight	Ears
4	Gray	12000	2
0	Red	2979	0
2	B&W	51	2
4	Gray	12000	2
0	Y&W	.08	0
0	Y&W	.08	0
2	Brown	12	2

Cuando se crea un modelo predictivo, estas características son las variables de predicción: las entradas que el modelo usa para determinar la salida. Es decir, las características pueden ser cualquier medida que ayude a distinguir las distintas observaciones.

Features

Legs	Color	Weight	Ears
4	Gray	12000	2
0	Red	2979	0
2	B&W	51	2
4	Gray	12000	2
0	Y&W	.08	0
0	Y&W	.08	0
2	Brown	12	2

12000



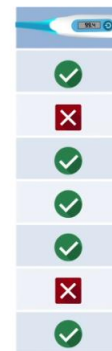
.08



Para muchos problemas, los datos tomarán esta forma naturalmente. Por ejemplo, predecir si un paciente pudiera tener una enfermedad concreta. Cada observación es un paciente. Las características serían datos biométricos y diagnósticos como: la edad, la presión arterial, si son fumadores, los resultados del análisis de sangre, etc.

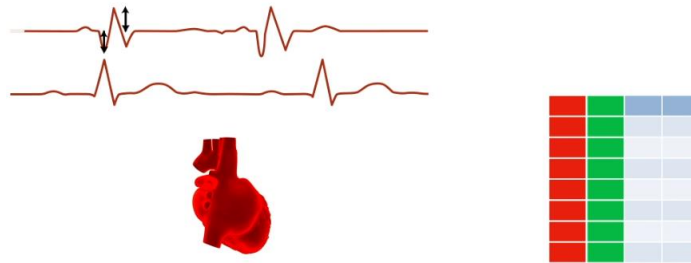
Features

Age	Sex	Smoker	Pulse
42	F	No	72

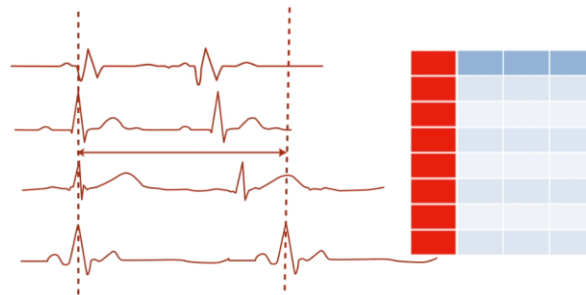


Pero las entradas no siempre tienen esa forma. ¿Qué pasa si el diagnóstico se basa en una imagen, como una radiografía o una resonancia? Ahora, la entrada es un arreglo de píxeles. ¿O si se basa en un sensor, como un electroencefalograma o un electrocardiograma? Ahora la entrada es un conjunto de señales a lo largo del tiempo.

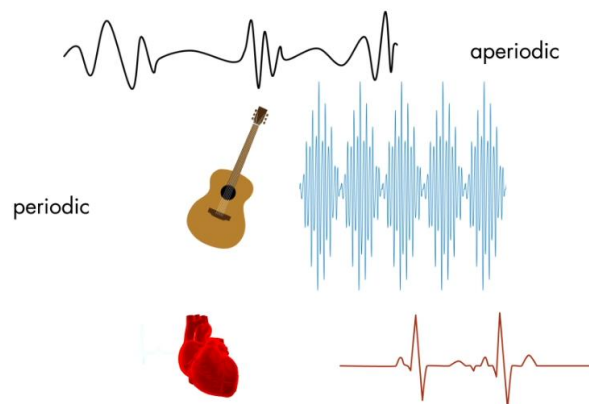
Para aplicar algoritmos de machine learning a una señal, antes es necesario transformar los datos sin procesar en un conjunto de características. ¿Cómo se obtienen estas características? Si ya conoce el sistema físico subyacente a las señales, puede usar este conocimiento para determinar las características distintivas que se deben buscar.



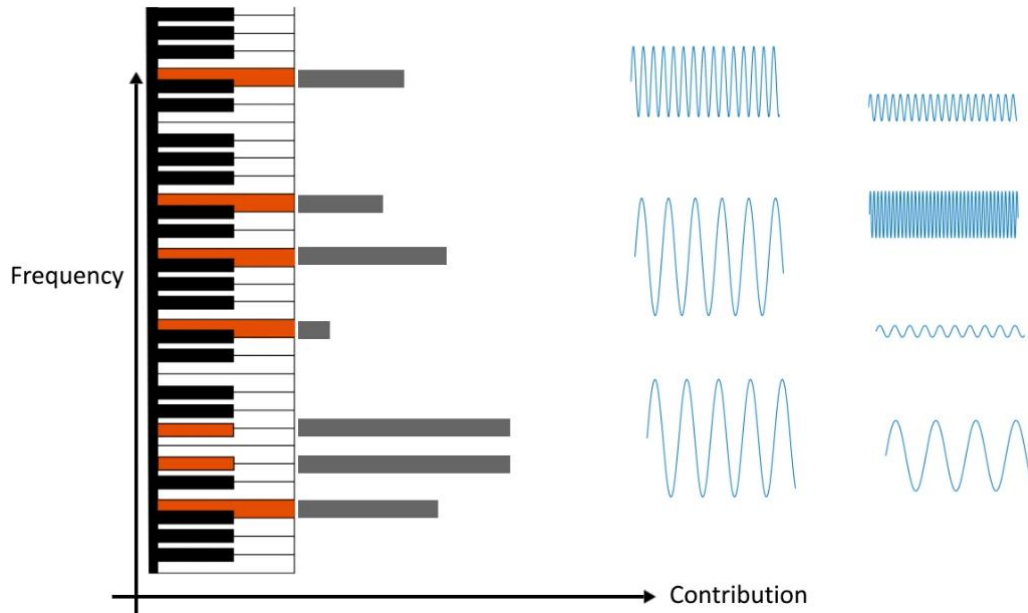
Pero, si no dispone de ese conocimiento, no pasa nada. Siempre es buena idea observar el conjunto de datos para ver si se le ocurren algunas ideas.



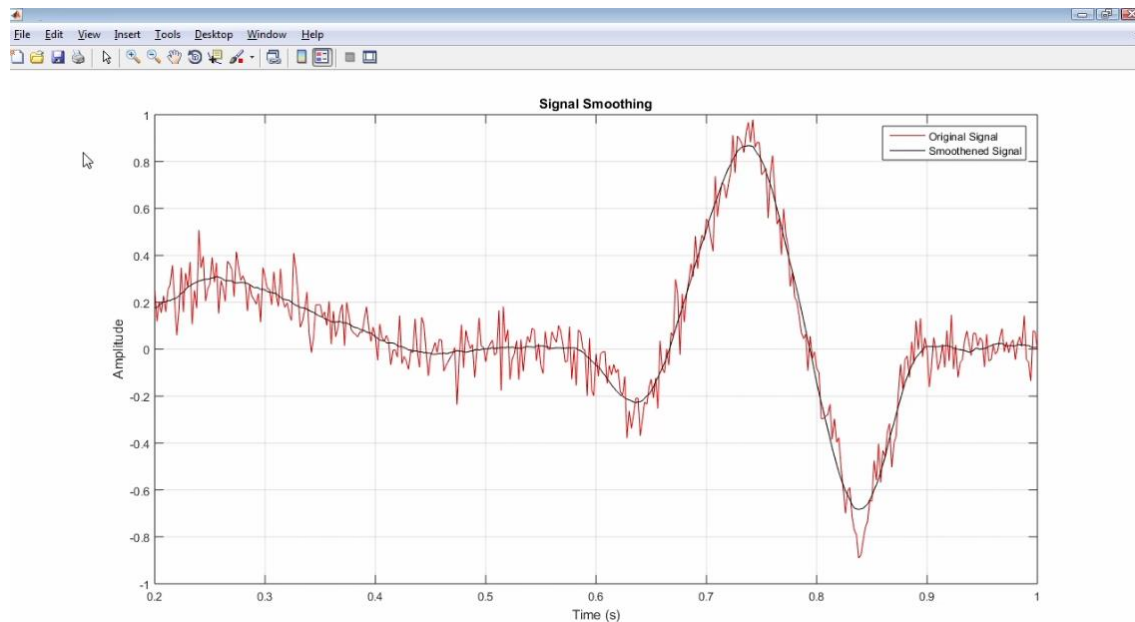
Una característica puede ser una simple medida estadística de toda la señal, una medida de las formas, como el número de máximos locales, o la anchura del pico, una medida de correlación o cualquier cantidad de otras medidas. Estas son características generales que podrían funcionar para cualquier tipo de señal. También hay técnicas específicas diseñadas para señales periódicas, como el sonido, la vibración, o el latido del corazón.



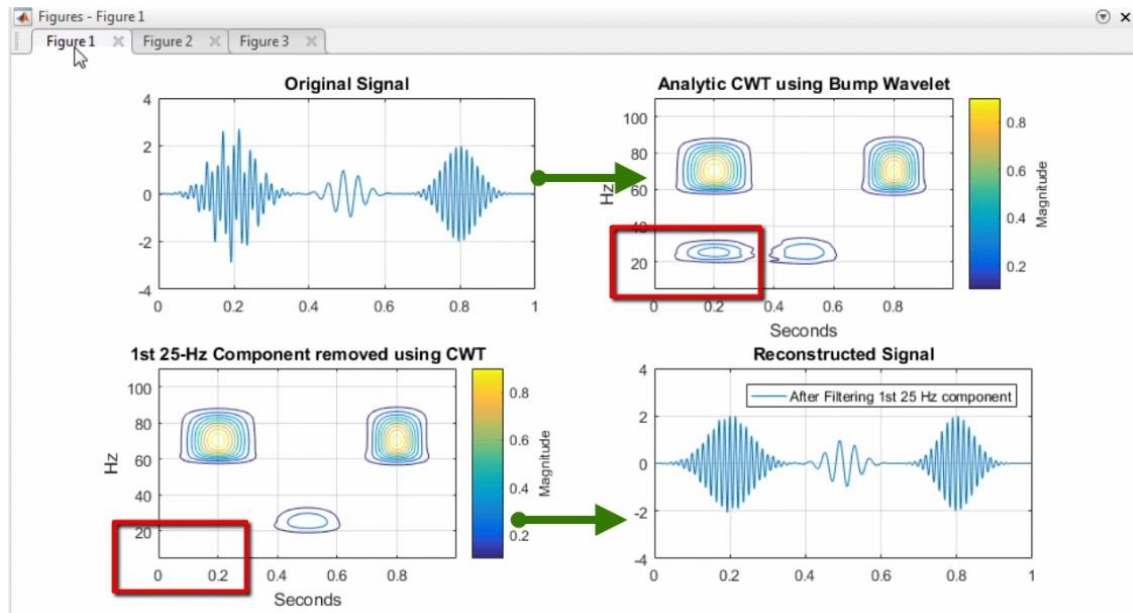
Normalmente, estas técnicas descomponen una señal en componentes simples y determinan en qué medida los distintos componentes contribuyen a la señal completa. Puede representar muchas señales periódicas, como el sonido, como combinaciones de senos y cosenos de distintas frecuencias.



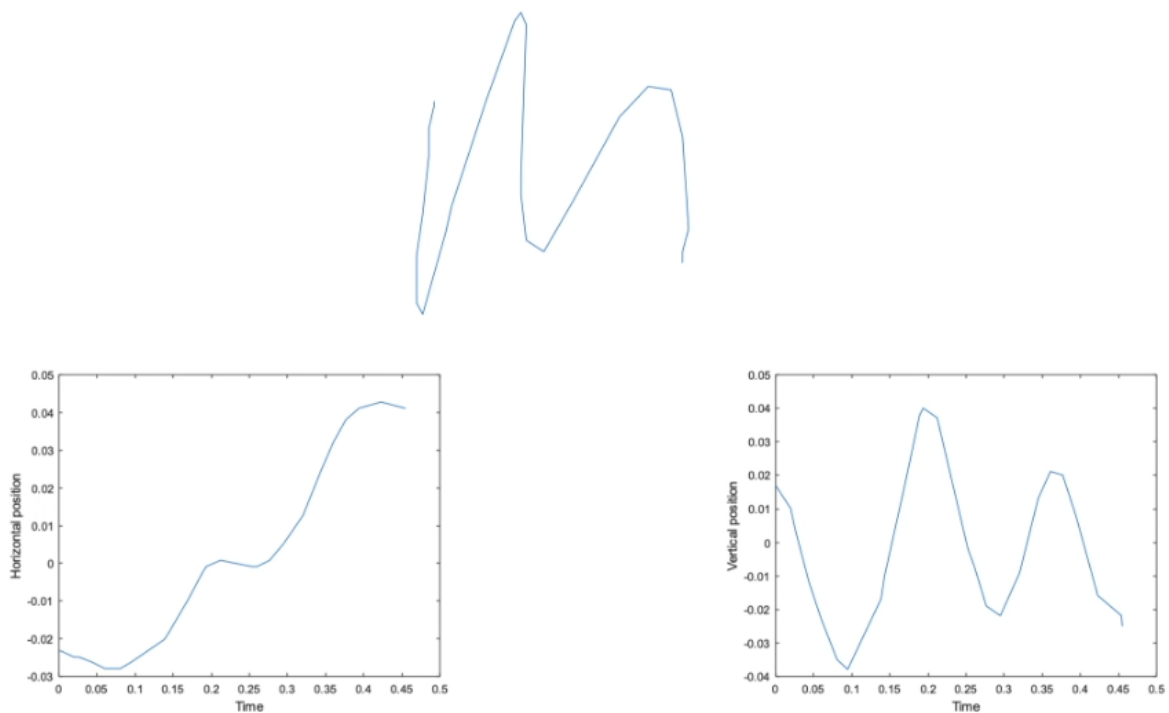
MATLAB y Signal Processing Toolbox tienen herramientas para realizar este tipo de análisis, así como para diseñar filtros para suavizar y limpiar las señales, además de otras técnicas de procesamiento de señales.



Otras señales periódicas, como el latido del corazón, se modelan mejor como combinaciones de pulsos cortos que se conocen como *wavelets*. Wavelet Toolbox amplía las capacidades de procesamiento de señales de MATLAB a este tipo de señales.



Las señales de escritura manual no son periódicas, así que no es necesario conocer estas técnicas para este curso.



En las siguientes secciones, aprenderá a utilizar las funciones estándar de MATLAB para extraer algunas características básicas de estas señales basadas en el tiempo. Después, utilizará la tabla de características para entrenar modelos de clasificación.