

Pontificia Universidad Católica de Chile
Escuela de Ingeniería
Departamento de Ciencia de la Computación



Sistemas Urbanos Inteligentes

Introducción a Redes Neuronales Parte 1

Hans Löbel

Dpto. Ingeniería de Transporte y Logística
Dpto. Ciencia de la Computación

Consideremos un problema de estimación de tiempo de viaje en un contexto urbano



- Hora del día
- Ubicación
- Inicio y fin (hora, ruta)
- Demanda según origen y hora
- Caracterización del instante
- Información reciente e histórica

¿Qué información podríamos/deberíamos utilizar?

Consideremos un problema de estimación de tiempo de viaje en un contexto urbano

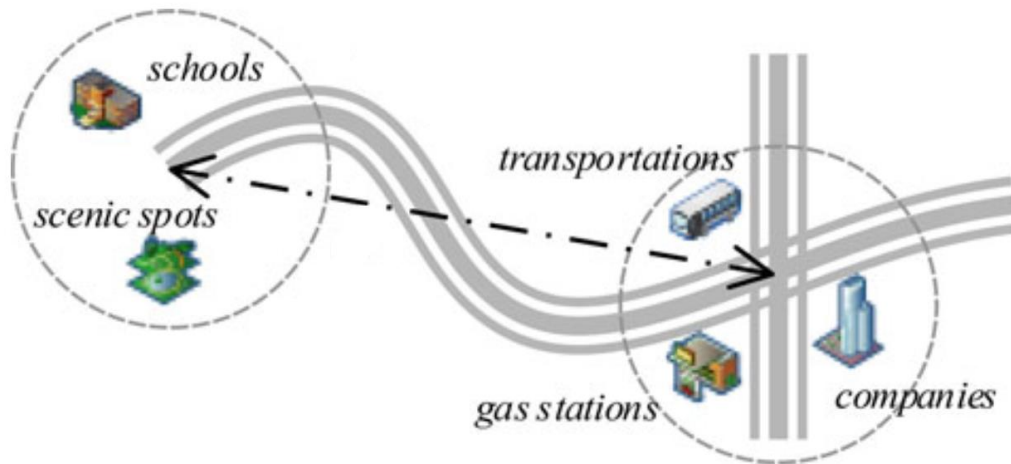


Figura tomada de Tang et al, "Deep Architecture for Citywide Travel Time Estimation Incorporating Contextual Information", JITS (2019)

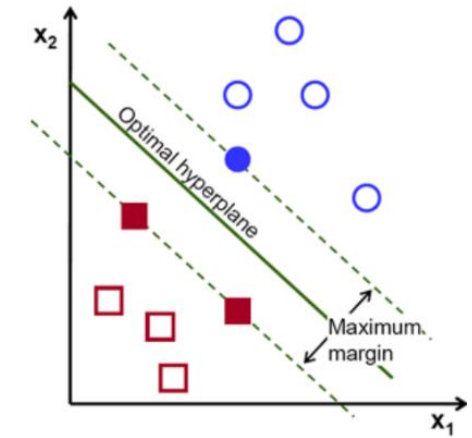
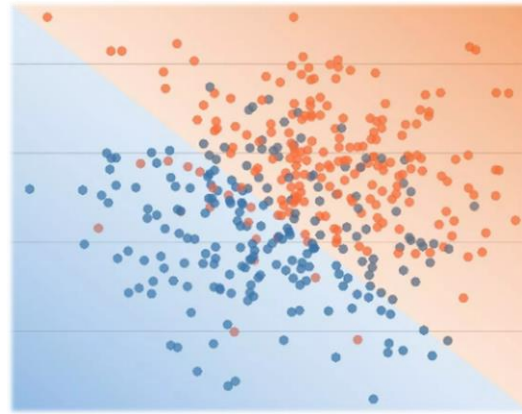
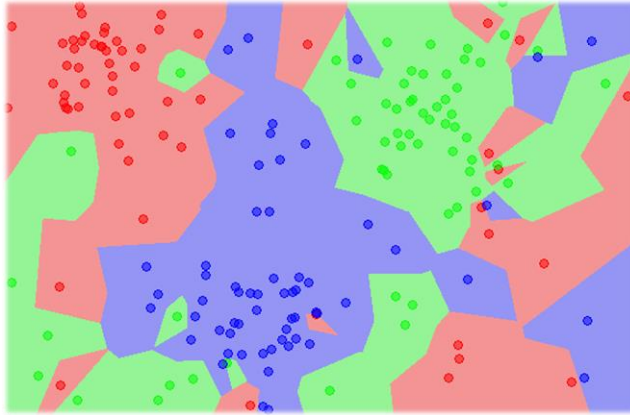
- Ruta
- Modo – tipo vehículo
- Contexto ruta (figura izq.)
- GPS
- Bluetooth
- BIP (fare cards)
- Uso de suelo
- Clima
- Estado ruta
- Apps

La pregunta es ahora, ¿cómo puedo combinar esta información de manera adecuada?

Ejemplo muestra los problemas esenciales de los enfoques estadísticos tradicionales (incluyendo ML)

- Hasta ahora, hemos estado usando características (*features*) de los datos para clasificar/predecir.
- Estas tienen que ver más con el conocimiento experto que tengamos (o no) del problema.
- Esto no es ideal por tres razones principales:
 - i. un experto es caro
 - ii. nada asegura que este conocimiento sea el mejor.
 - iii. la mayoría de las veces no tenemos idea como combinar las features

¿Qué es lo que realmente están aprendiendo estos algoritmos?



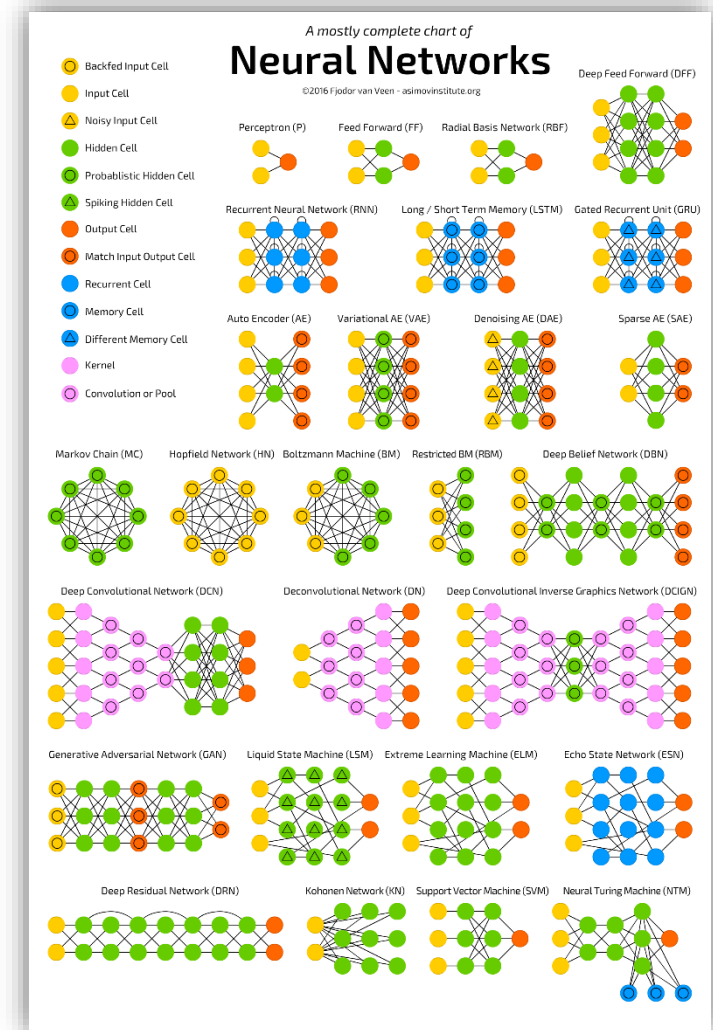
Hasta ahora no estamos “realmente” aprendiendo de los datos, sólo como caracterizarlos en base a *features* hechas a manos

Buscamos un mecanismo de aprendizaje, suficientemente general, para aprender aspectos complejos y eventualmente desconocidos de los datos, con el fin de mejorar el rendimiento en las tareas



Redes neuronales son capaces de aprender distintas **capas de representaciones de datos**

- Método altamente práctico y general para aprender funciones continuas y discretas.
- Brillan en la presencia de grandes volúmenes de datos.
- Existen redes para prácticamente cualquier problema.
- Difíciles de entrenar, interpretabilidad es compleja (caja negra) y sufren de serios problemas de sobreajuste (esto último ya no es tan claro).



A mostly complete chart of Neural Networks

©2016 Fjodor van Veen - asimovinstitute.org

-  Backfed Input Cell
-  Input Cell
-  Noisy Input Cell
-  Hidden Cell
-  Probabilistic Hidden Cell
-  Spiking Hidden Cell
-  Output Cell
-  Match Input Output Cell
-  Recurrent Cell
-  Memory Cell
-  Different Memory Cell
-  Kernel
-  Convolution or Pool

Perceptron (P)



Feed Forward (FF)



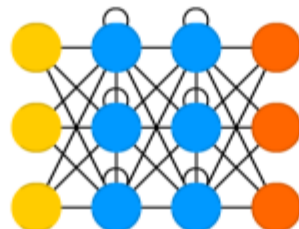
Radial Basis Network (RBF)



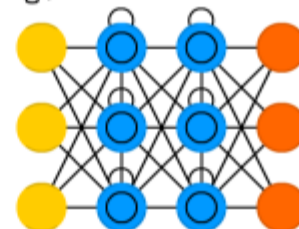
Deep Feed Forward (DFF)



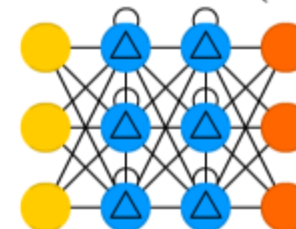
Recurrent Neural Network (RNN)



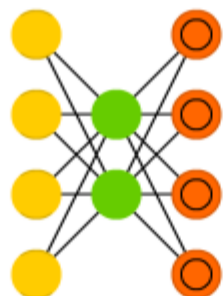
Long / Short Term Memory (LSTM)



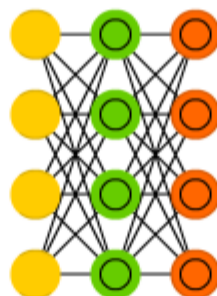
Gated Recurrent Unit (GRU)



Auto Encoder (AE)



Variational AE (VAE)

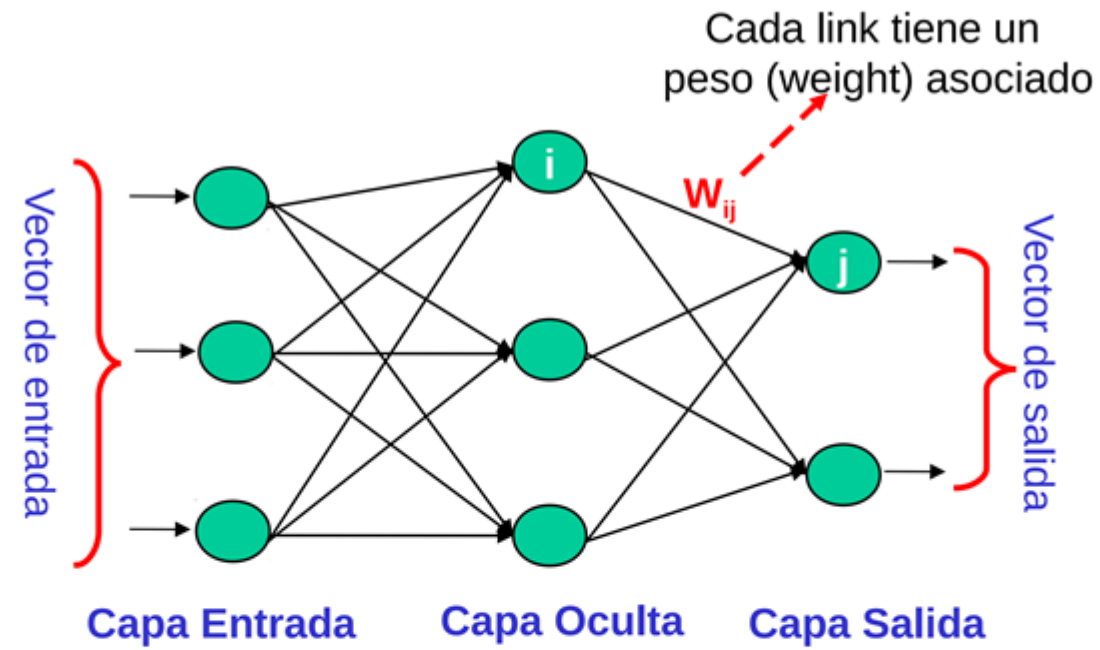


Denoising AE (DAE)



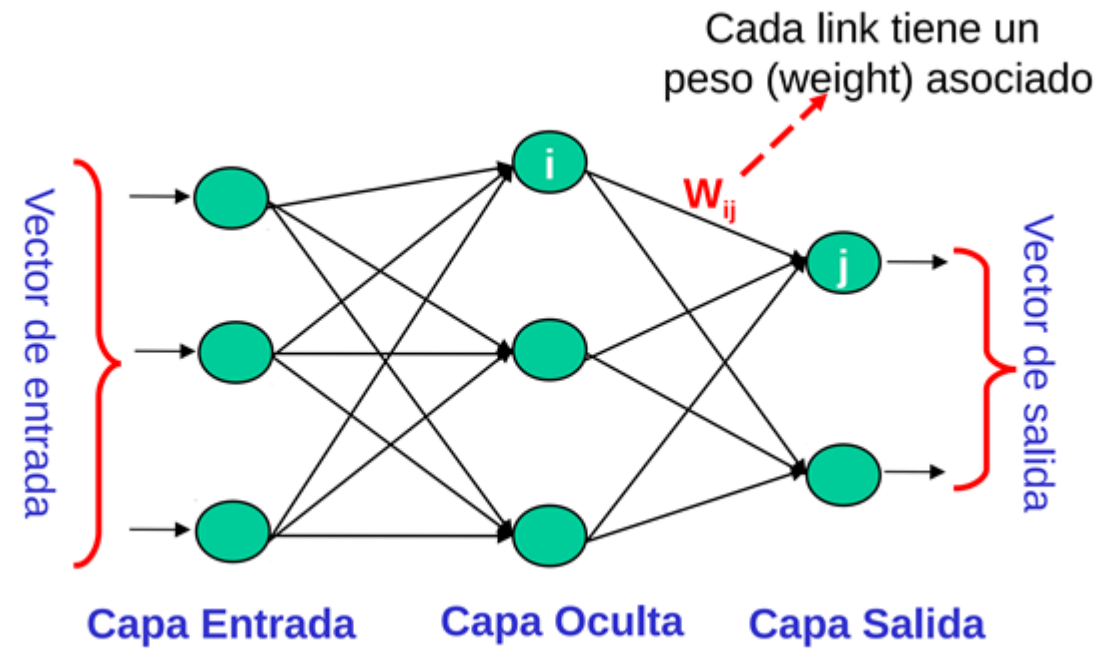
Sparse AE (SAE)





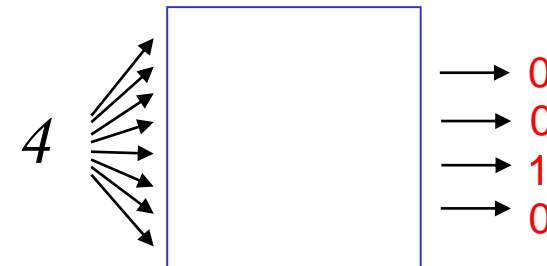
Una FFN es caracterizada por:

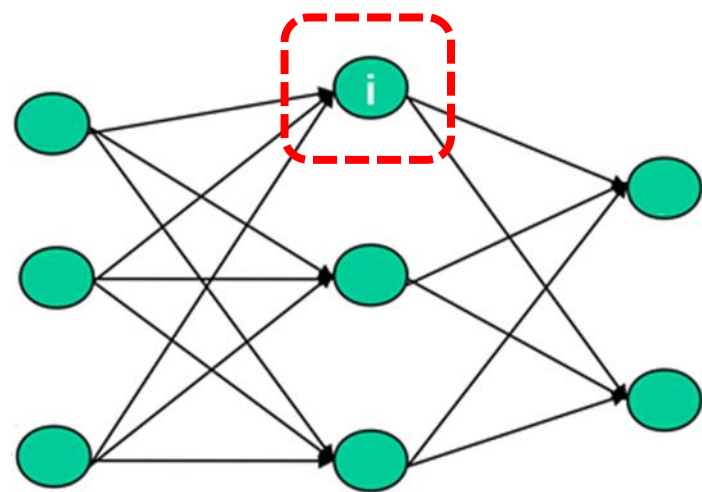
1. Estructura de la red, i.e., número de unidades (neuronas en cada) en cada capa y número de capas ocultas.
2. Método para determinar el valor de los pesos w_{ij} que conectan a una neurona i con una neurona j .
3. Función de activación que modula el impulso que envía una neurona a otra.

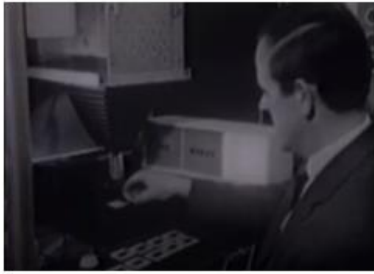


Dada la estructura de la red, ¿cómo podemos ajustar los pesos para que al ingresar un vector de entrada determinado obtengamos la salida deseada?

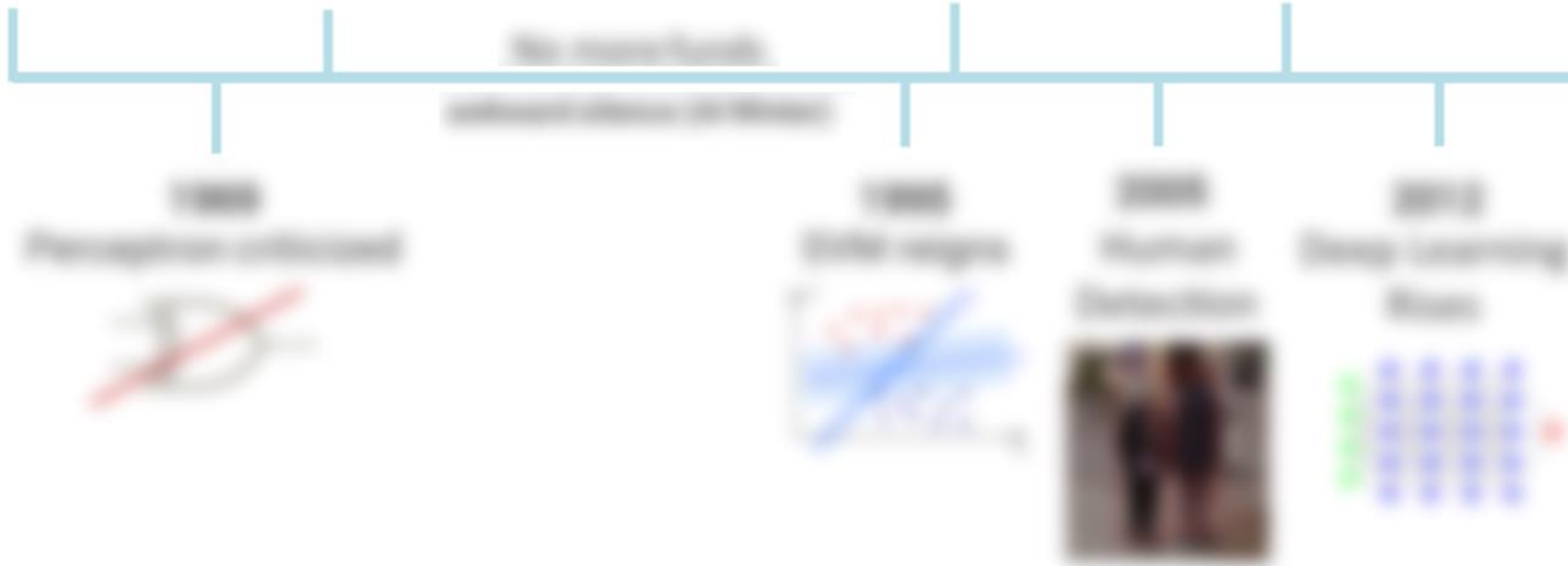
Ejemplo: cuando la red recibe una imagen de un 4, su codificación de salida debe indicar un 4.



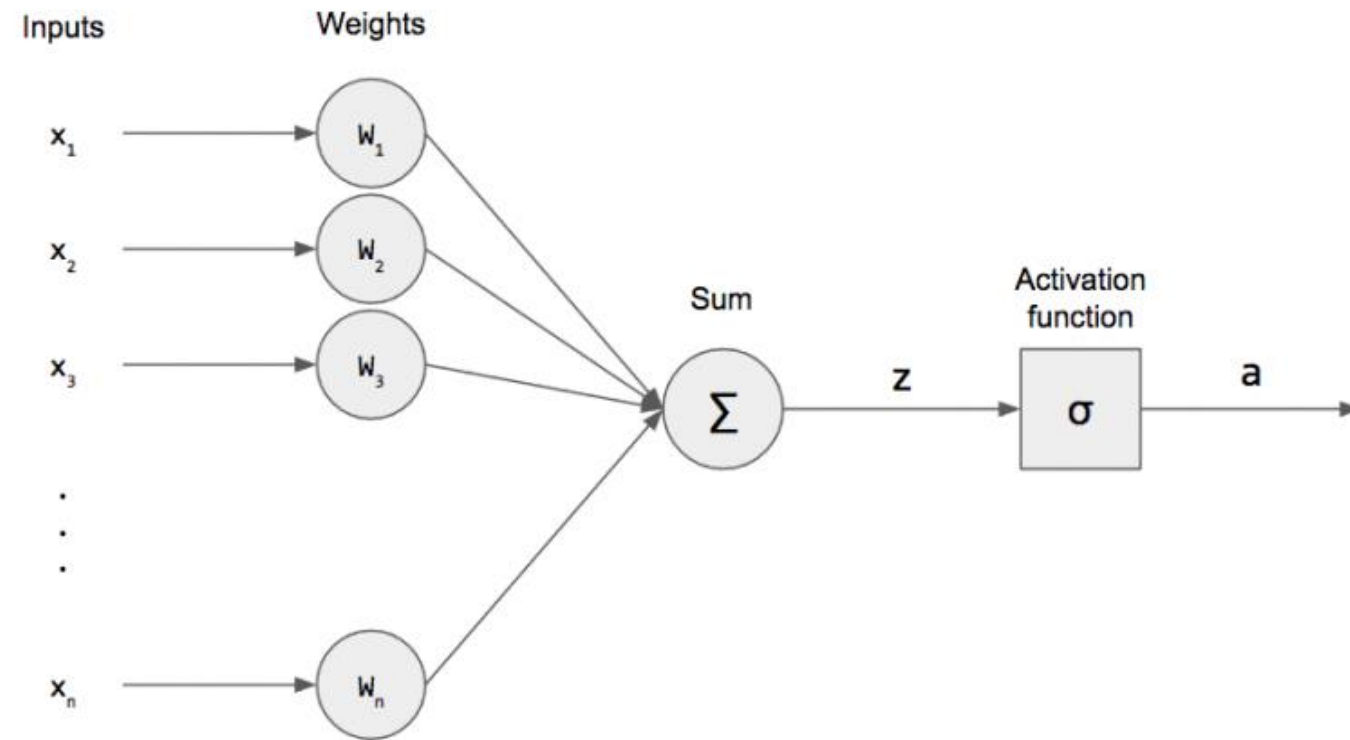




1958 Perceptron



La unidad fundamental de una red neuronal se conoce como **Perceptron**



Pontificia Universidad Católica de Chile
Escuela de Ingeniería
Departamento de Ciencia de la Computación



Sistemas Urbanos Inteligentes

Introducción a Redes Neuronales Parte 1

Hans Löbel

Dpto. Ingeniería de Transporte y Logística
Dpto. Ciencia de la Computación