

TÉCNICAS

TIPOS

CLASIFICACIÓN

de un determinado cliente (o segmento) en los próximos meses.

Nos permiten predecir un valor. Por ejemplo, cual es el beneficio

(en forma de 0 y 1) con su grado de probabilidad.

Los resultados del modelo son binarios, o un sí o un no

entrada como el riesgo de crédito, respuestas a comunicaciones

más propensos a abandonar. Para ello se establecen variables de

sí tratamos de clasificar entre nuestros clientes quiénes son

Permiten predecir la pertenencia a una clase. Por ejemplo

reconocimiento de patrones, graph analytics, etc.

en datos pasados. La variedad de técnicas empleadas incluye minería de datos, modelado

técnicas computacionales con el objetivo de realizar predicciones sobre el futuro basándonos

es un término paraguas para referirnos al conjunto de procesos que implican aplicar diferentes

para conseguir los resultados deseados.

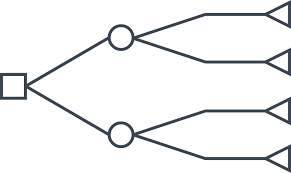
que se desea cumplir. y las funciones de minería cuanto mejor sea esta formulación inicial

más claras serán las directrices acerca de los datos y las funciones de minería que se utilizan

Los analistas de la empresa definen el problema a resolver y el objetivo concreto de empresa

El proyecto de minería de datos comienza con un plan bien definido de inteligencia comercial.

Minera de datos



**MODELO PREDICTIVO**

REGRESIÓN V

ÁRBOLES DE DECISIÓN

precisión.

clasificar casos antiguos o recientes con la máxima

los datos para generar reglas que pueda usar para

frente a no votantes o tipos de bacterias), puede usar

riesgo, suscriptores frente a no suscriptores, votantes

préstamos de alto riesgo frente a préstamos de bajo

datos divididos en clases que le interesan (por ejemplo

en un conjunto de reglas de decisión. sí dispone de

que predicen o clasifican observaciones futuras basándose

Se utilizan para desarrollar sistemas de clasificación



REGRESIÓN LINEAL Y LOGISTOCA

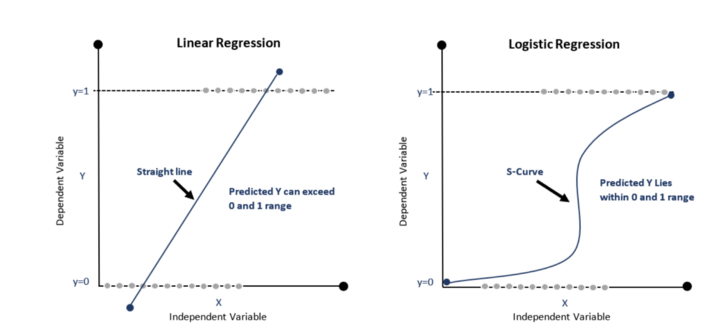
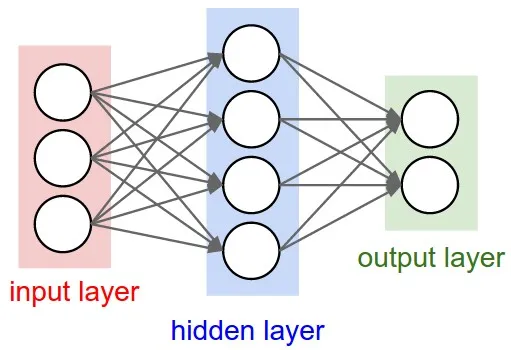
los años de experiencia. Por su parte el algoritmo de Regresión Logística se utiliza

La Regresión Lineal, se utiliza para proyectos de es

por ejemplo, queremos predecir el precio

de una casa o el decir cuando queremos predecir un valor numérico,

sueldo de una persona de acuerdo a



REDES NEURONALES

Consisten en técnicas sofisticadas capaces de modelar

sencillas como las regresiones o los árboles de decisión.

confirmar descubrimientos de otras técnicas más

de big data. Otras veces se usan simplemente para

cuantos más datos manejemos, ideal para el análisis

no lineales de los datos, lo que lo hace muy interesante

mismo tiempo. Tienen la capacidad de manejar relaciones

populares porque son muy potentes pero flexibles al

relaciones extremadamente complejas. Se han hecho

del Titanic, o si una acción de la bolsa de valores va a subir o no.

para proyectos de clasificación, si queremos conocer si una

persona vio o no el naufragio

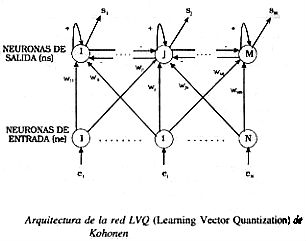


rendimiento al extraer datos históricos para buscar las razones detrás del éxito o el fracaso del

abordar el futuro. La analítica descriptiva examina el rendimiento pasado y entiende ese

La analítica descriptiva examina los datos y analiza los acontecimientos pasados ​​para saber cómo

**MODELO DESCRIPTIVO**



entre el número total de datos.

MEDIANA ARITMÉTICA: Es el valor obtenido al sumar todos los datos y dividir el resultado entre el

ordenados de menor a mayor. Variables cuantitativas.

MEDIANA: Es el valor que ocupa el lugar central de todos los datos cuando estos están

MODA: Es el valor que tiene mayos frecuencia absoluta. Variables cualitativas y cuantitativas.

K) MEDIDAS

Las más comunes son las siguientes:

Nos indican en torno a qué valor (centro) se distribuyen los datos.

EL MODELO DE KOHONEN

ANALITICAS

bidimensionales.

se representan internamente en forma de mapas

captadas del entorno a través de los órganos sensoriales

organizan en muchas zonas, de forma que las informaciones

como ocurre en el cerebro. En éste hay neuronas que se

para formar mapas de características de manera similar

Se trataba de un modelo de red neuronal con capacidad

TECNICAS

TIPOS

GRÁFICAS

en los datos.

características de las variables toda la información contenida

Su fin es resumir en forma de "parámetros" o valores de ciertas

pero no otro.

Dependiendo del tipo de variable se utilizará un tipo de gráfico

Histograma.

Existen tres tipos básicos de gráficos: Sectores, Barras e

en los datos.

Su fin es representar de forma gráfica la información contenida

utilizan este tipo de análisis post-mortem.

pasado. Casi todos los informes de gestión, tales como ventas, marketing, operaciones y finanzas

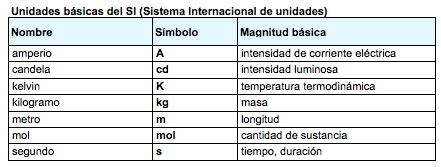
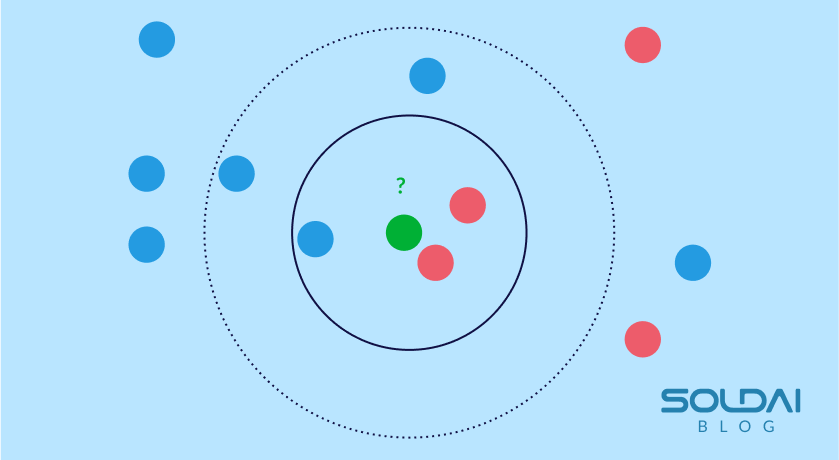


posición central respecto a los datos.

MEDIDAS DE DISPOSICIÓN: Sirven para medir el grado de representativo de las medidas de

muestra.

MEDIDAS DE POSICIÓN: Nos dan un valor promedio o característica representativa de la



de los k vecinos más cercanos al objeto xᵩ.

Es un método no paramétrico. Se usará la notación NN (K, xᵩ) para denotar el conjunto

Es un método sencillo de clasificación y regresión es el de los K vecinos más cercanos

VECINOS MAS PROXIMOS