

SUBDIRECCIÓN ACADÉMICA

DEPARTAMENTO DE SISTEMAS Y COMPUTACIÓN

ENERO - JUNIO 2020

INGENIERÍA INFORMÁTICA

MATERIA

DATOS MASIVOS

CATEDRÁTICO:

JOSE CHRISTIAN ROMERO HERNANDEZ

UNIDAD 2

# TAREA 1

ALUMNO

López Valencia Luis Daniel #15212159

### **1. Algoritmos de regresión** En las tareas de regresión, el programa de aprendizaje automático debe estimar y comprender las relaciones entre las variables. El análisis de regresión se enfoca en una variable dependiente y una serie de otras variables cambiantes, lo que lo hace particularmente útil para la predicción y el pronóstico.

### **2. Algoritmos bayesianos** Este tipo de algoritmos por clasificación están basados en el teorema de Bayes y clasifican cada valor como independiente de cualquier otro. Lo que permite predecir una clase o categoría en función de un conjunto dado de características, utilizando la probabilidad. A pesar de su simplicidad, el clasificador funciona sorprendentemente bien y se usa a menudo porque supera a los métodos de clasificación más sofisticados.

### **3. Algoritmos de agrupación** Se utilizan en el aprendizaje no supervisado, y sirven para categorizar datos no etiquetados, es decir, datos sin categorías o grupos definidos. El algoritmo funciona mediante la búsqueda de grupos dentro de los datos, con el número de grupos representados por la variable K. A continuación, funciona de manera iterativa para asignar cada punto de datos a uno de los K grupos según las características proporcionadas.

### **4. Algoritmos de árbol de decisión** Un árbol de decisión es una estructura de árbol similar a un diagrama de flujo que utiliza un método de bifurcación para ilustrar cada resultado posible de una decisión. Cada nodo dentro del árbol representa una prueba en una variable específica, y cada rama es el resultado de esa prueba.

### **5. Algoritmos de redes neuronales** Una red neuronal artificial (RNA) comprende unidades dispuestas en una serie de capas, cada una de las cuales se conecta a las capas anexas. Las RNA se inspiran en los sistemas biológicos, como el cerebro, y en cómo procesan la información. Por lo tanto, son esencialmente un gran número de elementos de procesamiento interconectados, que trabajan al unísono para resolver problemas específicos. También aprenden con el ejemplo y la experiencia, y son extremadamente útiles para modelar relaciones no lineales en datos de alta dimensión, o donde la relación entre las variables de entrada es difícil de entender.

### **6. Algoritmos de reducción de dimensión** La reducción de dimensión reduce el número de variables que se consideran para encontrar la información exacta requerida.

### **7. Algoritmos de Aprendizaje Profundo** Los algoritmos de aprendizaje profundo ejecutan datos a través de varias capas de algoritmos de redes neuronales, las cuales pasan a una representación simplificada de los datos a la siguiente capa. La mayoría funciona bien en conjuntos de datos que tienen hasta unos cientos de características o columnas. Sin embargo, un conjunto de datos no estructurado, como el de una imagen, tiene una cantidad tan grande de características que este proceso se vuelve engorroso o completamente inviable.