Aprendizaje No Supervisado

El aprendizaje no supervisado ahora lo que busca es que la computadora en vez de tomar una columna objetivo clasifique los datos, es decir que ahora ella va a definir las etiquetas.

Método de K – medias

Lo primero de este método consiste en definir el número de clusterizaciones quieres., a este número de clusterizaciones se les asigna el hiperparámetro K. De acuerdo al número de K dado, la computadora va a decidir cómo clasificar las etiquetas.

El siguiente paso es definir el número de centroides de cada cluster, este valor centroide es el valor más representativo de cada clusterización.

El tercer paso es calcular las distancias entre puntos del cluster.

El paso 4 es calcular la inercia, la cual es una medida de que tan bueno es la inercia del método.

¿Cómo se puede interpretar la distorsión?