

Detección de anomalías basada en técnicas de ensembles

Ignacio Aguilera Martos

5 de septiembre de 2019

Trabajo Fin de Grado

Código disponible en [GitHub](#)

1. Concepto de anomalía basado en distancias

Concepto de anomalía basado en distancias

Tukey's Fences

Tukey's Fences

Pensado para el caso uno-dimensional.

Tukey's Fences

Pensado para el caso uno-dimensional.

Tukey's Fences

Valores fuera del rango $[Q_1 - k(Q_3 - Q_1), Q_3 + k(Q_3 - Q_1)]$ con $k = 1,5$

Tukey's Fences

Pensado para el caso uno-dimensional.

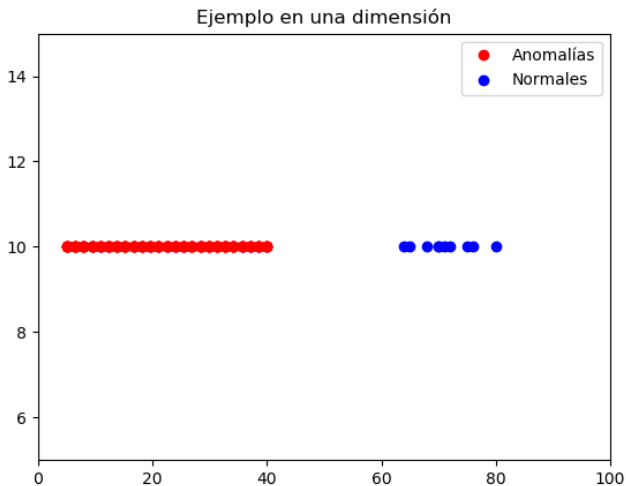
Tukey's Fences

Valores fuera del rango $[Q_1 - k(Q_3 - Q_1), Q_3 + k(Q_3 - Q_1)]$ con $k = 1,5$

La propuesta de $k = 1,5$ es arbitraria.

Tukey's Fences

Tukey's Fences



Extensión al caso de mayor dimensionalidad

Extensión al caso de mayor dimensionalidad

Criterio

Aplicar el criterio de Tukey a cada una de las características.

Extensión al caso de mayor dimensionalidad

Criterio

Aplicar el criterio de Tukey a cada una de las características. **Trivial**.

Extensión al caso de mayor dimensionalidad

Criterio

Aplicar el criterio de Tukey a cada una de las características. **Trivial**.

Criterio de clusters

1. Agrupamos los datos por clusters.

Extensión al caso de mayor dimensionalidad

Criterio

Aplicar el criterio de Tukey a cada una de las características. **Trivial**.

Criterio de clusters

1. Agrupamos los datos por clusters.
2. Encontramos el cluster más cercano para cada instancia.

Extensión al caso de mayor dimensionalidad

Criterio

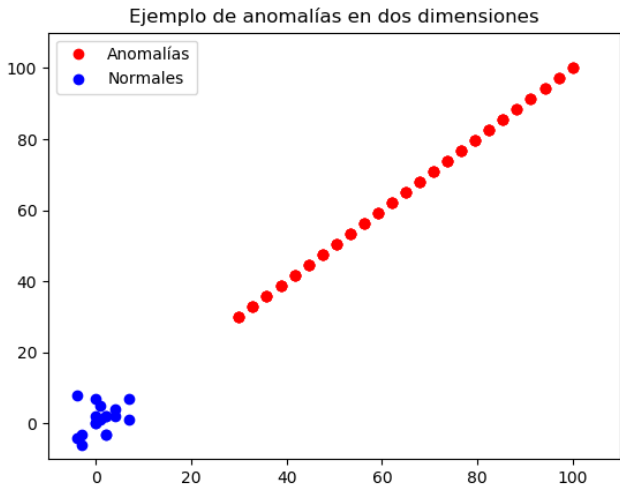
Aplicar el criterio de Tukey a cada una de las características. **Trivial**.

Criterio de clusters

1. Agrupamos los datos por clusters.
2. Encontramos el cluster más cercano para cada instancia.
3. Si la distancia del objeto al centroide del cluster es mayor que 1,5 veces la mayor distancia intercluster entonces es una anomalía.

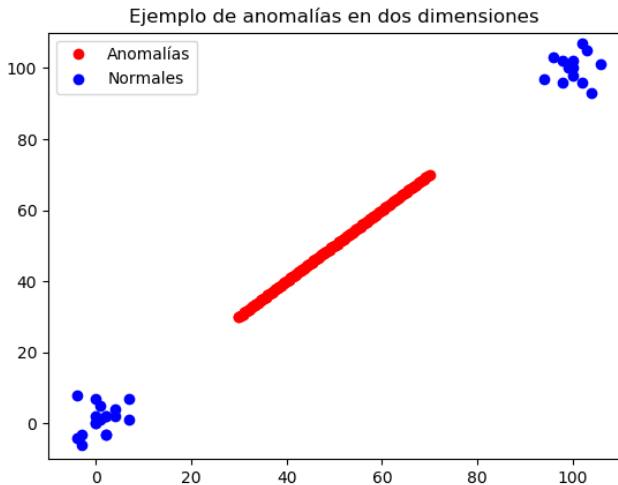
Ejemplo 1

Ejemplo 1



Ejemplo 2

Ejemplo 2



¿Preguntas?