



# Machine Learning DB-outlier

Prof. Hugo de Paula

# Análise de anomalias

## ***Outliers* (valores extremos)**

- “Um outlier é um fato que desvia tanto de outros fatos a ponto de gerar suspeitas de que foi gerado por um mecanismo diferente”.

# Análise de outlier

A identificação de outliers pode levar à descoberta de conhecimento inesperado em aplicações como:

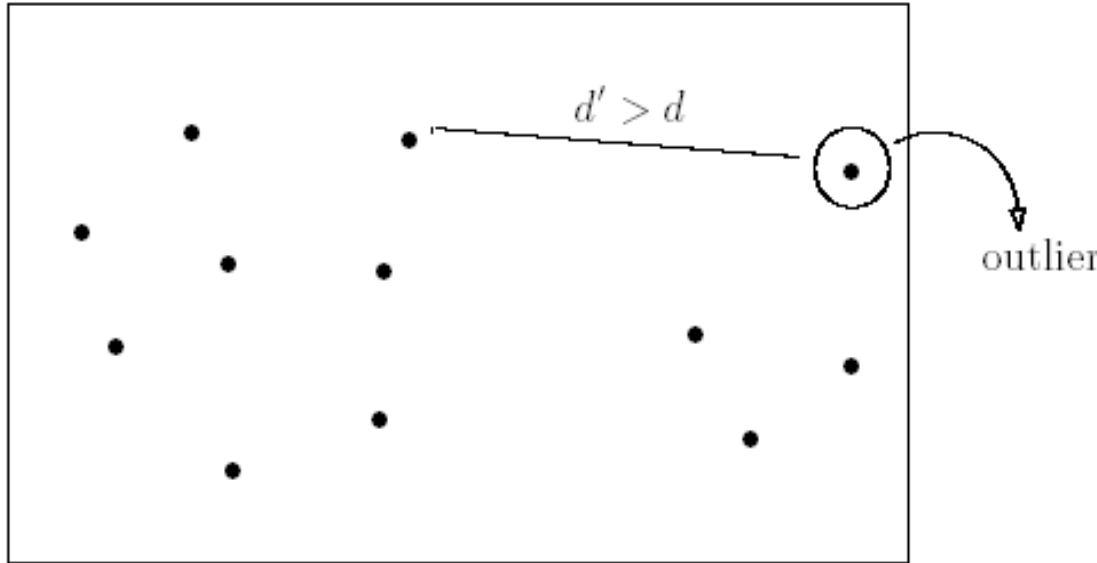
- Fraude telecomunicação.
- Fraude cartões de créditos.
- Segmentação de clientes.
- Performance de atletas profissionais (casos de dopagem, por exemplo).

# Análise de outlier

## Definição:

- Um objeto **O** de um banco de dados **D** é dito um **DB(p,d)-outlier** se pelo menos uma fração **p** ( $0 < p < 1$ ) dos objetos de **D** estão a uma distância maior do que **d** de **O**.

# Análise de outlier



Banco de Dados  $D$

objeto marcado é um  $DB(p,d)$ -outlier, para  $p = 2/3$  e  $d > d'$  (8 dos 12 objetos de  $D$  estão a uma distância maior do que  $d'$  deste objeto).

# DB-outlier: processo

- Sejam uma base **D** com **N** elementos, **p** um número entre 0 e 1 e **d** > 0.
- Considere `dist()` a função distância considerada.
- **d-vizinhança(O)** é o conjunto de pontos cujas distâncias a **O** são no máximo **d**.
- **p** é a fração mínima de objetos de **D** que devem ficar fora da **d-vizinhança** de um outlier.
- O número máximo de objetos dentro da vizinhança de um outlier, **M** é:  
**M = N(1 - p).**

# DB-outlier: processo

Encontrar todos os DB( $p$ ;  $d$ )-outliers se reduz a encontrar, para cada objeto **O** de **D**, uma vizinhança contendo no máximo **M** elementos.

- Executa-se uma busca dentro de um raio **d** para cada objeto **O**.
- Se **M + 1** elementos são encontrados nesta vizinhança, a busca termina e **O** é declarado um não-outlier. Os objetos que sobram são os outliers.

# DB-outlier: exemplo

Se considerarmos  $M=2$  e  $d=3$

Nos dados abaixo, as linhas 6 e 7 são outliers

	L1	L2	L3	L4	L5	L6	L7
L1	0	1	3	2	5	2,5	6
L2	1	0	4	3	2	6	1
L3	3	4	0	2	3	4	5
L4	2	3	2	0	7	8	5
L5	5	2	3	7	0	8	9
L6	2,5	6	4	8	8	0	9
L7	6	1	5	5	9	9	0



# Aviso legal

O material presente nesta apresentação foi produzido a partir de informações próprias e coletadas de documentos obtidos publicamente a partir da Internet. Este material contém ilustrações adquiridas de bancos de imagens de origem privada ou pública, não possuindo a intenção de violar qualquer direito pertencente à terceiros e sendo voltado para fins acadêmicos ou meramente ilustrativos. Portanto, os textos, fotografias, imagens, logomarcas e sons presentes nesta apresentação se encontram protegidos por direitos autorais ou outros direitos de propriedade intelectual.

Ao usar este material, o usuário deverá respeitar todos os direitos de propriedade intelectual e industrial, os decorrentes da proteção de marcas registradas da mesma, bem como todos os direitos referentes a terceiros que por ventura estejam, ou estiveram, de alguma forma disponíveis nos slides. O simples acesso a este conteúdo não confere ao usuário qualquer direito de uso dos nomes, títulos, palavras, frases, marcas, dentre outras, que nele estejam, ou estiveram, disponíveis.

É vedada sua utilização para finalidades comerciais, publicitárias ou qualquer outra que contrarie a realidade para o qual foi concebido. Sendo que é proibida sua reprodução, distribuição, transmissão, exibição, publicação ou divulgação, total ou parcial, dos textos, figuras, gráficos e demais conteúdos descritos anteriormente, que compõem o presente material, sem prévia e expressa autorização de seu titular, sendo permitida somente a impressão de cópias para uso acadêmico e arquivo pessoal, sem que sejam separadas as partes, permitindo dar o fiel e real entendimento de seu conteúdo e objetivo. Em hipótese alguma o usuário adquirirá quaisquer direitos sobre os mesmos.

O usuário assume toda e qualquer responsabilidade, de caráter civil e/ou criminal, pela utilização indevida das informações, textos, gráficos, marcas, enfim, todo e qualquer direito de propriedade intelectual ou industrial deste material.



**PUC Minas**  
**Virtual**

© PUC Minas • Todos os direitos reservados, de acordo com o art. 184 do Código Penal e com a lei 9.610 de 19 de fevereiro de 1998.  
Proibidas a reprodução, a distribuição, a difusão, a execução pública, a locação e quaisquer outras modalidades de utilização sem a devida autorização da Pontifícia Universidade Católica de Minas Gerais.