Aprendizado Descritivo

Aula 01 – Introdução: aprendizado descritivo x preditivo

Professor Renato Vimieiro

DCC/ICEx/UFMG

Introdução

- Quando falamos sobre aprendizado de máquina e mineração de dados, frequentemente associamos essas expressões a predição de valores
- Mais especificamente, temos a ideia de que aprendizado de máquina (AM) se resume a, dada uma entrada X, encontrar uma função f(X) que retorne o valor de uma variável alvo Y
  - Quando a variável alvo Y é categórica, chamamos o problema de **classificação**
  - Quando a variável alvo Y é contínua, chamamos o problema de **regressão**
- Assim, o senso comum define AM como aprender uma função f capaz de predizer valores para dados ainda não coletados
- Essa definição, embora restritiva, é correta para a classe de tarefas elencadas acima, chamadas de **aprendizado preditivo**

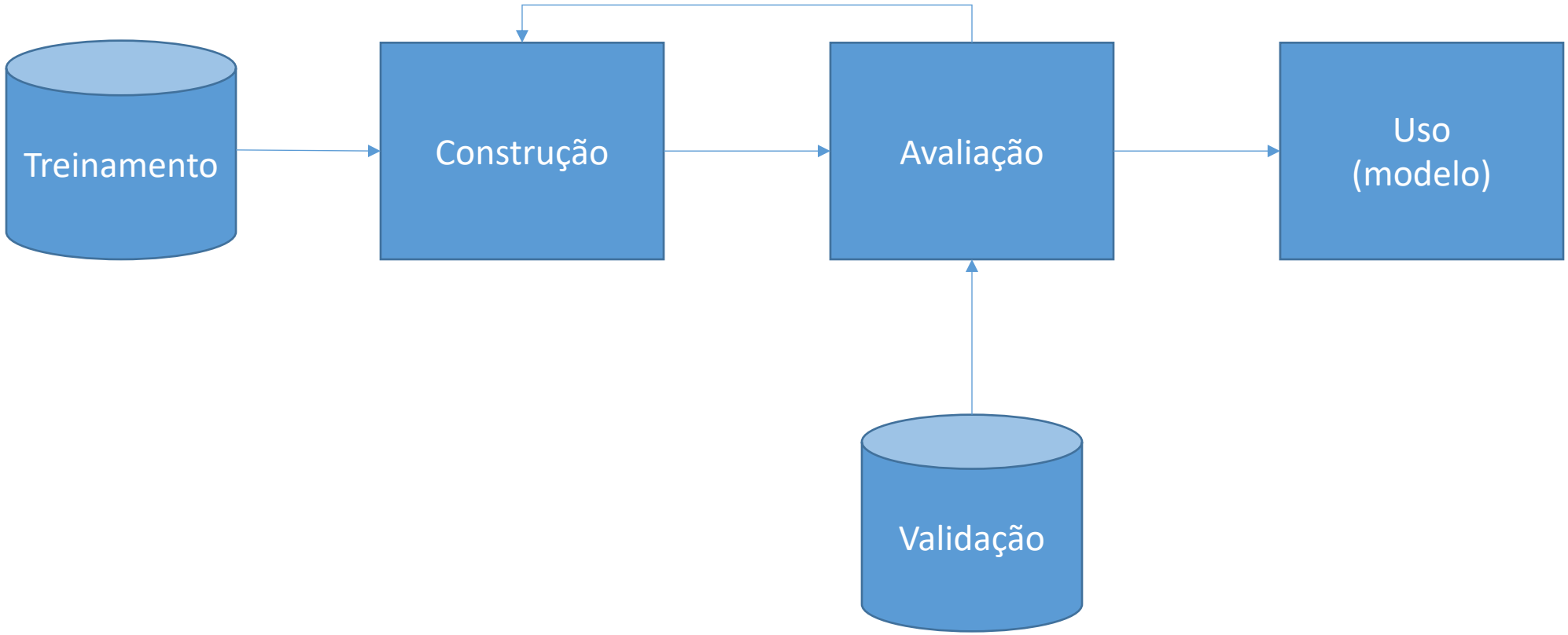
Aprendizado Preditivo

- Mais especificamente ainda, esse imaginário popular define o que conhecemos por aprendizado supervisionado de modelos preditivos
- De maneira mais formal, o objetivo dessas técnicas é aprender uma função cujo domínio é um conjunto de instâncias $\mathcal{X}$, e contradomínio um conjunto de saídas $\mathcal{Y}$
- Para tal, o algoritmo recebe um conjunto de pares (x, l(x)) $\in \mathcal{X} \times \mathcal{L}$
  - A função l(x) retorna o valor de saída esperado para a instância x
- Como o algoritmo obtém o modelo guiado por esse conjunto de entrada (treinamento), a tarefa é classificada como 'supervisionada', já que l(x) faz o papel de 'professor'
- Da mesma forma, como o modelo aprendido é usado para predizer valores de saída de novas instâncias, ele é chamado de preditivo

Aprendizado Preditivo

- Como queremos que o modelo seja fiel à realidade e, portanto, consiga capturar a relação entre instância (entrada) e saída, é comum, nesse cenário, avaliarmos sua qualidade comparando os valores obtidos com os verdadeiros para um conjunto de instâncias chamado de conjunto de validação
- Ou seja, nesse tipo de tarefa temos o seguinte fluxo de trabalho:

Aprendizado Preditivo



Aprendizado Preditivo

- Embora métodos de regressão e classificação (aprendizado supervisionado) sejam os mais populares em AM, existem outras abordagens preditivas que não requerem a variável alvo para ajustarem modelos
- Técnicas que não utilizam essas variáveis são classificadas como **não-supervisionadas**
- Uma tarefa de aprendizado não-supervisionado bastante popular é a de agrupamento (clustering)
- Essa tarefa consiste em encontrar subgrupos de elementos homogêneos nos dados

Aprendizado Preditivo

- Em outras palavras, a tarefa de agrupamento consiste em detectar grupos de instâncias que sejam mais parecidas entre si do que com as de outros grupos
- Sob a ótica de aprendizado preditivo, a tarefa consiste em encontrar uma função q: $\mathcal{X} \rightarrow \mathcal{C}$, cujo domínio $\mathcal{X}$ é um conjunto de instâncias, e o contradomínio $\mathcal{C}$ um conjunto de grupos (clusters)
- Note que a tarefa se assemelha à de classificação (já que os rótulos dos grupos são categóricos), porém o ajuste do modelo não leva em consideração rótulos pré-definidos
- Um exemplo de método aprendizado preditivo não-supervisionado é o K-Means
  - O modelo são os centroides e a distância euclidiana
  - Novas instâncias são alocadas nos clusters de cujos centroides elas sejam mais próximas de acordo com a distância euclidiana

Aprendizado Descritivo

- Aprendizado descritivo tem como objetivo central obter uma descrição para os dados
  - Isto é, o objetivo é encontrar um **modelo descritivo** para os dados
- Dessa forma, podemos apontar a primeira diferença para aprendizado preditivo:
  - Não temos mais a necessidade de dividir o conjunto de instâncias em treinamento e validação
- A divisão entre treinamento e validação não faz mais sentido, pois queremos obter um modelo para os dados que temos em mãos
- Consequentemente, a avaliação dos resultados (modelos) se torna mais difícil, já que não temos mais uma ‘verdade absoluta’ para compararmos as saídas

Aprendizado Descritivo

- Por outro lado, segundo Flach (2012), *o aprendizado descritivo leva à descoberta genuína de novos conhecimentos, e, dessa forma, está situado entre as áreas de mineração de dados e aprendizado de máquina*
- O objetivo de se buscar um modelo descritivo dos dados se justifica nas situações em que se quer responder perguntas do tipo “o quê aconteceu?”
- Ou seja, esses modelos descrevem situações passadas e, assim, auxiliam no processo de tomada de decisão

Aprendizado Descritivo

- Considere a seguinte situação em que um grande *Market place* deseja reduzir os custos de distribuição dos produtos que ele vende
- Uma prática muito utilizada atualmente é manter centros de distribuição regionais para estocar produtos vendidos frequentemente, reduzindo o custo e tempo de transporte, e, consequentemente, aumentando a satisfação dos clientes
- Apesar da estocagem de produtos populares nas regionais ser uma decisão trivial, ela pode ser aprimorada analisando-se o histórico de vendas
- Nesse histórico, podemos encontrar itens menos populares que, com certa frequência, são adquiridos junto com os mais populares
  - Isso nos permite decidir estocar também esses produtos menos populares, em menor quantidade, mas evitando, assim, um custo maior de se enviar tais produtos individualmente de centros mais distantes

Aprendizado Descritivo

- Considere um segundo caso real em que executivos do Wal-Mart utilizaram de AM para aumentar as vendas diante da ameaça do furação Frances em 2004
- Enquanto o furacão atravessava o Caribe, os executivos queriam prever os produtos que seus clientes consumiam diante de catástrofes

www.nytimes.com/2004/11/14/business/yourmoney/what-walmart-knows-about-customers-habits.html

The New York Times

What Wal-Mart Knows About Customers' Habits

By Constance L. Hays

Nov. 14, 2004

Correction Appended

HURRICANE FRANCES was on its way, barreling across the Caribbean, threatening a direct hit on Florida's Atlantic coast. Residents made for higher ground, but far away, in Bentonville, Ark., executives at Wal-Mart Stores decided that the situation offered a great opportunity for one of their newest data-driven weapons, something that the company calls predictive technology.

Aprendizado Descritivo

- O objetivo dos executivos do Wal-Mart era abastecer as lojas da Flórida, que estava no caminho do furacão, e assim aumentar as vendas
- Novamente, a decisão trivial seria aumentar o estoque de pilhas, água mineral, lanternas, e produtos não perecíveis
- A análise do histórico de vendas, nesse caso, poderia retornar que as lojas venderam todo o estoque de DVDs de um gênero específico de filmes
- No entanto, ao analisar os dados, eles chegaram à conclusão de que havia um aumento de 7x nas vendas de Pop Tarts sabor morango, e o campeão do aumento de vendas era cerveja

Aprendizado Descritivo

- Apesar do caso ter sido abordado como um exemplo de aprendizado preditivo na matéria, ele é um exemplo claro de aprendizado descritivo supervisionado
- Especificamente, ele é um caso claro da aplicação de excepcional model mining
  - Queremos encontrar padrões de vendas não usuais, isto é, detectar aumento da venda de produtos que esteja correlacionado positivamente a um evento de interesse
- O único porém é que essa tarefa foi inicialmente proposta somente em 2008, 4 anos após o evento!

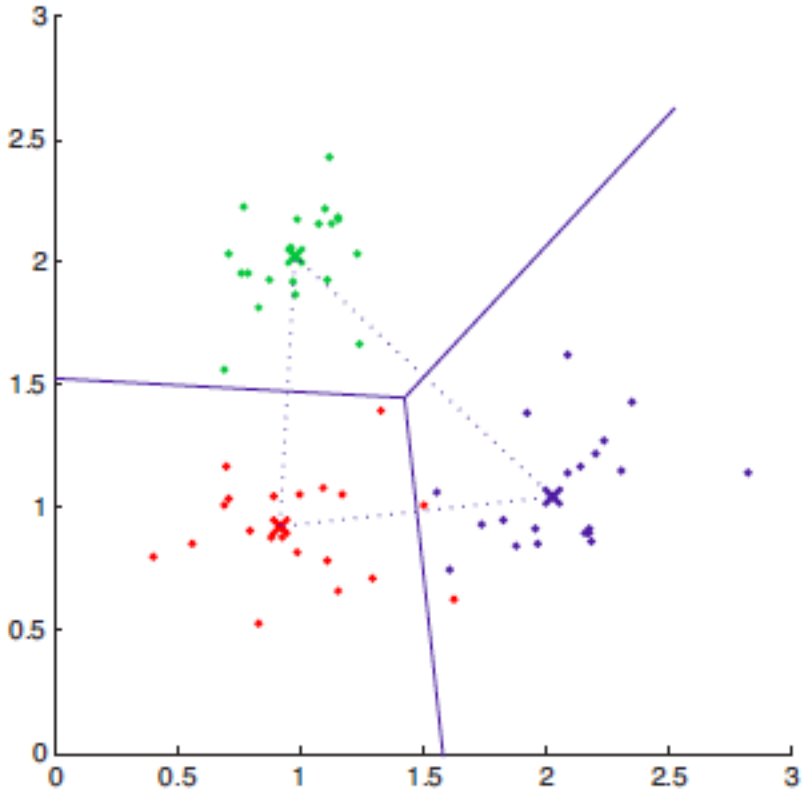
Aprendizado Descritivo x Preditivo

- Os exemplos ilustram bem a aplicabilidade do aprendizado descritivo:
  - Como dito, eles revelam o que ocorreu no passado e auxiliam na tomada de decisão de eventos futuros
- De forma análoga, os modelos preditivos utilizam dados históricos para prever o comportamento de dados futuros
- Note que a diferença, portanto, é bem tênue entre as duas abordagens
  - A diferença mais aparente é a intervenção humana no primeiro, contra um certo automatismo da segunda

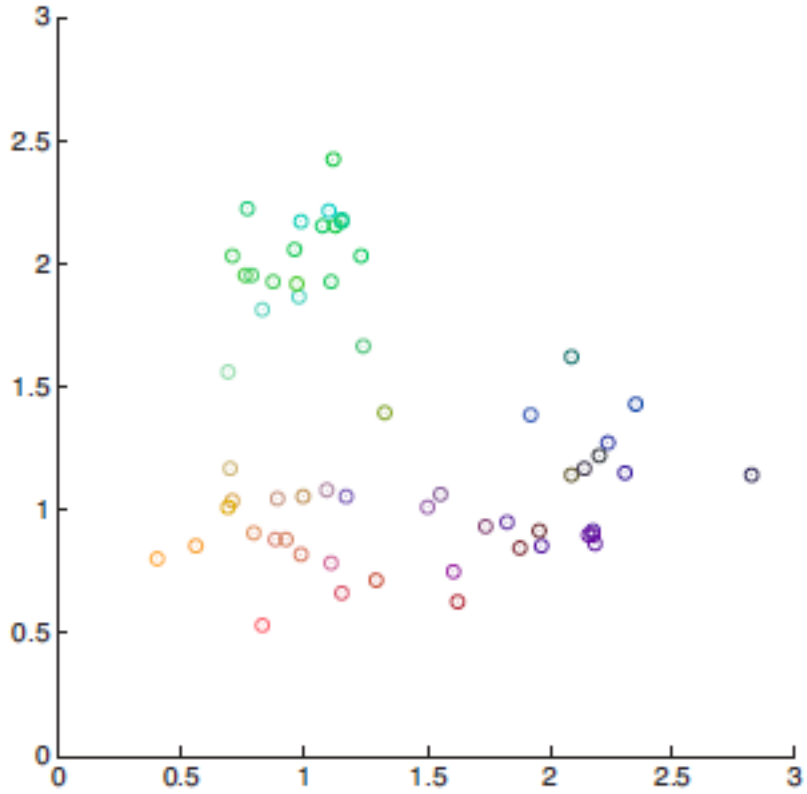
Aprendizado Descritivo x Preditivo

- Um exemplo dessa diferença tênue entre as duas abordagens é justamente a tarefa de agrupamento
- Incialmente apresentamos clustering como uma tarefa de aprendizado preditivo não-supervisionado
- A mesma tarefa, porém, pode ser apresentada como descritiva
- O objetivo no agrupamento descritivo é obter uma função q: $\mathcal{D} \rightarrow \mathcal{C}$, que mapeia as instâncias coletadas no conjunto de dados $\mathcal{D}$ a grupos específicos (clusters) $\mathcal{C}$
  - A diferença aqui é que assumimos como domínio apenas o conjunto de instâncias em mãos, e não toda população de instâncias possíveis
  - Ou seja, o resultado do nosso agrupamento é ‘apenas’ uma divisão das instâncias em grupos, permitindo a análise de similaridade entre elas; não estamos interessados em alocar novas instâncias aos grupos encontrados

Aprendizado Descritivo x Preditivo



Agrupamento preditivo

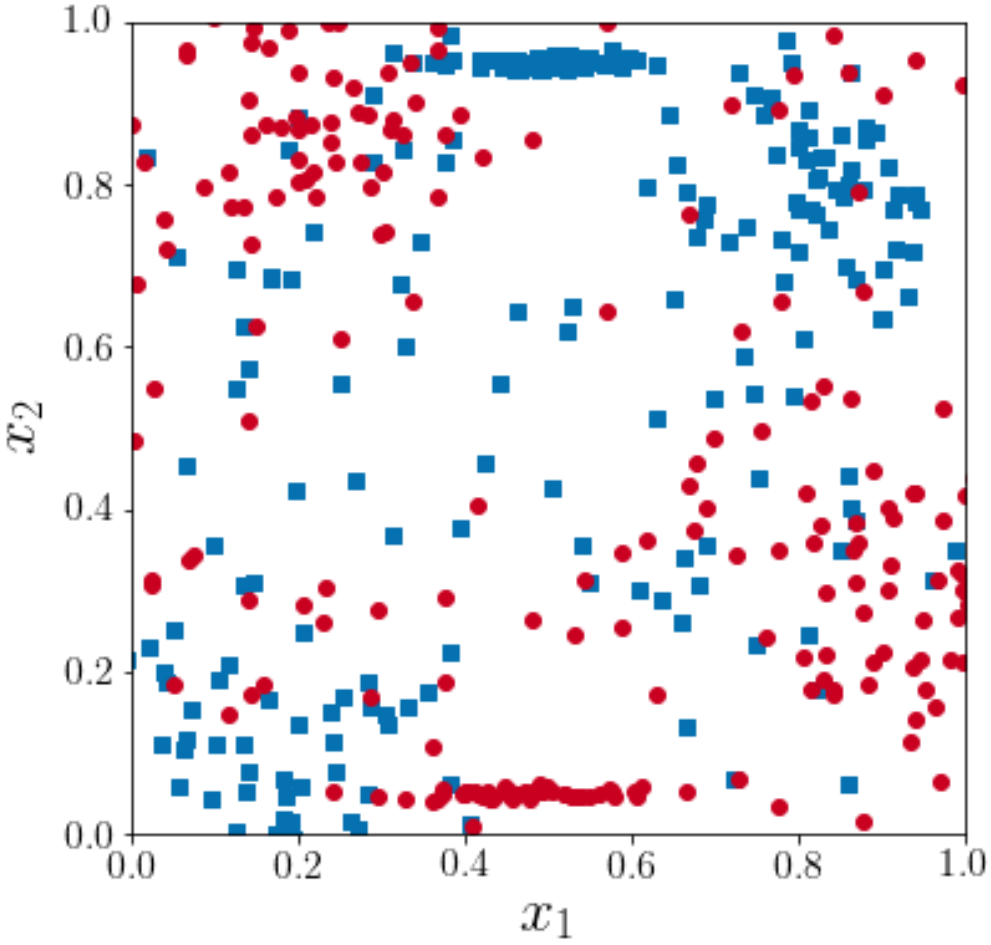


Agrupamento descritivo

Figura retirada de Flach (2012)

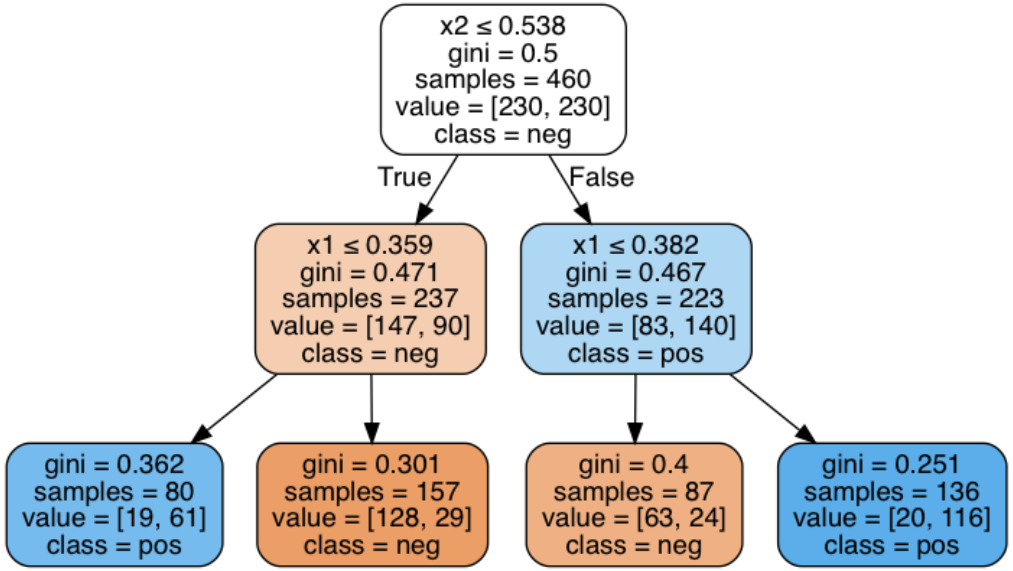
Aprendizado Descritivo x Preditivo

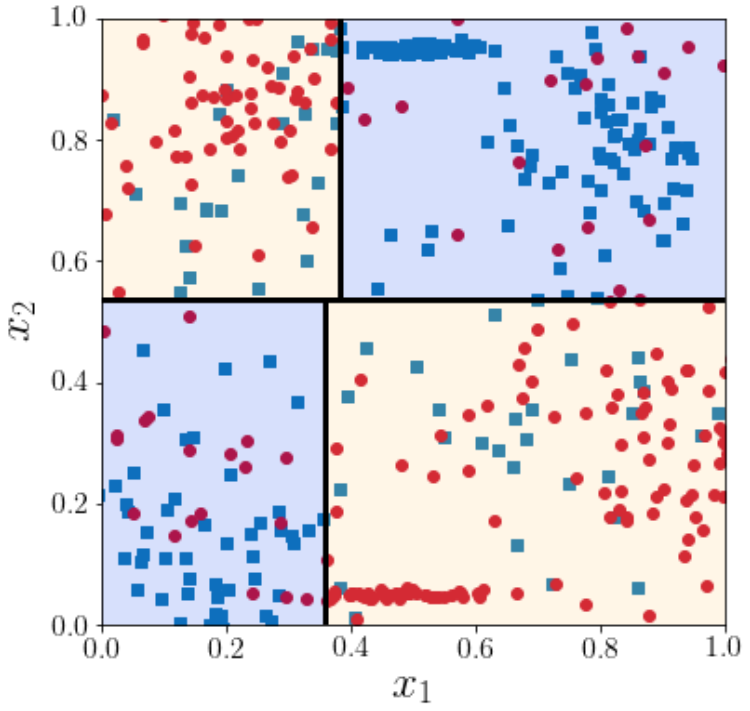
- Considere agora um exemplo mais relacionado ao segundo estudo de caso discutido anteriormente
- Suponha que nosso conjunto de entrada rotulado seja o seguinte
- Podemos traçar dois objetivos aqui:
  - Separar círculos de quadrados, para classificar novos pontos como um ou outro
  - Buscar padrões nos dados para entender as diferenças entre círculos e quadrados



Aprendizado Descritivo x Preditivo

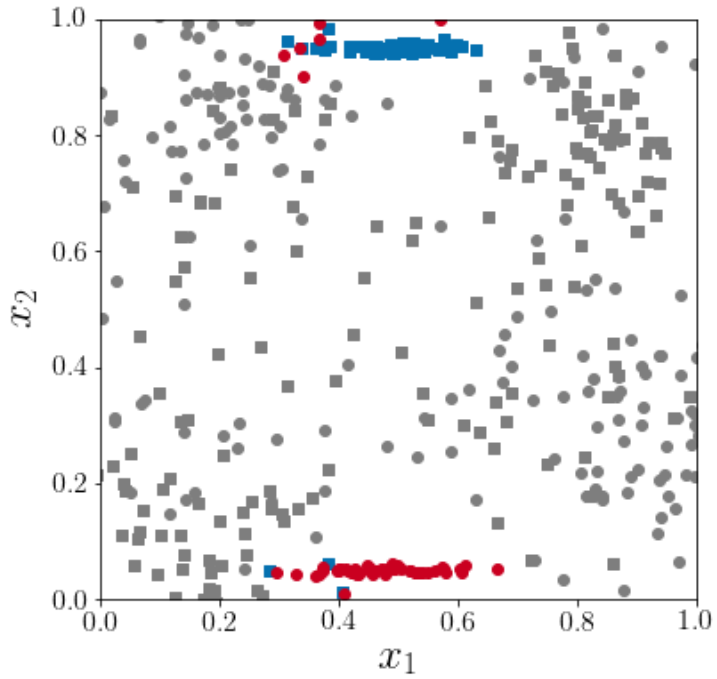
- O primeiro objetivo induz uma abordagem preditiva
- Logo, o abordamos como um problema de classificação





Aprendizado Descritivo x Preditivo

- O segundo objetivo induz a uma abordagem descritiva
- Logo, abordamos o problema como uma tarefa de descoberta de subgrupos



$$\sigma_1 \equiv x_1 \in [0.3, 0.7) \wedge x_2 \geq 0.9$$

$$\sigma_2 \equiv x_1 \in [0.275, 0.7) \wedge x_2 \leq 0.1$$

Aprendizado Descritivo x Preditivo

- Esse último exemplo mostra uma característica que frequentemente diferencia as duas abordagens
- As abordagens preditivas buscam ajustar modelos que aprendam as regularidades globais dos dados
  - No exemplo, encontrar as fronteiras que separam círculos de quadrados
- As abordagens descritivas, por outro lado, ajustam modelos que aprendam regularidades locais dos dados, i.e., padrões válidos apenas a subgrupos específicos
  - No exemplo, intervalos das variáveis que descrevem subgrupos com uma distribuição não usual de círculos e quadrados

Leitura

- Capítulo 3, Flach (2012)

Aprendizado Descritivo

Aula 01 – Introdução: aprendizado descritivo x preditivo

Professor Renato Vimieiro

DCC/ICEx/UFMG