

APRENDIZADO DE MÁQUINA - AM

Definições Básicas

Aprendizado de Máquina

2

- Desde que os computadores foram inventados, é perguntado se eles podem aprender
- Se fosse possível entender como programá-los para aprender, o impacto seria enorme
 - ▣ Poderiam melhorar automaticamente seu desempenho com a experiência

Aprendizado de Máquina

3

- Computadores capazes de aprender poderiam:
 - ▣ Utilizando registros médicos, definir quais tratamentos são mais eficientes para determinadas doenças
 - e determinados perfis de pacientes
 - ▣ Por meio de padrões de uso dos ocupantes de uma casa, definir como reduzir o consumo de energia
 - e melhorar o conforto
 - ▣ Dada a ordem de leitura de um jornal eletrônico, definir como salientar um determinado conjunto de notícias na melhor ordem

Aprendizado de Máquina

4

Na medida que os computadores se tornam mais sofisticados, parece inevitável que AM exercerá um papel central em Ciência da Computação e tecnologia de computadores

□ Tom Mitchell

Objetivos do Estudo de AM

5

- Entender melhor os mecanismos de aprendizado humano pode levar a um melhor entendimento das capacidades (e incapacidades) humanas
- Automatizar processo de aquisição de conhecimento
 - ▣ Modelar computacionalmente processo de aprendizado humano
 - ▣ Abre novas oportunidades para uso de computadores

Aprendizado de Máquina

6

□ Quadro atual

- Capacidade de aprendizado dos computadores de maneira geral ainda é muito inferior àquela dos seres humanos
- Porém, algoritmos de AM eficientes têm sido propostos para várias tarefas
- Entendimento teórico de aprendizado começa a se consolidar

Aplicações de AM

7

- Algoritmos de AM têm sido bem sucedidos, p. ex., para:
 - ▣ Identificar genes associados a determinadas doenças
 - ▣ Discriminar tecidos (saudáveis e doentes), objetos celestiais, ...
 - ▣ Identificar nichos de mercado
 - ▣ Prever a vazão de rios e nível de represas
 - ▣ Detectar uso fraudulento de cartões de crédito
 - ▣ Otimizar ações de controle em processos de produção
 - ▣ Reconhecimento de faces, de voz, de assinaturas ...
 - ▣ e tantas outras...

Exemplo: ALVINN

8



Dean Pomerleau
CMU



ALVINN

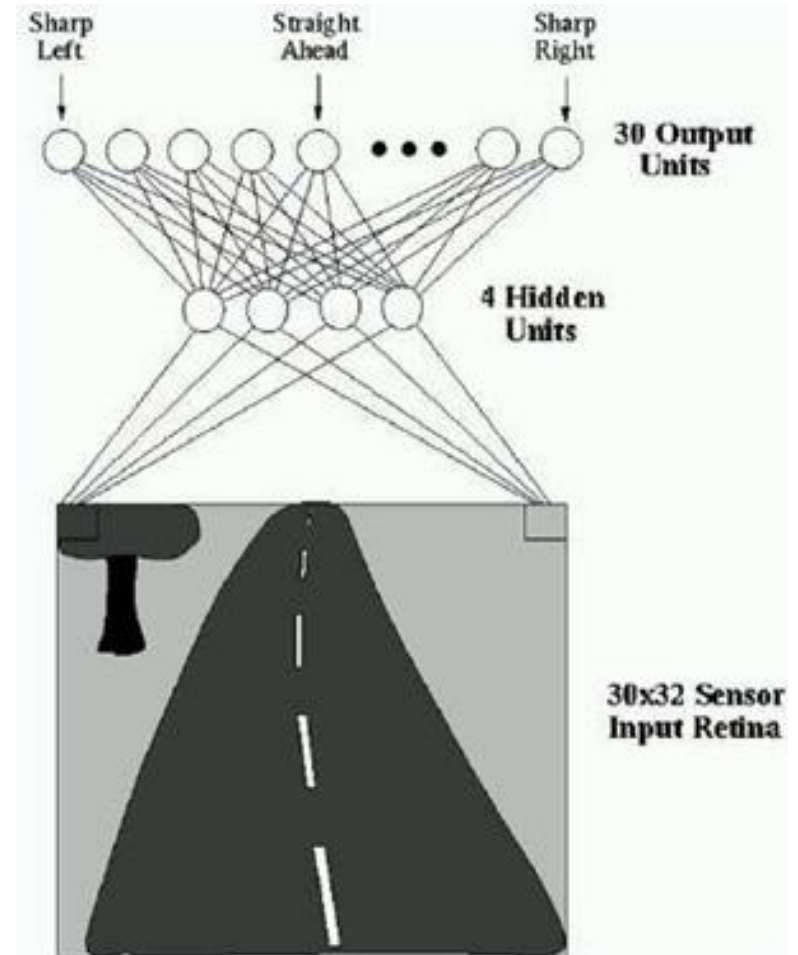
9

- Sistema automático de navegação para automóveis
 - ▣ Baseado em uma câmera montada no veículo
 - ▣ Dirigiu por 90 milhas de forma totalmente autônoma, a 70M/h (110 Km/h) e entre outros carros, em uma rodovia pública americana
 - ver breve descrição e refs. em (Mitchell, 1997) Cap. 1.
- Final da década de 80!
 - ▣ atualmente existem competições entre grandes universidades (Stanford, Carnegie Mellon, ...)

ALVINN

10

- Utiliza uma Rede Neural
 - ▣ 960 entradas
 - Matriz 30x32 derivada dos pixels de uma imagem
 - ▣ 4 unidades intermediárias
 - ▣ 30 unidades de saída
 - Cada uma representando um comando para a direção



Aplicações de AM

11

- <http://blogs.estadao.com.br/link/inteligencia-artificial-e-a-grande-aposta-de-empresas-de-tecnologia/>
- <http://admin-apps.webofknowledge.com/JCR/JCR?RQ=HOME>
- <http://aquare.la/pt/>
- <http://datastorm.com.br/>
- <http://convergenciadigital.uol.com.br/cgi/cgilua.exe/sys/start.htm?UserActiveTemplate=site&infoid=37042&sid=10>

Aplicações de AM

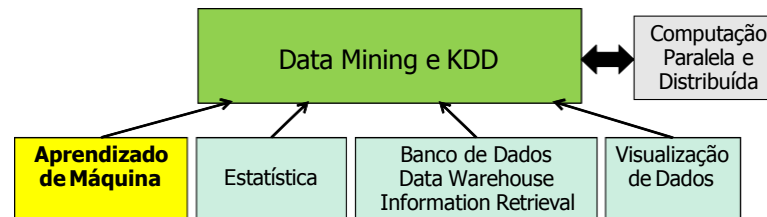
12

- Uma das áreas de aplicação mais importantes de AM têm sido a descoberta de conhecimento em grandes bases de dados (KDD)
 - ▣ Mineração de Dados (Data Mining – DM)
 - ▣ Exemplos:
 - Registros de compras em grandes supermercados
 - Registros de empréstimos financeiros
 - Registros de transações financeiras
 - Registros médicos
 - Projetos genoma
 - ...

AM, DM e KDD

13

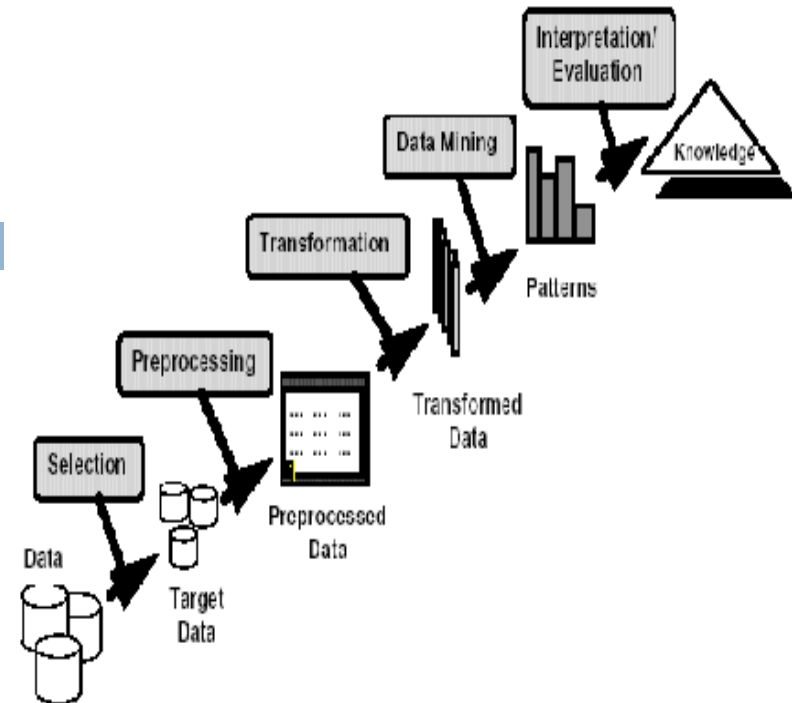
- Muitos utilizam os termos Aprendizado de Máquina e Data Mining de maneira indiscriminada, mas eles se referem a conceitos bem diferentes:



- DM é usualmente feito utilizando AM, mas nem sempre.

AM, DM e KDD

14



□ Já KDD é usualmente definido como um ciclo que envolve DM:

- Formalmente, é o processo não trivial de identificar padrões válidos, novos, potencialmente úteis e compreensíveis em dados:
 - Advances in Knowledge Discovery and Data Mining, Fayyad, Piatetsky-Shapiro, Smyth, and Uthurusamy, (Chapter 1), MIT Press, 1996

□ Mas AM não é um sub-conjunto de DM ou KDD!!!

- Reconhecimento de fala, voz, assinaturas, controle, robótica, ...

AM x Reconhecimento de Padrões

15

- Denotam áreas de pesquisa muito semelhantes
 - ▣ Diferenças principais estão na terminologia e nas áreas de origem (motivações, aplicações, ...)

AM	RP
Computação	Engenharia
Atributo	Característica
Exemplo	Padrão
...	...

Reconhecimento de Padrões

16

- Definição: técnica de tomada de decisões a partir de padrões complexos de informações
- Objetivo Principal:
 - ▣ Desenvolvimento de sistemas capazes de lidar com tarefas realizadas por seres humanos
- Função:
 - ▣ Reconhecimento de padrões de entrada como pertencentes ou não a uma ou mais dentre possíveis categorias, que podem ser previamente conhecidas ou não
 - Reconhecimento de voz, face, assinatura, escrita, ...
 - Mapeamento robótico autônomo, ...

Por que AM é importante?

17

- Algumas tarefas não podem ser bem executadas sem que seja por meio de exemplos
 - ▣ Ex.: Reconhecer pessoas
- Ser humano não é capaz de explicar (e portanto de programar) sua habilidade para executar alguns tipos de tarefas
 - ▣ Ex.: Andar de bicicleta
- Quantidade de conhecimento disponível pode ser muito grande para ser descrito (e portanto programado) por humanos
 - ▣ Ex.: Diagnóstico médico

Por que AM é importante?

18

- Ser humano não é capaz de executar algumas tarefas que demandam quantidades grandes de cálculos complexos, passíveis apenas de execução em computador:
 - ▣ P. ex., detectar inter-relacionamentos / correlações escondidas em grandes quantidades de dados
 - AM e MD podem encontrar essas relações
- Ambientes (e perfis dos problemas) podem mudar com o tempo
 - ▣ Capacidade de adaptação automática é fundamental

Por que estudar AM?

19

- Esse é o momento
 - ▣ Vários algoritmos efetivos e eficientes estão disponíveis
- Grande quantidade de dados disponíveis online
- Elevada capacidade dos recursos computacionais disponíveis

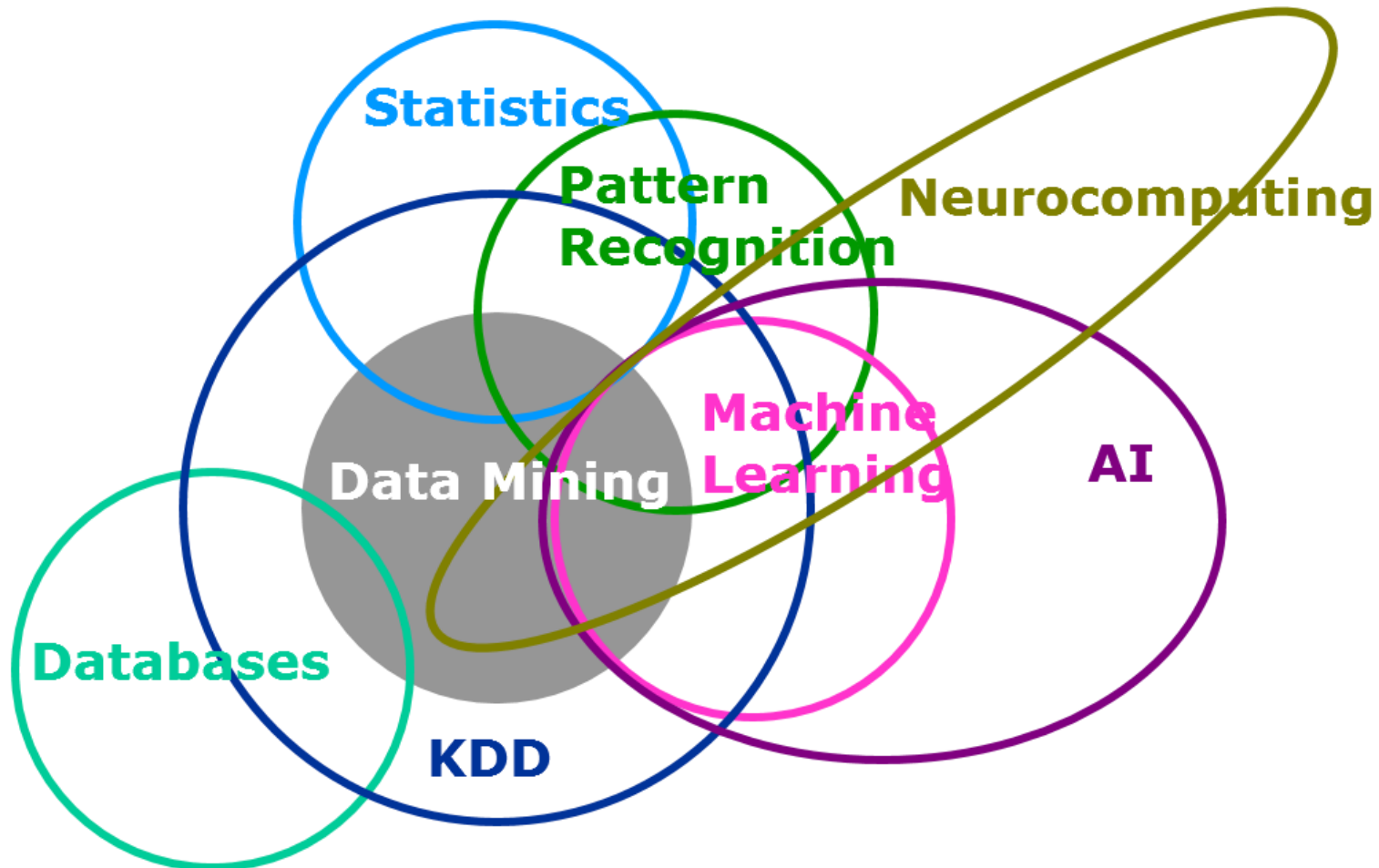
Disciplinas Relacionadas

20

- Inteligência Artificial
- Mineração de Dados
- Teoria da Informação
- Estatística (Bayesiana e Frequentista)
- Otimização
- Estimação, Filtragem e Controle
- Teoria da Complexidade Computacional
- Neurociências
- ...

Disciplinas Relacionadas

21



Definição de AM

22

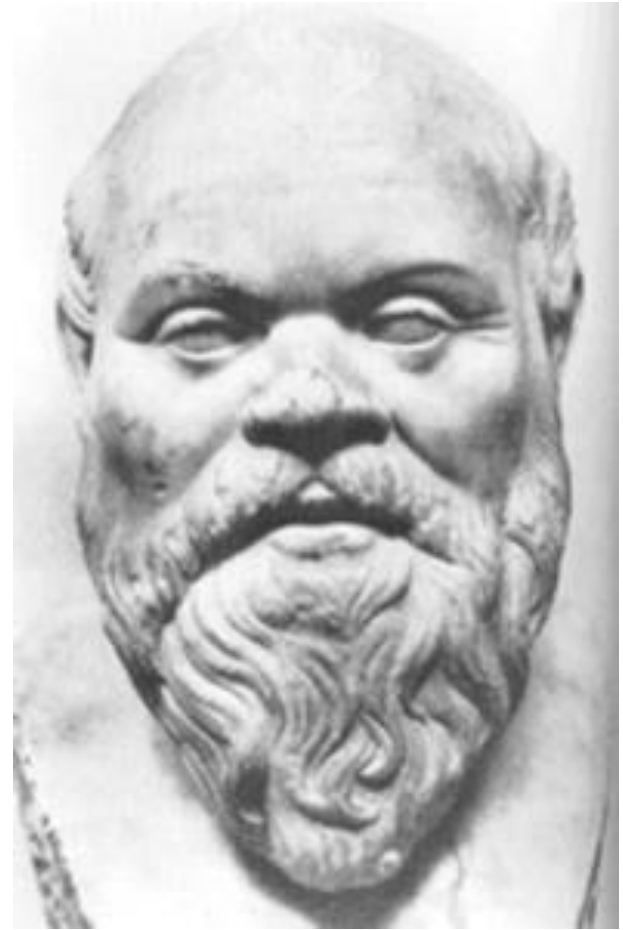
- Uma máquina de aprendizado, definida de maneira ampla, é qualquer dispositivo cujas ações são influenciadas por experiências anteriores (Nilsson 1965)
- Qualquer mudança em um sistema que o permita ter um melhor desempenho na segunda vez em que ele repita uma mesma tarefa ou outra retirada da mesma população (Simon 1983)
- Modificação de uma tendência comportamental por meio de experiência (Webster 1984)
- Uma melhoria na capacidade de processar informação a partir da atividade de processar informação (Tanimoto 1990)

Definição de Aprendizado

23

Aprender é Recordar

(Sócrates: Diálogos de Platão)



Definição de AM

24

□ Definição formal (Mitchell, 1997):

Um programa de computador é dito **aprender** a partir de uma experiência ***E*** com respeito a alguma classe de tarefas ***T*** e medida de desempenho ***P***, se seu desempenho em tarefas de ***T***, medido por ***P***, melhora com a experiência ***E***

Problemas Bem Definidos

25

Um programa de computador é dito **aprender** a partir de uma experiência ***E*** com respeito a alguma classe de tarefas ***T*** e medida de desempenho ***P***, se seu desempenho em tarefas de ***T***, medido por ***P***, melhora com a experiência ***E***

Um programa que aprende a jogar xadrez podia melhorar seu desempenho, *medido pela capacidade de vencer* na classe de tarefas envolvendo *jogar xadrez*, pela experiência obtida jogando *xadrez contra ele mesmo*

Problemas Bem Definidos

26

- Três características devem ser identificadas para um problema ser bem definido:
 - ▣ A classe de tarefas
 - ▣ A medida de desempenho a ser melhorada
 - ▣ A origem da experiência

Exemplo 1

27

- Problema de Aprender Xadrez:
 - ▣ **Tarefa T:** jogar xadrez
 - ▣ **Medida de desempenho P:** porcentagem dos jogos vencidos contra adversários
 - ▣ **Experiência de treinamento E:** praticar jogando contra si próprio ou contra adversários humanos (p.ex. internet)

Exemplo 2

28

- Problema de Filtrar SPAMs:
 - ▣ **Tarefa T:** categorizar mensagens de e-mail como spam ou legítima
 - ▣ **Medida de desempenho P:** porcentagem de mensagens de e-mail corretamente classificadas
 - ▣ **Experiência de treinamento E:** conjunto de e-mails rotulados por seres humanos

Exercício 1

29

- Problema de Reconhecimento de Escrita Manual:
 - ▣ **Tarefa T:**
 - ▣ **Medida de desempenho P:**
 - ▣ **Experiência de treinamento E:**

Exercício 1

30

- Problema de Reconhecimento de Escrita Manual:
 - ▣ **Tarefa T:** reconhecer e classificar palavras manuscritas dentro de imagens
 - ▣ **Medida de desempenho P:** porcentagem de palavras corretamente classificadas
 - ▣ **Experiência de treinamento E:** base de dados de palavras manuscritas com as respectivas classificações

Exercício 2

31

- Problema de um Robô Aprender a Dirigir
 - ▣ **Tarefa T:**
 - ▣ **Medida de desempenho P:**
 - ▣ **Experiência de treinamento E:**

Exercício 2

32

- Problema de um Robô Aprender a Dirigir
 - ▣ **Tarefa T:** dirigir em uma rodovia pública usando sensores de visão
 - ▣ **Medida de desempenho P:** distância média percorrida antes de um erro (definido por um observador humano)
 - ▣ **Experiência de treinamento E:** uma sequência de imagens e comandos de direção registrados observando um motorista humano

Exercício 3

33

- Problema de Diagnóstico Médico
 - ▣ Tarefa T:
 - ▣ Medida de desempenho P:
 - ▣ Experiência de treinamento E:

Exercício 3

34

- Problema de Diagnóstico Médico
 - ▣ **Tarefa T:** diagnosticar o estado de um paciente dado um conjunto de sintomas
 - ▣ **Medida de desempenho P:** porcentagem de pacientes corretamente diagnosticados
 - ▣ **Experiência de treinamento E:** base de dados de pacientes com seus respectivos diagnósticos corretos

Projetando um Sistema de Aprendizado

35

- Escolha da experiência de treinamento
 - ▣ Pode oferecer dois tipos de retorno (feedback)
 - Direto
 - Sistema aprende diretamente dos exemplos de treinamento
 - Ex.: aprender a jogar xadrez diretamente dos estados dos tabuleiros e seus movimentos corretos
 - Indireto
 - Informação é disponibilizada de forma indireta
 - Ex.: sequências de movimentos e resultado final de várias partidas de xadrez

Projetando um Sistema de Aprendizado

36

- Problema de atribuição de crédito (credit assignment)
 - ▣ Comum em aprendizado indireto
 - ▣ Definir a culpa ou crédito para um dado resultado
 - ▣ Ex.: Como definir o movimento que levou à vitória / derrota em um jogo de xadrez?
 - Jogo pode ser perdido mesmo que as jogadas iniciais sejam ótimas
- Aprendizado direto é geralmente mais fácil

Projetando um Sistema de Aprendizado

37

- Principais paradigmas de treinamento
 - ▣ Supervisionado
 - ▣ Semi-supervisionado
 - ▣ Não supervisionado
 - ▣ Reforço

Treinamento Supervisionado

38

- Guiado por um “professor” externo
 - ▣ “Professor” possui conhecimento sobre o ambiente
 - Representado por conjunto de pares (x, d)
 - ▣ Parâmetros do modelo são ajustados por apresentações sucessivas de pares (x, d)
 - Modelo procura reproduzir comportamento do “professor”

Treinamento Supervisionado

39

- Após seu treinamento, desempenho do sistema deve ser testado com novos dados:
 - ▣ Dados de teste
 - ▣ $D_{\text{teste}} \cap D_{\text{treinamento}} = \emptyset$

Treinamento por Reforço

40

- Guiado por um “crítico” externo
 - ▣ Processo de tentativa e erro
 - ▣ Procura maximizar sinal de reforço

- Se ação tomada por sistema é seguida por estado satisfatório, sistema é fortalecido, caso contrário, sistema é enfraquecido (lei de Thorndike)

Treinamento por Reforço

41

- Tipos de reforço
 - ▣ Positivo = recompensa
 - ▣ Negativo = punição
 - ▣ Nulo

Supervisionado X Reforço

42

Treinamento Supervisionado	Treinamento por Reforço
Professor	Crítico
Sistema de feedback É dito o que fazer Mais rápido	Sistema de feedback Faz e vê o que acontece Mais lento

Treinamento Não Supervisionado

43

- Não tem crítico ou professor externo
 - ▣ Extração de propriedades estatisticamente relevantes
 - ▣ Exemplos:
 - Clustering: descobre categorias automaticamente
 - Quantização: sumariza dados em grânulos automaticamente

Treinamento Semi-Supervisionado

44

- Tem um professor externo apenas para parte dos exemplos de treinamento
 - ▣ Exemplo:
 - Web mining: usuários podem fornecer alguns exemplos de páginas similares, pertencentes a uma determinada categoria, mas uma parcela ínfima de web pages teria essa informação associada

Projetando um Sistema de Aprendizado

45

- Representatividade dos exemplos
 - ▣ Aprendizado é mais confiável quando exemplos de treinamento seguem uma distribuição representativa (semelhante) à da população
 - ▣ Ex.: Jogo de xadrez
 - Conjunto de treinamento formado apenas por exemplos de jogo contra o próprio programa
 - Experiência de treinamento pode não ser representativa

Projetando um Sistema de Aprendizado

46

- Tipo exato de conhecimento a ser aprendido
 - ▣ Função alvo
 - Estabelece qual conhecimento será aprendido e permite verificar quão bem ele foi aprendido
 - Exemplos:
 - Função discriminante entre classes
 - Função de similaridade intra grupos
 - ...

Escolha da Função Alvo

47

- Exemplo:
 - ▣ Aprender a diagnosticar paciente de diabetes
 - Função = mapeamento entre características dos pacientes e os valores “diabético” e “não diabético”
 - ▣ Como aprender a função?
 - Ajustá-la aos dados disponíveis
 - ▣ Como determinar o desempenho da função aprendida?
 - Verificar quantos pacientes ela diagnostica corretamente

Modelos de Representação

48

- Modelos Matemáticos
 - ▣ Regressão linear,
 - ▣ Redes neurais (paradigma conexionista / bioinspirado),
 - ▣ Máquinas de vetores de suporte, ...
- Modelos Simbólicos
 - ▣ Árvores de decisão,
 - ▣ Regras em lógica proposicional ou de 1ª ordem,
 - ▣ Redes semânticas, ...

Modelos de Representação

49

- Modelos “Lazy” (paradigma baseado em instâncias)
 - ▣ K-NN,
 - ▣ Raciocínio Baseado em Casos (CBR), ...
- Modelos Probabilísticos (paradigma probabilístico)
 - ▣ Naive Bayes,
 - ▣ Redes Bayesianas,
 - ▣ Misturas de Gaussianas,
 - ▣ Modelos de Markov Escondidos (HMMs), ...

Técnicas de Aprendizado

50

- Dado um tipo de modelo, uma função alvo e um conjunto de exemplos de treinamento, é preciso algum mecanismo para obter, a partir dos exemplos, um modelo específico daquele tipo que represente bem a função alvo
- Esse mecanismo, denominado mecanismo de aprendizagem, consiste fundamentalmente de uma técnica de busca.
 - Busca-se no espaço dos modelos plausíveis por aquele modelo específico que melhor represente a função alvo

Técnicas de Aprendizado

51

- Algoritmos Baseados em Gradiente
 - ▣ Regressão Linear
 - ▣ Redes Neurais
 - ▣ ...
- Programação Dinâmica
 - ▣ Aprendizado para HMMs
 - ▣ ...

Técnicas de Aprendizado

52

- Dividir para Conquistar
 - ▣ Indução de árvores e regras de decisão

- Estimação de Probabilidades (paradigma probabilístico)
 - ▣ Naive Bayes, Redes Bayesianas, ...

- Computação Evolutiva (paradigma evolucionário / bioinspirado)
 - ▣ Aplicável a vários modelos...

Modelos e Técnicas

53

- Cada tipo de modelo é mais apropriado para uma determinada classe de problemas
 - ▣ Assim como cada técnica de aprendizado é mais apropriada para um tipo de modelo
 - ▣ É parte importante do estudo de AM aprender a identificar os cenários mais apropriados para cada modelo e técnica de aprendizado
- O modelo e a técnica estabelecem algo fundamental a ser formalizado mais adiante no curso:
 - ▣ o **bias indutivo** do aprendizado

Bias Indutivo

54

- Informalmente, o bias indutivo de um sistema de AM é uma tendência a privilegiar um dado conjunto de hipóteses e/ou determinadas hipóteses de um conjunto
 - ▣ Podemos ver uma hipótese como uma realização particular de um modelo para aproximar uma dada função alvo específica
- Definiremos bias indutivo de maneira mais formal em um momento posterior do curso

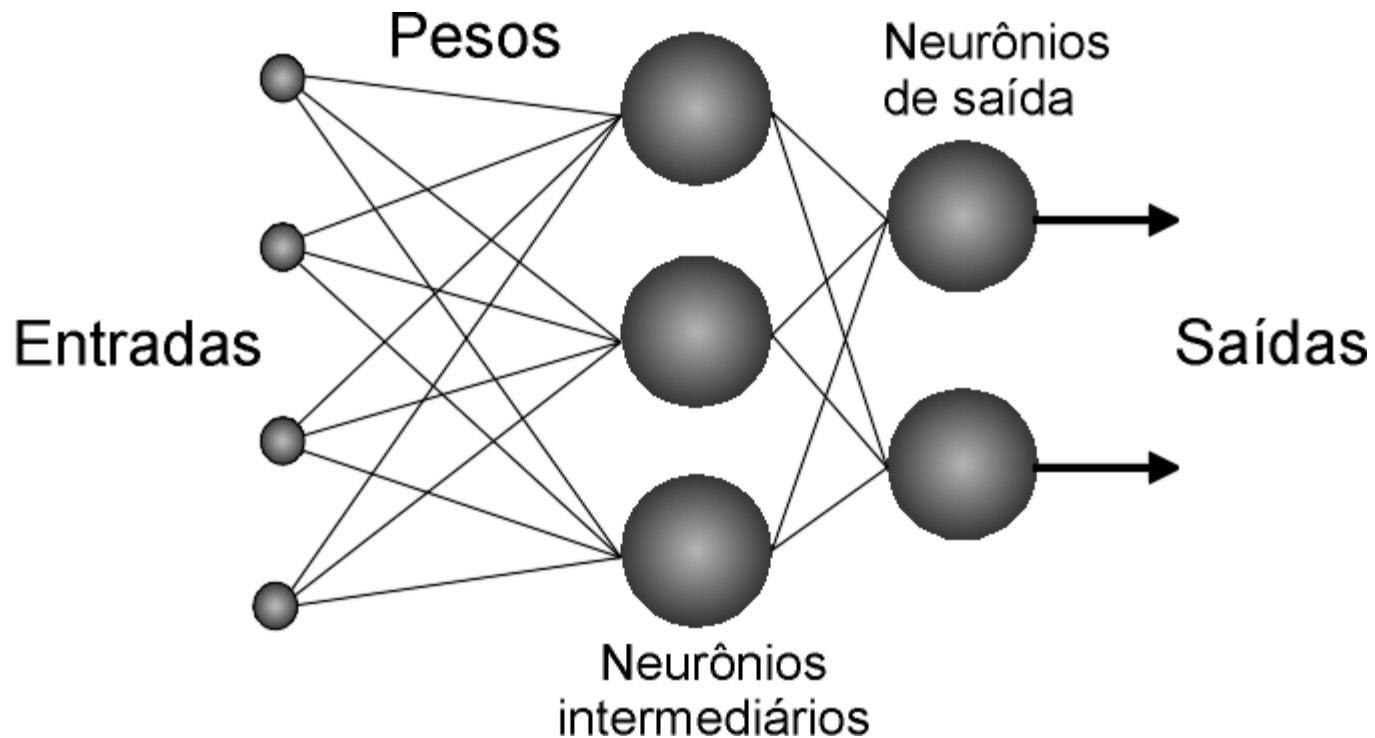
Exemplos de Modelos

55

- Regra de Decisão (paradigma simbólico):
 - ▣ **Se** Faixa_Etária = Centenário **Então** Rejeitar_Empréstimo
- Redes Bayesianas (paradigma probabilístico):
 - ▣ $P(\text{Bom_Pagador} \mid \text{salário, histórico_de_crédito, ...}) = 0.7$
- Rede Neurais (paradigma conexionista / bioinspirado):
 - ▣ Modelo matemático conexionista
- Vizinhos mais Próximos (paradigma baseado em instâncias):
 - ▣ Padrão pertence à classe majoritária do(s) padrão(ões) mais semelhante(s) a ele

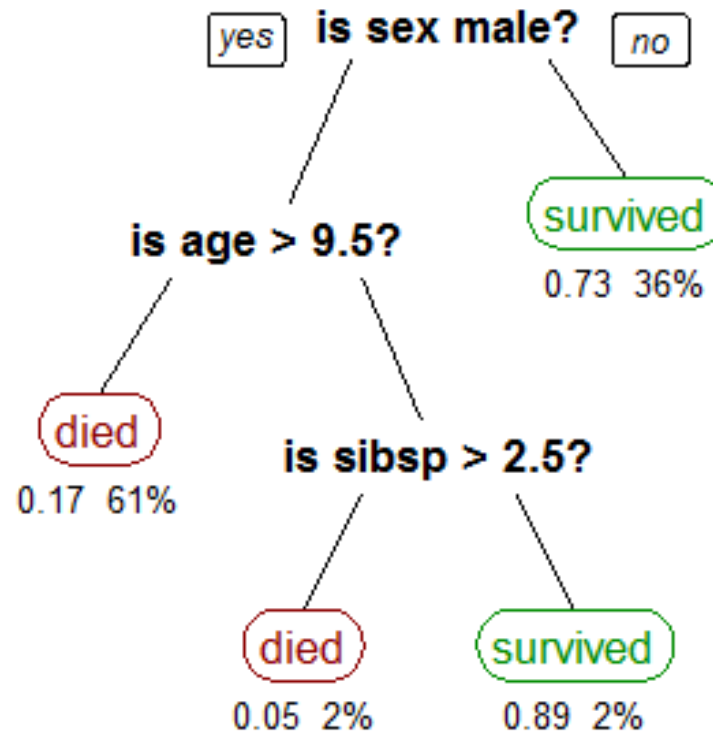
Exemplos de Modelos

56

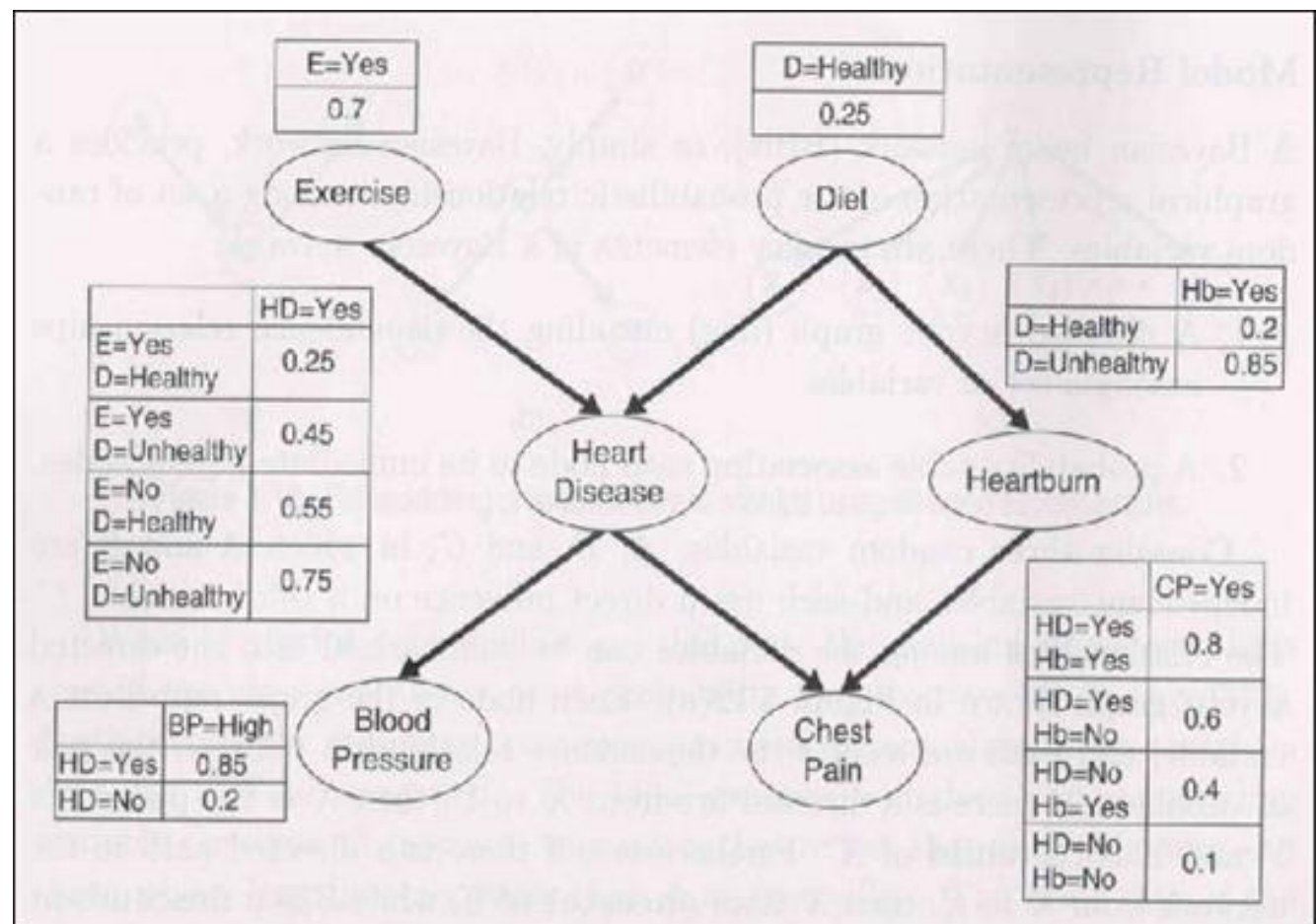


Exemplos de Modelos

57



□ Rede Bayesiana



Avaliação de AM

59

- Uma vez obtido um modelo a partir de exemplos de treinamento e de uma técnica de aprendizado, é preciso **avaliar** a eficácia / eficiência deste modelo / técnica para resolver a tarefa em questão
 - ▣ Em outras palavras, é preciso **validar** ou não o modelo obtido
- A validação de um modelo deve seguir uma metodologia rigorosa, seja sob a ótica **experimental** ou **teórica**

Avaliação de AM

60

- Avaliação Experimental
 - ▣ Conduzir experimentos controlados
 - Dados reais representativos em aplicações práticas
 - Dados benchmark em estudos acadêmicos e comparações
 - ▣ Extrair resultados de desempenho
 - Ex.: Acurácia de teste, tempos de treinamento e teste, etc
 - ▣ Analisar resultados e diferenças com rigor estatístico

Avaliação de AM

61

□ Avaliação Teórica

▣ Analisar algoritmos matematicamente e provar:

- Complexidade computacional
- Habilidade para ajustar dados de treinamento
- Habilidade para generalizar dados de treinamento
- Complexidade da amostra
 - Ordem de grandeza do no. de exemplos de treinamento necessários para aprender uma função com dada acurácia

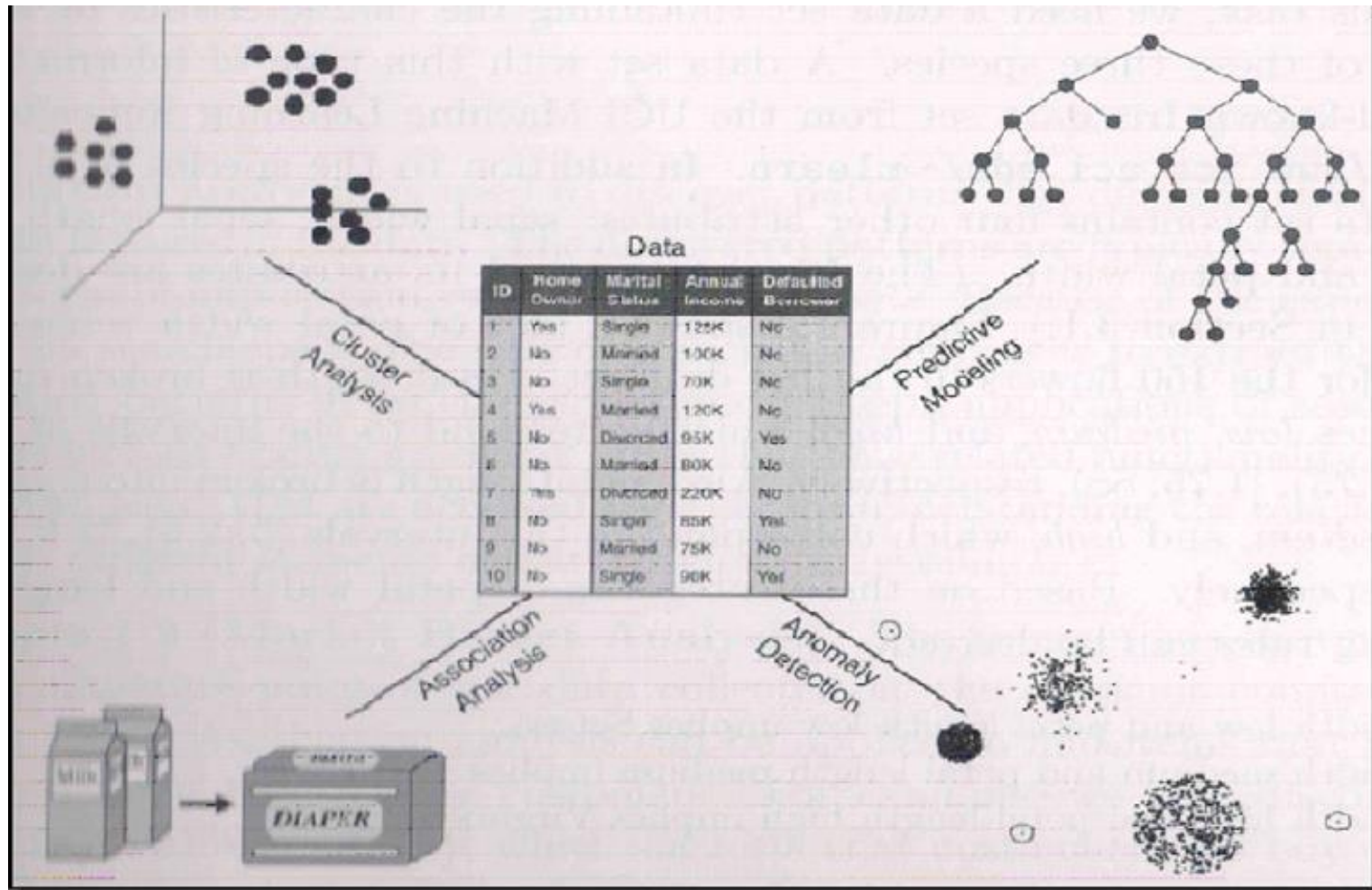
Grandes Áreas de Aplicação de AM

62

- Predição
 - ▣ Classificação de Padrões
 - ▣ Regressão
 - ▣ Detecção de Anomalias
- Mineração de Dados Descritiva
 - ▣ Análise de Agrupamentos
 - ▣ Análise de Associação
 - ▣ Indução de Regras
- Otimização
- Automação e Processamento de Sinais

Tarefas Fundamentais de DM e AM

63



Tarefas Preditivas

64

- Sob o ponto de vista de aproximação de funções, significa encontrar uma estimativa F' de uma função desconhecida F
 - ▣ Conjunto de pares de entrada-saída conhecido $\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$, onde $y = F(x)$
 - ▣ Paradigma de treinamento supervisionado

- Tarefas fundamentais:
 - ▣ Classificação (função discreta)
 - ▣ Regressão (função contínua)

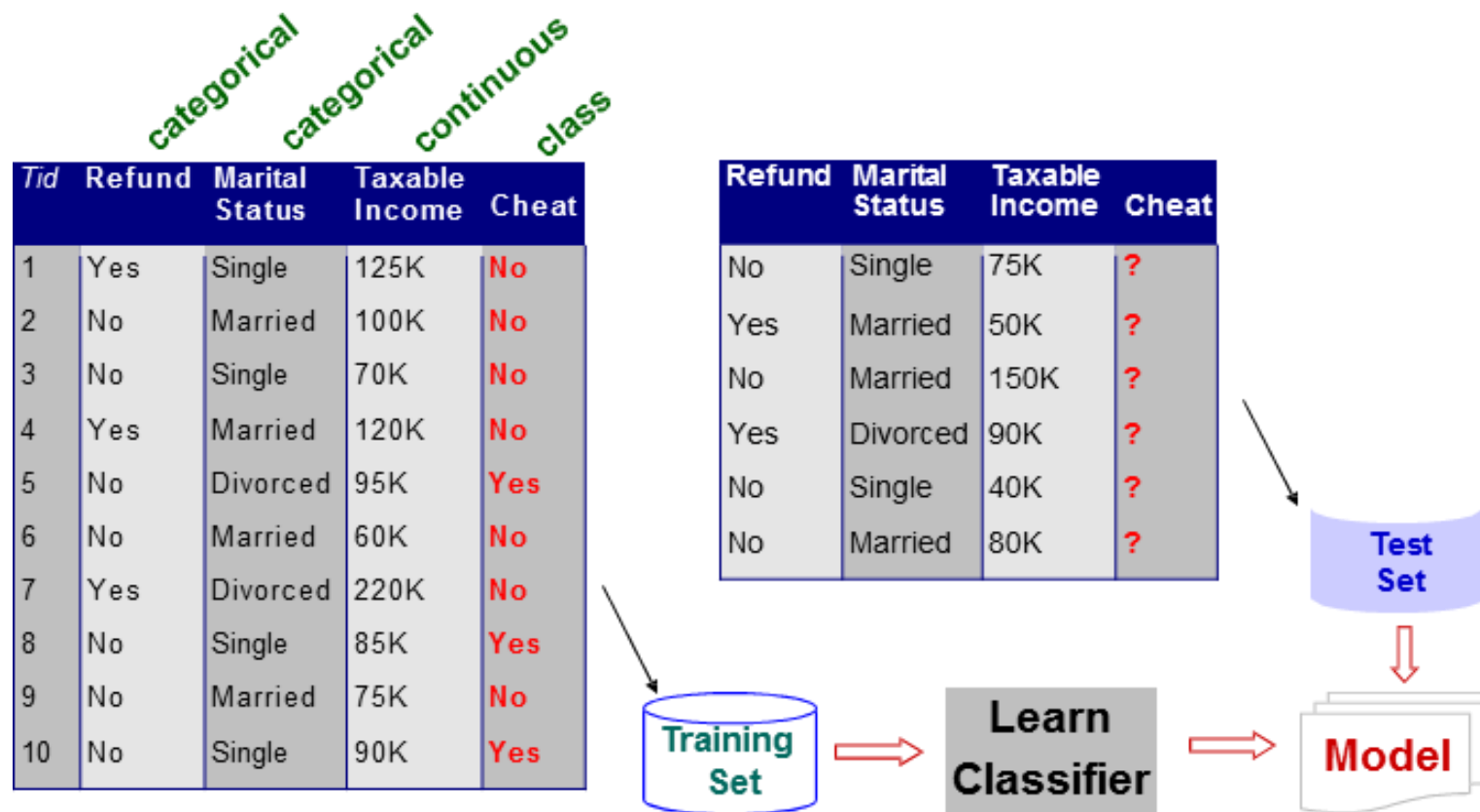
Classificação

65

- Técnica classifica novas entradas (padrões) em uma dentre várias classes discretas
 - ▣ Número definido de classes
- Exemplos
 - ▣ Diagnóstico
 - ▣ Análise de crédito
 - ▣ ...
- Se regras com semântica são extraídas que explicam como os padrões são classificados, também pode ser vista como uma tarefa descritiva, além de preditiva

Classificação

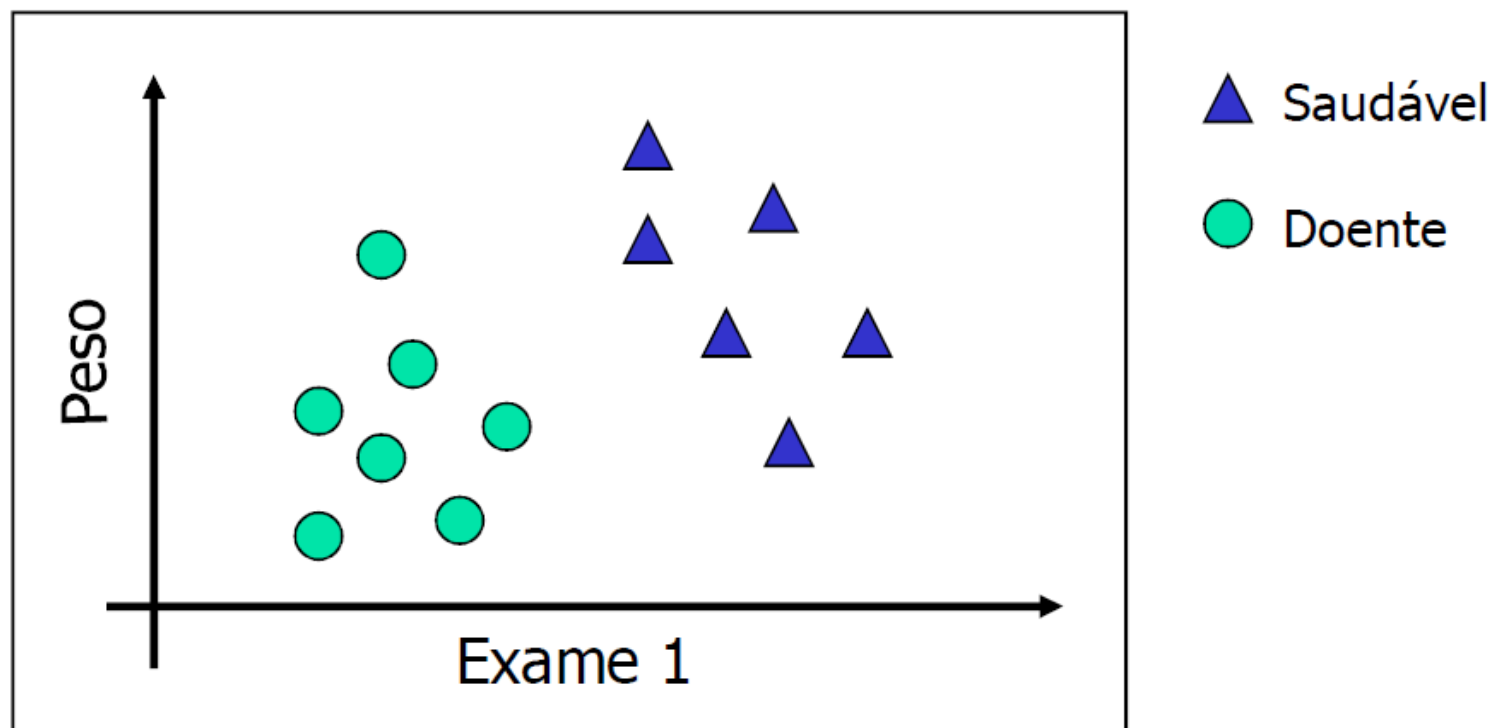
66



Classificação

67

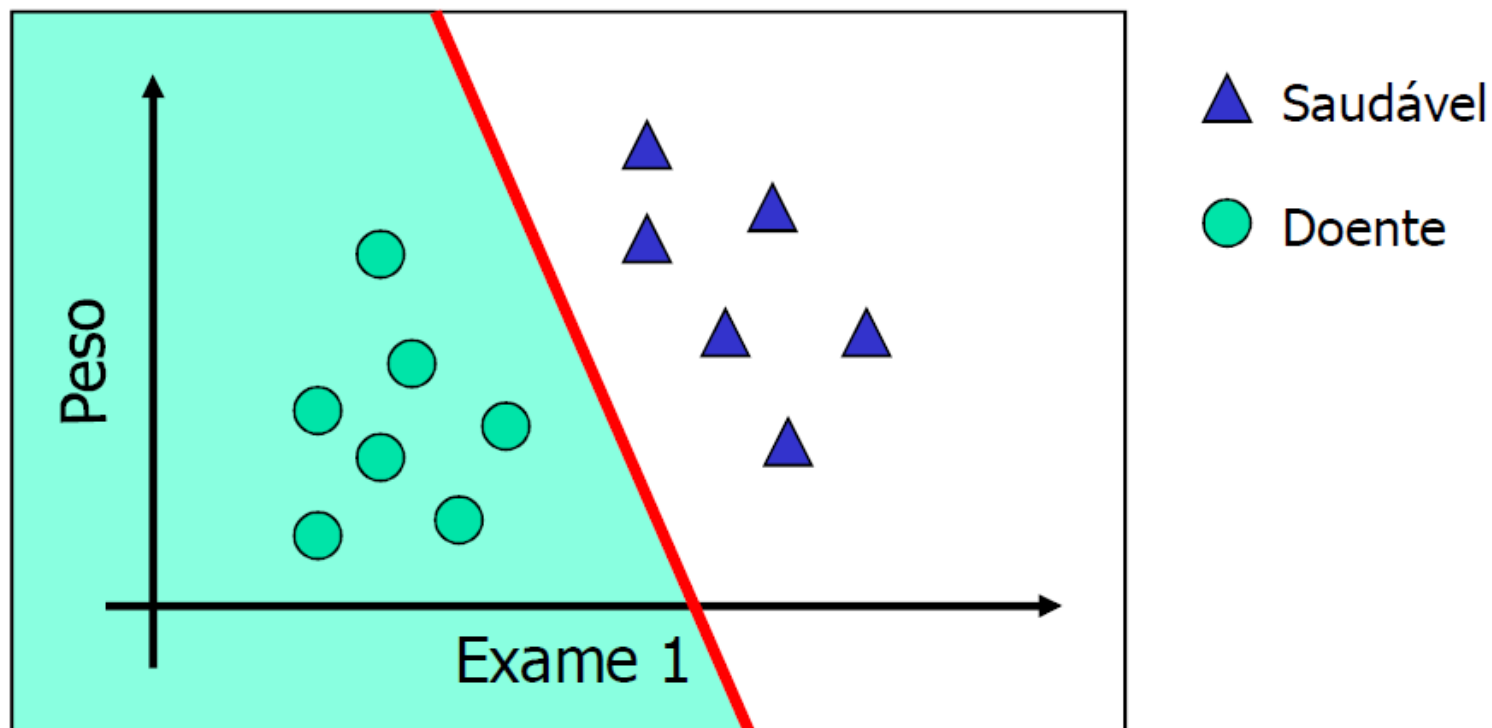
□ Como classificar?



Classificação

68

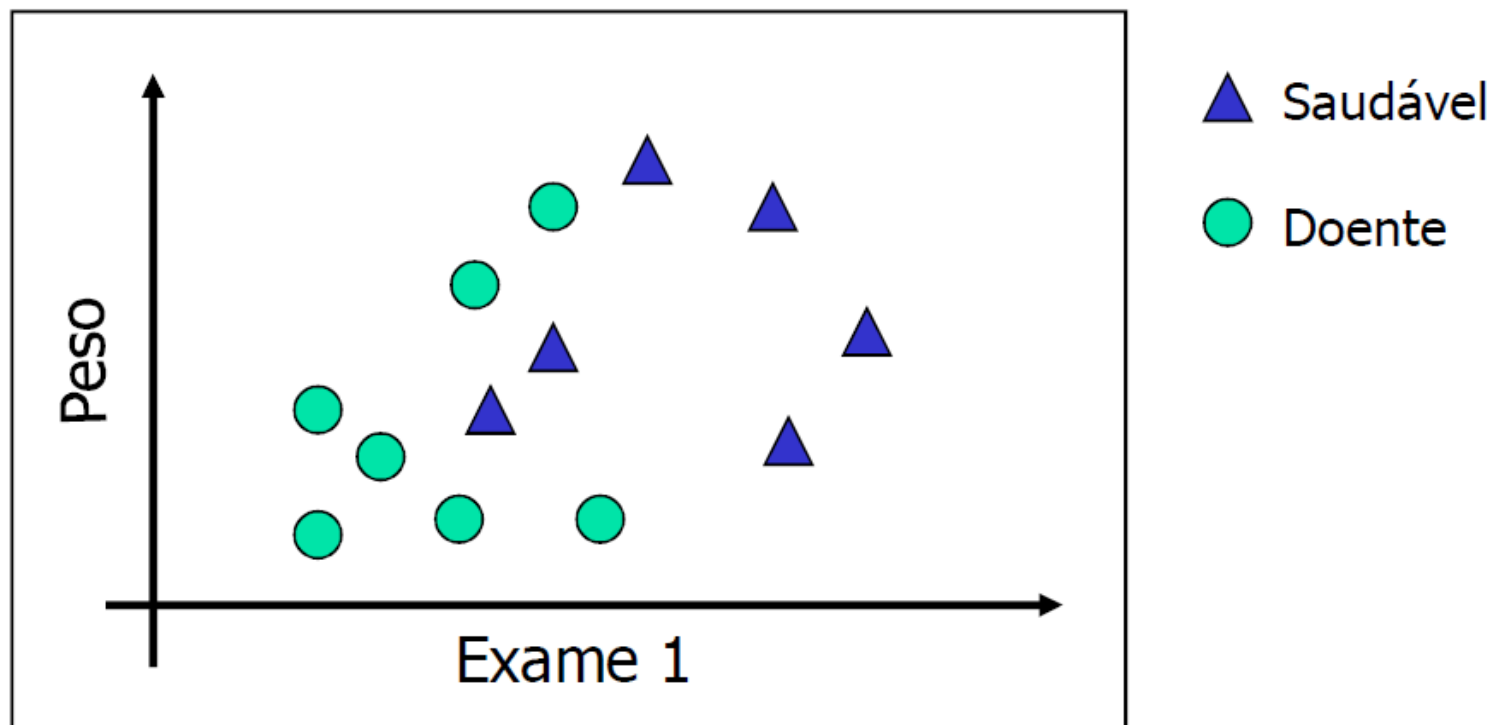
□ Problema linear



Classificação

69

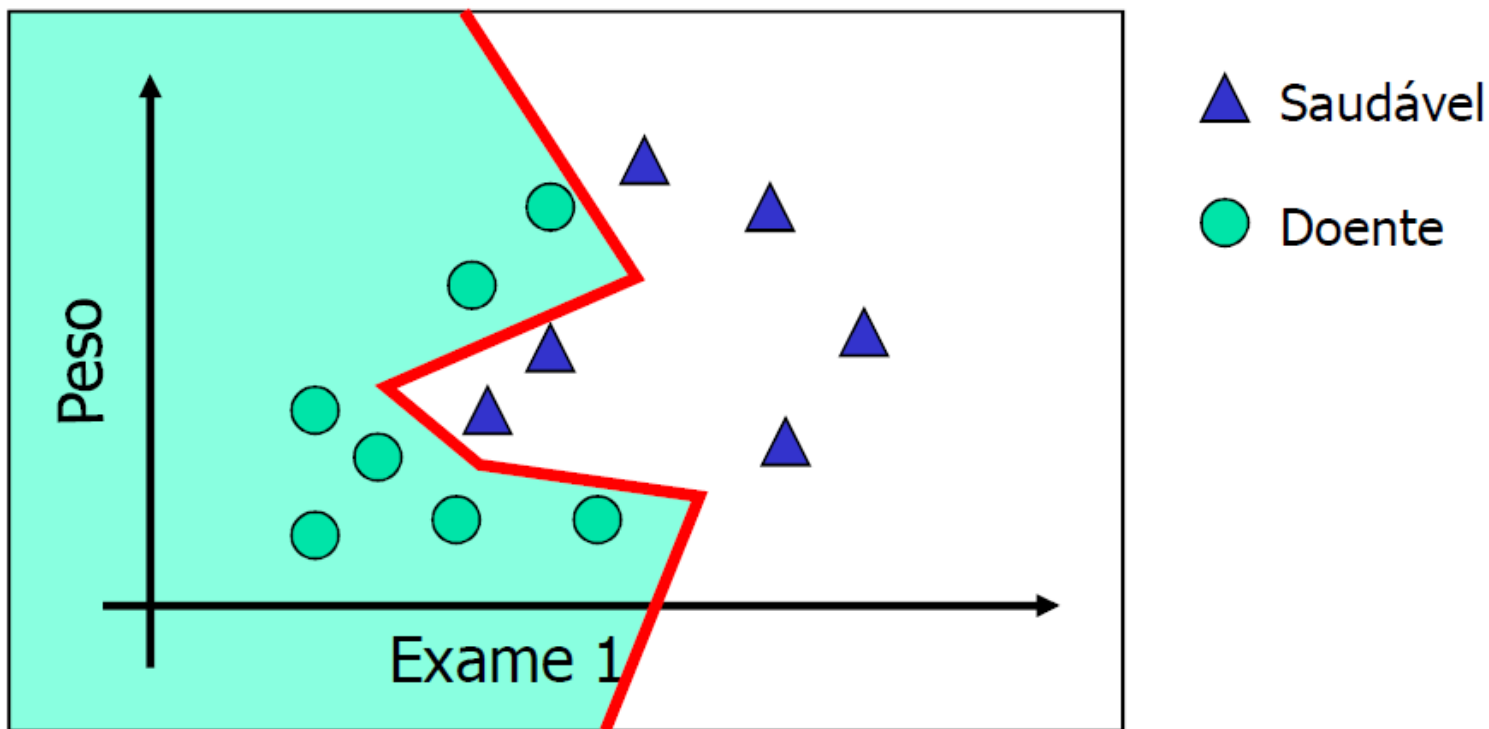
□ Como classificar?



Classificação

70

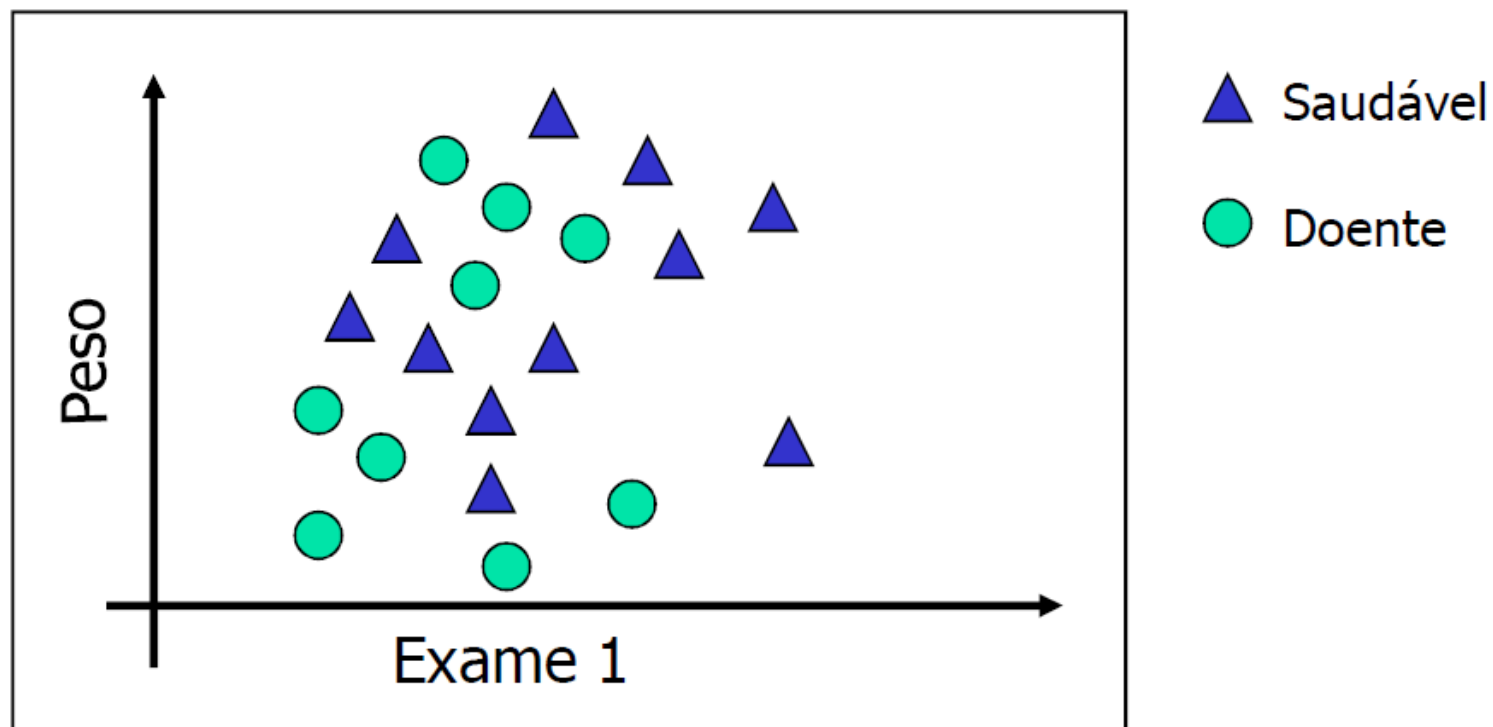
□ Problema não linear



Classificação

71

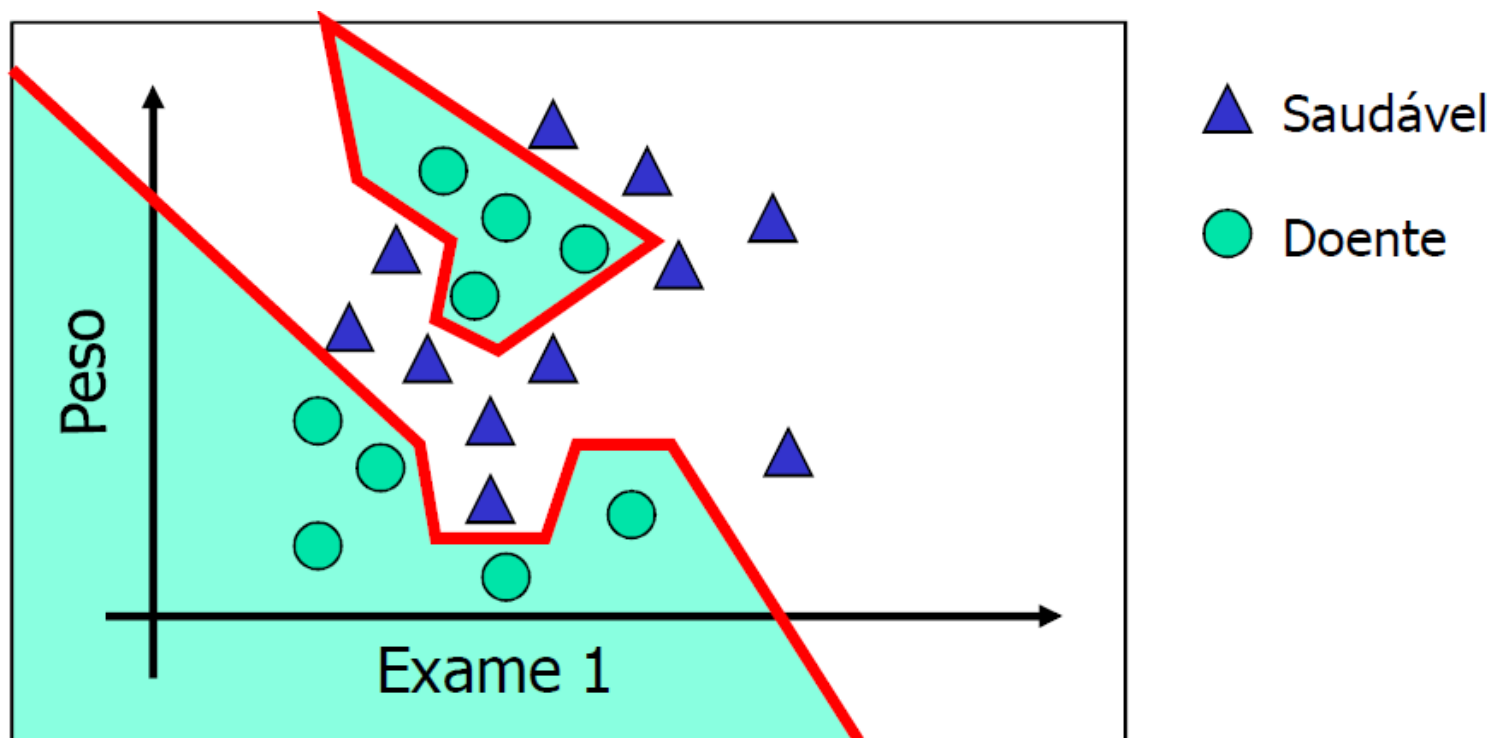
□ Como classificar



Classificação

72

□ Problema não linear



Regressão

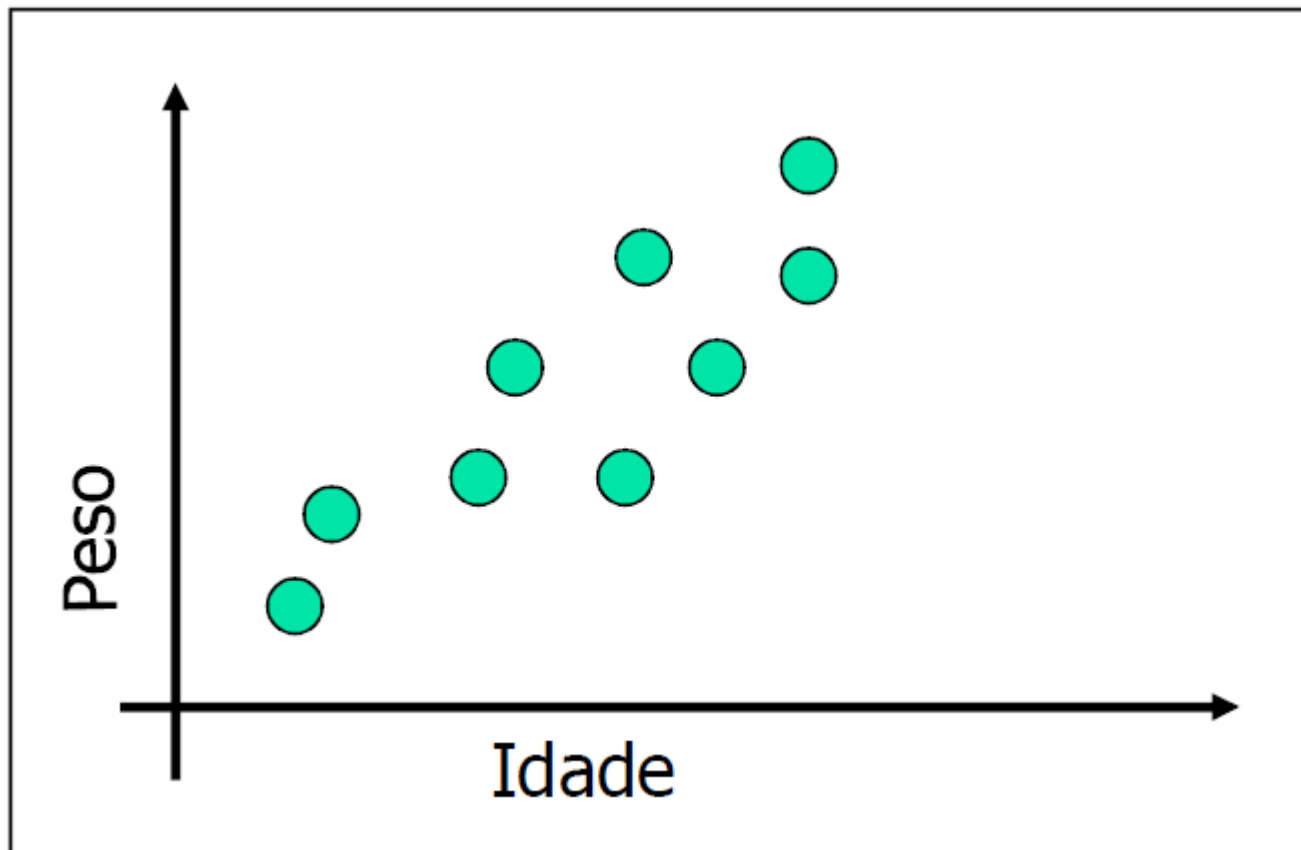
73

- Ao contrário de classificação, a saída em regressão não é limitada a uma quantidade pré-definida de valores
 - ▣ Aprende uma função definida em termos de alguma extrapolação dos exemplos conhecidos
 - ▣ Embora contínuas, variáveis podem ser discretas no tempo
- Exemplos
 - ▣ Previsão de séries temporais
 - ▣ Previsão de sistemas dinâmicos
 - ▣ Fusão de sensores
 - ▣ ...

Regressão

74

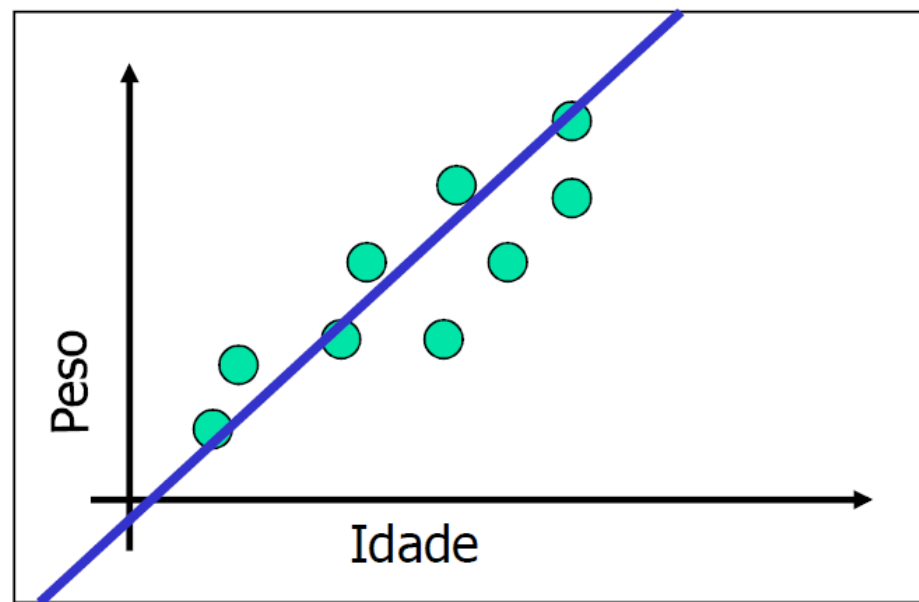
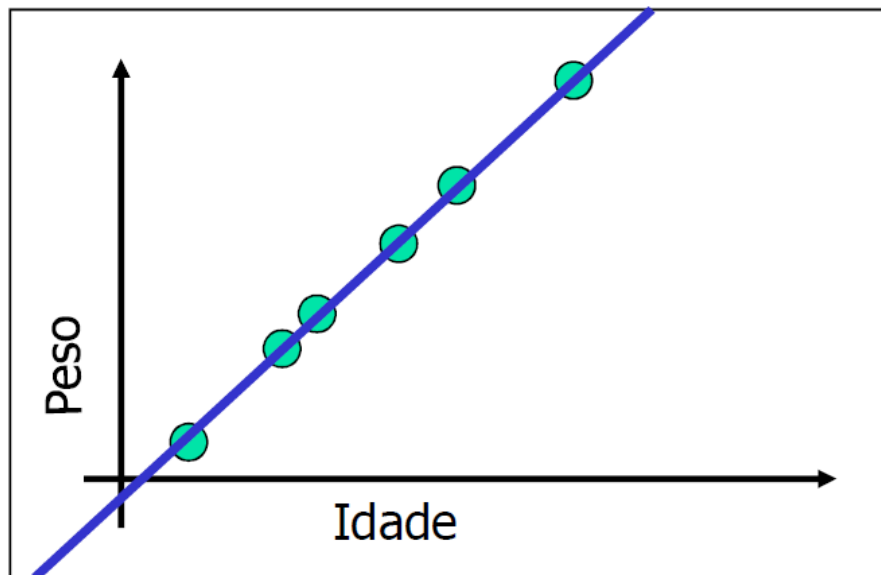
- Como definir o peso em função da idade?



Regressão

75

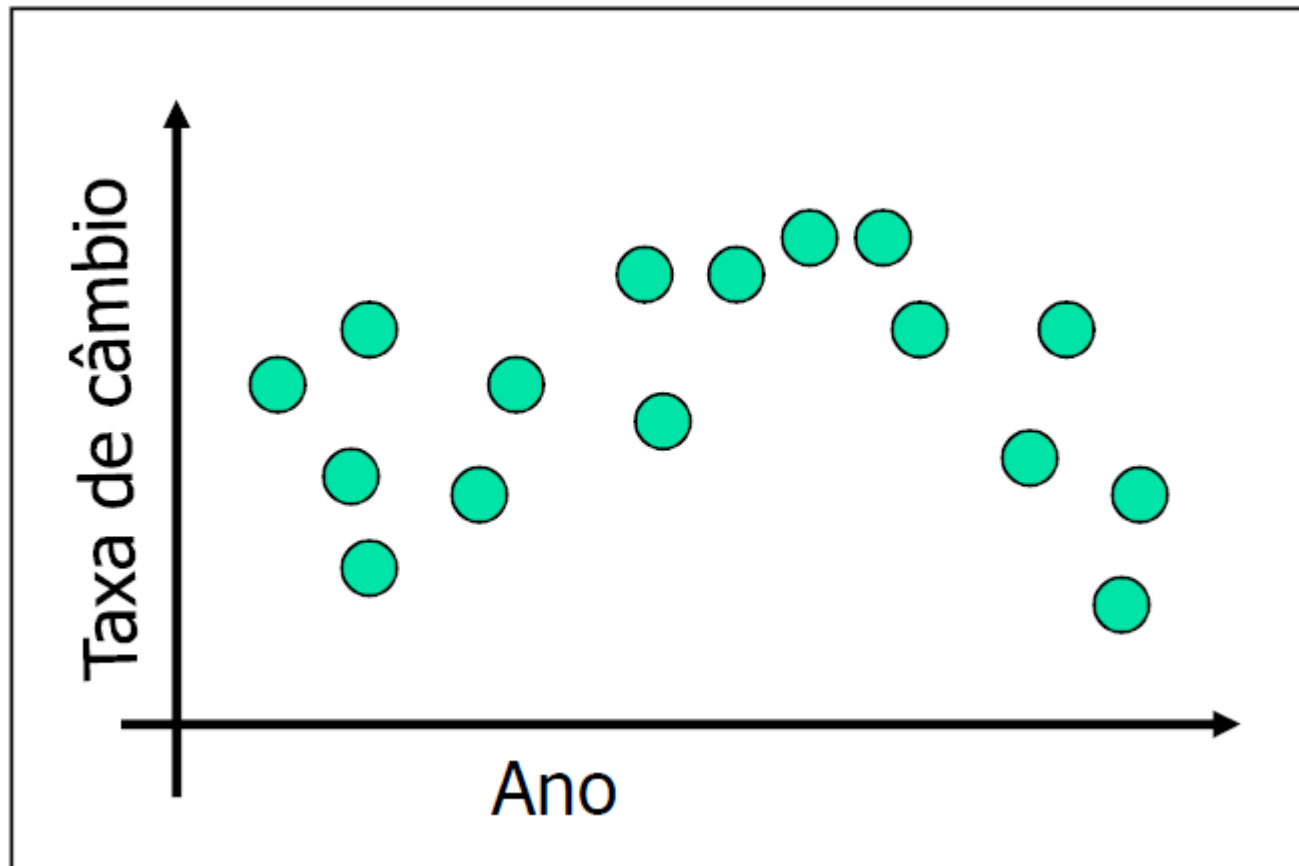
□ Problema linear



Regressão

76

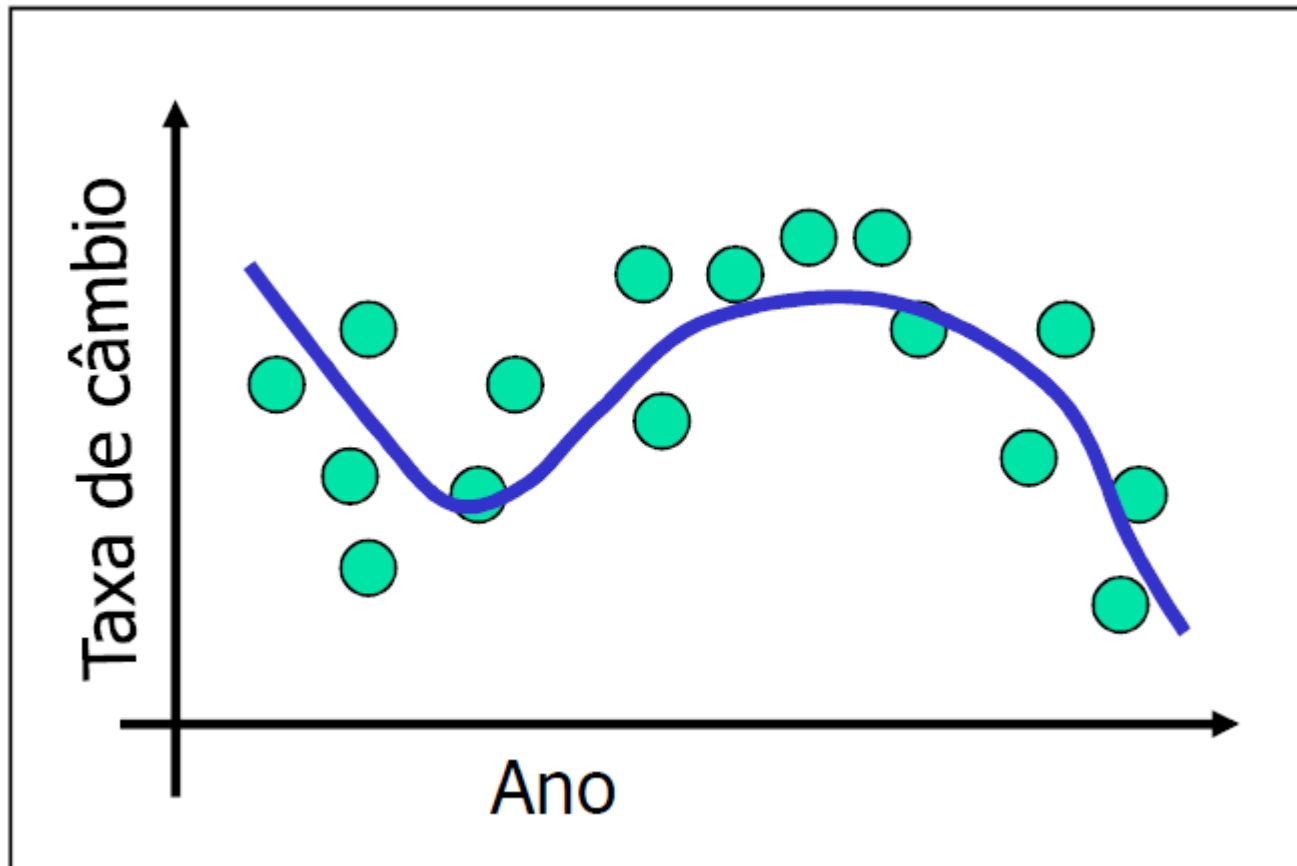
- Como definir taxa de câmbio de um ano?



Regressão

77

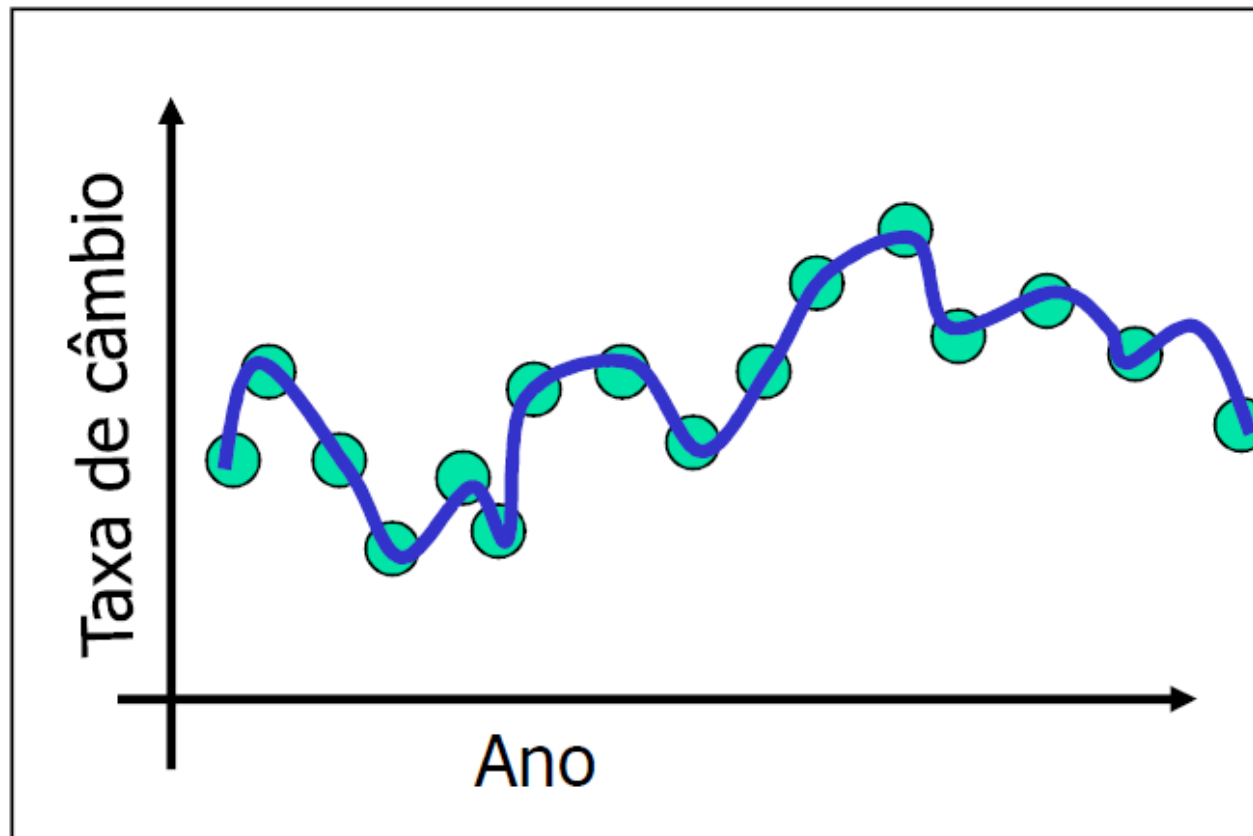
□ Problema não linear



Regressão

78

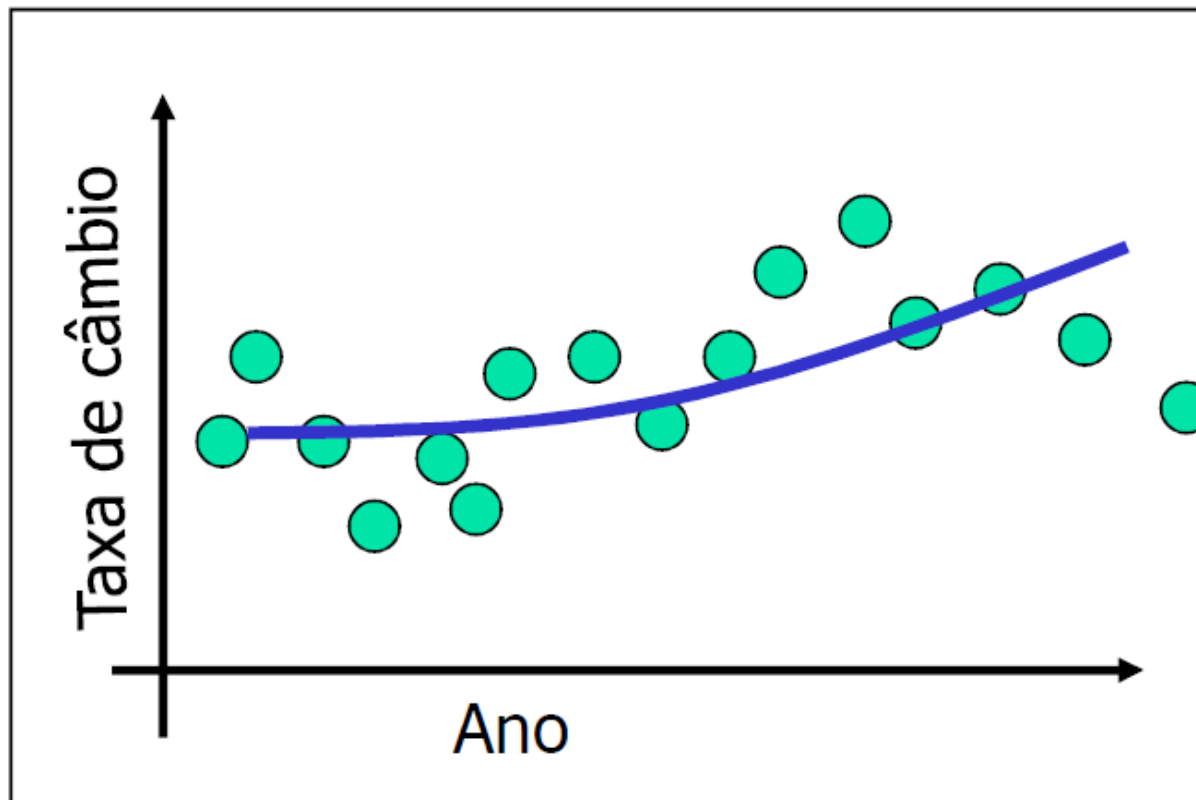
- Overfitting
 - ▣ Também presente em classificação



Regressão

79

- Underfitting
 - ▣ Também presente em classificação



Detecção de Anomalias

80

- Técnica identifica padrões cujas características são significativamente diferentes dos demais padrões
 - ▣ Exemplos de aplicação
 - Detecção de fraudes
 - Intrusões em redes de computadores
 - Perturbações em ecossistemas
- Como busca prever se um padrão novo é uma anomalia ou não, é vista como uma tarefa preditiva. No entanto, não é tratada necessariamente sob a ótica de aproximação de funções...
 - ▣ Pode fazer uso, p. ex., de clustering (detecção de outliers)

Tarefas Descritivas

81

- Encontrar padrões estruturais (correlações, grupos, outliers,...) em conjuntos de dados
- Tarefas fundamentais:
 - ▣ Análise de Agrupamento
 - ▣ Análise de Associação
- Usualmente associadas com o paradigma de treinamento não supervisionado

Agrupamento (clustering)

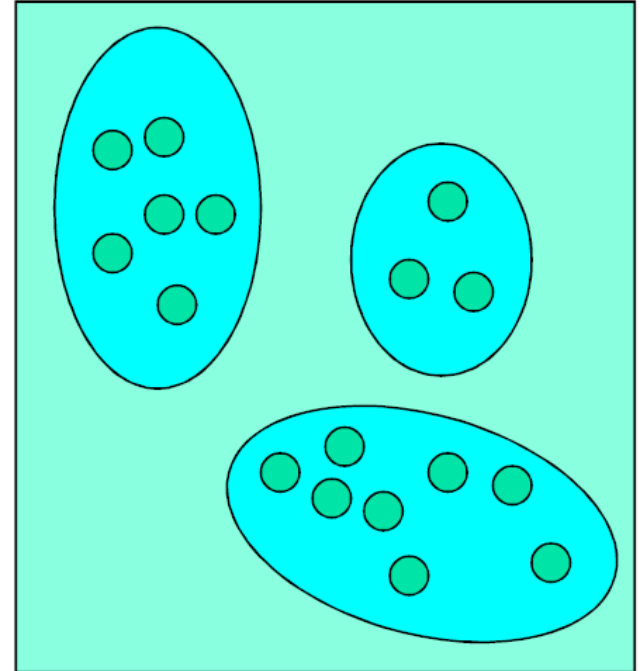
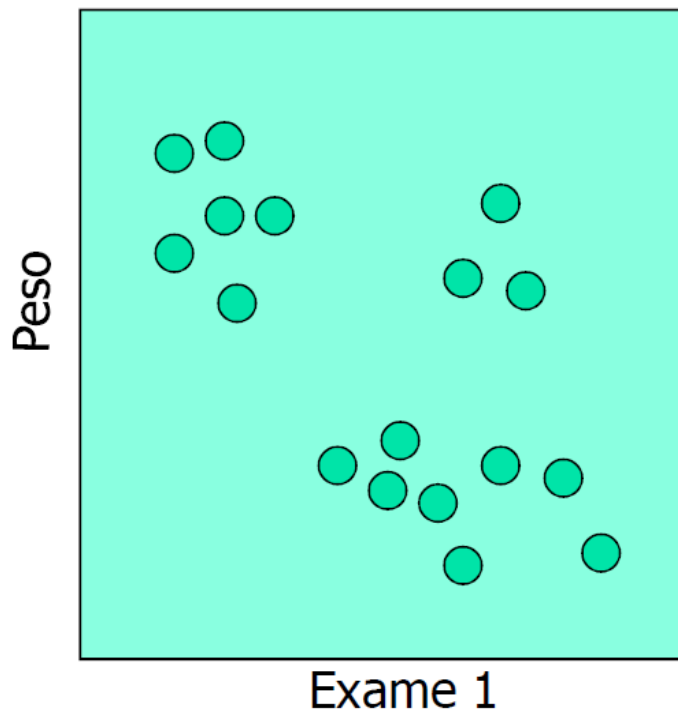
82

- Técnica explora semelhanças entre padrões para agrupar padrões parecidos entre si
 - ▣ Padrões em diferentes grupos devem ser dissimilares

- Exemplos
 - ▣ Segmentação de mercado (consumidores), de clientes, ...
 - ▣ Agrupamento de genes
 - ▣ ...

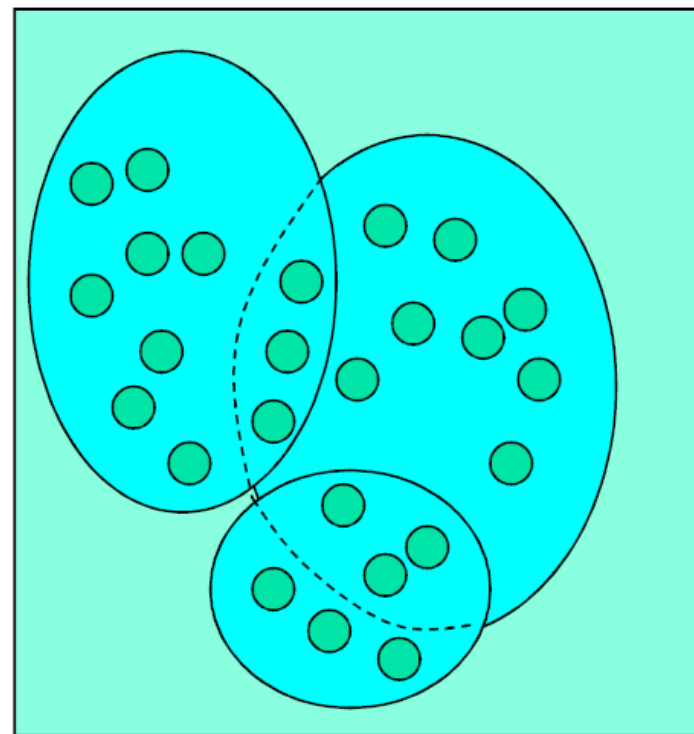
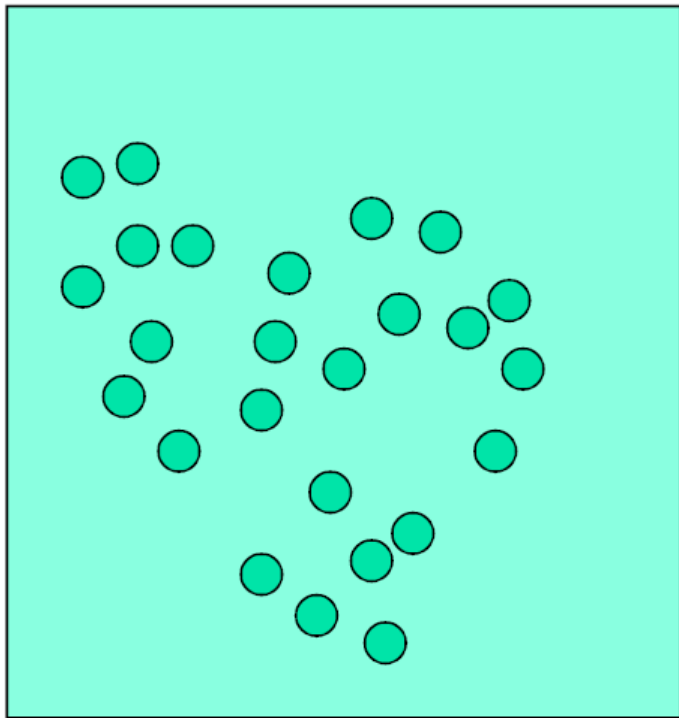
Agrupamento

83



Agrupamento

84



Associação

85

- Técnica descobre relações simétricas ou assimétricas entre conjuntos de padrões
- Exemplos de regras de associação
 - ▣ $\{\text{Fraldas}\} \Rightarrow \{\text{Leite}\}$ (útil, porém previsível)
 - ▣ $\{\text{Fraldas}, \text{Leite}\} \Rightarrow \{\text{Cerveja}\}$ (útil e inovadora)

Links Interessantes: Tutoriais, Pacotes, Conjuntos de Dados, ...

86

- Machine Learning Data Repository UC Irvine
 - ▣ <http://archive.ics.uci.edu/ml/>
- Free Matlab Toolbox for Pattern Recognition
 - ▣ <http://www.prtools.org>
- Tom Mitchell's webpage
 - ▣ <http://www.cs.cmu.edu/~tom/>
- Weka Software
 - ▣ <http://www.cs.waikato.ac.nz/ml/weka/>