

Redes Neurais (Conceitos Fundamentais)

Germano C. Vasconcelos
Centro de Informática - UFPE

Interdisciplinaridade



Redes Neurais é uma área interdisciplinar:

- Neurofisiologia
- Psicologia Cognitiva
- Matemática
- Física
- Medicina
- Percepção Humana
- Engenharias
- Ciência da Computação



Não é uma Área Nova ...



Trabalhos iniciais:

- McCulloch & Pitts (1943) - modelo de neurônio
- Donald Hebb (1949) - regra de aprendizagem
- Frank Rosenblatt (1962) - topologia de rede



Popularidade Mais Recente?



- Aparecimento de novos paradigmas (Deep Learning)
- Análise fundamentada em métodos científicos
- Avanço tecnológico dos computadores
- Sucesso em aplicações do mundo real
- Grandes players no mercado (Google, Facebook, Nvidia, Apple, etc. etc etc)

O que são Redes Neurais?

Sistemas paralelos distribuídos compostos por unidades de processamento simples interligadas entre si e com o ambiente por um número de conexões

De outra maneira...

São modelos inspirados na estrutura paralela do cérebro e que buscam reter algumas de suas propriedades...

Unidades



Neurônios

Interconexão



Redes Neurais

O que são Redes Neurais ?



Geralmente conexões estão associadas a pesos que armazenam o conhecimento da rede e servem para ponderar a entrada recebida pelo neurônio

Conhecimento



Aprendizagem



Semelhança ao Cérebro



- Elementos básicos (neurônios)
- Rede de processadores interconectados
- Conhecimento armazenado em conexões
- Conhecimento adquirido através de aprendizagem
- Estrutura inerentemente paralela
- Representações distribuídas



Elementos Básicos



- Vasta quantidade de modelos
- Alguns elementos presentes em todos
 - A função das unidades de processamento
 - A topologia da rede
 - A estratégia ou algoritmo de aprendizagem



A Função das Unidades



Caracterizado por três elementos básicos:

- Um conjunto de conexões de entrada
- Um estado de ativação
- Um valor de saída (resposta)

O comportamento do neurônio

- Ocorre um estímulo como entrada
- Ocorre uma computação do estado de ativação em função do estímulo
- Ocorre uma resposta em função da ativação



Modelos para o Neurônio

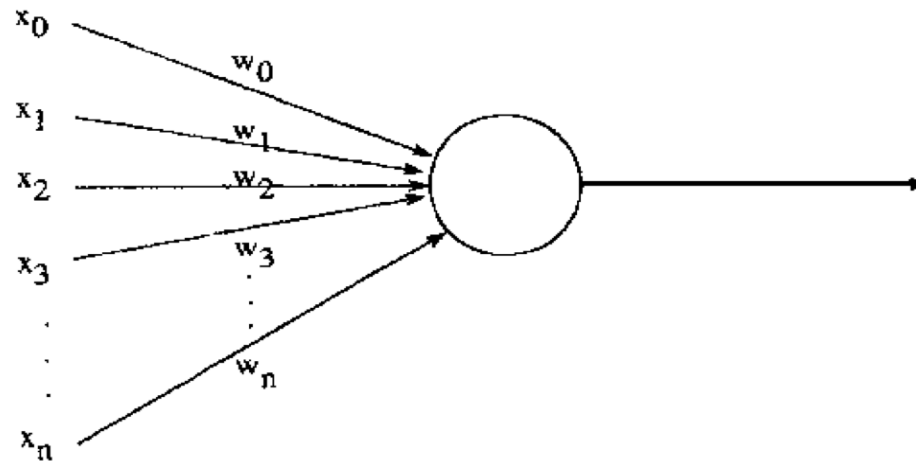


A definição da unidade de processamento caracteriza-se por:

- Regra de Propagação (estado de ativação)
- Função de Ativação (resposta do neurônio)



Exemplo: Neurônio de McCulloch & Pitts (MCP)

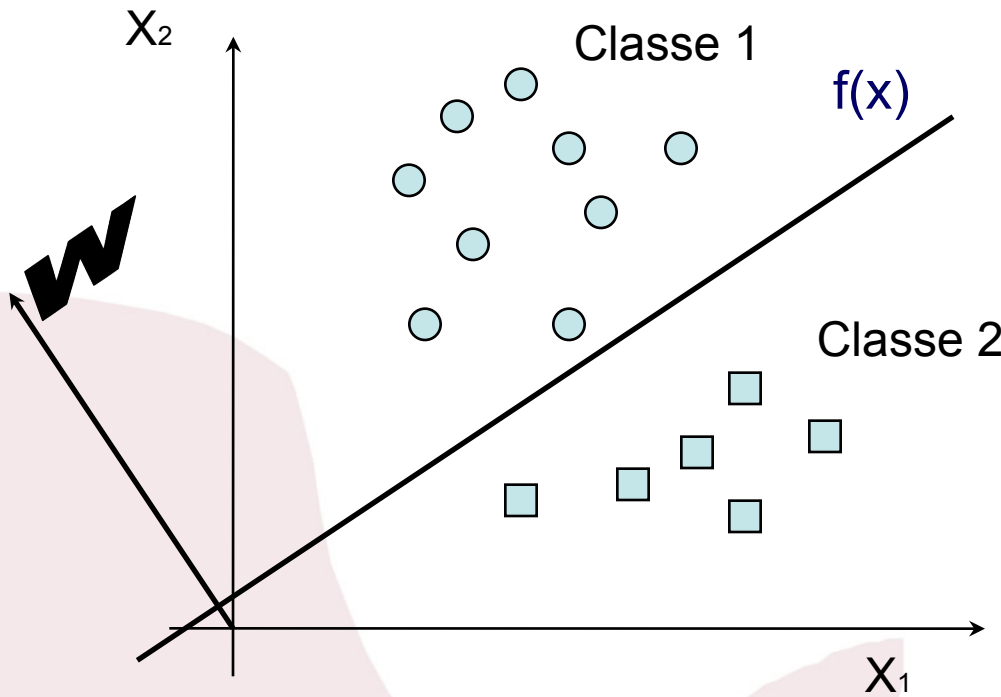


$$y = f_h \left[\sum_{i=1}^n w_i x_i - \theta \right]$$

Função de Ativação

Regra de Propagação

Visão Matemática do MCP



$$f(x) = \sum w_i \cdot x_i - \theta$$

$$f(x) = (|W| \cdot |X| \cos \Phi) - \theta$$

Considere o ponto onde

$$f(x) = 0:$$

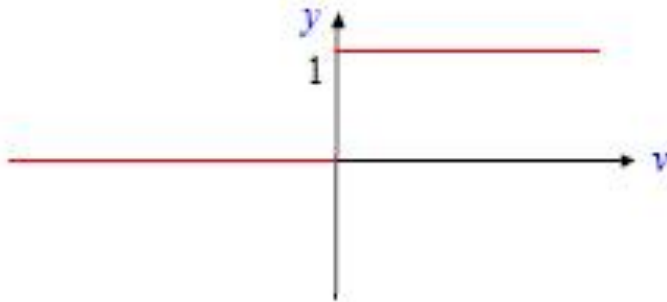
$$w_1 \cdot x_1 + w_2 \cdot x_2 - \theta = 0$$

$$x_2 = -w_1/w_2 \cdot x_1 + \theta/w_2$$

$$(y = m \cdot x + c)$$

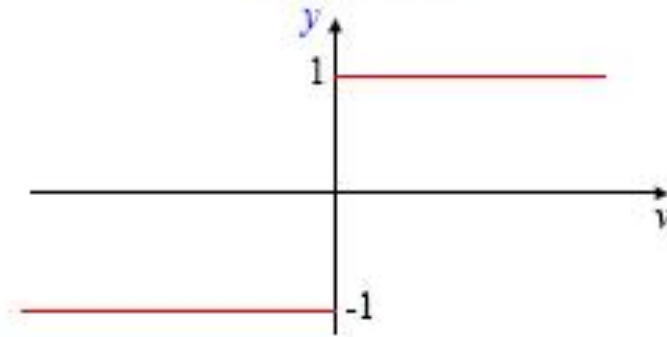
Funções de Ativação I

Função Limiar



$$y = \begin{cases} 1, & v \geq 0 \\ 0, & v < 0 \end{cases}$$

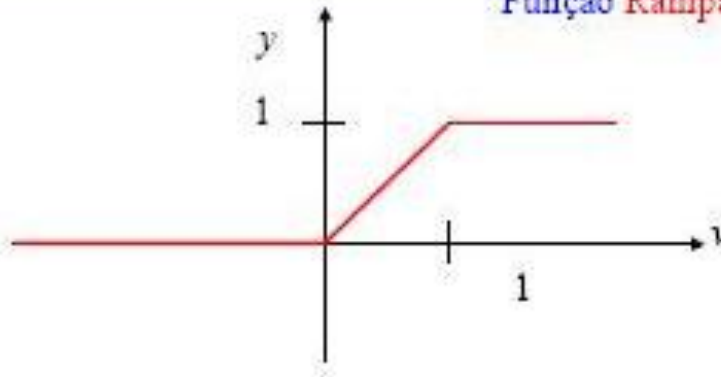
Função Sinal



$$y = \begin{cases} +1, & v \geq 0 \\ -1, & v < 0 \end{cases}$$

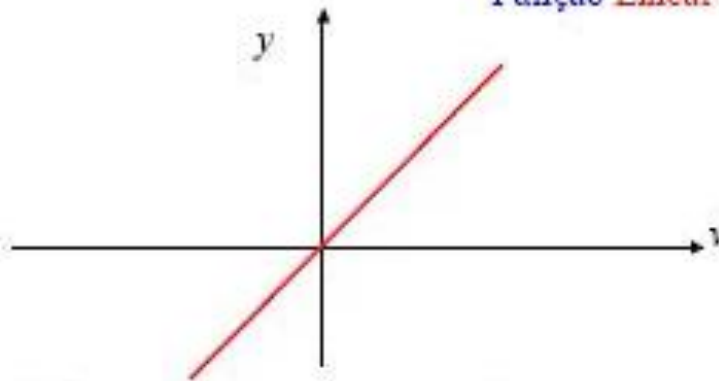
Funções de Ativação II

Função Rampa



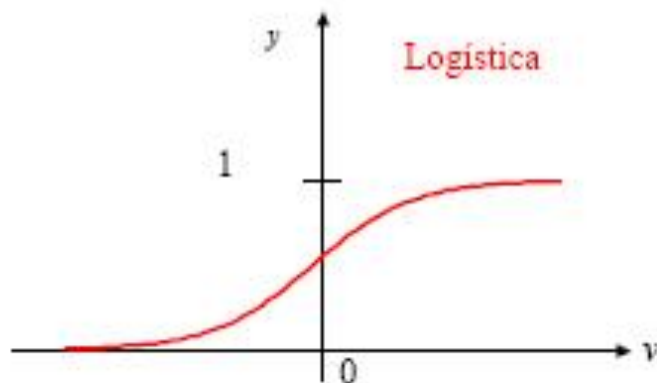
$$y = \begin{cases} 0, & v < 0 \\ v, & 0 \leq v < 1 \\ 1, & v > 1 \end{cases}$$

Função Linear

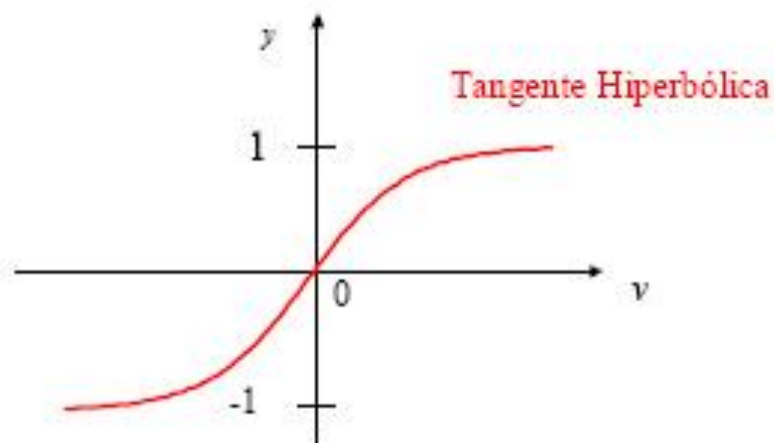


$$y = v$$

Funções de Ativação III - Sigmoids



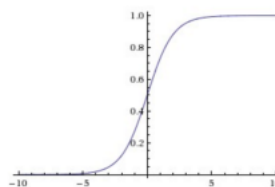
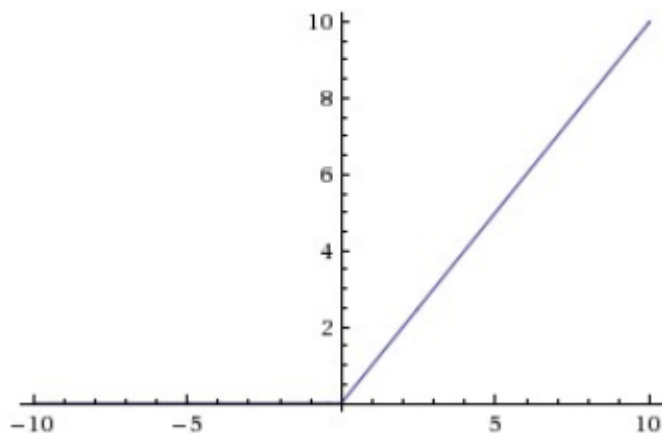
$$y = \frac{1}{(1 + e^{-v})}$$



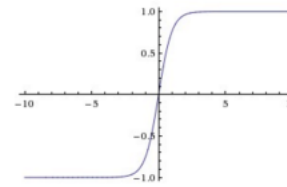
$$y = \tanh\left(\frac{v}{\alpha}\right) = \frac{1 - e^{-v/\alpha}}{1 + e^{-v/\alpha}}$$

Funções de Ativação IV – ReLU (Rectified Linear Unit)

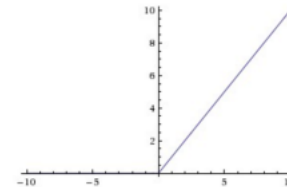
$$f(x) = \max(0, x)$$



Sigmoid

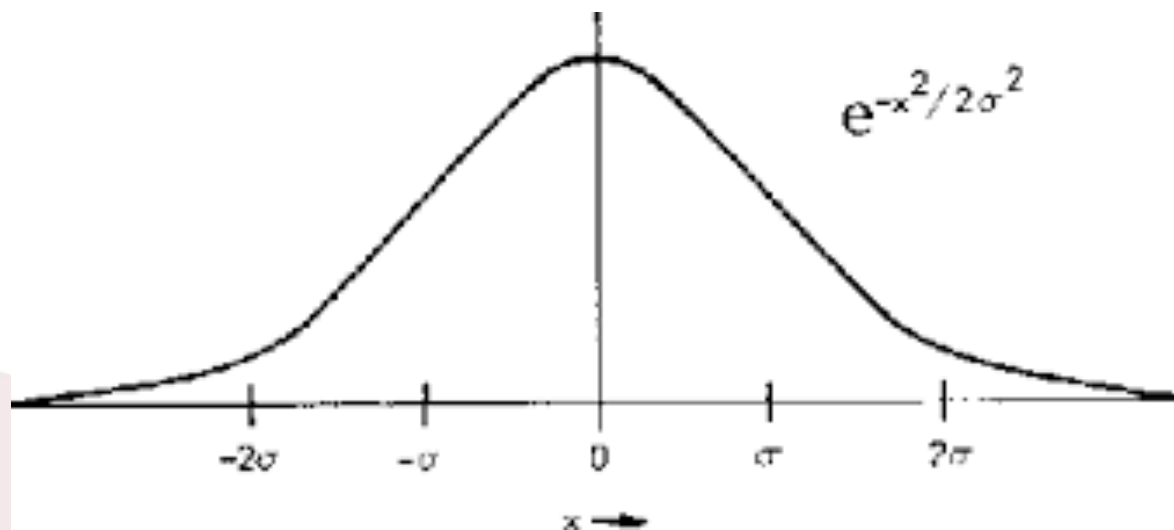


tanh

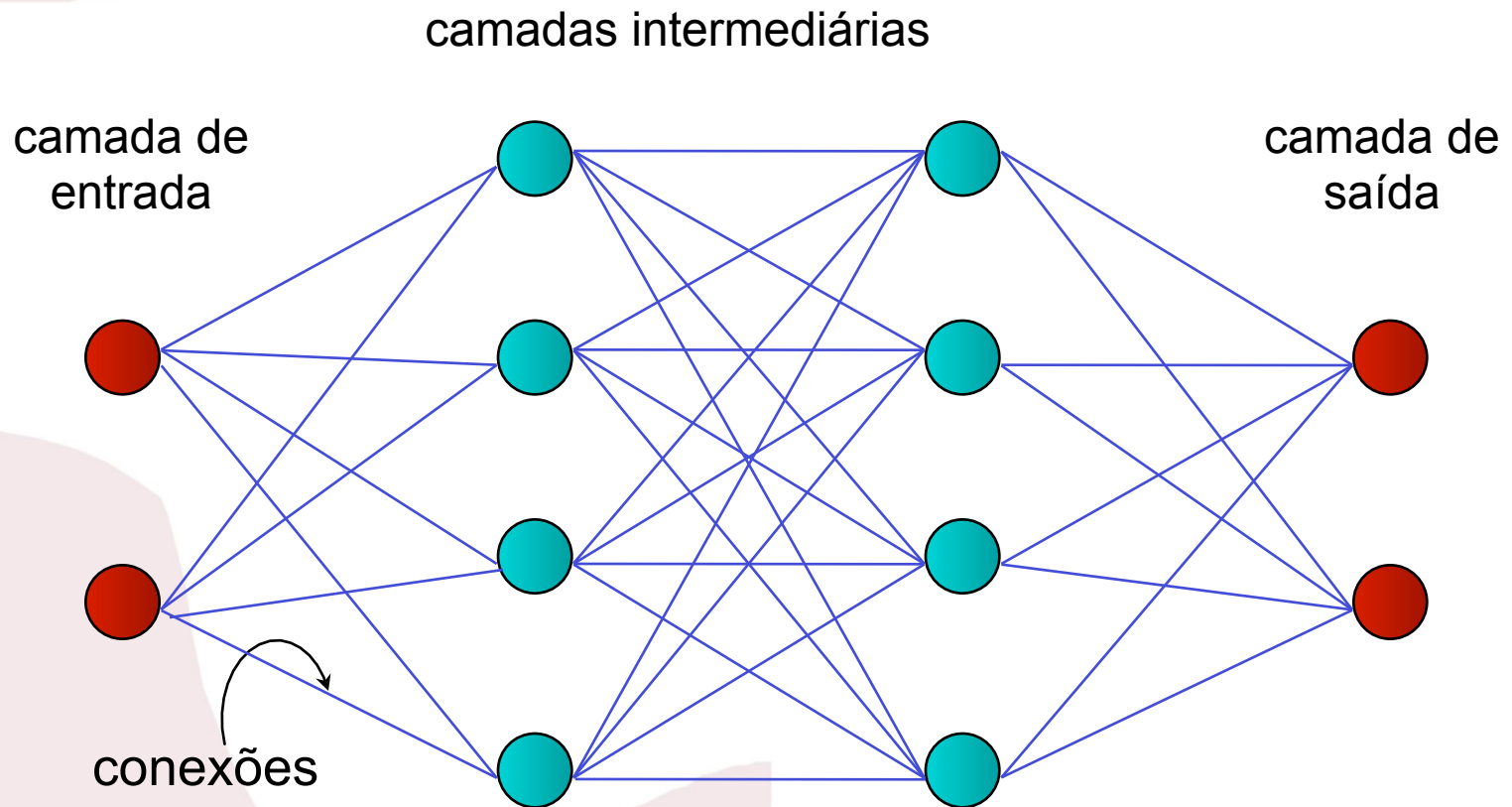


ReLU

Funções de Ativação V – Gaussiana

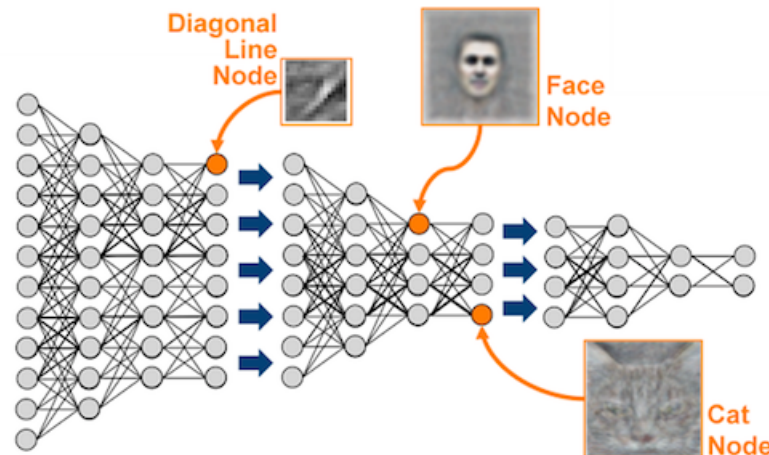
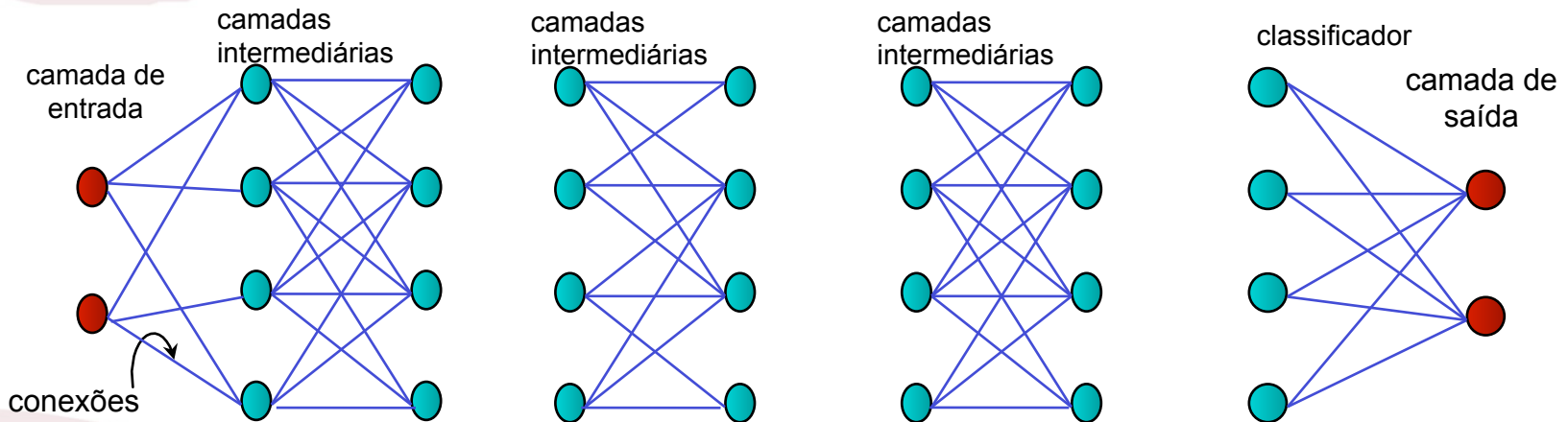


Topologias de Redes: Feedforward 1 (Rasas - Shalow)



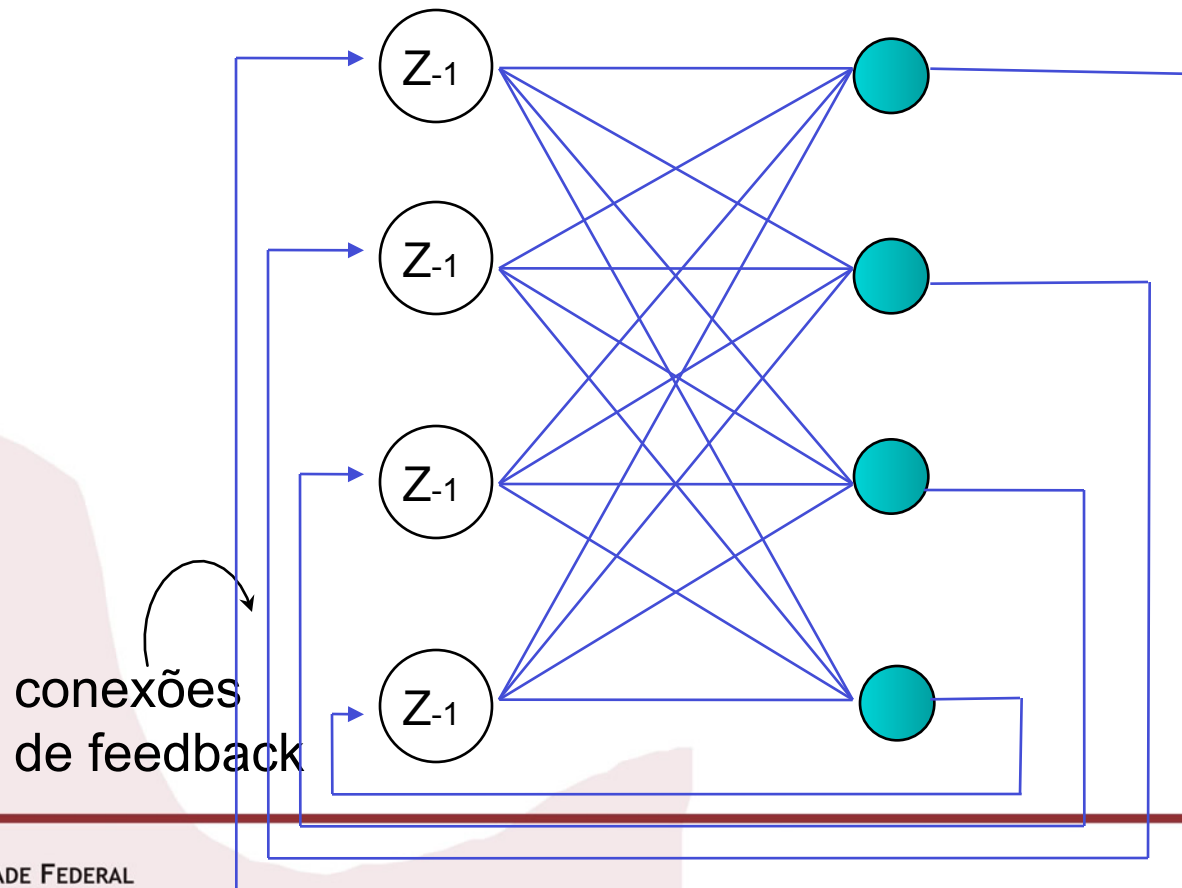
Topologias de Redes:

Feedforward 2 (Profundas - Deep)

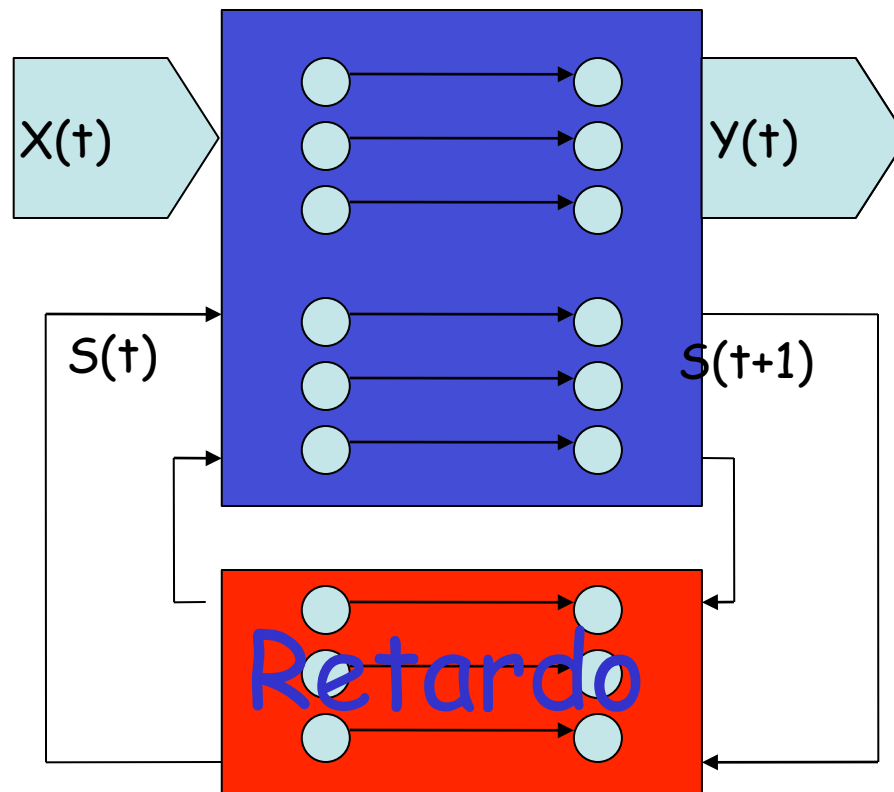


Topologias de Redes: Recorrente I

camada de
processamento

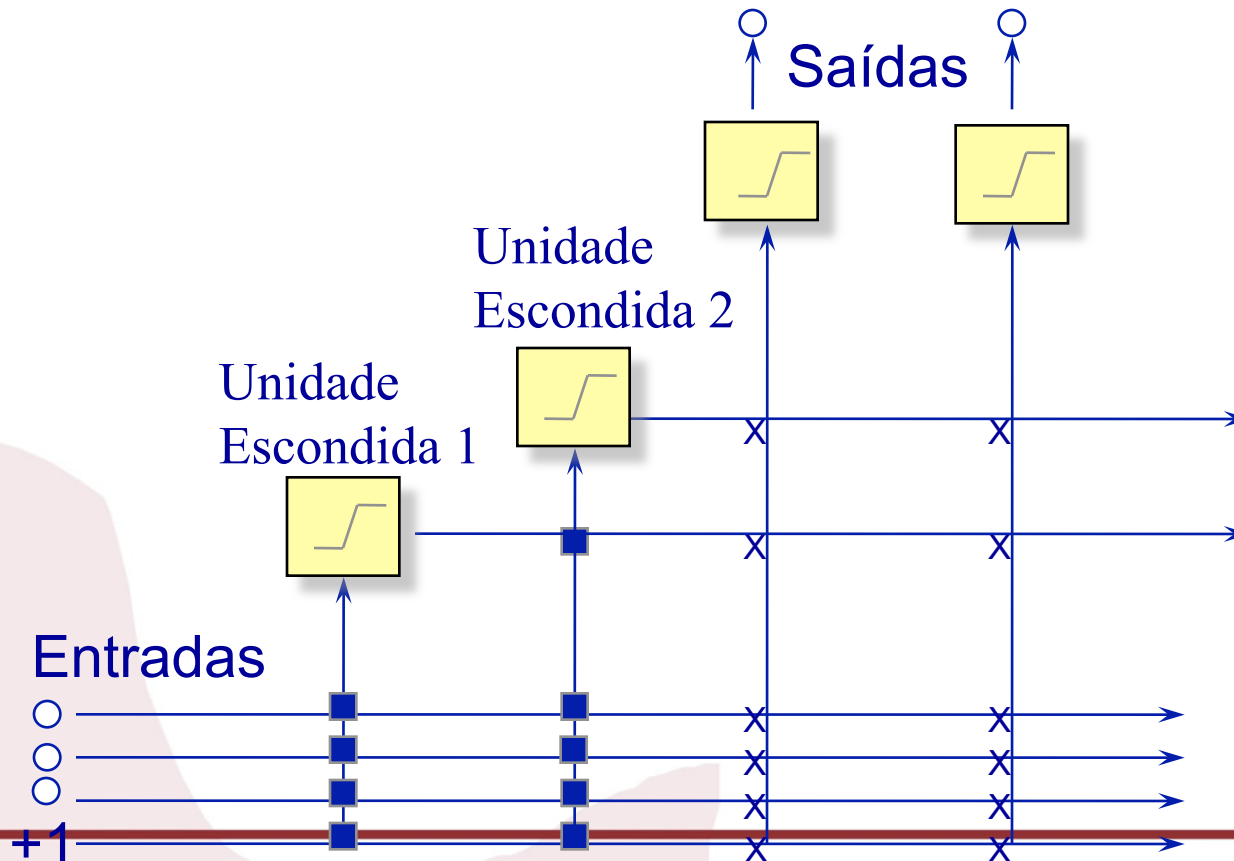


Topologias de Redes: Recorrente II



Topologias de Redes: Construtiva

$$C = \sum_o |\sum_p (y_{po} - y_o)(e_{po} - e_o)| \quad \text{—}$$



Aprendizagem



- Aprendizagem é aquisição de conhecimento
- “Conhecimento” ?
 - Termo vago. Depende de sob que ponto de vista é considerado



Diferentes Pontos de Vista



- Psicologia Cognitiva
- Psicologia Experimental
- Ciências Exatas



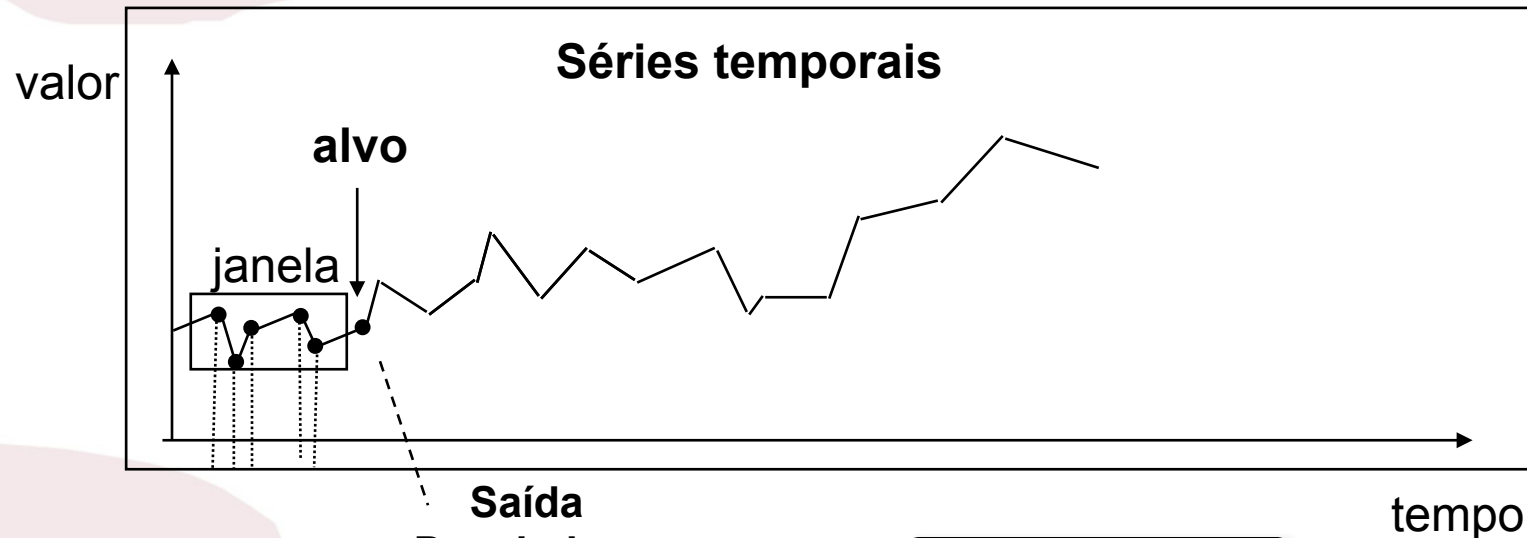
Aprendizagem nas Ciências Exatas



- Aprendizagem é aquisição e processamento de dados para aproximar e/ou representar funções (conhecimento)
- Aproximação de funções :
 - Respostas no domínio discreto: Classificação (e. g. decisão de conceder ou não crédito ao consumidor)
 - Respostas no domínio contínuo: Regressão (e. g. previsão do limite de crédito a ser concedido)

Aprendizagem

- Previsão: estimar valores futuros de uma série baseado em valores do presente e do passado;



Entradas
da rede =
 n valores
passados

Saída
Desejada =
valor da série
 k passos à
frente

Definição da
janela de entrada

Definição da
janela de saída

Ex: valor um
passo à frente

26

Computability x Learnability



- **Computability (Computabilidade):** é o poder computacional de um sistema, expresso em termos do universo de funções que ele é capaz de computar. (Requisitos)
 - MLP, RBF
 - pRAM
- **Learnability:** é o poder computacional que os algoritmos de aprendizagem oferecem a um sistema adaptativo expresso em termos do universo de funções que eles são capazes de fazer o sistema efetivamente computar. (Requisitos)
 - MLP, RBF
 - pRAM



Tipos de Aprendizagem I

(pelo grau de feedback)



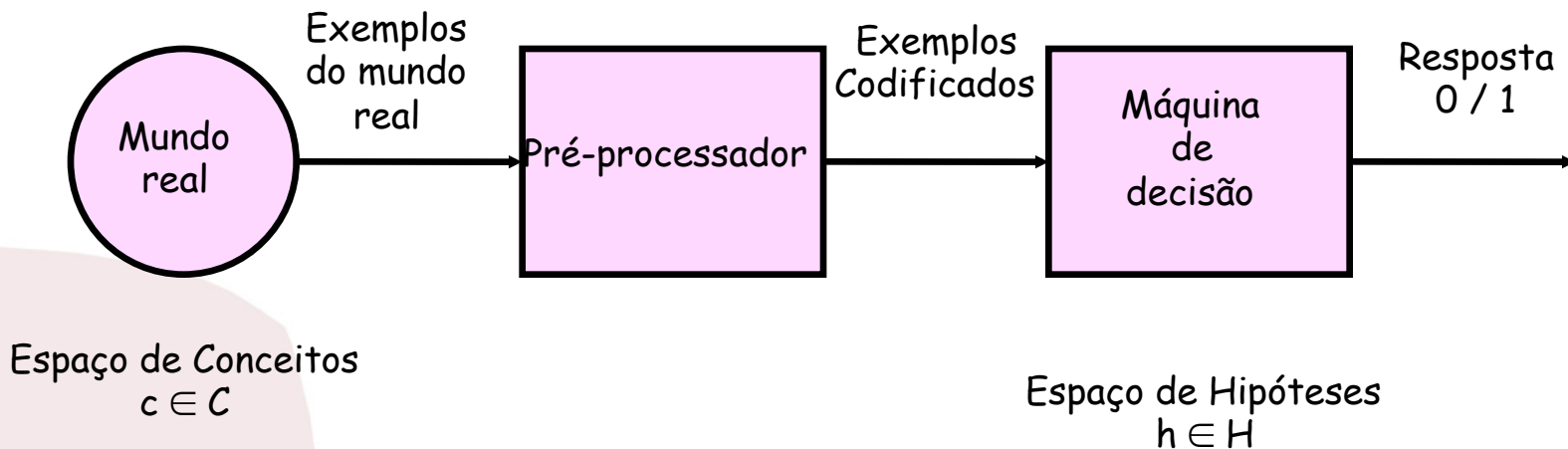
- **Supervisionada:** um “professor” diz quanto a resposta fornecida pelo sistema se aproxima da resposta desejada (e. g. nota de um aluno numa prova)
- **Por Reforço:** um “crítico” diz apenas se a resposta fornecida está certa ou errada (e. g. punição/recompensa no treinamento de animais)
- **Não-Supervisionada:** o sistema tenta se auto-organizar baseado nas similaridades entre os exemplos apresentados (e. g. desenvolvimento das células simples do córtex visual estriado)

Tipos de Aprendizagem II

(pelo grau de feedback)

- Supervisionada:
 - Conjunto de treinamento $s = \{(x_1, f(x_1)), (x_2, f(x_2)), \dots, (x_n, f(x_n))\}$
 - Convergência rápida
- Por Reforço:
 - Conjunto de treinamento $s = \{(x_1, \text{sgn}[f(x_1)]), (x_2, \text{sgn}[f(x_2)]), \dots, (x_n, \text{sgn}[f(x_n)])\}$
 - Convergência média
- Não-Supervisionada:
 - Conjunto de treinamento $s = \{(x_1,), (x_2,), \dots, (x_n,)\}$
 - Convergência lenta

Exemplo de Classificação Supervisionada



Exemplo de Classificação Supervisionada



- CASO ESPECÍFICO

Aprendizagem é o processo de modificação do estado da máquina M com base nos exemplos apresentados visando a deixá-la apta a classificar exemplos não-vistos do mesmo problema

- CASO GERAL

Aprendizagem é o processo de escolha da função $h \in H$ que melhor aproxime a função $c \in C$ a partir dos exemplos contidos no conjunto de treinamento

- DEFINIÇÃO

Algoritmo de aprendizagem é a função L que, a partir dos exemplos da amostra s , seleciona uma função $h \in H$ para aproximar o conceito $c \in C$

- CRITÉRIOS DE AVALIAÇÃO

Consistência:

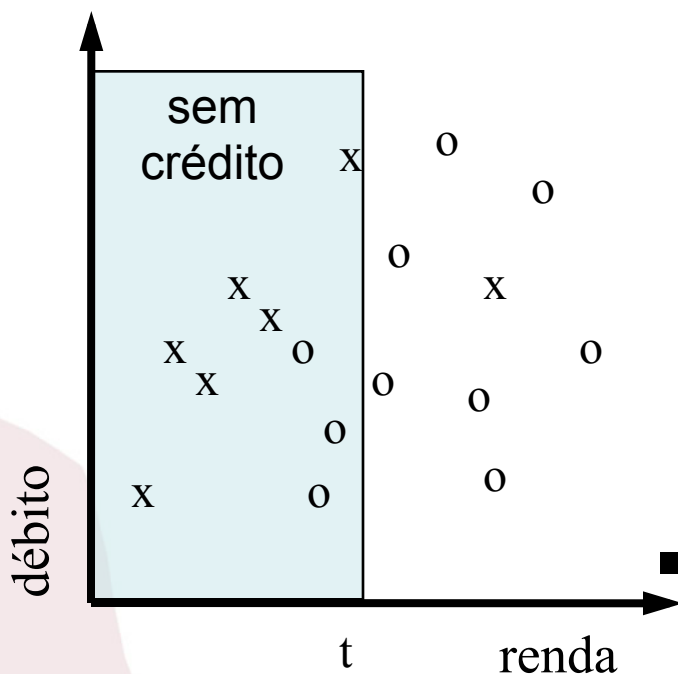
$$h(x_i) = f(x_i) \quad \forall (x_i, f(x_i)) \in s$$

Convergência:

Melhoria da aproximação com o aumento do número de exemplos no conjunto de treinamento

Exemplo de Análise de Crédito (I)

Análise de crédito

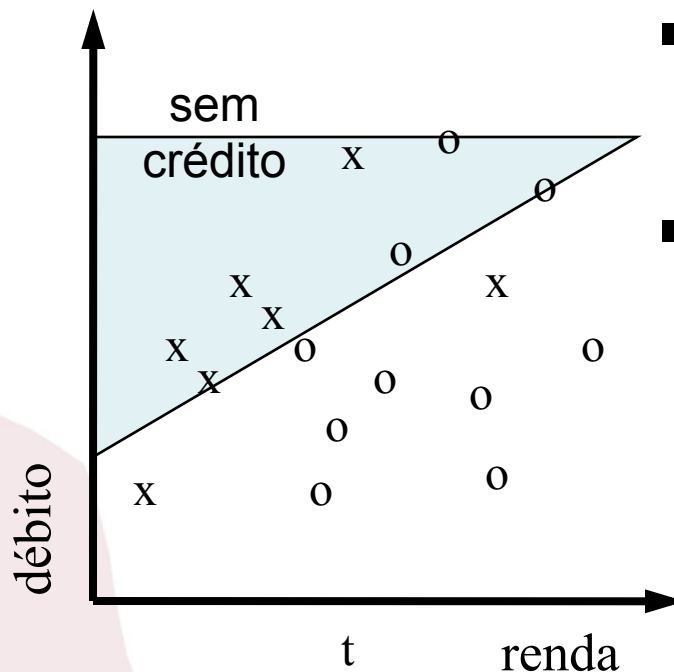


x: exemplo recusado
o: exemplo aceito

- Hiperplano paralelo: pode ser interpretado diretamente como uma regra:
 - se a renda é menor que t , então o crédito não deve ser liberado
- Exemplo:
 - árvores de decisão
 - indução de regras

Exemplo de Análise de Crédito (II)

Análise de crédito



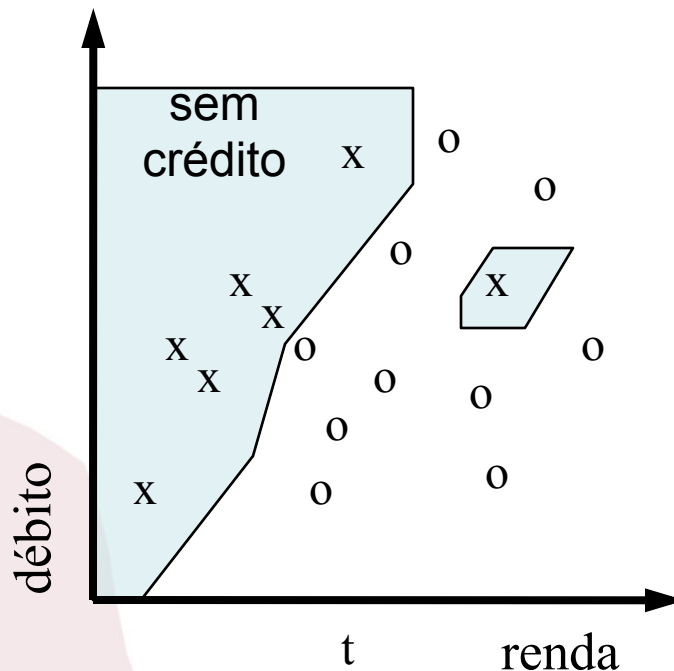
- Hiperplano oblíquo: melhor separação:
- Exemplos:
 - Perceptron
 - regressão linear

x: exemplo recusado

o: exemplo aceito

Exemplo de Análise de Crédito (III)

Análise de crédito

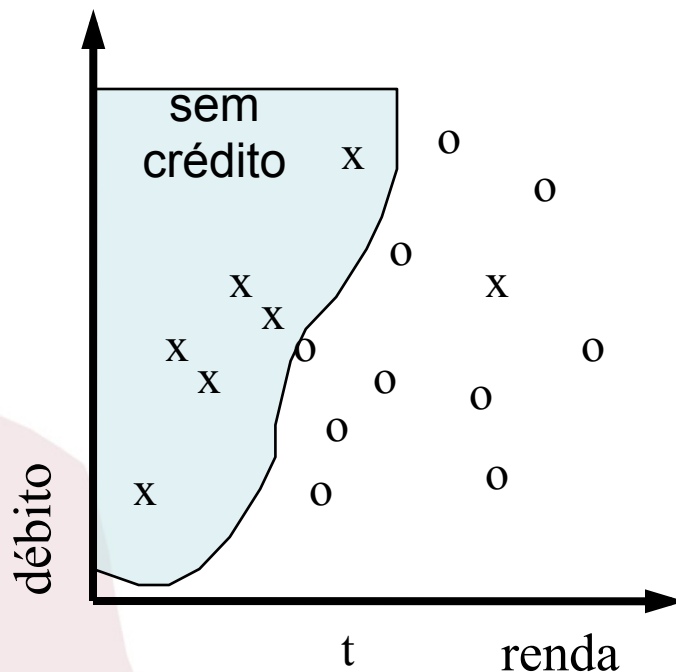


x: exemplo recusado
o: exemplo aceito

- Superfícies lineares por partes não contíguas
- Exemplos:
 - raciocínio baseado em casos
 - MLP-perceptrons multicamadas com função de saída sigmóide (squashing)
 - regressão linear por partes

Exemplo de Análise de Crédito (IV)

Análise de crédito



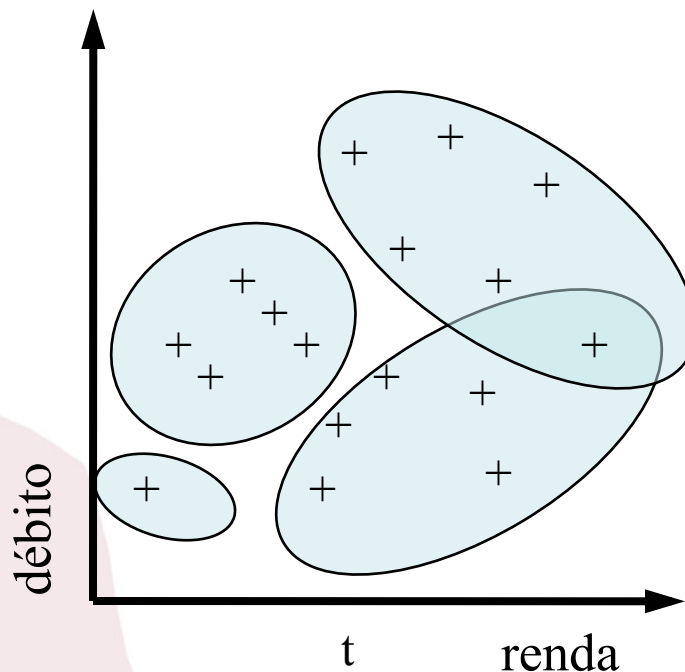
x: exemplo recusado

o: exemplo aceito

- Superfície não linear: melhor poder de classificação, pior interpretação
- Exemplos:
 - MLP perceptrons
 - Cascade Correlation
 - k-vizinhos mais próximos
 - regressão não-linear

Exemplo de Análise de Crédito (V)

Análise de crédito



+ : exemplo

- Superfícies Elípticas
- Exemplos:
 - RBF-Radial Basis Functions
- Agrupamento
- Exemplos:
 - Kohonen-vector quantization
 - Adaptive Resonance Theory (ART)